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

Ziemlańska, M.A.

Author: Ziemlańska, M.A.
Title: Approach to Markov Operators on spaces of measures by means of equicontinuity
Issue Date: 2021-02-10

Approach to Markov Operators on Spaces of Measures by Means of Equicontinuity

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof.dr.ir. H. Bijl,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 10 februari 2021
klokke 11:15 uur

door

Maria Aleksandra Ziemlańska

geboren te Gdańsk, Polen
in 1985

Samenstelling van de promotiecommissie:

Promotor: Prof. dr. A. Doelman

Copromotor: Dr. S.C. Hille

Promotiecommissie:

Prof.dr. E.R. Eliel

Prof.dr. W.T.F. den Hollander

Dr.ir. O.W. van Gaans

Prof.dr. J.M.A.M. van Neerven, TU Delft

Prof. dr K. Horbacz, Silesian University in Katowice

Approach to Markov Operators on Spaces of Measures
by Means of Equicontinuity

Email: maja.ziemlanska@gmail.com

Printed in the Netherlands

To Jos and Julian

Contents

Notation

Here we state some conventions regarding mathematical notation that we will use throughout the thesis.

- $\mathbb{N}$ denotes the set of natural numbers$\{1, 2, 3, \cdots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$
- $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$
- $\mathcal{M}(S)$ is the real vector space of finite signed measures on S
- $\mathcal{M}^+(S)$ is the cone of positive measures in $\mathcal{M}(S)$
- $\mathcal{P}(S)$ is the set of probability measures in $\mathcal{M}^+(S)$
- $\|\cdot\|_T V$ denotes the total variation norm on $\mathcal{M}(S)$. $\|\mu\|_{TV} = \mu^+(S) + \mu^-(S)$
- $\mathbb{1}_E$ is the indicator function of $E \subset S$
- For a measurable function $f : S \to \mathbb{R}$ and $\mu \in \mathcal{M}(S)$ we denote

$$\langle \mu, f \rangle = \int_S f d\mu$$

- $P : \mathcal{M}(S) \to \mathcal{M}(S)$ denotes Markov operator with a dual operator U
- $B(x, r)$ denotes the open ball of radius r centered at x
- In a metric space (S, d), if $A \subset S$ is nonempty, we denote by $A^\epsilon := \{x \in S : d(x, A) \leq \epsilon\}$ the closed ϵ-neighbourhood of A
- If S is a topological space, $C_b(S)$ is the Banach space of bounded continuous functions from S to $\mathbb{R}$, endowed with the supremum norm $\|\cdot\|_\infty$.
- $\langle \mu, f \rangle := \int_\Omega f d\mu$
- Markov operator is a map $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ such that:

(MO1) P is additive and $\mathbb{R}_+$ homogeneous;

(MO2) $\|P\mu\|_{TV} = \|\mu\|_{TV}$ for all $\mu \in \mathcal{M}^+(S)$;

P extends to a positive bounded linear operator on $(\mathcal{M}(S), \|\cdot\|_{TV})$ by $P\mu := P\mu^+ - P\mu^-$.

- We say that Markov process is stationary if its moments do not depend on the time shift.

Motivation

The subject of this thesis, *'Approach to Markov Operators on Spaces of Measures by Means of Equicontinuity'*, combines an analytical and probabilistic approach to Markov operators. The combination of both has yielded various novel results whose proofs are facilitated by the use of analytical concepts like equicontinuity, measures of non-compactness and attractors and probabilistic arguments.

Markov operators come naturally from Markov processes, hence stochastic processes whose future values are determined by most recent values, without the necessity to take into account the past.

We intentionally work with Markov operators on spaces of finite signed Borel measures on the underlying Polish state space. Other researchers have looked at the setting of such operators on continuous bounded functions or subspaces thereof (the 'dual picture' from our perspective) or spaces of integrable functions with respect to an invariant measure.

We start by motivating why we think the space of measures as a state space is a more suitable setting then the spaces of integrable functions.

Why work with measures as a state space?

There are two main reasons for working with a space of measures as a state space. Let us start with the approach coming from deterministic systems.

A deterministic perspective

Let us show examples of deterministic models with randomness (in their initial conditions, random interventions) which motivate us to use a space of measures as a state space. A first example is that of sustainable harvesting. In Example 0.0.1 we introduce a fishery model with randomness in the size of a catch. The same idea can be extended to other types of harvesting, i.e. crop harvesting, where random interventions could be weather conditions such as the amount of rainfall. We first describe the setting. The model and use therein of a measure formulation is discussed afterwards.

Example 0.0.1. *[Sustainable harvesting, [AHvG13, AHG12], Figure 0.0.1]*

One of the problems of fisheries is determining the quota: That is, the amount of fish which can be caught without extinction of the fish species. Fish population is not distributed homogeneously. Hence, the size of an intervention- the size of a single catch- can be considered random within certain limits, as may be the time between successive harvesting events. Between interventions the growth of the fish population may be modelled deterministically. The main purpose of sustainable harvesting is to catch as much as possible, without causing the extinction of the population with high probability.

Example 0.0.2 is another type of real life application. In this case one wants to determine the amount of medicine, antibiotics in this case, necessary and sufficient to cure an illness. The same idea can be applied to a broader class of medical treatments, but also to the optimal use of pesticides, water, the use of artificial light in greenhouses etc.

Example 0.0.2. *[Antibiotic treatment, Figure 0.0.2]*

Another example of a deterministic process with random interventions is the antibiotic treatment of bacterial infections. A common way of treating such infections is by giving doses of antibiotics in the form of injections or orally at certain moments in time. These medicines either kill the bacteria or prevent them from reproducing. We assume randomness in the amount of bacteria that are killed or influenced by a single dose of antibiotics. In the time between doses the number of bacteria will increase. The growth of the colony may be modelled deterministically. The main question is how to determine the right dose of antibiotics such that the bacterial population goes extinct - almost surely. Too small a

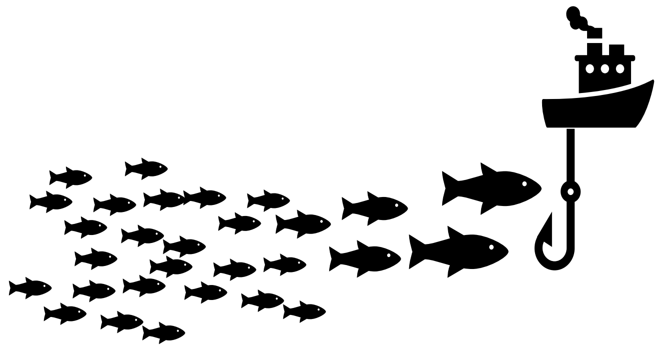

Figure 1: ***Sustainable harvesting**: A marine ecosystem from which fish are harvested. The interventions will interfere with the further growth of the population. The main question is to quantify the impact of this sampling process on the population. Intensive fishery reduces the fish population drastically. The catch size may be considered random (with certain limits) as the fish population is not homogeneously distributed.*

dose will not treat the illness and too big a dose can cause unwanted side effects to the patient.

Main question

> The main question in both examples is how to decide on the (maximal) size of interventions and the time intervals between them so that we get to the required results?

We shall now show how the above real life examples can be formalized in a mathematical model in the language of measures.

Mathematical description

Mathematically the above processes can be modelled as follows. The dynamics of population growth can be modelled deterministically when numbers of individuals are sufficiently large (e.g. bacteria colonies grow between antibiotic doses; fish populations grow between fishing periods). Abstractly, this can be formalized using a deterministic *dynamical system*: we have a state space S (nonempty). An element of the state space characterizes the state of the population eg. the number of fish or bacteria in the population, or their spatial distribution. Let

$$\phi_t : x_0 \mapsto \phi_t(x_0)$$

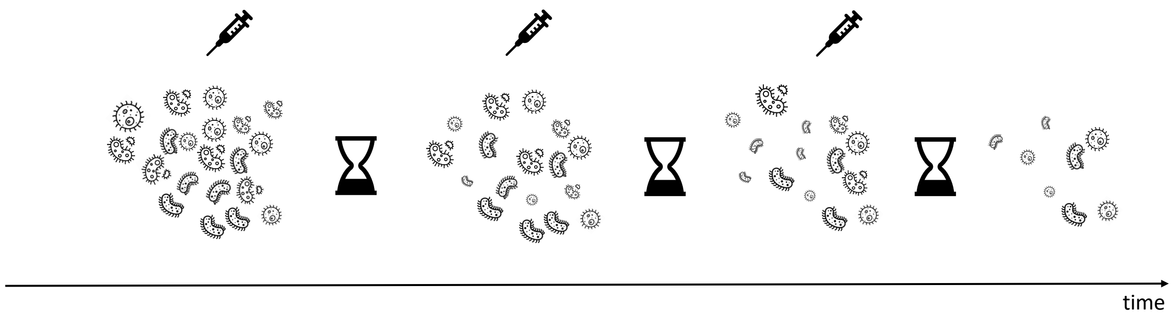

Figure 2: ***Antibiotic treatment***. *The question is how to determine the right dose of medicine.*

be the deterministic law that prescribes the state of the system at time t after it was in state x_0. The family of *flow maps* $(\phi_t)_{t\geq 0}$ has a *semigroup property*, i.e. for all $t, s \geq 0$ and $x_0 \in S$

$$\phi_t(\phi_s(x_0)) = \phi_{t+s}(x_0); \quad \phi_0(x_0) = x_0. \quad (*)$$

At discrete points in time we have random interventions (e.g. the impact of a dose of antibiotics in the population of bacteria or the size of a catch in a net). The position of the system in state space immediately after the intervention is given by a probability law which depends on the state of the system just before the intervention. Examples of such models can be found in [LM99] and [HHS16]. In such models one way of analysis is as follows.

The evolution of the system between interventions is given by the deterministic system $(\phi_t)_{t\geq 0}$ on S, where S is a Polish space. The population size just before the intervention will be $x' = \phi_{\Delta t}(x)$, where $x \in S$ is the state of the population just after the previous jump and Δt is the time between two interventions. For simplicity sake we can assume that Δt is fixed, non-random. At each point $x' \in S$ one has a probability distribution $Q_{x'}$ on S. $Q_{x'}(E)$ is the probability that the system state will be in $E \subset S$ immediately after an intervention, when the state just before an intervention was x'.

If μ is the probability distribution for the state of the system immediately after an intervention (or at $t = 0$), hence a measure, then the population state probability distribution after the n-th intervention is given by:

$$(P\mu)(A) := \int_S Q_{\phi_{\Delta t}(x)}(A) d\mu(x).$$

Here, P is a *Markov operator* that is P:

- maps positive Borel measures to positive Borel measures;
- is additive and positively homogenous;
- conserves mass.

A specific, more elaborate case of such a model can be found in [AHG12].

Other interesting applications of the measure-theoretical framework in modelling can be found in [EHM15], [EHM16]. These papers present applications to crowd dynamics. See also [AI05] for measure-formulation in population dynamics. As we see, measures naturally occur from these deterministic models.

Let us now go to the second type of models, probabilistic ones, which motivate the usage of measure spaces.

A probabilistic perspective

Let $(X_t^x)_{t\geq 0}$ be a family of stochastic processes in continuous time on a Polish space S with the Markov property. Here the superscript x indicates that $(X_t^x)_{t\geq 0}$ starts at $t = 0$ at x almost surely. For f a continuous and bounded function on S, i.e. $f \in \mathcal{C}_b(S)$ and μ a Borel probability measure describing the distribution for the start position x of the process define

$$\langle P_t \mu, f \rangle := \int_S \mathbb{E}[f(X_t^x)] d\mu(x).$$

To $f \in C_b(S)$ one can associate a function $U_t f$ given by

$$U_t f(x) := \mathbb{E}[f(X_t^x)].$$

Under conditions on the processes (being Feller), $U_t f \in C_b(S)$, in which case one obtains a semigroup of positive operators $(U_t)_{t\geq 0}$ on $C_b(S)$, such that $U_t \mathbb{1} = \mathbb{1}$. Then P_t is a Markov operator and $(P_t)_{t\geq 0}$ a *Markov semigroup*. That is, the operators satisfy a semigroup

$$\begin{array}{cccccccc}
f(X_1(\omega_1)) & f(X_2(\omega_1)) & \cdots & f(X_n(\omega_1)) & \cdots & \xrightarrow{\text{CLT, SLLN}} & \langle \mu^*, f\rangle \\
f(X_1(\omega_2)) & f(X_2(\omega_2)) & \cdots & f(X_n(\omega_2)) & \cdots & & \\
 & & \ddots & & & & \\
f(X_1(\omega_n)) & f(X_2(\omega_n)) & \cdots & f(X_n(\omega_n)) & \cdots & & \\
\downarrow \text{(CLT, SLLN)} & \downarrow & \cdots & \downarrow & & & \\
\langle P\mu_0, f\rangle & \langle P^2\mu_0, f\rangle & \cdots & \langle P^n\mu_0, f\rangle & \cdots & \xrightarrow{\text{Asymptotic stability}} & \langle \mu^*, f\rangle
\end{array}$$

Figure 3: *Sample trajectories*

property similar to (*) This semigroup in $C_b(S)$ is dual to $(P_t)_{t\geq 0}$:

$$\langle P_t\mu, f\rangle = \langle \mu, U_t f\rangle$$

for all $f \in \mathcal{C}_b(S)$.

There is a vast mathematical literature on Markov processes and semigroups. The interested reader may start in e.g. [KP80, MT09, LM00].

Let us consider a process X_n on $(\Omega, \mathcal{F}, \mathbb{P})$, $X_n : \Omega \to S$ and let us consider its realizations/sample trajectories $(f(X_n(\omega)))_{n\in\mathbb{N}}, \omega \in \Omega$ (see Figure 3). One of the fundamental problems of classical probability theory is the question about the asymptotic behaviour of the functional $f(X_n)$ as $n \to \infty$, where $f : S \to \mathbb{R}$ is a Borel measurable function, called an observable, for S Polish.

One of the question is whether the **Strong Law of Large Numbers** (SLLN) holds, i.e. whether time averages $\frac{1}{n}\sum_{m=1}^{n} f(X_m)$ converge in some sense to a constant, say C_f^*. If this is the case, then another question concerns fluctuations around C_f^*. Typically, if the observable is not 'unusually large', after proper rescaling fluctuations can be described by a Gaussian random variable. Here we see the **Central Limit Theorem** (CLT), which states that the random variable $\frac{1}{\sqrt{n}}\sum_{m=0}^{n}[f(X_m) - C_f^*]$ converges in law as $n \to \infty$ to a finite variance, centred normal variable. Put differently, roughly speaking, the time averages $\frac{1}{n}\sum_{m=0}^{n} f(X_m(\omega))$ of a sample trajectory will converge to C_f^* at a rate $\frac{1}{\sqrt{n}}$, as $n \to \infty$.

Central limit theorems proven for stationary Markov processes can be traced back to 1938 article [Doe38], in which Doeblin proved the central limit theorem for discrete time, countable Markov chains. Nowadays a sufficient condition for geometric ergodicity of an ergodic Markov chain is called the *Doeblin condition*, see [Lot86].

For stationary and ergodic Markov processes central limit theorems has been proven using different techniques throughout the years in e.g. [GH04, Eag75, DM02, MW00, Wu07] for discrete cases and in e.g. [Bha82, Hol05] for the continuous cases.

For non-stationary Markov processes we can find Central Limit Theorem results in [KW12]. Though, additional assumptions of spectral gap in the Wasserstein metric are needed to get the required results. In Chapter 5 we provide a new result of the validity of the Central Limit Theorem for a class of non-stationary Markov processes.

Asymptotic stability

Let us recall the definition of asymptotic stability of a Markov operator P on measures.

First let us introduce the definition of weak convergence of measures. Following [Bog07a] we say that a net $\{\mu_\alpha\}$ of measures is weakly convergent to a measure μ if for every continuous bounded real function f on S, one has

$$\lim_\alpha \int_S f(x)\mu_\alpha(dx) = \int_S f(x)\mu(dx).$$

Weak convergence can be defined by a topology. The weak topology on the space of finite signed Borel measures on S is the topology $\sigma(\mathcal{M}(S), \mathcal{C}_b(S))$: the weakest locally convex topology on $\mathcal{M}(S)$ such that the linear functions $\mu \to \int_S f d\mu$ are continuous, for every $f \in C_b(S)$. For more details see [Bog07a], Chapter 8.

Definition 0.0.3. *A measure μ^* is called invariant for the Markov operator P if $P\mu^* = \mu^*$. A Markov operator P is asymptotically stable if there exists an invariant measure $\mu^* \in \mathcal{P}(S)$ such that $P^n\mu \to \mu^*$ weakly as $n \to \infty$ for every $\mu \in \mathcal{P}(S)$.*

Note that the invariant measure of an asymptotically stable Markov operator is necessarily unique.

We can see that asymptotic stability examines the properties of the limit of $\langle P^n\mu_0, f\rangle$. Natural questions one may ask is how can we examine properties of the process P by analyzing properties of sample trajectories.

As we see in [LM99] or [HHSWS15, HS16, Hor06] asymptotic stability is the main tool for proving Central Limit Theorems and the Strong Law of Large Numbers. The existence of asymptotically stable, unique invariant measures for some classes of Markov processes, including those which the state space need not be locally compact, was obtained in [DX11, HM08, Sza08, KPS10]

Issues with L^1 as a state space

In the literature one can find multiple approaches to Markov semigroups. Many authors use an L^1 space as a state space. That is, an L^1 space with respect to a suitable (invariant) measure related to the Markov semigroup. Rudnicki in [Rud97] and [Rud00] works with Markov operators on an L^1space giving interesting examples of applications of Markov operators to diffusion processes and population dynamics. Also Rudnicki, Pichor and Tyran-Kamińska in [RPTK02] examine asymptotic properties of Markov operators and semigroups on L^1. In the book of Emelianov [Eme07] the L^1 setting is described which is motivated by applications to the probability theory and dynamical systems of Markov semigroups. Also Lasota and Mackey in [LM94] describe applications of Markov semigroups on L^1 spaces to the theory of fractals.

On the other hand authors like Szarek in eg. [SW12], [Sza97], [SM03] and Komorowski, Peszat, Szarek in [KPS10] work in spaces of measures instead of L^1 space. Let us show why we choose to work in this setting of measures too and what advantages it gives to work in spaces of measures.

Let us show the example, based on [GLMC10], how the measure-approach mitigates one issue, which is the *inconsistency of the L^1 norm with empirical data.*

Example 0.0.4. *(based on [DGMT98]) In observing populations in biology, social sciences and life sciences one often encounters the following situation. Individuals are characterised by states in a state space S. One splits these states into disjoint classes, e.g. age groups, length intervals, weight, etc. [Web08]:*

$$S_n : \quad S = \bigcup_{n=1}^{N} S_n \quad \text{where } N \text{ may be } \infty.$$

At specific times one observes - ideally - the total number of individuals with state in each class. For simplicity of exposition, take $S = \mathbb{R}_+$ and $S_n^h := S_n = [\,nh, (n+1)h\,)$. In modelling a population, the population state is described by a density function $F(x)$. So, the number of individuals with a state in a set $E \subset S$ is given by $\int_E F(x)dm(x)$ where m is Lebesgue measure on $\mathbb{R}_+$. Observations will be the total count of individuals in a class, i.e. values

$$a_n^h = \int_{S_n^h} F(x)dm(x).$$

Hence, the observed data does not approximate the density function F itself, but the integral of the density over state classes.

Let us now see what happens if we make our classes (age, weight, height) smaller, i.e. h becomes $h' < h$. Then observations $a_n^{h'}$ for the associate h' will give a better 'estimate' for $F(x)$.

Indeed, if F is continuous, then the Mean Value Theorem for Integrals implies that

$$F(x) \approx \frac{a_n^h}{h}$$

with $n = n(h,x)$ such that $x \in [nh,(n+1)h)$ for h sufficiently small. This is a pointwise estimate. That means that the rate of convergence of $\frac{a^h_{n(x,h)}}{h} \to F(x)$ as $h \downarrow 0$ can (and will typically) vary with x.

If one considers instead the estimation of F in $L^1(\mathbb{R}_+)$, then for a given size $h > 0$ of the class, the set

$$A^h_{n,L^1} := \left\{ f \in L^1(\mathbb{R}_+) : f \geq 0, \int_{[nh,(n+1)h)} f dm = a_n^h \right\}$$

consists of all distribution functions in $L^1(\mathbb{R}_+)$ that yield the observed numbers a_n^h in the classes S_n^h.

The size of this set in $L^1(\mathbb{R}_+)$ can be characterized by its diameter. We have for $f,g \in A^h_{n,L^1}$ that $\|f-g\|_{L^1} \leq \|f\|_{L^1} + \|g\|_{L^1} \leq 2\sum_{n=0}^\infty a_n^h$. On the other hand, for any $f \in A^h_{n,L^1}$, $g = \left\{ \frac{2a_n^h}{h} - f \text{ on } [nh,(n+1)h) \right\} \in A^h_{L^1}$ and $\|f-g\|_{L^1} = 2\sum_{n=1}^\infty a_n^h$. Hence,

$$\mathrm{diam} A^h_{n,L_1} := \sup\{\|f-g\|_{L^1} : f,g \in A_{L^1}\} = 2\sum_{n=0}^\infty a_n^h = 2\int_S F(x) dm.$$

Thus, $\mathrm{diam} A^h_{n,L_1}$ is independent of h. In other words, with the decreasing size of the classes, the set of possible distributions that are constant with the observations does not shrink in size. The L^1-distance between functions f and g equals the total variation distance (see Section 1.1) between the measures fdm and gdm.

$$\|f-g\|_{L^1(R_+,dm)} = \|fdm - gdm\|_{TV}.$$

The weak topology on measures, when restricted to the positive measures, is measurable, i.e. by means of the so-called Dudley metric, derived from the dual bounded Lipschitz norm $\|\cdot\|^*_{BL}$, see Section 1.1.

According to [GLMC10] in the Dudley metric $\|\cdot\|_{BL}^*$

$$\mathrm{diam}_{\|\cdot\|_{BL}^*}\left(\{f dm : f \in A_{n,L^1}^h\}\right) \leq h \cdot \sum_{n=1}^{\infty} a_n^h = h \int_{\mathbb{R}_+} F(x) dm.$$

So, in the Dudley metric, the set of all distributions that are consistent with observations does shrink to the actual distribution $F(x)dm$ *when the size of the classes decreases to zero. This shows that considering* L^1 *for equations describing processes based on empirical data may not be an optimal choice.*

Switching of dynamics

A common approach when it comes to constructing a new dynamical system from known ones is by *perturbation*. One approach, commonly employed in the field of differential equations, is adding new processes to the system at infinitesimal small time intervals. That is, one adds what is often called 'reaction terms' to the vector field that defines the dynamics. Another approach is that of switching between dynamics.

Let us present a few examples of mixing perturbations and different types of dynamics. Let A and B be $n \times n$ matrices and consider the linear system of ODEs in $\mathbb{R}^n$:

$$\frac{dx}{dt}(t) = Ax(t) + Bx(t). \tag{1}$$

The solution operator to (1) is given by the matrix expansion $e^{(A+B)t}$. In the sense described above, this 'model' describes two systems defined individually by

$$\frac{dx}{dt} = Ax, \quad \frac{dx}{dt} = Bx$$

combined together through infinitesimal superposition.

Alternatively, one may consider *switching* between the dynamics defined by A and that by B after time intervals Δt. That is, the trajectory defined inductively by $x_0 \in S$,

$$x_n := \begin{cases} e^{A\Delta t} x_{n-1}, & \text{if } n \text{ is even} \\ e^{B\Delta t} x_{n-1}, & \text{if } n \text{ is odd} \end{cases}$$

and

$$x_{\Delta t}(t, x_0) := \begin{cases} e^{A(t-\Delta t)} x_n, & \text{if } t \in [n\Delta t, (n+1)\Delta t), n \text{ is even} \\ e^{B(t-\Delta t)} x_n, & \text{if } t \in [n\Delta t, (n+1)\Delta t), n \text{ is odd} \end{cases}$$

Example 0.0.5 (*Lie product formula*, see [LE70])**.** *The Lie product formula, named after Sophus Lie, is the simplest, most basic formula showing that switching scheme for matrices A and B yields the same, that in the limit of infinitely fast switching*

$$e^{(A+B)t} = \lim_{n\to\infty} \left(e^{\frac{At}{n}} e^{\frac{Bt}{n}} \right)^n .$$

That is, the trajectory of the switched system will converge to that defined by infinitesimal superposition, in the limit of the infinitely fast switching.

Another example of switching different types of dynamics is Iterated Function Systems.

Example 0.0.6 (*Iterated Function Systems*)**.** *The iteration of a map Φ that maps the state space S into itself yields a dynamical system on S in discrete time. If one has N such maps $\Phi_i : S \to S, i = 1, \cdots, N$, one may alternate the application of the various Φ_i.*

This can be done probabilistically: with probability p_i one chooses map Φ_i (without memory of the map that has been applied in the previous step).

If the system is located at $x_0 \in S$, then the probability distribution for the location after the application of one of the maps Φ_i is

$$\sum_{i=1}^{N} p_i \delta_{\Phi_i(x_0)} \in \mathcal{P}(S),$$

where $\delta_{x'}$ denotes the Dirac or point mass located at x':

$$\delta_{x'}(E) = \begin{cases} 1, & \text{if } x' \in E \\ 0, & \text{otherwise} \end{cases}$$

Such a combination of a set of maps Φ_i and probabilities p_i by which one applies these maps constitutes the simplest example of an Iterated Function System (IFS).

Each of the maps Φ_i defines a (deterministic) Markov operator P_{Φ_i} by means of push-forward:

$$P_{\Phi_i}\mu(E) := \mu(\Phi_i^{-1}(E)), \quad \mu \in \mathcal{M}(S).$$

The Markov operator associated to the IFS (or the Markov chain associated to the IFS) is

$$P = \sum_{i=1}^{N} p_i P_{\Phi_i}.$$

More complicated versions (in particular for analysis of their behaviour) include e.g. de-

pendence of the map selection probabilities p_i *on states:*

$$p_i = p_i(x_0).$$

Iterated Function Systems are an important tool in the study of fractals and generalized fractals [LM00, LY94, HUT81, Bar12, LM94]

Example 0.0.7 (*Piecewise Deterministic Markov Processes* originated with [Dav84])**.** *Constructing a Piecewise Deterministic Markov Process (PDMP) is another way of getting a new dynamical system. PDMPs are a family of Markov processes involving a deterministic motion perturbed by a random jump.*

In Figure 4 we see a graphical presentation of an example of a PDMP. Motion starts at some point X_0 *and then* X_t *is given by a deterministic flow* $\phi_t(X_0)$ *until the first jump. Jumps occur spontaneously, for example in a Poisson-like fashion, with a certain rate. After a jump we land at* X_{t_1} *and motion restarts as before, that is, according to the fixed deterministic dynamical system* $(\phi_t)_{t\geq 0}$ *in* S*.*

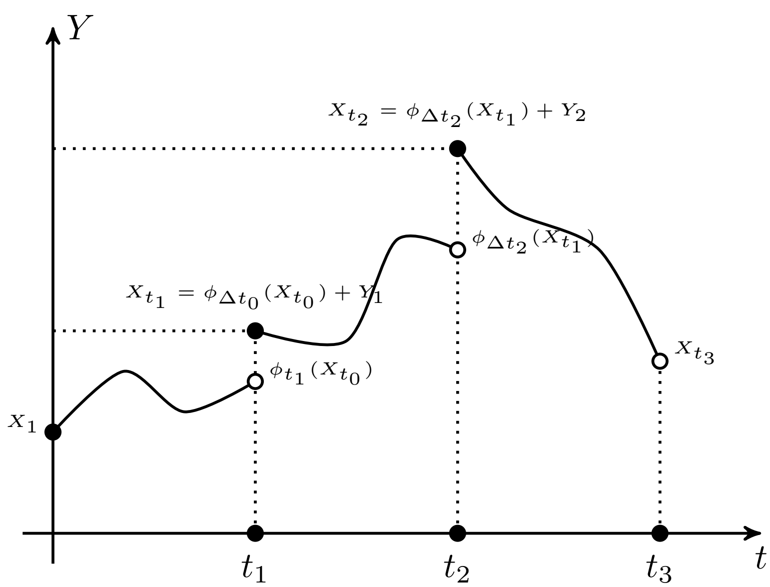

Figure 4: *Piecewise Deterministic Markov Process starting at* $t_0 = 0$ *with value* $X_0 \in Y$*. The motion until time* t_1*, the time of the first jump, is given by* $\phi_t(X_{t_0})$*. At time* t_1 *we have the first jump* Y_1*. Hence,* $X_{t_1} = \phi_{\Delta t_0}(X_1) + Y_1$ *and* X_{t_2} *becomes the 'new' starting point for the next deterministic evolution on the interval* $\Delta t_2 = t_2 - t_1$*.*

The more precise description of construction of PDMP can be found in [HADD84].

Many well-known examples fall into the framework of PDMP. In [HADD84] we can find descriptions of multiple models, both theoretical and applied, where PDMPs play a crucial role. Let us show one of these examples, the so called M/G/1 Queue (Example 0.0.8).

Example 0.0.8. *[M/G/1 Queue, [HADD84]] Customers arrive at a single-served queue according to a Poisson process with rate μ, and have independent identically distributed (i.i.d.) service time requirements with distribution function F. The virtual waiting time (VWT) is the time a customer arriving at time t would have to wait for service. This decreases at a unit rate between arrivals- see Figure 5. The queue has two states, "busy" and "empty". Hence, when VWT reaches 0, we get transition from state 1 ("busy"), to state 0 ("empty").*

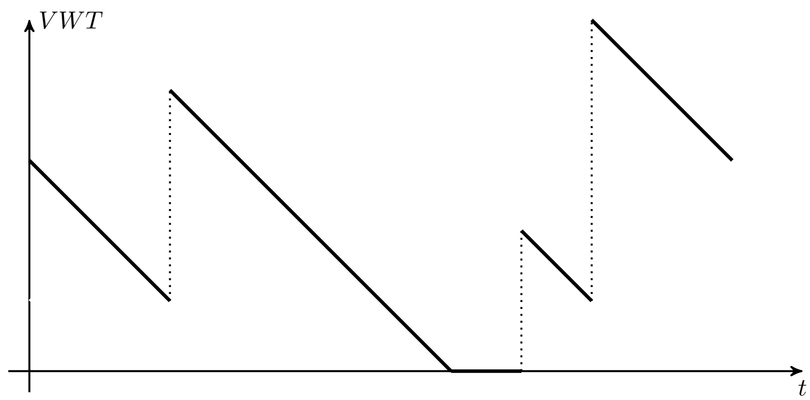

Figure 5: *M/G/1 Queue. A queue model, where arrivals are Markovian (modulated by a Poisson process), service times have a General distribution and there is a single server. VWT is the virtual waiting time*

Example 0.0.9 (Random dynamical systems, [HCWS17])**.** *In Figure 6 we show a more complicated example of a PDMP $(\bar{Y}(t))_{t\geq 0}$ from [HCWS17]. The deterministic component of the process evolves according to a finite number of semiflows, which are chosen with certain probabilities at switching times $\tau_1, \tau_2, \ldots$. Here we get additional randomness in the position after jumps. Hence, we "land" in an ϵ-neighbourhood of the state after the jump.*

Stability and ergodicity of PDMPs can be found in the work of Costa and Dufour [CD08, CD09, CD10]. All these results concern PDMPs for which the state space S is locally compact and Hausdorff. There are almost no results for PDMPs on Polish spaces, even

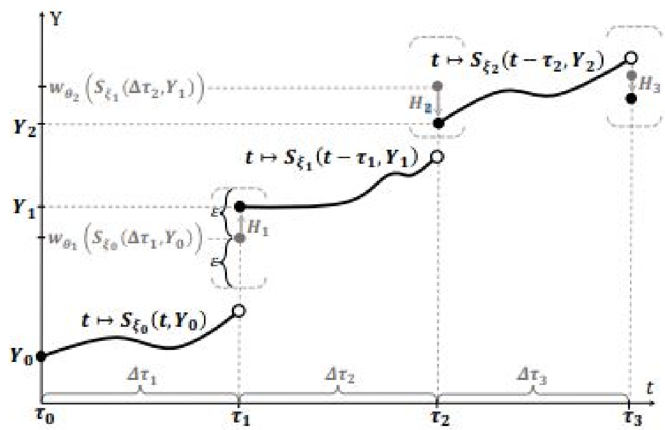

Figure 6: *[HCWS17] More general Piecewise Deterministic Markov Process. The deterministic component of the system evolves according to a finite collection of semiflows (randomly switched with time). Randomness of post-jump location comes from a selected semiflow and a random shift within an ϵ-neighbourhood.*

though there are strong examples showing that choosing a Polish space to work on is the right choice. In [GRTW11] we can find an analysis of PDMPs, on non-locally compact state space. This setting we shall call *infinite-dimensional*, because the state space is (a part of) an infinite dimensional Banach space. In [RTT16] the infinite-dimensional case of PDMPs is applied to neuron models.

Switching systems-different approaches

Switching schemes like the Lie-Trotter research presented in Chapter 3 were motivated by the idea of applying such schemes in the analysis of the long-term dynamics of complex deterministic dynamical systems. It relates to so-called *operator splitting techniques* which date back to the 1950s and found ample applications in Numerical Analysis. The classical splitting methods are the Lie-Trotter splitting, the Strang splitting [DHZ01, Str68, FH07] and the symmetrically weighted splitting method [Str63, CFH05]. The research in Chapter 3 was motivated to extend these approaches to the setting of Markov semigroups.

Originally splitting schemes applied to semigroups of strongly continuous linear operators, so-called C^0-semigroups [EBNHM13, HP57] and there were attempts to extend it to semigroups of non-linear operators with mixed success [CG12, KP84]. Our case of interest is that of Markov semigroups. There are several issues when working with Markov semigroups on spaces of measures. Although Markov semigroup are linear in the space of measures, they need not be strongly continuous operator semigroups, for the Dudley norm for example: the operators P_t that constitute the semigroup need not be continuous on the vector space of measures for the relevant topology, but only on the cone of positive measures. See Chapter 3 for more details.

Our models of interest are described by Markov operators. The objective is to provide conditions for convergence that are trackable in concrete models coming from applications. The theory of strongly continuous semigroups does not apply to these cases. Hence, the existing results for strongly continuous semigroups cannot be applied in our setting.

The connection between switching systems and their limit in the case of 'infinitely fast' switching - if it exists - can be exploited in two ways:

1. The first way of approaching switched systems is the so-called "divide and conquer" method [HKLR10, HP18]. The idea is to start from a known complicated system and split it into 'easier' systems to get a solution. Examples of 'divide and conquer' methods are:

    - 'Classical' Lie-Trotter [Tro59]

    - Convergence Rates of the Splitting Scheme [CvN10, GLMC10]

2. The second approach is to start from a switched system which is difficult to analyse. If we know that the limit of the system is close to the system itself we can analyse the limit instead of the switched system. This works well if one is able to identify the limit of the switched system. Here the natural question is what can we say about the limit of the switching system. Can we identify the generator of the limit semigroup? What can we say about the properties like continuity? Which properties are inherited by the limit from switching semigroups?

Focus on equicontinuity

Let us show now how working with equicontinuous families of Markov operators can lead to a generalization of existing concepts of contractive or non-expansive Markov operators.

A Markov operator P defined on a Polish state space S in a natural way defines by iterations a dynamical system on the space of probability measures. Natural questions occurring in the theory of dynamical systems are the ones describing the behaviour of the system. Hence, we are looking for example for steady states, which in the space of measures would be invariant measures, i.e. $\mu^* \in \mathcal{P}(S)$ such that $P\mu^* = \mu^*$.

We say that a Markov operator P is **strictly contractive** for the metric d on $\mathcal{P}(S)$ if

$$d(P\mu, P\nu) < d(\mu, \nu) \quad \text{for every} \quad \mu, \nu \in \mathcal{P}(S).$$

For strictly contractive Markov operators the natural tool to use is the Banach Fixed Point Theorem, which yields the existence of a unique invariant measure μ^*, provided $(\mathcal{P}(S), d)$ is complete. Moreover, this invariant measure is then automatically globally stable, as $d(P^n\mu, \mu^*) \to 0$ as $n \to \infty$ for every $\mu \in \mathcal{P}(X)$.

However, Markov operators are in general not strictly contractive.

We say that a Markov operator P is **non-expansive** for the metric d on $\mathcal{P}(S)$ if

$$d(P\mu, P\nu) \le d(\mu, \nu) \quad \text{for every} \quad \mu, \nu \in \mathcal{P}(S).$$

Szarek shows results of existence and uniqueness of invariant measures for non-expansive Markov operators that are non-expansive in a Fortet-Mourier norm [Sza03] .

Definition 0.0.10 ([Sza03] restricted to $\mathcal{P}(S)$)**.** *A Markov operator P is non-expansive for $\|\cdot\|_{FM,\rho}$, where ρ is some admissible metric in S, if*

$$\|P\mu_1 - P\mu_2\|_{FM,\rho} \le \|\mu_1 - \mu_2\|_{FM,\rho} \quad \textit{for} \quad \mu_1, \mu_2 \in \mathcal{P}(S),$$

where

$$\|\nu\|_{FM,\rho} = \sup\{|\langle f, \nu\rangle| : f \in \mathcal{C}(S), |f(x)| \le 1, |f(x) - f(y)| \le \rho(x,y)\}. \tag{2}$$

Any metric ρ that metrizes the topology of S such that (S, ρ) is separable and complete is called *admissible*. We will denote by $\mathcal{D}(S)$ the family of all admissible metrics on S. By $\mathrm{BL}(S,\rho)$ we will denote the space of bounded Lipschitz functions, hence

$$\mathrm{BL}(S,\rho) := \{f \in \mathcal{C}(S) : \|f\|_\infty < \infty, |f|_L < \infty\}.$$

Non-expansiveness is in principal dependent on a metric d, in particular on the choice of metric ρ on the underlying state space S if $d(\mu,\nu) = \|\mu - \nu\|_{FM,\rho}$. Markov operator may be non-expansive according to Definition 0.0.10 for an admissible metric ρ, but not for another admissible metric ρ'.

Let us now look at the family of iterates of Markov operator $\{P^n : n \in \mathbb{N}\}$. For P non-expansive this family is **equicontinuous** in the sense of the following definition.

Definition 0.0.11. *Let T be a topological space and (S,d) a metric space. We say that the family of continuous maps $\mathcal{E} \subset \mathcal{C}(T,S)$ is* ***equicontinuous*** *at $t_0 \in T$ if for every $\epsilon > 0$ there exists an open neighbourhood U_ϵ of t_0 such that*

$$d(f(t_0), f(t)) < \epsilon \quad \text{for all} \quad f \in \mathcal{E}, t \in U_\epsilon.$$

$\mathcal{E}$ is equicontinuous if it is equicontinuous at every point $t \in T$.

The equicontinuity of a family of iterates of a non-expansive Markov operator motivates the investigation of the class of Markov operators for which the family of its iterates is equicontinuous.

In the literature we can find a few concepts related to equicontinuity of families of Markov operators. In 1964 Jamison [Jam64] described the asymptotic behaviour of iterates of Markov operators on a compact metric space where he assumed equicontinuity of the family of (dual) Markov operators. For such operators he got the following results:

Theorem 0.0.12. *Let P be a regular Markov operator on a compact metric space X. Let U be a dual operator for P. Let P be a Feller operator, i.e. U maps $\mathcal{C}_b(X)$ into itself. Then the following conditions are equivalent:*

(i) P has a unique invariant measure.

(ii) For every $f \in \mathcal{C}(X)$ the sequence $U^{(n)}f := \frac{1}{n}\sum_{k=0}^{n-1} U^k f$ converges uniformly to a constant.

(iii) For every $f \in \mathcal{C}(X)$ the sequence $U^{(n)}f := \frac{1}{n}\sum_{k=0}^{n-1} U^k f$ converges pointwise to a constant.

The equivalence of (i) and (ii) is Theorem 2.1 from [Jam64] and the equivalence of (ii) and (iii) is Theorem 2.3 from [Jam64].

List of chapters and related works

- *Chapter 1 - Fundamental concepts and results*
- *Chapter 2 - On a Schur like property for spaces of measures and its consequences*, based on the work Sander C. Hille, Tomasz Szarek, Daniel T.H. Worm, Maria Ziemlańska. On a Schur-like property for spaces of measures. Preprint available at https://arxiv.org/abs/1703.00677. Main results published in *Statistics and Probability Letters*, Volume 169, 2021, https://doi.org/10.1016/j.spl.2020.108964.
- *Chapter 3 - Lie-Trotter product formula for locally equicontinuous and tight Markov operators*, based on the work Sander C. Hille, Maria A. Ziemlanska. Lie-Trotter product formula for locally equicontinuous and tight Markov semigroup. Preprint available at https://arxiv.org/abs/1807.07728
- *Chapter 4 - Equicontinuous families of Markov operators in view of asymptotic stability*, based on the work Sander C. Hille, T. Szarek, Maria A. Ziemlanska. Equicontinuous families of Markov operators in view of asymptotic stability. *Comptes Rendus Mathematique*, Volume 355, Number 12, Pages 1247-1251, 2017
- *Chapter 5 - Central Limit Theorem for some non-stationary Markov chains*, based on the work Jacek Gulgowski, Sander C. Hille, Tomasz Szarek, Maria A. Ziemlańska. Central Limit Theorem for some non-stationary Markov chains. *Studia Mathematica*, Number 246 (2019), Pages 109-131

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 $\Phi_t : S \to S$ be a family of measurable maps, $t \in \mathbb{R}_+$ such that $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_0 = \mathrm{Id}_S$. 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 $f : S \to \mathbb{R}$ is (globally) Lipschitz if there exists $L \geq 0$ such that*

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

Let $\mathrm{Lip}(S, d)$ (or $\mathrm{Lip}(S)$ for shorter notation) denote the vector space of Lipschitz functions on (S, d). The *Lipschitz constant* of $f \in \mathrm{Lip}(S)$ 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 $\mathrm{BL}(S, d)$ will denote the Banach space of all bounded Lipschitz functions f on S with the bounded Lipschitz or Dudley norm

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

Proposition 1.1.3 ([Wor10], Proposition 2.2.7)**.** $\mathrm{BL}(S, d)$ *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

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

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 $\|f\|_{BL(S,d,p)} := (|f|^p + \|f\|^p_\infty)^{\frac{1}{p}}$, $1 < p < \infty$.

The space $\mathcal{M}(S)$ embeds into $\mathrm{BL}(S)^*$ 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 $\mathrm{BL}(S,d)^*$ with the norm

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

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 $f : X \to Y$. 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 ecolution $\delta_x, x \in X$ 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 $f_\alpha(x) \to f(x)$ for each $x \in X$. Note that the topology of X does not play a role in the results on the topology of poinwise 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 $f : X \to Y$. Thus, for fixed f the map $X \to Y : x \mapsto f(x)$ is continuous. We look for a topology on $\mathcal{F}$ such that the map $\mathcal{F} \times X \to Y : (f,x) \mapsto f(x)$ 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 $f : X \to S$.

Definition 1.1.5. *The family $\mathcal{F}$ of functions $f : X \to S$ 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

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

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 unifornly 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 $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ such that

(i) P is additive and $\mathbb{R}_+$ homogenous: $P(\lambda_1\mu_1 + \lambda_2\mu_2) = \lambda_1 P\mu_1 + \lambda_2 P\mu_2$; 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 $U : \mathrm{BM}(S) \to \mathrm{BM}(S)$ is called dual to P if

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

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 $P_0 = Id$. 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 $\mathrm{BM}(S)$ 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 $\mu_n \to \mu$ weakly, then $\|\mu_n - \mu\|_{BL}^* \to 0$;

(ii) $\|\mu_n - \mu\|_{BL}^* \to 0$ implies $\mu_n \to \mu$ 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 $\mu_n \to \mu$ 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 $\|\mu_n - \mu\|_{BL}^* \to 0$ as $n \to \infty$, then $\mu_n \to \mu$ weakly.

We provide conditions on subsets of (signed) mesures $\mathcal{M}(S)$ 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 $f \in BL(S, d)$, $(\mu_n)_n$ 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 $x^* e_n \to 0$ 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

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

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 $w_i : S \to S, i \in I$, where (S, d) is a complete separable metric space with a family of continuous functions $p_i : S \to [0,1], i \in I$ 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. PDMSs 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 $q : \mathbb{R} \to \mathbb{R}$ a measurable and locally integrable function, where $M_{iq} f := iq \cdot f$ (on f in a smaller domain):

$$\begin{aligned} T(t)f(x) &:= e^{itq(x)} f(x) \\ S(t)f(x) &:= f(x+t) \\ \text{Further, define} & \\ U(t)f(x) &:= e^{\left(i\int_0^t q(x+s)ds\right)} f(x+t). \end{aligned}$$

For $f \in X$ we can compute products

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

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

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

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 $\eta > 0$:

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

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

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

Also

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

Let $\varepsilon > 0$. Take $t_0 \in \mathbb{R}, \delta > 0$ and let $\hat{I} \coloneqq \bigcup_{0\leq s\leq t} I + s$. 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,

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

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:

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

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

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

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

For $f\in C_c(\mathbb{R})$ and again using Corollary 2.2.6 (Bogachev [Bog07b], p.113) we get that $E_n f \to Ef$ 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

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

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

$$\begin{aligned}\|E_n f - Ef\|_{L^1} &\le \|E_n f - E_n f_0\|_{L^1} + \|E_n f_0 - Ef_0\|_{L^1} + \|Ef_0 - E_n f_0\|_{L^1} \le \\ &\le \|f-f_0\|_{L^1} + \|E_n f_0 - Ef_0\|_{L^1} + \|f-f_0\|_{L^1} \le 3\varepsilon\end{aligned}$$

for a sufficiently large n.
So we get in $L^1(\mathbb{R})$ that

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

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 $\mathcal{D}(A+B)=\{0\}$.

Chapter 2

On a Schur-like property for spaces of measures and its consequences

This chapter is based on:

Sander C. Hille, Tomasz Szarek, Daniel T.H. Worm, Maria Ziemlańska. On a Schur-like property for spaces of measures. Preprint available at https://arxiv.org/abs/1703.00677

Abstract:

A Banach space has the Schur property when every weakly convergent sequence converges in norm. We prove a Schur-like property for measures: if a sequence of finite signed Borel measures on a Polish space is such that it is bounded in total variation norm and such that for each bounded Lipschitz function the sequence of integrals of this function with respect to these measures converges, then the sequence converges in dual bounded Lipschitz norm or Fortet-Mourier norm to a measure. Two main consequences result: the first is equivalence of concepts of equicontinuity in the theory of Markov operators in probability theory and the second concerns conditions for the coincidence of weak and norm topologies on sets of measures that are bounded in total variation norm that satisfy additional properties. Finally, we derive weak sequential completeness of the space of signed Borel measures on Polish spaces from the Schur-like property.

2.1 Introduction

The mathematical study of dynamical systems in discrete or continuous time on spaces of probability measures has a long-lasting history in probability theory (as Markov operators and Markov semigroups, see e.g. [MT09]) and the field of Iterated Function Systems [BDEG88, LY94] in particular. In analysis there is a growing interest in solutions to evolution equations in spaces of positive or signed measures, e.g. in the study of structured population models [AI05, CCC13, CCGU12], crowd dynamics [PT11] or interacting particle systems [EHM16]. Although an extensive body of functional analytic results have been obtained within probability theory (e.g. see [Bil99, Bog07a, Dud66, LeC57]), there is still need for further results, driven for example by the topic of evolution equations in space of measures, in which there is no conservation of mass.

