Lineare Evolutionsgleichungen

Delio Mugnolo

(delio.mugnolo@uni-ulm.de)

Ulm, July 12, 2011.

Contents

Chapter 1. Introduction 5

Chapter 2. The Spectral Theorem 15
2.1. Adjoint operators 15
2.2. Spectral theory for unbounded operators 21
2.3. The Spectral Theorem for bounded self-adjoint operators 24
2.4. The Spectral Theorem for bounded normal operators 28
2.5. The Spectral Theorem for unbounded operators and the functional calculus 31
2.6. An application: the Dirac equation 37

Chapter 3. General semigroup theory 41
3.1. Generation results 41
3.2. Long-time results for C_0-semigroups 48

Bibliography 57

CHAPTER 1

Introduction

Consider the partial differential equation

$$\frac{\partial u}{\partial t}(t,x) = \Delta u(t,x) := \operatorname{div}\operatorname{grad} u(t,x) = \sum_{k=1}^{d} \frac{\partial^2 u}{\partial x^2}(t,x)$$

for $t \geq 0$ and $x \in \mathbb{R}^d$. It models the diffusion of heat or material in a medium – the so-called *heat equation*. This equation can be explicitly solved: in fact, it is well-known that for all "reasonable" initial value $u_0 := u(0,\cdot)$ (e.g., $u_0 \in L^1(\mathbb{R}^d)$ or else $u_0 \in C_0(\mathbb{R}^d) := \{v : \mathbb{R}^d \to \mathbb{C} : \lim_{|x|\to\infty} v(x) = 0\}$ will do)

$$u(t,x;u_0) := \frac{1}{(4\pi t)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-\frac{\|x-y\|^2}{4t}} u_0(y)dy, \qquad t > 0,\ x \in \mathbb{R}^d,$$

yields a *solution*, i.e., a function that is once continuously diffeentiable with respect to the variable t and twice continuously differentiable (and bounded) with respect to the variable x and that satisfies the heat equation pointwise for all $t > 0$ (and has a limit for $t \to 0+$ that coincides with the given initial value). Moreover, this solution is unique, as can be easily seen (cf. [**4**, Thm. 2.3.10]). Finally, a direct computation shows that u depends continuously on the initial data, i.e., for all $t > 0$ and $x \in \mathbb{R}^d$

$$u(t,x;\cdot) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$$

is continuous along with

$$u(t,x;\cdot) : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d).$$

In fact, the heat equation can be solved for nearly all initial data one may think of.

This is a radically smoothing property of a very special class of partial differential equations. It is more natural to expect that in order for a partial differential equation to be solvable – hence, in particular, to admits a solution that is suitably diffentiable – it is necessary that the initial data are just as smooth.

Example 1.1. *The prototypical case is that of the (1-dimensional)* transport equation

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial u}{\partial x}(t,x)$$

for $t > 0$ and $x \in (0,1)$ with boundary condition

$$u(t,0) = 0$$

for $t > 0$. Its solution is necessarily given by

$$u(t,x;u_0) = \begin{cases} u_0(t+x), & t+x \leq 1, \\ 0, & t+x > 1. \end{cases}$$

Hence, one sees that the transport equation is satisfied pointwise by u if and only if u_0 is differentiable.

However, even requiring that the initial value is as smooth as the solution should be is possibly not sufficient: this happens even if one thinks of equations that look just as natural. Take for example the *wave equation*

$$\frac{\partial^2 u}{\partial t^2}(t,x) = \Delta u(t,x)$$

or the *Schrödinger equation*

$$i\frac{\partial u}{\partial t}(t,x) = \Delta u(t,x).$$

Let $p \in (1,\infty)$. Then, a classical result obtained by Littman in 1963 (resp., by Hörmander in 1960), generalised by Hieber in 1991, show that one has a solution (by this we mean an L^p-function that is twice differentiable with respect to x, at least in a weak sense) only if $u_0 \in L^p(\mathbb{R}^d)$ is at least $2\big((n-1)|2^{-1}-p^{-1}|+1\big)$-times (resp., at least $2\big(n|2^{-1}-p^{-1}|+1\big)$-times) differentiable.

This shows that the choice of a right function space is fundamental to ensure existence and uniqueness of solutions to a given problem. It is therefore useful to proceed to a reformulation of a general initial value problem. Let us consider again the case of the heat equation. We are looking for solutions, i.e., for functions $u : \mathbb{R}_+ \times \mathbb{R}^d \ni (t,x) \mapsto \mathbb{C}$ that satisfy the above equation pointwise, including the initial and (possibly) the boundary conditions. However, this can be rephrased in a different way if we introduce a vector-valued function $U(t) := u(t,\cdot)$. In other words, we are looking for a function Banach space, say X, and a differentiable function $U : \mathbb{R}_+ \to X$ such that $U(t)$ satisfies the differential equation pointwise: this means that U has to be continuously differentiable with respect to t (its only variable), and smooth enough that it can be differentiated with respect to x. The decisive point now is to look at the differential expression Δ as an operator A that maps functions from $\mathbb{R}^d$ to $\mathbb{C}$ into functions from $\mathbb{R}^d$ to $\mathbb{C}$. We want to incorporate the boundary conditions in the domain $D(A)$ of A: hence, we are requiring $U(t)$ to be in $D(A)$ for all t. In other words, we have re-written the heat equation as an *abstract Cauchy problem*, i.e., an *ordinary* differential equation with values in a (possibly infinite dimensional) vector space with an initial condition of the form

$$\text{(ACP)} \qquad \begin{cases} U'(t) &= AU(t), \quad t \geq 0, \\ U(0) &= U_0 \in X. \end{cases}$$

Definition 1.2. *A* solution to (ACP) *is a function* $U \in C^1(\mathbb{R}_+;X)\cap C(\mathbb{R};D(A))$ *that satisfies* (ACP) *pointwise.*

This motivates the following definition, which has been introduced by Jacques Hadamard in 1898.

Definition 1.3. *Let X be a Banach space and A a linear operator on X. We call (ACP)* well-posed *if each of the following conditions hold.*

- *For all $U_0 \in X$ there exists a solution to (ACP).*
- *For all $U_0 \in X$ there is at most a solution to (ACP).*
- *Let $(U_{0n})_{n\in\mathbb{N}}$ be a sequence in X with associated solutions $(U_n)_{n\in\mathbb{N}}$ of the corresponding (ACP). Let additionally $U_0 \in X$ and U be the associated solution. If $(U_{0n})_{n\in\mathbb{N}}$ converges to U_0 in X, then $(U_n(t))_{n\in\mathbb{N}}$ converges to $U(t)$ in X for all $t \geq 0$.*

As a rule of thumb, all “natural” evolution equations are well-posed. However, this does not necessarily mean that they can be written as abstract Cauchy problems in an obvious way.

Example 1.4. *Consider the wave equation*

$$\frac{\partial^2 u}{\partial t^2}(t,x) = \Delta u(t,x)$$

for $t \geq 0$ and $x \in \mathbb{R}^d$. By an easy trick which is already common in the theory of ordinary differential equations, this equations can be re-written as

$$U'(t) = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix} U(t).$$

We will see that this first order reduction leads to a well-posed abstract Cauchy problem only if the space of admissible initial conditions is suitably small.

Example 1.5. *Sometimes, boundary conditions that are given by further differential equations are considered: one usually terms them* dynamic boundary conditions*. In most cases, the right idea is to transform them in an* *abstract Cauchy problem in a product space (that may be particularly involved). For example, consider the heat equation*

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) &= \Delta u(t,x), & t \geq 0,\ x \in \Omega, \\ \frac{\partial w}{\partial t}(t,z) &= -\frac{\partial w}{\partial n}(t,z), & t \geq 0,\ z \in \partial\Omega, \\ w(t,z) &= u(t,z), & t \geq 0,\ z \in \partial\Omega, \end{cases}$$

where $-\frac{\partial}{\partial n}$ *denotes the outer normal derivative. This can be written as*

$$U'(t) = \begin{pmatrix} \Delta & 0 \\ -\frac{\partial}{\partial n} & 0 \end{pmatrix} U(t),$$

where

$$U = \begin{pmatrix} u \\ u_{|\partial\Omega} \end{pmatrix}.$$

Observe that a relationship between the coordinates of U *is intrinsic in the case of the wave equation (there, the first coordinate is the time derivative of the second one) but has to be* imposed *in the case of the heat equation with dynamic boundary conditions.*

Example 1.6. *Consider the* delayed *partial differential equation*

$$\frac{\partial u}{\partial t}(t,x) = \Delta u(t,x) + u(t-1,x)$$

for $t \geq 0$ *and* $x \in \mathbb{R}^d$*. The variation of* u *at time* t *also depends on what has happened to* u *at time* $t-1$ *(think of a population model including pregnancy effects). Finding the right framework for such equations was an open problem for a long time. Finally, it was understood that the above equation cannot be written as an* (ACP) *on a natural function space: indeed, a solution cannot be uniquely determined by a function in* $C_0(\mathbb{R}^d)$*, but rather needs specification of an initial condition in the following form:*

$$u(s,x) = h(s,x),$$

for $s \in [-1,0]$ *and* $x \in \mathbb{R}^d$*. Then, this initial value problem can be written as*

$$U'(t) = AU(t)$$

on the Banach space $X := C([-1,0];C_0(\mathbb{R}^d)$*, where* $Af = f'$ *and the domain of* A *consists of those continuously differentiable functions* $f : [-1,0] \to C_0(\mathbb{R}^d)$ *such that*

- $f(0) \in C_0(\mathbb{R}^d)$ *s.t.* $\Delta f(0) \in C_0(\mathbb{R}^d)$ *and*
- $f'(0) = \Delta f(0) + f(-1)$.

Exercise 1.7. *Consider the* Volterra integro-differential equation

$$\frac{\partial u}{\partial t}(t,x) = \Delta u(t,x) + \int_0^t a(t-s)\Delta u(s)ds,$$

where $a : \mathbb{R}_+ \to \mathbb{R}$ *is a test function (i.e., a* C^∞*-function with compact support). Show that this equation can be formally written as an* (ACP) *with* $X = L^2(\mathbb{R}^d) \times L^1(\mathbb{R}_+; L^2(\mathbb{R}^d))$*.*

Remark 1.8. *The notion of abstract Cauchy problem can be generalised in several directions. For example, one may be interested in considering systems that are governed by physical laws that change in time or depend on the solution itself, like the heat-like equations*

$$\frac{\partial u}{\partial t}(t,x) = \operatorname{div}(a(t) \operatorname{grad} u)(t,x) := \sum_{k=1}^d \frac{\partial^2 u}{\partial x^2}(t,x),$$

or

$$\frac{\partial u}{\partial t}(t,x) = \operatorname{div}(b(\operatorname{grad} u) \operatorname{grad} u)(t,x),$$

for $a : \mathbb{R}_+ \to (0,\infty)$ and $b : \mathbb{R}^d \to (0,\infty)$. In many applications there are good reasons to consider such equations, which are called non-autonomous *and* non-linear*, respectively. However, their analysis is much more delicate and will not be considered in this course.*

Once a partial differential equation has been transformed into a (vector-valued) linear ordinary differential one, it may be tempting to apply known techniques that are already known from the theory of systems of (scalar-valued) ordinary differential equations. In particulat, it is known that if a system of ordinary differential equations can be written as (ACP) for a certain $m \times m$ matrix A, then the solution is given via the *exponential* of A, i.e.,

$$U(t) = e^{tA}U_0, \qquad t \geq 0,$$

where

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}, \qquad t \geq 0.$$

Computing explicitly the power of a matrix is seldom easy, but several numerical methods have been developed to help in this task. Moreover, the exponential function is a power series, so that it directly yields approximated solutions within a desired error range. A similar approach can be used when A is a *bounded* linear operator on X.

Proposition 1.9. *Let X be a Banach space and A a bounded linear operator on X. Then the solution U to (ACP) is given by*

$$U(t) = e^{tA}U_0, \qquad t \geq 0,$$

where

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}, \qquad t \geq 0. \tag{1.1}$$

This series converges uniformly in X. Moreover, $U(t) = e^{tA}U_0$ is a solution of (ACP) also for negative t and in fact for all $t \in \mathbb{C}$.

Proof. It suffices to shows that the partial sums

$$S_m(t) := \sum_{k=0}^{m} \frac{t^k A^k}{k!}, \qquad t \in \mathbb{R}.$$

of the series form a Cauchy sequence (with respect to the operator norm on X): in fact, if $m \geq n$, then for all $t \in \mathbb{R}$

$$\|S_m(t) - S_n(t)\| = \left\| \sum_{k=n}^{m} \frac{t^k A^k}{k!} \right\| \leq \sum_{k=n}^{m} \frac{|t|^k \|A^k\|}{k!} \leq \sum_{k=n}^{m} \frac{|t|^k \|A\|^k}{k!},$$

the Cauchy sequence associated to the partial sums of the series for the real number $e^{|t|\|A\|}$. □

In certain cases, this result is indeed interesting.

Example 1.10. *Let X be the Hilbert space of all functions over $\mathbb{Z}^n$, i.e., of all sequences with indices in $\mathbb{Z}^n$,*

$$X := \{x = (x_{v_1 \ldots v_n}) : v_1, \ldots, v_n \in \mathbb{Z}\}.$$

Each element v of the lattice is considered connected to all those further elements w such that $|v - w| = 1$. In this case, we write $v \sim w$. We want to model the following process: if a particle is in v, then it has probability $\frac{1}{2n}$ to reach any $w \sim v$. Thus, our diffusion operator can is given by

$$A : x \mapsto \sum_{v \sim w} (x_v - x_w).$$

Show that A is bounded on X. We can see this as a diffusion model on a lattice or, more generally, on a graph. In this case, v are the nodes of the (infinite) graph and two nodes v, w are adjacent if and only if $v \sim w$. What happens if we delete edges of $\mathbb{Z}^n$, i.e., if we consider diffusion on a subgraph of $\mathbb{Z}^n$? Is the corresponding (ACP) well-posed?

Unfortunately, in most cases the above Proposition will not help. The vast majority of interesting evolution equations are partial differential equations, and in particular they are associated with unbounded operators and (1.1) does in general not make sense. In fact, every time we take a further power of A, its domain becomes smaller and smaller – in fact, it can tend to $\{0\}$, as we see next.

Exercise 1.11. *Let $X = L^2(0,1)$ and $A : f \mapsto f'$ with domain $D(A) = \{f \in H^1(0,1) : f(0) = 0\}$. Show that if $e^{tA} f$ converges for all $t \in \mathbb{R}$, then $f = 0$.*

Remark 1.12. *The above exercise may look like the end of the story, but is not. In fact, the (ACP) associated with A on X as in Exercise 1.11 is well-posed, as we have seen in Example 1.1. Setting*

$$U(t) := T(t)U_0 = \begin{cases} U_0(t+\cdot), & \text{on } (0, 1-t), \\ 0, & \text{on } (1-t, 1), \end{cases}$$

we obtain a family $(T(t))_{t\geq 0}$ that enjoys most of the nice properties of the operator family defined in (1.1). In particular, $T(t)$ is a bounded linear operators for all $t \geq 0$ and it satisfies a certain functional equation that is typical of the exponential function. Most important, $\mathbb{R}_+ \ni t \mapsto T(t)U_0 \in L^2(0,1)$ is a mapping that solves the (ACP) associated with the operator considered in Exercise 1.11.

We are thus led to the following.

Definition 1.13. *Let X be a Banach space. A family $(T(t))_{t\geq 0}$ of bounded linear operators on X is called a* strongly-continuous semigroup *or* C_0-semigroup *if*

(SEMIGR) $$T(t)T(s) = T(t+s) \quad \text{for all } t, s \geq 0 \qquad \text{and} \qquad T(0) = \mathrm{Id},$$

and moreover if the orbit *$\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is continuous for all $x \in X$.*

Example 1.14. *Let X be a Banach space and let $(T(t))_{t\geq 0}$ be a C_0-semigroup on X. Let $\lambda \in \mathbb{C}$. Then the family*

$$\tilde{T}(t) := e^{\lambda t} T(t), \qquad t \geq 0,$$

defines a C_0-semigroup on X. In fact, $(SEMIGR)$ is satisfied owing to the elementary properties of the exponential function, while strong continuity holds since the product of two convergent functions converges to the product of the two limits. Moreover, it is straightforward to check that $A - \lambda$ is the generator of $(\tilde{T}(t))_{t\geq 0}$.

While the first two conditions in Definition 1.13 are somehow natural, since we want to reproduce the behaviour of the exponential function, one may be surprised of the weakness of the third condition. If A is a bounded linear operator, then it follows directly from the proof of Proposition 1.9 that the mapping $t \mapsto e^{tA}$ is continuous with respect to the operator norm (we say the semigroup is *norm continuous*). The following shows that the converse is also true.

Proposition 1.15. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a norm continuous C_0-semigroup. Then there exists a bounded linear operator A on X such that $T(t) = e^{tA}$ for all $t \geq 0$. Moreover, in this case $(T(t))_{t\geq 0}$ extends by continuity to a family $(T(t))_{t\in\mathbb{C}}$ that also satisfies the first two conditions of Definition 1.13.*

Proof. Set

$$V(t) := \int_0^t T(s)ds, \qquad t \geq 0.$$

Now, due to norm continuity these Bochner integrals converge in operator norm, hence define a family of bounded linear operators $(V(t))_{t\geq 0}$. Due to the fundamental theorem of (vector-valued) calculus, the mapping $\mathbb{R}_+ \ni t \mapsto V(t) \in \mathcal{L}(X)$ is continuously differentiable and there holds

$$\frac{d}{dt}V(t) = T(t), \qquad t \geq 0.$$

In particular,

$$\frac{d}{dt}V(0) = T(0) = \mathrm{Id}, \qquad t \geq 0.$$

Since $\frac{d}{dt}V(0) = \lim_{t\to 0+} \frac{1}{t}V(t)$, this shows that each $\frac{1}{t}V(t)$ and hence each $V(t)$ is invertible, for t small enough. Now, take $h > 0$ small enough that $V(h)$ is indeed invertible. Then

$$\begin{aligned} T(t) &= V(h)^{-1}V(h)T(t) \\ &= V(h)^{-1}\int_0^h T(s)T(t)ds \\ &= V(h)^{-1}\int_0^h T(s+t)ds \\ &= V(h)^{-1}\int_t^{t+h} T(s)ds \\ &= V(h)^{-1}\left(V(t+h) - V(t)\right). \end{aligned}$$

Now, $t \mapsto T(t)$ is a composition of differentiable functions, hence differentiable itself. More precisely,

$$\begin{aligned} \frac{d}{dt}T(t) &= \lim_{t_0\to 0+} \frac{T(t+t_0) - T(t)}{t_0} \\ &= \lim_{t_0\to 0+} \frac{T(t_0) - \mathrm{Id}}{t_0}T(t) \\ &= \frac{d}{dt}T(0)T(t) =: AT(t) \end{aligned}$$

holds for all $t \geq 0$. Since A is the limit (in operator norm!) of the incremental quotient, it is bounded as well. This concludes the proof. □

Definition 1.16. *Let X be a Banach space. A family $(T(t))_{t\geq 0}$ of bounded linear operators on X is called a* strongly-continuous group *or* C_0-group *if*

(GR) $\qquad T(t)T(s) = T(t+s)$ *for all* $t, s \in \mathbb{R}$ *and* $T(0) = \mathrm{Id}$,

and moreover if $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is continuous for all $x \in X$.

In particular, each operator $T(t)$ of a C_0-group is invertible with inverse $T(-t)$. In fact, C_0-groups are very special objects. They often have very good properties, as the following shows.

Definition 1.17. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a C_0-semigroup on X. The* generator *of $(T(t))_{t\geq 0}$ is the (possibly unbounded) linear operator on X defined by*

$$\begin{aligned} D(A) &:= \left\{x \in X : \lim_{t\to 0+} \frac{T(t)x - x}{t} \text{ exists}\right\}, \\ Ax &:= \lim_{t\to 0+} \frac{T(t)x - x}{t}. \end{aligned}$$

We also say that A generates *$(T(t))_{t\geq 0}$.*

Exercise 1.18. *Let H be a Hilbert space and A be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ on H. Define the* adjoint semigroup *as the operator family $(S(t))_{t\geq 0}$ such that*

$$S(t) := T(t)^*, \qquad t \geq 0.$$

Show that the adjoint semigroup is a C_0-semigroup on H whose generator is A^*.

Exercise 1.19. *Show that the generator of the C_0-semigroup introduced in Remark 1.12 is the operator introduced in Example 1.11.*

The reason why generators are important is that, just like in the finite dimensional case, the associated semigroup yields always the solution of an abstract Cauchy problem.

Theorem 1.20. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a C_0-semigroup on X. Denote by A its generator. Then the following assertions hold.*

(1) There exist constants $\omega \in \mathbb{R}$ and $M \geq 1$ such that

$$\|T(t)\|_{\mathcal{L}(X)} \leq M e^{\omega t} \qquad \textit{for all } t \geq 0.$$

In particular, for all $t_0 > 0$ there exists $M \geq 1$ such that

$$\|T(t)\|_{\mathcal{L}(X)} \leq M \qquad \textit{for all } t \in [0, t_0].$$

(2) If $x \in D(A)$, then $T(t)x \in D(A)$ for all $t \geq 0$ and moreover $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is continuously differentiable with

$$\frac{d}{dt}T(t)x = T(t)Ax = AT(t)x \qquad \textit{for all } t \geq 0.$$

In particular, $t \mapsto T(t)x$ is the solution of the abstract Cauchy problem associated with A with initial data x.

(3) If $x \in X$, then for all $t \geq 0$

$$\int_0^t T(s)x ds \in D(A),$$

so that

$$T(t)x - x = A \int_0^t T(s)x ds.$$

(4) If moreover $x \in D(A)$, then for all $t \geq 0$

$$T(t)x - x = \int_0^t T(s)Ax ds$$

(5) $D(A)$ is dense in X.

Definition 1.21. *Under the assumptions and with the notations of Theorem 1.20, the semigroup is called* uniformly bounded *if $\omega = 0$, and is called* contractive *if $\omega = 0$ and additionally $M = 1$.*

Proof. (1) Assume that there exists a sequence $(t_n)_{n\in\mathbb{N}} \subset [0, \infty)$ such that $\lim_{n\to\infty} t_n = 0$ but $\lim_{n\to\infty} \|T(t_n)\|_{\mathcal{L}(X)} = \infty$. If follows from the uniform boundedness priciple that also the numerical sequence $(\|T(t_n)x\|_X)_{n\in\mathbb{N}}$ is unbounded, in spite of the fact that $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is continuous at 0.

Hence, there exists $M \geq 1$ such that $\|T(s)\| \leq M$ for all $s \in [0, 1]$. In order to prove the general estimate, we perform a typical trick allowed by the semigroup law: we take $t \geq 0$ and split it into its natural part $n \in \mathbb{N}$ and its decimal part $s \in [0, 1)$, $t = s + n$. Then, we estimate via the submoltiplicativity of the norm

$$\|T(t)\| = \|T(s+n)\| = \|T(s)T(n)\| \leq \|T(s)\|\|T(n)\| = \|T(s)\|\|T(1)^n\| \leq \|T(s)\|\|T(1)\|^n \leq M^{n+1} \leq Me^{n\log M} \leq Me^{t\omega},$$

for all $t \geq 0$, where we have set $\omega := \log M$, since $t \geq n$.

