Semigroup theory in quantum mechanics

Jesús Aguado López
Universidad Autónoma de Madrid
jesagulop@gmail.com

Antonio Ismael Cano Mármol
Universidad de Murcia
ismael.cano@icmat.es

Álvaro Carballido Costas
Universidade de Santiago de Compostela
alvaro.carballido@rai.usc.es

Miguel García Fernández
Universitat de Barcelona
mgarcife42@alumnes.ub.edu

Francesc Gómez Marín
Universitat de València
gomafran@alumni.uv.es

Héctor Jardón-Sánchez
Universidad de Oviedo
h.jardonsanchez@lancaster.ac.uk

Antsa Ratsimanetrimanana
Basque Center for Applied Mathematics
University of the Basque Country
aratsimanetrimanana@bcamath.org

Course instructor

Jean-Bernard Bru
Basque Center for Applied Mathematics
jb.bru@ikerbasque.org

Abstract: In mathematics, the concept of strongly continuous one-parameter semigroup (C^0-semigroup) appears intuitively to be the generalization of the (usual) exponential function. Roughly speaking, this class of semigroups provides solutions of linear ordinary differential equations with constant coefficients in Banach spaces, see *Schrödinger equation* (1)–(2). Semigroup theory turns out to be fundamental in order to understand the time evolution in quantum mechanics, and is necessary in order to generate the dynamics of both well-known formulations (*Shrödinger picture* and *Heisenberg picture*). Within this paper, the main result that we present is the *Hille–Yosida theorem*, see section 5, which characterizes the generators of C^0-semigroups of linear operators on Banach spaces. It is named after the mathematicians Einar Hille and Kōsaku Yosida who independently stated it around 1948. This manuscript is highly inspired by Engel and Nagel's notes [2].

Resumen: En matemáticas, el concepto de semigrupo uniparamétrico fuertemente continuo (C^0-semigrupo) puede entenderse intuitivamente como generalización de la función exponencial. A grandes rasgos, esta clase de semigrupos ofrece soluciones a ecuaciones diferenciales ordinarias con coeficientes constantes en espacios de Banach, véase la *ecuación de Schrödinger* (1)–(2). La teoría de semigrupos resulta fundamental a la hora de comprender la evolución temporal en mecánica cuántica, y es necesaria para generar la dinámica de ambas formulaciones conocidas (*imagen de Schrödinger* e *imagen de Heisenberg*). En este artículo, el principal resultado presentado es el *teorema de Hille-Yosida*, ver sección 5, que caracteriza los generadores de los C^0-semigrupos de operadores lineales sobre espacios de Banach. Este teorema debe su nombre a los matemáticos Einar Hille y Kōsaku Yosida, quienes lo enunciaron independientemente en torno a 1948. El presente texto se inspira altamente en el libro de Engel y Nagel [2].

Keywords: Hille-Yosida theorem, semigroup theory, quantum mechanics.

MSC2010: 46N50, 47D03, 47D06, 58D25, 82C10.

Acknowledgements: The authors would like to thank the ANEM for the promotion of this journal as well as the editorial board. The Network on Functional Analysis and Applications organized the VIII School in Functional Analysis at the Basque Center for Applied Mathematics, where this project originated under the supervision of Jean-Bernard Bru, who helped the authors with revisions on style and content.

Antonio Ismael Cano Mármol was supported by the Department of Mathematics of the University of Murcia.

Héctor Jardón-Sánchez was supported by the María Cristina Masaveu Peterson Scholarship for Academic Excellence.

Reference: Aguado López, Jesús; Cano Mármol, Antonio Ismael; Carballido Costas, Álvaro; García Fernández, Miguel; Gómez Marín, Francesc; Jardón-Sánchez, Héctor; Ratsimanetrimanana, Antsa, and Bru, Jean-Bernard. "Semigroup theory in quantum mechanics". In: *TEMat monográficos*, 1 (2020): *Artículos de las VIII y IX Escuela-Taller de Análisis Funcional*, pp. 17-31. ISSN: 2660-6003. URL: https://temat.es/monograficos/article/view/vol1-p17.

This work is distributed under a Creative Commons Attribution 4.0 International licence
https://creativecommons.org/licenses/by/4.0/

Notations

- For a normed generic vector space $\mathcal{X}$, its norm is denoted by $\|\cdot\|_{\mathcal{X}}$.
- The identity element of a generic vector space $\mathcal{X}$ is denoted by $\mathbf{1}_{\mathcal{X}}$.
- The set of linear operators from $\mathcal{X}$ into $\mathcal{X}$ is denoted by $\mathcal{L}(\mathcal{X})$.
- The set of all bounded linear operators on $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is denoted by $\mathcal{B}(\mathcal{X})$. For an operator $A \in \mathcal{B}(\mathcal{X})$, its norm is defined by

$$\|A\|_{\mathcal{B}(\mathcal{X})} := \sup_{u \in \mathcal{X}} \frac{\|Au\|_{\mathcal{X}}}{\|u\|_{\mathcal{X}}}.$$

- If $\mathcal{X}$ is a Hilbert space, then its norm is associated to a scalar product denoted by $\langle \cdot, \cdot \rangle_{\mathcal{X}}$.
- For all $A, B \in \mathcal{B}(\mathcal{X})$, we define

$$[A, B] := AB - BA \qquad \text{and} \qquad \{A, B\} := AB + BA.$$

- For any complex number z, its conjugate is denoted by $\overline{z}$.

1. Introduction

The foundations of quantum mechanics were established during the first half of the 20th century. In the mid-twenties, two main formulations of quantum physics appeared, both meant to establish the principles of quantum theory. These two directions were taken by W. K. Heisenberg and by E. Schrödinger, respectively. After being in opposition, they turned out to be equivalent after several contributions of J. von Neumann on the foundation of quantum mechanics in the following years. Both formulations are currently used in any standard textbook on quantum physics. For the sake of clarity, we will first set the so-called *Schrödinger picture* of quantum mechanics. Indeed, it is widely known, used and commented in fields such as that of partial differential equations (PDEs), for instance, through the celebrated *Schrödinger equation*.

2. Schrödinger picture of quantum mechanics

In 1925, following de Broglie's hypothesis on wave property of matter, E. Schrödinger derived his celebrated equation, describing a time-dependent wave behavior of quantum objects. In fact, the state of the quantum system is completely described by a family of time dependent wave functions $\{\psi(t)\}_{t\in\mathbb{R}}$ within a Hilbert space $\mathcal{H}$. For instance, for the one-particle case, one generally considers the case $\mathcal{H} := L^2(\mathbb{R}^3)$ or $\mathcal{H} := \ell^2(\mathbb{Z}^3)$, respectively, for the continuum quantum system or the discrete one. This time evolution is fixed by a self-adjoint operator H acting on $\mathcal{H}$. Indeed, for any time $t \in \mathbb{R}$, the wave function is determined by the well-known *Schrödinger equation*:

$$(SE) \quad \begin{cases} \mathrm{i}\partial_t \psi(t) = H\psi(t), \\ \psi(0) = \psi_0 \in \mathcal{H}. \end{cases} \tag{1}$$

This implies that

$$\psi(t) = \mathrm{e}^{-\mathrm{i}tH}\psi_0, \qquad t \in \mathbb{R}. \tag{2}$$

Note that the fact that H is self-adjoint is important to give a sense to equations (1) and (2). It is described through Stone's theorem, see theorem 19, which sets that having a self-adjoint operator, acting on some Hilbert space, is a sufficient condition in order to define a strongly continuous one-parameter group (also denoted C_0-group). We will say some words on them later but, at this point, the aim is to give an intuition to the reader about the different ways to formulate quantum mechanics. A standard example taught to every student in quantum mechanics is brought by the case where $\mathcal{H} := L^2(\mathbb{R}^3)$ and $\|\psi(t)\|_{\mathcal{H}} = \|\psi_0\|_{\mathcal{H}} = 1$. Then, $|\psi(t,x)|^2$ is interpreted as the probability for the particle to be at a position $x \in \mathbb{R}^3$ at time $t \in \mathbb{R}$. As mentioned in the introduction above, a widely studied standard example in the field of PDEs is given by the case where the operator $H := -\Delta$ (the usual Laplacian operator). The same interpretation can be done on the lattice $\mathbb{Z}^3$, instead of taking $\mathbb{R}^3$.

