Self-Adjoint Operators in Hilbert Spaces*

Trabajo para la obtención del título de Graduado en Matemáticas por la UCM

by Ana Cañizares García†

March 2012

Abstract

We present the spectral theory for self-adjoint – bounded and unbounded – operators in Hilbert spaces. In the section dedicated to bounded operators we give two formulations of the spectral theorem: the first one involves the notion of spectral family and gives a representation of the operator as a spectral integral with respect to this family; the second one asserts that a self-adjoint bounded operator is essentially a multiplication (unitarily equivalent to one). In the last section we extend this theory to unbounded operators, the key for this construction being the Cayley transform and the spectral theory for unitary operators, which we also briefly discuss.

1 Bounded Operators

1.1 Spectrum and general results

In this introductory part we will deal with bounded linear operators defined on Banach spaces over the complex field $\mathbb{C}$.

Consider a bounded operator $A: X \to X$ (meaning $\operatorname{Dom} A = X$ and $\operatorname{Im} A \subseteq X$), where X is a Banach space. We say that A is **invertible** if there exists an operator $B\,(=: A^{-1})$ defined on X such that $B\,A = I$ and $A\,B = I$ (in finite dimensional Banach spaces one identity implies the other but not in the infinite dimensional case).

A complex number $\lambda \in \mathbb{C}$ is called a **regular point** of A if the operator $(A - \lambda\,I)^{-1}$ defined on X exists and is bounded. If $\lambda \in \mathbb{C}$ is not a regular point, we say it is a **spectrum point** and the set of all such points is called the **spectrum** $\sigma(A)$ of A. We can decompose the spectrum as follows:

(i) The **point spectrum** $\sigma_p(A)$ is the set of eigenvalues of A; that is, $\lambda \in \sigma_p$ if and only if $\ker(A - \lambda\,I) \neq 0$.

(ii) The **continuous spectrum** $\sigma_c(A)$ is the set of all $\lambda \in \mathbb{C}$ such that $\ker(A - \lambda I) = 0$ and $\operatorname{Im}(A - \lambda\,I)$ is dense in X.

(iii) The **residual spectrum** $\sigma_r(A)$ is the set of all $\lambda \in \mathbb{C}$ such that $\ker(A - \lambda I) = 0$ and $\overline{\operatorname{Im}(A - \lambda\,I)} \neq X$.

The spectrum is indeed the union of these three disjoint sets since λ is a regular point if and only if $\ker(A - \lambda\,I) = 0$ and $\operatorname{Im}(A - \lambda\,I) = X$, for if $(A - \lambda\,I): X \to X$ is bounded and bijective, its inverse is also bounded, by the open map theorem.

We now give some useful results about invertibility and properties of the spectrum of bounded operators.

∗. This document has been written using the GNU $\mathrm{T_EX_{MACS}}$ text editor (see `www.texmacs.org`).

†. Under the supervision of Prof. Fernando Bombal.

If $\|A\| = q < 1$, *then* $(I - A)$ *is invertible and* $(I - A)^{-1} = \sum_0^\infty A^k$. *Moreover,* $\|(I - A)^{-1}\| \leqslant 1/(1 - \|A\|)$.

Indeed, $\|A^k\| \leqslant \|A\|^k = q^k \to 0$ as $k \to \infty$. The sequence $S_n = \sum_0^n A^k$ is Cauchy, so it converges. Finally, $(I - A)\, S_n = I - A^{n+1} = S_n\, (I - A) \to I$ as $n \to \infty$. Therefore, $\lim S_n$ is the inverse operator of $(I - A)$.

Let A *be an invertible operator and let* B *be such that* $\|A - B\| < 1/\|A\|^{-1}$. *Then* B *is also invertible.*

Write $B = A\,(I - A^{-1}\,(A - B))$. Since $\|A^{-1}\,(A - B)\| < 1$ the operator $I - A^{-1}\,(A - B)$ is invertible. Now notice that the product of two invertible operators is invertible.

$\sigma(A)$ *is compact and is bounded by* $\|A\|$.

For if $|\lambda| > \|A\|$, then $I - \lambda^{-1} A$ is invertible and so is $A - \lambda\, I = -\lambda\,(I - \lambda^{-1} A)$; hence, $\lambda \notin \sigma(A)$. Now we prove that $\mathbb{C}\backslash\sigma(A)$ is open. Let $\lambda \notin \sigma(A)$. Then $A - \lambda\, I$ is invertible and, if $|\mu - \lambda| < \delta$, $\|(A - \lambda\, I) - (A - \mu\, I)\| = \|(\mu - \lambda)\, I\| < \delta$. Choosing $\delta = 1/\|(A - \lambda\, I)^{-1}\|$, we have that $A - \mu\, I$ is invertible, so the open ball with center λ and radius δ is contained in $\mathbb{C}\backslash\sigma(A)$.

The next result is a nice application of the Liouville theorem for holomorphic entire functions. We just outline the proof; for a complete one see [Rud86] (page 359).

Let $B(X)$ *be the set of all linear bounded operators on the Banach space* X. *Then, for every* $A \in B(X)$, $\sigma(A)$ *is not empty.*

Let $A \in B(X)$ and fix $\lambda_0 \notin \sigma(A)$. Then $(A - \lambda_0\, I)^{-1} \neq 0$ and the Hahn-Banach theorem implies the existence of a bounded linear functional Φ on $B(X)$ (since $B(X)$ is a Banach space with the operator norm) such that $\Phi((A - \lambda_0\, I)^{-1}) \neq 0$. Define, for $\lambda \notin \sigma(A)$, $f(\lambda) = \Phi((A - \lambda\, I)^{-1})$. It turns out that f is holomorphic on $\mathbb{C}\backslash\sigma(A)$ and $f(\lambda) \to 0$ as $\lambda \to \infty$. If $\sigma(A)$ were empty, then f would be an entire function and, by Liouville's theorem, it would be $f \equiv 0$, which contradicts $f(\lambda_0) \neq 0$.

1.2 Self-adjoint operators on Hilbert spaces

Let now $\mathcal{H}$ be a Hilbert space and let $A: \mathcal{H} \to \mathcal{H}$ be a bounded operator. For each $y \in \mathcal{H}$, by the Riesz representation theorem, there exists a unique $y^* \in \mathcal{H}$ such that $\langle Ax, y\rangle = \langle x, y^*\rangle$ for all $x \in \mathcal{H}$. The correspondence $y \mapsto y^*$ defines a bounded operator on $\mathcal{H}$ that is called the **adjoint** of A and is denoted by A^*.

We say that A is **self-adjoint** or **symmetric** if, for every $x, y \in \mathcal{H}$, $\langle Ax, y\rangle = \langle x, Ay\rangle$, that is, if $A = A^*$. In this case, we have that $\langle Ax, x\rangle \in \mathbb{R}$ for any $x \in \mathcal{H}$ (which is in fact a characterization of symmetric operators[1]). If $P: \mathcal{H} \to \mathcal{H}$ is a **projection** ($P^2 = P$) and, furthermore, $P = P^*$, we call P an **orthoprojection**.
In the space of self-adjoint operators we define the following order: A is called **positive** and we write $A \geqslant 0$ if $\langle Ax, x\rangle \geqslant 0$ for all $x \in \mathcal{H}$. If B is also self-adjoint and $B - A \geqslant 0$, we write $A \leqslant B$. We have:

$-M\, I \leqslant A \leqslant M\, I$ *implies* $\|A\| \leqslant M$.

Indeed, a self-adjoint operator satisfies $\|A\| = \sup_{\|x\| \leqslant 1} |\langle Ax, x\rangle|$, and this supremum is less than or equal to M. We also obtain $-I\, \|A\| \leqslant A \leqslant I\, \|A\|$.

1. Express $\langle Ax, y\rangle$ as an appropiate combination of quadratic forms:

$$4\,\langle Ax, y\rangle = \langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle + i\,\langle A(x+i\,y), x+i\,y\rangle - i\,\langle A(x-i\,y), x-i\,y\rangle.$$

If $m\,I \leqslant A \leqslant M\,I$, then $\sigma(A) \subset [m, M]$.

First consider the case $m = -M$. Then $M \geqslant \|A\|$, and every $|\lambda| > M$ is a regular point. For the general case, observe that $\frac{M-m}{2}\,I \leqslant A - \frac{M+m}{2}\,I \leqslant \frac{M-m}{2}\,I$, and one can proceed as before.

We will need the next two results to define functions of self-adjoint operators. The proofs are omitted and can be found in section 6.3 of [EMT04].

Let $A_0 \leqslant A_1 \leqslant \ldots \leqslant A_n \leqslant \ldots \leqslant A$. Then there exists the ***strong limit*** *of (A_n); i.e., there exists some B such that $A_n\,x \to Bx$ for every $x \in \mathcal{H}$.* (The strong limit of bounded operators is always bounded, by the uniform boundedness principle, and the strong limit of symmetric operators is clearly symmetric).

Let $m\,I \leqslant A \leqslant M\,I$ and let P be a polynomial satisfying $P(\lambda) \geqslant 0$ for all $\lambda \in [m, M]$. Then $P(A) \geqslant 0$.

1.2.1 Functions of operators

Let $m\,I \leqslant A \leqslant M\,I$. Let $a, b \in \mathbb{R}$ be such that $a < m \leqslant M < b$ and let $K = K[a, b]$ be the set of piecewise continuous bounded functions which are monotone decreasing limits of continuous functions. We start defining an operator $\varphi(A)$ for each function φ in K to extend aferwards this concept to a larger set of functions. It follows from the Weierstraß theorem that the functions in K are also monotone decreasing limits of polynomials. This enables us to give the following definition.

Let P_n be a sequence of polynomials decreasing to φ, that is, $P_n(t) \searrow \varphi(t)$ for all $t \in [m, M]$. Since the sequence $P_n(A) \geqslant P_{n+1}(A) \geqslant \ldots$ is bounded (φ is bounded), the strong limit of $P_n(A)$ exists and we call it $\varphi(A)$.

For fixed A, the correspondence $\varphi \in K \mapsto \varphi(A)$ is well defined since the limit does not depend on the sequence P_n but only on the function φ. In fact,

if Q_n and P_n are sequences of polynomials decreasing to ψ, $\varphi \in K$ respectively, and if $\psi(t) \leqslant \varphi(t)$ for all $t \in [m, M]$, then $\lim Q_n(A) \leqslant \lim P_n(A)$.

So, if $\psi = \varphi$, the limits are the same. To prove this result, fix $t \in [m, M]$ and fix n. There exists $N_0(t)$ such that $Q_N(t) < \psi(t) + 1/n$ for all $N \geqslant N_0(t)$. Since $\psi(t) + 1/n \leqslant \varphi(t) + 1/n \leqslant P_n(t) + 1/n$, we get $Q_N(t) < P_n(t) + 1/n$. This last relation is still true on some open interval $I(t)$ around t and such intervals form an open cover of $[m, M]$. Let $\{I(t_i)\}_{i \in F}$ be a finite subcover and set $N_0 = \max_{i \in F} N_0(t_i)$. Then $Q_N(t) < P_n(t) + 1/n$ for all $N \geqslant N_0$ and all $t \in [m, M]$, which means that $Q_N(A) \leqslant P_n(A) + \frac{1}{n}\,I$. Let first $N \to \infty$ and then let $n \to \infty$ to obtain the desired result.

Since $P_n(A)$ is self-adjoint for each real polynomial $P_n(t)$, $\varphi(A)$ is self-adjoint for every function $\varphi \in K$. The mapping $\varphi \mapsto \varphi(A)$ has the expected linar, multiplicative and "order-preserving" properties: $\varphi_1 + \varphi_2 \mapsto \varphi_1(A) + \varphi_2(A)$; for $c > 0$, $c\,\varphi_1 \mapsto c\,\varphi_1(A)$; if $\varphi_1 \cdot \varphi_2 \in K$, then $\varphi_1 \cdot \varphi_2 \mapsto \varphi_1(A)\,\varphi_2(A)$ – these are easily deduced from the definition – and, as it was proved before, $\varphi_1 \leqslant \varphi_2$ implies $\varphi_1(A) \leqslant \varphi_2(A)$.

We now extend this definition to functions of the form $\psi = f - \varphi \in K - K$ as $\psi(A) = f(A) - \varphi(A)$. It is well defined, for if $\psi = f_1 - \varphi_1 = f_2 - \varphi_2$, then $f_1 + \varphi_2 = \varphi_1 + f_2$, which implies $f_1(A) + \varphi_2(A) = \varphi_1(A) + f_2(A)$. Finally, for complex-valued functions $g = \varphi + i\,\psi$ with $\varphi, \psi \in K - K$, we set $g(A) = \varphi(A) + i\,\psi(A)$.