This chapter provides such functional analytic results in two directions: one concerning properties of families of Markov operators on the space of finite signed Borel measures $\mathcal{M}(S)$ on a Polish space S that satisfy equicontinuity conditions (Theorem 2.3.5). The other provides conditions on subsets of $\mathcal{M}(S)$, where S is a Polish space, such that weak topology on $\mathcal{M}(S)$ coincides with the norm topology defined by the Fortet-Mourier or dual bounded Lipschitz norm $\|\cdot\|_{\mathrm{BL}}^*$ (Theorem 2.3.7 and similar results in Section 2.3.2).

Both are built on Theorem 2.3.1, which states that if a sequence of signed measures is bounded in total variation norm and has the property that all real sequences are convergent that result from pairing the given sequence of measures by means of integration to each function in the space of bounded Lipschitz functions, $\mathrm{BL}(S)$, then the sequence is convergent for the $\|\cdot\|_{\mathrm{BL}}^*$-norm. This is a Schur-*like* property. Recall that a Banach space X has the Schur property if every weakly convergent sequence in X is norm convergent (e.g. [AJK06], Definition 2.3.4). For example, the sequence space ℓ^1 has the Schur property (cf. [AJK06], Theorem 2.3.6). Although the dual space of $(\mathcal{M}(S), \|\cdot\|_{\mathrm{BL}}^*)$ is isometrically isomorphic to $\mathrm{BL}(S)$ (cf. [HW09b], Theorem 3.7), the (completion of the) space $(\mathcal{M}(S), \|\cdot\|_{\mathrm{BL}}^*)$ is not a Schur space, generally (see Counterexample 2.3.2). The condition of bounded total variation cannot be omitted.

Properties of the space of Borel probability measures on S for the weak topology induced by pairing with $C_b(S)$ have been widely studied in probability theory, e.g. consult [Bog07a] for an overview. Dudley [Dud66] studied the pairing between *signed measures* and the space of bounded Lipschitz functions, $\mathrm{BL}(S)$, in further detail. Pachl investigated extensively the related pairing with $\mathrm{U}_b(S)$, the space of uniformly continuous and bounded functions [Pac79, Pac13]. See also [Kal04]. Because of our interest in equicontinuous families of

Markov operators on the one hand, which is intimately tied to 'test functions' in the space $\mathrm{BL}(S)$, and to dynamical systems in spaces of measures equipped with the $\|\cdot\|_{\mathrm{BL}}^*$-norm, or flat metric, on the other hand, we consider novel functional analytic properties of the space of finite signed Borel measures $\mathcal{M}(S)$ for the $\mathrm{BL}(S)$-weak topology in relation to the $\|\cdot\|_{\mathrm{BL}}^*$-norm topology.

Equicontinuous families of Markov operators were introduced in relation to asymptotic stability: the convergence of the law of stochastic Markov process to an invariant measure (e.g. e-chains [MT09], e-property [CH14, KPS10, LS06, Sza10], Cesaro-e-property [Wor10], Ch.7; see also [Jam64]). Hairer and Mattingly introduced the so-called asymptotic strong Feller property for that purpose [HM06]. Theorem 2.3.5 rigorously connects two dual viewpoints – concerning equicontinuity: Markov operators acting on measures (laws) and Markov operators acting on functions (observables). In dynamical systems theory too, there is special interest in ergodicity properties of maps with equicontinuity properties (e.g. [LTY15]).

The structure of the chapter is as follows. After having introduced some notation and concepts in Section 2.2 we provide in Section 2.3 the main results of the chapter. The delicate and rather technical proof of the Schur-like property, Theorem 2.3.1, is provided in Section 2.4. It uses a kind of geometric argument, inspired by the work of Szarek (see [KPS10, LS06]), that enables a tightness argument essentially. Note that our approach yields a new, independent and self-contained proof of the $\mathrm{U}_b(S)$-weak sequential completeness of $\mathcal{M}(S)$ (cf. [Pac79], or [Pac13], Theorem 5.45) as corollary. Section 2.5 shows that the Schur-like property also implies – for Polish spaces – the well-known fact of $\sigma(\mathcal{M}(S), C_b(S))$-weakly sequentially completeness of $\mathcal{M}(S)$. It uses a type of argument that is of independent interest.

2.2 Preliminaries

We start with some preliminary results on Lipschitz functions on a metric space (S,d). We denote the vector space of all real-valued Lipschitz functions by $\mathrm{Lip}(S)$. The Lipschitz constant of $f \in \mathrm{Lip}(S)$ is

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

$\mathrm{BL}(S)$ is the subspace of bounded functions in $\mathrm{Lip}(S)$. It is a Banach space when equipped with the bounded Lipschitz or Dudley norm

$$\|f\|_{\mathrm{BL}} := \|f\|_\infty + |f|_L.$$

The norm $\|f\|_{\mathrm{FM}} := \max(\|f\|_\infty, |f|_L))$ is equivalent. $\mathrm{BL}(S)$ is partially ordered by pointwise ordering.

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

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

The norms on $\mathrm{BL}(S)^*$ dual to either $\|\cdot\|_{\mathrm{BL}}$ or $\|\cdot\|_{\mathrm{FM}}$ introduce equivalent norms on $\mathcal{M}(S)$ through the map $\mu \mapsto I_\mu$. These are called the bounded Lipschitz norm, or Dudley norm, and Fortet-Mourier norm on $\mathcal{M}(S)$, respectively. $\mathcal{M}(S)$ equipped with the norm topology induced by either of these norms is denoted by $\mathcal{M}(S)_{\mathrm{BL}}$. It is not complete generally. We write $\|\cdot\|_{\mathrm{TV}}$ for the total variation norm on $\mathcal{M}(S)$:

$$\|\mu\|_{\mathrm{TV}} = |\mu|(S) = \mu^+(S) + \mu^-(S),$$

where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ. $\mathcal{M}^+(S)$ is the convex cone of positive measures in $\mathcal{M}(S)$. One has

$$\|\mu\|_{\mathrm{TV}} = \|\mu\|^*_{\mathrm{BL}} = \|\mu\|^*_{\mathrm{FM}} \qquad \text{for all } \mu \in \mathcal{M}^+(S). \tag{2.1}$$

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

A finite signed Borel measure μ is *tight* if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset S$ such that $|\mu|(S \setminus K_\varepsilon) < \varepsilon$. A family $M \subset \mathcal{M}(S)$ is *tight* or *uniformly tight* if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset S$ such that $|\mu|(S \setminus K_\varepsilon) < \varepsilon$ for all $\mu \in M$. According to Prokhorov's Theorem (see [Bog07a], Theorem 8.6.2), if (S, d) is a complete separable metric space, a set of Borel probability measures $M \subset \mathcal{P}(S)$ is tight if and only if it is precompact in $\mathcal{P}(S)_{\mathrm{BL}}$. Completeness of S is an essential condition for this theorem to hold.

In a metric space (S, d), if $A \subset S$ is nonempty, we write

$$A^\varepsilon := \{x \in S : d(x, A) \le \varepsilon\}$$

for the closed ε-neighbourhood of A.

2.3 Main results

A fundamental result on the weak topology on signed measures induced by the pairing with $BL(S)$ is the following fundamental result that provides a '*weak-implies-strong-convergence*' property for this pairing on which we build our main results:

Theorem 2.3.1 (Schur-like property)**.** *Let (S,d) be a complete separable metric space. Let $(\mu_n) \subset \mathcal{M}(S)$ be such that $\sup_n \|\mu_n\|_{\mathrm{TV}} < \infty$. If for every $f \in \mathrm{BL}(S)$ the sequence $\langle \mu_n, f\rangle$ converges, then there exists $\mu \in \mathcal{M}(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$ as $n \to \infty$.*

A self-contained, delicate proof of this result is deferred to Section 2.4. The condition that the sequence of measures must be bounded in total variation norm cannot be omitted as the following counterexample indicates.

Counterexample 2.3.2. *Let $S = [0,1]$ with the Euclidean metric. Let $d\mu_n := n\sin(2\pi nx)\,dx$, where dx is Lebesgue measure on S. Then $\|\mu_n\|_{\mathrm{TV}}$ is unbounded. Let $g \in \mathrm{BL}(S)$ with $|g|_L \leq 1$. According to Rademacher's Theorem, g is differentiable Lebesgue almost everywhere. Since $|g|_L \leq 1$, there exists $f \in L^\infty([0,1])$ such that for all $0 \leq a < b \leq 1$,*

$$\int_a^b f(x)\,dx = g(b) - g(a).$$

This yields

$$\langle \mu_n, g\rangle = \frac{1}{2\pi}\int_0^1 \cos(2\pi nx) f(x)\,dx.$$

Since $f \in L^2([0,1])$, it follows from Bessel's Inequality that

$$\lim_{n\to\infty}\int_0^1 \cos(2\pi nx) f(x)\,dx = 0.$$

So $\langle \mu_n, g\rangle \to 0$ for all $g \in \mathrm{BL}(S)$. Now, let $g_n \in \mathrm{BL}(S)$ be the piecewise linear function that satisfies $g_n(0) = 0 = g_n(1)$,

$$g_n\left(\tfrac{1+4i}{4n}\right) = \tfrac{1}{4n}, \qquad g_n\left(\tfrac{3+4i}{4n}\right) = -\tfrac{1}{4n}, \qquad \text{for } i \in \mathbb{N},\ 0 \leq i \leq n-1.$$

Then $|g|_L = 1$ and $\|g_n\|_\infty = \frac{1}{4n}$. An easy calculation shows that $\langle \mu_n, g_n\rangle = \frac{1}{\pi^2}$ for all $n \in \mathbb{N}$. Therefore $\|\mu_n\|^*_{\mathrm{BL}}$ cannot converge to zero as $n \to \infty$.

Theorem 2.3.1 has the following corollary. Here we denote by $\mathrm{U}_b(S)$ the Banach space of uniformly continuous bounded functions on S, equipped with the $\|\cdot\|_\infty$-norm. This result was originally obtained by Pachl [Pac79], see also [Pac13], Theorem 5.45.

Corollary 2.3.3. *$\mathcal{M}(S)$ is $\mathrm{U}_b(S)$-weakly sequentially complete.*

Proof. Let $(\mu_n) \subset \mathcal{M}(S)$ be such that $\langle \mu_n, f\rangle$ is Cauchy for every $f \in \mathrm{U}_b(S)$. Then it follows from the Uniform Boundedness Principle that the sequence (μ_n) is bounded in $\mathrm{U}_b(S)^*$. Consequently, $\sup_n \|\mu_n\|_{\mathrm{TV}} = M < \infty$. Theorem 2.3.1 implies that there exists $\mu \in \mathcal{M}(S)$ such that $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for every $f \in \mathrm{BL}(S)$. Since $\mathrm{BL}(S)$ is dense in $\mathrm{U}_b(S)$ ([Dud66], Lemma 8) and $\|\mu_n\|_{\mathrm{TV}} \leq M$ for all n, the convergence result holds for every $f \in \mathrm{U}_b(S)$. □

Remark 2.3.1. Theorem 2.3.1 is related to results on asymptotic proximity of sequences of distributions, e.g. see [DR09], Theorem 4. In that setting $\mu_n = P_n - Q_n$, where P_n and Q_n are probability measures. These are asymptotically proximate (for the $\|\cdot\|^*_{\mathrm{BL}}$-norm; other norms are considered as well) if $\|P_n - Q_n\|^*_{\mathrm{BL}} \to 0$. So one knows in advance that $\langle \mu_n, f\rangle \to 0$. That is, the limit measure μ exists: $\mu = 0$. Combining such a result with the $\mathrm{U}_b(S)$-weak sequential completeness of $\mathcal{M}(S)$ implies Theorem 2.3.1. We present, in Section 2.4, an independent proof using completely different methods, that results in both the completeness result and a particular case of the mentioned asymptotic proximity result. The limit measure is there obtained through a delicate tightness argument, essentially.

The statement of the particular case in which all measures are positive seems novel too:

Theorem 2.3.4. *Let (S,d) be a complete separable metric space. Let $(\mu_n) \subset \mathcal{M}^+(S)$ be such that for every $f \in \mathrm{BL}(S)$, $\langle \mu_n, f\rangle$ converges. Then $\langle \mu_n, f\rangle$ converges for every $f \in C_b(S)$. In particular, there exists $\mu \in \mathcal{M}^+(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$.*

Its proof is simpler compared to that of Theorem 2.3.1. In Section 2.4 we shall present a self-contained proof of this result as well, based on a 'set-geometric' argument that is (essentially) also used to prove Theorem 2.3.1.

As it turned out, the proof for signed measures cannot be reduced straightforwardly to the result for positive measures. This is mainly caused by the complication, that for a sequence (μ_n) of signed measures such that $\langle \mu_n, f\rangle$ that is convergent for every $f \in \mathrm{BL}(S)$, it need not hold that $\langle \mu_n^+, f\rangle$ and $\langle \mu_n^-, f\rangle$ converge for every $f \in \mathrm{BL}(S)$. Take for example on $S = \mathbb{R}$ with the usual Euclidean metric $\mu_n := \delta_n - \delta_{n+\frac{1}{n}}$. Then $\langle \mu_n, f\rangle \to 0$ for every $f \in \mathrm{BL}(\mathbb{R})$. However, $\mu_n^+ = \delta_n$ and $\mu_n^- = \delta_{n+\frac{1}{n}}$, so $\langle \mu_n^\pm, f\rangle$ will not converge for every $f \in \mathrm{BL}(\mathbb{R})$. Thus, an immediate reduction to positive measures is not possible.

The pairing of measures with bounded Lipschitz functions is precisely what is important for the study of Markov operators and semigroups that have particular equicontinuity properties, as we shall discuss next.

2.3.1 Equicontinuous families of Markov operators

A *Markov operator* on (measures on) S is a map $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ such that:

1. $P(\mu + \nu) = P\mu + P\nu$ and $P(r\mu) = rP\mu$ for all $\mu, \nu \in \mathcal{M}^+(S)$ and $r \geq 0$,
2. $(P\mu)(S) = \mu(S)$ for all $\mu \in \mathcal{M}^+(S)$.

In particular, a Markov operator leaves invariant the convex set $\mathcal{P}(S)$ of probability measures in $\mathcal{M}^+(S)$. Let $\mathrm{BM}(S)$ be the vector space of bounded Borel measurable real-valued functions on S. A Markov operator is called *regular* if there exists a linear map $U : \mathrm{BM}(S) \to \mathrm{BM}(S)$, the *dual operator*, such that

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

A regular Markov operator P is *Feller* if its dual operator maps $C_b(S)$ into itself. Equivalently, P is continuous for the $\|\cdot\|^*_{\mathrm{BL}}$-norm topology (cf. e.g. [HW09a] Lemma 3.3 and [Wor10] Lemma 3.3.2).

Regular Markov operators on measures appear naturally e.g. in the theory of Iterated Function Systems [BDEG88, LY94] and the study of deterministic flows by their lift to measures [PT11, EHM15]. Dual Markov operators on $C_b(S)$ (or a suitable linear subspace) are encountered naturally in the study of stochastic differential equations [DPZ14, KPS10]. Which specific viewpoint in this duality is used, is often determined by technical considerations and the mathematical problems that are considered.

Markov operators and semigroups with equicontinuity properties (called the '*e-property*') have convenient properties concerning existence, uniqueness and asymptotic stability of invariant measures, see e.g. [HHS16, KPS10, Sza10, SW12, Wor10]. After having defined these properties precisely below, we show by means of Theorem 2.3.1 that a dual viewpoint exists for equicontinuity too, in Theorem 2.3.5. In subsequent work further consequences of this result for the theory and application of equicontinuous families of Markov operators are examined. Some results in this direction were also discussed in parts of [Wor10], Chapter 7.

Let T be a topological space and (S', d') a metric space. A family of functions $\mathcal{E} \subset C(T, S')$ is *equicontinuous at* $t_0 \in T$ if for every $\varepsilon > 0$ there exists an open neighbourhood U_ε of t_0

such that
$$d'(f(t), f(t_0)) < \varepsilon \quad \text{for all } f \in \mathcal{E},\ t \in U_\varepsilon.$$

$\mathcal{E}$ is *equicontinuous* if it is equicontinuous at every point of T.

Following Szarek *et al.* [KPS10, Sza10], a family $(P_\lambda)_{\lambda\in\Lambda}$ of regular Markov operators has the *e-property* if for each $f \in \mathrm{BL}(S)$ the family $\{U_\lambda f : \lambda \in \Lambda\}$ is equicontinuous in $C_b(S)$. In particular one may consider the family of iterates of a single Markov operator P: $(P^n)_{n\in\mathbb{N}}$, or Markov semigroups $(P_t)_{t\in\mathbb{R}^+}$, where each P_t is a regular Markov operator and $P_0 = I$, $P_t P_s = P_{t+s}$.

Our main result on equicontinuous families of Markov operators is

Theorem 2.3.5. *Let $\{P_\lambda : \lambda \in \Lambda\}$ be a family of regular Markov operators on a complete separable metric space (S, d). Let U_λ be the dual Markov operator of P_λ. The following statements are equivalent:*

1. *$\{U_\lambda f : \lambda \in \Lambda\}$ is equicontinuous in $C_b(S)$ for every $f \in \mathrm{BL}(S)$.*
2. *$\{P_\lambda : \lambda \in \Lambda\}$ is equicontinuous in $C(\mathcal{M}^+(S)_{\mathrm{BL}}, \mathcal{M}^+(S)_{\mathrm{BL}})$,*
3. *$\{P_\lambda : \lambda \in \Lambda\}$ is equicontinuous in $C(\mathcal{P}(S)_{\mathrm{weak}}, \mathcal{P}(S)_{\mathrm{BL}})$*

Proof. (i) ⇒ (ii). Assume on the contrary that $\{P_\lambda : \lambda \in \lambda\}$ is not an equicontinuous family of maps. Then there exists a point $\mu_0 \in \mathcal{M}^+(S)$ at which this family is not equicontinuous. Hence there exists $\varepsilon_0 > 0$ such that for every $k \in \mathbb{N}$ there are $\lambda_k \in \Lambda$ and $\mu_k \in \mathcal{M}^+(S)$ such that
$$\|\mu_k - \mu_0\|_{\mathrm{BL}}^* < \tfrac{1}{k} \quad \text{and} \quad \|P_{\lambda_k}\mu_k - P_{\lambda_k}\mu_0\|_{\mathrm{BL}}^* \geq \varepsilon_0 \qquad \text{for all } k \in \mathbb{N}. \tag{2.2}$$

Because the measures μ_k are positive and the $\|\cdot\|_{\mathrm{BL}}^*$-norm metrizes the $C_b(S)$-weak topology on $\mathcal{M}^+(S)$ (cf. [Dud66], Theorem 18), $\langle \mu_k, f\rangle \to \langle \mu_0, f\rangle$ for every $f \in C_b(S)$. According to [Dud66], Theorem 7, this convergence is uniform on any equicontinuous and uniformly bounded subset $\mathcal{E}$ of $C_b(S)$. By assumption, $\mathcal{M}_f := \{U_{\lambda_k} f : k \in \mathbb{N}\}$ is such a family for every $f \in \mathrm{BL}(S)$. Therefore

$$|\langle P_{\lambda_k}\mu_k - P_{\lambda_k}\mu_0, f\rangle| = |\langle \mu_k - \mu_0, U_{\lambda_k} f\rangle| \to 0 \tag{2.3}$$

as $k \to \infty$ for every $f \in \mathrm{BL}(S)$. Since for positive measures μ one has $\|\mu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{BL}}^*$, one obtains
$$\big|\|\mu_k\|_{\mathrm{TV}} - \|\mu_0\|_{\mathrm{TV}}\big| \leq \|\mu_k - \mu_0\|_{\mathrm{BL}}^* \to 0.$$

So $m_0 := \sup_{k\geq 1} \|\mu_k\|_{\mathrm{TV}} < \infty$. Moreover,

$$\|P_{\lambda_k}\mu_k - P_{\lambda_k}\mu_0\|_{\mathrm{TV}} \leq \|P_{\lambda_k}\mu_k\|_{\mathrm{TV}} + \|P_{\lambda_k}\mu_0\|_{\mathrm{TV}} \leq \|\mu_k\|_{\mathrm{TV}} + \|\mu_0\|_{\mathrm{TV}} \leq m_0 + \|\mu_0\|_{\mathrm{TV}}.$$

Theorem 2.3.1 and (2.3) yields that $\|P_{\lambda_k}\mu_k - P_{\lambda_k}\mu_0\|_{\mathrm{BL}}^* \to 0$ as $k \to \infty$. This contradicts the second property in (2.2).

(ii) ⇒ (iii). Follows immediately by restriction of the Markov operators P_Λ to $\mathcal{P}(S)$.

(iii) ⇒ (i). Let $f \in \mathrm{BL}(S)$ and $x_0 \in S$. Let $\varepsilon > 0$. Since $\{P_\lambda : \lambda \in \Lambda\}$ is equicontinuous at δ_{x_0} there exists an open neighbourhood V of δ_{x_0} in $\mathcal{P}(S)_{\mathrm{weak}}$ such that

$$\|P_\lambda\delta_{x_0} - P_\lambda\mu\|_{\mathrm{BL}}^* < \varepsilon/(1 + \|f\|_{\mathrm{BL}}) \quad \text{for all } \lambda \in \Lambda \text{ and } \mu \in U_0.$$

Since the map $x \mapsto \delta_x : S \to \mathcal{P}(S)_{\mathrm{weak}}$ is continuous, there exists an open neighbourhood V_0 of x_0 in S such that $\delta_x \in V$ for all $x \in V_0$. Then

$$|U_\lambda f(x) - U_\lambda f(x_0)| = |\langle P_\lambda\delta_x - P_\lambda\delta_{x_0}, f\rangle| \leq \frac{\varepsilon}{1 + \|f\|_{\mathrm{BL}}} \cdot \|f\|_{\mathrm{BL}} < \varepsilon$$

for all $x \in V_0$ and $\lambda \in \Lambda$. □

A particular class of examples of Markov operators and semigroups is furnished by the lift of a map or semigroup $(\phi_t)_{t\geq 0}$ of measurable maps $\phi_t : S \to S$ to measures on S by means of push-forward:

$$P_t^\phi\mu(E) := \mu\big(\phi_t^{-1}(E)\big)$$

for every Borel set E of S and $\mu \in \mathcal{M}^+(S)$. A consequence of Theorem 2.3.5 is:

Proposition 2.3.6. *Let (S, d) be a complete separable metric space and let $(\phi_t)_{t\geq 0}$ be a semigroup of Borel measurable transformations of S. Then P_t^ϕ is a regular Markov operator for each $t \geq 0$. Moreover, $(P_t^\phi)_{t\geq 0}$ is equicontinuous in $C(\mathcal{M}^+(S)_{\mathrm{BL}}, \mathcal{M}^+(S)_{\mathrm{BL}})$ if and only if $(\phi_t)_{t\geq 0}$ is equicontinuous in $C(S, S)$.*

Proof. The regularity of P_t^ϕ is immediate, as $U_t^\phi f = f \circ \phi_t$.

'⇒': Let $x_0 \in S$ and $\varepsilon > 0$. Define $h(x) := 2x/(2 + x)$ and put $\varepsilon' := h(\varepsilon)$. By equicontinuity of $(P_t^\phi)_{t\geq 0}$ at δ_{x_0}, there exists and open neighbourhood U of δ_{x_0} in $\mathcal{M}^+(S)_{\mathrm{BL}}$ such that

$$\|P_t^\phi\mu - P_t^\phi\delta_{x_0}\|_{\mathrm{BL}}^* < \varepsilon'$$

for all $t \geq 0$ and $\mu \in U$. Because the map $\delta : x \mapsto \delta_x : S \to \mathcal{M}^+(S)_{\mathrm{BL}}$ is continuous,

$U_0 := \delta^{-1}(U)$ is open in S. It contains x_0. Moreover,

$$\|P_t^\phi \delta_x - P_t^\phi \delta_{x_0}\|_{\mathrm{BL}}^* = \|\delta_{\phi_t(x)} - \delta_{\phi_t(x_0)}\|_{\mathrm{BL}}^* = h\big(d(\phi_t(x), \phi_t(x_0)\big) < \varepsilon'$$

for all $x \in U_0$ and $t \geq 0$ (see [HW09b] Lemma 3.5). Because h is monotone increasing,

$$d\big(\phi_t(x), \phi_t(x_0)\big) < \varepsilon \qquad \text{for all } x \in U_0,\ t \geq 0.$$

'$\Leftarrow$': This part involves Theorem 2.3.5. Let $f \in \mathrm{BL}(S)$. Let U_t be the dual operator of P_t. Then for all $x, x_0 \in S$,

$$|U_t f(x) - U_t f(x_0)| = |f(\phi_t(x)) - f(\phi_t(x_0))| \leq |f|_L d(\phi_t(x), \phi_t(x_0)),$$

from which the equicontinuity of $\{U_t f : t \geq 0\}$ follows. The result is obtained by applying Theorem 2.3.5. □

2.3.2 Coincidence of weak and norm topologies

A further consequence of Theorem 2.3.1 is

Theorem 2.3.7. *Let (S,d) be a complete separable metric space and let $M \subset \mathcal{M}(S)$ be such that $m := \sup_{\mu \in M} \|\mu\|_{\mathrm{TV}} < \infty$. If the restriction of the $\sigma(\mathcal{M}(S), \mathrm{BL}(S))$-weak topology to M is first countable, then this topology coincides with the restriction of the $\|\cdot\|_{\mathrm{BL}}^*$-norm topology to M.*

Proof. We have to show that for any $\|\cdot\|_{\mathrm{BL}}^*$-norm closed set C, $C \cap M$ is closed in the restriction of the $\sigma(\mathcal{M}(S), \mathrm{BL}(S))$-weak topology to M. Since the latter is first countable, $C \cap M$ is relatively $\sigma(\mathcal{M}(S), \mathrm{BL}(S))$-weak closed if and only if for every $\sigma(\mathcal{M}(S), \mathrm{BL}(S))$-weakly converging *sequence* $\mu_n \to \mu$ in $\mathcal{M}(S)$ with $\mu_n \in C$, one has $\mu \in C$ (cf. [Kel55] Theorem 2.8, p. 72). Let (μ_n) be such a sequence. Because $\sup_{\mu \in M} \|\mu\|_{\mathrm{TV}} < \infty$ by assumption, Theorem 2.3.1 implies that there exists $\mu' \in \mathcal{M}(S)$ such that $\|\mu_n - \mu'\|_{\mathrm{BL}}^* \to 0$. Since C is relatively $\|\cdot\|_{\mathrm{BL}}^*$-norm closed in M, $\mu' \in C$. Moreover, $\langle \mu, f \rangle = \langle \mu', f \rangle$ for every $f \in \mathrm{BL}(S)$, so $\mu = \mu' \in C$. □

The following technical result provides a tractable condition that ensures first countability of the relative weak topology on the set M, as we shall show after having proven the result. We need to introduce some notation. For $\lambda > 0$ and $C \subset S$ closed and nonempty, define

$$h_{\lambda,C}(x) := \left[1 - \tfrac{1}{\lambda} d(x, C)\right]^+.$$

Then $h_{\lambda,C} \in \mathrm{BL}(S)$, $|h_{\lambda,C}|_L = \frac{1}{\lambda}$, $0 \le h_{\lambda,C} \le 1$ and $h_{\lambda,C} \downarrow \mathbb{1}_C$ pointwise as $\lambda \downarrow 0$. Moreover $h_{\lambda,C} = 0$ on $S \setminus C^\lambda$. We can now state the result.

Lemma 2.3.8. *Let $M \subset \mathcal{M}(S)$ be such that $m := \sup_{\mu \in M} \|\mu\|_{\mathrm{TV}} < \infty$. If for every $\mu \in M$ and every $\varepsilon > 0$ there exist $K_1, \dots, K_n \subset S$ compact such that for $K = \bigcup_{i=1}^n K_i$:*

1. *$|\mu|(S \setminus K) < \varepsilon$,*
2. *There exists $0 < \lambda_0 \le \varepsilon$ such that for all $0 < \lambda \le \lambda_0$ there exists $\delta_1, \dots, \delta_n > 0$ such that the following statement holds:*

   If $\nu \in M$ satisfies $|\langle \mu - \nu, h_{\lambda,K_i} \rangle| < \delta_i$ for all $i = 1, \dots, n$,
   then $|\nu|(S \setminus K^\lambda) < \varepsilon$.

Then the relative $\sigma(\mathcal{M}(S), \mathrm{BL}(S))$-weak topology on M is first countable.

Proof. We first define a countable family $\mathcal{F}$ of functions in $\bar{B} := \{g \in \mathrm{BL}(S) : \|g\|_\infty \le 1\}$ that is dense in $\bar{B}$ for the compact-open topology, i.e. the topology of uniform convergence on compact subsets of S. Let D be a countable dense subset of S. The family of *finite* subsets of D is countable. Let $I_{\mathbb{Q}} := \mathbb{Q} \cap [0,1]$. For a finite subset $F \subset D$, $\lambda \in I_{\mathbb{Q}} \setminus \{0\}$ and function $a : F \to I_{\mathbb{Q}}$ define

$$f_{F,a}^\lambda(x) := \bigvee_{y \in F} \big[a(y)(1 - \tfrac{1}{\lambda} d(x,y))^+ \big].$$

Here $\vee$ denotes the maximum, as before. Then $f_{F,a}^\lambda \in \mathrm{BL}(S)$, $|f_{F,a}^\lambda|_L \le \max_{y \in F} \frac{a(y)}{\lambda} \le \frac{1}{\lambda}$. Moreover, $f_{F,a}^\lambda$ vanishes outside $F^\lambda = \bigcup_{y \in F} B(y, \lambda)$. For a finite subset $F \subset D$ the family $\mathcal{F}_F$ of all such functions $f_{F,a}^\lambda$ with a and λ as indicated is countable. So the union $\mathcal{F}^+$ of all sets $\mathcal{F}_F$ over all finite $F \subset D$ is countable too. It is quickly verified that on any compact subset K of S any positive $h \in \bar{B}$ can be uniformly approximated by $f \in \mathcal{F}^+$. Consequently, $\mathcal{F} = \mathcal{F}^+ - \mathcal{F}^+ \subset \mathrm{BL}(S)$ is countable and any $h \in \bar{B}$ can be approximated uniformly on compact sets by means of $f \in \mathcal{F}$.

Now let $\mu \in M$ and consider the open neighbourhood

$$U_\mu(h, r) := \big\{ \nu \in M : |\langle \mu - \nu, h \rangle| < r \big\},$$

with $r > 0$ and $h \in \mathrm{BL}(S)$. Without loss of generality we can assume that $\|h\|_{\mathrm{BL}} = 1$. We shall prove that there exist $f_0,\dots,f_n \in \mathcal{F}$ and $q_0,\dots,q_n > 0$ in $\mathbb{Q}$ such that

$$\bigcap_{i=0}^{n} \left\{\nu \in M : \left|\langle \mu-\nu, h_i\rangle\right| < q_i\right\} \subset U_\mu(h,r). \tag{2.4}$$

Then the relative weak topology on M is first countable.

Let $\varepsilon \in \mathbb{Q}$ such that $0 < \varepsilon \le \frac{1}{6}r$ and let $K_i, K \subset S$ be compact and $0 < \lambda_0 \le \varepsilon$ as in the conditions of the lemma. There exists $f_0 \in \mathcal{F}$ such that $\sup_{x\in K} |h(x) - f_0(x)| \le \frac{1}{4m}\varepsilon$. Then for any $0 < \lambda \le \lambda_0$, $x \in K^\lambda$ and $x_0 \in K$,

$$\begin{aligned}|h(x) - f_0(x)| &\le |h(x) - h(x_0)| \;+\; |h(x_0) - f_0(x_0)| \;+\; |f_0(x_0) - f_0(x)|\\ &\le (1 + |f_0|_L) d(x, x_0) + \tfrac{1}{4m}\varepsilon.\end{aligned}$$

Hence

$$\sup_{x\in K^\lambda} |h(x) - f_0(x)| \le (1 + |f_0|_L)\lambda + \tfrac{1}{4m}\varepsilon.$$

Let $0 < \lambda_0' \le \lambda_0$ be such that $(1 + |f_0|_L)\lambda_0' \le \frac{1}{4m}\varepsilon$. Now one has, using property (i),

$$\begin{aligned}|\langle\mu-\nu,h\rangle| &\le |\langle\mu-\nu,h-f_0\rangle| + |\langle\mu-\nu,f_0\rangle|\\ &\le \int_{K^\lambda} |h-f_0|\, d|\mu-\nu| \;+\; 2|\mu|(S\smallsetminus K^\lambda) + 2|\nu|(S\smallsetminus K^\lambda) + |\langle\mu-\nu,f_0\rangle|\\ &\le \tfrac{1}{2m}\varepsilon\cdot 2m + 2\varepsilon + 2|\nu|(S\smallsetminus K^\lambda) + |\langle\mu-\nu,f_0\rangle| \end{aligned} \tag{2.5}$$

for all $0 < \lambda \le \lambda_0'$. Fix $\lambda \in \mathbb{Q}$ with $0 < \lambda \le \lambda_0'$ and let $\delta_1,\dots,\delta_n$ be as in property (ii).

The Hausdorff semidistance on closed and bounded subsets of S is given by

$$\delta(C,C') := \sup_{x\in C} d(x,C').$$

The Hausdorff distance is defined by

$$d_H(C,C') := \max\bigl(\delta(C,C'),\delta(C',C)\bigr).$$

The collection of finite subsets of D form a separable dense subset of the set of compact subsets of S, $\mathcal{K}(S)$, for d_H. If $F \subset D$ is finite and $K' \in \mathcal{K}(S)$, then by the Birkhoff

Inequalities

$$\begin{aligned}|h_{\lambda,K'} - h_{\lambda,F}| &= \left|\left[1 - \tfrac{1}{\lambda}d(x,K')\right]^+ - \left[1 - \tfrac{1}{\lambda}d(x,F)\right]^+\right| \\ &\leq \left|\left[1 - \tfrac{1}{\lambda}d(x,K')\right] - \left[1 - \tfrac{1}{\lambda}d(x,F)\right]\right| \\ &= \tfrac{1}{\lambda}|d(x,K') - d(x,F)| \leq \tfrac{1}{\lambda}\cdot d_H(K',F).\end{aligned}$$

Let $F_i \subset D$ be finite such that $d_H(K_i,F_i) \leq \frac{1}{4m}\lambda\delta_i$. Then $h_{\lambda,F_i} = f^{\lambda}_{F_i,\mathbb{1}} \in \mathcal{F}$. Put $f_i := h_{\lambda,F_i}$. Let $q_i \in \mathbb{Q}$ be such that $0 < q_i < \frac{1}{2}\delta_i$. If $\nu \in M$ is such that $|\langle \mu - \nu, f_i\rangle| < q_i$ for $i = 1,\dots,n$, then

$$|\langle \mu - \nu, h_{\lambda,K_i}\rangle| \leq \|h_{\lambda,K_i} - h_{\lambda,F_i}\|_\infty \cdot \|\mu - \nu\|_{\mathrm{TV}} + |\langle \mu - \nu, f_i\rangle| < \tfrac{1}{2}\delta_i + \tfrac{1}{2}\delta_i = \delta_i$$

According to condition (ii) one has $|\nu|(S \smallsetminus K^\lambda) < \varepsilon$. Put $q_0 = \varepsilon$. Inequality (2.5) then yields (2.4), as desired. □

Because conditions (i) and (ii) in Lemma 2.3.8 are immediately satisfied when M is uniformly tight, we obtain

Corollary 2.3.9. *Let (S,d) be a complete separable metric space and let $M \subset \mathcal{M}(S)$ such that $\sup_{\mu\in M}\|\mu\|_{\mathrm{TV}} < \infty$ and M is uniformly tight. Then the $\sigma(\mathcal{M}(S),\mathrm{BL}(S))$-weak topology coincides with the $\|\cdot\|^*_{\mathrm{BL}}$-norm topology on M.*

Remark 2.3.2. Gwiazda *et al.* [GLMC10] state at p. 2708 that the topology of narrow convergence in $\mathcal{M}(S)$, i.e. that of convergence of sequences of signed measures paired with $f \in C_b(S)$, is metrizable on tight subsets that are uniformly bounded in total variation norm. In fact it can be metrized by the norm $\|\cdot\|^*_{\mathrm{BL}}$.

A second case, more involved, in which the conditions of Lemma 2.3.8 are satisfied, is:

Proposition 2.3.10. *Let (S,d) be a complete separable metric space and let*

$$M := \{\mu \in \mathcal{M}(S) : \|\mu\|_{\mathrm{TV}} = \rho\}, \quad (\rho > 0).$$

Then condition (i) and (ii) of Lemma 2.3.8 hold. In particular, the relative $\sigma(\mathcal{M}(S),\mathrm{BL}(S))$-weak topology and relative $\|\cdot\|^*_{\mathrm{BL}}$-norm topology on M coincide.

Proof. Take $\varepsilon > 0$, $\mu \in M$ and let μ^+ and μ^- be the positive and negative part of μ, i.e. $\mu = \mu^+ - \mu^-$. Since $\mu^\pm$ are disjoint and tight, by Ulam's Lemma, there exist compact sets $K_\pm \subset S$ such that $K_+ \cap K_- = \varnothing$, $\mu^\pm(K_\mp) = 0$ and

$$\mu^+(S) - \mu^+(K_+) < \varepsilon/8 \quad \text{and} \quad \mu^-(S) - \mu^-(K_-) < \varepsilon/8. \tag{2.6}$$

In particular,

$$|\mu|(S \setminus (K_+ \cup K_-)) \leq \mu^+(S \setminus K_+) + \mu^-(S \setminus K_-) < \tfrac{1}{8}\varepsilon + \tfrac{1}{8}\varepsilon < \varepsilon,$$

so condition (i) of Lemma 2.3.8 is satisfied for $K = K_+ \cup K_-$.

Because K_+ and K_- are compact, there exists $\lambda_0 > 0$ such that $K_+^{\lambda_0} \cap K_-^{\lambda_0} = \varnothing$. Then $K_+^{\lambda} \cap K_-^{\lambda} = \varnothing$ for all $0 < \lambda \leq \lambda_0$. Without loss of generality we can assume that $\lambda_0 \leq \varepsilon$. Fix $0 < \lambda \leq \lambda_0$.

Let us assume for the moment that $\delta_\pm > 0$ have been selected. At the end we will then see how to choose these, such that condition (ii) will be satisfied. If $\nu \in M$ satisfies

$$|\langle \mu - \nu, h_{\lambda,K_+}\rangle| < \delta_+ \quad \text{and} \quad |\langle \mu - \nu, h_{\lambda,K_-}\rangle| < \delta_-, \tag{2.7}$$

then

$$\langle \mu - \nu^+, h_{\lambda,K_+}\rangle \leq \langle \mu - \nu^+ + \nu^-, h_{\lambda,K_+}\rangle \leq |\langle \mu - \nu, h_{\lambda,K_+}\rangle| < \delta_+.$$

Consequently, since $\mathbb{1}_{K_+} \leq h_{\lambda,K_+} \leq \mathbb{1}_{K_+^\lambda}$,

$$\mu^+(K_+) - \mu^-(K_+^\lambda) - \nu^+(K_+^\lambda) \leq \langle \mu - \nu^+, h_{\lambda,K_+}\rangle < \delta_+.$$

We obtain

$$\begin{aligned}\nu^+(K_+^\lambda) &> \mu^+(K_+) - \mu^-(K_+^\lambda) - \delta_+ \;\geq\; \mu^+(K_+) - \mu^-(S \setminus K_-) - \delta_+\\ &> \mu^+(K_+) - \tfrac{1}{8}\varepsilon - \delta_+.\end{aligned}$$

In a similar way,

$$\langle -\mu - \nu^-, h_{\lambda,K_-}\rangle \leq \langle \nu - \mu, h_{\lambda,K_-}\rangle < \delta_-,$$

whence

$$\nu^-(K_-^\lambda) > \mu^-(K_-) - \tfrac{1}{8}\varepsilon - \delta_-.$$

Therefore, using (2.6),

$$\begin{aligned}\nu^+(K_+^\lambda) + \nu^-(K_-^\lambda) &> \mu^+(K_+) + \mu^-(K_-) - \tfrac{1}{4}\varepsilon - (\delta_+ + \delta_-)\\ &> \mu^+(S) + \mu^-(S) - \tfrac{1}{2}\varepsilon - (\delta_+ + \delta_-) = \rho - (\delta_+ + \delta_- + \tfrac{1}{2}\varepsilon).\end{aligned}$$

Note that in this last step the assumption that M is a total variation sphere is used in an essential manner. The last inequality implies that

$$|\nu|(S \smallsetminus K^\lambda) = |\nu|(S) - |\nu|(K_+^\lambda) - |\nu|(K_-^\lambda) \leq \rho - \nu^+(K_+^\lambda) - \nu^-(K_-^\lambda) < \delta_+ + \delta_- + \tfrac{1}{2}\varepsilon.$$

Thus, if we take $K_1 = K_+$, $K_2 = K_-$, $\delta_+ = \delta_- = \delta_i = \frac{1}{4}\varepsilon$, we see that condition (ii) in Lemma 2.3.8 is satisfied. Theorem 2.3.7 then yields the final statement. □

Remark 2.3.3. 1.) In [Pac13], Theorem 5.38 and Corollary 5.39 come close to Theorem 2.3.7. A technical condition seems to prevent deriving our new result on coincidence of topologies from the results in [Pac13].
2.) The result stated in Proposition 2.3.10 can be found in [Pac13], Corollary 5.39. There, a proof of this result is provided using completely different techniques. Concerning coincidence of these topologies on total variation spheres, see some further notes in [Pac13], indicating e.g. [GL81].

In view of Corollary 2.3.9 and Proposition 2.3.10 one might be tempted to conjecture that the weak and norm topologies would coincide on sets of measures with uniformly bounded total variation. This does not hold however, as the following counterexample illustrates.

Counterexample 2.3.11. *Let* (S,d) *be the natural numbers* $\mathbb{N}$ *equipped with the restriction of the Euclidean metric on* $\mathbb{R}$*. Now,* $\mathrm{BL}(\mathbb{N})$ *is linearly isomorphic to* ℓ^∞*: the map* $f \mapsto (f(n))_{n\in\mathbb{N}}$ *is bijective and continuous. Hence it is a linear isomorphism by Banach's Isomorphism Theorem. Observe that* $|f|_L \leq 2\|f\|_\infty$*. Since* $(\mathbb{N},d)$ *is uniformly discrete, the norms* $\|\cdot\|_{\mathrm{BL}}^*$ *and* $\|\cdot\|_{\mathrm{TV}}$ *on* $\mathcal{M}(\mathbb{N})$ *are equivalent (cf. [HW09b], proof of Theorem 3.11). So* $\mathcal{M}(\mathbb{N})_{\mathrm{BL}}$ *is linearly isomorphic to* ℓ^1 *under the map* $\mu \mapsto (\mu(\{n\}))_{n\in\mathbb{N}}$*. One has* $\|\mu\|_{\mathrm{TV}} = \|(\mu)\|_{\ell^1}$*. Moreover, the duality between* $\mathcal{M}(\mathbb{N})$ *and* $\mathrm{BL}(\mathbb{N})$ *is precisely the duality between* ℓ^1 *and* ℓ^∞ *under the given isomorphisms. Consider now* $M := \{(\mu) \in \ell^1 : \|(\mu)\|_{\ell^1} \leq 1\}$*. It represents a set of measures that is uniformly bounded in total variation norm. Let* $S := \{(\mu) \in \ell^1 : \|(\mu)\|_{\ell^1} = 1\}$*. Then* S *is a* $\|\cdot\|_{\mathrm{TV}}$*-closed subset of* M*. The weak closure of* S *equals* M *however (cf. [Con85], Section V.1, Ex. 10). Therefore, the* $\|\cdot\|_{\mathrm{BL}}^*$ *(i.e.* $\|\cdot\|_{\mathrm{TV}}$*) and weak topologies cannot coincide on* M*.*

2.4 Proof of the Schur-like property

We provide a self-contained proof of the Schur-like property for spaces of measures, Theorem 2.3.1, using a 'set-geometric' argument. See Remark 2.4.2 below for alternative approaches.

We first introduce various technical lemmas that enable our set-geometric argument. Then we start with a complete proof of the particular case of positive measures, Theorem 2.3.1, as it will aid the reader in getting introduced to the type of argument employed, based on Lemma 2.4.3, and the complications that arise when proving the result for general signed measures in the section that follows.

2.4.1 Technical lemmas

The following lemmas are needed in the proof of the fundamental result.

Lemma 2.4.1. *Let $A \subset \mathrm{BL}(S)$ be such that $\sup_{f\in A}\|f\|_{\mathrm{BL}} < \infty$. Then $\sup(A)$ exists in $\mathrm{BL}(S)$ and $|\sup(A)|_L \le \sup_{f\in A}|f|_L$. In particular, $\|\sup(A)\|_{\mathrm{BL}} \le 2\sup_{f\in A}\|f\|_{\mathrm{BL}}$.*

Proof. Put $L := \sup_{f\in A}|f|_L$ and let $g = \sup(A)$, i.e. $g(x) := \sup\{f(x) : f \in A\}$ for every $x \in S$. Let $x, y \in S$. We may assume $g(x) \ge g(y)$. Let $\varepsilon > 0$. There exists $f \in A$ such that $g(x) < f(x) + \varepsilon$. By definition $g(y) \ge f(y)$. Hence

$$|g(x) - g(y)| \le g(x) - f(x) \ + \ f(x) - f(y) < \varepsilon + |f(x) - f(y)| \le \varepsilon + L\, d(x,y).$$

Since ε is arbitrary, we obtain that $|g(x) - g(y)| \le L d(x,y)$. Thus $g \in \mathrm{Lip}(S)$ and $|g|_L \le L$. Clearly, $\|g\|_\infty \le \sup_{f\in A}\|f\|_\infty < \infty$, so $g \in \mathrm{BL}(S)$ and $\|g\|_{\mathrm{BL}} \le 2\sup_{f\in A}\|f\|_{\mathrm{BL}}$. □

The *support* of $f \in C(S)$, denoted by $\operatorname{supp} f$, is the closure of the set of points where f is nonzero. Lemma 2.4.1 implies the following

Lemma 2.4.2. *Let $(f_k) \subset \mathrm{BL}(S)$ be such that $\sup_{k\ge1}\|f_k\|_{\mathrm{BL}} < \infty$. Assume that their supports are pairwise disjoint. Then the series $f(x) := \sum_{k=1}^\infty f_k(x)$ converges pointwise and $f \in \mathrm{BL}(S)$. In particular,*

$$\|f\|_\infty \le \sup_{k\ge1}\|f_k\|_\infty, \qquad |f|_L \le 2\sup_{k\ge1}|f_k|_L. \tag{2.8}$$

Proof. Because the sets $\operatorname{supp} f_k$ are pairwise disjoint, $f(x) = f_k(x)$ if $x \in \operatorname{supp} f_k$. So the positive part f^+ and negative part f^- of f satisfy $f^\pm = \sum_{k=1}^\infty f_k^\pm$ and it suffices to prove the result for $f \ge 0$. In that case, $f = \sup_{k\ge1} f_k$, and the first estimate in (2.8) follows immediately. The second follows from Lemma 2.4.1. □

Lemma 2.4.3. *Let (S,d) be a complete separable metric space. Let $\mu_n \in \mathcal{M}^+(S)$, $n \in \mathbb{N}$. Assume that $\{\mu_n : n \ge 1\}$ is not tight. Then there exists $\varepsilon > 0$, an increasing sequence (n_k)*

of positive integers and a sequence of compact sets (K_{n_k}) *such that*

$$\mu_{n_k}\left(K_{n_k}\right) \geq \varepsilon \qquad \textit{for all } k \geq 1$$

and

$$\operatorname{dist}(K_{n_k}, K_{n_m}) := \min\{d(x,y) \mid x \in K_{n_k},\ y \in K_{n_m}\} > \varepsilon \qquad \textit{for all } k \neq m.$$

This result was originally stated in [KPS10], Lemma 1, p. 1410, for a sequence (μ_n) of probability Borel measures with a proof in [LS06] (proof of Theorem 3.1, p. 517-518), but it is also valid for (positive) measures.

