SEMIGROUPS ON SPACES OF MEASURES

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 16 september 2010
klokke 15:00 uur
door

Daniël Theodorus Hendrikus Worm

geboren te Warnsveld
in 1983.

Samenstelling van de promotiecommissie:

Promotor:	Prof. dr. S.M. Verduyn Lunel
Copromotor:	Dr. S.C. Hille
Overige leden:	Prof. dr. O. Diekmann (Universiteit Utrecht)
	Dr. ir. O. van Gaans
	Prof. dr. F.H.J. Redig (Radboud Universiteit Nijmegen)
	Prof. dr. P. Stevenhagen
	Prof. dr. T. Szarek (Uniwersytet Gdański, Polen)

Semigroups on spaces of measures

Email: daniel.worm@gmail.com

Printed by Ipskamp Drukkers, Enschede
Cover design by Bernd Worm

Thomas Stieltjes Institute
for Mathematics

CONTENTS

NOTATIONAL CONVENTIONS

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

- $\mathbb{N}$ denotes the set of natural numbers $\{1, 2, 3, ...\}$. $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
- $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$.
- (Ω, Σ) is a measurable space.
- $\mathcal{M}(\Omega)$ is the real vector space of all signed finite measures on Ω.
- $\mathcal{M}^+(\Omega)$ is the cone of positive measures in $\mathcal{M}(\Omega)$.
- $\mathcal{P}(\Omega)$ consists of the probability measures in $\mathcal{M}^+(\Omega)$.
- The total variation norm on $\mathcal{M}(\Omega)$ is given by $\|\mu\|_{\mathrm{TV}} = \mu^+(\Omega) + \mu^-(\Omega)$.
- $\mathrm{BM}(\Omega)$ is the real vector space of all bounded measurable functions from Ω to $\mathbb{R}$.
- $\mathbb{1}_E$ is the indicator function of $E \subset \Omega$ and $\mathbb{1} := \mathbb{1}_\Omega$.
- If $f : \Omega \to \mathbb{R}$ is measurable and $\mu \in \mathcal{M}(\Omega)$, then

$$\langle \mu, f \rangle := \int_\Omega f \, d\mu.$$

- 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 $\| \cdot \|_\infty$.
- If (S, d) is a metric space, $C_{ub}(S)$ is the Banach space of uniformly continuous and bounded functions from S to $\mathbb{R}$, endowed with the supremum norm $\| \cdot \|_\infty$.
- If S is a locally compact Hausdorff space, $C_0(S)$ is the Banach space of bounded continuous functions f from S to $\mathbb{R}$ that *vanish at infinity*, i.e. for all $\epsilon > 0$ there is a compact $K \subset S$ such that $|f(x)| < \epsilon$ whenever $x \notin K$. $C_0(S)$ is endowed with the supremum norm $\| \cdot \|_\infty$.

- Let (S, d) be a metric space and $\epsilon > 0$. For $D \subset S$,

$$D^\epsilon := \{x \in S : d(x, D) < \epsilon\},$$

  and for $x \in S$, $B_x(\epsilon) = \{z \in S : d(z, x) < \epsilon\}$.
- For $x, y \in \mathbb{R}$, $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$.

CHAPTER

ONE

INTRODUCTION

The subject of this thesis, *semigroups on spaces of measures*, is located in a part of mathematics where (abstract) analysis and probability theory meet. The relevant questions in this field are therefore motivated partially by analysis, partially by probability theory. Answers to these questions might thus be obtained by arguments that originate from either of these two fields in mathematics or a suitable combination of both. Our personal background in analysis and the initial (analytical) questions that motivated investigation into this subject has resulted in dominance of the analytical viewpoint in what follows. The results thus obtained so far, presented in this thesis, and which shall be introduced further on, suggest further exploration of this approach. More emphasis in future research on the relationship with probability theory and the probabilistic viewpoint, with a suitable mixture of analytical and probabilistic techniques, will lead to further valuable results, we expect.

The research was in fact motivated by applications in analysis: the analysis of long term behaviour of so-called kinetic models for chemotaxis [14, 67, 60, 61], which is an example of a structured population model, be it without birth and death processes. The natural framework for formulating such models seems to be a cone of (positive) measures, rather than a Banach space of integrable functions (densities) of varying regularity. The dynamics in this cone is captured by a semigroup of transformation operators. Deliberately, we have chosen to investigate properties of these types of *dynamical systems* in a general abstract setting, to some extent similar in 'flavour' to the general abstract approach in [54, 108, 116] towards attractors of dynamical systems. That is, to establish properties of the semigroup and the (metric) state space it acts on, as general as possible, that imply strong conclusions on the long term dynamics. As a consequence, this approach has led us farther from specific applications than anticipated at the start. The results should be applicable in a

broad setting though. Let us consider these topics and results in more mathematical detail and introduce the main themes and structure of this thesis.

The field of deterministic dynamical systems has become an important field in mathematics with applications in many other sciences, where it is used to model all kinds of dynamic behaviour. Examples are population dynamics in biology [33, 96, 117], mechanics in physics [116] and chemical kinetics in chemistry [16]. It involves the motion of a system in time. An important class consists of the autonomous (or time-homogeneous) dynamical systems. One can express these in a very general way (see [35, Epilogue] for an interesting philosophical essay on this subject): the state of the system at any time can be characterised by a point in a set Ω, the *state space* of the system. The evolution of the system in time (discrete time, $T = \mathbb{N}_0$ or continuous time, $T = \mathbb{R}_+$) is then represented by a map $T \to \Omega : t \mapsto x_t$. In an autonomous deterministic dynamical system it is assumed that there exists a *semigroup* of evolution operators $\Phi_t : \Omega \to \Omega$, such that $\Phi_0 = \mathrm{Id}_\Omega$, $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $x_t = \Phi_t(x_0)$. In order to obtain a rich theory, more structure must be put on both Ω and $(\Phi_t)_{t\in T}$. For instance, one can take a measurable space (Ω, Σ) and $\Phi_t : \Omega \to \Omega$ measurable, or Ω a topological space, with the Φ_t Borel measurable or continuous. If $T = \mathbb{R}_+$, often some assumptions (measurability or continuity) on the map $t \mapsto \Phi_t(x_0)$ are made.

Even further specialising, we can consider as state space a Banach space X and assume $\Phi_t : X \to X$ to be bounded linear operators. If $T = \mathbb{R}_+$ and $t \mapsto \Phi_t(x)$ is continuous for all $x \in X$, this brings us in the realm of strongly continuous one-parameter semigroups, or C_0-semigroups, of bounded linear operators. Research in this field was initiated in the first half of the twentieth century, with a cornerstone result being the Hille-Yosida Generation Theorem in 1948, developed independently by Hille [58] and Yosida [128], giving necessary and sufficient conditions for a linear operator to be the generator of a C_0-semigroup. A great deal of theory and results have been developed for this well-studied field, see e.g. the seminal work of Hille and Phillips [59] and the more recent book by Engel and Nagel [35]. It has been extremely useful in the field of partial differential equations in which it has formed a functional analytic way to look at solutions to evolutionary equations (see e.g. [48, 94, 108]), and it also has valuable applications in, for instance, delay differential equations, control theory and Volterra equations [35, Chapter VI], [22].

A common approach for the construction of new dynamical systems from known ones is *perturbation*. The Trotter Product Formula [35, Section III.5] relates this to *switching*: when two semigroups are alternated sufficiently fast, the resulting motion in state space can be approximated by a perturbation of each by the generator of the other. *Stochastic* switching at fixed times between multiple dynamical systems provides a simple example how analysis and probability theory may meet in the setting of semigroups on spaces of measures. Let us consider discrete time, for simplicity of exposition. Assume Ω to be an arbitrary non-empty set, let $\Phi_1, \ldots, \Phi_N$ be maps from Ω to Ω and let $p_1, \cdots, p_N$ be such that p_i is the probability that at switching Φ_i is selected. Say we start at initial state $x_0 \in \Omega$. For the next step

we choose map Φ_i with probability p_i, which would give the state $\Phi_i(x_0)$. This is known as an *iterated function system*, which plays an important role in the theory of fractals [7, 66] and has for instance applications in image compression [42].

Such an iterated function system defines a semigroup on the space of measures as follows. Suppose (Ω, Σ) is a measurable space, and the Φ_i are measurable maps. At each switching time we consider the distribution of the possible states just after switching, and describe the evolution of the system on the set of probability measures $\mathcal{P}(\Omega)$: if we start with an initial distribution $\mu_0 \in \mathcal{P}(\Omega)$ (for instance $\mu_0 = \delta_{x_0}$), then the probability that the next state is in $E \in \Sigma$ is given by

$$P\mu_0(E) := \sum_{i=1}^{N} p_i \mu_0(\Phi_i^{-1}(E)).$$

Iterating, $P^n\mu_0$ is the probability measure that describes the distribution after n times of switching. If Ω is a complete separable metric space and the Φ_i are strict contractions, then there exists in fact a unique *invariant probability measure* μ_* ($P\mu_* = \mu_*$), and for every probability measure μ the iterates $P^n\mu$ converge, in some sense, to μ_*. The support of this invariant measure is exactly the fractal associated to the iterated function system [66].

Another way to add stochasticity to a dynamical system is to apply some fixed transformation (perturbation) at random times. An example of this is given by a cell-cycle model [23] (see also [46]): in this model, the state space Ω consists of the possible sizes of a cell, and there is a deterministic dynamical system $(\Phi_t)_{t\geq 0}$ describing the growth of a cell. A probability distribution over the cell size prescribes whether a cell of that size will divide into two smaller ones of halve size. There is also a probability distribution describing whether a cell of particular size will die. This defines a stochastic model for the life history of individuals in subsequent generations, yielding a stochastic process for the population state, being a measure over the individual state space Ω. In [23] a deterministic description is obtained for the evolution of the *expected* population composition at each time, which is a deterministic dynamical system $(P(t))_{t\geq 0}$ on the space of finite measures $\mathcal{M}(\Omega)$ (and in particular the cone of positive finite measures $\mathcal{M}^+(\Omega)$). In the cell-cycle model the individuals evolve independently from each other or the environment, ensuring that the maps $P(t)$ are linear on $\mathcal{M}(\Omega)$. Whenever there is interaction with/through the environment, one typically obtains nonlinear semigroups; see e.g. [60, 61].

The iterated function system we considered above is an example of a particular class of stochastic processes, the *time-homogeneous Markov processes* on a measurable space (Ω, Σ) [38, 86, 104]. An important concept in Markov processes that allows us to study them using (functional) analytic methods is the *transition function* $p(t, x, E)$, giving the probability that at time $t \in T$ the state is in $E \in \Sigma$, given that the state is x at time 0. As in the setting of an iterated function system, we can naturally associate a family of linear positive operators $(P(t))_{t\in T}$ on the space of

measures $\mathcal{M}(\Omega)$, by defining

$$P(t)\mu(E) := \int_\Omega p(t,x,E)\,d\mu(x).$$

If the transition function satisfies the Chapman-Kolmogorov identity, i.e.

$$p(t+s,x,E) = \int_\Omega p(s,y,E)\,p(t,x,dy),$$

then the operators $(P(t))_{t\in T}$ form a so-called *Markov semigroup*, because the operators $P(t)$ are Markov operators: positive linear operators on $\mathcal{M}(\Omega)$ that preserve mass on the cone of positive measures, i.e. $P(t)\mu(\Omega) = \mu(\Omega)$. The transition function defines a *dual* semigroup $(U(t))_{t\in T}$ on the space of bounded measurable real-valued functions $\mathrm{BM}(\Omega)$ (also known as the *transition semigroup* associated to the Markov process)

$$U(t)f(x) := \int_\Omega f(y)\,dp(t,x,dy).$$

The two semigroups obviously satisfy a duality relation:

$$\int_\Omega U(t)f\,d\mu = \int_\Omega f\,dP(t)\mu. \tag{1.1}$$

Both these semigroups are deterministic dynamical systems. A deterministic dynamical system $(\Phi_t)_{t\in T}$ on Ω is a specific example of a Markov process with transition function $p(t,x,E) = \delta_{\Phi_t(x)}(E)$. The associated Markov semigroup is then defined by $P_\Phi(t)\mu := \mu\circ\Phi_t^{-1}$. Markov semigroups form the main theme of our thesis. A natural and quite general assumption that we often impose on a Markov semigroup $(P(t))_{t\in T}$ is *regularity*, i.e. the existence of a semigroup $(U(t))_{t\in T}$ on $\mathrm{BM}(\Omega)$ such that (1.1) is satisfied. This is equivalent with the property that $(P(t))_{t\in T}$ can be represented by a transition function. Regular Markov semigroups and their duals are an important tool in the analysis of Markov processes [38, 40, 86].

Markov semigroups are also naturally associated with stochastic partial differential equations, because solutions to these equations are often given by time-homogeneous Markov processes on infinite dimensional Hilbert spaces or Banach spaces (see e.g. [19, 20]). When such solutions exist, interesting questions then concern existence, uniqueness and stability of invariant measures. Similar questions appear in the setting of *random dynamical systems*, to which also Markov operators and semigroups can be associated [64, 65].

The results that we have obtained in this thesis suggest that natural conditions on Markov operators and semigroups in order to obtain a sufficiently rich and general theory, for instance on long-term behaviour, seem to be at least regularity and *joint measurability*:

$$(t,x)\mapsto P(t)\delta_x(E) \text{ is jointly measurable for all } E\in\Sigma.$$

Interestingly, for many of our results we do not require any *continuity* dependence of the Markov semigroup on time. A continuity assumption on Markov semigroups that is often considered (see e.g. [20, 77]) is *strong stochastic continuity* (at zero), i.e. the map

$$t \mapsto \int_\Omega f \, dP(t)\mu$$

is continuous (at zero) for all $\mu \in \mathcal{M}^+(\Omega)$ and $f \in C_b(S)$, the space of bounded continuous functions. Joint measurability is more general than strong stochastic continuity, if the state space Ω is perfectly normal (which holds for instance if Ω is a metric space).

We now briefly mention some other ways in which to obtain semigroups on spaces of measures. Starting with a Markov semigroup $(P(t))_{t\geq 0}$, one can construct a new one by "perturbation", for instance by employing a Variation of Constants formula:

$$\mu_t = P(t)\mu_0 + \int_0^t P(t-s)F(\mu_s)\, ds, \tag{1.2}$$

for certain $F : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ linear with an appropriate interpretation of the integral, for instance in a set-wise manner (see Chapter 3), for which joint measurability of $(P(t))_{t\geq 0}$ is required. Depending on F, this would define a semigroup $(Q(t))_{t\geq 0}$ on $\mathcal{M}(\Omega)$ by $Q(t)\mu_0 := \mu_t$. A simple example of this, with $(P(t))_{t\geq 0}$ trivial, has been studied by Lasota and Mackey in the setting of L^1-spaces [78]. Another example is given by the above mentioned cell-cycle model developed in [23], where the perturbation map F has to do with division of cells into smaller ones.

The cell-cycle model is a particular example of a *structured population model*: a biological model dealing with the evolution of a structured population interacting with the environment. The population consists of individuals (e.g. animals, bacteria, cells) who are distinguished by their *state* (e.g. position, age, size, velocity), where Ω is the set of possible states, and dynamics at the individual level is then lifted to dynamical behaviour of the population. In this setting, the given Markov semigroup $(P(t))_{t\geq 0}$ describes the evolution of the individuals through state space without interaction with each other or the environment, and the F describes the influence of the environment on the motion of the individual. For instance, $(P(t))_{t\geq 0}$ might come from an underlying deterministic dynamical system $(\Phi_t)_{t\geq 0}$. In these models there will also be a description of the dynamics of the environment, which might be based on the dynamics of the population. This often forces F to become a nonlinear mapping, implying that the semigroup $(Q(t))_{t\geq 0}$ consists of *nonlinear* operators on the cone of positive measures. Therefore it is not a Markov semigroup. But whenever there is no birth or death of individuals in the model, the conservation of mass *does* hold, implying that $Q(t)\mu(\Omega) = \mu(\Omega)$.

In our thesis we focus on semigroups of *linear* operators on spaces of measures that conserve mass. Applying our results, that we shall mention below, to semigroups of nonlinear operators is an interesting next step. One way to obtain linear semigroups

associated to (1.2) is by supposing that the environment is constant (in time), which often ensures that F is linear, which is a useful way of studying the behaviour of the complete system by bootstrapping methods (see e.g. [21, 24, 25]). If birth or death does occur in the model, we lose conservation of mass. Some of our results still hold in this more general setting (see e.g. Remark 4.4.14). In some models the dynamics of the population are described using partial differential equations, where the unknown is the size and composition of the population at each time given by a density function, i.e. a function in $L^1(\Omega, \mu)$ for a certain measure μ on Ω. For instance, if Ω is (a subset of) $\mathbb{R}^n$, then often μ is chosen to be (a restriction of) the Lebesgue measure on $\mathbb{R}^n$. However, this view is somewhat restrictive, because not all realistic population states can be described by such a density function; for instance, it may be necessary to assume that all individuals have the same state $x \in \Omega$, which implies that the state of the population is given by (a multiple of) the Dirac measure δ_x. Thus the space of measures seems to be the more natural state space for structured population models (see e.g. [24, 25]).

A particular example of a structured population model without birth or death is given by a so-called *kinetic chemotaxis model* [14, 41, 60]. *Chemotaxis* is a process in biology in which moving organisms, like bacteria and amoebae, react to an external chemical signal. The kinetic chemotaxis model is a mesoscopic model, consisting of a pair of partial differential equations describing the evolution of the position-velocity distribution $f = f(x, v, t)$ of cells at position $x \in \mathbb{R}^n$ with velocity $v \in V \subset \mathbb{R}^n$ at time t and of the distribution $S = S(x, t)$ of the chemical signal at position $x \in \mathbb{R}^n$. The chemical signal plays the role of the environment. Using semigroup theory we obtained conditions for global existence of positive mild solutions in certain intersections of L^p-spaces (see our preprint [61]), generalising known results in the literature [14, 67]. Particular choices for p need to be made for the quite technical proof of the global existence results. These choices need not make particular sense in view of the biological interpretation of the results, and it seems to be more natural to be able to have solutions in the cone of all positive measures. Under certain conditions this can be done using (1.2), which will yield Markov semigroups if the chemical signal S is assumed to be constant.

Overview of this thesis

The examples and applications mentioned above give an indication why operators and semigroups on spaces of measures are of interest. In the study of these concepts, we envision a two-step approach. In the first step we consider general Markov operators and semigroups, and prove various interesting results on the structure and behaviour of these semigroups when they satisfy certain properties. This is the step we focus on in this thesis. The logical next step would be to provide means to determine whether Markov operators and semigroups that are constructed in various ways, motivated by applications mentioned above, have these properties, allowing one to apply results obtained in the first step. This second step is beyond the scope of our thesis.

We can divide the material roughly in two parts: the first deals with topologies on spaces of measures and continuity of Markov semigroups and the second is concerned with invariant and ergodic measures for Markov operators and semigroups.

First we give some comments on the generality of the state space Ω. In some parts (Chapters 3 and 4) we can assume the full generality of Ω being a measurable space. In other parts we require that Ω is a metric space (Chapter 2) or a *Polish space* (Chapters 5–8), i.e. a separable topological space that is metrisable by a complete metric. These state spaces are quite general. In many places in the literature on Markov operators or semigroups, Ω is assumed to be compact or locally compact, which has the practical advantage that the space of measures $\mathcal{M}(\Omega)$ can be identified with the dual of the space of continuous functions that vanish at infinity, by the Riesz Representation Theorem. One of the disadvantages of these assumptions is that it does not allow the state space to be an infinite dimensional Banach space, since a Banach space is locally compact if and only if it is finite dimensional. For our results we do not require (local) compactness of the state space, which ensures, among other things, that our results also hold for Markov semigroups associated to stochastic partial differential equations on Hilbert spaces or Banach spaces.

The space $\mathcal{M}(\Omega)$ is a Banach space for the total variation norm $\|\cdot\|_{\mathrm{TV}}$. The operators $P(t), t \in T$, are bounded and linear on $\mathcal{M}(\Omega)$, but in general Markov semigroups will not be strongly continuous (if $T = \mathbb{R}_+$), because the topology given by the total variation norm is too strong. An easy illustration of this is as follows: if x and y are two distinct elements of Ω, then $\|\delta_x - \delta_y\|_{\mathrm{TV}} = 2$. Thus if we have a deterministic dynamical system $(\Phi_t)_{t\geq 0}$, then the associated Markov semigroup $(P_\Phi(t))_{t\geq 0}$ will only be strongly continuous if $(\Phi_t)_{t\geq 0}$ is constant. In order to exploit results on C_0-semigroups, various approaches have been developed to circumvent this problem. Some of these are focused on the dual semigroup $(U(t))_{t\geq 0}$. For instance, under certain conditions this semigroup might be strongly continuous on particular invariant subspaces of $\mathrm{BM}(\Omega)$. Suppose Ω is a topological space endowed with its Borel σ-algebra, then a natural assumption on the Markov semigroup is the *Markov-Feller property*, i.e. the dual $(U(t))_{t\geq 0}$ leaves the space of bounded continuous functions $C_b(\Omega)$ invariant; however, even the restriction of $(U(t))_{t\geq 0}$ to $C_b(\Omega)$ (endowed with the supremum norm) is hardly ever strongly continuous. Under the extra assumption that Ω is also a locally compact Hausdorff space and $(U(t))_{t\geq 0}$ leaves the space of continuous functions that vanish at infinity, $C_0(\Omega)$, invariant, then it is much more common that this restriction is strongly continuous, see e.g. [38]. In this case $(P(t))_{t\geq 0}$ is the adjoint semigroup of a C_0-semigroup.

There are also approaches involving weaker topologies, such that the restriction of $(U(t))_{t\geq 0}$ to $C_b(\Omega)$ is continuous, see e.g. the method of π-semigroups by Priola [102], application of the theory of bi-continuous semigroups to transition semigroups by Farkas [39], and results by Kunze exploiting duality of the two semigroups [76]. Other approaches concentrate on the semigroup $(P(t))_{t\geq 0}$, see e.g. results by Lant and Thieme involving the theory of integrated semigroups [77]. Of the two semigroups, we focus on $(P(t))_{t\geq 0}$ as well, because we are interested in the dynamical behaviour

on the space of measures.

In **Chapter 2** we investigate weaker topologies on spaces of measures. We assume the state space to be a metric space (S, d) and consider the so-called *weak topology* $\sigma(\mathcal{M}(S), C_b(S))$ on the space of finite Borel measures $\mathcal{M}(S)$. In probability theory one often works with this topology. However, while the weak topology is locally convex, it is not in general given by a metric. Varadarajan [119] showed that this topology restricted to the cone of positive separable Borel measures $\mathcal{M}_s^+(S)$ *is* metrisable, by a complete metric if S is complete. Dudley [28, 29] later showed that on $\mathcal{M}_s^+(S)$ the weak topology is in fact equal to the topology given by a *norm* $\|\cdot\|_{\mathrm{BL}}^*$ on $\mathcal{M}_s(S)$. Note that the positive cone $\mathcal{M}_s^+(S)$ is the relevant part of $\mathcal{M}_s(S)$ in many applications, most notably those associated to structured population models and the Markov operators and semigroups that we mentioned above. The norm is defined using the dual of the Banach space of bounded Lipschitz functions $\mathrm{BL}(S)$. We will show that $\mathrm{BL}(S)$ is the dual of a Banach space $\mathcal{S}_{\mathrm{BL}}$, equal to the norm closure of $\mathcal{M}_s(S)$ in $\mathrm{BL}(S)^*$. If S is complete, $\mathcal{M}_s^+(S)$ is a closed convex cone in $\mathcal{S}_{\mathrm{BL}}$ with empty interior. While Markov operators need not be extendable to bounded linear operators on $\mathcal{S}_{\mathrm{BL}}$, this Banach space is still useful in the study of Markov operators and semigroups on metric spaces. For instance, the integral in (1.2) can be interpreted as Bochner integral in $\mathcal{S}_{\mathrm{BL}}$. We give equivalent conditions for elements in $\mathcal{S}_{\mathrm{BL}}$ to be in $\mathcal{M}_s(S)$.

Furthermore, we will consider a class of other Lipschitz spaces $\mathrm{Lip}_{e,h}(S)$, containing locally Lipschitz functions that need not be bounded but have some restrictions on the "growth" of the local Lipschitz constants, indicated by a function $h : \mathbb{R}_+ \to [1, \infty)$. These spaces are also dual spaces, and their preduals $\mathcal{S}_{e,h}$ contain certain spaces of measures $\mathcal{M}_h(S)$ densely, such that the positive measures in $\mathcal{M}_h(S)$ form a closed convex cone in $\mathcal{S}_{e,h}$. By making particular choices for h we obtain spaces $\mathcal{M}_h(S)$ consisting exactly of measures with finite k-th moment. Finally we show that, under mild conditions, a semigroup of Lipschitz transformations $(\Phi_t)_{t\geq 0}$ on the metric space embeds into strongly continuous semigroups of positive bounded linear operators on some of these Banach spaces generated by measures, even *isometrically* in one case.

Chapter 3 introduces some preliminaries needed for the subsequent chapters. Most of the results there are not new. We start by introducing the so-called set-wise integral for measure-valued functions and give its relations with a Bochner integral in the Banach space $\mathcal{S}_{\mathrm{BL}}$ as defined in Chapter 2 when the state space is a metric space. After that we introduce several concepts associated to Markov operators and semigroups on spaces of measures and relate them to properties in terms of the Banach space $\mathcal{S}_{\mathrm{BL}}$.

If $(P(t))_{t\geq 0}$ has an invariant measure μ, then it leaves the subspace of all finite measures absolutely continuous to μ invariant. By the Radon-Nikodym Theorem, we can identify this subspace with $L^1(\mu)$, and the restriction of the total variation norm to this subspace equals the L^1-norm. Under mild conditions on the Markov semigroup it is actually strongly continuous on $L^1(\mu)$. We mention a simple example that will

illustrate some of the results to follow: let $\Omega = \mathbb{R}$, $\Phi_t(x) := x + t$ and $(P_\Phi(t))_{t\geq 0}$ be the associated Markov semigroup. Then $(P_\Phi(t))_{t\geq 0}$ is strongly continuous on the invariant subspace $L^1(m)$, where m is the Lebesgue measure on $\mathbb{R}$, while the Markov semigroup is not strongly continuous on all of $\mathcal{M}(\Omega)$. Often in the literature Markov semigroups are defined and studied on L^1-spaces, where they usually *are* assumed to be strongly continuous, see e.g. [34, 78, 99, 100]. In **Chapter 4** we will relate Markov semigroups on spaces of measures to strongly continuous Markov semigroups on invariant L^1-spaces. For the existence of an invariant subspace $L^1(\mu)$, μ need not be invariant. We will give conditions on μ for which this holds.

We address two important issues:

(1) If $\mu \in \mathcal{M}^+(\Omega)$ is such that $L^1(\mu)$ is invariant under the Markov semigroup $(P(t))_{t\in\mathbb{R}_+}$, then $(P(t))_{t\geq 0}$ induces a semigroup on $L^1(\mu)$. We will give various equivalent conditions for the restricted semigroup to be strongly continuous, one of them being

$$\mathbb{R}_+ \to (\mathcal{M}(\Omega), \|\cdot\|_{\mathrm{TV}}) : t \mapsto P(t)\mu \text{ is continuous.} \tag{1.3}$$

(2) We characterise the subspace of strong continuity $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$, consisting of those μ that satisfy (1.3).

We will obtain various equivalent conditions for a measure to be in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$. A classical result by Plessner [101] implies that in the setting of our simple example $(P_\Phi(t))_{t\geq 0}$ the subspace of strong continuity is exactly $L^1(m)$. By the Lebesgue-Radon-Nikodym Theorem every μ in $\mathcal{M}(\mathbb{R})$ can be uniquely decomposed into $\mu_a + \mu_s$, where $\mu_a \in L^1(m)$, and μ_s is singular with respect to m, hence with respect to all elements in $L^1(m)$. A similar result holds in the general case: we show that the subspace of strong continuity is a projection band in the Banach lattice $(\mathcal{M}(\Omega), \|\cdot\|_{\mathrm{TV}})$, which yields a direct sum decomposition

$$\mathcal{M}(\Omega) = \mathcal{M}(\Omega)^0_{\mathrm{TV}} \oplus \mathcal{M}(\Omega)^\perp_{\mathrm{TV}}.$$

We characterise the complement $\mathcal{M}(\Omega)^\perp_{\mathrm{TV}}$ and prove a Wiener-Young type theorem.

Our main line of investigation in the last part of the thesis (Chapters 5–8) deals with invariant measures of Markov operators and semigroups on a Polish space S. A special role is played by the *ergodic measures* $\mathcal{P}_{\mathrm{erg}}(S)$, which are the extreme points of the convex set of invariant probability measures $\mathcal{P}_{\mathrm{inv}}(S)$, and for which many different characterisations exist. These measures may thus be viewed as the "indecomposable" invariant measures. A classical result (see e.g. [121, Chapter 6]) asserts that any invariant measure μ can be obtained as integral over the set of ergodic measures,

$$\mu(E) = \int_{\mathcal{P}_{\mathrm{erg}}(S)} \nu(E)\, d\rho_\mu(\nu),$$

when $\mathcal{P}_{\mathrm{erg}}(S)$ is considered as a measurable space in a suitable manner. If $\mathcal{P}_{\mathrm{inv}}(S)$ would be compact in $\mathcal{S}_{\mathrm{BL}}$ this would follow from Choquet theory (see e.g. [98]). In

general $\mathcal{P}_{\text{inv}}(S)$ need not be compact; however, we will give conditions in Chapter 8 which ensure that $\mathcal{P}_{\text{erg}}(S)$, hence $\mathcal{P}_{\text{inv}}(S)$, is compact in $\mathcal{S}_{\text{BL}}$. In our results, we obtain a parametrisation of the ergodic measures via a subset of state space, and use that to get an integral decomposition over this subset of invariant measures into ergodic measures. This extends known results [56, 134] from the setting of a locally compact separable metric space to that of a Polish space (see Section 5.6 for further connections to the literature of ergodic decompositions).

Let us illustrate our results by the most simple case. Let P_{Id} be the identity operator on $\mathcal{M}(S)$. Then P_{Id} is a Markov operator and every measure is invariant. The set of extreme points of $\mathcal{P}(S)$ consists exactly of all Dirac measures. Thus $\mathcal{P}_{\text{erg}}(S)$ can be parametrised by S through the map $x \mapsto \delta_x$. For every $E \subset S$ Borel, the map $x \mapsto \delta_x(E)$ is measurable, and

$$\mu(E) = \int_S \delta_x(E)\, d\mu(x) = \left[\int_S \delta_x \, d\mu(x)\right](E),$$

where the latter integral is a Bochner integral in $\mathcal{S}_{\text{BL}}$. Furthermore, we can decompose S into disjoint measurable invariant sets $S = \cup_{x\in S}\{x\}$, such that each ergodic measure is concentrated on exactly one of these sets.

In **Chapter 5** we generalise this: we consider a regular Markov operator P on a Polish space S and give a parametrisation of the ergodic measures associated with this operator in terms of a particular subset Γ^P_{cpie} of the state space: Γ^P_{cpie} is measurable and there is a surjective measurable map $x \mapsto \epsilon_x : \Gamma^P_{cpie} \to \mathcal{P}_{\text{erg}}(S)$ (not injective in general). This set Γ^P_{cpie} consists exactly of those points $x \in S$ for which the Cesàro averages

$$P^{(n)}\delta_x = \frac{1}{n}\sum_{k=0}^{n-1} P^k \delta_x$$

converge in $\mathcal{S}_{\text{BL}}$ to an ergodic measure, and for these points we define $\epsilon_x = \lim_{n\to\infty} P^{(n)}\delta_x$. We use this map to prove an integral decomposition of every invariant probability measure in terms of the ergodic measures over the set Γ^P_{cpie}, i.e.

$$\mu(E) = \int_{\Gamma^P_{cpie}} \epsilon_x(E)\, d\mu(x) = \left[\int_{\Gamma^P_{cpie}} \epsilon_x \, d\mu(x)\right](E) \text{ for all Borel sets } E \subset S,$$

and we give an "explicit" decomposition of the state space based on the convergence properties of the Cesàro averages of Dirac measures. From this we obtain a full Yosida-type decomposition of the state space, by showing the existence of a collection of disjoint invariant sets, such that each ergodic measure is concentrated on exactly one of these sets, and such that the complement of the union of these sets is measurable and thus a null set for each invariant measure.

Our main objective in **Chapter 6** is to show that analogous results to those achieved in Chapter 5 hold for regular jointly measurable Markov semigroups $(P(t))_{t\geq 0}$ on

Polish spaces, which extends known results [18] in the setting of locally compact separable metric spaces. Our approach is centered around the reduction to and relationship with the case of a single regular Markov operator associated to the Markov semigroup, the *resolvent operator*

$$R\mu = \int_{\mathbb{R}_+} e^{-t} P(t)\mu \, dt,$$

which enables us to exploit results from Chapter 5.

In the previous chapters we assumed the existence of invariant measures. It is also of interest to provide conditions for this to hold. In order to obtain this, we need more structure on the Markov operators and semigroups. We will assume certain equicontinuity properties are satisfied: we say a regular Markov operator P with dual operator U has the *e-property* if the family of iterates $(U^n f)_{n\in\mathbb{N}_0}$ is equicontinuous for all bounded Lipschitz f, and the weaker *Cesàro e-property* if the family of Cesàro averages $(U^{(n)} f)_{n\in\mathbb{N}}$ is equicontinuous for all bounded Lipschitz f, with analogous definitions for regular jointly measurable Markov semigroups. These are more general than the well-studied strong Feller property, which is often assumed in order to prove uniqueness of invariant measures (see e.g. [20]). Properties of Markov semigroups with the e-property and their applications to stochastic partial differential equations have been the subject of recent research (see e.g. [37, 70, 73, 81, 114]).

In **Chapter 7**, we show that the (Cesàro) e-property has several implications on the Yosida-type ergodic decomposition of state space given in Chapter 5 and Chapter 6: for instance, the map $x \mapsto \epsilon_x$ is actually continuous as map from S to $\mathcal{S}_{\mathrm{BL}}$ and the various measurable sets involved in this decomposition are actually closed. Using these additional properties of the ergodic decomposition, we obtain several new results on existence, uniqueness and stability of invariant measures.

In **Chapter 8**, we study the set of ergodic measures for a Markov semigroup with the (Cesàro) e-property on a Polish space S. We show that this set is closed in the weak topology or, equivalently, in $\mathcal{S}_{\mathrm{BL}}$. We introduce a weak concentrating condition around a compact set K and show that this condition has several implications on $\mathcal{P}_{\mathrm{erg}}(S)$, one of them being the existence of a Borel subset K_0 of K with a bijective map from K_0 to $\mathcal{P}_{\mathrm{erg}}(S)$, by sending a point x in K_0 to ϵ_x, the limit of the Cesàro averages of δ_x. Another implication is the compactness of $\mathcal{P}_{\mathrm{erg}}(S)$. We also give sufficient conditions for $\mathcal{P}_{\mathrm{erg}}(S)$ to be countable or even finite. Finally, we give quite general conditions that are necessary and sufficient for the Cesàro averages of any measure to converge to an invariant measure. These will imply necessary and sufficient conditions for a Markov semigroup with the Cesàro e-property to be weakly mean ergodic and asymptotically stable.

From the material in Chapter 2 only the definitions and some of the results surrounding $\mathcal{S}_{\mathrm{BL}}$ will be needed in subsequent chapters. The content of Chapter 3 will be used in all subsequent chapters. Chapter 4 and Chapters 5–8 can be read inde-

pendently from each other. Finally, from the set of Chapters 5–8, each chapter will build upon the theory and definitions of the previous one(s).

Six of the chapters are based on papers:

- Chapter 2 is mainly based on the paper *Embedding of semigroups of Lipschitz maps into positive linear semigroups on ordered Banach spaces generated by measures* [62], which is joint work with Sander Hille and has been published in Integral Equations and Operator Theory. Some additional results have been added to the chapter.
- Chapter 4 is a generalisation of the paper *Continuity properties of Markov semigroups and their restrictions to invariant L^1-spaces* [63], which is joint work with Sander Hille and has been published in Semigroup Forum. In the paper we considered a complete separable metric space as state space. In Chapter 4 we extend many of the results from the paper to the full generality of a measurable space.
- Chapter 5 is based on the paper *Ergodic decompositions associated with regular Markov operators on Polish spaces* [125] (with minor modifications), which is joint work with Sander Hille , which has been accepted by Ergodic Theory and Dynamical Systems, and has appeared online there (doi:10.1017/S0143385710000039).
- Chapter 6 is based on the paper *An ergodic decomposition associated to regular jointly measurable Markov semigroups on Polish spaces* (with minor modifications), which is joint work with Sander Hille (with minor modifications) and has been submitted. The paper can be found as Report 2010-02 on www.math.leidenuniv.nl.
- Chapter 7 is based on the paper *Equicontinuous families of Markov operators on complete separable metric spaces with applications to ergodic decompositions and existence, uniqueness and stability of invariant measures* (with minor modifications), which is joint work with Sander Hille and has been submitted. The paper can be found as Report 2010-03 on www.math.leidenuniv.nl.
- Chapter 8 is based on the paper *Ergodic measures of Markov semigroups with the e-property*, which is joint work with Tomasz Szarek and has been submitted. The paper can be found as Report 2010-09 on www.math.leidenuniv.nl. Some of the results in this chapter are slightly more general than the corresponding results in the paper.

CHAPTER

TWO

BANACH SPACES GENERATED BY CLASSES OF MEASURES ON A METRIC SPACE

2.1 Introduction

On a measurable space (S, Σ) we can consider the space $\mathcal{M}(S)$ of finite signed measures on S. This space is a Banach lattice when endowed with the total variation norm $\|\cdot\|_{\mathrm{TV}}$. However, the topology given by the total variation norm is often too strong in applications. As an illustration, we consider a family of measurable maps $\Phi_t : S \to S$, parametrised by the non-negative real numbers $t \in \mathbb{R}_+$, that satisfy the semigroup properties: $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_0 = \mathrm{Id}_S$. We can view this as a continuous-time deterministic or causal dynamical system in S. Then each Φ_t induces a linear operator $T_\Phi(t)$ on the space of signed measures $\mathcal{M}(S)$ on Σ by means of

$$T_\Phi(t)\mu := \mu \circ \Phi_t^{-1}. \tag{2.1}$$

The family of operators $(T_\Phi(t))_{t\geq 0}$ leaves the cone of positive measures $\mathcal{M}^+(S)$ invariant. It constitutes a positive linear semigroup in $\mathcal{M}(S)$ and Φ_t can be recovered from $T_\Phi(t)$ through the relation $T_\Phi(t)\delta_x = \delta_{\Phi_t(x)}$. In this sense, any semigroup of measurable maps on a measurable space (S, Σ) embeds into a positive linear semigroup on the space of signed measures on S.

However, the semigroup $(T_\Phi(t))_{t\geq 0}$ is only strongly continuous with respect to $\|\cdot\|_{\mathrm{TV}}$ if $(T_\Phi(t))_{t\geq 0}$ is constant, since $\|\delta_x - \delta_y\|_{\mathrm{TV}} = 2$ whenever $x \neq y$. Moreover, in general $t \mapsto T_\Phi(t)\delta_x = \delta_{\Phi_t(x)}$ will not even be strongly measurable, because its range will not be separable, which makes $(\mathcal{M}(S), \|\cdot\|_{\mathrm{TV}})$ not a suitable Banach space for studying

a variation of constants formula

$$\mu_t = T_\Phi(t)\mu_0 + \int_0^t T_\Phi(t-s)F(\mu_s)ds, \tag{2.2}$$

because the Banach space structure of $(\mathcal{M}(S), \|\cdot\|_{\mathrm{TV}})$ is poorly related to any way (e.g. Bochner or Pettis) of interpreting the integral in (2.2).

In this chapter we will consider the more specific case when S is a metric space with the Borel σ-algebra. The *weak topology* $\sigma(\mathcal{M}(S), C_b(S))$ on $\mathcal{M}(S)$ is often used in probability theory. However, it is inconvenient for perturbation theory: while it is locally convex, it is not given by a norm on $\mathcal{M}(S)$. There is an important result by Varadarajan [119] however, that the restriction to $\mathcal{M}^+(S)$ of this weak topology is metrisable (when S is separable, or when one restricts to separable positive measures), by a complete metric if S is complete ([119, Theorem 13 and Theorem 18]). Later Dudley [28] showed that a metric coming from a norm may be used: when considering the Banach space $\mathrm{BL}(S)$, given by all bounded Lipschitz functions on S, with norm $\|\cdot\|_{\mathrm{BL}} = |\cdot|_{\mathrm{Lip}} + \|\cdot\|_\infty$, it can be shown that $\mathcal{M}(S)$ can be embedded into its dual $\mathrm{BL}(S)^*$, and by [28, Theorem 9 and Theorem 18] it follows that the norm topology on $\mathrm{BL}(S)^*$ and the weak topology *coincide* on $\mathcal{M}^+(S)$, the relevant *cone* from a probabilistic and population dynamics point of view. We want to study these spaces and topologies further in this chapter.

For instance, we will show that $\mathrm{BL}(S)$ is actually the dual of a Banach space, $\mathcal{S}_{\mathrm{BL}}$, that contains the measures densely and in which $\mathcal{M}^+(S)$ is a closed convex cone. This space seems to be the natural one for studying e.g. (2.2). We will also consider a whole class of other Lipschitz spaces $\mathrm{Lip}_{e,h}(S)$, containing locally Lipschitz functions that need not be bounded but have some restrictions on the "growth" of the local Lipschitz constants, indicated by a function $h : \mathbb{R}_+ \to [1, \infty)$. These spaces are also dual spaces, and their preduals $\mathcal{S}_{e,h}$ contain certain spaces of measures $\mathcal{M}_h(S)$ densely, such that the positive measures in $\mathcal{M}_h(S)$ form a closed convex cone in $\mathcal{S}_{e,h}$. By making particular choices for h we obtain spaces $\mathcal{M}_h(S)$ consisting exactly of measures with finite k-th moment.

We will prove a characterisation of those elements in $\mathcal{S}_{\mathrm{BL}}$ that can be represented by measures. For this we first prove Theorem 2.3.24, which establishes an interesting relationship between weak convergence and norm convergence in $\mathcal{S}_{\mathrm{BL}}$. This result follows from a reinterpretation of a result by Pachl [92, Theorem 3.2] in view of our Banach space $\mathcal{S}_{\mathrm{BL}}$. Pachl's Theorem is formulated in the context of the Banach space $(C_{ub}(S), \|\cdot\|_\infty)$ of uniformly continuous, bounded functions of S, and basically says, among others, that a subset $M \subset \mathcal{M}(S)$ that is bounded on the unit ball in $C_{ub}(S)$ is relatively compact in $\mathcal{S}_{\mathrm{BL}}$ if and only if it is relatively $\sigma(\mathcal{M}(S), C_{ub}(S))$-countably compact.

Cones of positive measures define an ordering on $\mathcal{S}_{\mathrm{BL}}$ and $\mathcal{S}_{e,h}$. We will discuss the relationship between this ordering, the natural pointwise ordering on Lipschitz functions, and positive functionals on $\mathrm{BL}(S)$ and $\mathrm{Lip}_{e,h}(S)$, and give conditions for positive functionals on $\mathrm{BL}(S)$ to be representable by measures.

Finally we show that for semigroups of Lipschitz transformations $(\Phi_t)_{t\geq 0}$, we can extend the associated semigroup $(T_\Phi(t))_{t\geq 0}$ on $\mathcal{M}(S)$ to positive linear semigroups of bounded operators on the spaces $\mathcal{S}_{\mathrm{BL}}$ and $\mathcal{S}_e$ and we give sufficient conditions for strong continuity of these semigroups. The space $\mathcal{S}_e = \mathcal{S}_{e,1}$ is particularly interesting from this point of view, because we show that $(\Phi_t)_{t\geq 0}$ embeds *isometrically* into its associated semigroup on $\mathcal{S}_e$.

The outline of the chapter is as follows: In Section 2.2 and 2.3, we introduce Banach spaces of (locally) Lipschitz functions on S, $\mathrm{BL}(S)$ and $\mathrm{Lip}_{e,h}(S)$, investigate their dual spaces and introduce preduals for both, $\mathcal{S}_{\mathrm{BL}}$ and $\mathcal{S}_{e,h}$ respectively. The latter are closed subspaces of $\mathrm{BL}(S)^*$ and $\mathrm{Lip}_{e,h}(S)$. In Section 2.4 we consider positivity on the various spaces, and give conditions for positive functionals on $\mathrm{BL}(S)$ to be representable by measures. In Section 2.5 we present results on the embedding of a semigroup of Lipschitz transformations Φ_t on S into positive linear semigroups on $\mathcal{S}_{\mathrm{BL}}$ and $\mathcal{S}_e$.

Unless otherwise mentioned, we assume (S, d) to be a metric space consisting of at least two points.

2.2 Banach spaces of Lipschitz functions

$\mathrm{Lip}(S)$ denotes the vector space of *real-valued* Lipschitz functions on S. We only consider real-valued functions, because ordering will play a role in the results to follow. Moreover, it seems that real-valued functions are more 'natural' in the theory of spaces of Lipschitz functions (see [122, p. 13]). The Lipschitz seminorm $|\cdot|_{\mathrm{Lip}}$ is defined on $\mathrm{Lip}(S)$ by means of

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

Clearly, $|f|_{\mathrm{Lip}} = 0$ if and only if f is constant, so $|\cdot|_{\mathrm{Lip}}$ does not define a norm on $\mathrm{Lip}(S)$.

We shall write $\mathrm{LocLip}(S)$ to denote the vector space of real-valued *locally Lipschitz* functions on S , i.e., functions $f : S \to \mathbb{R}$, such that $f : B \to \mathbb{R}$ is Lipschitz continuous for each bounded $B \subset S$. Clearly $\mathrm{Lip}(S) \subset \mathrm{LocLip}(S)$. Note that some authors use the term 'locally Lipschitz' to denote real-valued functions f such that for all $x \in S$ there exists a neighbourhood on which f is Lipschitz continuous. This definition is more general than ours.

We start with some basic facts on Lipschitz functions that we will use repeatedly. First, the distance function is a Lipschitz function:

Lemma 2.2.1. *Let E be a non-empty subset of S. Then $x \mapsto d(x, E)$ is in* $\mathrm{Lip}(S)$. *If $\overline{E} = S$, then $d(\cdot, E) \equiv 0$ and if $\overline{E}$ is a proper subset of S, then $|d(\cdot, E)|_{\mathrm{Lip}} = 1$.*

This follows from the triangle inequality and the fact that $d(x, E) = d(x, \overline{E})$. In particular Lemma 2.2.1 implies that $x \mapsto d(x, y) \in \mathrm{Lip}(S)$ for all $y \in S$.

Second, the pointwise minima and maxima of a finite number of Lipschitz functions are again Lipschitz functions:

Lemma 2.2.2. ([28, Lemma 4]) *Given* $f_1, \cdots, f_n \in \mathrm{Lip}(S)$ *we define*

$$g(x) := \min(f_1(x), ..., f_n(x)) \text{ and } h(x) := \max(f_1(x), ..., f_n(x)).$$

Then $g, h \in \mathrm{Lip}(S)$ *and*

$$\max(|g|_{\mathrm{Lip}}, |h|_{\mathrm{Lip}}) \leq \max(|f_1|_{\mathrm{Lip}}, ..., |f_n|_{\mathrm{Lip}}).$$

In the sequel various normed spaces of (locally) Lipschitz functions on S and their Banach space properties will be the central objects of study. First, for each distinguished point $e \in S$ we introduce the norm $\|\cdot\|_e$ on $\mathrm{Lip}(S)$ by

$$\|f\|_e := |f(e)| + |f|_{\mathrm{Lip}}, f \in \mathrm{Lip}(S). \tag{2.3}$$

If e' is another element in S, then

$$\begin{aligned} \|f\|_e &\leq |f(e')| + |f(e) - f(e')| + |f|_{\mathrm{Lip}} \leq |f(e')| + |f|_{\mathrm{Lip}}(d(e, e') + 1) \\ &\leq \|f\|_{e'}(d(e, e') + 1). \end{aligned}$$

Thus $\|\cdot\|_e$ and $\|\cdot\|_{e'}$ are equivalent norms on $\mathrm{Lip}(S)$.

For the rest of the chapter, we fix an element $e \in S$ and write $\mathrm{Lip}_e(S)$ for the normed vector space $\mathrm{Lip}(S)$ with norm $\|\cdot\|_e$.

We can generalise this definition: let $h : \mathbb{R}_+ \to [1, \infty)$ be a non-decreasing function and $\theta > 0$. We define for an $f \in \mathrm{LocLip}(S)$

$$|f|_{\mathrm{Lip},\theta} := \sup\left\{\frac{|f(x) - f(y)|}{d(x, y)} : x, y \in \overline{B_e(\theta)}, x \neq y\right\}, \tag{2.4}$$

where we let $|f|_{\mathrm{Lip},\theta} = 0$ if there are no $x, y \in \overline{B_e(\theta)}$ such that $x \neq y$. Note that $|\cdot|_{\mathrm{Lip},\theta}$ is a seminorm on $\mathrm{LocLip}(S)$ for all $\theta \geq 0$.

We define for $f \in \mathrm{LocLip}(S)$:

$$\|f\|_{e,h} := |f(e)| + \sup_{\theta \geq 0} \frac{|f|_{\mathrm{Lip},\theta}}{h(\theta)} \in [0, \infty].$$

It induces a seminorm on the vector space

$$\mathrm{Lip}_{e,h}(S) := \{f \in \mathrm{LocLip}(S) : \|f\|_{e,h} < \infty\}. \tag{2.5}$$

Thus $\mathrm{Lip}_{e,h}(S)$ consists of those locally Lipschitz functions f whose local Lipschitz constant $|f|_{\mathrm{Lip},\theta}$ is of order $\mathcal{O}(h(\theta))$. If we choose $h \equiv 1$, then $\mathrm{Lip}_{e,h}(S) = \mathrm{Lip}_e(S)$, with $\|\cdot\|_{e,h} = \|\cdot\|_e$.

The following property is straightforward and implies that $\|\cdot\|_{e,h}$ is a norm on $\mathrm{Lip}_{e,h}(S)$:

Lemma 2.2.3. *If $f \in \mathrm{Lip}_{e,h}(S)$ and $x \in S$ then*

$$|f(x)| \leq \max(1, h(d(x,e))d(x,e))\,\|f\|_{e,h}. \tag{2.6}$$

Proof.

$$\begin{aligned} |f(x)| &\leq |f(x) - f(e)| + |f(e)| \leq |f|_{\mathrm{Lip},d(x,e)} d(x,e) + |f(e)| \\ &\leq h(d(x,e))d(x,e)\frac{|f|_{\mathrm{Lip},d(x,e)}}{h(d(x,e))} + |f(e)| \leq \max(1, h(d(x,e))d(x,e))\,\|f\|_{e,h}. \end{aligned}$$

□

Theorem 2.2.4. *If $h : \mathbb{R}_+ \to [1,\infty)$ is non-decreasing, then $\mathrm{Lip}_{e,h}(S)$ is a Banach space with respect to $\|\cdot\|_{e,h}$.*

Proof. Let $(f_n)_n$ be a Cauchy sequence in $\mathrm{Lip}_{e,h}(S)$. According to (2.6), $(f_n(x))_n$ is a Cauchy sequence for every $x \in S$. For $x \in S$, put $f(x) := \lim_{n\to\infty} f_n(x)$. Let $\epsilon > 0$. There is an $N \in \mathbb{N}$, such that $|f_n - f_m|_{\mathrm{Lip},\theta} \leq \epsilon h(\theta)$ for all $n, m \geq N$, $\theta \geq 0$. Now let $\theta \geq 0$. Then for $x, y \in \overline{B_e(\theta)}$ and $n, m \geq N$ we get that

$$\begin{aligned} |(f_n - f_m)(x) - (f_n - f_m)(y)| &\leq |f_n - f_m|_{\mathrm{Lip},\theta}\, d(x,y) \\ &\leq \epsilon h(\theta) d(x,y) \end{aligned}$$

Therefore,

$$|((f - f_m)(x)) - (f - f_m)(y))| \leq \epsilon h(\theta) d(x,y),$$

for $x, y \in \overline{B_e(\theta)}$, $m \geq N$. Hence $\frac{|f - f_m|_{\mathrm{Lip},\theta}}{h(\theta)} < \epsilon$ for all $m \geq N$. This holds for all $\theta \geq 0$, thus $f \in \mathrm{Lip}_{e,h}(S)$ and $\|f - f_m\|_{e,h} \to 0$ as $m \to \infty$. So $\mathrm{Lip}_{e,h}(S)$ is complete. □

If $h_i : \mathbb{R}_+ \to [1,\infty)$ is non-decreasing for $i \in \{1,2\}$, and $h_1 \leq h_2$, then $\mathrm{Lip}_{e,h_1}(S) \subset \mathrm{Lip}_{e,h_2}(S)$ and $\|f\|_{e,h_2} \leq \|f\|_{e,h_1}$ for every $f \in \mathrm{Lip}_{e,h_1}(S)$, so $\mathrm{Lip}_{e,h_1}(S) \hookrightarrow \mathrm{Lip}_{e,h_2}(S)$.

