Semigroups of Contractions on Banach Spaces and Some Applications

Tony Masso-Rivetti

Department of Mathematics

University of Notre Dame

Professor Liviu Nicolaescu
Advisor

[1]

2024

SEMIGROUPS OF CONTRACTIONS ON BANACH SPACES AND SOME APPLICATIONS

Contents

Date: Started August 30, 2023. Completed on . Last revised March 26, 2024.

Notations

$\mathbb{1}$	The identity operator.
$\boldsymbol{B}(X,Y)$	The space of continuous linear operators $X \to Y$, X, Y are normed vector spaces.
$\boldsymbol{B}(X)$	$= \boldsymbol{B}(X,X)$.
$C_0(\mathbb{R})$	The space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{\pm\infty} f(x) = 0$.
$C_c(\mathbb{R})$	The space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ with compact support.
$C_b(\mathbb{R})$	The space of bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$.
$\mathbf{cl}(S)$	The closure of the set S.
$\mathbb{E}[X]$	Expected value of random variable X.
$\Phi_{\#}\mu$	The pushforward of measure μ by measurable map Φ.
$IG(T_t)$	The infinitesimal generator of the C_0-semigroup $(T_t)_{t \geq 0}$.
$\mathrm{Prob}(\mathbb{R})$	The space of Borel probability measures on $\mathbb{R}$.
$\hat{\mu}$	The Fourier transform of the probability measure $\mu \in \mathrm{Prob}(\mathbb{R})$.
$\mu * \nu$	The convolution of the probability measures $\mu, \nu \in \mathrm{Prob}(\mathbb{R})$.
$\mathbb{P}_X$	The distribution of a random variable X.
$\mathbb{R}^*$	$= \mathbb{R} \setminus \{0\}$.
$\mathbb{R}_{\geq 0}$	$= [0, \infty)$.
$D(A)$	The domain of the linear operator A.
$\boldsymbol{R}(A)$	The range of map A.
$\bar{A}$	The closure of the unbounded operator A.
$\rho(A)$	The resolvent set of the linear operator A.
$R(\lambda, A)$	The resolvent of A at $\lambda \in \rho(A)$: $(\lambda - A)^{-1}$, for $\lambda \in \rho(A)$.
X^*	The topological dual of the normed vector space X.

Introduction

One of the important motivations for the study of operator semigroups is the following example from differential equations. For any $a, x_0 \in \mathbb{R}$, consider the one dimensional Cauchy problem

$$\begin{cases} x' = ax \\ x(0) = x_0 \end{cases}.$$

It follows from elementary calculus that the solution to such a problem is $x(t) = x_0 e^{at}$. G. Peano observed that this generalizes as follows. For any fixed $n \in \mathbb{N}$, $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, and $y_0 \in \mathbb{R}^n$, the solution of the linear Cauchy problem

$$\begin{cases} y' = Ay \\ y(0) = y_0 \end{cases}$$

can be described as $x(t) = e^{tA}x_0$, where

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

We note that the family $T_t = e^{tA} : \mathbb{R}^n \to \mathbb{R}^n$ satisfy the following conditions:

$$T_0 = \mathbb{1},\ T_{t+s} = T_t T_s,\ \forall s, t \geq 0$$

and the map $t \mapsto T_t y_0$ is continuous for any $y_0 \in \mathbb{R}^n$, and $d_t T_t|_0 = A$.

Next, consider the heat equation on $\mathbb{R}^n$, $\partial_t u = \Delta u$ with initial condition $u(0, x) = u_0(x) \in C_b(\mathbb{R}^n)$. This is formally similar to the above two Cauchy problem since it takes the form $f' = Lf$, where L is a linear operator. This form motivates one to seek solutions in the form $u(t) = \mathscr{T}_t u_0$, where $\mathscr{T}_t$ satisfies the same conditions as T_t. However, Δ is a linear operator on an infinite dimensional vector spaces and this causes complications. The exponential $e^{t\Delta}$ is not defined for $t < 0$ or for every $u \in C_b(\mathbb{R}^n)$, or even everywhere on $C_c^2(\mathbb{R}^n)$. As a result, defining what $\mathscr{T}_t$ should be, or even if such a family of operators exists satisfying $\mathscr{T}_t u_0$ solves $u' = \Delta u$ is subtle, and requires more technical machinery than solving the first two Cauchy problems. The theory of operator semigroups is what allows for determining when such families of operators exist, and how to describe their behavior.

The first part of the thesis focuses on the theory of operator semigroups on Banach spaces. Section 1.1 covers the foundations of the subject matter, including the relevant definitions. In particular, in infinite dimensions there are two forms of continuity involving the maps $\mathscr{T}_t$. This is reflected in the nature of the generator $A = \frac{d}{dt}\big|_{t=0} \mathscr{T}_t$: its is bounded iff $t \mapsto \mathscr{T}_t$ is continuous in the norm topology,

Section 1.2 covers the Hille-Yosida theorem on generation of semigroups describing necessary and sufficient conditions for an unbounded operator to be the generator of a continuous semigroup.

In section Section 1.3 we prove The Lumer-Philips theorem that expresses the Hille-Yosida conditions in termes of the more convenient concept of dissipative operators. .

Section 1.4 covers the Trotter-Kato approximation theorems. One consequence of these results is the Chernoff product formula which we present in Section 1.5 covers.

The second part of the Thesis focuses on applications of operator semigroups to probability.

Section 2.1 introduces basic properties of probability measures on $\mathbb{R}^n$, and introduces the convolution of probability measures. Convolution produces bounded operators spaces of continuous functions. This is is what allows for the theory of operator semigroups to be applied.

Section 2.2 introduces convolution semigroups of probability measures, which form an important class of operator semigroups. This section also contains many enlightening examples. Sections 2.3 and 2.4 are dedicated to the examples of the translation semigroup and heat semigroup respectively. In particular, we give an operator theoretic proof of the central limit theorem.

The appendix contains a brief overview of Bochner integrals. The theory of semigroups is not restricted to separable Banach Spaces; however, the Bochner integrals which arise in the subject are always of continuous functions defined on an interval, which greatly simplify the difficulties of integration functions valued in a possibly nonseparable Banach space.

1. Strongly continuous semigroups on Banach spaces

1.1. **Basic concepts.** For a Banach space X we denote by $\boldsymbol{B}(X)$ the space of bounded linear operators $X \to X$. A (algebraic) semigroup of bounded operators operators on X is a one-parameter family $(T_t)_{t\geq 0}$ in $\boldsymbol{B}(X)$ such that

$$T_0 = \mathbb{1}, \quad T_{s+t} = T_s T_t, \quad \forall s, t \geq 0.$$

Equivalently, $(T_t)_{t\geq 0}$ is a semigroup of bounded operators if it is a morphism of semigroups $(\mathbb{R}_{\geq 0}, +) \to (\boldsymbol{B}(X), \cdot)$.

Definition 1.1 (Uniformly continuous semigroups)**.** Let X be a Banach space and $(T_t)_{t\geq 0}$ a semigroup of bounded operators on X. T_t is called uniformly continuous, or norm continuous, if

$$\lim_{t\searrow 0} \|T_t - \mathbb{1}\| = 0.$$

□

A direct consequence of the above definition is that $\forall x \in X$,

$$\lim_{t\searrow 0} T_t x = x,$$

uniformly for $x \in B_1(0)$ in the unit ball of X. Indeed, for any $\|x\| \leq 1$,

$$\|T_t x - x\| \leq \|T_t - \mathbb{1}\| \cdot \|x\| \leq \|T_t - \mathbb{1}\|.$$

The above definition is equivalent to the map $t \mapsto T_t$ being continuous with respect to the norm topology on $\boldsymbol{B}(X)$.

Definition 1.2 (C_0-semigroups of operators)**.** A semigroup $(T_t)_{t\geq 0}$ of bounded operators on a Banach space is called *strongly continuous* or *C_0-semigroup* if

$$\forall x \in X, \quad \lim_{t\searrow 0} \|T_t x - x\| = 0.$$

□

The above definition is equivalent to the map $t \mapsto T_t$ being continuous with respect to the strong operator topology. Clearly, any uniformly continuous semigroup is a C_0-semigroup.

Lemma 1.3. *Let $(T_t)_{t\geq 0}$ be a C_0-semigroup of bounded operators on a Banach space X. Then, for any $\tau > 0$*

$$M(\tau) := \sup_{t\in[0,\tau]} \|T_t\| < \infty.$$

Proof. We follow the proof of Proposition 1.3 on page 4 of [4]. By the sub-multiplicity of the operator norm, we have

$$\|T_t\| = \|T_{t/n}^n\| \leq \|T_{t/n}\|^n, \quad \forall t > 0, \quad n \in \mathbb{N}.$$

This proves that

$$M(\tau) \leq M(\tau/n)^n.$$

Thus, if $M(\tau) = \infty$, then $M(\tau/n) = \infty$, $\forall n \in \mathbb{N}$. In particular, if $M(\tau/n) = \infty$ for all $n \in \mathbb{N}$, then there exists a non-negative sequence $t_n \searrow 0$ such that $\|T_{t_n}\| \to \infty$. This contradicts the uniform boundedness principle since

$$\lim_{n\to\infty} T_{t_n}x = x, \quad \forall x \in X, \quad \text{ which implies } \sup_{n\in\mathbb{N}} \|T_{t_n}x\| < \infty, \quad \forall x \in X.$$

□

Proposition 1.4. *Let $(T_t)_{t\geq 0}$ be a C_0-semigroup of bounded operators on a Banach space X. Then there exist $M > 0$ and $\omega \in \mathbb{R}$ such that*

$$\|T_t\| \leq Me^{\omega t}, \quad \forall t \geq 0.$$

Proof. We follow the proof of [1, Thm. 12.8]. Set

$$C := \|T_1\|, \quad K := \sup_{t\in[0,1]} \|T_t\|.$$

For any $t \geq 0$ we have $t = \lfloor t \rfloor + r$, for some $r \in [0, 1)$, and it follows that

$$\|T_t\| \leq \|T_1\|^{\lfloor t\rfloor} \cdot \|T_r\| \leq C^{\lfloor t\rfloor} K.$$

Note that

$$C^{\lfloor t\rfloor} \leq \begin{cases} C^t, & C \geq 1, \\ C^{t-1}, & C < 1. \end{cases}$$

Thus,

$$C^{\lfloor t\rfloor} \leq aC^t, \quad a := \max(1, C^{-1}).$$

By choosing $\omega := \log C$ and $M = aK$, we deduce

$$\|T_t\| \leq aKe^{\omega t} = Me^{\omega t}, \quad \forall t \geq 0.$$

□

Definition 1.5 (Infinitesimal generator of a C_0-semigroup)**.** Let (T_t) be a C_0-semigroup. Define the set

$$D(A) := \left\{ x \in X : \quad \lim_{t\searrow 0} \frac{T_t x - x}{t} \text{ exists} \right\}.$$

Define the infinitesimal generator of T_t to be the (unbounded) linear operator

$$A : D(A) \subset X \to X$$

given by

$$Ax = \lim_{t \searrow 0} \frac{T_t x - x}{t}.$$

□

Lemma 1.6. *Let A be the generator of a strongly continuous semigroup of contractions $(T_t)_{t \geq 0}$ defined on a Banach space X. Then the following hold.*

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

(ii) *For any $x \in X$, $T_t x \in D(A)$, $\forall t \geq 0$, the function*

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

is differentiable and

$$\frac{d}{dt} T_t x = A T_t x = T_t A x, \quad \forall t \geq 0.$$

(iii) *For any $x \in X$*

$$\int_0^t T_s x ds \in D(A) \quad \text{and} \quad T_t x - x = A \int_0^t T_s x ds.$$

(iv) *If $x \in D(A)$ then, $\forall t \geq 0$*

$$T_t x - x = \int_0^t T_s A x ds.$$

Proof. We follow the proof of [9, Thm. 2.4]. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ be the field of scalars for X.
(i) Let $\lambda \in \mathbb{K}$ and $x, y \in D(A)$. For any $t \geq 0$, compute the following:

$$\frac{T_t(\lambda x + y) - ax - y}{t} = \lambda \frac{T_t x - x}{t} + \frac{T_t y - y}{t}.$$

This implies that $\lim_{t \searrow 0} \frac{T_t(\lambda x + y) - \lambda x - y}{t}$ exists and is $\lambda Ax + Ay$. Thus $D(A)$ is a linear subspace, and A is a linear operator.
(ii) For any $x \in D(A)$, $t \geq 0$ with $T_t x \in D(A)$, and $h \in \mathbb{R}^\times$, we note the following equalities:

$$\frac{T_{t+h} x - T_t x}{h} = T_t \left(\frac{T_h x - x}{h} \right) = \frac{T_h (T_t x) - T_t x}{h}. \tag{1.6a}$$

Because T_t is continuous, by the first equality in (1.6a)

$$\lim_{h \searrow 0} \frac{T_{t+h} - T_t x}{h} = \lim_{h \searrow 0} T_t \left(\frac{T_h x - x}{h} \right) = T_t \left(\lim_{h \searrow 0} \frac{T_h x - x}{h} \right) = T_t A x.$$

Because $T_t x \in D(A)$, by the second equality in 1.6a:

$$\lim_{h \searrow 0} \frac{T_{t+h} - T_t x}{h} = \lim_{h \searrow 0} \frac{T_h (T_t x) - T_t x}{h} = A T_t x$$

Thus, it follows that $A T_t x = T_t A x$. It remains to be shown that

$$\lim_{h \nearrow 0} \frac{T_{t+h} x - T_t x}{h}$$

exists. For $s < 0$, set $h := -s$ and we have

$$\frac{T_{t+s}x - T_t x}{s} - T_t Ax = \frac{T_t x - T_{t-h}x}{h} - T_t Ax$$
$$= T_{t-h}\left(\frac{T_h x - x}{h} - Ax\right) + T_{t-h}Ax - T_t Ax.$$

Since T_t is continuous, we have $T_{t-h}Ax - T_t Ax \to 0$ as $h \searrow 0$. Additionally, since $\|T_{t-h}\| \leq 1$ for $h \in [0, t]$ we deduce

$$\lim_{h \searrow 0} \left\| T_{t-h}\left(\frac{T_h x - x}{h} - Ax\right)\right\| \leq \lim_{h \searrow 0} \left\| \frac{T_h x - x}{h} - Ax \right\| = \|Ax - Ax\| = 0$$

Thus the derivative of $t \mapsto T_t x$ exists and additionally, we have:

$$\frac{d}{dx} T_t x = AT_t x = T_t Ax.$$

(iii) Let $x \in X$, $t \geq 0$, and $h \in \mathbb{R}^*$. By the commutativity of the Bochner integral with continuous linear operators:

$$\frac{T_h - \mathbb{1}}{h} \int_0^t T_s x ds = \frac{1}{h} \int_0^t T_{s+h}x - T_s x ds = \frac{1}{h} \int_t^{t+h} T_s x ds - \frac{1}{h} \int_0^h T_s x ds$$

Since $t \mapsto T_t x$ is continuous, it follows from the fundamental theorem of calculus for Bochner integrals that

$$\lim_{h \to 0} \frac{T_h - \mathbb{1}}{h} \int_0^t T_s x ds = \lim_{h \to 0} \frac{1}{h} \int_t^{t+h} T_s x ds - \frac{1}{h} \int_0^h T_s x ds = T_t x - x$$

Since the above limit exists, $\int_0^t T_s x ds \in D(A)$.

(iv) Let $x \in D(A)$ and $t \geq 0$. We note that the proof of part (ii) implies that for any $x \in D(A)$ we have

$$\frac{d}{dt} T_t x = T_t Ax.$$

After integrating the above equality, by the fundamental theorem of calculus, we have

$$T_t x - x = \int_0^t \frac{d}{ds} T_s x ds = \int_0^t T_s Ax ds.$$

□

Remark 1.7. Lemma 1.6(ii) shows that for any $x_0 \in X$, the initial value

$$x'(t) = Ax(t), \quad t > 0, \quad x(0) = x_0, \tag{1.1}$$

has at least one solution $x \in C^0\big([0, \infty), X\big) \cap C^1\big((0, \infty), X\big)$ when A is the generator of a C_0-semigroup. As explained in [9, Sec.4.1] this solution is unique. □

Theorem 1.8. *let $(T_t)_{t \geq 0}$ be a strongly continuous semigroup of contractions on the Banach space X. Then its generator A is closed and densely defined and uniquely determines the semigroup. More precisely, this means that different semigroups have different generators.*

Proof. We follow the proof of Corollary 2.5 on page 6 of [9]. First we will prove the density of $D(A)$. Let $x \in X$. By Lemma 1.6 (ii), $\forall \varepsilon > 0$,

$$x_\varepsilon := \frac{1}{\varepsilon}\int_0^\varepsilon T_s x ds \in D(A)$$

Furthermore, since $t \mapsto T_t x$ is continuous by 1.6 (ii), it follows that $x_\varepsilon \to x$ as $\varepsilon \searrow 0$. Thus $D(A)$ is dense.