3. Heisenberg picture of quantum mechanics

Physical quantities such as position, speed, energy, etc., are self-adjoint operators acting on the Hilbert space $\mathcal{H}$. They are called *observables*, being all quantities of the physical system that can be measured. For instance, one of the most important observables is the celebrated self-adjoint *Hamiltonian* H that describes the time evolution of the wave function in the Schrödinger equation (1)–(2). This Hamiltonian is associated with the energy observable.

The measurement of a physical quantity (observable) has, from this point of view, a random character. The statistical distribution of its value is described by the family of wave functions $\{\psi(t)\}_{t\in\mathbb{R}}$ (see equation (1)). The expectation value of any observable B acting on $\mathcal{H}$ is given by

$$\langle \psi(t), B\psi(t)\rangle_{\mathcal{H}}.$$

By equation (2), it equals

$$\langle \psi(t), B\psi(t)\rangle_{\mathcal{H}} = \left\langle \psi_0, \mathrm{e}^{\mathrm{i}tH} B \mathrm{e}^{-\mathrm{i}tH}\psi_0\right\rangle_{\mathcal{H}}. \tag{3}$$

At this point, it turns out that, instead of considering the wave functions as being time-dependent, like in the Schrödinger picture of quantum mechanics, one can take them as fixed in time and assume a time evolution of the so-called observables. Both methods lead to the same statistical distribution as one can see in equation (3). Indeed, for the time evolution of any observable B, we apply on it the map $\tau_t(B) := \mathrm{e}^{\mathrm{i}tH} B \mathrm{e}^{-\mathrm{i}tH}$ for $t \in \mathbb{R}$. For an operator H acting on the Hilbert space $\mathcal{H}$, the family $\{\tau_t\}_{t\in\mathbb{R}}$ defines a strongly continuous group acting on $\mathcal{B}(\mathcal{H})$ and satisfies the following evolution equation for all $t \in \mathbb{R}$:

$$\partial_t \tau_t = \tau_t \circ \delta = \delta \circ \tau_t, \qquad \tau_0 = \mathbf{1}_{\mathcal{B}(\mathcal{H})}, \tag{4}$$

where $\mathbf{1}_{\mathcal{B}(\mathcal{H})}$ is the identity operator on $\mathcal{B}(\mathcal{H})$ and the generator δ is defined on some dense subset $\mathcal{D}$ of $\mathcal{B}(\mathcal{H})$. Note that, if H is a bounded operator on $\mathcal{H}$, then $\mathcal{D} = \mathcal{B}(\mathcal{H})$ and

$$\delta(B) := i\,[H, B], \qquad B \in \mathcal{B}(\mathcal{H}).$$

$\{\tau_t\}_{t\in\mathbb{R}}$ is a family of isomorphisms of $\mathcal{B}(\mathcal{H})$ and, for all $A, B \in \mathcal{D}$, one has

$$\delta(A^*) = \delta(A)^* \qquad \text{and} \qquad \delta(AB) = \delta(A)B + A\delta(B). \tag{5}$$

An operator satisfying (5) is called a *symmetric derivation* or $*$*-derivation*. A^* is the usual adjoint operator of A. Once again, more precise definitions of the mathematical tools that are involved to formulate quantum time evolution will be given later, since it is not necessary for the moment. Indeed, the aim of this section is to give the readers intuition about the different approaches that can be taken. At this point, the knowledge of semigroup properties turns out to be fundamental in order to understand the dynamics in quantum mechanics. Observe that, in the Schrödinger picture, one has a semigroup acting on a Hilbert space, while in the case of the Heisenberg picture, the semigroup acts on a Banach space. Within the next sections, we introduce the main results in relation to semigroup theory.

4. Semigroups and generators

First of all, let us give the definitions and basic results of semigroup theory as they are given in Engel and Nagel's notes [2]. These will provide the basis required to prove the main theorems studied in this article. Let X be a Banach space.

Definition 1. A **strongly continuous one-parameter semigroup**, also called C_0-semigroup, is a family $(T(t))_{t\geq 0}$ of bounded operators $T(t): X \to X$ satisfying the functional equation

$$\begin{cases} T(t+s) = T(t)T(s) & \text{for all } t, s \geq 0, \\ T(0) = \mathbf{1}_X \end{cases}$$

and the strong continuity property, which is nothing else but the continuity of the orbit maps

$$\begin{array}{rcl} \xi_x : \mathbb{R}^+ & \longrightarrow & X \\ t & \longmapsto & \xi_x(t) := T(t)x \end{array}$$

for each $x \in X$. If these properties hold not only in $\mathbb{R}^+$ but also in $\mathbb{R}$, we call $(T(t))_t$ a **strongly continuous group**, or C_0-group. ◂

Lemma 2. *Let $(T(t))_{t\geq 0}$ be a C_0-semigroup. Then, there exist $\omega \in \mathbb{R}$ and $M \geq 1$ such that, for all $t \geq 0$,*

$$\|T(t)\|_{\mathcal{B}(X)} \leq M\,\mathrm{e}^{\omega t}.$$

Proof. From the uniform boundedness, there exists $M \geq 1$ such that $\|T(s)\| \leq M$ for all $0 \leq s \leq 1$. Writing any $t \geq 0$ as $t = s + n$ with $n \in \mathbb{N}$ and $s \in [0, 1]$,

$$\|T(t)\|_{\mathcal{B}(X)} \leq \|T(s)\|_{\mathcal{B}(X)}\, \|T(1)\|^n_{\mathcal{B}(X)} \leq M^{n+1} = M\mathrm{e}^{n \log M} \leq M\mathrm{e}^{wt}$$

holds for $\omega := \log M$ and $t \geq 0$. ■

Definition 3. If lemma 2 holds for $\omega = 0$ and $M = 1$, the semigroup is called **contractive**. It means that $\|T(t)\|_{\mathcal{B}(X)} \leq 1$ for all $t \geq 0$. ◂

Example 4. Let $\mathcal{H}$ be a Hilbert space, $A \in \mathcal{B}(\mathcal{H}) := X$. It can be easily shown that the series

$$\mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{1}{n!}(tA)^n$$

converges and that $T(t) := \mathrm{e}^{tA}$ defines a C_0-group. From the triangle inequality, we deduce that

$$\|T(t)\|_{\mathcal{B}(X)} \leq \mathrm{e}^{t\|A\|_X},$$

and therefore lemma 2 holds for $M = 1$ and $\omega = \|A\|_X \in \mathbb{R}$. ◂

Remark 5 (abstract Cauchy problem). In example 4, we have been able to define a C_0-group from a bounded operator. This group satisfies

$$\begin{cases} \dot{T}(t) = AT(t) & \text{for all } t \geq 0, \\ T(0) = \mathbf{1}_X. \end{cases} \tag{6}$$