Lemma 2.2.5. *If $e' \in S$ is an element different from e, then $\mathrm{Lip}_{e,h}(S)$ and $\mathrm{Lip}_{e',h}(S)$ contain the same functions and the norms $\|\cdot\|_{e,h}$ and $\|\cdot\|_{e',h}$ are equivalent.*

Proof. Let $f \in \mathrm{Lip}_{e,h}(S)$, then

$$\begin{aligned}
\|f\|_{e',h} &\leq \|f - f(e')\mathbb{1}\|_{e',h} + \|f(e')\mathbb{1}\|_{e',h} \leq \|f\|_{e,h} + |f(e')| \\
&\leq \|f\|_{e,h} + |f(e)| + |f(e) - f(e')| \\
&\leq \|f\|_{e,h} + |f(e)| + d(e,e')h(d(e,e'))\|f\|_{e,h} \\
&\leq (2 + d(e,e')h(d(e,e')))\|f\|_{e,h}.
\end{aligned}$$

□

As we will see later, natural choices for h are: $h(\theta) = h_k(\theta) := \max(1, \theta^k)$, where $k \in \mathbb{R}_{\geq 0}$. Then $\mathrm{Lip}_{e,h_0}(S) = \mathrm{Lip}_e(S)$ with $\|\cdot\|_{e,h_0} = \|\cdot\|_e$. Furthermore, the following holds:

Lemma 2.2.6. *The map $x \mapsto d(x,e)^{k+1}$ is in $\mathrm{Lip}_{e,h_k}(S)$, with $\|d(\cdot,e)^{k+1}\|_{e,h_k} \leq k+1$.*

Proof. Straightforward computation shows that for $0 \leq a, b \leq \theta$, $|a^{k+1} - b^{k+1}| \leq (k+1)\theta^k|a-b|$. This yields $|h(a) - h(b)| \leq (k+1)\theta^k|a-b|$. Hence for $x, y \in \overline{B_e(\theta)}$ we have

$$|d(x,e)^{k+1} - d(y,e)^{k+1}| \leq (k+1)\theta^k|d(x,e) - d(y,e)| \leq (k+1)\theta^k d(x,y).$$

Thus $|d(\cdot,e)^{k+1}|_{\mathrm{Lip},\theta} \leq (k+1)\theta^k \leq (k+1)h(\theta)$, and so $\|d(\cdot,e)\|_{e,h} \leq k+1$. □

Another important space we consider is $\mathrm{BL}(S)$: the vector space of bounded Lipschitz functions from S to $\mathbb{R}$. For $f \in \mathrm{BL}(S)$ we define: $\|f\|_{\mathrm{BL}} := \|f\|_\infty + |f|_{\mathrm{Lip}}$. This defines a norm on $\mathrm{BL}(S)$.

Proposition 2.2.7. $\mathrm{BL}(S)$ *is complete with respect to* $\|\cdot\|_{\mathrm{BL}}$.

The proof of this proposition proceeds in a similar way to that of Proposition 2.2.4. See also [122, Proposition 1.6.2 (a)]. There, completeness is proved for the alternative (but equivalent) norm $\|f\|_{\mathrm{BL,max}} := \max(\|f\|_\infty, |f|_{\mathrm{Lip}})$. This norm is known as the Fortet-Mourier norm (see [45]) and is also referred to as the Wasserstein norm (for instance in [12]), though the latter term might also be used to denote a related norm on the space of measures with finite first moment associated to the Wasserstein-1 metric (see Remark 2.3.20).

In the rest of this section, we fix a non-decreasing $h : \mathbb{R}_+ \to [1, \infty)$.

If $f \in \mathrm{BL}(S)$, then $f \in \mathrm{Lip}_{e,h}(S)$, so there is a canonical embedding

$$j_h : \mathrm{BL}(S) \hookrightarrow \mathrm{Lip}_{e,h}(S),$$

with $\|j_h(f)\|_{e,h} \leq \|f\|_{\mathrm{BL}}$. Thus $\mathrm{BL}(S)$ embeds continuously into $\mathrm{Lip}_{e,h}(S)$. If S has finite diameter, then $\mathrm{BL}(S) = \mathrm{Lip}_{e,h}(S)$, and it is easy to see that in this case the norms $\|\cdot\|_{\mathrm{BL}}$ and $\|\cdot\|_{e,h}$ are equivalent. Otherwise we can consider the closure of $\mathrm{BL}(S)$ in $\mathrm{Lip}_{e,h}(S)$ with respect to $\|\cdot\|_{e,h}$:

Proposition 2.2.8. *Let S be a metric space with infinite diameter. Then*

$$\mathrm{BL}(S) \subsetneq \overline{\mathrm{BL}(S)}^{\|\cdot\|_{e,h}} \subsetneq \mathrm{Lip}_{e,h}(S).$$

Proof. Define $f(x) := \sqrt{d(x,e)+1}$. Then

$$|f(x) - f(y)| = \frac{|d(x,e) - d(y,e)|}{\sqrt{d(x,e)+1} + \sqrt{d(y,e)+1}} \leq \frac{d(x,y)}{\sqrt{d(x,e)+1} + \sqrt{d(y,e)+1}}.$$

So f is in $\mathrm{Lip}_e(S) \subset \mathrm{Lip}_{e,h}(S)$, but not in $\mathrm{BL}(S)$, since S has infinite diameter. We will show that $f \in \overline{\mathrm{BL}(S)}^{\|\cdot\|_e} \subset \overline{\mathrm{BL}(S)}^{\|\cdot\|_{e,h}}$. Let $f_n(x) := \min(f(x), n)$. Then $f_n \in \mathrm{BL}(S)$ by Lemma 2.2.2. Let $g_n := f - f_n$. Now let $x, y \in S$, $x \neq y$. If $f(x) \leq n$ and $f(y) \leq n$, $|g_n(x) - g_n(y)| = 0$. If $f(x) > n$ and $f(y) > n$, then $|g_n(x) - g_n(y)| = |f(x) - f(y)| \leq \frac{d(x,y)}{2n}$. If $f(x) > n$ and $f(y) \leq n$, then

$$|g_n(x) - g_n(y)| = |f(x) - n| \leq |f(x) - f(y)| \leq \frac{d(x,y)}{n+1}.$$

So $|f - f_n|_{\mathrm{Lip}} = |g_n|_{\mathrm{Lip}} \leq \frac{1}{n+1}$. Therefore $\|f - f_n\|_e \leq \frac{1}{n+1}$ for every $n \in \mathbb{N}$ thus $f \in \overline{\mathrm{BL}(S)}^{\|\cdot\|_e}$.

Now define $g(x) = d(x,e)$. Then g is in $\mathrm{Lip}_e(S) \subset \mathrm{Lip}_{e,h}(S)$, but not in $\mathrm{BL}(S)$. Suppose that $g \in \overline{\mathrm{BL}(S)}^{\|\cdot\|_{e,h}}$, then there is a $k \in \mathrm{BL}(S)$, with $\|g - k\|_{e,h} < \frac{1}{2}$. Moreover, Lemma 2.2.3 yields

$$|g(x) - k(x)| \leq \frac{1}{2}\max(1, h(d(x,e))d(x,e)).$$

This implies that

$$|k(x)| \geq |g(x)| - |g(x) - k(x)| \geq \frac{1}{2}\max(h(d(x,e))d(x,e), 1) - \frac{1}{2}.$$

Because S has infinite diameter, this contradicts that k is bounded. □

We shall write $\|\cdot\|^*_{\mathrm{BL}}$ and $\|\cdot\|^*_{e,h}$ to denote the dual norm on $\mathrm{BL}(S)^*$ and $\mathrm{Lip}_{e,h}(S)^*$ respectively. Note that the adjoint map $j_h^* : \mathrm{Lip}_e(S)^* \to \mathrm{BL}(S)^*$, which restricts a $\varphi \in \mathrm{Lip}_{e,h}(S)^*$ to $\mathrm{BL}(S)$, is continuous, with $\|j_h^*(\varphi)\|^*_{BL} \leq \|\varphi\|^*_e$.

Whenever S has infinite diameter, $\overline{\mathrm{BL}(S)}^{\|\cdot\|_{e,h}} \subsetneq \mathrm{Lip}_{e,h}(S)$, by Proposition 2.2.8. From this and the Hahn-Banach Theorem it follows that there exists a non-zero $\varphi \in \mathrm{Lip}_{e,h}(S)^*$ such that $\varphi|_{\mathrm{BL}(S)} = 0$, hence j_h^* is not injective.

We will use the term *Lipschitz spaces* to refer to $\mathrm{BL}(S)$ and the family of spaces $\mathrm{Lip}_{e,h}(S)$.

Remark. Various authors consider other Banach spaces of Lipschitz functions, such as e.g. Weaver [122], looking at $\mathrm{Lip}_0(S)$ consisting of all Lipschitz functions on S that vanish at some distinct point $e \in S$. On this subspace of $\mathrm{Lip}(S)$, $|\cdot|_{\mathrm{Lip}}$ is a norm for which $\mathrm{Lip}_0(S)$ is complete. Peng and Xu [95] for example, perform the standard construction of dividing out the constant functions in $\mathrm{Lip}(S)$. Then this space of equivalence classes of Lipschitz functions $\mathrm{Lip}(S)/\mathbb{R}\mathbb{1}$ is complete with respect to the norm $|\cdot|_{\mathrm{Lip}}$ and it is isometrically isomorphic to $\mathrm{Lip}_0(S)$.

2.3 Embedding of measures in dual Lipschitz spaces and the spaces $\mathcal{S}_{\mathrm{BL}}$ and $\mathcal{S}_{e,h}$

We will assume $h : \mathbb{R}_+ \to [1, \infty)$ to be a non-decreasing function.

In this section we are concerned with embedding measures into $\mathrm{BL}(S)^*$ and $\mathrm{Lip}_{e,h}(S)^*$.

Let $\mathcal{M}(S)$ be the space of all signed finite Borel measures on S and $\mathcal{M}^+(S)$ the convex cone of positive measures in $\mathcal{M}(S)$.

The Baire σ-algebra is the smallest σ-algebra on S for which all continuous real-valued functions on S are measurable. Since S is a metric space, the Baire and Borel σ-algebras coincide, because for any closed $C \subset S$, $f_C : x \mapsto d(x, C)$ is Lipschitz continuous by Lemma 2.2.1. Therefore we can apply some of the results from Dudley [28] on Baire measures.

Each $\mu \in \mathcal{M}(S)$ defines a linear functional I_μ on $\mathrm{BL}(S)$, by means of $I_\mu(f) := \int_S f \, d\mu$. Then

$$\begin{aligned} \|I_\mu\|^*_{\mathrm{BL}} &= \sup\left\{ \left| \int_S f \, d\mu \right| : \|f\|_{\mathrm{BL}} \le 1 \right\} \\ &\le \sup\left\{ \int_S |f| \, d|\mu| : \|f\|_{\mathrm{BL}} \le 1 \right\} \le |\mu|(S) = \|\mu\|_{\mathrm{TV}}, \end{aligned} \tag{2.7}$$

thus $I_\mu \in \mathrm{BL}(S)^*$. Moreover, one has

Lemma 2.3.1. *Let $\mu \in \mathcal{M}^+(S)$. Then $\|I_\mu\|^*_{\mathrm{BL}} = \|\mu\|_{\mathrm{TV}}$.*

Proof. Suppose $\mu \in \mathcal{M}^+(S)$. From (2.7) it follows that $\|I_\mu\|^*_{\mathrm{BL}} \le \|\mu\|_{\mathrm{TV}}$. Clearly the constant function $\mathbb{1}$ is in $\mathrm{BL}(S)$, with $\|\mathbb{1}\|_{\mathrm{BL}} = 1$. Then $\|\mu\|_{\mathrm{TV}} = \mu(S) = \int_S \mathbb{1} \, d\mu \le \|I_\mu\|^*_{\mathrm{BL}}$. Hence $\|I_\mu\|^*_{\mathrm{BL}} = \|\mu\|_{\mathrm{TV}}$. □

Lemma 2.3.2. ([28, Lemma 6])
The linear map $\mu \mapsto I_\mu : \mathcal{M}(S) \to \mathrm{BL}(S)^*$ is injective.

Thus we can continuously embed $\mathcal{M}(S)$ into $\mathrm{BL}(S)^*$ and identify $\mu \in \mathcal{M}(S)$ with $I_\mu \in \mathrm{BL}(S)^*$. When a functional $\varphi \in \mathrm{BL}(S)^*$ can be represented by a measure in this way, we shall write $\varphi \in \mathcal{M}(S)$.

For $\mu \in \mathcal{M}(S)$ we define

$$|\mu|_{e,h} := \sup\left\{\int_S |f(x)|\, d|\mu|(x) : f \in \mathrm{Lip}_{e,h}(S), \|f\|_{e,h} \leq 1\right\} \in [0,\infty].$$

Let us define the following subspace of $\mathcal{M}(S)$:

$$\mathcal{M}_h(S) := \{\mu \in \mathcal{M}(S) : |\mu|_{e,h} < \infty\}.$$

Note that $\mathcal{M}_h(S)$ does not depend on e by Lemma 2.2.5.

And we put $\mathcal{M}_h^+(S) := \mathcal{M}_h(S) \cap \mathcal{M}^+(S)$.

Lemma 2.3.3. *Let $\mu \in \mathcal{M}_h(S)$. Then $I_\mu \in \mathrm{Lip}_{e,h}(S)^*$ and the linear map $\mu \mapsto I_\mu : \mathcal{M}_h(S) \to \mathrm{Lip}_e(S)^*$ is injective.*

Proof. $\mathcal{M}_h(S)$ is a subspace of $\mathcal{M}(S)$ and thus embeds into $\mathrm{BL}(S)^*$. The image of $\mu \in \mathcal{M}_h(S)$ in $\mathrm{BL}(S)^*$ coincides with the one obtained by mapping $\mathcal{M}_h(S)$ into $\mathrm{Lip}_e(S)^*$ and then restricting to $\mathrm{BL}(S)$. Therefore $\mu \mapsto I_\mu$ is injective. □

Thus we can identify $\mu \in \mathcal{M}_h(S)$ with $I_\mu \in \mathrm{Lip}_{e,h}(S)^*$, and embed $\mathcal{M}_h(S)$ into $\mathrm{Lip}_{e,h}(S)^*$. When a functional $\varphi \in \mathrm{Lip}_{e,h}(S)^*$ can be represented by a measure in $\mathcal{M}_h(S)$, we shall write $\varphi \in \mathcal{M}_h(S)$.

If we take $h = h_k = \max(1, d(e,x)^k)$, then we obtain exactly the measures with finite $k+1$-th moment:

Proposition 2.3.4. *Let $k \in \mathbb{R}_{\geq 0}$. Then*

$$\mathcal{M}_{h_k}(S) = \left\{\mu \in \mathcal{M}(S) : \int_S d(x,e)^{k+1} d|\mu|(x) < \infty\right\}.$$

Proof. Let $\mu \in \mathcal{M}_{h_k}(S)$. By Lemma 2.2.6, $x \mapsto d(x,e)^{k+1}$ is in $\mathrm{Lip}_{e,h}(S)$, hence $\int_S d(x,e)^{k+1} d|\mu|(x) < \infty$.

Now let $\int_S d(x,e)^{k+1} d|\mu|(x) < \infty$. If $f \in \mathrm{Lip}_{e,h_k}(S)$, then by Lemma 2.2.3,

$$|f(x)| \leq \max(1, h_k(d(x,e))d(x,e))\|f\|_{e,h_k} = \max(1, d(x,e)^{k+1})\|f\|_{e,h_k}.$$

Therefore,

$$\begin{aligned}\int_S |f(x)| d|\mu|(x) &\leq \|f\|_{e,h} \int_S \max(1, d(x,e)^{k+1}) d|\mu|(x) \\ &\leq \left(|\mu|(S) + \int_S d(x,e)^{k+1} d|\mu|(x)\right)\|f\|_{e,h},\end{aligned}$$

hence $\mu \in \mathcal{M}_{h_k}(S)$. □

We can embed S into $\mathcal{M}(S)$ or $\mathcal{M}_h(S)$, by sending x to the Dirac measure δ_x. This embedding is not continuous in general with respect to the total variation norm, since $\|\delta_x - \delta_y\|_{\mathrm{TV}} = 2$ whenever $x \neq y$. However, we do have a continuous embedding into $\mathrm{Lip}_{e,h}(S)^*$, which is *isometric* if $h \equiv 1$.

Lemma 2.3.5. *Let $x \in S$, then δ_x is in $\mathrm{Lip}_{e,h}(S)^*$ with*

$$\max(1, d(x,e))d(x,e) \leq \|\delta_x\|^*_{e,h} \leq \max(1, h(d(x,e))d(x,e))$$

and $\|\delta_x\|^*_e = \max(1, d(x,e))$. For $x, y \in S$

$$d(x,y) \leq \|\delta_x - \delta_y\|^*_{e,h} \leq h(\max(d(x,e), d(y,e)))d(x,y). \tag{2.8}$$

Consequently, $x \mapsto \delta_x$ is a continuous embedding from S into $\mathrm{Lip}_{e,h}(S)^*$. In particular, $x \mapsto \delta_x$ is an isometric embedding from S into $\mathrm{Lip}_e(S)^*$.

Proof. Let $f \in \mathrm{Lip}_{e,h}(S)$ and $x \in S$. Then Lemma 2.2.3 implies that $\|\delta_x\|^*_{e,h} \leq \max(1, h(d(x,e))d(x,e))$. Thus $\|\delta_x\|^*_e \leq \max(1, d(x,e))$.

Consider $f(x) := d(x,e)$. Then $f \in \mathrm{Lip}_{e,h}(S)$ and $|f|_{\mathrm{Lip}} = 1$, according to Lemma 2.2.1. Hence $\|f\|_{e,h} \leq 1$, and $|\delta_x(f)| = d(x,e)$ for every $x \in S$. Also, the constant function $\mathbb{1} \in \mathrm{Lip}_e(S)$ and $\|\mathbb{1}\|_{e,h} = 1$. Furthermore, $|\delta_x(\mathbb{1})| = 1$. Hence $\|\delta_x\|^*_{e,h} \geq \max(1, d(x,e))$ and thus $\|\delta_x\|^*_e = \max(1, d(x,e))$.

Now, let $x, y \in S$ and $f \in \mathrm{Lip}_{e,h}(S)$. Then

$$\begin{aligned} |(\delta_x - \delta_y)(f)| &= |f(x) - f(y)| \leq |f|_{\mathrm{Lip},\max(d(x,e),d(y,e))} d(x,y) \\ &\leq \|f\|_{e,h} h(\max(d(x,e), d(y,e)))d(x,y), \end{aligned}$$

thus $\|\delta_x - \delta_y\|^*_{e,h} \leq h(\max(d(x,e), d(y,e)))d(x,y)$. So S embeds continuously into $\mathrm{Lip}_{e,h}(S)^*$. In particular, $\|\delta_x - \delta_y\|^*_e \leq d(x,y)$.

Let $f(z) := d(x,z) - d(x,e)$. Then $|f|_{\mathrm{Lip}} = |d(x,\cdot)|_{\mathrm{Lip}} = 1$, $\|f\|_{e,h} \leq 1$ and $|\delta_x(f) - \delta_y(f)| = d(x,y)$. Hence $\|\delta_x - \delta_y\|^*_{e,h} \geq d(x,y)$ and we obtain (2.8). Moreover, $x \mapsto \delta_x$ is an isometric embedding from S into $\mathrm{Lip}_e(S)^*$. □

The situation for the embedding of S into $\mathrm{BL}(S)^*$ is similar:

Lemma 2.3.6. *For every $x \in S$, δ_x is in $\mathrm{BL}(S)^*$, and $\|\delta_x\|^*_{\mathrm{BL}} = 1$. Furthermore for every $x, y \in S$,*

$$\|\delta_x - \delta_y\|^*_{\mathrm{BL}} = \frac{2d(x,y)}{2 + d(x,y)} \leq \min(2, d(x,y)). \tag{2.9}$$

Proof. Let $x \in S$ and $f \in \mathrm{BL}(S)$. Then $|\delta_x(f)| = |f(x)| \leq \|f\|_{\mathrm{BL}}$, hence $\|\delta_x\|^*_{\mathrm{BL}} \leq 1$. The constant function $\mathbb{1}$ is in $\mathrm{BL}(S)$ and $|\delta_x(\mathbb{1})| = 1 = \|\mathbb{1}\|_{\mathrm{BL}}$, so $\|\delta_x\|^*_{\mathrm{BL}} = 1$.

If $x = y$, then (2.9) is satisfied. Suppose $x \neq y$. Let $f \in \mathrm{BL}(S)$. Then

$$|f(x) - f(y)| \leq \min(|f|_{\mathrm{Lip}} d(x,y), 2\|f\|_\infty).$$

Hence

$$(2 + d(x, y))|f(x) - f(y)| \leq 2d(x, y)\|f\|_{\mathrm{BL}},$$

so

$$\|\delta_x - \delta_y\|^*_{\mathrm{BL}} = \sup_{\|f\|_{\mathrm{BL}} \leq 1} |f(x) - f(y)| \leq \frac{2d(x, y)}{2 + d(x, y)}.$$

Define $f(z) := \frac{d(z,y)-d(z,x)}{2+d(x,y)}$. Then

$$|f|_{\mathrm{Lip}} \leq \frac{1}{2 + d(x, y)}|d(\cdot, y) - d(\cdot, x)|_{\mathrm{Lip}} \leq \frac{2}{2 + d(x, y)},$$

where we use that $|d(\cdot, x)|_{\mathrm{Lip}} = 1$, by Lemma 2.2.1. Since $|d(z, y) - d(z, x)| \leq d(x, y)$ for all $z \in S$, we can conclude that $\|f\|_\infty \leq \frac{d(x,y)}{2+d(x,y)}$. Hence $\|f\|_{\mathrm{BL}} \leq 1$. Furthermore

$$|\delta_x(f) - \delta_y(f)| = |f(x) - f(y)| = \frac{2d(x, y)}{2 + d(x, y)}.$$

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

Remark. Instead of the norms $\|\cdot\|_{\mathrm{BL}}$ and $\|\cdot\|_e$, we could also consider the equivalent norms $\|\cdot\|_{\mathrm{BL,max}} = \max(\|f\|_\infty, |f|_{\mathrm{Lip}})$ and $\|f\|_{e,\max} := \max(|f(e)|, |f|_{\mathrm{Lip}})$. Then the isometric embedding statement from Lemma 2.3.5 holds with $\|\cdot\|^*_e$ replaced by $\|\cdot\|^*_{e,\max}$. The corresponding statement to (2.9) in Lemma 2.3.6 for the $\|\cdot\|^*_{\mathrm{BL,max}}$-norm is that $\|\delta_x - \delta_y\|^*_{\mathrm{BL,max}} = \min(2, d(x, y))$, which can be shown using the function $f(z) := \min(-1 + d(x, z), 1)$ if $d(x, y) < 2$ and $f(z) := \min(-1 + \frac{2d(x,z)}{d(x,y)}, 1)$ if $d(x, y) \geq 2$.

Let

$$D := \mathrm{span}\{\delta_x | x \in S\} = \left\{ \sum_{k=1}^{n} \alpha_k \delta_{x_k} : n \in \mathbb{N}, \alpha_k \in \mathbb{R}, x_k \in S \right\}. \tag{2.10}$$

Definition 2.3.7. *(i) $\mathcal{S}_{\mathrm{BL}}$ is the closure of D in $\mathrm{BL}(S)^*$.*

(ii) $\mathcal{S}_{e,h}$ is the closure of D in $\mathrm{Lip}_{e,h}(S)^*$.

(iii) $\mathcal{S}_e$ is the closure of D in $\mathrm{Lip}_e(S)^*$.

2.3.1 Identification of $\mathcal{S}_{\mathrm{BL}}$

A Borel measure $\mu \in \mathcal{M}(S)$ is called *separable* if there is a separable Borel measurable subset E of S, such that μ is concentrated on E, i.e. $|\mu|(S \backslash E) = 0$. Let $\mathcal{M}_s(S)$ be the separable Borel measures on S, and $\mathcal{M}^+_s(S)$ the set of positive, finite and separable Borel measures on S. If S is separable, $\mathcal{M}_s(S) = \mathcal{M}(S)$. It is easy to see that $\mathcal{M}_s(S)$ is a closed subspace of $\mathcal{M}(S)$ with respect to $\|\cdot\|_{\mathrm{TV}}$.

Let

$$D^+ := \left\{ \sum_{i=1}^{n} \alpha_i \delta_{x_i} : n \in \mathbb{N}, \alpha_i \in \mathbb{R}_+, x_i \in S \right\}.$$

We define $\mathcal{S}_{\mathrm{BL}}^+$ to be the closure of D^+ with respect to $\|\cdot\|_{\mathrm{BL}}^*$. Notice that $\mathcal{S}_{\mathrm{BL}}^+ \subset \mathcal{S}_{\mathrm{BL}}$ and all $\varphi \in \mathcal{S}_{\mathrm{BL}}^+$ are positive: $\langle \varphi, f \rangle \geq 0$ for all $0 \leq f \in \mathrm{BL}(S)$.

We will need the following theorem, which is based on a result from [28]:

Theorem 2.3.8. *$\mathcal{M}_s^+(S)$ is norm closed in $\mathrm{BL}(S)^*$ if and only if S is complete.*

Proof. If S is complete, then $\mathcal{M}_s^+(S)$ is norm closed in $\mathrm{BL}(S)^*$ by [28, Theorem 9]. Suppose S is not complete. Then there exists a Cauchy sequence $(x_n)_n$ in S that does not converge to an element in S. Then $(x_n)_n$ cannot have a convergent subsequence. This implies that for every $x \in S$ there must be an $\epsilon > 0$ and an $M \in \mathbb{N}$, such that $d(x, x_m) \geq \epsilon$ for all $m \in \mathbb{N}, m \geq M$, otherwise $(x_n)_{n\in\mathbb{N}}$ has a subsequence that converges to x.

We will show that $\mathcal{M}_s^+(S)$ cannot be norm closed in $\mathrm{BL}(S)^*$. By Lemma 2.3.6 $(\delta_{x_n})_n$ is a Cauchy sequence in $\mathrm{BL}(S)^*$. Now assume there is a $\mu \in \mathcal{M}_s^+(S)$, such that $\|\delta_{x_n} - \mu\|_{\mathrm{BL}}^* \to 0$. Then

$$\|\mu\|_{\mathrm{BL}}^* = \lim_{n\to\infty} \|\delta_{x_n}\|_{\mathrm{BL}}^* = 1.$$

We will show that μ must be zero, which gives a contradiction. We can assume, by taking a subsequence, that $\|\delta_{x_n} - \mu\|_{\mathrm{BL}}^* < \frac{1}{n^2}$. Now define $f_n(x) := \min(1, nd(x, x_n))$. Then $f_n \in \mathrm{BL}(S)$, with $|f_n|_{\mathrm{Lip}} \leq n$ and $\|f_n\|_\infty \leq 1$. Hence

$$\int_S f_n \, d\mu = \left| \delta_{x_n}(f_n) - \int_S f_n \, d\mu \right| < \frac{n+1}{n^2} \to 0 \text{ as } n \to \infty.$$

Now let $x \in S$. Then there exists an $\epsilon > 0$ and an $M \in \mathbb{N}$, such that $d(x, x_m) \geq \epsilon$ for all $m \in \mathbb{N}, m \geq M$. This implies that $f_n(x) \to 1$ as $n \to \infty$. Hence, by the Dominated Convergence Theorem,

$$\int_S \mathbb{1} \, d\mu = \left| \lim_{n\to\infty} \int_S f_n \, d\mu \right| = 0,$$

which implies that μ is zero. □

Our main result in this section is the following theorem:

Theorem 2.3.9. *$\mathcal{M}_s^+(S) \subset \mathcal{S}_{\mathrm{BL}}^+$. Furthermore, $\mathcal{S}_{\mathrm{BL}}^+ = \mathcal{M}_s^+(S)$ if and only if S is complete.*

Proof. First we show that $\mathcal{M}_s^+(S) \subset \mathcal{S}_{\mathrm{BL}}^+$. Let $\mu \in \mathcal{M}_s^+(S)$, and let E be a measurable separable subset of S on which μ is concentrated. We want to show that there

is an element $\varphi \in \mathcal{S}_{\mathrm{BL}}^+$ such that $\varphi(f) = \int_S f\, d\mu$ for all $f \in \mathrm{BL}(S)$. If $\mu(S) = 0$ this is clear, so we assume $\mu(S) > 0$.

We define the map $\delta : S \to \mathcal{S}_{\mathrm{BL}}$, sending x to δ_x. Then δ is Lipschitz continuous by Lemma 2.3.6. Also, since E is separable and δ is continuous, $\delta(E)$ is a separable subset of $\mathcal{S}_{\mathrm{BL}}$. Because $\mu(S \backslash E) = 0$, δ is μ-essentially separably valued. For any $\phi \in \mathcal{S}_{\mathrm{BL}}^*$ the function $x \mapsto \phi(\delta_x)$ is continuous, hence Borel measurable, so $x \mapsto \delta_x$ is weakly measurable. By the Pettis Measurability Theorem (e.g. [26, Theorem II.2]), δ is strongly μ-measurable. Furthermore,

$$\int_S \|\delta_x\|_{\mathrm{BL}}^*\, d\mu(x) = \int_S \mathbb{1}\, d\mu < \infty,$$

therefore $\delta : S \to \mathcal{S}_{\mathrm{BL}}$ is μ-Bochner integrable and $\int_S \delta_x\, d\mu(x)$ defines an element in $\mathcal{S}_{\mathrm{BL}}$. By [26, Corollary II.8] we get that

$$\frac{1}{\mu(S)} \int_S \delta_x\, d\mu(x) \in \overline{\mathrm{conv}}\{\delta_x : x \in E\} \subset \mathcal{S}_{\mathrm{BL}}^+.$$

Hence $\int_S \delta_x\, d\mu(x) \in \mathcal{S}_{\mathrm{BL}}^+$. Furthermore, by [26, Theorem 6] we obtain for all $f \in \mathrm{BL}(S)$ that $\langle \int_S \delta_x\, d\mu(x), f\rangle = \int_S \langle \delta_x, f\rangle\, d\mu(x) = \int_S f\, d\mu$. This implies $\int_S \delta_x\, d\mu(x)$ is a functional in $\mathcal{S}_{\mathrm{BL}}^+$ represented by μ. Thus $\mathcal{M}_s^+(S) \subset \mathcal{S}_{\mathrm{BL}}^+$.

Now assume S is complete. It is clear that for all $x \in S$, $\delta_x \in \mathcal{M}_s^+(S)$. Hence $D^+ \subset \mathcal{M}_s^+(S)$. From Theorem 2.3.8 we obtain that $\mathcal{M}_s^+(S)$ is norm closed in $\mathrm{BL}(S)^*$, hence $\mathcal{S}_{\mathrm{BL}}^+ \subset \mathcal{M}_s^+(S)$. If S is not complete, then by Theorem 2.3.8, $\mathcal{M}_s^+(S)$ is not norm closed in $\mathrm{BL}(S)^*$, which implies that $\mathcal{M}_s^+(S) \subsetneq \mathcal{S}_{\mathrm{BL}}^+$. □

The crucial observation towards identification of $\mathcal{S}_{\mathrm{BL}}$ is the following:

Corollary 2.3.10. *$\mathcal{M}_s(S)$ is a $\|\cdot\|_{\mathrm{BL}}^*$-dense subspace of $\mathcal{S}_{\mathrm{BL}}$.*

A well-studied topology on the space of measures $\mathcal{M}_s(S)$ is the weak topology $\sigma(\mathcal{M}_s(S), C_b(S))$. In general this topology need not be metrisable on $\mathcal{M}_s(S)$. However, its restriction to $\mathcal{M}_s^+(S)$ *is* metrisable, and in fact it follows from [28, Theorem 18] that:

Theorem 2.3.11. *The restriction of $\sigma(\mathcal{M}_s(S), C_b(S))$ to $\mathcal{M}_s^+(S)$ equals the restriction of the norm topology of $\mathcal{S}_{\mathrm{BL}}$ to $\mathcal{M}_s^+(S)$.*

In particular, the following holds:

Lemma 2.3.12. *Let $\mu_n, \mu \in \mathcal{M}_s^+(S)$. Then $\|\mu_n - \mu\|_{\mathrm{BL}}^* \to 0$ if and only if $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f \in C_b(S)$.*

One might ask when $\mathcal{S}_{\mathrm{BL}} = \mathcal{M}_s(S)$. To answer this question we need the notion of a uniformly discrete metric space. S is *uniformly discrete* if there is an $\epsilon > 0$ such that $d(x, y) > \epsilon$ for all $x, y \in S$, $x \neq y$. The following theorem settles our question:

Theorem 2.3.13. *$\mathcal{M}_s(S)$ is norm closed in $\mathrm{BL}(S)^*$ if and only if S is uniformly discrete. In this case the norms $\|\cdot\|_{\mathrm{TV}}$ and $\|\cdot\|_{\mathrm{BL}}^*$ are equivalent.*

Proof. Suppose $\overline{\mathcal{M}_s(S)}^{\|\cdot\|^*_{\mathrm{BL}}} = \mathcal{M}_s(S)$. Then $(\mathcal{M}_s(S), \|\cdot\|^*_{\mathrm{BL}})$ is a Banach space. Let I be the identity map from $(\mathcal{M}_s(S), \|\cdot\|_{\mathrm{TV}})$ to $(\mathcal{M}_s(S), \|\cdot\|^*_{\mathrm{BL}})$. Then, since $\|\mu\|^*_{\mathrm{BL}} \leq \|\mu\|_{\mathrm{TV}}$, I is a bounded linear map. Clearly, I is bijective, hence by the Inverse Mapping Theorem the inverse of I is a bounded linear map. Assume S is not uniformly discrete, then there are $x_n, y_n \in S$, such that $0 < d(x_n, y_n) < \frac{1}{n}$. Let $\mu_n = \delta_{x_n} - \delta_{y_n}$. Then $\|\mu_n\|_{\mathrm{TV}} = 2$, while $\|\mu_n\|^*_{\mathrm{BL}} \leq d(x_n, y_n) < \frac{1}{n}$, for all $n \in \mathbb{N}$. This implies I^{-1} cannot be bounded, which gives us a contradiction. Hence S must be uniformly discrete.

Now suppose S is uniformly discrete. Then there is an $\epsilon > 0$ such that $d(x, y) > \epsilon$ for all $x, y \in S$, $x \neq y$. Let $\mu \in \mathcal{M}_s(S)$. Let $S = P \cup N$ be the Hahn decomposition of S corresponding to μ, then $\mu^+ = \mu|_P$ and $\mu^- = \mu|_N$. Define

$$f(x) := \begin{cases} \min(\epsilon/4, 1/2) & \text{if } x \in P; \\ -\min(\epsilon/4, 1/2) & \text{if } x \in N. \end{cases}$$

Then $\|f\|_\infty \leq 1/2$ and

$$|f|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} \leq \frac{\epsilon/2}{\epsilon} = \frac{1}{2}.$$

Hence $\|f\|_{\mathrm{BL}} \leq 1$. Furthermore,

$$\begin{aligned} \left| \int_S f \, d\mu \right| &= \left| \int_P \min(\epsilon/4, 1/2) \, d\mu - \int_N \min(\epsilon/4, 1/2) \, d\mu \right| \\ &= |\mu^+(S) + \mu^-(S)| \min(\epsilon/4, 1/2) = \|\mu\|_{\mathrm{TV}} \min(\epsilon/4, 1/2). \end{aligned}$$

Hence

$$\|\mu\|_{\mathrm{TV}} \leq \|\mu\|^*_{\mathrm{BL}} \frac{1}{\min(\epsilon/4, 1/2)},$$

for all $\mu \in \mathcal{M}_s(S)$. Also, $\|\mu\|^*_{\mathrm{BL}} \leq \|\mu\|_{\mathrm{TV}}$ for all $\mu \in \mathcal{M}_s(S)$, hence the norms $\|\cdot\|^*_{\mathrm{BL}}$ and $\|\cdot\|_{\mathrm{TV}}$ are equivalent on $\mathcal{M}_s(S)$. This implies that

$$\overline{\mathcal{M}_s(S)}^{\|\cdot\|^*_{\mathrm{BL}}} = \overline{\mathcal{M}_s(S)}^{\|\cdot\|_{TV}} = \mathcal{M}_s(S).$$

□

Remark. Note that all the arguments in the proof of Theorem 2.3.13 hold when we replace $\mathcal{M}_s(S)$ by $\mathcal{M}(S)$. Hence $\overline{\mathcal{M}(S)}^{\|\cdot\|^*_{\mathrm{BL}}} = \mathcal{M}(S)$ if and only if S is uniformly discrete.

Corollary 2.3.14. *If S is not uniformly discrete, there are elements in $\mathcal{S}_{\mathrm{BL}}$, hence in $\mathrm{BL}(S)^*$, that cannot be represented by a measure in $\mathcal{M}(S)$.*

2.3.2 Identification of $\mathcal{S}_{e,h}$

We start with the observation that each $\varphi \in \mathcal{S}_{e,h}$ is completely determined by its restriction to $\mathrm{BL}(S)$; more precise:

Lemma 2.3.15. *Let $\varphi \in \mathcal{S}_{e,h}$ and $f \in \mathrm{Lip}_{e,h}(S)$. Define*

$$f_n(x) := \max(\min(f(x), n), -n). \tag{2.11}$$

Then $\lim_{n\to\infty} \varphi(f_n) = \varphi(f)$.

Proof. Obviously, $\|f_n\|_{e,h} \leq \|f\|_{e,h}$ for all $n \in \mathbb{N}$. Let $\epsilon > 0$. Then there is a $d \in D$ (with D as in (2.10)) such that $\|\varphi - d\|^*_{e,h} < \frac{\epsilon}{2(\|f\|_{e,h}+1)}$. There exists N_d such that $d(f - f_n) = 0$ for all $n \geq N_d$. Then for $n \geq N_d$ we have

$$\begin{aligned} |\varphi(f) - \varphi(f_n)| &\leq |\varphi(f) - d(f)| + |d(f) - d(f_n)| + |\varphi(f_n) - d(f_n)| \\ &\leq 2\|\varphi - d\|^*_{e,h}\|f\|_{e,h} < \epsilon. \end{aligned}$$

Hence $\lim_{n\to\infty} \varphi(f_n) = \varphi(f)$. □

Recall the natural embedding $j_h : \mathrm{BL}(S) \hookrightarrow \mathrm{Lip}_{e,h}(S)$, and the adjoint $j_h^* : \mathrm{Lip}_{e,h}(S)^* \to \mathrm{BL}(S)^*$, given by restriction. Then, as a consequence of Proposition 2.2.8, j_h^* is not injective whenever S has infinite diameter. Consider however the restriction $\hat{j}_h^*$ of j_h^* to $\mathcal{S}_{e,h}$.

Lemma 2.3.16. *$\hat{j}_h^*$ maps $\mathcal{S}_{e,h}$ injectively and densely into $\mathcal{S}_{\mathrm{BL}}$.*

Proof. Let $\varphi \in \mathcal{S}_{e,h}$ be such that $\hat{j}_h^*(\varphi) = 0$. Then $\varphi(f) = 0$ for all $f \in \mathrm{BL}(S)$. Lemma 2.3.15 implies that $\varphi(f) = 0$ for all $f \in \mathrm{Lip}_{e,h}(S)$, hence $\varphi = 0$. So $\hat{j}_h^*$ is injective. By continuity of $\hat{j}_h^*$,

$$\hat{j}_h^*(\mathcal{S}_{e,h}) = \hat{j}_h^*(\overline{D}^{\|\cdot\|^*_{e,h}}) \subset \overline{\hat{j}_h^*(D)}^{\|\cdot\|^*_{\mathrm{BL}}} = \overline{D}^{\|\cdot\|^*_{\mathrm{BL}}} = \mathcal{S}_{\mathrm{BL}}.$$

So we can continuously embed $\mathcal{S}_{e,h}$ into $\mathcal{S}_{\mathrm{BL}}$ and $\hat{j}_h^*(\mathcal{S}_{e,h})$ is dense in $\mathcal{S}_{\mathrm{BL}}$, since $\hat{j}_h^*(D) = D$ is dense in $\mathcal{S}_{\mathrm{BL}}$. □

Just as before, we restrict to the separable Borel measures: Let $\mathcal{M}_{s,h}(S) := \mathcal{M}_h(S) \cap \mathcal{M}_s(S)$, and $\mathcal{M}^+_{s,h}(S) := \mathcal{M}_h(S) \cap \mathcal{M}^+_s(S)$. Similar to $\mathcal{S}^+_{\mathrm{BL}}$, we define $\mathcal{S}^+_{e,h}$ to be the closure of D^+ with respect to $\|\cdot\|^*_{e,h}$ and $\mathcal{S}^+_e$ to be the closure of D^+ with respect to $\|\cdot\|^*_e$.

Now we can prove the analogue to Theorem 2.3.9:

Theorem 2.3.17. *$\mathcal{M}^+_{s,h} \subset \mathcal{S}^+_{e,h}$. Furthermore, $\mathcal{M}^+_{s,h} = \mathcal{S}^+_{e,h}$ if and only if S is complete.*

Proof. First we show that $\mathcal{M}^+_{s,h}(S) \subset \mathcal{S}^+_{e,h}$. To that end, let $\mu \in \mathcal{M}^+_{s,h}(S)$, and let E be a measurable separable subset of S on which μ is concentrated. We want to show there is an element $\varphi \in \mathcal{S}^+_{e,h}$ such that $\varphi(f) = \int_S f\, d\mu$ for all $f \in \mathrm{Lip}_{e,h}(S)$. If $\mu(S) = 0$ this is clear, so we assume $\mu(S) > 0$.

We define the map $\delta : S \to \mathrm{Lip}_{e,h}(S)^*, x \mapsto \delta_x$. Then δ is locally Lipschitz continuous, hence continuous, by Lemma 2.3.5. For all $f \in \mathrm{Lip}_{e,h}(S)$, the function $x \mapsto \delta_x(f) = f(x)$ is continuous, hence Borel measurable. Also, since E is separable and δ is continuous, $\delta(E)$ is a separable subset of $\mathrm{BL}(S)^*$. Because $\mu(S \backslash E) = 0$, δ is μ-separably valued. Hence by the Pettis Measurability Theorem, δ is strongly μ-measurable. Furthermore, by Lemma 2.3.5, we have for $x, y \in \overline{B_e(\theta)}$:

$$|\|\delta_x\|^*_{e,h} - \|\delta_y\|^*_{e,h}| \leq \|\delta_x - \delta_y\|^*_{e,h} \leq h(\theta) d(x, y).$$

The map $\gamma : x \mapsto \|\delta_x\|^*_{e,h}$ therefore is in $\mathrm{Lip}_{e,h}(S)$, with $\|\gamma\|_{e,h} \leq 2$.

Consequently, $\delta : S \to \mathrm{Lip}_{e,h}(S)^*$ is μ-Bochner integrable and $\int_S \delta_x\, d\mu(x)$ defines an element in $\mathrm{Lip}_{e,h}(S)^*$. Similar as in the proof of Theorem 2.3.9 we obtain that $\int_S \delta_x\, d\mu = \mu$ is contained in $\mathcal{S}^+_{e,h}$. Thus $\mathcal{M}^+_{s,h}(S) \subset \mathcal{S}^+_{e,h}$.

Now suppose that S is complete. It is clear that $D^+ \subset \mathcal{M}^+_{s,h}(S)$. Let $\varphi \in \mathcal{S}^+_{e,h}$, then there are $d_n \in D_+$ such that $\|\varphi - d_n\|^*_{e,h} \to 0$. Hence, by Lemma 2.3.16, $\|\varphi - d_n\|^*_{\mathrm{BL}} \leq \|\varphi - d_n\|^*_{e,h} \to 0$, so $\varphi \in \mathcal{M}^+_s(S)$ by Theorem 2.3.8 . Hence there is a $\mu \in \mathcal{M}^+_s(S)$ such that $\varphi(f) = \int_S f\, d\mu$ for all $f \in \mathrm{BL}(S)$. We need to show that $\mu \in \mathcal{M}^+_{s,h}(S)$ and $\varphi(f) = \int_S f\, d\mu$ for all $f \in \mathrm{Lip}_{e,h}(S)$.

Let $f \in \mathrm{Lip}_{e,h}(S)$, $f \geq 0$. Define f_n as in (2.11). Then

$$\varphi(f) = \lim_{n\to\infty} \varphi(f_n) = \lim_{n\to\infty} \int_S f_n\, d\mu = \int_S f\, d\mu < \infty$$

by Lemma 2.3.15 and the Monotone Convergence Theorem. Using $f = f^+ - f^-$ for general $f \in \mathrm{Lip}_{e,h}(S)$, where $f^+ = \max(f, 0)$ and $f^- = -\min(f, 0)$, we find that $f \in L^1(\mu)$ and $\varphi(f) = \int_S f\, d\mu$ for every $f \in \mathrm{Lip}_{e,h}(S)$.

In particular,

$$\int_S |f|\, d\mu = \varphi(|f|) \leq \||f|\|_{e,h} \|\varphi\|^*_{e,h} \leq \|f\|_{e,h} \|\varphi\|^*_{e,h},$$

hence $\mu \in \mathcal{M}^+_{s,h}(S)$ and we obtain $\mathcal{S}^+_{e,h} \subset \mathcal{M}^+_{s,h}(S)$.

Now suppose S is not complete. Then there is a Cauchy sequence $(x_n)_n$ in S that does not converge to an element in S. This sequence must be bounded in S. This implies by Lemma 2.3.5 that $(\delta_{x_n})_n$ is a Cauchy sequence in $\mathrm{Lip}_{e,h}(S)^*$. Suppose that $\mu \in \mathcal{M}^+_{s,h}(S)$ is such that $\|\delta_{x_n} - \mu\|^*_{e,h} \to 0$. Then $\|\delta_{x_n} - \mu\|^*_{\mathrm{BL}} \to 0$, but from the proof of Theorem 2.3.8 it follows that this is not possible. Hence $\mathcal{M}^+_{s,h}(S)$ is not norm closed in $\mathrm{Lip}_{e,h}(S)$, and since $\mathcal{M}^+_{s,h}(S) \subset \mathcal{S}^+_{e,h}$, this implies that $\mathcal{M}^+_{s,h}(S) \subsetneq \mathcal{S}^+_{e,h}$. $\square$

The following corollaries follow easily from Theorem 2.3.17:

Corollary 2.3.18. *$\mathcal{M}_{s,h}(S)$ is a $\|\cdot\|^*_{e,h}$-dense subspace of $\mathcal{S}_{e,h}$.*

Corollary 2.3.19. *$\mathcal{M}^+_{s,h}(S)$ is norm closed in $\mathcal{S}_{e,h}$ if and only if S is complete.*

Remark 2.3.20. In [74, Theorem 4.2] it is shown that the metric space S is complete if and only if the set of separable probability measures of finite first moment, $\mathcal{P}_{s,1}(S)$, is complete with respect to the metric H, where

$$H(\mu,\nu) = \sup_{\substack{f\in \mathrm{Lip}(S)\\ |f|_{\mathrm{Lip}}\leq 1}} \left| \int_S f\, d\mu - \int_S f d\nu \right|.$$

From Corollary 2.3.19 we can also conclude this theorem: it follows that when S is complete, the subset of separable probability measures of finite first moment $\mathcal{P}_{s,1}(S)$ is also a closed set of $\mathcal{S}_e$, hence complete with respect to $\|\cdot\|^*_e$. Let $\mu,\nu \in \mathcal{P}_{s,1}(S)$, then $\|\mu-\nu\|^*_e$ is equal to $H(\mu,\nu)$, since for $f \in \mathrm{Lip}(S)$ with $|f|_{\mathrm{Lip}} \leq 1$ we have

$$\left| \int_S f\, d\mu - \int_S f\, d\nu \right| = \left| \int_S f - f(e)\mathbb{1}\, d\mu - \int_S f - f(e)\mathbb{1}\, d\mu \right|,$$

and $g(x) := f(x) - f(e)$ satisfies: $\|g\|_e = |f|_{\mathrm{Lip}} \leq 1$. Furthermore, when S is not complete, $\mathcal{M}^+_{s,1}(S)$ is not complete with respect to $\|\cdot\|^*_e$, which implies that $\mathcal{P}_{s,1}(S)$ also cannot be complete with respect to $\|\cdot\|^*_e$, hence it is not complete with respect to H. It follows from the Kantorovich-Rubinstein Duality Theorem that H, and thus the metric induced by $\|\cdot\|^*_e$, equals the Wasserstein-1 metric (see e.g. [30, Theorem 11.8.2]).

2.3.3 $\mathcal{S}_{e,h}$ and $\mathcal{S}_{\mathrm{BL}}$ as predual of $\mathrm{Lip}_{e,h}(S)$ and $\mathrm{BL}(S)$.

Various spaces of Lipschitz functions have been shown to be isometrically isomorphic to the dual of a Banach space. For instance, $\mathrm{Lip}_0(S)$ is the dual of the so-called Arens-Eells space (see [5] and [122, Section 2.2]). It is also known that $\mathrm{BL}(S)$ endowed with the norm $\|f\|_{\mathrm{BL,max}} = \max(\|f\|_\infty, |f|_{\mathrm{Lip}})$ is isometrically isomorphic to the dual of a Banach space. For instance in [68, Theorem 4.1] the more general result is proven for $\mathrm{BL}(S, E^*)$, where E^* is the dual of a Banach space. Our aim in this section is to show that $\mathrm{BL}(S)$ is the dual of $\mathcal{S}_{\mathrm{BL}}$ and that $\mathrm{Lip}_{e,h}(S)$ is also the dual of a Banach space, namely $\mathcal{S}_{e,h}$.

Theorem 2.3.21. *$\mathcal{S}^*_{e,h}$ is isometrically isomorphic to $\mathrm{Lip}_{e,h}(S)$ under the map $\psi \mapsto T\psi$, where $T\psi(x) := \psi(\delta_x)$.*

Proof. We define $R : \mathrm{Lip}_{e,h}(S) \to \mathcal{S}^*_{e,h}$ such that $Rf(\varphi) := \varphi(f)$ for all $\varphi \in \mathcal{S}_{e,h} \subset \mathrm{Lip}_{e,h}(S)^*$. Clearly $|Rf(\varphi)| = |\varphi(f)| \leq \|\varphi\|^*_{e,h}\|f\|_{e,h}$, hence $\|Rf\|_{\mathcal{S}^*_{e,h}} \leq \|f\|_{e,h}$.