1.2.2 Spectral integral and spectral theorem (first version)

We have the ingredients for the construction of what we will call the spectral family of A. For $\lambda \in \mathbb{R}$, consider the function

$$e_\lambda(t) = \begin{cases} 1, & \text{for } t \leqslant \lambda, \\ 0, & \text{for } t > \lambda. \end{cases}$$

This function belongs to K and defines $E_\lambda = e_\lambda(A)$.

We deduce from the definition of e_λ that the operators E_λ are orthoprojections, they satisfy $E_\lambda = 0$ for $\lambda < m$, $E_\lambda = I$ for $\lambda \geqslant M$, and $E_\lambda\, E_\mu = E_\lambda$ for $\lambda \leqslant \mu$. The latter is equivalent to saying that $E_\lambda \leqslant E_\mu$ for $\lambda \leqslant \mu$. Moreover, E_λ is continuous from the right with respect to λ in the strong sense: let $P_n(t) \geqslant e_{\lambda+1/n}(t)$ such that $P_n(t) \searrow e_\lambda(t)$. Then $P_n(A) \geqslant E_{\lambda+1/n} \geqslant E_{\lambda+\alpha_n} \geqslant E_\lambda$, so $E_{\lambda+\alpha_n} \to E_\lambda$ as $\alpha_n \to 0$. A family of operators with these properties is called a spectral family. The general definition follows.

A family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ of orthoprojections is a **spectral family** if $E_\lambda \leqslant E_\mu$ for $\lambda \leqslant \mu$ and, in the strong sense, $E_\lambda \to 0$ as $\lambda \to -\infty$, $E_\lambda \to I$ as $\lambda \to \infty$ and $E_{\lambda+0} = E_\lambda$ (continuity from the right).

Let us study how these E_λ are related to A. For $\lambda_1 < \lambda_2$, we introduce the notation $E_{\lambda_1\lambda_2} = E_{\lambda_2} - E_{\lambda_1}$ and $H_{\lambda_1\lambda_2} = \operatorname{Im} E_{\lambda_1\lambda_2}$. Since

$$\lambda_1\,(e_{\lambda_2}(t) - e_{\lambda_1}(t)) \leqslant t\,(e_{\lambda_2}(t) - e_{\lambda_1}(t)) \leqslant \lambda_2\,(e_{\lambda_2}(t) - e_{\lambda_1}(t)),$$

we have that

$$\lambda_1\, E_{\lambda_1\lambda_2} \leqslant A\, E_{\lambda_1\lambda_2} \leqslant \lambda_2\, E_{\lambda_1\lambda_2}. \tag{1}$$

Two particular cases of (1) will be later useful: taking $\lambda_1 < m$ and $\lambda_2 = \lambda$, we get

$$A\, E_\lambda \leqslant \lambda\, E_\lambda, \tag{2}$$

and, for $\lambda_1 = \lambda$ and $\lambda_2 > M$, we obtain

$$\lambda\,(I - E_\lambda) \leqslant A\,(I - E_\lambda). \tag{3}$$

Observe now that (1) restricted to the subspace $H_{\lambda_1\lambda_2}$ yields $\lambda_1\, I_{H_{\lambda_1\lambda_2}} \leqslant A|_{H_{\lambda_1\lambda_2}} \leqslant \lambda_2\, I_{H_{\lambda_1\lambda_2}}$. Hence, for any $\lambda \in [\lambda_1, \lambda_2]$,

$$-(\lambda_2 - \lambda_1)\, I_{H_{\lambda_1\lambda_2}} \leqslant (A - \lambda\, I)|_{H_{\lambda_1\lambda_2}} \leqslant (\lambda_2 - \lambda_1)\, I_{H_{\lambda_1\lambda_2}},$$

which means that

$$\|A - \lambda\, I\|_{H_{\lambda_1\lambda_2}} \leqslant \lambda_2 - \lambda_1. \tag{4}$$

Thus, when λ_2 and λ_1 are close to each other, A appears to be "almost" a constant operator on this subspace. With this in mind we construct the integral representation of A.

Consider a partition of (a, b): $a < \lambda_0 < m \leqslant \lambda_1 < \ldots < \lambda_{n-1} \leqslant M < \lambda_n < b$ with norm $\Delta = \max|\lambda_{i+1} - \lambda_i| < \varepsilon$. Choose any $\mu_i \in [\lambda_i, \lambda_{i+1}]$. We write ΔE_{λ_i} for $E_{\lambda_{i+1}} - E_{\lambda_i}$. From (1) we conclude that

$$(\lambda_i - \mu_i)\, \Delta E_{\lambda_i} \leqslant A\, \Delta E_{\lambda_i} - \mu_i\, \Delta E_{\lambda_i} \leqslant (\lambda_{i+1} - \mu_i)\, \Delta E_{\lambda_i}$$

for each i. Adding this inequality for i from 0 to $n-1$, we get

$$-\varepsilon\, I \leqslant \sum_{i=0}^{n-1} \big(\lambda_i - \mu_i\big)\, \Delta E_{\lambda_i} \leqslant A - \sum_{i=0}^{n-1} \mu_i\, \Delta E_{\lambda_i} \leqslant \sum_{i=0}^{n-1} (\lambda_{i+1} - \mu_i)\, \Delta E_{\lambda_i} \leqslant \varepsilon\, I,$$

which implies

$$\left\| A - \sum_{i=0}^{n-1} \mu_i\, \Delta E_{\lambda_i} \right\| \leqslant \varepsilon$$

for any partition of the interval with norm less than ε and any choice of $\mu_i \in [\lambda_i, \lambda_{i+1}]$. We then define the **spectral integral** as the limit (in the operator norm) of the sum when $\varepsilon \to 0$ and we write

$$A = \int_a^b \lambda\, \mathrm{d}E_\lambda = \int_{m-0}^M \lambda\, \mathrm{d}E_\lambda = \int_{-\infty}^\infty \lambda\, \mathrm{d}E_\lambda.$$

With $m-0$ we mean any lower integration point $a < m$, and by the continuity from the right of E_λ we may take the upper bound to be exactly M. Since $\mathrm{d}E_\lambda$ is identically zero outside $[m, M)$ we may formally write $\int_{-\infty}^{\infty} \lambda \mathrm{d}E_\lambda$.

Theorem 1. (Spectral theorem) *For every self-adjoint bounded operator A on a Hilbert space $\mathcal{H}$ such that $m\ I \leqslant A \leqslant M\ I$, there exists a spectral family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ such that $A = \int \lambda \mathrm{d}E_\lambda$, with the integral converging in the operator norm, E_λ are strong limits of polynomials of A (so they commute with any operator B which commutes with A), and $\|Ax\|^2 = \int \lambda^2 \,\mathrm{d}\langle E_\lambda x, x\rangle$.*

Proof. We only need to prove the last statement since the others follow from the previous comments.

Since $\langle E_\lambda x, x\rangle$ is a monotone function of λ for any x we may understand $\int \lambda^2 \mathrm{d}\langle E_\lambda x, x\rangle$ as a Riemann-Stieltjes integral. Notice that ΔE_{λ_i} are pairwise ortogonal orthoprojections, i.e., for $i \neq j$, $\Delta E_{\lambda_i} \Delta E_{\lambda_j} = 0$. Then,

$$\begin{aligned} \left\|\sum \mu_i \,\Delta E_{\lambda_i} x\right\|^2 &= \left\langle \sum \mu_i \,\Delta E_{\lambda_i} x, \sum \mu_j \,\Delta E_{\lambda_j} x\right\rangle \\ &= \sum \mu_i \,\mu_j \,\langle \Delta E_{\lambda_i} x, \Delta E_{\lambda_j} x\rangle \\ &= \sum \mu_i^2 \,\langle \Delta E_{\lambda_i} x, x\rangle. \end{aligned}$$

As the norm of the partition approaches zero, the first expression tends to $\|Ax\|^2$ and the last expression to the Riemann-Stieltjes integral $\int \lambda^2 \,\mathrm{d}\langle E_\lambda x, x\rangle$. □

A spectral family satisfying the conditions of the theorem above is unique, but we will not prove it here.

We now wish to extend the notion of spectral integral for continuous functions.

Define the integral $\int_{m-0}^{M} \varphi(\lambda)\ \mathrm{d}E_\lambda$ for a continuous function φ on $[a,\ b]$ as the limit of the sums $\sum \varphi(\mu_i)\ \Delta E_{\lambda_i}$ when the norm of the partition tends to zero. This integral converges in the operator norm to the operator $\varphi(A)$. Hence, $\varphi(A) = \int_{m-0}^{M} \varphi(\lambda) \,\mathrm{d}E_\lambda$ and $\|\varphi(A)\,x\|^2 = \int \varphi^2(\lambda) \,\mathrm{d}\langle E_\lambda x, x\rangle$.

Indeed, let $\varphi(t) = t^k$. Then

$$A^k \leftarrow \left(\sum \mu_i \,\Delta E_{\lambda_i}\right)^k = \sum \mu_i^k \,\Delta E_{\lambda_i} \rightarrow \int \lambda^k \,\mathrm{d}E_\lambda,$$

that is, $A^k = \int_{m-0}^{M} \lambda^k \,\mathrm{d}E_\lambda$, so we conclude the statement for polynomials. Let now φ be a real valued continuous function on $[a, b]$ and, for a given $\varepsilon > 0$, find a polynomial $P(t)$ such that $|\varphi(t) - P(t)| < \varepsilon$ for all $t \in [a, b]$. This implies that $-\varepsilon\, I \leqslant \varphi(A) - P(A) \leqslant \varepsilon\, I$, so

$$\|\varphi(A) - P(A)\| \leqslant \varepsilon. \tag{5}$$

Then for any partition $\{\lambda_i\}$ with norm below some $\delta > 0$ (take $\delta < \varepsilon$), and any choice of $\mu_i \in [\lambda_i, \lambda_{i+1}]$, we have that

$$\left\|P(A) - \sum P(\mu_i) \,\Delta E_{\lambda_i}\right\| \leqslant \varepsilon. \tag{6}$$

Let

$$T = \sum P(\mu_i) \,\Delta E_{\lambda_i} - \sum \varphi(\mu_i) \,\Delta E_{\lambda_i}.$$

Since $-\varepsilon \leqslant P(\mu_i) - \varphi(\mu_i) \leqslant \varepsilon$, it holds $-\varepsilon\, I \leqslant T \leqslant \varepsilon\, I$, hence,

$$\|T\| \leqslant \varepsilon. \tag{7}$$

It follows now from (5), (6) and (7) that

$$\left\| \varphi(A) - \sum \varphi(\mu_i) \, \Delta E_{\lambda_i} \right\| \leqslant 3\,\varepsilon,$$

which completes the proof. (For complex valued continuous functions, just note that $(f + i\,g)(A) = f(A) + i\,g(A) = \int f(\lambda)\,\mathrm{d}E_\lambda + i \int g(\lambda)\,\mathrm{d}E_\lambda$).

We can determine the spectrum of A by studying the local behaviour of E_λ. This is stated in the next three results.

Theorem 2. *The point λ_0 is an eigenvalue of A if and only if λ_0 is a point of discontinuity of E_λ in the strong sense.*

Proof. Let λ_0 be a point of discontinuity of E_λ, that is, $\lim_{\lambda \nearrow \lambda_0} E_\lambda \neq E_{\lambda_0}$. The sequence $E_{\lambda\lambda_0}$ converges in the strong sense to an orthoprojection P_0 as $\lambda \nearrow \lambda_0$ – the strong limit exists, since the sequence is monotone and bounded, and the strong limit of orthoprojections is an orthoprojection –. Since $P_0 \neq 0$, there exists x_0 such that $P_0\,x_0 = x_0 \neq 0$. Let us prove that $x_0 \in H_{\mu\lambda_0}$ for any $\mu < \lambda_0$. For $\mu < \lambda < \lambda_0$, it is $E_{\mu\lambda_0}\,E_{\lambda\lambda_0} = E_{\lambda\lambda_0}$, so we have

$$E_{\mu\lambda_0}\,x_0 = E_{\mu\lambda_0} \lim_{\lambda \nearrow \lambda_0} E_{\lambda\lambda_0}\,x_0 = \lim_{\lambda \nearrow \lambda_0} E_{\mu\lambda_0}\,E_{\lambda\lambda_0}\,x_0 = x_0,$$

which implies $x_0 \in H_{\mu\lambda_0}$. From (4) we conclude that $\|(A - \lambda_0 I)x_0\| \leqslant |\mu - \lambda_0| \; \|x_0\|$ and, letting μ approach to λ_0, we get $A\,x_0 = \lambda_0\,x_0$.