(2) Let $x \in D(A)$ and $t > 0$. We want to show that $T(t)x \in D(A)$. In fact,

$$\lim_{h\to 0+} \frac{T(h)T(t)x - T(t)x}{h} = \lim_{h\to 0+} T(t)\frac{T(h)x - x}{h} = T(t) \lim_{h\to 0+} \frac{T(h)x - x}{h} = T(t)Ax.$$

In particular, this limit exists, which is all we need to show by definition of $D(A)$. Moreover, there also holds

$$AT(t)x = \lim_{h\to 0+} \frac{T(h)T(t)x - T(t)x}{h} = \lim_{h\to 0+} \frac{T(h) - \mathrm{Id}}{h} T(t)x = AT(t)x.$$

(3) Let $x \in X$ and $t > 0$. Then,

$$\begin{aligned}
\lim_{h\to 0+} \frac{T(h) - \mathrm{Id}}{h} \int_0^t T(s)xds &= \lim_{h\to 0} \frac{\int_0^t T(s+h)xds - \int_0^t T(s)xds}{h} \\
&= \lim_{h\to 0+} \frac{\int_h^{t+h} T(s)xds - \int_0^t T(s)xds}{h} \\
&= \lim_{h\to 0+} \frac{\int_t^{t+h} T(s)xds - \int_0^h T(s)xds}{h} \\
&= \lim_{h\to 0+} \frac{T(t)\int_0^h T(s)xds}{h} - \lim_{h\to 0} \frac{\int_0^h T(s)xds}{h} \\
&= T(t)x - x,
\end{aligned}$$

hence this limit exists and by definition

$$\int_0^t T(s)xds \in D(A)$$

and moreover

$$A\int_0^t T(s)xds = T(t)x - x.$$

(4) Let $x \in D(A)$ and $t > 0$. Set

$$f_h(s) := T(s)\frac{T(h)x - x}{h}, \qquad s \in [0,t],\ h > 0,$$

and

$$f(s) := T(s)Ax, \qquad s \in [0,t].$$

Then each $f_h : [0,t] \to X$ is continuous and one has

$$\lim_{h\to 0+} f_h = f \qquad in C([0,t];X),$$

since

$$\|f - f_h\|_\infty = \max_{s\in[0,t]} \|T(s)\|_{\mathcal{L}(X)} \left\| Ax\frac{T(h)x - x}{h} \right\|_X = 0,$$

by (1), and hence

$$\lim_{h\to 0+} \frac{T(h) - \mathrm{Id}}{h} \int_0^t T(s)xds = \lim_{h\to 0+} \int_0^t T(s)\frac{T(h)x - x}{h} ds = \int_0^t T(s)Axds.$$

(5) Let $x \in X$. Then

$$x = T(0)x = \lim_{t\to 0+} \frac{1}{t} \int_0^t T(s)xds$$

by the fundamental theorem of calculus. It follows from (2) that the RHS is the limit of a family of elements in $D(A)$. This yields the claim. □

Now, observe that to each C_0-semigroup corresponds a unique generator, by definition. Moreover, the generator is never trivial, by Theorem 1.20.(3) below). On the other hand, one can prove that to each generator corresponds a unique C_0-semigroup.

Lemma 1.22. *Let X be a Banach space and A be a linear operator on X. Then there is at most one continuously differentiable mapping $\mathbb{R}_+ \ni t \mapsto T(t) \in \mathcal{L}(X)$ (that is, with respect to the operator norm) such that*

$$\frac{d}{dt}T(t) = AT(t) \quad \text{for all } t \geq 0 \qquad \text{and} \qquad T(0) = \mathrm{Id}\,.$$

If in particular A is a generator, than the mapping whose existence is stated above is just the generated C_0-semigroup.

Proof. Let $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ two such families and define for a given $t \geq 0$

$$Q(s) := T(s)S(t-s), \qquad s \in [0,t].$$

Then $Q(\cdot)$ is continuously differentiable with

$$\frac{dQ}{dt}(s) = AT(s)S(t-s) + T(s)(-A)S(t-s) = 0,$$

hence

$$T(t) = Q(t) = Q(0) = S(t).$$

This completes the proof. ☐

Exercise 1.23. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup with generator A. A vector $x \in \bigcap_{k=1}^{\infty}$ is called* entire *for $(T(t))_{t\geq 0}$ if*

$$e^{tA}x := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k x$$

converges for all $t \in \mathbb{R}$.

(1) Show that if a vector is entire for $(T(t))_{t\geq 0}$, then $T(t)x = e^{tA}x$ for all $t \geq 0$.

(2) Let $(T(t))_{t\in\mathbb{R}}$ be a C_0-group on a Banach space X. For a given $x \in X$, set

$$x_n(z) := \sqrt{\frac{n}{2\pi}} \int_{\mathbb{R}} e^{-\frac{n(t-z)^2}{2}} T(t)x dt, \qquad n \in \mathbb{N},\ z \in \mathbb{C}.$$

Show that these integrals converge in X (strongly, i.e., for each given x), that each x_n is an entire function (i.e., it is holomorphic over $\mathbb{C}$), that $x_n(0) \to x$ in X, and finally that $x_n(s) = T(s)x_n$ for all $n \in \mathbb{N}$ and all $s \in \mathbb{R}$. Deduce that $\mathbb{R} \ni s \mapsto T(s)x_n \in X$ extends to an entire function.

(3) Conclude that if $(T(t))_{t\in\mathbb{R}}$ is a C_0-group, then the set of entire vectors is dense in X and $T(t)x = e^{tA}x$ for all $t \in \mathbb{R}$ and all entire vectors x.

CHAPTER 2

The Spectral Theorem

All in all, we have seen that checking the generator property of a linear operator A is a most efficient way to discuss well-posedness of (ACP). When is an operator a generator, then? A general, complete answer is known (the Theorem of Hille–Yosida, see Chapter 3), but it yields a characterisation which is almost impossible to check, in many situations. This is why we start with a special case – that of self-adjoint operators – which is however sufficient in many relevant applications to evolution equations, ranging from the Schrödinger to the heat equation, from the wave to the Dirac equation.

2.1. Adjoint operators

According to many theoretical physicists, the rule of thumb holds that an (ACP) is "natural" (a notion that is as elusive as that of "natural behaviour" of an animal or "natural ingredient" of a dish) if and only if A is self-adjoint. We will see that for self-adjoint operators it is particularly easy to prove well-posedness of (ACP) and, in some lucky cases, even give a semi-explicit formula for the solution.

Proposition 2.1. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. Then there exists exactly one bounded linear operator T^* from H_2 to H_1 such that*

$$(Tx|y)_{H_2} = (x|T^*y)_{H_1} \qquad \text{for all } x \in H_1 \text{ and } y \in H_2. \tag{2.1}$$

holds. Moreover,

$$\|T\| = \|T^*\| \qquad \text{and} \qquad \|T^*T\| = \|T\|^2 \qquad \text{for all } T \in \mathcal{L}(H). \tag{2.2}$$

Proof. Let $y \in H_2$. Then $\phi_y : H_1 \ni x \mapsto (Tx|y)_{H_2} \in \mathbb{K}$ defines a bounded linear functional, since by the Cauchy–Schwarz inequality $|\phi_y(x)| \le \|T\|\|x\|_{H_1}\|y\|_{H_2}$. Therefore, by the Representation Theorem of Riesz–Fréchet there exists a vector $T^*y \in (H_1)' \cong H_1$ such that $(Tx|y)_{H_2} = \langle \phi_y, x\rangle = (x|T^*y)_{H_1}$. This defines an operator $T^* : H_2 \ni y \mapsto T^*y \in H_1$.

To check linearity of T^*, take $y_1, y_2 \in H_2$ and observe that for all $x \in H_1$

$$\begin{aligned}(x|T^*(y_1+y_2) - T^*y_1 - T^*y_2)_{H_1} &= (x|T^*(y_1+y_2))_{H_2} - (x|T^*y_1)_{H_1} - (x|T^*y_2)_{H_1}\\ &= (Tx|y_1+y_2)_{H_1} - (Tx|y_1)_{H_2} - (Tx|y_2)_{H_2} = 0.\end{aligned}$$

Accordingly, $T^*(y_1+y_2) - T^*y_1 - T^*y_2$ belongs to $H^\perp = \{0\}$ for all y_1, y_2. Similarly, take $y \in H_2$ and $\lambda \in \mathbb{K}$ and observe that for all $x \in H_1$

$$\begin{aligned}(x|T^*(\lambda y) - \lambda T^*y)_{H_1} &= (x|T^*(\lambda y))_{H_2} - (x|\lambda T^*y)_{H_1}\\ &= (Tx|\lambda y)_{H_1} - \overline{\lambda}(Tx|y) = 0,\end{aligned}$$

i.e., $T^*(\lambda y) - \lambda T^*y$ belongs to $H^\perp = \{0\}$ for all $y \in H_2$ and all $\lambda \in \mathbb{K}$.

Boundedness of T^* follows by boundedness of T, since for all $y \in H_2$

$$\|T^*y\|_{H_1} \le \|\phi_y\| \le \|T\|\|y\|_{H_2}.$$

This shows that $\|T^*\| \le \|T\|$. Conversely, take $x \in H_1$ with $\|x\|_{H_1} \le 1$ and observe that

$$\|Tx\|^2_{H_2} = (Tx|Tx)_{H_2} = (T^*Tx|x)_{H_1} \le \|T^*Tx\|_{H_1}\|x\|_{H_1} \le \|T^*T\| \le \|T^*\|\|T\|.$$

Accordingly, $\|T\|^2 \leq \|T^*\|\|T\|$ and in particular $\|T\| \leq \|T^*\|$.

Finally, for all $x \in H_1$ with $\|x\|_{H_1} \leq 1$ one sees as above that

$$\|T^*Tx\| \geq |(T^*Tx|x)| = (Tx|Tx) = \|Tx\|^2,$$

and accordingly

$$\|T\|^2 \leq \|T^*T\| \leq \|T^*\|\|T\| = \|T\|\|T\| = \|T\|^2.$$

This completes the proof. □

Exercise 2.2. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2. Show that $T^{**} = T$.*

Definition 2.3. *Let H_1, H_2 be Hilbert spaces and $T \in \mathcal{L}(H_1, H_2)$. The unique operator $T^* \in \mathcal{L}(H_2, H_1)$ that satisfies* (2.1) *is called the* adjoint *of T.*

If $H_1 = H_2$ and $T^* = T$, then the operator T is called self-adjoint.

Example 2.4. *For any continuous $k : [0,1] \times [0,1] \to \mathbb{R}$ that is symmetric (i.e., such that $k(x,y) = k(y,x)$ for all $x, y \in [0,1]$) the* Fredholm operator *F_k defined by*

$$F_k f(x) := \int_0^1 k(x,y) f(y) dy$$

is self-adjoint. More generally, for any continuous $k : [0,1] \times [0,1] \to \mathbb{C}$ the adjoint of F_k is given by $F_k^* = F_{k^*}$, where k^* is defined by $k^*(x,y) := \overline{k(y,x)}$ a.e.

Example 2.5. *If $H = \mathbb{C}^d$ and $A \in \mathcal{L}(H) = M_d(\mathbb{C})$, then $A^* = \overline{A^T}$, where A^T denotes the transposed matrix of A.*

Example 2.6. *The operator $\lambda := \lambda \operatorname{Id}$ is self-adjoint for all $\lambda \in \mathbb{R}$. More generally, the adjoint of $\lambda \operatorname{Id}$ is $\overline{\lambda} \operatorname{Id}$ for all $\lambda \in \mathbb{C}$.*

Exercise 2.7. *Let H be a Hilbert space and $P \in \mathcal{L}(H)$ be a projector, i.e., $P^2 = P$. Show that P is the orthogonal projector of H onto some closed subspace if and only if P is self-adjoint.* (By definition, P is an orthogonal projector if and only if its null space and its range are orthogonal to each other).

Definition 2.8. *A linear operator U on a Hilbert space H is called* unitary *if U is invertible and $U^{-1} = U^*$. It is called* involutory *if $U^2 = \operatorname{Id}$. It is called* normal *if $UU^* = U^*U$.*

Lemma 2.9. *Let H be a Hilbert space and A be a bounded linear operator on H. Then $\|Ax\| = \|A^*x\|$ for all $x \in H$ if and only if A is normal.*

Proof. One has

$$\|Ax\|^2 = (Ax|Ax) = (A^*Ax|x) \stackrel{!}{=} (AA^*x|x) = (A^*x|A^*x) = \|A^*x\|^2$$

if and only if A is normal. □

Exercise 2.10. *(1) Show that a unitary operator is necessarily bounded and determine its norm.*

(2) Show that if an operator possesses the any two of the properties of being unitary, involutory, self-adjoint, then it possesses also the third.

(3) Show that if U is a normal operator, then so is $U + \lambda$ for any $\lambda \in \mathbb{C}$.

(4) Let H be a Hilbert space and E be a closed subspace of H. Show that if U is normal and leaves both E and $E^\perp$ invariant, then the adjoint of the restriction of U to E agrees with the restriction to E of the adjoint, and in particular the restriction of U to E is normal.

Definition 2.11. *Two linear operators A, B on a Hilbert space H are called* unitarily equivalent *if there exists a unitary operator U such that $A = U^{-1}BU$.*

It is apparent that unitary equivalence is an equivalence relation.

Beside the already discussed vector space structure, the set $\mathcal{L}(H)$ of all bounded linear operators on a Hilbert space H has also other significant properties. In particular, since it is closed under composition of its elements it qualifies as an algebra. We have seen that $\mathcal{L}(H)$ is even closed under a further operation, the so-called *involution* $A \mapsto A^*$. This leads us to introduce the following definition, first proposed by Israel Gelfand and Mark Neumark in 1943.

Definition 2.12. *A* complex Banach algebra $\mathcal{A}$ *is a complex algebra endowed with a submultiplicative mapping* $\|\cdot\| : \mathcal{A} \to \mathbb{R}$, *with respect to which* $\mathcal{A}$ *is a complete normed space and such that if* $\mathcal{A}$ *contains a unity* I, *then* $\|I\| = 1$. *A mapping* $* : \mathcal{A} \ni A \mapsto A^* \in \mathcal{A}$ *is called an* involution *if*

$$A^{**} = A, \quad (AB)^* = B^*A^* \quad \text{and} \quad (\lambda A + B)^* = \overline{\lambda} A^* + B^* \qquad \text{for all } A, B \in \mathcal{A} \text{ and all } \lambda \in \mathbb{C}.$$

A complex Banach algebra endowed with an involution $*$ *is called a* C^*-algebra *if additionally* $\|A^*A\| = \|A\|^2$ *for all* $A \in \mathcal{A}$. *(We have already shown that* $\mathcal{L}(H)$ *is a* C^**-algebra for each Hilbert space* H*).*

In particular, observe that all notions related to adjointness (and in particular those of self-adjoint, unitary and normal operators) can be naturally defined for elements of a C^*-algebra as well.

Exercise 2.13. *Let* $q : \Xi \to \mathbb{C}$ *be a bounded measurable function, where* Ξ *is a* σ*-finite measure space. Consider the linear operator* M_q *defined by*

$$M_q u := q \cdot u, \qquad u \in L^2(\Xi).$$

Show that the mapping $\mathcal{F} : B(\mathbb{R}) \ni f \mapsto f(M_q) \in \mathcal{L}(L^2(\Xi))$ *is a* $*$*-homomorphism, i.e., a mapping such that*

$$\mathcal{F}(f \cdot g) = \mathcal{F}(f)\mathcal{F}(g) \qquad \text{and} \qquad \mathcal{F}(f^*) = \mathcal{F}(f)^* \qquad \text{for all } f, g \in B(\mathbb{R}).$$

Remark 2.14. *One one hand, self-adjoint (bounded) operators have very special properties. On the other hand, it is possible to extend to general bounded operators many results that are first proved for self-adjoint operators by observing that each operator* A *can be written as*

$$A = S_1 + iS_2,$$

where

$$S_1 = \frac{1}{2}(A + A^*) \qquad \text{and} \qquad S_2 = \frac{1}{2i}(A - A^*),$$

with both S_1 *and* S_2 *self-adjoint. The operator* iS_2 *is* skew-adjoint*, i.e., it is self-adjoint after multiplication by* i[1]. *Skew-adjoint operators play an important rôle in the mathematical formulation of quantum mechanics and will be encountered later in connection with Stone's Theorem (Theorem 3.11).*

As we have already seen, in most cases the partial differential equations we are interested in are not associated with bounded operators. We therefore generalise the notion of self-adjoint operator as follows.

Definition 2.15. *Let* H_1, H_2 *be Hilbert spaces and consider the (possibly unbounded)* $A : H_1 \supset D(A) \to H_2$ *and* $B : H_1 \supset D(B) \to H_2$. *We say that* B contains A *and write*

$$A \subset B,$$

if $G(A) \subset G(B)$, *i.e., if* $D(A) \subset D(B)$ *and* $Ax = Bx$ *for all* $x \in D(A)$. *The operator* B *is called the* closure *of* A, *and we write* $B = \overline{A}$, *if it is closed, it contains* A *and moreover it is contained in any other linear operator that contains* A.

Definition 2.16. *Let* H *be a Hilbert space and* A *a (possibly unbounded, but densely defined) linear operator on* H. *Then the* adjoint A^* *of* A *is the linear operator defined by*

$$\begin{aligned} D(A^*) &:= \{y \in H : \exists z \in H : (Ax|y) = (x|z)\ \forall x \in D(A)\} \\ Ay &:= z. \end{aligned}$$

[1] Equivalently (why?), an operator A is skew-adjoint if $A + A^* = 0$.

The operator A is called symmetric *if $(Ax|y) = (x|Ay)$ for all $x, y \in D(A)$, that is, if $A \subset A^*$, and* self-adjoint *if $A = A^*$.*

Exercise 2.17. *Let H_1, H_2, H_3 be Hilbert spaces and consider densely defined operartors $T : H_1 \supset D(T) \to H_2$, $S : H_1 \supset D(S) \to H_2$, $R : H_2 \supset D(R) \to H_3$. Prove the following assertions.*

(a) $(\lambda T)^* = \overline{\lambda} T$ *for all* $\lambda \in \mathbb{C}$.
(b) $T^* + S^* \subset (T + S)^*$ *if* $T + S$ *is densely defined, with equality if* T, S *are bounded.*
(c) $T^* R^* \subset (RT)^*$, *with equality if* T, S *are bounded.*
(d) $S^* \subset T^*$ *if* $T \subset S$.
(e) $\operatorname{Ker}(\lambda - T^*)^\perp = \operatorname{Rg}(\overline{\lambda} - T)$ *for all* $\lambda \in \mathbb{C}$, *provided that* $H_1 = H_2$.
(f) $\operatorname{Ker}(T^*) = (\operatorname{Rg} T)^\perp$ *and* $\overline{\operatorname{Rg}(T^*)} = (\operatorname{Ker} T)^\perp$.

Exercise 2.18. *Let H be a Hilbert space and A be an unbounded operator on H. Let B be a bounded self-adjoint operator on H. Show that A with domain $D(A)$ is self-adjoint if and only if $A + B$ with domain $D(A)$ is self-adjoint.*

Lemma 2.19. *Let H_1, H_2 be Hilbert spaces and A be a bounded linear operator from H_1 to H_2. Then the following assertions hold.*

(a) A has closed range if and only if A^* has closed range.
(b) A is an isomorphism if and only if there exists $c > 0$ such that

$$\|Ax\| \geq c\|x\| \qquad \text{and} \qquad \|A^*y\| \geq c\|y\| \qquad \text{for all } x \in H_1,\ y \in H_2.$$

Proof. (a) If A^* has closed range, then $\operatorname{Rg}(T^*) = (\operatorname{Ker} T)^\perp$ and therefore $\overline{\operatorname{Rg}(T^*)}$ is an orthogonal subspace, hence it is closed. Since A is bounded, $A^{**} = A$ and the same argument yields the converse implication.

(b) If A is an isomorphism, then it is a classical consequence of the open mapping theorem that the former inequality holds. Furthermore, $\operatorname{Rg} A = H_2$ ☐

Exercise 2.20. *Let H be a Hilbert space and A a linear symmetric operator on H. Show that*

$$\|(T - \lambda)x\|^2 = \|(T - \operatorname{Re} \lambda)x\|^2 + |\operatorname{Im} \lambda|^2 \|x\|^2$$

holds for all $x \in D(A)$ and all $\lambda \in \mathbb{C}$.

For bounded operators it is clear that symmetry is the same as self-adjointness. In general, however, this is not true: it may happen that an operator is symmetric but it is "too small" to be self-adjoint, as in the case of the Laplacian on $L^2(0,1)$ whose domain agrees with the test functions on $(0,1)$. Let us consider another example in more detail.

Exercise 2.21. *Consider the operators A, B on $L^2(0,1)$ defined by*

$$Af := if', \qquad f \in D(A) = \{u \in H^1(0,1) : u(0) = 0\}$$

and

$$Bf := if', \qquad f \in D(B) = \{u \in H^1(0,1) : u(0) = 0 = u(1)\}.$$

(a) Show that A is self-adjoint.
(b) Show that B is closed and symmetric.
(c) Using the fact that if $g \in L^2(0,1)$ is orthogonal to f' for all test functions f on $(0,1)$, then g is constant a.e., prove that the adjoint of B has domain $H^1(0,1)$. Hence, B is not self-adjoint.
(d) Find (further) self-adjoint operators between B and B^*.

Definition 2.22. *Let X be a Banach space and A a (possibly unbounded) linear operator on X. Then A is called* closed *if its* graph *$G(A) := \{(x, Ax) \in D(A) \times X\}$ is closed in $X \times X$, i.e., if*

$$x \in D(A) \qquad \text{and} \qquad Ax = y$$

whenever

$$\lim_{n\to\infty} x_n = x \qquad \text{and} \qquad \lim_{n\to\infty} Ax_n = y$$

for a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$. *It is called* densely defined *if* $D(A)$ *is dense in* X.

Observe that if A is bounded (and in particular $D(A) = X$), then it is automatically closed. Conversely, if it is closed and $D(A) = X$, then A is bounded – this is nothing but the statement of one of the many consequences of the theorem of Banach–Steinhaus, which we recall next.

Theorem 2.23 (Closed Graph Theorem)**.** *Let* X, Y *be Banach spaces. Then a linear operator* $A : X \to Y$ *with* $D(A) = X$ *is bounded if and only if its graph* $G(A) = \{(x, Ax) \in X \times Y\}$ *is closed in* $X \times Y$.

Exercise 2.24. *Prove the* Hellinger–Toeplitz-Theorem*:*
Let H *be a Hilbert space and* $A : H \to H$ *with* $D(A) = H$ *be symmetric. Then* A *is bounded.*

Exercise 2.25. *Let* X *be a Banach space and* A *a (possibly unbounded) linear operator on* X. *Show that* A *is closed if and only if the vector space* $D(A)$ *is complete with respect to the so-called* graph norm

$$\|x\|_A := \|x\| + \|Ax\|, \qquad x \in X.$$

Lemma 2.26. *Let* H *be a Hilbert space and* A *a linear, densely defined operator on* H. *Then* A^* *is closed. In particular, each self-adjoint operator is closed.*

PROOF. Consider the operator matrix

$$U := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix},$$

which is unitary on $H \times H$. For all $x \in D(A)$ and $y \in D(A^*)$ one has

$$\Big((y, A^*y)|U(x, Ax)\Big) = \Big((y, A^*y)|(-Ax, x)\Big) = -(y|Ax) + (A^*y, x) = 0,$$

by definition of A^*. This shows that $G(A^*) \subset U(G(A))^\perp$, where $G(A)$ denotes the graph of A. If conversely $(y, z) \in U(G(A))^\perp$, then for all $x \in D(A)$ the vectors (y, z) and $U(x, Ax)$ are orthogonal, hence

$$0 = \Big((y, z)|(-Ax, x)\Big) = -(y|Ax) + (z|x),$$

and therefore $(Ax|y) = (x|z)$, i.e., $z = A^*y$, $(y, z) = (y, A^*y) \in G(A^*)$, and $G(A^*) \supset U(G(A))^\perp$.
This concludes the proof, because $U(G(A))^\perp$ is an orthogonal subspace, hence it is closed. □

Usually, checking symmetry of an operator is quite easy, but self-adjointness is much more delicate.

