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

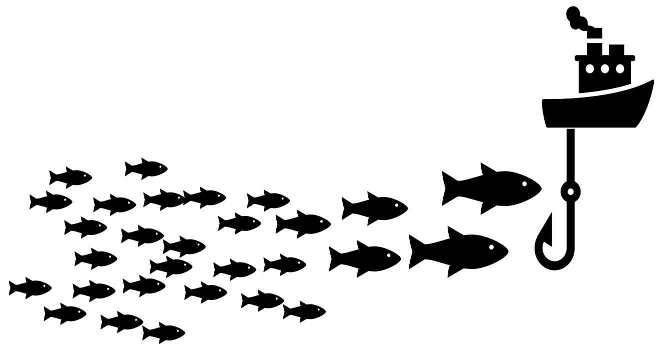

Figure 1: ***Sustainable harvesting***: *A marine ecosystem from which fish are harvested. The interventions will interfere with the further growth of the population. The main question is to quantify the impact of this sampling process on the population. Intensive fishery reduces the fish population 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

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

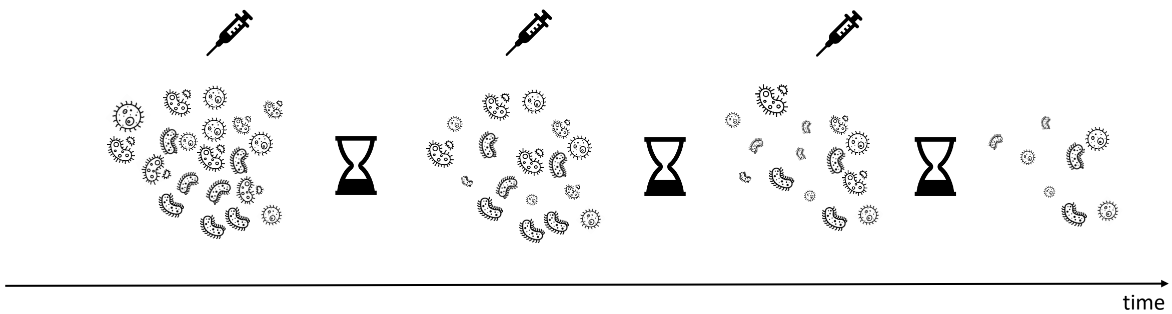

Figure 2: ***Antibiotic treatment****. The question is how to determine the right dose of medicine.*

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

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

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]:*

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

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 $\|\cdot\|_{BL}^*$

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

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*.*

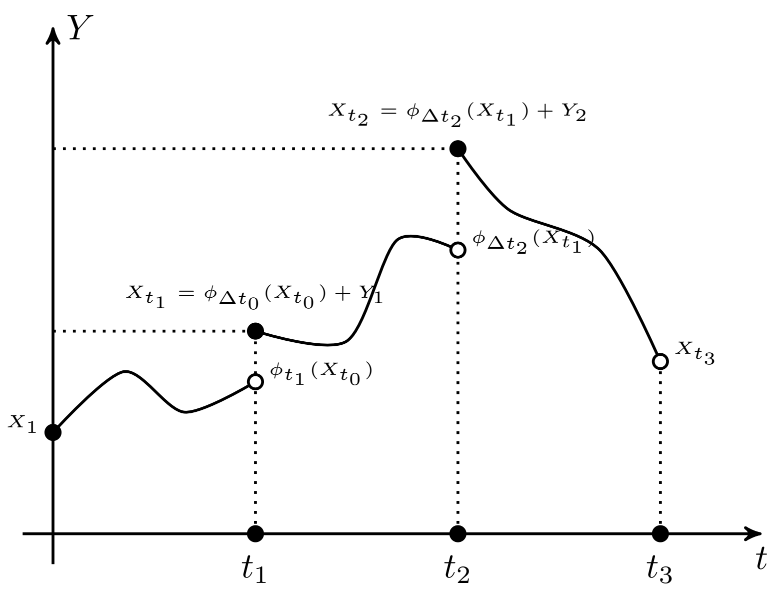

Figure 4: *Piecewise Deterministic Markov Process starting at* $t_0 = 0$ *with value* $X_0 \in Y$*. The motion until time* t_1*, the time of the first jump, is given by* $\phi_t(X_{t_0})$*. At time* t_1 *we have the first jump* Y_1*. Hence,* $X_{t_1} = \phi_{\Delta t_0}(X_1) + Y_1$ *and* X_{t_2} *becomes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

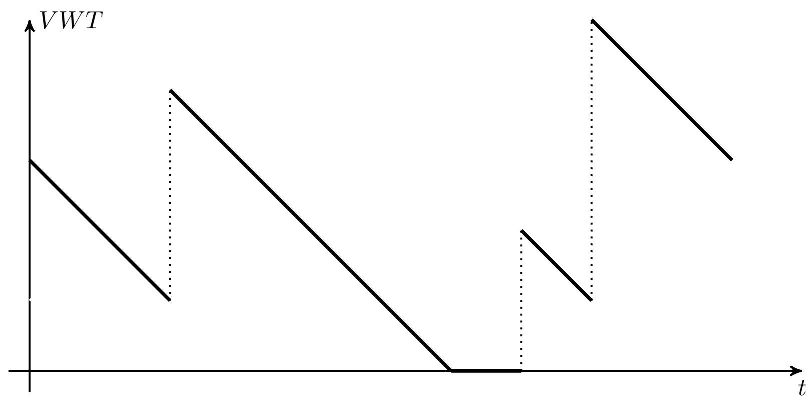

Figure 5: *M/G/1 Queue. A queue model, where arrivals are Markovian (modulated by a Poisson process), service times have a General distribution and there is a single server. VWT is the virtual waiting time*

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

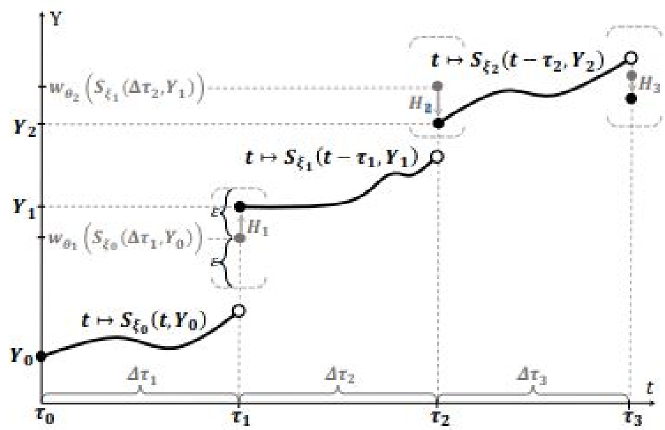

Figure 6: *[HCWS17] More general Piecewise Deterministic Markov Process. The deterministic component of the system evolves according to a finite collection of semiflows (randomly switched with time). Randomness of post-jump location comes from a selected semiflow and a random shift within an ϵ-neighbourhood.*

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

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

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*

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

where

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

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