Next we will prove that A is a closed operator. Consider any sequence $(x_n)_{n\in\mathbb{N}} \in D(A)$ such that $x_n \to 0$ and $Ax_n \to y$, for some $y \in X$. Firstly, by Lemma 1.6 (iv), for any $h > 0$ and $n \in \mathbb{N}$, we have

$$T_h x_n - x_n = \int_0^h T_s A x_n ds$$

Secondly, using the fact that $\|Ax_n - y\| \to 0$, we note that for any $h > 0$,

$$\int_0^h T_s A x_n ds \to \int_0^h T_s y ds \quad \text{as } n \to \infty \text{ because}$$

$$\left\| \int_0^h T_s A x_n ds - \int_0^h T_s y ds \right\| \leq \int_0^h |T_s(Ax_n - y)\| ds \leq h\|Ax_n - y\|.$$

Thirdly, by the above and Lemma 1.6 (iv)

$$\left(T_h x_n - x_n \right) = \int_0^h T_s A x_n ds.$$

On the other hand,

$$\left\| \left(\int_0^h T_s A x_n ds - \int_0^h T_s y ds \right) \right\| \leq \int_0^h \|Ax_n - y\| ds \leq h\|Ax_n - y\|$$

Hence

$$T_h x - x = \lim_{n\to\infty} (Tx_n - x_n) = \int_0^h T_s y ds$$

This shows that

$$\lim_{h\searrow 0} \frac{1}{h}\left(T_h x - x \right) = \lim_{h\searrow 0} \frac{1}{h}\int_0^h T_s y ds = y.$$

This proves that $x \in D(A)$ and $Ax = y$. This implies that A is a closed operator. □

✍*Notation:* We will use the notation $A = \mathrm{IG}(T_t)$ to indicate that A is the generator of the strongly continuous semigroup $(T_t)_{t\geq 0}$.

Theorem 1.9. *Let $A : D(A) \subset X \to X$ be an unbounded linear operator. Then the following are equivalent.*

(i) *The operator A is bounded.*
(ii) *The operator A is the infinitesimal generator of a uniformly continuous semigroup of linear operators.*

Proof. (i)⇒(ii) Assume that $A \in B(X)$. Define

$$T_t := \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}.$$

Since $B(X)$ is a Banach space it suffices to show that the above series is absolutely convergent. This indeed the case because

$$\sum_{n\geq 0} \left\| \frac{(tA)^n}{n!} \right\| \leq \sum_{n\geq 0} \frac{|t|^n \|A\|^n}{n!} = e^{|t|\cdot\|A\|} < \infty.$$

For any $u \in \mathbb{R}$ and any $N \geq 0$ we set

$$S_N(u) := \sum_{n=0}^{N} \frac{1}{n!}(uA)^n.$$

Then $S_N(u) \to T_u$ in $B(X)$ as $N \to \infty$. We will show that

$$\lim_{N\to\infty} \| S_{2N}(s+t) - S_N(s)S_N(t) \| = 0, \quad \forall s,t \in \mathbb{R}.$$

We deduce from the binomial theorem that

$$S_{2N}(s+t) = \sum_{n=0}^{2N} \sum_{k=0}^{n} \frac{(sA)^k (tA)^{n-k}}{k!(n-k)!} = \sum_{0\leq l+k\leq 2N} \frac{(sA)^l}{l!} \frac{(tA)^k}{k!}.$$

On the other hand,

$$S_N(s)S_N(t) = \left(\sum_{l=0}^{N} \frac{(sA)^l}{l!} \right) \cdot \left(\sum_{k=0}^{N} \frac{(tA)^k}{k!} \right) = \sum_{0\leq l,k\leq n} \frac{(sA)^l}{l!} \frac{(tA)^k}{k!}.$$

From the two statements above, it follows that

$$S_{2N}(s+t) - S_N(s)S_N(t) = \underbrace{\sum_{\substack{0\leq l+k\leq 2N \\ l>N}} \frac{(sA)^l}{l!} \frac{(tA)^k}{k!}}_{P_N(s,t)} + \underbrace{\sum_{\substack{0\leq l+k\leq 2N \\ k>N}} \frac{(sA)^l}{l!} \frac{(tA)^k}{k!}}_{Q_N(s,t)}.$$

Note that $P_N(s,t) = Q_N(t,s)$. It suffices to find an upper bound of $\|P_N(s,t)\|$ that is symmetric in s and t. Set

$$M = M(s,t) := \max(1, \|sA\|, \|tA\|).$$

Note that $M(s,t) = M(t,s)$ and

$$\|P_N(s,t)\| = \left\| \sum_{\substack{0\leq l+k\leq 2N \\ l>N}} \frac{(sA)^l}{l!} \frac{(tA)^k}{k!} \right\| \leq M^{2N} \sum_{\substack{0\leq l+k\leq 2N \\ l>N}} \frac{1}{(l+k)!} \binom{l+k}{l}$$

$$\leq \frac{M^{2N}}{(N+1)!} \sum_{\substack{0\leq l+k\leq 2N \\ l>N}} \binom{l+k}{l} \leq \frac{M^{2N}}{(N+1)!} \underbrace{\sum_{m=N+1}^{2N} 2^m}_{\leq 2^{2N+1}} \leq \frac{2(2M)^{2N}}{(N+1)!}$$

Hence
$$\|S_{2N}(s+t) - S_N(s)S_N(t)\| \le \frac{4(2M)^{2N}}{(N+1)!} \to 0 \quad \text{as } N \to \infty.$$
Let us prove that A generates T_t. Indeed, for any $x \in X$ we have
$$\begin{aligned}\frac{T_t x - x}{t} - Ax &= \frac{1}{t}\left(tAx + x + \sum_{n=2}^{\infty} \frac{(tA)^n}{n!}x - x\right) - Ax \\ &= \sum_{n=2}^{\infty} \frac{t^{n-1}A^n}{n!}x - Ax = A\sum_{n=2}^{\infty} \frac{t^{n-1}A^{n-1}}{n!}x\end{aligned}$$
so, by setting $n = m+1$,
$$\left\|\frac{T_t x - x}{t} - Ax\right\| \le \|A\| \sum_{m=1}^{\infty} \frac{(|t|\|A\|)^m}{(m+1)!}$$
$$\le \|A\| \sum_{m=1}^{\infty} \frac{(|t|\|A\|)^m}{m!} = \|A\|\big(e^{|t|\|A\|} - 1\big) \to 0 \quad \text{as } t \searrow 0.$$
(i) $\Leftarrow$ (ii) We follow the proof of [1, Thm. 1.4.21]. Assume that T_t is a uniformly continuous semigroup on X. We are required to show that the generator of T_t is bounded. By the continuity of $t \mapsto T_s$, guaranteed by uniform continuity, we have
$$\frac{1}{t}\int_0^t T_s ds \to \mathbb{1}.$$
Thus, $\exists \varepsilon > 0$ such that
$$\left\|\mathbb{1} - \frac{1}{\varepsilon}\int_0^{\varepsilon} T_s ds\right\| < 1.$$
This implies that $\varepsilon^{-1}\int_0^{\varepsilon} T_s ds$ is invertible, thus $W := \int_0^{\varepsilon} T_s ds$ is also invertible. We note that for any $t > 0$
$$\begin{aligned}W(T_t - \mathbb{1}) &= \int_0^{\varepsilon} T_s ds(T_t - \mathbb{1}) = \int_t^{t+\varepsilon} T_s ds - \int_0^{\varepsilon} T_s ds \\ &= \int_{\varepsilon}^{t+\varepsilon} T_s ds - \int_0^t T_s ds = (T_\varepsilon - \mathbb{1})\int_0^t T_s ds.\end{aligned}$$
Set $V := W^{-1}(T_\varepsilon - \mathbb{1})$. This allows for the following computation, for any $t > 0$:
$$\int_0^t VT_s ds = V\int_0^t T_s ds = W^{-1}(T_h - \mathbb{1})\int_0^t T_s ds = T_t - \mathbb{1}$$
Thus, for all $t > 0$, we have
$$\frac{T_t - \mathbb{1}}{t} = \frac{1}{t}\int_0^t VT_s ds$$
By the continuity of $t \mapsto VT_s$, it follows that $VT_0 = V$, is the generator of T_t, and V is continuous by construction.

□

1.2. **Hille-Yosida Theorem.**

Definition 1.10. Let T_t be a C_0-semigroup on a Banach space X. The semigroup is called a *contraction semigroup* if $\|T_t\| \leq 1$, $\forall t \geq 0$. □

Definition 1.11. Let X be a Banach space and $A : D(A) \subset X \to X$ be an unbounded linear operator on X.

(i) The *resolvent set* of A is

$$\rho(A) := \{\lambda \in \mathbb{C} : (\lambda - A) \quad \text{is invertible with bounded inverse}\}.$$

(ii) The family of bounded operators $R(\lambda, A) : X \to X$, $\lambda \in \rho(A)$,

$$R(\lambda, A) = (\lambda - A)^{-1},$$

is called the *resolvent* of A.

(iii) The *Hille-Yosida approximations* of A are the bounded operators.

$$A(\lambda) := AR(\lambda, A) = R(\lambda, A)A = \lambda^2 R(\lambda, A) - \lambda.$$

□

Remark 1.12. Let $A : D(A) \to X$ be a closed linear operator on Banach space X. The closed graph theorem shows that $\lambda \in \rho(A)$ if and only if $(\lambda - A)$ is invertible. □

Lemma 1.13. *Let A be a closed operator and $\rho(A)$ its resolvent set. Then for all $\lambda, \mu \in \rho(A)$, we have the following identities.*

$$AR(\lambda, A) = \lambda R(\lambda, A) - \mathbb{1}, \tag{1.12a}$$

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A). \tag{1.12b}$$

Proof. The first identity is immediate, as for any $\lambda \in \rho(A)$ we have

$$AR(\lambda, A) = (-\lambda + A)(\lambda - A)^{-1} + \lambda R(\lambda, A) = \lambda R(\lambda, A) - \mathbb{1}.$$

We also have the following identity, for all $\lambda \in \rho(A)$,

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

Additionally, the above implies that $\mathbb{1} = (\lambda - A)R(\lambda, A) = R(\lambda, A)(\lambda - A)$, for any $\lambda \in \rho(A)$. It follows that we have the following equality, for any $\lambda, \mu \in \rho(A)$:

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

□

Lemma 1.14. *Let A be an unbounded linear operator on the Banach space X and assume that A satisfies the HY conditions. Then the resolvent of A satisfies the following.*

(i) $\forall x \in X$

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A)x = x, \quad \forall x \in X. \tag{1.2}$$

(ii) $\forall x \in D(A)$,

$$\lim_{\lambda \to \infty} A(\lambda)x = Ax. \tag{1.3}$$

Proof. We follow the proofs of [9, Lemmas 3.2, 3.3].

(i) Suppose first that $x \in D(A)$. Then, as $\lambda \to \infty$,

$$\big\|\lambda R(\lambda, A)x - x\big\| = \big\|\lambda(\lambda - A)^{-1} - \mathbb{1})x\big\| = \big\|\,(\lambda - A)(\lambda - A)^{-1} + A(\lambda - A)^{-1} - \mathbb{1})x\,\big\|$$

$$= \|(\mathbb{1} + AR(\lambda, A) - \mathbb{1})x\| = \|AR(\lambda, A)x\| = \|R(\lambda, A)Ax\| \le \frac{1}{|\lambda|}\|Ax\| \to 0.$$

This proves (1.2) for $x \in D(A)$. Suppose now that $x \in X$. Since $D(A)$ is dense, $\forall \varepsilon > 0$ there exists $x_\varepsilon \in D(A)$ such that

$$\|x - x_\varepsilon\| < \frac{\varepsilon}{3}.$$

The HY conditions imply that $\|\lambda R(\lambda, A)\| \le 1$ so that

$$\|\lambda R(\lambda, A)x - \lambda R(\lambda, A)x_\varepsilon\| < \frac{\varepsilon}{3}.$$

Since $\lambda R(\lambda, A)x_\varepsilon \to x_\varepsilon$ as $\lambda \to \infty$, there exists $N_\varepsilon > 0$ such that

$$\forall \lambda > N_\varepsilon : \quad \|\lambda R(\lambda, A)x_\varepsilon - x_\varepsilon\| < \frac{\varepsilon}{3}.$$

Then, $\forall \lambda > N_\varepsilon$ we have

$$\|\lambda R(\lambda, A)x - x\| \le \|\lambda R(\lambda, A)x - \lambda R(\lambda, A)x_\varepsilon\| + \|\lambda R(\lambda, A)x_\varepsilon - x_\varepsilon\| + \|x_\varepsilon - x\| < \varepsilon.$$

(ii) We deduce from (i) that

$$\lim_{\lambda \to \infty} A(\lambda)x = \lim_{\lambda \to \infty} \lambda R(\lambda, A)Ax = Ax.$$

□

Lemma 1.15. *If A satisfies the HY estimates, and let $A(\lambda)$ be the Yosida approximations, then the following hold.*

(i) *For any $\lambda > 0$ the operator $A(\lambda)$ generates a uniformly continuous semigroups of contractions which is given by $e^{tA(\lambda)}$.*

(ii) *$\forall x \in X$, $\lambda > 0$, $\mu > 0$*

$$\|e^{tA(\lambda)}x - e^{tA(\mu)}x\| \le t\|A(\lambda)x - A(\mu)x\|.$$

Proof. We follow the proof of Lemma 3.4 on page 10 of [9].

(i) The Yosida approximation, $A(\lambda) = \lambda^2 R(\lambda, A) - \lambda$, is a bounded linear operator. Theorem 1.9 implies that $e^{tA(\lambda)}$ is a uniformly continuous semigroup of operators. Additionally:

$$\|e^{tA(\lambda)}\| = e^{-t\lambda}\|e^{t\lambda^2 R(\lambda,A)}\| \le e^{-t\lambda}e^{t\lambda^2\|R(\lambda,A)\|} \le 1.$$

Thus, $e^{tA(\lambda)}$ is a contraction semigroup $\forall \lambda > 0$.

(ii) For any $\lambda, \mu > 0$, consider the operators $A(\lambda)$ and $A(\mu)$, and we deduce from the product rule that and any $x \in X$:

$$\|e^{tA(\lambda)}x - e^{tA(\mu)}x\| = \left\|\int_0^1 \frac{d}{ds}e^{tsA(\lambda)}e^{t(1-s)A(\mu)}x)ds\right\|$$

$$\leq \int_0^1 t\left\|e^{tsA(\lambda)}e^{t(1-s)A\mu}(A(\lambda)x - A(\mu)x)\right\|ds \leq \|A(\lambda)x - A(\mu)\|.$$

□

Theorem 1.16 (Hille-Yosida)**.** *Let X be a Banach Space. A linear operator A on X is the generator of a contraction semigroup if and only if it satisfies the* Hille-Yosida conditions*.*

(i) *A is closed and densely defined.*
(ii) *The resolvent set $\rho(A)$ contains $(0, \infty)$ and*

$$\forall \lambda > 0 : \quad \|R(\lambda, A)\| < \frac{1}{\lambda}$$

.

For the sake of brevity the Hille-Yosida conditions will be referred to as the HY conditions

Proof. Proof that the HY conditions are necessary to generate a contraction semigroup. Assume that A is the infinitesimal generator of a contraction semigroup $(T_t)_{t\geq 0}$. We are required to show that it satisfies the Hille -Yosida conditions.

Theorem 1.8 implies that A is densely defined. It suffices to prove the second of the HY-conditions.

Proposition A.6 implies immediately the following result.

Lemma 1.17. *Let X, Y be Banach spaces, I an interval of the real axis and $f : I \to X$ a continuous function such that*

$$\int_I \|f(t)\|dt < \infty$$

and thus f is Bochner integrable. Suppose that $A : D(A) \subset X \to Y$ closed linear operator, and assume that $f(t) \in D(A)$, $\forall t \in I$. Then the function Af is also Bochner integrable. Then

$$\int_I Af(t)dt = A\int_I f dt.$$

□

The following lemma will be needed.