Conversely, assume that λ_0 is an eigenvalue of A and x_0 an eigenvector corresponding to λ_0. Let $\varepsilon > 0$. From (2) and (3) we obtain

$$A\,E_{\lambda_0 - \varepsilon} \leqslant (\lambda_0 - \varepsilon)\,E_{\lambda_0 - \varepsilon}$$

and

$$(\lambda_0 + \varepsilon)\,(I - E_{\lambda_0 + \varepsilon}) \leqslant A\,(I - E_{\lambda_0 + \varepsilon}),$$

which, for the eingenvector x_0, yield

$$\lambda_0\,\langle E_{\lambda_0 - \varepsilon}\,x_0, x_0 \rangle \leqslant (\lambda_0 - \varepsilon)\,\langle E_{\lambda_0 - \varepsilon}\,x_0, x_0 \rangle$$

and

$$(\lambda_0 + \varepsilon)\,\langle (I - E_{\lambda_0 + \varepsilon})\,x_0, x_0 \rangle \leqslant \lambda_0\,\langle (I - E_{\lambda_0 + \varepsilon})\,x_0, x_0 \rangle.$$

Since both $E_{\lambda_0 - \varepsilon}$ and $I - E_{\lambda_0 + \varepsilon}$ are positive, we have

$$\varepsilon\,\langle E_{\lambda_0 - \varepsilon}\,x_0, x_0 \rangle = \varepsilon\,\langle (I - E_{\lambda_0 + \varepsilon})x_0, x_0 \rangle = 0.$$

Now use that for any orthoprojection P it is $\langle Px, x \rangle = \langle P^2 x, x \rangle = \langle Px, Px \rangle = \|Px\|^2$ to conclude that $E_{\lambda_0 - \varepsilon}\,x_0 = 0$ and $E_{\lambda_0 + \varepsilon}\,x_0 = x_0 \neq 0$. Thus, λ_0 is a point of discontinuity of E_λ. □

Theorem 3. *Assume that E_λ is continuous in the strong sense at λ_0 and there exist sequences $\lambda_n \nearrow \lambda_0$ and $\mu_n \searrow \lambda_0$ such that $P_n = E_{\mu_n} - E_{\lambda_n}$ are not degenerate, i.e., $P_n \neq 0$ for every n. Then λ_0 is a point of continuous spectrum of A.*

Proof. Set $H_n = \operatorname{Im} P_n$. Since E_λ is strongly continuous at λ_0, $P_n \to 0$ in the strong sense. But $H_n \neq 0$ for any n; so, taking $x_n \in H_n$ with $\|x_n\| = 1$ and using (4), we have that $\|A x_n - \lambda_0\,x_n\| \leqslant \mu_n - \lambda_n \to 0$ as $n \to \infty$. Observe that it must be $\lambda_0 \in \sigma(A)$: if otherwise $A - \lambda_0$ were invertible, it would be $0 < \|A - \lambda_0\,I\| = \|A - \lambda_0\,I\| \; \|x_n\| \leqslant \|A x_n - \lambda_0\,x_n\|$ for every n, which leads to a contradiction. Since λ_0 is not a point of discontinuity of E_λ, $\lambda_0 \notin \sigma_p(A)$. Furthermore, the spectrum of a self-adjoint operator consists only of the point spectrum and the continuous spectrum (the proof is analogous to the one we give later for self-adjoint – posibly unbounded – operators). Hence, $\lambda_0 \in \sigma_c(A)$. □

Theorem 4. *Let $E_\mu - E_\lambda = 0$, with $\mu > \lambda$. Then every $\lambda_0 \in (\lambda, \mu)$ is a regular point of A.*

Proof. From (2), (3) and the fact that $E_\mu = E_\lambda$, we obtain

$$A\,E_\lambda \leqslant \lambda\,E_\lambda,\ \ \mu\,(I - E_\lambda) \leqslant A\,(I - E_\lambda). \tag{8}$$

Set $H_1 = \mathrm{Im}\ E_\lambda$ and $H_2 = \mathrm{Im}\ (I - E_\lambda)$. The subspaces H_1, H_2 are invariant under A and $H_1 \oplus H_2 = \mathcal{H}$. Let $\lambda_0 \in (\lambda, \mu)$. Using (8) we get that the operator A on H_1 satisfies $A|_{H_1} \leqslant \lambda\ I$; therefore, $(\lambda_0 - \lambda)\ I \leqslant (\lambda_0\ I - A)|_{H_1}$. Similarly on H_2, we get $\mu\ I \leqslant A|_{H_2}$; hence, $(\mu - \lambda_0)I \leqslant (A - \lambda_0 I)|_{H_2}$. Since $\lambda_0 - \lambda > 0$ and $\mu - \lambda_0 > 0$, the operators $(A - \lambda_0 I)|_{H_1}$ and $(A - \lambda_0 I)|_{H_2}$ are invertible. This implies that $A - \lambda_0 I$ is invertible. Indeed, for $x_1 \in H_1$, $x_2 \in H_2$, the inverse operator is given by $(A - \lambda_0 I)^{-1}(x_1 + x_2) = ((A - \lambda_0 I)|_{H_1})^{-1} x_1 + ((A - \lambda_0 I)|_{H_2})^{-1} x_2$. □

1.2.3 Multiplication version of the spectral theorem

In this section we give an equivalent formulation of the spectral theorem given above. We first define the concepts that will appear in the statement.

Let f be a complex-valued bounded measurable function on a measure space Λ with measure σ. Let A be defined on the Hilbert space $L^2(\sigma)$ by $(A\varphi)(\lambda) = f(\lambda)\ \varphi(\lambda)$, $\lambda \in \Lambda$. This operator is called the **multiplication** induced by f. It turns out that A^* is the multiplication induced by the complex conjugate $\bar{f}$, so a multiplication is self-adjoint if and only if the function that induces it is real.

An operator $U\colon \mathcal{H} \to \mathcal{H}$ is called an **isometry** if for all $x, y \in \mathcal{H}$ holds $\langle Ux, Uy \rangle = \langle x, y \rangle$, which is equivalent to saying that for all $x \in \mathcal{H}$ is $\langle Ux, Ux \rangle = \langle x, x \rangle$. If also $\mathrm{Im}\, U = \mathcal{H}$, then U is called **unitary**. Notice that in this case U is invertible and $U^* = U^{-1}$. We say that two operators A_1, A_2 are **unitarily equivalent** if there exists U unitary such that $A_1 = U^{-1}\, A_2\, U$. We use this same notion for operators $A_1\colon \mathcal{H}_1 \to \mathcal{H}_1$ and $A_2\colon \mathcal{H}_2 \to \mathcal{H}_2$, with $U\colon \mathcal{H}_1 \to \mathcal{H}_2$ a surjective isometry. We shall deal later with unitary operators and their spectral theory (section 2.2.2).

Theorem 5. (Spectral theorem, multiplication version) *Every self-adjoint bounded operator is unitarily equivalent to a multiplication.*

Concretely, if A is a self-adjoint bounded operator on a Hilbert space $\mathcal{H}$, then there exists a (real valued) bounded measurable function f on some measure space Λ with measure σ, and there exists an isometry U from $L^2(\sigma)$ onto $\mathcal{H}$, such that,

$$(U^{-1}\, A U \varphi)(\lambda) = f(\lambda)\ \varphi(\lambda),$$

for each $\varphi \in L^2(\sigma)$ and $\lambda \in \Lambda$.

In this formulation one can easily recognize the most basic version of the spectral theorem, which says that every self-adjoint matrix is diagonalizable (unitarily equivalent to a diagonal matrix). For this matter, let us consider the particular case where Λ is a finite set with n elements and σ is the counting measure in Λ. Then $L^2(\sigma)$ is the n-dimensional complex Euclidean space and the operator multiplication induced by f is

$$U^{-1}\, A U \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} = (f_1\, \varphi_1, \ldots, f_n\, \varphi_n),$$

whose matricial expression with respect to the canonical basis is the diagonal matrix $\mathrm{diag}(f_1, \ldots, f_n)$.

We say that A has a **simple spectrum** if there exists $x_0 \in \mathcal{H}$, called a **generator** or a **cyclic vector**, such that the closure of the subspace $\langle E_{\lambda_1\lambda_2}\, x_0 : \lambda_2 > \lambda_1\rangle$ is $\mathcal{H}$. This is equivalent to the fact that $\{\varphi(A)\, x_0\}$ is a dense set in $\mathcal{H}$ when $\varphi(t)$ runs over the family of all continuous functions on $[a, b]$. The proof of Theorem 5 requires A to have a generator, but a transfinite argument (see the end of the section) shows that $\mathcal{H}$ is always the direct sum of a family of subspaces, each of which reduces A, and such that the restriction of A to each of them has a generator. So once we prove the theorem for each such restriction, it follows for A itself, since the measure spaces corresponding to the direct summands of $\mathcal{H}$ have a natural direct sum, which corresponds to $\mathcal{H}$. Hence, we may assume that A has a generator, say x_0.

Proof. Let σ be the Lebesgue-Stieltjes measure in $\mathbb{R}$ given by the function $\sigma(\lambda) = \langle E_\lambda\, x_0, x_0\rangle$. Note that $\sigma \equiv 0$ outside $[m, M) \subset [a, b]$, so we may take the measure space to be $[a, b]$. Define the operator $T\colon L^2(\sigma) \to L^2(\sigma)$ by $T\varphi(\lambda) = \lambda\varphi(\lambda)$ (T is the multiplication induced by the identity function). We prove that A is unitarily equivalent to T.

First consider any continuous function $\varphi(\lambda) \in \mathcal{C}[a, b]$ and define the map $U\colon \varphi \mapsto U\varphi = \varphi(A)\, x_0 = \int\, \varphi(\lambda)\, \mathrm{d}E_\lambda\, x_0$. Then,

$$\|U\varphi\|^2 = \int_a^b \varphi^2(\lambda)\, \mathrm{d}\langle E_\lambda\, x_0, x_0\rangle = \int_a^b \varphi^2(\lambda)\, \mathrm{d}\sigma(\lambda) = \|\varphi\|^2_{L^2(\sigma)}.$$

The set of continuous functions is dense in $L^2(\sigma)$. Thus, the linear map U is uniquely extended to an isometry $U\colon L^2(\sigma) \to \mathcal{H}$ (surjective, because the image of an isometry is a complete space and the set $\{U\varphi\}$ is dense in $\mathcal{H}$). We check now that $U^{-1} AU\varphi = T\varphi$. Since

$$A\varphi(A)\, x_0 = \int \lambda\varphi(\lambda)\, \mathrm{d}E_\lambda\, x_0,$$

we have that $(U^{-1} AU\varphi)(\lambda) = U^{-1}(A\varphi(A)\, x_0) = \lambda\varphi(\lambda)$. □

Now we prove the assertion by which we can assume A to have a generator in the proof of Theorem 5:

Let A be a self-adjoint bounded operator on a Hilbert space $\mathcal{H}$. Let H_x be the closure of $\langle E_{\lambda_1\lambda_2}\, x\colon \lambda_2 > \lambda_1\rangle$, for $x \in \mathcal{H}$. Then H_x reduces A, each restriction A_x of A to H_x has a simple spectrum and there exists $L \subset \mathcal{H}$ orthonormal such that $\mathcal{H} = \bigoplus_{x\in L} H_x$.

Notice that, for $\mu_1 < \mu_2$, $\lambda_1 < \lambda_2$, we have $E_{\mu_1\mu_2}\, E_{\lambda_1\lambda_2} = E_{\nu_1\nu_2}$, where $(\nu_1, \nu_2] = (\mu_1, \mu_2] \cap (\lambda_1, \lambda_2]$ and $E_\emptyset = 0$. Then it is clear that $E_{\mu_1\mu_2}\, E_{\lambda_1\lambda_2}\, x \in H_x$; hence, $A E_{\lambda_1\lambda_2}\, x \in H_x$, since $A E_{\lambda_1\lambda_2}\, x$ is the limit of the sums $\sum\, \mu_i\, \Delta E_{\lambda_i}\, E_{\lambda_1\lambda_2}\, x \in H_x$. That is, $A(H_x) \subset H_x$. Furthermore, if $y \in H_x^\perp$, then $H_y \perp H_x$, since $\langle E_{\mu_1\mu_2}\, y, E_{\lambda_1\lambda_2}\, x\rangle = \langle y, E_{\nu_1\nu_2}\, x\rangle = 0$.