Theorem 2.27. *Let* H *be a Hilbert space and* A *a linear symmetric operator on* H. *Then the following assertions are equivalent.*

(i) A *is self-adjoint.*
(ii) A *is closed and both* $A^* + i$ *and* $A^* - i$ *are injective.*
(iii) Both $A + i$ *and* $A - i$ *are surjective.*

We first need to establish the following.

Lemma 2.28. *Let* H *be a Hilbert space and* A *a linear symmetric operator on* H. *Then the following assertions hold.*

(a) If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, *then* $A - \lambda$ *is injective.*
(b) If A *is closed and* $\lambda \in \mathbb{C} \setminus \mathbb{R}$, *then* $A - \lambda$ *has closed range.*
(c) If $\lambda \in \mathbb{C} \setminus \mathbb{R}$ *and* $\mu \in \mathbb{C}$ *with* $\mu \in B_{\mathrm{Im}\,\lambda}(\lambda)$, *then* $\mathrm{Ker}(A^* - \mu)$ *and* $\mathrm{Ker}(A^* - \lambda)^\perp$ *have trivial intersection.*
(d) If A *is closed, then the spaces* $\mathrm{Ker}(A^* - \lambda)$ *have the same dimension for all* $\lambda \in \mathbb{C}$ *with* $\mathrm{Im}\,\lambda > 0$, *and they have the same dimension for all* $\lambda \in \mathbb{C}$ *with* $\mathrm{Im}\,\lambda < 0$.

Proof. (a) It follow from Exercise 2.20 that

$$\|(A-\lambda)x\| \geq |\operatorname{Im}\lambda|\|x\| \qquad \text{for all } x \in D(A),$$

hence $A-\lambda$ is injective.

(b) Let $(y_n)_{n\in\mathbb{N}} \subset \operatorname{Rg}(A-\lambda)$, say $y_n := (A-\lambda)x_n$ for some $(x_n)_{n\in\mathbb{N}} \subset D(A)$. Let $y \in H$ such that $\lim_{n\to\infty} y_n = y$. Now, observe that for all $n, m \in \mathbb{N}$ it follows from (a) that

$$\|y_n - y_m\| = \|(A-\lambda)(x_n - x_m)\| \geq |\operatorname{Im}\lambda|\|x_n - x_m\|,$$

hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence because so is $(y_n)_{n\in\mathbb{N}}$. It follows that $\lim_{n\to\infty} x_n = x$ for some $x \in H$. But then closedness of $A-\lambda$ implies that $x \in D(A)$ and $Ax - \lambda x = y$, i.e., $y \in \operatorname{Rg}(A-\lambda)$.

(c) Let $x \in \operatorname{Ker}(A^*-\mu) \cap \operatorname{Ker}(A^*-\lambda)^\perp$, $x \neq 0$ – hence, we may normalise it. Since $x \in \operatorname{Ker}(A^*-\lambda)^\perp = \operatorname{Rg}(A-\overline{\lambda})$ by Exercise 2.17.(e), there exists $y \in D(A)$ with $x = Ay - \overline{\lambda}y$. Therefore, since $A^*y - \mu y$ and $x = Ay - \overline{\lambda}y$ are orthogonal, it follows that

$$0 = (A^*x - \mu x|y) = (x|Ay - \overline{\mu}y) = (x|Ay - \overline{\lambda}y + (\overline{\lambda}-\overline{\mu})y) = (x|(A-\overline{\lambda})^*(A-\overline{\lambda})x) + (x|(\overline{\lambda}-\overline{\mu})y).$$

Now, by (2.2) we obtain

$$0 = \|x\|^2 + (\lambda-\mu)(x|y),$$

and therefore by the Cauchy–Schwarz inequality and (a)

$$1 = \|x\|^2 = (\mu-\lambda)(x|y) \leq |\mu-\lambda|\|y\| \leq \frac{|\mu-\lambda|}{|\operatorname{Im}\lambda|}\|Ay - \overline{\lambda}y\| = \frac{|\mu-\lambda|}{|\operatorname{Im}\lambda|}\|x\| < 1.$$

This contradiction completes the proof.

(d) It follows from (c) that for f $\lambda \in \mathbb{C}\setminus\mathbb{R}$ and $\mu \in \mathbb{C}$ with $\mu \in B_{\operatorname{Im}\lambda}(\lambda)$ the dimension of $\operatorname{Ker}(A^*-\lambda)$ is larger than the dimension of $\operatorname{Ker}(A^*-\mu)$ (why?). If in particular $\mu \in B_{\frac{1}{2}\operatorname{Im}\lambda}(\lambda)$, then we also have that $\lambda \in B_{\operatorname{Im}\mu}(\mu)$, and accordingly the dimension of $\operatorname{Ker}(A^*-\lambda)$ is *smaller* than the dimension of $\operatorname{Ker}(A^*-\mu)$, i.e., the dimension of each $\operatorname{Ker}(A^*-\lambda)$ is constant in a suitably small neighbourhood of $\lambda \in \mathbb{C}\setminus\mathbb{R}$. It follows that the dimension of the spaces $\operatorname{Ker}(A^*-\lambda)$ is constant in either connected component of $\mathbb{C}\setminus\mathbb{R}$. ☐

Proof of Theorem 2.27. (i)⇒(ii) Since A is self-adjoint, it is closed by Lemma 2.26 and moreover $\operatorname{Ker}(A^* \pm i) = \operatorname{Ker}(A \pm i) = \{0\}$ by Lemma 2.28.(a).

(ii)⇒(iii) Both $A+i$ and $A-i$ have closed range by Lemma 2.28.(b), and because $\operatorname{Rg}(A^*\pm i)^\perp = \operatorname{Ker}(A\mp i) = \{0\}$ by Exercise 2.17.(e) and (ii), we deduce that $\operatorname{Rg}(A^* \pm i) = H$.

(iii)⇒(i) Due to symmetry of A we know that $A \subset A^*$, hence it suffices to prove that $D(A^*) \subset D(A)$. To this aim, let $y \in D(A^*)$. Due to surjectivity of $A-i$, there exists $x \in D(A) \subset D(A^*)$ such that $Ax - ix = A^*y - iy$. Since in particular $y - x \in D(A^*)$, we see that $(A^*-i)(y-x) = 0$. But by assumption $\operatorname{Rg}(A+i) = H$, hence $\operatorname{Ker}(A^*-i) = \operatorname{Rg}(A+i)^\perp = H^\perp = \{0\}$, hence $x = y$ and therefore $y \in D(A)$, i.e., $D(A^*) \subset D(A)$. ☐

Exercise 2.29. *Let Ξ be a σ-finite measure space and $q : \Xi \to \mathbb{C}$ a measurable function. Show that the associated* multiplication operator *defined by*

$$\begin{aligned} D(M_q) &:= \{u \in L^2(\Xi) : q \cdot u \in L^2(\Xi)\}, \\ M_q u &:= q \cdot u, \end{aligned}$$

is closed and densely defined in $L^2(\Xi)$. Moreover, show that the following assertions are equivalent.

(i) M_q is symmetric.
(ii) M_q is self-adjoint.
(iii) q is real-valued.

More generally, prove that for any measurable $q : \Xi \to \mathbb{C}$ the adjoint of M_q is given by $M_q^* = M_{\overline{q}}$, with maximal domain as above, where $\overline{q}$ is defined by $\overline{q}(x) := \overline{q(x)}$ a.e. Finally, show that it is a bounded operator if and only if $q \in L^\infty(\Xi)$ and in this case

$$\|M_q\|_{\mathcal{L}(L^2)} = \|q\|_\infty.$$

2.2. Spectral theory for unbounded operators

Unlike in the case of bounded linear operators, an unbounded linear operator may be invertible but fail to have an inverse that is bounded on H.

Definition 2.30. *Let X be a Banach space and A a closed linear operator on X. The* resolvent set $\rho(A)$ *is the set of all $\lambda \in \mathbb{C}$ such that $\lambda - A$ is invertible with inverse $R(\lambda, A)$ bounded on X. Its complement $\sigma(A)$ is called* spectrum *of A and the* resolvent *of A at $\lambda \in \rho(A)$ is the bounded linear operator $R(\lambda, A)$.*

Observe that if X is infinite dimensional, failure of $\lambda - A$ to be invertible does not imply that the equation $\lambda x = Ax$ has a nontrivial solution. This motivates the following.

Definition 2.31. *Let X be a Banach space and A a closed linear operator on X. The* point spectrum *of A is the set $\sigma_p(A)$ of all $\lambda \in \mathbb{C}$ such that $\lambda - A$ is not injective. Such λ are called* eigenvalues *of X.*

The approximate point spectrum *of A is the set $\sigma_{ap}(A)$ of all $\lambda \in \mathbb{C}$ such that $\lambda - A$ is not injective* or $\mathrm{Rg}(\lambda - A)$ *is not closed in X. Such λ are called* approximate eigenvalues *of X.*

Exercise 2.32. *Work out the details to prove the assertions of Example 2.29 and determine the spectrum of a multiplication operator M_q.* (Hint: Find out under which conditions the formal inverse $\frac{1}{q}$ of q is well-defined and gives rise to a bounded operator.)

Exercise 2.33. *Let X be a Banach space and A be a closed linear operator on X with nonempty resolvent set. Let $\lambda_0 \in \rho(A)$. Show that*

$$\sigma(R(\lambda_0, A)) = \left\{ \frac{1}{\lambda_0 - \lambda} : \lambda \in \sigma(A) \right\}.$$

Exercise 2.34. *Consider a σ-finite measure space (Ξ, μ) and define the* essential range *of a measurable function $q : \Xi \to \mathbb{C}$ by*

$$q_{ess}(\Xi) := \{z \in \mathbb{C} : \mu(\{x \in \Xi : |q(x) - z| < \epsilon) \neq 0 \text{ for all } \epsilon > 0\}.$$

Prove the following assertions concerning the multiplication operator M_q.

(1) M_q has bounded inverse if and only if $0 \notin q_{ess}(\Xi)$, and in this case $M_q^{-1} = M_{q^{-1}}$, where

$$q^{-1}(x) := \begin{cases} \frac{1}{q(x)} & \text{if } q(x) \neq 0, \\ 0 & \text{if } q(x) = 0. \end{cases}$$

(2) $\sigma(M_q) = q_{ess}(\Xi)$.

Exercise 2.35. *Let X be a Banach space and A a closed linear operator on X. Show that for each $\lambda \in \mathbb{C}$ the following assertion are equivalent:*

(i) $\lambda \in \sigma_{ap}(A)$.

(ii) There exists a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ such that $\|x_n\|_X = 1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \|\lambda x_n - Ax_n\|_X = 0$.

Exercise 2.36. *Let $H = \ell^2(\mathbb{Z})$ and $A(x_n)_{n\in\mathbb{Z}} := (x_{n+1})_{n\in\mathbb{Z}}$. Show that A has no eigenvalues but each λ with $|\lambda| = 1$ is an approximate eigenvalue.*

In the special case of bounded operators, all spectral notions and results can be formulated in a natural way in the more general context of C^*-algebras with unity, since this already permits to define the notion of invertibility. We neglect the obvious details.

Exercise 2.37. *Let H be a Hilbert space and A a closed, symmetric operator on H. Define its* Cayley transformed *$C(A)$ by*

$$C(A) := (A - i)(A + i)^{-1} = (A - i)R(i, -A).$$

Show that $C(A)$ is an isometry between $\mathrm{Rg}(A+i)$ and $\mathrm{Rg}(A-i)$, and that moreover A is self-adjoint if and only if $C(A)$ is unitary. Conclude that every self-adjoint operator is the Cayley transformed of a uniquely determined unitary operator.

Proposition 2.38. *Let $\mathcal{A}$ be a Banach algebra with unity I. Then the following assertions hold.*

(1) If $x \in \mathcal{A}$ is such that $\|x\| < 1$, then $\operatorname{Id} - x$ *is invertible and the* Neumann series

$$(\operatorname{Id} - x)^{-1} = \sum_{n=0}^{\infty} x^n$$

holds.

(2) The spectrum of each $x \in \mathcal{A}$ is a compact set contained in

$$\{z \in \mathbb{C} : |z| \leq \|x\|\}.$$

Proof. (1) Due to submultiplicativity of the norm and by convergence of the geometric series (in $\mathbb{R}$), one sees that

$$\sum_{n=0}^{\infty} \|x^n\| \leq \sum_{n=0}^{\infty} \|x\|^n = \frac{1}{1 - \|x\|}. \tag{2.3}$$

Since $\mathcal{A}$ is a Banach space, each absolutely convergent series is also convergent. This implies convergence of $\sum_{n=0}^{m} x^n =: S_m$ towards $\sum_{n=0}^{\infty} x^n =: S \in \mathcal{A}$ as $m \to \infty$. Hence,

$$S_m(\operatorname{Id} - x) = S_m - S_m x = S_m - x \sum_{n=0}^{m} x^n = S_m - \sum_{n=1}^{m+1} x^n = S_m - (S_{m+1} - I) = \operatorname{Id} - x^{m+1}$$

and accordingly

$$S(\operatorname{Id} - x) = \lim_{m \to \infty} S_m(\operatorname{Id} - x) = \operatorname{Id} - \lim_{m \to \infty} x^{m+1} = I,$$

since $\lim_{m\to\infty} \|x^{m+1}\| \leq \lim_{m\to\infty} \|x\|^{m+1} = 0$ because $\|x\| < 1$. One proves likewise that $(\operatorname{Id} - x)S = I$. We conclude that $(\operatorname{Id} - x)^{-1} = S = \sum_{n=0}^{\infty} x^n$.

(2) In order to see the boundedness of $\sigma(x)$, observe that for $|\lambda| \geq \|x\|$ one has

$$\|\operatorname{Id} - (\operatorname{Id} - \lambda^{-1} x)\| < 1,$$

hence by (1) $\operatorname{Id} - \lambda^{-1} x$ and therefore $\lambda - x$ are invertible, i.e., $\lambda \in \rho(x)$.

Closedness of the spectrum of an element of a C^*-algebra can be proved as in the case of an unbounded operator, see Proposition below. ☐

Proposition 2.39. *Let X be a Banach space and A a closed linear operator on X. Then $\sigma(A)$ is a closed set. In particular, for all $\mu \in \rho(A)$ and all λ sufficiently close to μ one has the power series expansion*

$$R(\lambda, A) = \sum_{n=0}^{\infty} (\mu - \lambda)^n R(\mu, A)^{n+1}. \tag{2.4}$$

Moreover, $R(\cdot, A)$ is a holomorphic mapping $\rho(A) \to \mathcal{L}(X)$ with

$$\frac{d^n}{d\lambda^n} R(\lambda, A) = (-1)^n n! R(\lambda, A)^{n+1}, \qquad \mu \in \rho(A).$$

Finally, the resolvent equation

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A), \qquad \lambda, \mu \in \rho(A)$$

holds.

Proof. We prove that the resolvent set is open. To see this, take $\mu \in \rho(A)$ and observe that for all $\lambda \in \mathbb{C}$ one has

$$\lambda - A = \lambda - \mu + \mu - A = (\operatorname{Id} + (\lambda - \mu) R(\mu, A))(\mu - A),$$

whence

$$\lambda \in \rho(A) \qquad \text{if and only if} \qquad 1 \in \rho((\mu - \lambda) R(\mu, A)) :$$

but by Proposition 2.38 this is the case if and only if $\|(\mu-\lambda)R(\mu,A)\| < 1$, i.e., if and only if $\|R(\mu,A)\|^{-1} > |\mu-\lambda|$. In other words, all λ in a small neighbourhood of μ belong to $\rho(A)$, too.

The claimed expression for the power series expansion of the resolvent follows from the Neumann series applied to

$$(\lambda - A)^{-1} = (\mu - A)^{-1}(\mathrm{Id} + (\lambda-\mu)R(\mu,A))^{-1}.$$

The same expression also shows that $R(\cdot,A)$ is complex analytic, the Taylor series being given by (2.4). It follows in particular that the n^{th} coefficient of the Taylor series

$$\sum_{n=0}^{\infty}(\mu-\lambda)^n R(\mu,A)^{n+1} = \sum_{n=0}^{\infty} R(\mu,A)^{n+1}(-1)^n(\lambda-\mu)^n,$$

i.e.

$$\frac{1}{n!}\frac{d^n}{d\lambda^n}R(\cdot,A)(\mu) = R(\mu,A)^{n+1}(-1)^n.$$

Finally, in order to prove the resolvent equation observe that

$$\lambda R(\lambda,A) - AR(\lambda,A) = \mathrm{Id},$$

hence

$$\big(\lambda R(\lambda,A) - AR(\lambda,A)\big)R(\mu,A) = R(\mu,A),$$

and similarly

$$R(\lambda,A)\big(\mu R(\mu,A) - AR(\mu,A)\big) = R(\lambda,A),$$

whence

$$\begin{aligned} R(\lambda,A) - R(\mu,A) &= R(\lambda,A)\big(\mu R(\mu,A) - AR(\mu,A)\big) - \big(\lambda R(\lambda,A) - AR(\lambda,A)\big)R(\mu,A) \\ &= \mu R(\lambda,A)R(\mu,A) - AR(\lambda,A)R(\mu,A) - \lambda R(\lambda,A)R(\mu,A) + AR(\lambda,A)R(\mu,A) \\ &= \mu R(\lambda,A)R(\mu,A) - \lambda R(\lambda,A)R(\mu,A). \end{aligned}$$

This concludes the proof. □

Exercise 2.40. *Let X be a Banach space and $B \in \mathcal{L}(X)$. Show that for all $\lambda \in \rho(A)$ one has*

$$\lambda \in \rho(A+B) \qquad \textit{if and only if} \qquad 1 \in \rho(R(\lambda,A)B) \qquad \textit{if and only if} 1 \in \rho(BR(\lambda,A)),$$

and in this case

$$R(\lambda,A+B) = R(1,R(\lambda,A)B)R(\lambda,A) = R(\lambda,A)R(1,BR(\lambda,A)).$$

Exercise 2.41. *Let H be a Hilbert space and A a linear, closed, symmetric operator on H.*

Show that then either the spectrum $\sigma(A)$ of A is contained in $\mathbb{R}$, or one of the following cases hold.

- $\sigma(A) = \mathbb{C}$,
- $\sigma(A) = \{\lambda \in \mathbb{C} : \operatorname{Im}\lambda \geq 0\}$,
- $\sigma(A) = \{\lambda \in \mathbb{C} : \operatorname{Im}\lambda \leq 0\}$.

Prove that furthermore $\sigma(A) \subset \mathbb{R}$ if and only if A is self-adjoint.

Conclude that any linear, closed, symmetric operator whose resolvent set is not disjoint from $\mathbb{R}$ is self-adjoint.

Exercise 2.42. *Let H be a Hilbert space and $A \in \mathcal{L}(H)$ such that*

$$|(Ax|x)_H| \geq \alpha\|x\|^2 \qquad \textit{for all } x \in H$$

for some $\alpha > 0$.

(1) Show that $\|Ax\| \geq \alpha\|x\|$ for all $x \in H$. (Hint: use the Cauchy–Schwarz inequality.)
(2) Show that A is injective and $A(H)$ is closed.
(3) Show that $\overline{A(H)} = H$. (Hint: Show that $(A(H))^\perp = \{0\}$.)
(4) Conclude that A is an isomorphism.

Remark 2.43. *Observe that the resolvent*

$$(\lambda - x)^{-1} = \frac{1}{\lambda}(1 - \frac{x}{\lambda})^{-1} = \sum_{n=0}^{\infty} \frac{x^n}{\lambda^{n+1}}$$

exists whenever $|\lambda| > \|x\|$*, and the general formula for the convergence radius of a power series yields*

$$\sup\{|\lambda| \in \mathbb{C} : \lambda \in \sigma(A)\} = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

Hence the spectral radius

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

which can be defined for all closed operators A*, is in fact finite and agrees with*

$$\lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$$

for all bounded operators and, more generally, for all elements x *of a* C^**-algebra.*

Exercise 2.44. *Let* X *be a Banach space and* $A : X \supset D(A) \to X$ *be a closed operator. Let* $(\lambda_n)_{n\in\mathbb{N}}$ *be a convergent sequence with* $\lim_{n\to\infty} \lambda_n =: \lambda_0$*. Show that* $\lambda_0 \in \sigma(A)$ *if and only if* $\lim_{n\to\infty} \|R(\lambda_n, A)\| = \infty$*.*

2.3. The Spectral Theorem for bounded self-adjoint operators

We are finally in the position to prove the first main result of this chapter.

Theorem 2.45 (Spectral Theorem for bounded self-adjoint operators)**.** *Let* A *be a bounded self-adjoint operator on a separable Hilbert space* H*. Then* A *is unitarily equivalent to a (necessarily bounded) multiplication operator* M_q *on* $L^2(\Xi)$ *for some finite measure space* Ξ *and some measurable function* $q : \Xi \to \mathbb{R}$*.*

This beautiful and natural formulation of the Spectral Theorem is one of the main legacies of Paul Halmos. The more usual version, which is plagued by obscure notions and technical details both in its statement and its proof (see e.g. the discussion in [**11**]), is more common in the mathematical literature and is still, 50 years after Halmos' article, the most popular one.

As pointed out in [**5**], the proof of the Spectral Theorem is elementary as soon as a few slightly more advanced results are available. We assume that two of them (the Riesz representation theorem for linear functionals on $C(K;\mathbb{R})$, K a compact set, mapping positive-valued functions into positive numbers; and the Stone–Weierstraß Theorem on approximation of continuous functions) are already known from a course in real analysis and only recall the spectral mapping theorem along with the following.

Lemma 2.46. *Let* $\mathcal{A}$ *be a* C^**-algebra with unity* I*. Then the spectral radius of each* $x \in \mathcal{A}$ *agrees with its norm whenever* x *is self-adjoint.*

Observe in particular that this implies that the spectrum of a self-adjoint element of a C^*-algebra is always non-empty.

Proof. We have already seen in Remark 2.43 that $r(x) \le \|x\|$ for all $x \in \mathcal{A}$. For $x \in \mathcal{A}$ self-adjoint $(x^*x)^m = x^{*m}x^m$, and therefore

$$\|x\|^{2^n} = \|x^*x\|^{2^{n-1}} = \|(x^*x)^{2^{n-1}}\|$$

by definition of C^*-algebra. Thus,

$$\|x\|^{2^n} = \|(x^*)^{2^{n-1}}x^{2^{n-1}}\| = \|(x^{2^{n-1}})^*x^{2^{n-1}}\| = \|x^{2^{n-1}}\|^2.$$

We conclude that

$$r(x) = \lim_{n\to\infty} \|x^{2^{n-1}}\|^{2^{1-n}} = \lim_{n\to\infty} \|x\|^{2^{n-1}2^{1-n}} = \|x\|.$$

This concludes the proof. ☐

Theorem 2.47 (Spectral mapping theorem, SMT)**.** *Let $\mathcal{A}$ be a Banach algebra with unity I. Let $x \in \mathcal{A}$ and let π be a polynomial with complex coefficients. Then*

$$\sigma(\pi(x)) = \pi(\sigma(x)) := \{\pi(\lambda) \in \mathbb{C} : \lambda \in \sigma(x)\}.$$

Proof. Let us first show that $\sigma(\pi(x)) \subset \pi(\sigma(x))$. Take $\mu \in \sigma(\pi(x))$ and because

$$\pi(z) - \mu = \alpha \prod_{i=1}^{n} (z - \lambda_i), \qquad z \in \mathbb{C},$$

for some $\alpha \in \mathbb{C}$ (where n is the degree of π and $\lambda_1, \ldots, \lambda_n$ its zeroes), one also has

$$\pi(x) - \mu = \alpha \prod_{i=1}^{n} (x - \lambda_i).$$

Hence, if all λ_i would belong to the resolvent set of x, then $\pi(x) - \mu$ would be invertible, a contradiction. Accordingly, at least one λ_j has to belong to the spectrum of x, and then $\pi(\lambda_j) - \mu = 0$, i.e., $\mu \in \{\pi(\lambda) \in \mathbb{C} : \lambda \in \sigma(A)\}$.