Lemma 1.18. *For any $\lambda > 0$ and $x \in X$ define*

$$R(\lambda)x := \int_0^\infty e^{-\lambda t}T_t x dt. \tag{1.4}$$

Then the map $x \to R(\lambda)x$ defines a bounded linear operator which satisfies,

$$\|R(\lambda)\| \leq \frac{1}{\lambda}$$

and additionally, $R(\lambda) = R(\lambda, A)$.

Proof of Lemma 1.18. We follow the proof of [1, Thm. 1.5.25]. The integral in the right-had-side of (1.4) is well defined since $e^{-\lambda t}T_t x$ is continuous and Bochner integrable since:

$$\int_0^\infty \|e^{-\lambda t}T_t x\|dt \leq \left(\int_0^\infty e^{-\lambda t}dt\right)\|x\| = \frac{1}{\lambda}\|x\| < \infty.$$

This implies that $\|R(\lambda)\| \leq \frac{1}{\lambda}$. Fior any $h > 0$ we have

$$\frac{T_h - I}{h} R(\lambda)x = \frac{1}{h}\int_0^\infty e^{-\lambda t}(T_{t+h}x - T_t x)dt = \frac{1}{h}\int_h^\infty e^{-\lambda(t-h)}T_t x dt - \frac{1}{h}\int_0^\infty e^{-\lambda t}T_t x dt$$
$$= \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}T_t x t - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t}T_t x dt.$$

Thus, by continuity:

$$\lim_{h\searrow 0} \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}T_t x t - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t}T_t x dt = (\lambda R(\lambda) - \mathbb{1})x.$$

This implies that $R(\lambda)x \in D(A)$, $\forall x \in X$ and $AR(\lambda)x = \lambda R_y - \mathbb{1}$, so $AR(\lambda) = \lambda R(\lambda) - \mathbb{1}$ and $\mathbb{1} = (\lambda - A)R(\lambda)$. Additionally, $\forall x \in D(A)$, since A is closed, we deduce from Lemma 1.17 that

$$R(\lambda)Ax = \int_0^\infty e^{-\lambda}T_t A x dt = \int_0^\infty e^{-\lambda} A T_t x dt = A\int_0^\infty e^{-\lambda}T_t x dt = AR(\lambda)x.$$

This implies that $R(\lambda)(\lambda - A)x = x$, $\forall x \in D(A), \lambda > 0$ and furthermore $R(\lambda) = (\lambda - A)^{-1}$, $\forall \lambda > 0$. This implies that A satisfies the second HY condition, as desired. ☐

For the proof that the HY conditions are sufficient for the generation of a contraction semigroup, we follow the proofs of [4, Thm. 3.5] and Theorem 2.6 in [7, Thm. 2.6].

Let us assume that A satisfies the Hille-Yosida conditions. We want to show that A is the generator of a strongly continuous semigroup of contractions. Recall for any $\lambda > 0$, the Yosida approximations,

$$A(\lambda) := \lambda A R(\lambda, A) = \lambda^2 R(\lambda, A) - \lambda,$$

are bounded operators. Theorem 1.9 implies that for any $\lambda > 0$, $T_\lambda(t) := \exp(tA(\lambda))$ are uniformly continuous semigroups. It suffices to show the following three statements are true:

(a) For all $x \in X$, $T_t x := \lim_{n\to\infty} T_n(t)x$ exists.
(b) T_t is a contraction C_0-semigroup.
(c) The generator of T_t is A.

(a) First, we note that $\exp(tA(\lambda)$ is a contraction semigroup, $\forall \lambda > 0$, because

$$\|\exp(tA(\lambda)\| = \|\exp(t(\lambda^2 R(\lambda, A) - \lambda))\| \leq e^{-\lambda t}\exp(\|\lambda^2 R(\lambda, A)\|t) \leq e^{-\lambda t}e^{\lambda t} = 1.$$

Next, using the result from functional analysis that on bounded subsets of $B(X)$, convergence with respect to the strong operator topology is equivalent to pointwise convergence on a dense subset, it suffices to show that $e^{tA(n)}x$ converges $\forall x \in D(A)$. We will show that it is Cauchy.

Indeed, for any $x \in D(A)$, we deduce from Proposition A.6(ii) and the fundamental theorem of calculus that, given any $\lambda, \mu > 0$

$$\exp(tA(\lambda)x - \exp(tA(\mu)x = \int_0^t \frac{d}{ds}\big(\exp((t-s)A(\mu))\exp(sA(\lambda))x\big)ds$$
$$= \int_0^t \exp((t-s)A(\mu)\exp(sA(\lambda))\big(A(\lambda)x - A(\mu)x\big)ds.$$

Since $T(\eta, \cdot)$ is a contraction semigroup for any $\eta > 0$, we deduce from the above that

$$\|\exp(tA(\lambda)x - \exp(sA(\lambda)x\| \leq t\|A(\lambda)x - A(\mu)x\|$$

Thus $(\exp(tA(n)x)_{n\in\mathbb{N}}$ is a Cauchy sequence for any $t \geq 0$, because $(tA(n)x)_{n\in\mathbb{N}}$ is a Cauchy sequence.

(b) It is clear that $\|T_t\| \leq 1$, $\forall t \geq 0$, because $T_n(t)$ are contraction semigroups $\forall n \in \mathbb{N}$. Additionally for any $x \in X$,

$$T_{s+t}x := \lim_{n\to\infty} T_n(s+t)x = \lim_{n\to\infty} T_n(s)T_n(t)x$$

The above limit is equal to T_sT_tx. Indeed, given any $\varepsilon > 0$, $\exists N_1, N_2 \in \mathbb{N}$ such that

$$\forall n \geq N_1, \quad \|T_n(t)x - T_tx\| < \frac{\varepsilon}{2M}, \quad \text{where } M := \sup_{n\in\mathbb{N}} \|T(n,s)\|,$$

and

$$\forall n \geq N_2, \quad \|T_ns)T_tx - T_sT_sx\| < \frac{\varepsilon}{2}.$$

Thus, given any $n \geq \max(N_1, N_2)$, we have:

$$\|T_n(s)T_nt)x - T_sT_tx\| \leq \|T_n(s)\| \cdot \|T_n(t)x - T_tx\| + \|T_n(s)T_tx - T_sT_sx\| < \varepsilon.$$

Additionally, for any $x \in X$, the map $t \mapsto T_tx$ is the uniform limit of continuous maps and is therefore continuous. Thus, T_t is a C_0-semigroup of contractions.

(c) Let $B : D(B) \to X$ denote the generator of T_t. Fix an arbitrary $\tau > 0$ and $x \in D(A)$ and consider the maps $\xi_n, \xi : [0,\tau] \to X$ given by $\xi_n(t) = T_n(t)x$, for any $n \in \mathbb{N}$ and $\xi(t) = T_tx$. By part (a), ξ_n converge uniformly to ξ. Additionally, $\forall n \in \mathbb{N}$, ξ_n is differentiable, with $\xi_n'(t) = T_n(t)A(n)x$. We note that the ξ_n' converge uniformly to $\eta : [0,\tau] \to X$, given by $\eta(t) = T_tAx$. This implies that ξ is differentiable and that $\xi = \eta$. Thus B is an extension of A.

By assumption, $\lambda \in \rho(A)$, so $\lambda - A$ is a bijection from $D(A)$ to X. Additionally, by the forwards implication $\lambda \in \rho(B)$, so $\lambda - B$ is a bijection from $D(B)$ to X. However, since $A \subseteq B$, it follows that $A = B$.

□

1.3. **Dissipative operators and cores.** We want to describe an equivalent formulation of Hille-Yosida's theorem more convenient to use in applications.

Suppose that $A : D(A) \subset X \to X$ is an unbounded operator. We say that A is *closable* if the closure in $X \times X$ of its graph $G(A)$ is the graph of another unbounded operator $\bar{A}$, i.e.,

$$\mathbf{cl}\big(G(A)\big) = G(\bar{A}).$$

It is easily seen that A is closable if and only if

$$\big(\{0\} \times X\big) \cap \mathbf{cl}\big(G(A)\big) = \{(0,0)\}.$$

In other words, if $(x_n)_{n\geq 1}$ is a sequence in $D(A)$ such that $x_n \to 0$ and $Ax_n \to y$, then $y = 0$.

Definition 1.19. Let $A : D(A) \subset X \to X$ be an unbounded operator on the Banach space X. We say that A is *dissipative* if

$$\|\lambda x - Ax\| \geq \lambda\|x\|, \quad \forall x \in D(A), \quad \lambda > 0.$$

□

Proposition 1.20. *Let $A : D(A) \subset X \to X$ be a dissipative operator on a Banach space X. Then the following are equivalent.*

(i) *The operator A is closed.*
(ii) *There exists $\lambda > 0$ such that the image of $\lambda - A$ is closed.*
(iii) *For any $\lambda > 0$ the image of $\lambda - A$ is closed.*

Proof. We follow the proof of Lemma 2.2. on page 12 of [5]. Obviously (iii) $\Rightarrow$ (ii).

(i) $\Rightarrow$ (iii) Assume that A is closed. Fix $\lambda > 0$. Let $y \in \mathbf{cl}(\boldsymbol{R}(\lambda - A))$. There exist $x_n \in D(A)$ such that $y_n := (\lambda - A)x_n \to y$. The dissipativity of A implies that

$$\|y_n - y_m\| \geq \frac{1}{\lambda}\|x_n - x_m\|,$$

so x_n is Cauchy and thus converges to some some x. On the other hand $Ax_n = \lambda x_n - y_n$ is also convergent.

Since A is closed we deduce $x \in D(A)$ and

$$Ax = \lim Ax_n = \lambda x - y \Rightarrow y = (\lambda - A)x \in \boldsymbol{R}(\lambda - A).$$

(ii) $\Rightarrow$ (i) Assume that $\boldsymbol{R}(\lambda - A)$ is closed for some $\lambda > 0$. Consider any $(x_n)_{n\in\mathbb{N}} \in D(A)$ such that $x_n \to 0$ and $Ax_n \to y$. Consider the sequence $(\lambda - A)x_n$. Since $\boldsymbol{R}(\lambda - A)$ is closed, it follows that $(\lambda - A)x_n \to 0 - y \in \boldsymbol{R}(\lambda - A)$. Thus $\exists x \in D(A)$ such that $(\lambda - A)x_n \to (\lambda - A)x$. Since A is dissipative, $\forall n \in \mathbb{N}$:

$$\|(\lambda - A)(x_n - x)\| \geq \lambda\|x_n - x\|.$$

Thus $x_n \to x$, so $x = 0$, implying $0 - y = (\lambda - A)x = 0$, so A is closed.

□

Proposition 1.21. *Let $A : D(A) \subset X \to X$ be a densely defined dissipative operator on a Banach space X. Then A is closable, its closure is dissipative and*

$$\boldsymbol{R}(\lambda - \bar{A}) = \mathbf{cl}\left(\ \boldsymbol{R}(\lambda - A)\ \right), \quad \forall\lambda > 0.$$

Proof. We follow the proof of Lemma 2.11 on page 16 of [5]. Let $(y_n)_{n\in\mathbb{N}} \in D(A)$ converge to y. Thus by continuity of the norm, $\forall\lambda > 0$:

$$\|(\lambda - A)y_n - \lambda y\| = \|(\lambda - A)y_n - \lambda \lim_{k\to\infty} Ax_k\| = \lim_{k\to\infty} \|(\lambda - A)y_n - \lambda Ax_k\|.$$

Because $x_k \to 0$ and A is dissipative, we have

$$\lim_{k\to\infty} \|(\lambda - A)y_n - \lambda Ax_k\| = \lim_{k\to\infty} \|(\lambda - A)(y_n - \lambda x_k)\| \geq \lim_{k\to\infty} \lambda\|y_n + \lambda x_k\| = \lambda\|y_n\|.$$

Thus, $\forall\lambda > 0$:

$$\frac{1}{\lambda}\|(\lambda - A)y_n - \lambda y\| = \|(1 - \lambda^{-1}A)y_n - y\| \geq y \geq \|y_n\|.$$

Since $\|y_n - y\| \to 0$, $\|y_n\| = 0$, so $y = 0$. Thus A is closable.

By definition of closure of a subset, $\boldsymbol{R}(\lambda - \bar{A}) \supset \mathbf{cl}(\lambda - A)$, since $(\lambda - \bar{A})$ is closed becuase $\bar{A}$ is closed and dissipative. Additionally, $\boldsymbol{R}(\lambda - \bar{A}) \subset \mathbf{cl}(\lambda - A)$ follows from the definition of closure of an operator.

□

Proposition 1.22. *Let $A : D(A) \subset X \to X$ be a densely defined dissipative operator on a Banach space X. Denote by $\bar{A}$ its closure. Then the following are equivalent.*

(i) $\rho(\bar{A}) \cap (0, \infty) \neq \emptyset$.
(ii) $(0, \infty) \subset \rho(\bar{A})$.
(iii) *There exists $\lambda > 0$ such that the range of $\lambda - A$ is dense in X.*
(iv) *For any $\lambda > 0$ the range of $\lambda - A$ is dense in X.*

Proof. We follow the proof of Lemma 2.3 on page 12 of [5] and Theorem 4.3 on page 14 of [9]. Clearly (ii) $\Rightarrow$ (i) and (iv) $\Rightarrow$ (iii).

To show (iii) $\Rightarrow$ (i), let $\lambda > 0$. In view of Proposition 1.21, the range of $\lambda - A$ is dense in X iff and only if $\boldsymbol{R}(\lambda \mathbb{1} - \bar{A}) = X$. Since $\bar{A}$ is dissipative we deduce that

$$\lambda \mathbb{1} - \bar{A} : D(\bar{A}) \to X$$

is bijective and

$$\|(\lambda - \bar{A})^{-1} x\| \leq \frac{1}{\lambda}\|x\|, \quad \forall x \in X.$$

Hence, the inverse $((\lambda - \bar{A})^{-1}$ is bounded, i.e., $\lambda \in \rho(\bar{A})$. Conversely, if $\lambda \in \rho(\bar{A})$, then $\boldsymbol{R}(\lambda - \bar{A}) = X$. Thus (iii) $\iff$ (i) and (iv) $\iff$ (ii).

It suffices to prove that (iii) $\Rightarrow$ (iv). To show this assume $\exists \lambda_0 > 0$ such that $\boldsymbol{R}(\lambda_0 - A)$ is dense in X, then by Proposition 1.21 $\boldsymbol{R}(\lambda_0 - \bar{A}) = X$. Let

$$\Lambda := \{\lambda > 0 : \lambda \in \rho(\bar{A})\}.$$

The set Λ is nonempty because $\lambda_0 \in \Lambda$. Additionally, using the functional analysis result that the resolvent set of any densely defined closed operator is open in $\mathbb{C}$, we get that Λ is open in $(0, \infty)$. To show Λ is closed in $(0, \infty)$, consider any sequence $(\lambda_n)_{n \in \mathbb{N}} \in \Lambda$ such that $\lambda_n \to \lambda$ for some $\lambda \in (0, \infty)$. Since each $\lambda_n \in \Lambda$, for any $y \in X$, there exists $x_n \in D(\bar{A})$ such that $(\lambda_n - \bar{A})x_n = y$. Thus, by the dissipativity of $\bar{A}$, we have for all $n \in \mathbb{N}$:

$$\|y\| = \|(\lambda_n - A)x_n\| \geq \lambda_n \|x_n\|.$$

In particular, since $(\lambda_n)_{n \in \mathbb{N}}$ is bounded, we have $\|x_n\| \leq C\|y\|$, for some $C > 0$. Additionally, given any $n, m \in \mathbb{N}$, calculate

$$\lambda_m \|x_n - x_m\| \leq \|(\lambda_m - \bar{A})(x_n - x_m) = \|(\lambda_m - \lambda_n + \lambda_n - \bar{A})x_n - y\|$$

$$= |\lambda_m - \lambda_n| \cdot \|x_n\| \leq C|\lambda_m - \lambda_n| \cdot \|y\|$$

Since $(\lambda_n)_{n \in \mathbb{N}}$ does not converge to zero, this implies that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. Let x be the limit of $(x_n)_{n \in \mathbb{N}}$. Since $\bar{A}x_n = \lambda_n x_n - y$, it follows that $\bar{A}x_n \to \lambda x - y$. Thus by closedness we have $(\lambda - \bar{A})x = y$. Since y is arbitrary, it follows that $\boldsymbol{R}(\lambda - \bar{A}) = X$ implying $\lambda \in \Lambda$, so Λ is closed. Thus $\Lambda = (0, \infty)$, as desired.

