



Universiteit
Leiden
The Netherlands

Approach to Markov operators on spaces of measures by means of equicontinuity

Ziemlańska, M.A.

Citation

Ziemlańska, M. A. (2021, February 10). *Approach to Markov operators on spaces of measures by means of equicontinuity*. Retrieved from <https://hdl.handle.net/1887/3135034>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3135034>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3135034> holds various files of this Leiden University dissertation.

Author: Ziemiańska, M.A.

Title: Approach to Markov Operators on spaces of measures by means of equicontinuity

Issue Date: 2021-02-10

Chapter 1

Fundamental concepts and results

1.1 Measures as functionals

Let us consider a measurable space (S, Σ) . We will denote $S := (S, \Sigma)$. On S we consider the space $\mathcal{M}(S)$ of finite signed measures. A typical example of a signed measure is the difference of two probability measures. Every signed measure is a difference of two nonnegative measures. Hence, for every $\mu \in \mathcal{M}(S)$ we have the equality $\mu = \mu^+ + \mu^-$. The measures μ^+ and μ^- are called positive and negative part of μ respectively. Such decomposition is called the *Jordan* or *Jordan-Hahn* decomposition. Following [Bog07b], there exist S^+ and S^- such that for all $A \in \mathcal{A}$ one has $\mu(A \cap S^-) \leq 0$ and $\mu(A \cap S^+) \geq 0$. We define the *total variation norm* on $\mathcal{M}(S)$ by $\|\mu\|_{TV} := |\mu|(S) = \mu^+(S) + \mu^-(S) = \sup_{B \in \Sigma, B \subset S} \mu(B) - \inf_{B \in \Sigma, B \subset S} \mu(B)$. $\mathcal{M}(S)$ endowed with $\|\cdot\|_{TV}$ is a Banach lattice. However, the topology given by $\|\cdot\|_{TV}$ norm is often too strong for applications. Let us show this in the following example, following [Wor10].

Example 1.1.1. Let $\Phi_t : S \rightarrow S$ be a family of measurable maps, $t \in \mathbb{R}_+$ such that $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_0 = \text{Id}_S$. Each Φ_t induces a linear operator $T_\Phi(t)$ on $\mathcal{M}(S)$ by

$$T_\Phi(t)\mu := \mu \circ \Phi_t^{-1}.$$

We get the following properties of the family $\mathbb{T} := (T_\Phi(t))_{t \geq 0}$:

- (+) $\mathbb{T}$ leaves the cone $\mathcal{M}^+(S)$ invariant
- (+) $T_\Phi(t)\delta_x = \delta_{\Phi_t(x)}$
- (-) $\mathbb{T}$ is strongly continuous with respect to $\|\cdot\|_{TV}$ only if it is constant, as $\|\delta_x - \delta_y\|_{TV} = 2$ whenever $x \neq y$.
- (-) In general $t \mapsto T_\Phi(t)\delta_t = \delta_{\Phi_t(x)}$ will not be strongly measurable as its range will not be separable. This makes $(\mathcal{M}(S), \|\cdot\|_{TV})$ not suitable to study a variation of constants formula

$$\mu_t = T_\Phi(t)\mu_0 + \int_0^t T_\Phi(t-s)F(\mu_s)ds,$$

as the integral is hard to interpret.

Throughout the thesis we will assume that S is a Polish space. Hence, S is separable and completely metrizable. Any metric d that metrizes the topology of S such that (S, d) is separable and complete is called *admissible*. We will denote by $\mathcal{D}(S)$ the family of all admissible metrics on S . We will consider $\mathcal{C}_b(S)$, the Banach space of continuous bounded

functions on S , with the supremum norm

$$\|f\|_\infty = \sup\{|f(x)| : x \in S\}.$$

Definition 1.1.2. *A function $f : S \rightarrow \mathbb{R}$ is (globally) Lipschitz if there exists $L \geq 0$ such that*

$$|f(x) - f(y)| \leq Ld(x, y) \quad \text{for all } x, y \in S. \quad (1.1)$$

Let $\text{Lip}(S, d)$ (or $\text{Lip}(S)$ for shorter notation) denote the vector space of Lipschitz functions on (S, d) . The *Lipschitz constant* of $f \in \text{Lip}(S)$ on (S, d) is

$$|f|_L := \sup \left\{ \frac{|f(x) - f(y)|}{d(x, y)} : x \neq y, x, y \in S \right\},$$

which is the best (i.e. smallest) constant L that can be used in (1.1). Following Dudley [Dud66], then $\text{BL}(S, d)$ will denote the Banach space of all bounded Lipschitz functions f on S with the bounded Lipschitz or Dudley norm

$$\|f\|_{BL, d} = |f|_L + \|f\|_\infty.$$

Proposition 1.1.3 ([Wor10], Proposition 2.2.7). *$\text{BL}(S, d)$ is complete with respect to $\|\cdot\|_{BL, d}$.*

We will denote $\|\cdot\|_{BL, d}$ by $\|\cdot\|_{BL}$ if no ambiguity occurs.