In addition to Lemma 2.4.3 the following observation is made:

Lemma 2.4.4. *Let* $(\mu_n) \subset \mathcal{M}^+(S)$ *be such that* $\sup_n \mu_n(S) < \infty$ *and let* (E_n) *be a sequence of pairwise disjoint Borel measurable subsets of* S*. Then for every* $\varepsilon > 0$ *there exists a strictly increasing subsequence* (n_i) *of* $\mathbb{N}$ *such that for every* $i \geq 1$*,*

$$\mu_{n_i}\left(\bigcup_{j \neq i} E_{n_j}\right) < \varepsilon. \tag{2.9}$$

Proof. Let us first prove that for every $\eta > 0$ there exists a strictly increasing subsequence (m_i) such that

$$\mu_{m_1}\left(\bigcup_{i>1} E_{m_i}\right) < \eta \tag{2.10}$$

and

$$\mu_{m_i}(E_{m_1}) < \eta \qquad \text{for all } i \geq 2. \tag{2.11}$$

Fix $\eta > 0$. Set $C := \sup_n \mu_n(S)$ and let $N \geq 1$ be such that $N\eta > C$. Since for every $n \geq 1$ we have $\sum_{m=1}^{N} \mu_n(E_m) = \mu_n\left(\bigcup_{m=1}^{N} E_m\right) \leq \mu_n(S) \leq C < N\eta$, there exists $m \in \{1, \ldots, N\}$ such that

$$\mu_n(E_m) < \eta. \tag{2.12}$$

Thus there exists $m_1 \in \{1, \ldots, N\}$ and an infinite set $\mathcal{S}$ such that condition (2.12) holds for all $n \in \mathcal{S}$. Let us split $\mathcal{S}$ into N disjoint infinite subsets $\mathcal{S}_1, \ldots, \mathcal{S}_N$.

Since

$$\bigcup_{n \in \mathcal{S}_i} E_n \cap \bigcup_{n \in \mathcal{S}_j} E_n = \varnothing \qquad \text{for } i, j \in \{1, \ldots, N\},\ i \neq j,$$

we have

$$\sum_{i=1}^{N}\mu_{m_1}\left(\bigcup_{n\in\mathcal{S}_i}E_n\right)=\mu_{m_1}\left(\bigcup_{i=1}^{N}\bigcup_{n\in\mathcal{S}_i}E_n\right)=\mu_{m_1}\left(\bigcup_{n\in\mathcal{S}}E_n\right)\le\mu_{m_1}(S)\le C<N\eta,$$

which, in turn, yields

$$\mu_{m_1}\left(\bigcup_{n\in\mathcal{S}_p}E_n\right)<\eta$$

for some $p\in\{1,\dots,N\}$. Now let $m_2,m_3,\dots$ be an increasing sequence of elements from the set $\mathcal{S}_p$.

By induction we shall define the sequences (m_i^k) for $k\ge 1$ in the following way. First set $m_i^1=m_i$ for $i=1,2,\dots$, where (m_i) is an increasing sequence satisfying conditions (2.10) and (2.11) with $\eta=\varepsilon/2$. Now if (m_i^{k-1}) is given, by what we have already proven, we may find its subsequence (m_i^k), $m_1^k>m_1^{k-1}$, satisfying conditions (2.10) and (2.11) with $\eta=\varepsilon/2^k$.

Now set $n_i:=m_1^i$ for $i=1,2,\dots$ and observe that

$$\mu_{n_i}\left(\bigcup_{j\neq i}E_{n_j}\right)=\sum_{j<i}\mu_{n_i}(E_{n_j})+\mu_{n_i}\left(\bigcup_{j>i}E_{n_j}\right)\le\sum_{j<i}\varepsilon/2^j+\varepsilon/2^i<\varepsilon.$$

The first term evaluation follows from (2.11), by the fact that n_i is an element of the sequences (m_n^j) for $j<i$. Similarly, the second term is evaluated by inequality (2.10). □

2.4.2 Proof of Theorem 2.3.4

Proof. (Theorem 2.3.4). Let $(\mu_n)\subset\mathcal{M}^+(S)$. At the beginning we show that it is enough to prove the claim for $(\mu_n)\subset\mathcal{P}(S)$. In fact, from the assumption that $\lim_{n\to\infty}\langle\mu_n,f\rangle$ exists for every $f\in\mathrm{BL}(S)$, in particular for $f\equiv 1$, we obtain that $\lim_{n\to\infty}\mu_n(S)$ also exists. Set $c=\lim_{n\to\infty}\mu_n(S)$ and observe that $c<\infty$, by the fact that $\sup_{n\ge1}\|\mu_n\|_{TV}<\infty$. If $c=0$, then we immediately see that $\mu\equiv 0$ fulfills the requirements of our theorem. On the other hand, if $c>0$, then, we can replace μ_n with $\tilde{\mu}_n:=\mu_n/\mu_n(S)$, which is a probability measure. If the theorem is proven to hold for $(\tilde{\mu}_n)$, then it holds for the (μ_n) as well.

To prove the theorem it suffices to show that the family $\{\tilde{\mu}_n : n\ge 1\}$ is tight, by the following argument. By Prokhorov's Theorem (see [Bog07a], Theorem 8.6.2) there exists some measure $\mu_*\in\mathcal{P}(S)$ and a subsequence (n_m) such that $\tilde{\mu}_{n_m}\to\mu_*$ weakly. Further, due to the fact that $\lim_{n\to\infty}\langle\tilde{\mu}_n,f\rangle$ exists for any $f\in\mathrm{BL}(S)$, we obtain that $\lim_{n\to\infty}\langle\tilde{\mu}_n,f\rangle=\langle\mu_*,f\rangle$ for $f\in\mathrm{BL}(S)$. This in turn, together with the tightness of $\{\tilde{\mu}_n : n\ge 1\}$, implies that $\tilde{\mu}_n\to\mu_*$ $C_b(S)$-weakly, as $n\to\infty$. Indeed, the tightness allows restricting (approximately) to a compact subset K. The continuous bounded function on S, when restricted to K can be approximated uniformly by a function in $\mathrm{BL}(K)$, since $\mathrm{BL}(K) \subset C(K)$ is $\|\cdot\|_\infty$-dense. The Metric Tietze Extension Theorem (cf. [McS34]) allows to extend the function in $\mathrm{BL}(K)$ to one in $\mathrm{BL}(S)$ without changing uniform norm and Lipschitz constant. The claim then follows. The C_b-weak convergence of $\tilde{\mu}_n$ to μ_* is equivalent to $\|\tilde{\mu}_n - \mu_*\|_{\mathrm{BL}}^* \to 0$, as $n \to \infty$, because the latter norm metrises C_b-weak convergence on $\mathcal{M}^+(S)$ (cf. [Dud66], Theorem 6 and Theorem 8). For $\mu = c\mu_*$ we obtain that $\|\mu_n - \mu\|_{\mathrm{BL}}^* \to 0$, as $n \to \infty$.

To complete the proof, we have to prove the claim that the family $\{\mu_n : n \geq 1\} \subset \mathcal{P}(S)$ is uniformly tight. Assume, contrary to our claim, that it is not tight. By Lemma 2.4.3, passing to a subsequence if necessary, we may assume that there exists $\varepsilon > 0$ and a sequence of compact sets (K_n) satisfying

$$\mu_n(K_n) \geq \varepsilon \quad \text{for every } n \geq 1 \tag{2.13}$$

and

$$\operatorname{dist}(K_n, K_m) := \min\{\rho(x,y) : x \in K_n \text{ and } y \in K_m\} > \varepsilon \quad \text{for } m \neq n. \tag{2.14}$$

From Lemma 2.4.4, with $E_n := K_n^{\varepsilon/3}$, it follows that there exists a subsequence (n_i) such that for every $i \geq 1$ we have

$$\mu_{n_i}\left(\bigcup_{j \neq i} K_{n_j}^{\varepsilon/3}\right) < \varepsilon/2. \tag{2.15}$$

Note that $\operatorname{dist}(K_{n_i}^{\varepsilon/3}, K_{n_j}^{\varepsilon/3}) > \varepsilon/3$ for $i \neq j$.

We define the function $f : X \to [0,1]$ by the formula

$$f(x) = \sum_{i=1}^{\infty} f_i(x),$$

where f_i are arbitrary Lipschitz functions with Lipschitz constant $3/\varepsilon$ satisfying

$$f_{i|K_{n_{2i}}} = 1 \quad \text{and} \quad 0 \leq f_i \leq \mathbf{1}_{K_{n_{2i}}^{\varepsilon/3}}.$$

According to Lemma 2.4.2, $f \in \mathrm{BL}(S)$ (with $\|f\|_\infty \leq 1$ and $|f|_L \leq 6/\varepsilon$).

To finish the proof it is enough to observe that for every $i \geq 1$ we have

$$\langle \mu_{n_{2i}}, f \rangle = \sum_{j=1}^{\infty} \langle \mu_{n_{2i}}, f_j \rangle \geq \mu_{n_{2i}}\left(K_{n_{2i}}\right) \overset{(2.13)}{\geq} \varepsilon$$

and

$$\langle \mu_{n_{2i+1}}, f\rangle = \sum_{j=1}^{\infty} \langle \mu_{n_{2i+1}}, f_j\rangle \le \sum_{j=1}^{\infty} \mu_{n_{2i+1}}\left(K_{n_{2j}}^{\varepsilon/3}\right) \le \mu_{n_{2i+1}}\left(\bigcup_{j\neq 2i+1} K_{n_j}^{\varepsilon/3}\right) \overset{(2.15)}{<} \varepsilon/2,$$

which contradicts the assumption that $\lim_{n\to\infty}\langle \mu_n, f\rangle$ exists for every $f \in \mathrm{BL}(S)$. Thus the family $\{\mu_n : n \ge 1\}$ is tight and we are done. □

Remark 2.4.1. 1.) An alternative proof is feasible, based upon the elaborate theory presented in [Pac13]. By taking $f = \mathbb{1}$, one finds that $\sup_n \|\mu_n\|_{\mathrm{TV}} < \infty$. Since $\mathrm{BL}(S)$ is dense in the space $\mathrm{U}_b(S)$ of uniformly continuous bounded functions on S for the supremum norm (cf. [Dud66], Lemma 8), one finds that $\langle \mu_n, f\rangle$ is Cauchy for every $f \in \mathrm{U}_b(S)$. According to [Pac13], Theorem 5.45, there exists $\mu \in \mathcal{M}(S)^+$ such that $\mu_n \to \mu$, $\mathrm{U}_b(S)$-weakly. Then [Pac13] Theorem 5.36 yields that $\|\mu_n - \mu\|_{\mathrm{BL}}^* \to 0$.
2.) In the proof we show that if (μ_n) is a sequence of positive Borel measures such that $\langle \mu_n, f\rangle$ converges for every $f \in \mathrm{BL}(S)$, then (μ_n) is uniformly tight in $\mathcal{M}^+(S)$. See [Bog07a], Corollary 8.6.3, p. 204, for results in this direction when $\langle \mu_n, f\rangle$ converges for every $f \in C_b(S)$. Under the additional condition that there exists $\mu_* \in \mathcal{M}^+(S)$ such that $\langle \mu_n, f\rangle \to \langle \mu_*, f\rangle$ for every $f \in C_b(S)$, tightness results appeared already in e.g. [LeC57], Theorem 4 for positive measures or [Bil99], Appendix III, Theorem 8 for probability measures.

2.4.3 Proof of Theorem 2.3.1

Proof. (Theorem 2.3.1). Let $(\mu_n) \subset \mathcal{M}(S)$ be signed measures such that $\sup_n \|\mu_n\|_{\mathrm{TV}} < \infty$. Denote by μ_n^+ and μ_n^- the positive and negative part of μ_n, $n \ge 1$, respectively. We consider the following set

$$\begin{aligned}\mathcal{C} := \Big\{\big(\beta, (m_n), (\nu_{m_n}), (\vartheta_{m_n})\big) : \beta \ge 0,\ (m_n) \subset \mathbb{N} - \text{an increasing sequence,}\\ \nu_{m_n}, \vartheta_{m_n} \in \mathcal{P}(S),\ \lim_{n\to\infty} \|\nu_{m_n} - \vartheta_{m_n}\|_{\mathrm{BL}}^* = 0\\ \text{and}\ \ \mu_{m_n}^+ \ge \beta\nu_{m_n},\ \mu_{m_n}^- \ge \beta\vartheta_{m_n}\Big\}.\end{aligned}$$

We first observe that $\mathcal{C} \neq \varnothing$, which follows from the fact that $\big(0, (m_n), (\nu_{m_n}), (\vartheta_{m_n})\big) \in \mathcal{C}$ for arbitrary (m_n) and $\nu_{m_n}, \vartheta_{m_n} \in \mathcal{P}(S)$ such that $\lim_{n\to\infty} \|\nu_{m_n} - \vartheta_{m_n}\|_{\mathrm{BL}}^* = 0$. Moreover, since $\bar{c} := \sup_{n\ge 1} \|\mu_n\|_{TV} < \infty$, we obtain that $0 \le \beta \le \bar{c}$ for every β for which there are some (m_n) and $\nu_{m_n}, \vartheta_{m_n}$ such that $\big(\beta, (m_n), (\nu_{m_n}), (\vartheta_{m_n})\big) \in \mathcal{C}$. We can therefore introduce

$$\alpha = \sup\Big\{\beta : \big(\beta, (m_n), (\nu_{m_n}), (\vartheta_{m_n})\big) \in \mathcal{C}\Big\}.$$

From the definition of α it follows that there exists a subsequence (m_n) of positive integers and an increasing sequence (α_n) of nonnegative constants satisfying $\lim_{n\to\infty}\alpha_n = \alpha$ and

$$\mu^+_{m_n} \geq \alpha_n \nu_{m_n} \quad \text{and} \quad \mu^-_{m_n} \geq \alpha_n \vartheta_{m_n},$$

where $\nu_{m_n}, \vartheta_{m_n} \in \mathcal{P}(S)$ are such that $\|\nu_{m_n} - \vartheta_{m_n}\|^*_{\mathrm{BL}} \to 0$ as $n \to \infty$.

To finish the proof it is enough to show that both the sequences $(\mu^+_{m_n} - \alpha_n\nu_{m_n})$ and $(\mu^-_{m_n} - \alpha_n\vartheta_{m_n})$ are tight. Indeed, then, by the Prokhorov Theorem ([Bog07a], Theorem 8.6.2) there exists a subsequence (m_{n_k}) of (m_n) and two measures μ^1 and μ^2 such that the sequences $(\mu^+_{m_{n_k}} - \alpha_{n_k}\nu_{m_{n_k}})$ and $(\mu^-_{m_{n_k}} - \alpha_{n_k}\vartheta_{m_{n_k}})$ converge $C_b(S)$-weakly to the positive measure μ^1 and μ^2, respectively. Hence also in $\|\cdot\|^*_{\mathrm{BL}}$-norm, according to Theorem 2.3.4. Consequently, $\|\mu_{m_{n_k}} - (\mu^1 - \mu^2)\|^*_{\mathrm{BL}} \to 0$ as $k \to \infty$, by the fact that $\|\nu_{m_{n_k}} - \vartheta_{m_{n_k}}\|^*_{\mathrm{BL}} \to 0$ as $k \to \infty$. This will complete the proof of the theorem. Indeed, if we know that the sequence (and also any subsequence) has a convergent subsequence (in the dual bounded Lipschitz norm), then the sequence is also convergent due to the fact that the limit of all convergent subsequences is the same, by the assumption that $\lim_{n\to\infty}\langle \mu_n, f\rangle$ exists for any $f \in \mathrm{BL}(S)$.

Assume now, contrary to our claim, that at least one of the families $(\mu^+_{m_n} - \alpha_n\nu_{m_n})$ or $(\mu^-_{m_n} - \alpha_n\vartheta_{m_n})$, say the first one, is not tight. By Lemma 2.4.3, passing to a subsequence if necessary, we may assume that there exists $\varepsilon > 0$ and a sequence of compact sets (K_n) satisfying

$$(\mu^+_{m_n} - \alpha_n\nu_{m_n})(K_n) \geq \varepsilon \tag{2.16}$$

and

$$\mathrm{dist}(K_i, K_j) \geq \varepsilon \quad \text{for } i, j \in \mathbb{N},\ i \neq j.$$

Set

$$\tilde{\mu}_n := \mu^+_{m_n} - \alpha_n\nu_{m_n} \quad \text{and} \quad \hat{\mu}_n := \mu^-_{m_n} - \alpha_n\vartheta_{m_n}.$$

Claim: *For any $0 < \eta \leq 1$ there exist j, as large as we wish, and $\tau_j, \chi_j \in \mathcal{P}(S)$ satisfying*

$$\tilde{\mu}_j \geq (\varepsilon/2)\tau_j, \qquad \hat{\mu}_j \geq (\varepsilon/2)\chi_j \quad \textit{and} \quad \|\tau_j - \chi_j\|^*_{\mathrm{BL}} \leq \eta.$$

Consequently, there will exist a subsequence (m_{j_n}) such that

$$\mu^+_{m_{j_n}} = \alpha_{j_n}\nu_{m_{j_n}} + \tilde{\mu}_{j_n} \geq \alpha_{j_n}\nu_{m_{j_n}} + (\varepsilon/2)\tau_{j_n},$$

$$\mu^-_{m_{j_n}} \geq \alpha_{j_n}\vartheta_{m_{j_n}} + (\varepsilon/2)\chi_{j_n} \qquad \text{and} \qquad \|\tau_{j_n} - \chi_{j_n}\|^*_{\mathrm{BL}} \to 0 \text{ as } n \to \infty.$$

Now, if we define probability measures $\varrho_{m_{j_n}}, \varsigma_{m_{j_n}}$ as follows

$$\varrho_{m_{j_n}} := (\alpha_{j_n}\nu_{m_{j_n}} + (\varepsilon/2)\tau_{j_n})(\alpha_{j_n} + \varepsilon/2)^{-1}, \qquad \varsigma_{m_{j_n}} := (\alpha_{j_n}\vartheta_{m_{j_n}} + (\varepsilon/2)\chi_{j_n})(\alpha_{j_n} + \varepsilon/2)^{-1},$$

we will obtain

$$\mu^+_{m_{j_n}} \geq (\alpha_{j_n} + \varepsilon/2)\varrho_{m_{j_n}}, \qquad \mu^-_{m_{j_n}} \geq (\alpha_{j_n} + \varepsilon/2)\varsigma_{m_{j_n}}$$

and $\lim_{n\to\infty} \|\varrho_{m_{j_n}} - \varsigma_{m_{j_n}}\|^*_{\mathrm{BL}} = 0$, which is impossible, because it contradicts the definition of α, since $\lim_{n\to\infty}(\alpha_{j_n} + \varepsilon/2) > \alpha$.

Let us prove the claim. Set $\xi_n := \tilde{\mu}_n + \hat{\mu}_n$ for $n \geq 1$ and let $C := \sup_{n\geq1} \xi_n(S)$. Observe that $C \leq \sup_{n\geq1} \|\mu_n\|_{\mathrm{TV}} < \infty$. Fix $0 < \eta \leq 1$ and let $\kappa \in (0, \varepsilon/6)$ be such that $6\kappa(1/\varepsilon + 2/\varepsilon^2) < \eta$. Lemma 2.4.4 yields an increasing sequence $(j_n) \subset \mathbb{N}$ such that

$$\xi_{j_n}\left(\bigcup_{l\neq n} K^{\varepsilon/3}_{j_l}\right) < \kappa/4 \tag{2.17}$$

and hence

$$\tilde{\mu}_{j_n}\left(\bigcup_{l\neq n} K^{\varepsilon/3}_{j_l}\right) < \kappa/4 \quad \text{and} \quad \hat{\mu}_{j_n}\left(\bigcup_{l\neq n} K^{\varepsilon/3}_{j_l}\right) < \kappa/4$$

for all $n = 1, 2, \ldots$.

Choose $N \geq 1$ such that $N\kappa/4 > C$ and set $W^p_{j_n} := K^{p\varepsilon/(3N)}_{j_n} \setminus K^{(p-1)\varepsilon/(3N)}_{j_n}$ for $p = 1, \ldots, N$. Observe that $W^p_{j_n} \cap W^q_{j_n} = \emptyset$ for $p \neq q$. Since $\sum_{p=1}^N \xi_{j_n}(W^p_{j_n}) = \xi_{j_n}(\bigcup_{p=1}^N W^p_{j_n}) \leq C$, $n \geq 1$, for every n there exists $p_n \in \{1, \ldots, N\}$ such that

$$\xi_{j_n}\left(W^{p_n}_{j_n}\right) < \kappa/4. \tag{2.18}$$

Now we are in a position to define a sequence (f_n) of functions from S to $[-1, 1]$. The construction is as follows. For $n = 2k + 1$ for $k \geq 1$, we set $f_n \equiv 0$. On the other hand, to define functions f_n for $n = 2k$ we introduce the measures

$$\tilde{\mu}'_{j_n}(\cdot) = \tilde{\mu}_{j_n}\left(\cdot \cap K^{(p_n-1)\varepsilon/(3N)}_{j_n}\right)$$

and

$$\hat{\mu}'_{j_n}(\cdot) = \hat{\mu}_{j_n}\left(\cdot \cap K^{(p_n-1)\varepsilon/(3N)}_{j_n}\right).$$

Further, there exists a Lipschitz function $\tilde{f}_n : K^{(p_n-1)\varepsilon/(3N)}_{j_n} \to [-1, 1]$ with $|\tilde{f}_n|_L \leq 1$ such that $\left\langle \tilde{\mu}'_{j_n} - \hat{\mu}'_{j_n}, \tilde{f}_n \right\rangle \geq \frac{1}{2}\|\tilde{\mu}'_{j_n} - \hat{\mu}'_{j_n}\|^*_{\mathrm{BL}}$. Let f_n be a Lipschitz extension of the function $\tilde{f}_n$ to

S such that $f_n(x) = \tilde{f}_n(x)$ for $x \in K_{j_n}^{(p_n-1)\varepsilon/(3N)}$ and $f_n(x) = 0$ for $x \notin K_{j_n}^{p_n\varepsilon/(3N)}$. We may assume that $|f_n|_L \leq 3N/\varepsilon$. The existence of the extension function follows from McShane's formula (see [McS34]). Let $f = \sum_{k=1}^{\infty} f_{2n}$. Since $\operatorname{dist}(\operatorname{supp} f_i, \operatorname{supp} f_j) > \varepsilon/3$ for $i, j \geq 1$, $i \neq j$, f is a bounded Lipschitz function, by Lemma 2.4.2.

We show that $\langle \mu_{m_{j_i}}, f\rangle \leq \kappa/2$ for $i = 2k+1$. Indeed, for k sufficiently large we have

$$\begin{aligned}\left\langle \mu_{m_{j_{2k+1}}}, f\right\rangle &= \sum_{n=1}^{\infty} \left\langle \mu_{m_{j_{2k+1}}}, f_{2n}\right\rangle \leq \sum_{n=1}^{\infty} \xi_{j_{2k+1}}\left(K_{j_{2n}}^{\varepsilon/3}\right) + \alpha_{j_{2k+1}}\|\nu_{m_{j_{2k+1}}} - \vartheta_{m_{j_{2k+1}}}\|_{\mathrm{BL}}^{*}\\ &\leq \xi_{j_{2k+1}}\left(\bigcup_{l\neq 2k+1} K_{j_l}^{\varepsilon/3}\right) + \alpha_{j_{2k+1}}\|\nu_{m_{j_{2k+1}}} - \vartheta_{m_{j_{2k+1}}}\|_{\mathrm{BL}}^{*}\\ &\overset{(2.17)}{<} \kappa/4 + \alpha_{j_{2k+1}}\|\nu_{m_{j_{2k+1}}} - \vartheta_{m_{j_{2k+1}}}\|_{\mathrm{BL}}^{*} < \kappa/2,\end{aligned}$$

by the properties of the measures $\nu_{m_{j_{2k+1}}}, \vartheta_{m_{j_{2k+1}}}$ and the definition of the functions f_{2n}. Therefore

$$\lim_{i\to\infty}\left\langle \mu_{m_{j_i}}, f\right\rangle = \lim_{k\to\infty}\left\langle \mu_{m_{j_{2k+1}}}, f\right\rangle \leq \kappa/2,$$

because we assume that the limit of $\langle \mu_m, f\rangle$ exists.

On the other hand, for $i = 2k$ we have

$$\begin{aligned}\left\langle \mu_{m_{j_{2k}}}, f\right\rangle &= \sum_{n=1}^{\infty}\left\langle \mu_{m_{j_{2k}}}, f_{2n}\right\rangle \geq -\sum_{n\neq k}^{\infty} \xi_{j_{2k}}\left(K_{j_{2n}}^{\varepsilon/3}\right) + \left\langle \mu_{m_{j_{2k}}}, f_{2n},\right\rangle\\ &\geq -\sum_{n\neq k}^{\infty} \xi_{j_{2k}}\left(K_{j_{2n}}^{\varepsilon/3}\right) - \xi_{j_{2k}}\left(W_{j_{2k}}^{p_{2k}}\right) + \left\langle \tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}, \tilde{f}_{2k}\right\rangle\\ &\geq -\kappa/4 - \kappa/4 + \frac{1}{2}\|\tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}\|_{\mathrm{BL}}^{*},\end{aligned}$$

by the fact that $\|f_{2n}\|_\infty \leq 1$. Since $\lim_{i\to\infty}\left\langle \mu_{m_{j_i}}, f\right\rangle \leq \kappa/2$, by the estimation obtained for $i = 2k+1$ and the assumption that the limit exists, we have

$$-\kappa/4 - \kappa/4 + \frac{1}{2}\|\tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}\|_{\mathrm{BL}}^{*} \leq 3\kappa/4$$

for k sufficiently large and consequently

$$\|\tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}\|_{\mathrm{BL}}^{*} \leq 3\kappa$$

for all k sufficiently large. Thus

$$\hat{\mu}'_{j_{2k}}(S) \geq \tilde{\mu}'_{j_{2k}}(S) - 3\kappa \geq \varepsilon - \varepsilon/2 = \varepsilon/2.$$

Hence, for probability measures

$$\tilde{\nu}_{j_{2k}} := \tilde{\mu}'_{j_{2k}}/\tilde{\mu}'_{j_{2k}}(S) \quad \text{and} \quad \hat{\nu}_{j_{2k}} := \hat{\mu}'_{j_{2k}}/\hat{\mu}'_{j_{2k}}(S)$$

we have for k sufficiently large

$$\tilde{\mu}_{j_{2k}} \geq \tilde{\mu}'_{j_{2k}} \geq (\varepsilon/2)\tilde{\nu}_{j_{2k}} \quad \text{and} \quad \hat{\mu}_{j_{2k}} \geq \hat{\mu}'_{j_{2k}} \geq (\varepsilon/2)\hat{\nu}_{j_{2k}}.$$

Finally, observe that for k sufficiently large,

$$\begin{aligned}
\|\tilde{\nu}_{j_{2k}} - \hat{\nu}_{j_{2k}}\|^*_{\mathrm{BL}} &\leq \|\tilde{\mu}'_{j_{2k}}/\tilde{\mu}'_{j_{2k}}(S) - \hat{\mu}'_{j_{2k}}/\tilde{\mu}'_{j_{2k}}(S)\|^*_{\mathrm{BL}} + \|\hat{\mu}'_{j_{2k}}\|^*_{\mathrm{BL}}|1/\tilde{\mu}'_{j_{2k}}(S) - 1/\hat{\mu}'_{j_{2k}}(S)| \\
&\leq (1/\tilde{\mu}'_{j_{2k}}(S))\|\tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}\|^*_{\mathrm{BL}} + 1/(\tilde{\mu}'_{j_{2k}}(S)\hat{\mu}'_{j_{2k}}(S))|\tilde{\mu}'_{j_{2k}}(S) - \hat{\mu}'_{j_{2k}}(S)| \\
&\leq 6\kappa/\varepsilon + 12\kappa/\varepsilon^2 < \eta,
\end{aligned}$$

by the fact that $\tilde{\mu}'_{j_{2k}}(S), \hat{\mu}'_{j_{2k}}(S) \geq \varepsilon/2$ and $|\tilde{\mu}'_{j_{2k}}(S) - \hat{\mu}'_{j_{2k}}(S)| \leq \|\tilde{\mu}'_{j_{2k}} - \hat{\mu}'_{j_{2k}}\|^*_{\mathrm{BL}} \leq 3\kappa$. This completes the proof of the claim, hence the theorem. □

Remark 2.4.2. It is possible to prove Theorem 2.3.1 by means of a reduction-to-ℓ^1-trick, inspired by ideas in [Pac79, Pac13], cf. [Hil14]. Another proof is feasible, starting from [Pac79], Theorem 3.2, see [Wor10]. However, here we prefer to present an independent, 'set-geometric' proof that is self-contained and founded on the well-established result for the case of positive measures, Theorem 2.3.4.

2.5 Further consequence: an alternative proof for weak sequential completeness

Theorem 2.3.1 allows – in the case of a Polish space – to give an alternative proof of the well-known fact that $\mathcal{M}(S)$ is $C_b(S)$-weakly sequentially complete, that goes back to Alexandrov [Ale43] and Varadarajan [Var61], see. e.g. [Dud66], Theorem 1 or [Bog07a], Theorem 8.7.1 for a more general topological setting. We include our proof based on Theorem 2.3.1 here, because it employs an argument for reduction to functions in $\mathrm{BL}(S)$, which by itself is an interesting result.

This reduction is based on the following observation. Let $\mathcal{D}_S$ be the set of all metrics on S that metrize the topology of S as a complete separable metric space. We need to stress the dependence of the space $\mathrm{BL}(S)$ on the chosen metric on S. So for $d \in \mathcal{D}_S$ we write $\mathrm{BL}(S, d)$ for the space of bounded Lipschitz functions on (S, d). The key observation is, that

$$C_b(S) = \bigcup_{d\in\mathcal{D}_S} \mathrm{BL}(S,d). \tag{2.19}$$

In fact, fix $d_0 \in \mathcal{D}_S$. If $f \in C_b(S)$, then

$$d_f(x,y) := d_0(x,y) \vee |f(x) - f(y)|$$

is a metric on S such that $d_f \in \mathcal{D}_S$ and $f \in \mathrm{BL}(S,d_f)$. Here $\vee$ denotes the maximum.

The precise statement we consider is the following:

Theorem 2.5.1 (Weak sequential completeness)**.** *Let S be a Polish space. Let $(\mu_n) \subset \mathcal{M}(S)$ be such that $\langle \mu_n, f\rangle$ converges for every $f \in C_b(S)$. Then there exists $\mu_* \in \mathcal{M}(S)$ such that $\langle \mu_n, f\rangle \to \langle \mu_*, f\rangle$ for every $f \in C_b(S)$.*

Proof. The norm of μ_n viewed as a continuous linear functional on $C_b(S)$ is its total variation norm. Hence, according to the Banach-Steinhaus Theorem, $\sup_{n\geq 1} \|\mu_n\|_{\mathrm{TV}} < \infty$. For any $d \in \mathcal{D}_S$, $\langle \mu_n, f\rangle$ converges for every $f \in C_b(S)$, so in particular for every $f \in \mathrm{BL}(S,d)$. The sequence (μ_n) is bounded in total variation norm, so Theorem 2.3.1 implies there exists $\mu_*^d \in \mathcal{M}(S)$ such that $\langle \mu_n, f\rangle \to \langle \mu_*^d, f\rangle$ for every $f \in \mathrm{BL}(S,d)$. We proceed to show that the limit measure μ_*^d is independent of d.

Let $d' \in \mathcal{D}_S$. Put

$$\bar{d}(x,y) := d(x,y) \vee d'(x,y).$$

Then $\bar{d} \in \mathcal{D}_S$ and $\mathrm{BL}(S,\bar{d})$ contains both $\mathrm{BL}(S,d)$ and $\mathrm{BL}(S,d')$. Let $C \subset S$ be closed. There exist sequences (h_n) and (h_n') in $\mathrm{BL}(S,d)$ and $\mathrm{BL}(S,d')$ respectively, such that $h_n \downarrow \mathbb{1}_C$ and $h_n' \downarrow \mathbb{1}_C$ pointwise. Both these sequences are in $\mathrm{BL}(S,\bar{d})$, so

$$\mu_*^d(C) = \lim_{k\to\infty} \left\langle \mu_*^d, h_k \right\rangle = \lim_{k\to\infty} \lim_{n\to\infty} \langle \mu_n, h_k\rangle = \lim_{k\to\infty} \left\langle \mu_*^{\bar{d}}, h_k \right\rangle = \mu_*^{\bar{d}}(C).$$

A similar argument applies to $\mu_*^{d'}$, using the sequence (h_n') in $\mathrm{BL}(S,d')$ instead of (h_n). So μ_*^d and $\mu_*^{d'}$ (and $\mu_*^{\bar{d}}$) agree on the π-system consisting of closed sets, which generate the Borel σ-algebra. Hence these measures are equal on all Borel sets. That is, there exists $\mu_* \in \mathcal{M}(S)$ such that $\langle \mu_n, f\rangle \to \langle \mu_*, f\rangle$ for every $f \in \mathrm{BL}(S,d)$ *for every* $d \in \mathcal{D}_S$. Thus for every $f \in C_b(S)$ in view of (2.19). □

Chapter 3

Lie-Trotter product formula for locally equicontinuous and tight Markov operators

This chapter is based on:

Sander C. Hille, Maria A. Ziemlanska. Lie-Trotter product formula for locally equicontinuous and tight Markov semigroup. Preprint available at https://arxiv.org/abs/1807.07728

Abstract:

In this chapter we prove a Lie-Trotter product formula for Markov semigroups in spaces of measures. We relate our results to "classical" results for strongly continuous linear semigroups on Banach spaces or Lipschitz semigroups in metric spaces and show that our approach is an extension of existing results. As Markov semigroups on measures are usually neither strongly continuous nor bounded linear operators for the relevant norms, we prove the convergence of the Lie-Trotter product formula assuming that the semigroups are locally equicontinuous and tight. A crucial tool we use in the proof is a Schur-like property for spaces of measures.

3.1 Introduction

The main purpose of this chapter is to generalize the Lie-Trotter product formula for strongly continuous linear semigroups in a Banach space to Markov semigroups on spaces of measures. The Lie-Trotter formula asserts the existence and properties of the limit

$$\lim_{n\to\infty}\left[S^1_{\frac{t}{n}}S^2_{\frac{t}{n}}\right]^n x =: S_t x,$$

where $(S^1_t)_{t\geq 0}$ and $(S^2_t)_{t\geq 0}$ are strongly continuous semigroups of bounded linear operators. It may equally be viewed as a statement considering the convergence of a *switching scheme*. The key challenge is to overcome the difficulties that result from the observation that 'typically' Markov semigroups do not consist of bounded linear operators (in a suitable norm on the signed measures) nor need to be strongly continuous. Therefore, the available results do not apply.

The Lie-Trotter product formula originated from Trotter [Tro59] in 1959 for strongly continuous semigroups, for which the closure of the sum of two generators was a generator of a semigroup given by the limit of the Lie-Trotter scheme, and generalized i.a. by Chernoff [Che74] in 1974. This approach does not seem to be general enough to be applicable in various numerical schemes however. As shown by Kurtz and Pierre in [KP80], even if the sum of two generators is again a generator of a strongly continuous semigroup, this semigroup may not be given by the limit of Lie-Trotter product formula as it may not converge. Consequently, the analysis of generators of semigroups can lead to non-convergent numerical splitting schemes. Hence, a different approach is needed. The analysis of commutator type conditions as in [KW01, CC04] avoids considering generators and their domains and may be easier to verify.

Splitting schemes were applied and played a very important role in numerical analysis and recently in the theory of stochastic differential equations to construct solutions of differential equations, e.g. the work of Cox and Van Neerven [Cox12]. It was shown by Carrillo, Gwiazda and Ulikowska in [CGU14] that properties of complicated models, like structured population models, can be obtained by splitting the original model into simpler ones and analyzing them separately, which also leads to switching schemes of a Lie-Trotter form. Bátkai, Csomós and Farkas investigated Lie-Trotter product formulae for abstract nonlinear evolution equations with a delay in [BCF17], a general product formula for the solution of nonautonomous abstract delay equations in [BCFN12] and analyzed the convergence of operator splitting procedures in [BCF13].

Our starting point are the conditions for convergence of the Lie-Trotter product formula formulated by Kühnemund and Wacker in [KW01]. This result appears to be a very useful tool in proving the convergence of the Lie-Trotter scheme without the need to have knowledge about generators of the semigroups involved. However, the semigroups considered by Kühnemund and Wacker are assumed to be strongly continuous. We extend Kühnemund and Wacker's case to semigroups of Markov operators on spaces of measures and present weaker sufficient conditions for convergence of the switching scheme. Our method of proof builds on [KW01], while the specific commutator condition that we employ (assumption 3) is motivated by [CC04].

The theory of Markov operators and Markov semigroups was studied by Lasota, Mackey, Myjak and Szarek in the context of fractal theory [SM03, LM94], iterated function systems and stochastic differential equations [LS06]. Markov semigroups acting on spaces of (separable) measures are usually not strongly continuous. The local equicontinuity (in measures) and tightness assumptions we employ are less restrictive and follow from strong continuity. The concept of equicontinuous families of Markov operators can be found in e.g. Meyn and Tweedie [MT09]. Also, Worm in [Wor10] extends the results of Szarek to families of equicontinuous Markov operators.

The outline of the chapter is as follows: in Section 3.2 we present the main results of this chapter. Theorem 3.2.2 in Section 3.2 is the convergence theorem and is the most important result in the chapter. The other important and non-trivial result is Theorem 3.2.1. Section 3.3 introduces Markov operators and Markov-Feller semigroups on a space of signed Borel measures $\mathcal{M}(S)$, investigates their topological properties and the consequences of equicontinuity and tightness of a family of Markov operators. In Section 3.4 we provide the tools to prove Theorem 3.2.1, i.e. that a composition of equicontinuous and tight families of Markov operators is again an equicontinuous and tight family. This result is quite delicate and seems like it was not considered in the literature before. We also provide a proof of the observation in Lemma 3.4.3 which says that a family of equicontinuous and tight family of Markov operators on a precompact subset of positive measures is again precompact. The proof of Theorem 3.2.1 can be found in Appendix 3.4.

In Section 3.5 we prove the convergence of the Lie-Trotter product formula for Markov operators. We provide more general assumptions then those provided in the Kühnemund-Wacker chapter (see [KW01]). As our semigroups are not strongly continuous and usually not bounded, we use the concept of (local) equicontinuity (see e.g. Chapter 7 in [Wor10]). This allows us to define a new admissible metric $d_{\mathcal{E}}$ and a new $\|\cdot\|_{\mathrm{BL},d_{\mathcal{E}}}$-norm dependent on the operators and the original metric d on S. The crucial assumption is the Commutator Condition Assumption 3.

To prove the convergence of our scheme under Assumptions 1-4 we use a Schur-like property for signed measures, see [HSWZ17], which allows us to prove weak convergence of the formula and conclude the strong/norm convergence. In Section 3.5 we show crucial technical lemmas. The proofs of most lemmas from Section 3.5 can be found in the Appendices 3.8.1 - 3.8.2. In Section 3.5 several useful properties of the limit operators that result from the converging Lie-Trotter formula are derived.

Section 3.7 shows that our approach is a generalization of Kühnemund-Wacker [Kuh01] and Colombo-Corli [CC04] cases. We show that if we consider Markov semigoups coming from lifts of deterministic operators, then the Kühnemund-Wacker and Colombo-Corli assumptions imply our assumptions and their convergence results of the Lie-Trotter formula or switching scheme follows from our main convergence result.

3.2 Main theorems

Let S be a Polish space, i.e. a separable completely metrizable topological space, see [Wor10]. Any metric d that metrizes the topology of S such that (S,d) is separable and complete is called *admissible*. Let d be an admissible metric on S. Following [Dud66], we denote the vector space of all real-valued Lipschitz functions on (S,d) by $\mathrm{Lip}(S,d)$. For $f\in\mathrm{Lip}(S,d)$ we denote the Lipschitz constant of f by

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

$\mathrm{BL}(S,d)$ is the subspace of bounded functions in $\mathrm{Lip}(S,d)$. Equipped with the bounded Lipschitz norm

$$\|f\|_{\mathrm{BL},d} := \|f\|_\infty + |f|_{L,d}$$

it is a Banach space, see [Dud66]. The vector space of finite signed Borel measures on S, $\mathcal{M}(S)$, embeds into the dual of $(\mathrm{BL}(S),\|\cdot\|_{\mathrm{BL},d})$, see [Dud66], thus introducing the dual bounded Lipschitz norm $\|\cdot\|^*_{\mathrm{BL},d}$ on $\mathcal{M}(S)$

$$\|\mu\|^*_{\mathrm{BL},d} := \sup\left\{|\langle\mu,f\rangle| : f\in\mathrm{BL}(S,d), \|f\|_{\mathrm{BL},d} = \|f\|_\infty + |f|_{L,d}\leq 1\right\}, \tag{3.1}$$

for which the space becomes a normed space. It is not complete unless (S,d) is uniformly discrete (see [Wor10], Corollary 2.3.14). The cone $\mathcal{M}^+(S)$ of positive measures in $\mathcal{M}(S)$ is closed [Wor10, Dud66]. $\mathcal{P}(S)$ is the convex subset of $\mathcal{M}^+(S)$ of probability measures. The topology on $\mathcal{M}(S)$ induced by $\|\cdot\|^*_{\mathrm{BL},d}$ is weaker then the norm topology associated

with the total variation norm $\|\mu\|_{TV} := \mu^+(S) + \mu^-(S)$, where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ (see [Bog07b], p.176).
We define a Markov operator on S to be a map $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ such that

(i) P is additive and $\mathbb{R}_+$-homogeneous;

(ii) $\|P\mu\|_{TV} = \|\mu\|_{TV}$ for all $\mu \in \mathcal{M}^+(S)$.

Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of Markov operators.

Following Lasota and Szarek [LS06], and Worm [Wor10], we say that $(P_\lambda)_{\lambda\in\Lambda}$ is *equicontinuous* at $\mu \in \mathcal{M}^+(S)$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|P_\lambda\mu - P_\lambda\nu\|^*_{\mathrm{BL},d} < \varepsilon$ for every $\nu \in \mathcal{M}^+(S)$ such that $\|\mu - \nu\|^*_{\mathrm{BL},d} < \delta$ and for every $\lambda \in \Lambda$. $(P_\lambda)_{\lambda\in\Lambda}$ is called equicontinuous if it is equicontinuous at every $\mu \in \mathcal{M}^+(S)$. We will examine properties of space of bounded Lipschitz functions is Section 3.3.

Let $\Theta \subset \mathcal{P}(S)$. Following [Bog07a] we call Θ *uniformly tight* if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset S$ such that $\mu(K_\epsilon) \geq 1 - \epsilon$ for all $\mu \in \Theta$.

The following theorem is a crucial tool for proving convergence of the Lie-Trotter scheme for Markov semigroups, and also an important and non-trivial result on its own. Proof of Theorem 3.2.1 can be found in Section 3.4.

Theorem 3.2.1. *Let $(P_\lambda)_{\lambda\in\Lambda}$, $(Q_\gamma)_{\gamma\in\Gamma}$ be equicontinuous families of Markov operators on (S,d). Assume that $(Q_\gamma)_{\gamma\in\Gamma}$ is tight. Then the family $\{P_\lambda Q_\gamma : \lambda \in \Lambda, \gamma \in \Gamma\}$ is equicontinuous on (S,d). Moreover, if $(P_\lambda)_{\lambda\in\Lambda}$ is tight, then the family $\{P_\lambda Q_\gamma : \lambda \in \Lambda, \gamma \in \Gamma\}$ is tight on (S,d).*

We now present assumptions under which we prove the convergence of the Lie-Trotter scheme. Even though they may seem technical, they are motivated by existing examples of convergence of Lie-Trotter schemes with weaker assumptions then those in [KW01, CC04] (see Section 3.7).

Let $(P^1_t)_{t\geq 0}$ and $(P^2_t)_{t\geq 0}$ be Markov semigroups. Let $\delta > 0$. Define

$$\mathcal{P}^i(\delta) := \{P^i_t : t \in [0,\delta]\} \text{ for } i = 1,2,$$

$$\mathcal{F}(\delta) := \left\{\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n : n \in \mathbb{N}, t \in [0,\delta]\right\}.$$

Let d be an admissible metric on S such that the following assumptions hold:

Assumption 1. *There exists $\delta_1 > 0$ such that $\mathcal{P}^1(\delta_1)$ and $\mathcal{P}^2(\delta_1)$ are equicontinuous and tight families of Markov operators on (S,d).*

Assumption 2 (Stability condition)**.** *There exists $\delta_2 > 0$ such that $\mathcal{F}(\delta_2)$ is an equicontinuous family of Markov operators on (S,d).*

Under Assumption 1, the operators $P^i_t, 0 \le t \le \delta$, are Feller: there exist $U^i_t : C_b(S) \to C_b(S)$ such that $\langle P^i_t \mu, f\rangle = \langle \mu, U^i_t f\rangle$ for every $f \in C_n(S), \mu_0 \in \mathcal{M}^+(S)$, $0 \le t \le \delta$.

Let $f \in \mathrm{BL}(S,d)$ and consider

$$\mathcal{E}(f) \coloneqq \left\{ U^2_s U^1_{s'} \left[U^2_{\frac{t}{n}} U^1_{\frac{t}{n}} \right]^n f : n \in \mathbb{N}, s, s', t \in [0,\delta] \right\}. \tag{3.2}$$

By Theorem 7.2.2 in [Wor10] or Theorem 3.4.2 below, equicontinuity of the family $(P_\lambda)_{\lambda\in\Lambda}$ is equivalent to equicontinuity of the family $(U_\lambda f)_{\lambda\in\Lambda}$ for every $f \in \mathrm{BL}(S,d)$. Then, as we will show in Lemma 3.5.4, $\mathcal{E}(f)$ is an equicontinuous family if $\delta \le \min(\delta_1, \delta_2)$. It defines a new admissible metric on S:

$$d_{\mathcal{E}(f)}(x,y) \coloneqq d(x,y) \vee \sup_{g\in\mathcal{E}(f)} |g(x) - g(y)|, \quad \text{for} \quad x, y \in S. \tag{3.3}$$

Assumption 3 (Commutator condition)**.** *There exists a dense convex subcone M_0 of $\mathcal{M}^+(S)_{\mathrm{BL},d}$ that is invariant under $(P^i_t)_{t\ge0}$ for $i = 1,2$ and for every $f \in \mathrm{BL}(S,d)$ there exists $\delta_{3,f} > 0$ such that for the admissible metric $d_{\mathcal{E}(f)}$ on S there exists $\omega_f : [0,\delta_{3,f}] \times M_0 \to \mathbb{R}_+$ continuous, non-decreasing in the first variable, such that the Dini-type condition holds*

$$\int_0^{\delta_{3,f}} \frac{\omega_f(s,\mu_0)}{s} ds < +\infty \quad \textit{for all} \quad \mu_0 \in M_0, \textit{ and} \tag{3.4}$$

$$\left\| P^1_t P^2_t \mu_0 - P^2_t P^1_t \mu_0 \right\|^*_{\mathrm{BL},d_{\mathcal{E}(f)}} \le t\omega_f(t,\mu_0)$$

for every $t \in [0,\delta_{3,f}], \mu_0 \in M_0$.