□

Theorem 1.23. *Let $A : D(A) \subset X \to X$ be a closed and densely defined unbounded operator. Then A the following.*

(i) *The operator A satisfies the Hille-Yosida conditions.*
(ii) *The operator A is dissipative and $\lambda - A$ is surjective for some/all $\lambda > 0$.*

Proof. We follow the proofs of [4, Thm 3.5, Prop. 3.14].

(i) ⇒ (ii) We assume that A satisfies the HY conditions; in particular $(0,\infty) \subset \rho(A)$ and $\lambda\|R(\lambda, A)\| < 1$. We note that $\lambda - A$ is surjective because $(0,\infty) \subset \rho(A)$. Additionally, for any $x \in D(A)$ and $\lambda > 0$, we set $y := \lambda x - Ax$. Thus

$$\|\lambda x - Ax\| = \|y\| > \lambda\|R(\lambda, A)y\| = \lambda\|x\|,$$

so A is dissipative.

(ii) ⇒ (i) We assume that A is dissipative and that $\lambda - A$ is surjective for some/all $\lambda > 0$. We first need to prove that $\lambda - A$ is injective. Consider any $x \in D(A)$ such that $\lambda x - Ax = 0$. By the dissipativity of A, for any $\lambda > 0$ we have

$$0 = \|\lambda x - Ax\| > \lambda\|x\|$$

Thus $x = 0$ and $\lambda - A$ is invertible for any $\lambda > 0$, so $(0,\infty) \subset \rho(A)$. Next we seek to show that A satisfies the HY estimate $\lambda\|R(\lambda, A)\| < 1$ for any $\lambda > 0$. For any $\lambda > 0$ and $y \in D(A)$ with $\|y\|$, the following calculation holds:

$$\lambda\|R(\lambda, A)y\| \leq \|(\lambda - A)R(\lambda, A)y\| = 1$$

Thus, we have $\lambda\|R(\lambda, A)\| \leq 1$, for all $\lambda > 0$, so A satisfies the HY conditions. □

Corollary 1.24 (Lummer-Phillips)**.** *Let $A : D(A) \subset X \to X$ be a densely defined dissipative operator. Then its closure is the generator of a semigroup of contractions if and only if the range of $\lambda - A$ is dense in X for some/all $\lambda > 0$.*

Proof. We follow the proof of [4, Thm. 3.5]. Assume first that $\bar{A}$ is the generator of a contraction C_0-semigroup. By Theorem 1.16 $(0,\infty) \subset \rho(A)$. For any $\lambda > 0$, by Lemma 1.22 $R(\lambda - A)$ is dense in X.

We assume that $\boldsymbol{R}(\lambda - A)$ is dense for some/all $\lambda > 0$. By Lemma 1.22, we have $(0,\infty) \subset \rho(\bar{A})$, so in particular $\lambda - \bar{A}$ is surjective. Since $\bar{A}$ is also dissipative, by Theorem 1.23 the operator $\bar{A}$ satisfies the HY conditions, and is therefore the generator of a C_0-semigroup of contractions.

□

Definition 1.25. Suppose that $A : D(A) \subset X \to X$ is a closed unbounded operator. A *core* of A is a subspace $D \subset D(A)$ such that $G(A) = \mathbf{cl}\big(G(A|_D)\big)$, where $A|_D$ denotes the restriction of A to D. In other words, $\forall x \in D(A)$, there exists a sequence $(x_n) \in D$ such that $x_n \to x$ and $Ax_n \to Ax$ as $n \to \infty$. □

From Proposition 1.22 we deduce the following characterization of cores of generators of semigroups of contraction.

Corollary 1.26. *Suppose that $A : D(A) \subset X \to X$ is the generator of a semigroups of (linear) contractions on the Banach space X. A subspace $D \subset D(A)$ is a core of A if and only is D is dense in X and $\big(\lambda - A\big)(D)$ is dense in X for some/all $\lambda > 0$.* □

Proposition 1.27. *Suppose that $A : D(A) \subset X \to X$ is the generator of a semigroup of (linear) contractions on Banach space X. If a linear subspace $D \subseteq D(A)$ is $\|\cdot\|$-dense in X and furthermore $T_t(D) \subseteq D$, for any $t \geq 0$, then D is a core for A.*

Proof. We follow the proof of [4, Prop. 1.7]. By assumption, for any $x \in D(A)$, we have a sequence $(x_n)_{n\in\mathbb{N}} \in D$ such that $x_n \to x$. By continuity of $s \mapsto T_s x$ with respect to $\|\cdot\|_{G(A)}$, and the continuity of the Bochner integral, we have for any $t > 0$:

$$\left\| \frac{1}{t}\int_0^t T_s x_n ds - \frac{1}{t}\int_0^t T_s x ds \right\|_{G(A)} \to 0 \quad \text{as } n \to \infty.$$

Additionally, by continuity of $s \mapsto T_s x_n$ with respect to $\|\cdot\|_{G(A)}$, we have $\forall n \in \mathbb{N}$,

$$\left\| \frac{1}{t}\int_0^t T_s x_n ds - \frac{1}{t} T_s x ds \right\|_{G(A)} \to 0 \quad \text{as } t \searrow 0.$$

The above two statements imply that give any $\varepsilon > 0$, $\exists N \in \mathbb{N}$ and $\delta > 0$ such that for all $n \geq N$ and $0 < t < \delta$, we have

$$\left\| \frac{1}{t}\int_0^t T_s x_n ds - x \right\| < \varepsilon.$$

Finally, by continuity of $s \mapsto T_s x_n$ with respect to $\|\cdot\|_{G(A)}$, we have that

$$\int_0^t T_s x_n ds \in \mathbf{cl}(G(A|_D)) \quad \forall n \in \mathbb{N}, \forall t > 0$$

This implies that x is in the $\|\cdot\|_{G(A)}$-closure of D. ☐

Proposition 1.28. *Let $(T_t)_{t\geq 1}$ be a semigroup of contractions on the Banach space X with generator A. Define inductively*

$$D(A^n) := \left\{ x \in D(A^{n-1}); \quad Ax \in D(A^{n-1}) \right\}.$$

Set

$$D(A^\infty) = \bigcap_{n\geq 1} D(A^n).$$

Then $D(A^\infty)$ is a core of A.

Proof. We follow the proof of [4, Prop. 1.8]. By definition and Lemma 1.6, $D(A^\infty)$ is a $(T_t)_{t\geq 0}$ invariant subspace. Thus by Proposition 1.27, it suffices to show that $D(A^\infty)$ is dense in X. To show this, we construct a subspace of $D(A^\infty)$ and show that it is dense in X.

Let

$$K := \{\varphi \in C^\infty(\mathbb{R}) : \operatorname{Supp}\varphi \text{ is compact and } \operatorname{Supp}\varphi \subset (0,\infty)\}.$$

For any $x \in X$ and $\varphi \in K$, define

$$x_\varphi := \int_0^\infty \varphi(s) T_s x ds.$$

We now seek to show that $x_\varphi \in D(A^\infty)$, for all $x \in X, \varphi \in K$. More precisely we will show that $x_\varphi \in D(A)$ and $Ax_\varphi = x_{-\varphi}$. This implies inductively that $x_\varphi \in D(A^\infty)$.

For any $h > 0$, $x \in X$ and $\varphi \in K$we have

$$\begin{aligned}\frac{T_h - \mathbb{1}}{h} x_\varphi &= \frac{1}{h}\int_0^\infty \varphi(s)(T_{s+h} - T_s) x ds = \frac{1}{h}\int_h^\infty \varphi(s-h) T_s x ds - \frac{1}{h}\int_0^\infty \varphi(s) T_s x ds \\ &= \frac{1}{h}\int_h^\infty (\varphi(s-h) - \varphi(s)) T_s x ds - \frac{1}{h}\int_0^h \varphi(s) T_s x ds.\end{aligned}$$

Next, using the fact that $\varphi(s) = 0$, for all $s \leq 0$, we conclude

$$\frac{1}{h}\int_h^\infty (\varphi(s-h) - \varphi(s))T_s x ds - \frac{1}{h}\int_0^h \varphi(s)T_s x ds = \int_0^\infty \frac{1}{h}(\varphi(s-h) - \varphi(s))T_s x ds.$$

We note that because $\frac{1}{h}(\varphi(s-h) - \varphi(s))$ converges to $\varphi'(s)$, and because φ has compact support, $\|\frac{1}{h}(\varphi(s-h) - \varphi(s))T_s x\|$ is bounded by $B(s)\|T_s x\|$, for some $B : \mathbb{R} \to [0, \infty)$. We can assume without loss of generality that $\operatorname{Supp} B = \operatorname{Supp} \varphi$. Thus B has compact support, so $\int_0^\infty B(s)\|T_s x\| ds < \infty$. Using the Proposition A.6 (ii) and the dominated convergence theorem we deduce

$$\lim_{h \searrow 0} \int_0^\infty \frac{1}{h}(\varphi(s-h) - \varphi(s))T_s x ds = \int_0^\infty -\varphi'(s)T_s x ds = x_{-\varphi'}.$$

This proves that $x_\varphi \in D(A)$ and $Ax_\varphi = x_{-\varphi'}$.

Set

$$D := \operatorname{span}\{x_\varphi \in X : x \in X, \varphi \in K\}.$$

We will show that D is dense in X. Assume for the sake of contradiction that D is not dense in X. Hahn-Banach theorem implies that there exists $x^* \in X^*$ such that $D \subset \ker(x^*)$ and $x^* \neq 0$. Usiung Proposition A.6(ii) we deduce that for any $x \in X$ and $\varphi \in K$, we have:

$$0 = x^*\left(\int_0^\infty \varphi(s)T_s x ds\right) = \int_0^\infty \varphi(s)x^*(T_s x) ds.$$

Because the above is true for any $\varphi \in K$, and $s \mapsto x^*(T_s x)$ is continuous, we must have $x^*(T_s x) = 0$ for all $t \geq 0$ and for any $x \in X$. By setting $s = 0$, we have that $x^*(x) = 0$, for all $x \in X$, which is a contradiction. Thus D is dense in X. □

1.4. **Trotter-Kato Approximation Theorems.**

Definition 1.29 (Pseudo-Resolvent)**.** Let $\Lambda \subset \mathbb{C}$ and X be a Banach Space. A family of operators $(\mathscr{J}(\lambda))_{\lambda \in \Lambda} \in B(X)$ is called a *pseudo-resolvent* if $\forall \lambda, \mu \in \Lambda$:

$$\mathscr{J}(\lambda) - \mathscr{J}(\mu) = (\mu - \lambda)\mathscr{J}(\lambda)\mathscr{J}(\mu).$$

□

Proposition 1.30. *Let $(T_n(t))_{n \in \mathbb{N}, t \geq 0}$ be a sequence of contraction semigroups on the Banach space X. Denote by A_n the generator of T_n. Assume that $\exists \lambda_0 > 0$ such that*

$$\lim_{n \to \infty} R(\lambda_0, A_n)x \quad \textit{exists} \quad \forall x \in X.$$

Then $\forall \lambda > 0$, and more generally any $\lambda \in \mathbb{C}$ with $\mathbf{Re}(\lambda) > 0$

$$R(\lambda)x := \lim_{n \to \infty} R(\lambda, A_n)x \;\; \textit{exists} \; \forall x \in X$$

and the family $(R(\lambda))_{\lambda > 0}$ is a pseudo-resolvent.

Proof. We follow the proof of [4, Prop. 1.4]. Define

$$\Omega := \left\{\lambda \in \mathbb{C} : \mathbf{Re}(\lambda) > 0, \text{ and } \lim_{n \to \infty} R(\lambda, A_n)x \text{ exists } \forall x \in X\right\}.$$

We note that $\lambda_0 \in \Omega$, so it is known that Ω is nonempty. We will prove that Ω is both a closed and open subset of $\{\mathbf{Re}\,\lambda > 0\}$.

Fix $\mu \in \Omega$ and $\alpha \in (0,1)$ and define

$$U_\alpha(\mu) := \left\{ \lambda \in \mathbb{C} : \quad \mathbf{Re}\,\lambda > 0,\ \frac{|\mu - \lambda|}{\mathbf{Re}(\mu)} < \alpha \right\}.$$

Note that $U_\alpha(\mu)$ is an open subset of $\mathbb{C}$ containing μ. Moreover

$$|\mu - \lambda| \| R(\mu, A_n)\| < \alpha, \quad \forall n \in \mathbb{N},\ \forall \lambda \in U_\alpha(\mu).$$

Since $\mathbf{Re}(\lambda) > 0$ and A_n generates a semigroup of contractions we deduce from the Hille-Yosida theorem that $\lambda - A_n$ is invertible. Moreover, from the equality

$$\lambda - A_n = \mu - A_n + \lambda - \mu = \big(\, 1 - (\mu - \lambda) R(\mu, A_n) \,\big)\big(\, \mu - A_n \,\big)$$

we deduce that

$$R(\lambda, A_n) = (\lambda - A_n)^{-1} = R(\mu, A_n)(\mathbb{1} - (\mu - \lambda) R(\lambda, A_n))^{-1}$$

$$= \sum_{k=1}^{\infty} (\mu - \lambda)^k \left(R(\mu, A_n)\right)^{k+1}.$$

Observe that for $\lambda \in U_\alpha$ we have

$$|\mu - \lambda|^k \left\| (R(\mu, A_n))^{k+1} \right\| \leq \frac{1}{\mathbf{Re}(\mu)} \alpha^k$$

Hence the above convergence is uniform for $\lambda \in U_\alpha$ and $n \in \mathbb{N}$. This implies that $R(\lambda, A_n)x$ converges for any $x \in X$ as $n \to \infty$, $\forall \lambda \in U_\alpha$. Thus, μ has an open neighborhood in Ω, so Ω is open.

To prove that Ω is also closed consider a cluster point λ of Ω, with $\mathbf{Re}(\lambda) > 0$. We note that $\forall \alpha \in (0,1)$, can find $\mu \in \Omega$ such that $\lambda \in U_\alpha(\mu)$ for some $\alpha \in (0,1)$. The above argument shows that $\lambda \in \Omega$, so Ω is closed in $H := \{ z \in \mathbb{C} : \mathbf{Re}(z) > 0 \}$. Since $\Omega \neq \varnothing$ is both closed and open in H, we have that $\Omega \cap H = H$ by contentedness.

Since $R(\lambda, A_n)$ all satisfy the pseudo-resolvent equation for all n, it follows that the limits also satisfy the pseudo-resolvent equation, so $R(\lambda)$ are a pseudo-resolvent. □

Lemma 1.31. *Let X be a Banach space, $\Lambda \subseteq \mathbb{C}$, and $(\mathscr{J}(\lambda))_{\lambda \in \Lambda}$ a pseudo-resolvent*

(i) *The following hold $\forall \lambda, \mu \in \Lambda$:*
 (a) $\mathscr{J}(\lambda)\mathscr{J}(\mu) = \mathscr{J}(\mu)\mathscr{J}(\lambda)$
 (b) $\ker \mathscr{J}(\lambda) = \ker \mathscr{J}(\mu)$
 (c) $\boldsymbol{R}\,\mathscr{J}(\lambda) = \boldsymbol{R}\,\mathscr{J}(\mu)$

(ii) *The following are equivalent*
 (a) $\exists A : D(A) \to X$ *densely defined and closed operator such that* $\Lambda \subset \rho(A)$ *and* $\mathscr{J}(\lambda) = R(\lambda, A)$, $\forall \lambda \in \Lambda$.
 (b) *For some (or all)* $\lambda \in \Lambda$, $\ker(\mathscr{J}(\lambda)) = 0$ *and and it is dominant, i.e.,*

$$\overline{\boldsymbol{R}(\mathscr{J}(\lambda))} = X.$$

(iii) *Assume* $\exists(\lambda_n)_{n\in\mathbb{N}} \in \Lambda$ *is unbounded sequence. If* $\forall x \in X$

$$\lim_{n\to\infty} \lambda_n \mathscr{J}(\lambda_n)x = x \tag{1.5}$$

Then $\mathscr{J}(\lambda)$ *is the resolvent of some densely defined closed operator.*

Alternatively, if $\overline{\boldsymbol{R}(\mathscr{J}(\lambda))} = X$ *and* $\|\lambda_n \mathscr{J}_{\lambda_n}\| \leq 1$, $\forall n \in \mathbb{N}$, *then (1.5) holds.*

Proof. We follow the proofs of Lemma 1.5 on page 139, Proposition 1.6, and Corollary 1.7 of [4].