We can equip the space $\mathcal{M}(S)$ with different equivalent norms. Zaharopol [Zah00], Lasota and Szarek [LS06], Lasota and Yorke [LY94] use the Fortet-Mourier norm of the form

$$\|\mu\|_{FM}^* = \sup \left\{ \int_S f d\mu : f \in \text{BL}(S, d), \|f\|_{FM} = \max(\|f\|_\infty, |f|_L) \leq 1 \right\}.$$

The name Fortet-Mourier can be misleading though, as in the original paper Fortet and Mourier [[FM53], p.277] construct the bounded Lipschitz/Dudley norm $\|\cdot\|_{BL, d}$, not the Fortet-Mourier norm.

The norm $\|\cdot\|_{FM}$ is equivalent to $\|\cdot\|_{BL, d}$ and to all the norms of the form $\|f\|_{BL(S, d, p)} := (|f|^p + \|f\|_\infty^p)^{\frac{1}{p}}$, $1 < p < \infty$.

The space $\mathcal{M}(S)$ embeds into $\text{BL}(S)^*$ by means of integration $\mu \mapsto I_\mu$, where

$$I_\mu(f) = \langle \mu, f \rangle := \int_S f d\mu.$$

Each element in $\mathcal{M}(S)$ defines an element of the dual space $\text{BL}(S, d)^*$ with the norm

$$\|\mu\|_{BL, d}^* = \sup\{\langle \mu, f \rangle : f \in \text{BL}(S, d), \|f\|_{BL, d} \leq 1\}.$$

Let us recall few useful facts about the space $(\mathcal{M}(S), \|\cdot\|_{BL, d}^*)$.

Lemma 1.1.4 ([Wor10], Lemma 2.3.6). *For every $x \in S$, δ_x is in $(\mathcal{M}(S), \|\cdot\|_{BL, d}^*)$ and $\|\delta_x\|_{BL, d}^* = 1$. Moreover, for $x, y \in S$,*

$$\|\delta_x - \delta_y\|_{BL, d}^* = \frac{2d(x, y)}{2 + d(x, y)} \leq \min(2, d(x, y)).$$

By $\mathcal{M}^+(S)$ we denote the convex cone of positive measures in $\mathcal{M}(S)$. One has

$$\|\mu\|_{TV} = \|\mu\|_{BL}^* = \|\mu\|_{FM}^* \text{ for all } \mu \in \mathcal{M}^+(S).$$

In general, for $\mu \in \mathcal{M}(S)$, $\|\mu\|_{BL}^* \leq \|\mu\|_{FM}^* \leq \|\mu\|_{TV}$.

1.1.1 Some topologies on spaces of maps

Let us outline the main topologies we are interested in. To describe the topologies consider topological spaces X and Y and a collection $\mathcal{F}$ of maps $f : X \rightarrow Y$. Let us show different ways to provide $\mathcal{F}$ with a topology.

- **Topology of pointwise convergence** [[Kel55], p.88] on $\mathcal{F}$:

The topology of pointwise convergence is of importance as this is the smallest topology for $\mathcal{F}$ for which each point evolution $\delta_x, x \in X$ is continuous on $\mathcal{F}$, see [Kel55] p.217. A net of function $(f_\alpha)_{\alpha \in A}$ converges to f if and only if $f_\alpha(x) \rightarrow f(x)$ for each $x \in X$. Note that the topology of X does not play a role in the results on the topology of pointwise convergence on $\mathcal{F}$.