Conversely, let $\lambda \in \sigma(x)$ and take $\pi(\lambda) \in \mathbb{C}$. Take a polynomial $\tilde{\pi}$ with complex coefficients such that the polynomial $\pi(\cdot) - \pi(\lambda)$ agrees with $\tilde{\pi}(\cdot)(\cdot - \lambda)$ – this is surely possible, since λ is a zero if $\pi(\cdot) - \pi(\lambda)$. Accordingly,

$$\pi(x) - \pi(\lambda) = \tilde{\pi}(x)(x - \lambda).$$

We want to show that $\pi(\lambda) \in \sigma(\pi(x))$. In fact, if $\pi(x) - \pi(\lambda)$ would be invertible, then

$$\mathrm{Id} = \Big(\pi(x) - \pi(\lambda)\Big)^{-1} \tilde{\pi}(x)(x - \lambda) = (x - \lambda)\Big(\pi(x) - \pi(\lambda)\Big)^{-1} \tilde{\pi}(x),$$

hence $x - \lambda$ would be invertible, too – a contradiction. □

Definition 2.48. *Let A be a self-adjoint operator on a Hilbert space H. An element $\mathbf{x} \in H$ is called a* cyclic vector *for A if*

$$\{\pi(A)\mathbf{x} \in H : \pi \textit{ is a polynomial with complex coefficients}\}$$

is dense in H, i.e., if the space spanned by $\{A^k \mathbf{x} : k \in \mathbb{N}\}$ is dense in H.

Proof of the Spectral Theorem for bounded self-adjoint operators. The proof consists of two parts: we first prove the assertion in the case that A is self-adjoint and H contains a cyclic vector for A, then extend it to the case of a self-adjoint operator on a general Hilbert space.

(1) First of all, assume H to contain a cyclic vector $\mathbf{x}$ for A. Define a mapping L by setting

$$L(\pi) := (\pi(A)\mathbf{x}|\mathbf{x}), \qquad \pi \text{ polynomial with real coefficients},$$

and extending by density to a bounded linear functional $L : C(\sigma(A)) \to \mathbb{R}$ (linearity of L is clear and boundedness follows by observing that

$$\begin{aligned} |L(\pi)| &= |(\pi(A)\mathbf{x}|\mathbf{x})| \\ &\leq \|\pi(A)\|\|\mathbf{x}\|^2 \\ &= r(\pi(A))\|\mathbf{x}\|^2 \\ &= \sup\{|\lambda| \in \mathbb{R} : \lambda \in \sigma(\pi(A))\}\|\mathbf{x}\|^2 \\ &= \sup\{|\pi(\lambda)| \in \mathbb{R} : \lambda \in \sigma(A)\}\|\mathbf{x}\|^2, \end{aligned}$$

where the last identity follows from the Spectral Mapping Theorem). Now we want to apply the Riesz Representation Theorem: to this aim we still have to show positivity of L. In fact, positivity of $L(\pi)$ is clear if π is the square of a polynomial $\tilde{\pi}$ with real coefficients, since then $\pi(A) = \tilde{\pi}(A)^2$ is self-adjoint and therefore

$$L(\pi) = (\tilde{\pi}(A)^2\mathbf{x}|\mathbf{x}) = (\tilde{\pi}(A)\mathbf{x}|\tilde{\pi}(A)\mathbf{x}) = \|\tilde{\pi}(A)\mathbf{x}\|^2.$$

If instead f is a general positive-valued function, then we approximate its square root uniformly by a sequence of polynomials $(\pi_n)_{n\in\mathbb{N}}$ with real coefficients (by the Stone–Weierstraß Theorem) and obtain that also

$$L(f) = \lim L(\pi_n^2) \geq 0.$$

We can finally apply the Riesz Representation Theorem and deduce that L is represented by a finite measure, i.e., there exists a Borel measure μ on $\Xi := \sigma(A)$ such that

$$L(f) = \int_\Xi f d\mu.$$

We now introduce a new operator $U : L^2(\Xi) \to H$, letting first

$$U : L^2(\Xi) \ni \tilde{\pi} \mapsto \tilde{\pi}(A)\mathbf{x} \in H$$

for all polynomials $\tilde{\pi}$ with complex coefficients and then extending by density. We can now use the fact that, by self-adjointness of A, $\tilde{\pi}(A), \tilde{\pi}(A)^*$ and $\tilde{\pi}(A)^*\tilde{\pi}(A)$ is a polynomial in A. Since

$$\|\tilde{\pi}\|_2^2 = \int_\Xi \overline{\tilde{\pi}}\tilde{\pi} d\mu = L(\overline{\tilde{\pi}}\tilde{\pi}) = (\tilde{\pi}(A)^*\tilde{\pi}(A)\mathbf{x}|\mathbf{x}) = (\tilde{\pi}(A)\mathbf{x}|\tilde{\pi}(A)\mathbf{x}) = \|\tilde{\pi}(A)\mathbf{x}\|^2 = \|U\tilde{\pi}\|^2,$$

we see that U is an isometry when restricted to the dense subspace of polynomials with complex coefficients, hence by density it is an isometry from $L^2(\Xi)$ to H. Finally, by assumption $\mathbf{x}$ is a cyclic vector for A, hence

$$\operatorname{Rg} U = \{\tilde{\pi}(A)\mathbf{x} \in H : \tilde{\pi} \text{ is a polynomial with complex coefficients } \}$$

is dense in H; but an isometry has closed range, hence necessarily $\operatorname{Rg} U = H$.

We finally show that $U^{-1}AU : H \to H$ is a multiplication operator: Let ϕ be the identity function on $\sigma(A)$ and for each polynomial with complex coefficients $\tilde{\pi}$ set

$$\hat{\pi} := \phi\tilde{\pi} : \lambda \mapsto \lambda\tilde{\pi}(\lambda).$$

We conclude that

$$U^{-1}AU\tilde{\pi} = U^{-1}A\tilde{\pi}(A)\mathbf{x} = U^{-1}\hat{\pi}(A)\mathbf{x} = U^{-1}U\hat{\pi} = \hat{\pi} = \phi\tilde{\pi}$$

for all polynomials $\tilde{\pi}$ with complex coefficients, so that

$$U^{-1}AU = M_\phi$$

by density.

(2) In the general case of a Hilbert space without a cyclic vector for A, we first want to prove that

$$H = \bigoplus_n H_n, \tag{2.5}$$

where the direct sum is over a countable (possibly finite) family of *cyclic subspaces*, i.e., of subspaces H_n of H that are invariant under A and such that each H_n has a cyclic vector for $A_n := A_{|H_n}$.

First of all, define $\mathcal{H}$ as the set of countable (possibly finite) orthogonal families of cyclic subspaces of H. Clearly, $\mathcal{H}$ is inductively ordered by $\subset$, hence if $\mathcal{K} \subset \mathcal{H}$ is a chain[2], then $\mathcal{K}$ has an upper bound $\tilde{K} := \bigcup_{K\in\mathcal{K}} K$. By Zorn's Lemma, $\mathcal{H}$ has (at least) one maximal element $\hat{K}$, i.e., a countable (possibly finite) orthogonal family (H_n) of cyclic subspaces of H. It remains to prove that

$$H = \bigoplus_n H_n.$$

If this would not be the case, there would exist $x \in H$ such that $x \in H_n^\perp$ for all n, and accordingly the closure K of the span of $\{A^k x : k \in \mathbb{N}\}$ would be orthogonal to all H_n, hence $K \oplus \bigoplus_n H_n$ would be a maximal element larger than $\hat{K}$, a contradiction. Summing up, (2.5) is proved, hence $\bigcup_n H_n$ is total in H and in particular for

[2] That is, $K_1, K_2 \in \mathcal{K}$ implies that $K_1 \subset K_2$ or $K_2 \subset K_1$.

all $x \in H$ there exists a unique sequence (x_n) such that $x = \sum_n x_n$ and moreover $(x|y)_H = \sum_n (x_n|y_n)$ for all $x, y \in H$.

Hence, by (1) we deduce that the spectral theorem holds on each H_n, i.e., there exist a Borel measure μ_n on $\Xi_n := \sigma(A_n)$, where $A_n := A_{|H_n}$, a measurable function $\phi_n : \Xi_n \to \mathbb{R}$ and a unitary operator $U_n : L^2(\Xi_n) \to H_n$ such that

$$U_n^{-1} A_n U_n = M_{\phi_n},$$

where each M_{ϕ_n} acts by

$$(M_{\phi_n} f)(\xi) = \xi f(\xi), \qquad \xi \in \Xi_n.$$

Now, it suffices to consider $(\Xi, \mu) := \bigoplus_n (\Xi_n, \mu_n)$, i.e.,

$$\Xi := \bigcup_n (\Xi_n \times \{n\}) = \bigcup_n (\sigma(A_n) \times \{n\}) \subset \mathbb{R}^2.$$

Then (Ξ, μ) is a σ-finite measure space defined as follows: The Borel σ-algebra on Ξ consists of sets of the form

$$\bigcup_n O_n \times \{n\}, \qquad O_n \text{ element of the Borel } \sigma\text{-algebra on } \Xi_n,$$

and μ acts on such sets by

$$\mu(\bigcup_n O_n \times \{n\}) := \sum_n \mu_n(O_n).$$

We can finally show the spectral theorem fore general H. Define a function $\phi : \Xi \to \mathbb{R}$ by $\phi(\xi, n) = \xi$. Then ϕ is measurable and bounded and

$$(M_\phi f)(\xi, n) = \phi(\xi, n) f(\xi, n) = \xi f(\xi, n).$$

Moreover, consider $U := \mathrm{diag}(U_n) : L^2(\Xi) \to H$ defined by

$$(U^{-1}x)(\xi, n) = \sum_m (U_m^{-1} x_m)(\xi).$$

This operator is unitary, since

$$(U^{-1}x|U^{-1}y) = \sum_n \int_{\Xi_n} (U_n^{-1} y_n)(\xi) \overline{(U_n^{-1} x_n)(\xi)} d\mu_n(\xi) = \sum_n (x_n|y_n)_{H_n} = (x|y).$$

It remains to prove that

$$U^{-1} A U = M_\phi :$$

in fact,

$$(U^{-1}Af)(\xi, n) = (U_n^{-1} A x_n)(\xi) = \xi (U^{-1} x_n)(\xi) = \phi(\xi, n)(U^{-1}x)(\xi, n), \qquad \text{for all } x = \sum_n x_n \in H = \bigoplus_n H_n.$$

This concludes the proof of the claim in the self-adjoint case. □

Remark 2.49. *The above theorem can also be extended to self-adjoint operators on non-separable Hilbert spaces, but part (2) of the proof is more delicate, since one also needs an argument based on transfinite induction. Since essentially all Hilbert space that naturally appear in the theory of evolution equations are separable, we omit this extension.*

Exercise 2.50. *Use the spectral theorem to prove the following assertion: Let H be a Hilbert space and A be a self-adjoint operator on H. If $\sigma(A) = \{0, 1\}$, then A is an orthogonal projector.*

2.4. The Spectral Theorem for bounded normal operators

Our aim in this section is to prove that the Spectral Theorem for bounded self-adjoint operators prevails for operators that are merely normal. The proof of this generalisation requires a number of technical lemmata.

Lemma 2.51. *Let H be a Hilbert space and A a bounded linear operator on H. Then the following assertions hold.*

(a) Let $\lambda \in \mathbb{C}$. If $\lambda \in \sigma(A) \setminus \sigma_{ap}(A)$, then $\overline{\lambda} \in \sigma_p(A^*)$.
(b) If A is normal, then $\sigma_p(A^*) = \{\overline{\lambda} \in \mathbb{C} : \lambda \in \sigma_p(A)\}$.
(c) If A is normal, then $\sigma(A) = \sigma_{ap}(A)$.

Proof. (a) Assume without loss of generality that $0 = \lambda \in \sigma(A) \setminus \sigma_{ap}(A)$, i.e., assume that A is injective and its range is closed. It follows from Exercise 2.17.(f) applied to A^* that $\operatorname{Rg} A = \overline{\operatorname{Rg} A} = \operatorname{Ker}(A^*)$. Hence, injectivity of A^* would imply that A is surjective, a contradiction to the fact that $A \notin \sigma(A)$.

(b) It suffices to observe that for all $\lambda \in \mathbb{C}$ one has $(T - \overline{\lambda})^* = A^* - \lambda$, hence $\|A - \overline{\lambda}\| = \|A^* - \lambda\|$.

(c) By definition, $\sigma(A) \supset \sigma_{ap}(A)$. Let now $\lambda \in \sigma(A)$ and assume that $\lambda \notin \sigma_{ap}(A)$. Then by (a) $\overline{\lambda} \in \sigma_p(A^*)$, hence (b) $\lambda \in \sigma_p(A) \subset \sigma_{ap}(A)$, a contradiction. □

Lemma 2.52. *Let H be a Hilbert space and A a bounded linear operator on H. Then the following assertions hold.*

(a) Let $\lambda \in \mathbb{C}$ and $(x_n)_{n\in\mathbb{N}} \subset H$. If $\lim_{n\to\infty}(A - \lambda)x_n = 0$, then $\lim_{n\to\infty}(A^k - \lambda^k)x_n = 0$ for all $k \in \mathbb{N}$.
(b) Let $B \in \mathcal{L}(H)$, $\lambda, \mu \in \mathbb{C}$, and $(x_n)_{n\in\mathbb{N}} \subset H$. If $\lim_{n\to\infty}(A - \lambda)x_n = 0$ and $\lim_{n\to\infty}(B - \mu)x_n = 0$, then $\lim_{n\to\infty}(A^k B^\ell - \lambda^k \mu^\ell)x_n = 0$ for all $k, \ell \in \mathbb{N}$.

Proof. (a) The assertion is a direct consequence of the factorisation $(A^k - \lambda^k) = (A^{k-1} + \lambda A^k + \cdots + \lambda^{k-1})(A - \lambda)$.

(b) The assertion follows factorising (how?). □

Lemma 2.53. *Let H be a Hilbert space and A be a bounded normal operator on H such that $0 \in \sigma(A)$. Let $\epsilon > 0$. Then there exists a closed subspace $E \subset H$, $E \neq \{0\}$, such that*

- *$\|S_{|E}\| \leq \epsilon$ and*
- *E is left invariant under each $B \in \mathcal{L}(H)$ that commutes with both A and A^*.*

Proof. First of all, observe that the operator A^*A, which is self-adjoint and by assumption agrees with A^*A, cannot be bijective; since otherwise $Ax = 0$ and hence $A^*Ax = 0$ would imply $x = 0$ for all $x \in H$, i.e., A would be injective; and similarly for all $x \in H$ there would exist $y \in H$ with $AA^*y = A(A^*y) = x$, i.e., A would be surjective.

The main idea of the proof is that by the spectral theorem for self-adjoint operators A^*A is unitarily equivalent to a multiplication operator M_q on a space $L^2(\Xi)$, hence we can consider $\phi(A^*A)$ for any bounded continuous function ϕ, which is unitarily equivalent to the multiplication operator $M_{\phi\circ q}$.

Let now $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions defined by

- $f(x) = f(-x)$ for all $x \in \mathbb{R}$,
- $f(x) := 1$ for all $x \in [0, \frac{\epsilon^2}{2}]$,
- $f(x) := \frac{2}{\epsilon^2}(\frac{\epsilon^2}{2} - x) + 1$ for all $x \in [\frac{\epsilon^2}{2}, \epsilon^2]$,
- $f(x) := 0$ for all $x \in [\epsilon^2, \infty)$ (and hence $\operatorname{supp}(f) \subset [-\epsilon^2, \epsilon^2]$),
- $g(x) := f(2x)$ for all $x \in \mathbb{R}$.

In particular, observe that $\operatorname{supp}(g) \subset [-\frac{\epsilon^2}{2}, \frac{\epsilon^2}{2}] = f^{-1}(\{1\})$, hence $(1-f)(x)g(x) = 0$ for all $x \in \mathbb{R}$ and accordingly

$$(\operatorname{Id} - f(A^*A))g(A^*A) = 0. \tag{2.6}$$

Since A^*A is self-adjoint, we can plug it into f, using the Spectral Theorem 2.45, and define $E := \operatorname{Ker}(\operatorname{Id} - f(A^*A))$. To begin with, we check that E is non-trivial. By (2.6) $\operatorname{Rg} g(A^*A)$ is contained in $\operatorname{Ker}(\operatorname{Id} - f(A^*A))$, i.e., in E, hence it suffices to prove that $\operatorname{Rg} g(A^*A)$ is non-trivial. For all spectral arguments, and in particular for all our purposes, we can identify A^*A with M_q: since $\mu\{\xi \in \Xi : |q(\xi)| < \frac{\epsilon^2}{4}\} = 0$ would imply the bijectivity of A^*A, we necessarily have $\mu(\{\xi \in \Xi : |q(\xi)| < \frac{\epsilon^2}{4}\}) > 0$. Since however

$$\{\xi \in \Xi : |q(\xi)| < \frac{\epsilon^2}{4}\} \subset \{\xi \in \Xi : \phi(q(\xi)) = 1\}$$

we deduce that also

$$\mu(\{\xi \in \Xi : \phi(q(\xi)) = 1\}) > 0,$$

hence $\phi \circ q \not\equiv 0$, and in particular $\operatorname{Rg} A^*A \equiv \operatorname{Rg} M_{\phi\circ q} \neq \{0\}$, as we wanted to show.

Now, observe that for all $x \in E$ with $\|x\| \leq 1$, hence such that $f(A^*A)x = x$, one has

$$\begin{aligned}
\|Ax\|^2 &= (Ax|Ax) = (A^*Ax|x) \\
&= (A^*Af(A^*A)x|x) \\
&\leq \|A^*Af(A^*A)x\|\|x\| \\
&\leq \|(id \cdot f)(A^*A)x\| \\
&\overset{\text{Lemma 2.46}}{=} \sup_{\lambda\in\mathbb{R}}\{|(id \cdot f)(\lambda)| : \lambda \in \sigma(A^*A)\} \\
&\overset{\text{SMT}}{=} \sup_{\lambda\in\sigma(A^*A)} |\lambda f(\lambda)| \\
&\leq \epsilon^2,
\end{aligned}$$

by construction of f.

Finally, let B be a bounded linear operator that commutes with both A and A^*. Then B also commutes with A^*A and with all of its powers, hence in particular (by the Stone–Weierstraß Theorem) with $f(A^*A)$. Let now $x \in E$. Then $(\operatorname{Id} - f(A))x = 0$ and therefore

$$(\operatorname{Id} - f(A^*A))Bx = B(\operatorname{Id} - f(A^*A))x = 0,$$

i.e., $Bx \in E$, too. Similarly, B^* leaves E invariant, which means that for all $x \in E^\perp$ and all $y \in E$

$$(Bx|y) = (x|B^*y) = 0,$$

i.e., B leaves $E^\perp$ invariant. □

We are finally in the position to prove a kind of Spectral Mapping Theorem for polynomial of two variables, plus some miscellaneous properties of normal operators that will be essential in the proof of the Spectral Theorem.

Lemma 2.54. *Let H be a Hilbert space and A a bounded normal operator on H. Let moreover $\pi : \mathbb{C}^2 \to \mathbb{C}$ be a polynomial. Then the following assertions hold.*

(a) Let $(x_n)_{n\in\mathbb{N}} \subset H$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$. Let $\lambda \in \mathbb{C}$. If $\lim_{n\to\infty}(A - \lambda)x_n = 0$, then also

$$\lim_{n\to\infty} (\pi(A, A^*) - \pi(\lambda, \lambda^*))x_n = 0.$$

(b) $\{\pi(\mu, \overline{\mu}) \in \mathbb{C} : \mu \in \sigma(A)\} = \sigma(\pi(A, A^*))$.

(c) $\|A\| = \sup\{|\lambda| : \lambda \in \sigma(A)\}$.

(d) $\|\pi(A, A^*)\| = \sup\{|\pi(z, \overline{z})| : z \in \sigma(A)\}$.

(e) Let $\pi' : \mathbb{C}^2 \to \mathbb{C}$ be a further polynomial. If $\pi(z, \overline{z}) = \pi'(z, \overline{z})$ for all $z \in \sigma(A)$, then $\pi(A, A^*) = \pi'(A, A^*)$.

Proof. (a) By assumption, $\lim_{n\to\infty} \|(A^* - \overline{\lambda})x_n\| = \lim_{n\to\infty} \|(A - \lambda)x_n\| = 0$. Then the assertion follows by Lemma 2.52.(b) if p is a monomial, i.e., $p(z_1, z_2) = z_1^k z_2^j$ for some $j, k \in \mathbb{N}$ and all $z_1, z_2 \in \mathbb{C}$; and by linearity in the general case.

(b) Let us first prove that $\{\pi(\mu, \overline{\mu}) \in \mathbb{C} : \mu \in \sigma(A)\} \subset \sigma(\pi(A, A^*))$ holds. Let $\lambda \in \{\pi(\mu, \overline{\mu}) \in \mathbb{C} : \mu \in \sigma(A) = \sigma_{ap}(A)\}$. Take $\mu \in \sigma_{ap}(A)$ such that $\lambda = \pi(\mu, \overline{\mu})$. In order to prove that $\lambda \in \sigma(\pi(A, A^*))$, i.e., that

$\pi(\mu,\overline{\mu}) \in \sigma_{ap}(\pi(A,A^*))$ (why?), we need by Exercise 2.35 to show that there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset H$ with $\|x_n\| = 1$ such that

$$\lim_{n\to\infty} \|\pi(\mu,\overline{\mu})x_n - \pi(A,A^*)x_n\| = 0. \tag{2.7}$$

Because $\mu \in \sigma_{ap}(A)$, again by Exercise 2.35 there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset H$ with $\|x_n\| = 1$ such that $\|\lambda x_n - Ax_n\| = 0$, and we conclude by (a) that the claimed inclusion holds.

Conversely, let $\mu \in \sigma(\pi(A,A^*))$. Define $B = \pi(A,A^*) - \mu$, which is a normal operator with $0 \in \sigma(B)$. Moreover, B commutes with both A and A^*. Then by Lemma 2.53 for all $n \in \mathbb{N}$ there exists a closed subspace $E_n \neq \{0\}$ such that $\|S_{|E_n}\| \leq \epsilon_n$ and such that both E_n and $E_n^\perp$ are left invariant under A. In particular, the restriction $A_{|E_n}$ is a bounded normal operator on E_n, since

$$(\cdot|\cdot)_H = (\cdot|\cdot)_{E_n} + (\cdot|\cdot)_{E_n^\perp},$$

and hence $(A_{|E_n})^* = (A^*)_{|E_n}$. Take now some $\lambda_n \in \sigma(A_{|E_n})$ which, by Lemma 2.51.(c), is also an approximate eigenvalue. Hence, by Exercise 2.35 there exists a sequence of vectors $(x_n)_{n\in\mathbb{N}}$ with $x_n \in E_n$ and $\|x_n\| = 1$ and such that $\|(A-\lambda_n)x_n\| < \frac{1}{n}$. Because $\lambda_n \in \sigma(A_{|E_n})$ and hence $|\lambda_n| \leq \|A_{|E_n}\| \leq \|A\|$ for all $n \in \mathbb{N}$, the sequence $(\lambda_n)_{n\in\mathbb{N}}$ is bounded and hence convergent (up to taking a subsequence) towards some λ, which then belongs to $\sigma_{ap}(A)$, too, since for the same sequence $(x_n)_{n\in\mathbb{N}}$ one has

$$\|(A-\lambda)x_n\| \leq \|(A-\lambda_n)x_n\| + \|(\lambda_n - \lambda)x_n\| \leq \frac{1}{n} + |\lambda - \lambda_n| \to 0.$$

By (a) one then has

$$\lim_{n\to\infty} (\pi(A,A^*) - \pi(\lambda,\lambda^*))x_n = 0.$$

Since moreover

$$\|(\pi(A,A^*) - \mu)x_n\| < \frac{1}{n} \qquad \text{for all } n \in \mathbb{N},$$

it follows that

$$\lim_{n\to\infty} (\mu - \pi(\lambda,\lambda^*))x_n = (\mu - \pi(\lambda,\lambda^*)) \lim_{n\to\infty} x_n = 0,$$

whence $\mu = \pi(\lambda,\lambda^*)$, because $\|x_n\| = 1$ for all n.