(i) By the pseudo-resolvent equation, $\forall \lambda, \mu \in \Lambda$:

$$\mathscr{J}(\lambda) = \mathscr{J}(\mu) + (\mu - \lambda)\mathscr{J}(\lambda)\mathscr{J}(\mu) = \big(\mathbb{1} + (\mu-\lambda)\mathscr{J}(\lambda)\big)\mathscr{J}(\mu).$$

Thus, $\boldsymbol{R}(\mathscr{J}(\lambda)) \subseteq \boldsymbol{R}(\mathscr{J}(\mu))$ and $\ker(\mathscr{J}(\mu)) \subseteq \ker(\mathscr{J}(\lambda))$. Similarly:

$$\mathscr{J}(\mu) = \mathscr{J}(\lambda) + (\lambda - \mu)\mathscr{J}(\mu)\mathscr{J}(\lambda) = \big(\mathbb{1} + (\lambda-\mu)\mathscr{J}(\mu)\big)\mathscr{J}(\lambda).$$

Thus, $\boldsymbol{R}(\mathscr{J}(\lambda)) \supseteq \boldsymbol{R}(\mathscr{J}(\mu))$ and $\ker(\mathscr{J}(\mu)) \supseteq \ker(\mathscr{J}(\lambda))$, implying equality. In particular, this implies that for any pseudo-resolvent, the statements 'for some' and 'for all' are equivalent when dealing with the range and kernel.

For commutativity, by the pseudo-resolvent equation, $\forall \lambda, \mu \in \Lambda$:

$$\mathscr{J}(\lambda)\mathscr{J}(\mu) = \frac{1}{\mu-\lambda}\big(\mathscr{J}(\lambda) - \mathscr{J}(\mu)\big) = \frac{1}{\lambda-\mu}\big(\mathscr{J}(\mu) - \mathscr{J}(\lambda)\big) = \mathscr{J}(\mu)\mathscr{J}(\lambda).$$

(ii) (b) $\Rightarrow$ (a) Assume that for some (or all) $\lambda \in \Lambda$, $\mathscr{J}(\lambda)$ is injective and has dense range. Define the unbounded linear operator $A : D(A) \to X$ by $A := \mu - \mathscr{J}(\mu)^{-1}$, for fixed $\mu \in \Lambda$. By assumption A is densely defined, because $D(A) = \boldsymbol{R}(\mathscr{J}(\mu))$.

We next seek to show that A is closed. Indeed, for any sequence $(x_n)_{n\in\mathbb{N}} \in D(A)$ that satisfies $x_n \to 0$ and $Ax_n \to y$ for some $y \in X$, the sequence must also satisfy the following:

$$\mathscr{J}(\mu)(Ax_n) = \mathscr{J}(\mu)(\mu x_n - \mathscr{J}(\mu)^{-1}x_n) = \mu\mathscr{J}(\mu)x_n - x_n \to 0$$

Since $\mathscr{J}(\mu)$ is injective and continuous, we have $Ax_n \to 0$. Thus A is closed, as desired. Additionally, by the construction of A, we have $\forall \lambda \in \Lambda$

$$\begin{aligned}(\lambda - A)\mathscr{J}(\lambda) &= ((\lambda-\mu) + (\mu - A))\mathscr{J}(\lambda)\\ &= ((\lambda-\mu) + (\mu - A))\big(\mathbb{1} + (\mu-\lambda)\mathscr{J}(\lambda)\big)\mathscr{J}(\mu) =\\ &= \mathbb{1} + (\lambda-\mu)\big(\mathscr{J}(\mu) - \mathscr{J}(\lambda) - (\lambda-\mu)\mathscr{J}(\lambda)\mathscr{J}(\mu)\big) = \mathbb{1}.\end{aligned}$$

The last two equalities follow from the pseudo-resolvent equation. Similarly, $\forall \lambda \in \Lambda$:

$$\begin{aligned}\mathscr{J}(\lambda)(\mu - A) &= \mathscr{J}(\lambda)\big((\lambda-\mu) + (\mu - A)\big)\\ &= \big(\mathbb{1} + (\mu-\lambda)\mathscr{J}(\lambda)\big)\mathscr{J}(\mu)\big((\lambda-\mu) + (\mu-A)\big) =\\ &= \mathbb{1} + (\lambda-\mu)\big(\mathscr{J}(\mu) - \mathscr{J}(\lambda) - (\lambda-\mu)\mathscr{J}(\lambda)\mathscr{J}(\mu)\big) = \mathbb{1}.\end{aligned}$$

Thus, $\mathscr{J}(\lambda) = R(\lambda, A)$, $\forall \lambda \in \Lambda$.

(a) $\Rightarrow$ (b) The proof is by contradiction. Assume that for some (or all) $\lambda \in \Lambda$, $\mathscr{J}(\lambda)$ is not injective nor $\mathscr{J}(\lambda)$ is not dominant.

If $\mathscr{J}(\lambda)$ is not injective $\exists x \in X$, $x \neq 0$ such that $\mathscr{J}(\lambda)x = 0$. Then for any densely defined closed operator $A : D(A) \to X$ such that $(\lambda - A)^{-1} \in B(X)$, we must have$(\lambda - A)0 = 0 \neq x$. This implies that $(\lambda - A)^{-1}x \neq 0 = \mathscr{J}(\lambda)x$. Thus $\mathscr{J}(\lambda) \neq R(\lambda, A)$.

If $\mathscr{J}(\lambda)$ is not dominant, then for any densely defined closed operator $A : D(A) \to X$ such that $(\lambda - A)^{-1} \in B(X)$, $(\lambda - A)^{-1}$ is dominant, so $(\lambda - A)^{-1} \neq \mathscr{J}(\lambda)$.

(iii) Firstly note that if (1.5) holds, then

$$X = \overline{\bigcup_{n\in\mathbb{N}} \boldsymbol{R}(\mathscr{J}(\lambda_n))} = \overline{\boldsymbol{R}(\mathscr{J}(\lambda))}, \qquad \forall \lambda \in \Lambda.$$

Additionally, if (1.5) holds, then $\ker \mathscr{J}(\lambda) = 0$. Indeed, if $x \in \ker \mathscr{J}(\lambda)$, then $x \in \ker \mathscr{J}(\lambda_n)$, $\forall n$ we must have

$$x = \lim_{n\to\infty} \lambda_n \mathscr{J}(\lambda_n) x = 0.$$

Thus, by (ii), $\mathscr{J}(\lambda)$ is the resolvent of some densely defined closed operator.

Assume that $\overline{\boldsymbol{R}(\mathscr{J}(\lambda))} = X$ and $\|\lambda_n \mathscr{J}(\lambda_n)\| \le 1$, $\forall n \in \mathbb{N}$. It follows that

$$\|\mathscr{J}(\lambda_0)\| \le \frac{1}{|\lambda_n|}, \qquad \forall n \in \mathbb{N}.$$

This implies, $\forall \mu \in \Lambda$ we have

$$\lim_{n\to\infty} \left\|(\lambda_n \mathscr{J}(\lambda_n) - \mathbb{1})\mathscr{J}(\mu)\right\| = \lim_{n\to\infty} \left\|\lambda_n \mathscr{J}(\lambda_n)\mathscr{J}_\nu - \mathscr{J}(\mu)\right\|$$
$$= \lim_{n\to\infty} \left\|\mu \mathscr{J}(\lambda_n)\mathscr{J}(\mu) - \mathscr{J}(\lambda_n)\right\| = 0.$$

Thus (1.5) holds.

□

Theorem 1.32 (First Trotter-Kato approximation theorem)**.** *Let X be a Banach space and let $(T(t))_{t\ge0}$ and $(T_n(t))_{t\ge0}$, $\forall n \in \mathbb{N}$, be contraction semigroups on X. Let A and A_n denote the infinitesimal generators of $T(t)$ and $T_n(t)$ respectively. Let $D \subset X$ be a core of A. Then we have the sequence of implications*

$$\text{(i)} \implies \text{(ii)} \iff \text{(iii)} \iff \text{(iv)}$$

involving the statements below.

(i) *$D \subset D(A_n)$, $\forall n \in \mathbb{N}$ and $A_n x \to Ax$, $\forall x \in D$.*

(ii) *$\forall x \in D$, $\exists x_n \in D(A_n)$, $\forall n \in \mathbb{N}$ such that*

$$x_n \to x \ \text{ and } \ A_n x_n \to Ax.$$

(iii) *$R(\lambda, A_n)x \to R(\lambda, A)x$, $\forall x \in X$ for some (all) $\lambda > 0$.*

(iv) *$T_n(t)x \to T(t)x$ $\forall x \in X$, and uniformly for t on any compact interval.*

Proof. We follow the proof of Theorem 1.8 on page 141 of [4].

(i) ⇒ (ii) Assume that $D \subset D(A_n)$, $\forall n \in \mathbb{N}$ and that $A_n x \to Ax$, $\forall x \in D$. Thus (ii) holds with (x_n) the constant sequence $x_n = x \in D(A_n)$, $\forall n \in \mathbb{N}$.

(ii) ⇒ (iii) Assume that for all $x \in D$ and $n \in \mathbb{N}$, $\exists x_n \in D(A_n)$ such that

$$x_n \to x \ \text{ and } \ A_n x_n \to Ax.$$

Let $\lambda > 0$. We note that $\|R(\lambda, A_n)\| \le 1$, for all $n \in \mathbb{N}$, by the Hille-Yosida theorem, since A_n generate contraction semigroups. It suffices to show that $\forall y \in (\lambda - A)(D)$, $R(\lambda, A_n)y \to$

$R(\lambda, A)y$, by the uniform boundedness principle. Take any $x \in D$ and set $y = (\lambda - A)x$. By assumption, $\exists x_n \in D(A_n)$ such that $x_n \to x$ and $A_n x_n \to Ax$. We set

$$y_n := (\lambda - A_n)x_n \to x - Ax = y.$$

This implies:

$$\begin{aligned}\|R(\lambda, A_n)y - R(\lambda, A)y\| &= \|R(\lambda, A_n)y - R(\lambda, A_n)y_n + R(\lambda, A_n)y_n - R(\lambda, A)y\| \\ &\leq \|R(\lambda, A_n)y - R(\lambda, A_n)y_n\| + \|R(\lambda, A_n)y_n - R(\lambda, A)y\|.\end{aligned}$$

By choice of $y = (\lambda - A)x$, it follows that:

$$\begin{aligned}&\|R(\lambda, A_n)y - R(\lambda, A_n)y_n\| + \|R(\lambda, A_n)y_n - R(\lambda, A)y\| \\ &\leq \|R(\lambda, A_n)\| \cdot \|y - y_n\| + \|x_n - x\| \to 0.\end{aligned}$$

Thus

$$\lim_{n\to\infty} \|R(\lambda, A_n)y - R(\lambda, A)y\| = 0.$$

(iii) $\Rightarrow$ (ii) Assume that $R(\lambda, A_n)x \to R(\lambda, A)x$, For all $x \in X$ and some $\lambda > 0$. For any $x \in D$, pick y such that $x = R(\lambda, A)y$ and define the sequence $x_n := R(\lambda, A_n)y$, $\forall n \in \mathbb{N}$. Thus, by construction, we have

$$A_n x_n = A_n R(\lambda, A_n) = \lambda R(\lambda, A_n)y - y \to \lambda R(\lambda, A)y - y = Ax.$$

(iv) $\Rightarrow$ (iii) Assume that $T(n, t)x \to T_t x$ for all $x \in X$, and uniformly on any compact interval. By the integral representation of the resolvent, $\forall \lambda > 0$, $\forall x \in X$ we have

$$\begin{aligned}\|R(\lambda, A_n)x - R(\lambda, A_n)x\| &= \left\| \int_0^\infty e^{-\lambda t}\big(T(t)x - T_n(t)x\big)dt \right\| \\ &\leq \int_0^\infty e^{-\lambda t}\|T(t)x - T_n(t)x\|dt.\end{aligned}$$

By the dominated convergence theorem, this converges to 0, so $R(\lambda, A_n)x \to R(\lambda, A)x$.

(iii) $\Rightarrow$ (iv) Assume that $R(\lambda, A_n)x \to R(\lambda, A)x$, $\forall x \in X$ for some $\lambda > 0$. Fix some $t_0 > 0$. For all $x \in X$ and $\forall t \in [0, t_0]$, we have, using the fact that $R(\lambda, A)$ and T_t commute and $R(\lambda, A_n)$ and $T(n, t)$ commute,

$$\begin{aligned}&\|(T_n(t)) - T(t))R(\lambda, A)x\| = \\ &\big\|T_n(t)(R(\lambda, A) - R(\lambda, A_n))x + R(\lambda, A_n)(T_n(t) - T(t))x + (R(\lambda, A_n) - R(\lambda, A))x\big\| \\ &\leq \underbrace{\|T_n(t)(R(\lambda, A) - R(\lambda, A_n))x\|}_{D_1(n)} + \underbrace{\|R(\lambda, A_n)(T_n(t) - T(t))x\|}_{D_2(n)} + \underbrace{\|(R(\lambda, A_n) - R(\lambda, A))x\|}_{D_3(n)}.\end{aligned}$$

The goal is now to show that $D_i(n) \to 0$ as $n \to \infty$, for $i = 1, 2, 3$.

Firstly, we note that $\|T_n(t)\| \leq 1$, $\forall n \in \mathbb{N}$ and $t \in [0, t_0]$. This implies by the uniform boundedness principle and continuity that:

$$\lim_{n\to\infty} D_1(n) = \lim_{n\to\infty} \|T_n(t)(R(\lambda, A) - R(\lambda, A_n))x\| = 0.$$

Additionally, this convergence is uniform, by the uniform boundedness principle.

Secondly, note that

$$D_3(n) = \|(R(\lambda, A_n) - R(\lambda, A))x\| \to 0$$

By assumption that $R(\lambda, A_n)x \to R(\lambda, A)x$. Additionally, this convergence is uniform since $[0, t_0]$ is compact and $t \mapsto T(t)x$ is uniformly continuous.

Thirdly, we note that

$$D_2(n) = \|R(\lambda, A_n)(T_n(t) - T(t))x\| \leq \|R(\lambda, A_n)\big(T(t) - T_n(t)\big)R(\lambda, A)x\|,$$

because $\|R(\lambda, A)\| \leq 1$. Next, we seek to show that above expression can be represented as an integral. We define

$$G(s) := T_n(t-s)R(\lambda, A_n)T(s)R(\lambda, A)x,$$

so that

$$G(t) - G(0) = R(\lambda, A_n)\big(T(t) - T_n(t)\big)R(\lambda, A)x.$$

Then

$$\begin{aligned}\frac{d}{ds}G(s) &= T_n(t-s)\Big(- A_nR(\lambda, A_n)T(s) + R(\lambda, A_n)T(s)A\Big)R(\lambda, A)x \\ &= T_n(t-s)\Big((\mathbb{1} - \lambda R(\lambda, A_n)R(\lambda, A) + R(\lambda, A_n)(-\mathbb{1} + \lambda R(\lambda, A)\Big)T(s)x \\ &= T_n(t-s)\big(R(\lambda, A) - R(\lambda, A_n)\big)T(s)x.\end{aligned}$$

Using the fact that $\|T_n(t-s)\| \leq 1$, this allows for the following bound on $D_2(n)$:

$$\begin{aligned}D_3(n) \leq \|R(\lambda, A_n)(T(t) - T_n(t))R(\lambda, A)x\| &\leq \int_0^t \|(R(\lambda, A) - R(\lambda, A_n)T(s)x\|ds \\ &\leq \sup_{s\in[0,t_0]} \|(R(\lambda, A) - R(\lambda, A_n)T(s)x\|.\end{aligned}$$

By repeating the same argument which showed $D_3(n) \to 0$, we have $D_2(n) \to 0$, uniformly on $[0, t_0]$, as $n \to \infty$

Thus, for any $x \in X$, we have that $\|(T(n,t) - T_t)R(\lambda, A)x\| \to 0$, uniformly on $[0, t_0]$. Since any $y \in D(A)$ can be written as $R(\lambda, A)x$, for some $x \in X$, it follows that $\|T(n,t)y - T_ty\| \to 0$ for all $y \in D(A)$ and uniformly on $[0, t_0]$. Thus, by the uniform boundedness principle and the fact that $\|T(n,t)y - T_ty\| \leq 2$, we have that $T(n,t)x \to T_tx$, for any $x \in X$ and uniformly on $[0, t_0]$. □

Theorem 1.33 (Second Trotter-Kato approximation theorem)**.** *Let X be a Banach space and suppose that for any $n \in \mathbb{N}$ $(T_n(t))_{t\geq 0}$ is a contraction semigroup on X with generator A_n. For $\lambda_0 > 0$ we have the implications*

$$\text{(i)} \implies \text{(ii)} \iff \text{(iii)}$$

involving statements listed below. Additionally if (i) holds, then $G = \bar{A}$.