- **Compact open topology** on $\mathcal{F}$:

The other topology of interest, which does depend on the topology of X , is the compact-open topology. Let $\mathcal{F}$ be a collection of continuous maps $f : X \rightarrow Y$. Thus, for fixed f the map $X \rightarrow Y : x \mapsto f(x)$ is continuous. We look for a topology on $\mathcal{F}$ such that the map $\mathcal{F} \times X \rightarrow Y : (f, x) \mapsto f(x)$ is (jointly) continuous. Here the compact open topology plays a role. Let us define

$$\mathcal{F}(A, B) := \{f \in \mathcal{F} | f[A] \subset B\}$$

for $A \subset X, B \subset Y$.

The sets $\mathcal{F}(K, U)$ such that $K \subset X$ compact and $U \subset Y$ open are a subbase for the compact open topology. For more details see [Kel55]. In the main part of this thesis- concerning Markov operators- equicontinuous functions of maps play a central role.

Equicontinuous families of maps

Let X be a topological space and (S, d) a metric space. Let $\mathcal{F}$ be a family of maps $f : X \rightarrow S$.

Definition 1.1.5. *The family $\mathcal{F}$ of functions $f : X \rightarrow S$ is equicontinuous at $x \in X$ if for every $\epsilon > 0$ there exists an open neighbourhood U_ϵ of x such that*

$$d(f(x), f(y)) < \epsilon \text{ for } y \in U_\epsilon, f \in \mathcal{F}.$$

Family $\mathcal{F}$ is equicontinuous if it is equicontinuous at every point of X .

Let us now recall Theorem 15, p.232 from [Kel55].

Theorem 1.1.6. *If $\mathcal{F}$ is an equicontinuous family, then the topology of pointwise convergence is jointly continuous. Therefore it coincides with the compact open topology.*

1.1.2 Tight sets of measures

A finite signed Borel measure μ is called *tight* (see eg. Dudley [Dud66]) if for every $\epsilon > 0$ there exists a compact set $K \subset S$ such that $|\mu|(S \setminus K) < \epsilon$. The class of all tight measures is denoted by $\mathcal{M}_t(S)$.

A family $M \subset \mathcal{M}(S)$ is *uniformly tight* (Abbrev. *tight*) if for every $\epsilon > 0$ there exists a compact set $K \subset S$ such that $|\mu|(S \setminus K) < \epsilon$ for all $\mu \in M$.

A sequence of measures $(\mu_n)_{n \in \mathbb{N}} \subset \mathcal{M}(S)$ is *weakly convergent* to a measure $\mu \in \mathcal{M}(S)$ if for every bounded continuous real function f on S one has

$$\lim_{n \rightarrow \infty} \langle \mu_n, f \rangle = \langle \mu, f \rangle.$$

A frequent problem is the following. Can one select a weakly convergent subsequence (hence in the weak topology $\sigma(\mathcal{M}, C_b(X))$) in a given sequence? It turns out that the problem can be reduced to analyzing the uniform tightness of the family of measures.

Hence (uniform) tightness of measures is a key to understanding the weak convergence of sequences of measures.

Theorem 1.1.7. *[Prokhorov Theorem, [Bog07a] Theorem 8.6.2] Let X be a complete metric space and let $\mathcal{M}$ be a family of Borel measures on X . Then the following conditions are equivalent:*

- (i) *every sequence $\{\mu_n\} \subset \mathcal{M}$ contains a weakly convergent subsequence;*
- (ii) *the family $\mathcal{M}$ is uniformly tight and uniformly bounded in the total variation norm.*

The above conditions are equivalent for any complete metric space X if $\mathcal{M} \subset \mathcal{M}_t(S)$.

Tightness of sets of measures is a tool used in analyzing the existence of invariant measures for Markov operators, e.g. by Szarek in [Sza03]. By Proposition 5.1 in [Sza03] we get that a continuous (in a weak topology) Markov operator which is tight admits an invariant distribution.

1.2 Markov operators on spaces of measures and semigroups of Markov operators

Markov operators occur in diverse branches of pure and applied mathematics. They are used in studying dynamical systems and dynamical systems with stochastic perturbations. Semigroups of Markov operators are generated by e.g. stochastic differential equations or deterministic partial differential equations. Transport equations, which are generating Markov semigroups, appear in the theory of population dynamics [Hei86, Rud00, Rud97]. Such processes were also extensively studied in close connection to fractals and semifractals [BD85, BEH89, DF99, LM94, LM00].