(c) Set $\pi(z_1,z_2) := z_1 z_2$ and therefore $\pi(A,A^*) = AA^* = A^*A$. Then

$$\begin{aligned}(\sup\{|\lambda| : \lambda \in \sigma(A)\})^2 &= \sup\{\lambda\overline{\lambda} : \lambda \in \sigma(A)\}\\ &= \sup\{|\pi(\lambda,\overline{\lambda})| : \lambda \in \sigma(A)\}\\ &\overset{(b)}{=} \sup\{|\lambda| : \lambda \in \sigma(\pi(A,A^*))\}\\ &= \sup\{|\lambda| : \lambda \in \sigma(A^*A)\}\\ &= \|A^*A\| = \|A\|^2,\end{aligned}$$

where the second-to-last identity follows from the self-adjointness of A^*A and Lemma 2.46.

(d) One has

$$\|\pi(A,A^*)\| \overset{(c)}{=} \sup\{|z| : z \in \sigma(\pi(A,A^*))\} \overset{(b)}{=} \sup\{|\pi(z,\overline{z})| : z \in \sigma(A)\}.$$

(e) One has for the polynomial $\pi - \pi'$

$$\|\pi(A,A^*) - \pi'(A,A^*)\| \overset{(d)}{=} \sup\{|\pi(z,\overline{z}) - \pi'(z,\overline{z})| : z \in \sigma(A)\}.$$

This finally concludes the proof. □

We are now in the position to prove the following.

Theorem 2.55 (Spectral Theorem for bounded normal operators)**.** *Let A be a bounded normal operator on a separable Hilbert space H. Then A is unitarily equivalent to a (necessarily bounded) multiplication operator M_q on $L^2(\Xi)$ for some finite measure space Ξ and some measurable function $q : \Xi \to \mathbb{C}$.*

Definition 2.56. *Let A be a bounded normal operator on a Hilbert space H. An element $\mathbf{x} \in H$ is called a* cyclic vector for A *if*

$$\{\pi(A, A^*)\mathbf{x} \in H : \pi \textit{ is a polynomial with complex coefficients}\}$$

is dense in H, i.e., if the space spanned by $\{A^k(A^*)^j\mathbf{x} : k, j \in \mathbb{N}\}$ is dense in H.

Proof. The proof of the theorem mimicks that of the self-adjoint case, and we will not repeat it entirely. We will only observe that, after introducing the definition of cyclic vector for a bounded normal operator, all the ingredients of the proof in the self-adjoint case have been recovered in the normal case in the last few pages. In fact, both the Stone–Weierstraß Theorem is valid for complex-valued functions, and so is the Riesz Representation Theorem (for *bounded* linear functionals on $C(K;\mathbb{C})$ rather than for positive ones on $C(K;\mathbb{R})$); while both Lemma 2.46 and the Spectral Mapping Theorem have been extended to the normal case in Lemma 2.54.

To fix the ideas, let A be a bounded normal operator on a Hilbert space H and let

$$\mathcal{V} := \{f \in C(\sigma(A)) : \exists \pi_f : \mathbb{C}^2 \to \mathbb{C} \text{ polynomial s.t. } f(z) = \pi_f(z, \overline{z}) \text{ for all } z \in \sigma(A)\},$$

where the polynomial is uniquely determined by Lemma 2.54.(e). Then, $\mathcal{V}$ is a complex Banach algebra, and in fact a C^*-algebra with respect to the involution $f \mapsto \overline{f}$. It plays the rôle that was played by the space of real polynomials in the proof of the Spectral Theorem for bounded self-adjoint operators, and in particular one begins by assuming that H contains a cyclic vector $\mathbf{x}$ for A and then defines

$$L(f) := (\pi_f(A, A^*)\mathbf{x}|\mathbf{x}), \qquad f \in \mathcal{V},$$

which can be extended by density (due to the Stone–Weierstraß Theorem) to a bounded linear functional on $C(\sigma(A))$. In fact, it is possible to define

$$f(A) := \pi_f(A, A^*),$$

which is a bounded linear operator with

$$\|f(A)\| = \|\pi_f(A, A^*)\| = \sup_{z\in\sigma(A)} |\pi_f(z, \overline{z})| = \sup_{z\in\sigma(A)} |f(z)|,$$

by Lemma 2.54.(d): If $f(z) = z^k\overline{z}^j$ and therefore $f(A) = A^k(A^*)^j$, then $\overline{f}(z) = z^j\overline{z}^k$ and therefore

$$\overline{f}(A) = A^j(A^*)^k = f(A)^* \qquad \text{for all } f \in \mathcal{V}.$$

The remainder of the proof can be repeated verbatim. □

Exercise 2.57. *Deduce from the above Spectral Theorem that all bounded normal operators can be (not necessarily uniquely) factorised as a product of a self-adjoint and a unitary operator.*

2.5. The Spectral Theorem for unbounded operators and the functional calculus

The main result of this chapter is an extension of the Spectral Theorem to general, possibly unbounded selfadjont operators. Again, it is shown that such operators can be identified with multiplication operators.

Unbounded multiplication operators are the handy source for a manifold of examples and counterexamples. For example, they can be used to prove that each closed subset of $\mathbb{C}$ can be the spectrum of a linear operator.

Exercise 2.58. *Consider a measure space (Ξ, μ) and define the* essential range *of a measurable function $q : \Xi \to \mathbb{C}$ as*

$$q_{ess}(\Xi) := \{z \in \mathbb{C} : \mu(\{\xi \in \Xi : |q(\xi) - z| < \epsilon) \neq 0 \textit{ for all } \epsilon > 0\}.$$

Prove the following assertions concerning the multiplication operator M_q.

(1) M_q has bounded inverse if and only if $0 \notin q_{ess}(\Xi)$, and in this case $M_q^{-1} = M_{q^{-1}}$, where

$$q^{-1}(x) := \begin{cases} \frac{1}{q(\xi)} & \textit{if } q(\xi) \neq 0, \\ 0 & \textit{if } q(\xi) = 0. \end{cases}$$

(2) $\sigma(M_q) = q_{ess}(\Xi)$.

Theorem 2.59 (Spectral Theorem for unbounded self-adjoint operators)**.** *Let A be a (possibly unbounded) self-adjoint operator on a separable Hilbert space H. Then A is unitarily equivalent to a multiplication operator M_q on $L^2(\Xi)$ for some σ-finite measure space Ξ and some measurable function $q : \Xi \to \mathbb{R}$.*

Proof. Let A be a self-adjoint operator. Therefore, by Theorem 2.27 the operators $A \pm i$ are invertible with bounded inverse. First of all, observe that

$$((A \pm i)^{-1})^* = (A \mp i)^{-1}. \tag{2.8}$$

In fact, for all $x, y \in D(A)$ one has

$$((A - i)x|y) = (x|(A + i)y),$$

or rather

$$((A - i)x|(A + i)^{-1}(A + i)y) = ((A - i)^{-1}(A - i)x|(A + i)y).$$

Setting $u := (A - i)x$ and $v := (A - i)y$ yields

$$(u|(A + i)^{-1}v) = ((A - i)^{-1}u|v)$$

for all $u \in \mathrm{Rg}(A - i) = H$ and all $v \in \mathrm{Rg}(A + i) = H$, which yields (2.8). Now,

$$(A + i)^{-1}((A + i)^{-1})^* = (A + i)^{-1}(A - i)^{-1} = \frac{1}{2i}((A - i)^{-1} - (A + i)^{-1}),$$

where the last identity follows from the resolvent equation, and similarly

$$((A + i)^{-1})^*(A + i)^{-1} = (A - i)^{-1}(A + i)^{-1} = \frac{1}{-2i}((A + i)^{-1} - (A - i)^{-1}),$$

i.e., the bounded operator $(A+i)^{-1}$ is normal and, likewise, so is $(A-i)^{-1}$. Accordingly, by the Spectral Theorem for bounded normal operators there exists a σ-finite measure space Ξ and a measurable function $q : \Xi \to \mathbb{C}$ and a unitary operator $U : L^2(\Xi) \to H$ such that

$$U^{-1}(A + i)^{-1}U = M_q.$$

In particular, the fact that $U^{-1}(A + i)^{-1}U$ and hence M_q is injective yields that q is essentially non-vanishing, i.e.,

$$\mu(\{\xi \in \Xi : q(\xi) = 0\}) = 0.$$

Define a new measurable function $p : \Xi \to \mathbb{C}$ by

$$p(\xi) := \begin{cases} \frac{1}{q(\xi)} - i & \text{if } q(\xi) \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

Then by definition $p + i = \frac{1}{q}$. In order to conclude the proof, it suffice to show that

$$U^{-1}AU = M_p,$$

with equality of domains, i.e.,

$$D(A) = \{x \in H : p \cdot (U^{-1}x) \in L^2(\Xi)\}.$$

Let us first prove the condition on the domains. In fact, we have that if $x \in D(A)$, and letting

$$y := (A + i)x,$$

yields

$$U^{-1}x = U^{-1}(A + i)^{-1}y = q \cdot (U^{-1}y) = \frac{1}{p + i} \cdot (U^{-1}y),$$

whence

$$p \cdot (U^{-1}x) = \frac{p}{p + i} \cdot (U^{-1}y) \in L^2(\Xi),$$

because $\frac{p}{p+i} \in L^\infty(\Xi)$.

Let conversely $pU^{-1}x \in L^2(\Xi)$. Then also $(p+i)\cdot U^{-1}x \in L^2(\Xi)$, hence there exists $y \in H$ with

$$U^{-1}y = (p+i)\cdot U^{-1}x,$$

hence

$$U^{-1}(A+i)^{-1}y = q \cdot U^{-1}y = U^{-1}x,$$

and we finally obtain

$$x = (A+i)^{-1}y \in D(A).$$

Finally, let $x \in D(A)$ and let

$$y := (A+i)x.$$

Then

$$U^{-1}x = U^{-1}(A+i)^{-1}y = M_q U^{-1}y = \frac{1}{p+i}\cdot U^{-1}y.$$

It follows that

$$U^{-1}Ax + iU^{-1}x = U^{-1}(A+i)x = U^{-1}y = (p+i)\cdot U^{-1}x = p\cdot U^{-1}x + iU^{-1}x,$$

i.e.,

$$U^{-1}Ax = p\cdot U^{-1}x.$$

Finally observe that while it is a priori not clear that (2.9) defines a real-valued function, self-adjointness of the multiplication operator M_p and Exercise 2.29 yield the claim. □

Remark 2.60. *The proof shows that the σ-finite measure space Ξ and the unitary operator U associated with A are the same as those associated with the normal operator $(A+i)^{-1} = R(i,-A)$. By Exercise 2.33 we deduce that again, like in the bounded case, Ξ is related to the spectrum of A.*

In fact, assume H to contain a cyclic vector for A and A to have compact resolvent (this is equivalent to the fact that $D(A)$ is compactly embedded in H, as is the case, e.g., for differential operators defined on Sobolev spaces on bounded domains with smooth boundary). Then it is well-known that A has discrete spectrum, i.e., $\sigma(A) = \sigma_p(A)$ is a countable set and

$$\Xi = \sigma((A+i)^{-1}) = \left\{ \frac{1}{i+\lambda} : \lambda \in \sigma_p(A) \right\}.$$

Hence, the measurable function $p : \Xi \to \mathbb{C}$ is in fact a sequence $(p_n)_{n\in\mathbb{N}}$, and the functional calculus acts on A by plugging this sequence in a given function entrywise, i.e.,

$$f(A) = UM_{(f(p_n))_{n\in\mathbb{N}}}U^{-1}.$$

Remark 2.61. *We emphasise that, just like in the finite dimensional case, even if each self-adjoint operator can be individually diagonalised this is not simultaneously true for families of self-adjoint operators, i.e., there need not exist a unitary operator U such that each of the self-adjoint operators is equivalent to some multiplication operators with respect to the same operator U.*

What is the spectral theorem good for? The fact of the matter is that a multiplication operator can be plugged in many functions (of a real variable!) in order to obtain solutions to partial differential equations.

Let Ξ be a σ-finite measure space and $q : \Xi \to \mathbb{C}$ a measurable function. We define a linear *multiplication operator* M_q by

$$\begin{aligned} D(M_q) &:= \{u \in L^2(\Xi) : q\cdot u \in L^2(\Xi)\}, \\ M_q u &:= q\cdot u. \end{aligned}$$

Such a multiplication operator is always closed and densely defined; by Hölder's inequality, it is bounded if and only if $q \in L^\infty(\Xi)$, and in this case $\|M_q\|_{\mathcal{L}(L^2(\Xi))} = \|q\|_\infty$. If M_q is indeed bounded, then it possible to define $f(M_q)$ for any real analytic function f, just as we have done when we defined the exponential function in Chapter 1.

We can finally define the functional calculus for a self-adjoint operator, and thus to take advantage of the spectral theorem. However, we can show that $f(M_q)$ can be given a meaning in a more general context. Before introducing a functional calculus for multiplication operators, observe that the space $B(\mathbb{R})$ of all bounded measurable functions over $\mathbb{R}$ is a Banach algebra with respect to the sup-norm, and in fact even a C^*-algebra (why?).

Proposition 2.62. *Let $f \in B(\mathbb{R})$ and $p : \Xi \to \mathbb{C}$ be a measurable function, where Ξ is a σ-finite measure space Ξ. Then*

$$\big(f(M_p)u\big)(x) := f(p(x))u(x), \qquad u \in L^2(\Xi),\ x \in \Xi,$$

i.e.,

$$f(M_p)u = (f \circ p)u, \qquad u \in L^2(\Xi).$$

defines a linear operator on $L^2(\Xi)$ which is bounded is additionally f is bounded.

Proof. Linearity of $f(M_p)$ is clear. Its boundedness is a direct consequence of the fact that

$$\|f(M_p)\|_{\mathcal{L}(L^2(\Xi))} = \|f \circ p\|_\infty \le \|f\|_\infty.$$

This is all we need to prove. □

Using the representation of self-adjoint operators as multiplication operators, the following holds. We do not prove the uniqueness part of its assertion.

Theorem 2.63 (Measurable functional calculus for self-adjoint operators)**.** *Let A be a linear self-adjoint operator on a Hilbert space H. Then there exists exactly one linear mapping, called* functional calculus

$$\mathcal{F} : B(\mathbb{R}) \ni f \mapsto f(A) \in \mathcal{L}(H)$$

such that all the following properties hold.

- $\|\mathcal{F}(f)\|_{\mathcal{L}(H)} \le \|f\|_\infty$ *for all* $f \in B(\mathbb{R})$.
- $\mathcal{F}(f \cdot g) = \mathcal{F}(f)\mathcal{F}(g)$ *for all* $f, g \in B(\mathbb{R})$.
- $\mathcal{F}(\overline{f}) = \mathcal{F}(f)^*$ *for all* $f \in B(\mathbb{R})$.
- *If* $(f_n)_{n\in\mathbb{N}} \subset B(\mathbb{R})$ *such that* $\lim f_n = id$ *pointwise and* $|f_n(r)| \le |r|$ *for all* $r \in \mathbb{R}$, *then* $\lim_{n\to\infty} f_n(A)x = Ax$ *for all* $x \in D(A)$.
- *If* $(f_n)_{n\in\mathbb{N}} \subset B(\mathbb{R})$ *and* $f \in B(\mathbb{R})$ *such that* $\lim f_n = f$ *pointwise and* $\|f_n\|_\infty$ *is uniformly bounded, then* $\lim_{n\to\infty} f_n(A)x = f(A)x$ *for all* $x \in H$.

Remark 2.64. *The assumption on self-adjointness is essential to guarantee that $f(A)$ defines a bounded linear operator for all bounded measurable functions f. However, this depends on our construction. It turns out that there exist several more functional calculi, where the rule of thumb is that* the rougher the operator A (and hence its spectrum), the narrower the choice of functions one can plug A in.

Example 2.65. *A self-adjoint operator A on a Hilbert space H is called* bounded from above, *or* semibounded, *if there exists $\omega \in \mathbb{R}$ such that*

$$\mathrm{Re}(Au|u)_H \le \omega \|u\|_H^2, \qquad u \in D(A).$$

Show that a self-adjoint operator generates a C_0-semigroup of bounded *linear operators on H if and only if A is semibounded.*

Now, the idea is that many partial differential equations, even some of those that are not in the form of (ACP), can be *formally* written as a vector-valued ordinary differential equation associated with a linear operator A: As a rule of thumb, if this ODE admits a solution, then the solution of the partial differential equation is given by plugging A into the corresponding function. The prototypical case is that of the abstract Cauchy problem (ACP), which is solved by the C_0-semigroup generated by A, formally obtained plugging A into the exponential function. Several properties of this C_0-semigroup can be obtained applying the spectral theorem.

Exercise 2.66. *Let H be a separable Hilbert space and consider a self-adjoint operator A on H such that $(Au|u)_H \leq 0$ for all $u \in D(A)$. Show that the semigroup $(T(t))_{t\geq 0}$ generated by A satisfies*

$$\|AT(t)\|_{\mathcal{L}(H)} \leq \frac{1}{te}, \qquad t > 0.$$

In particular, $T(t)$ is a bounded linear operator from H to $D(A)$ for all $t > 0$, and in fact one sees similarly that $T(t)$ maps H into $D(A^k)$ for arbitrary k. Are you able to find an analogous estimate for

$$\|t^k A^k T(t)\|_{\mathcal{L}(H)}, \qquad t > 0?$$

For a further example, consider the wave equation on an open domain Ω of $\mathbb{R}^d$ with Dirichlet boundary conditions, which can be formally written down as

$$U''(t) = AU(t),$$

where $\mathcal{A}$ is the self-adjoint operator Δ on $L^2(\Omega)$. Since the corresponding ODE is solved by

$$U(t) = \cos(t\sqrt{-A})U(0) + (\sqrt{-A})^{-1}\sin(t\sqrt{-A})U'(0),$$

it suffices to apply the corresponding formula to the multiplication operator associated with $\mathcal{A} = \Delta$ in order to have a solution formula for the partial differential equation. In the same way we can introduce the solution of the Schrödinger equation or find the generator of a semigroup $(T(t))_{t\geq 0}$ by formally taking its logarithm at $t = 1$.

Example 2.67. *We have already emphasised that an essentially bounded function gives rise to a bounded multiplication operator: hence, closed operators associated via the spectral theorem to bounded multiplication operators are necessarily bounded, hence not very interesting. The relevance of the spectral theorem consists in its assertion in the unbounded case. Now, observe that the proof of the spectral theorem is in generally not constructive, but it becomes so if a cyclic vector is known. A fundamental example is the case of the Laplacian on $\mathbb{R}$ for which the spectral theorem yields an explicit representation as the multiplication operator by the function $q(x) = x^2$, $x \in \mathbb{R}$, up to a unitary transformation which is, in fact, the Fourier transfomation (cf.* [**3**, §VI.5] *for an extension to a much more general case).*

Remark 2.68. *Like in the bounded case, the spectral theorem formulated above can be extended to the class of normal operators, i.e., of operators that commute with their adjoint. The proof is more technical and there is no abundance of evolution equations associated with normal operators: we therefore neglect this extension, but see* [**5**] *for a sketch of the general proof and some interesting general considerations about the theorem.*

Exercise 2.69. *Fill the details in the discussion of Example 2.67, i.e., show directly the self-adjointness of the Laplacian on $\mathbb{R}^d$.*

Exercise 2.70. *Let $H = L^2(0,1)$ and consider the operator*

$$\begin{aligned} D(A) &:= \{u \in H^1(0,1) : u(0) = ku(1)\}, \\ Au &:= iu'. \end{aligned}$$

Show that A is self-adjoint for all $k \in \mathbb{C}$ such that $|k| = 1$ and conclude that the evolution equation

$$\frac{\partial u}{\partial t}(t,x) + \frac{\partial u}{\partial x}(t,x) = 0, \qquad t \in \mathbb{R},\ x \in (0,1),$$

has a unique solution u for all initial data $u_0 \in D(A)$, which satisfies $\|u(t)\|_{L^2} = \|u_0\|$ for all $t \in \mathbb{R}$.

Exercise 2.71. *Consider the convection-reaction-diffusion equation*

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x) + k\frac{\partial u}{\partial x}(t,x) + qu(t,x), \qquad x \in (0,1),\ t > 0,$$

with Dirichlet boundary conditions. Show that the associated operator A is not selfadjoint on $L^2(0,1)$ with the standard inner product, but prove that it becomes self-adjoint with respect with a suitable, modified inner product that is equivalent to the original one.

Remark 2.72. *In quantum mechanics there are three fundamental equations, each valid for describing the evolution of the* wavefunction *of particular kinds of particles. The wavefunction is by definition a complex-valued measurable function whose L^2-norm over a certain domain yields the probability that the given particle is found in said domain. All these three equations are partial differential equations on L^2-spaces, whose elements (commonly called* spinors *in mathematical physics) have a different number of vectors in order to describe properties of different particles. These three equations*[3] *are the* Klein–Gordon equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c^2\Delta - \frac{m^2c^4}{\hbar^2}u(t,x), \qquad t\in\mathbb{R},\ x\in\mathbb{R}^3,$$

on $L^2(\mathbb{R}^3;\mathbb{C})$, valid for pions *(like the infamous, hypothetical Higgs boson); the* Schrödinger equation

$$i\frac{\partial u}{\partial t}(t,x) = c^2\Delta - \frac{m^2c^4}{\hbar^2}u(t,x), \qquad t\in\mathbb{R},\ x\in\mathbb{R}^3,$$

and its vector valued counterpart, the Pauli equation *(identical to the Schrödinger equation, besides the fact that the spinors are vectors in $L^2(\mathbb{R}^3;\mathbb{C}^2)$ instead of $L^2(\mathbb{R}^3;\mathbb{C})$), which are valid for general elementary particles, like electrons, depending of whether they have spin* 0 *or rather $\frac{1}{2}$; and the* Dirac equation *on $L^2(\mathbb{R}^3;\mathbb{C}^4)$ for* Dirac fermions *(i.e., for fermions, like quarks, that do not coincide with their own anti-particle), which we discuss in the next section. The differential operator appearing on the right of these equations is usually called the* Hamiltonian *of the system. The above Hamiltonians are* free*, in the sense that they describe particles that have no interaction whatsoever with other particles. While these equations are first defined on the whole space $\mathbb{R}^3$, one is usually interested in the time evolution of wavefunctions in smaller domains and/or that are subject to interactions with other particles. This is usually modelled by considering a* potential*, i.e., adding one or more terms to the free Hamiltonian, usually in form of multiplication operators by an unbounded function. As explained before, it is fundamental in quantum mechanics not only that these evolution equations are well-posed – which can be characterised by their being governed by a C_0-group, as we will see later – but also that the operators of the group are unitary, for the reasons explained above. While showing such a behaviour for the equation associated with the free Hamiltonian is easy, proving this for physically relevant potentials is much more involved. For example, it was Kato the first who was able to discuss in full generality the simplest model of quantum mechanics, the Schrödinger equation for the hydrogen atom which is described by a free Hamiltonian perturbed by a multiplication operator M_q, where q is a* Coulomb potential *of the form*

$$q(x) := \frac{1}{\|x\|}, \qquad x\in\mathbb{R}^3,$$

(the singularity in 0 represents the atomic nucleus).