Now we prove the existence of such L using Zorn's Lemma. The set $\mathcal{L} = \{K \subset H\colon K$ orthonormal and $H_x \perp H_{x'}, \forall x,\, x' \in K, x \neq x'\}$ is partially ordered with the set inclusion. Since $\{x\} \in L$ for any $x \in H$ with $\|x\| = 1$, the set is not empty. Let $\mathcal{K} \subset \mathcal{L}$ be totally ordered. Then $K_0 = \bigcup \mathcal{K}$ is an upper bound for $\mathcal{K}$ in $\mathcal{L}$, since for x, $x' \in K_0$ there exist K, $K' \in \mathcal{K}$ such that $x \in K$, $x' \in K'$, which we can suppose to be $K \subset K'$, so $\{x, x'\}$ is an orthonormal set with $H_x \perp H_{x'}$ if $x \neq x'$. Furhermore is $K \subset K_0$ for every $K \in \mathcal{K}$. The Zorn Lemma now yields the existence of a maximal element L in $\mathcal{L}$. We conclude that $\mathcal{H} = \bigoplus_{x\in L} H_x$, for if that were not the case, there would exist $y \perp \bigoplus_{x\in L} H_x$ with norm one, we would have $L \subsetneq L \cup \{y\} \in \mathcal{L}$ and L would not be maximal.

2 Unbounded Operators

2.1 Introduction

Operators in Hilbert spaces are frequently used for studying the solutions of linear differential equations and are a fundamental piece in quantum mechanics. But the operators appearing in these contexts are rarely defined on the whole Hilbert space and are almost never bounded. In this section we will extend the theory from section 1 to posibly unbounded operators defined on subspaces of Hilbert spaces.

Even though the operators in this section are generally not bounded, most of them present a weaker but related property: they are closed. We say that A defined on $\mathcal{D}_A = \operatorname{Dom} A \subset \mathcal{H}$ is **closed** if the set $\Gamma(A) = \{(x, Ax) : x \in \mathcal{D}_A\}$ (its **graph**) is closed in $\mathcal{H} \times \mathcal{H}$. A bounded operator is always closed, and the closed graph theorem asserts that every closed operator defined on the whole $\mathcal{H}$ is bounded.

We wish to define for $A: \mathcal{D}_A \to \mathcal{H}$ an adjoint operator A^* such that $\langle Ax, y\rangle = \langle x, A^* y\rangle$ for all $x \in \mathcal{D}_A$ and all y in the largest possible subspace of $\mathcal{H}$. In order for the element $A^* y$ to be uniquely determined by the relation above it is necessary and sufficient that $\mathcal{D}_A$ be dense in $\mathcal{H}$. From now on we shall assume our operators to be densely defined on $\mathcal{H}$. Then the largest possible subset of $\mathcal{H}$ on which A^* can be defined is the subspace

$$\mathcal{D}_{A^*} = \{y \in \mathcal{H} : \langle Ax, y\rangle = \langle x, y^*\rangle \text{ for all } x \in \mathcal{D}_A \text{ and some } y^* \in \mathcal{H}\}$$

and, since $\mathcal{D}_A$ is dense, the element y^* is unique for each $y \in \mathcal{D}_{A^*}$. We may then define $A^* y = y^*$. Some properties for adjoint operators are:

A^* is a closed. if $y_n \to y$, $y_n^* \to y^*$ and $\langle Ax, y_n\rangle = \langle x, y_n^*\rangle$ for every $x \in \mathcal{D}_A$, then $\langle Ax, y\rangle = \langle x, y^*\rangle$ for every $x \in \mathcal{D}_A$; that is, $y \in \mathcal{D}_{A^*}$ and $A^* y = y^*$.

We say that A_2 **extends** A_1 and write $A_1 \subset A_2$ if $\mathcal{D}_{A_1} \subset \mathcal{D}_{A_2}$ and $A_2|_{\mathcal{D}_{A_1}} = A_1$. Then $A_1 \subset A_2$ implies $A_1^* \supset A_2^*$.

Let us consider some examples. Let $\mathcal{H} = L^2([0,1])$ and $A_1 x = A_2 x = A_3 x = i\, x'$ with

$$\begin{aligned}
\mathcal{D}_{A_1} &= \{x \in \mathcal{C}([0,1]) : x \text{ absolutely continuous and } x' \in L^2([0,1])\},\\
\mathcal{D}_{A_2} &= \{x \in \mathcal{D}_{A_1} : x(0) = x(1)\},\\
\mathcal{D}_{A_3} &= \{x \in \mathcal{D}_{A_1} : x(0) = 0 = x(1)\}.
\end{aligned}$$

Here $A_3 \subset A_2 \subset A_1$. Moreover, since polynomials are dense in $L^2([0,1])$ – polynomials are dense in $\mathcal{C}([0,1])$ with the supremum norm, $\mathcal{C}([0,1])$ is dense in $L^2([0,1])$ and the measure of $[0,1]$ is finite –, polynomials taking fixed values at 0 and 1 are also dense. Therefore A_1, A_2 and A_3 are densely defined. Observe also that these are unbounded operators: consider the set $\{x_n = t^n - t\}$, which is contained in $\mathcal{D}_{A_3}$ and hence in the two other domains. This set is bounded in $L^2([0,1])$, but $\|i\, x_n'\|_2 = \|n\, t^{n-1} - 1\|_2 \to \infty$ as $n \to \infty$. Now, for $x, y \in \mathcal{D}_{A_1}$, we have

$$\begin{aligned}
\langle i\, x', y\rangle &= i \int_0^1 x'(t)\, \overline{y}\,(t)\, \mathrm{d}t = i \left[x(1)\, \overline{y}\,(1) - x(0)\, \overline{y}\,(0) - \int_0^1 x(t)\, \overline{y}\,'(t)\, \mathrm{d}t \right]\\
&= i\, [x(1)\, \overline{y}\,(1) - x(0)\, \overline{y}\,(0)] + \langle x, i\, y'\rangle.
\end{aligned} \tag{9}$$

This means that $A_3 \subset A_1^*$, $A_2 \subset A_2^*$ and $A_1 \subset A_3^*$. Let now $y \in \mathcal{D}_{A_3^*}$ and set $z(t) = \int_0^t A_3^*\, y(s)\, \mathrm{d}s$. Then we have

$$\begin{aligned}
\langle A_3\, x, y\rangle &= \langle x, A_3^*\, y\rangle = \int_0^1 x(t)\, \overline{A_3^*\, y}\,(t)\, \mathrm{d}t = \int_0^1 x(t)\, \overline{z}\,'(t)\, \mathrm{d}t\\
&= x(1)\, \overline{z}\,(1) - x(0)\, \overline{z}\,(0) - \int_0^1 x'(t)\, \overline{z}\,(t)\, \mathrm{d}t = \langle A_3\, x, -i\, z\rangle.
\end{aligned}$$

From this identity one concludes that $A_1 \supset A_3^*$ and, using (9) again for $y \in \mathcal{D}_{A_2^*}$ and $y \in \mathcal{D}_{A_1^*}$, we get that $A_2 \supset A_2^*$ and $A_3 \supset A_1^*$. So we have $A_1 = A_3^*$, $A_2 = A_2^*$ and $A_3 = A_1^*$.

Since the operators A_1, A_2 and A_3 are adjoints of some operators in $\mathcal{H}$, we deduce that they are closed. The next result shows that for a closed (densely defined) operator A, the adjoint operator A^* is also densely defined, and provides a way to prove the existence and uniqueness of solutions of certain differential equations as we later show.

Proposition 6. *Let A be a closed operator in $\mathcal{H}$. Then*

(i) $\mathcal{D}_{A^*}$ is dense in $\mathcal{H}$ and $A^{**} = A$.

(ii) Given z and w in $\mathcal{H}$, there exist unique $x \in \mathcal{D}_A$ and $y \in \mathcal{D}_{A^*}$ such that

$$Ax + y = z \quad \text{and} \quad x - A^* y = w.$$

(iii) Given $w \in \mathcal{H}$, there is a unique $x \in \mathcal{D}_{A^*A}$ such that

$$x + A^* A x = w.$$

Proof. Consider the Hilbert space $\mathcal{H} \times \mathcal{H}$ with the scalar product $\langle (x_1, x_2), (y_1, y_2) \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle$. Since A is closed, the set $F = \{(x, Ax) : x \in \mathcal{D}_A\}$ is a closed subspace in $\mathcal{H} \times \mathcal{H}$. Hence, we have $\mathcal{H} \times \mathcal{H} = F \oplus F^\perp$ and it is easy to see that $F^\perp = \{(-A^* y, y) : y \in \mathcal{D}_{A^*}\}$. We now prove (i).

Let $u \perp \mathcal{D}_{A^*}$. Then we have that $(0, u) \in (F^\perp)^\perp = F$, which means $u = A(0) = 0$. Hence, $\mathcal{D}_{A^*}$ is dense in $\mathcal{H}$. For the second part, let $x \in \mathcal{D}_A$. Then we have $\langle A^* y, x \rangle = \langle y, Ax \rangle$ for every $y \in \mathcal{D}_{A^*}$. Hence, $x \in \mathcal{D}_{A^{**}}$ and $A^{**} x = Ax$. Conversely, if $z \in \mathcal{D}_{A^{**}}$, then $(z, A^{**} z) \perp (-A^* y, y)$ for every $y \in \mathcal{D}_{A^*}$; that is, $(z, A^{**} z) \in (F^\perp)^\perp = F$. Thus, $z \in \mathcal{D}_A$ and $Az = A^{**} z$.

For (ii), since $(w, z) \in \mathcal{H} \times \mathcal{H} = F \oplus F^\perp$, there are unique $x \in \mathcal{D}_A$ and $y \in \mathcal{D}_{A^*}$ such that $(w, z) = (x, Ax) + (-A^* y, y)$; that is, $w = x - A^* y$ and $z = Ax + y$.

For (iii), letting $z = 0$ in (ii), we get that $Ax = -y$ and $w = x - A^* y = x + A^* A x$. $\square$

Let us go back to the operators A_1, A_2, A_3 defined above. We can now conclude that given z and w in $L^2([0,1])$, there are unique absolutely continuous functions x and y on $[0,1]$ with x' and y' in $L^2([0,1])$, $y(0) = 0 = y(1)$ such that $i\,x' + y = z$ and $x - i\,y' = w$. Here we considered the operators A_1 and $A_1^* = A_3$. With the other two possibilities we see that the conditions $y(0) = 0 = y(1)$ can be either replaced by the conditions $x(0) = x(1)$, $y(0) = y(1)$, or by the conditions $x(0) = 0 = x(1)$. Part (iii) of the last result yields that given $w \in L^2([0,1])$, there exists a unique absolutely continuous function x on $[0,1]$ with x' absolutely continuous, $x'' \in L^2([0,1])$, $x'(0) = 0 = x'(1)$, such that $-x'' + x = w$. As before, the conditions $x'(0) = 0 = x'(1)$ can be replaced by $x(0) = x(1), x'(0) = x'(1)$ or by $x(0) = 0 = x(1)$.

We say that a densely defined operator A in $\mathcal{H}$ is **symmetric** if, for every $x, y \in \mathcal{D}_A$, holds $\langle Ax, y \rangle = \langle x, Ay \rangle$ or, equivalently, if A^* extends A. If A is symmetric and $\mathcal{D}_A = \mathcal{H}$, then A is in fact bounded. Indeed, consider the image E of the closed unit ball in $\mathcal{H}$ by A. For a fixed $y \in \mathcal{H}$, we have $|\langle Ax, y \rangle| = |\langle x, Ay \rangle| \leqslant \|x\| \, \|Ay\| \leqslant \|Ay\|$ for all $x \in \mathcal{H}$ with $\|x\| \leqslant 1$. Then, by the uniform boundedness principle, the set E is bounded in $\mathcal{H}$ and hence A is bounded. This result is known as the *Hellinger-Toeplitz theorem*.