Assumption 4 (Extended Commutator Condition)**.** *Assume that Assumption 3 holds and, in addition, for every $f \in \mathrm{BL}(S,d)$, there exists $\delta_{4,f} > 0$ and for $\mu_0 \in M_0$ there exists $C_f(\mu_0) > 0$ such that for every $t \in [0,\delta_{4,f}]$,*

$$\omega_f(t, P\mu_0) \le C_f(\mu_0)\omega_f(t,\mu_0)$$

for all $P \in \mathcal{P}^2(\delta_{4,f}) \cdot \mathcal{F}(\delta_{4,f}) \cdot \mathcal{P}^1(\delta_{4,f})$.

Now we can formulate the main theorem of this chapter, which is the strong convergence of the Lie-Trotter scheme. The proof of Theorem 3.2.2 can be found in Section 3.5.

Theorem 3.2.2. *Let $(P^1_t)_{t\ge0}$ and $(P^2_t)_{t\ge0}$ be semigroups of Markov operators. Assume* *that Assumptions 1-4 hold. Then for every $t \geq 0$ there exists a unique Markov operator $\overline{\mathbb{P}}_t : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ such that for every $\mu \in \mathcal{M}^+(S)$:*

$$\left\| \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^n \mu - \overline{\mathbb{P}}_t \mu \right\|^*_{\mathrm{BL},d} \to 0 \text{ as } n \to \infty \tag{3.5}$$

If, additionally, a single $\delta_{3,f}$, $\delta_{4,f}$, $C_f(\mu_0)$ and $\omega_f(\cdot, f)$ can be chosen in (A3) and (A4) to hold uniformly for $f \in \mathrm{BL}(S,d)$, $\|f\|_{\mathrm{BL},d} \leq 1$, then convergence in (3.5) is uniform for t in compact subsets of $\mathbb{R}_+$.

3.3 Preliminaries

3.3.1 Markov operators and semigroups

We start with some preliminary results on Markov operators on spaces of measures, see [Wor10, EK86, LM00]. Let S be a Polish space, $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ a Markov operator. We extend P to a positive bounded linear operator on $(\mathcal{M}(S), \|\cdot\|_{TV})$ by $P\mu := P\mu^+ - P\mu^-$. P is a bounded linear operatos on $\mathcal{M}(S)$ for $\|\cdot\|_{TV}$. 'Typically' it is not bounded for $\|\cdot\|^*_{\mathrm{BL},d}$. Denote by $\mathrm{BM}(S)$ the space of all bounded Borel measurable functions on S. Following [HW09b], Definition 3.2 or [SM03] we will call a Markov operator P *regular* if there exists $U : \mathrm{BM}(S) \to \mathrm{BM}(S)$ such that

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

Let (S, Σ) be a measurable space. According to [Wor10], Proposition 3.3.3, P is regular if and only if

(i) $x \mapsto P\delta_x(E)$ is measurable for every $E \in \Sigma$ and

(ii) $P\mu(E) = \int_S P\delta_x(E) d\mu(x)$ for all $E \in \Sigma$.

We call the operator $U : \mathrm{BM}(S) \to \mathrm{BM}(S)$ the *dual operator* of P. The Markov operator P is a *Markov-Feller operator* if it is regular and the dual U maps $C_b(S)$ into itself. A Markov semigroup $(P_t)_{t\geq 0}$ on S is a semigroup of Markov operators on $\mathcal{M}^+(S)$. The Markov semigroup is regular (or Feller) if all the operators P_t are regular (or Feller). Then $(U_t)_{t\geq 0}$ is a semigroup on $\mathrm{BM}(S)$, which we call the *dual semigroup*.

3.3.2 Topological preliminaries

Following [Kel55], p.230, a topological space X is a *k-space* if for any subset A of X holds that if A intersects each closed compact set in a closed set, then A is closed. According to [Eng77], Theorem 3.3.20 every first-countable Hausdorff space is a k-space. Every metric space is first countable, hence also a k-space. In particular $(\mathcal{M}^+(S), \|\cdot\|^*_{\mathrm{BL},d})$ is a k-space. Let $\mathcal{F}$ be a family of continuous maps from a topological space X to a metric space (Y, d_Y). $\mathcal{F}$ is *equicontinuous at point* $x \in X$ if for every $\varepsilon > 0$ there exists an open neighbourhood U_ε of X in X such that

$$d_Y(f(x), f(x')) < \varepsilon \text{ for all } x' \in U_\varepsilon, \forall f \in \mathcal{F}.$$

A family $\mathcal{F}$ of maps is *equicontinuous* if and only if it is *equicontinuous at every point.* A family $\mathcal{F}$ of maps from a metric space (X, d_X) to a metric space (Y, d_Y) is *uniformly equicontinuous* if for every $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$d_Y(f(x), f(x')) < \varepsilon \text{ for all } x, x' \in X \text{ such that } d_X(x, x') < \delta_\varepsilon \text{ for all } f \in \mathcal{F}.$$

Lemma 3.3.1. *Let (K, d) be a compact metric space and (Y, d_Y) a metric space. An equicontinuous family $\mathcal{F} \subset \mathcal{C}(K, Y)$ is uniformly equicontinuous.*

Proof. Let $\varepsilon > 0$. For each $x \in K$ there exists an open ball $B_x(\delta_x)$, $\delta_x > 0$ such that $d_Y(f(f), f(x')) < \varepsilon$ for every $x' \in B_x(\delta_x)$ and $f \in \mathcal{F}$. By compactness of K, it is covered by finitely many balls, say $B_{x_i}(\delta_{x_i}/2), i = 1, \cdots, n$. Let $\delta := \min_i \frac{\delta_{x_i}}{2}$. If $x, x' \in K$ are such that $d(x, x') < \delta$, then there exists x_{i_0} such that $x \in B_{x_{i_0}}(\delta_{x_{i_0}}/2)$. Necessarily,

$$d(x', x_{i_0}) \leq d(x', x) + d(x, x_{i_0}) < \delta + \delta_{x_{i_0}}/2 < \delta_{x_{i_0}}.$$

Thus, $d_Y(f(x), f(x')) < \varepsilon$, proving the uniform equicontinuity on K. □

For a family of maps $\mathcal{F}$ on X and $x \in X$ we write $\mathcal{F}[x] := \{f(x) : f \in \mathcal{F}\}$. Following [Kel55] we introduce the compact-open topology. Let X, Y be topological spaces. Let F denote a non-empty set of functions from X to Y. For each subset K of X and each subset U of Y, define $W(K, U)$ to be the set of all members of F which carry K into U; that is $W(K, U) := \{f : f[K] \subset U\}$. The family of all sets of the form $W(K, U)$, for K a compact subset of X and U open in Y, is a subbase for the compact-open topology for F. The family of finite intersections of sets of the form $W(K, U)$ is then a base for the compact open topology. We write co-topology as abbreviation for compact-open topology. For two topological spaces T and T, $C(T,T')$ is the set of continuous maps from T to T'. The following generalized Arzela-Ascoli type theorem is based on [Kel55], Theorem 7.18.

Theorem 3.3.2. *Let $\mathcal{C}$ be the family of all continuous maps from a k-space X which is either Hausdorff or regular to a metric space (Y,d), and let $\mathcal{C}$ have the co-topology. Then a subfamily $\mathcal{F}$ of $\mathcal{C}$ is compact if and only if:*

(a) $\mathcal{F}$ is closed in $\mathcal{C}$;

(b) the closure of $\mathcal{F}[x]$ in Y is compact for each x in X;

(c) $\mathcal{F}$ is equicontinuous on every compact subset of X.

Theorem 3.3.3. *[Bargley and Young [RJ66], Theorem 4] Let X be a Hausdorff k-space and Y a Hausdorff uniform space. Let $\mathcal{F} \subset C(X,Y)$. Then $\mathcal{F}$ is compact in the co-topology if and only if*

(a) $\mathcal{F}$ is closed;

(b) $\mathcal{F}[x]$ has compact closure for each $x \in X$;

(c) $\mathcal{F}$ is equicontinuous.

This is a generalization of Theorem 8.2.10 in [Eng77]. This yields the conclusion that for a closed family of continuous functions $\mathcal{F}$ such that $\mathcal{F}[x]$ is precompact for every x, equicontinuity on compact sets is equivalent to continuity.
Moreover, Theorem 3.3.3 can be rephrased for a family $\mathcal{F}$ that is relatively compact in $\mathcal{C}$, meaning that its (compact-open) closure is compact:

Theorem 3.3.4. *Let X be a Hausdorff k-space and Y a metric space. Let $\mathcal{C} = C(X,Y)$, equipped with the co-topology. A subset $\mathcal{F}$ of $\mathcal{C}$ is relatively compact iff:*

(a) The closure of $\mathcal{F}[x] := \{f(x) : f \in \mathcal{F}\}$ in Y is compact for every $x \in X$.

(b) $\mathcal{F}$ is equicontinuous on every compact subset of X.

Statement (b) can be replaced by

(b') $\mathcal{F}$ is equicontinuous on X.

Proof. Let $\overline{\mathcal{F}}$ be the closure of $\mathcal{F}$ in $\mathcal{C}$. Assume it is compact, then according to Theorem 3.3.2, the closure of $\overline{\mathcal{F}}[x]$ in Y is compact for every $x \in X$. Hence the closure of $\mathcal{F}[x]$, which is contained in the closure of $\overline{\mathcal{F}}[x]$, will be compact too. The family $\mathcal{F}$ is equicontinuous on X for every compact subset of X, because it is a subset of $\overline{\mathcal{F}}$ that has his property.
On the other hand, if $\mathcal{F}$ satisfies (a) and (b), or (b'), then $\overline{\mathcal{F}}$ obviously satisfies condition

(a) in Theorem 3.3.2. Now let $f \in \overline{\mathcal{F}}$. Then there exists a net $(f_\nu) \subset \mathcal{F}$ such that $f_\nu \to f$. Point evaluation at x is continuous for the co-topology, so $f_\nu(x) \to f(x)$ in Y. Since $f_\nu(x)$ is contained in a compact set in Y for every ν, $f(x)$ will be contained in this compact set too. So (b) holds in Theorem 3.3.2 for $\overline{\mathcal{F}}$. In a similar way one can show (c) in Theorem 3.3.2. Let $K \subset X$ be compact. The co-topology on $C(X,Y)$ is identical to the topology of uniform convergence on compact subsets ([Kel55], Theorem 7.11). So if $f_* \in \overline{\mathcal{F}}$ and $(f_\nu) \subset \mathcal{F}$ is a net such that $f_\nu \to f_*$, then $f_\nu|_K \to f_*|_K$ uniformly. If $x_0 \in K$, then for every $\varepsilon > 0$ there exists an open neighbourhood U of x_0 in K such that

$$d_Y(f(x), f(x_0)) < \tfrac{1}{2}\varepsilon \qquad \text{for all } f \in \mathcal{F},\ x \in U.$$

Consequently,

$$d_Y(f_*(x), f_*(x_0)) = \lim_\nu d_Y(f_\nu(x), f_\nu(x_0)) \leq \tfrac{1}{2}\varepsilon < \varepsilon$$

for all $x \in U$. So $\overline{\mathcal{F}}$ is equicontinuous on K too. Theorem 3.3.2 then yields the compactness of $\overline{\mathcal{F}}$ in $\mathcal{C}$, hence the relative compactness of $\mathcal{F}$. □

In [Wor10] and in [HSWZ17] we can find the following result, which will be crucial in the proving norm convergence of the Lie-Trotter product formula.

Theorem 3.3.5. *Let S be complete and separable. Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}_s(S)$ and $N \geq 0$ be such that $\langle \mu_n, f\rangle$ converges as $n \to \infty$ for every $f \in \mathrm{BL}(S) \simeq \mathcal{M}(S)^*_{\mathrm{BL}}$ and*

$$\|\mu_n\|_{TV} \leq N \quad \textit{for every } n \in \mathbb{N}.$$

Then there exists $\mu \in \mathcal{M}(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$ as $n \to \infty$.

3.3.3 Tight Markov operators

Let us now introduce the concept of tightness of sets of measures and families of Markov operators. According to [Bog07a], Theorem 7.1, all Borel measures on a Polish space are Radon i.e. locally finite and inner regular. Also, by Definition 8.6.1 in [Bog07a] we say that a family of Radon measures $\mathcal{M}$ on a topological space S is called *uniformly tight* if for every $\varepsilon > 0$, there exists a compact set K_ε such that $|\mu|(S \setminus K_\varepsilon) < \varepsilon$ for all $\mu \in \mathcal{M}$. Moreover, we say that a family $(P_\lambda)_{\lambda\in\Lambda}$ of Markov operators is *tight* if for each $\mu \in \mathcal{M}^+(S)_{\mathrm{BL}}$, $\{P_\lambda\mu : \lambda \in \Lambda\}$ is *uniformly tight*. The following theorem, which is a rephrased version of Theorem 8.6.2 in [Bog07a], due to Prokhorov shows that in our case tightness of the $\|\cdot\|_{TV}$-uniformly bounded family is equivalent to precompactness of $\{P_\lambda\mu \,|\, \lambda \in \Lambda\}$ in $\mathcal{M}^+(S)_{\mathrm{BL}}$.

Theorem 3.3.6 (Prokhorov theorem). *Let S be a complete separable metric space and let M be a family of finite Borel measures on S. The following conditions are equivalent:*

(i) Every sequence $\{\mu_n\} \subset M$ contains a weakly convergent subsequence.

(ii) The family M is uniformly tight and uniformly bounded in total variation norm.

3.4 Equicontinuous families of Markov operators

Let S be a Polish space and consider a semigroup $(P_t)_{t\geq 0}$ of Markov operators. We will examine the properties of equicontinuous families of Markov operators. An equicontinuous family of Markov operators must consist of $\|\cdot\|^*_{\mathrm{BL},d}$-continuous operators. These are Feller ([Wor10], Lemma 7.2.1). Due to Theorem 3.3.2, a closed subset F of the mappings from $\mathcal{M}^+(S)_{\mathrm{BL}}$ to $\mathcal{M}^+(S)_{\mathrm{BL}}$ with the co-topology is compact if and only if $F|_K$ is equicontinuous for each compact $K \subset \mathcal{M}^+(S)$ and the set $\{P_t\mu : P_t \in F\} \subset \mathcal{M}^+(S)$ has a compact closure for every $\mu \in \mathcal{M}^+(S)$. A continuous function on a compact metric space is uniformly continuous. A similar statement holds for equicontinuous families.

Lemma 3.4.1. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of Markov operators on S. If $(P_\lambda)_{\lambda\in\Lambda}$ is an equicontinuous family on the compact set $K \subset \mathcal{M}^+(S)$, then $(P_\lambda)_{\lambda\in\Lambda}$ is uniformly equicontinuous on K.*

The following result, found in [HSWZ17] and based on [Wor10], Theorem 7.2.2, gives equivalent conditions for a family of regular Markov operators to be equicontinuous:

Theorem 3.4.2. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of regular Markov operators on the complete separable metric space (S,d). Let U_λ be the dual operator of P_λ. Then the following statements are equivalent:*

(i) $(P_\lambda)_{\lambda\in\Lambda}$ is an equicontinuous family;

(ii) $(U_\lambda f)_{\lambda\in\Lambda}$ is an equicontinuous family in $C_b(S)$ for all $f \in \mathrm{BL}(S,d)$;

(iii) $\{U_\lambda f | f \in B, \lambda \in \Lambda\}$ is an equicontinuous family for every bounded set $B \subset \mathrm{BL}(S,d)$.

In the next part of this section we show results which allow us to prove Theorem 3.2.1, that is that the composition of an equicontinuous family of Markov operators with an equicontinuous and tight family of Markov operators is equicontinuous. Additionally, if both families are tight, the composition is also tight. One can find an example of equicontinuous and tight families of Markov operators in [Sza03].

Let us first prove the following crucial observation.

Lemma 3.4.3. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be an equicontinuous and tight family of Markov operators on (S,d) and let $K \subset \mathcal{M}^+(S)_{\mathrm{BL}}$ be precompact. Then $\{P_\lambda\mu \,|\, \mu \in K, \lambda \in \Lambda\} \subset \mathcal{M}^+(S)_{\mathrm{BL}}$ is precompact.*

Proof. As K is precompact, then $\overline{K}$ is compact in $\mathcal{M}^+(S)_{\mathrm{BL}}$. So $(P_\lambda|_{\overline{K}}) \subset C(\overline{K}, \mathcal{M}^+(S)_{\mathrm{BL}})$ is equicontinuous and for each $\mu \in \bar{K}$, $\{P_\lambda\mu|\lambda \in \Lambda\}$ is precompact, by tightness of the family $(P_\lambda)_{\lambda\in\Lambda}$. Hence, by Theorems 3.3.2 - 3.3.3, $\{P_\lambda|_{\overline{K}}\} \subset C(\overline{K}, \mathcal{M}^+(S)_{\mathrm{BL}})$ is relatively compact for the compact-open topology, which is the $\|\cdot\|_\infty$-norm topology in this case. Let us consider the evaluation map

$$\begin{array}{rcl} ev : C(\overline{K}, \mathcal{M}^+(S)_{\mathrm{BL}}) \times \overline{K} & \to & \mathcal{M}^+(S)_{\mathrm{BL}} \\ (F,\mu) & \mapsto & F(\mu). \end{array}$$

Theorem 5, [Kel55], p.223 yields that this map is jointly continuous if $C(\overline{K}, \mathcal{M}^+(S)_{\mathrm{BL}})$ is equipped with the co-topology. So

$$K' = \{F(\mu) \,|\, F \in \mathrm{Cl}(\{P_\lambda|_{\overline{K}} : \lambda \in \Lambda\}), \mu \in \overline{K}\}$$

is compact in $\mathcal{M}^+(S)_{\mathrm{BL}}$. □

To prove Theorem 3.2.1, we will need the following result.

Proposition 3.4.4. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a tight family of regular Markov operator on S. If $(P_\lambda)_{\lambda\in\Lambda}$ is equicontinuous for one admissible metric on S, then it is equicontinuous for any admissible metric.*

The key point in the proof of Proposition 3.4.4 is a series of results on characterisation of compact sets in the space of continuous maps when equipped with the co-topology. These can be stated in quite some generality, originating in [Kel55, Eng77, RJ66].

Proof. Let d be the admissible metric on S for which (P_λ) is equicontinuous in $\mathcal{C}_d := C(\mathcal{P}(S)_{\mathrm{weak}}, \mathcal{P}(S)_{\mathrm{BL},d})$. Let d' be any other admissible metric on S. We must show that (P_λ) is an equicontinuous family in $\mathcal{C}_{d'} := C(\mathcal{P}(S)_{\mathrm{weak}}, \mathcal{P}(S)_{\mathrm{BL},d'})$.

By assumption, $\{P_\lambda\mu : \lambda \in \Lambda\}$ is tight for every $\mu \in \mathcal{P}(S)$. By Prokhorov's Theorem (see [Bog07a], Theorem 8.6.2), it is relatively compact in $\mathcal{P}(S)_{\mathrm{BL},d}$, because the $\|\cdot\|_{\mathrm{BL},d}$-norm topology coincides with the weak topology on $\mathcal{M}^+(S)$. Because (P_λ) is equicontinuous in $\mathcal{C}_d$, Theorem 3.3.4 yields that (P_λ) is relatively compact in $\mathcal{C}_d$, for the co-topology. Since the topologies on $\mathcal{P}(S)$ defined by the norms $\|\cdot\|_{\mathrm{BL},d'}$, d' admissible, all coincide with the weak topology, (P_λ) is relatively compact in $\mathcal{C}_{d'}$ for any admissible metric d'.

Again the application of Theorem 3.3.4, but now in opposite direction, yields that (P_λ) is equicontinuous in $\mathcal{C}_{d'}$. □

Proposition 3.4.5. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of Markov operators on (S,d). If $(P_\lambda)_{\lambda\in\Lambda}$ is tight, then the following are equivalent:*

(i) For every $K \subset \mathcal{M}(S)^+_{\mathrm{BL}}$ precompact, $(P_\lambda|_K)_{\lambda\in\Lambda}$ is equicontinuous on K.

(ii) $(P_\lambda)_{\lambda\in\Lambda}$ is equicontinuous (on S).

To prove Proposition 3.4.5 we apply Theorem 3.3.2 and Theorem 3.3.3 to the k-space $(\mathcal{M}^+(S)_{\mathrm{BL}}, \|\cdot\|^*_{\mathrm{BL},d})$.

Now we are in a position to prove Theorem 3.2.1.

Proof. (Theorem 3.2.1) Let $(P_\lambda)_{\lambda\in\Lambda}$ and $(Q_\gamma)_{\gamma\in\Gamma}$,with families of dual operators $(U_\lambda)_{\lambda\in\Lambda}$ and $(V_\gamma)_{\gamma\in\Gamma}$ respectively, be equicontinuous. Let $f \in \mathrm{BL}(S,d)$. Then $\{U_\lambda f|\lambda \in \Lambda\} = \mathcal{E}$ is equicontinuous. Let $d_{\mathcal{E}}$ be the associated admissible metric as defined in (3.3) with $\mathcal{E}(f)$ replaced by $\mathcal{E}$. Then $\mathcal{E}$ is contained in the unit ball $B_{\mathcal{E}}$ of $(\mathrm{BL}(S,d_{\mathcal{E}}), \|\cdot\|_{\mathrm{BL},d_{\mathcal{E}}})$. As $(Q_\gamma)_{\gamma\in\Gamma}$ is an equicontinuous family for d, by Proposition 3.4.4 it is equicontinuous for any admissible metric on S. Hence, it is equicontinuous for $d_{\mathcal{E}}$. Then, by Theorem 3.4.2 (iii)

$$\mathcal{F} = \{V_\gamma g : g \in B_{\mathcal{E}}, \gamma \in \Gamma\} \text{ is equicontinuous in } C_b(S).$$

In particular, as subset of $\mathcal{F}$,

$$\{V_\gamma U_\lambda f : \gamma \in \Gamma, \lambda \in \Lambda\} \text{ is equicontinuous in } C_b(S).$$

Hence, by Theorem 3.4.2, $(P_\lambda Q_\gamma)_{\lambda\in\Lambda,\gamma\in\Gamma}$ is equicontinuous for d. If $(P_\lambda)_{\lambda\in\Lambda}$ is an equicontinuous and tight family, then Lemma 3.4.3 implies that for any $K \subset \mathcal{M}^+(S)_{\mathrm{BL}}$ compact, $K_Q := \{Q_\gamma\nu|\gamma \in \Gamma, \nu \in K\}$ is precompact. Thus, $\{P_\lambda\mu|\lambda \in \Lambda, \mu \in K_Q\} = \{P_\lambda Q_\gamma\nu|\lambda \in \Lambda, \gamma \in \Gamma, \nu \in K\} \subset \mathcal{M}^+(S)_{\mathrm{BL}}$ is precompact. In particular, this holds for for $K = \{\nu_0\}$. □

In the above proof of Theorem 3.2.1 we only need assumption, that the family $(Q_\gamma)_{\gamma\in\Gamma}$ is tight. In case both $(P_\lambda)_{\lambda\in\Lambda}$ and $(Q_\gamma)_{\gamma\in\Gamma}$ are tight, there is an alternative way of proving Theorem 3.2.1 using Lemma 3.4.3.

As a consequence of Theorem 3.2.1 we get the following Corollary.

Corollary 3.4.6. *The composition of a finite number of equicontinuous and tight families of Markov operators is equicontinuous and tight.*

3.5 Proof of convergence of Lie-Trotter product formula

Throughout this section we assume that $(P_t^1)_{t\geq 0}$ and $(P_t^2)_{t\geq 0}$ are Markov-Feller semigroups on S with dual semigroups $(U_t^1)_{t\geq 0}$, $(U_t^2)_{t\geq 0}$, respectively.

We start by examining some consequences of Assumptions 1 - 4 formulated in Section 3.2. Introduce

$$\mathcal{F}_<(\delta) := \left\{ \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^i : n \in \mathbb{N}, i \leq n-1, t \in [0,\delta] \right\}.$$

Lemma 3.5.1. *The following statements hold:*

(i) If Assumption 1 holds, then $\mathcal{P}^1(\delta)$ and $\mathcal{P}^2(\delta)$ are equicontinuous and tight for every $\delta > 0$.

(ii) If $\mathcal{F}(\delta_2)$ is equicontinuous then $\mathcal{F}_<(\delta_2)$ is equicontinuous.

(iii) $\mathcal{F}_<(\delta_2)$ is equicontinuous and tight iff $\mathcal{F}(\delta_2)$ is equicontinuous and tight.

Proof. (i) Is an immediate consequence of Theorem 3.2.1 and the semigroup property of $(P_t^i)_{t\geq 0}$.

(ii) Let $t \in [0,\delta_2]$ and $i, n \in \mathbb{N}$ such that $i \leq n-1$. Observe that $\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^i = \left[P^1_{\frac{1}{i}\frac{it}{n}} P^2_{\frac{1}{i}\frac{it}{n}} \right]^i$ with $\frac{it}{n} \in [0,\delta_2]$. Hence $\mathcal{F}_<(\delta_2) \subset \mathcal{F}(\delta_2)$. A subset of an equicontinuous family of maps is equicontinuous.

(iii) The following subsets of $\mathcal{F}_<(\delta_2)$,

$$\mathcal{F}^1_<(\delta) := \left\{ P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} : n \in \mathbb{N}, t \in [0,\delta] \right\}$$

and

$$\mathcal{F}^*_<(\delta) := \left\{ \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^{n-1} : n \in \mathbb{N}, t \in [0,\delta] \right\}$$

are equicontinuous and tight, because $\mathcal{F}_<(\delta_2)$ is. Note that $\mathcal{F} \subset \mathcal{F}^1_<(\delta_2) \cdot \mathcal{F}^*_<(\delta_2)$. According to Theorem 3.2.1 the latter product is equicontinuous and tight. Hence $\mathcal{F}$ is equicontinuous and tight. In part (ii) we observe that $\mathcal{F}_<(\delta_2) \subset \mathcal{F}(\delta_2)$, so equicontinuity and tightness of $\mathcal{F}(\delta_2)$ implies that of $\mathcal{F}_<(\delta_2)$. $\square$

Lemma 3.5.2 (Eventual equicontinuity)**.** *If Assumptions 1 and 2 hold, then for each* *compact* $\Gamma \subset \mathbb{R}_+$ *there exists* $N = N_\Gamma$ *such that*

$$\mathcal{F}_\Gamma^N := \left\{ \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^n : n \in \mathbb{N}, n \geq N, t \in \Gamma \right\}$$

is equicontinuous.

Proof. Let $N \in \mathbb{N}$ be such that $\frac{t}{N} \leq \min(\delta_1, \delta_2) =: \delta$ for all $t \in \Gamma$. For $n \geq N$ we have, with $k := n - N$

$$\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^n = \left[P^1_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} P^2_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} \right]^{k+N} = \left[P^1_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} P^2_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} \right]^k \left[P^1_{\frac{t}{N+k}} P^2_{\frac{\cdot t}{N+k}} \right]^N .$$

Since $\frac{t}{N+k} \in [0, \delta]$ for $k \in \mathbb{N}_0$ and $\mathcal{P}^1(\delta)$ and $\mathcal{P}^2(\delta)$ are equicontinuous and tight (by assumption), the family $\left\{ \left[P^1_{\frac{t}{N+k}} P^2_{\frac{\cdot t}{N+k}} \right]^N : k \in \mathbb{N}_0, t \in \Gamma \right\}$ is equicontinuous and tight according to Theorem 3.2.1. The family $\left\{ \left[P^1_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} P^2_{\frac{1}{k} \cdot \frac{k \cdot t}{N+k}} \right]^k : k \in \mathbb{N}, t \in \Gamma \right\} \subset \mathcal{F}(\delta_2)$ is equicontinuous by Assumption 2. Hence Theorem 3.2.1 yields equicontinuity of $\mathcal{F}_\Gamma^N$. $\square$

Lemma 3.5.3. *If Assumptions 1 and 2 hold and, additionally,* $\mathcal{F}(\delta)$ *is a tight family for some* $\delta = \delta_2 > 0$*, then* $\mathcal{F}(\delta)$ *is equicontinuous and tight for any* $\delta > 0$*.*

Proof. Let $\delta_2 > 0$ such that Assumption 2 holds for δ_2. Let

$$\begin{aligned} \mathcal{F}(2\delta_2) &:= \left\{ \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^n : t \in [0, 2\delta_2], n \in \mathbb{N} \right\} \\ &= \underbrace{\left\{ \left[P^1_{\frac{t'}{m}} P^2_{\frac{t'}{m}} \right]^{2m} : t' := \frac{t}{2} \in [0, \delta_2], m \in \mathbb{N} \right\}}_{\mathcal{F}_m^{\text{even}}} \cup \underbrace{\left\{ \left[P^1_{\frac{t'}{2m+1}} P^2_{\frac{t'}{2m+1}} \right]^{2m+1} : t' \in [0, \delta_2], m \in \mathbb{N} \right\}}_{\mathcal{F}_m^{\text{odd}}} \end{aligned}$$

Due to Theorem 3.2.1, $\mathcal{F}_m^{\text{even}}(\delta_2)$ is an equicontinuous and tight family as a product of equicontinuous and tight families.

$$\begin{aligned} \mathcal{F}_m^{\text{odd}}(\delta_2) &= \left\{ \left[P^1_{\frac{t_m}{m}} P^2_{\frac{t_m}{m}} \right]^{2m+1} : t_m = t \cdot \tfrac{m}{2m+1}, t \in [0, \delta_2], m \in \mathbb{N} \right\} \\ &\subset \left\{ \left[P^1_{\frac{t_m}{m}} P^2_{\frac{t_m}{m}} \right] \left[P^1_{\frac{t_m}{m}} P^2_{\frac{t_m}{m}} \right]^m \left[P^1_{\frac{t_m}{m}} P^2_{\frac{t_m}{m}} \right]^m : t_m = t \cdot \tfrac{m}{2m+1}, t \in [0, \delta_2], m \in \mathbb{N} \right\} \end{aligned}$$

Hence, due to Theorem 3.2.1, $\mathcal{F}_m^{\text{odd}}(\delta_2)$ is an equicontinuous and tight family. $\square$

Lemma 3.5.4. *Let $f \in \mathrm{BL}(S,d)$ and $\delta = \min(\delta_1, \delta_2)$. If Assumptions 1 and 2 hold, then $\mathcal{E}(f)$ defined by (3.2) is equicontinuous in $C_b(S)$.*

Note that $\mathcal{E}(f)$ depends on the choice of f. Lemma 3.5.4 is a consequence of Assumptions 1 and 2 and Theorem 3.4.2.

Remark 3.5.5. *Technically, one requires that particular subsets of $\mathcal{E}(f)$ are equicontinuous. Namely, that*

$$\mathcal{E}_k(f) = \left\{ U^2_{\frac{lt}{kn}} U^1_{\frac{jt}{kn}} \left[U^2_{\frac{t}{n}} U^1_{\frac{t}{n}} \right]^n f : n, j, l, i \in \mathbb{N}, j \le kn, i \le n-1, l \le kn, t \in [0, \delta_2] \right\}$$

is equicontinuous for every k. This seems to be quite too technical a condition.

Remark 3.5.6. *The commutator condition that we propose in Assumption 3 is weaker than the commutator conditions in [Kuh01], conditions (C) and (C^*) in [CC04] and commutator condition in Proposition 3.5 in [Col09].*

For later reference, we present some properties of function $t \mapsto \omega(t) := \omega_f(t, \mu_0)$, that occurs in Assumptions 3 and 4.

Lemma 3.5.7. *Let $\omega = \omega_f(\cdot, \mu_0) : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous, nondecreasing function such that Dini condition (3.4) in Assumption 3 holds. Then $\lim_{t \to 0^+} \omega(t) = 0$ and for any $0 < a < 1$.*

(a) $\sum_{n=1}^{\infty} \omega(a^n t) < \infty$ *for all $t > 0$;*

(b) $\lim_{t \to 0} \sum_{n=1}^{\infty} \omega(a^n t) = 0$.

Proof. For (a) Suppose that $\inf_{0<t<1} \omega(t) = m > 0$. Then by 3.4 in Assumption 3 we get

$$\int_0^1 \frac{\omega(s)}{s} ds \ge \int_0^1 \frac{m}{s} ds = +\infty.$$

So $m = 0$. From the fact that $\int_0^\sigma \frac{\omega(t)}{t} dt < +\infty$ we have

$$\begin{aligned} \infty &> \textstyle\sum_{n=0}^{\infty} \int_{a^{n+1}t}^{a^n t} \frac{\omega(s)}{s} ds \ge \sum_{n=0}^{\infty} \frac{\omega(a^{n+1}t)}{a^n t} (a^n t - a^{n+1} t) = \\ &= \textstyle\sum_{n=0}^{\infty} \omega(a^{n+1}t) \left[1 - \frac{a^{n+1}t}{a^n t} \right] = (1-a) \sum_{n=1}^{\infty} \omega(a^n t) \end{aligned}$$

This proves (a).

For (b) let $\varepsilon > 0$. According to (a) there exists $n_0 \in \mathbb{N}$ such that

$$\sum_{n=n_0}^{\infty} \omega(a^n) < \frac{\varepsilon}{2}.$$

Moreover, because $\lim_{t\to 0^+} \omega(t) = 0$, there exists $t_0 \leq 1$ such that $\omega(at_0) < \frac{\varepsilon}{2n_0}$. Then for every $0 < t \leq t_0$ and $n \in \mathbb{N}$, $1 \leq n \leq n_0$, $\omega(a^n t) \leq \omega(at_0) \leq \frac{\varepsilon}{2n_0}$. So

$$\sum_{n=1}^{\infty} \omega(a^n t) < \sum_{n=1}^{n_0-1} \omega(a^n t) + \sum_{n=n_0}^{\infty} \omega(a^n t) < \frac{\varepsilon(n_0-1)}{2n_0} + \frac{\varepsilon}{2} < \varepsilon.$$

□

To show our main result we need technical lemmas which we present in this section. Proofs of results from this section can be found in Appendix 3.8.1.

Lemma 3.5.8. *The following identities hold: for fixed $k \in \mathbb{N}$, $m := kn$ and $j \leq m$.*

(a) $P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} = \sum_{l=0}^{j-1} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-1-l)t}{m}}$

(b) $P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^k = \sum_{j=1}^{k-1} P^1_{\frac{tj}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{k-1-j}$

(c) $\left(P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right)^n - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^m = \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} \right)^n - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{n\cdot k} =$
$= \sum_{i=0}^{n-1} \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} \right)^i \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^k \right) \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{k\cdot(n-1-i)}.$

Combining Lemma 3.5.8 (a) - (c) we get the following Corollary.

Corollary 3.5.9. *For any $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $m := kn$ one has*

$$\left(P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right)^n - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^m =$$
$$= \sum_{i=0}^{n-1} \sum_{j=1}^{k-1} \sum_{l=0}^{j-1} \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} \right)^i P^1_{\frac{jt}{m}} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-l)t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{k(n-i)-j-1}$$

Lemma 3.5.10. *Let $f \in \mathrm{BL}(S,d)$ and $\mu_0 \in M_0$. Assume that Assumptions 1 - 4 hold and put $\delta_f = \min(\delta_1, \delta_2, \delta_{3,f}, \delta_{4,f})$. Then for all $t \geq 0$ and $n, k \in \mathbb{N}$ such that $\frac{t}{nk} \in [0, \delta_f]$:*

$$\left| \left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}} \right]^n \mu_0 - \left[P^1_{\frac{t}{kn}} P^2_{\frac{t}{kn}} \right]^{n\cdot k} \mu_0, f \right\rangle \right| \leq C_f(\mu_0) \frac{k-1}{2} t\omega_f \left(\frac{t}{nk}, \mu_0 \right).$$

We can now finally get to the proof of our main result, Theorem 3.2.2, i.e. the convergence of the Lie-Trotter product formula for Markov operators. We need the lemma that yields the convergence of the subsequence of the form $\left\langle \left[P^1_{\frac{t}{2^n}} P^1_{\frac{t}{2^n}}\right]^{2^n} \mu_0, f \right\rangle$ for $\mu_0 \in M_0$ and for every $f \in \mathrm{BL}(S,d)$. Then, using this result, we will show that the sequence $\left\langle \left[P^1_{\frac{t}{n}} P^1_{\frac{t}{n}}\right]^{n} \mu_0, f \right\rangle$ also converges for every $f \in \mathrm{BL}(S,d)$. From that we can extend from $\mu_0 \in M_0$ to $\mu \in \mathcal{M}^+(S)$. Recall that $\delta_f := \min(\delta_1, \delta_2, \delta_{3,f}, \delta_{4,f})$.

Remark 3.5.11. *The "weak" convergence in our setting is a convergence of a sequence of measures paired with a bounded Lipschitz function. Hence it differs from the "standard" definition of weak convergence (see [Bog07a] Definition 8.1.1), where the sequence of measures is paired with continuous bounded functions. However, since* $\mathrm{BL}(S,d) \simeq \mathcal{M}(S)^*_{\mathrm{BL}}$ *(see [HW09b], Theorem 3.7) our terminology is proper from a functional analytical perspective.*

Lemma 3.5.12. *Let* $(P^1_t)_{t\geq 0}$ *and* $(P^2_t)_{t\geq 0}$ *be Markov semigroups such that Assumptions 1 - 4 hold. Let* $\mu_0 \in M_0$ *and* $f \in \mathrm{BL}(S,d)$*. Then the sequence* $(r_n)_{n\in\mathbb{N}}$ *where* $r_n := \left\langle \left[P^1_{\frac{t}{2^n}} P^1_{\frac{t}{2^n}}\right]^{2^n} \mu_0, f \right\rangle$ *converges for every* $t \geq 0$*, uniformly for* t *in compact subsets of* $\mathbb{R}_+$*.*

Proof. The case $t = 0$ is trivial. So fix $t > 0$. Let $f \in \mathrm{BL}(S,d)$. There exists $N \in \mathbb{N}$ such that $\frac{t}{2^N} \in [0, \delta_f]$. Let $i, j \in \mathbb{N}, i > j \geq N$. Then $2^i = 2^j \cdot 2^l$ with $l = i - j < i$. Lemma 3.5.10 yields for any $\mu_0 \in M_0$, that

$$\begin{aligned}
&\left| \left\langle \left(\left[P^1_{\frac{t}{2^i}} P^2_{\frac{t}{2^i}}\right]^{2^i} - \left[P^1_{\frac{t}{2^j}} P^2_{\frac{t}{2^j}}\right]^{2^j} \right) \mu_0, f \right\rangle \right| \\
\leq &\sum_{l=j}^{i-1} \left| \left\langle \left(\left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} - \left[P^1_{\frac{t}{2^{l+1}}} P^2_{\frac{t}{2^{l+1}}}\right]^{2^{l+1}} \right) \mu_0, f \right\rangle \right| \\
\leq &C_f(\mu_0) \frac{t}{2} \sum_{l=j}^{i-1} \omega_f\left(\frac{t}{2^{l+1}}, \mu_0\right),
\end{aligned} \tag{3.6}$$

with ω_f as in Assumption 3. According to Lemma 3.5.7 (a), $\sum_{l=0}^{\infty} \omega_f\left(\frac{t}{2^{l+1}}, \mu_0\right) < +\infty$. So for every $\varepsilon > 0$ there exists $N' \in \mathbb{N}, N' \geq N$ such that $\sum_{l=j}^{i-1} \omega_f\left(\frac{t}{2^{l+1}}, \mu_0\right) < \varepsilon$ for every $i, j \geq N$. Also, by property b) in Lemma 3.5.7, $\omega_f\left(\frac{t}{2^{l+1}}, \mu_0\right)$ can be made uniformly small, when t is in a compact subset of $\mathbb{R}_+$. Hence the sequence $(r_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$, hence convergent. □

Observe that a measure $\mu \in \mathcal{M}^+(S)$ is uniquely defined by its values on $f \in \mathrm{BL}(S,d)$. Lemma 3.5.12 and the Banach-Steinhaus Theorem (see [Bog07b], Theorem 4.4.3) allow us to define a positively homogeneous map $\mathbb{P}_t : M_0 \to \mathrm{BL}(S,d)^*$ by means of

$$\langle \mathbb{P}_t \mu_0, f\rangle := \lim_{n\to\infty} \left\langle \left[P^1_{\frac{t}{2^n}} P^2_{\frac{t}{2^n}}\right]^{2^n} \mu_0, f\right\rangle.$$

However, according to Theorem 3.3.5, $\mathbb{P}_t\mu_0 \in \mathcal{M}^+(S)$ for every $\mu_0 \in M_0$ and

$$\left[P^1_{\frac{t}{2^n}} P^2_{\frac{t}{2^n}}\right]^{2^n} \mu_0 \to \mathbb{P}_t\mu_0 \tag{3.7}$$

strongly, in $\|\cdot\|^*_{\mathrm{BL},d}$-norm.

Proposition 3.5.13. *Let $(P^1_t)_{t\ge 0}$ and $(P^2_t)_{t\ge 0}$ be Markov semigroups such that Assumptions 1 - 4 hold. If $\mu_0 \in M_0$, then for every $f \in \mathrm{BL}(S,d)$ and for all $t \ge 0$, $\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \mu_0, f\right\rangle$ converges to $\langle \mathbb{P}_t\mu_0, f\rangle$.*

Proof. Let $f \in \mathrm{BL}(S)$, $t \ge 0$ and fix $\varepsilon > 0$. Put $\delta_f = \min(\delta_1, \delta_2, \delta_{3,f}, \delta_{4,f})$. For any $l \in \mathbb{N}$, using Lemma 3.5.10, one has

$$\begin{aligned}
\left|\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| \quad \le \quad & \left|\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \mu_0 - \left[P^1_{\frac{t}{n2^l}} P^1_{\frac{t}{n2^l}}\right]^{n2^l} \mu, f\right\rangle\right| \\
+ \quad & \left|\left\langle \left[P^1_{\frac{t}{n2^l}} P^1_{\frac{t}{n2^l}}\right]^{n2^l} \mu_0 - \left[P^1_{\frac{t}{2^l}} P^1_{\frac{t}{2^l}}\right]^{2^l} \mu_0, f\right\rangle\right| \\
+ \quad & \left|\left\langle \left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right|.
\end{aligned}$$

Pick N such that for $n \ge N$ one has $\frac{t}{n} \in [0, \delta_f]$. Then

$$\begin{aligned}
\left|\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| \quad \le \quad & \textstyle\sum_{i=0}^{l-1} \left|\left\langle \left[P^1_{\frac{t}{2^i n}} P^2_{\frac{t}{2^i n}}\right]^{2^i n} \mu_0 - \left[P^1_{\frac{t}{2^{i+1}n}} P^2_{\frac{t}{2^{i+1}n}}\right]^{2^{i+1}n} \mu_0, f\right\rangle\right| \\
& + C_f(\mu_0)\tfrac{n-1}{2} t\omega_f\left(\tfrac{t}{n2^l}, \mu_0\right) \\
& + \left|\left\langle \left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| \\
\le \quad & \textstyle\sum_{i=0}^{l-1} C_f(\mu_0)\tfrac{1}{2} t\omega_f\left(\tfrac{t}{2^i n}, \mu_0\right) + C_f(\mu_0)\tfrac{n-1}{2} t\omega_f\left(\tfrac{t}{n2^l}, \mu_0\right) \\
& + \left|\left\langle \left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| \\
= \quad & \tfrac{1}{2} C_f(\mu_0) t\left[\textstyle\sum_{i=0}^{l} \omega_f\left(\tfrac{t}{2^i n}, \mu_0\right) + (n-1)\omega_f\left(\tfrac{t}{n2^l}, \mu_0\right)\right] \\
& + \left|\left\langle \left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right|.
\end{aligned}$$

According to Proposition 3.5.13 there exists N_0 such that for any $l \geq N_0$

$$\left|\left\langle \left[P^1_{\frac{t}{2^l}} P^2_{\frac{t}{2^l}}\right]^{2^l} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| < \frac{\varepsilon}{3}.$$

Lemma 3.5.7 (b) yields $N_1 \in \mathbb{N}$, $N_1 \geq N$ such that for every $n \geq N_1$ and $l \in \mathbb{N}$,

$$\sum_{i=0}^{l} \omega_f\left(\frac{t}{2^i n}, \mu_0\right) \leq \sum_{i=0}^{\infty} \omega_f\left(\frac{t}{2^i n}, \mu_0\right) < \left(1 + \frac{1}{2}C_f(\mu_0)t\right)^{-1} \frac{\varepsilon}{3}.$$

Since $\omega_f(s, \mu_0) \downarrow 0$ as $s \downarrow 0$, for every $n \geq N_1$, there exists $l_n \geq N_0$ such that

$$\omega_f\left(\frac{t}{n 2^{l_n}}, \mu_0\right) < \frac{1}{n-1}\left(1 + \frac{1}{2}tC_f(\mu_0)\right)^{-1} \frac{\varepsilon}{3}.$$

So by choosing $l = l_n$ in the above derivation, we get that

$$\left|\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^{n} \mu_0 - \mathbb{P}_t\mu_0, f\right\rangle\right| < \varepsilon \text{ for every } n \geq N_1.$$

□

The next lemma shows that once the convergence of $\left\langle \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^{n} \mu_0, f\right\rangle$ is established for $\mu_0 \in M_0$ then we have convergence for all $\mu \in \mathcal{M}^+(S)$.

Lemma 3.5.14. *Assume that Assumptions 1 - 4 hold. Then for every* $\mu \in \mathcal{M}^+(S)$ *and* $t \geq 0$, $\left(\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^{n} \mu\right)_{n\in\mathbb{N}}$ *is a Cauchy sequence in* $\mu \in \mathcal{M}^+(S)$ *for* $\|\cdot\|^*_{\mathrm{BL},d}$.

Proof. Let $\mu \in \mathcal{M}^+(S)$. Let $\epsilon > 0$. By Assumption 2, $\mathcal{F}(\delta)$ is an equicontinuous family. Thus there exists $\delta_\epsilon > 0$ such that

$$\left\|\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^{n} \mu - \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^{n} \nu\right\|^*_{\mathrm{BL},d} < \epsilon/3$$

for all $\nu \in \mathcal{M}^+(S)$ such that $\|\mu - \nu\|^*_{\mathrm{BL},d} < \delta_\epsilon$. As $M_0 \subset \mathcal{M}^+(S)$ dense, there exists $\mu_0 \in M_0$

such that $\|\mu - \mu_0\|^*_{\mathrm{BL},d} < \delta_\epsilon$. Then

$$\begin{aligned}\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu\right\|^*_{\mathrm{BL},d} \leq\ & \left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu - \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_0\right\|^*_{\mathrm{BL},d} \\ &+ \left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_0 - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu_0\right\|^*_{\mathrm{BL},d} \\ &+ \left\|\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu_0 - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu\right\|^*_{\mathrm{BL},d}\end{aligned} \tag{3.8}$$

According to Proposition 3.5.13 and Theorem 3.3.5, there exists $N \in \mathbb{N}$ such that for $n, m \geq N$,

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_0 - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu_0\right\|^*_{\mathrm{BL},d} < \epsilon/3.$$

Hence for $n, m \geq N$, we obtain for (3.8) that

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^m \mu\right\|^*_{\mathrm{BL},d} < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

which proves that $\left(\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu\right)_n$ is a Cauchy sequence. □

Lemma 3.5.14 allows us to define for $\mu \in \mathcal{M}^+(S)$ and $t \in [0, \delta]$

$$\bar{\mathbb{P}}_t\mu := \lim_{n\to\infty}\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu$$

as a limit in $\mathcal{M}^+(S)_{\mathrm{BL}}$. Then $\bar{\mathbb{P}}_t\mu_0 = \mathbb{P}_t\mu_0$ for $\mu_0 \in M_0$, according to Proposition 3.5.13.