Now define $T : \mathcal{S}^*_{e,h} \to \mathrm{Lip}_{e,h}(S)$ such that $T\psi(x) := \psi(\delta_x)$ for all $x \in S$. It can easily be verified that $T\psi$ is indeed in $\mathrm{Lip}_{e,h}(S)$, and that T is linear. Now we want

to show that $\|T\psi\|_{e,h} \leq \|\psi\|_{\mathcal{S}^*_{e,h}}$. Let $\theta \geq 0$, $x, y \in \overline{B_e(\theta)}, x \neq y$. Then

$$\begin{aligned}
|\psi(\delta_e)| + \frac{|\psi(\delta_x - \delta_y)|}{h(\theta)d(x,y)} &\leq \max\left(\left|\psi(\delta_e) - \frac{\psi(\delta_x - \delta_y)}{h(\theta)d(x,y)}\right|, \left|\psi(\delta_e) + \frac{\psi(\delta_x - \delta_y)}{h(\theta)d(x,y)}\right|\right) \\
&\leq \max\left(\left|\psi(\delta_e - \frac{\delta_x - \delta_y}{h(\theta)d(x,y)})\right|, \left|\psi(\delta_e + \frac{\delta_x - \delta_y}{h(\theta)d(x,y)})\right|\right) \\
&\leq \|\psi\|_{\mathcal{S}^*_{e,h}} \max\left(\left\|\delta_e - \frac{\delta_x - \delta_y}{h(\theta)d(x,y)}\right\|^*_e, \left\|\delta_e + \frac{\delta_x - \delta_y}{h(\theta)d(x,y)}\right\|^*_e\right)
\end{aligned}$$

Now for all $f \in \mathrm{Lip}_{e,h}(S)$, with $\|f\|_{e,h} \leq 1$, we have

$$\begin{aligned}
\left|(\delta_e - \frac{\delta_x - \delta_y}{h(\theta)d(x,y)})(f)\right| &= \left|f(e) - \frac{f(x) - f(y)}{h(\theta)d(x,y)}\right| \\
&\leq |f(e)| + \frac{|f(x) - f(y)|}{h(\theta)d(x,y)} \\
&\leq |f(e)| + \frac{|f|_{\mathrm{Lip},\theta}}{h(\theta)} \leq 1.
\end{aligned}$$

So $\|\delta_e - \frac{\delta_x - \delta_y}{h(\theta)d(x,y)}\|^*_{e,h} \leq 1$. Since we can reverse x and y, we also obtain that $\|\delta_e + \frac{\delta_x - \delta_y}{h(\theta)d(x,y)}\|^*_{e,h} \leq 1$ Hence we can conclude that for all $x, y \in \overline{B_e(\theta)}$, $x \neq y$, we have

$$|\psi(\delta_e)| + \frac{|\psi(\delta_x) - \psi(\delta_y)|}{h(\theta)d(x,y)} \leq \|\psi\|_{\mathcal{S}^*_{e,h}}.$$

Hence for all $\theta \geq 0$:

$$|\psi(\delta_e)| + \sup_{\substack{x,y \in \overline{B_e(\theta)} \\ x \neq y}} \frac{|\psi(\delta_x - \delta_y)|}{h(\theta)d(x,y)} \leq \|\psi\|_{\mathcal{S}^*_{e,h}},$$

thus

$$\|T\psi\|_{e,h} \leq \|\psi\|_{\mathcal{S}^*_{e,h}}, \text{ for all } \psi \in \mathcal{S}^*_{e,h}.$$

Now we need to show R and T are each others' inverses. Let $f \in \mathrm{Lip}_{e,h}(S)$, then

$$T(Rf)(x) = Rf(\delta_x) = f(x), \text{ for all } x \in S.$$

Hence $T \circ R = \mathrm{Id}_{\mathrm{Lip}_{e,h}(S)}$. Now let $\psi \in \mathcal{S}^*_{e,h}$, and let $d \in D$, then $d = \sum_{k=1}^n \alpha_k \delta_{x_k}$, for certain $\alpha_k \in \mathbb{R}$ and $x_k \in S$. Then

$$R(T\psi)(d) = \sum_{k=1}^n \alpha_k T\psi(x_k) = \sum_{k=1}^n \alpha_k \psi(\delta_{x_k}) = \psi(d).$$

Hence $R(T\psi) = \psi$ on a dense subset of $\mathcal{S}_{e,h}$, so $R(T\psi) = \psi$ on $\mathcal{S}_{e,h}$. Hence $R \circ T = \mathrm{Id}_{\mathcal{S}^*_{e,h}}$. Consequently we get that for all $f \in \mathrm{Lip}_{e,h}(S) : \|Rf\|^*_{e,h} \leq \|f\|_{e,h} =$

$\|T(Rf)\|_{e,h} \leq \|Rf\|^*_{e,h}$, hence R is an isometric isomorphism from $\mathrm{Lip}_{e,h}(S)$ to $\mathcal{S}^*_{e,h}$, with T as its inverse.

□

A similar result holds for $\mathrm{BL}(S)$:

Theorem 2.3.22. *$\mathcal{S}^*_{\mathrm{BL}}$ is isometrically isomorphic to $\mathrm{BL}(S)$ under the map $\psi \mapsto T\psi$, where $T\psi(x) := \psi(\delta_x)$.*

Proof. We define $R : \mathrm{BL}(S) \to \mathcal{S}^*_{\mathrm{BL}}$ such that $Rf(\varphi) := \varphi(f)$ for all $\varphi \in \mathcal{S}_{\mathrm{BL}} \subset \mathrm{BL}(S)^*$. And we define $T : \mathcal{S}^*_{\mathrm{BL}} \to \mathrm{BL}(S)$ such that $T\psi(x) := \psi(\delta_x)$ for all $x \in S$. Then analogous to the proof of Theorem 2.3.21 we can show that $\|Rf\|_{\mathcal{S}^*_{\mathrm{BL}}} \leq \|f\|_e$, that $\|T\psi\|_{\mathrm{BL}} \leq \|\psi\|_{\mathcal{S}^*_{\mathrm{BL}}}$ and that R and T are each others' inverses. Hence R is an isometric isomorphism from $\mathrm{BL}(S)$ to $\mathcal{S}^*_{\mathrm{BL}}$, with T as its inverse. □

2.3.4 A result on weak convergence in $\mathcal{S}_{\mathrm{BL}}$

Let $C_{ub}(S)$ denote the Banach space of bounded uniformly continuous real-valued functions on S, with the supremum norm $\|\cdot\|_\infty$. A Borel measure μ on S is tight if for every $\epsilon > 0$ there exists a compact $K \subset S$ such that $|\mu|(S \backslash K) < \epsilon$. Pachl [92] considered the space of tight Borel measures $\mathcal{M}_t(S)$ as subspace of $C_{ub}(S)^*$ and defined the following norm on $\mathcal{M}_t(S)$:

$$\|\mu\|_d = \sup\{|\langle \mu, f\rangle| : f \in \mathrm{BL}(S) \text{ such that } \|f\|_\infty \leq 1 \text{ and } |f|_{\mathrm{Lip}} \leq 1\}.$$

This norm also makes sense on the larger space $\mathcal{M}(S)$ and clearly $\|\mu\|^*_{\mathrm{BL}} \leq \|\mu\|_d \leq 2\|\mu\|^*_{\mathrm{BL}}$ for all $\mu \in \mathcal{M}(S)$, so $\|\cdot\|_d$ and $\|\cdot\|^*_{\mathrm{BL}}$ are equivalent.

In order to state Pachl's main result, we first need some definitions: A subset M of a topological space T is *relatively compact* if its closure in T is compact. M is *sequentially compact* (respectively, *relatively sequentially compact*) if every sequence in M has a subsequence convergent to a point in M (respectively, to a point in T) and M is *countably compact* (respectively, *relatively countably compact*) if every sequence in M has a cluster point in M (respectively, in T). If T is a metric space, the concepts of compactness, sequential compactness and countable compactness coincide. In general this need not be the case, but both compactness and sequential compactness imply countable compactness.

Pachl proved the following result [92, Theorem 3.2]:

Theorem 2.3.23. *Let S be complete.*

(a) Let $(\mu_n)_n \subset \mathcal{M}_t(S)$ be such that $(\langle \mu_n, g\rangle)_n$ converges as $n \to \infty$ for all $g \in C_{ub}(S)$. Then there is a $\mu \in \mathcal{M}_t(S)$ for which

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

(b) *Let a set $M \subset \mathcal{M}_t(S)$ be bounded on the unit ball of $C_{ub}(S)$. The following conditions are equivalent:*

(i) *M is relatively $\|\cdot\|_d$-compact.*

(ii) *M is relatively $\sigma(\mathcal{M}_t(S), C_{ub}(S))$-countably compact.*

From the corollary of [28, Theorem 2] it follows that

$$\mathcal{M}_s(S) = \mathcal{M}_t(S)$$

whenever S is complete.

From Theorem 2.3.23 we obtain the following result, which has some interesting consequences and will play an important role in Chapter 7.

Theorem 2.3.24. *Let S be complete. Let $(\mu_n)_n \subset \mathcal{M}_s(S)$ and $N \geq 0$ be such that $\langle \mu_n, f\rangle$ converges as $n \to \infty$ for every $f \in \mathrm{BL}(S) \cong \mathcal{S}^*_{\mathrm{BL}}$ and*

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

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

Proof. Observe that the dual norm on $C_{ub}(S)^*$ of an element in $\mathcal{M}_s(S)$ equals the total variation norm of that element. By Theorem 2.3.23 and the equivalence of $\|\cdot\|^*_{\mathrm{BL}}$ and $\|\cdot\|_d$ it follows that the set $\{\mu_n : n \in \mathbb{N}\}$ is relatively $\|\cdot\|^*_{\mathrm{BL}}$-compact whenever $\langle \mu_n, g\rangle$ converges as $n \to \infty$ for every $g \in C_{ub}(S)$.

Fix $g \in C_{ub}(S)$ and $\epsilon > 0$. By [28, Lemma 8] $\mathrm{BL}(S)$ is dense in $C_{ub}(S)$, so there exists an $f \in \mathrm{BL}(S)$ with $\|f - g\|_\infty \leq \epsilon$. For every $m, n \in \mathbb{N}$

$$\begin{aligned} |\langle \mu_m, g\rangle - \langle \mu_n, g\rangle| &\leq |\langle \mu_m, g - f\rangle| + |\langle \mu_n, g - f\rangle| + |\langle \mu_m, f\rangle - \langle \mu_n, f\rangle| \\ &\leq 2N\epsilon + |\langle \mu_m, f\rangle - \langle \mu_n, f\rangle|. \end{aligned}$$

Since $\langle \mu_n, f\rangle$ converges as $n \to \infty$, this implies that $(\langle \mu_n, g\rangle)_n$ is Cauchy and thus converges for every $g \in C_{ub}(S)$. So by Theorem 2.3.23 (a) there is a $\mu \in \mathcal{M}_s(S)$ such that

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

for every $g \in C_{ub}(S)$.

Then $\{\mu_n : n \in \mathbb{N}\}$ is indeed relatively $\|\cdot\|^*_{\mathrm{BL}}$-compact in $\mathcal{S}_{\mathrm{BL}}$. Suppose that μ_n does not converge to μ in $\mathcal{S}_{\mathrm{BL}}$, then there is a subsequence $(\mu_{n_k})_k$ and an $\epsilon > 0$ such that $\|\mu_{n_k} - \mu\|^*_{\mathrm{BL}} \geq \epsilon$ for every $k \in \mathbb{N}$. By relative compactness, there exists a subsequence of $(\mu_{n_k})_k$, that we will also denote by $(\mu_{n_k})_k$, that converges in $\mathcal{S}_{\mathrm{BL}}$, say to $\varphi \in \mathcal{S}_{\mathrm{BL}}$. Then for every $f \in \mathrm{BL}(S)$

$$\langle \mu, f\rangle = \lim_{k\to\infty} \langle \mu_{n_k}, f\rangle = \varphi(f).$$

Since an element in $\mathcal{S}_{\mathrm{BL}}$ is determined uniquely by pairings with all elements in $\mathrm{BL}(S) = \mathcal{S}^*_{\mathrm{BL}}$, $\varphi = \mu$. Thus $\|\mu_{n_k} - \mu\|^*_{\mathrm{BL}} \to 0$ which gives a contradiction. □

Theorem 2.3.24 yields a characterisation of the elements in $\mathcal{S}_{\mathrm{BL}}$ that can be represented by measures:

Theorem 2.3.25. *Let S be complete and $\varphi \in \mathcal{S}_{\mathrm{BL}}$. Then $\varphi \in \mathcal{M}_s(S)$ if and only if there is an $M > 0$ such that $|\varphi(f)| \leq M\|f\|_\infty$ for all $f \in \mathrm{BL}(S)$.*

Proof. Suppose $\varphi = \mu \in \mathcal{M}(S)$. Then for every $f \in \mathrm{BL}(S)$,

$$|\varphi(f)| = \left|\int_S f\,d\mu\right| \leq \|f\|_\infty \|\mu\|_{\mathrm{TV}}.$$

Now let $\varphi \in \mathcal{S}_{\mathrm{BL}}$ and $M > 0$ such that $|\varphi(f)| \leq M\|f\|_\infty$ for all $f \in \mathrm{BL}(S)$.

Let $B := \{\mu \in \mathcal{M}_s(S) : \|\mu\|_{\mathrm{TV}} \leq M\}$. We will show that B is closed in $\mathcal{S}_{\mathrm{BL}}$: let $\psi \in \mathcal{S}_{\mathrm{BL}}$ and $(\mu_n)_n \subset B$ be such that $\|\mu_n - \psi\|^*_{\mathrm{BL}} \to 0$ as $n \to \infty$. By Theorem 2.3.24 this implies that there is a $\mu \in \mathcal{M}(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$, i.e. $\psi = \mu$, which implies that $\mu \in \mathcal{M}_s(S)$. Now let $f \in \mathrm{BL}(S)$, then

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

Since $\mathrm{BL}(S)$ is dense in $C_{ub}(S)$, this statement holds for all $f \in C_{ub}(S)$, so in particular $\|\mu\|_{\mathrm{TV}} \leq M$. Thus B is a closed convex subset of $\mathcal{S}_{\mathrm{BL}}$.

Suppose that $\varphi \notin B$. Then φ is *strictly separated* from B by [17, Corollary IV.3.10]: there is an $f \in \mathcal{S}^*_{\mathrm{BL}} = \mathrm{BL}(S)$ and an $\alpha \in \mathbb{R}$, such that $\langle \mu, f\rangle < \alpha$ for all $\mu \in B$, and $\varphi(f) > \alpha$. For every $x \in S$, both $M\delta_x$ and $-M\delta_x$ are in B. Thus $Mf(x) = \langle M\delta_x, f\rangle < \alpha$ and $-Mf(x) < \alpha$. Thus $\alpha > 0$ and $\|f\|_\infty \leq \frac{\alpha}{M}$. Thus $\varphi(f) \leq |\varphi(f)| \leq M\|f\|_\infty \leq \alpha$, which is a contradiction. So $\varphi \in B \subset \mathcal{M}(S)$. □

A natural question that arises from Theorem 2.3.24 is whether the uniform bound on the total variation norm of the measures $(\mu_n)_n$ is necessary. Suppose we could remove that condition, and still have the same conclusion as Theorem 2.3.24 for a complete metric space S. This would imply the *Schur property* for $\mathcal{S}_{\mathrm{BL}}$, i.e. any weakly convergent sequence is norm convergent. E.g. ℓ^1 has the Schur property [1, Theorem 2.3.6]. In fact, if $(\varphi_n)_n \subset \mathcal{S}_{\mathrm{BL}}$ is such that there exists $\varphi \in \mathcal{S}_{\mathrm{BL}}$ to which $(\varphi_n)_n$ converges weakly, then by density of $\mathcal{M}_s(S)$ in $\mathcal{S}_{\mathrm{BL}}$ there exists a sequence $(\mu_n)_n \subset \mathcal{M}_s(S)$ such that $\mu_n \to \varphi$ weakly in $\mathcal{S}_{\mathrm{BL}}$. The modified version of Theorem 2.3.24 would imply that $\mu_n \to \varphi$ in $\mathcal{S}_{\mathrm{BL}}$. Hence $\varphi_n \to \varphi$ in $\mathcal{S}_{\mathrm{BL}}$.

We now give an example and sketch a proof showing that in general the Schur property will not hold for $\mathcal{S}_{\mathrm{BL}}$, not even when S is a compact metric space.

Example. Let $S = [0,1]$ with the Euclidean metric. Let $\mu_n = n\sin(2\pi nx)\,dx$. Let $g \in \mathrm{BL}(S)$ be such that $\|g\|_{\mathrm{BL}} \leq 1$. Rademacher's Theorem (see e.g. [89]) yields that g is Lebesgue almost everywhere differentiable. Since $|g|_{\mathrm{Lip}} \leq 1$, this implies that there is a $f \in L^\infty([0,1])$ such that for all $0 \leq a < b \leq 1$

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

From this one can show that

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

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

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

Thus $\lim_{n\to\infty}\langle \mu_n, g\rangle = 0$ for all $g \in \mathrm{BL}(S)$. For $n \in \mathbb{N}$ define $g_n(0) = g_n(1) = 0$,

$$g_n(\tfrac{1+4i}{4n}) = \tfrac{1}{4n} \text{ for } i \in \{0,1,...,n-1\},$$

$$g_n(\tfrac{3+4i}{4n}) = \tfrac{-1}{4n} \text{ for } i \in \{0,1,...,n-1\},$$

and let g_n be the piecewise linear function whose graph consists of linear segments between these points. Then $|g_n|_{\mathrm{Lip}} = 1$ and $\|g_n\|_\infty = \frac{1}{4n}$, so $\|g_n\|_{\mathrm{BL}} = 1 + \frac{1}{4n}$. An easy calculation shows that $\langle \mu_n, g_n\rangle = \frac{1}{\pi^2}$ for all $n \in \mathbb{N}$, so in particular $\|\mu_n\|^*_{\mathrm{BL}}$ does not converge to zero as $n \to \infty$.

2.4 Positive functionals on Lipschitz spaces

We can endow $\mathrm{BL}(S)$ and $\mathrm{Lip}_{e,h}(S)$ with pointwise ordering, so $f \geq g$ if $f(x) \geq g(x)$ for all $x \in S$. From Lemma 2.2.2 it follows that $\mathrm{BL}(S)$ and $\mathrm{Lip}_{e,h}(S)$ are Riesz spaces with respect to this ordering. However, $\|\cdot\|_{\mathrm{BL}}$ and $\|\cdot\|_{e,h}$ are not Riesz norms, since $|f| \leq |g|$ need not imply that $|f|_{\mathrm{Lip}} \leq |g|_{\mathrm{Lip}}$. We are interested in the question which positive functionals in

$$\mathrm{BL}(S)^*_+ := \{\phi \in \mathrm{BL}(S)^* : \phi(f) \geq 0 \text{ for all } f \in \mathrm{BL}(S), f \geq 0\}$$

can be represented by measures on S.

We obtained in Corollary 2.3.14 that whenever S is not uniformly discrete, there are $\varphi \in \mathcal{S}_{\mathrm{BL}} \subset \mathrm{BL}(S)^*$ that cannot be represented by measures. However, since $\mathrm{BL}(S)$ is not a normed Riesz space, $\mathrm{BL}(S)^*$ need not be a Riesz space at all, and in particular it may be the case that the vector space generated by $\mathrm{BL}(S)^*_+$ is a proper subspace of $\mathrm{BL}(S)^*$. So it still might be true that $\mathrm{BL}(S)^*_+$ consists of measures.

We first make a few observations: if ϕ is a *positive* linear functional on $\mathrm{BL}(S)$ (not necessarily continuous), then for every $f \in \mathrm{BL}(S)$ we obtain

$$|\phi(f)| = |\phi(f^+) - \phi(f^-)| \leq \phi(|f|) \leq \phi(\|f\|_\infty \mathbb{1}) = \|f\|_\infty \phi(\mathbb{1}). \tag{2.12}$$

$\mathrm{BL}(S) \subset C_{ub}(S)$ is dense [28, Lemma 8]. According to (2.12) any positive linear functional on $\mathrm{BL}(S)$ extends to a (positive) continuous linear functional on $C_{ub}(S)$. Moreover, any linear functional ϕ on $\mathrm{BL}(S)$ that satisfies

$$|\phi(f)| \leq M\|f\|_\infty \text{ for every } f \in \mathrm{BL}(S) \tag{2.13}$$

for some $M > 0$ extends to a bounded linear functional on $C_{ub}(S)$. On the other hand, given a bounded linear functional ψ on $C_{ub}(S)$, we obtain for every $f \in \mathrm{BL}(S)$ that

$$|\psi(f)| \leq \|\psi\|\|f\|_\infty \leq \|\psi\|\|f\|_{\mathrm{BL}}.$$

So the restriction of ψ to $\mathrm{BL}(S)$ defines an element of $\mathrm{BL}(S)^*$. This gives a bijective correspondence between $C_{ub}(S)^*$ and those elements of $\mathrm{BL}(S)^*$ satisfying (2.13). Thus we have

Proposition 2.4.1. *$C_{ub}(S)^*$ embeds continuously into $\mathrm{BL}(S)^*$ by restriction. With this identification, $C_{ub}(S)^*_+ = \mathrm{BL}(S)^*_+$.*

In view of Proposition 2.4.1 we may rephrase Theorem 2.3.25 as

Corollary 2.4.2. *Let S be complete. Then $C_{ub}(S)^* \cap \mathcal{S}_{\mathrm{BL}} = \mathcal{M}_s(S)$.*

The following result characterises the situation in which *all* positive linear functionals can be represented by measures:

Theorem 2.4.3. *Every $\phi \in \mathrm{BL}(S)^*_+$ can be represented by a $\mu \in \mathcal{M}^+(S)$ if and only if S is compact.*

Proof. If S is not complete, the proof of Theorem 2.3.8 can be used to show that $\mathcal{M}^+(S)$ is not closed in $\mathrm{BL}(S)^*$. Since $\mathcal{M}^+(S) \subset \mathrm{BL}(S)^*_+$ and $\mathrm{BL}(S)^*_+$ *is* closed in $\mathrm{BL}(S)^*$, this implies that $\mathcal{M}^+(S) \subsetneq \mathrm{BL}^*_+(S)$.

So from now on we assume that S is complete. As we stated above, we can uniquely extend every $\phi \in \mathrm{BL}(S)^*_+$ to an element of $C_{ub}(S)^*_+$ and we obtain all of $C_{ub}(S)^*_+$ in this way. It is not hard to show that if $\phi \in \mathrm{BL}(S)^*_+$ is represented by a measure μ, then its extension $\bar{\phi}$ satisfies: $\bar{\phi}(f) = \int_S f\, d\mu$ for every $f \in C_{ub}(S)$. So it remains to show when all of $C_{ub}(S)^*_+$ can be represented by measures.

Since $C_{ub}(S)$ is a Banach lattice, hence a normed Riesz space, [85, Proposition 1.3.7] implies that $C_{ub}(S)^*$ is a Banach lattice, and in particular $C_{ub}(S)^* = C_{ub}(S)^*_+ - C_{ub}(S)^*_+$.

If S is compact, then $C_{ub}(S) = C(S)$, and by the Riesz Representation Theorem, every $\phi \in C_{ub}(S)^*$ can be represented by a measure.

If S is not compact, but locally compact, then $C_0(S) \subsetneq C_{ub}(S)$, since the constant function $\mathbb{1}$ is in $C_{ub}(S)$, but not in $C_0(S)$ (otherwise S would be compact). By the Hahn-Banach Theorem, we can find a non-trivial $\phi \in C_{ub}(S)^*$ such that $\phi|_{C_0(S)} \equiv 0$. Clearly this implies that ϕ cannot be represented by a measure in $\mathcal{M}(S)$, since this measure would have to be zero by the Riesz Representation Theorem.

Now suppose that S is not locally compact. We will prove the following claim:

Claim: There exists a sequence $(x_n)_n \subset S$ and an $\epsilon > 0$, such that $d(x_n, x_m) \geq \epsilon$ whenever $n \neq m$.
Proof of claim: Suppose that for every sequence $(x_n)_n \subset S$ and every $\epsilon > 0$, there is a subsequence $(x_{n_k})_k$ and an $x \in S$, such that $(x_{n_k})_k \subset B_x(\epsilon)$. Then we can

construct a subsequence which is Cauchy, hence convergent to some $z \in S$, and this would imply that S is compact. This is by assumption not possible, so there is a sequence $(x_n)_n \subset S$ and an $\epsilon > 0$ such that for every $n \in \mathbb{N}$ there are only finitely many $m \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$, and then we can easily extract a subsequence that satisfies the condition we want.

Now let $(x_n)_n \subset S$ be a sequence satisfying the condition in the claim, and let $T = \{x_n : n \in \mathbb{N}\}$. Then T is a closed subset of S that is locally compact, but not compact. Then there is a non-zero $\phi \in C_{ub}(T)^*$ such that $\phi(g) = 0$ for all $g \in C_0(T)$.

We can define $\psi : C_{ub}(S) \to \mathbb{R}$, by $\psi(f) = \phi(f|_T)$. Then ψ is non-zero and $|\psi(f)| \leq \|\phi\|\|f|_T\|_\infty \leq \|\phi\|\|f\|_\infty$. So $\psi \in C_{ub}(S)^*$. Suppose there exists a $\nu \in \mathcal{M}(S)$ such that $\psi(f) = \int_S f\, d\nu$ for all $f \in C_{ub}(S)$.

Fix $f \in C_{ub}(S)$ and define $f_n(x) := (1 - nd(x,T))^+ \cdot f(x)$ for $n \in \mathbb{N}$. Then $f_n \in C_{ub}(S)$, $f_n|_T = f|_T$ and $f_n \to 1\!\!1_T \cdot f$ pointwise. Thus the Dominated Convergence Theorem implies that

$$\begin{aligned} \int_T f\, d\nu &= \int_S 1\!\!1_T \cdot f\, d\nu = \lim_{n\to\infty} \int_S f_n\, d\nu \\ &= \lim_{n\to\infty} \psi(f_n) = \phi(f|_T) = \psi(f) = \int_S f\, d\nu. \end{aligned}$$

This holds for all $f \in C_{ub}(S)$, thus $\nu = \nu(\cdot \cap T)$. Since T is countable and ν is non-zero, there must be an $n \in \mathbb{N}$ for which $\nu(\{x_n\}) \neq 0$. Define $g := (1 - \frac{1}{\epsilon} d(x, x_n))^+$, then $g \in C_{ub}(S)$, $g(x_n) = 1$ and $g(x_m) = 0$ whenever $n \neq m$. Hence $g|_T \in C_0(T)$, so we obtain

$$\nu(\{x_n\}) = \int_T g\, d\nu = \phi(g|_T) = 0,$$

which is a contradiction. □

The closed convex cone $\mathcal{S}^+_{\mathrm{BL}}$ defines a partial ordering '$\geq$' on $\mathcal{S}_{\mathrm{BL}}$ by means of $\phi \geq \psi$ if and only if $\phi - \psi \in \mathcal{S}^+_{\mathrm{BL}}$. Then $(\mathcal{S}_{\mathrm{BL}}, \geq)$ is an ordered Banach space. In a similar fashion, $\mathcal{S}^+_{e,h}$ introduces a partial ordering on $\mathcal{S}_{e,h}$. Theorem 2.4.4, Theorem 2.4.6 and Proposition 2.4.8 below show that the various orderings are compatible.

Note that $\mathcal{S}^+_{\mathrm{BL}}$ is not a generating cone in $\mathcal{S}_{\mathrm{BL}}$, unless S is uniformly discrete (Theorem 2.3.13).

Theorem 2.4.4. $\mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)^*_+ = \mathcal{S}^+_{\mathrm{BL}}$.

Proof. Clearly, $\mathcal{S}^+_{\mathrm{BL}} \subset \mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)^*_+$. Suppose that there exists a $\phi \in \mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)^*_+$ such that $\phi \notin \mathcal{S}^+_{\mathrm{BL}}$. If $\phi(1\!\!1) = 0$, then $\phi(f) = 0$ for every $f \in \mathrm{BL}(S)$, by positivity of ϕ and (2.12), hence $\phi \in \mathcal{S}^+_{\mathrm{BL}}$. So $\phi(1\!\!1) > 0$. Let

$$M := \left\{ \sum_{i=1}^n \alpha_i \delta_{x_i} : n \in \mathbb{N}, 0 \leq \alpha_i \leq \phi(1\!\!1), x_i \in S, \text{ for } i = 1, ..., n \right\},$$

then $M \subset \mathcal{S}_{\mathrm{BL}}^{+}$. Let $\overline{M}$ be the closure of M in $\mathcal{S}_{\mathrm{BL}}^{+}$ with respect to $\|\cdot\|_{\mathrm{BL}}^{*}$. By assumption, ϕ is not in $\overline{M}$. Since M is convex, $\overline{M}$ is a closed convex subset of $\mathcal{S}_{\mathrm{BL}}$. Thus ϕ is strictly separated from $\overline{M}$ by [17, Corollary IV.3.10]: there is an $f \in \mathcal{S}_{\mathrm{BL}}^{*} = \mathrm{BL}(S)$ and an $\alpha \in \mathbb{R}$, such that $\langle m, f\rangle < \alpha$ for all $m \in \overline{M}$, and $\langle \phi, f\rangle = \phi(f) > \alpha$. Clearly $\phi(\mathbb{1})\delta_x \in \overline{M}$ for all $x \in S$, hence

$$\langle \phi(\mathbb{1})\delta_x, f\rangle = \phi(\mathbb{1})f(x) < \alpha \text{ for all } x \in S.$$

So $f < \frac{\alpha}{\phi(\mathbb{1})}$ and by positivity of ϕ,

$$\phi(f) < \phi(\frac{\alpha\mathbb{1}}{\phi(\mathbb{1})}) = \alpha,$$

which is a contradiction. So $\mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)_{+}^{*} = \mathcal{S}_{\mathrm{BL}}^{+}$. □

From Theorem 2.3.9 and Theorem 2.4.4 we get the following result:

Corollary 2.4.5. *$\mathcal{M}_{s}^{+}(S) \subset \mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)_{+}^{*}$, and $\mathcal{S}_{\mathrm{BL}} \cap \mathrm{BL}(S)_{+}^{*} = \mathcal{M}_{s}^{+}(S)$ if and only if S is complete.*

The following theorem can be proved similarly to Theorem 2.4.4:

Theorem 2.4.6. $\mathcal{S}_{e,h} \cap \mathrm{Lip}_{e,h}(S)_{+}^{*} = \mathcal{S}_{e,h}^{+}$.

And the following corollary follows from Theorem 2.3.17 and Theorem 2.4.6:

Corollary 2.4.7. *$\mathcal{M}_{s,h}^{+}(S) \subset \mathcal{S}_{e,h} \cap \mathrm{Lip}_{e,h}(S)_{+}^{*}$, and $\mathcal{S}_{e,h} \cap \mathrm{Lip}_{e,h}(S)_{+}^{*} = \mathcal{M}_{s,h}^{+}(S)$ if and only if S is complete.*

We have seen in Lemma 2.3.16 that $\mathcal{S}_{e,h}$ can be considered as a dense subspace of $\mathcal{S}_{\mathrm{BL}}$. The closed convex cones $\mathcal{S}_{e,h}^{+}$ and $\mathcal{S}_{\mathrm{BL}}^{+}$ in both spaces relate as follows:

Proposition 2.4.8. $\mathcal{S}_{\mathrm{BL}}^{+} \cap \mathcal{S}_{e,h} = \mathcal{S}_{e,h}^{+}$.

Proof. Using Theorem 2.4.4, we obtain

$$\begin{aligned} \mathcal{S}_{\mathrm{BL}}^{+} \cap \mathcal{S}_{e,h} &= \mathrm{BL}(S)_{+}^{*} \cap \mathcal{S}_{e,h} \\ &= \{\varphi \in \mathcal{S}_{e,h} : \varphi(f) \geq 0, \text{ for all } 0 \leq f \in \mathrm{BL}(S)\} =: P. \end{aligned}$$

Now, if $\varphi \in \mathcal{S}_{e,h}$ is such that $\varphi(f) \geq 0$ for all positive $f \in \mathrm{BL}(S)$, then, by Lemma 2.3.15, $\varphi(g) \geq 0$ for all positive $g \in \mathrm{Lip}_{e,h}(S)$. Hence $\varphi \in \mathcal{S}_{e,h}^{+}$ and $P \subset \mathcal{S}_{e,h}^{+}$. Clearly $\mathcal{S}_{e,h}^{+} \subset P$. □

2.5 Embedding of Lipschitz semiflows into positive linear semigroups

Let $\mathrm{Lip}(S,S)$ be the space of Lipschitz maps from S to S. For $T \in \mathrm{Lip}(S,S)$, we define the Lipschitz constant

$$|T|_{\mathrm{Lip}} := \sup\left\{\frac{d(T(x),T(y))}{d(x,y)} : x,y \in S, x \neq y\right\}.$$

Definition 2.5.1. A family of maps $(\Phi_t)_{t\geq 0}$ from S into S is a *Lipschitz semigroup on S* if

(i) for all $t \geq 0$, $\Phi_t \in \mathrm{Lip}(S,S)$,

(ii) for all $s,t \geq 0$, $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_0 = \mathrm{Id}_S$.

A Lipschitz semigroup $(\Phi_t)_{t\geq 0}$ on S is called *strongly continuous* if $t \mapsto \Phi_t(x)$ is continuous at $t = 0$ for all $x \in S$.

In this section we will show that we can "embed" in a natural way strongly continuous Lipschitz semigroups, with an additional assumption on the Lipschitz constant, into strongly continuous semigroups on Banach spaces, namely the Banach spaces $\mathcal{S}_e$ and $\mathcal{S}_{\mathrm{BL}}$ defined in Section 2.3, even isometrically in the case of $\mathcal{S}_e$.

Lemma 2.5.2. *Let $T \in \mathrm{Lip}(S,S)$. For any $f \in \mathrm{Lip}_e(S)$,*

$$\|f \circ T\|_e \leq \max(1, d(e,T(e)) + |T|_{\mathrm{Lip}})\|f\|_e,$$

and for $g \in \mathrm{BL}(S)$,

$$\|g \circ T\|_{\mathrm{BL}} \leq \max(1, |T|_{\mathrm{Lip}})\|g\|_{\mathrm{BL}}.$$

Proof. It is easy to check that for $f \in \mathrm{Lip}_e(S)$, $|f \circ T|_{\mathrm{Lip}} \leq |f|_{\mathrm{Lip}}|T|_{\mathrm{Lip}}$, hence we have

$$\begin{aligned}
\|f \circ T\|_e &\leq |f(T(e))| + |f|_{\mathrm{Lip}}|T|_{\mathrm{Lip}} \\
&\leq |f(e)| + |f|_{\mathrm{Lip}} d(e,T(e)) + |f|_{\mathrm{Lip}}|T|_{\mathrm{Lip}} \\
&\leq \max(1, d(e,T(e)) + |T|_{\mathrm{Lip}})\|f\|_e.
\end{aligned}$$

And for $g \in \mathrm{BL}(S)$, we have

$$\begin{aligned}
\|g \circ T\|_{\mathrm{BL}} &\leq \|g \circ T\|_\infty + |g|_{\mathrm{Lip}}|T|_{\mathrm{Lip}} \\
&\leq \|g\|_\infty + |g|_{\mathrm{Lip}}|T|_{\mathrm{Lip}} \leq \max(1, |T|_{\mathrm{Lip}})\|g\|_{\mathrm{BL}}.
\end{aligned}$$

□

Let $(\Phi_t)_{t\geq 0}$ be a Lipschitz semigroup on S. Then we define a semigroup of operators on $\mathrm{Lip}_e(S)$. Let $f \in \mathrm{Lip}_e(S)$ and $t \geq 0$, and let $S_\Phi(t)f := f \circ \Phi_t$. Then $S_\Phi(t)$ is a bounded linear operator on $\mathrm{Lip}_e(S)$, by Lemma 2.5.2, and $\|S_\Phi(t)\|_{\mathcal{L}(\mathrm{Lip}_e(S))} \leq \max(1, d(e, \Phi_t(e)) + |\Phi_t|_{\mathrm{Lip}})$. Hence $(S_\Phi(t))_{t\geq 0}$ is a semigroup of bounded linear operators on $\mathrm{Lip}_e(S)$.

So the dual operators $(S^*_\Phi(t))_{t\geq 0}$ form a semigroup of bounded linear operators on $\mathrm{Lip}_e(S)^*$.

Lemma 2.5.3. $S^*_\Phi(t)(\mathcal{S}_e) \subset \mathcal{S}_e$.

Proof. Let $f \in \mathrm{Lip}_e(S)$. Then

$$(S^*_\Phi(t)\delta_x)(f) = \delta_x(S_\Phi(t)f) = \delta_x(f \circ \Phi_t) = f(\Phi_t(x)) = \delta_{\Phi_t}(f), \tag{2.14}$$

for all $x \in S, t \geq 0$. Thus $S^*_\Phi(t)(D) \subset D$. Hence, by continuity of $S^*_\Phi(t)$, $S^*_\Phi(t)(\mathcal{S}_e) \subset \mathcal{S}_e$. □

Thus we can define a semigroup $(\hat{T}_\Phi(t))_{t\geq 0}$ of bounded linear operators on $\mathcal{S}_e$ by setting

$$\hat{T}_\Phi(t)\varphi := S^*_\Phi(t)\varphi, \text{ for all } \varphi \in \mathcal{S}_e, t \geq 0.$$

Theorem 2.5.4. *For all $x, y \in S$ and $s, t \geq 0$,*

$$d(\Phi_s(x), \Phi_t(y)) = \|\hat{T}_\Phi(s)\delta_x - \hat{T}_\Phi(t)\delta_y\|^*_e. \tag{2.15}$$

Furthermore, the following are equivalent:

(i) $(\hat{T}_\Phi(t))_{t\geq 0}$ *is a strongly continuous semigroup on* $\mathcal{S}_e$.

(ii) $(\Phi_t)_{t\geq 0}$ *is strongly continuous and* $\limsup_{t\downarrow 0} |\Phi_t|_{\mathrm{Lip}} < \infty$.

(iii) $(\Phi_t)_{t\geq 0}$ *is strongly continuous and there exist* $M \geq 1$ *and* $\omega \in \mathbb{R}$ *such that* $|\Phi_t|_{\mathrm{Lip}} \leq Me^{\omega t}$ *for all* $t \geq 0$.

Proof. From Lemma 2.3.5 and (2.14) we get that for every $x, y \in S$ and $t, s \geq 0$

$$\begin{aligned} \|\hat{T}_\Phi(s)\delta_x - \hat{T}_\Phi(t)\delta_y\|^*_e &= \|\delta_{\Phi_s(x)} - \delta_{\Phi_t(y)}\|^*_e \\ &= d(\Phi_s(x), \Phi_t(y)). \end{aligned}$$

$(i) \Rightarrow (iii)$: There exist $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|\hat{T}_\Phi(t)\|_{\mathcal{L}(\mathcal{S}_e)} \leq Me^{\omega t}$ for all $t \geq 0$. Hence it follows from (2.15) that for $x, y \in S$ and $t \geq 0$,

$$\begin{aligned} d(\Phi_t(x), \Phi_t(y)) &= \|\hat{T}_\Phi(t)\delta_x - \hat{T}_\Phi(t)\delta_y\|^*_e \\ &\leq Me^{\omega t}\|\delta_x - \delta_y\|^*_e = Me^{\omega t}d(x, y). \end{aligned}$$

Hence $|\Phi_t|_{\mathrm{Lip}} \leq Me^{\omega t}$ for all $t \geq 0$. From (2.15) and strong continuity of $(\hat{T}_\Phi(t))_{t\geq 0}$ it follows that $(\Phi_t)_{t\geq 0}$ is strongly continuous.

$(iii) \Rightarrow (ii)$: This is trivial.

$(ii) \Rightarrow (i)$: We want to show that there is a $\delta > 0$ and an $M \geq 1$ such that $\sup_{0\leq t\leq\delta} \|\hat{T}_\Phi(t)\|_{\mathcal{L}(\mathcal{S}_e)} \leq M$, and that $(\hat{T}_\Phi(t))_{t\geq 0}$ is strongly continuous on D. Then we can conclude by [35, Proposition 5.3] that $(\hat{T}_\Phi(t))_{t\geq 0}$ is strongly continuous on $\mathcal{S}_e$, since D is dense in $\mathcal{S}_e$ by definition. Since $\limsup_{t\downarrow 0} |\Phi_t|_{\mathrm{Lip}} < \infty$, there exist $M_1, \delta > 0$ such that $|\Phi_t|_{\mathrm{Lip}} \leq M_1$ for all $0 \leq t \leq \delta$. We know that

$$\begin{aligned} \|\hat{T}_\Phi(t)\|_{\mathcal{L}(\mathcal{S}_e)} &\leq \|S^*_\Phi(t)\|_{\mathcal{L}(\mathrm{Lip}_e(S)^*)} \\ &= \|S_\Phi(t)\|_{\mathcal{L}(\mathrm{Lip}_e(S))} \leq \max(1, d(e, \Phi_t(e)) + |\Phi_t|_{\mathrm{Lip}}). \end{aligned}$$

Now, since $[0,\delta]$ is compact, $\Phi_{[0,\delta]}(e)$ is compact, hence bounded, in S, so there is an $M_2 > 0$ such that $d(e, \Phi_t(e)) \leq M_2$ for all $0 \leq t \leq \delta$. Hence $\sup_{0\leq t\leq\delta} \|\hat{T}_\Phi(t)\|_{\mathcal{L}(\mathcal{S}_e)} \leq \max(1, M_1 + M_2) =: M < \infty$.

By (2.15) and strong continuity of $(\Phi_t)_{t\geq 0}$ we have for every $x \in S$ that

$$\|\hat{T}_\Phi(t)\delta_x - \delta_x\|^*_e = d(\Phi_t(x), x) \to 0$$

as $t \downarrow 0$. Hence by linearity $\lim_{t\downarrow 0} \|\hat{T}_\Phi(t)d - d\|^*_e = 0$ for all $d \in D$. □

Remarks. (1) Notice that for all $\varphi \in \mathcal{S}_e$, $f \in \mathrm{Lip}_e(S)$ and $t \geq 0$, we have

$$f(\hat{T}_\Phi(t)\varphi) = (\hat{T}_\Phi(t)\varphi)(f) = (S^*_\Phi(t)\varphi)(f) = \varphi(S_\Phi(t)(f)) = (S_\Phi(t)f)(\varphi).$$

Therefore $\hat{T}^*_\Phi(t)f = S_\Phi(t)f$ for all $f \in \mathrm{Lip}_e(S)$ and under the equivalent conditions of Theorem 2.5.4, $(S_\Phi(t))_{t\geq 0}$ is the dual semigroup of a strongly continuous semigroup. As $\mathcal{S}_e$ is not reflexive in general, $(S_\Phi(t))_{t\geq 0}$ cannot be expected to be strongly continuous. It *is* on the smaller space $\mathcal{S}_e^{\odot}$, the sun dual space, by definition. It would be interesting to be able to identify the latter space.

(2) In [95, Corollary 3 and Remark 4] a result similar to Theorem 2.5.4 is proven, but in less generality, since there S is taken to be a closed subset of a Banach space. In [95] the duality of spaces of Lipschitz functions is also exploited to show this result, but there the Banach space $\mathrm{Lip}_0(S)$ is used, consisting of the Lipschitz functions vanishing at some distinct point e in S. Since the semigroup $T_\Phi(t)$ will in general not map $\mathrm{Lip}_0(S)$ into itself, unless e is a fixed point of $(\Phi_t)_{t\geq 0}$, the proof in [95] needs to make use of the Banach space $\mathrm{Lip}(S)/\mathbb{R}\mathbb{1}$. By making use of the space $\mathrm{Lip}_e(S)$, we have no such difficulties.

Notice that the semigroup $(S_\Phi(t))_{t\geq 0}$ defined above is also a semigroup of bounded linear operators on $\mathrm{BL}(S)$, by Lemma 2.5.2. Then $(S^*_\Phi(t))_{t\geq 0}$ is a semigroup of bounded linear operators on $\mathrm{BL}(S)^*$. Using similar techniques as above, we can show that $S^*_\Phi(t)(\mathcal{S}_{\mathrm{BL}}) \subset \mathcal{S}_{\mathrm{BL}}$ for all $t \geq 0$. Hence we can define a semigroup $(T_\Phi(t))_{t\geq 0}$ on $\mathcal{S}_{\mathrm{BL}}$ by restricting $S^*_\Phi(t)$ to $\mathcal{S}_{\mathrm{BL}}$. Under the equivalent conditions of Theorem 2.5.4 this semigroup is strongly continuous:

Theorem 2.5.5. *For all $x, y \in S, s, t \geq 0$,*

$$\begin{aligned} \|T_\Phi(s)(\delta_x) - T_\Phi(t)(\delta_y)\|^*_{\mathrm{BL}} &= \frac{2d(\Phi_s(x), \Phi_t(y))}{2 + d(\Phi_s(x), \Phi_t(y))} \\ &\leq \min(2, d(\Phi_s(x), \Phi_t(y))). \end{aligned}$$

If $\limsup_{t\downarrow 0} |\Phi_t|_{\mathrm{Lip}} < \infty$ and $(\Phi_t)_{t\geq 0}$ is strongly continuous, then $(T_\Phi(t))_{t\geq 0}$ is a strongly continuous semigroup on $\mathcal{S}_{\mathrm{BL}}$.

The proof is similar to the proof of Theorem 2.5.4, but here the equality follows from Lemma 2.3.6.

Let $t \geq 0$. Then $\hat{T}_\Phi(t)(D^+) \subset D^+$ and $T_\Phi(t)(D^+) \subset D^+$, hence by the continuity of $\hat{T}_\Phi(t)$ and $T_\Phi(t)$ we can conclude that $\hat{T}_\Phi(t)(\mathcal{S}_e^+) \subset \mathcal{S}_e^+$ and $T_\Phi(t)(\mathcal{S}_{\mathrm{BL}}^+) \subset \mathcal{S}_{\mathrm{BL}}^+$. Thus $(\hat{T}_\Phi(t))_{t\geq 0}$ and $(T_\Phi(t))_{t\geq 0}$ are positive semigroups.

Thus, if S is complete,

$$\hat{T}_\Phi(t)(\mathcal{M}_{s,1}^+(S)) \subset \mathcal{M}_{s,1}^+(S)$$

and

$$T_\Phi(t)(\mathcal{M}_s^+(S)) \subset \mathcal{M}_s^+(S).$$

In the following proposition we will show that this also holds if S is not complete.

Proposition 2.5.6. *Let $t \geq 0$. Then $T_\Phi(t)$ and $\hat{T}_\Phi(t)$ leave $\mathcal{M}_s(S)$ and $\mathcal{M}_{s,1}(S)$ invariant, respectively. Moreover, they are given by* (2.1).

Proof. Let $\mu \in \mathcal{M}_s(S)$. Then for all $f \in \mathrm{BL}(S)$ and $t \geq 0$ we have:

$$T_\Phi(t)(\mu)(f) = \mu(S_\Phi(t)f) = \int_S f \circ \Phi_t \, d\mu = \int_S f d(\mu \circ \Phi_t^{-1}),$$

where $\mu \circ \Phi_t^{-1}$ is again a Borel measure, since Φ_t is continuous on S. Hence $T_\Phi(t)(\mu)$ is represented by the measure $\mu \circ \Phi_t^{-1}$. We now want to show that $\mu \circ \Phi_t^{-1}$ is a separable measure. Since μ is separable, there is a separable Borel measurable subset E of S, such that $|\mu|(S \backslash E) = 0$. By continuity of Φ_t, $\Phi_t(E)$ is separable, and so is $\overline{\Phi_t(E)}$. For any Borel measurable $A \subset S \backslash \overline{\Phi_t(E)}$, $\mu \circ \Phi_t^{-1}(A) = 0$. Therefore $|\mu \circ \Phi_t^{-1}|(S \backslash \overline{\Phi_t(E)}) = 0$, so $\mu \circ \Phi_t^{-1}$ is separable.

Similarly we get that for $\mu \in \mathcal{M}_{s,1}(S)$ and $t \geq 0$, $\hat{T}_\Phi(t)(\mu)$ is represented by the separable Borel measure $\mu \circ \Phi_t^{-1}$. Then, by Lemma 2.3.3, $\mu \circ \Phi_t^{-1} \in \mathcal{M}_1(S)$, hence in $\mathcal{M}_{s,1}(S)$. So $\hat{T}_\Phi(t)(\mathcal{M}_{s,1}(S)) \subset \mathcal{M}_{s,1}(S)$. □

Corollary 2.5.7. *Let $t \geq 0$. Then $T_\Phi(t)$ and $\hat{T}_\Phi(t)$ leave $\mathcal{M}_s^+(S)$ and $\mathcal{M}_s^+(S)$ invariant, respectively.*

So we see that the strongly continuous semigroup $(T_\Phi(t))_{t\geq 0}$ on $\mathcal{S}_{\mathrm{BL}}$, when restricted to $\mathcal{M}_s(S)$, is the semigroup defined by (2.1).

2.6 Notes

This chapter is an extended version of the paper [62]. First of all, the class of spaces of locally Lipschitz spaces $\mathrm{Lip}_{e,h}(S)$ have not been considered in [62], except for the space $\mathrm{Lip}_e(S)$. The first part of Section 2.3.4 appeared in [127] and Theorem 2.3.25 is new. We stated Theorem 2.4.3 in [62] without proof.

Varadarajan already showed in 1961 that the restriction of the weak topology $\sigma(\mathcal{M}_s(S), C_b(S))$ to $\mathcal{M}_s^+(S)$ is metrisable [119]. In 1966 Dudley showed that the metric can be chosen to be a norm [28]. These facts seem to be better known among probabilists than among analysts. See [92] for some further interesting properties of convergence of measures with respect to Dudley's norm.

As we stated before, various spaces of Lipschitz functions have been shown to be isometrically isomorphic to the dual of a Banach space. See e.g. Weaver's monograph on Lipschitz algebra's [122] and [68, Theorem 4.1]. But the result that, in the case of $\mathrm{BL}(S)$, this Banach space can be taken to be $\mathcal{S}_{\mathrm{BL}}$, which comes down to the closure of $\mathcal{M}_s(S)$ with respect to Dudley's bounded Lipschitz norm, seems to be new, just as the identification of the predual of $\mathrm{Lip}_{e,h}(S)$ as $\mathcal{S}_{e,h}$.

In [95, Corollary 3 and Remark 4] a result similar to Theorem 2.5.4 is proven, but in less generality, since there S is taken to be a closed subset of a Banach space.

CHAPTER

THREE

MARKOV OPERATORS AND SEMIGROUPS

3.1 Introduction

Markov operators on the cone of positive finite measures are additive and positively homogeneous operators on this cone that preserve mass, i.e. the total variation norm. A Markov semigroup is a semigroup of Markov operators. They naturally occur in probability theory and the theory of Markov processes [38, 77, 86], as well as the theory of stochastic partial differential equations [20].

In this chapter we introduce several concepts related to (regular) Markov operators and Markov semigroups on measurable spaces (Ω, Σ) and relate them to properties in terms of the Banach space $\mathcal{S}_{\mathrm{BL}}$ as defined in Chapter 2 when Ω is a metric space. We start in Section 3.2 by defining the so-called set-wise integral for functions with values in spaces of measures and give its relations with a Bochner integral in $\mathcal{S}_{\mathrm{BL}}$. In Section 3.3 we discuss Markov operators, and Section 3.4 is focused on Markov semigroups. Most of the results gathered here are not new, but form a basis for the upcoming chapters.

3.2 Measure-valued integration

Let (Ω, Σ) and (Ω', Σ') be measurable spaces. For $\mu \in \mathcal{M}(\Omega)$ and $f \in \mathrm{BM}(\Omega)$ we define $\langle \mu, f \rangle = \int_\Omega f \, d\mu$.

We will make use of the Monotone Class Theorem for functions, which we state here for convenience (see e.g. [124, Theorem II.4]). Recall that a π-*system* on Ω is a

family $\mathcal{E}$ of subsets of Ω such that $\mathcal{E}$ is non-empty and $E \cap F \in \mathcal{E}$ for all $E, F \in \mathcal{E}$ [124].

Theorem 3.2.1. *Let $\mathcal{E}$ be a π-system on Ω and let $\mathcal{H}$ be a vector space of functions from Ω to $\mathbb{R}$ such that*

(i) $\mathcal{H}$ contains the indicator function $\mathbb{1}_E$ of every $E \in \mathcal{E}$, and $\mathcal{H}$ contains $\mathbb{1}_\Omega$,

(ii) if $(f_n)_n$ is a sequence of elements of $\mathcal{H}$ with $f_n \geq 0$ and $f_n \uparrow f$, where f is bounded, then $f \in \mathcal{H}$.

Then $\mathcal{H}$ contains every bounded real-valued function which is measurable with respect to the σ-algebra generated by $\mathcal{E}$.

From this we obtain the following result on measurability of measure-valued functions:

Proposition 3.2.2. *Let $p : \Omega \to \mathcal{M}(\Omega')$. Let $\mathcal{E} \subset \Sigma'$ be a π-system that generates Σ'. Then the following conditions are equivalent:*

(i) For each $E \in \mathcal{E} \cup \{\Omega'\}$, the map $\Omega \to \mathbb{R} : \omega \mapsto p(\omega)(E)$ is measurable.

(ii) For each $f \in \mathrm{BM}(\Omega')$, the map $\Omega \to \mathbb{R} : \omega \mapsto \langle p(\omega), f\rangle$ is measurable.

Proof. (i)⇒(ii): Let $\mathcal{H}$ the vector space of functions $h \in \mathrm{BM}(\Omega')$, for which $\omega \mapsto \langle p(\omega), h\rangle$ is measurable from Ω to $\mathbb{R}$. Our aim is to show that $\mathcal{H}$ satisfies the conditions of the Monotone Class Theorem (Theorem 3.2.1). Then it follows that $\mathcal{H} = \mathrm{BM}(\Omega')$.

Theorem 3.2.1 *(i)* is satisfied by assumption. Suppose $h_n \in \mathcal{H}$ such that $0 \leq h_n \uparrow h \leq M$, for some function $h : \Omega \to \mathbb{R}$. Then $h \in \mathrm{BM}(\Omega')$. Then $H_n : \omega \mapsto \langle p(\omega), h_n\rangle$ is measurable for all $n \in \mathbb{N}$. Fix $\omega \in \Omega$. By the Dominated Convergence Theorem $\lim_{n\to\infty} H_n(\omega) = \langle p(\omega), h\rangle$. This implies that the function $\omega \mapsto \langle p(\omega), h\rangle$ is the pointwise limit of measurable functions, hence measurable. So $h \in \mathcal{H}$ and the conditions of Theorem 3.2.1 are satisfied.