Observe that any symmetric extension B of a symmetric operator A is between A and A^*; that is, $A \subset B \subset A^*$, since $B \subset B^* \subset A^*$. We say A is **self-adjoint** if $A = A^*$. Then a self-adjoint operator has no symmetric extensions and is obviously closed. Now, if A is symmetric and $\operatorname{Im} A = \mathcal{H}$, then A is self-adjoint. Indeed, since $\operatorname{Im} A = \mathcal{H}$, for every $y \in \mathcal{D}_{A^*}$ there exists $x \in \mathcal{D}_A$ such that $Ax = A^* y$. Let us show that $x = y$, which would mean $A = A^*$. For every $z \in \mathcal{D}_A$, is $\langle Az, y \rangle = \langle z, A^* y \rangle = \langle z, Ax \rangle = \langle Az, x \rangle$. Thus, we have $\langle u, y \rangle = \langle u, x \rangle$ for every $u \in \operatorname{Im} A = \mathcal{H}$ and hence $x = y$.

Next we wish to illustrate the role of unbounded operators in quantum mechanics. For this purpose let us consider the **multiplication operator** $Ax = t\,x$ and the **differentiation operator** $Bx = i\,x'$ in $\mathcal{H} = L^2(\mathbb{R})$ defined on the largest suitable domains $\mathcal{D}_A = \{x \in \mathcal{H}: Ax \in \mathcal{H}\}$, $\mathcal{D}_B = \{x \in \mathcal{H}: Bx \in \mathcal{H}\}$. There is a strong connection between these operators; namely, the the Fourier transform $\mathcal{F}$ on $L^2(\mathbb{R})$.

The Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ are defined on $L^1(\mathbb{R})$ by

$$\begin{aligned}\mathcal{F}x(t) &= \hat{x}(t) = \frac{1}{\sqrt{2\,\pi}}\int_{-\infty}^{\infty} e^{-ist}\,x(s)\,\mathrm{d}s, \quad t \in \mathbb{R}\\ \mathcal{F}^{-1}x(t) &= \check{x}(t) = \frac{1}{\sqrt{2\,\pi}}\int_{-\infty}^{\infty} e^{ist}\,x(s)\,\mathrm{d}s, \quad t \in \mathbb{R}.\end{aligned}$$

For $x \in L^1(\mathbb{R})$, we have $\hat{x}, \check{x} \in L^1(\mathbb{R})$. Moreover, by the Plancherel's theorem, if $x \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then $\hat{x}, \check{x} \in L^2(\mathbb{R})$ and $\|\hat{x}\|_2 = \|\check{x}\|_2 = \|x\|_2$. This equality allows us to define the Fourier transform for $x \in L^2(\mathbb{R})$ as the limit in $L^2(\mathbb{R})$ of $(\hat{x}_n)$, where the sequence $(x_n) \subset L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ converges to x in $L^2(\mathbb{R})$; the inverse Fourier transform may be defined similarly. A detailed exposition can be founded, for instance, in section 4.3.1 of [Eva10].

For each smooth function x with compact support we get, integrating by parts, $\widecheck{i\,x'} = t\,\check{x}$, and by approximation the same is true for every x in $L^2(\mathbb{R})$. Hence, $\mathcal{F}^{-1}\,\mathcal{D}_B = \mathcal{D}_A$ and $\mathcal{F}^{-1}\,Bx = A\mathcal{F}^{-1}\,x$, which means $\mathcal{F}^{-1}\,B\,\mathcal{F} = A$. We now prove that A and B are densely defined unbounded self-adjoint operators in $L^2(\mathbb{R})$. In view of the last equality and the fact that $\mathcal{F}$ is a unitary operator on $L^2(\mathbb{R})$ ($\mathcal{F}$ is bounded and $\mathcal{F}^* = \mathcal{F}^{-1}$), it is enough to prove it for A. Let us denote by χ_E the characteristic function of a subset $E \subset \mathbb{R}$. To show that $\mathcal{D}_A$ is dense, let $x \in L^2(\mathbb{R})$. For each n, the function $x_n = \chi_{[-n,n]}\,x$ is in $\mathcal{D}_A$ and $\|x_n - x\|_2 \to 0$, by Lebesgue's dominated convergence theorem. A is unbounded since the functions $x_n = \chi_{[n,n+1]}$ satisfy $\|x_n\|_2 = 1$, but $\|Ax_n\|_2^2 = \int_n^{n+1} t^2\,\mathrm{d}t = n^2 + n + \frac{1}{3} \to \infty$ as $n \to \infty$. Next, for $x, y \in \mathcal{D}_A$, we have $\langle Ax, y\rangle = \int t\,x(t)\,\overline{y(t)}\,\mathrm{d}t = \int x(t)\,\overline{t\,y(t)}\,\mathrm{d}t = \langle x, Ay\rangle$. Hence, A is symmetric. It remains to prove that A is self-adjoint. Let then $y \in \mathcal{D}_{A^*}$. This means that there is $z \in L^2(\mathbb{R})$ such that $\langle Ax, y\rangle = \langle x, z\rangle$ for all $x \in \mathcal{D}_A$; that is,

$$\int_{-\infty}^{\infty} x(t)\,[t\,\overline{y(t)} - \overline{z(t)}]\,\mathrm{d}t = 0$$

for all $x \in \mathcal{D}_A$. Let $x_n(t) = \chi_{[-n,n]}(t)\,[t\,y(t) - z(t)]$ for each $n \in \mathbb{N}$. Then $x_n \in \mathcal{D}_A$ and

$$\int_{-n}^{n} |x_n|^2 = \int_{-n}^{n} x_n(t)[t\,\overline{y(t)} - \overline{z(t)}]\,\mathrm{d}t = 0,$$

which implies that $t\,y(t) = z(t)$ for almost all $t \in [-n, n]$. Since this is true for each n, we conclude that $t\,y(t) = z(t) \in L^2(\mathbb{R})$, so $y \in \mathcal{D}_A$. Therefore, A is a densely defined unbounded self-adjoint operator in $L^2(\mathbb{R})$ and so is $\mathcal{F}\,A\,\mathcal{F}^{-1} = B$.

Now consider the extremely simplified physical system consisting of one single particle of mass $m = 1$ constrained to move in one dimension ($\mathbb{R}$) at an arbitrary but fixed time. In classical mechanics, the dynamics of the system are completely determined by two real numbers: the position and the momentum of the particle. In quantum mechanics the description of the system is not deterministic as in classical mechanics, but probabilistic.

Let ψ be a complex valued measurable function on $\mathbb{R}$, and let $\int_E |\psi|^2$ represent the probability that the particle lies in a measurable set E. For this to make sense we need to assume that $\int_{\mathbb{R}} |\psi|^2 = 1$, since the particle must be somewere in $\mathbb{R}$. In particular, $\psi \in L^2(\mathbb{R})$. Such a function is called a **pure state** of the system or a **wave function**. Then, if the system is in state ψ, the average position of the particle, or **mean value**, is given by

$$\mu_\psi = \int_{\mathbb{R}} s\,|\psi(s)|^2,$$

and the expectation of the squared deviation of its position from the mean value, that is, the **variance**, is given by

$$\mathrm{var}_\psi = \int_{\mathbb{R}} (s-\mu_\psi)^2\,|\psi(s)|^2.$$

Let $Q\psi(s) = s\,\psi(s)$ with $\mathcal{D}_Q = \{\psi \in L^2(\mathbb{R}) : Q(\psi) \in L^2(\mathbb{R})\}$. Q is the multiplication operator considered earlier and we have $\mu_\psi = \langle Q\psi, \psi\rangle$ and $\mathrm{var}_\psi = \langle (Q-\mu_\psi I)^2\,\psi, \psi\rangle$, for all $\psi \in \mathcal{D}_Q$. In this context the most natural name for Q is that of **position operator**.

The position is said to be an "observable quantity" in quantum mechanics. An **observable** is a densely defined and posibly unbounded self-adjoint operator in $L^2(\mathbb{R})$, as is the case with the position operator. The mean value and the variance of an observable A in state ψ are defined by $\mu_\psi(A) = \langle A\psi, \psi\rangle$ and $\mathrm{var}_\psi(A) = \langle (A-\mu_\psi I)^2\,\psi, \psi\rangle$.

Another important observable quantity is the momentum, whose corresponding **momentum operator** is given by $P\psi = \frac{h}{2\pi i}\,\psi'$ with $\mathcal{D}_P = \{\psi \in L^2(\mathbb{R}) : P\psi \in L^2(\mathbb{R})\}$, where h is Planck's constant. Note that the momentum operator is a scalar multiple of the differentiation operator considered above and hence an observable indeed.

Next, the commutator $PQ - Q\,P$ on $\mathcal{D}_{PQ} \cap \mathcal{D}_{QP}$ is

$$(PQ - Q\,P)\,\psi(s) = \frac{h}{2\pi i}\,[(s\,\psi(s))' - s\,\psi'(s)] = \frac{h}{2\pi i}\,\psi(s),$$

it equals a scalar multiple of the identity operator on the domain $\mathcal{D}_{PQ} \cap \mathcal{D}_{QP}$, which can be proved to be dense in $L^2(\mathbb{R})$ observing that it includes the functions $\psi_n(s) = s^n\,e^{-s^2/2}$, for $n \in \mathbb{N} \cup \{0\}$. We apply here a result for commutators of symmetric operators in $\mathcal{H}$ (for a proof see [Lim96], page 587), which yields, for $\psi \in \mathcal{D}_{PQ} \cap \mathcal{D}_{QP}$,

$$\left|\left\langle \frac{h}{2\pi i}\,\psi, \psi\right\rangle\right| = |\langle (PQ - Q\,P)\,\psi, \psi\rangle| \leqslant 2\,\|P\psi - \langle P\psi, \psi\rangle\,\psi\|\,\|Q\psi - \langle Q\psi, \psi\rangle\,\psi\|.$$

If $\langle \psi, \psi\rangle = 1$, it follows that

$$\frac{h}{2\pi} \leqslant (\mathrm{var}_\psi(P))^{1/2}\,(\mathrm{var}_\psi(Q))^{1/2},$$

which is the famous *Heisenberg uncertainty principle*. This means that the position and the momentum of the particle cannot be simultaneously determined with arbitrarily small error, since the product of the standard deviations is always greater than the positive constant $\frac{h}{2\pi}$.

Note that in the proof of this principle it is essential that the commutator of the position and momentum operators be a scalar multiple of the identity operator. It turns out that it is not possible for bounded operators A, B on $\mathcal{H}$ to satisfy $A\,B - B\,A = I$. If it were so, then we would have $n\,B^{n-1} = A\,B^n - B^n\,A$, for $n \geqslant 1$, and hence

$$n\,\|B^{n-1}\| \leqslant 2\,\|A\|\,\|B^n\| \leqslant 2\,\|A\|\,\|B\|\,\|B^{n-1}\|.$$

If $\|B^{n-1}\| \neq 0$, then $n \leqslant 2\,\|A\|\,\|B\|$, so it must be $\|B^m\| = 0$ for some positive integer m. But then we have $m\,B^{m-1} = 0$, $(m-1)\,B^{m-2} = 0$, ... , and we finally conclude that $B = 0$, which is a contradiction.

2.2 Spectral theory

2.2.1 Spectrum; previous results

We continue with the spectral theory for densely defined unbounded operators in Hilbert spaces. Let $A\colon \mathcal{D}_A \to \mathcal{H}$ with $\overline{\mathcal{D}_A} = \mathcal{H}$. As in the case of bounded operators, $\lambda \in \mathbb{C}$ is a regular point if and only if $(A - \lambda\,I)^{-1}\colon \mathcal{H} \to \mathcal{D}_A$ exists and is bounded. The same definitions for spectrum and its three subsets – point, continuous and residual spectrum – make sense here.

If A is closed, the spectrum is the union $\sigma = \sigma_p \cup \sigma_c \cup \sigma_r$, since $\ker\,(A - \lambda\,I) = 0$ and $\mathrm{Im}(A - \lambda\,I) = \mathcal{H}$ imply that the inverse $(A - \lambda\,I)^{-1}$ is closed, defined on the entire space $\mathcal{H}$ and also, by the closed graph theorem, bounded.

Let us study some properties. We use the notation $\Delta_A(\lambda) = \mathrm{Im}(A - \lambda\,I)$.

$\Delta_A(\lambda)$ is not dense in $\mathcal{H}$ if and only if $\overline{\lambda} \in \sigma_p(A^*)$.