Thus, as a consequence of Lemma 3.5.14 we have proven the first part of Theorem 3.2.2.

Concerning the second part of the proof: the arguments in the proofs of the lemmas and propositions that together finish the proof of Theorem 3.2.2, show upon inspection that in case where stronger versions of Assumptions 3 and 4 hold, then immediately $\|\cdot\|^*_{\mathrm{BL},d}$-norm estimates can be obtained. That is, if in Assumptions 3 and 4 a single $\delta_{3,f}$, $\delta_{4,f}$. $C_f(\mu_0)$ and $\omega_f(\cdot, \mu_0)$ can be chosen to hold uniformly for f in the unit ball of $\mathrm{BL}(S, d)$, then one obtains Theorem 3.2.2 (i.e. norm-convergence of the Lie-Trotter product) without the need of Theorem 3.3.5. Then one easily checks that convergence is uniform in t in compact subsets of $\mathbb{R}_+$. In fact for $\mu \in M_0$ this result is captured in the preceding remarks. Let $\Gamma \subset \mathbb{R}_+$ be compact. According to Lemma 3.5.2, $\mathcal{F}^N_\Gamma$ is equicontinuous for N sufficiently large. Then all estimates in the proof of Lemma 3.5.14 can be made uniformly in $t \in \Gamma$.

Moreover, in the situation described above, the rate of convergence of the Lie-Trotter product is controlled by properties of $\omega(\cdot, \mu_0)$, according to the proof of Proposition 3.5.13.

3.6 Properties of the limit

Let us now analyse the properties of the limit operator family $(\overline{\mathbb{P}}_t)_{t\geq 0}$ as obtained by the Lie-Trotter product formula. First we show that $\overline{\mathbb{P}}_t$ is a Feller operator, i.e. it is continuous on $\mathcal{M}^+(S)$ for $\|\cdot\|^*_{\mathrm{BL,d}}$.

3.6.1 Feller property

Lemma 3.6.1. *Let $(P^1_t)_{t\geq 0}$ and $(P^2_t)_{t\geq 0}$ be semigroups of regular Markov-Feller operators that satisfy Assumptions 1 - 4. Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{M}^+(S)$ and $\mu^* \in \mathcal{M}^+(S)$ be such that $\mu_n \to \mu^*$ in $\mathcal{M}^+(S)_{\mathrm{BL}}$ as $n\to\infty$. Then $\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_n \to \overline{\mathbb{P}}_t\mu^*$ in $\mathcal{M}^+(S)_{\mathrm{BL}}$ for $t\in[0,\delta_2]$.*

Proof. Let $\epsilon > 0$. From Assumption 2 (stability) we get that there exists $\delta_\epsilon > 0$ such that

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu - \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu^*\right\|^*_{\mathrm{BL},d} < \varepsilon/2$$

for every $\nu \in \mathcal{M}^+(S)$ such that $\|\mu-\mu^*\|^*_{\mathrm{BL},d} < \delta_\epsilon$ for all $t\in[0,\delta_2]$. Since $\mu_n \to \mu^*$, there exists $N_0\in\mathbb{N}$ such that

$$\|\mu_n-\mu^*\|^*_{\mathrm{BL},d_{\mathcal{E}(f)}} < \delta_\epsilon$$

for all $n\geq N_0$. From Theorem 3.2.2 we know that there exists $N_1\in\mathbb{N}$ such that for every $n\geq N_1$

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu^* - \overline{\mathbb{P}}_t\mu^*\right\|_{\mathrm{BL},d} < \epsilon/2.$$

Then for $n\geq N := \max(N_0,N_1)$,

$$\begin{aligned}\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_n - \overline{\mathbb{P}}_t\mu^*\right\|^*_{\mathrm{BL},d} &\leq \left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu_n - \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu^*\right\|_{\mathrm{BL},d} \\ &+ \left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \mu^* - \overline{\mathbb{P}}_t\mu^*\right\|^*_{\mathrm{BL},d_{\mathcal{E}(f)}} < \epsilon. \end{aligned}$$ □

Proposition 3.6.2. *If Assumptions 1 - 4 then for all $k\in\mathbb{N}, t\geq 0$*

$$\overline{\mathbb{P}}_{kt}\mu = \overline{\mathbb{P}}^k_t\mu \text{ for all } \mu\in\mathcal{M}^+(S).$$

In particular, $\overline{\mathbb{P}}_t\overline{\mathbb{P}}_s\mu = \overline{\mathbb{P}}_{t+s}\mu$ for all $t,s\geq 0$ such that $\frac{t}{s}\in\mathbb{Q}$.

Proof. Let $\mu\in\mathcal{M}^+(S)$. Let $\epsilon>0$. Without loss of generality we can assume that $t\in[0,\delta_2]$. For $k=1$ the statement is obviously true. Assume it has been proven for k. We now show it holds for $k+1$ as well. As we know that the limit of the Lie-Trotter product exists (Theorem 3.2.2), we can consider in the limit any subsequence. Take $n=(k+1)m$, $m\to\infty$:

$$\overline{\mathbb{P}}_{(k+1)t}\mu = \lim_{m\to\infty}\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{(k+1)m}\mu = \lim_{m\to\infty}\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\left(\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\right).$$

Hence there exists $N_0\in\mathbb{N}$ such that for all $m>N_0$,

$$\left\|\overline{\mathbb{P}}_{(k+1)t}\mu - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\left(\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\right)\right\|^*_{BL,d} < \frac{\epsilon}{3}.$$

Since by assumption $\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\to\overline{\mathbb{P}}_{kt}\mu$, Lemma 3.6.1 yields that there exists $N_1\geq N_0$ such that for $m\geq N_1$:

$$\left\|\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\left(\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\right) - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\overline{\mathbb{P}}_{kt}\mu\right\|^*_{BL,d} < \frac{\epsilon}{3}.$$

Also, by Theorem 3.2.2 we get $N_2\geq N_1$ such that for every $m\geq N_2$

$$\left\|\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\overline{\mathbb{P}}_{kt}\mu - \overline{\mathbb{P}}^{k+1}_t\mu\right\|^*_{BL,d} < \frac{\epsilon}{3}.$$

Hence for $m\geq N_2$,

$$\begin{aligned}\left\|\overline{\mathbb{P}}_{(k+1)t}\mu - \overline{\mathbb{P}}^{k+1}_t\mu\right\|^*_{BL,d} &\leq \left\|\overline{\mathbb{P}}_{(k+1)t}\mu - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\left(\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\right)\right\|^*_{BL,d}\\ &+ \left\|\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\left(\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{km}\mu\right) - \left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\mathbb{P}_{kt}\mu\right\|^*_{BL,d}\\ &+ \left\|\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{m}\mathbb{P}_{kt}\mu - \overline{\mathbb{P}}^{k+1}_t\mu\right\|^*_{BL,d} < \epsilon.\end{aligned}$$

If $t,s>0$ are such that $\frac{t}{s}\in\mathbb{Q}$, then there exist $m,r\in\mathbb{N}$: $rt=ms$. Hence, by the first part,

$$\overline{\mathbb{P}}_{t+s}\mu = \overline{\mathbb{P}}_{(m+r)\cdot\frac{s}{r}}\mu = \overline{\mathbb{P}}^{(m+r)}_{\frac{s}{r}}\mu = \overline{\mathbb{P}}^{m}_{\frac{s}{r}}\overline{\mathbb{P}}^{r}_{\frac{s}{r}}\mu = \overline{\mathbb{P}}_t\overline{\mathbb{P}}_s\mu.$$

□

Proposition 3.6.3. $\overline{\mathbb{P}}_t:\mathcal{M}^+(S)_{\mathrm{BL}}\to\mathcal{M}^+(S)_{\mathrm{BL}}$ *is continuous for all* $t\geq 0$.

Proof. First we will get the result for $t\in[0,\delta_2]$. Let $\mu\in\mathcal{M}^+(S)$ and $\epsilon>0$. By Assumption 2, there exists $\delta_\epsilon>0$ such that

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^{n}\mu - \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^{n}\nu\right\|^*_{\mathrm{BL},d} < \frac{\epsilon}{2} \tag{3.9}$$

for every $\nu \in \mathcal{M}^+(S)$ such that $\|\mu-\nu\|^*_{\mathrm{BL},d} < \delta_\epsilon$ and all $n \in \mathbb{N}, t \in [0,\delta_2]$. Then, by taking the limit $n \to \infty$ in (3.9), using Theorem 3.2.2,

$$\left\|\overline{\mathbb{P}}_t\mu - \overline{\mathbb{P}}_t\nu\right\|^*_{\mathrm{BL},d} \leq \frac{\epsilon}{2} < \epsilon$$

for all $\mu,\nu \in \mathcal{M}^+(S)$ such that $\|\mu-\nu\|^*_{\mathrm{BL},d} < \delta_\epsilon$. So $\overline{\mathbb{P}}_t$ is continuous for all $t \in [0,\delta_2]$. Now we can use Proposition 3.6.2 to extend the result to all $t \geq 0$. □

In the proof we actually show more, which we formulate as a corollary.

Corollary 3.6.4. *The family $\overline{\mathcal{P}}(\delta) = \left\{\overline{\mathbb{P}}_t : t \in [0,\delta]\right\}$ is equicontinuous for every $0 < \delta \leq \delta_2$.*

3.6.2 Semigroup property

Let us now analyze the full semigroup property of the limit. Recall Proposition 3.6.2. The extension to all pairs $t,s \in \mathbb{R}_+$ of the semigroup property is not obvious. We do not assume any continuity of Markov semigroups. However, let us show the following:

Proposition 3.6.5. *Assume that Assumptions 1-4 hold and additionally that $t \mapsto P^i_t\mu : \mathbb{R}_+ \to \mathcal{M}^+(S)_{\mathrm{BL}}$ are continuous for $i = 1,2$ and all $\mu \in \mathcal{M}^+(S)$. Then $(\overline{\mathbb{P}}_t)_{t\geq 0}$ is strongly continuous and it is a semigroup.*

Proof. Put $\mathbb{Q}^n_t := \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n$. If $\mu_0 \in M_0$, then by the strong continuity of the semigroup $(P^i_t)_{t\geq 0}$ on $\mathcal{M}^+(S)$, we obtain that $F_n : \mathbb{R}_+ \to \mathbb{R} : t \mapsto \langle \mathbb{Q}^n_t\mu_0, f\rangle$ is continuous for all $n \in \mathbb{N}$. According to Lemma 3.5.12, F_{2^N} converges uniformly on compact subsets of $\mathbb{R}_+$ to $t \mapsto \langle \overline{\mathbb{P}}\mu_0, f\rangle$. Hence the latter function is continuous on $\mathbb{R}_+$.

Now, first take $t^* \in [0,\delta_2)$ and $(t_k)_k \subset [0,\delta_2)$ such that $(t_k)_k \to t^*$. Let $\mu \in \mathcal{M}^+(S)$ and $\epsilon > 0$. Since the family $\overline{\mathcal{P}}(\delta_2)$ is equicontinuous (Corollary 3.6.4), there exists $\delta_\epsilon > 0$ such that for all $\nu \in \mathcal{M}^+(S)$ with $\|\mu-\nu\|^*_{\mathrm{BL},d} < \delta_\epsilon$,

$$\left\|\overline{\mathbb{P}}_t\mu - \overline{\mathbb{P}}_t\nu\right\|^*_{\mathrm{BL},d} < \frac{\epsilon}{3(1+\|f\|_{\mathrm{BL},d})} \qquad \text{for all } t \in [0,\delta_2].$$

M_0 is dense in $\mathcal{M}^+(S)$. So there exists $\nu_0 \in M_0$ such that $\|\mu-\mu_0\|^*_{\mathrm{BL},d} < \delta_\epsilon$. Then

$$\begin{aligned}\left|\left\langle \overline{\mathbb{P}}_{t^*}\mu - \overline{\mathbb{P}}_{t_k}\mu, f\right\rangle\right| &\leq \left\|\overline{\mathbb{P}}_{t^*}\mu - \overline{\mathbb{P}}_{t^*}\mu_0\right\|^*_{\mathrm{BL},d}\cdot\|f\|_{\mathrm{BL},d} \\ &\quad + \left|\left\langle \overline{\mathbb{P}}_{t^*}\mu_0 - \overline{\mathbb{P}}_{t_k}\mu_0, f\right\rangle\right| + \left\|\overline{\mathbb{P}}_{t_k}\mu_0 - \overline{\mathbb{P}}_{t_k}\mu\right\|^*_{\mathrm{BL},d}\cdot\|f\|_{\mathrm{BL},d} \\ &< \frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3} = \epsilon,\end{aligned}$$

when $k \geq N$ such that $\left|\left\langle \overline{\mathbb{P}}_{t^*}\mu_0 - \overline{\mathbb{P}}_{t_k}\mu_0, f\right\rangle\right| < \frac{\epsilon}{3}$ for all $k \geq N$. So, by Theorem 3.3.5, $t \mapsto \overline{\mathbb{P}}_t\mu$ is continuous on $[0, \delta_2)$.

Now we show that the continuity of $t \mapsto \overline{\mathbb{P}}_t\mu$ on $[0, m\delta_2)$ implies continuity on $[0, (m+1)\delta_2)$. Let $t^* \in [0, (m+1)\delta_2)$ and $t_k \in [0, (m+1)\delta_2)$ such that $t_k \to t^*$. According to Proposition 3.6.2,

$$\overline{\mathbb{P}}_{t_k}\mu = \overline{\mathbb{P}}_{\frac{t_k}{m+1}}\left(\overline{\mathbb{P}}_{\frac{mt_k}{m+1}}\mu\right) = \overline{\mathbb{P}}_{\frac{t_k}{m+1}}\left[\overline{\mathbb{P}}_{\frac{mt_k}{m+1}}\mu - \overline{\mathbb{P}}_{\frac{mt^*}{m+1}}\mu\right] + \overline{\mathbb{P}}_{\frac{t_k}{m+1}} \cdot \overline{\mathbb{P}}_{\frac{mt^*}{m+1}}\mu.$$

Because $\frac{t_k}{m+1} \in [0, \delta_0)$, $\overline{\mathcal{P}}(\delta_2)$ is equicontinuous and $\overline{\mathbb{P}}_{\frac{mt_k}{m+1}}\mu \to \overline{\mathbb{P}}_{\frac{mt^*}{m+1}}\mu$ by assumption, the first term can be made arbitrarily small for sufficiently large k. The second term converges to $\overline{\mathbb{P}}_{\frac{t^*}{m+1}} \cdot \overline{\mathbb{P}}_{\frac{mt^*}{m+1}}\mu$, which equals $\overline{\mathbb{P}}_{t^*}\mu$ by Proposition 3.6.2. So indeed, $t \mapsto \overline{\mathbb{P}}_t\mu$ is continuous on $[0, (m+1)\delta_2)$. We conclude that $t \mapsto \overline{\mathbb{P}}_t\mu$ is continuous on $\mathbb{R}_+$. According to Proposition 3.6.2, $\overline{\mathbb{P}}_t\overline{\mathbb{P}}_s\mu = \overline{\mathbb{P}}_{t+s}\mu$ for all $t, s \in \mathbb{R}_+$ such that $\frac{t}{s} \in \mathbb{Q}$. Because $t \mapsto \overline{\mathbb{P}}_t\mu$ is continuous, the semigroup property must hold for all $t, s \in \mathbb{R}_+$. □

We say that a Markov semigroup is **stochastically continuous at 0** if $\lim_{h \searrow 0} P_h\mu = \mu$ for every $\mu \in \mathcal{M}^+(S)_{\mathrm{BL}}$. Stochastic continuity at 0 implies **right-continuity** at every $t_0 \geq 0$, but not left-continuity. The next result shows together with equicontinuity, that stochastic continuity at 0 implies strong continuity.

Proposition 3.6.6. *Let $(P_t)_{t\geq 0}$ be a Markov-Feller semigroup. Assume that there exists $\delta > 0$ such that $(P_t)_{t\in[0,\delta]}$ is equicontinuous. If $(P_t)_{t\geq 0}$ is stochastically continuous at 0, then it is strongly continuous.*

Proof. $(P_t)_{t\in[0,\delta]}$ is equicontinuous and $P_{t'}$ is Feller for all $t' \geq 0$. Consequently, $(P_t)_{t\in[t',t'+\delta]}$ is an equicontinuous family for every $t' \in \mathbb{R}_+$. Hence $(P_t)_{t\in[0,T]}$ is equicontinuous for every $T \in \mathbb{R}_+$. So, if $\varepsilon > 0$, there exists an open neighbourhood U in $\mathcal{M}^+(S)$ of μ such that

$$\|P_t\nu - P_t\mu\|^*_{\mathrm{BL}} < \varepsilon$$

for every $\nu \in U$. Let $t_0 > 0$. From the fact, that $(P_t)_{t\geq 0}$ is (strongly) stochastically continuous at 0, there exists $\delta > 0$ such that for every $0 < h < \delta$, $P_h\mu \in U$. Then, from the fact that

$$\|P_{t_0}\mu - P_{t_0-h}\mu\|^*_{\mathrm{BL}} = \|P_{t_0-h}\mu - P_{t_0-h}P_h\mu\|^*_{\mathrm{BL}},$$

we get

$$\|P_{t_0-h}\mu - P_{t_0}\mu\|^*_{\mathrm{BL}} < \varepsilon \text{ for all } 0 < h < \delta.$$

So $t \mapsto P_t\mu$ is also left-continuous at every $t_0 > 0$. □

Corollary 3.6.7. *If $(P_t)_{t\geq 0}$ is stochastically continuous and $(P_t)_{t\in[0,\delta]}$ is equicontinuous, then $(P_t)_{t\in[0,T]}$ is tight for every $T > 0$.*

Remark 3.6.8. *From Proposition 3.6.6 we can conclude that a Markov semigroup that is stochastically continuous at 0 but not strongly continuous, cannot be equicontinuous.*

3.6.3 Symmetry

We prove that, if the family $\mathcal{P}^1(\delta)$ is tight - as we assume in Assumption 1 - then the limit does not depend on the order in which we start switching semigroups $(P_t^1)_{t\geq 0}$ and $(P_t^2)_{t\geq 0}$.

Now let us prove the following lemma.

Lemma 3.6.9. *Let $(P_t^1)_{t\in\mathcal{T}}$ and $(P_t^2)_{t\in\mathcal{T}}$ be semigroups of regular Markov-Feller operators. Let $n \in \mathbb{N}$, $t \in \mathbb{R}_+$. Then*

$$\left(P_t^1 P_t^2\right)^n - \left(P_t^2 P_t^1\right)^n = \sum_{i=0}^{n-1} (P_t^2 P_t^1)^{n-i-1} C_{t,t}^{1,2} (P_t^1 P_t^2)^i \tag{3.10}$$

$$= \sum_{i=0}^{n-1} (P_t^1 P_t^2)^{n-i-1} C_{t,t}^{1,2} (P_t^2 P_t^1)^i \tag{3.11}$$

where $C_{s,t}^{i,j} = P_s^i P_t^j - P_t^j P_s^i$.

Proof. We prove (3.10) by induction. Let L_n denote the left-hand side in equality (3.10), R_n the right-hand side. Obviously $L_1 = R_1$. Assume that $L_{n-1} = R_{n-1}$. Then:

$$\begin{aligned}
L_n &= (P_s^1 P_s^2)^n - (P_s^2 P_s^1)^n = \\
&= \left[(P_s^1 P_s^2)^{n-1} - (P_s^2 P_s^1)^{n-1}\right] P_s^1 P_s^2 + \left(P_s^2 P_s^1\right)^{n-1} P_s^1 P_s^2 - \left(P_s^2 P_s^1\right)^n = \\
&= \left[\textstyle\sum_{i=0}^{n-2} (P_s^2 P_s^1)^{n-i-2} C_{s,s}^{1,2} (P_s^1 P_s^2)^i\right] P_s^1 P_s^2 + \left(P_s^2 P_s^1\right)^{n-1} \left(P_s^1 P_s^2 - P_s^2 P_s^1\right) = \\
&= \textstyle\sum_{i=0}^{n-2} (P_s^2 P_s^1)^{n-i-2} C_{s,s}^{1,2} (P_s^1 P_s^2)^{i+1} + \left(P_s^2 P_s^1\right)^{n-1} C_{s,s}^{1,2} = \\
&= \textstyle\sum_{i=0}^{n-1} (P_s^2 P_s^1)^{n-i-1} C_{s,s}^{1,2} (P_s^1 P_s^2)^i = R_n. \qquad \square
\end{aligned}$$

Next we prove that the limit of the switching scheme does not depend on the order of switched semigroups in the product formula.

Proposition 3.6.10. *Let $(P_t^1)_{t\geq 0}$ and $(P_t^2)_{t\geq 0}$ be semigroups of Markov operators for which Assumptions 1 - 4 hold, and additionally that Assumption 2 holds for $(P_t^1)_{t\geq 0}$ and $(P_t^2)_{t\geq 0}$*

swapped. Let $\mu \in \mathcal{M}^+(S)$. Then

$$\lim_{n\to\infty}\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\mu = \lim_{n\to\infty}\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^n\mu.$$

Proof. Let $t \in \mathbb{R}_+$, $\mu_0 \in M_0$, $f \in \mathrm{BL}(S,d)$ and fix $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $\frac{t}{N} \le \delta$, where $\delta = \min(\delta_{3,f}, \delta_{4,f})$. Since $(P^1_t)_{t\ge0}$ and $(P^2_t)_{t\ge0}$ are equicontinuous, they consist of Feller operators necessarily. According to Lemma 3.6.9, for $n \ge N$

$$\begin{aligned}
\left|\left\langle\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\mu_0 - \left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^n\mu_0, f\right\rangle\right| &= \left|\left\langle\sum_{i=0}^{n-1}\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^{n-i-1}C^{1,2}_{\frac{t}{n},\frac{t}{n}}\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^i\mu_0, f\right\rangle\right| \\
&\le \sum_{i=0}^{n-1}\left|\left\langle C^{1,2}_{\frac{t}{n},\frac{t}{n}}\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^i\mu_0, \left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^{n-i-1}f\right\rangle\right| \\
&\le \sum_{i=0}^{n-1}\left\|C^{1,2}_{\frac{t}{n},\frac{t}{n}}\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^i\mu_0\right\|^*_{\mathrm{BL},d_{\mathcal{E}(f)}}\cdot\left\|\left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^{n-i-1}f\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}} \\
&\le \sum_{i=0}^{n-1}\frac{t}{n}\omega_f\left(\frac{t}{n},\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^i\mu_0\right) \\
&\le C_f(\mu_0)t\omega_f\left(\frac{t}{n},\mu_0\right),
\end{aligned}$$

because $\left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^{n-i-1}f \in \mathcal{E}(f)$.

As t is fixed and $\lim_{s\to0}\omega_f(s,\mu_0) = 0$, we obtain for every $f \in \mathrm{BL}(S,d)$ and $\mu_0 \in M_0$

$$\lim_{n\to\infty}\left|\left\langle\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\mu_0 - \left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^n\mu_0, f\right\rangle\right| = 0.$$

Then, by Theorem 3.3.5, it also converges in norm. Hence,

$$\left\|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\mu_0 - \left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^n\mu_0\right\|^*_{\mathrm{BL}} \to 0 \text{ as } n\to\infty.$$

Define $\hat{\mathbb{P}}_t\mu := \lim_{n\to\infty}\left[P^2_{\frac{t}{n}}P^1_{\frac{t}{n}}\right]^n\mu$, for $\mu \in \mathcal{M}^+(S)$. Since by assumption Assumption 2 holds with P^1_t and P^2_t swapped, Proposition 3.6.3 holds for $\hat{\mathbb{P}}_t$ as well: both $\overline{\mathbb{P}}_t$ and $\hat{\mathbb{P}}_t$ are continuous on $\mathcal{M}^+(S)$. Since M_0 is a dense subset of $\mathcal{M}^+(S)_{\mathrm{BL}}$ and $\overline{\mathbb{P}}_t\mu_0 = \hat{\mathbb{P}}_t\mu_0$ for $\mu_0 \in M_0$, we obtain $\overline{\mathbb{P}}_t = \hat{\mathbb{P}}_t$ on $\mathcal{M}^+(S)$. □

3.7 Relation to literature

We shall now show that Theorem 3.2.2 is a generalization of existing results. We start with the approach of Kühnemund and Wacker [KW01] and show in detail that their result follows from Theorem 3.2.2. Then we provide proof that also the Proposition 3.5 in Colombo-Guerra [Col09] follows from Theorem 3.2.2.

3.7.1 Kühnemund-Wacker

Kühnemund and Wacker [KW01] provided conditions for C_0-semigroups that ensure convergence of the Lie-Trotter product. Their setting is the following: Let $(T(t))_{t\geq 0}, (S(t))_{t\geq 0}$ be strongly continuous linear semigroups on a Banach space $(E, \|\cdot\|)$ that consists of bounded linear operators. Let $F \subset E$ be a dense linear subspace, equipped with a norm $|||\cdot|||$, such that both $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ leave F invariant.

Assumption KW 1. *$(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ are **exponentially bounded** on $(F, |||\cdot|||)$, so there exist $M_T, M_S \geq 1$, and $\omega_T, \omega_S \in \mathbb{R}$ such that*

$$|||T(t)||| \leq M_T e^{\omega_T t}, \quad |||S(t)||| \leq M_S e^{\omega_S t}$$

for all $t \geq 0$.

Assumption KW 2. *$(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ are **locally Trotter stable** on both $(E, \|\cdot\|)$ and $(F, |||\cdot|||)$. There exists $\delta > 0$ and $M_E^\delta, M_F^\delta \geq 1$ such that*

$$\left\| \left[T\left(\tfrac{t}{n}\right) S\left(\tfrac{t}{n}\right) \right]^n \right\| \leq M_E^\delta$$

$$\left|\left|\left| \left[T\left(\tfrac{t}{n}\right) S\left(\tfrac{t}{n}\right) \right]^n \right|\right|\right| \leq M_F^\delta$$

for all $t \in [0, \delta]$ and $n \in \mathbb{N}$.

Assumption KW 3. *(Commutator condition) There exists $\alpha > 1$, $\delta' > 0$ and $M_1 \geq 0$ such that*

$$\|T(t)S(t)f - S(t)T(t)f\| \leq M_1 t^\alpha |||f|||$$

for all $f \in F$, $t \in [0, \delta]$.

Theorem 3.7.1 (Kühnemund and Wacker, [KW01], Theorem 1)**.** *Let $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$ be strongly continuous semigroups satisfying Assumptions* KW1 - KW3*. Then the Lie-Trotter product formula holds, i.e.*

$$\mathbb{P}_t x := \lim_{n\to\infty} \left[T\left(\tfrac{t}{n}\right) S\left(\tfrac{t}{n}\right) \right]^n x$$

exists in $(E, \|\cdot\|)$ for every $x \in X$, and convergence is uniform for every t in compact intervals in $\mathbb{R}_+$. Moreover, $(\mathbb{P}(t))_{t\geq 0}$ is a strongly continuous semigroup in E.

We shall now show that Theorem 3.7.1 follows from our result. Note that in Theorem 3.7.1 there is no assumption that $(E,\|\cdot\|)$ should be separable, while we assume that (S,d) is separable. This issue can be overcome as follows. Fix $x\in E$. Define $T^1_t := T(t)$, $T^2_t := S(t)$ and

$$E_x = \mathrm{Cl}_E\left(\mathrm{span}_{\mathbb{R}}\left\{T^{i_N}_{t_N}\cdot T^{i_{N-1}}_{t_{N-1}}\cdot\ddots\cdot T^{i_1}_{t_1} : N\in\mathbb{N}, i_k\in\{1,2\}, k=1,2,\cdots,N\right\}\right).$$

Then $E_x\subset E$ is the smallest separable closed subspace that contains x and is both $(T(t))_{t\geq0}$ and $(S(t))_{t\geq0}$-invariant. Let $S=E_x$ with metric $d(y,y') := \|y-y'\|$. Then (S,d) is separable and complete.

Lifts

Let $(P^1_t)_{t\geq0}$ be the lift of $T(t)$ to $\mathcal{M}^+(S)$ and $(P^2_t)_{t\geq0}$ be the lift of $S(t)$ to $\mathcal{M}^+(S)$. That is, for $\mu\in\mathcal{M}^+(S)$,

$$P^1_t\mu := \int_S \delta_{T(t)x}\mu(dx),\quad P^2_t\mu := \int_S \delta_{S(t)x}\mu(dx), \tag{3.12}$$

where the integrals are considered as Bochner integrals in $\overline{\mathcal{M}(S)}_{\mathrm{BL}}$, the closure of $\mathcal{M}(S)_{\mathrm{BL}}$ in $\mathrm{BL}(S,d)^*$. Since $\mathcal{M}^+(S)\subset\overline{\mathcal{M}(S)}_{\mathrm{BL}}$ is closed, $P^i_t\mu\in\mathcal{M}^+(S)$. So

$$P^1_t\delta_x := \delta_{T(t)x},\quad P^2_t\delta_x := \delta_{S(t)x}. \tag{3.13}$$

We show that $(P^i_t)_{\geq0}$, $i=1,2$, defined by (3.12) satisfy Assumptions 1 - 4.

First consider Assumption 1. We discuss $(P^1_t)_{t\geq0}$ only; the argument for $(P^2_t)_{t\geq0}$ is similar. The map $t\mapsto P^1_t\mu:\mathbb{R}_+\to\mathcal{M}^+(S)_{\mathrm{BL}}$ is continuous if and only if $t\mapsto\langle P^1_t\mu,f\rangle$ is continuous for every $f\in C_b(S)$. Clearly, $\langle P^1_t\mu,f\rangle = \int_S\langle\delta_{T(t)x},f\rangle\mu(dx) = \int_S f(T(t)x)\mu(dx)$. Using the strong continuity of $(T(t))_{t\geq0}$ and Lebesgue's Dominated Convergence Theorem we see that $t\mapsto\langle P^1_t\mu,f\rangle$ is indeed continuous on $\mathbb{R}_+$. Thus, $\{P^1_t\mu : t\in[0,\delta]\}$ is compact in $\mathcal{M}^+(S)_{\mathrm{BL}}$, that is: tight.

Let $\phi\in\mathrm{BL}(S,d)$ and $x_0\in S$. Let U^1_t be dual operators to P^1_t. Then:

$$\begin{aligned}
&|U^1_t\phi(x)-U^1_t\phi(x_0)| = |\langle P^1_t\delta_x - P^1_t\delta_{x_0},\phi\rangle|\\
&\underbrace{=}_{def}|\langle\delta_{T(t)x}-\delta_{T(t)x_0},\phi\rangle| = |\phi(\delta_{T(t)x})-\phi(\delta_{T(t)x_0})| \underbrace{\leq}_{\phi\in\mathrm{BL}(S,d)} |\phi|_L\cdot\|T(t)x-T(t)x_0\|\\
&\leq|\phi|_L\cdot\|T(t)\|\cdot\|x-x_0\|\underbrace{\leq}_{KW1}|\phi|_l\cdot M_Te^{\omega_T t}\cdot\|x-x_0\|
\end{aligned}$$

So there exists δ_T such that $\{U_t^1\phi : t \in [0,\delta_T]\}$ is equicontinuous in $C_b(S)$. Hence, $\{P_t^1 : t \in [0,\delta_T]\}$ forms an equicontinuous family, according to Theorem 3.4.2.

The stability condition in Assumption 2 can be shown as follows. Let $\phi \in \mathrm{BL}(S,d)$, $x_0 \in S$.

$$\begin{aligned}
\left|\left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^n \phi(x) - \left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^n \phi(x_0)\right| &= \left|\left\langle \delta_x - \delta_{x_0}, \left[U^2_{\frac{t}{n}}U^1_{\frac{t}{n}}\right]^n \phi\right\rangle\right| \\
&= \left|\left\langle \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \delta_x - \left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n \delta_{x_0}, \phi\right\rangle\right| \\
&= \left|\left\langle \delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n x} - \delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n x_0}, \phi\right\rangle\right| \\
&= \left|\phi\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n x - \phi\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n x_0\right| \\
&\le |\phi|_L \left\|\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n (x - x_0)\right\| \\
&\le |\phi|_L \cdot \left\|\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n\right\| \cdot \|x - x_0\| \\
&\le |\phi|_L \cdot M_E^\delta \cdot \|x - x_0\|
\end{aligned}$$

by Assumption KW3, for $t \in [0,\delta]$, $n \in \mathbb{N}$. Theorem 3.4.2 again implies equicontinuity of $\mathcal{F}(\delta)$. Let $\phi \in F \subset E$. We define

$$M_0 := \mathrm{span}_{\mathbb{R}_+}\{\delta_\phi \,|\, \phi \in F\} \subset \mathcal{M}^+(S).$$

Then M_0 is dense in $\mathcal{M}^+(S)$ and $(P_t^i)_{t\ge 0}$-invariant, $i = 1,2$.
Moreover, define

$$|\mu_0|_{M_0} := \int_F \|\phi\|\mu_0(d\phi). \tag{3.14}$$

So

$$\left|\sum_{k=1}^N a_k\delta_{\phi_k}\right|_{M_0} = \sum_{k=1}^N a_k\|\phi_k\|.$$

To check the commutator condition in Assumption 3, let $f \in \mathrm{BL}(S,d)$ and $\mu_0 \in M_0$. We define a new admissible metric $d_{\mathcal{E}(f)}$ as in (3.3). Then for $y, y' \in E_x = S$,

$$d_{\mathcal{E}(f)}(y,y') = \|y - y'\| \vee \sup_{g\in\mathcal{E}(f)} |h(y) - h(y')|.$$

For $h \in \mathcal{E}(f)$ there exist s, s' and $t \in [0,\delta]$, with $\delta = \min(\delta_1,\delta_2)$, such that

$$\begin{aligned}
|h(y) - h(y')| &= \left|f\left(\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n T(s')S(s)y\right) - f\left(\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n T(s')S(s)y'\right)\right| \\
&\le |f|_{L,d} \cdot \left\|\left[T\left(\tfrac{t}{n}\right)S\left(\tfrac{t}{n}\right)\right]^n T(s')S(s)\right\| \cdot \|y - y'\| \\
&\le M \cdot |f|_{L,d} \cdot \|y - y'\|
\end{aligned}$$

for some constant $M > 0$, according to Assumptions KW 1 - 2.

$$\left\|P_t^1 P_t^2 \mu_0 - P_t^2 P_t^1 \mu_0\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}}^* \le \int_S \left\|P_t^1 P_t^2 \delta_\phi - P_t^2 P_t^1 \delta_\phi\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}}^* \mu_0(d\phi)$$

Let $B_{\mathcal{E}(f)}$ be the unit ball in $\mathrm{BL}(S, d_{\mathcal{E}(f)})$ for $\|\cdot\|_{\mathrm{BL},d_{\mathcal{E}(f)}}$. By the Commutator Condition KW 3 we get the following:

$$\begin{aligned}\|P_t^1 P_t^2 \delta_\phi - P_t^2 P_t^1 \delta_\phi\|_{\mathrm{BL},d_{\mathcal{E}(f)}}^* &= \sup_{g\in B_{\mathcal{E}(f)}} |g(T(t)S(t)\phi) - g(S(t)T(t)\phi)| \\ &\le \sup_{g\in B_{\mathcal{E}(f)}} |g|_{L,d_{\mathcal{E}(f)}} \cdot d_{\mathcal{E}(f)} \left(T(t)S(t)\phi, S(t)T(t)\phi\right) \\ &\le \max(1, |f|_{L,d} M) \|T(t)S(t)\phi - S(t)T(t)\phi\| \\ &\le \max(1, |f|_{L,d} M) M_1 t^\alpha \||\phi\||.\end{aligned}$$

Define

$$\omega_f(t, \mu_0) := \max(1, |f|_{L,d} M) M_1 t^{\alpha-1} |\mu_0|_{M_0}.$$

Since $\alpha > 1$, $\omega_f : \mathbb{R}_+ \times M_0 \to \mathbb{R}_+$ is continuous, non-decreasing and for every $\delta > 0$

$$\begin{aligned}\int_0^\delta \frac{\omega_f(t, \mu_0)}{t} dt &= \max(1, |f|_{L,d} M) |\mu_0|_{M_0} M_1 \int_0^\delta t^{\alpha-2} dt \\ &= \max(1, |f|_{L,d} M) M_1 \frac{\delta^{\alpha-1}}{\alpha - 1} < +\infty.\end{aligned}$$

Moreover, for $\mu_0 \in M_0$,

$$\begin{aligned}\left\|P_t^1 P_t^2 \mu_0 - P_t^2 P_t^1 \mu_0\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}}^* &\le \int_S \left\|P_t^1 P_t^2 \delta_\phi - P_t^2 P_t^1 \delta_\phi\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}}^* \mu_0(d\phi) \\ &\le \max\left(1, |f|_{L,d} M\right) M_1 t^{\alpha-1} \int_S \||\phi\|| \mu_0(d\phi) \\ &= t\omega_f(t, \mu_0).\end{aligned}$$

Hence, we get Assumption 3 for all $\mu_0 \in M_0$ and $\delta_{3,f} = \delta'$.

Let us now check Assumption 4. First, for any $\phi \in F$,

$$\left|\left[P_{\frac{t}{n}}^1 P_{\frac{t}{n}}^2\right]^n \delta_\phi\right|_{M_0} = \left|\delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n \phi}\right|_{M_0} = \left\|\left|\left[T\left(\tfrac{t}{n}\right) S\left(\tfrac{t}{n}\right)\right]^n \phi\right\|\right| \le M_F^\delta \||\phi\||.$$

For $\mu_0 \in M_0$ we get

$$
\begin{aligned}
\left|\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \mu_0\right|_{M_0} &= \left|\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \left(\sum_k a_k \delta_{\phi_k}\right)\right|_{M_0} \\
&= \left|\sum_k a_k \delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n \phi_k}\right|_{M_0} \\
&= \sum_k a_k \left|\delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n \phi_k}\right|_{M_0} \\
&\leq \sum_k a_k M_F^\delta \|\|\phi_k\|\| \\
&= M_F^\delta |\mu_0|_{M_0}.
\end{aligned}
$$

Furthermore,

$$
\left|P^1_t \delta_\phi\right|_{M_0} = \left|\delta_{T(t)\phi}\right|_{M_0} = \|\|T(t)\phi\|\| \leq M_T e^{\omega_T t}\|\|\phi\|\| \leq M_T e^{\omega_T \delta}\|\|\phi\|\|
$$

and similarly

$$
\left|P^2_t \delta_\phi\right|_{M_0} \leq M_S e^{\omega_S \delta}\|\|\phi\|\|.
$$

Then for $0 \leq t \leq \delta$

$$
|P^1_t \mu_0|_{M_0} \quad \leq \quad M_T e^{\omega_T \delta} |\mu_0|_{M_0}
$$

and

$$
|P^2_t \mu_0|_{M_0} \quad \leq \quad M_S e^{\omega_T \delta} |\mu_0|_{M_0}.
$$

Thus,

$$
\left|P^2_s \left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n P^1_{t'} \mu_0\right|_{M_0} \leq M_T M_S M_F^\delta e^{(\omega_T + \omega_s)\delta} \cdot |\mu_0|_{M_0}
$$

and with $C_f(\mu_0) \coloneqq M_T M_S M_F^\delta e^{(\omega_T+\omega_s)\delta}$ (independent of f and μ_0) and $\delta_{4,f} = \min(\delta, \delta')$, we see that Assumptions 1 - 4 hold.

Hence, we conclude that the Lie-Trotter formula holds for $(P^i_t)_{t\geq 0}$, $i = 1, 2$. Moreover, as $\delta_{3,f}$, $\delta_{4,f}$, $C_f(\mu_0)$ and ω_f can be chosen uniformly for f in the unit ball in $(\mathrm{BL}(S,d), \|\cdot\|_{\mathrm{BL},d})$, the convergence is uniform in f in compact subsets of $\mathbb{R}_+$. Furthermore, for every $y \in E_x$, $\left[P^1_{\frac{t}{n}} P^2_{\frac{t}{n}}\right]^n \delta_y = \delta_{[T(\frac{t}{n})S(\frac{t}{n})]^n y} \to \overline{\mathbb{P}}_t \delta_y$ in $\mathcal{M}^+(S)_{\mathrm{BL}}$ as $n \to \infty$.

The set of Dirac measures is closed in $\mathcal{M}^+(S)_{\mathrm{BL}}$. To show this, let $(\delta_{x_n})_n$ be a sequence of Dirac measures such that $\delta_{x_n} \to \mu$ for some $\mu \in \mathcal{M}^+(S)$. Then $(\delta_{x_n})_n$ is a Cauchy sequence, and

$$
\|\delta_{x_n} - \delta_{x_m}\|^*_{BL,d} = \frac{2d(x_n, x_m)}{2 + d(x_n, x_m)}
$$

([HW09b] Lemma 2.5). Then also $(x_n)_{n\in\mathbb{N}} \subset S$ is a Cauchy sequence. As S is complete, $(x_n)_{n\in\mathbb{N}}$ is convergent. Hence, there exists $x^* \in S$ such that $x_n \to x^*$ as $n \to \infty$ and

$$\|\delta_{x_n} - \delta_{x^*}\|^*_{BL,d} = \frac{2d(x_n, x^*)}{2 + d(x_n, x^*)} \to 0 \text{ as } n \to \infty.$$

Hence, $\overline{\mathbb{P}}_t\delta_y = \delta_{\mathbb{P}^x_t y}$ for a specific $\mathbb{P}^x_t \subset E$ (as in statement Theorem 3.7.1). Because the $(P^i_t)_{t\geq 0}$, $i = 1, 2$, are strongly continuous in this setting, $(\overline{\mathbb{P}}_t)_{t\geq 0}$ is a semigroup by Proposition 3.6.5. Therefore, $(\mathbb{P}^x_t)_{t\geq 0}$ is a strongly continuous semigroup on E_x. The operators $\mathbb{P}_t$ are linear and continuous:

Let $y_n \in E_x$ such that $Y_n \to y$ in E. Then

$$\begin{aligned}\|\mathbb{P}^x_t y_n - \mathbb{P}^x_t y\|^*_{\mathrm{BL},d} &= \frac{2\|\delta_{\mathbb{P}_t y_n} - \delta_{\mathbb{P}_t y}\|^*_{\mathrm{BL},d}}{2 + \|\delta_{\mathbb{P}_t y_n} - \delta_{\mathbb{P}_t y}\|^*_{\mathrm{BL},d}} \\ &= \frac{2\|\overline{\mathbb{P}}_t\delta_{y_n} - \overline{\mathbb{P}}_t\delta_y\|^*_{\mathrm{BL},d}}{2 + \|\overline{\mathbb{P}}_t\delta_{y_n} - \overline{\mathbb{P}}_t\delta_y\|^*_{\mathrm{BL},d}} \to 0.\end{aligned}$$

Moreover, $E = \bigcup_{x\in E} E_x$, and the semigroups $(\mathbb{P}^x_t)_{t\geq 0}$ and $(\mathbb{P}^{x'}_t)_{t\geq 0}$ agree on $E_x \cap E_{x'}$. This allows us to define a strongly continuous semigroup $(\mathbb{P}_t)_{t\geq 0}$ of bounded linear operators on E that agrees with $(\mathbb{P}^x_t)_{t\geq 0}$ on E_x.

3.7.2 Colombo-Guerra

Colombo and Guerra in [Col09], generalizing Colombo and Corli [CC04], also established conditions that ensure the convergence of the Lie-Trotter formula for linear semigroups in a Banach space that do not involve the domains of their generators. Instead, like in the results of Kühnemund and Wacker [KW01], they build on a commutator condition (Assumption CG 3 stated below) that is weaker than that in [KW01]. It is this condition that motivated our Assumption 3.

The situation in [Col09] is as follows. Let $S^1, S^2 : \mathbb{R}_+ \times X \mapsto X$ be strongly continuous semigroups on a Banach space X. Assume that there exists a normed vector space Y which is densely embedded in X and invariant under both semigroups such that:

Assumption CG 1. *The two semigroups are* ***locally Lipschitz in time in*** Y*, i.e. there exists a compact map* $K : Y \to \mathbb{R}$ *such that for* $i = 1, 2$

$$\left\|S^1_t u - S^i_{t'} u\right\|_X \leq K(u)|t - t'| \quad \text{for all } u \in Y,\ t, t' \in I.$$

Assumption CG 2. *The two semigroups are* ***exponentially bounded on*** F *and* ***locally Trotter stable on*** X ***and*** Y*, i.e. there exists a constant* H *such that for all* $t \in [0,1]$*,* $n \in \mathbb{N}$

$$\|S_t^1\|_Y + \|S_t^2\|_Y + \left\|\left(S^1_{\frac{t}{n}} S^2_{\frac{t}{n}}\right)^n\right\|_X + \left\|\left(S^1_{\frac{t}{n}} S^2_{\frac{t}{n}}\right)^n\right\|_Y \le H.$$

Assumption CG 3 (Commutator condition)**.**

$$\left\|S_t^1 S_t^2 u - S_t^2 S_t^1 u\right\|_X \le t\omega(t)\|u\|_Y$$

is satisfied for all $u \in Y$ *and* $t \in [0,\delta]$ *with some* $\delta > 0$*, and for a suitable* $\omega : [0,\delta] \to \mathbb{R}^+$ *with* $\int_0^\delta \frac{\omega(\tau)}{\tau} d\tau < +\infty$*.*

Theorem 3.7.2. *Under Assumptions CG1-CG3 there exists a global semigroup* $Q : [0,+\infty) \times X \to X$ *such that for all* $u \in Y$*, there exists a constant* C_u *such that for* $t > 0$

$$\frac{1}{t}\left\|Q(t)u - S_t^1 S_t^2 u\right\|_X \le C_u \int_0^t \frac{\omega(\xi)}{\xi} d\xi.$$

In fact, [Col09] Proposition 3.5 also includes a statement of convergence of so-called Euler polygonals to orbits of Q. The interested reader should consult [Col09] for further details on this topic.

It is the construction in this case that allows us to conclude that Theorem 3.7.2 and Theorem 3.2.2 are highly similar to the Kühnemund-Wacker case discussed in the previous section. Therefore we state the main reasoning and give the immediate results.