(ii)⇒(i): This is trivial. □

Proposition 3.2.3. *Let μ be a σ-finite measure on (Ω, Σ). Suppose $p : \Omega \to \mathcal{M}(\Omega')$ satisfies the conditions of Proposition 3.2.2 and suppose there is a $g \in L^1(\mu)$ such that*

$$\|p(\omega)\|_{\mathrm{TV}} \leq g(\omega) \text{ for } \mu\text{-a.e. } \omega \in \Omega.$$

Then

$$\nu(E) := \int_\Omega p(\omega)(E)\, d\mu(\omega), \;\; E \in \Sigma' \tag{3.1}$$

defines a measure $\nu \in \mathcal{M}(\Omega')$ and for every $f \in \mathrm{BM}(\Omega')$

$$\int_{\Omega'} f\, d\nu = \int_\Omega \langle p(\omega), f\rangle\, d\mu(\omega). \tag{3.2}$$

Proof. Clearly $\nu(\emptyset) = 0$. Also, for every $E \in \Sigma'$,

$$|\nu(E)| = \left| \int_\Omega p(\omega)(E)\, d\mu(\omega) \right| \leq \int_\Omega |p(\omega)(E)|\, d\mu(\omega) \leq \int_\Omega g(\omega)\, d\mu(\omega) < \infty,$$

so it remains to show that ν is σ-additive. Let $(E_n)_n$ be a sequence of pairwise disjoint sets in Σ', and let $E := \cup_{n\in\mathbb{N}} E_n$.

For every $N \in \mathbb{N}$, $\omega \mapsto p(\omega)(\cup_{n=1}^N E_n)$ is measurable, and

$$\lim_{N\to\infty} \sum_{n=1}^N p(\omega)(E_n) = p(\omega)(E)$$

for every $\omega \in \Omega$, thus by the Dominated Convergence Theorem

$$\nu(E) = \lim_{N\to\infty} \sum_{n=1}^N \int_\Omega p(\omega)(E_n)\, d\mu(\omega) = \sum_{n=1}^\infty \nu(E_n).$$

Thus $\nu \in \mathcal{M}(\Omega')$.

By definition, (3.2) holds for all $f = \mathbb{1}_E$, $E \in \Sigma'$, hence for all simple functions $f : \Omega' \to \mathbb{R}$. Now let $f \in \mathrm{BM}(\Omega)_+$. Then there is a sequence $(f_n)_n$ of simple functions from $\Omega' \to \mathbb{R}$, such that $0 \leq f_n \uparrow f$ pointwise. Then by the Dominated Convergence Theorem

$$\lim_{n\to\infty} \int_{\Omega'} f_n\, d\nu = \int_{\Omega'} f\, d\nu$$

and for every $\omega \in \Omega$,

$$0 \leq \langle p(\omega), f_n \rangle \uparrow \langle p(\omega), f \rangle,$$

thus

$$\lim_{n\to\infty} \int_\Omega \langle p(\omega), f_n \rangle\, d\mu(\omega) = \int_\Omega \langle p(\omega), f \rangle\, d\mu(\omega).$$

Hence (3.2) holds for all $f \in \mathrm{BM}(\Omega')_+$, and by linearity for all $f \in \mathrm{BM}(\Omega)$. □

We call ν in (3.1) the *set-wise integral* of p with respect to μ.

Let $p : \Omega' \mapsto \mathcal{M}(\Omega')$ be defined as follows: $p(x) := \delta_x$ for every $x \in \Omega'$. Then for every $E \in \Sigma'$, $x \mapsto p(x)(E) = \mathbb{1}_E(x)$ is measurable. Moreover, for every $\mu \in \mathcal{M}(\Omega')$,

$$\int_\Omega \delta_x(E)\, d\mu(x) = \mu(E) \text{ for every } E \in \Sigma'. \tag{3.3}$$

In the remainder of this section we show that if Ω' is a complete separable metric space (S, d), the set-wise integral coincides with a Bochner integral in the Banach space $\mathcal{S}_{\mathrm{BL}}$ associated to S. These results are partly based on [63, Section 2].

Consider a complete separable metric space (S, d) with its Borel σ-algebra. Recall the Banach space $\mathcal{S}_{\mathrm{BL}}$ we introduced in Chapter 2.

Proposition 3.2.4. *Let $p : \Omega \to \mathcal{M}(S)$. Then the following conditions are equivalent:*

(i) *p is strongly measurable as map from Ω to $\mathcal{S}_{\mathrm{BL}}$.*

(ii) *For each $f \in \mathrm{BM}(S)$, the map $\Omega \to \mathbb{R} : x \mapsto \langle p(\omega), f\rangle$ is measurable.*

(iii) *For each Borel measurable $E \subset S$, the map $\Omega \to \mathbb{R} : \omega \mapsto p(\omega)(E)$ is measurable from Ω to $\mathbb{R}$.*

Proof. *(i)*$\Rightarrow$ *(ii)*: Let $\mathcal{H}$ the vector space of measurable functions h from S to $\mathbb{R}$, such that $\omega \mapsto \langle p(\omega), h\rangle$ is measurable from Ω to $\mathbb{R}$. Let $\mathcal{C}$ be the π-system of closed sets in S.

Since p is strongly measurable, it is weakly measurable. Let C be a closed set in S and let $g_n(x) := \max(1 - nd(x, C), 0)$. Then $g_n \in \mathrm{BL}(S) = \mathcal{S}^*_{\mathrm{BL}}$, hence $G_n : \omega \mapsto \langle p(\omega), g_n\rangle$ is measurable from Ω to $\mathbb{R}$. Since C is closed, $g_n(x) \to \mathbb{1}_C(x)$ for every $x \in S$. Fix $\omega \in \Omega$. Then all g_n are in $L^1(p(\omega))$, thus $\mathbb{1}_C$ is in $L^1(p(\omega))$ and

$$\lim_{n\to\infty} G_n(\omega) = \langle p(\omega), \mathbb{1}_C\rangle$$

by the Dominated Convergence Theorem. So the function

$$\omega \to p(\omega)(C) = \langle p(\omega), \mathbb{1}_C\rangle$$

is the pointwise limit of measurable functions, hence measurable. Since the σ-algebra generated by $\mathcal{C}$ is the Borel σ-algebra, Proposition 3.2.2 implies that $\Omega \to \mathbb{R} : \omega \mapsto \langle p(\omega), f\rangle$ is measurable for every $f \in \mathrm{BM}(S)$.

(ii)$\Leftrightarrow$ *(iii)*: Follows from Proposition 3.2.2

(ii)$\Rightarrow$*(i)*: Since $\mathrm{BL}(S) \subset \mathrm{BM}(S)$, $\omega \mapsto \langle p(\omega), h\rangle$ is measurable for every $h \in \mathrm{BL}(S)$. So p is weakly measurable. Since $\mathcal{S}_{\mathrm{BL}}$ is separable (by separability of S), p is strongly measurable by the Pettis Measurability Theorem. $\square$

Let $p : \Omega \to \mathcal{S}_{\mathrm{BL}}$ be Bochner integrable with respect to a σ-finite measure μ on Ω. If $p(\Omega) \subset \mathcal{M}^+(S)$, then $\int_\Omega p(\omega)\, d\mu(\omega) \in \mathcal{M}^+(S)$, since $\mathcal{M}^+(S) = \mathcal{S}^+_{\mathrm{BL}}$ is a closed convex cone of $\mathcal{S}_{\mathrm{BL}}$. If $p(\Omega) \subset \mathcal{M}(S)$, then $\int_\Omega p(\omega)\, d\mu(\omega)$ need not be in $\mathcal{M}(S)$:

Example. Let $\Omega = [0, 1]$ with Lebesgue measure μ. Let $S = \mathbb{R}$ with Euclidean metric. We define $p(0) = 0$ and

$$p(t) = \frac{1}{t}\delta_t - \frac{1}{t}\delta_{-t} \text{ for } t \in (0, 1].$$

Then $p : \Omega \to \mathcal{M}(S)$, and p is Bochner integrable with respect to μ as map from Ω to $\mathcal{S}_{\mathrm{BL}}$, since p is strongly measurable and $\|p(t)\|^*_{\mathrm{BL}} \leq 2$ for all $t \in \Omega$. Let $\varphi = \int_\Omega p(t)\, d\mu(t)$. For $n \in \mathbb{N}$ we define

$$f_n(x) = \begin{cases} 0 & \text{if } x \leq 0; \\ nx & \text{if } 0 < x < 1/n; \\ 1 & \text{if } x \geq 1/n. \end{cases}$$

Then $f_n \in \mathrm{BL}(S)$, $\|f_n\|_\infty \leq 1$ and

$$\langle \varphi, f_n \rangle = 1 + \log n.$$

Then Theorem 2.3.25 implies that $\varphi \notin \mathcal{M}(S)$.

Proposition 3.2.5. *Let $p : \Omega \to \mathcal{S}_{\mathrm{BL}}$ be Bochner integrable with respect to a σ-finite measure μ on Ω such that $p(\Omega) \subset \mathcal{M}(S)$. Suppose there is a $g \in L^1(\mu)$ such that*

$$\|p(\omega)\|_{\mathrm{TV}} \leq g(\omega) \text{ for } \mu\text{-a.e. } \omega \in \Omega.$$

Then $\nu := \int_\Omega p(\omega)\, d\mu(\omega) \in \mathcal{M}(S)$ and

$$\int_S f\, d\nu = \int_\Omega \langle p(\omega), f \rangle d\mu(\omega),$$

for any $f \in \mathrm{BM}(S)$. Thus for any Borel set $E \subset S$,

$$\left[\int_\Omega p(\omega) d\mu(\omega)\right](E) = \int_\Omega p(\omega)(E) d\mu(\omega).$$

Proof. The Bochner integrability of p implies that there exists a strongly measurable $\hat{p} : \Omega \to \mathcal{S}_{\mathrm{BL}}$ that is equal to p μ-a.e. and such that $\hat{p}(\Omega) \subset \mathcal{M}(S)$ as well. In particular, Proposition 3.2.4 implies that $t \mapsto \hat{p}(\omega)(E)$ is measurable for every Borel set E. Applying Proposition 3.2.3 yields that we can define the set-wise integral

$$\nu'(E) = \int_\Omega \hat{p}(\omega)(E)\, d\mu(\omega) \text{ for all Borel sets } E,$$

such that

$$\int_S f\, d\nu' = \int_\Omega \langle \hat{p}(\omega), f \rangle d\mu(\omega) \text{ for all } f \in \mathrm{BM}(S).$$

In particular, $\langle \nu', f \rangle = \langle \nu, f \rangle$ for all $f \in \mathrm{BL}(S)$, so $\nu = \nu' \in \mathcal{M}(S)$. □

Corollary 3.2.6. *Let $p : \Omega \to \mathcal{S}^+_{\mathrm{BL}}$ be Bochner integrable with respect to a σ-finite measure μ on Ω, and define $\nu := \int_\Omega p(\omega) d\mu(\omega)$. Then*

$$\int_S f\, d\nu = \int_\Omega \langle p(\omega), f \rangle d\mu(\omega),$$

for any bounded measurable $f : S \to \mathbb{R}$. Thus for any Borel set $E \subset S$,

$$\left[\int_\Omega p(\omega) d\mu(\omega)\right](E) = \int_\Omega p(\omega)(E) d\mu(\omega).$$

Proof. Let $p : \Omega \to \mathcal{S}^+_{\mathrm{BL}}$ be Bochner integrable with respect to μ in $\mathcal{M}^+(\Omega)$. Then there exists a strongly measurable $\hat{p} : \Omega \to \mathcal{S}^+_{\mathrm{BL}}$ that is equal to p μ-a.e. In particular, Proposition 3.2.4 implies that $t \mapsto \hat{p}(\omega)(E)$ is measurable for every Borel set E. Now, $\|p(\omega)\|_{\mathrm{TV}} = \|p(\omega)\|^*_{\mathrm{BL}}$, so Bochner integrability of p with respect to μ implies that $\|\hat{p}(\cdot)\|_{\mathrm{TV}} \in L^1(\mu)$. Now we can apply Proposition 3.2.5. □

Corollary 3.2.7. *Let $p : \Omega \to \mathcal{M}(S)$ and μ a σ-finite measure on Ω such that p and μ satisfy the conditions of Proposition 3.2.3. Then p is Bochner integrable with respect to μ as map from Ω to $\mathcal{S}_{\mathrm{BL}}$ and the Bochner integral equals the set-wise integral.*

Proof. For every $f \in \mathrm{BL}(S) \subset \mathrm{BM}(S)$, $\omega \mapsto \langle p(\omega), f\rangle$ is measurable, so p is weakly measurable as map from Ω to $\mathcal{S}_{\mathrm{BL}}$. By separability of $\mathcal{S}_{\mathrm{BL}}$, p is strongly measurable by the Pettis Measurability Theorem. Separability of $\mathcal{S}_{\mathrm{BL}}$ also implies that there is a countable subset $D \subset \mathcal{S}_{\mathrm{BL}}^* \cong \mathrm{BL}(S)$ such that

$$\|\varphi\|_{\mathrm{BL}}^* = \sup\{|\langle \varphi, f\rangle| : f \in D\},$$

so $\omega \mapsto \|p(\omega)\|_{\mathrm{BL}}^*$ is measurable and $\int_S \|p(\omega)\|_{\mathrm{BL}}^* \, d\mu(\omega) \leq \int_S g(\omega)\, d\mu(\omega) < \infty$ by assumption. Thus p is Bochner integrable with respect to μ, and the Bochner integral equals the set-wise integral since they agree on all measurable sets, according to Proposition 3.2.5. □

Corollary 3.2.8. *For any $\mu \in \mathcal{M}^+(S)$, as a Bochner integral in $\mathcal{S}_{\mathrm{BL}}$,*

$$\int_S \delta_x \, d\mu(x) = \mu.$$

Proof. This follows from Corollary 3.2.7 and (3.3). □

Note that we proved and used a similar result in $\mathcal{S}_{e,h}$ in the proof of Theorem 2.3.17.

Remark. Suppose that S is a Polish space and let d be a complete metric on S metrising the given topology. We will show that most of the concepts above do not depend on the particular metric we choose. Since the norm $\|\cdot\|_{\mathrm{BL}}^*$ depends on the metric d, the completion of $\mathcal{M}(S)$ with respect to $\|\cdot\|_{\mathrm{BL}}^*$ (which is $\mathcal{S}_{\mathrm{BL}}$) may depend on the metric d as well. However $\mathcal{S}_{\mathrm{BL}}^+ = \mathcal{M}^+(S)$ does *not* depend on d. Also, the restriction of the norm topology on $\mathcal{S}_{\mathrm{BL}}$ to $\mathcal{M}^+(S)$ equals the restriction of the weak topology $\sigma(\mathcal{M}(S), C_b(S))$ to $\mathcal{M}^+(S)$, and thus does not depend on the specific metric d. If (Ω, Σ) is a measurable space and $p : \Omega \to \mathcal{M}^+(S)$, then by Proposition 3.2.4 p is strongly measurable as map from Ω to $\mathcal{S}_{\mathrm{BL}}$ if and only if $\omega \mapsto p(\omega)(E)$ is measurable for every $E \in \Sigma$, so this property does not depend on the metric d. Note that $\|\nu\|_{\mathrm{BL}}^* = \|\nu\|_{\mathrm{TV}}$ for every $\nu \in \mathcal{M}^+(S)$, thus the property that a map $p : \Omega \to \mathcal{M}^+(S)$ is Bochner integrable (to the Banach space $\mathcal{S}_{\mathrm{BL}}$) with respect to a σ-finite measure $\mu \in \mathcal{M}^+(\Omega)$ does not depend on the specific metric. Furthermore, it follows from Corollary 3.2.6 that the value of the Bochner integral also does not depend on d.

3.3 Markov operators

Let (Ω, Σ) be a measurable space.

Following [78, Section 12.3], we define a *Markov operator* to be a map $P : \mathcal{M}^+(\Omega) \to \mathcal{M}^+(\Omega)$, such that

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

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

P extends to a positive *bounded* linear operator on $(\mathcal{M}(\Omega), \|\cdot\|_{\mathrm{TV}})$ given by $P\mu := P\mu^+ - P\mu^-$. The operator norm of this extension is 1.

Proposition 3.3.1. *Let P be a Markov operator. The following are equivalent:*

(i) There exists $U : \mathrm{BM}(\Omega) \to \mathrm{BM}(\Omega)$ such that

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

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

(b) $P\mu = \int_\Omega P\delta_x \, d\mu(x)$ set-wise.

Moreover, if (Ω', Σ, ν) is a finite measure space and $h : \Omega' \to \mathcal{M}^+(\Omega)$ is such that $\omega' \mapsto h(\omega')(E)$ is measurable for every $E \in \Sigma$ and $h(\cdot)(\Omega)$ is ν-integrable, then

$$P \int_{\Omega'} h(\omega') d\nu(\omega') = \int_{\Omega'} Ph(\omega') \, d\nu(\omega')$$

If (Ω, Σ) equals a complete separable metric space with its Borel σ-algebra, then all the integrals can be considered as Bochner integrals in $\mathcal{S}_{\mathrm{BL}}$.

Proof. (i)⇒(ii): Let $E \in \Sigma$. Then the map $x \mapsto P\delta_x(E) = U\mathbb{1}_E(x)$ is in $\mathrm{BM}(\Omega)$, hence measurable. Furthermore

$$P\mu(E) = \langle \mu, U\mathbb{1}_E\rangle = \int_\Omega U\mathbb{1}_E(x) \, d\mu(x) = \int_\Omega P\delta_x(E) \, d\mu(x).$$

(ii)⇒(i): For $f \in \mathrm{BM}(\Omega)$ and $x \in S$ we define

$$Uf(x) := \langle P\delta_x, f\rangle.$$

From *(ii)(a)* and by approximating f by step functions it follows that $Uf \in \mathrm{BM}(\Omega)$, so for $\mu \in \mathcal{M}^+(\Omega)$ we obtain, by Proposition 3.2.3 and *(ii)(b)*,

$$\langle \mu, Uf\rangle = \int_\Omega \langle P\delta_x, f\rangle \, d\mu(x) = \left\langle \int_\Omega P\delta_x \, d\mu(x), f \right\rangle = \langle P\mu, f\rangle.$$

Let (Ω', Σ, ν) be a finite measure space and $h : \Omega' \to \mathcal{M}^+(\Omega)$ such that $\omega \mapsto h(\omega')(E)$ is measurable for every $E \in \Sigma$. Then by *(i)* and Proposition 3.2.3 we have for all

$f \in \mathrm{BM}(\Omega)$

$$\begin{aligned}\left\langle P\int_{\Omega'} h(\omega')\,d\nu(\omega'), f\right\rangle &= \left\langle \int_{\Omega'} h(\omega')\,d\nu(\omega'), Uf\right\rangle \\ &= \int_{\Omega'} \langle h(\omega'), Uf\rangle\,d\nu(\omega') \\ &= \left\langle \int_{\Omega'} Ph(\omega')\,d\nu(\omega'), f\right\rangle.\end{aligned}$$

The final statement follows from Corollary 3.2.7. □

Following [47, 87], we will call a Markov operator P *regular* if it satisfies the equivalent conditions of Proposition 3.3.1, and we call the operator $U : \mathrm{BM}(\Omega) \to \mathrm{BM}(\Omega)$ appearing in Proposition 3.3.1 the *dual* of P.

A map $p : \Omega \times \Sigma \to \mathbb{R}$ is a *transition probability* (as in e.g. [86, Section 3.4.1], [38, Section 4.1]) if:

(T1) For every $x \in \Omega$ the map $p_x : \Sigma \to \mathbb{R}$, defined by $p_x(E) := p(x, E)$ for every $E \in \Sigma$, is a probability measure.

(T2) For every $E \in \Sigma$ the function $g_E : \Omega \to \mathbb{R}$ defined by $g_E(x) := p(x, E)$ for every $x \in \Omega$, is measurable.

Every transition probability p *generates* a Markov operator P, by

$$P\mu(E) := \int_\Omega p(x, E)\,d\mu(x).$$

By Proposition 3.3.1 P is regular and the dual is given by $Uf(x) = \int_\Omega f(y)\,dp(x, \cdot)$. Moreover, every regular Markov operator P is generated by the transition probability $p(x, E) := P\delta_x(E)$ for all $x \in \Omega, E \in \Sigma$. An important class of examples of regular Markov operators is given by measurable maps:

Example. Let $\Phi : \Omega \to \Omega$ be measurable. Then

$$P_\Phi \mu := \mu \circ \Phi^{-1}$$

defines a regular Markov operator with dual given by

$$U_\Phi f(x) := f(\Phi(x)).$$

Example. Let $N \in \mathbb{N}$ and $\Phi_i : \Omega \to \Omega$ and $p_i : \Omega \to \mathbb{R}_+$ be measurable for $i = 1, \cdots, N$ such that

$$\sum_{i=1}^N p_i(x) = 1 \text{ for all } x \in \Omega.$$

The pair of sequences $(\Phi_1, \cdots, \Phi_N; p_1, \cdots, p_N)$ is called an *iterated function system.* It defines a Markov operator as follows:

$$P\mu(E) := \sum_{i=1}^{N} \int_{\Phi^{-1}(E)} p_i(x) d\mu(x).$$

P is regular with dual

$$Uf(x) := \sum_{i=1}^{N} p_i(x) f(\Phi_i(x)).$$

Iterated function systems and their connections with fractals are well-studied, especially in the specific setting of a complete metric space where the mappings Φ_i are contractions and the p_i are constant. See the seminal paper by Hutchinson [66] and Barnsley's book on fractals [7]. For more recent work, see e.g. [78, 87, 111, 112, 133].

Note that not every Markov operator is regular. We now give an explicit example of a non-regular Markov operator:

Example. Let $\Omega = [0, 1]$ with the Borel σ-algebra and let m denote the Lebesgue measure restricted to Ω. By the Lebesgue-Radon-Nikodym Decomposition Theorem, every μ in $\mathcal{M}(\Omega)$ can be uniquely decomposed into $\mu_a + \mu_s$, where μ_a is absolutely continuous with respect to m and μ_s is singular with respect to m. For $\mu \in \mathcal{M}^+(\Omega)$ we define

$$P\mu = \mu_a(\Omega)m + \mu_s.$$

Then P defines a Markov operator. Note that $P\delta_x = \delta_x$ for all $x \in \Omega$. Let $\mu \in \mathcal{P}(\Omega)$ be absolutely continuous with respect to μ, but unequal to m, then

$$P\mu = m \neq \mu = \int_\Omega P\delta_x \, d\mu(x),$$

thus μ is not regular by Proposition 3.3.1.

For $n \in \mathbb{N}$ we define the *Cesàro averages* of a Markov operator P

$$P^{(n)} := \frac{1}{n} \sum_{k=0}^{n-1} P^k.$$

If P is regular with dual U, we define

$$U^{(n)} := \frac{1}{n} \sum_{k=0}^{n-1} U^k.$$

If P is a regular Markov operator with dual U, then $P^{(n)}$ is also a regular Markov operator with dual $U^{(n)}$.

A measure $\mu \in \mathcal{M}(\Omega)$ is an *invariant measure* (with respect to P) if $P\mu = \mu$.

A Markov operator P on a topological space S is a *Markov-Feller operator* if there is a map $V : C_b(S) \to C_b(S)$ such that for every $f \in C_b(S)$, $\langle P\mu, f\rangle = \langle \mu, Vf\rangle$.

Let S be a Polish space. Let d be a complete metric on S that metrises its topology and $\mathcal{S}_{\mathrm{BL}}$ the associated Banach space from Chapter 2.

Proposition 3.3.2. *Let P be a Markov operator. Then the following are equivalent:*

(i) P is a Markov-Feller operator.

(ii) $P : \mathcal{S}^+_{\mathrm{BL}} \to \mathcal{S}^+_{\mathrm{BL}}$ is continuous.

(iii) P is regular and the dual U leaves $C_b(S)$ invariant.

Proof. *(i)⇒(ii)*: Let $\mu_n, \mu \in \mathcal{M}^+(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$, then by Lemma 2.3.12 $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. Thus

$$\langle P\mu_n, f\rangle = \langle \mu_n, Vf\rangle \to \langle \mu, Vf\rangle = \langle P\mu, f\rangle,$$

for every $f \in C_b(S)$. Thus, again by Lemma 2.3.12, $\|P\mu_n - P\mu\|^*_{\mathrm{BL}} \to 0$.

(ii)⇒(iii): For $f \in \mathrm{BM}(S)$ define $Uf(x) := \langle P\delta_x, f\rangle, x \in S$. Since $x \mapsto \delta_x$ is a continuous embedding from S into $\mathcal{S}^+_{\mathrm{BL}}$, $x \mapsto P\delta_x$ is also continuous, hence strongly measurable, from S to $\mathcal{S}^+_{\mathrm{BL}}$. So by Proposition 3.2.4 the map $x \mapsto \langle P\delta_x, f\rangle$ is measurable. It is also bounded by *(MO2)* and boundedness of f. So U maps $\mathrm{BM}(S)$ into itself. Let $g \in C_b(S) \subset \mathrm{BM}(S)$. Using Lemma 2.3.12 and continuity of P, it can be shown that $x \mapsto \langle P\delta_x, g\rangle$ is continuous from S to $\mathbb{R}$, hence $Ug \in C_b(S)$.

Let $\mu \in \mathcal{M}^+(S)$ and $f \in \mathrm{BM}(S)$. By Corollary 3.2.6

$$\langle \mu, Uf\rangle = \int_S \langle P\delta_x, f\rangle \, d\mu(x) = \int_S f \, d\tilde{P}\mu,$$

where $\tilde{P}\mu := \int_S P\delta_x \, d\mu(x)$. Now it remains to prove that $\tilde{P}\mu = P\mu$. Clearly $\tilde{P}$ is positively homogeneous and additive from $\mathcal{M}^+(S)$ to $\mathcal{M}^+(S)$. Also, $\tilde{P}\delta_y = P\delta_y$ for all $y \in S$. So P and $\tilde{P}$ coincide on

$$D^+ = \{\sum_{i=1}^n \alpha_i \delta_{x_i} : n \in \mathbb{N}, \alpha_i \in \mathbb{R}_+, x_i \in S\},$$

which is dense in $\mathcal{S}^+_{\mathrm{BL}}$. Let $\mu_n \to \mu$ in $\mathcal{S}^+_{\mathrm{BL}}$, then for all $g \in C_b(S)$,

$$\int_S g \, d\tilde{P}\mu_n = \langle \mu_n, Ug\rangle \to \langle \mu, Ug\rangle = \int_S g \, d\tilde{P}\mu,$$

by Lemma 2.3.12, thus $\tilde{P}$ is also continuous from $\mathcal{S}^+_{\mathrm{BL}}$ to $\mathcal{S}^+_{\mathrm{BL}}$. Hence $P = \tilde{P}$ on $\mathcal{M}^+(S)$ by density of D^+ in $\mathcal{S}^+_{\mathrm{BL}}$.

(iii)⇒(i): This is trivial. □

Let $\Phi : S \to S$ be measurable. The associated Markov operator P_Φ is Markov-Feller if and only if Φ is continuous.

We now give a necessary and sufficient condition for a Markov operator to be extendable to a bounded linear operator on $\mathcal{S}_{\mathrm{BL}}$:

Proposition 3.3.3. *Let $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ be a regular Markov operator. Then the following statements are equivalent:*

(i) P can be extended to a bounded linear operator on $\mathcal{S}_{\mathrm{BL}}$.

(ii) The dual U leaves $\mathrm{BL}(S)$ invariant and there is a $C > 0$ such that

$$\|Uf\|^*_{\mathrm{BL}} \leq C\|f\|^*_{\mathrm{BL}} \text{ for all } f \in \mathrm{BL}(S). \tag{3.4}$$

Proof. (i)⇒(ii): Let $\mu \in \mathcal{M}^+(S)$. Then for all $f \in \mathrm{BL}(S) \cong \mathcal{S}^*_{\mathrm{BL}}$,

$$\langle \mu, P^* f\rangle = \langle P\mu, f\rangle = \langle \mu, Uf\rangle.$$

Since this holds for all $\mu \in \mathcal{M}^+(S)$, P^* is the restriction of U to $\mathrm{BL}(S)$. Consequently, U maps $\mathrm{BL}(S)$ to $\mathrm{BL}(S)$ and (3.4) is satisfied with $C = \|P^*\|$.

(ii)⇒(i): Let $\varphi \in \mathcal{S}_{\mathrm{BL}}$, then there exist $(\mu_n)_n \subset \mathcal{M}(S)$ such that $\|\mu_n - \varphi\|^*_{\mathrm{BL}} \to 0$. For all $m, n \in \mathbb{N}$

$$\|P\mu_n - P\mu_m\|^*_{\mathrm{BL}} = \sup\{|\langle \mu_n - \mu_m, Uf\rangle| : f \in \mathrm{BL}(S), \|f\|^*_{\mathrm{BL}} \leq 1\} \leq C\|\mu_n - \mu_m\|^*_{\mathrm{BL}},$$

hence $(P\mu_n)_n$ is Cauchy in $\mathcal{S}_{\mathrm{BL}}$ and thus we can define $P\varphi := \lim_{n\to\infty} P\mu_n$. Observe that $P\varphi$ does not depend on the converging $(\mu_n)_n$. Now P is a linear operator mapping $\mathcal{S}_{\mathrm{BL}}$ into itself and $\|P\varphi\|^*_{\mathrm{BL}} \leq C\|\varphi\|^*_{\mathrm{BL}}$ for all $\varphi \in \mathcal{S}_{\mathrm{BL}}$. □

There are examples of Markov-Feller operators that cannot be extended to bounded linear operators on $\mathcal{S}_{\mathrm{BL}}$:

Example. Let $S = [0, 1]$ with Euclidean metric and $\Phi(x) := \sqrt{x}$. Then $\Phi : S \to S$ is continuous, so the associated Markov operator P_Φ is Markov-Feller. We define $f(x) := x$, an element of $\mathrm{BL}(S)$. But $U_\Phi f(x) = f(\Phi(x)) = \sqrt{x}$ is not Lipschitz continuous, thus U_Φ does not map $\mathrm{BL}(S)$ into itself. Thus P_Φ does not extend to a bounded linear operator on $\mathcal{S}_{\mathrm{BL}}$ by Proposition 3.3.3

3.4 Markov semigroups

Let (Ω, Σ) be a measurable space.

A *Markov semigroup* on Ω is a semigroup $(P(t))_{t\geq 0}$ of Markov operators on $\mathcal{M}^+(\Omega)$. $(P(t))_{t\geq 0}$ is called *regular* if $P(t)$ is regular for all $t \geq 0$. In this case we obtain the *dual semigroup* $(U(t))_{t\geq 0}$ on $\mathrm{BM}(\Omega)$, where $U(t)$ is the dual of $P(t)$ for every $t \in \mathbb{R}_+$.

We call a Markov semigroup $(P(t))_{t\geq 0}$ *jointly measurable* if the function $(t,x) \mapsto P(t)\delta_x(E)$ is jointly measurable from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}$ for every $E \in \Sigma$.

Lemma 3.4.1. *Let $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup. Then for every $\mu \in \mathcal{M}^+(\Omega)$, $t \mapsto P(t)\mu(E)$ is measurable for every $E \in \Sigma$.*

Proof. Let $\mu \in \mathcal{M}^+(\Omega)$ and $E \in \Sigma$. Since $P(t)$ is regular for every $t \in \mathbb{R}_+$,

$$P(t)\mu(E) = \int_S P(t)\delta_x(E)\, d\mu(x),$$

so it follows from the joint measurability and Tonelli's Theorem that $t \mapsto P(t)\mu(E)$ is measurable from $\mathbb{R}_+$ to $\mathbb{R}$. □

A map $p : \mathbb{R}_+ \times \Omega \times \Sigma \to \mathbb{R}$ is a *Markov transition function* (as in [38, Section 4.1]) if:

(TF1) $p(t,x,\cdot)$ is a probability measure for every $t \in \mathbb{R}_+$, $x \in \Omega$ and $p(0,x,\cdot) = \delta_x$ for every $x \in \Omega$.

(TF2) $p(\cdot,\cdot,E)$ is jointly $\mathcal{B}(\mathbb{R}_+) \times \Sigma$-measurable for every $E \in \Sigma$.

(TF3) p satisfies the Chapman-Kolmogorov equation

$$p(t+s,x,E) = \int_S p(s,y,E)p(t,x,dy)$$

for every $s,t \in \mathbb{R}_+, x \in \Omega, E \in \Sigma$.

Remark. The definition of the term "(Markov) transition function" varies in the literature. For instance, in [77] it is not assumed that $p(\cdot,\cdot,E)$ is *jointly* measurable, and $p(t,x,\cdot)$ is only assumed to be a positive measure.

Every Markov transition function p *generates* a regular jointly measurable Markov semigroup: $P(t)\mu(E) := \int_\Omega p(t,x,E)\, d\mu(x)$. The semigroup property follows from *(TF3)*.

Let $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup and define $p(t,x,E) := P(t)\delta_x(E)$. By regularity and Proposition 3.3.1,

$$P(t)\mu(E) = \int_\Omega P(t)\delta_x(E)\, d\mu(x).$$

Obviously p satisfies *(TF1)* and *(TF2)*. Let $s,t \in \mathbb{R}_+, x \in \Omega, E \in \Sigma$, then

$$\begin{aligned} p(t+s,x,E) &= \langle P(t+s)\delta_x, \mathbb{1}_E\rangle = \langle P(t)\delta_x, U(s)\mathbb{1}_E\rangle = p(t+s,x,E) \\ &= \int_\Omega p(s,y,E)p(t,x,dy). \end{aligned}$$

So a regular jointly measurable Markov semigroup corresponds to a Markov transition function and vice versa.

Time-homogeneous Markov processes can be defined using Markov transition functions (see e.g. [38, Section 4.1]), so regular jointly measurable Markov semigroups are a natural object of interest when studying these processes.

A measure $\mu \in \mathcal{M}(\Omega)$ is an *invariant measure* (with respect to $(P(t))_{t\geq 0}$) if $P(t)\mu = \mu$ for all $t \in \mathbb{R}_+$.

A class of examples of regular Markov semigroups is given by deterministic dynamical systems, induced by semigroups of measurable maps on the state space Ω. A *semigroup of measurable maps* on Ω is a family of maps $(\Phi_t)_{t\geq 0}$, such that $\Phi_t : \Omega \to \Omega$ is measurable, $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_0 = \mathrm{Id}_\Omega$ for all $s, t \in \mathbb{R}_+$. $(\Phi_t)_{t\geq 0}$ is called *jointly measurable* if $(t, x) \mapsto \Phi_t(x)$ is measurable from $\mathbb{R}_+ \times \Omega$ to Ω.

Proposition 3.4.2. *Let $(\Phi_t)_{t\geq 0}$ be a semigroup of measurable maps on S. Then*

(i) $P_\Phi(t)\mu := \mu \circ \Phi_t^{-1}$ defines a regular Markov semigroup $(P_\Phi(t))_{t\geq 0}$ with dual $(U_\Phi(t))_{t\geq 0}$ given by $U_\Phi(t)f(x) := f(\Phi_t(x))$.

(ii) If $(\Phi_t)_{t\geq 0}$ is jointly measurable, then $(P_\Phi(t))_{t\geq 0}$ is a jointly measurable Markov semigroup.

Proof. (i) is easy to verify. Assume that $(\Phi_t)_{t\geq 0}$ is jointly measurable. Then for every $E \in \Sigma$

$$(t, x) \mapsto P_\Phi(t)\delta_x(E) = \mathbb{1}_E(\Phi_t(x))$$

is measurable, thus $(P_\Phi(t))_{t\geq 0}$ is jointly measurable. □

For the remainder we assume that $\mathbf{P} = (P(t))_{t\geq 0}$ is a regular jointly measurable Markov semigroup with dual $\mathbf{U} = (U(t))_{t\geq 0}$.

In Section 3.3 we defined and considered properties of Cesàro averages of Markov operators. We will do so now for Markov semigroups: For $\mu \in \mathcal{M}^+(\Omega)$ and $t > 0$ we define the Cesàro averages

$$\mathbf{P}^{(t)}\mu := \frac{1}{t}\int_0^t P(s)\mu\, ds, \qquad P^{(0)}\mu := \mu,$$

where the integral is defined set-wise.

Lemma 3.4.3. *For every $t > 0$, $\mathbf{P}^{(t)}$ is a regular Markov operator with dual*

$$\mathbf{U}^{(t)}f(x) := \frac{1}{t}\int_0^t U(s)f(x)\, ds.$$

Proof. Let $f \in \mathrm{BM}(\Omega), \mu \in \mathcal{M}^+(\Omega)$ and $t > 0$. Then it follows from Lemma 3.4.1 and Proposition 3.2.3 that

$$\begin{aligned}\langle \mathbf{P}^{(t)}\mu, f\rangle &= \frac{1}{t}\int_0^t \langle P(s)\mu, f\rangle\, ds = \frac{1}{t}\int_0^t \langle \mu, U(s)f\rangle\, ds \\ &= \frac{1}{t}\int_0^t \int_S U(s)f(x)\, d\mu(x)\, ds.\end{aligned}$$

Since $(P(t))_{t\geq 0}$ is jointly measurable, Proposition 3.2.3 implies that the map

$$(s,x) \mapsto U(s)f(x) = \langle P(s)\delta_x, f\rangle$$

is measurable from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}$. Then Fubini's Theorem yields that $\mathbf{U}^{(t)}f : x \mapsto \frac{1}{t}\int_0^t U(s)f(x)\,ds$ is measurable. This map is bounded by $\|f\|_\infty$. So we can again apply Fubini's Theorem to obtain:

$$\begin{aligned}\frac{1}{t}\int_0^t \int_S U(s)f(x)\,d\mu(x)\,ds &= \frac{1}{t}\int_S \int_0^t U(s)f(x)\,ds\,d\mu(x)\\ &= \langle \mu, \mathbf{U}^{(t)}f\rangle,\end{aligned}$$

which proves the statement. □

Now we consider a topological space S. On a topological space there are two natural σ-algebra's: the Borel σ-algebra $\mathcal{B}(S)$ and the Baire σ-algebra, $\mathcal{B}_{ba}(S)$, which is the smallest σ-algebra such that every continuous function is measurable. Clearly $\mathcal{B}_{ba}(S) \subset \mathcal{B}(S)$, but they need not be the same. A topological space S is *normal* if it is Hausdorff and any two disjoint closed subsets of S can be separated by neighbourhoods, and *perfectly normal* if it is normal and every closed set is a G_δ-set, i.e. a countable intersection of open sets. In a perfectly normal space, the Baire σ-algebra and the Borel σ-algebra coincide [13, Proposition 6.3.4]. Any metric space is perfectly normal.

Let S be a topological space with either the Baire σ-algebra or the Borel σ-algebra. We call a Markov semigroup $(P(t))_{t\geq 0}$ on S *Markov-Feller* if $P(t)$ is Markov-Feller for every $t \in \mathbb{R}_+$, *strongly stochastically continuous* when $t \mapsto \langle P(t)\mu, f\rangle$ is continuous for all $\mu \in \mathcal{M}^+(S)$ and $f \in C_b(S)$, and *strongly stochastically continuous at zero* when $t \mapsto \langle P(t)\mu, f\rangle$ is continuous at zero for all $\mu \in \mathcal{M}^+(S)$ and $f \in C_b(S)$.

Lemma 3.4.4. *Let $(P(t))_{t\geq 0}$ be a regular Markov semigroup. Then the following are equivalent:*

(i) $(P(t))_{t\geq 0}$ is strongly stochastically continuous at zero,

(ii) $t \mapsto \langle P(t)\mu, f\rangle$ is continuous at zero for all $\mu \in \mathcal{M}(S)$ and $f \in C_b(S)$.

(iii) $t \mapsto \langle P(t)\mu, f\rangle$ is right continuous for all $\mu \in \mathcal{M}(S)$ and $f \in C_b(S)$.

Proof. *(i)⇒(ii)*: This follows from $\langle P(t)\mu, f\rangle = \langle P(t)\mu^+, f\rangle - \langle P(t)\mu, f\rangle$. *(ii)⇒(iii)*: Let $s, t \in \mathbb{R}_+$. Then

For every $x \in S$, $U(s)f(x) = \langle P(s)\delta_x, f\rangle \to f(x)$ as $s \downarrow 0$. Thus by the Dominated Convergence Theorem we have for all $t \in \mathbb{R}_+$

$$\lim_{s\downarrow 0}\langle P(t+s)\mu, f\rangle = \lim_{s\downarrow 0}\langle P(t)\mu, U(s)f\rangle = \langle P(t)\mu, f\rangle.$$

(iii)⇒(i): This is trivial. □

We will say $(P(t))_{t\geq 0}$ is right strongly stochastically continuous if it satisfies Lemma 3.4.4 *(iii)*.

Proposition 3.4.5. *Let $(P(t))_{t\geq 0}$ be a regular Markov semigroup on a perfectly normal space S that is strongly stochastically continuous at zero. Then it is jointly measurable.*

Proof. By Lemma 3.4.4 $(P(t))_{t\geq 0}$ is a right strongly stochastically continuous Markov semigroup.

For every $f \in C_b(S)$ consider the map $H_f : (t,x) \mapsto \langle P(t)\delta_x, f\rangle$. Then $t \mapsto H_f(t,x)$ is right continuous for every $x \in S$, and for every fixed $t \in \mathbb{R}_+$, $x \mapsto H_f(t,x) = U(t)f(x)$ is measurable by regularity. So $(t,x) \mapsto \langle P(t)\delta_x, f\rangle$ is jointly measurable by [13, Lemma 6.4.6]. Now let C be closed in S. Then there exists a sequence $(U_n)_n$ of open subsets S, such that $C = \cap_{n=1}^{\infty} U_n$. By the Tietze Extension Theorem (see e.g. [44, Theorem 4.16]), there exist $f_n : S \to [0,1]$ continuous, such that $f_n|_C = \mathbb{1}_C$ and $f_n|_{S\setminus U_n} = 0$. Then $f_n \to \mathbb{1}_C$ pointwise, so by the Dominated Convergence Theorem, $\langle P(t)\delta_x, f_n\rangle \to P(t)\delta_x(C)$ for all $t \in \mathbb{R}_+$. So $(t,x) \mapsto P(t)\delta_x(C)$ is a pointwise limit of measurable functions, hence (jointly) measurable. By Proposition 3.2.2, $(t,x) \mapsto P(t)\delta_x(E)$ is jointly measurable for all Borel sets E. □

Let S be a topological space. A semigroup $(\Phi_t)_{t\geq 0}$ of measurable maps on S is called *strongly continuous (at zero)* if the map $\mathbb{R}_+ \to S : t \mapsto \Phi_t(x)$ is continuous (at zero) for all $x \in S$.

Lemma 3.4.6. *Let $(\Phi_t)_{t\geq 0}$ be a semigroup of measurable maps on S and $(P(t))_{t\geq 0}$ its associated Markov semigroup. If $(\Phi_t)_{t\geq 0}$ is strongly continuous (at zero) then $(P(t))_{t\geq 0}$ is strongly stochastically continuous (at zero). The converse holds if S is a normal space.*

Proof. Let $t \mapsto \Phi_t(x)$ be continuous at $s \in \mathbb{R}_+$ for all $x \in S$, then $t \mapsto U(t)f(x)$ is continuous at $s \in \mathbb{R}_+$ for all $x \in S$ and $f \in C_b(S)$. Using the Dominated Convergence Theorem, it can be shown that $t \mapsto \langle P(t)\mu, f\rangle = \langle \mu, U(t)f\rangle$ is continuous at s for all $f \in C_b(S), \mu \in \mathcal{M}(S)$. This proves one direction.

Now assume S is a normal space and that $t \mapsto \langle P(t)\delta_x, f\rangle = f(\Phi_t(x))$ is continuous at $s \in \mathbb{R}_+$ for all $f \in C_b(S)$. Suppose that $t \mapsto \Phi_t(x)$ is not continuous at s, then there is a neighbourhood U of x and a sequence $\Phi_{t_n}(x)$ such that $t_n \in \mathbb{R}_+$, $t_n \to s$ and $\Phi_{t_n}(x) \notin U$. By the Tietze Extension Theorem there is a $f \in C_b(S)$ such that $f|_{S\setminus U} \equiv 0$ and $f(\Phi_s(x)) = 1$, since $S\setminus U$ and $\{\Phi_s(x)\}$ are closed. But then $f(\Phi_{t_n}(x))$ does not converge to $f(\Phi_s(x))$. □

Let (S,d) be a metric space and $(\Phi_t)_{t\geq 0}$ a Lipschitz semigroup on S defined as in Section 2.5. Since $(\Phi_t)_{t\geq 0}$ is a semigroup of measurable maps on S, it generates a regular Markov semigroup $(P(t))_{t\geq 0}$ on S. This Markov semigroup is Markov-Feller and equals the $(T_\Phi(t))_{t\geq 0}$ from Section 2.5 and, as shown there, can be extended to a positive linear semigroup on $\mathcal{S}_{\mathrm{BL}}$. Note that Markov semigroups will not in

general be extendable to semigroups on $\mathcal{S}_{\mathrm{BL}}$, even when they are Markov-Feller. Easy counterexamples can be constructed using Proposition 3.3.3.

3.5 Notes

The notion of set-wise integral of functions mapping into a space of measures is not new; it has been used often (implicitly) in the literature in various settings. E.g. in [135] this topic is also discussed, in a more specific setting (Ω' is taken to be a locally compact separable metric space). The connections between the set-wise integral and the Bochner integral on $\mathcal{S}_{\mathrm{BL}}$ seem to be new. Part of these results have appeared in [63, Section 2].

There is an intimate connection between Markov operators and semigroups on a space of measures and (discrete-time or continuous-time) Markov processes [38, 77, 86]. They have been widely studied over the last few decades, in a varied range of state spaces, e.g. finite and countable spaces [15], compact (metric) spaces [9, 75, 105, 130], locally compact separable metric spaces [57, 86, 130, 132], Polish spaces [20, 78, 112] and general measurable spaces [69, 86, 104] (the list of references is far from complete).

There exist also other notions of Markov operators and semigroups, for instance Markov operators and semigroups on L^1-spaces [34, 43, 78, 99, 100] or (more generally) on Banach lattices [34].

CHAPTER

FOUR

CONTINUITY PROPERTIES OF MARKOV SEMIGROUPS AND THEIR RESTRICTIONS TO INVARIANT L^1-SPACES

4.1 Introduction

Markov semigroups $(P(t))_{t\geq 0}$ on measurable spaces (Ω, Σ) that one obtains in practice are hardly ever continuous for the total variation norm $\|\cdot\|_{\mathrm{TV}}$ on the space of finite measures $\mathcal{M}(\Omega)$ on the underlying measurable space (Ω, Σ). Notable exceptions are Markov jump processes [38, 40], which yield strongly continuous semigroups in $\mathcal{M}(\Omega)$ for $\|\cdot\|_{\mathrm{TV}}$ [118]. This may have motivated other researchers to consider the more restrictive setting of strongly continuous Markov semigroups on L^1-spaces with respect to particular positive measures (see e.g. [34, 78, 99, 100]).

In this chapter we address two closely related questions. In the case that $(P(t))_{t\geq 0}$ leaves invariant a cone $\Gamma \subset \mathcal{M}^+(\Omega)$ such that the measures in Γ are all absolutely continuous with respect to a single measure μ, i.e. $\Gamma = L^1_+(\Omega, \mu)$, it induces a semigroup of non-expansive linear operators on $L^1(\Omega, \mu)$ that are isometries on $L^1_+(\Omega, \mu)$. The first question is then to characterise when this induced semigroup is strongly continuous. This is the main topic of Section 4.3 and the question is answered in Theorem 4.3.6, under a measurability assumption and the assumption that for each $\mu \in \mathcal{M}^+(\Omega)$, the map $t \mapsto \int_\Omega f\, d[P(t)\mu]$ is continuous at zero for all f in a prescribed subset X of the bounded measurable functions. E.g., when Ω is a metric space endowed with its Borel σ-algebra, then one can take $X = C_b(\Omega)$.

Also under these assumptions and the additional assumption that $(P(t))_{t\geq 0}$ is regular, we then deal with the second question in Section 4.4, which is to characterise

the subspace $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ of $\mathcal{M}(\Omega)$ that consists of all measures μ that are continuous (C^0) for the total variation norm topology, i.e. all μ for which $t \mapsto P(t)\mu$ is continuous for $\|\cdot\|_{\mathrm{TV}}$. This subspace contains in particular all invariant measures. The characterisation is given in Theorem 4.4.6 and Theorem 4.4.10, and exploits results from [51] on modules of Banach algebras with approximate identity and properties of the set-wise integral as defined in Section 3.2.

In Section 4.5 we prove some more properties of $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$. This subspace turns out to be a *projection band* in the Banach lattice $\mathcal{M}(\Omega)$ (Proposition 4.5.1), hence it is complemented. This complement is characterised and will not be $(P(t))_{t\geq 0}$-invariant in general (unfortunately). An additional result of our approach is a generalisation of a classical result by Wiener and Young [123] to general Markov semigroups (Theorem 4.5.9).

4.2 Space of measures viewed as Banach lattice

$\mathcal{M}(\Omega)$ is a Banach space endowed with the total variation norm $\|\cdot\|_{\mathrm{TV}}$. Let $\mu, \nu \in \mathcal{M}(\Omega)$. μ is absolutely continuous with respect to ν, $\mu \ll \nu$, if $|\mu|(E) = 0$ for every $E \in \Sigma$ for which $|\nu|(E) = 0$.

Let $\mu \in \mathcal{M}(\Omega), \nu \in \mathcal{M}^+(\Omega)$, then $\mu \ll \nu$ if and only if $\mu(E) = 0$ for every $E \in \Sigma$ such that $\nu(E) = 0$, which is easy to prove.

We refer to [3], [85] and [131] for the basic theory on Riesz spaces and Banach lattices.

$\mathcal{M}(\Omega)$ is an ordered vector space for the partial ordering defined by

$$\mu \leq \nu \text{ whenever } \mu(E) \leq \nu(E) \text{ for all } E \in \Sigma.$$

$\mathcal{M}(\Omega)$ is a Riesz space, where the least upper bound of μ and ν is given by

$$\mu \vee \nu(E) := \sup\{\mu(A) + \nu(E \setminus A) \,|\, A \in \Sigma,\ A \subset E\},$$

and the greatest lower bound is given by

$$\mu \wedge \nu(E) := \inf\{\mu(A) + \nu(E \setminus A) \,|\, A \in \Sigma,\ A \subset E\}.$$

The positive and negative part of $\mu \in \mathcal{M}(\Omega)$ as introduced in measure theory, μ^+ and μ^-, correspond to the concepts of positive and negative part in a Riesz space: $\mu^+ = \mu \vee 0$, $\mu^- = (-\mu)^+$ and $|\mu| = \mu^+ + \mu^-$. $\mu, \nu \in \mathcal{M}(\Omega)$ are mutually singular, $\mu \perp \nu$, if there is a $U \in \Sigma$, such that $\mu(E) = \mu(E \cap U)$ and $\nu(E) = \nu(E \setminus U)$ for every $E \in \Sigma$. Note that $|\mu| \leq |\nu|$ implies $\mu \ll \nu$. Mutual singularity of $\mu, \nu \in \mathcal{M}(\Omega)$ corresponds to the concept of disjointness in a Riesz space: μ and ν are disjoint, $\mu \perp \nu$, whenever $|\mu| \wedge |\nu| = 0$. $\mathcal{M}(\Omega)$ is in fact a Dedekind complete Riesz space [85, 1.1 Example vi)].

$\mathcal{M}(\Omega)$ is a Banach lattice for the total variation norm since $\|\mu\|_{\mathrm{TV}} = |\mu|(\Omega) = \||\mu|\|_{\mathrm{TV}}$. $\|\cdot\|_{\mathrm{TV}}$ is an L-norm: $\|\mu+\nu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{TV}} + \|\nu\|_{\mathrm{TV}}$ for all $\mu,\nu \in \mathcal{M}^+(\Omega)$, hence $\mathcal{M}(\Omega)$ is an L-space . This also implies that $\|\mu+\nu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{TV}} + \|\nu\|_{\mathrm{TV}}$ for all $\mu,\nu \in \mathcal{M}(\Omega)$ such that $\mu \perp \nu$. As in all Banach lattices, the lattice operations are continuous for the norm topology (see e.g. [85, Proposition 1.1.6]).

We will now recall some concepts in Riesz spaces that we will need later on: Let X be a Riesz space. A subspace I of X is an *ideal* of X if $|x| \leq |y|$ for some $y \in I$ implies $x \in I$. An ideal B of X is a *band* of X if $\sup(A) \in B$ for every subset $A \subset B$ which has a supremum in X. A band B of X is a *projection band* if there exists a +linear projection $P : X \to B$, such that $0 \leq Px \leq x$ for all $x \in X_+$. In this case $X = B \oplus B^\perp$, where $B^\perp := \{x \in X : x \perp y \text{ for all } y \in B\}$.