Markov operator P is defined as a map $P : \mathcal{M}^+(S) \rightarrow \mathcal{M}^+(S)$ such that

- (i) P is additive and $\mathbb{R}_+$ homogenous: $P(\lambda_1\mu_1 + \lambda_2\mu_2) = \lambda_1P\mu_1 + \lambda_2P\mu_2$; for $\mu_i \in \mathcal{M}^+(S)$, $\lambda_i \geq 0$.
- (ii) $\|P\mu\|_{TV} = \|\mu\|_{TV}$ for all $\mu \in \mathcal{M}^+(S)$.

Every Markov operator can be extended to the space of all signed measures. Namely, for every $\mu \in \mathcal{M}(S)$, we get a decomposition $\mu = \mu_1 - \mu_2$, where $\mu_1, \mu_2 \in \mathcal{M}^+(S)$. We set $P\mu = P\mu_1 - P\mu_2$.

The decomposition is not unique, but different decompositions vary by a positive measure,

such that $P\mu$ does not depend on the decomposition chosen.

A linear operator $U : \text{BM}(S) \rightarrow \text{BM}(S)$ is called dual to P if

$$\langle P\mu, f \rangle = \langle \mu, Uf \rangle \text{ for all } \mu \in \mathcal{M}^+(S), f \in \text{BM}(S).$$

If such an operator U exists, it is unique and we call the Markov operator P *regular*. U is positive and satisfies $U\mathbb{1} = \mathbb{1}$. The Markov operator P is a *Markov-Feller operator* if it is regular and the dual operator U maps the space of continuous bounded functions $C_b(S)$ into itself. A Markov semigroup $(P_t)_{t \geq 0}$ on S is a semigroup of Markov operators on $\mathcal{M}^+(S)$. The semigroup property entails that $P_t \circ P_s = P_{t+s}$ and $P_0 = \text{Id}$. Markov semigroup is *regular* (or *Feller*) if all operators P_t are regular (or Feller). Then $(U_t)_{t \geq 0}$ is a semigroup on $\text{BM}(S)$ which is called a *dual semigroup*.

1.3 Convergence of sequences of measures

Dudley analyzed the relation between- in our terminology- weak and $\|\cdot\|_{BL}^*$ convergence of measures. In Theorems 6, 7 and 11 [Dud66], Dudley showed the following for pseudo-metric spaces slightly adapted to our terminology. Any metric space is a pseudo-metric space.

Theorem 1.3.1. *Let (S, d) be a pseudo-metric space, $\mu_n \in \mathcal{M}_s(S)$: Then:*

- (i) *if $\mu_n \rightarrow \mu$ weakly, then $\|\mu_n - \mu\|_{BL}^* \rightarrow 0$;*
- (ii) *$\|\mu_n - \mu\|_{BL}^* \rightarrow 0$ implies $\mu_n \rightarrow \mu$ weakly for any sequence in $\mathcal{M}_s(S)$ if and only if S is uniformly discrete;*
- (iii) *if S is a topological space, $\mu_n \in \mathcal{M}_s^+(S)$ and μ_n converges to μ weakly, then $\mu_n \rightarrow \mu$ uniformly on any equicontinuous and uniformly bounded class of functions on S ;*
- (iv) *if (S, d) is a metric space, $\mu_n, \mu \in \mathcal{M}^+(S)$ and $\|\mu_n - \mu\|_{BL}^* \rightarrow 0$ as $n \rightarrow \infty$, then $\mu_n \rightarrow \mu$ weakly.*

We provide conditions on subsets of (signed) measures $\mathcal{M}(S)$ such that the weak topology on $\mathcal{M}(S)$ coincides with the norm topology defined by the dual bounded Lipschitz norm $\|\cdot\|_{BL}$ or by Fortet-Mourier norm $\|\cdot\|_{FM}^*$, see Theorem 2.3.5, Theorem 2.3.7 and similar results in Section 2.3.2. These build on Theorem 2.3.1 which states that for a bounded (in total variation norm) sequence of signed measures (μ_n) that converges weakly, that is $\langle \mu_n, f \rangle$ is convergent for any $f \in \text{BL}(S, d)$, $(\mu_n)_n$ converges in $\|\cdot\|_{BL, d}$ norm.

This is not precisely the *Schur-property* for $(\mathcal{M}(S), \|\cdot\|_{BL}^*)$, but can be considered a *Schur-like* property, further discussed in Chapter 2.

Let us recall definition (Definition 2.3.4., [NJKAKK06]) of the *Schur property*.