The main topic studied in this article is the existence and properties of such an A for a general C_0-semigroup $(T(t))_{t\geq 0}$ by using the abstract Cauchy problem (6). ◂

Definition 6. A C_0-semigroup $(T(t))_{t\geq 0}$ is called **uniformly continuous** if the map

$$\begin{array}{rcl} \mathbb{R}^+ & \longrightarrow & \mathcal{B}(X) \\ t & \longmapsto & \|T(t)\|_{\mathcal{B}(X)} \end{array}$$

is continuous. ◂

Proposition 7. *Let $(T(t))_{t\geq 0}$ be a uniformly continuous semigroup. Then, there exists a bounded operator A on X such that $T(t) = \mathrm{e}^{tA}$ for all $t \geq 0$.*

For more details, see theorem 2.12 in Engel and Nagel's notes [2]. Within this article, we focus our study on the general case of strong continuity. In this case, the existence of such a bounded operator A requires a deeper study of operator semigroups, see remark 5. We start by defining the generator of a C_0-semigroup.

Definition 8 (generator). The **generator** $A : D(A) \subseteq X \to X$ of a C_0-semigroup $(T(t))_{t\geq 0}$ is the operator

$$Ax := \dot{\xi}_x(0) = \lim_{h\downarrow 0} \frac{1}{h}(T(h)x - x)$$

with domain

$$D(A) = \left\{ x \in X : \lim_{h\downarrow 0} \frac{1}{h}(T(h)x - x) \text{ exists} \right\}.$$ ◂

Note that the orbit map ξ_x is differentiable on $\mathbb{R}^+$ if and only if it is right-differentiable at $t = 0$. Indeed, the derivative of $\xi_x(t)$ at any t depends only on the derivative at $t = 0$ in the following way:

$$\dot{\xi}_x(t) = T(t)\dot{\xi}_x(0).$$

The following lemma summarizes some (basic) properties of the generator. They will be used throughout the proofs of the upcoming results.

Lemma 9. *Let $(A, D(A))$ be the generator of a C_0-semigroup $(T(t))_{t\geq 0}$. Then,*

(i) *$A : D(A) \to X$ is a linear operator.*

(ii) *If $x \in D(A)$, then $T(t)x \in D(A)$ and, for all $t \geq 0$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}T(t)x = T(t)Ax = AT(t)x.$$

(iii) *For all $t \geq 0$ and $x \in X$,*

$$\int_0^t T(s)x \,\mathrm{d}s \in D(A).$$

(iv) *For all $t \geq 0$,*

$$T(t)x - x = \begin{cases} A\displaystyle\int_0^t T(s)x\,\mathrm{d}s & \textit{if } x \in X, \\ \displaystyle\int_0^t T(s)Ax\,\mathrm{d}s & \textit{if } x \in D(A). \end{cases}$$

Proof. (i) From definition 8, it is clear that A is a linear operator and that $D(A)$ is a linear subspace of X.

(ii) Let $x \in D(A)$. Since $T(t)$ is bounded for all $t \geq 0$,

$$T(t)Ax = T(t)\lim_{h\downarrow 0}\frac{1}{h}(T(h)x - x) = \lim_{h\downarrow 0}\frac{1}{h}\left(T(h)T(t)x - T(t)x\right) = AT(t)x$$

with

$$\frac{\mathrm{d}}{\mathrm{d}t}(T(t)x) \coloneqq T(t)\lim_{h\downarrow 0}\frac{1}{h}(T(h)x - x),$$

and hence

$$T(t)Ax = \frac{\mathrm{d}}{\mathrm{d}t}(T(t)x) = AT(t)x.$$

(iii) For all $t \geq 0$ and $x \in X$,

$$\begin{aligned} \lim_{h\downarrow 0}\frac{1}{h}\left(T(h)\int_0^t T(s)x\,\mathrm{d}s - \int_0^t T(s)x\,\mathrm{d}s\right) &= \lim_{h\downarrow 0}\left(\frac{1}{h}\int_0^t T(s+h)x\,\mathrm{d}s - \frac{1}{h}\int_0^t T(s)x\,\mathrm{d}s\right) \\ &= \lim_{h\downarrow 0}\left(\frac{1}{h}\int_t^{t+h} T(s)x\,\mathrm{d}s - \frac{1}{h}\int_0^h T(s)x\,\mathrm{d}s\right) \\ &= T(t)x - x \end{aligned}$$

(note that the last limit holds from the fundamental theorem of calculus).

(iv) Note that from what we have just seen, for any $x \in X$,

$$Tx - x = \lim_{h\downarrow 0}\frac{1}{h}\left(T(h)\int_0^t T(s)x\,\mathrm{d}s - \int_0^t T(s)x\,\mathrm{d}s\right) = A\int_0^t T(s)x\,\mathrm{d}s.$$

Moreover, if $x \in D(A)$, note that

$$\left\|T(s)\frac{T(h)x - x}{h} - T(s)Ax\right\|_X \leq \|T(s)\|_{\mathcal{B}(X)}\left\|\frac{T(h)x - x}{h} - Ax\right\|_X.$$

Hence, on $s \in [0, t]$, for any $t \geq 0$ we have the following convergence, which is uniform with respect to s:

$$T(s)\frac{T(h)x - x}{h} \xrightarrow[h\downarrow 0]{u} T(s)Ax.$$

Therefore, for any $x \in D(A)$,

$$\lim_{h\downarrow 0} \frac{1}{h}(T(h) - \mathbf{1}_X) \int_0^t T(s)x \, \mathrm{d}s = \int_0^t T(s) \lim_{h\downarrow 0} \left(\frac{1}{h}T(h) - \mathbf{1}_X\right) x \, \mathrm{d}s = \int_0^t T(s)Ax \, \mathrm{d}s.$$

This concludes the proof. ■

The following theorems give us further properties of the generator.

Theorem 10. *The generator A of a C^0-semigroup $(T(t))_{t\geq 0}$ is closed, densely defined and it determines the semigroup uniquely.*

Proof. Let us prove A is closed. Suppose there is a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ such that $x_n \to x \in X$, for $n \to \infty$. Suppose that $Ax_n \to y \in X$, for $n \to \infty$. It suffices, by the characterisation of closed operators, to show that $x \in D(A)$ and $Ax = y$. Since $x_n \in D(A)$, for $t > 0$, one has (see lemma 9)

$$T(t)x_n - x_n = \int_0^t T(s)Ax_n \, \mathrm{d}s. \tag{7}$$

We check now that $\int_0^t T(s)Ax_n \, \mathrm{d}s$ converges to $\int_0^t T(s)y \, \mathrm{d}s$ as $n \to \infty$. By strong continuity, the map $s \in [0, t] \mapsto T(s)y$ is integrable over $[0, t]$. By the triangular inequality of integrals,

$$\begin{aligned}\left\|\int_0^t T(s)Ax_n \, \mathrm{d}s - \int_0^t T(s)y \, \mathrm{d}s\right\|_X &= \left\|\int_0^t T(s)(Ax_n - y) \, \mathrm{d}s\right\|_X \leq \int_0^t \|T(s)(Ax_n - y)\|_X \, \mathrm{d}s \\ &\leq \int_0^t \|T(s)\|_{\mathcal{B}(X)} \|Ax_n - y\|_X \mathrm{d}s \leq \left(\int_0^t M\mathrm{e}^{\omega s} \, \mathrm{d}s\right) \|Ax_n - y\|_X.\end{aligned}$$