In a remark in [3] (under definition 4.20) it is shown that every L-space has order continuous norm as a consequence of [85, Theorem 2.4.2]. Furthermore, in a Banach lattice with order continuous norm, every closed ideal is a projection band [85, Corollary 2.4.2]. These statements imply

Theorem 4.2.1. *Every closed ideal in $\mathcal{M}(\Omega)$ is a projection band.*

4.3 Restriction to invariant L^1-spaces

Let $\mu \in \mathcal{M}^+(\Omega)$. For $f \in L^1(\mu)$ we define the measure $j_\mu(f)(E) := \int_E f\,d\mu$ for all $E \in \Sigma$. Then j_μ is a linear map from $L^1(\mu)$ into $\mathcal{M}(\Omega)$, and the Radon-Nikodym Theorem implies that $j_\mu(L^1(\mu))$ consists exactly of the measures in $\mathcal{M}(\Omega)$ that are absolutely continuous with respect to μ.

Lemma 4.3.1. *j_μ is an isometric embedding of $L^1(\mu)$ into $\mathcal{M}(\Omega)$, i.e.*

$$\|j_\mu(f)\|_{\mathrm{TV}} = \|f\|_1 \textit{ for all } f \in L^1(\mu).$$

The proof is straightforward.

Let $P : \mathcal{M}^+(\Omega) \to \mathcal{M}^+(\Omega)$ be a Markov operator. Then the following lemma holds:

Lemma 4.3.2. *If $\mu,\nu \in \mathcal{M}^+(\Omega)$ satisfy $\mu \ll \nu$, then $P\mu \ll P\nu$.*

Proof. There exists $f \in L^1_+(\nu)$, such that $j_\nu(f) = \mu$. There are $f_n \in L^\infty_+(\mu)$ with $\|f_n - f\|_1 \to 0$. According to Lemma 4.3.1,

$$\|Pj_\nu(f_n) - Pj_\nu(f)\|_{\mathrm{TV}} \leq \|j_\nu(f_n) - j_\nu(f)\|_{\mathrm{TV}} = \|f_n - f\|_1 \to 0.$$

Furthermore, $0 \leq j_\nu(f_n) \leq \|f_n\|_\infty \nu$. Hence by positivity of P, $0 \leq Pj_\nu(f_n) \leq \|f_n\|_\infty P\nu$. Therefore $Pj_\nu(f_n) \ll P\nu$, hence $Pj_\nu(f_n) \in L^1_+(P\nu)$ for all $n \in \mathbb{N}$. Because $L^1_+(P\nu)$ is closed in $\mathcal{M}^+(\Omega)$, $Pj_\nu(f) \in L^1_+(P\nu)$ as well, thus $P\mu \ll P\nu$. $\square$

Corollary 4.3.3. *P leaves $j_\mu(L^1_+(\mu))$ invariant if and only if $P\mu \ll \mu$.*

Proof. Clearly, if P leaves $j_\mu(L^1_+(\mu))$ invariant, then in particular $P\mu \ll \mu$. The proof in the opposite direction follows from Lemma 4.3.2: if $f \in L^1_+(\mu)$, then $0 \leq j_\mu(f) \ll \mu$, hence $Pj_\mu(f) \ll P\mu \ll \mu$. □

Suppose that P leaves $j_\mu(L^1_+(\mu))$ invariant. Then P induces an additive and positively homogeneous map $T : L^1_+(\mu) \to L^1_+(\mu)$:

$$Tf := j_\mu^{-1} \circ P \circ j_\mu(f).$$

Because $L^1(\mu)$ is a Banach lattice, T extends to a positive bounded linear operator on $L^1(\mu)$, which we will also denote by T, and $\|Tf\|_1 = \|f\|_1$ for every $f \in L^1_+(\mu)$ by Lemma 4.3.1 and *(MO2)*. So

$$\|T\| = \sup\{\|Tf\|_1 \,:\, f \in L^1_+(\mu),\ \|f\|_1 \leq 1\} = 1 \tag{4.1}$$

T will be called *the operator* (on $L^1(\mu)$) *induced by* P.

If $(P(t))_{t\geq 0}$ is a Markov semigroup that leaves $j_\mu(L^1_+(\mu))$ invariant, then it induces a semigroup $(T(t))_{t\geq 0}$ on $L^1(\mu)$ which is positive and consists of isometries on $L^1_+(\mu)$, and $\|T(t)\| = 1$ for every $t \in \mathbb{R}_+$. Our aim is to find conditions on the Markov semigroup and on μ such that $(T(t))_{t\geq 0}$ is strongly continuous.

Crucial in our approach is the following characterisation of relatively weakly compact subsets of L^1 (e.g. [1], Theorem 5.2.9, p. 109):

Theorem 4.3.4 (Dunford-Pettis)**.** *Let (Ω, Σ, μ) be a σ-finite measure space. In addition let $\mathcal{F}$ be a bounded set in $L^1(\mu)$. Then the following conditions on $\mathcal{F}$ are equivalent:*

(i) $\mathcal{F}$ is relatively weakly compact.

(ii) For every sequence $(A_n)_n$ of disjoint measurable sets

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F}} \int_{A_n} |f|\, d\mu = 0.$$

We also need the following sufficient condition for strong continuity of semigroups on $L^1(\mu)$. Its proof is inspired by that of [35, Theorem I.5.8], where it is shown that a weakly continuous semigroup on a Banach space is strongly continuous.

Proposition 4.3.5. *Let (Ω, Σ, μ) be a σ-finite measure space and $(T(t))_{t\geq 0}$ a semigroup of bounded linear operators on $L^1(\mu)$. Suppose that there exist $\delta > 0$ and $M \geq 1$, such that $\|T(t)\| \leq M$ for all $t \in [0, \delta]$, and a weakly dense subset $D \subset L^1(\mu)$ such that the map*

$$t \mapsto T(t)f, \qquad \mathbb{R}_+ \to L^1(\mu)$$

is weakly measurable and weakly continuous at zero for every $f \in D$. Then $(T(t))_{t\geq 0}$ is a C_0-semigroup.

Proof. We will show that there is a norm dense subspace $\hat{D}$ of $L^1(\mu)$, such that $t \mapsto T(t)f$ is norm continuous at zero for $f \in \hat{D}$. Then $(T(t))_{t\geq 0}$ is strongly continuous on $L^1(\mu)$ (e.g. [35, Proposition I.5.3]).

Fix $f \in D$. Let $E \in \Sigma$, then $\mathbb{1}_E \in L^\infty(\mu)$, so the map

$$t \mapsto j_\mu(T(t)f)(E) = \int_E T(t)f \, d\mu = \langle T(t)f, \mathbb{1}_E \rangle$$

is measurable by assumption.

Let $0 < r < \delta$. Since $t \mapsto \|j_\mu(T(t)f)\|_{\mathrm{TV}} = \|T(t)f\|_1$ is bounded on $[0, \delta]$, we can define by Proposition 3.2.3:

$$\nu_r(E) := \frac{1}{r} \int_0^r j_\mu(T(t)f)(E) \, dt \qquad \text{for all } E \in \Sigma.$$

Then $\nu_r \in \mathcal{M}^+(\Omega)$.

Note that $\nu_r \ll \mu$: if $\mu(E) = 0$ for some $E \in \Sigma$, then $j_\mu(T(t)f)(E) = 0$ for all $t \in \mathbb{R}_+$, and thus $\nu_r(E) = 0$. So $\nu_r \in j_\mu(L^1(\mu))$. Let $f_r \in L^1(\mu)$ be such that $\nu_r = j_\mu(f_r)$.

Proposition 3.2.3 implies that for every $g \in \mathrm{BM}(\Omega)$

$$\langle f_r, g \rangle = \langle \nu_r, g \rangle = \frac{1}{r} \int_0^r \int_\Omega T(t)f \cdot g \, d\mu \, dt.$$

Let $g \in L^\infty(\mu)$, then by assumption, $\langle T(t)f, g \rangle \to \langle f, g \rangle$ as $t \downarrow 0$, thus

$$\lim_{r \downarrow 0} \langle f_r, g \rangle = \langle f, g \rangle.$$

Hence $f_r \to f$ weakly as $r \downarrow 0$. This implies that

$$\hat{D} := \operatorname{span}\{f_r : f \in D, r > 0\}$$

is weakly dense in $L^1(\mu)$. Norm closure and weak closure agree on convex sets. Therefore $\hat{D}$ is a norm dense subspace of $L^1(\mu)$. Now, fix $f \in D$, $r > 0$ and let $0 \leq s \leq r$. By the semigroup property

$$M' := \sup\{\|T(t)\| : 0 \leq t \leq 2r\} < \infty.$$

Now for every $g \in L^\infty(\mu)$ with $\|g\|_\infty \leq 1$ we have

$$\begin{aligned}
|\langle T(s)f_r - f_r, g \rangle| &= \frac{1}{r} \left| \int_s^{s+r} \langle T(t)f, g \rangle \, dt - \int_0^r \langle T(t)f, g \rangle \, dt \right| \\
&= \frac{1}{r} \left| \int_r^{s+r} \langle T(t)f, g \rangle \, dt - \int_0^s \langle T(t)f, g \rangle \, dt \right| \\
&\leq \frac{1}{r} \int_r^{s+r} |\langle T(t)f, g \rangle| \, dt + \frac{1}{r} \int_0^s |\langle T(t)f, g \rangle| \, dt \leq M' \|f\|_1 \, \frac{2s}{r}.
\end{aligned}$$

Thus $\|T(s)f_r - f_r\|_1 \to 0$ as $s \downarrow 0$. By linearity $t \mapsto T(t)g$ is norm continuous at zero for every $f \in \hat{D}$. Since $\hat{D}$ is a norm dense subspace of $L^1(\mu)$, $(T(t))_{t\geq 0}$ is strongly continuous on $L^1(\mu)$. □

An $\mathcal{M}(\Omega)$-*separating set* is a set $X \subset \mathrm{BM}(S)$ such that if $\mu, \nu \in \mathcal{M}(\Omega)$ satisfy

$$\langle \mu, \phi \rangle = \langle \nu, \phi \rangle \text{ for all } \phi \in X,$$

then $\mu = \nu$.

The fundamental result of this section is:

Theorem 4.3.6. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup on Ω and $\mu \in \mathcal{M}^+(\Omega)$, such that*

(a) $t \mapsto P(t)\nu(E)$ is measurable for every $E \in \Sigma, \nu \in \mathcal{M}^+(\Omega)$,

(b) There is an $\mathcal{M}(\Omega)$-separating set $X \subset \mathrm{BM}(\Omega)$ such that for every $\nu \in \mathcal{M}(\Omega)$ and $\phi \in X$,

$$\lim_{t\downarrow 0} \langle P(t)\nu, \phi \rangle = \langle \nu, \phi \rangle.$$

(c) $j_\mu(L^1_+(\mu))$ is $(P(t))_{t\geq 0}$-invariant.

Let $(T(t))_{t\geq 0}$ be the semigroup on $L^1(\mu)$ induced by $(P(t))_{t\geq 0}$. Then the following statements are equivalent:

(i) $(T(t))_{t\geq 0}$ is strongly continuous.

(ii) The map $t \mapsto P(t)\mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$.

(iii) For all $\tau > 0$ and for any $f \in L^1(\mu)$ the partial orbit $\mathcal{O}_\tau := \{T(t)f : 0 \leq t \leq \tau\}$ is weakly compact.

(iv) There exists $\tau > 0$ such that for any sequence A_n of disjoint sets in Σ,

$$\lim_{n\to\infty} \sup_{0\leq t\leq \tau} P(t)\mu(A_n) = 0. \tag{4.2}$$

Proof. (i) ⇒(ii). From Lemma 4.3.1 it follows that for every $s, t \in \mathbb{R}_+$,

$$\|P(t)\mu - P(s)\mu\|_{\mathrm{TV}} = \|P(t)j_\mu(\mathbb{1}) - P(s)j_\mu(\mathbb{1})\|_{\mathrm{TV}} = \|T(t)\mathbb{1} - T(s)\mathbb{1}\|_1. \tag{4.3}$$

By assumption $t \mapsto T(t)\mathbb{1}$ is continuous from $\mathbb{R}_+$ to $L^1(\mu)$, hence $t \mapsto P(t)\mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$.

(ii)⇒(iii). Equation (4.3) and assumption *(ii)* yield that $t \mapsto T(t)\mathbb{1}$ is continuous from $\mathbb{R}_+$ to $L^1(\mu)$.

Let $\tau > 0$. By continuity the partial orbit $\{T(t)\mathbb{1} : 0 \leq t \leq \tau\}$ is norm compact, hence weakly compact.

(iii)⇒(iv) According to Theorem 4.3.4, for any sequence of disjoint measurable sets A_n,

$$0 = \lim_{n\to\infty} \sup_{0\le t\le\tau} \int_{A_n} |T(t)\mathbb{1}| d\mu = \lim_{n\to\infty} \sup_{0\le t\le\tau} P(t)\mu(A_n).$$

(iv)⇒(i). Recall (4.1), which implies $\|T(t)\| = 1$ for all $t \ge 0$.
Let $f \in L^1(\mu)$, $g \in L^\infty(\mu)$ and $\hat{g}$ a representative of g in $\mathrm{BM}(\Omega)$. By Proposition 3.2.2, the map

$$t \mapsto \langle T(t)f, g\rangle = \langle P(t)j_\mu(f), \hat{g}\rangle$$

is measurable, thus $t \mapsto T(t)f$ is weakly measurable.

We will show that for every $f \in L^\infty(\mu)$, $t \mapsto T(t)f : \mathbb{R}_+ \to L^1(\mu)$ is weakly continuous at zero. Since $L^\infty(\mu)$ is norm dense in $L^1(\mu)$, this implies that all the conditions of Proposition 4.3.5 are satisfied, hence $(T(t))_{t\ge0}$ is a C_0-semigroup on $L^1(\mu)$.

First we consider $f \in L^\infty_+(\mu)$ and we will prove that the partial orbit $\mathcal{O}_\tau = \{T(t)f : 0 \le t \le \tau\}$ is relatively weakly compact in $L^1(\mu)$. According to Theorem 4.3.4 it suffices to show that for any sequence of disjoint measurable subsets A_n of S,

$$\lim_{n\to\infty} \sup_{0\le t\le\tau} \int_{A_n} T(t)f d\mu = 0. \tag{4.4}$$

Since

$$0 \le j_\mu(f) \le j_\mu(\|f\|_\infty \mathbb{1}) = \|f\|_\infty \mu,$$

we obtain

$$\int_{A_n} T(t)f d\mu = P(t)j_\mu(f)(A_n) \le \|f\|_\infty P(t)\mu(A_n).$$

Thus (4.4) holds by (4.2). By the Eberlein-Šmulian Theorem (see e.g. [31, Theorem V.6.1]) C is relatively weakly sequentially compact in $L^1(\mu)$.

Suppose that $t \mapsto T(t)f : \mathbb{R}_+ \to L^1(\mu)$ is not weakly continuous at zero. Then there exists $g \in L^\infty(\mu)$, $\epsilon > 0$ and a sequence $(t_n)_n \subset [0, \tau]$, such that $t_n \downarrow 0$ and $|\langle T(t_n)f - f, g\rangle| \ge \epsilon$ for every $n \in \mathbb{N}$. Since C is relatively weakly sequentially compact in $L^1(\mu)$, there is a subsequence $(t_{n_k})_k$ and a $h \in L^1(\mu)$, such that $T(t_{n_k})f \to h$ weakly in $L^1(\mu)$. In particular for every $\phi \in X \subset \mathrm{BM}(S)$, $\langle T(t_{n_k})f, \phi\rangle \to \langle h, \phi\rangle$. But

$$\langle T(t_{n_k})f, \phi\rangle = \langle P(t_{n_k})j_\mu(f), \phi\rangle \to \langle j_\mu(f), \phi\rangle,$$

for every $\phi \in X$ by assumption (b). Thus $\langle j_\mu(f), \phi\rangle = \langle j_\mu(h), \phi\rangle$ for every $\phi \in X$. So $f = h$, which yields a contradiction.

For general $f \in L^\infty(\mu)$, we can write $f = f^+ - f^-$, thus $t \mapsto T(t)f$ is weakly continuous at zero as well.

□

Remark. (on the proof of Theorem 4.3.6)
Let μ be as in the theorem. Then $\|P(t)j_\mu(f) - P(s)j_\mu(f)\|_{\mathrm{TV}} = \|T(t)f - T(s)f\|_1$ for every $t, s \in \mathbb{R}_+$ and $f \in L^1(\mu)$. This is what we use to prove *(i)*$\Rightarrow$*(ii)*. However it does not seem to be possible to prove *(ii)* $\Rightarrow$ *(i)* directly using this identity, since the continuity of $t \mapsto P(t)j_\mu(f)$ for general $f \in L^1(\mu)$ is not an easy consequence of the continuity of $t \mapsto P(t)\mu$. For this the 'detour' we take via *(iii)* and *(iv)* seems to be necessary.

Example. Let $(P(t))_{t\geq 0}$ be a Markov semigroup on Ω for which conditions (a) and (b) of Theorem 4.3.6 hold. If $\mu \in \mathcal{M}^+(\Omega)$ is an invariant measure of $(P(t))_{t\geq 0}$, then $(P(t))_{t\geq 0}$ leaves $j_\mu(L^1_+(\mu))$ invariant by Corollary 4.3.3. Also, condition *(ii)* of Theorem 4.3.6 is satisfied, so the induced semigroup on $L^1(\mu)$ is strongly continuous.

We cannot simply drop the continuity-at-zero condition (b), as the following example shows:

Example. Let (Ω, Σ) be a measurable space and fix a $\mu \in \mathcal{P}(\Omega)$. For every $\nu \in \mathcal{P}(\Omega)$ we define

$$P(t)\nu = \begin{cases} \nu & \text{if } t = 0; \\ \mu & \text{if } t > 0. \end{cases}$$

Then we can extend $P(t)$ to a positively homogeneous and additive map from $\mathcal{M}^+(\Omega)$ to $\mathcal{M}^+(\Omega)$ and it is straightforward to show that $P(t)$ is a regular Markov operator for every $t \in \mathbb{R}_+$. $(P(t))_{t\geq 0}$ is a Markov semigroup and for every $E \in \Sigma$, $t \mapsto P(t)\mu(E)$ is measurable. Let X be any $\mathcal{M}(\Omega)$-separating subset of $\mathrm{BM}(\Omega)$ and let $\nu \in \mathcal{P}(\Omega)$ be unequal to μ, then there must be an $f \in X$ such that $\langle \nu, f\rangle \neq \langle \mu, f\rangle$, so $\langle P(t)\nu, f\rangle = \langle \nu, f\rangle$ cannot converge to $\langle \mu, f\rangle$ as $t \downarrow 0$. Since $P(t)\mu = \mu$ for all $t \in \mathbb{R}_+$, Lemma 4.3.3 implies $(P(t))_{t\geq 0}$ leaves $j_\mu(L^1(\mu))$ invariant. Moreover, $t \mapsto P(t)\mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$. However, if there exists any $\nu \in \mathcal{P}(\Omega)$ with $\nu \ll \mu$ and $\nu \neq \mu$, then clearly the induced semigroup on $L^1(\mu)$ is not strongly continuous.

We will show in the next section that the continuity-at-zero condition (b) *can* be dropped when we consider Markov semigroups coming from an underlying semigroup of measurable maps on Ω.

In the remainder of this section we consider some applications of Theorem 4.3.6 in certain topological spaces. Recall the definitions of Baire σ-algebra, normal and perfectly normal spaces from Section 3.4.

We define an $\mathcal{M}(\Omega)$-*norming set* to be a set $X \subset \mathrm{BM}(\Omega)$, such that

$$\|\mu\|_{\mathrm{TV}} = \sup\{|\langle \mu, f\rangle| : f \in X, \|f\|_\infty \leq 1.\}$$

Clearly every $\mathcal{M}(\Omega)$-norming set is an $\mathcal{M}(\Omega)$-separating set. The following result can be found in [77, Corollary 3.9]:

Lemma 4.3.7. *Let S be a normal space endowed with its Baire σ-algebra. Then $C_b(S)$ is an $\mathcal{M}(S)$-norming set.*

Corollary 4.3.8. *Let S be a normal space endowed with its Baire σ-algebra, $(P(t))_{t\geq 0}$ a Markov semigroup on S and $\mu \in \mathcal{M}^+(S)$. If*

$$t \mapsto P(t)\mu(E) \text{ is measurable for every Baire measurable } E \subset S,$$

$(P(t))_{t\geq 0}$ is strongly stochastically continuous at zero and $j_\mu(L^1_+(\mu))$ is $(P(t))_{t\geq 0}$-invariant. Then the conclusion in Theorem 4.3.6 on the equivalence of statements (i)-(iv) holds.

Proof. Conditions $(a), (b)$ and (c) in Theorem 4.3.6 hold by Lemma 4.3.7. □

Corollary 4.3.9. *Let S be a perfectly normal space, $(P(t))_{t\geq 0}$ a Markov semigroup on S and $\mu \in \mathcal{M}^+(S)$. If $(P(t))_{t\geq 0}$ is right strongly stochastically continuous and $j_\mu(L^1_+(\mu))$ is $(P(t))_{t\geq 0}$-invariant. Then the conclusion in Theorem 4.3.6 on the equivalence of statements (i)-(iv) holds.*

Proof. The statement follows from Corollary 4.3.8 once we can show that $t \mapsto P(t)\nu(E)$ is measurable for every Borel (= Baire) measurable subset E and $\nu \in \mathcal{M}^+(S)$. Fix $\nu \in \mathcal{M}^+(S)$. By right continuity, $t \mapsto \langle P(t)\nu, f\rangle$ is measurable for every $f \in C_b(S)$. Let C be closed in S. Then there exists a sequence $(U_n)_n$ of open subsets of S, such that $C = \cap_{n=1}^\infty U_n$. By the Tietze Extension Theorem (see e.g. [44, Theorem 4.16]), there exist $f_n : S \to [0,1]$ continuous, such that $f_n|_C = 1\!\!1_C$ and $f_n|_{S\setminus U_n} = 0$. Then $f_n \to 1\!\!1_C$ pointwise, so by the Dominated Convergence Theorem, $\langle P(t)\nu, f_n\rangle \to P(t)\nu(C)$ for all $t \in \mathbb{R}_+$. So $t \mapsto P(t)\nu(C)$ is a pointwise limit of measurable functions, hence measurable. By Proposition 3.2.2, $t \mapsto P(t)\nu(E)$ is measurable for all Borel sets E. □

Corollary 4.3.10. *Let S be a perfectly normal space, $(P(t))_{t\geq 0}$ a regular Markov semigroup that is strongly stochastically continuous at zero and $\mu \in \mathcal{M}^+(S)$ such that $j_\mu(L^1_+(\mu))$ is $(P(t))_{t\geq 0}$-invariant. Then the conclusion in Theorem 4.3.6 on the equivalence of statements (i)-(iv) holds.*

Proof. Follows from Corollary 4.3.9 and Lemma 3.4.4. □

In particular, the previous corollaries all hold on a metric space.

Not every strongly stochastically continuous Markov semigroup on a metric space which leaves $j_\mu(L^1_+(\mu))$ invariant for some $\mu \in \mathcal{M}^+(S)$ satisfies the equivalent conditions *(i) – (iv)* of Theorem 4.3.6, as the following example will show.

Example. Let m denote the Lebesgue measure on $\mathbb{R}^n$. The diffusion semigroup $(T_d(t))_{t\geq 0}$ on $L^1(\mathbb{R}^n) = L^1(\mathbb{R}^n, m)$ is given by

$$T_d(t)f(x) := \int_{\mathbb{R}^n} h_d(x-y,t)f(y)dm(y), \text{ for } t > 0,$$

where the diffusion kernel h_d is given by

$$h_d(x,t) = (4\pi dt)^{-n/2} e^{-|x|^2/4dt}.$$

Let $\mu \in \mathcal{M}(\mathbb{R}^n)$, then one can show that $x \mapsto g_\mu(x) = \int_{\mathbb{R}} h_d(x-y,t)f(y)d\mu(y)$ is in $L^1(\mathbb{R}^n)$, and hence defines a measure $g_\mu dm$. We can extend $T_d(t)$ to a map $P_d(t) : \mathcal{M}(\mathbb{R}^n) \to \mathcal{M}(\mathbb{R}^n)$, by defining $P_d(t)\mu$ to be $g_\mu dm$. Then $P_d(t)$ is linear, leaves $\mathcal{M}^+(\mathbb{R}^n)$ invariant, and $\|P_d(t)\mu\|^*_{\mathrm{BL}} \leq \|\mu\|^*_{\mathrm{BL}}$ for all $\mu \in \mathcal{M}(\mathbb{R}^n)$, so $P_d(t)$ can be extended to a bounded linear operator $(\overline{P_d(t)})_{t\geq 0}$ on $\mathbb{R}^n_{\mathrm{BL}}$. Moreover, $(\overline{P_d(t)})_{t\geq 0}$ is strongly continuous on $\mathbb{R}^n_{\mathrm{BL}}$. Hence $(P_d(t))_{t\geq 0}$ is a strongly stochastically continuous Markov semigroup by Lemma 3.4.4. Note that $P_d(t)\mu \ll m$ for every $\mu \in \mathcal{M}(\mathbb{R}^n)$ and $t > 0$. Now let $f \in L^1(\mathbb{R}^n)$ such that $f > 0$ almost everywhere, and set $\mu = f\,dm + \delta_0$. Then $P_d(t)(\mu) \ll \mu$ for all $t \geq 0$, so $(P_d(t))_{t\geq 0}$ leaves $L^1_+(\mu)$ invariant by Corollary 4.3.3. But $P_d(t)\mu \in L^1(\mathbb{R}^n)$ for all $t > 0$, hence $t \mapsto P_d(t)\mu$ cannot be continuous from $\mathbb{R}_+$ to $\mathcal{M}(\mathbb{R}^n)$, since $L^1(\mathbb{R}^n)$ is closed in $\mathcal{M}(\mathbb{R}^n)$ and $\mu \notin L^1(\mathbb{R}^n)$, so condition *(ii)* of Theorem 4.3.6 is not satisfied. Note that the diffusion semigroup is strongly continuous on $L^1(\mathbb{R}^n)$, so $t \mapsto P_d(t)\mu$ *is* continuous from $(0,\infty)$ to $\mathcal{M}(\mathbb{R}^n)$.

The final observation of the example above holds in greater generality:

Proposition 4.3.11. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup on Ω and $\mu \in \mathcal{M}^+(\Omega)$ such that the conditions (a) and (c) are satisfied, and such that $L^1(\mu)$ is separable. Then $t \mapsto P(t)\mu$ is continuous from $(0,\infty)$ to $\mathcal{M}(\Omega)$.*

Proof. Let $(T(t))_{t\geq 0}$ be the induced semigroup on $L^1(\mu)$. Then $(T(t))_{t\geq 0}$ is weakly measurable by the same argument as the one stated in the first part of the proof of Theorem 4.3.6*(iv)⇒(i)*, thus strongly measurable by the Pettis Measurability Theorem. Then [59, Theorem 10.2.3] implies that, for every $f \in L^1(\mu)$, $t \mapsto T(t)f$ is continuous from $(0,\infty)$ to $L^1(\mu)$, so $t \mapsto P(t)\mu = j_\mu(T(t)\mathbb{1})$ is continuous from $(0,\infty)$ to $\mathcal{M}(\Omega)$. □

Note that $L^1(\mu)$ need not be separable, even if μ is finite. An example is given by uncountable copies of the unit interval with Lebesgue measure. See [55, Section 42] for a chacterisation when $L^1(\mu)$ is separable in terms of properties of the measure μ. Further on we will show that a similar statement as that of Proposition 4.3.11 holds without the separability assumption, but with some other assumptions.

4.4 Strong continuity for total variation norm

Let $(P(t))_{t\geq 0}$ be a jointly measurable Markov semigroup on Ω. It extends to a positive semigroup of bounded linear operators on $\mathcal{M}(\Omega)$ as we have seen. Typically the latter is not strongly continuous. In this section we will give several characterisations of the closed invariant subspace of $\mathcal{M}(\Omega)$ on which $(P(t))_{t\geq 0}$ *is* strongly continuous, i.e. the space

$$\mathcal{M}(\Omega)^0_{\mathrm{TV}} := \{\mu \in \mathcal{M}(\Omega) : t \mapsto P(t)\mu \text{ is continuous from } \mathbb{R}_+ \text{ to } \mathcal{M}(\Omega)\}$$

of C^0-vectors in $\mathcal{M}(\Omega)$ for the $\|\cdot\|_{\mathrm{TV}}$-topology.

Our approach is inspired by that of Gulick *et al.* [51]. There the following situation is considered: A locally compact *group* G acts as a group of homeomorphisms $(\Phi_g)_{g\in G}$ on a locally compact Hausdorff space X, sending $x \in X$ to $\Phi_g(x)$. This induces an action $(P(g))_{g\in G}$ on the Banach space of bounded Radon measures on X, $\mathcal{M}(X)$, endowed with total variation norm, given by $P(g)\mu(E) := \mu(\Phi_{g^{-1}}E)$. The subspace of $\mathcal{M}(X)$, consisting of measures μ such that $g \mapsto P(g)\mu$ is continuous from G to $\mathcal{M}(X)$ is then identified using convolution of certain functions on G with Radon measures on X, and this identification is used to provide several characterisations of this subspace (see also [82]).

Adopting this approach to our setting is not straightforward: instead of a group G as in [51], we consider a (specific) semigroup $\mathbb{R}_+$, which implies that actions need not be invertible. Also, in [51] an action of the group on the underlying space X is considered, which induces an action on $\mathcal{M}(X)$. While we look, more generally, at actions of $\mathbb{R}_+$ on $\mathcal{M}(\Omega)$ directly, that contain those coming from an underlying action on Ω by Proposition 3.4.2, which need not be continuous, only measurable. Furthermore, in [51] X must be locally compact, since measures on X are defined there by constructing certain functionals on $C_0(X)$; in a large part of our setting we can consider the generality of a measurable space (Ω, Σ). Our approach is based on the theory of integrating functions with values in $\mathcal{M}(\Omega)$ as given in Section 3.2 and on Theorem 4.3.6. It provides a characterisation of $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ analogous to those in [51] and [82].

These characterisations will help in identifying when the restriction of $(P(t))_{t\geq 0}$ to invariant L^1-spaces is strongly continuous.

The fundamental result from [51] that we need is a result on Banach modules. Let A be a Banach algebra with multiplication $*$. A net (e_α) in A is an *approximate identity* of A, if $\lim_\alpha e_\alpha * f = f$ and $\lim_\alpha f * e_\alpha = f$ for all $f \in A$. It is a *bounded* approximate identity if the net is bounded. A Banach space M is a *Banach module* over A if there exists a bilinear map $\star : A \times M \to M$ having the following properties:

(BM1) $(f * g) \star m = f \star (g \star m)$ for all $f, g \in A, m \in M$.

(BM2) $\|f \star m\|_M \leq \|f\|_A \|m\|_M$ for all $f \in A, m \in M$.

Proposition 4.4.1. ([51, Corollary 2.3]) *Let A be a Banach algebra with bounded approximate identity (e_α). If M is a Banach module over A, then $A \star M := \{a \star m : a \in A, m \in M\}$ is a closed subspace of M. In particular, $m \in A \star M$ if and only if $\lim_\alpha e_\alpha \star m = m$.*

The latter characterisation of elements in $A \star M$ shows that $A \star M$ is indeed a vector subspace of M.

Proposition 4.4.2. *The Banach space $L^1(\mathbb{R}_+)$ is a commutative Banach algebra with multiplication defined by convolution:*

$$f * g(t) := \int_0^t f(t-s)g(s)\,ds,$$

with bounded approximate identity (e_n) *given by* $e_n = n\mathbb{1}_{[0,\frac{1}{n}]}$.

The proof is straightforward, observing that $L^1(\mathbb{R}_+)$ is canonically contained as closed subspace in the commutative Banach algebra $L^1(\mathbb{R})$ with convolution.

By joint measurability of the Markov semigroup and Proposition 3.2.3, we can define for $f \in L^1(\mathbb{R}_+)$ and $\mu \in \mathcal{M}(\Omega)$ the set-wise integral

$$f *_P \mu := \int_{\mathbb{R}_+} f(s)P(s)\mu \, ds \in \mathcal{M}(\Omega),$$

since $\|f(\cdot)P(\cdot)\mu\|_{\mathrm{TV}} \leq |f(\cdot)|$ (Lebesgue) almost everywhere. Clearly $(f,\mu) \mapsto f *_P \mu$ is a bilinear map from $L^1(\mathbb{R}_+) \times \mathcal{M}(\Omega)$ into $\mathcal{M}(\Omega)$.

The right translation semigroup $(R_+(t))_{t\geq 0}$ on $L^1(\mathbb{R}_+)$ is given by:

$$R_+(t)f(s) := \begin{cases} f(s-t), & \text{if } s \geq t, \\ 0, & \text{if } 0 \leq s \leq t. \end{cases}.$$

It is a strongly continuous positive semigroup on $L^1(\mathbb{R}_+)$.

Proposition 4.4.3. *Let* $P(t)$ *be regular for some* $t \in \mathbb{R}_+$. *Then*

$$P(t)(f *_P \mu) = f *_P (P(t)\mu) = (R_+(t)f) *_P \mu$$

for all $f \in L^1(\mathbb{R}_+), \mu \in \mathcal{M}(\Omega)$.

Proof. Since $P(t)$ is regular, there exists an $U(t) : \mathrm{BM}(\Omega) \to \mathrm{BM}(\Omega)$ such that $\langle P(t)\mu, h\rangle = \langle \mu, U(t)h\rangle$ for all $\mu \in \mathcal{M}(\Omega), h \in \mathrm{BM}(\Omega)$.

Let $h \in \mathrm{BM}(S)$.

$$\langle P(t)(f *_P \mu), h\rangle = \langle f *_P \mu, U(t)h\rangle = \int_\Omega U(t)h \, d\left[\int_{\mathbb{R}_+} f(s)P(s)\mu \, ds\right].$$

By Proposition 3.2.3 we get

$$\begin{aligned} \int_\Omega U(t)h \, d\left[\int_{\mathbb{R}_+} f(s)P(s)\mu \, ds\right] &= \int_{\mathbb{R}_+} \langle P(s)\mu, U(t)h\rangle f(s) \, ds \\ &= \int_{\mathbb{R}_+} \langle P(t+s)\mu, h\rangle f(s) \, ds. \end{aligned} \tag{4.5}$$

Again applying Proposition 3.2.3 yields that (4.5) equals

$$\int_\Omega h \, d\left[\int_{\mathbb{R}_+} f(s)P(t+s)\mu \, ds\right] = \langle f *_P (P(t)\mu), h\rangle.$$

Hence $P(t)(f *_P \mu) = f *_P (P(t)\mu)$.

For every $E \in \Sigma$, the map $s \mapsto f(s)P(t+s)\mu(E)$ is integrable from $\mathbb{R}_+$ to $\mathbb{R}$. Using the fact that Lebesgue measure on $\mathbb{R}$ is invariant under translation,

$$\int_{\mathbb{R}_+} f(s)P(t+s)\mu(E)\,ds = \int_{\mathbb{R}_+} (R_+(t)f)(s)P(s)\mu(E)\,ds = (R_+(t)f) *_P \mu(E).$$

□

From this point on we will implicitly assume that $(P(t))_{t\geq 0}$ is regular. Because we also assume that $(P(t))_{t\geq 0}$ is jointly measurable, condition (a) of Theorem 4.3.6 is automatically satisfied by Lemma 3.4.1, i.e. $t \mapsto P(t)\nu(E)$ is measurable for all $E \in \Sigma$ and $\nu \in \mathcal{M}^+(\Omega)$.

Proposition 4.4.4. *Let $f, g \in L^1(\mathbb{R}_+)$ and $\mu \in \mathcal{M}(\Omega)$, then*

(i) $(f * g) *_P \mu = f *_P (g *_P \mu)$

(ii) $\|f *_P \mu\|_{\mathrm{TV}} \leq \|f\|_1 \|\mu\|_{\mathrm{TV}}$.

Consequently, $\mathcal{M}(\Omega)$ is a Banach module over $L^1(\mathbb{R}_+)$.

Proof. We first prove *(i)*. We use Fubini's Theorem and Proposition 4.4.3: For every $E \in \Sigma$

$$\begin{aligned}
(f * g) *_P \mu(E) &= \int_{\mathbb{R}_+} \left\{ \int_{\mathbb{R}_+} f(s)g(t-s)\,ds \right\} P(t)\mu(E)\,dt \\
&= \int_{\mathbb{R}_+} f(s) \left\{ \int_{\mathbb{R}_+} g(t-s)P(t)\mu(E)\,dt \right\} ds \\
&= \int_{\mathbb{R}_+} f(s)(R_+(s)g) *_P \mu(E)\,ds \\
&= \int_{\mathbb{R}_+} f(s)[P(s)(g *_P \mu)](E)\,ds = f *_P (g *_P \mu)(E).
\end{aligned}$$

For part *(ii)* let first $f \in L^1_+(\mathbb{R}_+)$ and $\nu \in \mathcal{M}^+(\Omega)$. Then $f *_P \nu \in \mathcal{M}^+(\Omega)$ and

$$\|f *_P \nu\|_{\mathrm{TV}} = f *_P \nu(\Omega) = \int_{\mathbb{R}_+} f(t)P(t)\nu(\Omega)\,dt = \|\nu\|_{\mathrm{TV}}\|f\|_1$$

by property *(MO2)*. For general $f \in L^1(\mathbb{R}_+)$ and $\mu \in \mathcal{M}(S)$ we then obtain

$$\begin{aligned}
\|f *_P \mu\|_{\mathrm{TV}} &= \|(f^+ - f^-) *_P (\mu^+ - \mu^-)\|_{\mathrm{TV}} \\
&\leq (\|f^+\|_1 + \|f^-\|_1)\|\mu^+\|_{\mathrm{TV}} + (\|f^+\|_1 + \|f^-\|_1)\|\mu^-\|_{\mathrm{TV}} \\
&= \|f\|_1(\|\mu^+\|_{\mathrm{TV}} + \|\mu^-\|_{\mathrm{TV}}) = \|f\|_1\|\mu\|_{\mathrm{TV}},
\end{aligned}$$

by using bilinearity and the fact that $\mathcal{M}(\Omega)$ and $L^1(\mathbb{R}_+)$ are L-spaces.

So $\mathcal{M}(\Omega)$ is a Banach module over $L^1(\mathbb{R}_+)$. □

Put $L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega) := \{f *_P \mu : f \in L^1(\mathbb{R}_+), \mu \in \mathcal{M}(\Omega)\}$. Then we have, by Proposition 4.4.2, Proposition 4.4.4 and Proposition 4.4.1, the following result:

Corollary 4.4.5. *$L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$ is a non-trivial closed subspace of $\mathcal{M}(\Omega)$.*

This closed subspace equals the subspace of strong continuity of $P(t)$ with respect to $\|\cdot\|_{\mathrm{TV}}$:

Theorem 4.4.6. *For $\mu \in \mathcal{M}(\Omega)$ the following are equivalent:*

(i) $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, *i.e.* $t \mapsto P(t)\mu : \mathbb{R}_+ \to \mathcal{M}(\Omega)$ *is continuous.*

(ii) $\mu \in L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$.

Proof. (i)⇒ (ii): Let $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$. By Proposition 4.4.1 it is sufficient to show that $e_n * \mu \to \mu$. Let $\epsilon > 0$. Since $t \mapsto P(t)\mu : \mathbb{R}_+ \to \mathcal{M}(\Omega)$ is continuous, there exists an $N \in \mathbb{N}$, such that $\|P(t)\mu - \mu\|_{\mathrm{TV}} \leq \epsilon$ for all $t \in [0, \frac{1}{N}]$. For $n \in \mathbb{N}$

$$e_n * \mu - \mu = n \int_0^{\frac{1}{n}} P(t)\mu - \mu \, dt$$

is defined set-wise.

By continuity, $t \mapsto P(t)\mu - \mu : [0, \frac{1}{n}] \to \mathcal{M}(\Omega)$ is strongly measurable and bounded, hence Bochner integrable, so we can also view the integral $n \int_0^{\frac{1}{n}} P(t)\mu - \mu \, ds$ as a Bochner integral in $\mathcal{M}(\Omega)$ (endowed with the total variation norm). Bounded linear operators on $\mathcal{M}(\Omega)$ can be brought inside of the Bochner integral (see e.g. [26, Theorem II.2.6]), so in particular the Bochner integral agrees with the set-wise integral.

Moreover,

$$\|e_n *_P \mu - \mu\|_{\mathrm{TV}} \leq n \int_0^{\frac{1}{n}} \|P(t)\mu - \mu\|_{\mathrm{TV}} \, dt \leq \epsilon,$$

for all $n \geq N$.

(ii)⇒ (i): Let $\mu = f *_P \nu \in L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$. It suffices to prove continuity at zero. According to Proposition 4.4.3 and Proposition 4.4.4,

$$\begin{aligned} \|P(t)\mu - \mu\|_{\mathrm{TV}} &= \|(R_+(t)f) *_P \nu - f *_P \nu\|_{\mathrm{TV}} \\ &\leq \|R_+(t)f - f\|_1 \|\nu\|_{\mathrm{TV}}. \end{aligned}$$

Since $(R_+(t))_{t\geq 0}$ is strongly continuous on $L^1(\mathbb{R}_+)$, $\|R_+(t)f - f\|_1 \to 0$ as $t \downarrow 0$ and thus $\|P(t)\mu - \mu\|_{\mathrm{TV}} \to 0$. So $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$. □

Thus from now on we can identify $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ with $L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$.

Example 4.4.7. Let $(\Phi_t)_{t\geq 0}$ be a jointly measurable semigroup of measurable maps on Ω and $(P_\Phi(t))_{t\geq 0}$ the associated Markov semigroup. Let μ be a σ-finite measure on Ω such that

$$m(\{t \in \mathbb{R}_+ : \Phi_t(x) \in N\}) = 0 \text{ for all } \mu\text{-null sets } N \in \Sigma \tag{4.6}$$

where m is the Lebesgue measure on $\mathbb{R}$. Then $\mathcal{M}(\Omega)^0_{\mathrm{TV}} \subset j_\mu(L^1(\mu))$: let $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, then $\nu = f *_P \rho$ for some $f \in L^1(\mathbb{R}_+)$ and $\rho \in \mathcal{M}(\omega)$. Let N be a μ-null set, then

$$\begin{aligned}\nu(N) &= \int_{\mathbb{R}_+} f(t)P_\Phi(t)\rho(N)\,dt = \int_\Omega \int_{\mathbb{R}_+} f(t)P_\Phi(t)\delta_x(N)\,dt\,d\rho(x)\\ &= \int_\Omega \int_{\mathbb{R}_+} f(t)\delta_{\Phi_t(x)}(N)\,dt\,d\rho(x) = 0\end{aligned}$$

by (4.6), thus $\nu \ll \mu$. A simple example of a semigroup satisfying the above conditions is given by $\Omega = \mathbb{R}$ and $\Phi_t(x) = x + tv$ for some $v \in \mathbb{R}, v \neq 0$. So in this case $\mathcal{M}(\mathbb{R})^0_{\mathrm{TV}} \subset j_m(L^1(m))$. In fact, equality holds, as we shall see in Example 4.4.12.

We can use Theorem 4.4.6 to show that part of the conclusion of Theorem 4.3.6 holds for certain Markov semigroups that need not have the continuity-at-zero condition (b) from that theorem:

Proposition 4.4.8. *Let $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup, with dual semigroup $(U(t))_{t\geq 0}$, such that*

$$U(t)|f| = |U(t)f| \text{ for every } f \in \mathrm{BM}(\Omega), t \in \mathbb{R}_+. \tag{4.7}$$

Let $\mu \in \mathcal{M}^+(\Omega)$ be such that $j_\mu(L^1_+(\mu))$ is $(P(t))_{t\geq 0}$-invariant. Let $(T(t))_{t\geq 0}$ be the induced semigroup on $L^1(\mu)$. Then the following statements are equivalent:

(i) $(T(t))_{t\geq 0}$ is strongly continuous.

(ii) The map $t \mapsto P(t)\mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$.

Proof. The proof of *(i)⇒(ii)* goes as in the proof of Theorem 4.3.6. It remains to show that *(ii)⇒(i)*. As in the proof of Theorem 4.3.6*(iv)⇒(i)* we want to apply Proposition 4.3.5.

As we remarked in (4.1), $\|T(t)\| = 1$ for all $t \geq 0$. Let $f \in L^1(\mu)$, $g \in L^\infty(\mu)$ and $\hat{g}$ a representative of g in $\mathrm{BM}(\Omega)$. By Proposition 3.2.2 and joint measurability of $(P(t))_{t\geq 0}$,

$$t \mapsto \langle T(t)f, g\rangle = \langle P(t)j_\mu(f), \hat{g}\rangle$$

is measurable, thus $t \mapsto T(t)f$ is weakly measurable.

We will show that for every $f \in L^\infty(\mu)$, $t \mapsto T(t)f : \mathbb{R}_+ \to L^1(\mu)$ is weakly continuous at zero. Since $L^\infty(\mu)$ is norm dense in $L^1(\mu)$, this implies that all the conditions of Proposition 4.3.5 are satisfied, hence $(T(t))_{t\geq 0}$ is a C_0-semigroup on $L^1(\mu)$.

Let $f \in L^\infty_+(\mu)$ and $\nu := j_\mu(f) \in \mathcal{M}^+(\Omega)$. Then $0 \le \nu \le \|f\|_\infty \mu$. Let $g \in L^\infty(\mu)$ and $\hat{g}$ a representative of g in $\mathrm{BM}(\Omega)$. By assumption, $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, so Theorem 4.4.6 implies that there exists a $\rho \in \mathcal{M}(\Omega)$ and $h \in L^1(\mathbb{R}_+)$, such that $\mu = h *_P \rho$. Let $\tau \in \mathbb{R}_+$. By Proposition 3.2.3 we have

$$\begin{aligned}\int_\Omega |U(t)\hat{g} - U(\tau)\hat{g}|\, d\mu &= \int_{\mathbb{R}_+} \langle h(s)P(s)\rho, |U(t)\hat{g} - U(\tau)\hat{g}|\, ds \\ &= \int_{\mathbb{R}_+} h(s)\langle \rho, U(s)|U(t)\hat{g} - U(\tau)\hat{g}|\, ds \\ &=^{(4.7)} \int_{\mathbb{R}_+} h(s)\langle \rho, |U(s+t)\hat{g} - U(s+\tau)\hat{g}|\, ds.\end{aligned}$$

Now we can apply Fubini's Theorem, by joint measurability, to obtain

$$\int_\Omega |U(t)\hat{g} - U(\tau)\hat{g}|\, d\mu = \int_\Omega \int_{\mathbb{R}_+} h(s)|U(s+t)\hat{g}(x) - U(s+\tau)\hat{g}(x)|\, ds\, d\rho(x).$$

Now fix $x \in \Omega$. Then $k(s) := U(s)\hat{g}(x)$ is a bounded measurable function on $\mathbb{R}_+$ and the Dominated Convergence Theorem yields

$$\int_{\mathbb{R}_+} h(s)|k(s+t) - k(s+\tau)|\, ds \to 0 \text{ as } t \to \tau.$$

Also, $\int_{\mathbb{R}_+} |h(s)||U(s+t)\hat{g}(x) - U(s+\tau)\hat{g}(x)|\, ds \le 2\|h\|_\infty \|\hat{g}\|_1$. So again by the Dominated Convergence Theorem,

$$\int_\Omega |U(t)\hat{g} - U(\tau)\hat{g}|\, d\mu \to 0 \text{ as } t \to \tau. \tag{4.8}$$

From (4.8) it now simply follows that

$$\int_\Omega |U(t)\hat{g} - U(\tau)\hat{g}|\, d\nu \to 0 \text{ as } t \to \tau.$$

In particular, $t \mapsto \langle \nu, U(t)\hat{g}\rangle = \langle T(t)f, g\rangle$ is continuous from $\mathbb{R}_+$ to $\mathbb{R}$. This holds for all $g \in L^\infty(\mu)$. For general $f \in L^\infty(\mu)$, we can write $f = f^+ - f^-$ and obtain that $t \mapsto T(t)f$ is weakly continuous at zero as well. □

This result has the following important application:

Corollary 4.4.9. *Let $(\Phi_t)_{t\ge0}$ be a jointly measurable semigroup of measurable maps on Ω, let $(P_\Phi(t))_{t\ge0}$ be the associated Markov semigroup on Ω and let $\mu \in \mathcal{M}^+(\Omega)$ such that $j_\mu(L^1_+(\mu))$ is $(P_\Phi(t))_{t\ge0}$-invariant. Let $(T(t))_{t\ge0}$ be the induced semigroup on $L^1(\mu)$. Then the following statements are equivalent:*

(i) $(T(t))_{t\ge0}$ is strongly continuous.

(ii) The map $t \mapsto P_\Phi(t)\mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$.

Proof. By Proposition 3.4.2 $(P_\Phi(t))_{t\geq 0}$ is a regular jointly measurable Markov semigroup on Ω. So in order to apply Proposition 4.4.8, we only need to show (4.7). Let $f \in \mathrm{BM}(\Omega)$ and $x \in \Omega$, then

$$U_\Phi(t)|f|(x) = |f|(\Phi_t(x)) = |f(\Phi_t(x))| = |U_\Phi(t)f(x)|.$$

□

Now we provide several equivalent conditions for a measure to be in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$. As in Theorem 4.3.6, we require the following assumption:

(AS) There is an $\mathcal{M}(\Omega)$-separating set $X \subset \mathrm{BM}(\Omega)$ such that for every $\nu \in \mathcal{M}(\Omega)$ and $\phi \in X$,

$$\lim_{t\downarrow 0}\langle P(t)\nu, \phi\rangle = \langle \nu, \phi\rangle.$$

Theorem 4.4.10. *Suppose (AS) or* (4.7) *holds. Let $\mu \in \mathcal{M}(\Omega)$. Then the following are equivalent:*

(i) $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$.

(ii) $t \mapsto P(t)\mu(E)$ *is continuous for every* $E \in \Sigma$.

(iii) If $E \in \Sigma$ is such that $P(t)\mu(E) = 0$ for (Lebesgue) almost every $t \in \mathbb{R}_+$, then $\mu(E) = 0$.

(iv) There is a $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ such that $\mu \ll \nu$.

(v) There is a $\nu \in \mathcal{M}^+(\Omega)^0_{\mathrm{TV}}$ such that $j_\nu(L^1(\nu))$ is $(P(t))_{t\geq 0}$-invariant and $\mu \in j_\nu(L^1(\nu))$.

Proof. *(i)⇒ (ii)*: This follows from the fact that $|\nu(E)| \leq \|\nu\|_{\mathrm{TV}}$ for all $\nu \in \mathcal{M}(\Omega)$ and $E \in \Sigma$.
(ii)⇒(iii): This is trivial.
(iii)⇒ (iv): Let $f \in L^1(\mathbb{R}_+)$, such that $f(t) > 0$ for almost every $t \in \mathbb{R}_+$. Define $\nu := f *_P |\mu|$. Let $E \in \Sigma$ be such that $\nu(E) = 0$, then $P(t)|\mu|(E) = 0$ for almost every $t \in \mathbb{R}_+$. By positivity of $P(t)$, $|P(t)\mu|(E) \leq P(t)|\mu|(E) = 0$ for almost every $t \in \mathbb{R}_+$, hence $\mu(E) = 0$. Thus $\mu \ll \nu$.
(iv)⇒(v): Let $f \in L^1(\mathbb{R}_+)$, such that $f(t) > 0$ for almost every $t \in \mathbb{R}_+$. Define $\rho := f *_P |\nu| \in \mathcal{M}^+(\Omega) \cap L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$. Now, let $E \in \Sigma$ be such that $\rho(E) = 0$. Then $P(t)|\nu|(E) = 0$ for almost every $t \in \mathbb{R}_+$. By positivity of $P(t)$,

$$|P(t)\nu(E)| \leq P(t)|\nu|(E) = 0,$$

for almost every $t \in \mathbb{R}_+$. Since $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, $t \mapsto P(t)\nu(E)$ is continuous, so $\nu(E) = 0$. So $\mu \ll \nu \ll \rho$.