(i) *There exists a densely defined unbounded operator $A : D(A) \subset X \to X$, a core D of A and $\lambda_0 > 0$ such that $\boldsymbol{R}(\lambda_0 - A)$ is dense in X and $A_nx \to Ax$, $\forall x \in D$.*

(ii) *The operators $R(\lambda_0, A_n)$ converge strongly to some $R \in B(X)$ with dense image (R is dominant).*

(iii) *The semigroups $(T_n(t))_{t\geq 0}$ converge strongly and uniformly on compact interval as $n \to \infty$, to a contraction semigroup $(T_t)_{t\geq 0}$ C_0-semigroup on X with infinitesimal generator G such that $R(\lambda, A) = R$*

Proof. Similarly to the first Trotter-Kato approximation theorem, we will prove for contraction semigroups. We follow the proof of Theorem 1.9 on page 144 of [4].

(i) $\Rightarrow$ (ii) Because each A_n generate a semigroup of contractions, they satisfy the HY-conditions. Thus, by Theorem 1.23, we have that each A_n is dissipative and $\lambda - A_n$ is surjective for all $\lambda > 0$ and $n \in \mathbb{N}$. This implies that A is also dissipative since

$$\|\lambda x - Ax\| = \lim_{n\to\infty} \|\lambda x - A_n x\| \geq \lambda \|x\|, \quad \forall x \in D.$$

Corollary 1.24 shows that A is closable and its closure $\bar{A}$ generates a contraction semigroup. In particular $\bar{A}$ satisfies the Hille-Yosida conditions so $(0, \infty) \subset \rho(\bar{A})$.

Set $R = R(\lambda_0, \bar{A})$. Its image is the domain of $\bar{A}$ so it is dense. We will show that $R = \lim_{n\to\infty} R(\lambda_0, A_n)$ with respect to the strong operator topology.

For any $x \in (\lambda_0 - \bar{A})(D)$, there exists $y \in D$ such that $(\lambda_0 - \bar{A})y = x$; in particular, $y = Rx$. This allows for the following calculation, for any $x \in (\lambda_0 - \bar{A})(D)$,

$$R(\lambda_0, A_n)x = R(\lambda_0, A_n)\big(\lambda_0 - A_n - (\lambda_0 - A_n) + (\lambda_0 - \bar{A})\big)y$$

$$= y + R(\lambda_0, A_n)(A_n y - \bar{A}y) \to y = Rx.$$

By the Hille-Yosida theorem, we have $\|R(\lambda_0, A_n)\| \leq \frac{1}{\lambda_0}$, for all $n \in \mathbb{N}$. This allows for the following computation

$$\|R(\lambda_0, A_n)(A_n y - \bar{A}y)\| \leq \|R(\lambda_0, A_n)\| \cdot \|A_n y - \bar{A}y\| \leq \frac{1}{\lambda_0}\|A_n y - \bar{A}y\| \to 0.$$

Thus $R(\lambda_0, A_n) \to R$ with respect to the strong operator topology because $(\lambda_0 - \bar{A})(D)$ is dense.

(iii) $\Rightarrow$ (ii) Follows directly from the first Trotter-Kato Approximation theorem.

(ii) $\Rightarrow$ (iii) Assume that $R(\lambda_0, A_n)$ converge strongly to some $R \in B(X)$ with dense image. Define the following family of operators on X:

$$R(\lambda), \lambda > 0 \quad \text{by} \quad R(\lambda)x := \lim_{n\to\infty} R(\lambda, A_n)x.$$

$R(\lambda)$ form a pseudo-resolvent since $R(\lambda, A_n)$ are a pseudo-resolvent for each $n \in \mathbb{N}$. Note that $\forall \lambda > 0$, $\|\lambda R(\lambda)\| \leq 1$ by the uniform boundedness principle, and has dense image. By Lemma 1.31 implies there exists $B : D(B) \to X$ densely defined closed operator such that $R(\lambda) = R(\lambda, B)$, $\forall \lambda > 0$. By the Hille-Yosida theorem, this generates a C_0-semigroup. Then, by the first Trotter-Kato Theorem, the semigroups converge.

Additionally, if (i) holds, then $R(\lambda_0) = R(\lambda_0, G)$. Since D is a core of $\bar{A}$, we have

$$R(\lambda_0)(\lambda_0 - A)x = x, \qquad \forall x \in D.$$

Additionally because $\boldsymbol{R}(\lambda_0 - \bar{A})$ is X, we also have

$$R(\lambda)(\lambda_0 - \bar{A})x = x, \qquad \forall x \in X.$$

Next, we note that because $\lambda_0 \in \rho(\bar{A})$, $R(\lambda, G) = R(\lambda, \bar{A})$ because both operators are continuous and agree on a dense subset. Thus $G = \bar{A}$. ☐

1.5. **Chernoff product formula.**

Lemma 1.34. *Let $A \in B(X)$ satisfy $\|A^n\| \leq M$, for all $n \in \mathbb{N}$. Then*

$$\| \exp(n(A - \mathbb{1}))x - A^n x\| \leq \sqrt{n} M \|Ax - x\|.$$

Proof. We follow the proof of Lemma 2.1 on page 149 of [4]. Fix any $n \in \mathbb{N}$. Then we have the following algebraic manipulation,

$$\exp(n(A - \mathbb{1})) - A^n = e^{-n}(e^{nA} - e^n A^n) = e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!}(A^k - A^n).$$

Additionally, for any $k \in \mathbb{N}$, with $k > n$, we have

$$A^k - A^n = \sum_{j=n}^{k-1} A^{j+1} - A^j = \sum_{j=n}^{k-1} A^j(A - \mathbb{1}).$$

This implies the following bound for any $k > n$, $k \in \mathbb{N}$,

$$\|A^k x - A^n x\| \leq (k - n)M\|Ax - x\|.$$

This bound yields

$$\| \exp(n(A - \mathbb{1}))x - A^n x\| \leq e^{-n} M \|Ax - x\| \sum_{k=0}^{\infty} \left(\frac{n^k}{k!}\right)^{1/2} \left(\frac{n^k}{k!}\right)^{1/2} (k - n),$$

and by the ℓ^2 Cauchy Schwarz inequality,

$$\leq e^{-n} M \|Ax - x\| \Bigg(\underbrace{\sum_{k=0}^{\infty} \frac{n^k}{k!}}_{=:\alpha} \Bigg)^{1/2} \Bigg(\underbrace{\sum_{k=0}^{\infty} \frac{n^k}{k!}(k - n)^2}_{=:\beta} \Bigg)^{1/2}.$$

We have $\alpha = e^n$ and

$$\beta = \sum_{k \geq 0} k^2 \frac{n^k}{k!} - 2n \sum_{k \geq 0} k \frac{n^k}{k!} + n^2 \sum_{\geq 0} \frac{n^k}{k!}$$

(use (2.5) and (2.6) with $\lambda = n$)

$$= e^n(n + n^2) - 2n^2 e^n + n^2 e^n = n e^n.$$

We deduce

$$\| \exp(n(A - \mathbb{1}))x - A^n x\| \leq e^{-n} M \|Ax - x\| e^{n/2} (n e^n)^{1/2} = \sqrt{n} M \|Ax - x\|.$$

□

Theorem 1.35 (Chernoff product formula)**.** *Suppose that $\big(V(t)\big)_{t \geq 0}$ be a family of linear contractions on the Banach space X with there following properties.*

(i) $V(0) = \mathbb{1}$.

(ii) *There exists a dense subspace $D \subset X$ such that for any $x \in D$ the limit*

$$\lim_{h \searrow 0} \frac{1}{h}(V(h)x - x)$$

exists. We denote it by Ax.

(iii) *There exists* $\lambda_0 > 0$ *such that* $(\lambda_0 - A)(D)$ *is dense in* X.

Then A *is closable, its closure* $\bar{A}$ *generates a semigroup of contractions* $(T_t)_{t\geq 0}$ *and*

$$T_t x = \lim_{n\to\infty} V\left(\frac{t}{n}\right)^n x, \quad \forall x \in X. \tag{1.6}$$

Additionally, this convergence is uniform on compact intervals.

Proof. We follow the proof of [4, Thm.2.2]. Define

$$A_n(s) = \frac{n}{s}(V(s/n) - \mathbb{1}) \in B(X), \quad \forall s > 0.$$

Clearly $A_n(s)x \to Ax$ for all $x \in D$ as $n \to \infty$, and uniformly on $s \in [s_0, s_1]$. Note that

$$\|e^{tA_n(s)}\| \leq e^{-tn/s} \left\|\exp\left(\frac{tn}{s}V(s/n)\right)\right\| \leq e^{-tn/s} \sum_{m=0}^{\infty} \frac{(tn)^m \|V(s/n)^m\|}{s^m m!} \leq 1.$$

Because both D and $(\lambda - A)(D)$ are dense in X, we can apply the second Trotter-Kato Theorem 1.33 to the sequence of contraction semigroups $\exp(tA_n(s))$, to get that $\bar{A}$ generates a contraction semigroup, which will be called T_t. Additionally, by Theorem 1.33, T_t also satisfies

$$\|T_t x - e^{tA_n(s)}x\| \to 0 \quad \text{ as } n \to \infty \text{ and uniformly for } s \in [s_0, s_1].$$

Next, we seek to show the equality in 1.6. By Lemma 1.34, we calculate the following, for any $n \in \mathbb{N}$, $x \in X$ and $s > 0$,

$$\left\|\exp\left(sA_n(s)\right)x - V\left(\frac{s}{n}\right)^n x\right\| = \left\|\exp\left(n\left(V\left(\frac{s}{n}\right) - \mathbb{1}\right)\right)x - V\left(\frac{s}{n}\right)^n x\right\|$$
$$\leq \sqrt{n}\left\|V\left(\frac{s}{n}\right)x - x\right\| = \frac{s}{\sqrt{n}}\|A_n(s)x\|.$$

Finally we note that $\frac{s}{\sqrt{n}}\|A_n(s)x\| \to 0$ as $n \to \infty$, for all $x \in D$, and uniformly on any compact $s \in [s_0, s_1]$. Because D is dense, and $\left\|\exp\left(sA_n(s)\right)x - V\left(\frac{s}{n}\right)^n x\right\| \leq 2$ for all $s \geq 0$, $n \in \mathbb{N}$ and $x \in X$, we have the desired convergence for any $x \in X$ by the uniform boundedness principle. □

Example 1.36 (Yosida's approximation)**.** Suppose that $(T_t)_{t\geq 0}$ is a semigroup of contractions on the Banach space X with generator A. For any $t > 0$ we set

$$V(t) = \big(1 - tA\big)^{-1} = t^{-1}\big(t^{-1} - A\big)^{-1} = t^{-1}R(t^{-1}, A).$$

SSet $\lambda := t^{-1}$. Since $\|R(\lambda, A)\| \leq \frac{1}{\lambda}$ we deduce $\|V(t)\| \leq 1$. Note that

$$\frac{1}{t}\big(V(t) - 1\big) = \lambda\big(\lambda R(\lambda, A) - 1\big) = \lambda^2 R(\lambda, A) - \lambda = A(\lambda).$$

Lemma 1.14 shows that

$$\lim_{\lambda\to\infty} A_\lambda x = Ax, \quad \forall x \in D(A).$$

Hence

$$\lim_{t\searrow 0} \frac{1}{t}\big(V(t) - 1\big)x = Ax, \quad \forall x \in D(A).$$

We deduce from Chernoff's product formula

$$\left(1 - \frac{t}{n}A\right)^{-n} x \to T_t x, \quad \forall x \in X.$$

□

2. Semigroups of probability measures

In this section we will describe some probabilistic applications of the theory of semigroups we have developed in the previous section.

2.1. **The space of Borel probability measures on $\mathbb{R}$.**

Denote by $\mathrm{Prob}(\mathbb{R})$ the space of probability measures on $\mathbb{R}$. Denote by $\mathscr{B}$ the sigma-algebra of Borel subsets of $\mathbb{R}$, by $C_b(\mathbb{R})$ the space of bounded continuous functions $\mathbb{R} \to \mathbb{R}$, and by $C_0(\mathbb{R})$ the space of continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that

$$\lim_{x \to \pm\infty} f(x) = 0.$$

Denote by $\| - \|$ the sup-norm on $C_b(\mathbb{R})$. For $f \in C_b(\mathbb{R})$ and $\mu \in \mathrm{Prob}(\mathbb{R})$ we set

$$\mu\big[\, f \,\big] := \int_{\mathbb{R}} f(x)\mu[dx].$$

The set of atoms of $\mu \in \mathrm{Prob}(\mathbb{R})$ is the collection

$$\mathscr{A}_\mu = \big\{ x \in \mathbb{R}; \quad \mu\big[\, \{x\} \,\big] > 0 \,\big\}.$$

Clearly the set of atoms is at most countable since for any $n \in \mathbb{N}$ the collection

$$\big\{ x \in \mathbb{R}; \quad \mu\big[\, \{x\} \,\big] > 1/n \,\big\}$$

has cardinality $< n$.

Definition 2.1. A sequence $(\mu_n)n \in \mathbb{N}$ in $\mathrm{Prob}(\mathbb{R})$ is said to converge weakly to $\mu \in \mathrm{Prob}(\mathbb{R})$, and we write this $\mu_n \Rightarrow \mu$, if

$$\lim_{n \to \infty} \mu_n\big[\, f \,\big] = \mu\big[\, f \,\big], \quad \forall f \in C_b(\mathbb{R}).$$

□

The Fourier transform of a measure $\mu \in \mathrm{Prob}(\mathbb{R})$ is the function

$$\widehat{\mu} : \mathbb{R} \to \mathbb{C}, \quad \widehat{\mu}(\xi) = \int_{\mathbb{R}} e^{\boldsymbol{i}\xi x} \mu[dx].$$

For a proof of the following result we refer to [8, Sec. 2.2].

Theorem 2.2. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in $\mathrm{Prob}(\mathbb{R})$ and $\mu \in \mathrm{Prob}(\mathbb{R})$. The following statements are equivalent.*

(i) *The sequence (μ_n) converges weakly to μ.*

(ii) *For any $f \in C_0(\mathbb{R})$*

$$\mu_n\big[\, f \,\big] \to \mu\big[\, f \,\big].$$

(iii) *For any $a, b \in \mathbb{R} \setminus \mathscr{A}_\mu$, $a < b$*

$$\lim_{n \to \infty} \mu_n\big[\, (a, b) \,\big] = \mu\big[\, (a, b) \,\big].$$

(iv) *For any* $\xi \in \mathbb{R}$
$$\lim_{n\to\infty} \widehat{\mu}_n(\xi) = \widehat{\mu}(\xi).$$
□

For any $\mu \in \mathrm{Prob}(\mathbb{R})$ and $f \in C_b(\mathbb{R})$ we denote by $T_\mu f$ the function $\mathbb{R} \to \mathbb{R}$ defined by
$$T_\mu f(x) = \int_{\mathbb{R}} f(x+y)\mu[dy].$$
Clearly $T_\mu f \in C_b(\mathbb{R})$, $\forall f \in C_b(\mathbb{R})$. The dominated convergence theorem implies that
$$T_\mu f \in C_0(\mathbb{R}), \quad \forall f \in C_0(\mathbb{R}).$$
Note that
$$\|T_\mu f\| \leq \|f\|, \ \forall f \in C_b(\mathbb{R}), \ \forall \mu \in \mathrm{Prob}(\mathbb{R}).$$
For any random variable Y we set $T_Y = T_{\mathbb{P}_Y}$, where $\mathbb{P}_Y \in \mathrm{Prob}(\mathbb{R})$ is the distribution of Y. Note that for any $f \in C_b(\mathbb{R})$ we have
$$T_Y f(x) = \mathbb{E}\big[\, f(x+Y)\,\big], \quad x \in \mathbb{R},$$
where $\mathbb{E}\big[\, - \,\big]$ denotes the expectation of a random variable.