Observe that, for any B, we have $y \in (\mathrm{Im}\ B)^{\perp}$ if and only if for every $x \in \mathcal{D}_B$ holds $0 = \langle Bx, y\rangle = \langle x, B^*\,y\rangle$, which is equivalent to $B^*\,y = 0$, since $\mathcal{D}_B$ is dense in $\mathcal{H}$. That is,

$$(\mathrm{Im}\,B)^{\perp} = \ker B^*. \tag{10}$$

Applying (10) to $A - \lambda\,I$ yields $(\Delta_A(\lambda))^{\perp} = \ker\,(A^* - \overline{\lambda} I)$, and hence the desired result.

Let A be symmetric. Then $\sigma_p(A) \subset \mathbb{R}$.

It follows immediately from $\lambda\ \langle x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\lambda}\ \langle x, x\rangle$. However, this last property may not be true for the whole spectrum. We will give later an example of a densely defined closed symmetric operator whose spectrum is not contained in $\mathbb{R}$.

Let $z = \lambda + i\,\mu$ for $\lambda, \mu \in \mathbb{R}$ with $\mu \neq 0$ and let A be closed and symmetric. Then $\Delta_A(z)$ is a closed subspace.

Let $A_\lambda = A - \lambda\,I$, which is also symmetric for $\lambda \in \mathbb{R}$. For $x \in \mathcal{D}_A$, we have

$$\begin{aligned} \|Ax - z\,x\|^2 &= \|A_\lambda\,x\|^2 - \langle A_\lambda\,x, i\,\mu\,x\rangle - \langle i\,\mu\,x, A_\lambda\,x\rangle + \mu^2\,\langle x, x\rangle \\ &= \|A_\lambda\,x\|^2 + \mu^2\,\|x\|^2, \end{aligned}$$

and then

$$\|Ax - z\,x\| \geqslant |\mu|\,\|x\|. \tag{11}$$

Now assume that $y_n = Ax_n - z\,x_n \to u$. (11) implies that $\|y_n - y_m\| \geqslant |\mu|\,\|x_n - x_m\|$ and hence the sequence (x_n) converges to some x in $\mathcal{H}$. Since A is closed, so is $A - z\,I$; therefore, $x \in \mathcal{D}_A$ and $u = (A - z\,I)\,x \in \Delta_A(z)$.

Notice that the inequality (11) is true not only for closed, but for any symmetric operator. We conclude that for A symmetric the spectrum is also $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. Indeed, we have to show that if $\ker\,(A - \lambda\,I) = 0$ and $\mathrm{Im}(A - \lambda\,I) = \mathcal{H}$, then the inverse operator $(A - \lambda\,I)^{-1}$ is bounded. For $\lambda \notin \mathbb{R}$ it follows from (11). For $\lambda \in \mathbb{R}$, $A - \lambda\,I$ is symmetric and its image is the whole space $\mathcal{H}$; so, as it was proved in section 2.1, $A - \lambda\,I$ is self-adjoint. Hence both $A - \lambda\,I$ and $(A - \lambda\,I)^{-1}$ are closed, and we conclude, by the closed graph theorem, that the inverse is bounded.

If $A = A^*$, then $\sigma(A) \subset \mathbb{R}$ and $\sigma(A) = \sigma_p(A) \cup \sigma_c(A)$.

A is closed and symmetric, so $\Delta_A(z)$ is a closed subspace for $z \in \mathbb{C} \backslash \mathbb{R}$. Assume $\Delta_A(z) \neq \mathcal{H}$ (and hence not dense in $\mathcal{H}$). Then is $\overline{z} \in \sigma_p(A^*) = \sigma_p(A) \subset \mathbb{R}$, which is absurd. Thus, z is a regular point of A. For the second part, let $\lambda \in \mathbb{R}$ and $\overline{\Delta_A(\lambda)} \neq \mathcal{H}$. Then $\overline{\lambda} = \lambda \in \sigma_p(A)$, which implies $\sigma_r(A) = \emptyset$.

Proposition 7. *Let A be a closed symmetric operator. The numbers $n_z = \operatorname{codim} \Delta_A(z)$ are constant on the subsets $\{\operatorname{Im} z > 0\}$ and $\{\operatorname{Im} z < 0\}$.*

Let $n_1 = \operatorname{codim} \Delta_A(z)$ for $\operatorname{Im} z > 0$ and let $n_2 = \operatorname{codim} \Delta_A(z)$ for $\operatorname{Im} z < 0$. The numbers (n_1, n_2) are called the **indices of defect** of A. Observe that the indices of defect of a self-ajoint operator are $(0, 0)$. This concept will be useful to determine if a closed symmetric operator can be extended by another symmetric or self-adjoint operator (see section 2.2.3).

Proof. We will prove it for the subset $\{\operatorname{Im} z > 0\}$; the other case is similar. It is enough to prove that the function $n_z = \operatorname{codim} \Delta_A(z)$ is locally constant on $\{\operatorname{Im} z > 0\}$, i.e., that any z_0 with $\operatorname{Im} z_0 > 0$ has a neighborhood in $\{\operatorname{Im} z > 0\}$ on which $n_z \equiv n_{z_0}$. Indeed, given two points z_1, z_2 in $\{\operatorname{Im} z > 0\}$, consider the segment $[z_1, z_2]$. The neighborhoods corresponding to each point on the segment form an open cover, take a finite subcover and conclude that n_z is constant on $[z_1, z_2]$. So fix z_0 and assume that such neighborhood does not exist. Then there exists a sequence (z_m) in $\{\operatorname{Im} z > 0\}$ such that $\lim_{m \to \infty} z_m = z_0$ but $n_{z_m} \neq n_{z_0}$. There are two possible situations:

(i) either the sequence (z_m) has a subsequence – also denoted by (z_m) – such that $n_{z_m} < n_{z_0}$ for every m,

(ii) or $n_{z_m} > n_{z_0}$ for all m large enough.

Consider case (i). For any m there exists $g_m \neq 0$ such that $g_m \perp \Delta_A(z_0)$ and $g_m \in \Delta_A(z_m)$: indeed, let $N_0 = (\Delta_A(z_0))^{\perp}$ and $N_m = (\Delta_A(z_m))^{\perp}$, and let P_0 be the orthogonal projection on N_0. Since $\dim P_0(N_m) \leqslant \dim N_m = n_m < n_0 = \dim N_0$, we can take $g_m \in N_0$ with $g_m \neq 0$ and $g_m \perp P_0(N_m)$. Notice that $N_m \subset P_0(N_m) \oplus N_0^{\perp}$, which implies that $g_m \in N_m^{\perp}$, as desired. So there exists $f_m \in \mathcal{D}_A$, which can be chosen to be $\|f_m\| = 1$, such that $g_m = (A - z_m I)\, f_m$. It holds $\langle (A - z_m I)\, f_m, (A - z_0 I)\, f_m \rangle = 0$, which can be written as

$$\|(A - z_0 I)\, f_m\|^2 - (z_m - z_0)\, \langle f_m, (A - z_0 I)\, f_m \rangle = 0.$$

This yields

$$\|(A - z_0 I)\, f_m\|^2 \leqslant |z_m - z_0|\, \|(A - z_0 I)\, f_m\|\, \|f_m\|$$

and, since $(A - z_0 I)\, f_m \neq 0$ and $\|f_m\| = 1$, we conclude that $\|(A - z_0 I)\, f_m\| \leqslant |z_m - z_0|$, in contradiction with (11).

The proof in case (ii) is similar: take $g_m \neq 0$ with $g_m \perp \Delta_A(z_m)$ and $g_m \in \Delta_A(z_0)$, and choose f_m such that $g_m = (A - z_0 I)\, f_m$ and $\|f_m\| = 1$. As before, $\langle (A - z_m I)\, f_m, (A - z_0 I)\, f_m \rangle = 0$ leads to a contradiction. □

We now give a few examples which show that some of the results we proved for the spectrum of bounded operators are no longer true for unbounded ones. For the first example go back to the operator A_3 considered earlier. This operator is closed, densely defined and symmetric, but $\sigma(A_3) = \mathbb{C}$ since $\operatorname{Im}(A_3 - \lambda I) \neq \mathcal{H}$ for every $\lambda \in \mathbb{C}$. Every $y \in \operatorname{Im}(A_3 - \lambda I)$ satisfies $i\, x' - \lambda\, x = y$ for some absolutely continuous function x on [0, 1], such that $x' \in L^2([0,1])$ and $x(0) = 0 = x(1)$. Solving the differential equation we get

$$x(t) = c\, e^{-i\lambda t} - i\, e^{-i\lambda t} \int_0^t e^{i\lambda s}\, y(s)\, \mathrm{d}s,$$

for $t \in [0, 1]$ and some constant $c \in \mathbb{C}$. Now the conditions $x(0) = 0 = x(1)$ yield $c = 0$ and $\int_0^1 e^{i\lambda s}\, y(s)\, \mathrm{d}s = 0$. In particular, the function $y(s) = e^{-i\lambda s}$ for $s \in [0, 1]$ is not in $\mathrm{Im}(A_3 - \lambda I)$.

The previous example also shows that the spectrum of an operator in $\mathcal{H}$ may not be bounded. The bad news is that this may also occur for self-adjoint operators. For instance, the operator A_2 considered earlier is self-adjoint, but, for each $n \in \mathbb{Z}$, the operator $A_2 - 2\, n\, \pi\, I$ is not injective, since the function $x_n(t) = e^{-2n\pi i t}$ lies in $\ker(A_2 - 2\, n\, \pi\, I)$.

We finally point out that the spectrum of an operator in $\mathcal{H}$ may be empty. Consider the operator $A_4\, x = i\, x'$ in $\mathcal{H} = L^2([0, 1])$ with domain $\mathcal{D}_{A_4} = \{x \in \mathcal{D}_{A_1} : x(0) = 0\}$, and let $\lambda \in \mathbb{C}$. Then, for every $y \in L^2([0, 1])$, there is a unique $x \in \mathcal{D}_{A_4}$ such that $i\, x' - \lambda\, x = y$; namely, $x(t) = -i\, e^{-i\lambda t} \int_0^t e^{i\lambda s}\, y(s)\, \mathrm{d}s$. Furthermore, x depends continuously on y. Hence, the inverse operator $(A_4 - \lambda\, I)^{-1}$ is defined on $\mathcal{H}$ and is bounded, for all $\lambda \in \mathbb{C}$.

2.2.2 Spectral decomposition

Let $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ be a spectral family of orthoprojections. Let $a < b$ and let $\varphi(t)$ be a continuous function on $[a, b]$. As we did in the previous section, we define for $x \in \mathcal{H}$ the integral $\int_a^b \varphi(\lambda)\, \mathrm{d}E_\lambda\, x$ as the limit when the norm of the partition tends to zero of the sums $\sum \varphi(\mu_i)\, \Delta E_{\lambda_i} x$. Let us now define the improper integral $\int_{-\infty}^{\infty} \varphi(\lambda)\, \mathrm{d}E_\lambda\, x$ as the limit in $\mathcal{H}$ of the integral $\int_{-N}^{M} \varphi(\lambda)\, \mathrm{d}E_\lambda x$ when $M, N \to \infty$, for those x for which the limit exists. Again, we have

$$\left\| \int_a^b \varphi(\lambda)\, \mathrm{d}E_\lambda\, x \right\|^2 = \int_a^b |\varphi(\lambda)|^2\, \mathrm{d}\langle E_\lambda\, x, x\rangle.$$

Hence, we conclude that the improper integral exists if and only if $\int_{-\infty}^{\infty} |\varphi(\lambda)|^2\, \mathrm{d}\langle E_\lambda\, x, x\rangle < \infty$.

Next we give the converse statement of Theorem 1 in the context of unbounded operators.

Theorem 8. *Let $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ be a spectral family of orthoprojections and consider the operator A with*

$$\mathcal{D}_A = \left\{ x : \int_{-\infty}^{\infty} \lambda^2\, \mathrm{d}\langle E_\lambda\, x, x\rangle < \infty \right\} \text{ and } A = \int_{-\infty}^{\infty} \lambda \mathrm{d}E_\lambda,$$

which means that, for $x \in \mathcal{D}_A$, we have $Ax = \int_{-\infty}^{\infty} \lambda \mathrm{d}E_\lambda x$. Then A is a self-adjoint operator and

$$\|Ax\|^2 = \int \lambda^2\, \mathrm{d}\langle E_\lambda\, x, x\rangle.$$

We say that A has a **spectral decomposition** if there exists a spectral family $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ such that A satisfies the conditions of the theorem above with this family.