Also, by Proposition 4.4.3, we have for every $s \geq 0$

$$P(s)\rho(E) = \int_{\mathbb{R}_+} f(t)P(t+s)|\nu|(E)\,dt = 0,$$

since $P(t)|\nu|(E) = 0$ for every $t \geq 0$. So $P(t)\rho \ll \rho$ for all $t \geq 0$. According to Corollary 4.3.3 $(P(t))_{t\geq 0}$ leaves $j_\rho(L^1(\rho))$ invariant. We already showed that $\mu \ll \rho$, i.e. $\mu \in j_\rho(L^1(\rho))$.
(v)⇒(i): Since $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, $t \mapsto P(t)\nu : \mathbb{R}_+ \to \mathcal{M}(\Omega)$ is continuous. If *(AS)* holds, then Theorem 4.3.6 implies that the semigroup $(T(t))_{t\geq 0}$ in $L^1(\nu)$ induced by $(P(t))_{t\geq 0}$ is strongly continuous. If (4.7) holds, then the previous statement follows from Proposition 4.4.8 instead. By assumption there is an $f \in L^1(\nu)$ such that $j_\nu(f) = \mu$. Then

$$\|P(t)\mu - \mu\|_{\mathrm{TV}} = \|T(t)f - f\|_1 \to 0,$$

as $t \downarrow 0$. □

Remark. The condition in Theorem 4.4.10, on the existence of an $\mathcal{M}(\Omega)$-separating subset of $\mathrm{BM}(S)$ such that $t \mapsto \langle P(t)\mu, f\rangle$ is continuous at zero for every $f \in X$, is only used in the proof of the implication *(v)⇒(i)*, where we apply Theorem 4.3.6.

Corollary 4.4.11. *Suppose (AS) or* (4.7) *holds. Let* $\mu \in \mathcal{M}^+(\Omega)$. *If there is a* $\tau > 0$ *such that* $\mu \ll P(t)\mu$ *for all* $t \in [0,\tau]$, *then* $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$.

Proof. Let $E \in \Sigma$ be such that $P(t)\mu(E) = 0$ for almost every $t \in \mathbb{R}_+$. Then there is a $t \in [0,\tau]$ such that $P(t)\mu(E) = 0$, and then $\mu(E) = 0$, since $\mu \ll P(t)\mu$. Hence $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10. □

Example 4.4.12. Let $(\Phi_t)_{t\geq 0}$ be a jointly measurable semigroup of measurable maps on Ω and $(P_\Phi(t))_{t\geq 0}$ the associated Markov semigroup. Suppose there exists a σ-finite measure μ on Ω such that for all $E \in \Sigma$ and all $t \in \mathbb{R}_+$

$$\mu(\Phi_t^{-1}(E)) = 0 \text{ if and only if } \mu(E) = 0. \tag{4.9}$$

Since μ is σ-finite, we can construct an $f \in L^1_+(\mu)$ such that $f > 0$ μ-a.e. Let ρ be the measure defined by $\rho = j_\mu(f)$, then $\rho \ll \mu \ll \rho$ and $\rho \in \mathcal{M}^+(\Omega)$. From this and (4.9) it follows that $\rho(\Phi_t^{-1}(E)) = 0$ if and only if $\rho(E) = 0$, so $P_\Phi(t)\rho \ll \rho$ for all $t \in \mathbb{R}_+$. Corollary 4.4.10 implies that $\rho \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, and since $P_\Phi(t)\rho \ll \rho$ as well, $(P_\Phi(t))_{t\geq 0}$ leaves $L^1(\rho)$ invariant, and thus Corollary 4.4.9 yields the strong continuity of the induced semigroup on $L^1(\rho)$. If $\nu \ll \mu$, then $\nu \ll \rho$, thus $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ as well, thus the induced semigroup on $L^1(\mu)$ is also strongly continuous.

A simple example of a semigroup satisfying the above conditions is given by $\Omega = \mathbb{R}$ and $\Phi_t(x) = x + tv$ for some $v \in \mathbb{R}$. Then (4.9) is satisfied when we take μ to be the Lebesgue measure m on $\mathbb{R}$. Thus the induced semigroup on $L^1(\mu)$, given by $T(t)f(x) = f(x - vt)$, is strongly continuous by the comments above. In particular, $j_\mu(L^1(\mu)) \subset \mathcal{M}(\mathbb{R})^0_{\mathrm{TV}}$, thus by Example 4.4.7 we obtain that $j_\mu(L^1(\mu)) = \mathcal{M}(\mathbb{R})^0_{\mathrm{TV}}$.

If $\mu \in \mathcal{M}^+(\Omega)$ is an invariant measure of $(P(t))_{t\geq 0}$, then $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, since $t \mapsto P(t)\mu = \mu$ is continuous from $\mathbb{R}_+$ to $\mathcal{M}(\Omega)$. It would be interesting to be able to characterise the invariant measures among those in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$.

Whenever the Markov semigroup arises from a non-trivial underlying semigroup of measurable maps on Ω, $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ cannot be too large:

Proposition 4.4.13. *Let $(\Phi_t)_{t\geq 0}$ be a jointly measurable semigroup of measurable maps on Ω, and let $(P_\Phi(t))_{t\geq 0}$ be the associated Markov semigroup. Then $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = \mathcal{M}(\Omega)$ if and only if $\Phi_t = \mathrm{Id}$ for every $t \in \mathbb{R}_+$.*

Proof. Suppose $\Phi_t = \mathrm{Id}$ for every $t \in \mathbb{R}_+$. Then $P_\Phi(t)\mu = \mu$ for every $t \in \mathbb{R}_+$ and $\mu \in \mathcal{M}(\Omega)$, hence $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = \mathcal{M}(\Omega)$. Suppose $\mathcal{M}(\Omega) = \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, and let $x \in \Omega$. Then

$$\|\delta_{\Phi(t)x} - \delta_x\|_{\mathrm{TV}} = \|P_\Phi(t)\delta_x - \delta_x\|_{\mathrm{TV}} \downarrow 0,$$

as $t \downarrow 0$. Hence there is a $\tau > 0$ such that $\delta_{\Phi_t(x)} = \delta_x$ for all $t \in [0, \tau)$, and then by the semigroup law $\delta_{\Phi_t(x)} = \delta_x$ for all $t \in \mathbb{R}_+$, so $\Phi_t(x) = x$ for all $t \in \mathbb{R}_+$. □

However, there do exist non-trivial Markov semigroups $(P(t))_{t\geq 0}$ such that $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = \mathcal{M}(\Omega)$; in [118, Section 5] a C_0-semigroup on $\mathcal{M}(\Omega)$ is constructed, which under certain conditions is a Markov semigroup.

Remark 4.4.14. One might also consider semigroups $(P(t))_{t\geq 0}$ on $\mathcal{M}(\Omega)$ for which the linear positive operators $P(t)$ satisfy a more general condition than *(MO2)*:

$$\|P(t)\mu\|_{TV} \leq Me^{\lambda t}\|\mu\|_{\mathrm{TV}},$$

for certain $M \geq 1$ and $\lambda \geq 0$. In this case we can still achieve results similar to Theorem 4.4.10, using a weighted L^1-space instead of $L^1(\mathbb{R}_+)$

$$L^1_{\lambda,M}(\mathbb{R}_+) := \{f \in L^1(\mathbb{R}_+) : t \mapsto e^{\lambda t}|f(t)| \in L^1(\mathbb{R}_+), \}$$

with norm $\|f\|_{\lambda,M} := \int_{\mathbb{R}_+} Me^{\lambda t}|f(t)|\,dt$.

As before, we can consider a normal space S endowed with its Baire σ-algebra and $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup on S that is strongly stochastically continuous at zero. Then by Lemma 4.3.7 (AS) is satisfied with $X = C_b(S)$, so the results in this section all hold in this setting. In particular we can take S to be a perfectly normal (or metric) space and $(P(t))_{t\geq 0}$ a regular Markov semigroup on S that is strongly stochastically continuous at zero (by Proposition 3.4.5).

Now we consider S to be a Polish space. Let d be any complete metric on S that metrises its topology and let $\mathcal{S}_{\mathrm{BL}}$ be the Banach space associated to (S, d). Let $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup on S. Then the integrals $f *_P \mu = \int_{\mathbb{R}_+} f(s)P(s)\mu\,ds$ for $f \in L^1(\mathbb{R}_+)$ and $\mu \in \mathcal{M}(S)$ are actually Bochner integrals in $\mathcal{S}_{\mathrm{BL}}$ by Proposition 3.2.4.

We assume in addition that $(P(t))_{t\geq 0}$ is strongly stochastically continuous at zero. Then it is jointly measurable by Lemma 3.4.4 and Proposition 3.4.5.

We now want to give some apparently weaker conditions than those in Theorem 4.4.10, which turn out to be equivalent. These may be useful for showing that a particular measure μ is in $\mathcal{M}(S)^0_{\mathrm{TV}}$.

Lemma 4.4.15. *Let $\mu \in \mathcal{M}(S), \nu \in \mathcal{M}^+(S)$. Then the following are equivalent:*

(i) $\mu \ll \nu$

(ii) $\mu(K) = 0$ *for all compact K in S such that $\nu(K) = 0$.*

Proof. *(i)* $\Rightarrow$ *(ii)*: Trivial. *(ii)* $\Rightarrow$ *(i)*: Let E be a Borel set in S such that $\nu(E) = 0$. Then $\nu(K) = 0$ for all compact K such that $K \subset E$, hence $\mu^+(K) = \mu^-(K)$ for all compact $K \subset E$. Since S is a complete separable metric space, μ^+ and μ^- are inner regular, i.e. for every Borel set E in S, there are compact $K_n \subset E$, such that $\lim\limits_{n\to\infty} \mu^+(K_n) = \mu^+(E)$ and $\lim\limits_{n\to\infty} \mu^-(K_n) = \mu^-(E)$ (see e.g. [11, Theorem 1.1 and 1.3]). So $\mu^+(E) = \mu^-(E)$ and $\mu(E) = \mu^+(E) - \mu^-(E) = 0$. □

Theorem 4.4.16. *Let $\mu \in \mathcal{M}(S)$. Then the following are equivalent:*

(i) $\mu \in \mathcal{M}(S)^0_{\mathrm{TV}}$.

(ii) For all compact K in S, $t \mapsto P(t)\mu(K)$ is continuous.

(iii) If K in S compact and $P(t)\mu(K) = 0$ for almost every $t \in \mathbb{R}_+$, then $\mu(K) = 0$.

Proof. *(i)*$\Rightarrow$ *(ii)*: This is trivial.

(ii)$\Rightarrow$ *(iii)*: Let K in S be compact, such that $P(t)\mu(K) = 0$ for almost every $t \in \mathbb{R}_+$. Then, by continuity of $t \mapsto P(t)\mu(K)$, $\mu(K) = 0$.

(iii)$\Rightarrow$ *(i)*: Let $f \in L^1(\mathbb{R}_+)$, such that $f(t) > 0$ for almost every $t \in \mathbb{R}_+$. Define $\nu := f *_P |\mu|$. Let K in S be compact, such that $\nu(K) = 0$, then $P(t)|\mu|(K) = 0$ for almost every $t \in \mathbb{R}_+$. By positivity of $P(t)$, $|P(t)\mu|(K) \leq P(t)|\mu|(K) = 0$ for almost every $t \in \mathbb{R}_+$, hence $\mu(K) = 0$. Thus $\mu \ll \nu$ by Lemma 4.4.15. □

An important consequence of the characterisations in Theorem 4.4.10 is:

Proposition 4.4.17. *$\mathcal{M}(S)^0_{\mathrm{TV}}$ is dense in $\mathcal{S}_{\mathrm{BL}}$.*

Proof. Let $\mu \in \mathcal{M}(S)$ and $\epsilon > 0$. Then there is a $\tau > 0$ such that $\|P(t)\mu - \mu\|^*_{\mathrm{BL}} < \epsilon$ for all $t \in [0, \tau]$. By Theorem 4.4.10 $e_n *_P \mu \in \mathcal{M}(S)^0_{\mathrm{TV}}$.

$$\begin{aligned} \|e_n *_P \mu - \mu\|^*_{\mathrm{BL}} &= n\| \int_0^{\frac{1}{n}} P(t)\mu - \mu \, dt\|^*_{\mathrm{BL}} \\ &\leq n \int_0^{\frac{1}{n}} \|P(t)\mu - \mu\|^*_{\mathrm{BL}} \, dt < \epsilon, \end{aligned}$$

for all $t \in [0, \tau]$. So $\|e_n *_P \mu - \mu\|^*_{\mathrm{BL}} \to 0$. Since $\mathcal{M}(S)$ is dense in $\mathcal{S}_{\mathrm{BL}}$, $\mathcal{M}(S)^0_{\mathrm{TV}}$ is dense in $\mathcal{S}_{\mathrm{BL}}$ as well. □

4.5 Decomposition of the space of measures

4.5.1 Absolute continuous and singular measures

For $\mu \in \mathcal{M}(\mathbb{R})$, define $\mu_t(E) := \mu(E - t)$, $t \in \mathbb{R}$. It is a classical result by Plessner [101] that $\|\mu_t - \mu\|_{\mathrm{TV}} \to 0$ as $t \to 0$ if and only if μ is absolutely continuous with respect to the Lebesgue measure m. Then the Lebesgue-Radon-Nikodym Decomposition Theorem implies that every μ in $\mathcal{M}(\mathbb{R})$ can be uniquely decomposed into $\mu_a + \mu_s$, where $\mu_a \in j_m(L^1(\mathbb{R}, m))$, and μ_s is singular with respect to m.

We can translate this to our setting: let $\Phi_t(x) = x + t$, then $(\Phi_t)_{t\in\mathbb{R}}$ defines a strongly continuous group of continuous mappings $\Phi_t : \mathbb{R} \to \mathbb{R}$. This defines a strongly stochastically continuous Markov group $(P_\Phi(t))_{t\in\mathbb{R}}$ by $P_\Phi(t)\mu = \mu \circ \Phi_t^{-1}$ according to Proposition 3.4.2. Note that we only formulated Proposition 3.4.2 for semigroups, but it can easily be adapted for groups. Plessner's result implies that the subspace of strong continuity $\mathcal{M}(\mathbb{R})^0_{\mathrm{TV}}$ equals $j_m(L^1(\mathbb{R}, m))$ (see also Example 4.4.12), and every $\mu \in \mathcal{M}(\mathbb{R})$ can be uniquely decomposed into $\mu_a + \mu_s$, where $\mu_a \in \mathcal{M}(\mathbb{R})^0_{\mathrm{TV}}$ and μ_s is singular with respect to every $\nu \in \mathcal{M}(\mathbb{R})^0_{\mathrm{TV}}$. We will generalise this decomposition to our setting.

As in the previous section let $(P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup on a measurable space (Ω, Σ). Recall the notion of projection band from Section 4.2.

Proposition 4.5.1. *When* (*AS*) *or* (4.7) *holds,* $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ *is a projection band in* $\mathcal{M}(\Omega)$.

Proof. Let $\mu, \nu \in \mathcal{M}(\Omega)$ such that $0 \leq |\mu| \leq |\nu|$ and $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$. Then $|\nu| \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10. Since $\mu \ll |\nu|$, $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, again by Theorem 4.4.10. Thus $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ is an ideal in $\mathcal{M}(\Omega)$. Since $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ is closed, it is a projection band by Theorem 4.2.1. □

As a projection band, $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ is complemented in $\mathcal{M}(\Omega)$. One has

$$\mathcal{M}(\Omega) = \mathcal{M}(\Omega)^0_{\mathrm{TV}} \oplus (\mathcal{M}(\Omega)^0_{\mathrm{TV}})^\perp, \tag{4.10}$$

by [85, Theorem 1.2.9].

We will show that $(\mathcal{M}(\Omega)^0_{TV})^\perp = \mathcal{M}(\Omega)^s_{\mathrm{TV}}$, where

$$\mathcal{M}(\Omega)^s_{\mathrm{TV}} := \{\mu \in \mathcal{M}(\Omega) : \mu^+ \perp P(t)\mu^+, \mu^- \perp P(t)\mu^- \text{ for almost every } t \geq 0\}.$$

Our approach is inspired by that of Liu and Van Rooij [82].

Proposition 4.5.2. *Suppose (AS) or* (4.7) *holds. Let* $\mu \in \mathcal{M}(\Omega)$*. Then the following are equivalent:*

(i) $\mu \in \mathcal{M}(\Omega)^s_{\mathrm{TV}}$.

(ii) $\mu \perp \nu$ *for every* $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$

(iii) For all $\nu \in \mathcal{M}(\Omega)$, $\mu \perp P(t)\nu$ *for almost every* $t \in \mathbb{R}_+$.

Proof. (i)⇒ (ii): Let $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, then $|\nu| \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10. By the Lebesgue-Radon-Nikodym Theorem, there are unique $\mu_a^+, \mu_s^+ \in \mathcal{M}(\Omega)^+$, such that $\mu^+ = \mu_a^+ + \mu_s^+$, $\mu_a^+ \ll |\nu|$ and $\mu_s^+ \perp |\nu|$. Then $\mu_a^+ \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10. By assumption, $\mu^+ \perp P(t)\mu^+$ for almost every $t \in \mathbb{R}_+$. Let $t \in \mathbb{R}_+$ be such that $\mu^+ \perp P(t)\mu^+$. Then there is a set $U \in \Sigma$, such that $\mu^+(E) = \mu^+(E \cap U)$ and $P(t)\mu^+(U) = 0$ for all $E \in \Sigma$. So

$$0 \leq \mu_a^+(E \backslash U) \leq \mu^+(E \backslash U) = 0.$$

Consequently $\mu_a^+(E) = \mu_a^+(E \cap U)$ for all $E \in \Sigma$, and

$$0 \leq P(t)\mu_a^+(U) \leq P(t)\mu^+(U) = 0,$$

so $P(t)\mu_a^+ \perp \mu_a^+$.

Hence $\mu_a^+ \perp P(t)\mu_a^+$ for almost every $t \in \mathbb{R}_+$. $\{\mu_a^+\}^\perp$ is a band in $\mathcal{M}(\Omega)$. Therefore it is closed. Since $t \mapsto P(t)\mu_a^+ : \mathbb{R}_+ \to \mathcal{M}(\Omega)$ is continuous, $\mu_a^+ \in \{\mu_a^+\}^\perp$, thus $\mu_a^+ = 0$. This implies that $\mu^+ = \mu_s^+$, so $\mu^+ \perp |\nu|$, and therefore $\mu^+ \perp \nu$.

In a similar way we can prove that $\mu^- \perp \nu$, hence $\mu \perp \nu$.

(ii)⇒ (iii): Let $\nu \in \mathcal{M}(\Omega)$ and define $\rho := f *_P |\nu| \in L^1(\mathbb{R}_+) *_P \mathcal{M}(\Omega)$, where $f \in L^1(\mathbb{R}_+)$, such that $f(t) > 0$ for almost every $t \in \mathbb{R}_+$. Then $\rho \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10. By *(ii)* $\mu \perp \rho$, hence there is a $U \in \Sigma$, such that $\mu(E) = \mu(E \cap U)$ and $\rho(U) = 0$ for all $E \in \Sigma$. Hence $P(t)|\nu|(U) = 0$ for almost every $t \in \mathbb{R}_+$. Then positivity of $(P(t))_{t\geq 0}$ implies that for almost every $t \in \mathbb{R}_+$, $|P(t)\nu|(U) = 0$, hence $|P(t)\nu| \perp \mu$. So $P(t)\nu \perp \mu$ for almost every $t \in \mathbb{R}_+$.

(iii)⇒ (i): By assumption, $\mu \perp P(t)\mu^+$ and $\mu \perp P(t)\mu^-$ for almost every $t \in \mathbb{R}_+$. Hence $|\mu| \perp P(t)\mu^+$ and $|\mu| \perp P(t)\mu^-$, so $\mu^+ \perp P(t)\mu^+$ and $\mu^- \perp P(t)\mu^-$ for almost every $t \in \mathbb{R}_+$. □

Corollary 4.5.3. *When (AS) or* (4.7) *holds,* $\mathcal{M}(\Omega)^s_{\mathrm{TV}} = (\mathcal{M}(\Omega)^0_{\mathrm{TV}})^\perp$.

This implies that $\mathcal{M}(\Omega)^s_{\mathrm{TV}}$ is a projection band by [85, Proposition 1.2.7].

As in [82] we call $\mu \in \mathcal{M}(\Omega)$ *absolutely continuous* with respect to $(P(t))_{t\geq 0}$ if $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ and *singular* with respect to $(P(t))_{t\geq 0}$ if $\mu \in \mathcal{M}(\Omega)^s_{\mathrm{TV}}$. This terminology is based on the fact that $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ if and only if there is a $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ such that $\mu \ll |\nu|$ by Theorem 4.4.10, and $\mu \in \mathcal{M}(\Omega)^s_{\mathrm{TV}}$ if and only if μ and ν are singular for every $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ by Theorem 4.4.10.

An immediate consequence of (4.10) and Corollary 4.5.3 is the following:

Proposition 4.5.4. *When (AS) or* (4.7) *holds, every* $\mu \in \mathcal{M}(\Omega)$ *has a unique decomposition* $\mu = \mu_a + \mu_s$, *with* $\mu_a \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, *and* $\mu_s \in \mathcal{M}(\Omega)^s_{\mathrm{TV}}$.

We denote the band projections on $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ and $\mathcal{M}(\Omega)^s_{\mathrm{TV}}$ by P_0 and P_s respectively. Then P_0, P_s are positive bounded linear operators on $\mathcal{M}(\Omega)$, with $\|P_0\| \leq 1$ and $\|P_s\| \leq 1$, and $P_0\mu = \mu_a$, $P_s\mu = \mu_s$.

As before we can apply the results in this section to a normal space S with its Baire σ-algebra, and a regular jointly measurable Markov semigroup on S that is strongly stochastically continuous at zero. In particular, we can consider a perfectly normal (or metric) space S, with a regular Markov semigroup on S that is strongly stochastically continuous at zero.

While $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ is invariant under $(P(t))_{t\geq 0}$, $\mathcal{M}(\Omega)^s_{\mathrm{TV}}$ need not be, as the following example shows:

Example. Let $\Omega = \mathbb{R}_+$ with Euclidean metric. Define $\Phi_t(x) = \max(x - t, 0)$, for $t, x \in \mathbb{R}_+$. Then $(\Phi_t)_{t\geq 0}$ is a strongly continuous semigroup of continuous maps on S, hence it defines, by Proposition 3.4.2, a strongly stochastically continuous Markov semigroup $(P(t))_{t\geq 0}$ given by $P(t)\mu := \mu \circ \Phi_t^{-1}$. Let $x > 0$, then clearly $\delta_x \perp P(t)\delta_x$ for all $t > 0$, hence $\delta_x \in \mathcal{M}(\Omega)^s_{\mathrm{TV}}$. However, for $t \geq x$, $P(t)\delta_x = \delta_0$, and δ_0 is in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$, and not in $\mathcal{M}(\Omega)^s_{\mathrm{TV}}$, since $P(t)\delta_0 = \delta_0$ for all $t \in \mathbb{R}_+$.

For each $\mu \in \mathcal{M}(\Omega)$, we can define $d(\mu, \mathcal{M}(\Omega)^0_{\mathrm{TV}})$ to be the distance of μ to $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ with respect to $\|\cdot\|_{\mathrm{TV}}$. Clearly, $\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ if and only if $d(\mu, \mathcal{M}(\Omega)^0_{\mathrm{TV}}) = 0$.

Lemma 4.5.5. *Suppose (AS) or* (4.7) *holds. Let* $\mu \in \mathcal{M}(\Omega)$. *Then* $d(\mu, \mathcal{M}(\Omega)^0_{\mathrm{TV}}) = \|\mu_s\|_{\mathrm{TV}}$.

Proof. '$\leq$': $\mu = \mu_a + \mu_s$, so $\|\mu - \mu_a\|_{\mathrm{TV}} = \|\mu_s\|_{\mathrm{TV}}$. Hence

$$d(\mu, \mathcal{M}(\Omega)^0_{\mathrm{TV}}) = \inf_{\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}} \|\mu - \nu\|_{\mathrm{TV}} \leq \|\mu - \mu_a\|_{\mathrm{TV}} = \|\mu_s\|_{\mathrm{TV}}.$$

'$\geq$': Let $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$. Then

$$\|\mu_s\|_{\mathrm{TV}} = \|P_s\mu\|_{\mathrm{TV}} = \|P_s\mu - P_s\nu\|_{\mathrm{TV}} \leq \|\mu - \nu\|_{\mathrm{TV}},$$

which implies that $\|\mu_s\|_{\mathrm{TV}} \leq d(\mu, \mathcal{M}(\Omega)^0_{\mathrm{TV}})$. □

Lemma 4.5.6. *Suppose (AS) or* (4.7) *holds. Let* $\mu \in \mathcal{M}(\Omega)$. *The function* $t \mapsto \|P_sP(t)\mu\|_{\mathrm{TV}}$ *is non-increasing.*

Proof. It suffices to show that $\|P_sP(t)\mu\|_{\mathrm{TV}} \leq \|P_s\mu\|_{\mathrm{TV}}$ for all $t \in \mathbb{R}_+$.

Let $0 \leq t$. First assume $\mu \in \mathcal{M}^+(\Omega)$, then $0 \leq \mu_a \leq \mu$. Since $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ is invariant under $P(t)$, $P_0P(t)\mu_a = P(t)\mu_a$, hence

$$0 \leq P(t)\mu_a = P_0P(t)\mu_a \leq P_0P(t)\mu.$$

Then

$$0 \leq P_s P(t)\mu = P(t)\mu - P_0 P(t)\mu \leq P(t)\mu - P(t)\mu_a = P(t)\mu_s,$$

hence

$$\|P_s P(t)\mu\|_{\mathrm{TV}} \leq \|P(t)\mu_s\|_{\mathrm{TV}} \leq \|\mu_s\|_{\mathrm{TV}}. \tag{4.11}$$

Now let $\mu = \mu^+ - \mu^- \in \mathcal{M}(\Omega)$. Then $P_s\mu^+ \perp P_s\mu^-$, which implies that $\|P_s\mu\|_{\mathrm{TV}} = \|P_s\mu^+\|_{\mathrm{TV}} + \|P_s\mu^-\|_{\mathrm{TV}}$. By (4.11)

$$\begin{aligned} \|P_s P(t)\mu\|_{\mathrm{TV}} &\leq \|P_s P(t)\mu^+\|_{\mathrm{TV}} + \|P_s P(t)\mu^-\|_{\mathrm{TV}} \\ &\leq \|P_s\mu^+\|_{\mathrm{TV}} + \|P_s\mu^-\|_{\mathrm{TV}} = \|P_s\mu\|_{\mathrm{TV}}. \end{aligned}$$

□

It follows from Theorem 4.4.10 that whenever a jointly measurable regular Markov semigroup satisfies (AS) or (4.7), there exists for every element in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ a $\mu \in \mathcal{M}^+(\Omega)$ such that $j_\mu(L^1(\mu))$ is $(P(t))_{t\geq 0}$-invariant and contained in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$. We saw in Example 4.4.12 that there are Markov semigroups for which $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = j_\mu(L^1(\mu))$ for some $\mu \in \mathcal{M}^+(\Omega)$. In general this does not hold: for instance, when $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = \mathcal{M}(\Omega)$. We can use our results above to give a necessary and sufficient condition for this to hold.

Proposition 4.5.7. *Suppose* (AS) *or* (4.7) *holds. Then the following are equivalent:*

(i) There is a $\mu \in \mathcal{M}^+(\Omega)$ *such that for every* μ*-null set* N *and every* $\sigma \in \mathcal{M}(\Omega)$,

$$P(t)\sigma(N) = 0 \text{ for Lebesgue almost every } t \in \mathbb{R}_+.$$

(ii) There exists a $\nu \in \mathcal{M}^+(\Omega)$ *such that* $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = j_\nu(L^1(\nu))$.

Proof. (i)⇒(ii): Let $\rho \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$. Then there exists an $f \in L^1(\mathbb{R}_+)$ and a $\sigma \in \mathcal{M}(\Omega)$ such that $\rho = f *_P \sigma$. Let N be a μ-null set, then

$$\rho(N) = \int_{\mathbb{R}_+} f(t)P(t)\sigma(N)\,dt = 0,$$

since $P(t)\sigma(N) = 0$ for almost every $t \in \mathbb{R}_+$. Thus $\rho \ll \mu$, which implies that $\mathcal{M}(\Omega)^0_{\mathrm{TV}} \subset j_\mu(L^1(\mu))$. Since μ need not be in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$, we are not done yet. Since $\rho \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, $\mu \perp P_s\mu$. Because $\mu = P_0\mu + P_s\mu$ and $\rho \ll \mu$, this implies that $\rho \ll P_0\mu =: \nu$. Then $\nu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, thus $\mathcal{M}(\Omega)^0_{\mathrm{TV}} = j_\nu(L^1(\nu))$.

(ii)⇒(i): Let N be a ν-null set and $\sigma \in \mathcal{M}^+(\Omega)$, and let $f \in L^1(\mathbb{R}_+)$ be such that $f > 0$ almost everywhere. We define $\rho = f *_P \sigma$. By assumption, $\rho \ll \nu$, thus

$$\int_{\mathbb{R}_+} f(t)P(t)\sigma(N)\,dt = \rho(N) = 0.$$

Thus $P(t)\sigma(N) = 0$ for almost every $t \in \mathbb{R}_+$. If $\sigma \in \mathcal{M}(\Omega)$, then $P(t)\sigma(N) = P(t)\sigma^+(N) + P(t)\sigma^-(N) = 0$ almost everywhere. □

Recall Proposition 4.3.11, in which we show that, under certain conditions, whenever μ is such that $(P(t))_{t\geq 0}$ leaves $j_\mu(L^1(\mu))$ invariant, $P(t)\mu \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$ for all $t > 0$. We needed there the assumption that $L^1(\mu)$ is separable. Using our results in this section we can now show that there exists an invariant L^1-space contained in $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ such that $P(t)(j_\mu(L^1(\mu)))$ maps into this space for all $t > 0$. We do not need the separability condition here.

Theorem 4.5.8. *Suppose* (*AS*) *or* (4.7) *holds. Let $\mu \in \mathcal{M}^+(\Omega)$ be non-zero such that $(P(t))_{t\geq 0}$ leaves $j_\mu(L^1(\mu))$ invariant. Then the following statements hold:*

(i) $\mu_a = P_0\mu$ *is non-zero.*

(ii) $(P(t))_{t\geq 0}$ *leaves* $j_{\mu_a}(L^1(\mu_a))$ *invariant.*

(iii) For all $\nu \in j_\mu(L^1(\mu))$, $P(t)\nu \in j_{\mu_a}(L^1(\mu_a))$ *for all* $t > 0$.

Proof. (i) It follows from Theorem 4.5.2 that $P(t)\mu \perp \mu_s$ for almost every $t \in \mathbb{R}_+$. Since $P(t)\mu \ll \mu = \mu_a + \mu_s$, $P(t)\mu \ll \mu_a$ for almost every $t \in \mathbb{R}_+$, which implies that $\mu_a \neq 0$.

(ii) By Corollary 4.3.3 it suffices to show that $P(t)\mu_a \ll \mu_a$ for all $t \in \mathbb{R}_+$. Since $\mu \in \mathcal{M}^+(\Omega)$, $\mu_a \in \mathcal{M}^+(\Omega)$, and thus

$$P(t)\mu_a \ll P(t)\mu \ll \mu_a$$

for almost every $t \in \mathbb{R}_+$. Let $N \in \Sigma$ be a μ_a-null set, then $P(t)\mu_a(N) = 0$ for almost every $t \in \mathbb{R}_+$. Since $\mu_a \in \mathcal{M}(\Omega)^0_{\mathrm{TV}}$, $t \mapsto P(t)\mu_a(N)$ is continuous, thus $P(t)\mu_a(N) = 0$ for all $t \in \mathbb{R}_+$. Hence $P(t)\mu_a \ll \mu_a$ for all $t \in \mathbb{R}_+$.

(iii) Let $\nu \in j_\mu(L^1(\mu))$, then $P(t)\nu \ll \mu$ for all $t \in \mathbb{R}_+$. By Theorem 4.5.2, $P(t)\nu \perp \mu_s$ for almost every $t \in \mathbb{R}_+$, thus $P(t)\nu \ll \mu_a$ for almost every $t \in \mathbb{R}_+$. Since $j_{\mu_a}(L^1(\mu_a))$ is $(P(t))_{t\geq 0}$-invariant, $P(t)\nu \ll \mu_a$ for all $t > 0$. □

4.5.2 A Wiener-Young type theorem

Wiener and Young [123] extended the result by Plessner (see Section 4.5.1), by showing that for all $\mu \in \mathcal{M}(\mathbb{R})$, $\limsup_{t\to 0} \|P(t)\mu - \mu\|_{\mathrm{TV}} = 2\|\mu_s\|_{\mathrm{TV}}$, where $P(t)\mu = \mu \circ \Phi_t^{-1}$, with $\Phi_t(x) = x + t$, and μ_s is the singular component of μ with respect to the Lebesgue measure.

We generalise this result to a regular jointly measurable Markov semigroup $(P(t))_{t\geq 0}$ on a measurable space (Ω, Σ) with a slightly stronger assumption than *(AS)*:

(AN) There is an $\mathcal{M}(\Omega)$-norming set $X \subset \mathrm{BM}(\Omega)$ such that for every $\nu \in \mathcal{M}(\Omega)$ and $\phi \in X$,

$$\lim_{t\downarrow 0} \langle P(t)\nu, f\rangle = \langle \nu, f\rangle.$$

It has been generalised in several other directions: see for instance [90, 93] for a generalisation in the setting of adjoint semigroups of positive strongly continuous semigroups on Banach lattices. Note that the Markov semigroups we consider here are in general *not* adjoints of strongly continuous semigroups.

Theorem 4.5.9. *Suppose (AN) holds. Let $\mu \in \mathcal{M}(\Omega)$. Then*

$$\limsup_{t\downarrow 0} \|P(t)\mu - \mu\|_{\mathrm{TV}} = 2\|\mu_s\|_{\mathrm{TV}}.$$

Proof. Step 1. $\lim_{t\downarrow 0} \|P(t)\mu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{TV}}$ *for all* $\mu \in \mathcal{M}(\Omega)$.
Let $\epsilon > 0$. Since

$$\|\mu\|_{\mathrm{TV}} = \sup\{|\langle \mu, f\rangle| : f \in X, \|f\|_\infty \leq 1\},$$

there is an $f \in X$ with $\|f\|_\infty \leq 1$ and $|\|\mu\|_{\mathrm{TV}} - \langle \mu, f\rangle| < \epsilon/2$. Since $f \in X$, there exists a $\tau > 0$ such that

$$|\langle P(t)\mu, f\rangle - \langle \mu, f\rangle| < \epsilon/2 \text{ for all } t \in [0, \tau).$$

Thus for $t \in [0, \tau)$ we obtain

$$\|P(t)\mu\|_{\mathrm{TV}} \geq |\langle P(t)\mu, f\rangle| \geq \|\mu\|_{\mathrm{TV}} - \epsilon,$$

and by *(MO2)* $\|P(t)\mu\|_{\mathrm{TV}} \leq \|\mu\|_{\mathrm{TV}}$, hence the statement holds.

Step 2. $\limsup_{t\downarrow 0} \|P(t)\mu - \mu\|_{\mathrm{TV}} = 2\|\mu_s\|_{\mathrm{TV}}$ *for all* $\mu \in \mathcal{M}(\Omega)$.
Clearly $\|P(t)\mu_a - \mu_a\|_{\mathrm{TV}} \to 0$. This implies that $\limsup_{t\downarrow 0} \|P(t)\mu - \mu\|_{\mathrm{TV}} = \limsup_{t\downarrow 0} \|P(t)\mu_s - \mu_s\|_{\mathrm{TV}}$. By Proposition 4.5.2, $P(t)\mu_s \perp \mu_s$ for almost every $t \in \mathbb{R}_+$, say for all $t \in N$, where $\mathbb{R}_+ \setminus N$ has measure zero. Hence, for these t, $\|P(t)\mu_s - \mu_s\|_{\mathrm{TV}} = \|P(t)\mu_s\|_{\mathrm{TV}} + \|\mu_s\|_{\mathrm{TV}}$ and thus $\|P(t)\mu_s - \mu_s\|_{\mathrm{TV}} \to 2\|\mu_s\|_{\mathrm{TV}}$ as $t \downarrow 0$ in $\mathbb{R}_+ \setminus N$ by Step 1. Noting that $\|P(t)\mu_s - \mu_s\|_{\mathrm{TV}} \leq 2\|\mu_s\|_{\mathrm{TV}}$ by the triangle inequality and *(MO2)*, the proof of Step 2 is complete. □

Remark. As before, we can apply the above results to a normal space S with its Baire σ-algebra, and a regular jointly measurable Markov semigroup on S that is strongly stochastically continuous at zero. And in particular to a perfectly normal (or metric) space S and a regular Markov semigroup on S that is strongly stochastically continuous at zero.

4.6 Notes

This chapter is an extended version of our paper [63]. The state space considered there is a complete separable metric space. In this chapter we extend many of the results from [63] to the full generality of a measurable space. Apart from generalisations, there are some new results as well. We will list some of the important ones: Proposition 4.3.5 is new here. One of the main results in the paper is [63,

Theorem 4.6], where the Markov semigroup is assumed to be strongly stochastically continuous. Here we recover this result as Corollary 4.3.10 of Theorem 4.3.6, but now we only require strong stochastic continuity at zero. For the characterisation of $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ we do not need any continuity assumption: joint measurability suffices. Proposition 4.4.8, Corollary 4.4.9, Proposition 4.5.7 and Theorem 4.5.8 are also new.

Our results on the characterisation of $\mathcal{M}(\Omega)^0_{\mathrm{TV}}$ in Section 4.4 were inspired by analogous results in the setting of continuous group actions on locally compact Hausdorff spaces [51, 82].

Some of our results on the decomposition in Section 4.5 have their counterparts in the setting of adjoint semigroups of positive strongly continuous semigroups on Banach lattices [90, 93]. Note that the Markov semigroups we consider here are in general *not* adjoints of strongly continuous semigroups.

CHAPTER

FIVE

ERGODIC DECOMPOSITIONS ASSOCIATED TO REGULAR MARKOV OPERATORS ON POLISH SPACES

5.1 Introduction

In this chapter we study Markov operators on a Polish space S. There is much interest lately in Markov operators on non-locally compact Polish spaces, by e.g. Szarek and coworkers [79, 88, 110, 112, 113] and Ollivier [91]. More specific examples of Markov operators on Polish spaces are given by iterated function systems [87, 111], ARCH processes in econometrics [71] and random dynamical systems on separable Banach spaces [64, 65], which also have various applications in mathematical biology.

It is well-established that the set of invariant probability measures for a Markov operator – when it is non-empty – is convex with non-empty set of extreme points, the so-called *ergodic invariant measures*, denoted by $\mathcal{P}_{\mathrm{erg}}(S)$. Each invariant measure μ can be represented as an integral over the extreme points, in the sense that

$$\mu(E) = \int_{\mathcal{P}_{\mathrm{erg}}(S)} \nu(E)\, d\rho_\mu(\nu) \tag{5.1}$$

for each Borel set E in S, see e.g. [121, Chapter 6], and [32], where it has been obtained for standard spaces, using probabilistic arguments. In the pure topological setting such *ergodic integral decompositions* have been considered in somewhat different formulation on compact Hausdorff spaces using Choquet theory (see e.g. [72]).

Another notion of ergodic decomposition has been developed, going back to the pioneering works of Krylov and Bogolioubov [75], Beboutov [9] and Yosida [129, 130]. The main idea is as follows: Given a Markov operator P with an invariant probability measure μ, there exists a Borel set $\Gamma \subset S$, a surjective map $x \mapsto \mu_x$ from Γ onto $\mathcal{P}_{\mathrm{erg}}(S)$ and a decomposition of Γ into Borel sets Γ_x, such that $\mu_x(\Gamma_x) = 1$ and

$$\int_S f\, d\mu = \int_\Gamma \left(\int_{\Gamma_x} f\, d\mu_x \right) d\mu(x). \tag{5.2}$$

Moreover, each Γ_x contains an invariant set with respect to P such that μ_x is also concentrated on this smaller set. Such a result had been obtained by Yosida for the class of Markov operators whose dual operator maps the space of continuous functions with compact support into itself, defined on separable metric spaces in which the closed and bounded sets are compact ([129], [130, Chapter XIII.4]). Hernández-Lerma and Lasserre later covered the more general setting of regular Markov operators on a locally compact separable metric space ([56], [57, Chapter 5]). Zaharopol [132, 134] managed to extend and strengthen some of their results. In particular, he was able to obtain the existence of a set Γ and a decomposition $\{\Gamma_x\}_{x\in\Gamma}$ of Γ that do not depend on the particular invariant measure. These sets can be defined "explicitly", using convergence properties of Cesàro averages of Dirac measures. In [132, pp. 42–43] and [134, p. 50] he poses the open problem whether the decomposition can be extended to more general metric spaces, like Polish spaces.

In this chapter we present a solution to this problem, in which we are able to remove the local compactness condition and extend Zaharopol's results to the more general setting of regular Markov operators on Polish spaces, also obtaining an "explicit" description of the relevant sets in terms of convergence properties of Cesàro averages of Dirac measures. This solves Zaharopol's open problem in a satisfactory manner (also according to Zaharopol by personal communication).

Let us summarise our approach. We extend the definition of a suitable decreasing sequence of subsets of S, $\Gamma_t^P \supset \Gamma_{cp}^P \supset \Gamma_{cpi}^P \supset \Gamma_{cpie}^P$, which form the preliminary Yosida-type decomposition. These sets depend on the Markov operator P, *but not* on a pre-chosen invariant measure. However, any invariant measure is concentrated on the smallest set, Γ_{cpie}^P, which plays the role of the Γ mentioned above. We bring the (quite technical) proofs of these results together in Section 5.3. In Section 5.4 we show that an equivalence relation $\sim$ can be defined on Γ_{cpie}^P such that each equivalence class $[x]$ corresponds uniquely with an ergodic invariant measure ϵ_x. Moreover, we show that *every* ergodic invariant measure can be obtained in this way, which gives a bijection between $\Gamma_{cpie}^P/\sim$ and $\mathcal{P}_{\mathrm{erg}}(S)$. For each $x \in \Gamma_{cpie}^P$ we obtain an invariant set $S_{[x]}$ contained in $[x]$ on which ϵ_x is concentrated and such that ϵ_x is the only ergodic invariant measure of the restriction of P to $S_{[x]}$. This is the so-called Yosida-type ergodic decomposition of S. We show by analytic arguments that for every invariant probability measure μ

$$\mu = \int_{\Gamma_{cpie}^P} \epsilon_x\, d\mu(x),$$

interpreted as Bochner integral in the Banach space $\mathcal{S}_{\mathrm{BL}}$ (introduced in Chapter 2). This implies a result for μ evaluated at E similar to (5.1) by using results on Bochner integration in $\mathcal{S}_{\mathrm{BL}}$ (Section 3.2). In Section 5.5 we give results on convergence of Cesàro averages of measures, based on some of the sets we define in Section 5.3 and Section 5.4.

Let S be a Polish space and d a complete metric on S that metrises its topology. Let $\mathcal{S}_{\mathrm{BL}}$ be the separable Banach space associated to (S, d) (see Chapter 2).

5.2 Preliminaries

In order to arrive at the ergodic decompositions, we need to generalise results of Zaharopol [134] from the setting of locally compact separable metric spaces to the setting of Polish spaces, i.e. separable completely metrisable topological spaces. Note that a locally compact separable metric space need not be complete, for instance $(0, 1)$ with the Euclidean metric. However, it is a well-known result that every locally compact space with a countable base *is* Polish, see e.g. [8, Remark 5 in §29]. Since every locally compact separable metric space has a countable base, the following holds.

Proposition 5.2.1. *Every locally compact separable metric space is a Polish space.*

The following result will be crucial in several places where we need to prove convergence of probability measures and measurability of particular sets.

Proposition 5.2.2. *Let (X, d) be a separable metric space. There exists a countable* convergence determining *set D in $C_b(X)$ consisting of bounded Lipschitz functions with bounded support, i.e. if $\mu, \mu_1, \mu_2, \dots \in \mathcal{P}(X)$ are such that*

$$\langle \mu_n, f\rangle \to \langle \mu, f\rangle \text{ as } n \to \infty \text{ for all } f \in D$$

then $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. Consequently, if $\mu, \nu \in \mathcal{P}(S)$ satisfy $\langle \mu, f\rangle = \langle \nu, f\rangle$ for every $f \in D$, then $\mu = \nu$.

Proof. The existence of the countable convergence determining set D follows from the proof of [38, Proposition 3.4.4]: there it is shown that it suffices to check convergence for finite sums of certain bounded Lipschitz functions $f_{i,j}$ with bounded support, where i and j range over $\mathbb{N}$. There are countably many of such sums, which completes the proof. □

A collection of measures $M \subset \mathcal{P}(S)$ is *tight* if for every $\epsilon > 0$ there exists a compact $K \subset S$ such that $\mu(K) \geq 1 - \epsilon$ for every $\mu \in M$ and every $\epsilon > 0$. For $E \subset S$ and $\epsilon > 0$ we define $E^\epsilon := \{x \in S : d(x, E) < \epsilon\}$. Then the following holds:

Theorem 5.2.3. *Let $M \subset \mathcal{P}(S)$. The following statements are equivalent:*

(i) M is tight.

(ii) For each $\epsilon > 0$ there is a compact $K \subset S$ such that

$$\mu(K^\epsilon) \geq 1 - \epsilon \text{ for every } \mu \in M.$$

(iii) M is relatively compact in $\mathcal{S}_{\mathrm{BL}}$.

The equivalence between *(i)* and *(ii)* can be found in [38, Theorem 3.2.2]. And the equivalence between *(i)* and *(iii)* follows from Prokhorov's Theorem (see e.g. [38, Theorem 3.2.2]) and Lemma 2.3.12.

An important result that we will use several times is the following ergodic theorem proven by Kakutani [69, Theorem 1].

Theorem 5.2.4. *(Kakutani's Ergodic Theorem) Let P be a regular Markov operator with dual U and μ an invariant probability measure. If $f \in \mathrm{BM}(S)$, then there exists a $g \in \mathrm{BM}(S)$ such that the sequence of Cesàro averages $(U^{(n)} f)_n$ converges pointwise μ-a.e. to g and $\langle \mu, f \rangle = \langle \mu, g \rangle$.*

We will also need the following result on the dual U of a regular Markov operator P:

Lemma 5.2.5. *Let μ be an invariant probability measure.*

(i) If $f, g \in \mathrm{BM}(S)$ such that $f = g$ μ-a.e., then $Uf = Ug$ μ-a.e.

(ii) For every $f \in \mathrm{BM}(S)$, $\int_S |Uf| \, d\mu \leq \int_S |f| \, d\mu$.

Proof. (i) Let $f, g \in \mathrm{BM}(S)$ such that $f = g$ μ-a.e. Note that $|U(f-g)| \leq U|f-g|$, since U is positive. Therefore

$$\begin{aligned} \int_S |Uf - Ug| \, d\mu &= \langle \mu, |U(f-g)| \rangle \leq \langle \mu, U|f-g| \rangle \\ &= \langle P\mu, |f-g| \rangle = \langle \mu, |f-g| \rangle = 0. \end{aligned}$$

Thus $Uf = Ug$ μ-a.e.

(ii) Let $f \in \mathrm{BM}(S)$, then

$$\int_S |Uf| \, d\mu \leq \langle \mu, U|f| \rangle = \langle P\mu, |f| \rangle = \int_S |f| \, d\mu.$$

□

5.3 A preliminary Yosida-type decomposition

From now on, we let $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ be a regular Markov operator with dual U on a Polish space S. As before, we choose a complete metric d metrising the topology on S, so that we can make use of the associated Banach space $\mathcal{S}_{\mathrm{BL}}$.

However, we make certain that the specific sets and functions we define do not depend on the metric we choose.

In this section we will define and prove properties of certain subsets of S, based on convergence properties of the Cesàro averages $P^{(n)}\delta_x$ (see Section 3.3 for their definition). These are interesting in their own right, and turn out to be an important ingredient in proving the ergodic decompositions in Section 5.4. Some of these sets are generalisations to the setting of Polish spaces of those formulated by Yosida in [130] and [130, Chapter XIII, Section 4] in order to obtain an ergodic decomposition of state space. This motivates our terminology of "preliminary Yosida-type decomposition" for this collection of subsets. Zaharopol [134] extended this Yosida decomposition to the setting of regular Markov operators on locally compact separable metric spaces. He calls it the *KBBY-decomposition*, because of pioneering work on this decomposition by Krylov, Bogolioubov, Beboutov and Yosida. For the most part, we follow his notation for these sets.

The first set of interest is defined by

$$\Gamma_t^P := \{x \in S : (P^{(n)}\delta_x)_n \text{ is tight}\}.$$

So Γ_t^P consists of those $x \in S$ such that for every sequence $(n_k)_k \subset \mathbb{N}$ there is a subsequence $(n_{k_l})_l$ for which $P^{(n_{k_l})}\delta_x$ converges in $\mathcal{S}_{\mathrm{BL}}$ as $l \to \infty$. Thus it is natural to consider the subset

$$\Gamma_{cp}^P := \{x \in S : (P^{(n)}\delta_x)_n \text{ converges in } \mathcal{S}_{\mathrm{BL}}\}.$$

In general, Γ_{cp}^P need not be equal to Γ_t^P: if S is compact, $\Gamma_t^P = S$, while Γ_{cp}^P is not necessarily equal to S. See for instance Example 5.5.7 at the end of this chapter.

For $x \in \Gamma_{cp}^P$ we write ϵ_x to denote the limit of $(P^{(n)}\delta_x)_n$ in $\mathcal{S}_{\mathrm{BL}}$, which is a probability measure.

If P is Markov-Feller, then ϵ_x is an invariant probability measure for every $x \in \Gamma_{cp}^P$. This follows from the fact that by Proposition 3.3.2 for every $x \in \Gamma_{cp}^P$

$$\begin{aligned} P\epsilon_x &= P \lim_{n\to\infty} P^{(n)}\delta_x = \lim_{n\to\infty} PP^{(n)}\delta_x = \lim_{n\to\infty} \left[P^{(n)} + P^n/n - 1/n\right]\delta_x \quad (5.3) \\ &= \lim_{n\to\infty} P^{(n)}\delta_x = \epsilon_x. \end{aligned}$$

If P is not a Markov-Feller operator, then the measures ϵ_x may not be invariant for any $x \in \Gamma_{cp}^P$. An example is given in [57, Example 5.2.5]. Another relevant set therefore is

$$\Gamma_{cpi}^P := \{x \in \Gamma_{cp}^P : \epsilon_x \text{ is invariant}\}.$$

Example 5.3.1. Let (S, d) be a metric space consisting of N elements: $x_1, \cdots, x_N$. Then we can represent $\mathcal{M}(S)$ as $\mathbb{R}^N$. Because S is uniformly discrete, $\mathcal{S}_{\mathrm{BL}} = \mathcal{M}(S)$ by Theorem 2.3.13 and the norms $\|\cdot\|_{\mathrm{TV}}$ and $\|\cdot\|_{\mathrm{BL}}^*$ are equivalent on $\mathcal{M}(S)$. Let P be a Markov operator on S, then we can represent P as an $N \times N$ *stochastic matrix*

$Q = (q_{i,j})_{i,j=1}^N$, i.e. $q_{i,j} \geq 0$ for all $i, j \in \{1, \cdots, N\}$ and $\sum_{i=1}^N q_{i,j} = 1$. Observe that invariant measures of P correspond with eigenvectors of Q associated to eigenvalue 1. Since S is compact, $\Gamma_t^P = S$. Moreover, the Cesàro averages of the matrix Q converge to a matrix $\hat{Q} = (\hat{q}_{i,j})_{i,j=1}^N$ entry-wise as $n \to \infty$, see e.g. [15, Section I.6]. For $i \in \{1, \cdots, N\}$ let y_i be the column vector with a 1 on the i-th place, and zeroes elsewhere. Then $P^{(n)}\delta_{x_i}$ converges as $n \to \infty$ to $\sum_{j=1}^N \hat{q}_{i,j}\delta_{x_j}$. Thus $\Gamma_{cp}^P = S$ and $\epsilon_{x_i} = \sum_{j=1}^N \hat{q}_{i,j}\delta_{x_j}$. Note that P is Markov-Feller, thus $\Gamma_{cpi}^P = \Gamma_{cp}^P = S$ as well.

Lemma 5.3.2. *Let P be a Markov-Feller operator. If Γ_t^P is non-empty, then P has an invariant probability measure. If P has a unique invariant probability measure, then $\Gamma_t^P = \Gamma_{cp}^P = \Gamma_{cpi}^P$.*

Proof. Let $x \in \Gamma_t^P$. For every sequence $(n_k)_k \subset \mathbb{N}$ there is a subsequence $(n_{k_l})_l$ such that $P^{(n_{k_l})}\delta_x$ converges in $\mathcal{S}_{\mathrm{BL}}$ as $l \to \infty$. By the same argument as in (5.3) $\lim_{l\to\infty} P^{(n_{k_l})}\delta_x$ is an invariant probability measure.