Theorem 2.3. *Let* $(\mu_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathrm{Prob}(\mathbb{R})$ *and* $\mu \in \mathrm{Prob}(\mathbb{R})$. *The following statements are equivalent.*

(i) *The sequence* (μ_n) *converges weakly to* μ.
(ii) *For any* $f \in C_0(\mathbb{R})$
$$\lim_{n\to\infty} \|T_{\mu_n} f - T_\mu f\| = 0.$$

Proof. (i) $\Rightarrow$ (ii) Let $f \in C_0(\mathbb{R})$. For each $x \in \mathbb{R}$ we define
$$f_x : \mathbb{R} \to \mathbb{R}, \quad f_x(y) = f(x+y), \ \forall y \in \mathbb{R}.$$
Then
$$T_{\mu_n} f(x) = \mu_n\big[\, f_x \,\big].$$
Since f is uniformly continuous the map
$$\mathbb{R} \ni x \mapsto f_x \in C_0(\mathbb{R})$$
is also uniformly continuous with respect to the sup-norm.

Fix $\varepsilon > 0$. Since $\mu_n \Rightarrow \mu$ there exists $M > 0$ such that
$$\mu_n\big[\, \{|y| > M\} \,\big], \ \mu_n\big[\, \{|y| > M\} \,\big] < \varepsilon, \ \forall n \in \mathbb{N}$$
We can assume that $M, -M$ are not atoms of μ. We have
$$\begin{aligned} \big|\mu_n\big[\, f_x \,\big] - \mu\big[\, f \,\big]\big| &\leq \left| \int_{[-M,M]} f_x(y)\mu_n[dy] - \int_{[-M,M]} f_x(y)\mu[dy] \right| \\ &\quad + \int_{|y|>M} |f|\mu_n[dy] + \int_{|y|>M} |f|\mu[dy] \\ &\leq \left| \int_{[-M,M]} f_x(y)\mu_n[dy] - \int_{[-M,M]} f_x(y)\mu_n[dy] \right| + 2\varepsilon\|f\|. \end{aligned}$$

Hence

$$\sup_{x\in\mathbb{R}} \big|\mu_n[\,f_x\,] - \mu[\,f\,]\big| \le \sup_{x\in\mathbb{R}} \left| \int_{[-M,M]} f_x(y)\mu_n[dy] - \int_{[-M,M]} f_x(y)\mu_n[dy] \right| + 2\varepsilon\|f\|.$$

Since $f\in C_0(\mathbb{R})$, $\forall\varepsilon>0$ there exists $K>0$ such that

$$\sup_{y\in[-M,M]} |f_x(y)| < \varepsilon,\ \forall |x|>K.$$

Hence

$$\left| \int_{[-M,M]} f_x(y)\mu_n[dy] - \int_{[-M,M]} f_x(y)\mu_n[dy] \right| < 2\varepsilon,\ \forall |x|>K,\ \forall n\in\mathbb{N}. \tag{2.1}$$

We deduce from (2.1) that

$$\sup_{|x|>K} \big|\mu_n[\,f_x\,] - \mu[\,f\,]\big| \le 2\varepsilon + 2\varepsilon\|f\|. \tag{2.2}$$

Consider now the continuous functions

$$g, g_n : [-K,K]\to\mathbb{R},\ g_n(x) = \int_{[-M,M]} f_x(y)\mu_n[dy],\ g(x) = \int_{[-M,M]} f_x(y)\mu_n[dy].$$

Since $\mu_n \Rightarrow \mu$, and $M,-M$ are not atoms of μ we deduce

$$g_n(x)\to g(x),\quad \forall x\in[-K,K].$$

The sequence (g_n) is equicontinuous since $x\mapsto f_x$ is uniformly continuous with respect to the sup-norm. Hence g_n converges uniformly to g on $[-K,K]$, i.e.,

$$\lim_{n\to\infty} \sup_{|x|\le K} |g_n(x)-g(x)| = 0.$$

We have

$$\sup_{|x|\le K} \big|\mu_n[\,f_x\,] - \mu[\,f\,]\big| \le \sup_{|x|\le K} |g_n(x)-g(x)| + 2\varepsilon\|f\|.$$

Hence

$$\limsup_{n\to\infty} \sup_{|x|\le K} \big|\mu_n[\,f_x\,] - \mu[\,f\,]\big| \le 2\varepsilon\|f\|.$$

Using (2.2) we deduce that $\forall\varepsilon>0$ we have

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}} \big|\mu_n[\,f_x\,] - \mu[\,f\,]\big| \le 2\varepsilon + 2\varepsilon\|f\|.$$

This proves (ii). The implication (ii) ⇒ (i) is immediate since

$$\mu[\,f\,] = T_\mu f(0),\quad \forall\mu\in\mathrm{Prob}(\mathbb{R}),\ \forall f\in C_b(\mathbb{R}).$$

□

Definition 2.4 (Convolution of measures)**.** Let $\mu,\nu\in\mathrm{Prob}(\mathbb{R})$. Then the convolution of μ and ν is the Borel measure $\mu*\nu$ on $\mathbb{R}$ defined if

$$\mu*\nu[\,B\,] = \int_{\mathbb{R}} \nu[\,B-x\,]\mu[dx],$$

for any $B\in\mathscr{B}$. Clearly $\mu*\nu$ is a probability measure.

One can prove (see [8, Sec. 1.3.6]) that $\mu * \nu = \alpha_{\#}(\mu \otimes \nu)$-the pushforward of the product measure $\mu \otimes \nu$ on $\mathbb{R}^2$ via the addition map $\alpha : \mathbb{R}^2 \to \mathbb{R}$, $\alpha(x, y) = x + y$. More explicitly, this means that for any Borel set $B \subset \mathbb{R}$

$$\mu * \nu(B) = \mu \otimes \nu(\alpha^{-1}(B)) =: \alpha_{\#}(\mu \otimes \nu)(B). \tag{2.3}$$

This leads to the following probabilistic interpretation of the operation of convolution. Let

$$X, Y : (\Omega, \mathscr{S}, \mathbb{P}) \to \mathbb{R}$$

be two *independent* random variables with distributions $\mathbb{P}_X, \mathbb{P}_Y \in \mathrm{Prob}(\mathbb{R})$. Then

$$\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y. \tag{2.4}$$

In particular

$$T_{X+Y} = T_X \cdot T_Y.$$

For a proof of the following result we refer to [8].

Lemma 2.5. *The convolution* $* : \mathrm{Prob}(\mathbb{R}) \times \mathrm{Prob}(\mathbb{R}) \to \mathrm{Prob}(\mathbb{R})$ *is commutative and associative. The Dirac measure* δ_0 *is the identity element with respect to the convolution. In other words* $\big(\mathrm{Prob}(\mathbb{R}), *)$ *is a commutative semigroup with* 1.

Moreover, if μ *and* ν *are absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}$,

$$\mu[dx] = \rho_\mu dx, \quad \nu[dx] = \rho_\nu(x)dx,$$

then

$$\mu * \nu[dx] = \rho_\mu * \rho_\nu(x)dx,$$

where

$$\rho_\mu * \rho_\nu(x) = \int_{\mathbb{R}} \rho_\mu(x - y)\rho_\mu(y)dy.$$

□

From (2.3) one obtains immediately that

$$\widehat{\mu * \nu}(\xi) = \widehat{\mu}(\xi) \cdot \widehat{\nu}(\xi), \quad \forall \xi \in \mathbb{R}, \quad \forall \mu, \nu \in \mathbb{R}.$$

The following result follows immediately from the definition.

Proposition 2.6. *For any* $\mu, \nu \in \mathrm{Prob}(\mathbb{R})$

$$T_\mu T_\nu = T_{\mu * \nu}, \quad \forall \mu, \nu \in \mathrm{Prob}(\mathbb{R}).$$

In other words the correspondence $\mu \to T_\mu$ *is a continuous morphism from the semigroup* $(\mathrm{Prob}(\mathbb{R}), *)$ *to* $(B(X), \cdot)$. □

2.2. **Convolution semigroups of probability measures.** A family of Borel *probability measures* $(\mu_t)_{t \geq 0}$ on $\mathbb{R}$ is called a *convolution semigroup* if it satisfies the following conditions

(i) $\mu_0 = \delta_0$.
(ii) $\mu_{s+t} = \mu_s * \mu_t$.
(iii) $\mu_t \Rightarrow \mu_0$ as $t \searrow 0$.

Condition (iii) signifies that the measure μ_t converges weakly to the Dirac measure δ_0. It implies that the map

$$[0,\infty) \ni t \mapsto \mu_t \in \mathrm{Prob}(\mathbb{R})$$

is continuous with respect to the topology of weak convergence on $\mathrm{Prob}(\mathbb{R})$. We see that $(\mu_t)_{t\geq 0}$ is a convolution semigroup if $\forall t, s \geq 0$, $\xi \in \mathbb{R}$

$$\widehat{\mu}_{t+s}(\xi) = \widehat{\mu}_t(\xi)\cdot\widehat{\mu}_s(\xi),\ \ \widehat{\mu}_0(\xi) = 1,$$

and the map $t \to \widehat{\mu}_t(\xi)$ is continuous.

Theorem 2.3 has the following immediate consequence

Corollary 2.7. *Let $(\mu_t)_{t\geq 0}$ be a convolution semigroup of probability measures. Then the induced operator semigroup,*

$$T_{\mu_t} f(x) := \int_{\mathbb{R}} f(x+y)\mu_t[dy],$$

is a strongly continuous semigroup of contractions on $C_0(\mathbb{R})$. In other words,

$$\lim_{t\searrow 0} \|T_{\mu_t} f - f\| = 0, \quad \forall f \in C_0(\mathbb{R}),$$

where $\|-\|$ denotes the sup-norm on $C_0(\mathbb{R})$. □

Example 2.8 (The translation semigroup)**.** The family of Dirac measures $(\delta_t)_{t\geq 0}$ o is a convolution semigroup. Indeed

$$\widehat{\delta}_t(\xi) = e^{\boldsymbol{i}t\xi}$$

and obviously

$$\widehat{\delta}_t(\xi)\widehat{\delta}_t(\xi) = \widehat{\delta}_{t+s}(\xi).$$

□

Example 2.9 (The heat semigroup)**.** For each $t > 0$ we denote by $\gamma_t \in \mathrm{Prob}(\mathbb{R})$ the Gaussian measure with mean 0 and variance t. More precisely

$$\gamma_t[dx] = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \lambda[dx].$$

Its Fourier transform is

$$\widehat{\gamma}_t(\xi) = e^{-t\xi^2/2}.$$

We set $\gamma_0 := \delta_0$. The family $\big(\gamma_t\big)_{t\geq 0}$ is a convolution semigroup since

$$\widehat{\gamma}_t(\xi)\widehat{\gamma}_s(\xi) = \widehat{\gamma}_{t+s}(\xi), \quad \forall s, t \geq 0.$$

□

Example 2.10. For $\lambda > 0$ we define the Poisson measure $\mu_\lambda \in \mathrm{Prob}(\mathbb{R})$ to be

$$\mu_\lambda = \sum_{k\geq 0} e^{-\lambda}\frac{\lambda^k}{k!}\delta_k.$$

Its Fourier transform is

$$\widehat{\mu}_\lambda(\xi) = \sum_{n\geq 0} e^{-\lambda}\frac{e^{ik\xi}\lambda^k}{k!} = e^{e^{i\xi}\lambda-\lambda} = e^{\lambda(e^{i\xi}-1)}.$$

This proves that for any $\lambda > 0$ the family $\big(\mu_{t\lambda}\big)_{t\geq 0}$ is a convolution semigroup. This is called the *Poisson semigroup* with parameter λ.

Note that the function $x \mapsto e^{tx}$ is μ_λ-integrable for any $t \in \mathbb{R}$ and

$$M_\lambda(t) := \int_{\mathbb{R}} e^{tx}\mu_\lambda\big[\,dx\,\big] = \sum_{n\geq 0} e^{-\lambda}\frac{e^{kt}\lambda^k}{k!} = e^{\lambda(e^t-1)}.$$

We can derivate under the integral sign in the above equality and we deduce

$$M_\lambda'(0) = \int_{\mathbb{R}} x\mu_\lambda\big[\,dx\,\big] = \sum_{n\geq 0} e^{-\lambda}k\frac{\lambda^k}{k!}$$

Note that $M_\lambda'(t) = \lambda e^t e^{\lambda(e^t-1)}$ so

$$\lambda e^\lambda = e^\lambda M_\lambda'(0) = \sum_{k\geq 0} k\frac{\lambda^k}{k!}. \tag{2.5}$$

We have

$$M_\lambda''(t) = \lambda e^t e^{\lambda(e^t-1)} + \lambda^2 e^{2t} e^{\lambda(e^t-1)}$$

Hence $M_\lambda''(0) = \lambda + \lambda^2$ and we deduce

$$(\lambda+\lambda^2)e^\lambda = e^\lambda M_\lambda''(0) = \sum_{k\geq 0} k^2\frac{\lambda^k}{k!}. \tag{2.6}$$

□

2.3. **The translation semigroup.** The family of Dirac measures $(\delta_t)_{t\geq 0}$ ois a convolution semigroup. Moreover

$$T_{\delta_t} f(x) = f(x+t),\ \forall x \in \mathbb{R},,\ \forall t \geq 0.$$

We denote by A the infinitesimal generator of T_t. We define inductively

$$C_0^k(\mathbb{R}) := \big\{\, f \in C_0^{k-1}(\mathbb{R}) \cap C^2(\mathbb{R});\ f' \in C_0^{k-1}(\mathbb{R})\,\big\}.$$

Note that $C_0^k(\mathbb{R})$ is dense in $C_0(\mathbb{R})$ with respect to the sup-norm. Moreover

$$T_t\big(\,C_0^k(\mathbb{R})\,\big) \subset C_0^k(\mathbb{R}).$$

Proposition 2.11. *The space $C_0^2(\mathbb{R})$ is contained in $D(A)$ and*

$$Af = f',\ \forall f \in C_0^2(\mathbb{R}).$$

In particular, $C_0^2(\mathbb{R})$ is a core of A.

Proof. Let $f \in C_0^2(\mathbb{R})$. For any $t > 0$ we deduce from Taylor's formula with Lagrange remainder that

$$T_t f(x) = f(x+t) - f(x) = f'(x)t + \frac{t^2}{2} f''(\xi),\ \xi \in (x, x+t).$$

We deduce that for any $t > 0$ we have

$$\sup_{x\in\mathbb{R}} \left| \frac{f(x+t) - f(x)}{t} - f'(x) \right| \leq \frac{\|f''\|}{2} t$$

Hence

$$\lim_{t\searrow 0} \left\| \frac{1}{t}\big(T_t f - f \big) - f' \right\| = 0$$

This proves that $\frac{1}{t}\big(T_t f - f\big)$ converges in the sup-norm to f'. Hence $f \in D(A)$ and $Af = f'$. The fact that $C_0^2(\mathbb{R})$ is a core of A now follows from Proposition 1.27. □

For $h > 0$ define

$$\Delta_h : C_0(\mathbb{R}) \to C_0(\mathbb{R}), \quad \Delta_h f(x) = \frac{f(x+h) - f(x)}{h} = \frac{1}{h}\big(T_h f(x) - f(x) \big)$$

The first Trotter-Kato approximation theorem implies that

$$\lim_{h\searrow 0} e^{t\Delta_h} = T_t, \quad t \geq 0,$$

in the strong operator topology. Hence, for any $f \in C_0(\mathbb{R})$, $t > 0$ and any $x \in \mathbb{R}$

$$u(x+t) = \lim_{h\searrow 0} e^{t\Delta_h} f(x) \tag{2.7}$$

uniformy in x.

One should compare this with the Taylor expansion for a real analytic function. If we set $D = \frac{d}{dx}$, and f is real analytic, then we have a Taylor expansion

$$f(x+t) = \sum_{n\geq 0} \frac{t^n}{n!} D^n f(x) = e^{tD} f(x) \tag{2.8}$$

where the right-hand-side converges uniformly for small t. Formally, (2.8) is obtained from (2.7) by letting $h \searrow 0$ since $\Delta_h \to D$ as $h \searrow 0$.