Definition 1.3.2. *A Banach space X has the Schur property (or X is a Schur space) if weak and norm sequential convergence coincide in X , i.e. a sequence $(x_n)_n$ in X converges to 0 weakly if and only if $(x_n)_n$ converges to 0 in norm.*

By the following example (for details see Example 2.5.4, Megginson [MBA98]) we can see that the space l_2 does not have the Schur property. In general, none of the spaces l_p , $1 < p < \infty$ has the Schur's property.

Example 1.3.3. *Let (e_n) be the sequence of unit vectors in l_2 . Then $x^*e_n \rightarrow 0$ for each x^* in l^* , and so the sequence (e_n) converges to 0 with respect to the weak topology. Since $\|e_n\| = 1$ for each n , the sequence (e_n) cannot converge to 0 with respect to the norm topology. The norm and weak topologies of l_2 are therefore different, so it is possible for the weak topology of a normed space to be a proper subtopology of the norm topology.*

1.4 Lie-Trotter product formula

Chapter 3 of this thesis, the Lie-Trotter product formula for Markov operators, was motivated by the need to deal with more and more complicated models of physical phenomena. Citing [HKLR10] "A strategy to deal with complicated problems is to **“divide and conquer”**. (In the context of equations of evolution type) a rather successful approach in this spirit has been operator splitting." Let us show the simplest example of an operator splitting scheme (based on [HKLR10]). We want to solve the Cauchy problem

$$\frac{dU}{dt} + \mathcal{A}(U) = 0, \quad U(0) = U_0,$$

for an operator $\mathcal{A}$. Formally we get the solution of the form

$$U(t) = e^{t\mathcal{A}}U_0.$$

Though, here the information about the operator $\mathcal{A}$ is needed. If $\mathcal{A}$ is of some "complicated" form, we need to find a way to be able to compute this solution in an optimal way. Assuming we can write $\mathcal{A}$ as a sum $\mathcal{A}_1 + \mathcal{A}_2$ and solve problems

$$\frac{dU}{dt} + \mathcal{A}_1(U) = 0, \quad U(0) = U_0$$

and

$$\frac{dU}{dt} + \mathcal{A}_2(U) = 0, \quad U(0) = U_0$$

separately with solutions

$$U(t) = e^{t\mathcal{A}_1}U_0$$

and

$$U(t) = e^{t\mathcal{A}_2}U_0,$$

we get an operator splitting of the simplest form:

$$U(t_{n+1}) \approx e^{-\Delta t \mathcal{A}_2} e^{-\Delta t \mathcal{A}_1} U(t_n),$$

where $t_n := n\Delta t$.

If $\mathcal{A}_1$ and $\mathcal{A}_2$ would commute, we get $e^{-t\mathcal{A}_1}e^{-t\mathcal{A}_2} = e^{-t\mathcal{A}}$. Hence, the method is exact. For noncommuting operators we get the Lie-Trotter (or Lie-Trotter-Kato) formula of the form

$$U(t) = e^{t\mathcal{A}}U_0 = \lim_{\Delta t \rightarrow 0, t=n\Delta t} (e^{-\Delta t \mathcal{A}_2} e^{-\Delta t \mathcal{A}_1})^n U_0.$$

The questions which one wants to answer is if the above limit exists and, if yes, does it give the solution of an original problem. If the answers are positive, one can use the approximating scheme to analyze the more difficult original problem. Various conditions for convergence are stated and discussed in Chapter 3.

Hence, we see that operator splitting schemes can be a way to go when analyzing complicated models. Let us now show why we are interested in product formula for Markov operators. One way to construct a new dynamical system from a known one is by perturbing the original problem. One of the examples of such constructions is an *iterated function system* (IFS), which is analyzed in the theory of fractals [Bar88, BDEG88, LM94, MS03]. An IFS is an example of stochastic switching at fixed times between deterministic flows. An IFS $\{(w_i, p_i); i \in I\}$ with probabilities is given by a family of continuous functions $w_i : S \rightarrow S, i \in I$, where (S, d) is a complete separable metric space with a family of continuous functions $p_i : S \rightarrow [0, 1], i \in I$ s.t. $\sum_{i \in I} p_i(x) = 1$. Such IFS defines a Markov operator P acting on measures by

$$P\mu(A) = \sum_{i \in I} \int_X \mathbb{1}(w_i(x)) p_i(x) \mu(dx) \text{ for } \mu \in \mathcal{M}(S), A \in \mathcal{B}.$$

Such a Markov operator is also Feller, hence it seems natural to consider Markov-Feller operators.