Theorem 2.73. *Let H be separable Hilbert space and A be a self-adjoint operator. Then the initial value problem associated with the abstract Schrödinger equation*

$$i\frac{\partial U}{\partial t}(t) = AU(t), \qquad t\in\mathbb{R},$$

is well-posed and its solution is given by a C_0-group of unitary operators.

In fact, we will see later (Stone's Theorem) that the converse also holds.

Proof. Up to a unitary transformation, A agrees with a multiplication operator M_p on some Hilbert space $L^2(\Xi)$, for some measurable function p. Hence, the equation is unitarily equivalent to the ordinary (vector-valued) differential equation

$$i\frac{\partial V}{\partial t}(t) = M_pV(t), \qquad t\in\mathbb{R},$$

[3] It is interesting to observe that, unlike most other evolution equations of mathematical physics, these equations are not, strictly speaking, *derived* by physical considerations but rather introduced axiomatically, based only on an analogy with the Hamiltonian formulation of classical (i.e., Newtonian) mechanics (this analogy process is usually referred to as "first quantisation") and on mathematical properties.

which is solved by

$$V(t) := e^{-itM_p}V(0), \qquad t \in \mathbb{R}.$$

Now, we can give this formal expression a precise meaning owing to Theorem 2.63, i.e.,

$$(e^{-itM_p}V(0))(\xi) = e^{-itp(\xi)} \cdot V(0)(\xi), \qquad t \in \mathbb{R},\ \xi \in \Xi.$$

Furthermore, by Exercise 2.29

$$\|e^{-itM_p}\|_{\mathcal{L}(L^2)} = \|e^{-itp}\|_{L^\infty} = 1 \qquad \text{for all } t \in \mathbb{R},$$

since p is real-valued. □

Exercise 2.74. *Also the Klein–Gordon equation is governed by a unitary group. To see this, consider a first-order reduction similar to that in Example 1.4 and prove self-adjointness of the associated operator matrix in the Hilbert space $H^1_{en}(\mathbb{R}^3) \times L^2(\mathbb{R}^3)$, where $H^1_{en}(\mathbb{R}^3)$ is the closure of $H^1(\mathbb{R}^3)$ with respect to*

$$|||u||| := \|\nabla u\|, \qquad u \in H^1(\mathbb{R}^3).$$

2.6. An application: the Dirac equation

2.6.1. The 3-dimensional Dirac equation. The Dirac equation is formally similar to a Schrödinger equation, but with a Hamiltonian that is a first-order differential operator on the Hilbert space $H = L^2(\mathbb{R}^3; \mathbb{C}^4)$. More precisely, the equation is given by

$$i\hbar\frac{\partial u}{\partial t}(t,x) = \begin{pmatrix} mc^2\beta & -i\hbar c\sigma\cdot\nabla \\ -i\hbar c\sigma\cdot\nabla & -mc^2\beta \end{pmatrix} u(t,x), \qquad t \in \mathbb{R},\ x \in \mathbb{R}^3,$$

where the matrix-vector σ is given by

$$\sigma := (\sigma_1,\sigma_2,\sigma_3) := \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) \qquad \text{and} \qquad \beta := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

the vector consisting of the so-called *Pauli matrices*[4]. The associated abstract Cauchy problem is hence

$$i\frac{\partial U}{\partial t}(t) = AU(t), \qquad t \in \mathbb{R},$$

where A is given by the 4×4 operator matrix

$$A := \begin{pmatrix} \frac{mc^2}{\hbar}\operatorname{Id} & -ic\sigma\cdot\nabla \\ -ic\sigma\cdot\nabla & -\frac{mc^2}{\hbar}\operatorname{Id} \end{pmatrix}.$$

Now, the diagonal block matrices are bounded operators on the space $L^2(\mathbb{R}^3; \mathbb{C}^4)$, hence by Exercise 2.18 the operator A is self-adjoint if and only if

$$A_1 := \begin{pmatrix} 0 & -ic\sigma\cdot\nabla \\ -ic\sigma\cdot\nabla & 0 \end{pmatrix}$$

is. Moreover, this operator matrix has a special structure: it is unitarily equivalent to the operator matrix

$$\tilde{A}_1 := \begin{pmatrix} K & 0 \\ 0 & -K \end{pmatrix} := \begin{pmatrix} -ic\sigma\cdot\nabla & 0 \\ 0 & ic\sigma\cdot\nabla \end{pmatrix}$$

via the operator matrix

$$U := \frac{1}{\sqrt{2}}\begin{pmatrix} \operatorname{Id} & \operatorname{Id} \\ \operatorname{Id} & -\operatorname{Id} \end{pmatrix}.$$

---

[4] The reason why exactly these matrices play such an important rôle in quantum mechanics is that, together with the identity matrix, they span the space of all 2×2 Hermitian matrices.

Now, it is clear that A_1 is self-adjoint if and only if its upper-left block is so, i.e., if and only if

$$K := -ic\sigma \cdot \nabla = -c \begin{pmatrix} i\frac{\partial}{\partial x_3} & i\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} \\ i\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} & -i\frac{\partial}{\partial x_3} \end{pmatrix}$$

is self-adjoint.

Remark 2.75. *Observe that the operator A is not* bounded below*, i.e.,*

$$\mathrm{Re}(AU, U) \geq \omega \|U\|^2, \qquad U \in D(A), \tag{2.10}$$

does not hold for any $\omega \in \mathbb{R}$. This is why the more common formulation of the spectral theorem (cf. [**1**, App. B]*) cannot be applied in the case of the Dirac equation: this failure of* (2.10) *is the reason why the spectrum of A consists of arbitrarily large* and *arbitrarily small values. Since spectral values have the physical meaning of critical energies of particles, the existence of negative spectral values (only based on a formal mathematical model) was quite puzzling, until Richard Feynman proposed an explanation of this fact by describing negative values of a particle as positive particles of its antiparticle. In this approach, the Dirac equation effectively describes a pair particle/antiparticle, each corresponding to two components of the 4-spinor, one moving foreward in time, the other backward.*

Theorem 2.76. *The operator K is self-adjoint on $L^2(\mathbb{R}^3; \mathbb{C}^2)$, hence the operator A is self-adjoint on $L^2(\mathbb{R}^3; \mathbb{C}^4)$.*

It is not clear when self-adjointness of this operator has been proved for the first time: the above result may be due to Kato, whose approach in [**6**, § V.5.4] is the common one in the later literature. While his proof is certainly correct, his approach is based on directly diagonalising the operator via the Fourier transform, thus allowing for a direct application of the functional calculus, without any need to apply the spectral theorem in the first place: see e.g. [**13**, Thm. 1.1]. I am not aware of a proof of Theorem 2.76 that uses Theorem 2.27 instead.

2.6.2. The 1-dimensional Dirac equation. Let us now focus on a peculiar case – that of 1-dimensional Dirac equations. While one can think that the 1-dimensional Dirac equation is simply a special case of the 3-dimensional one, this is not completely true. On one hand, the equation dramatically simplifies (since one can drop the terms depending on $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}$) and this in turn permits to find simpler matrices that satisfy the necessary algebraic conditions: in particular, one can now chose the Dirac matrix

$$\sigma := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Therefore, we can consider 2-spinors (instead of 4-spinors like in the 3-dimensional case) and are eventually led to the equation

$$i\frac{\partial U}{\partial t}(t) = AU(t), \qquad t \in \mathbb{R},$$

where A is given by the 2×2 operator matrix

$$A := \begin{pmatrix} \frac{mc^2}{\hbar} & -ic\frac{\partial}{\partial x} \\ -ic\frac{\partial}{\partial x} & -\frac{mc^2}{\hbar} \end{pmatrix}. \tag{2.11}$$

Unlike in the 3-dimensional case, self-adjointness of A can be checked directly, applying Theorem 2.27.

Proposition 2.77. *The operator A with domain $H^1(\mathbb{R}; \mathbb{C}^2)$ in* (2.11) *is a self-adjoint operator on the Hilbert space $L^2(\mathbb{R}; \mathbb{C}^2)$.*

Proof. We can think of A as a bounded perturbation of

$$A_0 := \begin{pmatrix} 0 & -ic\frac{\partial}{\partial x} \\ -ic\frac{\partial}{\partial x} & 0 \end{pmatrix},$$

hence by Exercise 2.18 it suffices to check self-adjointness of A_0. By Theorem 2.27 we have to check that A_0 is symmetric and that both $A_0 + i$ and $A_0 - i$ are surjective. Symmetry follows by observing that, by integration by parts, for all $f \in D(A)$

$$\begin{aligned} (A_0 f|g) &= -\int_{\mathbb{R}} icf_2'\overline{g_1}dx - \int_{\mathbb{R}} icf_1'\overline{g_2}dx \\ &= \int_{\mathbb{R}} icf_2\overline{g_1'}dx + \int_{\mathbb{R}} icf_1\overline{g_2'}dx \\ &= -\int_{\mathbb{R}} cf_2\overline{ig_1'}dx - \int_{\mathbb{R}} cf_1\overline{ig_2'}dx = \big(f\big|A_0 g\big)\,. \end{aligned}$$

In order to show that $A_0 \pm i$ is surjective, take $h = (h_1, h_2) \in L^2(\mathbb{R};\mathbb{C}^2)$ and observe that finding $f \in D(A)$ such that $A_0 f \pm if = h$ amounts to finding $f = (f_1, f_2) \in H^1(\mathbb{R};\mathbb{C}^2)$ such that

$$\begin{cases} -icf_2'(x) \pm if_1(x) = h_1(x), & x \in \mathbb{R}, \\ -icf_1'(x) \pm if_2(x) = h_2(x), & x \in \mathbb{R}, \end{cases}$$

or rather

$$\begin{cases} f_1'(x) = \pm\frac{1}{c}f_2(x) + \frac{i}{c}h_2(x), & x \in \mathbb{R}, \\ f_2'(x) = \pm\frac{1}{c}f_1(x) + \frac{i}{c}h_1(x), & x \in \mathbb{R}, \end{cases}$$

which is a linear system of ordinary differential equations. We can re-write it as a vector-valued problem

$$f' = Cf + \tilde{h} := \pm\begin{pmatrix} 0 & \frac{1}{c} \\ \frac{1}{c} & 0 \end{pmatrix} f + \frac{i}{c}h$$

and is hence solved by the variation of parameter formula

$$f(x) = e^{xC}f(0) + \int_0^x e^{(x-y)C}\tilde{h}(y)dy, \qquad x \in \mathbb{R}.$$

This concludes the proof. □

Remark 2.78. *1-dimensional Dirac equations are realised, for instance, when the waveguide we are considering is (almost) 1-dimensional, like in the case of graphene. More precisely, consider a* quantum dot*, i.e., a (finite) carbon molecule that looks like a 3-regular tessellation of a small region of the plane (we are neglecting the boundary effects).*

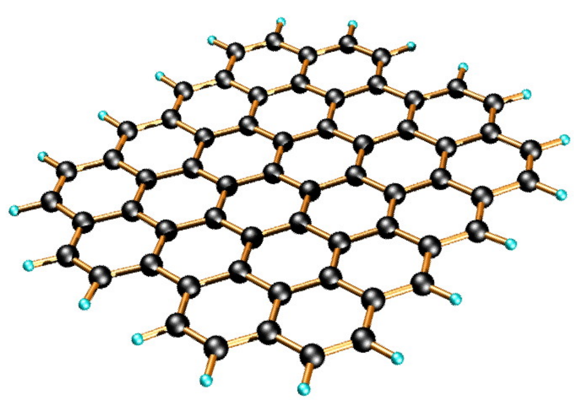

Figure 1. (courtesy of Zheng Yan and Andrew R. Barron)

For relevant molecules, and in particular for graphene, one has observed that the evolution of such quantum systems is described by a Dirac equation for an electron moving along the bonds (the atom being its ramification nodes). If the quantum dot contains m bonds, the relevant Hilbert space is instead $L^2(0,1;\mathbb{C}^{2m})$, and one has to impose suitable transmission conditions in the nodes. Interestingly enough, such conditions have not been identified yet: however, an algebraic parametrisation of all transmission conditions leading to a self-adjoint operator (and hence, by Theorem 2.73, to a C_0-group of unitary operators governing the Dirac equation) has been obtained in [**2**, §4.1].

Definition 2.79. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a C_0-semigroup on X. Then $(T(t))_{t\geq 0}$ is called* strongly stable *if*

$$\lim_{t\to\infty} T(t)x = 0 \qquad \text{for all } x \in X.$$

Proposition 2.80. *Let H be a Hilbert space and A be a self-adjoint operator bounded from above. Then the semigroup generated by A is strongly stable if and only if 0 is not an eigenvalue of A.*

Proof. By the Spectral Theorem, A is unitarily equivalent to a multiplication operator M_q on some $L^2(\Xi)$-space. Since stability is clearly invariant under unitary transformation, it suffices to check under which assumptions $(e^{tM_q})_{t\geq 0}$ is strongly stable, i.e., under which assumptions

$$\lim_{t\to\infty} e^{tM_q} f = \lim_{t\to\infty} e^{tq} f = 0 \qquad \text{for all } f \in L^2(\Xi).$$

A necessary condition for stability is hence clearly that the essential range of q is contained in $(-\infty, 0]$; and moreover 0 is in the essential range of q (i.e., $\{\xi \in \Xi : q(\xi) = 0\}$ has positive measure) if and only if 0 is not an eigenvalue of M_q and hence of A, since by Exercise 2.34 the spectrum of M_q agrees with the essential range of q (why does $0 \notin \sigma_p(A)$ already imply that $\{\xi \in \Xi : q(\xi) = 0\}$ has zero measure?). □

CHAPTER 3

General semigroup theory

What if the operator A that appears in an abstract Cauchy problem is not self-adjoint? Then, we cannot apply the spectral theorem applied above and we have no hope to define the solution by plugging A into the exponential function. There exist many interesting examples of non-selfadjoint operators, like the first derivative that acts on functions on $\mathbb{R}$ (which is associated with the transport equation) and the elliptic operators like

$$A := \nabla \cdot (\alpha \nabla),$$

acting on functions on some domain $\Omega \subset \mathbb{R}^d$, say with Dirichlet boundary conditions, whenever the matrix-valued coefficient α is such that $\alpha(x)$ is *not* hermitian for a.e. $x \in \Omega$. Moreover, the methods presented in the previous setting is restricted to abstract Cauchy problems on Hilbert spaces, whereas in many cases the relevant state space is not L^2 but, say, L^1 (in fact, the L^1-norm is the relevant one for a large class of problems, since it represents the total heat, mass, population, etc. of a system.

Another, more difficult but more general way to define such a solution, i.e., to check whether A generates a semigroup, is presented in this chapter.

3.1. Generation results

First of all, we notice the following properties of a semigroup's generator that complement those in Theorem 1.20.

Lemma 3.1. *Let X be a Banach space and let $(T(t))_{t\geq 0}$ be a C_0-semigroup on X with generator A. Using the same notations for the constants introduced in Theorem 1.20.(1), the following assertions hold.*

(1) A is closed, densely defined and determines $(T(t))_{t\geq 0}$ uniquely.
(2) If for some $\lambda \in \mathbb{C}$ the Laplace transform of the semigroup exists strongly, i.e., if $\int_0^\infty e^{-\lambda s}T(s)xds$ converges for all $x \in X$, then $\lambda \in \rho(A)$ and such a Laplace transform coincides with the resolvent operator of A at λ, i.e., $R(\lambda, A) = \int_0^\infty e^{-\lambda s}T(s)xds$.
(3) Conversely, if $\mathrm{Re}\lambda > \omega$, then $\lambda \in \rho(A)$ and the resolvent operator $R(\lambda, A)$ agrees with $\int_0^\infty e^{-\lambda s}T(s)ds$.
(4) The so-called Hille–Yosida-estimate

$$\|R(\lambda, A)\| \leq \frac{M}{\mathrm{Re}\lambda - \omega}$$

holds for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > \omega$.

The estimate in Theorem 1.20.(1) holds for infinitely many ω, but it is their infimum ($\neq -\infty$, since M has to stay ≥ 1) that is relevant. Such an infimum is called *growth bound* of $(T(t))_{t\geq 0}$. One of the most important features of semigroup theory is that even when it is impossible to find an explicit formula for the solution of an abstract Cauchy problem, it is sometimes possible to find interesting asymptotic properties of the semigroup (i.e., long-time behaviour of solutions) by careful investigations of its generator (in particular, of its spectrum). However, it is in general a very difficult question whether the growth bound is in fact also the minimum, i.e., whether it is attained. The main exception is the simple case of $\omega = 0$ (contractive C_0-semigroups).

PROOF. (1) Let $(x_n)_{n\in\mathbb{N}} \subset D(A)$ such that $x_n \to x$ and $Ax_n \to y$ for some $x, y \in X$. Then by Theorem 1.20

$$T(t)x_n - x_n = \int_0^t T(s)Ax_n ds, \qquad n \in \mathbb{N},$$

and passing to the limit for $n \to \infty$

$$T(t)x - x = \int_0^t T(s)yds.$$

Now, divide both sides by t and take the limit for $t \to 0+$: the LHS is just Ax (simply by definition of generator), and in particular $x \in D(A)$, whereas the RHS agrees with $T(0)y = y$ by the fundamental theorem of calculus.

(2) Due to Example 1.14, one can assume that $\lambda = 0$ (since we could otherwise repeat the same argument for the rescaled semigroup $(\tilde{T}(t))_{t\geq 0}$ and eventually use the fact that $0 \in \rho(\lambda - A)$ if and only if $\lambda \in \rho(A)$). Therefore

$$\begin{aligned}
\frac{T(h) - \mathrm{Id}}{h} R(0)x &= \frac{T(h) - \mathrm{Id}}{h} \int_0^\infty T(s)xds \\
&= \frac{1}{h}\int_0^\infty T(s+h)xds - \frac{1}{h}\int_0^\infty T(s)xds \\
&= \frac{1}{h}\int_h^\infty T(s)xds - \frac{1}{h}\int_0^\infty T(s)xds \\
&= -\frac{1}{h}\int_0^h T(s)xds.
\end{aligned}$$

Hence, letting $h \to 0+$, we obtain that $AR(0, A) = -\mathrm{Id}$ (and in particular $R(0, A) \subset D(A)$). Moreover, let $x \in D(A)$. Then

$$\lim_{t\to\infty} \int_0^\infty T(s)xds = R(0, A)x$$

as well as

$$\lim_{t\to\infty} A\int_0^\infty T(s)xds = \lim_{t\to\infty}\int_0^\infty T(s)Axds = R(0, A)Ax,$$

by Theorem 1.20.(4). By (2) A is closed, hence $R(0)Ax = AR(0)x = -x$, hence $R(0, A)$ is the inverse of A, i.e., $0 \in \rho(A)$ (boundedness of $R(0, A)$ follows from closedness of A).

(3) Let now $\lambda \in \mathbb{C}$ such that $\mathrm{Re}\lambda > \omega$. Then the Laplace transform of the semigroup at λ exists, since it satisfies the estimate

$$\|\int_0^\infty e^{-\lambda s}T(s)ds\| \leq \int_0^\infty e^{-\mathrm{Re}\lambda s}Me^{\omega s}ds = M\int_0^\infty e^{\omega - \mathrm{Re}\lambda s}ds = \lim_{t\to\infty} M\int_0^t e^{\omega-\mathrm{Re}\lambda s}ds = \frac{M}{\mathrm{Re}\lambda - \omega}.$$

By (3), one then necessarily has $\lambda \in \rho(A)$ and $R(\lambda, A)$ agrees with $\int_0^\infty e^{-\lambda s}T(s)ds$.

(4) The above estimate and the identity of $R(\lambda, A)$ and $\int_0^\infty e^{-\lambda s}T(s)ds$ yields the claim and concludes the proof. □

Let us summarise what we already know on generators of C_0-semigroups: They are closed, densely defined linear operators whose resolvent set contains a right half-plane of $\mathbb{C}$ and such that the Hille–Yosida-estimate is satisfied.

The first breakthrough in the theory of semigroups was obtained in 1948, when Einar Hille and Kosaku Yosida proved independently that the converse also holds. More precisely, the following holds.

Theorem 3.2 (Hille–Yosida 1948). *Let X be a Banach space and A be a (possibly unbounded) linear operator on X. Then the following assertions are equivalent.*

(a) A is the generator of a C_0-semigroup of contractions.

(b) A is densely defined, closed and moreover $\lambda \in \rho(A)$ and $\|\lambda R(\lambda, A)\| \leq 1$ for all real $\lambda > 0$.

Before proving this theorem, we need to show the following results.

Lemma 3.3. *Let X be a Banach space and A be a closed, densely defined operator on X. If there exist $\omega \in \mathbb{R}$ and $M \geq 1$ such that*

$$\lambda \in \rho(A) \quad \text{and} \quad \|\lambda R(\lambda, A)\| \leq M \qquad \text{for all real } \lambda \geq \omega.$$

Then the following assertions hold.

(1) $\lim_{\lambda\to\infty} \lambda R(\lambda, A)x = x$ *for all* $x \in X$.

(2) $\lim_{\lambda\to\infty} \lambda AR(\lambda, A)x = Ax$ *for all* $x \in D(A)$.

Proof. (1) One has that $\lambda R(\lambda, A)y = R(\lambda, A)Ay + y$ by definition of resolvent, for all $y \in D(A)$. Therefore,

$$\|\lambda R(\lambda, A)y - y\| \leq \|R(\lambda, A)Ay\| \leq \frac{M}{\lambda} \to 0,$$

i.e., $\lambda R(\lambda, A)y \to y$ as $\lambda \to \infty$. Since this holds on the dense subset $D(A)$ of X, it also holds on the whole X, i.e., one has proved the claimed strong convergence result.

(2) The assertion follows observing that $\lambda AR(\lambda, A)x = \lambda R(\lambda, A)Ax$ for all $x \in D(A)$ and applying (1) on $Ax \in X$. □

Lemma 3.4. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a family of bounded linear operators on X that satisfies* (SEMIGR). *Then the following assertions are equivalent.*

(i) The semigroup is strongly continuous.

(ii) The semigroup is strongly continuous at $0+$.

(iii) There exist $\tau > 0$ and $M \geq 1$, along with $D \subset X$ such that $\overline{D} = X$, for which $\|T(t)\| \leq M$ for all $t \in [0, \delta]$ and $\lim_{t\to 0} T(t)x = x$ for all $x \in D$.

Proof. (i) ⇒(iii) is Theorem 1.20.(1).

(iii)⇒(ii) The idea is to prove strong continuity at $t = 0+$ of the semigroup by proving its strong continuity with respect to arbitrary sequences that converge to $0+$. The equivalence of continuity and sequential continuity on Banach spaces will do the job.

Let $(t_n)_{n\in\mathbb{N}} \subset [0, \infty)$ be an arbitrary null sequence and denote by K the set of its values. Then $\{T(t_n) \in \mathcal{L}(X) : n \in \mathbb{N}\}$ is the set of values of a continuous function on a compact set, hence it is a bounded set and moreover by assumption $\lim_{t\to 0} T(t)x = x$ for all $x \in D$, hence in particular its restriction to the discrete set K is continuous, i.e., $K \ni t \mapsto T(\cdot)x \in X$ is continuous for all $x \in D$. We conclude from the forthcoming Exercise 3.5 that in fact $K \ni t \mapsto T(\cdot)x \in X$ is continuous for all $x \in X$, and since $(t_n)_{n\in\mathbb{N}} \subset [0, \infty)$ is a null sequence one has $\lim_{n\to\infty} T(t_n)x = x$, qed.