The equality (2.7) can be rewritten as

$$e^{\frac{t}{h}(T_h - 1)} f \to T_t f \quad \text{as } h \to 0.$$

Equivalently

$$e^{-\frac{t}{h}} e^{\frac{t}{h} T_h} f \to T_t f. \tag{2.9}$$

Let μ_λ be the Poisson measure parameter $\lambda > 0$ described in Example 2.10. Consider the continuous map

$$[0,\infty) \ni s \mapsto F_h(s) = T_{sh} f \in C_0(\mathbb{R}).$$

Define

$$F_{\lambda,h}(s) := \int_{[0,\infty)} F_h(s) \mu_\lambda\big[\, ds \,\big].$$

The right-hand-side of the above equality is an average of the bounded function $F_h(s)$ with respect to the measure μ_λ. Then

$$e^{-\frac{t}{h}}e^{\frac{t}{h}T_h}f = F_{t/h,h}(s). \tag{2.10}$$

The equality (2.9) shows that if $\lambda \nearrow \infty$

$$\lim_{\lambda\to\infty} F_{\lambda,t/\lambda}(1) = T_t f = F_t(1). \tag{2.11}$$

For a very interesting discussion of the probabilistic meaning of (2.10) and implications of the above "accident" we refer to [6, Sec. VII.5,VII.6, X.9].

2.4. **The heat semigroup.** Consider the heat semigroup (γ_t) described in Example 2.9. Set $H_t := T_{\gamma_t}$. The contraction semigroup $(H_t)_{t\geq 0}$ is also referred to as the *heat semigroup*. Denote by A the generator of the heat semigroup. Note that

$$H_t f(x) = \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} f(x+y)e^{-\frac{y^2}{2t}}dy = \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} f(z)e^{-\frac{(z-x)^2}{2t}}dz.$$

This proves that $H_t f\in C^\infty(\mathbb{R})$, $\forall f\in C_0(\mathbb{R})$. Note that $H_t C_0^k(\mathbb{R})\subset C_0^k(\mathbb{R})$, $\forall k$.

Proposition 2.12. *The subspace $C_0^3(\mathbb{R})$ is a core of A and*

$$Af = \frac{1}{2}f'',\quad \forall f\in C_0^3(\mathbb{R}).$$

Proof. Let $f\in C_0^3(\mathbb{R})$. We have

$$H_f fx) - f(x) = \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} f(x+y)e^{-\frac{y^2}{2t}}dy - f(x)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x+t^{1/2}z)e^{-z^2/2}dz - \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} f(x)e^{-z^2/2}dz$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\big(f(x+t^{1/2}z) - f(x) - f'(x)t^{1/2}z\big)e^{-z^2/2}dz$$

Using the Taylor expansion with Lagrange remainder we deduce that

$$f(x+t^{1/2}z) - f(x) - f'(x)t^{1/2}z = \frac{t}{2}f''(x)z^2 + \frac{1}{3!}f^{(3)}(\eta)t^{3/2}$$

for some $\eta\in(x, x+t^{1/2}z)$. Hence

$$\left|\frac{f(x+t^{1/2}z) - f(x) - f'(x)t^{1/2}z}{t} - \frac{1}{2}f''(x)z^2\right| \leq \frac{t^{1/2}|z|^3}{6}\|f^{(3)}\|.$$

Since

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} z^2e^{-z^2/2}dz = 1$$

we deduce that

$$\frac{1}{t}\big(H_t f(x) - f(x)\big) - \frac{1}{2}f''(x)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\frac{f(x+t^{1/2}z) - f(x) - f'(x)t^{1/2}z}{t} - \frac{1}{2}f''(x)z^2\right)e^{-z^2/2}dz$$

Hence, $\forall x \in \mathbb{R}$ we have

$$\left| \frac{1}{t}\big(H_t f(x) - f(x) \big) - \frac{1}{2} f''(x) \right|$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \frac{f(x+t^{1/2}z) - f(x) - f'(x)t^{1/2}z}{t} - \frac{1}{2} f''(x) z^2 \right| e^{-z^2/2} dz$$

$$\leq \frac{\|f^{(3)}\| t^{1/2}}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{|z|^3}{6} e^{-z^2/2} dz$$

□

Here is a remarkable consequence of Proposition 2.12 and Remark 1.7.

Corollary 2.13. *For any function $f_0 \in C_0(\mathbb{R})$ there exists a unique function $u \in C^0\big([0,\infty) \times \mathbb{R} \big) \cap C^\infty\big((0,\infty) \times \mathbb{R} \big)$ such that*

$$\begin{gathered} x \mapsto u(t,x) \in C_0(\mathbb{R}), \quad \forall t \geq 0, \\ \lim_{t \searrow 0} \sup_{x \in \mathbb{R}} |u(t,x) - f_0(x)| = 0, \end{gathered} \tag{2.12}$$

$$\partial_t u(t,x) - \frac{1}{2} \partial^2_{xx} u(t,x) = 0, \quad \forall (t,x) \in (0,\infty) \times \mathbb{R}. \tag{2.13}$$

□

Suppose now that $\mu \in \mathrm{Prob}(\mathbb{R})$ is a probability measure satisfying the conditions

$$\int_{\mathbb{R}} x \mu\big[dx \big] = 0, \quad \int_{\mathbb{R}} x^2 \mu\big[dx \big] = 1. \tag{2.14}$$

Fix a random variable X with distribution μ, $\mathbb{P}_X = \mu$. Then we can rewrite (2.14) as

$$\mathbb{E}\big[X \big] = 0, \quad \mathrm{Var}\big[X \big] = \mathbb{E}\big[X^2 \big] = 1. \tag{2.15}$$

For $t > 0$ we denote by $\mathscr{R}_t$ the rescaling map

$$\mathscr{R}_t : \mathbb{R} \to \mathbb{R}, \quad \mathscr{R}_t(x) = tx.$$

We set

$$\mu_t := (\mathscr{R}_t)_\# \mu \iff \mu_t\big[(a,b) \big] = \mu\big[(t^{-1}a, t^{-1}b) \big], \quad \forall a,b,\ t > 0.$$

Equivalently, μ_t is the distribution on tX. Note that $\mu_t \Rightarrow \delta_0$. Moreover

$$T_{\mu_t} f(x) = T_{tX} f(x) = \int_{\mathbb{R}} f(x+ty) \mu\big[dy \big].$$

We set $V(t) = T_{\mu_t}$.

Proposition 2.14. *For any $f \in C^3_0(\mathbb{R})$ we have*

$$\lim_{t \searrow 0} \left\| \frac{1}{t}\big(V(t) - \mathbb{1} \big) f - \frac{1}{2} f'' \right\| = 0.$$

Proof. Let $f \in C_0^3(\mathbb{R})$. Using (2.14) we deduce that

$$\big(V_\mu(t) - \mathbb{1}\big)f(x) = \int_{\mathbb{R}} \Big(f\big(x + t^{1/2}y\big) - f(x) - t^{1/2}yf'(x) \Big)\mu\big[\,dy\,\big],$$

$$\frac{1}{t}\big(V_\mu(t) - \mathbb{1}\big)f(x) - \frac{1}{2}f''(x)$$

$$= \int_{\mathbb{R}} \underbrace{\frac{1}{t}\Big(f\big(x + t^{1/2}y\big) - f(x) - f'(x)t^{1/2}y - \frac{1}{2}f''(x)ty^2 \Big)}_{=U_t(x,y)}\mu\big[\,dy\,\big].$$

Using Taylor's formula with Lagrange remainder

$$f\big(x + t^{1/2}y\big) - f(x) - f'(x)t^{1/2}y = \frac{1}{2}f''(\xi)ty^2$$

for some $\xi = \xi_{x,y} \in (x, x + t^{1/2}y)$. Hence

$$\Big| f\big(x + t^{1/2}y\big) - f(x) - f'(x)t^{1/2}y - \frac{1}{2}f''(x)ty^2 \Big| = \Big| \frac{1}{2}f''(\xi) - \frac{1}{2}f''(x) \Big| \cdot ty^2$$

$$\leq \min\Big(t\|f\|_{C^2}|y|^2, \frac{1}{2}\|f\|_{C^3}t^{3/2}|y|^3 \Big)$$

$$\leq t\|f\|_{C^3}\min\Big(|y|^2, \frac{1}{2}t^{1/2}|y|^3 \Big), \quad \forall t > 0,\ x, y \in \mathbb{R}.$$

Hence

$$0 \leq U_t(x,y) \leq \|f\|_{C^3}\min\Big(|y|^2, \frac{1}{2}t^{1/2}|y|^3 \Big), \quad \forall t > 0,\ x, y \in \mathbb{R}$$

We deduce that

$$\Big\|\frac{1}{t}\big(V_\mu(t) - \mathbb{1}\big)f - \frac{1}{2}f''\Big\| \leq \|f\|_{C^3}\int_{\mathbb{R}}\min\Big(|y|^2, \frac{1}{2}t^{1/2}|y|^3 \Big)\mu\big[\,dy\,\big]$$

$$\leq \frac{1}{2}\|f\|_{C^3}\int_{|y|<R} t^{1/2}|y|^3 + \|f\|_{C^3}\int_{|y|>R}|y|^2\mu\big[\,dy\,\big]$$

$$\leq \underbrace{\frac{t^{1/2}}{2}R^3\|f\|_{C^3}}_{=:A(R,t)} + \underbrace{\|f\|_{C^3}\int_{|y|>R}|y|^2\mu\big[\,dy\,\big]}_{=:B(R)}.$$

Since

$$\int_{\mathbb{R}} y^2\mu\big[\,dy\,\big] < \infty$$

we deduce that for any $\varepsilon > 0$ there exists $R(\varepsilon) > 0$ such that $B\big(R(\varepsilon)\big) < \frac{\varepsilon}{2}$. Next, choose $\delta(\varepsilon) > 0$ such that $A\big(R(\varepsilon), \delta(\varepsilon)\big) < \frac{\varepsilon}{2}$. Then for $t < \delta(\varepsilon)$

$$\Big\|\frac{1}{t}\big(V_\mu(t) - \mathbb{1}\big)f - \frac{1}{2}f''\Big\| < \varepsilon.$$

□

We deduce from Chernoff's product formula that

$$\lim_{n\to\infty} V_\mu(t/n)^n f \to H_t f, \quad \forall f \in C_0(\mathbb{R}). \tag{2.16}$$

Let us observe that if $\mu, \nu \in \mathrm{Prob}(\mathbb{R})$, then

$$(\mu * \nu)_t = \mu_t * \nu_t \quad \forall t > 0. \tag{2.17}$$

Indeed, if X and Y are independent random variables with distributions μ and respectively ν, we have

$$\mathbb{P}_{t^{1/2}(X+Y)} = \mathbb{P}_{t^{1/2}X} * \mathbb{P}_{t^{1/2}Y}.$$

Suppose now that $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed (i.i.d.) random variables such that

$$\mathbb{E}\big[\, X_n \,\big] = 0, \quad \mathbb{E}\big[\, X_n^2 \,\big] = 1, \quad \forall n \in \mathbb{N}.$$

Denote by μ the common distribution of these random variables. Set

$$Z_n := \frac{1}{\sqrt{n}}\big(\, X_1 + \cdots + X_n \,\big).$$

Observe that

$$\mathbb{E}\big[\, Z_n \,\big] = 0, \quad \mathbb{E}\big[\, Z_n^2 \,\big] = 1, \quad \forall n \in \mathbb{N},$$

and

$$T_{t^{1/2}Z_n} = T_{(t/n)^{1/2}X_1} \cdots T_{(t/n)^{1/2}X_n} = \big(\, T_{\mu_{t/n}} \,\big)^n = V_\mu(t/n)^n \to H_t \quad \text{as } n \to \infty.$$

We have thus proved the celebrated *Central Limit Theorem*

Corollary 2.15. *Suppose that $(X_n)_{n\in\mathbb{N}}$ a sequence of i.i.d. random variables such that*

$$\mathbb{E}\big[\, X_n^2 \,\big] = 1, \quad \mathbb{E}\big[\, X_n \,\big] = 0, \quad \forall n.$$

then the random variables

$$\frac{1}{\sqrt{n}}\big(\, X_1 + \cdots + X_n \,\big)$$

converge in distribution to a standard normal random variable. □

Appendix A. A brief introduction to Bochner integral

We survey here a few facts about the Bochner integral. For proofs and more details we refer to [2, Sec.7.5], [3, III.6] or [10, V.5, V.6].

Throughout this appendix $(\Omega, \mathscr{S}, \mu)$ will denotes a measured space, X a real Banach space, and X^* the topological dual of X. We denote by by $\langle -, - \rangle$ the natural pairing

$$\langle -, - \rangle : X^* \times X \to \mathbb{R}, \quad X^* \times X \ni (\xi, x) \mapsto \langle \xi, x \rangle := \xi(x).$$

A function, $f : \Omega \to X$, is called *simple* or *elementary* if there exist $S_n \in \mathscr{S}$ and $x_1, \dots, x_n \in X$ such that

$$f(\omega) = \sum_{i=1}^n I_{S_i}(\omega) x_i, \quad \forall \omega \in \Omega,$$

where I_{S_i} denotes the indicator function of S_i. We denote by $\mathrm{Elem}(\Omega, \mathscr{S}, X)$ the space of elementary functions from $(\Omega, \mathscr{S}) \to X$.

Definition A.1. A function $f : \Omega \to X$ is said to be *strongly measurable* if there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathrm{Elem}(\Omega, \mathscr{S}, X)$ such that

$$f(\omega) = \lim_{n\to\infty} f_n(\omega), \quad \forall \omega \in \Omega.$$

□

Definition A.2 (Bochner integrability)**.** A strongly measurable function $f : \Omega \to X$ is called *Bochner integrable* or *strongly integrable* if the non-negative function $\|f\| : \Omega \to \mathbb{R}$ is integrable with respect to the measure μ. □

Definition A.3. For $g \in \mathrm{Elem}(\Omega, \mathscr{S}, X)$

$$g = \sum_{i=1}^{n} I_{S_i} x_i,$$

we define the Bochner integral of g over Ω as:

$$\int_\Omega g d\mu = \sum_{i=1}^{n} \mu[S_i] x_i.$$

Clearly g is Bochner integrable iff $\mu[I_{S_i}] < \infty$, $\forall i = 1, \dots, n$.

Lemma A.4. *Suppose that $f : \Omega \to X$ is a Bochner integrable function. Let $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}} \in$ Elem$(\Omega, \mathscr{S}, X)$ be sequences of Bochner integral elementary functions s.t. $f_n, g_n \to f$. Then the limits*

$$\lim_{n\to\infty} \int_\Omega f_n d\mu \text{ and } \lim_{n\to\infty} \int_\Omega g_n d\mu$$

exists and are equal. □

Definition A.5 (The Bochner Integral)**.** Suppose that $f : \Omega \to X$ is a Bochner integrable function. Let $(f_n)_{n\in\mathbb{N}} \in \mathrm{Elem}(\Omega, \mathscr{S}, X)$ be a sequence of Bochner integral elementary functions s.t. $f_n \to f$. Define the Bochner integral of f over Ω as

$$\int_\Omega f d\mu := \lim_{n\to\infty} \int_\Omega f_n d\mu$$

□

In the remainder of this sequence we will focus exclusively on the special case when $(\Omega, \mathscr{S}, \mu)$ is an interval $\subset \mathbb{R}$ equipped with the Borel sigma-algebra and the Lebesgue measure. Theorems [2, Thm.7.5.4, 7.5.6] imply the following result.

Proposition A.6. *Suppose that $f : I \to X$ is a continuous function. Then the following hold.*

(i) *$f : I \to X$ is strongly integrable iff $\|f\|$ is integrable.*
(ii) $\forall \xi \in X^*$

$$\left\langle \xi, \int_I f(t)dt \right\rangle = \int_Y \langle\, \xi, f(t) \,\rangle dt.$$

□

References

[1] D. Applebaum: *Semigroups of Linear Operators with Applications to Probability and Physics*, Cambridge University Press, 2019.

[2] T. Bühler, D. A. Salamon: *Functional Analysis*, Amer. Math. Soc., 2018.

[3] N. Dunford, J.T. Schwartz: *Linear Operators. Part I: General Theory*, John Wiley & Sons, 1957.

[4] K.-J. Engel, R. Nagel: *A Short Course on Operator Semigroups*, Springer Verlag, 2006.

[5] S. N. Ethier, T. G. Kurz: *Markov Processes. Charachterization and Convergence*, John Wiley & Sons, 2015.

[6] W. Feller: *An Introduction to Probability Theory and Its Applications*, vol.2, John Wiley& Sons, 1970.

[7] J. A. Goldstein: *Semigroups of Linear operators and Applications*, Oxford University Press, 1985

[8] L. I. Nicolaescu: *A Graduate Course in Probability*, World Scientific, 2022. https://www3.nd.edu/~lnicolae/Grad_Prob_web.pdf

[9] A. Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer Ver;ag, 1983

[10] K.Yosida: *Functional Analysis*, 6th Edition, Springer Verlag, 1980.