Let $u \in X$. We take $S = X_u$, where the latter is the smallest separable Banach space in X that is invariant under $(S_t^i)_{t\ge0}$, $i = 1,2$, equipped with the metric induced by the norm on X. Let P_t^1 and P_t^2 be lifts of S_t^1 and S_t^2 to $\mathcal{M}^+(S)$:

$$P_t^i \delta_u := \delta_{S_t^i u},\ P_t^i \mu := \int_{\mathcal{U}} \delta_{S_t^i u} \mu(du), i = 1,2.$$

Now we check if P_t^1 and P_t^2 satisfy Assumptions 1 - 4.
As in Section 3.7.1, because $(S_t^1)_{t\ge0}$ and $(S_t^2)_{t\ge0}$ are strongly continuous semigroups, $(P_t^1)_{t\ge0}$ and $(P_t^2)_{t\ge0}$ are tight. Moreover, if $\phi \in \mathrm{BL}(S,d)$ and $v,w \in X_u$, and U_t^1 and U_t^2 are the dual operators of P_t^1 and P_t^2 respectively, then:

$$|U_t^1\phi(v) - U_t^1\phi(w)| \le |\phi|_L \cdot H \cdot \|v - w\|_X$$

This yields the equicontinuity condition for U_t^1. Similarly equicontinuity for U_t^2 is established. A similar computation yields Assumption 2:

$$\begin{aligned}\left|\left[U^1_{\frac{t}{n}}U^2_{\frac{t}{n}}\right]^n\phi(v)-\left[U^1_{\frac{t}{n}}U^2_{\frac{t}{n}}\right]^n\phi(w)\right| &= \left|\phi\left[\left(S^2_{\frac{t}{n}}S^1_{\frac{t}{n}}\right)^n v\right]-\phi\left[\left(S^2_{\frac{t}{n}}S^1_{\frac{t}{n}}\right)^n w\right]\right|\\ &\le |\phi|_L\cdot\left\|\left(S^2_{\frac{t}{n}}S^1_{\frac{t}{n}}\right)^n(v-w)\right\|_X \le |\phi|_L\cdot H\cdot\|v-w\|_X\end{aligned}$$

To check the Commutator Condition in Assumption 3, let $f\in \mathrm{BL}(S,d)$, put $M_0:=\operatorname{span}\{\delta_v|v\in Y\cap X_u\}$ and $|\mu_0|_{M_0}$ as in (3.14). Then define

$$\omega_f(t,\mu_0):=\max(1,|f|_{L,d}M)\omega(t)|\mu_0|_{M_0}.$$

Commutator Condition CG 3 yields

$$\left\|P^1_tP^2_t\delta_u-P^2_tP^1_t\delta_u\right\|^*_{\mathrm{BL},d_{\mathcal{E}(f)}}\le\max(1,|f|_{L,d}M)t\omega(t)\|u\|_Y$$

as before, which established Assumption 3. Note that ω_f can be chosen uniformly for f in the unit ball of $\mathrm{BL}(S,d)$.

Assumption 4 is obtained from the estimate

$$\left|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\delta_u\right|_{M_0}=\left|\delta_{\left[S^1_{\frac{t}{n}}S^2_{\frac{t}{n}}\right]^n u}\right|_{M_0}=\left\|\left[S^2_{\frac{t}{n}}S^1_{\frac{t}{n}}\right]^n u\right\|_X\le H\|u\|_X,$$

which yields

$$\left|\left[P^1_{\frac{t}{n}}P^2_{\frac{t}{n}}\right]^n\mu_0\right|_{M_0}\le H|\mu_0|_{M_0}.$$

and

$$\left|P^1_t\delta_\phi\right|_{M_0}=\left|\delta_{S^1_tu}\right|_{M_0}=\|S^1_tu\|_Y\le H\|u\|_Y,\quad\left|P^2_t\delta_\phi\right|_{M_0}\le H\|u\|_Y$$

which yields

$$\left|P^1_t\mu_0\right|_{M_0}\le H|\mu_0|_{M_0}\text{ and }\left|P^2_t\mu_0\right|_{M_0}\le H|\mu_0|_{M_0}.$$

Thus, the Lie-Trotter formula holds for $(P^1_t)_{t\ge0}$ and $(P^2_t)_{t\ge0}$. A similar argument as in Section 3.7.1 yields Theorem 3.7.2.

3.8 Appendices

3.8.1 Proof of Lemma 3.5.8

(a) We will check it by induction on j. Let $j=1$. Then the left hand side in the equation 3.5.8, (a) is of the form

$$L = P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}},$$

while the right hand side is

$$R = \sum_{l=0}^{0} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(1-1-l)t}{m}} = P^2_{\frac{0t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(1-1-0)t}{m}} = L.$$

Assume that (a) holds for $j-1$:

$$P^1_{\frac{t}{m}} P^2_{\frac{(j-1)t}{m}} - P^2_{\frac{(j-1)t}{m}} P^1_{\frac{t}{m}} = \sum_{l=0}^{j-2} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-2-l)t}{m}}.$$

Then for j:

$$\begin{aligned}
L &= P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} \\
&= \left(P^1_{\frac{t}{m}} P^2_{\frac{(j-1)t}{m}} - P^2_{\frac{(j-1)t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{t}{m}} + P^2_{\frac{(j-1)t}{m}} P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} \\
&= \left(\sum_{l=0}^{j-2} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-2-l)t}{m}} \right) P^2_{\frac{t}{m}} + P^2_{\frac{(j-1)t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) \\
&= \sum_{l=0}^{j-2} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-1-l)t}{m}} + P^2_{\frac{(j-1)t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) \\
&= \sum_{l=0}^{j-1} P^2_{\frac{lt}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{(j-1-l)t}{m}} = R.
\end{aligned}$$

(b) We will check it by induction on k. Let $k=2$.

$$L = P^1_{\frac{2t}{m}} P^2_{\frac{2t}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^2$$

$$\begin{aligned}
R &= \sum_{j=1}^{1} P^1_{\frac{tj}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{2-1-j} \\
&= P^1_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^2_{\frac{t}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{t}{m}} = L.
\end{aligned}$$

Assume that for $k-1$ we have:

$$P^1_{\frac{(k-1)t}{m}} P^2_{\frac{(k-1)t}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{k-1} = \sum_{j=1}^{k-2} P^1_{\frac{tj}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}} \right) P^2_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \right)^{k-2-j}.$$

Then for k we have:

$$\begin{aligned}
L &= P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k \\
&= \left[P^1_{\frac{(k-1)t}{m}} P^2_{\frac{(k-1)t}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k-1}\right] P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} - P^1_{\frac{(k-1)t}{m}} P^2_{\frac{(k-1)t}{m}} P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} + P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} \\
&= \left[\sum_{j=1}^{k-2} P^1_{\frac{tj}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}}\right) P^2_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k-2-j}\right] P^1_{\frac{t}{m}} P^2_{\frac{t}{m}} \\
&\quad - P^1_{\frac{(k-1)t}{m}} \left(P^2_{\frac{(k-1)t}{m}} P^1_{\frac{t}{m}} - P^1_{\frac{t}{m}} P^2_{\frac{(k-1)t}{m}}\right) P^2_{\frac{t}{m}} \\
&= \sum_{j=1}^{k-1} P^1_{\frac{tj}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}}\right) P^2_{\frac{t}{m}} \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k-1-j} \\
&= R.
\end{aligned}$$

(c) Let $n = 1$. Then

$$L = P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k$$

$$R = \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^0 \left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right] \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k\cdot(1-1-0)} = L$$

Now let's assume that

$$\begin{aligned}
&\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^{n-1} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{(n-1)\cdot k} = \\
&= \sum_{i=0}^{n-2} \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^i \left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right] \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k\cdot(n-2-i)}
\end{aligned}$$

and let us check for n:

$$
\begin{aligned}
L &= \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^n - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{n\cdot k} \\
&= \left(\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^{n-1} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{(n-1)\cdot k}\right)\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k - \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^{n-1}\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k \\
&+ \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^n \\
&= \left(\sum_{i=0}^{n-2}\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^i\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right)\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k\cdot(n-2-i)}\right)\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k - \\
&- \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^{n-1}\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k + \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^n \\
&= \sum_{i=0}^{n-2}\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^i\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right)\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k\cdot(n-1-i)} + \\
&+ \left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^{n-1}\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right] = \\
&= \sum_{i=0}^{n-1}\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right)^i\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^k\right]\left(P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right)^{k\cdot(n-1-i)} = R. \qquad \square
\end{aligned}
$$

3.8.2 Proof of Lemma 3.5.10

Let $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $m := kn$ be such that $\frac{t}{nk} \in [0, \delta_f]$. Then by Lemma 3.5.8 (c) we get

$$
\begin{aligned}
&\left|\left\langle\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right]^n \mu_0 - \left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{n\cdot k}\mu_0, f\right\rangle\right| \\
&= \left|\left\langle\sum_{i=0}^{n-1}\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right]^i\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^k\right)\left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{k\cdot(n-1-i)}\mu, f\right\rangle\right| \\
&\le \sum_{i=0}^{n-1}\left|\left\langle\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right]^i\left(P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}} - \left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^k\right)\left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{k\cdot(n-1-i)}\mu, f\right\rangle\right| = (**)
\end{aligned}
$$

by Lemma 3.5.8 (b)

$$
\begin{aligned}
(**) &= \sum_{i=0}^{n-1}\left|\left\langle\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right]^i\left(\sum_{j=1}^{k-1} P^1_{\frac{tj}{m}}\left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}}\right)P^2_{\frac{t}{m}} \times\right.\right.\right. \\
&\left.\left.\left.\times\left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{k-1-j}\right)\left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{k\cdot(n-1-i)}\mu_0, f\right\rangle\right| \\
&\le \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\left|\left\langle\left[P^1_{\frac{kt}{m}} P^2_{\frac{kt}{m}}\right]^i P^1_{\frac{tj}{m}}\left(P^1_{\frac{t}{m}} P^2_{\frac{jt}{m}} - P^2_{\frac{jt}{m}} P^1_{\frac{t}{m}}\right)P^2_{\frac{t}{m}} \times\right.\right. \\
&\left.\left.\times\left[P^1_{\frac{t}{m}} P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu_0, f\right\rangle\right| = (***)
\end{aligned}
$$

by Lemma 3.5.8 (a) we get

$$
\begin{aligned}
(***) &= \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\left|\left\langle\left[P^1_{\frac{kt}{m}}P^2_{\frac{kt}{m}}\right]^i P^1_{\frac{tj}{m}}\left(\sum_{l=0}^{j-1}P^2_{\frac{lt}{m}}\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)P^2_{\frac{(j-1-l)t}{m}}\right)P^2_{\frac{t}{m}}\times\right.\right.\\
&\left.\left.\times\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu_0, f\right\rangle\right|\\
&\le \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\left|\left\langle\left[P^1_{\frac{kt}{m}}P^2_{\frac{kt}{m}}\right]^i P^1_{\frac{tj}{m}}\left(P^2_{\frac{lt}{m}}\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)P^2_{\frac{(j-1-l)t}{m}}\right)P^2_{\frac{t}{m}}\times\right.\right.\\
&\left.\left.\times\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu_0, f\right\rangle\right|\\
&= \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\left|\left\langle\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)P^2_{\frac{(j-l)t}{m}}\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu_0,\right.\right.\\
&\left.\left.U^2_{\frac{lt}{m}}U^1_{\frac{tj}{m}}\left[U^2_{\frac{kt}{m}}U^1_{\frac{kt}{m}}\right]^i f\right\rangle\right|.
\end{aligned}
$$

For every $i,j,l\in\mathbb{N}$ we get

$$g^n_{i,j,l} := U^2_{\frac{lt}{m}}U^1_{\frac{tj}{m}}\left[U^2_{\frac{kt}{m}}U^1_{\frac{kt}{m}}\right]^i f\in\mathcal{E}(f).$$

Let $\nu^n_{i,j,l} := P^2_{\frac{(j-l)t}{m}}\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu$. Then $\nu^n_{i,j,l}\in M_0$. Note that $\left\|g^n_{i,j,l}\right\|_{\mathrm{BL},d_{\mathcal{E}(f)}}\le 1$.

Using Assumption 4 we get:

$$
\begin{aligned}
&\sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\left|\left\langle\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)\nu^n_{i,j,l}, g^n_{i,j,l}\right\rangle\right|\\
&\le \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\left|\left\langle\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)\nu^n_{i,j,l}, g^n_{i,j,l}\right\rangle\right|\\
&\le \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\left\|\left(P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}-P^2_{\frac{t}{m}}P^1_{\frac{t}{m}}\right)\nu^n_{i,j,l}\right\|^*_{\mathrm{BL},d_{\mathcal{E}()}}\cdot\left\|g^n_{i,j,l}\right\|_{\mathrm{BL},d_{\mathcal{E}()}}\\
&\le \sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}\frac{t}{m}\omega_f\left(\frac{t}{m}, P^2_{\frac{(j-l)t}{m}}\left[P^1_{\frac{t}{m}}P^2_{\frac{t}{m}}\right]^{k(n-i)-1-j}\mu_0\right)\\
&\le \frac{t}{m}\sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}C_2(\mu_0)\omega_f\left(\frac{t}{m},\mu_0\right)\\
&\le C_f(\mu_0)\frac{t}{m}\omega_f\left(\frac{t}{m},\mu_0\right)\sum_{i=0}^{n-1}\sum_{j=1}^{k-1}\sum_{l=0}^{j-1}1\\
&= C_f(\mu_0)\frac{t}{m}\omega_f\left(\frac{t}{m},\mu_0\right)\frac{n(k-1)k}{2}.
\end{aligned}
$$

So with $m = nk$ we get the result. $\square$

Chapter 4

Equicontinuous families of Markov operators in view of asymptotic stability

This chapter is based on:

Sander C. Hille, T. Szarek, Maria A. Ziemlanska. Equicontinuous families of Markov operators in view of asymptotic stability. based on the work Sander C. Hille, T. Szarek, Maria A. Ziemlanska. Equicontinuous families of Markov operators in view of asymptotic stability. *Comptes Rendus Mathematique*, Volume 355, Number 12, Pages 1247-1251, 2017.

Abstract:

The relation between equicontinuity – the so-called e–property and stability of Markov operators is studied. In particular, it is shown that any asymptotically stable Markov operator with an invariant measure such that the interior of its support is non-empty satisfies the e–property.

4.1 Introduction

This chapter is centered around two concepts of equicontinuity for Markov operators defined on probability measures on Polish spaces: the *e-property* and the *Cesàro e-property*. Both appeared as a condition (among others) in the study of ergodicity of Markov operators. In particular they are very useful in proving the existence of a unique invariant measure and its *asymptotic stability*: at whatever probability measure one starts, the iterates under the Markov operator will weakly converge to the invariant measure. The first concept appeared in [LS06, SW12] while the second was introduced in [Wor10] as a theoretical generalisation of the first. It allowed the author to extend various results by replacing the e-property condition by the apparently weaker Cesàro e-property condition.

Interest in equicontinuous families of Markov operators existed already before the introduction of the e-property. Jamison [Jam64], working on compact metric state spaces, introduced the concepts of (dual) Markov operators on the continuous functions that are 'uniformly stable' or 'uniformly stable in mean' to obtain a kind of asymptotic stability results in this setting. Meyn and Tweedie [MT09] introduced the so-called '*e-chains*' on locally compact Hausdorff topological state spaces, for similar purposes. See also [Zah14] for results in a locally compact metric setting. The above mentioned concepts were used in proving ergodicity for some Markov chains (see [Ste94, Cza12, CH14, ESvR12, GL15, KPS10]).

It is worth mentioning here that similar concepts appear in the study of mean equicontinuous dynamical systems mainly on compact spaces (see for instance [LTY15]). However it must be stressed here that our space of Borel probability measures defined on some Polish space is non-compact, typically, in the generality in which we consider the question.

While studying the e–property, the natural question arose whether any asymptotically stable Markov operator satisfies this property. Proposition 6.4.2 in [MT09] asserts this holds when the phase space is compact. In particular, the authors claimed that the stronger e–chain property is satisfied. Unfortunately, the proof contains a gap and an example can be constructed showing that some additional assumptions must be added for the claimed result to hold.

Striving to repair the gap of the Meyn-Tweedie result mentioned above, we show that any asymptotically stable Markov operator with an invariant measure such that the interior of its support is nonempty satisfies the e–property.

4.2 Some (counter) examples

Let (S,d) be a Polish space. By $B(x,r)$ we denote the open ball in (S,d) of radius r, centered at $x \in S$. Further $\overline{E}$, $\mathrm{Int}_S E$ denote the closure of $E \subset S$ and the interior of E, respectively. By $C_b(S)$ we denote the vector space of all bounded real-valued continuous functions on S and by $BM(S)$ all bounded real-valued Borel measurable functions, both equipped with the supremum norm $\|\cdot\|_\infty$. By $BL(S)$ we denote the subspace of $C_b(S)$ of all bounded Lipschitz functions (for the metric d on S). For $f \in BL(S)$, $|f|_L$ denotes the Lipschitz constant of f.

By $\mathcal{M}(S)$ we denote the family of all finite Borel measures on S and by $\mathcal{P}(S)$ the subfamily of all probability measures in $\mathcal{M}(S)$. For $\mu \in \mathcal{M}(S)$, its *support* is the set

$$\operatorname{supp}\mu := \{x \in S : \mu(B(x,r)) > 0 \text{ for all } r > 0\}.$$

Recall the concept of Markov operators on measures, see Section 1.2. A measure μ_* is called *invariant* if $P\mu_* = \mu_*$. A Markov operator P is *asymptotically stable* if there exists a unique invariant measure $\mu_* \in \mathcal{P}(S)$ such that $P^n\mu \to \mu_*$ weakly as $n \to \infty$ for every $\mu \in \mathcal{P}(S)$.

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

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

If such 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 Feller operator P satisfies the *e–property* at $z \in S$ if for any $f \in BL(S)$ we have

$$\lim_{x \to z} \sup_{n \geq 0, n \in \mathbb{N}} |U^n f(x) - U^n f(z)| = 0, \tag{4.1}$$

i.e. if the family of iterates $\{U^n f : n \in \mathbb{N}\}$ is equicontinuous at $z \in S$. We say that a Feller operator satisfies the *e–property* if it satisfies it at any $z \in S$.

D. Worm slightly generalized the e–property introducing the Cesàro e–property (see [Wor10]). Namely, a Feller operator P will satisfy the *Cesàro e–property* at $z \in S$ if for any $f \in BL(S)$

we have

$$\lim_{x\to z}\sup_{n\geq 0, n\in\mathbb{N}}\left|\frac{1}{n}\sum_{k=1}^{n}U^k f(x)-\frac{1}{n}\sum_{k=1}^{n}U^k f(z)\right|=0. \tag{4.2}$$

Analogously a Feller operator satisfies the *Cesàro e–property* if it satisfies this property at any $z\in S$.

Let us recall Proposition 6.4.2 in [MT09] that contains - informally - a gap in its proof (slightly reformulated):

Proposition 4.2.1. *Suppose that the Markov chain Φ has the Feller property, and that there exists a unique probability measure π such that for every x*

$$P^n(x,\cdot)\to\pi \quad \textit{weakly as } n\to\infty$$

Then Φ is an e-chain.

The following example shows that Proposition 6.4.2 fails.

Example 4.2.2. *Let $S=\{1/n : n\geq 1\}\cup\{0\}$ and let $T:S\to S$ be given by the following formula:*

$$T(0)=T(1)=0 \qquad \textit{and} \qquad T(1/n)=1/(n-1) \quad \textit{for } n\geq 2.$$

The operator $P:\mathcal{M}(S)\to\mathcal{M}(S)$ given by the formula $P\mu=T_*(\mu)$ (the pushforward measure) is asymptotically stable but it does not satisfy the e–property at 0.

For a Markov operator Jamison [Jam64] introduced the property of *uniform stability in mean* when $\{U^n f : n\in\mathbb{N}\}$ is an equicontinuous family of functions in the space of real-valued continuous function $C(S)$ for every $f\in C(S)$. Here S is a compact metric space. Since the space of bounded Lipschitz functions is dense for the uniform norm in the space of bounded uniformly continuous functions, this property coincides with the Cesàro e–property for compact metric spaces. Now, if the Markov operator P on the compact metric space is asymptotically stable, with the invariant measure $\mu_*\in P(S)$, then $\frac{1}{n}\sum_{k=1}^{n}U^k f\to\langle f,\mu_*\rangle$ pointwise, for every $f\in C(S)$. According to Theorem 2.3 in [Jam64] this implies that P is uniformly stable in mean, i.e. has the Cesàro e–property.

Example 4.2.3. *Let $(k_n)_{n\geq 1}$ be an increasing sequence of prime numbers. Set*

$$S:=\{(\overbrace{0,\ldots,0,}^{k_n^i-1-times}\ i/k_n,0,\ldots)\in l^\infty : i\in\{0,\ldots,k_n\}, n\in\mathbb{N}\}.$$

The set S endowed with the l^∞-norm $\|\cdot\|_\infty$ is a (noncompact) Polish space. Define $T:S\to S$

by the formula

$$T((0,\ldots)) = T((\overbrace{0,\ldots,0,}^{k_n^{k_n}-1-times}\ 1,0,\ldots)) = (0,\ldots,0,\ldots) \qquad \textit{for } n \in \mathbb{N}$$

and

$$T((\overbrace{0,\ldots,0,}^{k_n^{i}-1-times}\ i/k_n,0,\ldots)) = (\overbrace{0,\ldots,0,}^{k_n^{i+1}-1-times}\ (i+1)/k_n,0,\ldots) \quad \textit{for } i \in \{1,\ldots,k_n-1\}, n \in \mathbb{N}.$$

The operator $P : \mathcal{M}(S) \to \mathcal{M}(S)$ given by the formula $P\mu = T_*(\mu)$ is asymptotically stable but it does not satisfy the Cesàro e–property at 0. Indeed, if we take an arbitrary continuous function $f : S \to \mathbb{R}_+$ such that $f((0,\ldots,0,\ldots)) = 0$ and $f(x) = 1$ for $x \in S$ such that $\|x\|_\infty \geq 1/2$ we have

$$\frac{1}{k_n}\sum_{i=1}^{k_n} U^i f((\overbrace{0,\ldots,0}^{k_n-1},1/k_n,0,\ldots)) - \frac{1}{k_n}\sum_{i=1}^{k_n} U^i f((0,\ldots)) \geq 1/2.$$

4.3 Main result

We are in a position to formulate the main result of this chapter. Recall that a metric d is called admissible for the Polish space S if d metrizes the topology on S and the metric space (S,d) is separable and complete.

Theorem 4.3.1. *Let P be an asymptotically stable Feller operator and let μ_* be its unique invariant measure. If $\mathrm{Int}_S(\mathrm{supp}\,\mu_*) \neq \varnothing$, then P satisfies the e–property for any admissible metric d on S.*

Its proof involves the following two lemmas:

Lemma 4.3.2. *Let P be an asymptotically stable Feller operator and let μ_* be its unique invariant measure. Let U be dual to P. If $\mathrm{Int}_S(\mathrm{supp}\,\mu_*) \neq \varnothing$, then for every admissible metric d on S, $f \in C_b(S)$ and any $\varepsilon > 0$ there exists a ball $B \subset \mathrm{supp}\,\mu_*$ such that*

$$|U^n f(x) - U^n f(y)| \leq \varepsilon \qquad \textit{for any } x,y \in B,\ n \in \mathbb{N}. \tag{4.3}$$

Proof. Fix $f \in C_b(S)$ and $\varepsilon > 0$. Let W be an open set in S such that $W \subset \mathrm{supp}\,\mu_*$. Set $Y = \overline{W}$ and observe that the subspace Y is a complete metric space, hence a Baire space.

Set

$$Y_n := \{x \in Y : |U^m f(x) - \langle f, \mu_* \rangle| \le \varepsilon/2 \text{ for all } m \ge n\}$$

and observe that Y_n is closed and

$$Y = \bigcup_{n=1}^{\infty} Y_n.$$

By the Baire Category Theorem there exist $N \in \mathbb{N}$ such that $\mathrm{Int}_Y Y_N \neq \varnothing$. Thus there exists a set $V \subset Y_N$ open in the space Y and consequently, because of the construction of Y, an open ball $B = B(z, r_0)$ for the admissible metric d in S such that $B \subset Y_N \subset \operatorname{supp} \mu_*$. Since

$$|U^n f(x) - \langle f, \mu_* \rangle| \le \varepsilon/2 \qquad \text{for any } x \in B \text{ and } n \ge N,$$

condition (4.3) is satisfied for all $x, y \in B, n \ge N$. Since the $U^n f$, $n = 1, \cdots, N$ are continuous at z, there exists $r^\varepsilon \le r_0$ such that $|U^n f(z) - U^n f(x)| \le \frac{\varepsilon}{2}$ for all $x \in B(z, r^\varepsilon)$, $n = 1, \cdots, N$. Then condition (4.3) is satisfied for all $x, y \in B := B(z, r^\varepsilon)$ and $n \in \mathbb{N}$. □

Lemma 4.3.3. *Let $\alpha \ge 0$. If $\mu \in \mathcal{M}^+(S)$, $x_0 \in S$ and $r > 0$ are such that $\mu(B(x_0, r)) > \alpha$, then there exists $0 < r \le r$ such that $\mu((B(x_0, r')) > \alpha$ and $\mu(S(x_0, r)) = 0$.*

Proof. For any increasing sequence $(r_n) \subset (0, r]$ such that $r_n \uparrow r$, $\mu(B(x_0, r_n)) \to \mu(B(x_0, r)) > \alpha$. Hence there exists $n_0 \in \mathbb{N}$ such that: $\mu(B(x_0, r_{n_0})) > \alpha$.

Put $r_0 := r_{n_0}$. Then $r_0 > 0$ and $\mu(B(x_0, r')) > \alpha$ for all $r' \in [r_0, r]$. The map $\Psi : [r_0, r] \times S \mapsto \mathbb{R} : (r', x) \mapsto \frac{d(x, x_0)}{r'}$ is separately continuous in r' and x, so it is jointly Borel measurable ([Bog07a], Theorem 7.14.5, p.129).

$$\begin{aligned} \mu(B(x_0, r')) &= \int_S \mathbb{1}_{B(x_0, r')}(y) d\mu(y) \\ &= \int_S \mathbb{1}_{\left\{x : \frac{d(x, x_0)}{r'} < 1\right\}}(y) d\mu(y) \\ &= \int_S \mathbb{1}_{[0,1)}(\Psi(r', y)) d\mu(y). \end{aligned} \tag{4.4}$$

Since Ψ is jointly Borel measurable, $(r', y) \mapsto \mathbb{1}_{[0,1)}(\Psi(r', y))$ is jointly Borel measurable. By the Fubini-Tonelli Theorem (or [Bog07a], Lemma 7.6.4, p.93, or [Bog07b], Corollary 3.3.3, p.182), $\underline{\phi} : r' \mapsto \mu(B(x_0, r'))$ is Borel measurable on $[r_0, r]$. In a similar manner, one shows that $\overline{\psi} := \mu(\overline{B}(x_0, r))$ is Borel measurable, where $\overline{B}(x_0, r) := \{x \in S : d(x, x_0) \le r\}$. Put $\phi(r) := \overline{\phi}(r) - \underline{\phi}(r)$. According to Lusin's Theorem, there exists a compact subset K of $[r_0, r]$, of strictly positive Lebesgue measure, such that $\phi|_K$ is continuous. Put $S(x_0, r') := \overline{B}(x_0, r') \smallsetminus B(x_0, r') = \{x \in S : d(x, x_0) = r'\}$.

Since Lebesgue measure is non-atomic, K must have at least denumerably many distinct points. Let $(r_n)_{n\in\mathbb{N}}$ be a sequence in K that consists of distinct points. Since K is a compact space, there is a subsequence $(r_{n_k})_{k\in\mathbb{N}}$ that converges to an $r' \in K$ as $k \to \infty$.

We can construct a further subsequence from $(r_{n_k})_{k\in\mathbb{N}}$ (denoted the same for convenience), that is either strictly increasing, or strictly decreasing towards r'.

(1) $r_{n_k} \uparrow r'$:

Define $A_1 := B(x_0, r_{n_1}), A_k := B(x_0, r_{n_k}) \setminus \overline{B}(x_0, r_{n_{k-1}})$.

Then

$$B(x_0, r') = \dot{\bigcup_{k=1}^{\infty}} A_k \cup S(x_0, r_{n_k}).$$

So

$$\mu(B(x_0, r')) = \sum_{k=1}^{\infty} \mu(A_k) + \mu(S(x_0, r_{n_k})) < \infty.$$

Hence, $\lim_{k\to\infty} \mu(S(x_0, r_{n_k})) = 0$. Because $r_{n_k} \in K$ and $\phi|_K$ is continuous, we get

$$\mu(S(x_0, r)) = \lim_{k\to\infty} \mu(S(x_0, r_{n_k})) = 0.$$

(2) $r_{n_k} \downarrow r'$:

Now define $A_k := B(x_0, r_{n_k}) \setminus \overline{B(x_0, r_{n_{k+1}}}$ for $k = 1, 2, \cdots$. Then

$$B(x_0, r_{n_1}) = \bigcup_{k=1}^{\infty} [A_k \cup S(x_0, r_{n_{k+1}})] \cup \overline{B(x_0, r')}.$$

Hence, $\lim_{k\to\infty} \mu(S(x_0, r_{n_k})) = 0$, as above, yielding the conclusion that $\mu(S(x_0, r')) = 0$.

Since $\partial B(x_0, r') \subset S(x_0, r')$ we find $\mu(\partial B(x_0, r') = 0$.

□

We are now ready to prove Theorem 4.3.1.

Proof. (Theorem 4.3.1) Assume, contrary to our claim, that P does not satisfy the e–property for some admissible metric d on S. Therefore there exist a function $f \in BL(S, d) \subset C_b(S)$ and a point $x_0 \in S$ such that

$$\limsup_{x\to x_0} \sup_{n\geq 0, n\in\mathbb{N}} |U^n f(x) - U^n f(x_0)| > 0.$$

Hence, there exists $\epsilon > 0$ and $\delta_0 > 0$ such that for all $0 < \delta < \delta_0$,

$$\sup_{x \in B(x_0,\delta)} \sup_{n \geq 0, n \in \mathbb{N}} |U^n f(x) - U^n f(x_0)| \geq 4\varepsilon.$$

Thus, one has a sequence $(x_k)_{k \in \mathbb{N}}$ such that $x_k \in (B(x_0, \frac{\delta_0}{k}))$ and

$$\sup_{n \geq 0, n \in \mathbb{N}} |U^n f(x_k) - U^n f(x_0)| \geq 3\varepsilon \quad \text{for all } k \in \mathbb{N}.$$

Let $B_f = B(z, 2r)$ be an open ball contained in $\operatorname{supp} \mu_*$ such that

$$|U^n f(x) - U^n f(y)| \leq \varepsilon \qquad \text{for all } x, y \in B_f, n \in \mathbb{N}, \tag{4.5}$$

which exists according to Lemma 4.3.2. Since $B_f \subset \operatorname{supp} \mu_x$, one has $\gamma := \mu_*(B_f) > 0$. Choose $\alpha \in (0, \gamma)$ Because P is asymptotically stable, by the Alexandrov Theorem (eg. [EK86], Theorem 3.1) one has

$$\liminf_{n \to \infty} P^n \mu(B_f) \geq \mu_*(B_f) = \gamma > \alpha \qquad \text{for all } \mu \in \mathcal{P}(S), \tag{4.6}$$

Fix $N \in \mathbb{N}$ such that $2(1-\alpha)^N \|f\|_\infty < \epsilon$. Inductively we shall define measures $\nu_i^{x_0}, \mu_i^{x_0}, \nu_i^{x_k}, \mu_i^{x_k}$ and integers $n_i, i = 1, 2, \cdots, N$ in the following way:

Equation (4.6) allows us to choose $n_1 \geq 1$ such that

$$P^{n_1} \delta_{x_0}(B(z, r)) > \alpha. \tag{4.7}$$

According to Lemma 4.3.3 it is possible to choose $0 < r_1 \leq r$ such that

$$P^{n_1} \delta_{x_0}(B(z, r_1)) > \alpha \quad \text{and} \quad P^{n_1} \delta_{x_0}(S(z, r_1)) = 0.$$

Define

$$\nu_1^x(\cdot) = \frac{P^{n_1} \delta_x(\cdot \cap B(z, r_1))}{P^{n_1} \delta_x(B(z, r_1))}. \tag{4.8}$$

Because $P^{n_1} \delta_{x_0}(S(z, r_1)) = 0$ and P is Feller, $P^{n_1} \delta_x(B(z, r_1))$ converges to $P^{n_1} \delta_{x_0}(B(z, r_1)) > \alpha > 0$ if $x \to x_0$. So ν_1^x is a well-defined probability measure, concentrated on $B(z, r_1)$, for all x sufficiently close to x_0, say if $d(x, x_0) < d_1$, and $P^{n_1} \delta_x(B(z, r_1)) > \alpha$ for such x.

Define

$$\mu_1^x(\cdot) = \frac{1}{1-\alpha} \left(P^{n_1} \delta_x(\cdot) - \alpha \nu_1^x(\cdot) \right). \tag{4.9}$$

Then $\mu_1^x \in \mathcal{P}(S)$ for all $x \in S$: $d(x, x_0) < d_1$.

Since $x_k \to x_0$, there exists $N_1 \in \mathbb{N}$ such that $d(x_k, x_0) < d_1$ for all $k \geq N_1$. If $U \subset S$ is open, then by Alexandrov's Theorem,

$$\liminf_{k\to\infty} P^{n_1}\delta_{x_k}(U \cap B(z, r_1)) \geq P^{n_1}\delta_{x_0}(U \cap B(z, r)).$$

Consequently,

$$\liminf_{k\to\infty} \nu_1^{x_k}(U) = \liminf_{k\to\infty} \frac{P^{n_1}\delta_{x_k}(U \cap B(z, r_1))}{P^{n_1}\delta_{x_k}(B(z, r_1))} \geq \frac{P^{n_1}\delta_{x_0}(U \cap B(z, r_1))}{P^{n_1}\delta_{x_0}(B(z, r_1))} = \nu_1^{x_0}(U).$$

Thus, $\nu_1^{x_k} \to \nu_1^{x_0}$ weakly as $k \to \infty$. Then also $\mu_1^{x_k} \to \mu_1^{x_0}$.

Assume that we have defined $\nu_i^{x_0}, \mu_i^{x_0}, \nu_i^{x_k}, \mu_i^{x_k}$ and n_i for $i = 1, 2, \cdots, l$, for some $l < N$ such that $\nu_i^{x_k} \to \nu_i^{x_0}, \mu_i^{x_k} \to \mu_i^{x_0}$ weakly. Then, equation (4.6) allows to pick $n_{l+1} \in \mathbb{N}$ such that

$$P^{n_{l+1}}\mu_l^{x_0}(B(z, r)) > \alpha.$$

According to Lemma 4.3.3 one can select $0 < r_{l+1} \leq r$ such that $P^{n_{l+1}}\mu_l^{x_0}(B(z, r_{l+1})) > \alpha$ and $P^{n_{l+1}}\mu_l^{x_l}(S(z, r_{l+1})) = 0$. Define

$$\nu_{l+1}^{x_k} := \frac{P^{n_{l+1}}\mu_l^{x_k}(\cdot \cap B(z, r_{l+1}))}{P^{n_{l+1}}\mu_l^{x_k}(B(z, r_{l+1}))} \tag{4.10}$$

and

$$\mu_{l+1}^{x_k} := \frac{1}{1-\alpha}(P^{n_{l+1}}\mu_l^{x_k} - \alpha\nu_{l+1}^{x_k}). \tag{4.11}$$

Because $\mu_l^{x_k} \to \mu_l^{x_0}$ weakly, and $P^{n_{l+1}}\mu_l^{x_k}(\partial B(z, r_{l+1})) = 0$.

$P^{n_{l+1}}\mu_l^{x_k}(B(z, r_{l+1})) \to P^{n_{l+1}}\mu_l^{x_0}(B(z, r_{l+1})) > \alpha > 0$ as $k \to \infty$. Thus, $\nu_{l+1}^{x_k}$ is well defined for k sufficiently large and $\nu_{l+1}^{x_k} \to \nu_{l+1}^{x_0}$, weakly, by a similar argument as for $\nu_1^{x_k} \to \nu_1^{x_0}$. We conclude from (4.11), that $\mu_{l+1}^{x_k} \to \mu_{l+1}^{x_0}$ weakly too.

Moreover, the construction is such that we have

$$P^{n_1+n_2+\cdots+n_N}\delta_{x_k} = \alpha P^{n_2+\cdots+n_N}\nu_i^{x_k} + \alpha(1-\alpha)P^{n_3+\cdots+n_N}\nu_2^{x_k} + \cdots + \alpha(1-\alpha)^{N-1}\nu_N^{x_k} + (1-\alpha)^N\mu_N^{x_k}$$

for $k = 0$ and all $k \in \mathbb{N}$ sufficiently large. By construction, $\operatorname{supp}\nu_i^{x_k} \subset \overline{B(z, r)} \subset B(z, 2r) =$

B_f. So for all $n \in \mathbb{N}$, $i = 1, 2, \cdots, N$ and k sufficiently large

$$\begin{aligned} |\langle P^n \nu_i^{x_k}, f\rangle - \langle P^n \nu_i^{x_0}, f\rangle| &= \left| \int_S U^n f(x) \nu_i^{x_k}(dx) - \int_S U^n f(y) \nu_i^{x_0}(dy) \right| \\ &\leq \int_{B_f} \int_{B_f} |U^n f(x) - U^n f(y)| \, \nu_i^{x_k}(dx) \nu_i^{x_0}(dy) \\ &\leq \epsilon. \end{aligned}$$

Moreover, there exists $N_0 \in \mathbb{N}$ such that for all $k \geq N_0$,

$$|\langle P^n \delta_{x_k} - P^n \delta_{x_0}, f\rangle| < \epsilon$$

for all $0 \leq n < n_1 + n_2 + \cdots + n_N$. For $n \geq n_1 + n_2 + \cdots + n_N$ one has for k sufficiently large,

$$\begin{aligned} P^n \delta_{x_k} &= \alpha P^{n-n_1} \nu_1^{x_k} + \alpha(1-\alpha) P^{n-n_1-n_2} \nu_2^{x_k} + \cdots + \\ &\quad + \alpha(1-\alpha)^{N-1} P^{n-n_1-\cdots-n_N} \nu_N^{x_k} + (1-\alpha)^N P^{n-n_1-\cdots-n_N} \mu_N^{x_k}. \end{aligned}$$

Therefore, for these n and k,

$$\begin{aligned} |\langle P^n \delta_{x_n}, f\rangle - \langle P^n \delta_{x_0}, f\rangle| &\leq \varepsilon(\alpha + \alpha(1-\alpha) + \cdots + \alpha(1-\alpha)^{N-1}) + 2(1-\alpha)^N \|f\|_\infty \\ &\leq \varepsilon + \varepsilon = 2\varepsilon. \end{aligned}$$

Thus, the construction of the $(x_k)_{k \in \mathbb{N}}$ is such that for k sufficiently large

$$3\varepsilon \leq \sup_{n \geq 0, n \in \mathbb{N}} |U^n f(x_k) - U^n f(x_0)| = \sup_{n \geq 0} |\langle P^n \delta_{x_k}, f\rangle - \langle P^n \delta_{x_0}, f\rangle| \leq 2\varepsilon$$

which is impossible. This completes the proof.

□

Acknowledgements. We thank Klaudiusz Czudek for providing us with Example 4.2.2 (communicated through Tomasz Szarek).

Chapter 5

Central Limit Theorem for some non-stationary Markov chains

This chapter is based on:

Jacek Gulgowski, Sander C. Hille, Tomasz Szarek, Maria A. Ziemlańska. Central Limit Theorem for some non-stationary Markov chains. *Studia Mathematica*, Number 246 (2019), Pages 109-131.

Abstract:

Using the classical Central Limit Theorem for stationary Markov chains proved by M. I. Gordin and B. A. Lifšic in [GL78] we show that it also holds for non–stationary Markov chains provided the transition probabilities satisfy the spectral gap property in the Kantorovich–Rubinstein norm.

5.1 Introduction

In this chapter we aim at showing that the Central Limit Theorem (**CLT**) obtained for stationary Markov chains by Gordin and Lifšic in [GL78] (see also its extension due to M. Maxwell and M. Woodroofe in [MW00]) may be extended to non–stationary ones provided that their transition probabilities satisfy the Spectral Gap Property in the Kantorovich–Rubinstein norm (see condition **(A2)** below). This condition implies that the iteration of the Markov operator associated to these transition probabilities constitute an equicontinuous family of continuous maps on probability measures, equipped with the Dudley metric, see Chapter 1, Theorem 2.8.4. In this way we obtain the result stronger than the almost sure result known in the literature as a quenched Central Limit Theorems (see [Pel15, BI95]).

Recently the **CLT** was proved for various non–stationary Markov processes (see [KW12, Kuk02]). For more details we refer readers to the book by T. Komorowski *et al.* [KLO12], where a more detailed description of recent results on central limit theorems is provided. Our result is in the same spirit as the main theorem in [KW12]. However some delicate approximation allow us to obtain the **CLT** for initial distributions with 2-nd moment finite instead of $2+\delta$.

In this chapter we introduce notation that deviates from previous chapters, mainly because those used here are more common in the field of probability theory, while the topic considered is specially targeted to an audience from this field.

Suppose that (X,ρ) is a Polish space. By $\mathcal{B}(X)$ we denote the family of all Borel sets in X. Denote by $B_b(X)$ the set of all bounded Borel measurable functions equipped with the supremum norm and let $C_b(X)$ be its subset consisting of all bounded continuous functions.

By $\mathcal{M}_1$ and $\mathcal{M}$ we denote the spaces of all probability Borel measures and of all Borel measures on X, respectively. Let $\pi : X\times\mathcal{B}(X)\to[0,1]$ be a transition probability on X and let $U : B_b(X)\to B_b(X)$ be defined by $Uf(x)=\int_X f(y)\pi(x,\mathrm{d}y)$ for every $f\in B_b(X)$. The operator U is dual to the Markov operator P defined on $\mathcal{M}$ and given by the formula $P\mu(\cdot)=\int_X \pi(x,\cdot)\mu(\mathrm{d}x)$ for $\mu\in\mathcal{M}$, i.e.

$$\int_X f(x)P\mu(\mathrm{d}x)=\int_X Uf(x)\mu(\mathrm{d}x)\quad\text{for any } f\in B_b(X) \text{ and } \mu\in\mathcal{M}.$$

In particular, we have $P\delta_x(\cdot)=\pi(x,\cdot)$ for $x\in X$.

Suppose that (X_n) is an X–valued Markov chain given over some probability space $(\Omega,\mathcal{F},P)$, whose transition probability is π. If the distribution of X_0 is μ_0, then the distribution of X_n equals $P^n\mu_0$, $n \geq 1$. By $(\mathcal{F}_n)$ we denote the natural filtration of the chain, i.e. the increasing family of σ–algebras $\mathcal{F}_n := \sigma(X_i : i \leq n)$.

For any $x \in X$ and $n \geq 1$ we define the measure $\mathbf{P}_x^n$ on $(X^n, \mathcal{B}(X^n))$, where

$$\mathcal{B}(X^n) = \sigma(\{A_1 \times A_2 \times \cdots \times A_n : A_i \in \mathcal{B}(X), i = 1, \ldots, n\})$$

by the formula

$$\begin{aligned}
&\mathbf{P}_x^n(A_1 \times \cdots \times A_n) \\
&= \int_{X^n} \chi_{A_1 \times \cdots \times A_n}(x_1, \ldots, x_n) \pi(x_{n-1}, \mathrm{d}x_n) \cdots \pi(x_1, \mathrm{d}x_2) \delta_x(\mathrm{d}x_1) \\
&= \int_{X^n} \chi_{A_1 \times \cdots \times A_n}(x_1, \ldots, x_n) P\delta_{x_{n-1}}(\mathrm{d}x_n) \cdots P\delta_{x_1}(\mathrm{d}x_2) P\delta_x(dx_1).
\end{aligned}$$

Here $\chi_{A_1 \times \cdots \times A_n}$ denotes the characteristic function of $A_1 \times \cdots \times A_n$. Obviously, the distribution of the random vector $(X_1, \ldots, X_n)$ if $X_0 = x$ is given by $\mathbf{P}_x^n$. On the other hand, if the distribution of X_0 equals μ_0, then the distribution of the random vector $(X_1, \ldots, X_n)$ is equal to $\mathbf{P}_{\mu_0}^n(\cdot) = \int_X \mathbf{P}_x^n(\cdot)\mu_0(\mathrm{d}x)$.

By $\mathbf{P}_x$ we will denote the probability measure on the Borel σ-algebra of the trajectory space X^∞ associated with (X_n) with $X_0 = x$. Further, $\mathbb{E}_x$ denotes the expected value with respect to $\mathbf{P}_x$. Analogously, if X_0 is distributed with μ_0, then $\mathbf{P}_{\mu_0}(\cdot) = \int_X \mathbf{P}_x(\cdot)\mu_0(\mathrm{d}x)$. Similarly, we have then $\mathbb{E}_\mu(\cdot) = \int_X \mathbb{E}_x(\cdot)\mu(\mathrm{d}x)$.

5.2 Assumptions

Set

$$\mathcal{M}_1^p := \left\{\mu \in \mathcal{M}_1 \ : \ \int_X [\rho(x, x_0)]^p \, \mu(dx) < \infty\right\} \qquad \text{for } p \geq 1.$$

We equip the space $\mathcal{M}_1^1$ with the Kantorovich-Rubinstein distance

$$\|\mu - \nu\| := \sup \left| \int_X f \mathrm{d}\mu - \int_X f \mathrm{d}\nu \right| \qquad \text{for } \mu, \nu \in \mathcal{M}_1^1,$$

where the supremum is taken over all Lipschitz functions $f : X \to \mathbb{R}$ with the Lipschitz constant bounded by 1. If μ and ν are two Borel probability measures on some Polish spaces W and Z respectively, then by $\mathcal{C}(\mu, \nu)$ we denote the set of all joint Borel probability measures on $W \times Z$ whose marginals are μ and ν (the so–called *coupling*). If $\mu, \nu \in \mathcal{M}_1^1$, then according to the following Kantorovich-Rubinstein Theorem

Theorem 5.2.1. *[Kantorovich-Rubinstein Theorem [Bog07a], Theorem 8.10.45] The Kantorovich-Rubinstein distance $\|\mu-\nu\|$ between Radon probability measures $\mu, \nu \in \mathcal{M}_1^1$ can be represented in the form*

$$\|\mu - \nu\| = \inf_{\gamma \in \mathcal{C}(\mu,\nu)} \int_{X\times X} \rho(x,y)\,\gamma(\mathrm{d}x, \mathrm{d}y).$$

Moreover, there exists a measure $\lambda_0 \in C(\mu,\nu)$ at which the value $\|\mu-\nu\|$ is attained.

This immediately implies

$$\|\delta_x - \delta_y\| = \rho(x,y) \qquad \text{for } x, y \in X.$$

Assumptions:

(A1) We assume that P leaves $\mathcal{M}_1^1$ invariant;

(A2) there exist $C > 0$ and $0 < q < 1$ such that

$$\|P^n\mu - P^n\nu\| \le Cq^n\|\mu - \nu\| \tag{5.1}$$

for all $\mu, \nu \in \mathcal{M}_1^1$ and $n \in \mathbb{N}$;

(A3) we assume that there exists $\mu \in \mathcal{M}_1^2$ such that

$$\sup_{n\ge 1} \int_X [\rho(x,x_0)]^2\, P^n\mu(\mathrm{d}x) < +\infty \qquad \text{for some (thus all) } x_0 \in X. \tag{5.2}$$

The following theorem is standard but we provide its proof for completeness of our presentation.