(ii)⇒(i) As one may expect, the third implication relies on the semigroup law. Let $t_0 > 0$ and take $x \in X$. We want to prove that $\lim_{t\to t_0} T(t)x = T(t_0)x$. Of course, it suffices to prove right *and* left continuity at t_0. To begin with, we prove that $\lim_{t\to t_0+} T(t)x = T(t_0)x$: in fact,

$$\|T(t_0 + h)x - T(t_0)x\| \leq \|T(t_0)\|\|T(h)x - x\| \overset{h\to 0+}{\longrightarrow} 0.$$

Right continuity is a bit more delicate. First, observe that $T(\cdot)$ is uniformly bounded on $[0, t_0]$[1], with just the same proof as in Theorem 1.20.(1). This actually yields uniform boundedness on some small interval $[0, \delta]$; but then uniform boundedness on the whole interval $[0, t_0]$ is a consequence of the semigroup law: take $t \leq t_0$ and pick $m \in \mathbb{N}$ and $\delta_0 \in (0, \delta]$ in such a way that $t = m\delta_0$.

$$\|T(t)\| = \|T(m\delta_0)\| \leq \|T(\delta_0)\|^m \leq \sup_{s\in[0,\delta} \|T(s)\|^m < \infty.$$

Then for all $h \in (-t_0, 0)$

$$\|T(t_0 + h)x - T(t_0)x\| \leq \|T(t_0 + h)\|\|x - T(-h)x\| \overset{h\to 0+}{\longrightarrow} 0,$$

[1] Observe that this is always the case for C_0-semigroups, but we are *not* assuming strong continuity.

by uniform boundedness of $\|T(t_0 + h)\|$. □

Exercise 3.5. *Let X be a Banach space and $F : K \to \mathcal{L}(X)$, where K is a compact subset of $\mathbb{R}$. Prove that the following assertions are equivalent.*

(i) F is strongly continous.
(ii) F is uniformly bounded and there exists $D \subset X$ such that $\overline{D} = X$ and the orbits $K \ni t \mapsto F(t)x \in X$ are continuous for all $x \in D$.
(iii) The mapping

$$K \times C \ni (t, x) \mapsto F(t)x \in X$$

is (jointly) uniformly continuous for all compact subset C of X.

At least four different proofs of the Hille–Yosida Theorem exist: we will present the most famous and *most constructive* one, that due to Yosida.

Proof of the Hille–Yosida Theorem. It has already been proved that A is necessarily densely defined and closed and that it satisfies the Hille–Yosida-estimate (which is stronger than the claimed one, since by assumptions we can take $\omega = 0$ and $M = 1$ when applying Lemma 3.1). To prove the converse inequality, Yosida's essential idea was to define a family of *bounded* operators on X, which hence generate a family of C_0-semigroups that is proved to converge towards a C_0-semigroup that is in turn generated by A. One can therefore look at the Hille–Yosida Theorem as an approximation result for sequences of semigroups, but there is in fact not much choice about this, since no natural candidate is available as semigroup generated by A.

To begin with, we consider the *Yosida approximants*, the sequence of bounded linear operators

$$A_n := nAR(n, A) = n^2 R(n, A) - n\,\mathrm{Id}, \qquad n \in \mathbb{N}.$$

The reason we chose this sequence is that, by Lemma 3.3,

$$\lim_{n\to\infty} A_n x = Ax \qquad \text{for all } x \in D(A). \tag{3.1}$$

Furthermore, observe that the C_0-semigroup generated by A_n (by Proposition 1.9) is contractive, since

$$\|e^{tA_n}\| = \|e^{tn^2R(n,A)-tn\,\mathrm{Id}}\| = \|e^{tn^2R(n,A)}e^{-tn}\| = e^{-tn}\|e^{tn^2R(n,A)}\| \le e^{-tn}e^{t\|n^2R(n,A)\|} \le e^{-tn}e^{tn} = 1, \qquad t \ge 0.$$

In order to show that the sequence $(e^{tA_n}x)_{n\in\mathbb{N}}$ is convergent for all $t \ge 0$ and all $x \in X$, it suffices to show that $(e^{tA_n}x)_{n\in\mathbb{N}}$ is a Cauchy sequence for all $t \ge 0$ and all $x \in D(A)$. In fact, fix $t > 0$ and $x \in D(A)$ and consider the function Q defined by

$$Q : [0, t] \ni s \mapsto e^{(t-s)A_m}e^{sA_n} \in \mathcal{L}(X)$$

(a similar mapping has been useful in the proof of Lemma 1.22). One then has

$$\begin{aligned}
e^{tA_n}x - e^{tA_m}x &= Q(t)x - Q(0)x \\
&= \int_0^t \frac{d}{ds}Q(s)x\,ds \\
&= \int_0^t \frac{d}{ds}e^{(t-s)A_m}e^{sA_n}x\,ds \\
&= \int_0^t \big(-A_m e^{(t-s)A_m}e^{sA_n}x + e^{(t-s)A_m}A_n e^{sA_n}x\big)ds \\
&= \int_0^t e^{(t-s)A_m}e^{sA_n}(A_n - A_m)x\,ds,
\end{aligned}$$

and hence by contractivity of the semigroups

$$\begin{aligned}\|e^{tA_n}x - e^{tA_m}x\| &\leq \int_0^t \|e^{(t-s)A_m}e^{sA_n}(A_n - A_m)x\|ds \\ &\leq \int_0^t \|(A_n - A_m)x\|ds \\ &= t\|(nAR(n,A) - mAR(m,A))x\|.\end{aligned}$$

Letting $n, m \to \infty$, this expression tends to 0 by Lemma 3.3. Hence, by density of $D(A)$ one deduces that $(e^{tA_n})_{n\in\mathbb{N}}$ converges strongly, uniformly in t belonging to compact intervals. Let us call its strong limit $T(t)$, for $t \geq 0$.

Now, since $(T(t))_{t\geq 0}$ is the (strong) limit of a sequence of C_0-semigroups of contractions, and in particular of families of contractive operators satisfying the semigroup law (SEMIGR), also $(T(t))_{t\geq 0}$ is necessarily a family of contractive operators that satisfy the semigroup law (SEMIGR). Also observe that because for all $x \in X$ the mapping

$$[0, \infty) \ni t \mapsto T(t)x \in X$$

satisfies

$$\|T(t)x - e^{tA_m}x\| = \lim_{n\to\infty} \|e^{tA_n}x - e^{tA_m}x\| \leq \lim_{n\to\infty} t\|(nAR(n,A) - mAR(m,A))x\|,$$

it is uniform limit (for t in compact intervals) of a sequence of continuous functions, it is continuous. Accordingly, $(T(t))_{t\geq 0}$ is strongly continuous by Lemma 3.4.

Summing up, we have introduced a sequence of bounded operators and showed that the semigroups they generate converge to a C_0-semigroup. It remains to prove that the generator of this semigroup is exactly A. To this aim, we denote by B the generator of $(T(t))_{t\geq 0}$ and are going to prove that $A = B$. First of all, we prove that $A \subset B$. Take $x \in D(A)$: we have already observed that the orbit sequence

$$[0, \infty) \ni t \mapsto e^{tA_m}x \in X, \qquad m \in \mathbb{N},$$

converges uniformly (for t in compact intervals) to

$$[0, \infty) \ni t \mapsto T(t)x \in X.$$

Similarly, the sequence of differentiated orbits

$$[0, \infty) \ni t \mapsto e^{tA_m}A_m x \in X, \qquad m \in \mathbb{N},$$

converges uniformly (for t in compact intervals) to

$$[0, \infty) \ni t \mapsto T(t)Ax \in X,$$

by (3.1). Hence, the orbit of $(T(t))_{t\geq 0}$ is a function that is uniform limit of differentiable functions $[0, \infty) \to X$, hence itself differentiable, and in particular its derivative at 0 agrees with the limit of the derivatives of the approximating orbits, i.e., with

$$T(0)Ax = Ax.$$

We conclude that for all $x \in D(A)$ the orbit $(T(t)x)_{t\geq 0}$ is differentiable at 0 and

$$Bx = \frac{d}{dt}T(t)x|_{t=0} = Ax,$$

i.e., $D(A) \subset D(B)$ and $Ax = Bx$.

To prove the converse inclusion let $\lambda > 0$, hence in particular $\lambda \in \rho(A)$, so that $\lambda - A : D(A) \to X$ is bijective. Also $\lambda - B : D(B) \to X$ is a bijection: in fact, since B generates a C_0-semigroup of contractions, the assertion follows from Lemma 3.1.(4). We have already showed (while proving the converse inclusion) that $(\lambda - A)x = (\lambda - B)x$ for all $x \in D(A)$: i.e., $\lambda - A$ is a restriction of $\lambda - B$. But if two bijective operators, of which one is a restriction of the other, coincide on the domain of the smaller one, then they have to coincide as operators. In fact, if they would not, i.e., if there were $x \in D(B) \setminus D(A)$ (we can already rule out the case that

$D(B) \subset D(A)$ but $Bx \neq Ax$, by the proof of the converse inclusion), then $Bx \in X$ and therefore $Bx = Ay$ for some $y \in D(A)$. But since $A \subset B$, one has $Bx = By$, i.e., $(\lambda - B)x = (\lambda - B)y$ and therefore $x = y$ by the invertibility of $\lambda - B$. This is a contradiction to the fact that $y \in D(A)$ and $x \in D(B) \setminus D(A)$. □

A few years later, it became clear that an ingenious rescaling argument allows for a generalisation of the Hille–Yosida theorem to the general, non-contractive case. We present the statement but omit the rather technical proof, cf. [**3**, II.3.8].

Theorem 3.6 (Hille–Yosida Theorem (general case))**.** *Let X be a Banach space and A be a densely defined, closed linear operator on X such that $\lambda \in \rho(A)$ for all real $\lambda > \omega$. Then the following assertions are equivalent.*

(a) A is the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ such that $\|T(t)\| \leq Me^{\omega t}$ for all $t \geq 0$.

(b) A is densely defined, closed and moreover $\lambda \in \rho(A)$ and $\left\|\Big((\lambda-\omega)R(\lambda, A)\Big)^n\right\| \leq M$ for all real $\lambda > \omega$ and all $n \in \mathbb{N}$.

(c) A is densely defined, closed and moreover $\lambda \in \rho(A)$ and $\|R(\lambda, A)\| \leq \frac{M}{\mathrm{Re}(\lambda-\omega)}$ for all real $\lambda > \omega$.

We only mention the following step in the proof of the above theorem.

Exercise 3.7. *Let X be a Banach space and let $(T(t))_{t\geq 0}$ be a uniformly bounded C_0-semigroup on X. Then*

$$|\!|\!|x|\!|\!| := \sup_{t\geq 0} \|T(t)x\|, \qquad x \in X,$$

defines a norm on X that is equivalent to the original one, and with respect to which $(T(t))_{t\geq 0}$ becomes contractive.

Remark 3.8. *Hille's approach is also remarkable, in particular because it paved the road to the nonlinear extension of the semigroup theory. His idea was to use another approximation scheme, namely the so-called* backward Euler scheme, *which amounts to discretise the evolution equation by fixing $t \geq h > 0$ and imposing that A satisfies a difference equation of the form*

$$\frac{S(t)x - S(t-h)x}{h} \stackrel{!}{=} AS(t)x,$$

for some operator family $(S(t))_{t\geq 0}$. This yields

$$S(t) \stackrel{!}{=} (I - hA)^{-1} S(t-h).$$

For $h := \frac{t}{n}$ we obtain

$$S(t) \stackrel{!}{=} \left(I - \frac{t}{n}A\right)^{-1} S\left(\frac{n-1}{n}t\right).$$

Repeating the same procedure for the term $S\left(\frac{n-1}{n}t\right)$ yields

$$\frac{S(t-h)x - S(t-2h)x}{h} \stackrel{!}{=} AS(t)x,$$

whence

$$S(t-h) \stackrel{!}{=} (I - hA)^{-1} S(t-2h),$$

and chosing again $h := \frac{t}{n}$ we obtain

$$S\left(\frac{n-1}{n}t\right) \stackrel{!}{=} \left(I - \frac{t}{n}A\right)^{-1} S\left(\frac{n-2}{n}t\right),$$

hence

$$S(t) \stackrel{!}{=} \left(I - \frac{t}{n}A\right)^{-2} S\left(\frac{n-2}{n}t\right).$$

and proceding recursively we arrive at

$$S(t) \stackrel{!}{=} \left(I - \frac{t}{n}A\right)^{1-n} S\left(\frac{t}{n}\right),$$

and finally at

$$S(t) \stackrel{!}{=} \left(I - \frac{t}{n}A\right)^{-n} =: Z_n(t).$$

At this stage, this is still merely an Ansatz. *Hille's proof consistend in showing that in fact $Z_n(t)$ is for all $t \geq 0$ and all $n \in \mathbb{N}$ a contractive operator that converges strongly to a C_0-semigroup whose generator is A. Summing up, he proved that*

$$T(t)x = \lim_{n\to\infty} \left(I - \frac{t}{n}A\right)^{-n} x = \lim_{n\to\infty} \frac{n}{t} R\left(\frac{n}{t}, A\right)^n x, \qquad x \in X. \tag{3.2}$$

Theorem 3.9 (Lumer–Phillips 1961)**.** *Let H be a Hilbert space and A be a (possibly unbounded) linear operator on X. Then the following assertions are equivalent.*

(a) A is the generator of a C_0-semigroup of contractions.
(b) A is dissipative*, i.e.,*

$$\mathrm{Re}(x|Ax)_H \leq 0 \qquad \text{for all } x \in D(A)$$

holds, and $(\lambda - A)$ is surjective for all $\lambda > 0$.

Proof. We only have to prove that dissipativity is equivalent to the Hille–Yosida condition. Take $\lambda > 0$. Squaring the Hille–Yosida condition yields for all $\lambda > 0$ and all $y \in X$

$$\lambda^2 \|R(\lambda, A)y\|^2 \leq \|y\|^2,$$

or rather for $x := R(\lambda, A)y$

$$\begin{aligned} \lambda^2\|x\|^2 &\leq \|\lambda x - Ax\|^2 = (\lambda x - Ax|\lambda x - Ax)_H \\ &= \lambda^2\|x\|^2 - 2\lambda \mathrm{Re}\,(x|Ax)_H + \|Ax\|^2, \end{aligned}$$

i.e., validity of

$$2\mathrm{Re}\,(x|Ax)_H \leq \frac{1}{\lambda}\|Ax\|^2, \qquad x \in D(A),$$

for all $\lambda > 0$ is equivalent to the Hille–Yosida condition.

Hence, if the Hille–Yosida condition holds, dissipativity can be checked letting $\lambda \to \infty$. If conversely dissipativity holds, then

$$2\mathrm{Re}\,(x|Ax) \leq 0 \leq \frac{1}{\lambda}\|Ax\|^2, \qquad x \in D(A),$$

i.e., the Hille–Yosida condition is also checked repeating the above arguments. □

Remark 3.10. *The Lumer–Phillips Theorem can be extended to general Banach spaces X by replacing the inner product of H with the duality between X and X'; and $x \in X$ by all elements of the duality set $j(x) := \{x' \in X' : \|x'\|_X^2 = \|x'\|_{X'}^2 = \langle x, x'\rangle\}$, i.e., the Lumer–Phillips condition becomes*

$$\mathrm{Re}\,\langle x, Ax\rangle \leq 0 \qquad \text{for all } x \in D(A) \text{ and all } x' \in j(x).$$

Such a set is nonempty by the Hahn–Banach Theorem (why?) for all $x \in X$ – and a singleton if X is reflexive. We refer to [**3**, §II.3.b] *for the proof, which is longer and more technical than in the Hilbert space case.*

We have already seen that C_0-(semi)groups of unitary operators are particularly important for application in quantum mechanics.

Theorem 3.11 (Stone 1932)**.** *Let A be a densely defined operator on a Hilbert space H. Then iA is the generator of a C_0-group of unitary operators, i.e., the abstract Schrödinger equation* (2.10) *is well-posed, if and only if A is self-adjoint.*

Proof. Sufficiency has been proved in Theorem 2.73. In order to see that self-adjointness of A is also necessary, denote by $T(t))_{t\in\mathbb{R}}$ the C_0-group generated by iA. Since each $T(t)$ is unitary,

$$T(t)^* = T(t)^{-1} = T(-t)$$

by the semigroup law, which by Exercise 1.18 means that $(iA)^* = \bar{\imath}A^* = -iA^*$ is the generator of the C_0-group $(T(t)^*)_{t\in\mathbb{R}} = (T(-t))_{t\in\mathbb{R}}$. On the other hand, a direct computation shows that $-iA$ is the generator of $(T(-t))_{t\in\mathbb{R}}$, hence $-iA^* = -iA$. Accordingly, $A = A^*$. □

3.2. Long-time results for C_0-semigroups

While the Hille–Yosida theorem in its general version is a fundamental tool for the proofs of abstract generation results, it relies on a countable family of estimates which, in general, can only be checked in very special cases. One is – of course – the contractive case. Another one has been found in 1999 by Gomilko.

Theorem 3.12 (Gomilko 1999)**.** *Let X be a Banach space and A be a densely defined, closed operator on X. Assume that $\lambda \in \rho(A)$ for all real $\lambda > 0$ and that moreover the* Gomilko condition

$$\delta \int_{\delta+i\mathbb{R}} |\left\langle R(\lambda, A)^2 x, y\right\rangle| d\lambda \leq M,$$

holds for all $\delta > 0$, $x \in X$ and $y \in X'$ and some M that can depend on x, y but not on δ. Then A is the generator of a C_0-semigroup $(T(t))_{t\geq 0}$ such that

$$\|T(t)\| \leq \frac{Me}{2\pi}, \qquad t \geq 0.$$

(Here $\langle\cdot,\cdot\rangle$ denotes the duality between X and X').

Proof. First of all, observe that the Gomilko conditions is effectively a condition on the integrability of the *derivative* of the resolvent operator, due to the well-known relation

$$\frac{d}{d\lambda} R(\lambda, A) = -R(\lambda, A)^2, \qquad \lambda \in \rho(A).$$

Take $\delta > 0$, which we will fix later. It is well-known that the resolvent operator is holomorphic (hence weakly holomorphic) in the resolvent set, and moreover the Gomilko condition implies integrability of the complex function

$$\lambda \mapsto \left\langle R^2(\lambda + \delta, A)x, y\right\rangle$$

on the axis $\delta + i\mathbb{R}$ and in fact on the whole halfplane $\{z \in \mathbb{C} : \mathrm{Re} z \geq 0\}$, for all given $x \in X$ and $y \in X'$. Therefore, it is possible to apply Cauchy's integral formula and obtain for all $x \in X$ and all $y \in X'$

$$\left\langle R^2(\mu, A)x, y\right\rangle = \frac{1}{2\pi i}\int_{\delta+i\mathbb{R}} \frac{\left\langle R^2(\lambda, A)x, y\right\rangle}{\mu - \lambda} d\lambda, \qquad \mu > 0,\ \delta \in (0, \mu),$$

or rather, as already observed,

$$\left\langle \frac{d}{d\lambda} R(\mu, A)x, y\right\rangle = \frac{1}{2\pi i}\int_{\delta+i\mathbb{R}} \frac{d}{d\lambda}\left\langle R(\lambda, A)x, y\right\rangle \frac{d\lambda}{\lambda - \mu}, \qquad \mu > 0,\ \delta \in (0, \mu),$$

and in fact, by a well-known formula in function theory, for all $k \in \mathbb{N}$

$$\left\langle \frac{d^k}{d\lambda^k} R(\mu, A)x, y\right\rangle = \frac{1}{2\pi i k}\int_{\delta+i\mathbb{R}} \frac{d}{d\lambda}\left\langle R(\lambda, A)x, y\right\rangle \frac{d\lambda}{(\lambda - \mu)^k}, \qquad \mu > 0,\ \delta \in (0, \mu).$$

This was the main idea in the proof; the remainder mostly consists of convenient estimating. In fact, for all $\mu > 0$ and all $\delta \in (0, \mu)$

$$\begin{aligned}
\left|\left\langle \frac{d^k}{d\lambda^k} R(\mu, A)x, y\right\rangle\right| &\leq \frac{1}{2\pi k}\int_{\delta+i\mathbb{R}} \left|\frac{d}{d\lambda}\langle R(\mu, A)x, y\rangle\right| \frac{d\lambda}{|\lambda-\mu|^k} \\
&\leq \frac{1}{2\pi k(\mu-\delta)^k}\int_{\delta+i\mathbb{R}} \left|\frac{d}{d\lambda}\langle R(\lambda, A)x, y\rangle\right| d\lambda \\
&\leq \frac{M}{2\pi k\delta(\mu-\delta)^k},
\end{aligned}$$

where we have used the fact that $\mu - \delta \leq |\mu - \lambda|$ for all $\lambda \in \delta + i\mathbb{R}$. Since δ is still arbitrary, we can let $\delta = \frac{\mu}{k+1}$ and deduce

$$\begin{aligned}
\left|\left\langle \frac{d^k}{d\lambda^k} R(\mu, A)x, y\right\rangle\right| &\leq \frac{M}{2\pi\frac{\mu k}{k+1}(\mu - \frac{\mu}{k+1})^k} \\
&= \frac{M}{2\pi\frac{\mu k}{k+1}\mu^k(1-\frac{1}{k+1})^k} \\
&= \frac{M}{2\pi\mu^{k+1}\frac{k}{k+1}(\frac{k}{k+1})^k} \\
&= \frac{M}{2\pi\mu^{k+1}(\frac{k}{k+1})^{k+1}} \\
&= \frac{M}{2\pi\mu^{k+1}(1-\frac{1}{k+1})^{k+1}} \\
&= \frac{M}{2\pi\mu^{k+1}}\left(1-\frac{1}{k+1}\right)^{-(k+1)} \\
&\leq \frac{Me}{2\pi\mu^{k+1}}.
\end{aligned}$$

by monotony of the above sequence that converges to e. In other words

$$\left|\left\langle \mu^{k+1} R(\mu, A)^{k+1}x, y\right\rangle\right| \leq \frac{Me}{2\pi} \qquad \text{for all } x \in X,\ y \in X',$$

and hence by the uniform boundedness principle

$$\left\|\mu^{k+1} R(\mu, A)^{k+1}\right\| \leq \frac{Me}{2\pi} \qquad \text{for all } k = 1, 2, \ldots. \tag{3.3}$$

By the general Hille–Yosida Theorem, it is sufficient to check that this estimate holds also for $k = 0$. We make again use of (3.2) in order to obtain for all μ_1, μ_2 and all $x \in X$

$$R(\mu_1, A)x - R(\mu_2, A)x = \int_{\mu_2}^{\mu_1} \frac{d}{d\lambda} R(\lambda, A)x d\lambda = \int_{\mu_1}^{\mu_2} R(\lambda, A)^2 x d\lambda, \tag{3.4}$$

whence

$$\|R(\mu_1, A)x - R(\mu_2, A)x\| \leq \int_{\mu_1}^{\mu_2} \|R(\lambda, A)^2 x\| d\lambda \leq \frac{Me}{2\pi}\int_{\mu_1}^{\mu_2} \frac{1}{\lambda^2}\|x\| d\lambda = \frac{Me}{2\pi}\left(\frac{1}{\mu_1} - \frac{1}{\mu_2}\right)\|x\|$$

for all $x \in X$, which tends to 0 as $\mu_1, \mu_2 \to \infty$. Hence, by the Cauchy criterium we deduce that for all $x \in X$ $R(\mu, A)x$ converges as $\mu \to \infty$, and we denote by Rx this limit.