Suppose P has a unique invariant probability measure μ^*. Let $x \in \Gamma_t^P$, then for every sequence $(n_k)_k \subset \mathbb{N}$ there is a subsequence $(n_{k_l})_l$ such that $P^{(n_{k_l})}\delta_x$ converges in $\mathcal{S}_{\mathrm{BL}}$ to μ^*. This implies that $\lim_{n\to\infty} P^{(n)}\delta_x = \mu^*$, thus $x \in \Gamma_{cpi}^P$. □

Finally we will consider a further subset $\Gamma_{cpie}^P \subset \Gamma_{cpi}^P$. We postpone its definition here (see (5.9)), because it requires some concepts that will be defined later on. In Section 5.4.1 it will be shown that Γ_{cpie}^P consists of exactly those $x \in \Gamma_{cpi}^P$ for which ϵ_x is ergodic. Obviously

$$\Gamma_{cpie}^P \subset \Gamma_{cpi}^P \subset \Gamma_{cp}^P \subset \Gamma_t^P. \tag{5.4}$$

We will show in this section that these sets are all Borel measurable and – more importantly – that for *every* invariant probability measure μ,

$$\mu(\Gamma_t^P) = \mu(\Gamma_{cp}^P) = \mu(\Gamma_{cpi}^P) = \mu(\Gamma_{cpie}^P) = 1. \tag{5.5}$$

Of course, there need not be any invariant probability measure. In that case Γ_{cpi}^P and Γ_{cpie}^P are empty. There exist regular Markov operators however, for which there are no invariant probability measures, while Γ_{cp}^P and Γ_t^P are non-empty (see e.g. [132, Example 5.1]).

In order to prove (5.5) it would suffice to show that $\mu(\Gamma_{cpie}^P) = 1$. Technically we cannot achieve this directly, however. Instead we proceed stepwise in the chain (5.4) downwards: the result that an invariant probability measure is concentrated on the larger set is used in proving the concentration result of the set one step lower. An important ingredient in these results is Kakutani's Ergodic Theorem (Theorem 5.2.4).

In order to deal with Γ_t^P, we first show some – apparently new – equivalences for tightness of a collection of measures. We start by introducing some notation. For

$E \subset S$ and $\epsilon > 0$ let us define

$$f_E^\epsilon(x) := (1 - d(x, E)/\epsilon)^+ .$$

This function is in $\mathrm{BL}(S)$. In particular, $0 \le f_E^\epsilon \le 1$ and $|f_E^\epsilon|_{\mathrm{Lip}} = 1/\epsilon$. Also, $f(x) = 0$ for every $x \notin E^\epsilon$. If $E \subset F \subset S$ then $d(x, F) \le d(x, E)$, so $f_E^\epsilon \le f_F^\epsilon$. Moreover, if $\epsilon \le \epsilon'$ then $f_F^\epsilon \le f_F^{\epsilon'}$.

Theorem 5.3.3. *Let $D \subset S$ be dense, and let $M \subset \mathcal{P}(S)$. The following statements are equivalent:*

(i) M is tight.

(ii) For each $\epsilon > 0$ there is a finite subset $F \subset D$ such that

$$\mu(F^\epsilon) > 1 - \epsilon \text{ for every } \mu \in M.$$

(iii) For each $\epsilon > 0$ there is a finite subset $F \subset D$ such that

$$\langle \mu, f_F^\epsilon \rangle > 1 - \epsilon \text{ for every } \mu \in M.$$

(iv) For every $m \in \mathbb{N}$ there is a finite subset $F \subset D$ such that

$$\langle \mu, f_F^{1/m} \rangle > 1 - 1/m \text{ for every } \mu \in M.$$

Proof. *(ii)⇒(i)*: Follows from Theorem 5.2.3 since every finite set is compact.

(i)⇒(ii): Let $\epsilon > 0$. Then there is, by Theorem 5.2.3, a compact $K \subset S$, such that $\mu(K^{\epsilon/2}) \ge 1 - \epsilon/2$. Since K is compact and $D \subset S$ is dense, there exists a finite $F \subset D$, such that $K \subset \cup_{x \in F} B_x(\epsilon/2)$. Then $K^{\epsilon/2} \subset F^\epsilon$, hence for every $\mu \in M$

$$\mu(F^\epsilon) \ge \mu(K^{\epsilon/2}) \ge 1 - \epsilon/2 > 1 - \epsilon.$$

(ii)⇒(iii): One can easily verify that, for any $\epsilon > 0$, one can choose $0 < \delta < \epsilon$ such that

$$\delta/\epsilon + \delta - \delta^2/\epsilon < \epsilon. \tag{5.6}$$

By *(ii)* there exists a finite $F \subset D$, such that $\mu(F^\delta) > 1 - \delta$ for all $\mu \in M$. If $x \in F^\delta$, then

$$f_F^\epsilon(x) = 1 - d(x, F)/\epsilon > 1 - \delta/\epsilon > 0.$$

Thus, by (5.6) we obtain for every $\mu \in M$

$$\begin{aligned} \langle \mu, f_F^\epsilon \rangle &\ge (1 - \delta/\epsilon)\mu(F^\delta) > (1 - \delta/\epsilon)(1 - \delta) \\ &= 1 - \delta/\epsilon - \delta + \delta^2/\epsilon > 1 - \epsilon. \end{aligned}$$

(iii)⇒(ii): Let $\epsilon > 0$. Then there is a finite subset $F \subset D$ such that $\langle \mu, f_F^\epsilon \rangle > 1 - \epsilon$. Then $\mu(F^\epsilon) \ge \langle \mu, f_F^\epsilon \rangle > 1 - \epsilon$.

(iii)⇒(iv): This is trivial.

(iv)⇒(iii): Let $\epsilon > 0$. Then there is an $m \in \mathbb{N}$ such that $1/m < \epsilon$. By assumption there is a finite subset $F \subset D$ such that $\langle \mu, f_F^{1/m} \rangle > 1 - 1/m > 1 - \epsilon$. Now, $f_F^\epsilon \geq f_F^{1/m}$, thus $\langle \mu, f_F^\epsilon \rangle > 1 - \epsilon$. □

Proposition 5.3.4. *Γ_t^P is a Borel set and $\mu(\Gamma_t^P) = 1$ for every invariant probability measure μ.*

Proof. Let D be a countable dense subset of S. Let $\mathcal{F}$ be the collection of finite subsets of D. Then $\mathcal{F}$ is countable. For $F \in \mathcal{F}$ and $m, n \in \mathbb{N}$, we define

$$\begin{aligned} K_{F,m,n} &:= \{x \in S : \langle P^{(n)}\delta_x, f_F^{1/m} \rangle > 1 - 1/m\} \\ &= \{x \in S : U^{(n)} f_F^{1/m}(x) > 1 - 1/m\}. \end{aligned}$$

By Theorem 5.3.3 we have $\Gamma_t^P = \cap_{m\in\mathbb{N}} \cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}$.

Since P is regular, $K_{F,m,n}$ is Borel measurable for every $F \in \mathcal{F}$ and $m, n \in \mathbb{N}$, thus Γ_t^P is Borel measurable.

Let μ be an invariant probability measure. We will show that $\mu(\cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}) = 1$ for every $m \in \mathbb{N}$. Because μ is a probability measure, this implies that $\mu(\Gamma_t^P) = 1$.

Fix $m \in \mathbb{N}$ and $0 < \delta < 1$. Because (S, d) is a complete separable metric space, $\{\mu\}$ is tight by Theorem 5.2.3. Thus there exists an $F_0 \in \mathcal{F}$, depending on δ, m and μ, such that

$$\langle \mu, f_{F_0}^{1/m} \rangle \geq \langle \mu, f_{F_0}^{\delta/m} \rangle > 1 - \delta/m.$$

For convenience, put $f := f_{F_0}^{1/m}$.

By Kakutani's Ergodic Theorem there is a $g \in \mathrm{BM}(S)$ and a Borel set C such that $\mu(C) = 1$ and $U^{(n)} f(x) \to g(x)$ for every $x \in C$. Consequently, $0 \leq g(x) \leq 1$ for every $x \in C$. Moreover, $\langle \mu, f \rangle = \langle \mu, g \rangle$.

Now let $A := \{x \in C : g(x) < 1 - \frac{1}{2m}\}$. Then A is measurable and for every $x \in C$

$$g(x) \leq (1 - \frac{1}{2m}) 1\!\!1_A(x) + 1\!\!1_{C\setminus A}.$$

Therefore

$$1 - \delta/m < \langle \mu, f \rangle = \langle \mu, g \rangle \leq \mu(A)(1 - \frac{1}{2m}) + (1 - \mu(A)).$$

This implies that $\mu(A) \leq 2\delta$.

Let $B := C\setminus A$, then $\mu(B) = \mu(C) - \mu(A) \geq 1 - 2\delta$.

We will show that $B \subset \cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}$. Fix $x \in B$. Then

$$g(x) = \lim_{n\to\infty} U^{(n)} f(x) \geq 1 - \frac{1}{2m},$$

so there is an $N \in \mathbb{N}$ such that $U^{(n)} f(x) > 1 - 1/m$ for every $n > N$. The finite set of measures $\{\delta_x, P^{(1)}\delta_x, \cdots, P^{(N)}\delta_x\}$ is tight, so by Theorem 5.3.3*(iv)* there exists an $F_1 \in \mathcal{F}$ such that $U^{(n)} f_{F_1}^{1/m}(x) > 1 - 1/m$ for $1 \leq n \leq N$. Now put $F := F_0 \cup F_1$. Then $F \in \mathcal{F}$ and $f_F^{1/m} \geq f_{F_i}^{1/m}$ for $i = 0, 1$. Thus $U^{(n)} f_F^{1/m}(x) > 1 - 1/m$ for every $n \in \mathbb{N}$ and $x \in \cap_{n\in\mathbb{N}} K_{F,m,n}$. So indeed $B \subset \cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}$ and consequently

$$\mu(\cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}) \geq \mu(B) \geq 1 - 2\delta.$$

Since we can choose $0 < \delta < 1$ arbitrarily, we obtain that
$\mu(\cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}) = 1$. Thus $\mu(\Gamma_t^P) = 1$. □

We consider the set $\Gamma_{cp}^P = \{x \in S : P^{(n)}\delta_x \text{ converges in } \mathcal{S}_{\mathrm{BL}}\}$.

Lemma 2.3.12 implies that $x \in \Gamma_{cp}^P$ if and only if there is a $\mu \in \mathcal{M}^+(S)$ such that $\langle P^{(n)}\delta_x, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. So Γ_{cp}^P does not depend on the choice of the metric.

If S is a locally compact separable metric space, then the definition in [132, 134] of the set Γ_{cp}^P can be written as follows: $x \in \Gamma_{cp}^P$ if and only if there is a $\mu \in \mathcal{P}(S)$ such that $\langle P^{(n)}\delta_x, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_0(S)$. It then follows from [8, Theorem 30.8] that $\langle P^{(n)}\delta_x, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. This is equivalent to our definition for Γ_{cp}^P. In this setting for S, there is a countable convergence determining set in $C_0(S)$, which is crucial for proving measurability.

Our aim is to show that Γ_{cp}^P is a Borel set and that $\mu(\Gamma_{cp}^P) = 1$ for every invariant probability measure μ, which will extend [134, Proposition 5.4 and Proposition 5.9]. We first need some preliminary results.

Let $\{f_k : k \in \mathbb{N}\}$ be an enumeration of the countable subset of $\mathrm{BL}(S)$ from Proposition 5.2.2. Define

$$\Gamma_c^P = \{x \in S : \langle P^{(n)}\delta_x, f_k\rangle \text{ converges for every } k \in \mathbb{N}\}.$$

Proposition 5.3.5. *Γ_c^P is a Borel set and $\mu(\Gamma_c^P) = 1$ for every invariant probability measure μ.*

Proof. We can write

$$\Gamma_c^P = \{x \in S : (U^{(n)} f_k(x))_n \text{ converges for every } k \in \mathbb{N},\} = \bigcap_{m\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} H_{m,n,k}$$

where

$$H_{m,n,k} := \{x \in S : |U^{(n_1)} f_k(x) - U^{(n_2)} f_k(x)| \leq 1/m \text{ whenever } n_1, n_2 \geq n.\}$$

Now,

$$H_{m,n,k} = \bigcap_{n_1 \geq n} \bigcap_{n_2 \geq n} \{x \in S : |U^{(n_1)} f_k(x) - U^{(n_2)} f_k(x)| \leq 1/m\}.$$

Since P is regular, we can conclude that $H_{m,n,k}$ is Borel measurable for every $m, n, k \in \mathbb{N}$, thus Γ_c^P is Borel measurable as well.

Let μ be an invariant probability measure. By Kakutani's Ergodic Theorem there exists for every $k \in \mathbb{N}$ a Borel set C_k such that $\mu(C_k) = 1$ and $U^{(n)} f_k(x)$ converges for every $x \in C_k$ as $n \to \infty$. Define $C := \cap_{k \in \mathbb{N}} C_k$, then $\mu(C) = 1$ as well, and clearly $C \subset \Gamma_c^P$, so $\mu(\Gamma_c^P) = 1$. □

Note that $\Gamma_{cp}^P \subset \Gamma_c^P$, but they need not be equal, as the following example shows:

Example. Let $S = \mathbb{R}_+$ with Euclidean metric. Let $\Phi : S \to S, x \mapsto x + 1$, and let P_Φ be the Markov operator induced by Φ. Then for every $x \in S$

$$P_\Phi^{(n)} \delta_x = \frac{1}{n}(\delta_x + \delta_{x+1} + \cdots + \delta_{x+n-1}),$$

which does not converge to a measure as $n \to \infty$. So $\Gamma_{cp}^{P_\Phi} = \emptyset$. However, for every $f \in C_b(S)$ with bounded support, and $x \in S$,

$$\langle P_\Phi^{(n)} \delta_x, f \rangle = \frac{1}{n}(f(x) + f(x+1) + \cdots + f(x+n-1)) \to 0,$$

so $\Gamma_c = S$.

However, we do have the following:

Proposition 5.3.6. $\Gamma_{cp}^P = \Gamma_t^P \cap \Gamma_c^P$. *Consequently* Γ_{cp}^P *is Borel measurable and* $\mu(\Gamma_{cp}^P) = 1$ *for every invariant probability measure* μ.

Proof. It is obvious that $\Gamma_{cp}^P \subset \Gamma_t^P \cap \Gamma_c^P$.

Now take $x \in \Gamma_t^P \cap \Gamma_c^P$. Since $x \in \Gamma_t^P$, there is a subsequence $(P^{(n_m)} \delta_x)_m$ such that $(P^{(n_m)} \delta_x)_m$ converges in $\mathcal{S}_{\mathrm{BL}}$, say to $\mu \in \mathcal{M}^+(S)$. Then μ is a probability measure. By Lemma 2.3.12 $\langle P^{(n_m)} \delta_x, f \rangle \to \langle \mu, f \rangle$ for every $f \in C_b(S)$. Because $x \in \Gamma_c^P$ we know that for every $k \in \mathbb{N}$

$$\lim_{n \to \infty} \langle P^{(n)} \delta_x, f_k \rangle = \lim_{m \to \infty} \langle P^{(n_m)} \delta_x, f_k \rangle = \langle \mu, f_k \rangle.$$

Thus by Proposition 5.2.2 we know that $\|P^{(n)} \delta_x - \mu\|_{\mathrm{BL}}^* \to 0$ as $n \to \infty$, so $x \in \Gamma_{cp}^P$. So Proposition 5.3.4 and Proposition 5.3.5 imply that Γ_{cp}^P is Borel measurable and $\mu(\Gamma_{cp}^P) = 1$ for every invariant probability measure μ. □

For $f \in \mathrm{BM}(S)$, let

$$A_f := \{x \in S : (U^{(n)} f(x))_n \text{ converges.}\} \tag{5.7}$$

Then it can easily be shown that A_f is Borel measurable. Kakutani's Ergodic Theorem implies that $\mu(A_f) = 1$ for every invariant probability measure μ. We can define

$$f^*(x) := \begin{cases} \lim_{n\to\infty} U^{(n)} f(x) & \text{if } x \in A_f \cap \Gamma^P_{cp} \\ 0 & \text{if } x \notin A_f \cap \Gamma^P_{cp}. \end{cases}$$

Then f^* is measurable as the pointwise limit of a sequence of measurable functions. Observe that by Proposition 5.3.6, $U^{(n)} f \to f^*$ μ-a.e. for every invariant probability measure μ, thus f^* plays the role of g in Kakutani's Ergodic Theorem. Note that f^* does not depend on the specific invariant probability measure.

If $f \in C_b(S)$, then $\Gamma^P_{cp} \subset A_f$, since $U^{(n)} f(x) = \langle P^{(n)} \delta_x, f\rangle \to \langle \epsilon_x, f\rangle$, so in this case

$$f^*(x) = \begin{cases} \langle \epsilon_x, f\rangle & \text{if } x \in \Gamma^P_{cp} \\ 0 & \text{if } x \notin \Gamma^P_{cp}. \end{cases}$$

For $f \in \mathrm{BM}(S)$, we can define the function

$$f^\square(x) := \begin{cases} \langle \epsilon_x, f\rangle & \text{if } x \in \Gamma^P_{cp} \\ 0 & \text{if } x \notin \Gamma^P_{cp}. \end{cases}$$

Clearly if $f \in C_b(S)$ then $f^\square = f^*$.

In [134], f^* and $f^\square$ are also defined, though their definitions differ slightly from ours. As in [134], the functions f^* and $f^\square$ play an important role in the proof of the Borel measurability of Γ^P_{cpi}. We first show some properties of these functions, analogous to those shown in [134], namely that for every bounded Borel measurable function f, $f^\square$ is measurable and equals f^* μ-a.e. for every invariant probability measure μ. This implies that $f^\square$ can play the role of g in Kakutani's Ergodic Theorem for any invariant probability measure μ.

Let us define the following map from S to $\mathcal{S}_{\mathrm{BL}}$:

$$\eta(x) := \begin{cases} \epsilon_x & \text{if } x \in \Gamma^P_{cp} \\ 0 & \text{if } x \notin \Gamma^P_{cp}. \end{cases}$$

Lemma 5.3.7. *$\eta : S \to \mathcal{S}_{\mathrm{BL}}$ is strongly measurable. In particular, for every $f \in \mathrm{BM}(S)$, $f^\square$ is in $\mathrm{BM}(S)$.*

Proof. If η is strongly measurable, then it follows from Proposition 3.2.4 that $x \mapsto \langle \eta(x), f\rangle = f^\square(x)$ is measurable for every $f \in \mathrm{BM}(S)$.

First we define the map $\eta_n : S \to \mathcal{S}_{\mathrm{BL}}$ by

$$\eta_n(x) := \begin{cases} P^{(n)}\delta_x & \text{if } x \in \Gamma^P_{cp} \\ 0 & \text{if } x \notin \Gamma^P_{cp}. \end{cases}$$

We claim that η_n is strongly measurable. Since $\mathcal{S}_{\mathrm{BL}}$ is separable, it suffices to show that η_n is weakly measurable, by the Pettis Measurability Theorem.

Let $g \in \mathrm{BL}(S) \cong \mathcal{S}^*_{\mathrm{BL}}$, then $\langle \eta_n(x), g\rangle = \mathbb{1}_{\Gamma^P_{cp}}(x) U^{(n)} g(x)$, so from measurability of Γ^P_{cp} (Proposition 5.3.6) we obtain that the map $x \mapsto \langle \eta_n(x), g\rangle$ is Borel measurable. Thus η_n is indeed weakly measurable, hence strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$.

For every $x \in S$, $\|\eta_n(x) - \eta(x)\|^*_{\mathrm{BL}} \to 0$, so η is the pointwise limit of strongly measurable functions, hence strongly measurable. □

The proof of the following lemma is based on that of the locally compact version [134, Lemma 5.10].

Lemma 5.3.8. *Let μ be an invariant probability measure, $(f_n)_n$ a sequence in $\mathrm{BM}(S)$, $f \in \mathrm{BM}(S)$ and assume that*

(i) there exists an $M > 0$ such that $|f_n(x)| \leq M$ for every $n \in \mathbb{N}, x \in S$.

(ii) The sequence $(f_n)_n$ converges pointwise to f

(iii) $f^*_n = f^\square_n$ μ-a.e. for every $n \in \mathbb{N}$.

Then $f^* = f^\square$ μ-a.e.

Proof. Clearly $|f(x)| \leq M$ for every $x \in S$, thus $|f_n(x) - f(x)| \leq 2M$. Therefore $\int_S |f_n - f|\, d\mu \to 0$ by the Dominated Convergence Theorem. By Lemma 5.2.5

$$\int_S |U^{(m)} f_n - U^{(m)} f|\, d\mu \leq \int_S |f_n - f|\, d\mu \text{ for every } m, n \in \mathbb{N}. \tag{5.8}$$

Let $h \in \mathrm{BM}(S)$, then $\mu(A_h) = 1$, where A_h is the Borel set defined as in (5.7). So Proposition 5.3.6 implies that $A := \Gamma^P_{cp} \cap A_f \cap (\cap_{n\in\mathbb{N}} A_{f_n})$ is Borel measurable with $\mu(A) = 1$. And for every $x \in A$, $U^{(m)} f_n(x) \to f^*_n(x)$ as $m \to \infty$ for every $n \in \mathbb{N}$, and $U^{(m)} f(x) \to f^*(x)$.

By the Dominated Convergence Theorem we can conclude that $\int_S |U^{(m)} f - f^*|\, d\mu$ and $\int_S |U^{(m)} f_n - f^*_n|\, d\mu$ converge to zero as $m \to \infty$. From this and (5.8) it follows that $\int_S |f^*_n - f^*|\, d\mu \leq \int_S |f_n - f|\, d\mu \to 0$ as $n \to \infty$.

Also, since the f_n are uniformly bounded and $f_n \to f$ pointwise, the Dominated Convergence Theorem implies that $\langle \epsilon_x, f_n\rangle \to \langle \epsilon_x, f\rangle$ for every $x \in \Gamma^P_{cp}$. Thus $f^\square_n \to f^\square$ pointwise. Now,

$$|f^\square_n(x) - f^\square(x)| = |\langle \epsilon_x, f_n - f\rangle| \leq 2M\|\epsilon_x\|_{\mathrm{TV}} = 2M,$$

thus again by the Dominated Convergence Theorem $\int_S |f^\square_n - f^\square|\, d\mu \to 0$.

Note that $\int_S |f^*_n - f^\square_n|\, d\mu = 0$ for every $n \in \mathbb{N}$ by assumption. Thus

$$\int_S |f^* - f^\square|\, d\mu \leq \int_S |f^* - f^*_n|\, d\mu + \int_S |f^\square_n - f^\square|\, d\mu \to 0$$

as $n \to \infty$. Therefore $f^* = f^\square$ μ-a.e. $\square$

In order to prove that for general $f \in \mathrm{BM}(S)$, $f^* = f^\square$ μ-a.e. for every invariant probability measure μ, we will use a different approach than the one used in the proof of [134, Proposition 5.11], since that approach is based on a result, [134, Proposition 2.2], that uses the local compactness of the state space. Instead, we use the Monotone Class Theorem.

Proposition 5.3.9. *Let μ be an invariant probability measure. Then $f^* = f^\square$ μ-a.e. for every $f \in \mathrm{BM}(S)$.*

Proof. Step 1. Let $C \subset S$ be closed. Then $(\mathbb{1}_C)^* = (\mathbb{1}_C)^\square$ μ-a.e.

Let $f_n = (1 - nd(x, C))^+$, then $f_n \in \mathrm{BL}(S)$ and $f_n(x) \to \mathbb{1}_C(x)$ for every $x \in S$. Since $f_n \in C_b(S)$, $f_n^* = f_n^\square$ for every $n \in \mathbb{N}$. Also $|f_n(x)| \leq 1$ for every $x \in S, n \in \mathbb{N}$. Thus $(f_n)_n$ and $f = \mathbb{1}_C$ satisfy the conditions of Lemma 5.3.8, so $(\mathbb{1}_C)^* = (\mathbb{1}_C)^\square$ μ-a.e.

Step 2. $f^* = f^\square$ μ-a.e. for every $f \in \mathrm{BM}(S)$.

Let $\mathcal{H} = \{f \in \mathrm{BM}(S) : f^* = f^\square \ \mu\text{-a.e.}\}$. By Step 1 $\mathcal{H}$ contains $\mathbb{1}_C$ for every $C \subset S$ closed. Let $(f_n)_n$ be a sequence of elements of $\mathcal{H}$ with $f_n \geq 0$ and $f_n \uparrow f$, where f is bounded, then $f \in \mathcal{H}$ by Lemma 5.3.8. Therefore, since the collection of closed sets is a π-system for S and the σ-algebra generated by the closed sets is the Borel σ-algebra, application of the Monotone Class Theorem gives that $\mathcal{H} = \mathrm{BM}(S)$. $\square$

Proposition 5.3.10. *Γ^P_{cpi} is Borel measurable and $\mu(\Gamma^P_{cpi}) = 1$ for every invariant probability measure μ.*

Proof. Let $\{f_n : n \in \mathbb{N}\} \subset \mathrm{BL}(S)$ be an enumeration of the countable subset given by Proposition 5.2.2, and let $B_n := \{x \in \Gamma^P_{cp} : f_n^\square(x) = (Uf_n)^\square(x)\}$, then B_n is measurable, since $f_n^\square$ and $(Uf_n)^\square$ are measurable by Lemma 5.3.7. Let $x \in \Gamma^P_{cp}$. According to Proposition 5.2.2, $P\epsilon_x = \epsilon_x$ if and only if

$$\langle \epsilon_x, Uf_n \rangle = \langle P\epsilon_x, f_n \rangle = \langle \epsilon_x, f_n \rangle \text{ for every } n \in \mathbb{N}$$

Thus $x \in \Gamma^P_{cpi}$ if and only if $(Uf_n)^\square(x) = f_n^\square(x)$ for every $n \in \mathbb{N}$. So $\Gamma^P_{cpi} = \bigcap_{n=0}^\infty B_n$ is Borel measurable.

For any $x \in S$ and $a \in \mathbb{R}$, $U^{(m)}(Uf_n)(x) \to a$ as $m \to \infty$ if and only if $U^{(m)} f_n(x) \to a$ as $m \to \infty$, thus $f_n^* = (Uf_n)^*$. Since $f_n \in C_b(S)$, $f_n^* = f_n^\square$, so $(Uf_n)^* = f_n^\square$ for every $n \in \mathbb{N}$. Thus $B_n = \{x \in S : (Uf_n)^*(x) = (Uf_n)^\square(x)\} \cap \Gamma^P_{cp}$. By Proposition 5.3.6 and Proposition 5.3.9 we can conclude that $\mu(B_n) = 1$ for every $n \in \mathbb{N}$. Hence $\mu(\Gamma^P_{cpi}) = 1$. $\square$

Now we consider the following subset, technically defined by

$$\Gamma^P_{cpie} := \left\{ x \in \Gamma^P_{cpi} : \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) = 0 \text{ for every } f \in C_b(S). \right\} \quad (5.9)$$

Notice that Γ^P_{cpie} is well-defined, since f^* is a Borel measurable function and Γ^P_{cpi} is Borel measurable. Also, Γ^P_{cpi}, f^* and ϵ_x are all independent of the choice of the metric d, thus Γ^P_{cpie} also does not depend on the choice of the metric.

This set is similar to the set defined in [134, Section 6] in the setting of locally compact separable metric space, but with $C_b(S)$ replaced by $C_0(S)$. However, it can be shown using the Dominated Convergence Theorem that if $\int_{\Gamma^P_{cpi}}(f^*(y) - f^*(x))^2 d\epsilon_x(y) = 0$ for every $f \in C_0(S)$, then $\int_{\Gamma^P_{cpi}}(f^*(y) - f^*(x))^2 d\epsilon_x(y) = 0$ for every $f \in C_b(S)$, so our formulation of Γ^P_{cpie} generalises the one in [134] to the setting of Polish spaces.

We can reformulate Γ^P_{cpie}:

Lemma 5.3.11. *Let $\{f_n : n \in \mathbb{N}\} \subset \mathrm{BL}(S)$ be an enumeration of the countable subset from Proposition 5.2.2. Then*

$$\Gamma^P_{cpie} = \left\{ x \in \Gamma^P_{cpi} : \int_{\Gamma^P_{cpi}} (f^*_n(y) - f^*_n(x))^2 d\epsilon_x(y) = 0 \text{ for every } n \in \mathbb{N}. \right\}$$

Proof. Clearly '$\subseteq$' holds. Now let $x \in \Gamma^P_{cpi}$ be such that $\int_{\Gamma^P_{cpi}}(f^*_n(y) - f^*_n(x))^2 d\epsilon_x(y) = 0$ for every $n \in \mathbb{N}$. Then there are Borel sets $B_n \subset \Gamma^P_{cpi}$ with $\epsilon_x(B_n) = 1$, such that $f^*_n(y) = f^*_n(x)$ for every $y \in B_n$.

Let $B = \bigcap_{n=1}^{\infty} B_n$, then $\epsilon_x(B) = 1$. Let $n \in \mathbb{N}$. Since $f_n \in C_b(S)$, $f^*_n(y) = \langle \epsilon_y, f_n \rangle$ for every $y \in \Gamma^P_{cp}$, so for every $y \in B$ $\langle \epsilon_y, f_n \rangle = \langle \epsilon_x, f_n \rangle$. This holds for every $n \in \mathbb{N}$, thus by Proposition 5.2.2, $\epsilon_x = \epsilon_y$ for every $y \in B$. Thus $f^*(y) = f^*(x)$ for every $y \in B$ and $f \in C_b(S)$. Since $\epsilon_x(B) = 1$, $x \in \Gamma^P_{cpie}$. □

Lemma 5.3.12. *Let μ be an invariant probability measure. Then*

$$\int_{\Gamma^P_{cpi}} \left(\int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \right) d\mu(x) = 0$$

for every $f \in C_b(S)$.

Proof. Let $f \in C_b(S)$ and let $x \in \Gamma^P_{cpi}$, then $\epsilon_x(\Gamma^P_{cpi}) = 1$ by Proposition 5.3.10, so

$$\begin{aligned}
\int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) &= \int_S (f^*(y) - f^*(x))^2 d\epsilon_x(y) \\
&= \langle \epsilon_x, (f^*)^2 \rangle - 2f^*(x)\langle \epsilon_x, f^* \rangle + (f^*(x))^2 \epsilon_x(\Gamma^P_{cpi}) \\
&= ((f^*)^2)^{\square}(x) - 2f^*(x) f^{\square}(x) + (f^*)^2(x).
\end{aligned}$$

So the map from Γ^P_{cpi} to $\mathbb{R}$ given by $x \mapsto \int_{\Gamma^P_{cpi}}(f^*(y) - f^*(x))^2 d\epsilon_x(y)$ is Borel measurable since Γ^P_{cpi} is Borel measurable by Proposition 5.3.10, f^* is Borel measurable and $((f^*)^2)^{\square}$ and $f^{\square}$ are Borel measurable by Lemma 5.3.7.

By Proposition 5.3.9 $((f^*)^2)^\square = ((f^*)^2)^*$ μ-a.e. and $f^\square = f^*$ since $f \in C_b(S)$. Also note that by Kakutani's Ergodic Theorem

$$\langle \mu, ((f^*)^2)^* \rangle = \langle \mu, (f^*)^2 \rangle.$$

Thus, using the fact that $\mu(\Gamma^P_{cpi}) = 1$,

$$\begin{aligned}\int_{\Gamma^P_{cpi}} \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \, d\mu(x) &= \langle \mu, ((f^*)^2)^\square - 2f^* f^\square + (f^*)^2 \rangle \\ &= \langle \mu, ((f^*)^2)^* \rangle - 2\langle \mu, f^* f^* \rangle + \langle \mu, (f^*)^2 \rangle = 0.\end{aligned}$$

□

The following theorem generalises [134, Theorem 6.1 and Theorem 6.3].

Theorem 5.3.13. *Γ^P_{cpie} is Borel measurable in S and $\mu(\Gamma^P_{cpie}) = 1$ for every invariant probability measure μ.*

Proof. Let $x \in \Gamma^P_{cpi}$. Let $\{f_n : n \in \mathbb{N}\}$ be the countable set from Proposition 5.2.2. For every $n \in \mathbb{N}$ we define $g_n : \Gamma^P_{cpi} \to \mathbb{R}$ as follows:

$$g_n(x) := \int_{\Gamma^P_{cpi}} (f_n^*(y) - f_n^*(x))^2 d\epsilon_x(y).$$

We have shown in the proof of Lemma 5.3.12 that

$$g_n(x) = ((f_n^*)^2)^\square(x) - 2f_n^*(x)(f_n)^\square(x) + (f_n^*)^2(x),$$

and thus g_n is Borel measurable for every $n \in \mathbb{N}$. By Lemma 5.3.11 $\Gamma^P_{cpie} = \bigcap_{n=1}^\infty g_n^{-1}(\{0\})$, hence Γ^P_{cpie} is a Borel subset of Γ^P_{cpi}, thus Borel measurable.

Let μ be an invariant probability measure. By Lemma 5.3.12 $\int_{\Gamma^P_{cpi}} g_n \, d\mu = 0$ and since the g_n are positive, $g_n(x) = 0$ for μ-a.e. $x \in \Gamma^P_{cpi}$. So $\mu(g_n^{-1}(\{0\})) = \mu(\Gamma^P_{cpi}) = 1$ for all $n \in \mathbb{N}$ by Proposition 5.3.10. Thus $\mu(\Gamma^P_{cpie}) = 1$ as well. □

5.4 The ergodic decompositions

We start in Section 5.4.1 with the definition and some properties of ergodic measures. In Section 5.4.2 we show that the set Γ^P_{cpie} defined in Section 5.3 consists of exactly those $x \in \Gamma^P_{cpi}$ for which ϵ_x is ergodic. Moreover, we show that *every* ergodic measure is of the form ϵ_x for some $x \in \Gamma^P_{cpie}$. Using these results we prove an integral decomposition of invariant probability measures in terms of ergodic measures. In Section 5.4.3 we complete the Yosida-type ergodic decomposition of the state space S.

Let $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ be a regular Markov operator with dual U.

5.4.1 Ergodic measures

In the literature one can find various different (equivalent) definitions for ergodic measures. See e.g. the classical paper by Rosenblatt [106] about regular Markov operators on general measurable spaces. In this section we will use the definition from [56, 57, 132].

A Borel measurable subset E of S is an *invariant set* (or P-invariant set) if $P\delta_x(E) = 1$ for every $x \in E$. We define an *ergodic measure* μ (with respect to P) to be an invariant probability measure, such that $\mu(E) = 0$ or 1 whenever E is an invariant set. To P and $x \in S$ we may associate a Markov chain $(X_n^x)_n$ such that the law of X_n^x is given by $P^n\delta_x$. E is an invariant set if, for every $x \in E$, the associated Markov chain starting in $x \in E$ will remain in E with probability 1.

We shall write $\mathcal{P}_{\text{inv}}(S)$ to denote the convex set of invariant probability measures and $\mathcal{P}_{\text{erg}}(S)$ to denote the subset of ergodic measures, both with respect to P.

Note that a Borel set E is invariant if and only if $U\mathbb{1}_E \geq \mathbb{1}_E$. Let μ be an invariant probability measure. A Borel measurable subset E is a *μ-almost (P-)invariant set* if $U\mathbb{1}_E \geq \mathbb{1}_E$ μ-a.e. If E is μ-almost invariant, then

$$\begin{aligned} 0 &\leq \int_S |U\mathbb{1}_E(x) - \mathbb{1}_E(x)|\, d\mu(x) = \int_S U\mathbb{1}_E(x)\, d\mu(x) - \int_S \mathbb{1}_E(x)\, d\mu(x) \\ &= \langle P\mu - \mu, \mathbb{1}_E\rangle = 0. \end{aligned}$$

Thus $U\mathbb{1}_E(x) = \mathbb{1}_E(x)$ for μ-a.e. $x \in S$.

The following lemma follows easily from the definitions.

Lemma 5.4.1. *Let μ be an invariant probability measure. The following statements are equivalent for a Borel set $E \subset S$:*

(i) E is μ-almost invariant.

(ii) $P\delta_x(E) = 1$ for μ-almost every $x \in E$.

(iii) $\int_E P\delta_x(E)\, d\mu(x) = \mu(E)$.

Thus every invariant set is μ-almost invariant for every invariant probability measure μ, but not necessarily the other way around.

In some places in the literature, e.g. in [2, Definition 19.23], an ergodic measure is defined to be an invariant probability measure μ, such that $\mu(A) = 0$ or $\mu(A) = 1$ for every μ-almost invariant set A. We first show that this alternative definition is equivalent to our definition for ergodic measures.

Lemma 5.4.2. *Let μ be an invariant probability measure and let $B \subset S$ be a μ-almost P-invariant Borel set. Then there is a Borel measurable $C \subset B$ such that C is invariant and $\mu(C) = \mu(B)$.*

Proof. Let $B_0 = B$, $B_n := \{x \in B_{n-1} : U\mathbb{1}_{B_{n-1}}(x) = P\delta_x(B_{n-1}) = 1\}$. B_0 is measurable by assumption. Suppose B_{n-1} is measurable for some $n \in \mathbb{N}$, then $B_n = (U\mathbb{1}_{B_{n-1}})^{-1}(\{1\})$ is measurable, so by induction B_n is measurable for every $n \in \mathbb{N}$.

Now suppose that $\mu(B_{n-1}) = \mu(B)$ for some $n \in \mathbb{N}$. Since $B_{n-1} \subset B$, this implies that $\mathbb{1}_{B_{n-1}} = \mathbb{1}_B$ μ-a.e. So by Lemma 5.2.5 $U\mathbb{1}_{B_{n-1}} = U\mathbb{1}_B$ μ-a.e. Thus

$$\mu(\{x \in B : U\mathbb{1}_{B_{n-1}}(x) = 1\}) = \mu(\{x \in B : U\mathbb{1}_B(x) = 1\}) = \mu(B).$$

Now,

$$\begin{aligned} \mu(B_n) &= \mu(B_{n-1} \cap \{x \in B : U\mathbb{1}_{B_{n-1}}(x) = 1\}) = \mu(B_{n-1} \cap B) \\ &= \mu(B_{n-1}) = \mu(B), \end{aligned}$$

thus $\mu(B_n) = \mu(B)$ for every $n \in \mathbb{N}$. Let $C = \bigcap_{n=1}^{\infty} B_n$, then $\mu(C) = \lim_{n\to\infty} \mu(B_n) = \mu(B)$. Let $x \in C$, then $P\delta_x(C) = \lim_{n\to\infty} P\delta_x(B_n)$. Also, for every $n \in \mathbb{N}$, $x \in B_{n+1}$, so $P\delta_x(B_n) = 1$ and thus $P\delta_x(C) = 1$ for every $x \in C$.

Hence C is an invariant Borel set of S with $C \subset B$ and $\mu(C) = \mu(B)$. □

Corollary 5.4.3. *Let μ be an invariant probability measure and $B \subset S$ Borel such that $\mu(B) = 1$. Then there exists a Borel measurable $C \subset B$ such that C is invariant and $\mu(C) = 1$.*

Proof. $\mathbb{1}_B = \mathbb{1}_S$ μ-a.e., so by Lemma 5.2.5 $U\mathbb{1}_B = U\mathbb{1}_S$ μ-a.e. For every $x \in S$ $U\mathbb{1}_S(x) = P\delta_x(S) = 1 = \mathbb{1}_S(x)$. Thus $U\mathbb{1}_B = U\mathbb{1}_S = \mathbb{1}_S = \mathbb{1}_B$ μ-a.e. Application of Lemma 5.4.2 concludes the proof. □

A proof of the following result, for a more general state space, i.e. just metrisable, can be found in [2, Theorem 19.25].

Theorem 5.4.4. *The extreme points of $\mathcal{P}_{\text{inv}}(S)$ are exactly the ergodic measures.*

The following theorem gives an equivalent condition for invariant measures to be ergodic.

Theorem 5.4.5. *Let μ be an invariant probability measure. Then the following statements are equivalent:*

(i) μ is ergodic.

(ii) There exists a Borel subset B of S such that $\mu(B) = 1$ and such that $U^{(n)}f(x)$ converges to $\langle \mu, f\rangle$ as $n \to \infty$ for every $x \in B$ and $f \in C_b(S)$.

An analogous result has been proven by Zaharopol [132, Lemma 3.3.1] in the setting of a locally compact separable metric space, with $C_b(S)$ replaced by $C_0(S)$. A crucial ingredient in the proof of [132, Lemma 3.3.1] is the separability of the Banach space $C_0(S)$. In our Polish setting, $C_0(S)$ cannot play a role, since it need not contain any non-zero functions. The bigger space $C_b(S)$ is not separable in general, however

Proposition 5.2.2 will be exactly what we need in this situation. We first need some preliminary results.

Lemma 5.4.6. *Let μ be an ergodic measure and $f \in \mathrm{BM}(S)$. Then $U^{(n)}f(x) \to \langle \mu, f\rangle$ for μ-a.e. $x \in S$, i.e. $f^* = \langle \mu, f\rangle$ μ-a.e.*

Proof. In fact, the statement holds more generally for a regular Markov operator on a measurable space. For a proof see [57, Proposition 2.4.2]. □

Proof. (Theorem 5.4.5)

(i) ⇒ (ii): Let $\{f_n : n \in \mathbb{N}\}$ be an enumeration of the countable subset in $\mathrm{BL}(S)$ given by Proposition 5.2.2. Since μ is ergodic, Lemma 5.4.6 implies that for every $n \in \mathbb{N}$ there exists a Borel set B_n with $\mu(B_n) = 1$, such that $\lim_{m\to\infty} U^{(m)}f_n(x) \to \langle \mu, f_n\rangle$ for every $x \in B_n$. Set $B := \bigcap_{n=1}^{\infty} B_n$, then $\mu(B) = 1$. For every $x \in B$ and $n \in \mathbb{N}$ we know that

$$\langle P^{(m)}\delta_x, f_n\rangle = U^{(m)}f_n(x) \to \langle \mu, f_n\rangle \text{ as } m \to \infty,$$

thus by Proposition 5.2.2,

$$U^{(m)}f(x) = \langle P^{(m)}\delta_x, f\rangle \to \langle \mu, f\rangle \text{ as } m \to \infty,$$

for every $f \in C_b(S)$ and every $x \in B$.

(ii)⇒ (i): Let μ be an invariant probability measure such that there exists a Borel set B with $\mu(B) = 1$ and $U^{(n)}f(x) \to \langle \mu, f\rangle$ for every $f \in C_b(S)$ and $x \in B$. Suppose that $\mu = \lambda\mu_1 + (1-\lambda)\mu_2$ for some $0 < \lambda < 1$ and μ_1, μ_2 invariant probability measures, then $\mu_1(B) = \mu_2(B) = 1$. Let $f \in C_b(S)$. By Kakutani's Ergodic Theorem there is a $g \in \mathrm{BM}(S)$ such that $U^{(n)}f \to g$ μ_1-a.e. Since $\mu_1(B) = 1$, $U^{(n)}f \to \langle \mu, f\rangle$ μ_1-a.e., thus $g = \langle \mu, f\rangle$ μ_1-a.e. The same holds for μ_2, so in particular $\langle \mu_1, f\rangle = \langle \mu, f\rangle = \langle \mu_2, f\rangle$ for every $f \in C_b(S)$, thus $\mu_1 = \mu_2$. This implies that μ is an extreme point of the set of invariant probability measures, thus μ is ergodic by Theorem 5.4.4. □

5.4.2 An integral decomposition of invariant measures

It need not be true that ϵ_x is an ergodic measure whenever $x \in \Gamma^P_{cpi}$. A very simple example is given in [132, Example 2.2.4] which shows that even in a more restrictive setting of a Markov-Feller operator in a compact metric space this need not be the case. We will show that the set Γ^P_{cpie}, defined in Section 5.3, consists of exacty those $x \in \Gamma^P_{cpi}$ for which ϵ_x *is* ergodic. We will use this to give an integral decomposition of ergodic measures.

We define the equivalence relation $\sim$ on Γ^P_{cpie} as follows: $x \sim y$ if and only if $\epsilon_x = \epsilon_y$. Let $[x]$ be the equivalence class of $x \in \Gamma^P_{cpie}$ defined by $\sim$.

The following theorem extends [134, Lemma 6.4 and Theorem 6.5]:

Theorem 5.4.7. *(i) For every $x \in \Gamma^P_{cpie}$ the set $[x]$ is Borel measurable, ϵ_x is an ergodic measure and $\epsilon_x([x]) = 1$.*

(ii) Conversely, any ergodic measure μ is of the form $\mu = \epsilon_x$ for some $x \in \Gamma^P_{cpie}$.

Proof. (i) Let $x \in \Gamma^P_{cpie}$. Let $\{f_n : n \in \mathbb{N}\}$ be an enumeration of the countable subset of $\mathrm{BL}(S)$ from Proposition 5.2.2. We define

$$F_n = \{y \in \Gamma^P_{cpi} : f_n^*(y) = f_n^*(x)\} = (f_n^*)^{-1}(f_n^*(x))$$

and $E_n = \Gamma^P_{cpie} \cap F_n$. Since f_n^* is a Borel measurable function and Γ^P_{cpie} is a Borel set, E_n and F_n are also Borel sets. From Proposition 5.2.2 it follows that $x \sim y$ if and only if $\langle \epsilon_x, f_n \rangle = \langle \epsilon_y, f_n \rangle$ for every $n \in \mathbb{N}$, or equivalently $f_n^*(x) = f_n^*(y)$ for every $n \in \mathbb{N}$. Thus $[x] = \bigcap_{n=1}^{\infty} E_n$ is measurable.

Since $\Gamma^P_{cpie} \subset \Gamma^P_{cpi}$, ϵ_x is an invariant probability measure. For every $y \in [x]$ and $f \in C_b(S)$,

$$U^{(n)} f(y) \to \langle \epsilon_y, f \rangle = \langle \epsilon_x, f \rangle.$$

So if we can show that $\epsilon_x([x]) = 1$, then μ is ergodic by Theorem 5.4.5. Since $x \in \Gamma^P_{cpie}$ we know that $\int_{\Gamma^P_{cpi}} (f_n^*(y) - f_n^*(x))^2 \, d\epsilon_x(y) = 0$ for every $n \in \mathbb{N}$. This implies that $\epsilon_x(F_n) = 1$, since $\epsilon_x(\Gamma^P_{cpi}) = 1$ by Proposition 5.3.10. So $\epsilon_x(E_n) = 1$ as well, since $\epsilon_x(\Gamma^P_{cpie}) = 1$ by Theorem 5.3.13, thus $\epsilon_x([x]) = 1$.

(ii) Let μ be an ergodic measure. Then by Theorem 5.4.5 there is a measurable $A \subset S$, such that $\mu(A) = 1$ and $U^{(n)} f(x) \to \langle \mu, f \rangle$ for every $f \in C_b(S)$ and $x \in A$. Since μ is an invariant probability measure, it follows that $A \subset \Gamma^P_{cpi}$. A is not empty since $\mu(A) = 1$, so there is an $x \in A$ and then clearly $\mu = \epsilon_x$. Now, for every $y \in A$ and $f \in C_b(S)$, $f^*(y) = \langle \epsilon_x, f \rangle$, so since $\epsilon_x(\Gamma^P_{cpi} \backslash A) = 0$

$$\begin{aligned} \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) &= \int_A (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \\ &= \int_A (\langle \epsilon_x, f \rangle - \langle \epsilon_x, f \rangle)^2 \, d\epsilon_x(y) = 0, \end{aligned}$$

for every $f \in C_b(S)$. So $x \in \Gamma^P_{cpie}$. □

Corollary 5.4.8. *Let μ, ν be ergodic measures. Then either $\mu = \nu$ or μ and ν are mutually singular.*

Proof. By Theorem 5.4.7 there are $x, y \in \Gamma^P_{cpie}$ such that $\mu = \epsilon_x$ and $\nu = \epsilon_y$. If $y \sim x$ then $\epsilon_x = \epsilon_y$. If $x \not\sim y$ then $[x] \cap [y] = \emptyset$ and $\epsilon_x([x]) = 1$ and $\epsilon_y([y]) = 1$. Thus $\epsilon_x \perp \epsilon_y$. □

According to Theorem 5.3.13 and Theorem 5.4.7 the following holds:

Corollary 5.4.9. *The following statements are equivalent:*

(i) There exists an invariant probability measure.

(ii) Γ^P_{cpie} *is not empty.*

(iii) There exists an ergodic measure.

This implies

Corollary 5.4.10. *If there exists only one invariant probability measure* μ*, then* μ *is ergodic.*

The following theorem gives an integral decomposition of invariant probability measures into ergodic measures.

Theorem 5.4.11. *Let* μ *be an invariant probability measure. Then the map*

$$x \mapsto \begin{cases} \epsilon_x & \text{if } x \in \Gamma^P_{cpie} \\ 0 & \text{if } x \notin \Gamma^P_{cpie}. \end{cases} \tag{5.10}$$

is strongly measurable from S *to* $\mathcal{S}_{\mathrm{BL}}$ *and*

$$\mu = \int_{\Gamma^P_{cpie}} \epsilon_x \, d\mu(x),$$

as Bochner integral in $\mathcal{S}_{\mathrm{BL}}$*.*

Proof. Corollary 3.2.8 and Theorem 5.3.13 imply that

$$\mu = \int_{\Gamma^P_{cpie}} \delta_x \, d\mu(x).$$

By Proposition 3.3.1 and the invariance of μ we obtain for every $n \in \mathbb{N}$

$$\mu = P^{(n)}\mu = \int_{\Gamma^P_{cpie}} P^{(n)}\delta_x \, d\mu(x).$$

Now, $P^{(n)}\delta_x \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$ for every $x \in \Gamma^P_{cpie}$. So the measurability of Γ^P_{cpie} implies that the map defined in (5.10) is strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$, and by the Dominated Convergence Theorem we can conclude that

$$\mu = \lim_{n\to\infty} P^{(n)}\mu = \int_{\Gamma^P_{cpie}} \lim_{n\to\infty} P^{(n)}\delta_x \, d\mu(x) = \int_{\Gamma^P_{cpie}} \epsilon_x \, d\mu(x).$$

□

Corollary 5.4.12. *Suppose* P *has a unique ergodic measure* μ^**. Then* μ^* *is the only invariant probability measure.*

Proof. Let μ be an invariant probability measure. Then by Theorem 5.4.11 $\mu = \int_{\Gamma^P_{cpie}} \epsilon_x \, d\mu(x)$. ϵ_x is ergodic, so $\epsilon_x = \mu^*$ for every $x \in \Gamma^P_{cpie}$. The result now follows from $\mu(\Gamma^P_{cpie}) = 1$ (Theorem 5.3.13). □

Example. As an illustration, we apply some of the above results in the simple setting given in Example 5.3.1. Let $S = \{x_1, \cdots, x_N\}$ be a finite metric space and P a Markov operator on S. We can represent $\mathcal{M}(S)$ by $\mathbb{R}^N$. Let Q be the $N \times N$ stochastic matrix associated to P. We saw in Example 5.3.1 that $\Gamma_t^P = \Gamma_{cp}^P = \Gamma_{cpi}^P = S$. Consequently there are invariant measures, thus Corollary 5.4.9 implies that Γ_{cpie}^P is non-empty. However, Γ_{cpie}^P need not be equal to S in this setting: see [132, Example 2.2.4]. We can write $\Gamma_{cpie}^P = \{x_{i_1}, \cdots, x_{i_m}\}$ for some $1 \leq m \leq N$. By Theorem 5.3.13, every invariant probability measure is concentrated on Γ_{cpie}^P, so of the form

$$\sum_{k=1}^{m} \alpha_i \delta_{x_{i_k}} \text{ for some } \alpha_i \geq 0 \text{ such that } \sum_{k=1}^{m} \alpha_i = 1. \tag{5.11}$$

Clearly not every measure of the form (5.11) is invariant. Let μ be an ergodic measure. Since μ is invariant (and S is finite), $\operatorname{supp}(\mu)$ is P-invariant. Corollary 5.4.8 implies that $\operatorname{supp}(\mu) \cap \operatorname{supp}(\nu) = \emptyset$ for two different ergodic measures μ and ν. The sets $\operatorname{supp}(\mu)$ correspond to a set of coordinates in $\mathbb{R}^N$. Such a set is known in the literature as an ergodic set (see e.g. [15]). For $x \in \Gamma_{cpie}^P$ we can consider the equivalent class $[x]$. Note that this class may consist of more points than the ergodic measure given by ϵ_x: Take for instance $S = \{x_1, x_2\}$ and

$$Q = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then the only invariant probability measure is δ_{x_2}, but $\Gamma_{cpie}^P = S$. By Theorem 5.4.11 we can obtain every invariant probability measure as a finite sum of ergodic measures.