Here we have used lemmas 9 and 2. The sequence of inequalities follows from the triangular inequality for integrals, the boundedness of $T(s)$ (by strong continuity) and the exponential growth bound. Since Ax_n converges to y, we deduce that $\lim_{n\to\infty} \int_0^t T(s)Ax_n \, \mathrm{d}s = \int_0^t T(s)y \, \mathrm{d}s$. But the strong continuity of the semigroup yields that $\lim_{n\to\infty} T(t)x_n - x_n = T(t)x - x$. By uniqueness of limits, we conclude from equation (7) that, for all $t \geq 0$,

$$T(t)x - x = \int_0^t T(s)y \, \mathrm{d}s.$$

When t is taken to be positive, we have

$$\frac{1}{t}(T(t)x - x) = \frac{1}{t}\int_0^t T(s)y \, \mathrm{d}s.$$

When t approaches zero we are simply taking the derivative of $T(t)x$ at $t = 0$. That limit exists by the fundamental theorem of vector calculus (the integrand $T(s)y$ is continuous). This implies that

$$Ax = \frac{\mathrm{d}}{\mathrm{d}t}(T(t)x)\,|_{t=0} = T(0)y = y.$$

This means that $Ax = y$ is well-defined, thus $x \in D(A)$. Hence A is closed.

To see that A is densely defined, let us consider $x \in X$. By lemma 9,

$$\int_0^t T(s)x \, \mathrm{d}s \in D(A)$$

for all $t > 0$. Moreover, since T is strongly continuous, taking the limit $t \to 0$ is possible again by the fundamental theorem of vector calculus. This yields

$$\lim_{t\to 0} \frac{1}{t} \int_0^t T(s)x \, \mathrm{d}s = T(0)x = x.$$

Thus, $D(A)$ is a dense subspace of X.

Finally, to prove that A determines the semigroup uniquely, we suppose there is another strongly continuous semigroup $(S(t))_{t\geq 0}$ such that its generator is $A : D(A) \to X$. Let $x \in D(A)$ and $t \geq 0$ be fixed. To see that they are the same semigroup, we define the auxiliary function

$$s \in [0, t] \mapsto \psi_{t,x}(s) = T(t-s)S(s)x.$$

We differentiate this at $s \in [0, t]$. Consider the quotient

$$\begin{aligned}\frac{1}{h}(\psi_{t,x}(s+h) - \psi_{t,x}(s)) &= \frac{1}{h}(T(t-s-h)S(s+h)x - T(t-s)S(s)x) \\ &= \left[T(t-s-h)\frac{1}{h}(S(s+h)x - S(s)x)\right] + \left[\frac{1}{h}(T(t-s-h) - T(t-s))S(s)x\right].\end{aligned}$$

The second term converges to $-AT(t-s)S(s)x$ as $h \to 0$, since $S(s)x \in D(A)$ by lemma 9. The minus sign comes from the chain rule: $-A$ is the generator of $s \mapsto T(t-s)$.

The first term converges to $T(t-s)AS(s)x$ due to the fact that $||T(t-s-h)||_{\mathcal{B}(X)}$ is exponentially bounded by $M\mathrm{e}^{\omega(t-s-h)}$ and the strong continuity of the semigroups. See lemma A.19 in Engel and Nagel's notes [2], where this version of the product rule for semigroups composition is proved in detail. Therefore,

$$\frac{\mathrm{d}}{\mathrm{d}s}\psi_{t,x}(s) = T(t-s)AS(s)x + -AT(t-s)S(s)x.$$

Since semigroups and generators commute (here $-A$ is the generator of $T(t-s)$), we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}s}\psi_{t,x}(s) = 0$$

for all $s \in [0, t]$. Therefore, $\psi_{t,x}$ is constant:

$$T(t)x = \psi_{t,x}(0) = \psi_{t,x}(t) = S(t)x.$$

Hence, $T(t)$ and $S(t)$ agree on $D(A)$, which is dense on X. Thus, they agree on all X. ■

Now, we are going to see some definitions and properties to prove the Hille-Yosida theorem.

Definition 11. Let $\lambda \in \mathbb{C}$ and let A be a closed linear operator. The **resolvent set** of A is defined by

$$\rho(A) := \{\lambda \in \mathbb{C} : (\lambda \mathbf{1}_X - A) \text{ is bijective}\},$$

and $R(\lambda, A) := (\lambda \mathbf{1}_X - A)^{-1}$ is called the *resolvent map* of A. ◂

Remark 12. Let $(T(t))_{t\geq 0}$ be a semigroup. For $\mu \in \mathbb{C}$ and $\alpha > 0$, we define the rescaled semigroup $(S(t))_{t\geq 0}$ by

$$S(t) = \mathrm{e}^{\mu t} T(\alpha t), \qquad t \geq 0.$$

Note that, if $(A, D(A))$ is the generator of $(T(t))_{t\geq 0}$, then $(\alpha \mathbf{1}_X, D(\mu \mathbf{1}_X + \alpha A))$ is the generator of $(S(t))_{t\geq 0}$ and the resolvent map is $R(\lambda, \mu \mathbf{1}_X + \alpha A) = \frac{1}{\alpha} R(\frac{\lambda}{\alpha} - \frac{\mu}{\alpha}, A)$ for $\lambda \in \rho(\mu \mathbf{1}_X + \alpha A)$ ◂

Theorem 13. *Let $(T(t))_{t\geq 0}$ be a strongly continuous semigroup on the Banach space X, and let $(A, D(A))$ be its generator. If $\lambda \in \mathbb{C}$ is such that*

$$R(\lambda) := \int_0^{+\infty} \mathrm{e}^{-\lambda s} T(s)x \, \mathrm{d}s$$

is well-defined for all $x \in X$, then $\lambda \in \rho(A)$ and $R(\lambda) = R(\lambda, A)$.

Proof. Without loss of generality, we can assume that $\lambda = 0$. Therefore, one needs to show that $0 \in \rho(A)$. In particular, we will show that $R(0) = R(0, A) = (-A)^{-1}$. For all $x \in X$ and $h > 0$,

$$\frac{T(h) - \mathbf{1}_X}{h} R(0)x = \frac{T(h) - \mathbf{1}_X}{h} \int_0^{+\infty} T(s)x \, \mathrm{d}s = \frac{1}{h} \int_0^{+\infty} T(s+h)x \, \mathrm{d}s - \frac{1}{h} \int_0^{+\infty} T(s)x \, \mathrm{d}s$$
$$= \frac{1}{h} \int_h^{+\infty} T(s)x \, \mathrm{d}s - \frac{1}{h} \int_0^{+\infty} T(s)x \, \mathrm{d}s = -\frac{1}{h} \int_0^{h} T(s)x \, \mathrm{d}s.$$

Moreover,

$$\lim_{h \to 0} \left(\frac{1}{h} \int_0^h T(s)x \, \mathrm{d}s \right) = x.$$

Thus, $R(0)x \in D(A)$ and $A\,R(0) = -\mathbf{1}_X$. Furthermore, if $x \in D(A)$, we have that

$$\lim_{t \to +\infty} \int_0^t T(s)x \, \mathrm{d}s = R(0)x$$

and, by lemma 9,

$$\lim_{t \to +\infty} A \int_0^t T(s)x \, \mathrm{d}s = \lim_{t \to +\infty} \int_0^t T(s)Ax \, \mathrm{d}s = R(0)Ax.$$