Proposition 5.2.1. *If* **(A1)** *and* **(A2)** *hold, then P has a unique invariant measure $\mu_* \in \mathcal{M}_1^1$ and for any $\mu \in \mathcal{M}_1^1$ we have $\|P^n\mu - \mu_*\| \to 0$ as $n \to \infty$. Moreover, if* **(A3)** *holds, then $\mu_* \in \mathcal{M}_1^2$.*

Proof. For any $\mu \in \mathcal{M}_1^1(X)$ and any $m, n \in \mathbb{N}$, $m \ge n$ one obtains from (5.1) that

$$\|P^n\mu - P^m\mu\| \le \sum_{k=0}^{m-n-1} \|P^{n+k}\mu - P^{n+k+1}\mu\| \le \frac{Cq^n}{1-q}\|\mu - P\mu\|.$$

So $(P^n\mu)$ is a Cauchy sequence in the $\|\cdot\|$-complete space $\mathcal{M}_1^1$ ([Vil08], Theorem 6.18). Hence it converges to some invariant measure $\mu_* \in \mathcal{M}_1^1(X)$. Then again by (5.1) for any

$\mu \in \mathcal{M}_1^1(X)$,

$$\|P^n\mu - \mu_*\| = \|P^n\mu - P^n\mu_*\| \le Cq^n\|\mu - \mu_*\| \to 0 \quad \text{as } n \to \infty.$$

So μ_* must be unique. Now suppose that **(A3)** holds and let $\mu \in \mathcal{M}_1^2$ be such that (5.2) is satisfied. Let $x_0 \in X$. For any $K > 0$, one has

$$\sup_{n\ge1} \int_X [\rho(x,x_0) \wedge K]^2\, P^n\mu(\mathrm{d}x) \le \sup_{n\ge1} \int_X \rho(x,x_0)^2\, P^n\mu(\mathrm{d}x) =: M_0 < \infty.$$

Since $x \mapsto [\rho(x,x_0) \wedge K]^2$ is a Lipschitz function on X and $\|P^n\mu - \mu_*\| \to 0$, we have

$$\int_X [\rho(x,x_0) \wedge K]^2\, P^n\mu(\mathrm{d}x) \to \int_X [\rho(x,x_0) \wedge K]^2\, \mu_*(\mathrm{d}x).$$

So $\int_X [\rho(x,x_0) \wedge K]^2\, \mu_*(\mathrm{d}x) \le M_0$ for every K. By the Monotone Convergence Theorem we obtain that $\int_X [\rho(x,x_0)]^2\, \mu_*(\mathrm{d}x) \le M_0$ too. □

5.3 Gordin–Lifšic results for stationary case

We start with the following simple consequences of the Gordin and Lifšic result on the **CLT** for stationary Markov chains:

Proposition 5.3.1. *Let P be a Markov operator that satisfies* **(A1) - (A3)** *and let (X_n) be a stationary Markov chain corresponding to P. Then for any bounded Lipschitz function $g : X \to \mathbb{R}$ such that $\int_X g\mathrm{d}\mu_* = 0$, where μ_* is the unique invariant distribution for P, the limit*

$$\sigma^2 := \lim_{n\to\infty} \mathbb{E}_{\mu_*}\left(\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}}\right)^2$$

exists and is finite. Moreover, if $\sigma > 0$, then

$$\lim_{n\to\infty} P\left(\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} < a\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^a e^{-\frac{y^2}{2\sigma^2}}\mathrm{d}y \qquad \textit{for all } a \in \mathbb{R}.$$

Otherwise, if $\sigma = 0$, then the sequence

$$\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} \qquad \textit{for } n \ge 1$$

converges in distribution to 0.

Proof. Let a bounded Lipschitz function $g : X \to \mathbb{R}$ such that $\int_X g \mathrm{d}\mu_* = 0$ be given. With no loss of generality we may assume that the Lipschitz constant of g equals 1. From Gordin and Lifšic [GL78] it follows that to finish the proof it is enough to show that

$$\sum_{n=0}^{\infty} \|U^n g\|_{L^2(X,\mu_*)} < \infty,$$

where U is dual to P. In fact, since $\int_X U^i g(x)\mu_*(\mathrm{d}x) = \int_X g(x)\mu_*(\mathrm{d}x) = 0$ for all $i \geq 0$, we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \|U^n g\|_{L^2(X,\mu_*)} &= \sum_{n=1}^{\infty} \left(\int_X (U^n g(x))^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&= \sum_{n=1}^{\infty} \left(\int_X (U^n g(x) - \int_X U^n g(y)\mu_*(\mathrm{d}y))^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&\leq \sum_{n=1}^{\infty} \left(\int_X \left(\int_X |U^n g(x) - U^n g(y)| \mu_*(\mathrm{d}y) \right)^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&\leq \sum_{n=1}^{\infty} \left(\int_X \left(\int_X C q^n \rho(x,y)\mu_*(\mathrm{d}y) \right)^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&\leq C \sum_{n=1}^{\infty} q^n \left(\int_X \left(\int_X \rho(x,y)\mu_*(\mathrm{d}y) \right)^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&\leq C \sum_{n=1}^{\infty} q^n \left(\int_X \left(\rho(x,x_0) + \int_X \rho(x_0,y)\mu_*(\mathrm{d}y) \right)^2 \mu_*(\mathrm{d}x) \right)^{1/2} \\
&\leq C(1-q)^{-1} \left(2 \int_X [\rho(x,x_0)]^2 \mu_*(\mathrm{d}x) + 2 \left(\int_X \rho(x_0,y)\mu_*(\mathrm{d}y) \right)^2 \right)^{1/2} < \infty,
\end{aligned}$$

by the fact that $\mu_* \in \mathcal{M}_1^2$. The proof is complete. $\square$

5.4 Auxiliary lemmas

We start with a theorem on the existence of a suitable coupling for the trajectories of a given Markov chain.

Theorem 5.4.1. *Assume that a Markov operator P satisfies* **(A1)** *and* **(A2)**. *Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1). Then there exists a constant $\kappa > 0$ such that for every integers $l \geq l_0$ and $m \geq 1$ and every two points $x, y \in X$ we have a measure $\mathbf{P}_{x,y}^{ml} \in \mathcal{C}(\mathbf{P}_x^{ml}, \mathbf{P}_y^{ml})$ satisfying*

$$\sum_{i=1}^{m} \int_{X^{lm} \times X^{lm}} \rho(x_{il}, y_{il}) \mathbf{P}_{x,y}^{ml}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{ml}, \mathrm{d}y_1, \ldots, \mathrm{d}y_{ml}) \leq \kappa \rho(x,y). \tag{5.3}$$

Proof. Choose $\tilde{q} \in (Cq^{l_0}, 1)$. Fix $l \geq l_0$. By induction on m for every two points $x, y \in X$ we construct $\mathbf{P}^{ml}_{x,y} \in \mathcal{C}(\mathbf{P}^{ml}_x, \mathbf{P}^{ml}_y)$ and prove that

$$\int_{X^{lm}\times X^{lm}} \left(\sum_{i=1}^{m} \rho(x_{il}, y_{il}) \right) \mathbf{P}^{ml}_{x,y}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{ml}, \mathrm{d}y_1, \ldots, \mathrm{d}y_{ml}) \leq \sum_{j=1}^{m} \tilde{q}^j \rho(x, y). \tag{5.4}$$

Then the hypothesis will hold with $\kappa = \tilde{q}(1 - \tilde{q})^{-1}$. Fix $x, y \in X$ and let $m = 1$. Since

$$\|P^l \delta_x - P^l \delta_y\| \leq Cq^{l_0} \rho(x, y),$$

by the Kantorovich–Rubinstein theorem there exists $\mu^1_{x,y} \in \mathcal{C}(P^l\delta_x, P^l\delta_y)$ such that

$$\|P^l \delta_x - P^l \delta_y\| \leq \int_{X\times X} \rho(u, v) \mu^1_{x,y}(\mathrm{d}u, \mathrm{d}v) \leq \tilde{q}\rho(x, y).$$

From the proof of the Kantorovich–Rubinstein theorem it follows that the function $X \times X \ni (x, y) \to \mu^1_{x,y} \in \mathcal{C}(P^l\delta_x, P^l\delta_y)$ is measurable if the space $\mathcal{C}(P^l\delta_x, P^l\delta_y)$ is endowed with some metric of weak topology (see Theorem 11.8.2 in [Dud02]). We may assume that the measure $\mu^1_{x,y}$ is absolutely continuous with respect to the product measure $P^l\delta_x \otimes P^l\delta_y$. Let $g_{x,y} : X \times X \to [0, +\infty)$ be the Radon–Nikodem density, i.e.

$$g_{x,y}(u, v) = \frac{\mathrm{d}\mu^1_{x,y}}{\mathrm{d}(P^l\delta_x \otimes P^l\delta_y)}(u, v) \qquad u, v \in X.$$

Define

$$\begin{aligned}
&\mathbf{P}^l_{x,y}(A) \\
&= \int_{X^l} \cdots \int_{X^l} \chi_A(x_1, \ldots, x_l, y_1, \ldots, y_l) g_{x,y}(x_l, y_l) \pi(x_{l-1}, \mathrm{d}x_l) \pi(y_{l-1}, \mathrm{d}y_l) \\
&\cdots \pi(x_1, \mathrm{d}x_2) \pi(y_1, \mathrm{d}y_2) \delta_x(\mathrm{d}x_1) \delta_y(\mathrm{d}y_1).
\end{aligned}$$

Let $A = [A_1 \times \cdots \times A_l] \times X^l$ for some Borel sets $A_1, \ldots, A_l \subset X$. We have

$$\begin{aligned}
&\mathbf{P}^l_{x,y}(A) \\
&= \int_{X^l} \chi_{A_1\times\cdots\times A_l}(x_1, \ldots, x_l) \left(\int_{X^l} g_{x,y}(x_l, y_l) \pi(y_{l-1}, \mathrm{d}y_l) \cdots \pi(y_1, \mathrm{d}y_2) \delta_y(\mathrm{d}y_1) \right) \\
&\times \pi(x_{l-1}, \mathrm{d}x_l) \cdots \pi(x_1, \mathrm{d}x_2) \delta_x(\mathrm{d}x_1).
\end{aligned}$$

Moreover, the measure

$$\int_{X^{l-1}} \chi_{A_1\times\cdots\times A_{l-1}}(x_1, \ldots, x_{l-1}) \pi(x_{l-1}, \cdot) \cdots \pi(x_1, \mathrm{d}x_2) \delta_x(\mathrm{d}x_1)$$

is absolutely continuous with respect to $P^l\delta_x$. Obviously,

$$\int_{X^l} g_{x,y}(x_l,y_l)\pi(y_{l-1},\mathrm{d}y_l)\cdots\pi(y_1,\mathrm{d}y_2)\delta_y(\mathrm{d}y_1) = \int_X g_{x,y}(x_l,y_l)P^l\delta_y(\mathrm{d}y_l).$$

If we show that $\int_X g_{x,y}(\cdot,y_l)P^l\delta_y(\mathrm{d}y_l) = 1$, $P^l\delta_x$-a.s., then

$$\begin{aligned}
&\mathbf{P}^l_{x,y}(A)\\
&= \int_{X^l} \chi_{A_1\times\cdots\times A_l}(x_1,\ldots,x_l)\pi(x_{l-1},\mathrm{d}x_l)\cdots\pi(x_1,\mathrm{d}x_2)\delta_x(\mathrm{d}x_1)\\
&= \mathbf{P}^l_x(A_1\times\cdots\times A_l).
\end{aligned}$$

To do this set $h(x_l) := \int_X g_{x,y}(x_l,y_l)P^l\delta_y(\mathrm{d}y_l)$. From the definition of the coupling measure $\mu^1_{x,y}$ for any Borel set $B\subset X$ we have

$$\begin{aligned}
P^l\delta_x(B) = \mu^1_{x,y}(B\times X) &= \int_B\left(\int_X g_{x,y}(x_l,y_l)P^l\delta_y(\mathrm{d}y_l)\right)P^l\delta_x(\mathrm{d}x_l)\\
&= \int_B h(x_l)P^l\delta_x(\mathrm{d}x_l).
\end{aligned}$$

Since the above equality holds for an arbitrary Borel set $B\subset X$, we obtain that $h(x_l)=1$, $P^l\delta_x$-a.s., and consequently $\mathbf{P}^l_{x,y}([A_1\times\cdots\times A_l]\times X^l) = \mathbf{P}^l_x(A_1\times\cdots\times A_l)$. In the same way we show that $\mathbf{P}^l_{x,y}(X^l\times[B_1\times\cdots\times B_l]) = \mathbf{P}^l_y(B_1\times\cdots\times B_l)$ for Borel sets $B_1,\ldots,B_l\subset X$. Hence $\mathbf{P}^l_{x,y}\in\mathcal{C}(\mathbf{P}^l_x,\mathbf{P}^l_y)$. Finally, we have

$$\begin{aligned}
&\int_{X^l\times X^l}\rho(x_l,y_l)\mathbf{P}^l_{x,y}(\mathrm{d}x_1,\ldots\mathrm{d}x_l,\mathrm{d}y_1,\ldots,\mathrm{d}y_l)\\
&= \int_{X^l\times X^l}\rho(x_l,y_l)g_{x,y}(x_l,y_l)\pi(x_{l-1},\mathrm{d}x_l)\pi(y_{l-1},\mathrm{d}y_l)\cdots\pi(x_1,\mathrm{d}x_2)\pi(y_1,\mathrm{d}y_2)\delta_x(\mathrm{d}x_1)\delta_y(\mathrm{d}y_1)\\
&= \int_{X\times X}\rho(x_l,y_l)g_{x,y}(x_l,y_l)P^l\delta_x(\mathrm{d}x_l)P^l\delta_y(\mathrm{d}y_l) = \int_{X\times X}\rho(x_l,y_l)\mu^1(\mathrm{d}x_l,\mathrm{d}y_l)\le\tilde{q}\rho(x,y)
\end{aligned}$$

and the first step of the proof is finished.

Assume now that for $j=1,\ldots,m$ and arbitrary $x,y\in X$ there exists $\mathbf{P}^{jl}_{x,y}\in\mathcal{C}(\mathbf{P}^{jl}_x,\mathbf{P}^{jl}_y)$ such that condition (5.4) holds with m replaced by j.

Define the measure $\mathbf{P}^{(m+1)l}_{x,y}$ on $X^{(m+1)l}\times X^{(m+1)l}$ by the formula

$$\begin{aligned}
\mathbf{P}^{(m+1)l}_{x,y}(A) = &\int_{X^{(m+1)l}\times X^{(m+1)l}}\chi_A(x_1,\ldots,x_{(m+1)l},y_1,\ldots,y_{(m+1)l})\\
&\mathbf{P}^{ml}_{x_l,y_l}(\mathrm{d}x_{l+1},\ldots,\mathrm{d}x_{(m+1)l},\mathrm{d}y_{l+1},\ldots,\mathrm{d}y_{(m+1)l})\mathbf{P}^l_{x,y}(\mathrm{d}x_1,\ldots,\mathrm{d}x_l,\mathrm{d}y_1,\ldots,\mathrm{d}y_l)
\end{aligned}$$

for any Borel set $A \subset X^{(m+1)l} \times X^{(m+1)l}$. From the Markov property it easily follows that $\mathbf{P}_{x,y}^{(m+1)l} \in \mathcal{C}(\mathbf{P}_x^{(m+1)l}, \mathbf{P}_y^{(m+1)l})$. We also have

$$\begin{aligned}
&\int_{X^{(m+1)l}\times X^{(m+1)l}} \sum_{j=1}^{m+1} \rho(x_{jl}, y_{jl}) \mathbf{P}_{x,y}^{(m+1)l}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{(m+1)l}, \mathrm{d}y_1, \ldots, \mathrm{d}y_{(m+1)l}) \\
&\le \int_{X^l\times X^l} \left(\int_{X^{ml}\times X^{ml}} \sum_{j=2}^{m+1} \rho(x_{jl}, y_{jl}) \mathbf{P}_{x_l,y_l}^{ml}(\mathrm{d}x_{l+1}, \ldots, \mathrm{d}x_{(m+1)l}, \mathrm{d}y_{l+1}, \ldots, \mathrm{d}y_{(m+1)l}) \right) \\
&\quad \mathbf{P}_{x,y}^{l}(\mathrm{d}x_1, \ldots, \mathrm{d}x_l, \mathrm{d}y_1, \ldots, \mathrm{d}y_l) + \int_{X^l\times X^l} \rho(x_l, y_l) \mathbf{P}_{x,y}^{l}(\mathrm{d}x_1, \ldots \mathrm{d}x_l, \mathrm{d}y_1, \ldots, \mathrm{d}y_l) \\
&\le \sum_{j=1}^{m} \tilde{q}^j \int_{X^l\times X^l} \rho(x_l, y_l) \mathbf{P}_{x,y}^{l}(\mathrm{d}x_1, \ldots \mathrm{d}x_l, \mathrm{d}y_1, \ldots, \mathrm{d}y_l) + \tilde{q}\rho(x,y) \\
&\le \sum_{j=1}^{m} \tilde{q}^j \tilde{q}\rho(x,y) + \tilde{q}\rho(x,y) = \sum_{j=1}^{m+1} \tilde{q}^j \rho(x,y)
\end{aligned}$$

by the inductive hypothesis for 1 and m. This completes the proof. □

As a consequence of Proposition 5.3.1 and Theorem 5.4.1 we have the following:

Proposition 5.4.1. *Let P be a Markov operator that satisfies* **(A1) - (A3)** *and let (X_n) be a Markov chain corresponding to P. Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \ge l_0$ be given. Set $Y_n = X_{nl}$ for $n \ge 0$ and assume that $X_0 = Y_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \to \mathbb{R}$ such that $\int_X g \mathrm{d}\mu_* = 0$, where μ_* is the unique invariant distribution for P, the limit*

$$\tilde{\sigma}^2 = \lim_{n\to\infty} \mathbb{E}_{\mu_*} \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2$$

exists and is finite. Moreover, if $\tilde{\sigma}^2 > 0$, then the sequence of random vectors (Y_n) satisfies

$$\lim_{n\to\infty} P\left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} < a \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a} e^{-\frac{y^2}{2\tilde{\sigma}^2}} \mathrm{d}y$$

for all $a \in \mathbb{R}$.

Proof. Without loss of generality we may assume that $g : X \to \mathbb{R}$ is bounded by 1 and its Lipschitz constant is also bounded by 1. By φ_n^x, $x \in X$, we denote the characteristic function of the random variable

$$\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}},$$

where $Y_0 = x$. When the distribution of Y_0 is equal to μ_* the characteristic function is denoted by $\varphi_n^{\mu_*}$. Obviously,

$$\varphi_n^{\mu_*}(t) = \int_X \varphi_n^y(t)\mu_*(\mathrm{d}y) \qquad \text{for } t \in \mathbb{R}.$$

Moreover, from Proposition 5.3.1 it follows that

$$\lim_{n\to\infty} \varphi_n^{\mu_*}(t) = e^{-\frac{\tilde{\sigma}^2 t^2}{2}} \qquad \text{for } t \in \mathbb{R}.$$

Theorem 5.4.1 allows us to evaluate the difference of characteristic functions $\varphi_n^x(t)$ and $\varphi_n^y(t)$ for $x, y \in X$ and $t \in \mathbb{R}$. Fix $x, y \in X$. We have

$$\begin{aligned}
&|\varphi_n^x(t) - \varphi_n^y(t)| = |\int_{X^{ln}} \exp\left(it\frac{g(x_0) + g(x_l)\ldots + g(x_{nl})}{\sqrt{n}}\right)\mathbf{P}_x^{nl}(\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_{nl}) \\
&- \int_{X^{ln}} \exp\left(it\frac{g(y_0) + g(y_l)\ldots + g(y_{nl})}{\sqrt{n}}\right)\mathbf{P}_y^{nl}(\mathrm{d}y_1, \mathrm{d}y_2, \cdots, \mathrm{d}y_{nl})| \\
&\le \frac{|t|}{\sqrt{n}} \int_{X^{ln}\times X^{ln}} [\rho(x,y) + \rho(x_l, x_l) + \cdots + \rho(x_{nl}, y_{nl})]\, \mathbf{P}_{x,y}^{nl}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{nl}, \mathrm{d}y_1, \ldots, \mathrm{d}y_{nl}) \\
&\le \frac{|t|}{\sqrt{n}}(1+\kappa)\rho(x,y) \qquad \text{for } t \in \mathbb{R},
\end{aligned}$$

where $\mathbf{P}_{x,y}^{nl} \in \mathcal{C}(\mathbf{P}_x^{nl}, \mathbf{P}_y^{nl})$ is given by Theorem 5.4.1. Consequently, for $t \in \mathbb{R}$ we obtain

$$\begin{aligned}
&|\varphi_n^x(t) - \varphi_n^{\mu_*}(t)| = |\varphi_n^x(t) - \int_X \varphi_n^y(t)\mu_*(\mathrm{d}y)| \\
&\le \lim_{n\to\infty} \frac{|t|(1+\kappa)}{\sqrt{n}} \int_X \rho(x,y)\mu_*(\mathrm{d}y) \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Since $\lim_{n\to\infty} \varphi_n^{\mu_*}(t) = e^{-t^2\tilde{\sigma}^2/2}$, we obtain that $\lim_{n\to\infty} \varphi_n^x(t) = e^{-t^2\tilde{\sigma}^2/2}$ and the proof is complete. □

For any Markov chain (Z_n) with values in the space X and an arbitrary Lipschitz function $g : X \to \mathbb{R}$ we shall denote

$$S_m^n(g(Z_i)) := \sum_{i=m}^{n} g(Z_i) \qquad \text{for } n \ge m \ge 0.$$

The key point in proving the main theorem is the following lemma:

Lemma 5.4.2. *Let P be a Markov operator that satisfies* **(A1) - (A2)** *and let (X_n) be a*

Markov chain corresponding to P. Assume that $g : X \to \mathbb{R}$ is a bounded Lipschitz function. Then there exists a constant $M > 0$ such that for any $n \in \mathbb{N}$ the function

$$x \to \mathbb{E}_x \left(\frac{S_0^n(g(X_i))}{\sqrt{n}} \right)^2$$

has Lipschitz constant $\leq M$.

Proof. Again assume that $g : X \to \mathbb{R}$ is a 1–Lipschitz function bounded by 1. Let l_0 be a positive integer such that $Cq^{l_0} < 1$ with the constants C, q given by (5.1). We have

$$\begin{aligned}\mathbb{E}_x \left(\frac{S_0^n(g(X_i))}{\sqrt{n}} \right)^2 &= \frac{1}{n} \sum_{i=0}^{n} \mathbb{E}_x (g(X_i))^2 + \frac{2}{n} \sum_{i=0}^{n-1} \sum_{j:0<j-i<l_0, j\leq n} \mathbb{E}_x (g(X_i) g(X_j)) \\ &+ \frac{2}{n} \sum_{i=0}^{n} \sum_{j:0\leq j<l_0,\, l_0+j+i\leq n} \mathbb{E}_x [g(X_i)(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}))],\end{aligned}$$

where $k_{i,j}$ are maximal positive integers such that $k_{i,j} l_0 + j + i \leq n$. We show that all three terms are Lipschitzean with a Lipschitz constant independent of n. To do this fix $x, y \in X$. Since the function $(g(\cdot))^2$ is a ≤ 2–Lipschitz function, by (5.1) we have

$$|\mathbb{E}_x (g(X_i))^2 - \mathbb{E}_y (g(X_i))^2| \leq 2 \| P^i \delta_x - P^i \delta_y \| \leq 2C\rho(x, y) \qquad \text{for } i = 1, \ldots, n$$

and the first term

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_x (g(X_i))^2$$

is a $\leq 2C$–Lipschitz function.

Further, for $j > i$ we have

$$\mathbb{E}_z g(X_i) g(X_j) = \mathbb{E}_z (g(X_i) \mathbb{E}_{|X_i} g(X_j)) \qquad \text{for all } z \in X,$$

where $\mathbb{E}_{|X_i} g(X_j) := \mathbb{E}[g(X_j)|\mathcal{F}_i]$. Since $\mathbb{E}_{|X_i} g(X_j) = h(X_i)$, where h is a bounded by 1 a $\leq C$–Lipschitz function, we obtain that

$$\begin{aligned}&|\mathbb{E}_x g(X_i) g(X_j) - \mathbb{E}_y g(X_i) g(X_j)| = |\mathbb{E}_x (g(X_i) h(X_i)) - \mathbb{E}_y (g(X_i) h(X_i))| \\ &\leq 2C\rho(x, y)\end{aligned}$$

for $i = 1, \dots, n$ and consequently the Lipschitz constant of the second term

$$\frac{2}{n}\sum_{i=1}^{n-1}\sum_{j:0<j-i<l_0, j\le n}\mathbb{E}_x g(X_i)g(X_j)$$

is bounded from above by $4l_0C$.

Finally, we have

$$\mathbb{E}_x\left(g(X_i)(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}))\right) = \mathbb{E}_x\left(g(X_i)\mathbb{E}_{|X_i}(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}))\right).$$

Using Theorem 5.4.1 we show that the function

$$X \ni z \to \mathbb{E}_z(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j})) \qquad \text{for all } i, j \ge 0$$

has a Lipschitz constant $\le C\kappa q^j$, where κ is given by Theorem 5.4.1. Indeed, for any $i, j \ge 0$ we have

$$\mathbb{E}_z(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j})) = \mathbb{E}_z\mathbb{E}_{|X_j}(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j})).$$

In turn, from Theorem 5.4.1 we obtain that the function

$$X \ni u \to r(u) = \mathbb{E}_u(\sum_{m=1}^{k_{i,j}} g(X_{ml_0}))$$

has a Lipschitz constant $\le \kappa$. Indeed, fix $u, v \in X$. We have

$$\begin{aligned}
&|r(u) - r(v)| = |\mathbb{E}_u(\sum_{m=1}^{k_{i,j}} g(x_{ml_0})) - \mathbb{E}_v(\sum_{m=1}^{k_{i,j}} g(x_{ml_0}))| \\
&\le |\int_{(X^{k_{i,j}l_0})^2}\sum_{m=1}^{k_{i,j}}(g(x_{ml_0}) - g(y_{ml_0}))\mathbf{P}_{u,v}^{k_{i,j}l_0}(\mathrm{d}x_1, \dots, \mathrm{d}x_{k_{i,j}l_0}; \mathrm{d}y_1, \dots, \mathrm{d}y_{k_{i,j}l_0}) \\
&\le \int_{(X^{k_{i,j}l_0})^2}\sum_{m=1}^{k_{i,j}}\rho(x_{ml_0}, y_{ml_0})\mathbf{P}_{u,v}^{k_{i,j}l_0}(\mathrm{d}x_1, \dots, \mathrm{d}x_{k_{i,j}l_0}; \mathrm{d}y_1, \dots, \mathrm{d}y_{k_{i,j}l_0}) \\
&\le \kappa\rho(u, v),
\end{aligned}$$

by Theorem 5.4.1. Since

$$\mathbb{E}_z(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j})) = \int_X r(u)P^j\delta_z(\mathrm{d}u),$$

the function

$$X \ni z \to \mathbb{E}_z\Big(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j})\Big)$$

has a Lipschitz constant $\leq C\kappa q^j$, by (5.1). Moreover its supremum norm is bounded by n. Since the Lipschitz constant of the function $z \to g(z)r(z)$ is bounded by $n + C\kappa q^j$ we obtain that the function

$$X \ni x \to \mathbb{E}_x\left(g(X_i)\Big(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i})\Big)\right)$$

has a Lipschitz constant $\leq Cq^i(n + C\kappa q^j)$. Thus the third term

$$\frac{2}{n}\sum_{i=0}^{n}\sum_{j:0\leq j<l_0,\, l_0+j+i\leq n} \mathbb{E}_x\Big[g(X_i)\Big(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i})\Big)\Big]$$

has Lipschitz constant bounded by

$$\frac{2}{n}\sum_{i=0}^{n}\sum_{j=0}^{l_0-1} Cq^i(n + C\kappa q^j) \leq \frac{2C}{n}\sum_{i=0}^{n} l_0(n + C\kappa)q^i \leq \frac{2l_0C(1 + C\kappa)}{1-q}.$$

Thus the function

$$x \to \mathbb{E}_x\left(\frac{S_0^n(g(X_i))}{\sqrt{n}}\right)^2$$

has Lipschitz constant $\leq M$, where

$$M = 2C + 4l_0C + \frac{2l_0C(1 + C\kappa)}{1-q}.$$

This completes the proof. $\square$

Lemma 5.4.3. *Let P be a Markov operator that satisfies* **(A1) - (A3)** *and let (X_n) be a Markov chain corresponding to P. Assume that $g : X \to \mathbb{R}$ is a bounded Lipschitz function. Let $x \in X$ and $\varepsilon > 0$. Then there exists $K > 0$ and $N_0 \in \mathbb{N}$ such that*

$$\mathbb{E}_{P^n\delta_x}\left[\left(\frac{S_0^n(g(X_i))}{\sqrt{m}}\right)^2 \chi_{[K,+\infty)}\left(\left|\frac{S_0^n(g(X_i))}{\sqrt{m}}\right|\right)\right] \leq \varepsilon \tag{5.5}$$

for all $m \in \mathbb{N}$ and $n \geq N_0$.

Proof. Fix $x \in X$ and let $\varepsilon > 0$. Let g be given as above. Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \geq l_0$ be given. Set $Y_n = X_{nl}$ for $n \geq 0$. We first show that the hypothesis holds if we replace X_n with Y_n. Set

$$S_m := \frac{g(Y_0) + g(Y_1) + \ldots + g(Y_m)}{\sqrt{m}} \qquad \text{for } m \geq 0.$$

The Markov chain (Y_n) corresponds to the operator P^l and the assumptions of Lemma 5.4.2 are satisfied with P and (X_n) replaced by P^l and (Y_n), respectively. Thus the functions $y \to \mathbb{E}_y(S_m)^2$ are Lipschitzean with the Lipschitz constant bounded by, say $\tilde{M}$, independent of m. The Markov operator P satisfies **(A2)**, so that we may find $N_0 \in \mathbb{N}$ such that

$$|\mathbb{E}_{P^n\delta_x} S_m^2 - \mathbb{E}_{\mu_*} S_m^2| \leq \tilde{M} \| P^n \delta_x - \mu_* \| \leq \varepsilon/(9l^3) \quad \text{for } n \geq N_0 \text{ and } m \in \mathbb{N}. \tag{5.6}$$

Moreover, from Theorem 5.4.1 it follows that for any Lipschitz function $\Theta : \mathbb{R} \to \mathbb{R}$ with the Lipschitz constant $\hat{M} > 0$ such that we have for all $m \geq 1$

$$|\mathbb{E}_y \Theta(S_m) - \mathbb{E}_z \Theta(S_m)| \leq \frac{\kappa \hat{M}}{\sqrt{m}} \rho(y, z) \qquad \text{for } y, z \in X.$$

Consequently, for some $N_1 \in \mathbb{N}$ we have

$$|\mathbb{E}_{P^n\delta_x} \Theta(S_m) - \mathbb{E}_{\mu_*} \Theta(S_m)| \leq \varepsilon/(9l^3) \quad \text{for } n \geq N_1 \text{ and } m \in \mathbb{N}. \tag{5.7}$$

Let $\varphi_K : \mathbb{R} \to \mathbb{R}$, $K \geq 1$ be a Lipschitz function such that

$$\chi_{[0,K-1]}(y) \leq \varphi_K(y) \leq \chi_{[0,K]}(y) \qquad \text{for all } x \in \mathbb{R}.$$

We are going to show that there exists $K \geq 1$ such that

$$\mathbb{E}_{\mu_*}(S_m^2(1 - \varphi_K(|S_m|)) < \varepsilon/(9l^3) \qquad \text{for all } m \geq 1.$$

By Proposition 5.3.1 we have

$$\mathbb{E}_{\mu_*} S_m^2 \to \tilde{\sigma}^2 \qquad \text{as } m \to \infty$$

and

$$\mathbb{E}_{\mu_*}(S_m^2 \varphi_K(|S_m|)) \to \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^2 \varphi_K(|y|) e^{-y^2/(2\tilde{\sigma}^2)} \mathrm{d}y \qquad \text{as } m \to \infty.$$

In the case when $\tilde{\sigma} = 0$ we set $e^{-y^2/(2\tilde{\sigma}^2)} = 0$.

Since the integral on the right hand side converges to σ^2 as $K \to \infty$, we obtain that there exist constants m_0 and k_0 such that

$$\mathbb{E}_{\mu_*}(S_m^2(1-\varphi_{K_0}(|S_m|)) = \mathbb{E}_{\mu_*}S_m^2 - \mathbb{E}_{\mu_*}(S_m^2\varphi_{K_0}(|S_m|)) < \varepsilon/(9l^3) \quad \text{for all } m \geq m_0.$$

Enlarging K_0 if necessary we obtain

$$\mathbb{E}_{\mu_*}(S_m^2(1-\varphi_{K_0}(|S_m|)) < \varepsilon/(9l^3) \quad \text{for all } m \geq 1. \tag{5.8}$$

Observe that the function $y \to \Theta(y)$ given by the formula $\Theta(y) := y^2\varphi_{K_0}(|y|)$ is a Lipschitz function. Combining (5.6) - (5.8), we obtain that for $n \geq \max\{N_0, N_1\}$ we have

$$\begin{aligned}
&\mathbb{E}_{P^n\delta_x}[S_m^2\chi_{[K_0,+\infty)}(|S_m|)] \leq \mathbb{E}_{P^n\delta_x}S_m^2(1-\varphi_{K_0}(|S_m|)) \\
&\leq |\mathbb{E}_{P^n\delta_x}S_m^2 - \mathbb{E}_{\mu_*}S_m^2| + |\mathbb{E}_{P^n\delta_x}\Theta(S_m) - \mathbb{E}_{\mu_*}\Theta(S_m)| + \mathbb{E}_{\mu_*}(S_m^2(1-\varphi_{K_0}(|S_m|)) \\
&\leq \varepsilon/(9l^3) + \varepsilon/(9l^3) + \varepsilon/(9l^3) = \varepsilon/(3l^3).
\end{aligned}$$

Having this we may show that (5.5) holds with $K = lK_0$. Set $m_i = \max\{j : i + jl \leq m\}$ for $i = 0, \ldots, l-1$. We have

$$\begin{aligned}
&\mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right)^2 \chi_{[K,+\infty)}\left(\left|\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right|\right)\right] \\
&\leq l^2\mathbb{E}_{P^n\delta_x}\left[\max_{0\leq i\leq l-1}\left(\frac{g(X_i)+g(X_{i+l})+\ldots+g(X_{i+m_il})}{\sqrt{m_i}}\right)^2 \chi_{[K,+\infty)}\left(\left|\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right|\right)\right] \\
&\leq l^2\mathbb{E}_{P^n\delta_x}\max_{0\leq i\leq l-1}\left[\left(\frac{g(X_i)+g(X_{i+l})+\ldots+g(X_{i+m_il})}{\sqrt{m_i}}\right)^2 \chi_{[K_0,+\infty)}\left(\left|\frac{g(X_i)+\ldots+g(X_{i+m_il})}{\sqrt{m_i}}\right|\right)\right] \\
&\leq l^2\mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_0)+g(X_l)+\ldots+g(X_{m_0l})}{\sqrt{m_1}}\right)^2 \chi_{[K_0,+\infty)}\left(\left|\frac{g(X_0)+g(X_l)+\ldots+g(X_{m_0l})}{\sqrt{m_0}}\right|\right)\right] \\
&+ l^2\mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_1)+g(X_{1+l})+\ldots+g(X_{1+m_1l})}{\sqrt{m_1}}\right)^2 \chi_{[K_0,+\infty)}\left(\left|\frac{g(X_1)+g(X_{1+l})+\ldots+g(X_{1+m_1l})}{\sqrt{m_1}}\right|\right)\right] \\
&\qquad + \cdots + l^2\mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_{l-1})+g(X_{2l-1})+\ldots+g(X_{l-1+m_{l-1}l})}{\sqrt{m_{l-1}}}\right)^2\right. \\
&\left.\times\, \chi_{[K_0,+\infty)}\left(\left|\frac{g(X_{l-1})+g(X_{2l-1})+\ldots+g(X_{l-1+m_{l-1}l})}{\sqrt{m_{l-1}}}\right|\right)\right] \leq l^2 l\varepsilon/l^3 = \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the proof is complete. $\square$

We are now in a position to show the following:

Proposition 5.4.2. *Let P be a Markov operator that satisfies* **(A1) - (A3)** *and let (X_n) be a Markov chain corresponding to P. Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \geq l_0$ be given. Assume that $Y_n = X_{nl}$ for $n \geq 0$. If $g : X \to \mathbb{R}$ is a bounded Lipschitz function, then*

$$\lim_{m\to\infty} \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_m)}{\sqrt{m}} \right)^2 = \tilde{\sigma}^2,$$

where

$$\tilde{\sigma}^2 = \lim_{m\to\infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \ldots + g(X_n)}{\sqrt{n}} \right)^2.$$

Proof. First observe that for any $k \geq 1$ we have

$$\left| \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2 - \mathbb{E}_x \left(\frac{g(Y_k) + g(Y_{k+1}) + \ldots + g(Y_{n+k})}{\sqrt{n}} \right)^2 \right| \to 0$$

as $n \to \infty$. Further we have

$$\mathbb{E}_x \left(\frac{g(Y_k) + g(Y_{k+1}) + \ldots + g(Y_{n+k})}{\sqrt{n}} \right)^2 = \mathbb{E}_{P^k\delta_x} \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2,$$

by the Markov property. On the other hand, the Markov chain (Y_n) corresponds to the operator P^l and the assumptions of Lemma 5.4.2 are satisfied with P and (X_n) replaced with P^l and (Y_n), respectively. Thus the functions

$$X \ni x \to \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2 \qquad \text{for } n \geq 0$$

are Lipschitzean with the Lipschitz constant bounded by, say $\tilde{M}$, independent of n. Consequently, we obtain

$$\begin{aligned}&\left| \mathbb{E}_{P^k\delta_x} \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2 - \mathbb{E}_{\mu_*} \left(\frac{g(Y_0) + g(Y_1) + \ldots + g(Y_n)}{\sqrt{n}} \right)^2 \right| \\ &\leq \tilde{M} \|P^k\delta_x - \mu_*\| \leq \tilde{M}Cq^k \|\delta_x - \mu_*\|.\end{aligned}$$

Hence we finally have the desired convergence. □

5.5 The Central Limit Theorem

Now we are in a position to formulate and prove the main theorem of this paper.

Theorem 5.5.1. *(Central Limit Theorem) Let P be a Markov operator that satisfies* **(A1) - (A3)** *and let (X_n) be a Markov chain corresponding to P. Assume that $X_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \to \mathbb{R}$ such that $\int_X g \mathrm{d}\mu_* = 0$, where μ_* is the unique invariant distribution for P, the sequence of random vectors (X_n) satisfies*

$$\lim_{n\to\infty} P\left(\frac{g(X_0) + g(X_1) + \ldots + g(X_n)}{\sqrt{n}} < a\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{a} e^{-\frac{y^2}{2\sigma^2}} \mathrm{d}y \qquad \textit{for all } a \in \mathbb{R},$$

if

$$\sigma^2 = \lim_{n\to\infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \ldots + g(X_n)}{\sqrt{n}}\right)^2 > 0.$$

Otherwise, if $\sigma = 0$, the sequence

$$\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} \qquad \textit{for } n \geq 1$$

converges in distribution to 0.

Proof. To prove the theorem we show that for any $x \in X$

$$\lim_{n\to\infty} \mathbb{E}_x \exp\left[it\left(\frac{g(X_0) + g(X_1) + \ldots + g(X_n)}{\sqrt{n}}\right)\right] = e^{-\sigma^2 t^2/2} \quad \text{for } t \in \mathbb{R}.$$

Without loss of generality we may assume that g is bounded by 1 and its Lipschitz constant is also bounded by 1. We will make use of the following formula:

$$e^{ia} = 1 + ia - a^2/2 - R(a)a^2,$$

where $|R(a)| \leq 1$ and $\lim_{a\to 0} R(a) = R(0) = 0$.

Fix $x \in X$, $t \in \mathbb{R} \setminus \{0\}$ and $\varepsilon > 0$. Let $k_0 \in \mathbb{N}$ and $\eta > 0$ be such that

$$|(1 - \sigma_0^2 t^2/2k)^k - e^{-\sigma^2 t^2/2}| < \varepsilon \quad \text{for } |\sigma_0 - \sigma| < \eta \text{ and } k \geq k_0.$$

Set $D := \sup_{n\geq 1} \int_X \int_X \rho(u, z) P^n \delta_x(\mathrm{d}u) \mu_*(\mathrm{d}z)$ and observe that for any x_0 we have according to Proposition 5.2.1

$$D \le \sup_{n\ge1} \int_X \rho(u,x_0) P^n\delta_x(\mathrm{d}u) + \int_X \rho(z,x_0)\mu_*(\mathrm{d}z) < +\infty.$$

For given x, t, ε, by Lemma 5.4.3, we choose $K > 0$ and $N_0 \in \mathbb{N}$ such that

$$\mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right)^2 \chi_{[K,+\infty)}\left(\left|\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right|\right)\right] \le \varepsilon/(2t^2) \tag{5.9}$$

for $n \ge N_0$ and $m \in \mathbb{N}$.

Fix $l \ge N_0$ such that

$$q^l CD|t|((1-q)^{-1} + M|t|) < \varepsilon \qquad \text{and} \qquad q^l C < 1, \tag{5.10}$$

where the constants q, C are given by condition (5.1) and the constant M is given by Lemma 5.4.2. Since

$$\sup_{m,n\ge1} \mathbb{E}_{P^n\delta_x}\left[\left(\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right)^2\right] < \infty$$

one may choose $k \ge k_0$ such that

$$\begin{aligned}\mathbb{E}_{P^n\delta_x}&\left(\left(\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right)^2 \left|R\left(\frac{\theta t}{\sqrt{k}}\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right)\right|\right.\\ &\times \left.\chi_{[0,K]}\left(\left|\frac{g(X_0)+\ldots+g(X_m)}{\sqrt{m}}\right|\right)\right) \le \varepsilon/(2t^2) \qquad \text{for all } m,n \ge 1 \text{ and } \theta \in [0,1],\end{aligned} \tag{5.11}$$

by the fact that $|R(a)| \to 0$ as $a \to 0$ and $|\sigma_m - \sigma| < \eta$ for $m \ge k$, where

$$\sigma_m^2 = \mathbb{E}_{\mu_*}\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m}}\right)^2 = \mathbb{E}_{\mu_*}\left(\frac{g(X_l)+\ldots+g(X_m)}{\sqrt{m}}\right)^2.$$

This can be done due to the fact that σ_m tends to σ as $m \to +\infty$. For positive integers u, v and w, $v \ge l$, where l is such that condition (5.10) holds, we set

$$I_{u,v,w} = \mathbb{E}_x \exp\left[it\left(\frac{g(X_l)+\ldots+g(X_w)+\ldots+g(X_{(u-1)w+l})+\ldots+g(X_{uw})}{\sqrt{wv}}\right)\right]$$

and

$$I_w = \mathbb{E}_x \exp\left[it\left(\frac{g(X_0)+\ldots+g(X_w)}{\sqrt{w}}\right)\right].$$

Note that $\lim_{n\to\infty}|I_n - I_{m,m,[n/m]}| = 0$ for every $m \geq 1$. So let us consider in further detail the terms $I_{j,k,m}$ for given k, $j \leq k$ and $m \geq k$. In the following we will use the notation

$$\mathbb{E}_{|X_{(j-1)m}}[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_{(j-1)m}].$$

We have

$$\begin{aligned}
&I_{j,k,m} = \mathbb{E}_x \exp\left[it\left(\frac{g(X_l)+\ldots+g(X_m)+\ldots+g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{km}}\right)\right]\\
&= \mathbb{E}_x\left(\exp it\left[\left(\frac{g(X_l)+\ldots+g(X_m)+\ldots+g(X_{(j-2)m+l})+\ldots+g(X_{(j-1)m})}{\sqrt{km}}\right)\right]\right.\\
&\left.\times\, \mathbb{E}_{|X_{(j-1)m}} \exp it\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{km}}\right)\right)\\
&= \mathbb{E}_x\left(\exp it\left[\left(\frac{g(X_l)+\ldots+g(X_m)+\ldots+g(X_{(j-2)m+l})+\ldots+g(X_{(j-1)m})}{\sqrt{km}}\right)\right]\right.\\
&\left(1+\frac{it}{\sqrt{km}}\mathbb{E}_{|X_{(j-1)m}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm})) - \frac{t^2}{2k}\mathbb{E}_{|X_{(j-1)m}}\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2\right.\\
&\left.\left.-\frac{t^2}{k}\mathbb{E}_{|X_{(j-1)m}}\left[\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2 R\left(\frac{t}{\sqrt{km}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right)\right]\right).
\end{aligned} \tag{5.12}$$

Since $\mathbb{E}_{\mu_*}g(X_i) = 0$, by **(A2)** we obtain

$$\begin{aligned}
|\mathbb{E}_u(g(X_l)+\ldots+g(X_m))| &= |\mathbb{E}_{P^l\delta_u}(g(X_0)+\ldots+g(X_{m-l})) - \mathbb{E}_{\mu_*}(g(X_0)+\ldots+g(X_{m-l}))|\\
&\leq \sum_{i=l}^{m}|\mathbb{E}_{P^i\delta_u}(g(X_0) - \mathbb{E}_{\mu_*}(g(X_0))| \leq \sum_{i=l}^{m}\|P^i\delta_x - \mu_*\|\\
&\leq q^l C(1-q)^{-1}\|\delta_u - \mu_*\| \leq q^l C(1-q)^{-1}\int_X \rho(u,z)\mu_*(\mathrm{d}z)
\end{aligned}$$

and hence we have

$$\begin{aligned}
&\mathbb{E}_x\left|\frac{it}{\sqrt{km}}\mathbb{E}_{|X_{(j-1)m}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right| \leq q^l C(1-q)^{-1}\frac{|t|}{\sqrt{km}}\mathbb{E}_x\left[\int_X \rho(X_{(j-1)m},z)\mu_*(\mathrm{d}z)\right]\\
&\leq q^l C(1-q)^{-1}\frac{|t|}{\sqrt{km}}\int_X\int_X \rho(u,z)\mu_*(\mathrm{d}z)P^{(j-1)m}\delta_x(\mathrm{d}u) \leq q^l CD(1-q)^{-1}\frac{|t|}{\sqrt{km}}.
\end{aligned}$$

On the other hand, by Lemma 5.4.2 and **(A2)** we have

$$\begin{aligned}
&\left|\mathbb{E}_u\left(\frac{g(X_l)+\ldots+g(X_m)}{\sqrt{m}}\right)^2-\mathbb{E}_{\mu_*}\left(\frac{g(X_l)+\ldots+g(X_m)}{\sqrt{m}}\right)^2\right| \\
&=\left|\mathbb{E}_{P^l\delta_u}\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m}}\right)^2-\mathbb{E}_{\mu_*}\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m}}\right)^2\right| \\
&\le q^l CM\frac{m-l}{m}\|\delta_u-\mu_*\| \le q^l CM\int_X \rho(u,z)\mu_*(\mathrm{d}z),
\end{aligned}$$

where M is given in Lemma 5.4.2 and consequently we obtain

$$\begin{aligned}
&\mathbb{E}_x\left|\frac{t^2}{k}\mathbb{E}_{|X_{(j-1)m}}\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2-\frac{t^2}{k}\mathbb{E}_{\mu_*}\left(\frac{g(X_l)+\ldots+g(X_m)}{\sqrt{m}}\right)^2\right| \\
&\le q^l M\frac{t^2}{k}\mathbb{E}_x\left[\int_X \rho(X_{(j-1)m},z)\mu_*(\mathrm{d}z)\right] \le q^l CM\frac{t^2}{k}\int_X\int_X \rho(u,z)\mu_*(\mathrm{d}z)P^{(j-1)m}\delta_x(\mathrm{d}u) \\
&\le q^l CMD\frac{t^2}{k}.
\end{aligned}$$

Finally conditions (5.9) and (5.11) will allow us to evaluate the term

$$\mathbb{E}_x\left|\frac{t^2}{k}\mathbb{E}_{|X_{(j-1)m}}\left[\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2 R\left(\frac{t}{\sqrt{km}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right)\right]\right|.$$

Indeed, we have

$$\begin{aligned}
&\mathbb{E}_x\left|\frac{t^2}{k}\mathbb{E}_{|X_{(j-1)m}}\left[\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2 R\left(\frac{t}{\sqrt{km}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right)\right]\right| \\
&\le \frac{t^2}{k}\mathbb{E}_x\left[\mathbb{E}_{|X_{(j-1)m}}\left|\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2 R\left(\frac{t}{\sqrt{km}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right)\right|\right] \\
&= \frac{t^2}{k}\mathbb{E}_x\left[\left|\left(\frac{g(X_{(j-1)m+l})+\ldots+g(X_{jm})}{\sqrt{m}}\right)^2 R\left(\frac{t}{\sqrt{km}}(g(X_{(j-1)m+l})+\ldots+g(X_{jm}))\right)\right|\right]
\end{aligned}$$

$$= \frac{t^2}{k}\mathbb{E}_{P^{(j-1)m+l}\delta_x}\left[\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m}}\right)^2 \left| R\left(\frac{t}{\sqrt{km}}(g(X_0)+\ldots+g(X_{m-l}))\right)\right|\right]$$

$$\le \frac{t^2}{k}\mathbb{E}_{P^{(j-1)m+l}\delta_x}\left[\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m-l}}\right)^2 \chi_{[K,+\infty)}\left(\left|\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m-l}}\right|\right)\right]$$

$$+ \frac{t^2}{k}\mathbb{E}_{P^{(j-1)m+l}\delta_x}\left(\left(\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m-l}}\right)^2 \left|R\left(\sqrt{\frac{m-l}{m}}\frac{t}{\sqrt{k}}\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m-l}}\right)\right|\right.$$

$$\left.\times\, \chi_{[0,K]}\left(\left|\frac{g(X_0)+\ldots+g(X_{m-l})}{\sqrt{m-l}}\right|\right)\right) \overset{\text{by (5.9) and (5.11)}}{\le} \frac{t^2}{k}\varepsilon/(2t^2) + \frac{t^2}{k}\varepsilon/(2t^2) = \varepsilon/k.$$

Now from (5.12) it follows that

$$|I_{j,k,m} - I_{j-1,k,m}(1-\sigma_m^2 t^2/2k)| \le q^l CD(1-q)^{-1}\frac{|t|}{\sqrt{km}} + q^l CMD\frac{t^2}{k} + \frac{\varepsilon}{k} \le 2\varepsilon/k \quad \text{for } j=1,\ldots,k,$$

by condition (5.10) and the fact that $m \ge k$. Iterating this formula k-times we obtain

$$|I_{k,k,m} - (1-\sigma_m^2 t^2/2k)^k| \le 2\varepsilon$$

and consequently

$$|I_{k,k,m} - e^{-\sigma^2 t^2/2}| \le 3\varepsilon \quad \text{for all } m \text{ sufficiently large.}$$

Since $\varepsilon > 0$ was arbitrary, we obtain that $\lim_{n\to\infty} |I_{k,k,[n/k]} - I_n| = 0$, which completes the proof. □

5.6 Example

Let (X,ρ) be a Polish space and let $(T,\mathcal{A})$ be a measurable space. Let $\nu : \mathcal{A} \to [0,\infty)$ be some measure on T. Let $p : T \times X \to [0,\infty)$ be a measurable function such that

$$\int_T p(t,x)\nu(\mathrm{d}t) = 1 \qquad \text{for } x \in X.$$

We shall assume that $p(t,\cdot) : X \to [0,\infty)$ for $t \in T$ is a Lipschitzean function. Denote its Lipschitz constant by $k(t)$.