The next example which motivates analyzing switching schemes for Markov semigroup, is piecewise-deterministic Markov processes (PDMPs) where deterministic motion is punctuated by random jumps occurring according to a suitable distribution. PDMPs have wide applications e.g. to gene expression in the work of Hille, Horbacz and Szarek [HHS16] and Mackey, Tyran-Kaminska and Yvines [MTKY13]. The analysis of such processes is concentrated mostly on their long time behaviour. By analyzing Lie-Trotter formula in such a setting we may be able to extend the analysis of piecewise-deterministic Markov processes to switching between deterministic and stochastic models.

Intriguing example

Let us now consider an example of a convergent Lie-Trotter product formula for a right translation semigroup and a multiplication semigroup for which no assumption on generators is made. This example is originally from Goldstein [Gol85], p.56, without detailed proof though.

Let $X := L^1(\mathbb{R})$ (complex-valued) and let us consider the right translation semigroup $(S(t))_{t \geq 0}$ and a multiplication semigroup $(T(t))_{t \geq 0}$ generated by $B = M_{iq}$ for $q : \mathbb{R} \rightarrow \mathbb{R}$ a measurable and locally integrable function, where $M_{iq}f := iq \cdot f$ (on f in a smaller domain):

$$T(t)f(x) := e^{itq(x)}f(x)$$

$$S(t)f(x) := f(x+t)$$

Further, define

$$U(t)f(x) := e^{(i \int_0^t q(x+s)ds)}f(x+t).$$

For $f \in X$ we can compute products

$$\left[T\left(\frac{t}{n}\right) S\left(\frac{t}{n}\right) \right]^n f(x) = \exp\left(i \sum_{k=0}^{n-1} q(x+kt/n)t/n\right) \cdot f(x+t), \quad \text{for } t \geq 0, x \in \mathbb{R}.$$

We want to show that the product converges in Lebesgue-measure to $U(t)f(x)$.

First let us now proof the following lemma:

Lemma 1.4.1. *For every $g \in L^1_{loc}(\mathbb{R})$, $t > 0$,*

$$\sum_{j=0}^{n-1} g\left(\cdot + \frac{tj}{n}\right) \frac{t}{n} \rightarrow \int_0^t g(\cdot + s)ds, \quad \text{as } n \rightarrow \infty$$

in Lebesgue-measure, on every compact interval I .

Proof. Let μ be a Lebesgue measure on $\mathbb{R}$. We want to show that for all $\eta > 0$:

$$\lim_{n \rightarrow \infty} \mu \left(x \in I : \left| \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) \frac{t}{n} - \int_0^t g(x+s) ds \right| \geq \eta \right) = 0.$$

By Chebyshev inequality (see Bogachev [Bog07b], Theorem 2.5.3.) we get

$$\mu \left(x \in I : \left| \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) \frac{t}{n} - \int_0^t g(x+s) ds \right| \geq \eta \right) \leq \frac{1}{\eta} \int_I \left| \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) \frac{t}{n} - \int_0^t g(x+s) ds \right| dx$$

Also

$$\begin{aligned} & \int_I \left| \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) \frac{t}{n} - \int_0^t g(x+s) ds \right| dx = \int_I \left| \frac{t}{n} \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) - \int_0^t g(x+s) ds \right| dx = \\ & = \int_I \left| \frac{t}{n} \sum_{j=0}^{n-1} \left[g \left(x + \frac{tj}{n} \right) - \frac{n}{t} \int_{\frac{t(j-1)}{n}}^{\frac{tj}{n}} g(x+s) ds \right] \right| dx = \\ & = \int_I \left| \frac{t}{n} \sum_{j=0}^{n-1} \frac{n}{t} \int_{\frac{t(j-1)}{n}}^{\frac{tj}{n}} \left[g \left(x + \frac{tj}{n} \right) - g(x+s) \right] ds \right| dx \leq \\ & \leq \sum_{j=0}^{n-1} \int_I \int_{\frac{t(j-1)}{n}}^{\frac{tj}{n}} |g \left(x + \frac{tj}{n} \right) - g(x+s)| ds dx \end{aligned}$$