Proof. We first need to show that $\mathcal{D}_A$ is dense in $\mathcal{H}$ for A^* to be well defined. Indeed, consider the orthoprojection $E_{-N,M} = E_M - E_{-N}$. Then, for any $x \in \mathcal{H}$, we obtain

$$\int_{-\infty}^{\infty} \lambda^2\, \mathrm{d}\langle E_\lambda\, E_{-N,M}\, x, x\rangle = \int_{-N}^{M} \lambda^2\, \mathrm{d}\langle E_\lambda\, x, x\rangle < \infty.$$

So $E_{-N,M}\, x \in \mathcal{D}_A$ and $\mathcal{D}_A$ is dense in $\mathcal{H}$. Let us prove that A is symmetric. For $x \in \mathcal{D}_A$, we have

$$A E_{-N,M}\, x = \int_{-\infty}^{\infty} \lambda \mathrm{d}E_\lambda\, E_{-N,M}\, x = \int_{-N}^{M} \lambda \mathrm{d}E_\lambda\, x = E_{-N,M} \int_{-\infty}^{\infty} \lambda \mathrm{d}E_\lambda\, x = E_{-N,M}\, A x,$$

hence, for $x,y\in\mathcal{D}_A$,

$$\langle E_{-N,M}Ax,y\rangle=\left\langle\int_{-N}^{M}\lambda\mathrm{d}E_\lambda x,y\right\rangle=\left\langle x,\int_{-N}^{M}\lambda\mathrm{d}E_\lambda y\right\rangle=\langle x,E_{-N,M}Ay\rangle.$$

Letting M,N go to infinity, we obtain $\langle Ax,y\rangle=\langle x,Ay\rangle$, for all $x,y\in\mathcal{D}_A$. It remains to prove $A=A^*$. Let $y\in\mathcal{D}_{A^*}$. For any $x\in\mathcal{H}$, since $E_{-N,M}x\in\mathcal{D}_A$, we get

$$\begin{aligned}\langle x,E_{-N,M}A^*y\rangle &= \langle E_{-N,M}x,A^*y\rangle &&= \langle AE_{-N,M}x,y\rangle\\ &= \left\langle\int_{-N}^{M}\lambda\mathrm{d}E_\lambda x,y\right\rangle &&= \left\langle x,\int_{-N}^{M}\lambda\mathrm{d}E_\lambda y\right\rangle.\end{aligned}$$

Hence, $E_{-N,M}A^*y=\int_{-N}^{M}\lambda\mathrm{d}E_\lambda y$. In this last identity, the limit of the left-hand side when M,N tend to infinity exists, so the limit of the right-hand side exists as well, which means that $y\in\mathcal{D}_A$ and $A^*y=\int_{-\infty}^{\infty}\lambda\mathrm{d}E_\lambda y=Ay$. □

Cayley transform and spectral theory for unitary operators

Let A be a closed symmetric operator. For every $x\in\mathcal{D}_A$, define

$$y=(A-iI)x \text{ and } z=(A+iI)x.$$

Then $H_1:=\mathrm{Im}\,(A+iI)$ and $H_2:=\mathrm{Im}\,(A-iI)$ are closed subspaces of $\mathcal{H}$. Since $\pm i$ are not eigenvalues of a symmetric operator, both operators $A\pm iI$ are injective. Hence, each x defines y and z in a unique way and we may define the operator $V\colon H_1\to H_2$ by $z=Vy$. We easily check that V is an isometry:

$$\begin{aligned}\langle Vy,Vy\rangle &= \langle (A-iI)x,(A-iI)x\rangle &&= \|Ax\|^2+\|x\|^2\\ &= \langle (A+iI)x,(A+iI)x\rangle &&= \langle y,y\rangle.\end{aligned}$$

We call V the **Cayley transform** of A. From the definition of y and z one deduces that

$$x=\frac{1}{2i}(I-V)y \quad\text{and}\quad Ax=\frac{1}{2}(I+V)y.$$

Notice that 1 is not an eigenvalue of V since $z=y$ implies $x=0$, so $z=y=0$. Thus, $(I-V)^{-1}$ is formally defined and we have $Ax=i\,(I+V)\,(I-V)^{-1}x$. If A is also self-adjoint, then $\pm i\notin\sigma(A)$ and hence $H_1=H_2=\mathcal{H}$ (the indices of defect are $(0,0)$), so V is a unitary operator. The converse is also true:

If A is closed and symmetric, and the Cayley transform V is unitary (i.e., $H_1=H_2=\mathcal{H}$ or equivalently the indices of defect of A are $(0,0)$), then A is self-adjoint.

We will now be dealing with unitary operators and their spectral properties, which we will use in our proof of the previous statement and will also provide a spectral decomposition for the operator A which is claimed to be self-adjoint.

Our aim is to define a spectral integral for unitary operators. We proceed in a similar way as we did in the case of self-adjoint bounded operators.

Consider a Laurent polynomial $P(z)$ of z and z^{-1} for $|z|=1$; i.e., $z=e^{it}$ with $t\in\mathbb{R}$ and

$$P(z)=\sum_{k=-m}^{n}a_k e^{ikt}=\sum_{k=-m}^{n}a_k z^k.$$

Now define the operator

$$P(z)|_{z:=U}=P(U)=\sum_{k=-m}^{n}a_k U^k.$$

This correspondence between Laurent polynomials and operators has the property of linearity: $(P_1 + P_2)(U) = P_1(U) + P_2(U)$; the property of multiplicativity: $(P_1 \cdot P_2)(U) = P_1(U)\,P_2(U)$; the adjoint satisfies $P(U)^* = \sum \overline{a_k} U^{-k} = \overline{P}\,(U^{-1}) = \overline{P(z)}|_{z:=U}$; and if $P(z) \geqslant 0$ for $|z| = 1$, then $P(U) \geqslant 0$. The last one is not that obvious and can be proved using the fact that if $P(z) \geqslant 0$ for $|z| = 1$, then there exists another polynomial $Q(z)$ such that $P(z) = Q(z) \cdot \overline{Q(z)} = |Q(z)|^2$ (for a proof see [EMT04], Lemma 8.1.1). From this follows that $P(U) = Q(U)\,Q(U)^*$, which is always a positive operator.

We define now functions of a unitary operator as it was done for self-adjoint bounded operators, starting with the functions $\varphi(e^{it})$, $0 \leqslant t < 2\,\pi$, of the class K. Let $\varphi(z) = \varphi(e^{it}) \geqslant 0$ and let $P_n(z) \searrow \varphi(e^{it})$ for every t. Define $\varphi(U) \geqslant 0$ as the strong limit of the bounded sequence $P_n(U) \geqslant P_{n+1}(U) \geqslant \ldots \geqslant 0$. Repeating the same arguments we used for self-adjoint operators, we conclude that the operator $\varphi(U)$ is well defined and we can extend the definition for functions of the form $c_1\,\varphi_1 + c_2\,\varphi_2$ for any, real at first and in general complex, c_1 and c_2. The properties of linearity, multiplicativity and order-preserving for real-valued functions still hold. We also notice that $\varphi(U)^* = \overline{\varphi(z)}|_{z:=U}$, which follows from the fact that it is true for the approximating polynomials.

The next step is to build the spectral family of orthoprojections. Consider the functions

$$\psi_\lambda(e^{it}) = \begin{cases} 1, & \text{for}\, 0 < t \leqslant \lambda, \\ 0, & \text{for}\, \lambda < t \leqslant 2\,\pi, \end{cases} \qquad \text{for}\, \lambda \in (0, 2\,\pi),$$

$\psi_0(e^{it}) \equiv 0$ and $\psi_{2\pi}(e^{it}) \equiv 1$. Then $\psi_\lambda(U) = F_\lambda$ are orthoprojections. To prove the continuity from the right of F_λ, consider the function

$$\tilde{\psi}_0(e^{it}) = \begin{cases} 1, & \text{for}\, t = 0, \\ 0, & \text{for}\, 0 < t < 2\,\pi. \end{cases}$$

Then $\tilde{\psi}_0(U) = \tilde{F}_0$ is also an orthoprojection and the funcions $\psi_\lambda(e^{it}) + \tilde{\psi}_0(e^{it})$ are continuous from above, and hence in K. One can proceed as we did in the previous section to prove the continuity from the right of the family $\{F_\lambda + \tilde{F}_0\}$ and then return to $\{F_\lambda\}$.

For the construction of the spectral integral, consider a partition $\{\lambda_j\}$ of $[0, 2\,\pi]$ with $\mu_j \in [\lambda_j, \lambda_{j+1}]$, and define the function

$$e^{it} - \sum_{j=0}^{n-1} e^{i\mu_j}\,(\psi_{\lambda_{j+1}}(e^{it}) - \psi_{\lambda_j}(e^{it})) =: \chi(e^{it}).$$

For every t we find k such that $\lambda_k \leqslant t \leqslant \lambda_{k+1}$. Then, if $\max_j\, |\lambda_{j+1} - \lambda_j| < \varepsilon$, we have $|\chi(e^{it})| \leqslant |e^{it} - e^{i\mu_k}| \leqslant |t - \mu_k| \leqslant \varepsilon$. Therefore, $\overline{\chi(e^{it})} \cdot \chi(e^{it}) \leqslant \varepsilon^2$, which implies $\chi(U)^*\,\chi(U) \leqslant \varepsilon^2\,I$. It follows that

$$\|\chi(U)\| = \left\| U - \sum_{j=0}^{n-1} e^{i\mu_j}\,(F_{\lambda_{j+1}} - F_{\lambda_j}) \right\| \leqslant \varepsilon. \tag{12}$$

Thus, the sum approximates U with respect to the operator norm and we may write

$$U = \int_0^{2\pi} e^{i\lambda}\,\mathrm{d}F_\lambda.$$

We also have $\|\chi(U)^*\| = \left\|U^{-1} - \sum e^{-i\mu_j}\,(F_{\lambda_{j+1}} - F_{\lambda_j})\right\| \leqslant \varepsilon$, so $U^* = U^{-1} = \int_0^{2\pi} e^{-i\lambda}\,\mathrm{d}F_\lambda$. Moreover, $U^n = \int_0^{2\pi} e^{in\lambda}\,\mathrm{d}F_\lambda$, for $n \in \mathbb{Z}$. Finally, exactly as we did previously for self-adjoint operators, one proves that, for any continuous function $\varphi(e^{it})$,

$$\varphi(U) = \int_0^{2\pi} \varphi(e^{i\lambda})\,\mathrm{d}F_\lambda.$$

Lemma 9. *Let U be a unitary operator and let $\{F_\lambda\}$ be the spectral family defined above. The point $e^{i\lambda_0}$, $\lambda_0 \in [0, 2\pi]$, is an eigenvalue of U if and only if λ_0 is a point of discontinuity of F_λ in the strong sense.*

Proof. Let λ_0 be a point of discontinuity of F_λ. Consider, for $\lambda \leqslant \lambda_0$, the orthoprojections $F_{\lambda\lambda_0} = F_{\lambda_0} - F_\lambda$. In the same way as in the proof of Theorem 2, we conclude that the stong limit of $F_{\lambda\lambda_0}$ when $\lambda \nearrow \lambda_0$ is an orthoprojection $P_0 \neq 0$, and hence there exists x_0 such that $P_0\, x_0 = x_0 \neq 0$ and $x_0 \in \operatorname{Im} F_{\lambda\lambda_0}$, for every $\lambda < \lambda_0$. Using the fact that

$$|(e^{it} - e^{i\lambda_0})\,(\psi_{\lambda_0}(e^{it}) - \psi_\lambda(e^{it}))| \leqslant \max_{\lambda \leqslant t \leqslant \lambda_0} |e^{it} - e^{i\lambda_0}| \leqslant |\lambda - \lambda_0|,$$

we get

$$\|(U - e^{i\lambda_0} I)\, x_0\| \leqslant \|(U - e^{i\lambda_0} I)\, F_{\lambda\lambda_0}\|\, \|x_0\| \leqslant |\lambda - \lambda_0|\, \|x_0\|.$$

Thus, $U x_0 = e^{i\lambda_0} x_0$.