By theorem 10, we deduce that

$$R(0)Ax = AR(0)x = -x, \qquad \text{for } x \in D(A).$$

This concludes the proof. ∎

Corollary 14. *Let $(T(t))_{t \geq 0}$ be a strongly continuous semigroup such that*

$$\|T(t)\|_{\mathcal{B}(X)} \leq M\mathrm{e}^{\omega t} \qquad \omega \in \mathbb{R},\ M \geq 1.$$

If $\lambda \in \mathbb{C}$ and $\omega < \operatorname{Re}\lambda$, then

$$\|R(\lambda, A)\|_{\mathcal{B}(X)} \leq \frac{M}{\operatorname{Re}\lambda - \omega}.$$

Proof. For $t, t' \geq 0$,

$$\left\| \int_{t'}^t \mathrm{e}^{-\lambda s} T(s) \, \mathrm{d}s \right\|_{\mathcal{B}(X)} \leq M \int_{t'}^t \mathrm{e}^{(\omega - \operatorname{Re}\lambda)s} \, \mathrm{d}s.$$

By the Cauchy criterium, for $\omega < \operatorname{Re}\lambda$,

$$\int_0^\infty \mathrm{e}^{(\omega - \operatorname{Re}\lambda)s} \, \mathrm{d}s$$

exists. Therefore, by theorem 13, $\lambda \in \rho(A)$ and

$$R(\lambda, A) = \int_0^\infty \mathrm{e}^{-\lambda s} T(s) \, \mathrm{d}s.$$

Obviously,

$$\|R(\lambda, A)\|_{\mathcal{B}(X)} \leq M \int_0^\infty e^{(\omega - \operatorname{Re}\lambda)s} \, \mathrm{d}s = \frac{M}{\operatorname{Re}\lambda - \omega}.$$

This concludes the proof. ∎

5. Hille-Yosida generation theorem

So far, we have given necessary properties for an operator to be a generator of a strongly continuous semigroup on X. In particular, for a strongly continuous contraction semigroup $(T(t))_{t\geq 0}$, we know by theorem 10 that its generator $(A, D(A))$ is closed and densely defined. Moreover, because of corollary 14 and definition 3, for every $\lambda \in \mathbb{C}$ such that $\operatorname{Re}\lambda > 0$, $\lambda \in \rho(A)$ and

$$\|R(\lambda, A)\|_{\mathcal{B}(X)} \leq \frac{1}{\operatorname{Re}\lambda}.$$

Now we are going to show that these conditions are sufficient for contraction semigroup. First we recall a result that will be useful in the sequel.

Lemma 15 ([2, ch. I, §1, proposition 1.3]). *Let $(T(t))_{t\geq 0}$ be a semigroup. If there exist a dense subset $D \subset X$, $\delta > 0$ and $M \geq 1$ such that*

(i) $\|T(t)\|_{\mathcal{B}(X)} \leq M$ *for all* $t \in [0, \delta]$ *and*

(ii) $\lim_{t\downarrow 0} T(t)x = x$ *for all* $x \in D$,

then $(T(t))_{t\geq 0}$ is a strongly continuous semigroup.

Theorem 16 (Hille-Yosida, 1948). *Let $(A, D(A))$ be a linear operator on a Banach space X. The following statements are equivalent:*

(i) $(A, D(A))$ *generates a strongly continuous contraction semigroup.*

(ii) $(A, D(A))$ *is closed, densely defined, and for all* $\lambda > 0$, $\lambda \in \rho(A)$ *and* $\|\lambda R(\lambda, A)\|_{\mathcal{B}(X)} \leq 1$.

(iii) $(A, D(A))$ *is closed, densely defined, and for all* $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > 0$, $\lambda \in \rho(A)$ *and* $\|R(\lambda, A)\|_{\mathcal{B}(X)} \leq \frac{1}{\operatorname{Re}\lambda}$.

Proof. Note that (i) yields (iii) by an application of corollary 14. Moreover, (ii) is a straightforward conclusion of (iii). Thus, it remains to prove that (ii) implies (i).

To that purpose, we define the *Yosida approximants*

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

Note that, for each $n \in \mathbb{N}$,

$$\|A_n\|_{\mathcal{B}(X)} \leq n\|nR(n, A)\|_{\mathcal{B}(X)} + n \leq 2n.$$

Moreover, since

$$(n - A)(m - A) = (m - A)(n - A),$$

one has

$$R(n, A)R(m, A) = R(m, A)R(n, A) \qquad \text{and} \qquad [A_n, A_m] = 0.$$

These properties imply that the semigroups $(T_n(t))_{t\geq 0}$ given by $T_n(t) := \mathrm{e}^{tA_n}$, $t \geq 0$, are uniformly continuous, and mutually commute.

Because of the fact that $A_n x = nAR(n, A)x = n^2 R(n, A)x - nI$ converges to Ax for $x \in D(A)$ (see Engel and Nagel's notes [2, ch. II, §3, proposition 3.4]), we can anticipate the following properties:

(a) $T(t)x := \lim_{n\to\infty} T_n(t)x$ exists for each $x \in X$.

(b) $(T(t))_{t\geq 0}$ is a C_0-contraction semigroup on X.

(c) This semigroup has generator $(A, D(A))$.

In order to prove (a), we observe that $(T_n(t))_{t\geq 0}$ is a contraction semigroup for each $n \in \mathbb{N}$:

$$\|T_n(t)\|_{\mathcal{B}(X)} \leq \mathrm{e}^{-nt}\mathrm{e}^{\|n^2 R(n,A))\|_{\mathcal{B}(X)} t} \leq \mathrm{e}^{-nt}\mathrm{e}^{nt} = 1 \tag{8}$$

for $t \geq 0$, by assumption (ii).

Now, by using the mutual commutativity of the semigroups $(T_n(t))_{t\geq 0}$ for all $n \in \mathbb{N}$ and the vector-valued version of the fundamental theorem of calculus, for $x \in D(A)$, $t \geq 0$, $m, n \in \mathbb{N}$,

$$T_n(t)x - T_m(t)x = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}(T_m(t-s)T_n(s)x)\,\mathrm{d}s = \int_0^t T_m(t-s)T_n(s)(A_n x - A_m x)\,\mathrm{d}s.$$

By using the triangle inequality and (8), we obtain that

$$\|T_n(t)x - T_m(t)x\|_{\mathcal{B}(X)} \leq t\|A_n x - A_m x\|_{\mathcal{B}(X)}. \tag{9}$$

For $x \in D(A)$, since $(A_n x)_{n\in\mathbb{N}}$ is a Cauchy sequence, $(T_n(t)x)_{n\in\mathbb{N}}$ is a Cauchy sequence, i. e., $(T_n(t)x)_{n\in\mathbb{N}}$ converges to some $T(t)x$. Now let $x \in X$. Since $D(A)$ is dense in X, one has

$$\forall \varepsilon > 0, \exists y \in D(A) : \|x - y\|_X < \varepsilon.$$

Therefore,

$$\|T_n(t)x - T_m(t)x\|_{\mathcal{B}(X)} \leq \|T_n(t)(x-y)\|_{\mathcal{B}(X)} + \|T_n(t)y - T_m(t)y\|_{\mathcal{B}(X)} + \|T_m(t)(y-x)\|_{\mathcal{B}(X)}.$$

Observe that the right side of the inequality is arbitrarily small as n, m go to ∞ because $(T_n(t))_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence, $(T_n(t)(x))_{n\in N}$ is a Cauchy sequence, for all $t \geq 0$ and $x \in X$. Therefore, it converges to some $T(t)x$, for all $x \in X$.