Let us show that $Rx = 0$ for all $x \in X$. In fact, if $x \in D(A)$, then $R(\lambda, A)x = R(\lambda, A)^2(\lambda - A)x$, so that

$$\|Rx\| = \lim_{\lambda\to\infty} \|R(\lambda, A)x\| \leq \lim_{\lambda\to\infty} \frac{Me}{2\pi}\frac{1}{\lambda^2}\|(\lambda - A)x\| \leq \lim_{\lambda\to\infty} \frac{Me}{2\pi}\frac{1}{\lambda^2}(\lambda\|x\| + \|Ax\|) = 0.$$

Now, letting $\mu_2 \to \infty$ in (3.4) we deduce that

$$R(\mu_1, A)x = \lim_{\mu_2 \to \infty} R(\mu_1, A)x - R(\mu_2, A)x = \int_{\mu_1}^{\infty} R(\lambda, A)^2 x d\lambda,$$

and therefore

$$\|R(\mu_1, A)x\| \leq \int_{\infty}^{\mu_1} \|R(\lambda, A)^2 x\| d\lambda \leq \int_{\mu_1}^{\infty} \frac{Me}{2\pi\lambda^2} \|x\| d\lambda = \frac{Me}{2\pi\mu_1} \|x\|,$$

i.e.,

$$\|\lambda R(\lambda, A)\| \leq \frac{Me}{2\pi}.$$

This is the sought-after Hille–Yosida estimate for $k = 0$, and we have hence concluded the proof. ☐

Proposition 3.13. *Let X be a reflexive Banach space. Then to each uniformly bounded C_0-semigroup $(T(t))_{t\geq 0}$ on X we can associate a generalised limit, i.e., a bounded linear operator $P := \mathrm{LIM}_{t\to\infty} T(t)$ such that*

1. $\mathrm{LIM}_{t\to\infty} T(t) = \lim_{t\to\infty} T(t)$ *for all weakly convergent C_0-semigroups;*
2. $\mathrm{LIM}_{t\to\infty}(T+S)(t) = \mathrm{LIM}_{t\to\infty} T(t) + \mathrm{LIM}_{t\to\infty} S(t)$ *for all uniformly bounded C_0-semigroups $(T(t))_{t\geq 0}$, $(S(t))_{t\geq 0}$;*
3. $\mathrm{LIM}_{t\to\infty} T(t+s) = \mathrm{LIM}_{t\to\infty} T(t)$ *for all uniformly bounded C_0-semigroups $(T(t))_{t\geq 0}$.*
4. $\liminf_{t\to\infty} \langle T(t)x, x'\rangle \leq \mathrm{LIM}_{t\to\infty} \langle\langle T(t)x, x'\rangle \leq \limsup_{t\to\infty} \langle T(t)x, x'\rangle$ *for all uniformly bounded C_0-semigroups $(T(t))_{t\geq 0}$ and all $x \in X$, $x' \in X'$.*

Proof. Observe that $L^\infty(\mathbb{R}_+, X)$ is a Banach space. Hence, it is possible to define a Banach limit on $L^\infty(\mathbb{R}_+, X)$ by considering the real-valued mappings

$$\mathbb{R}_+ \ni t \mapsto \langle T(t)x, x'\rangle \in \mathbb{C}, \qquad x \in X,\ x' \in X',$$

cf. [**7**, § 4.2]. For fixed $x \in X$ this defines a functional on X', i.e., $\mathrm{LIM}_{t\to\infty} T(t)x$ is an element of $X'' = X$. ☐

Of course, being based on the Hahn–Banach theorem the Banach limit of a bounded semigroup is useful only at a theoretical level: one usually desires more (and more constructive) information.

Definition 3.14. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a C_0-semigroup on X. Define*

$$C(r)x := \frac{1}{r}\int_0^r T(s)x, \qquad x \in X,\ r > 0.$$

These operators are called Cesaro means *of $(T(t))_{t\geq 0}$.*

The idea is that while a C_0-semigroup may be oscillating even very fast, averaging it over time may "regularise" its behaviour and show a convergence pattern. This is related to old ideas in thermodynamics and statistical mechanics (think of the properties of a Brownian motion).

To begin with, we recall a result which is a consequence of an extension of the Spectral mapping theorem to the case of arbitrary generators of semigroups: this is a wide topic and we will not go into details and refer to [**3**, Cor. IV.3.8] instead.

Lemma 3.15. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a C_0-semigroup on X with generator A. Consider the* fixed space

$$\mathrm{fix}(T(t))_{t\geq 0} := \{x \in X : T(t)x = x \text{ for all } t \geq 0\}$$

of $(T(t))_{t\geq 0}$. Then

$$\mathrm{Ker}\, A = \mathrm{fix}(T(t))_{t\geq 0}.$$

We are ready to state some elementary properties of the Cesaro means.

Lemma 3.16. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a C_0-semigroup on X with generator A. Then the following assertions hold.*

(a) For all $r > 0$ and all $x \in X$ the vector $C(r)x$ belongs to the closed convex hull of the orbit of x, i.e., of $\{T(t)x : t \geq 0\}$.
(b) For all $r, t > 0$ one has

$$C(r)(\mathrm{Id} - T(t)) = (\mathrm{Id} - T(t))C(r) = \frac{1}{r}(\mathrm{Id} - T(r)) \int_0^t T(s)ds.$$

(c) Assume that $\lim_{r\to\infty} \|T(r)\| = 0$. Then $\lim_{r\to\infty} C(r)x \in \mathrm{fix}(T(t))_{t\geq 0}$, whenever such a limit exits.

Proof. In order to show (a), observe that this is just the vector-valued analogon of a classical property of Riemann integrals (it is essentially the integral version of the mean value theorem) and can be proved likewise; while (b) follows from (SEMIGR).

In order to prove (c), let $y := \lim_{r\to\infty} C(r)x$. We have to show that $T(t)y = y$, i.e., that $(\mathrm{Id} - T(t))y = 0$ for all $t \geq 0$. One has for all (fixed) $t \geq 0$

$$\begin{aligned}(\mathrm{Id} - T(t))y &= \lim_{r\to\infty} (\mathrm{Id} - T(t))C(r)x \\ &= \lim_{r\to\infty} \frac{1}{r}(\mathrm{Id} - T(r)) \int_0^t T(s)xds \\ &= \lim_{r\to\infty} \frac{1}{r} \int_0^t T(s)xds - \lim_{r\to\infty} \frac{1}{r} T(r) \int_0^t T(s)xds,\end{aligned}$$

whence

$$\|(\mathrm{Id} - T(t))y\| \leq \lim_{r\to\infty} \frac{1}{r} \left\| \int_0^t T(s)xds \right\| - \lim_{r\to\infty} \frac{\|T(r)\|}{r} \left\| \int_0^t T(s)xds \right\| = 0,$$

i.e,

$$(\mathrm{Id} - T(t))y = \lim_{r\to\infty} (\mathrm{Id} - T(t))C(r)x = 0, \tag{3.5}$$

as we wanted to prove. □

Definition 3.17. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a C_0-semigroup on X. Then $(T(t))_{t\geq 0}$ is called* mean ergodic *if the Cesaro means $C(r)x$ converge as $r \to \infty$ for all $x \in X$, and in this case the operator*

$$P : X \ni x \mapsto \lim_{r\to\infty} C(r)x \in X$$

is called mean ergodic projector *of the semigroup.*

Lemma 3.18. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a mean ergodic C_0-semigroup on X. If $\lim_{r\to\infty} \|T(r)\| = 0$, then the mean ergodic projector is an orthogonal projector that commutes with each operator of the semigroup. Its range is $\mathrm{Ker} A$ and its null space is $\overline{\mathrm{Rg}\, A}$.*

Proof. Let $t \geq 0$. One has

$$P - T(t)P = \lim_{r\to\infty} (\mathrm{Id} - T(t)C(r) = 0,$$

by Lemma 3.16.(b) and in particular by 3.5, and therefore also

$$P - PT(t) = \lim_{r\to\infty} C(r)(\mathrm{Id} - T(t)) = 0.$$

Accordingly, for all $r > 0$

$$P = T(r) = PC(r),$$

and hence

$$P = \lim_{r\to\infty} PC(r) = P^2,$$

by definition. Accordingly, P is a projector. Its range is $\mathrm{Ker}\, A$ by Lemma 3.16.(c). In order to show that P is an *orthogonal* projector, it suffices to show that its null space is orthogonal to $\mathrm{fix}(T(t))_{t\geq 0}$.

In fact, observe that by Lemma 3.16.(b) each orbit of $(\mathrm{Id} - T(t))_{t\geq 0}$ is contained in the null space of P, hence also the span of all orbits of $(\mathrm{Id} - T(t))_{t\geq 0}$, and (due to closedness of $\mathrm{Ker}\, P$) also the closure $\mathcal{O}$ of such a span, i.e., $\mathcal{O} \subset \mathrm{Ker}\, P$. We are going to show that the set $\mathcal{O}$ actually *agrees* with $\mathrm{Ker}\, P$, thus concluding the proof.

Assume to this aim that $\mathrm{Ker}\, P \not\subset \mathcal{O}$, i.e., assume that there exists $z \in \mathrm{Ker}\, P \setminus \mathcal{O}$. By the Hahn–Banach Theorem there exists some functional x' that separates $\mathcal{O}$ (and in particular the set of all elements of orbits of $(\mathrm{Id} - T(t))_{t\geq 0}$, i.e., of all vectors of the form $x - T(t)x$, for some $x \in X$ and some $t \geq 0$) from $\{z\}$.

However, if $\langle x - T(t)x, x'\rangle = 0$ for all $x \in X$ and all $t \geq 0$, then

$$0 = \langle x - T(t)x, x'\rangle = \langle x, x' - T(t)^*x'\rangle, \qquad x \in X,\ t \geq 0,$$

hence $x' = T(t)^*x'$ for all $t \geq 0$, i.e., x' is a fixed point of $(T(t)^*)_{t\geq 0}$, where T^* denotes the Banach space adjoint of an operator T. Accordingly,

$$\langle x, x'\rangle = \langle x, T(t)^*x'\rangle = \langle T(t)x, x'\rangle, \qquad x \in X,\ t \geq 0,$$

hence in particular

$$\langle x, x'\rangle = \frac{1}{r}\int_0^r \langle T(s)x, x'\rangle ds = \langle C(r)x, x'\rangle, \qquad x \in X,\ r \geq 0.$$

Thus, in particular $\langle x, x'\rangle = 0$ for all $x \in \mathrm{Ker}\, P$, i.e., for all x such that $Px = \lim_{r\to\infty} C(r)x = 0$ – and in particular for $x = z$. This yields a contradiction. □

Proposition 3.19. *Let X be a Banach space and $(T(t))_{t\geq 0}$ a uniformly bounded C_0-semigroup on X with generator A. Denote by $(C(r))_{r\geq 0}$ the associated Cesaro means. Then the following assertions are equivalent.*

(i) *$(T(t))_{t\geq 0}$ is mean ergodic (i.e., $(C(r))_{r\geq 0}$ converges strongly).*
(ii) *$(C(r))_{r\geq 0}$ converges weakly.*
(iii) *For each $x \in X$ there is a monotone increasing unbounded sequence $(r_n)_{n\in\mathbb{N}}$ such that $(C(r_n)x)_{n\in\mathbb{N}}$ has a weak accumulation point.*
(iv) *For each $x \in X$ the closed convex hull of the orbit $\{T(t)x : t \geq 0\}$ is not disjoint from the fixed space of $(T(t))_{t\geq 0}$.*
(v) *The fixed space of $(T(t))_{t\geq 0}$ separates the fixed space of the dual semigroup $(T(t)^*)_{t\geq 0}$, i.e, for any two different $x', y' \in \mathrm{Ker}\, A^*$ there exist $x \in \mathrm{Ker}\, A$ such that $\langle x, x'\rangle \neq \langle x, y'\rangle$.*

In the proof we will need the following elementary fact.

Lemma 3.20. *Let X be a Banach space. Then all convex closed subsets of X are also weakly closed.*

Proof. Let A be a convex closed subset of X. For any $x \in X \setminus A$ there is, by the Hahn–Banach Theorem, some $\phi_x \in X'$ that separates A and $\{x\}$, say, such that $\phi_x(A) \leq 0$ and $\phi(x) > 0$, i.e., there is a weakly closed set $\phi_x^{-1}(\infty, 0]$ such that $A \subset C_x$ but $x \notin C_x$. Then

$$A = \bigcap_{x\in X\setminus A} \phi_x^{-1}(\infty, 0]$$

is weakly closed. □

Proof of Proposition 3.19. The implications (i)⇒(ii)⇒(iii) are obvious.

To prove (iii)⇒(iv), let $x \in X$. Take a monotone unbounded sequence $(r_n)_{n\in\mathbb{N}}$ as in (ii) and take some y that is element of the weak closure of $\{C(r_n)x : n \geq m\}$ for all $m \in \mathbb{N}$. Due to weak-continuity of the semigroup, we also have that, for all $t \geq 0$, $y - T(t)y$ is element of the weak closure of

$$\{(\mathrm{Id} - T(t))C(r_n)x : n \geq m\} = \left\{\frac{1}{r_n}(\mathrm{Id} - T(t))\int_0^t T(s)xds : n \geq m\right\}$$

(the equality being valid due to Lemma 3.16) and hence in the ball with center 0 and radius

$$\frac{1}{r_m}(tM + tM^2)\|x\|,$$

for all $m \in \mathbb{N}$ (here we have denoted by M the uniform bound $\{T(t)x : t \geq 0\}$). In other words,

$$\|y - T(t)y\| \leq \frac{1}{r_m}(tM + tM^2)\|x\| \qquad \text{for all } m \in \mathbb{N},$$

for a divergent sequence $(r_m)_{m\in\mathbb{N}}$, i.e., $y = T(t)y$. We have proved that y belongs to the fixed space of $(T(t))_{t\geq 0}$: It remains to observe that because by Lemma 3.16.(a) $C(r_n)x$ belongs to the closed convex hull of $\{T(t)x : t \geq 0\}$, which (as a convex set) is therefore also *weakly* closed, by Lemma 3.20. Hence, also each element of the weak closure of $\{C(r_n)x : n \geq m\}$ for all $m \in \mathbb{N}$ (and in particular y) are element of this weakly closed set.

In order to show that (v) holds, take two different vectors $x', y' \in \operatorname{Ker} A'$ and let x_0 be a vector in X that separates them. In order to find a vector in $\operatorname{Ker} A$ that also separates them, observe that by (iv) we can take z that is both in the closed convex hull of the orbit $\{T(t)x_0 : t \geq 0\}$ and in $\operatorname{Ker} A$. Since x', y' are in the fixed space of $(T(t)^*)_{t\geq 0}$, one has

$$\langle T(t)x_0, x'\rangle = \langle x_0, T(t)^*x'\rangle = \langle x_0, x'\rangle \qquad \text{for all } t \geq 0,$$

and similarly $\langle T(t)x_0, y'\rangle = \langle x_0, y'\rangle$ for all $t \geq 0$, and hence x', y' are constant also on the closed convex hull of the orbit $\{T(t)x_0 : t \geq 0\}$, and in particular $\langle z, x'\rangle = \langle x_0, x'\rangle$ and $\langle z, y'\rangle = \langle x_0, y'\rangle$, hence $z \in KernA$ separates x', y' because so does x_0.

Finally, let us show that (v)⇒(i). First, observe that the Cesaro means $C(r)y$ converge as $r \to \infty$ both for y in the fixed space of $(T(t))_{t\geq 0}$ (of course) and for y in some orbit, i.e., for y of the form $y = x - T(t)x$, $t \geq 0$, $x \in X$, since by Lemma 3.16.(b)

$$\lim_{r\to\infty} \|C(r)(x - T(t)x)\| \leq \lim_{r\to\infty} \frac{1}{r}\|(\mathrm{Id} - T(t))\| \int_0^t \|T(s)x\|ds = 0,$$

by boundedness of $(T(t))_{t\geq 0}$. hence the Cesaro means $C(r)y$ converge as $r \to \infty$ for all $y \in X_0$, where X_0 is the direct sum of the fixed space of $(T(t))_{t\geq 0}$ and of the span of the set of all orbits of $(\mathrm{Id} - T(t))_{t\geq 0}$. Now, we are going to prove that X_0 is dense in X: this will yield that, by density, the Cesaro means $C(r)y$ converge as $r \to \infty$ for all $y \in X$, i.e., that $(T(t))_{t\geq 0}$ is mean ergodic.

To proved the claimed density result, take $x' \in X'$ such that

$$\langle x, x'\rangle = 0 \qquad \text{for all } x \in X_0.$$

If we can prove that then $x' = 0$, then the assertion follows. In fact, it suffices to observe that because

$$0 = \langle x - T(t)x, x'\rangle = \langle x, (\mathrm{Id} - T(t)^*)x'\rangle \qquad \text{for all } x \in X,\ t \geq 0,$$

x' belongs to the fixed space of $(T(t)^*)_{t\geq 0}$. Now, by assumption the fixed space of $(T(t))_{t\geq 0}$ separates the fixed space of $(T(t)^*)_{t\geq 0}$, and in particular x' from the functional of constant value 0: if there would hold $x' \neq 0$, then there should exist x in the fixed space of $(T(t))_{t\geq 0}$ such that $\langle x, x'\rangle \neq \langle x, 0\rangle = 0$. But this is impossible, because the fixed space of $(T(t))_{t\geq 0}$ is contained in X_0, which in turn is contained in the null space of x' by assumption. This concludes the proof. □

Corollary 3.21. *Let X be a reflexive Banach space and $(T(t))_{t\geq 0}$ a uniformly bounded C_0-semigroup on X. Then $(T(t))_{t\geq 0}$ is mean ergodic.*

Proof. By the Banach–Alaoglu Theorem, the bounded set $\{T(t) \in \mathcal{L}(X) : t \geq 0\}$ is relatively compact in the weak operator topology. Now, it suffices to observe that condition (iii) in Theorem 3.19 is satisfied. □

Exercise 3.22. *Let $\alpha > 0$. Show that a C_0-semigroup $(T(t))_{t\geq 0}$ on a Banach space X is α*-periodic*, i.e.,*

$$T(t+\alpha) = T(t) \qquad \text{for all } t \geq 0,$$

provided that each eigenvalue of its generator A is contained in the set $\frac{2\pi i}{\alpha}\mathbb{Z}$ and the correspondent eigenvectors are total in X, i.e., their span is dense in X. (Hint: Take $x \neq 0$, t and n such that $\frac{2\pi i n}{\alpha} \in \sigma_p(A)$, consider the function $Q : s \mapsto e^{\frac{2\pi i n}{\alpha}(t-s)} T(s)x$ and use an argument similar to that used to prove Lemma 1.22.)

Exercise 3.23. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a C_0-semigroup on X with generator A. Show that if $\lambda \in \sigma_p(A)$, then $e^{t\lambda} \in \sigma_p(T(t))$.* (Hint: Show that the semigroup rescaled by λ is constant.)

Semigroup theory is not convenient for applications that require an exact knowledge of solutions to evolution equations. Instead, one of its main features is the possibility of investigating several qualitative properties of solutions, like convergence to stationary solutions or invariance of relevant subsets of the state space X.

Proposition 3.24. *Let X be a Banach space and $(T(t))_{t\geq 0}$ be a C_0-semigroup of contractions on X with generator A. Let C be a closed conves subset of X. Then $T(t)C \subset C$ for all $t \geq 0$ if and only if $\lambda R(\lambda, A)C \subset C$ for all real $\lambda > \omega$.*

While we only state and prove this theorem in the contractive case, but refer to [**8**] for an extension to the general case.

Proof. Let $T(t)C \subset C$ for all $t \geq 0$. Let $\lambda > 0$ and $x \in X$. By Lemma 3.1 one has

$$\lambda R(\lambda, A)x = \int_0^\infty \lambda e^{-\lambda s} T(s)x ds.$$

If $\lambda R(\lambda, A)C \not\subset C$, then by the Hahn–Banach theorem C and $\lambda R(\lambda, A)C$ can be separated, i.e., there exist $\alpha \in \mathbb{R}$ and a functional $\phi \in X'$ such that

$$\mathrm{Re}\,\langle \lambda R(\lambda, A)x, \phi\rangle > \alpha \geq \mathrm{Re}\,\langle y, \phi\rangle \qquad \text{for all } y \in C.$$

Taking in particular $y := T(t)x$ we obtain

$$\begin{aligned} \mathrm{Re}\,\langle \lambda R(\lambda, A)x, \phi\rangle \quad > \alpha \quad &= \int_0^\infty \lambda e^{-\lambda s} x ds \\ &\geq \quad \mathrm{Re}\left\langle \int_0^\infty \lambda e^{-\lambda s} T(t)x, \phi \right\rangle \\ &\geq \quad \mathrm{Re}\,\langle \lambda R(\lambda, A)x, \phi\rangle . \end{aligned}$$

Conversely, let $\lambda > \omega$. If $\lambda R(\lambda, A)C \subset C$, then $\lambda^n R(\lambda, A)^n C \subset C$ for all $n \in \mathbb{N}$, and by (3.2) we obtain that also

$$T(t)x = \lim_{n\to\infty} \frac{n}{t} R\left(\frac{n}{t}, A\right)^n x \in C,$$

due to closedness of C. □

Exercise 3.25. *Let H_1, H_2 be Hilbert spaces and T be a bounded linear operator from H_1 to H_2.*
(1) Show that the graph *of T, i.e., $\{(f, g) \in H_1 \times H_2 : Tf = g\}$ is a closed subspace of $H_1 \times H_2$.*
(2) Prove that the subspace of $H_1 \times H_2$ orthogonal to the graph of T is given by $\{(f, g) \in H_1 \times H_2 : f = -T^* g\}$.
(3) Show that both $I + TT^*$ and $I + T^*T$ are invertible.
(4) Conclude that the orthogonal projector of $H_1 \times H_2$ onto the graph of T is given[2] by the operator matrix

$$\begin{pmatrix} (I + T^*T)^{-1} & T^*(I + TT^*)^{-1} \\ T(I + T^*T)^{-1} & \mathrm{Id} - (I + TT^*)^{-1} \end{pmatrix}.$$

[2]This has been first obtained in 1950 by mathematician, quantum physicist and early computer scientist John von Neumann ([**10**]) and can be even generalised to the case of an unbounded linear operator O, under the sole assumption that the graph of O be closed.

Exercise 3.26. *Let H be a Hilbert space and $(T(t))_{t\geq 0}$ be a C_0-semigroup on H with generator A and associated sesquilinear form (a, V). Let S be a further bounded linear operator on H. Show that $T(t)$ commutes with S for each $t \geq 0$ if and only if the semigroup*

$$\begin{pmatrix} T(t) & 0 \\ 0 & T(t) \end{pmatrix}, \qquad t \geq 0,$$

leaves the graph of S invariant. Deduce from Exercise 3.25 a criterion on A for commutation of $T(t)$ and S.

Bibliography

[1] Wolfgang Arendt, Charles J.K. Batty, Matthias Hieber, Frank Neubrander, *Vector-Valued Laplace Transforms and Cauchy Problems*, Birkhäuser, Basel, 2001.

[2] Jens Bolte, Jonathan Harrison, *Spectral statistics for the Dirac operator on graphs*, J. Phys. A **36** (2003), 2747–2769.

[3] Klaus-Jochen Engel, Rainer Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer-Verlag, Berlin, 2000.

[4] Lawrence C. Evans, *Partial Differential Equations*, AMS, Providence (RI), 1998.

[5] Paul R. Halmos, *What does the spectral theorem say?*, Am. Math. Mon. **70** (1963), 241–247.

[6] Tosio Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1995.

[7] Peter D. Lax, *Functional Analysis*, Wiley, Hoboken, NJ, 2002.

[8] Amir Manavi, Hendrik Vogt, Jürgen Voigt, *Domination of semigroups associated with sectorial forms*, J. Oper. Theory **54** (2005), 9–25.

[9] Martin Mathieu, *Funktionalanalysis*, Spektrum, Heidelberg, 1999.

[10] John von Neumann, *Functional Operators. Vol. II. The Geometry of Orthogonal Spaces*, Princeton University Press, Princeton, 1950.

[11] Walter Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

[12] Herbert Schröder, *Funktionalanalysis*, Akademie Verlag, Berlin, 1997.

[13] Berndt Thaller, *The Dirac Equation*, Springer-Verlag, 1992.