Now assume that $e^{i\lambda_0}$ is an eigenvalue of U. Then

$$0 = \|(U - e^{i\lambda_0} I)\, x_0\|^2 = \int_0^{2\pi} |e^{i\lambda} - e^{i\lambda_0}|^2\, \mathrm{d}\langle F_\lambda\, x_0, x_0\rangle.$$

For any $\varepsilon > 0$, $|e^{i\lambda} - e^{i\lambda_0}|^2$ is greater than some strictly positive constant on $\{\lambda \in [0, 2\pi]: |\lambda - \lambda_0| \geqslant \varepsilon\}$. Hence, if $0 < \lambda_0 < 2\pi$, we have

$$0 = \int_{\lambda_0 + \varepsilon}^{2\pi} \mathrm{d}\langle F_\lambda\, x_0, x_0\rangle = \langle x_0, x_0\rangle - \langle F_{\lambda_0+\varepsilon}\, x_0, x_0\rangle$$

and

$$0 = \int_0^{\lambda_0 - \varepsilon} \mathrm{d}\langle F_\lambda\, x_0, x_0\rangle = \langle F_{\lambda_0 - \varepsilon}\, x_0, x_0\rangle,$$

which imply that $F_{\lambda_0+\varepsilon}\, x_0 = x_0$ and $F_{\lambda_0-\varepsilon}\, x_0 = 0$, for any $\varepsilon > 0$. Therefore, λ_0 is a point of discontinuity of F_λ. If $\lambda_0 = 2\pi$, then $F_{\lambda_0} x_0 = x_0$ by definition and $F_{\lambda_0 - \varepsilon}\, x_0 = 0$ as before. □

Lemma 9 implies that $1 \notin \sigma_p(U)$ if and only if F_λ is continuous at $\lambda_0 = 2\pi$, which means that $F_\lambda \to I$ strongly when $\lambda \nearrow 2\pi$.

Also the analogous versions to theorems 3 and 4 hold for unitary operators, but we will not use them here. (See exercises 6 and 7 of chapter 8 from [EMT04]).

We are now ready to return to the Cayley transform for closed symmetric (probably unbounded) operators and prove the next theorem.

Theorem 10. *If A is closed and symmetric, and the Cayley transform V is unitary (i.e., $H_1 = H_2 = \mathcal{H}$ or equivalently the indices of defect are $(0,0)$), then A is self-adjoint with spectral decomposition $E_t = F_s$ for $t = -\cot\,(s/2)$, where $\{F_s\}_{s \in [0, 2\pi]}$ is the spectral decomposition of V.*

Proof. Clearly, $E_t = F_s$ are orthoprojections, $E_\tau \leqslant E_t$ for $\tau \leqslant t$, $E_{t+0} = E_t$ and $E_t \to 0$ as $t \to -\infty$. Furthermore, since $1 \notin \sigma_p(V)$, we conclude from Lemma 9 that $E_t \to I$ as $t \to \infty$. Hence, $\{E_t\}$ is a spectral family of orthoprojections. By Theorem 8, the relations $\mathcal{D}_B = \left\{x: \int_{-\infty}^{\infty} t^2\, \mathrm{d}\langle E_t\, x, x\rangle < \infty\right\}$ and $Bx = \int_{-\infty}^{\infty} t\, \mathrm{d}E_t\, x$ define a self-adjoint operator B. We now need to show that $A = B$.

Let $x \in \mathcal{D}_A$. We have that, for some $y \in \mathcal{H}$,

$$A x = \frac{1}{2}\,(I + V)\, y = \frac{1}{2} \int_0^{2\pi} (1 + e^{is})\, \mathrm{d}F_s\, y. \tag{13}$$

Similarly,

$$x = \frac{1}{2\,i}\,(I - V)\,y = \frac{1}{2\,i}\int_0^{2\pi}(1 - e^{i\tau})\,\mathrm{d}F_\tau\,y$$

and therefore

$$F_s\,x = \frac{1}{2\,i}\,(I - V)\,F_s\,y = \frac{1}{2\,i}\int_0^{s}(1 - e^{i\tau})\,\mathrm{d}F_\tau\,y. \tag{14}$$

From (13) and (14) we get

$$A\,x = \int_0^{2\pi} i\,\frac{1 + e^{is}}{1 - e^{is}}\,\mathrm{d}F_s\,x = -\int_0^{2\pi}\cot\left(\frac{s}{2}\right)\mathrm{d}F_s\,x = \int_{-\infty}^{\infty} t\,\mathrm{d}E_t\,x.$$

In particular, this shows that the improper integral $\int_{-\infty}^{\infty} t\,\mathrm{d}E_t\,x$ converges, which occurs if and only if $\int_{-\infty}^{\infty} t^2\,\mathrm{d}\langle E_t\,x, x\rangle < \infty$. Hence, for $x \in \mathcal{D}_A$, we have $x \in \mathcal{D}_B$ and $A\,x = B\,x$.

It remains to prove that $\mathcal{D}_B \subset \mathcal{D}_A$. That is, for x such that $\int_{-\infty}^{\infty} t^2\,\mathrm{d}\langle E_t\,x, x\rangle < \infty$, we need to find $y \in \mathcal{H}$ with $(I - V)\,y = 2\,i\,x$.

Consider the improper integral $\int_0^{2\pi} -e^{-is/2}\,(\sin(s/2))^{-1}\,\mathrm{d}F_s\,x$. We claim it converges in $\mathcal{H}$. Indeed,

$$\left\|\int_0^{2\pi}\frac{-e^{-is/2}}{\sin(s/2)}\,\mathrm{d}F_s\,x\right\|^2 = \int_0^{2\pi}\frac{|e^{-is/2}|^2}{\sin(s/2)^2}\,\mathrm{d}\langle F_s\,x, x\rangle \leqslant \int_0^{2\pi}\frac{1}{\sin(s/2)^2}\,\mathrm{d}\langle F_s\,x, x\rangle < \infty.$$

That the last integral is finite follows from our assumption of the integral $\int_{-\infty}^{\infty} t^2\,\mathrm{d}\langle E_t\,x, x\rangle = \int_0^{2\pi}\cot^2(s/2)\,\mathrm{d}\langle F_s\,x, x\rangle$ being finite and the fact that $\int_0^{2\pi}\mathrm{d}\langle F_s\,x, x\rangle < \infty$.

Let then $y = \int_0^{2\pi} -e^{-is/2}\,(\sin(s/2))^{-1}\,\mathrm{d}F_s\,x$. We have

$$F_s\,y = -\int_0^{s}\frac{e^{-i\tau/2}}{\sin(\tau/2)}\,\mathrm{d}F_\tau\,x.$$

Therefore,

$$(I - V)\,y = \int_0^{2\pi}(1 - e^{is})\,\mathrm{d}F_s\,y = -\int_0^{2\pi}\frac{(1 - e^{is})\,e^{-is/2}}{\sin(s/2)}\,\mathrm{d}F_s\,x = 2\,i\int_0^{2\pi}\mathrm{d}F_s\,x = 2\,i\,x$$

and $x \in \mathcal{D}_A$. □

2.2.3 Symmetric and self-adjoint extensions of a symmetric operator

In this last section we answer the questions: *when can a closed symmetric operator be extended preserving the symmetry? And, if there exists a symmetric extension, could it be self-adjoint?*

Assume A_1 is a symmetric extension of A – in the strict sense, with $A_1 \neq A$ –. Then there exists $x_1 \in \mathcal{D}_{A_1} \backslash \mathcal{D}_A$, which gives $(A_1 + i\,I)\,x_1 = y_1$ and $(A_1 - i\,I)\,x_1 = z_1 = V_1\,y_1$, where $y_1 \notin H_1$ and $z_1 \notin H_2$. Hence, the Cayley transform V_1 of A_1 is an isometric extension of V and does not coincide with V. We can now assert the following.

If the indices of defect are $(n, 0)$ or $(0, n)$ and $n \neq 0$, then A does not have any symmetric extension, i.e., A is maximal symmetric and not self-adjoint.

Indeed, either $H_1 = \mathcal{H}$ or $H_2 = \mathcal{H}$, so the Cayley transform does not have any extensions. Furthermore, the indices are not $(0, 0)$, which means that A is not self-adjoint.

Consider now the case when both indices are not equal to zero. Let $\mathcal{D}_{V_1} = H_1' = H_1 \oplus L_1$ and $\operatorname{Im} V_1 = H_2' = H_2 \oplus L_2$. V_1 is an isometry, its restriction to H_1 is V and its restriction to L_1 is an isometry between L_1 and L_2. Thus, $\dim L_1 = \dim L_2$ and we arrive at:

If the indices of A are (m, n) and $m \neq n$, then there is no self-adjoint extension of A.

Indeed, the Cayley transform of a self-adjoint operator satisfies $H_1 \oplus L_1 = \mathcal{H}$ and $H_2 \oplus L_2 = \mathcal{H}$. Hence, it must be $m = \operatorname{codim} H_1 = \operatorname{codim} H_2 = n$.

But we can build a symmetric extension of A from any isometric extension V_1 of V. Consider the operator A_1 with

$$\mathcal{D}_{A_1} = \left\{ x : x = \frac{1}{2\,i}\,(I - V_1)\,y, y \in \mathcal{D}_{V_1} \right\},$$

$$A_1\,x = \frac{1}{2}\,(I + V_1)\,y, \qquad \text{for } x = \tfrac{1}{2\,i}\,(I - V_1)\,y \in \mathcal{D}_{A_1}.$$

In order to show that each $x \in \mathcal{D}_{A_1}$ is uniquely determined (for A_1 to be well defined), we need to show that $(I - V_1)$ is injective, i.e., that $1 \notin \sigma_p(V_1)$. Let $y_0 \in \mathcal{D}_{V_1}$ such that $y_0 = V_1\,y_0$. For each $x \in \mathcal{D}_{A_1}$ there is a y such that $x = \frac{1}{2\,i}\,(I - V_1)\,y$. Then

$$\langle y_0, x \rangle = \left\langle y_0, \frac{1}{2\,i}\,(I - V_1)\,y \right\rangle = \frac{-1}{2\,i}(\langle y_0, y\rangle - \langle V_1\,y_0, V_1\,y\rangle) = 0.$$

Hence, $y_0 \perp \mathcal{D}_{A_1}$, which implies that $y_0 = 0$, since $\mathcal{D}_{A_1} \supset \mathcal{D}_A$ and $\mathcal{D}_A$ is dense in $\mathcal{H}$. That is, $1 \notin \sigma_p(V_1)$. It remains to show that A_1 is symmetric. For any $x_1, x_2 \in \mathcal{D}_{A_1}$, there exist y_1, $y_2 \in \mathcal{D}_{V_1}$ such that

$$\begin{aligned}
\langle A_1\,x_1, x_2\rangle &= -\frac{1}{4\,i}\,\langle (I + V_1)\,y_1, (I - V_1)\,y_2\rangle \\
&= -\frac{1}{4\,i}\,(\langle y_1, y_2\rangle + \langle V_1\,y_1, y_2\rangle - \langle y_1, V_1\,y_2\rangle - \langle V_1\,y_1, V_1\,y_2\rangle) \\
&= -\frac{1}{4\,i}\,(\langle V_1\,y_1, y_2\rangle - \langle y_1, V_1\,y_2\rangle).
\end{aligned}$$

For $\langle x_1, A_1\,x_2\rangle$ we obtain the same expresion. We can also conclude:

If the indices of A are (n, n), $n \neq 0$, then there exists a self-adjoint extension A_1 of A.

Indeed, we can decompose $\mathcal{H}$ as $\mathcal{H} = H_1 \oplus L_1$ and $\mathcal{H} = H_2 \oplus L_2$, with $\dim L_1 = \dim L_2 = n$. Let T be an isometry between the n-dimensional Hilbert spaces L_1 and L_2. Setting $V_1(u, v) = (Vu, Tv)$, for $u \in H_1$, $v \in L_1$, we obtain an extension of V to a unitary operator V_1: $\mathcal{H} \to \mathcal{H}$. The corresponding symmetric extension A_1 of A has indices $(0, 0)$ and is therefore self-adjoint.

Bibliography

[Cas11] Domenico P.L. Castrigiano. Spektraltheorie linearer operatoren. 2011.

[EMT04] Yuli Eidelman, Vitali Milman and Antonis Tsolomitis. *Functional Analysis. An Introduction*, volume 66 of *Graduate Studies in Mathematics*. American Mathematical Society, 2004.

[Eva10] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, second edition, 2010.

[Hal63] P. R. Halmos. What does the spectral theorem say? *The American Mathematical Monthly*, 70(3):241–247, March 1963.

[Lim96] Balmohan V. Limaye. *Functional Analysis*. New Age International, second edition, 1996.

[Rud86] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, third edition, 1986.