In (b), one needs to prove that the family of operators defined above is a strongly continuous contraction semigroup. First, observe that

$$x = \lim_{n\to\infty} T_n(0)x = T(0)x.$$

Hence,

$$T(0) = \mathbf{1}_X.$$

Moreover, for $t, s \geq 0$,

$$T(t+s)x = \lim_{n\to\infty} T_n(t+s)x = \lim_{n\to\infty} T_n(t)T_n(s)x. \tag{10}$$

Furthermore, for $t, s \geq 0$,

$$T_n(t)T_n(s)x = T_n(t)T(s)x + T_n(t)(T_n(s) - T(s))x.$$

Finally, observe that

$$\lim_{n\to\infty} \|T_n(t)(T_n(s) - T(s))x\|_X = 0$$

and

$$\lim_{n\to\infty} T_n(t)T(s)x = T(t)T(s)x.$$

By (10) we thus deduce the semigroup property.

To prove that the family $(T(t))_{t\geq 0}$ is strongly continuous, note that, by (9), for all $x \in D(A)$, $T(s)x$ is actually the uniform limit of $T_n(s)x$ on the interval $[0, t]$. The maps $s \in [0, t] \mapsto T_n(s)x$ are continuous. Hence, the uniform limit $s \in [0, t] \mapsto T(s)x$ is also continuous. From (8), $\|T(t)\| \leq 1$ for all $t \geq 0$. Thus, by using lemma 15 with $D = D(A)$ we conclude the family is strongly continuous.

To finish the proof, it remains to show that the generator of $(T(t))_{t\geq 0}$, namely $(B, D(B))$, is $(A, D(A))$. Fix any $x \in D(A)$. The orbit map

$$\xi_x : t \in [0, t_0] \mapsto \xi_x(t) = T(t)x$$

is the uniform limit of

$$\xi_x^n : t \in [0, t_0] \mapsto \xi_x^n(x) = T_n(t)x.$$

Also, their derivatives

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi_x^n : t \in [0, t_0] \mapsto T_n(t)A_n x$$

converge uniformly to

$$\eta_x : t \mapsto T(t)Ax.$$

Indeed, for $t \in [0, t_0]$

$$\|T_n(t)A_n x - T(t)Ax\|_{\mathcal{B}(X)} \leq \|T_n(t)(A_n x - Ax)\|_{\mathcal{B}(X)} + \|(T_n(t) - T(t))Ax\|_{\mathcal{B}(X)}$$

and the right hand side vanishes as n goes to ∞ uniformly with respect to $t \in [0, t_0]$. Since

$$\xi_x^n(t) = x + \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\xi_x^n(s)\,\mathrm{d}s = x + \int_0^t T_n(s)A_n x\,\mathrm{d}s,$$

by taking $n \to \infty$, we have

$$\xi_x(t) = \lim_{n\to\infty} \xi_x^n(t) = x + \int_0^t T(s)Ax\,\mathrm{d}s = x + \int_0^t \eta_x(s)\,\mathrm{d}s.$$

Thus, ξ_x is differentiable with $\frac{\mathrm{d}}{\mathrm{d}t}\xi_x(t)|_{t=0} = \eta(0) = Ax$, i. e., $D(A) \subseteq D(B)$ and $Ax = Bx$, for $x \in D(A)$.

Now let $\lambda > 0$. By hypothesis, $\lambda \in \rho(A)$. Since $(B, D(B))$ is the generator of the contraction semigroup $(T(t))_{t\geq 0}$, $\lambda \in \rho(B)$. Thus, both $(\lambda - A)$ and $(\lambda - B)$, possibly unbounded, admit a bounded inverse operator mapping the whole space onto the domain of the generator. Then, for every $y \in D(B)$, we get that

$$(\lambda - B)y = \mathbf{1}_X(\lambda - B)y = (\lambda - A)\underbrace{R(\lambda, A)(\lambda - B)y}_{\in D(A)}.$$

Moreover, since A and B agree on $D(A)$,

$$(\lambda - B)y = (\lambda - B)R(\lambda, A)(\lambda - B)y.$$

By applying $R(\lambda, B)$ on both sides we get

$$y = R(\lambda, A)(\lambda - B)y \in D(A).$$

This implies that $D(B) \subset D(A)$, thus $(A, D(A)) = (B, D(B))$. This concludes the proof. ∎

A generalization of the Hille-Yosida theorem was set in 1952 by Feller, Miyadera and Phillips. Its proof relies on the generation theorem proved by Hille and Yosida, which can be applied after a rescaling argument and a renormalization of the space.

Theorem 17 (general generation theorem, Feller-Miyadera-Phillips, 1952). *Let $(A, D(A))$ be a linear operator on a Banach space X and let $\omega \in \mathbb{R}$, $M \geq 1$ be constants. Then, the following properties are equivalent.*

(i) *$(A, D(A))$ generates a C_0-semigroup $(T(t))_{t\geq 0}$ satisfying that, for all $t \geq 0$,*

$$\|T(t)\| \leq M\mathrm{e}^{\omega t}.$$

(ii) *$(A, D(A))$ is closed, densely defined, and for all $\lambda > \omega$, $\lambda \in \rho(A)$ and*

$$\forall n \in \mathbb{N} \quad \|\,[(\lambda - \omega)R(\lambda, A)]^n\,\| \leq M.$$

(iii) *$(A, D(A))$ is closed, densely defined, and for all $\lambda \in \mathbb{C}$ such that $\operatorname{Re}\lambda > \omega$, $\lambda \in \rho(A)$ and*

$$\forall n \in \mathbb{N} \quad \|R(\lambda, A)^n\| \leq \frac{M}{(\operatorname{Re}\lambda - \omega)^n}.$$

Proof. The fact that (i) implies (iii) is proved in corollary 1.11 of Engel and Nagel's notes [2]. We shall omit this proof for a matter of space. Then, (iii) immediately implies (ii). Thus, we will detail the fact that (ii) implies (i).

We have already seen that, if A generates $(T(t))_{t\geq 0}$, then $A-\omega$ generates $\left(\mathrm{e}^{-\omega t}T(t)\right)_{t\geq 0}$. Furthermore, the resolvent satisfies

$$R(\lambda, A-\omega) = R(\lambda+\omega, A).$$

Hence, for any $\lambda > 0$, $\lambda \in \rho(A-\omega)$. One can assume without loss of generality that $\omega = 0$. Therefore, by hypothesis

$$\forall n \in \mathbb{N} \quad \|\lambda^n R(\lambda, A)^n\| \leq M. \tag{11}$$

Note that throughout the rest of the proof, as it has already been defined previously, $R(\lambda, A)$ is denoted by $R(\lambda)$.