Let $\varepsilon > 0$. Take $t_0 \in \mathbb{R}$, $\delta > 0$ and let $\hat{I} := \bigcup_{0 \leq s \leq t} I + s$. Then $\hat{I}$ is compact with a nonempty interior. There exists $h \in C_c(\hat{I})$ such that

$$\int_{\hat{I}} |g(x) - h(x)| dx < \frac{\varepsilon}{3}.$$

Then

$$\int_I |g(x+s) - h(x+s)| dx \leq \int_{\hat{I}} |g(y) - h(y)| dy < \frac{\varepsilon}{3}, \text{ for all } 0 \leq s \leq t$$

and for s_0 and s in $[0, t]$ sufficiently close,

$$\begin{aligned} \int_I |g(x+s_0) - g(x+s)| dx & \leq \int_I |g(x+s_0) - h(x+s_0)| dx + \int_I |h(x+s_0) - h(x+s)| dx + \\ & \int_I |h(x+s) - g(x+s)| dx \leq \varepsilon \end{aligned}$$

as h is uniformly continuous on $\hat{I}$. So for s sufficiently close to s_0 in $[0, t]$, $\int_I |g(x+s_0) - g(x+s)| dx$ can be made arbitrarily small. Using the above estimation we get for n sufficiently large that:

$$\begin{aligned} & \mu \left(x \in I : \left| \sum_{j=0}^{n-1} g \left(x + \frac{tj}{n} \right) \frac{t}{n} - \int_0^t g(x+s) ds \right| \geq \eta \right) \leq \\ & \leq \frac{1}{\eta} \sum_{j=0}^{n-1} \int_I \int_{\frac{t(j-1)}{n}}^{\frac{tj}{n}} |g \left(x + \frac{tj}{n} \right) - g(x+s)| ds dx < \\ & < \frac{1}{\eta} \sum_{j=0}^{n-1} \int_{\frac{t(j-1)}{n}}^{\frac{tj}{n}} \int_I \varepsilon ds = \frac{\varepsilon}{\eta} \sum_{j=0}^{n-1} \frac{t}{n} < \varepsilon \frac{t}{\eta} \end{aligned}$$

So indeed, for every $g \in L^1_{loc}(\mathbb{R})$ we get the convergence in measure. $\square$

Now let us apply the continuous map $\exp(i\cdot) : \mathbb{R} \rightarrow \mathbb{C}$. According to [Bog07b], Corollary 2.2.6, we get convergence in measure on a compact interval I of

$$E_n := \exp\left(i \sum_{j=1}^n g\left(\cdot + \frac{tj}{n}\right) \frac{t}{n}\right) \rightarrow \exp\left(i \int_0^t g(\cdot + s) ds\right) =: E.$$

For $f \in C_c(\mathbb{R})$ and again using Corollary 2.2.6 (Bogachev [Bog07b], p.113) we get that $E_n f \rightarrow E f$ in measure. Since $|E_n f| = |f| \in L^1$, by Dominated Convergence Theorem (cf. [Bog07a], Theorem 2.8.5, p.132) we get convergence in L^1 -norm, so

$$\|E_n f - E f\|_{L^1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

As $C_c(\mathbb{R}) \subset L^1(\mathbb{R})$ is $\|\cdot\|_{L^1}$ -dense, for $f \in L^1(\mathbb{R})$ we can find $f_0 \in C_c(\mathbb{R})$ such that

$$\|f - f_0\|_{L^1} < \varepsilon.$$

Then

$$\begin{aligned} \|E_n f - E f\|_{L^1} &\leq \|E_n f - E_n f_0\|_{L^1} + \|E_n f_0 - E f_0\|_{L^1} + \|E f_0 - E_n f_0\|_{L^1} \leq \\ &\leq \|f - f_0\|_{L^1} + \|E_n f_0 - E f_0\|_{L^1} + \|f - f_0\|_{L^1} \leq 3\varepsilon \end{aligned}$$

for a sufficiently large n .

So we get in $L^1(\mathbb{R})$ that

$$\lim_{n \rightarrow \infty} \left[T\left(\frac{t}{n}\right) S\left(\frac{t}{n}\right) \right]^n f = U(t)f.$$

The intriguing part of this example is the fact, that the Lie-Trotter product formula holds for T and S , but these semigroups do not satisfy common conditions for convergence of Lie-Trotter schemes (see Chapter 3). By Theorem 8.12 in [Gol85] if A is the generator of S and B is a generator of T then, if $\overline{A+B}$ is a generator, then it is a generator of U . However, it is possible that $\overline{A+B}$ need not be a generator; in fact, it can even happen that $\mathcal{D}(A+B) = \{0\}$.