Let $\pi_t : X \times \mathcal{B}(X) \to [0,1]$, $t \in T$, be a transition probability and let P_t and U_t denote the corresponding Markov operator and its dual, respectively. We shall consider the Markov chain (X_n) on some space $(\Omega,\mathcal{F},\mathbb{P})$ corresponding to the action of randomly chosen transition probabilities with probability depending on position, i.e. the Markov chain (X_n) given by the formula:

$$\mathbb{P}[X_{n+1} \in \cdot | X_n = x] = \pi(x, \cdot) = \int_T p(t,x)\pi_t(x,\cdot)\nu(\mathrm{d}t) \quad \text{for } x \in X \text{ and } n \geq 1. \tag{5.13}$$

This model generalizes some random dynamical systems studied in [HSl16].

If the distribution of X_0 is equal to μ, then the distribution of X_n is given by $P^n\mu$, where $P : \mathcal{M} \to \mathcal{M}$ is of the form

$$P\mu(A) = \int_T \int_X p(t,x)\pi_t(x,A)\mu(\mathrm{d}x)\nu(\mathrm{d}t) \qquad \text{for } A \in \mathcal{B}(X). \tag{5.14}$$

Theorem 5.6.1. *Assume that π_t for $t \in T$ are transition probabilities and there exists $q < 1$ such that for the Markov operators P_t corresponding to π_t we have*

$$\|P_t\mu_1 - P_t\mu_2\| \leq q\|\mu_1 - \mu_2\| \qquad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1^1.$$

Moreover, for some $x_0 \in X$ and the operator U_t dual to P_t, $t \in T$, we have

$$U_t[\rho(\cdot, x_0)]^2(x) \leq a[\rho(x,x_0)]^2 + b$$

for some $a < 1$ and $b > 0$. Finally set

$$\gamma(t) \coloneqq \sup_{x \in X} \int_X \rho(z,x_0)\pi_t(x,\mathrm{d}z)$$

for some $x_0 \in X$ and assume that

$$\int_T \gamma(t)k(t)\nu(\mathrm{d}t) < 1 - q$$

and

$$\int_T p(t,x)\gamma(t)\nu(\mathrm{d}t) < \infty \qquad \text{for } x \in X.$$

Let (X_n) be a Markov chain given by (5.13). Assume that $X_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \to \mathbb{R}$ such that $\int_X g\mathrm{d}\mu_* = 0$, where μ_* is the unique invariant distribution for P, the sequence of random vectors (X_n) satisfies

$$\lim_{n\to\infty} P\left(\frac{g(X_0) + g(X_1) + \ldots + g(X_n)}{\sqrt{n}} < v\right) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{v} e^{-\frac{y^2}{2\sigma^2}}\mathrm{d}y \quad \text{for } v \in \mathbb{R},$$

if

$$\sigma^2 = \lim_{n\to\infty} \mathbb{E}_{\mu_*}\left(\frac{g(X_0)+g(X_1)+\ldots+g(X_n)}{\sqrt{n}}\right)^2 > 0.$$

Otherwise, if $\sigma = 0$, *the sequence*

$$\frac{g(X_0)+g(X_1)+\cdots+g(X_n)}{\sqrt{n}} \qquad \text{for } n \geq 1$$

converges in distribution to 0.

Proof. From Theorem 5.5.1 it follows that to finish the proof it is enough to show that the Markov operator P given by (5.14) satisfies conditions **(A1)** - **(A3)**.

Observe that P is a Feller operator and its dual is of the form

$$Uf(x) = \int_T \int_X p(t,x) f(z) \pi_t(x, \mathrm{d}z)\nu(\mathrm{d}t) \qquad \text{for } f \in C_b(X). \tag{5.15}$$

The operator U may be extended to an arbitrary positive unbounded function f and then it is also given by formula (5.15). The operator U can be, in fact, extended to the space of all Lipschitz functions. To show this, observe that for $x_0 \in X$ we have

$$\begin{aligned} U\rho(\cdot, x_0)(x) &= \int_T p(t,x) \int_X \rho(z,x_0)\pi_t(x,\mathrm{d}z)\nu(\mathrm{d}t) \leq \int_T p(t,x)\gamma(t)\nu(\mathrm{d}t) \\ &\leq \int_T |p(t,x) - p(t,x_0)|\gamma(t)\nu(\mathrm{d}t) + \int_T p(t,x_0)\gamma(t)\nu(\mathrm{d}t) \\ &\leq \int_T k(t)\gamma(t)\nu(\mathrm{d}t) + \int_T p(t,x_0)\gamma(t)\nu(\mathrm{d}t) < \infty. \end{aligned}$$

Therefore for any Lipschitz function $f : X \to \mathbb{R}$ with Lipschitz constant L and $x \in X$ we have

$$|Uf(x)| \leq U|f|(x) \leq |f(x_0)| + LU\rho(\cdot, x_0)(x) < \infty,$$

by the fact that the Lipschitz constant of the function $|f|$ is bounded by L.

Set $\hat{q} := \int_T \gamma(t)k(t)\nu(\mathrm{d}t) + q < 1$ and let $f : X \to \mathbb{R}$ be an arbitrary Lipschitz function with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$. We show that then Uf is a

Lipschitz function with Lipschitz constant bounded by $\hat{q}$. Indeed, for any $x, y \in X$ we have

$$\begin{aligned}
|Uf(x) - Uf(y)| &= \left| \int_T \int_X p(t,x) f(z) \pi_t(x, \mathrm{d}z) \nu(\mathrm{d}t) - \int_T \int_X p(t,y) f(z) \pi_t(y, \mathrm{d}z) \nu(\mathrm{d}t) \right| \\
&\leq \int_T \int_X |p(t,x) - p(t,y)| |f(z)| \pi_t(x, \mathrm{d}z) \nu(\mathrm{d}t) \\
&+ \int_T p(t,y) \left| \int_X f(z) \pi_t(x, \mathrm{d}z) - \int_X f(z) \pi_t(y, \mathrm{d}z) \right| \nu(\mathrm{d}t) \\
&\leq \int_T k(t) \left(\int_X \rho(z, x_0) \pi_t(x, \mathrm{d}z) \right) \nu(\mathrm{d}t) \cdot \rho(x,y) + q \int_T p(t,y) \nu(\mathrm{d}t) \cdot \rho(x,y) \\
&\leq \left(\int_T k(t) \gamma(t) \nu(\mathrm{d}t) + q \right) \rho(x,y) = \hat{q} \rho(x,y).
\end{aligned}$$

To verify **(A1)** we compute

$$\begin{aligned}
\int_X \rho(x, x_0) P\mu(\mathrm{d}x) &= \int_X U\rho(\cdot, x_0)(x) \mu(\mathrm{d}x) \\
&= \int_X (U\rho(\cdot, x_0)(x) - U\rho(\cdot, x_0)(x_0)) \mu(\mathrm{d}x) + U\rho(\cdot, x_0)(x_0) \\
&\leq \hat{q} \int_X \rho(x, x_0) \mu(\mathrm{d}x) + U\rho(\cdot, x_0)(x_0) \leq \int_X \rho(x, x_0) \mu(\mathrm{d}x) + U\rho(\cdot, x_0)(x_0) < \infty
\end{aligned}$$

for $\mu \in \mathcal{M}_1^1$.

Observe that for any $\mu_1, \mu_2 \in \mathcal{M}_1^1$ we have

$$\|\mu_1 - \mu_2\| = \sup \left| \int_X f(x) \mu_1(\mathrm{d}x) - \int_X f(x) \mu_2(\mathrm{d}x) \right|,$$

where the supremum is taken over all Lipschitz functions $f : X \to \mathbb{R}$ with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$. Indeed for any $f : X \to \mathbb{R}$ with Lipschitz constant bounded by 1, we have

$$\left| \int_X f(x) \mu_1(\mathrm{d}x) - \int_X f(x) \mu_2(\mathrm{d}x) \right| = \left| \int_X (f(x) - f(x_0)) \mu_1(\mathrm{d}x) - \int_X (f(x) - f(x_0) \mu_2(\mathrm{d}x) \right|.$$

To show that **(A2)** holds fix $\mu_1, \mu_2 \in \mathcal{M}_1^1$. Then we have

$$\begin{aligned}
&\|P\mu_1 - P\mu_2\| = \sup \left| \int_X f(x) P\mu_1(\mathrm{d}x) - \int_X f(x) P\mu_2(\mathrm{d}x) \right| \\
&= \hat{q} \sup \left| \int_X \hat{q}^{-1} Uf(x)) \mu_1(\mathrm{d}x) - \int_X \hat{q}^{-1} Uf(x) \mu_2(\mathrm{d}x) \right| \leq \hat{q} \sup \left| \int_X f(x) \mu_1(\mathrm{d}x) - \int_X f(x) \mu_2(\mathrm{d}x) \right| \\
&= \hat{q} \|\mu_1 - \mu_2\|
\end{aligned}$$

Here the supremum is taken over all Lipschitz functions $f : X \to \mathbb{R}$ with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$, by the fact that $\hat{q}^{-1}Uf$ is a Lipschitz function with Lipschitz constant bounded by 1. Hence **(A2)** holds.

Finally, we show that **(A3)** holds with $\mu = \delta_{x_0}$. To do this it is enough to prove that

$$U[\rho(\cdot, x_0)(x)]^2 \le a[\rho(x, x_0)]^2 + b \tag{5.16}$$

for some $a < 1$ and $b > 0$. Indeed, iterating the above formula we have

$$U^n[\rho(\cdot, x_0)]^2(x) \le a^n[\rho(x, x_0)]^2 + b(1-a)^{-1} \qquad \text{for } n \ge 1 \text{ and } x \in X,$$

consequently,

$$\begin{aligned}
&\sup_{n\ge1} \int_X [\rho(x, x_0)]^2 P^n \delta_{x_0}(\mathrm{d}x) \\
&\le \sup_{n\ge1} \left(a^n \int_X [\rho(x, x_0)]^2 \delta_{x_0}(\mathrm{d}x) + b(1-a)^{-1} \right) = b(1-a)^{-1} < \infty.
\end{aligned}$$

and **(A3)** will hold. To verify condition (5.16) we compute

$$\begin{aligned}
U[\rho(\cdot, x_0)(x)]^2 &= \int_T p(t, x_0) \int_X [\rho(z, x_0)]^2 \pi_t(x_0, \mathrm{d}z)\nu(\mathrm{d}t) = \int_T p(t, x_0) U_t[\rho(\cdot, x_0)]^2(x)\nu(\mathrm{d}t) \\
&\le \int_T p(t, x_0)(a[\rho(x, x_0]^2 + b)\nu(\mathrm{d}t) = a[\rho(x, x_0]^2 + b.
\end{aligned}$$

This completes the proof. □

Index

Bibliography

[AHvG13] T. Alkurdi, S.C. Hille, and O. van Gaans. Ergodicity and stability of a dynamical system perturbed by impulsive random interventions. Journal of Mathematical Analysis and Applications, 407(2):480 – 494, 2013.

[AI05] A.S. Ackleh and K. Ito. Measure-valued solutions for a hierarchically size-structured population. Journal of Differential Equations, 217(2):431 – 455, 2005.

[AJK06] F. Albiac and N.J. Kalton. Topics in Banach Space Theory, volume 233, 2006.

[Ale43] A. Alexandrov. Additive set-functions in abstract spaces. Matematiceskij sbornik, 55(2-3):169–238, 1943.

[AHG12] T. Alkurdi, S.C. Hille, O. van Gaans. On metrization of unions of function spaces on different intervals. Journal of the Australian Mathematical Society, 92(3):281–297, 2012.

[Bar88] M.F. Barnsley. Fractals Everywhere. Academic Press Professional, Inc., San Diego, CA, USA, 1988.

[Bar12] M.F. Barnsley. Fractals Everywhere. Dover Books on Mathematics. Dover Publications, 2012.

[BCF13] A. Bátkai, P. Csomós, and B. Farkas. Operator splitting for nonautonomous delay equations. Computers and Mathematics with Applications, 65(3):315 – 324, 2013. Efficient Numerical Methods for Scientific Applications.

[BCF17] A. Bátkai, P. Csomós, and B. Farkas. Operator splitting for dissipative delay equations. Semigroup Forum, 95(2):345–365, Oct 2017.

[BCFN12] A. Bátkai, P. Csomós, B. Farkas, and G. Nickel. Operator splitting with spatial-temporal discretization. Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equations , 161–171, 2012.

[BD85] M. F. Barnsley and S. Demko. Iterated function systems and the global construction of fractals. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 399(1817):243–275, 1985.

[BDEG88] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. Annales de l'I.H.P. Probabilités et statistiques, 24(3):367–394, 1988.

[BEH89] M.F. Barnsley, J.H. Elton, and D.P. Hardin. Recurrent iterated function systems. Constructive Approximation, 5(1):3–31, Dec 1989.

[Ber16] B. Berckmoes. On the hausdorff measure of noncompactness for the parameterized Prokhorov metric. Journal of Inequalities and Applications, 2016(1):215, Sep 2016.

[Bha82] R. Bhattacharya. On the functional central limit theorem and the law of the iterated logrithm for Markov processes. Probability Theory and Related Fields, 60:185–201, 06 1982.

[BI95] A.N. Borodin and I.A. Ibragimov. Limit theorems for functionals of random walks. In Proc. Steklov Inst, Math. 195, 1994; English transl., 1995. Amer. Math. Soc.

[Bil99] P. Billingsley. Convergence of probability measures. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[Bog07a] V.I. Bogachev. Measure Theory. Number T. 2 in Measure Theory. Springer, 2007.

[Bog07b] V.I. Bogachev. Measure Theory. Number T. 1 in Measure Theory. Springer, 2007.

[CC04] R.M. Colombo and A. Corli. A semilinear structure on semigroups in a metric space. Semigroup Forum, 68(3):419–444, 2004.

[CCC13] J.A. Canizo, J.A. Carrillo, and S. Cuadrado. Measure solutions for some models in population dynamics. Acta Appl. Math, 123:141–156, 2013.

[CCGU12] J.A. Carrillo, R.M. Colombo, P. Gwiazda, and A. Ulikowska. Structured populations, cell growth and measure valued balance laws. Journal of Differential Equations, 252:3245–3277, 2012.

[CD08] O.L.V. Costa and F. Dufour. Stability and ergodicity of piecewise deterministic Markov processes. In 2008 47th IEEE Conference on Decision and Control, pages 1525–1530, Dec 2008.

[CD09] O.L.V. Costa and F. Dufour. The policy iteration algorithm for average continuous control of piecewise deterministic Markov processes. In Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, pages 506–511, Dec 2009.

[CD10] O.L.V. Costa and F. Dufour. Singular perturbation for the discounted continuous control of piecewise deterministic Markov processes. In 49th IEEE Conference on Decision and Control (CDC), pages 1436–1441, Dec 2010.

[CG12] C. Canzi and G. Guerra. A simple counterexample related to the Lie–Trotter product formula. Semigroup Forum, 84(3):499–504, Jun 2012.

[CGU14] J.A. Carrillo, P. Gwiazda, and A. Ulikowska. Splitting-particle methods for structured population models: Convergence and applications. Mathematical Models and Methods in Applied Sciences, 24(11):2171–2197, 2014.

[CH14] D. Czapla and K. Horbacz. Equicontinuity and stability properties of Markov chains arising from iterated function systems on Polish spaces. Stoch. Anal. Appl, 32:1–29, 2014.

[Che74] P.R. Chernoff. Product formulas, nonlinear semigroups, and addition of unbounded operators. Memoirs of the American Mathematical Society. American Mathematical Society, 1974.

[Col09] G. Colombo, R.M. Guerra. Differential equations in metric spaces with applications. Discrete and Continuous Dynamical Systems, 23(3):733–753, 2009.

[Con85] J.B. Conway. A Course in Functional Analysis. New York, Springer Verlag, 1985.

[Cox12] S. Cox. Stochastic differential equations in Banach spaces: Decoupling, delay equations, and approximations in space and time. PhD Thesis, Delft Universit, 12 March 2012.

[CFH05] P. Csomós and I. Faragó and Á. Havasi. Weighted sequential splittings and their analysis. Computers and Mathematics with Applications, 50(7):1017 – 1031, 2005. Numerical Methods and Computational Mechanics.

[CvN10] S. Cox and J. van Neerven. Convergence rates of the splitting scheme for parabolic linear stochastic Cauchy problems. SIAM Journal on Numerical Analysis, 48(2):428–451, 2010.

[Cza12] D. Czapla. A criterion of asymptotic stability for Markov-Feller e-chains on Polish spaces. Ann. Polon. Math, 105:267–291, 2012.

[Dar55] G. Darbo. Punti uniti in trasformazioni a codominio non compatto. Rendiconti del Seminario Matematico della Università di Padova, 24:84–92, 1955.

[Dav84] M.H.A. Davis. Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. J. Roy. Statist. Soc. Ser. B, number 3, vol.46:353–388, 1984.

[DF99] P. Diaconis and D. Freedman. Iterated random functions. Siam Review - SIAM REV, 41:45–76, 03 1999.

[DGMT98] O. Diekmann, M. Gyllenberg, J.A.J. Metz, and H. Thieme. On the formulation and analysis of general deterministic structured population models: I. linear theory. Journal of Mathematical Biology, 36(4):349–388, 3 1998.

[DHZ01] I. Dimov, Á. Havasi, and Z. Zlatev. L-commutativity of the operators in splitting methods for air pollution models. Annales Universitatis

Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Mathematica, 44, 01 2001.

[Die48] J. Dieudonne. On topological groups of homeomorphisms. American Journal of Mathematics, (3), 1948.

[DM02] J. Dedecker and F. Merlevède. Necessary and sufficient conditions for the conditional central limit theorem. Ann. Probab., 30(3):1044–1081, 07 2002.

[Doe38] W. Doeblin. Sur deux problèmes de m. Kolmogoroff concernant les chaînes dénombrables. Bulletin de la Société Mathématique de France, 66:210–220, 1938.

[DPZ14] G. Da Prato and J. Zabczyk. Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2014.

[DR09] Y. Davydov and V. Rotar. On asymptotic proximity of distributions. J. Theor. Probab, 22:82–99, 2009.

[Dud66] R.M. Dudley. Convergence of Baire measures. Studia Math, 27:251–268, 1966.

[Dud02] R.M. Dudley. Real Analysis and Probability. Cambridge University Press, 2002.

[DX11] Z. Dong and Y. Xie. Ergodicity of stochastic 2D Navier–Stokes equation with Lévy noise. Journal of Differential Equations, 251(1):196 – 222, 2011.

[Eag75] G.K. Eagleson. On Gordin's central limit theorem for stationary processes. Journal of Applied Probability, 12(1):176–179, 1975.

[EHM15] J.H.M. Evers, S.C. Hille, and A. Muntean. Mild solutions to a measure-valued mass evolution problem with flux boundary conditions. J. Differential Equations, 259(3):1068–1097, 2015.

[EHM16] J. Evers, S.C. Hille, and A. Muntean. Measure-valued mass evolution problems with flux boundary conditions and solution-dependent velocities. SIAM J. Math. Anal, 48:1929–1953, 2016.

[EK86] S.N. Ethier and T.G. Kurtz. Markov processes : characterization and convergence. Wiley series in probability and mathematical statistics. J. Wiley & Sons, New York, Chichester, 1986.

[Eme07] E.Y. Emel'yanov. Non-spectral Asymptotic Analysis of One-Parameter Operator Semigroups. Operator Theory: Advances and Applications. Birkhäuser Basel, 2007.

[EBNHM13] K.J. Engel, S. Brendle, R. Nagel, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi. One-Parameter Semigroups for Linear Evolution Equations, Graduate Texts in Mathematics, Springer New York, 2013

[Eng77] R. Engelking. General topology. Monografie matematyczne. PWN, 1977.

[ESvR12] A. Es-Sarhir and M.K. von Renesse. Ergodicity of stochastic curve shortening flow in the plane. SIAM J. Math. Anal, 44:224–244, 2012.

[FH07] I. Farago and A. Havasi. Consistency analysis of operator splitting methods for C_0-semigroups expression. Semigroup Forum, 74(1):125–139, Feb 2007.

[FM53] R. Fortet and E. Mourier. Convergence de la répartition empirique vers la répartition théorique. Annales scientifiques de l'École Normale Supérieure, 3e serie, 70(3):267–285, 1953.

[GH04] M.I. Gordin and H. Holzmann. The central limit theorem for stationary Markov chains under invariant splittings. Stochastics and Dynamics, 04(01):15–30, 2004.

[GL78] M.I. Gordin and B.A. Lifšic. The central limit theorem for stationary Markov processes. Soviet Math, 19:392–394, 1978.

[GL81] E.E. Granirer and M. Leinert. On some topologies which coincide on the unit sphere of the fourier-stieltjes algebra $b(g)$ and of the measure algebra $m(g)$. Rocky Mountain J. Math, 11:459–472, 1981.

[GL15] F.Z. Gong and Y. Liu. Ergodicity and asymptotic stability of feller semigroups on Polish metric spaces. Sci. China Math, 58, 2015.

[GLMC10] P. Gwiazda, Th. Lorenz, and A. Marciniak-Czochra. A nonlinear structured population model: Lipschitz continuity of measure-valued solutions with respect to model ingredients. Journal of Differential Equations, 248(11):2703 – 2735, 2010.

[Gol85] J.A. Goldstein. Semigroups of linear operators and applications:. Oxford mathematical monographs. Oxford University Press, 1985.

[GRTW11] M. Georg Riedler, M. Thieullen, and G. Wainrib. Limit theorems for infinite-dimensional piecewise deterministic Markov processes. applications to stochastic excitable membrane models. Electronic Journal of Probability, 17, 12 2011.

[HADD84] M.H.A. Davis and M. Dempster. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models [and discussion]. Journal of the Royal Statistical Society. Series B: Methodological, 46:353–388, 05 1984.

[Hal88] J.K. Hale. Asymptotic Behavior of Dissipative Systems. Mathematical surveys and monographs. American Mathematical Society, 1988.

[HCWS17] K. Horbacz, D. Czapla, and H. Wojewodka-Sciazko. Ergodic properties of some piecewise-deterministic Markov process with application to a gene expression model. 07 2017.

[Hei86] H.J.A.M. Heijmans. Markov semigroups and structured population dynamics. In R. Nagel, U. Schloterbeck, and M.P.H. Wolff, editors, Aspects of Positivity in Functional Analysis, volume 122 of North-Holland Mathematics Studies, pages 199 – 208. North-Holland, 1986.

[HHS16] S. Hille, K. Horbacz, and T. Szarek. Existence of a unique invariant measure for a class of equicontinuous Markov operators with application to a stochastic model for an autoregulated gene. Annales mathématiques Blaise Pascal, 23:171–217, 01 2016.

[HHSWS15] S. Hille, K. Horbacz, T. Szarek, and H. Wojewodka-Sciazko. Limit theorems for some Markov operators. 06 2015.

[Hil14] S.C. Hille. Dynamical systems in spaces of measures. Lecture Notes of PhD course given at Università degli Studi di Milano-Bicocca, Milano, 20, January 2014.

[HKLR10] H. Holden, K. Karlsen, K.A. Lie, and H. Risebro. Splitting methods for partial differential equations with rough solutions. Analysis and Matlab programs. 04 2010.

[HM06] M. Hairer and J.C. Mattingly. Ergodicity of the 2d Navier-Stokes equations with degenerate stochastic forcing. Ann. Math, 164:993–1032, 2006.

[HM08] M. Hairer and J.C. Mattingly. Spectral gaps in Wasserstein distances and the 2d stochastic Navier–Stokes equations. Ann. Probab., 36(6):2050–2091, 11 2008.

[Hol05] H. Holzmann. Martingale approximations for continuous-time and discrete-time stationary Markov processes. Stochastic Processes and their Applications, 115(9):1518 – 1529, 2005.

[Hor06] K. Horbacz. Asymptotic stability of a semigroup generated by randomly connected poisson driven differential equations. Bollettino della Unione Matematica Italiana. Serie VIII. Sezione B. Articoli di Ricerca Matematica, 3, 10 2006.

[HP57] E. Hille and R.S. Phillips. Functional Analysis and Semi-groups. Number t. 31 in Colloquium publications. American Mathematical Society, 1957.

[HP18] S. Hadfield and A. Papageorgiou. Divide and conquer approach to quantum Hamiltonian simulation. New Journal of Physics, 20(4):043003, apr 2018.

[HS16] K. Horbacz and M. Sleczka. Law of large numbers for random dynamical systems. Journal of Statistical Physics, 162(3):671–684, Feb 2016.

[HSl16] K. Horbacz and M. Ślęczka. Law of large numbers for random dynamical systems, 162:671–684, 2016.

[HSWZ17] S.C. Hille, T. Szarek, D.T.H. Worm, and M.A. Ziemlańska. On a Schur-like property for spaces of measures. arXiv, 2017.

[HSZ17] S.C. Hille, T.Szarek, and M.A. Ziemlańska. Equicontinuous families of Markov operators in view of asymptotic stability. Comptes Rendus Mathematique, 355(12):1247 – 1251, 2017.

[HUT81] J.E. Hutchinson. Fractals and self similarity. Indiana University Mathematics Journal, 30(5):713–747, 1981.

[HW09a] S.C. Hille and D.T.H. Worm. Continuity properties of Markov semigroups and their restrictions to invariant l^1-spaces. Semigroup Forum, 79:575–600, 2009.

[HW09b] S.C. Hille and D.T.H. Worm. Embedding of semigroups of Lipschitz maps into positive linear semigroups on ordered Banach spaces. Int. Equ. Oper. Theory, 63:351–371, 2009.

[Jam64] B. Jamison. Asymptotic behaviour of succesive iterates of continuous functions under a Markov operator. J. Math. Anal. Appl, 9:203–214, 1964.

[Kal04] N.J. Kalton. Spaces of Lipschitz and Hölder functions and their applications. Collect. Math, 55:171–217, 2004.

[Kel55] J.L. Kelley. General topology. Van Nostrand, 1955.

[KLO12] T. Komorowski, C. Landim, and S. Olla. Fluctuations in Markov processes. Time symmetry and martingale approximation. Springer-Verlag, Heidelberg, 2012.

[KP84] T. Kurtz and M. Pierre. A counterexample for the trotter product formula. Journal of Differential Equations - J DIFFERENTIAL EQUATIONS, 52:407–414, 05 1984.

[KP80] T.G. Kurtz, M. Pierre, A Counterexample for the Trotter Product Formula. MRC technical summary report. Defense Technical Information Center, 1980.

[KPS10] T. Komorowski, S. Peszat, and T. Szarek. On ergodicity of some Markov operators. Ann. Prob, 38:1401–1443, 2010.

[Kuh01] F. Kuhnemund. Bi-continuous semigroups on spaces with two topologies: Theory and applications. 2001.

[Kuk02] S.B. Kuksin. Ergodic theorems for 2D statistical hydrodynamics. Reviews in Mathematical Physics, 14:1–16, 2002.

[KW01] F. Kühnemund and M. Wacker. Commutator conditions implying the convergence of the Lie-Trotter products. Proceedings of the American Mathematical Society, 129(12):3569–3582, 2001.

[KW12] T. Komorowski and A. Walczuk. Central limit theorem for Markov processes with spectral gap in the Wasserstein metric. Stochastic Processes and Appl., 122:2155–2184, 2012.

[LE70] S. Lie, F. Engel. Theorie der Transformationsgruppen. AMS Chelsea Publishing Series, v.1, 1970.

[LeC57] L. LeCam. Convergence in distribution of stochastic processes. Univ. California Publ. Statistics, 2:207–236, 1957.

[LM94] A. Lasota and M.C. Mackey. Chaos, fractals, and noise : stochastic aspects of dynamics. Applied mathematical sciences. Springer-Verlag, New York, 1994. Rev. ed. of: Probabilistic properties of deterministic systems. 1985.

[LM99] A. Lasota and M. Mackey. Cell division and the stability of cellular population. Journal of Mathematical Biology, 38:241–261, 01 1999.

[LM00] A. Lasota and J. Myjak. Attractors of multifunctions. Bulletin of the Polish Academy of Sciences, Mathematics, 48, 01 2000.

[Lot86] H.P. Lotz. Positive linear operators on lp and the Doeblin condition. Aspects of Positivity in Functional Analysis, volume 122 of North-Holland Mathematics Studies, pages 137 – 156. North-Holland, 1986.

[LS06] A. Lasota and T. Szarek. Lower bound technique in the theory of a stochastic differential equations. J. Diff. Equ, 231:513–533, 2006.

[LTY15] J. Li, S. Tu, and X. Ye. Mean equicontinuity and mean sensitivity. Ergod. Th. & Dynam. Sys, 35:2587–2612, 2015.

[LvG09] B. Lemmens and O. van Gaans. Dynamics of non-expansive maps on strictly convex banach spaces. Israel Journal of Mathematics, 171(1):425–442, Jun 2009.

[LY94] A. Lasota and J. A. Yorke. Lower bound technique for Markov operators and iterated function systems. Random & Comp. Dynamics, 2:41–77, 1994.

[MBA98] R.E. Megginson, D.W. Brooks, S.J. Axler, F.W. Gehring, and P.R. Halmos. An Introduction to Banach Space Theory. Graduate Texts in Mathematics. Springer New York, 1998.

[McS34] E. J. McShane. Extension of range of functions. Bull. Amer. Math. Soc, 40 (12):837–842, 1934.

[MPN11] J. Mallet-Paret and R.D. Nussbaum. Inequivalent measures of noncompactness. Annali di Matematica Pura ed Applicata, 190(3):453–488, Sep 2011.

[MS03] J. Myjak and T. Szarek. Attractors of iterated function systems and Markov operators. Abstr. Appl. Anal., 2003(8):479–502, 04 2003.

[MT09] S. Meyn and R.L. Tweedie. Markov Chains and Stochastic Stability. Cambridge University Press, New York, NY, USA, 2nd edition, 2009.

[MTKY13] M. Mackey, M. Tyran-Kamińska, and R. Yvinec. Dynamic behavior of stochastic gene expression models in the presence of bursting. SIAM Journal on Applied Mathematics, 73(5):1830–1852, 2013.

[MW00] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. Ann. Probab., 28:713–724, 2000.

[MZ05] P. Magal and X.-Q. Zhao. Global attractors and steady states for uniformly persistent dynamical systems. SIAM J. Math. Analysis, 37(1):251–275, 2005.

[NJKAKK06] F.A. Nigel, J. Kalton, F. Albiac, N.J. Kalton, and N.J. Kalton. Topics in Banach Space Theory. Number t. 10 in Graduate Texts in Mathematics. Springer, 2006.

[Pac79] J.K. Pachl. Measures as functionals on uniformly continuous functions. Pacific J. Math, 82:515–521, 1979.

[Pac13] J.K. Pachl. Uniform Spaces and Measures, volume 30. New York, Fields Institute Monographs, Springer, 2013.

[Pel15] M. Peligrad. Quenched invariance principle via martingale approximation. asymptotic laws and methods in stochastics. Fields Inst, 76:149–165, 2015.

[PT11] B. Piccoli and A. Tosin. Time-evolving measures and macroscopic modeling of pedestrian flows. Arch. Rational Mech. Anal, 199:707–738, 2011.

[RJ66] R.W. Bargley and J.S. Yang. On k-spaces and function spaces. Proc. AMS, (17(3)), 1966.

[RPTK02] R. Rudnicki, K. Pichor, and M. Tyran-Kamińska. Markov semigroups and their applications. Lecture Notes in Physics, 597:215–238, 11 2002.

[RTT16] V. Renault, M. Thieullen, and E. Trélat. Optimal control of infinite-dimensional piecewise deterministic Markov processes and application to the control of neuronal dynamics via optogenetics. Networks and Heterogeneous Media, 12, 07 2016.

[Rud97] R. Rudnicki. Asymptotic stability for Markov operators: A counter-example. Bulletin of the Polish Academy of Sciences Mathematics, 45:1–5, 01 1997.

[Rud00] R. Rudnicki. Markov operators: applications to diffusion processes and population dynamics. Applicationes Mathematicae, 27:67–79, 01 2000.

[SM03] T. Szarek and J. Myjak. Attractors of iterated function systems and Markov operators. Abstract and Applied Analysis, 2003(8):479–502, 2003.

[Ste94] W.J. Runggaldier, L. Stettner. Approximations of discrete time partially observed control problems Applied Mathematics Monographs CNR, 1994

[Str63] G. Strang. Accurate partial difference methods I: Linear cauchy problems. Archive for Rational Mechanics and Analysis, 12(1):392–402, Jan 1963.

[Str68] G. Strang. On the construction and comparison of difference schemes. SIAM Journal on Numerical Analysis, 5(3):506–517, 1968.

[SW12] T. Szarek and D. T. H. Worm. Ergodic measures of Markov semigroups with the e-property. Ergod. Th. & Dynam. Sys, 32:1117–1135, 2012.

[SY02] G.R. Sell and Y. You. Dynamics of Evolutionary Equations. Springer, 2002.

[Sza97] T. Szarek. Markov operators acting on polish spaces. Annales Polonici Mathematici, 1997.

[Sza03] T. Szarek. Invariant measures for nonexpansive Markov operators on polish spaces. Dissertationes Math., (415), 2003.

[Sza08] T. Szarek. The uniqueness of invariant measures for Markov operators. Studia Mathematica - STUD MATH, 189:225–233, 01 2008.

[Sza10] T. Szarek, M. Ślęczka and M. Urbański . On stability of velocity vectors for some passive tracer models. Bulletin of the London Mathematical Society, 42:923–936, 2010.

[Tem88] R. Temam. Infinite-Dimensional Dynamical Systems in Mechanics and Physics, volume 68. 1988.

[Tro59] H. F. Trotter. On the product of semi-groups of operators. Proceedings of the American Mathematical Society, 10(4):545, April 1959.

[Var61] V.S. Varadarajan. Measures on topological spaces (russ.). Mat. Sbornik, 55:161–228, 1961.

[Vil08] C. Villani. Optimal transport, old and new. Springer-Verlag, Berlin, 2008.

[Web08] G.F. Webb. Population Models Structured by Age, Size, and Spatial Position, pages 1–49. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[Wor10] D.T.H. Worm. Semigroups on spaces of measures. PhD thesis, Leiden University, 2010.

[Wu07] W.B. Wu. Strong invariance principles for dependent random variables. Ann. Probab., 35(6):2294–2320, 11 2007.

[Zah00] R. Zaharopol. Fortet–Mourier norms associated with some iterated function systems. Statistics and Probability Letters, 50(2):149 – 154, 2000.

[Zah14] R. Zaharopol. Invariant Probabilities of Transition Functions. Springer, in the series ‘Probability and Its Applications’, Cham, 2014.

Samenvatting

Het onderwerp van dit proefschrift 'Aanpak van Markov Operatoren op Ruimten van Maten door Middel van Equicontinuiteit', combineert een analytische en kanstheoretische aanpak van Markov operatoren. Wij beschouwen Markov operatoren die afkomstig zijn van deterministische dynamische systems en ook stochastische processen die afkomen van een kanstheoretische aanpak.

In de studie van Markov operatoren en Markov semigroepen zijn de centrale vragen het begrijpen van het gedrag van de processen en van de semigroepen. Het is van bijzonder belang om vast te stellen of er invariante maten bestaan, de eventuele uniciteit van deze en inzicht krijgen in het lange termijn gedrag van het proces en het dynamisch systeem dat gedefinieerd wordt door de geassocieerde Markov operator of semigroep. Onderzoek naar deze vragen gaat terug tot het werk van Andrey Markov, die een Markov eigenschap voor ketens beschreef. Een groot deel van de theory voor Markov ketens kan gevonden worden in het boek van Meyn en Tweedie, die een grote bijdrage hebben geleverd aan de theory van Markov ketens en die een bemerkenswaardige beschrijving hebben gegeven van 'e-chains'. Dit was motivatie voor veel onderzoekers om te werken met equicontinuiteitseigenschappen. Deze theorie is toepasbaar wanneer de onderliggende toestandsruimte locaal compact is. Als dat niet het geval is - in de algemeenheid van zogenaamde Poolse ruimten - is er theorie in ontwikkeling. Lasota en Szarek, en in recente jaren Worm, generaliseerden theorie van Markov operatoren en families van Markov operatoren naar deze situatie. De theorie werd ontwikkeld door te beginnen met contractieve Markov operatoren in de artikelen van Lasota, door niet-expansieve Markov operatoren in die van Szarek, en uiteindelijk door het beschouwen van equicontinue families van Markov operatoren in die van Szarek, Hille en Worm. Wij breiden resultaten van hen uit en schijnen nieuw licht op de reeds bestaande resultaten door deze geldig te laten zijn onder meer algemene condities.

Een verband tussen zwakke en sterke (norm-) convergentie van rijen van getekende maten is het eerste en fundamentele resutaat in dit proefschrift. Het cruciale onderdeel is de uitbreiding van de resultaten die geldig zijn voor positieve eindige maten naar ruimten van getekende maten. Dit is een heel algemeen middel dat niet alleen in de theorie van Markov operatoren gebruikt kan worden, maar ook in algemene (geavanceerde) maattheorie. Met behulp van dit resultaat, waarvoor wij een zelfstandig leesbaar en onafhankelijk bewijs geven, leiden wij een inzichtelijke correspondentie af tussen de equicontinuiteitseigenschap zoals die geïntroduceerd is door Komorowski, Peszat en Szarek in [KPS10] ('e-property') en de gebruikelijke notie van een equicontinue familie van afbeeldingen, namelijk die gegeven zijn door de de semigroep van Markov operatoren op maten. Met dit resultaat zijn wij in staat om een Lie-Trotter productformule te bewijzen voor Markov semigroepen.

De kernideeën in de generalisatie van de Lie-Trotter productformule tot Markov semigroepen is om realistische, verifieerbare, condities te geven en convergentie te bewijzen van de productformule in de relevante normen. De nieuwe cruciale aannames laten sterke continuiteit van de semigroep en begrensheid van de individuele operatoren vallen, aangezien Markov semigroepen op maten vaak noch sterk continu zijn, noch bestaan uit begrensde operatoren op getekende maten ten aanzien van de Dudley of Fortet-Mourier norm. Ook worden de eigenschappen van de limietsemigroep geanalyseerd die afkomt van de alternerende semigroepen waarmee men begint. Dit geeft een extra mogelijkheid om gecompliceerde problemen te benaderen, door hen op te slitsen in afwisselende 'eenvoudiger' problemen.

Het volgende deel van het proefschrift beschrijft de relatie tussen equicontinuiteit en stabiliteit van Markov operatoren. In het bijzonder wordt aangetoond, dat elke asymptotisch stabiele Markov operator met een invariante maat die zo is dat het inwendige van diens drager niet-leeg is, de e-eigenschap ('e-property') heeft. Deze resultaten zijn van belang aangezien zij vergelijkbare resultaten uitbreiden die geldig zijn op compacte ruimten naar de theorie van Poolse ruimten, waarin geen (locaal) compactheid wordt verondersteld.

Als laatste laten wij zien dat de Centrale Limietsteling (CLSt) geldt voor een klasse van niet-stationaire Markov ketens op Poolse ruimten. Recente resultaten ten aanzien van CLSt van Komorowski en medeauteurs voor niet-stationaire Markov processen laten het belang van dit onderwerp zien. In het bijzonder in toepassingen maken de geldigheid van de Wet van de Grote Aantallen en de Centrale Limietstelling het in principe mogelijk om informatie over de 'vorm' van de invariante maat te verkrijgen door het simuleren van (veel) individuele realisaties van de keten en dan het gemiddelde te nemen. De CLSt geeft de snelheid van convergentie van deze procedure. De uitbreiding van het resultaat van Gordin en Lifsic die wij bewijzen, is mogelijk vanwege de spectrale kloof van de Markov operator ten aanzien van de Kantorovich-Rubinstein norm. Enige delicate benaderingen staan ons toe een sterker resultaat af te leiden dan dat van Komorowski.

Summary

The subject of this thesis, ‘Approach to Markov Operators on Spaces of Measures by Means of Equicontinuity’, combines an analytical and probabilistic approach to Markov operators. We look at Markov operators coming from deterministic dynamical systems and also stochastic processes which come from a probabilistic approach.

In the study of Markov operators and Markov semigroups the central problems are to understand the behaviour of the processes and semigroups. Of particular interest is to identify the existence and uniqueness of invariant measures and long term behaviour of the process and dynamical system defined by the associated Markov operator or semigroup. Research on these questions dates back to the works of Andrey Markov, who described a Markov property for chains. A big part of theory for Markov chains can be found in the book by Meyn and Tweedie, who made a big contribution to the theory of Markov chains and gave a noteworthy description of e-chains, which was the motivation to working with equicontinuity properties for many authors. This theory is applicable when the underlying state space is locally compact. If it is not - in the generality of so-called Polish spaces - there is theory under development. Lasota and Szarek, and in recent years Worm generalized theory of Markov operators and families of Markov operators to this setting. In subsequent years, the theory was being developed starting with contractive Markov operators in the works of Lasota, through non-expansive Markov operators in Szarek’s,, and finally equicontinuous families of Markov operators in that of Szarek, Hille and Worm. We extend their results and give a new light to the existing ones by providing less restrictive conditions in cases.

A connection between weak and strong (norm) convergence of sequences of signed measures is a starting point of the thesis. The crucial thing is the extension of the results from positive measures to the spaces of signed measures. This is a very general tool which can be used not only in a theory of Markov operators, but also in general (advanced) measure theory. With the aid of this result, for which we give a self-contained proof, one gets an enlightening correspondence between the equicontinuity property as defined by Komorowski, Peszat and Szarek in [KPS10] (e-property) and the usual notion of an equicontinuous family of maps, furnished by the semigroup defined by the Markov operator on measures. With this result, we are able to prove a Lie-Trotter product formula for Markov semigroups.

The key ideas of the generalization of the Lie-Trotter product formula to Markov semigroups is to give realistic conditions and prove convergence in the relevant norms. The crucial assumptions are to drop strong continuity and boundedness of the semigroup, as

Markov semigroups are often neither strongly continuous nor consist of bounded operators on signed measures in the appropriate Dudley or Fortet-Mourier norm on signed measures. Also, the properties of the limit semigroups coming from the starting semigroups are being analysed. This gives an additional tool for dealing with complicated problems by splitting them into alternative "easier" ones.

The next part of this thesis describes the relation between equicontinuity and stability of Markov operators. In particular, it is shown that any asymptotically stable Markov operator with an invariant measure such that the interior of its support is non-empty, satisfies the e-property. These results are of importance as they extend similar, existing results on compact spaces to the theory of Polish spaces which does not assume compactness.

Finally, we show the Central Limit Theorem (CLT) for some non-stationary Markov chains on Polish spaces. Recent CLT results by Komorowski and co-workers for non-stationary Markov processes show the importance of the topic. In particular in applications, the validity of Law of Large Numbers and the CLT for non-stationary chains allows one in principle to get information on the 'shape' of the invariant measure by means of simulating sample trajectories of the chain and averaging. The CLT provides the rate of convergence of this procedure. The extension of Gordin and Lifšic that we prove is possible by the aid of the spectral gap property in the Kantorovich-Rubinstein norm. Some delicate approximation allows us to also obtain a stronger result then Komorowski's.

Acknowledgments

First, I would like to thank my supervisor Dr. Sander Hille. Without long hours of conversations and effort you put into my PhD project this thesis would never come to life. Second I would like to thank Prof. Dr. Tomasz Szarek, who recommended me for this position. And who, as co-author, by his expertise and passion for the topic, was a source of many inspiring results. My thanks go to my promotor Prof. Dr. Arjen Doelman for bringing the Analysis Group together. I would also like to thank my office mates Eric, Lotte, Bjorn, Corine, Leonardo and Claudiu. From among the staff I would like to thank Martina and Hermen-Jan for all the supportive discussions.

Finally I would like to thank the people closest to me. My parents, sisters and grandma for always believing in me. Iza and Pawel, who stayed there for me even though I moved far away. My parents-in-law, Eric and Pinar for becoming my family. And finally, Jos and Julian. I didn't know you before moving to The Netherlands but now I cannot imagine life without the two of you.

Curriculum vitae

Maria Aleksandra Ziemlańska was born in Gdańsk, Poland in 1985. After graduating from high school in 2003, she studied Biotechnology at Gdańsk Technical University. In 2006 she started her bachelor studies in pure mathematics at Gdańsk University and in 2009 she obtained a Bachelor of Science in Mathematics. She continued her master studies in financial Mathematics and in 2011 she obtained a Master of Science in Mathematics . She wrote her Master's thesis under the title *Option pricing in continuous time*. In 2014 she moved to the University of Leiden and started a PhD research under the supervision of dr. Sander C. Hille, funded by a Huygens Fellowship.