Now, we define, for any $\mu > 0$, the following norm on X:

$$\|x\|_\mu := \sup_{n\geq 0} \|\mu^n R(\mu)^n x\|_X, \tag{12}$$

which is equivalent to $\|\cdot\|_X$. In fact, the estimate $\|x\|_\mu \leq M\|x\|_X$ follows from equation (11). By taking $n=0$ in (12) we get the equivalence of norms:

$$\forall x \in X, \quad \|x\|_X \leq \|x\|_\mu \leq M\|x\|_X. \tag{13}$$

Moreover,

$$\|\mu R(\mu)x\|_\mu = \sup_{n\geq 1} \|\mu^n R(\mu)^n x\|_X \leq \sup_{n\geq 0} \|\mu^n R(\mu)^n x\|_X = \|x\|_\mu. \tag{14}$$

Let $0 < \lambda \leq \mu$ and fix some $x \in X$. Observe that, for $R(\lambda)x \in D(A)$ and $R(\mu)(\mu - A)$ acting as the identity on $D(A)$,

$$R(\lambda)x = R(\mu)(\mu-A)R(\lambda)x = R(\mu)(\mu-\lambda)R(\lambda)x + R(\mu)(\lambda-A)R(\lambda)x = R(\mu)(x+(\mu-\lambda)R(\lambda)x).$$

By the triangle inequality on μ-norms,

$$\|R(\lambda)x\|_\mu \leq \|R(\mu)x\|_\mu + \|(\mu-\lambda)R(\mu)R(\lambda)x\|_\mu,$$

and, by using equation (14), we obtain that

$$\|\lambda R(\lambda)x\|_\mu \leq \|x\|_\mu.$$

Together with the norm equivalence in (13), this inequality implies

$$\|\lambda^n R(\lambda)^n x\|_X \leq \|\lambda^n R(\lambda)^n x\|_\mu \leq \|x\|_\mu.$$

By considering the supremum over n of the left hand side, we obtain the following property of the μ-norms:

$$\forall x \in X, \quad \|x\|_\lambda \leq \|x\|_\mu \text{ for } 0 < \lambda \leq \mu.$$

Because of equation (13),

$$|||x||| := \sup_{\mu>0} \|x\|_\mu$$

is well-defined and actually defines another norm on X. Because of the equivalence relation of the μ-norms, the norm $|||\cdot|||$ satisfies

$$\forall x \in X, \quad \|x\|_X \leq |||x||| \leq M\|x\|_X. \tag{15}$$

One concludes that $|||\lambda R(\lambda)||| \leq 1$. Thus, $(A, D(A))$ satisfies the hypothesis of theorem 16 and generates a $|||\cdot|||$-contraction semigroup $(T(t))_{t\geq 0}$ in the Banach space $(X, |||\cdot|||)$. It follows from the equivalence of the $|||\cdot|||$-norm and the previous norm established in equation (15) that, for every $t \geq 0$,

$$\|T(t)\|_{\mathcal{B}(X)} \leq M.$$

This concludes the proof. ∎

6. Hilbert space generation theorems

In this section, let $\mathcal{H}$ be a Hilbert space. First of all, given a strongly continuous semigroup $(T(t))_{t\geq 0} \subset \mathcal{B}(\mathcal{H})$, one shall define its *adjoint semigroup* as $(T(t)^*)_{t\geq 0}$. Note that, since $T(t)^*T(s)^* = (T(s)T(t))^* = T(t+s)^*$, for $t, s \geq 0$, the adjoint semigroup is well-defined.

Proposition 18. *Let $(A, D(A))$ be the generator of the C_0-semigroup $(T(t))_{t\geq 0}$ acting on a Hilbert space $\mathcal{H}$. Then, its adjoint semigroup is strongly continuous with generator $(A^*, D(A^*))$.*

Proof. For $(T(t))_{t\geq 0}$ being strongly continuous there exist $M \geq 0$, $\omega \in \mathbb{R}$ such that the growth bound $\|T(t)\|_{\mathcal{B}(\mathcal{H})} \leq M\mathrm{e}^{\omega t}$ holds. Since $\|T(t)^*\|_{\mathcal{B}(\mathcal{H})} = \|T(t)\|_{\mathcal{B}(\mathcal{H})}$, the adjoint semigroup satisfies the same inequality. Let $x \in D(A)$ be a normalised vector and $z \in D(A^*)$. Then, by the properties stated in lemma 9,

$$\langle x, T(t)^*z - z\rangle = \langle T(t)x - x, z\rangle = \int_0^t \langle AT(\tau)x, z\rangle \,\mathrm{d}\tau = \int_0^t \langle x, T(\tau)^*A^*z\rangle \,\mathrm{d}\tau. \tag{16}$$

Thus, by the Cauchy-Schwarz and triangle inequalities[1]

$$|\langle x, T(t)^*z - z\rangle| \leq \int_0^t \|T(\tau)^*\|_{\mathcal{B}(\mathcal{H})}\|A^*z\|_{\mathcal{H}} \,\mathrm{d}\tau \leq Mt\mathrm{e}^{\omega t}\|A^*z\|_{\mathcal{H}}. \tag{17}$$

Since $D(A)$ is dense in $\mathcal{H}$, it follows from above that $\|T(t)^*z - z\|_{\mathcal{H}} \leq Mt\mathrm{e}^{\omega t}\|A^*z\|_{\mathcal{H}}$. Therefore, $\lim_{t\downarrow 0} T(t)^*z = z$ for every $z \in D(A)$. Moreover, for $t_0 > 0$ and $t \in [0, t_0]$, $\|T(t)^*\| \leq M\mathrm{e}^{\omega t_0}$ ($M = 1$, $\omega = 0$ in the contraction case). By lemma 15, the adjoint semigroup is strongly continuous.

Suppose that $(B, D(B))$ is the generator of the adjoint semigroup. Let $x \in D(A)$ and $y \in D(B)$. Observe that

$$\langle Ax, y\rangle = \lim_{t\downarrow 0} \frac{1}{t}\langle T(t)x - x, y\rangle = \lim_{t\downarrow 0} \frac{1}{t}\langle x, T(t)^*y - y\rangle = \langle x, By\rangle.$$

Therefore, $D(B) \subset D(A^*)$, by definition of $D(A^*)$. Moreover, since $D(A)$ is dense, (16) implies that, for $z \in D(A^*)$,

$$T(t)^*z - z = \int_0^t T(\tau)^*A^*z \,\mathrm{d}\tau.$$

Hence, for $z \in D(A^*)$,

$$Bz = \lim_{h\downarrow 0} \frac{1}{h}(T(h)^*z - z) = \lim_{h\downarrow 0} \frac{1}{h}\int_0^t T(\tau)^*A^*z \,\mathrm{d}\tau = A^*z$$

holds and $D(A^*) \subset D(B)$ and $A^* = B$. ∎

A (possibly unbounded) operator A acting on a Hilbert space is said to be *skew-adjoint* whenever $A^* = -A$. The next theorem, due to Stone, deals with generators satisfying this property. The proof we provide relies on the Hille-Yosida contraction generation theorem. As we will see in the next section, the generators of evolution groups in quantum mechanics are skew-adjoint.

Theorem 19 (Stone, 1932). *Let $(A, D(A))$ be an operator acting on a Hilbert space. Then, $(A, D(A))$ generates a unitary C_0-group $(U(t))_{t\in\mathbb{R}}$ if and only if A is skew-adjoint.*

Proof. If $(U(t))_{t\in\mathbb{R}}$ is a unitary C_0-group, then A^* is the generator for $U(t)^* = U(t)^{-1} = U(-t)$, as was shown in the previous theorem. Given any $x \in D(A)$,

$$\lim_{h\downarrow 0} \frac{1}{h}(U(h)^*x - x) = \lim_{h\downarrow 0} \frac{1}{h}(U(-h)x - x) = -Ax,$$

[1] Note that the exponential term in the right-hand-side of equation (17) should be omitted in the contraction case.

so $x \in D(A^*)$. Since the left hand side equals A^*x, $D(A) \subset D(A^*)$ and $-A$ agrees with A^* along its domain. One could repeat the same argument for arbitrary $x \in D(A^*)$, obtaining that $D(A^*) \subset D(A)$. Therefore, $D(A) = D(A^*)$ and A is skew-adjoint.

On the other hand, note that, if $(A, D(A))$ is skew-adjoint, then $(iA, D(A))$ is self-adjoint. Thus, both $(A, D(A))$ and $(A^*, D(A^*)) = (-A, D(A))$ have a purely imaginary spectrum (lying on $\mathrm{i}\mathbb{R}$). It follows that

$$\|\lambda R(\lambda, A)\|_{\mathcal{B}(\mathcal{H})} \leq \sup_{\mu \in \mathrm{i}\mathbb{R}} \frac{\lambda}{|\lambda + \mu|} \leq 1.$$

The same calculation is satisfied by $-A$ trivially. Therefore, because of the Hille-Yosida generation theorem in the contraction case, $(A, D(A))$, respectively $(-A, D(A))$, is the generator of the semigroup $(U(t)^+)_{t\geq 0}$, respectively $(U(t)^-)_{t\geq 0}$.

Now, we proceed to show that $(U(t))_{t\in\mathbb{R}}$ defined by

$$U(t) = \begin{cases} U(t)^+ & t \geq 0, \\ U(-t)^- & t < 0 \end{cases}$$

is a unitary C_0-group. Indeed, the strong continuity follows after its definition. All that is left to prove is that $(U(t))_{t\in\mathbb{R}}$, with composition as a product, is a unitary group.

We proceed to show that $U(t)$, $U(-t)$ are inverse elements with $U(0) = \mathbf{1}_{\mathcal{H}}$ as identity element. To this end, fix any $x \in D(A)$. For $t = 0$, $U(0)^+U(0)^-x = \mathbf{1}_{\mathcal{H}}x = x$. Then, for $t > 0$, because of the derivative properties of C_0-semigroups and the skew-adjointness of $(A, D(A))$,

$$\frac{\mathrm{d}}{\mathrm{d}t} U(t)^+U(t)^-x = [U(t)^+AU(t)^- + U(t)^+A^*U(t)^-]\, x = 0.$$

Thus, for $t > 0$, $U(t)^+U(t)^-x = x$, and $D(A)$ is dense in $\mathcal{H}$, so $U(t)U(-t) = \mathbf{1}_{\mathcal{H}}$.

In order to prove that $(U(t))_{t\in\mathbb{R}}$ is closed under composition, fix any $t, s > 0$. We have $U(t)U(s) = U(t+s)$ and $U(-t)U(-s) = U(-t-s)$, since $(U(t)^+)_{t\geq 0}$ and $(U(t)^-)_{t\geq 0}$ are semigroups. If $t < s$, then $U(t)U(-s) = U(t)U(-t)U(t-s) = U(t-s)$, and the $t > s$ case follows similarly. Since composition is associative, $(U(t))_{t\in\mathbb{R}}$ is a group.

In order for A to be skew-adjoint, $(U(t)^*)_{t\geq 0}$ must be generated by $A^* = -A$, as follows from proposition 18. However, $-A$ generates $(U(-t))_{t\geq 0}$ too, as follows from the construction above. Theorem 10 ensures the uniqueness of the semigroup generated by $-A$ so, for every $t \geq 0$, $U(t)^* = U(-t)$, i. e., the C_0-group $(U(t))_{t\in\mathbb{R}}$ is unitary. ∎

7. Back to quantum mechanics

In the setting of quantum mechanics, as explained in the introduction, the space of all possible states of the system is modelled by a Hilbert space $\mathcal{H}$. The energy of the system, described by the Hamiltonian H (self-adjoint operator), determines the evolution of the system via the Schrödinger equation, previously defined in (1).

As one can see in (2), the solution of the above system has the form $\psi(t,x) = U(t)\psi_0(x)$, where $U(t) \in \mathcal{B}(\mathcal{H})$ for $t \in \mathbb{R}$. Again, by (2), $U(t)$ satisfies

$$\begin{cases} \partial_t U(t) = -\mathrm{i}HU(t), \\ U(0) = \mathbf{1}_{\mathcal{H}}. \end{cases} \tag{18}$$

Since H is a self-adjoint operator, H is densely defined, and so is $-\mathrm{i}H$. Thus, Stone's theorem ensures that $-\mathrm{i}H$ will generate a strongly continuous unitary group $(U(t))_{t\in\mathbb{R}}$ satisfying the functional equation in (18).

In terms of the wave function interpretation, we need that the evolution semigroup preserves the norm of the original state ψ_0. Otherwise, there would be an undesirable loss (or gain) of probability if, for example,

$$\|\psi_t\|_{\mathcal{H}} < \|\psi_0\|_{\mathcal{H}} = 1.$$

Stone's theorem ensures this will not occur. Since the evolution operator is unitary, we are guaranteed that, in $\mathcal{H}$,

$$\|\psi_t\|_{\mathcal{H}} = \|U(t)\psi_0\|_{\mathcal{H}} = \|\psi_0\|_{\mathcal{H}} = 1.$$

In terms of the Heisenberg picture introduced in section 3, the time evolution of an observable B in a system determined by the Hamiltonian H is given by the action of a strongly continuous group $\{\tau_t\}_{t\in\mathbb{R}}$ on $\mathcal{B}(\mathcal{H})$. This time evolution is defined by $\tau_t(B) \coloneqq e^{itH}Be^{-itH}$, for $t \in \mathbb{R}$ and $B \in \mathcal{B}(\mathcal{H})$. These operators satisfy equation (4), where $\delta : D(\delta) \subset \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is a symmetric derivation defined on a dense subset $D(\delta)$ of $\mathcal{B}(\mathcal{H})$. It can be proved (in case of interest, see [1]) that these symmetric derivations satisfy the hypothesis of the Hille-Yosida generation theorem. Therefore, they are the generators of the C_0-group $\{\tau_t\}_{t\in\mathcal{R}}$ and determine the evolution of the physical system uniquely.

In fact, the δ operators described above belong to the class of *dissipative operators*, which are contained in the core of the Lumer-Phillips generation theorem [2, theorem 3.15]. This theorem allows to adapt the Hille-Yosida generation theorem to dissipative operators, in a similar way as the Stone theorem, which adjusts our main theorem to self-adjoint operators.

Bru and de Siqueira Pedra [1] show an example of application of symmetric derivations in quantum mechanics generating a C_0-group. These structures are associated to the behaviour of fermions in lattices.

References

[1] Bru, Jean-Bernard and Siqueira Pedra, Walter de. *Lieb-Robinson Bounds for Multi-Commutators and Applications to Response Theory*. SpringerBriefs in Mathematical Physics. Cham: Springer, 2016. https://doi.org/10.1007/978-3-319-45784-0.

[2] Engel, Klaus-Jochen and Nagel, Rainer. *A Short Course on Operator Semigroups*. Universitext. New York: Springer, 2006. https://doi.org/10.1007/0-387-36619-9.