Theorem 5.4.13. *The following statements hold:*

(i) $\mathcal{P}_{\text{inv}}(S)$ *is dense in the closed convex hull of* $\mathcal{P}_{\text{erg}}(S)$ *in* $\mathcal{S}_{\text{BL}}$.

(ii) If P *is a Markov-Feller operator then* $\mathcal{P}_{\text{inv}}(S)$ *equals the closed convex hull of* $\mathcal{P}_{\text{erg}}(S)$ *in* $\mathcal{S}_{\text{BL}}$.

Proof. If there exists no invariant probability measure then both $\mathcal{P}_{\text{inv}}(S)$ and $\mathcal{P}_{\text{erg}}(S)$ are empty and then *(i)* and *(ii)* hold. So suppose there exist invariant probability measures.

(i) Let μ be an invariant probability measure. By Theorem 5.4.11 $\mu = \int_{\Gamma_{cpie}^P} \epsilon_x \, d\mu(x)$. By [26, Corollary 8]

$$\frac{1}{\mu(\Gamma_{cpie}^P)} \int_{\Gamma_{cpie}^P} \epsilon_x \, d\mu(x) \in \overline{\operatorname{conv}}(\{\epsilon_x : x \in \Gamma_{cpie}^P\}).$$

By Theorem 5.3.13 $\mu(\Gamma_{cpie}^P) = 1$ and by Theorem 5.4.7 $\epsilon_x \in \mathcal{P}_{\text{erg}}(S)$ for every $x \in \Gamma_{cpie}^P$, thus

$$\mu = \int_{\Gamma_{cpie}^P} \epsilon_x \, d\mu(x) \in \overline{\operatorname{conv}}(\mathcal{P}_{\text{erg}}(S)).$$

(ii) From the linearity of P it follows that any convex combination of two invariant probability measures is again an invariant probability measures. Let $(\mu_n)_n$ be a Cauchy sequence of invariant probability measures with respect to $\|\cdot\|^*_{\mathrm{BL}}$. Since $\mathcal{P}(S)$ is closed in $\mathcal{S}_{\mathrm{BL}}$, there is a $\mu \in \mathcal{P}(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$. Then, since P is Markov-Feller, $\mu_n = P\mu_n \to P\mu$, so $\mu = P\mu$. Thus $\mathcal{P}_{\mathrm{inv}}(S)$ is closed and convex in $\mathcal{S}_{\mathrm{BL}}$. So the closed convex hull of $\mathcal{P}_{\mathrm{erg}}(S)$ is contained in $\mathcal{P}_{\mathrm{inv}}(S)$. □

If P is not a Markov-Feller operator, $\mathcal{P}_{\mathrm{inv}}(S)$ need not be closed in $\mathcal{S}_{\mathrm{BL}}$:

Example. Let $S = [0,1]$ with Euclidean metric and define $\Phi : S \to S$ by

$$\Phi(x) := \begin{cases} x & \text{if } x \neq 1; \\ 0 & \text{if } x = 1. \end{cases}$$

Then Φ is Borel measurable, thus, as remarked in Section 3.3, $P_\Phi \mu := \mu \circ \Phi^{-1}$ defines a regular Markov operator $P_\Phi : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$. Clearly $\delta_x \in \mathcal{P}_{\mathrm{inv}}(S)$ if and only if $x \in [0,1)$, but $\|\delta_x - \delta_1\|^*_{\mathrm{BL}} \leq |x-1| \to 0$ as $x \to 1$, so $\mathcal{P}_{\mathrm{inv}}(S)$ is not closed in $\mathcal{S}_{\mathrm{BL}}$.

5.4.3 Full Yosida-type ergodic decomposition

Let R be a Borel set in S. There is a natural bijection between $\mathcal{M}(R)$ and $\mathcal{M}_R(S) := \{\mu \in \mathcal{M}(S) : |\mu|(S \backslash R) = 0\}$: one can extend any finite Borel measure μ on R to a finite Borel measure $\overline{\mu}$ on S, by defining $\overline{\mu}(E) := \mu(E \cap R)$ for every Borel set E in S. Then clearly $|\overline{\mu}|(S \backslash R) = 0$. On the other hand, if ν is a finite Borel measure on S such that $|\nu|(S \backslash R) = 0$, then its restriction to R defines a Borel measure μ such that $\overline{\mu} = \nu$.

Let R be an invariant Borel set. Then P leaves $\mathcal{M}_R(S)$ invariant: if $\mu \in \mathcal{M}_R(S)$, then by Proposition 3.3.1

$$P|\mu|(R) = \int_S P\delta_x(R)\, d|\mu|(x) \geq \int_R d|\mu| = |\mu|(R) = |\mu|(S).$$

Thus $|P\mu|(S \backslash R) \leq P|\mu|(S \backslash R) = |\mu|(S \backslash R) = 0$. So we can restrict P to $\mathcal{M}_R(S)$. This gives a 'restriction' of P to a regular Markov operator on $\mathcal{M}(R)$.

The following theorem extends the Yosida-type ergodic decomposition by Hernández-Lerma and Lasserre [56, Proposition 4.5] for regular Markov operators on locally compact separable metric spaces.

Theorem 5.4.14. *Let S be a Polish space and P a regular Markov operator. If there exist ergodic measures or, equivalently, if Γ^P_{cpie} is not empty, then for every $[x] \in \Gamma^P_{cpie}/\sim$ the following statements hold:*

(i) There exists an invariant Borel set $S_{[x]} \subset [x]$ such that $\epsilon_x(S_{[x]}) = 1$.

(ii) ϵ_x is the unique invariant probability measure of $P_{[x]}$, where $P_{[x]}$ is the restriction of P to $\mathcal{M}(S_{[x]})$.

(iii) $P_{[x]}$ *is ergodic in the sense that* $S_{[x]}$ *cannot be written as the union of two disjoint* $P_{[x]}$*-invariant sets* A *and* B *with* $\epsilon_x(A) > 0$ *and* $\epsilon_x(B) > 0$.

Proof. Let $x \in \Gamma^P_{cpie}$.

(i) By Theorem 5.4.7 $\epsilon_x([x]) = 1$, so by Corollary 5.4.3 there is an invariant Borel set $S_{[x]} \subset [x]$ such that $\epsilon_x(S_{[x]}) = \epsilon_x([x]) = 1$.

(ii) Since $S_{[x]}$ is invariant, we can restrict P to a regular Markov operator $P_{[x]}$ on $\mathcal{M}(S_{[x]})$. Let μ be a $P_{[x]}$-invariant probability measure on $S_{[x]}$ and $\overline{\mu}$ the extension of μ to S. Then $\overline{\mu}$ is an invariant probability measure on S such that $\overline{\mu}(S_{[x]}) = 1$, thus $\overline{\mu}(S_{[x]}) = 1$. Now, by Theorem 5.4.11 and since $\overline{\mu}(S \backslash S_{[x]}) = 0$

$$\overline{\mu} = \int_{\Gamma^P_{cpie}} \epsilon_y \, d\overline{\mu}(y) = \int_{S_{[x]}} \epsilon_y \, d\mu(y) = \int_{S_{[x]}} \epsilon_x \, d\mu(y) = \epsilon_x,$$

thus μ is the restriction of ϵ_x to $S_{[x]}$.

(iii) Let A, B be disjoint $P_{[x]}$-invariant Borel subsets of $S_{[x]}$ such that $\epsilon_x(A) > 0$ and $\epsilon_x(B) > 0$. Then A, B are disjoint invariant Borel subsets of S, thus by ergodicity of ϵ_x, $\epsilon_x(A) = \epsilon_x(B) = 1$. But then $\epsilon_x(A \cup B) = 2$, which gives a contradiction, since ϵ_x is a probability measure. □

In the next section we will show that $\bigcup_{x \in \Gamma^P_{cpie}} S_{[x]}$ is an invariant Borel set.

5.5 Application to convergence of Cesàro averages

Section 5.3 dealt with convergence properties of Cesàro averages of iterations under P of Dirac measures δ_x. In this section we apply these results to show that if measures are concentrated on Γ^P_{cp}, Γ^P_{cpi} or $[z]$ for some $z \in \Gamma^P_{cpie}$, then the Cesàro averages of these measures converge in $\mathcal{S}_{\mathrm{BL}}$ to measures, invariant measures and ergodic measures respectively. Consequently, it is of interest to be able to specify these sets in particular cases. Throughout this section we assume that the Markov operator P is regular.

Let

$$\mathcal{P}^P_t := \{\mu \in \mathcal{P}(S) : (P^{(n)}\mu)_n \text{ is tight}\},$$

$$\mathcal{P}^P_{cp} := \{\mu \in \mathcal{P}(S) : (P^{(n)}\mu)_n \text{ converges in } \mathcal{S}_{\mathrm{BL}} \text{ as } n \to \infty\},$$

and define

$$\epsilon_\mu := \lim_{n \to \infty} P^{(n)}\mu$$

for every $\mu \in \mathcal{P}^P_{cp}$. Note that by definition $\delta_x \in \mathcal{P}^P_{cp}$ if and only if $x \in \Gamma^P_{cp}$, and $\epsilon_{\delta_x} = \epsilon_x$. Analogous to Γ^P_{cpi} and Γ^P_{cpie} we can define

$$\mathcal{P}^P_{cpi} := \{\mu \in \mathcal{P}^P_{cp} : \epsilon_\mu \text{ is invariant}\}$$

and

$$\mathcal{P}^P_{cpie} := \{\mu \in \mathcal{P}^P_{cpi} : \epsilon_\mu \text{ is ergodic.}\}$$

Proposition 5.5.1. *Let $\mu \in \mathcal{P}(S)$ such that $\mu(\Gamma^P_{cp}) = 1$. Then $\mu \in \mathcal{P}^P_{cp}$, $\epsilon_\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x)$ and for every $f \in C_b(S)$, $\langle \epsilon_\mu, f\rangle = \langle \mu, f^*\rangle$.*

Proof. By Proposition 3.3.1 and $\mu(\Gamma^P_{cp}) = 1$, we have

$$P^{(n)}\mu = \int_{\Gamma^P_{cp}} P^{(n)}\delta_x \, d\mu(x).$$

Now, for every $x \in \Gamma^P_{cp}$, $P^{(n)}\delta_x \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$, so by the Dominated Convergence Theorem we obtain that $P^{(n)}\mu$ converges, thus $\mu \in \mathcal{P}^P_{cp}$ and

$$\lim_{n\to\infty} P^{(n)}\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x).$$

According to Proposition 3.2.5, for every $f \in C_b(S)$,

$$\langle \epsilon_\mu, f\rangle = \int_{\Gamma^P_{cp}} \langle \epsilon_x, f\rangle \, d\mu(x) = \int_S f^*(x) \, d\mu(x).$$

□

Proposition 5.5.2. *Suppose $\mu \in \mathcal{P}(S)$ is such that $\mu(\Gamma^P_{cpi}) = 1$. Then $\mu \in \mathcal{P}^P_{cpi}$ and $\epsilon_\mu = \int_{\Gamma^P_{cpi}} \epsilon_x \, d\mu(x)$.*

Proof. Since $\mu(\Gamma^P_{cp}) \geq \mu(\Gamma^P_{cpi}) = 1$, Proposition 5.5.1 implies that $\mu \in \mathcal{P}^P_{cp}$ and

$$\epsilon_\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x) = \int_{\Gamma^P_{cpi}} \epsilon_x \, d\mu(x).$$

According to Proposition 3.3.1,

$$P\epsilon_\mu = \int_{\Gamma^P_{cpi}} P\epsilon_x \, d\mu(x) = \int_{\Gamma^P_{cpi}} \epsilon_x \, d\mu(x) = \epsilon_\mu.$$

So $\mu \in \mathcal{P}^P_{cpi}$. □

We now state some results on convergence of Cesàro averages of (possibly signed) finite Borel measures.

Corollary 5.5.3. *Let μ be a finite Borel measure on S such that $|\mu|(S\backslash\Gamma^P_{cp}) = 0$. Then there is a finite Borel measure μ^* such that the following statements holds:*

(i) $\|P^{(n)}\mu - \mu^*\|^*_{\mathrm{BL}} \to 0$ *as* $n \to \infty$.

(ii) $\langle \mu^*, f\rangle = \langle \mu, f^*\rangle$ *for every* $f \in C_b(S)$

(iii) If $|\mu|(S\backslash\Gamma^P_{cpi}) = 0$*, then* μ^* *is invariant.*

Proof. These results follow by writing $\mu = \mu^+ - \mu^-$ and applying Proposition 5.5.1 and Proposition 5.5.2 to scaled versions of μ^+ and μ^-. □

The following proposition gives a condition for stronger convergence of the Cesàro average of a measure. This generalises [56, Theorem 3.1(g)] from the locally compact setting to the Polish setting.

Proposition 5.5.4. *Let* ν *be an invariant probability measure and* $\mu \in \mathcal{M}(S)$ *such that* $\mu \ll \nu$*. Then there is an invariant probability measure* μ^* *such that* $\|P^{(n)}\mu - \mu^*\|_{\mathrm{TV}} \to 0$ *and* $\langle \mu^*, f\rangle = \langle \mu, f^*\rangle$ *for every* $f \in C_b(S)$*.*

Proof. By Proposition 5.3.10, $\nu(S\backslash\Gamma^P_{cpi}) = 0$, thus $|\mu|(S\backslash\Gamma^P_{cpi}) = 0$ as well. So Corollary 5.5.3 implies that there is a finite invariant Borel measure μ^* such that $\|P^{(n)}\mu - \mu^*\|^*_{\mathrm{BL}} \to 0$ and $\langle \mu^*, f\rangle = \langle \mu, f^*\rangle$ for every $f \in C_b(S)$.

From Lemma 4.3.2 we obtain that if $\mu_1, \mu_2 \in \mathcal{M}^+(S)$ are such that $\mu_1 \ll \mu_2$, then $P\mu_1 \ll P\mu_2$. Since $\mu \ll \nu$ and ν is invariant, we obtain $P^{(n)}\mu \ll \nu$ for every $n \in \mathbb{N}$. Now let j_ν be the isometric embedding from $L^1(\nu)$ into $(\mathcal{M}(S), \|\cdot\|_{\mathrm{TV}})$, where $j_\nu(f) = f\,d\nu$ for every $f \in L^1(\nu)$. Since ν is invariant, P induces a positive linear operator $T : L^1(\nu) \to L^1(\nu)$ such that $j_\nu(Tf) = Pj_\nu(f)$ for every $f \in L^1(\nu)$. By the Mean Ergodic Theorem (see e.g. [43, Theorem VII.A]) there exists, for every $f \in L^1(\nu)$, an $\hat{f} \in L^1(\nu)$ such that $\|1/n \sum_{k=0}^{n-1} T^k f - \hat{f}\|_1 \to 0$.

Let $f = \frac{d\mu}{d\nu} \in L^1(\nu)$. Note that $P^{(n)}\mu \in j_\nu(L^1(\nu))$ for every $n \in \mathbb{N}$, thus

$$\|P^{(n)}\mu - j_\nu(\hat{f})\|^*_{\mathrm{BL}} \leq \|P^{(n)}\mu - j_\nu(\hat{f})\|_{\mathrm{TV}} = \|T^{(n)}f - \hat{f}\|_1 \to 0.$$

so $j_\nu(\hat{f}) = \mu^*$. □

The previous results might suggest that $\mu \in \mathcal{P}^P_{cpie}$ whenever $\mu(\Gamma^P_{cpie}) = 1$. However, this is generally not true: $\mu(\Gamma^P_{cpie}) = 1$ for every invariant probability measure μ according to Theorem 5.3.13, and $\epsilon_\mu = \mu$ for these measures, while they obviously need not be ergodic. The following result does hold, however:

Proposition 5.5.5. *Let* $\mu \in \mathcal{P}(S)$*. If* $\mu([z]) = 1$ *for some* $z \in \Gamma^P_{cpie}$*, then* $\mu \in \mathcal{P}^P_{cpie}$ *and* $\epsilon_\mu = \epsilon_z$*.*

Proof. Suppose that $\mu([z]) = 1$ for some $z \in \Gamma^P_{cpie}$, then $\mu(\Gamma^P_{cpi}) = 1$, so $\mu \in \mathcal{P}^P_{cpi}$ by Proposition 5.5.2, and $\epsilon_\mu = \int_{\Gamma^P_{cpi}} \epsilon_x\,d\mu(x)$. Since $\mu([z]) = 1$ we have

$$\epsilon_\mu = \int_{\Gamma^P_{cpi}} \epsilon_x\,d\mu(x) = \int_{[z]} \epsilon_x\,d\mu(x) = \int_{[z]} \epsilon_z\,d\mu(x) = \epsilon_z,$$

so $\epsilon_\mu = \epsilon_z$ is ergodic, thus $\mu \in \mathcal{P}^P_{cpie}$. □

Recall the notation from Theorem 5.4.14. Consider the set $G := \bigcup_{x\in\Gamma^P_{cpie}} S_{[x]}$. We will show that this set is an invariant Borel set and that every invariant probability measure is concentrated on this set. Then Theorem 5.4.14 implies that we can decompose S into a Borel set $S\backslash G$, which is a μ-null set for every invariant probability measure μ, and a collection of invariant Borel sets such that every ergodic measure is concentrated on exactly one of these sets.

Proposition 5.5.6. *The set $G = \bigcup_{x\in\Gamma^P_{cpie}} S_{[x]}$ is an invariant Borel set and $\mu(G) = 1$ for every invariant probability measure μ.*

Proof. Let $C_0 := \Gamma^P_{cpie}$, $C_n := \{x \in C_{n-1} : P\delta_x(C_{n-1}) = 1\}$ and $C = \cap_{n\in\mathbb{N}} C_n$. It follows from the proof of Lemma 5.4.2 and Corollary 5.4.3 that C_n is measurable for all $n \in \mathbb{N}$, that C is an invariant Borel set and that, for all invariant probability measures μ, $\mu(C) = 1$, since $\mu(\Gamma^P_{cpie}) = 1$. We will show that $G = C$. Note that for all $x \in \Gamma^P_{cpie}$, $S_{[x]} = \cap_{n\in\mathbb{N}} B^{[x]}_n$, where $B^{[x]}_0 := [x]$ and

$$B^{[x]}_n := \{y \in B^{[x]}_{n-1} : P\delta_y(B^{[x]}_{n-1}) = 1\}.$$

Clearly $C_0 = \cup_{x\in\Gamma^P_{cpie}} B^{[x]}_0$. Now assume $C_{n-1} = \cup_{x\in\Gamma^P_{cpie}} B^{[x]}_{n-1}$ for some $n \in \mathbb{N}$. We will show that $C_n = \cup_{x\in\Gamma^P_{cpie}} B^{[x]}_n$.

Let $y \in B^{[x]}_n$ for some $x \in \Gamma^P_{cpie}$, then $y \in C_{n-1}$ and

$$P\delta_y(C_{n-1}) \geq P\delta_y(B^{[x]}_{n-1}) = 1,$$

thus $y \in C_n$.

Now let $y \in C_n$. Then $y \in C_{n-1}$, so there is an $x \in \Gamma^P_{cpie}$ such that $y \in B^{[x]}_{n-1} \subset [x]$, so $[y] = [x]$. Suppose that $y \notin B^{[x]}_n$. Then $P\delta_y(B^{[x]}_{n-1}) < 1$. We define the positive measures

$$\nu_1 := P\delta_y(\cdot \cap (S\backslash B^{[x]}_{n-1})) \text{ and } \nu_2 := P\delta_y(\cdot \cap B^{[x]}_{n-1}),$$

then $P\delta_y = \nu_1 + \nu_2$ and $\nu_1 \neq 0$. Now,

$$\lim_{k\to\infty} P^{(k)} P\delta_y = \lim_{k\to\infty} P^{(k)}\delta_y = \epsilon_y = \epsilon_x.$$

Let $\mu_1 = \frac{\nu_1}{\nu_1(S)}$. If $\nu_2 = 0$, then $\mu_1 = P\delta_y$, so $\lim_{k\to\infty} P^{(k)}\mu_1 = \epsilon_x$. If $\nu_2 \neq 0$, define $\mu_2 := \frac{\nu_2}{\nu_2(S)} \in \mathcal{P}(S)$. Then $\mu_2([x]) \geq \mu_2(B^{[x]}_{n-1}) = 1$, thus $\lim_{k\to\infty} P^{(k)}\mu_2 = \epsilon_x$ by Proposition 5.5.5. Thus $\lim_{k\to\infty} P^{(k)}\mu_1 = \epsilon_x$ as well. By assumption, $P\delta_y(C_{n-1}) = 1$, thus $\mu_1(C_{n-1}) = 1$, which implies that $\mu_1(C_{n-1}\backslash B^{[x]}_{n-1}) = 1$. Then

$$P^{(k)}\mu_1 = \int_S P^{(k)}\delta_z \, d\mu_1(z) = \int_{C_{n-1}\backslash B^{[x]}_{n-1}} P^{(k)}\delta_z \, d\mu_1(z).$$

By the Dominated Convergence Theorem, taking $k \to \infty$:

$$\epsilon_x = \int_{C_{n-1} \setminus B_{n-1}^{[x]}} \epsilon_z \, d\mu_1(z),$$

so

$$1 = \epsilon_x(B_{n-1}^{[x]}) = \int_{C_{n-1} \setminus B_{n-1}^{[x]}} \epsilon_z(B_{n-1}^{[x]}) \, d\mu_1(z) = 0,$$

since $(C_{n-1} \setminus B_{n-1}^{[x]}) \cap [x] = \emptyset$. This gives a contradiction, so $y \in B_n^{[x]}$.

Now we can conclude that $C = G$. □

The conditions in Proposition 5.5.1, Proposition 5.5.2 and Proposition 5.5.5 are not *necessary* in general for a measure μ to be in $\mathcal{P}_{cp}^P$, $\mathcal{P}_{cpi}^P$ and $\mathcal{P}_{cpie}^P$ respectively. In the following example we construct a regular Markov operator on a *compact* Polish space S, such that there exists a $\mu \in \mathcal{P}(S)$ for which $P^{(n)}\mu$ converges to an ergodic measure, but $\mu(\Gamma_{cp}^P) = 0$.

Example 5.5.7. Let $S = [0, 1]$ with Euclidean metric d. Let $(a_n)_n$ be a sequence in $\{0, 1\}$ such that $1/n \sum_{k=1}^n a_k$ does not converge. Define $b_n := a_n + (-1)^{a_n}/2^{n+1}$ and $c_n = 1 - b_n$. Then $b_n, c_n \in (0, 1)$.

Define $\Phi : S \to S$ as follows:

$$\Phi(x) := \begin{cases} b_{n+1} & \text{if } x = b_n \\ c_{n+1} & \text{if } x = c_n \\ 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ x & \text{else} \end{cases}$$

It is straightforward to prove that Φ is a well-defined, Borel measurable map. Let P_Φ be the regular Markov operator associated to Φ, i.e. $P_\Phi \mu(E) := \mu(\Phi^{-1}(E))$ for every $\mu \in \mathcal{M}^+(S)$ and $E \subset S$ Borel. Then $P_\Phi \delta_{b_n} = \delta_{b_{n+1}}$ and $P_\Phi \delta_{c_n} = \delta_{c_{n+1}}$ for every $n \in \mathbb{N}$.

We will show that $b_1 \notin \Gamma_{cp}^P$. To that end, let $f : S \to \mathbb{R} : x \mapsto x$. Then

$$\langle P^{(n)} \delta_{b_1}, f \rangle = \tfrac{1}{n} \sum_{k=1}^n \langle \delta_{b_k}, f \rangle = \tfrac{1}{n} \sum_{k=1}^n b_k.$$

Now $|\frac{1}{n} \sum_{k=1}^n b_k - 1/n \sum_{k=1}^n a_k| \leq \frac{1}{n} \sum_{k=1}^n 2^{k+1} \to 0$, so $\frac{1}{n} \sum_{k=1}^n b_k$ does not converge, thus $b_1 \notin \Gamma_{cp}^P$. By similar reasonings, $c_1 \notin \Gamma_{cp}^P$.

Let $\mu = \frac{1}{2}\delta_{b_1} + \frac{1}{2}\delta_{c_1}$. Then

$$P^{(n)}\mu = \tfrac{1}{2n} \sum_{k=1}^n \delta_{\frac{1}{2^{k+1}}} + \tfrac{1}{2n} \sum_{k=1}^n \delta_{1 - \frac{1}{2^{k+1}}} \to \tfrac{1}{2}\delta_0 + \tfrac{1}{2}\delta_1$$

in $\mathcal{S}_{\mathrm{BL}}$ as $n \to \infty$. Note that $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is an ergodic measure, thus $\mu \in \mathcal{P}^P_{cpie}$, but $\mu(\Gamma^P_{cp}) = 0$.

5.6 Notes

This chapter is based on the paper [125] with some minor modifications, and the addition of Proposition 5.5.6.

The notion of ergodic decomposition discussed in this chapter goes back to the pioneering works of Krylov and Bogolioubov [75], who considered Markov operators on a compact metric space coming from a continuous invertible map on that space. Their results yield a measurable map from state space to the set of ergodic measures, that can be used to decompose invariant probability measures into ergodic measures via an integral over state space. Moreover, a decomposition of state space into disjoint invariant measurable sets is achieved, such that for every such set there exists exactly one ergodic measure concentrated on this set. Motivated by their paper there have been basically two kinds of generalisations of their results.

The first focuses on extension to Markov operators. Beboutov [9] generalised the results of Krylov and Bogolioubov to arbitrary Markov-Feller operators on compact metric spaces. Yosida [129, 130] extended this to the class of Markov operators whose dual operator maps the space of continuous functions with compact support into itself, defined on separable metric spaces in which the closed and bounded sets are compact. More recently, Hernández-Lerma and Lasserre and Zaharopol covered the more general setting of regular Markov operators on a locally compact separable metric space [56, 57, 132, 134], where Zaharopol showed that the sets and maps in his decomposition do not depend on a particular chosen invariant measure. The extension to the setting of Polish spaces in this chapter answers an open question by Zaharopol.

The second class of generalisations of the results by Krylov and Bogolioubov remains in the setting of deterministic dynamical systems. Varadarajan [120] considered locally compact groups of invertible transformations acting measurably on a quite general class of measurable spaces, which encompass Polish spaces. His work has been further generalised by other authors, dealing with *quasi-invariant measures* instead of invariant measures (see e.g. [50, 103] and references therein). The proofs of these results are of a measure theoretic and probabilistic nature, centering around the existence of certain conditional expectations, proven using (generalised) Cesàro averages.

Note that these results can be applied to a regular Markov operator P on a Polish space S: one can associate to P an invertible measurable transformation on the trajectory space $S^{\mathbb{Z}}$ by shifting trajectories one step, and there is a correspondence between invariant and ergodic measures with respect to P and invariant and ergodic measures on the trajectory space. Using the theory on ergodic decomposition for

invertible measurable transformations, one can obtain an integral decomposition of invariant measures (with respect to P) in terms of ergodic measures (see e.g. [121, Chapter 6]). However, these results do not yield a parametrisation of the ergodic measures in terms of the *original* state space S or a decomposition of S into disjoint invariant measurable sets such that on each of these sets exactly one ergodic measure is concentrated, as we obtained in this chapter.

CHAPTER

SIX

ERGODIC DECOMPOSITIONS ASSOCIATED TO REGULAR JOINTLY MEASURABLE MARKOV SEMIGROUPS ON POLISH SPACES

6.1 Introduction

In this chapter we consider regular jointly measurable Markov semigroups on Polish spaces. There is much interest lately in continuous-time Markov processes (and their associated Markov semigroups) on non-locally compact Polish spaces, for instance those coming from stochastic partial differential equations in separable Hilbert spaces (see e.g. [10, 19, 20, 36, 83]) or in separable Banach spaces [49]. Other recent research on Markov semigroups on Polish spaces is performed by Szarek and coworkers [73, 81, 114].

Our aim is to show that results analogous to those achieved in Chapter 5 for regular Markov operators hold for regular jointly measurable Markov semigroups on Polish spaces as well. We obtain a Yosida-type ergodic decomposition of the state space. Moreover, we obtain a parametrisation of ergodic measures for such an operator in terms of classes of subsets of state space and an integral decomposition of invariant measures into ergodic measures using this parametrisation. This decomposition associates with every ergodic measure a Borel measurable subset of the state space on which it is contained, and is such that the corresponding subsets of distinct ergodic measures are disjoint. Such a subset furthermore contains a Borel set that is invariant under the Markov semigroup, and such that the corresponding ergodic measure is actually concentrated on this smaller subset. We call this the full Yosida-type ergodic decomposition associated to the Markov semigroup.

In view of the Markov operator setting, the main problem here is, of course, to deal with the *uncountable* family of Markov operators $(P(t))_{t\geq 0}$. We show that one can reduce this setting to the operator setting, by considering a single regular Markov operator that is associated to the Markov semigroup instead, the *resolvent operator* R. It turns out that $(P(t))_{t\geq 0}$ and R have the same invariant and ergodic measures. Moreover, the Cesàro averages for the semigroup and the resolvent,

$$\frac{1}{t}\int_0^t P(s)\mu\, ds, \qquad \text{respectively} \qquad \frac{1}{n}\sum_{k=0}^{n-1} R^k\mu,$$

have the same convergence properties (made precise in Theorem 6.2.7 and its corollaries). These results imply that the Yosida-type ergodic decomposition associated to R, from Chapter 5, actually works for the semigroup $(P(t))_{t\geq 0}$ as well. We obtain the full Yosida-type ergodic decomposition in Section 6.3.3, that generalises results by Costa and Dufour [18]. There, the authors consider Markov semigroups on locally compact separable metric spaces, and require more regularity. Our results do not require *continuity* assumptions on the Markov semigroup: the operators $P(t)$ need not be Markov-Feller, only regular, and the orbits $t \mapsto P(t)\mu$ need not be continuous in any sense: jointly measurable is sufficient. These are natural assumptions when considering Markov semigroups associated to Markov processes.

One of the results needed to establish the full Yosida-type ergodic decomposition for Markov semigroups is Theorem 6.3.3, which also yields the conclusion that the usual equivalent notions for ergodicity of an invariant measure μ of a Markov semigroup $(P(t))_{t\geq 0}$ (see Theorem 6.3.2) are equivalent to

$$\mu(E) = 0 \text{ or } 1 \text{ for every Borel set } E \text{ that is } (P(t))_{t\geq 0}\text{-invariant},$$

i.e. for every $t \geq 0$, $P(t)\delta_x(E) = 1$ for all $x \in E$. We could not retrieve this natural analogue of the definition of ergodicity for the operator case (see Section 5.4.1) from the Markov semigroup literature in the generality of a regular jointly measurable Markov semigroup on a Polish space.

In Section 6.2 we introduce the resolvent operator associated to a Markov semigroup. The main theorem of this section (Theorem 6.2.7) shows that convergence properties of the Cesàro averages of a Markov semigroup and its resolvent coincide. In Section 6.3.1 we define ergodicity of invariant measures for Markov operators and Markov semigroups and give several equivalent characterisations. We prove analogues of the ergodic decomposition results from Chapter 5 in the setting of Markov semigroups in Section 6.3.2, and give a full Yosida-type ergodic decomposition in Section 6.3.3.

Let S be a Polish space and d a complete metric on S that metrises its topology. Let $\mathcal{S}_{\mathrm{BL}}$ be the separable Banach space associated to (S, d) (see Chapter 2).

6.2 Resolvent operator of a regular jointly measurable Markov semigroup

Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup with dual semigroup $\mathbf{U} = (U(t))_{t\geq 0}$.

Lemma 6.2.1. *Let $\mu \in \mathcal{M}(S)$ and $f \in L^1(\mathbb{R}_+)$. Then $\int_{\mathbb{R}_+} f(s)P(s)\mu\, ds$ is well defined as Bochner integral in $\mathcal{S}_{\mathrm{BL}}$ and takes value in $\mathcal{M}(S)$. It satisfies:*

(i)
$$\left\| \int_{\mathbb{R}_+} f(s)P(s)\mu\, ds \right\|_{\mathrm{TV}} \leq \|f\|_1 \|\mu\|_{\mathrm{TV}}.$$

(ii) For every $t > 0$,
$$P(t) \int_{\mathbb{R}_+} f(s)P(s)\mu\, ds = \int_{\mathbb{R}_+} f(s)P(s+t)\mu\, ds.$$

(iii) If $f \in L^1_+(\mathbb{R}_+)$ and $\mu \in \mathcal{M}^+(S)$, then $\int_{\mathbb{R}_+} f(s)P(s)\mu\, ds \in \mathcal{M}^+(S)$.

Proof. First note that $s \mapsto f(s)P(s)\mu$ is Bochner integrable from $\mathbb{R}_+$ to $\mathcal{S}_{\mathrm{BL}}$ with respect to Lebesgue measure, and $\|f(s)P(s)\mu\|_{\mathrm{TV}} \leq \|\mu\|_{\mathrm{TV}}|f(s)|$ for every $s \in \mathbb{R}_+$. Thus $\nu := \int_{\mathbb{R}_+} f(s)P(s)\mu\, ds \in \mathcal{M}(S)$ by Proposition 3.2.5 and we can also obtain ν as set-wise integral. Then Proposition 4.4.4 implies that
$$\left\| \int_{\mathbb{R}_+} f(s)P(s)\mu\, ds \right\|_{\mathrm{TV}} \leq \|f\|_1 \|\mu\|_{\mathrm{TV}}.$$

(ii) also follows from Proposition 4.4.4. Finally, if $f \in L^1_+(\mathbb{R}_+)$ and $\mu \in \mathcal{M}^+(S)$, then $f(s)P(s)\mu \in \mathcal{M}^+(S)$ for almost every $s \in \mathbb{R}_+$, so $\int_{\mathbb{R}_+} f(s)P(s)\mu\, ds \in \mathcal{M}^+(S)$ as well. □

We define the *resolvent family* (associated with the Markov semigroup $(P(t))_{t\geq 0}$) to be the collection of operators $\{R_\lambda : \lambda > 0\}$:
$$R_\lambda \mu := \int_{\mathbb{R}_+} e^{-\lambda t} P(t)\mu\, dt \text{ for every } \mu \in \mathcal{M}^+(S).$$

We define the *resolvent operator* (associated with $(P(t))_{t\geq 0}$) to be $R := R_1$.

Proposition 6.2.2. *For every $\lambda > 0$, λR_λ is a regular Markov operator with dual operator λU_λ given by*
$$\lambda U_\lambda f(x) := \lambda \int_{\mathbb{R}_+} e^{-\lambda t} U(t) f(x)\, dt.$$

Proof. First of all, $g : t \mapsto e^{-\lambda t}$ is in $L^1_+(\mathbb{R}_+)$ and $\|g\|_1 = 1/\lambda$. Thus by Lemma 6.2.1 R_λ is a positively homogeneous and additive map from $\mathcal{M}^+(S)$ to $\mathcal{M}^+(S)$ and $\|\lambda R_\lambda \mu\|^*_{\mathrm{BL}} = \|\mu\|^*_{\mathrm{BL}}$ for every $\mu \in \mathcal{M}^+(S)$. Thus λR_λ is a Markov operator. Let $f \in \mathrm{BM}(S)$, then by applying Proposition 3.2.5 we get

$$\begin{aligned}
\langle \lambda R_\lambda \mu, f\rangle &= \lambda \int_{\mathbb{R}_+} e^{-\lambda t} \langle P(t)\mu, f\rangle \, dt \\
&= \lambda \int_{\mathbb{R}_+} e^{-\lambda t} \langle \mu, U(t) f\rangle \, dt \\
&= \lambda \int_{\mathbb{R}_+} e^{-\lambda t} \int_S U(t) f(x) \, d\mu(x) \, dt.
\end{aligned}$$

Since $(P(t))_{t\geq 0}$ is jointly measurable, Proposition 3.2.4 implies that $(t,x) \mapsto U(t)f(x) = \langle P(t)\delta_x, f\rangle$ is measurable from $\mathbb{R}_+ \times S$ to $\mathbb{R}$, so we can apply Fubini's Theorem and obtain

$$\langle \lambda R_\lambda \mu, f\rangle = \int_S \int_{\mathbb{R}_+} \lambda e^{-\lambda t} U(t) f(x) \, dt \, d\mu(x).$$

□

The *resolvent identity* holds for $(R_\lambda)_{\lambda>0}$:

Lemma 6.2.3. *Let $\lambda, \gamma > 0$ and $\mu \in \mathcal{M}^+(S)$. Then*

$$R_\lambda \mu - R_\gamma \mu = (\gamma - \lambda) R_\gamma R_\lambda \mu.$$

Proof. Note that for every $t > 0$

$$e^{-\lambda t} - e^{-\gamma t} = (\gamma - \lambda) e^{-\lambda t} \int_0^t e^{-(\gamma-\lambda)s} \, ds.$$

Now by using this equality and Fubini's Theorem, we obtain

$$\begin{aligned}
R_\lambda \mu - R_\gamma \mu &= \int_{\mathbb{R}_+} (e^{-\lambda t} - e^{-\gamma t}) P(t)\mu \, dt \\
&= \int_{\mathbb{R}_+} P(t)\mu (\gamma - \lambda) e^{-\lambda t} \int_0^t e^{-(\gamma-\lambda)s} \, ds \, dt \\
&= (\gamma - \lambda) \int_{\mathbb{R}_+} \int_0^t e^{-\lambda t} e^{-(\gamma-\lambda)s} P(t)\mu \, ds \, dt \\
&= (\gamma - \lambda) \int_{\mathbb{R}_+} e^{-\gamma s} \int_s^\infty e^{-\lambda(t-s)} P(t) \, \mu \, dt \, ds \\
&= (\gamma - \lambda) \int_{\mathbb{R}_+} e^{-\gamma s} \int_0^\infty e^{-\lambda u} P(u+s) \, \mu \, du \, ds.
\end{aligned}$$

By Lemma 6.2.1 the last expression equals

$$(\gamma - \lambda) \int_{\mathbb{R}_+} e^{-\gamma s} P(s) \int_{\mathbb{R}_+} e^{-\lambda u} P(u)\, \mu\, du\, ds = (\gamma - \lambda) R_\gamma R_\lambda \mu.$$

□

Lemma 6.2.4. *Let X be a set and $(\Phi_t)_{t\geq 0}$ a semigroup of maps $\Phi_t : X \to X$. Let $x^* \in X$ be such that $\Phi_t(x^*) = x^*$ for Lebesgue almost all $t \in \mathbb{R}_+$. Then $\Phi_t(x^*) = x^*$ for all $t \in \mathbb{R}_+$.*

Proof. There exists a Lebesgue null set $N \subset \mathbb{R}_+$ such that $\Phi_t(x^*) = x^*$ for every $t \in \mathbb{R}_+ \backslash N$. Let $t \in N$. Then there must be an $s \in \mathbb{R}_+ \backslash N$ such that $t + s \in \mathbb{R}_+ \backslash N$, otherwise $t + \mathbb{R}_+ \backslash N$ is contained in N and by translation invariance of Lebesgue measure m

$$m(N) \geq m(t + \mathbb{R}_+ \backslash N) = m(\mathbb{R}_+ \backslash N) = \infty,$$

a contradiction. Thus, with t and s as above,

$$\Phi_t(x^*) = \Phi_t(\Phi_s(x^*)) = \Phi_{t+s}(x^*) = x^*.$$

□

Proposition 6.2.5. *Let $\mu \in \mathcal{M}^+(S)$. Then the following are equivalent:*

(i) *μ is $(P(t))_{t\geq 0}$-invariant.*

(ii) *$\lambda R_\lambda \mu = \mu$ for some $\lambda > 0$.*

(iii) *$\lambda R_\lambda \mu = \mu$ for every $\lambda > 0$.*

Proof. *(i)⇒(ii)*: This is immediate.
(ii)⇒(iii): Suppose that $\lambda > 0$ is such that $\lambda R_\lambda \mu = \mu$. Let $\gamma > 0$, then by the resolvent identity, Lemma 6.2.3,

$$\frac{1}{\lambda}\mu - R_\gamma \mu = \frac{\gamma - \lambda}{\lambda} R_\gamma \mu.$$

So $\gamma R_\gamma \mu = \mu$.
(iii)⇒(i): By assumption we have for every $\lambda > 0$:

$$\int_{\mathbb{R}_+} e^{-\lambda t} P(t)\mu\, dt = R_\lambda \mu = \frac{1}{\lambda}\mu = \int_{\mathbb{R}_+} e^{-\lambda t} \mu\, dt.$$

Hence by [4, Theorem 1.7.3] $P(t)\mu = \mu$ for Lebesgue almost every $t \in \mathbb{R}_+$. By Lemma 6.2.4 $P(t)\mu = \mu$ for every $t \in \mathbb{R}_+$. □

Lemma 6.2.6. *For every $k \in \mathbb{N}, \lambda > 0$ and $\mu \in \mathcal{M}^+(S)$*

$$R_\lambda^k \mu = \int_0^\infty \frac{t^{k-1}}{(k-1)!} e^{-\lambda t} P(t)\mu\, dt,$$

as Bochner integral in $\mathcal{S}_{\mathrm{BL}}$.

Proof. Let $\lambda > 0$ and $\mu \in \mathcal{M}^+(S)$. We prove the statement by induction. Clearly it holds for $k = 1$. Suppose it holds for some $k \in \mathbb{N}$. By Lemma 6.2.1 we can bring $P(t)$ in and outside of integrals. Using this result and Fubini's Theorem (which we can apply because we are considering jointly measurable Markov semigroups) we obtain

$$\begin{aligned}
R_\lambda^{k+1}\mu &= \int_0^\infty e^{-\lambda t} P(t) \int_0^\infty \frac{s^{k-1}e^{-\lambda s}}{(k-1)!} P(s)\mu \, ds \, dt \\
&= \int_0^\infty \frac{s^{k-1}}{(k-1)!} \int_0^\infty e^{-\lambda(t+s)} P(t+s)\mu \, dt \, ds \\
&= \int_0^\infty \frac{s^{k-1}}{(k-1)!} \int_s^\infty e^{-\lambda\sigma} P(\sigma)\mu \, d\sigma \, ds \\
&= \int_0^\infty \left(\int_0^\sigma \frac{s^{k-1}}{(k-1)!} \, ds \right) e^{-\lambda\sigma} P(\sigma)\mu \, d\sigma \\
&= \int_0^\infty \frac{\sigma^k}{k!} e^{-\lambda\sigma} P(\sigma)\mu \, d\sigma.
\end{aligned}$$

So the statement holds for $k + 1$. □

For a regular jointly measurable Markov semigroup $\boldsymbol{P} = (P(t))_{t\geq 0}$, recall the notation $\mathbf{P}^{(t)}$ for Cesàro average of the Markov semigroup from Section 3.4. We write $\mathcal{M}(S)_{\mathrm{TV}}$ for the Banach space consisting of $\mathcal{M}(S)$ endowed with the total variation norm. We are now ready to state and prove the main result in this section:

Theorem 6.2.7. *Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup. Then*

$$\|R^{(n)} - \mathbf{P}^{(n)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = \mathcal{O}(1/\sqrt{n}) \text{ as } n \to \infty.$$

In particular, for every $\mu \in \mathcal{M}(S)$,

$$\lim_{n\to\infty} \|R^{(n)}\mu - \mathbf{P}^{(n)}\mu\|_{\mathrm{BL}}^* = \lim_{n\to\infty} \|R^{(n)}\mu - \mathbf{P}^{(n)}\mu\|_{\mathrm{TV}} = 0.$$

Proof. Note that

$$\|R^{(n)} - \frac{1}{n}\sum_{k=1}^n R^k\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = \frac{1}{n}\|R^n - I\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \leq \frac{2}{n}.$$

So the theorem is proven once we have shown that

$$\|\mathbf{P}^{(n)} - \frac{1}{n}\sum_{k=1}^n R^k\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = \mathcal{O}(1/\sqrt{n}) \text{ as } n \to \infty.$$

Note that in the following we want to estimate $\|\cdot\|_{\mathrm{TV}}$-norms of Bochner integrals in $\mathcal{S}_{\mathrm{BL}}$ that take values in $\mathcal{M}(S)$, which is possible by Lemma 6.2.1. Let $\mu \in \mathcal{P}(S)$.

By Lemma 6.2.6 we obtain

$$\begin{aligned}\left\|\mathbf{P}^{(n)}\mu - \frac{1}{n}\sum_{k=1}^{n} R^k\mu\right\|_{\mathrm{TV}} &= \frac{1}{n}\left\|\int_0^n P(t)\mu\,dt - \sum_{k=0}^{n-1}\int_0^\infty e^{-t}\frac{t^k}{k!}P(t)\mu\,dt\right\|_{\mathrm{TV}} \\ &\leq \frac{1}{n}\sum_{k=0}^{n-1}\left\|\int_n^\infty e^{-t}\frac{t^k}{k!}P(t)\mu\,dt\right\|_{\mathrm{TV}} \\ &\quad + \left\|\frac{1}{n}\int_0^n\left(1 - e^{-t}\sum_{k=0}^{n-1}\frac{t^k}{k!}\right)P(t)\mu\,dt\right\|_{\mathrm{TV}}.\end{aligned}$$

Since $e^{-t}\sum_{k=0}^{n-1}\frac{t^k}{k!} < e^{-t}e^t = 1$, we obtain by Lemma 6.2.1 and *(MO2)*

$$\left\|\mathbf{P}^{(n)}\mu - \frac{1}{n}\sum_{k=1}^{n} R^k\mu\right\|_{\mathrm{TV}} \leq \left(\frac{1}{n}\sum_{k=0}^{n-1}\int_n^\infty e^{-t}\frac{t^k}{k!}\,dt + \frac{1}{n}\int_0^n 1 - e^{-t}\sum_{k=0}^{n-1}\frac{t^k}{k!}\,dt\right).$$

Now,

$$\begin{aligned}\frac{1}{n}\int_0^n 1 - e^{-t}\sum_{k=0}^{n-1}\frac{t^k}{k!}\,dt &= \frac{1}{n}\sum_{k=0}^{n-1}\left[1 - \int_0^n e^{-t}\frac{t^k}{k!}\,dt\right] \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int_n^\infty e^{-t}\frac{t^k}{k!}\,dt\end{aligned}$$

since $\int_0^\infty e^{-t}\frac{t^k}{k!}\,dt = 1$ for every $k \in \mathbb{N}$, thus

$$\left\|\mathbf{P}^{(n)}\mu - \sum_{k=1}^{n} R^k\mu\right\|_{\mathrm{TV}} \leq \frac{2}{n}\sum_{k=0}^{n-1}\int_n^\infty e^{-t}\frac{t^k}{k!}\,dt$$

for every $n \in \mathbb{N}$.

By elementary calculation we obtain for every $k, n \in \mathbb{N}$,

$$\int_n^\infty e^{-t}\frac{t^k}{k!}\,dt = e^{-n}\sum_{i=0}^{k}\frac{n^i}{i!},$$

thus

$$\begin{aligned}\sum_{k=0}^{n-1}\int_n^\infty e^{-t}\frac{t^k}{k!}\,dt &= e^{-n}\sum_{k=0}^{n-1}\sum_{i=0}^{k}\frac{n^i}{i!} = e^{-n}\sum_{i=0}^{n-1}\frac{n^i(n-i)}{i!} \\ &= e^{-n}\left(\sum_{i=1}^{n-1}\left(\frac{n^{i+1}}{i!} - \frac{n^i}{(i-1)!}\right) + n\right) = e^{-n}\frac{n^n}{(n-1)!}.\end{aligned}$$

So

$$\left\| \mathbf{P}^{(n)}\mu - \frac{1}{n}\sum_{k=1}^{n} R^k \mu \right\|_{\mathrm{TV}} \leq \frac{2n^n}{e^n n!} = \frac{\sqrt{2\pi n} n^n}{e^n n!} \cdot \frac{2}{\sqrt{2\pi n}}.$$

This holds for every $\mu \in \mathcal{P}(S)$. Since $\mathcal{M}(S)_{\mathrm{TV}}$ is an *L-space* (see Section 4.2), the norm of a bounded linear operator T on $\mathcal{M}(S)_{\mathrm{TV}}$ equals $\sup\{\|T\mu\|_{\mathrm{TV}} : \mu \in \mathcal{P}(S)\}$. So

$$\left\| \mathbf{P}^{(n)} - \frac{1}{n}\sum_{k=1}^{n} R^k \right\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \leq \frac{\sqrt{2\pi n} n^n}{e^n n!} \cdot \frac{2}{\sqrt{2\pi n}}.$$

By Stirling's Formula, $\lim_{n\to\infty} \frac{\sqrt{2\pi n} n^n}{e^n n!} = 1$. Thus

$$\left\| \mathbf{P}^{(n)} - \frac{1}{n}\sum_{k=1}^{n} R^k \right\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = \mathcal{O}(1/\sqrt{n}) \text{ as } n \to \infty.$$

The final statement follows from the continuity of the embedding $\mathcal{M}(S)_{\mathrm{TV}} \hookrightarrow \mathcal{S}_{\mathrm{BL}}$. □

We now state some further convergence properties of $(P(t))_{t\geq 0}$ that will be useful later on:

Lemma 6.2.8. *Let $s, t \in \mathbb{R}_{>0}$. Then*

$$\|\mathbf{P}^{(s)} - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \leq \frac{2|t-s|}{\max(s,t)}.$$

Proof. Let $\mu \in \mathcal{P}(S)$. Suppose $0 < s < t$. Then by Lemma 6.2.1

$$\begin{aligned} \|\mathbf{P}^{(s)}\mu - \mathbf{P}^{(t)}\mu\|_{\mathrm{TV}} &\leq \left\| (1/s - 1/t) \int_0^s P(r)\mu \, dr \right\|_{\mathrm{TV}} + \left\| \frac{1}{t} \int_s^t P(r)\mu \, dr \right\|_{\mathrm{TV}} \\ &\leq \frac{2|t-s|}{t} \end{aligned}$$

So

$$\|\mathbf{P}^{(s)} - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \leq \frac{2|t-s|}{\max(s,t)}.$$

□

This implies the following results:

Corollary 6.2.9. *The map $t \mapsto \mathbf{P}^{(t)} : (0,\infty) \to \mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})$ is continuous. In particular, $t \mapsto \mathbf{P}^{(t)}\mu$ is continuous as map from $(0,\infty)$ into $\mathcal{M}(S)_{\mathrm{TV}}$ and into $\mathcal{S}_{\mathrm{BL}}$, for every $\mu \in \mathcal{M}(S)$.*

Corollary 6.2.10. *Let $s_n, t_n \in \mathbb{R}_+$, such that $M := \sup_n |t_n - s_n| < \infty$ and $\lim_{n\to\infty} s_n = \lim_{n\to\infty} t_n = \infty$. Then*

$$\lim_{n\to\infty} \|\mathbf{P}^{(s_n)} - \mathbf{P}^{(t_n)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = 0.$$

Thus also

$$\lim_{n\to\infty} \|\mathbf{P}^{(s_n)}\mu - \mathbf{P}^{(t_n)}\mu\|^*_{\mathrm{BL}} = \|\mathbf{P}^{(s_n)}\mu - \mathbf{P}^{(t_n)}\mu\|_{\mathrm{TV}} = 0$$

for every $\mu \in \mathcal{M}^+(S)$.

Theorem 6.2.7 and Corollary 6.2.10 imply the following:

Corollary 6.2.11. *Let $0 < t_n \uparrow \infty$. Then for every $\mu \in \mathcal{M}(S)$,*

$$\lim_{n\to\infty} \|\mathbf{P}^{(t_n)}\mu - R^{(\lfloor t_n \rfloor)}\mu\|^*_{\mathrm{BL}} = \lim_{n\to\infty} \|\mathbf{P}^{(t_n)}\mu - R^{(\lfloor t_n \rfloor)}\mu\|_{\mathrm{TV}} = 0.$$

Lemma 6.2.12. *Let $s \in \mathbb{R}_+$ and $t > 0$. Then*

$$\|\mathbf{P}^{(t)}P(s) - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \le \frac{2s}{t}.$$

In particular

$$\lim_{t\to\infty} \|\mathbf{P}^{(t)}P(s) - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = 0$$

for every $s \in \mathbb{R}_+$.

Proof. Let $\mu \in \mathcal{P}(S)$. Then

$$\begin{aligned}
\|\mathbf{P}^{(t)}P(s)\mu - \mathbf{P}^{(t)}\mu\|_{\mathrm{TV}} &= \frac{1}{t}\left\|\int_s^{s+t} P(r)\mu\,dr - \int_0^t P(r)\mu\,dr\right\|_{\mathrm{TV}} \\
&= \frac{1}{t}\left\|\int_t^{s+t} P(r)\mu\,dr - \int_0^s P(r)\mu\,dr\right\|_{\mathrm{TV}} \\
&\le \frac{1}{t}\left\|\int_t^{s+t} P(r)\mu\,dr\right\|_{\mathrm{TV}} + \frac{1}{t}\left\|\int_0^s P(r)\mu\,dr\right\|_{\mathrm{TV}}.
\end{aligned}$$

So by Lemma 6.2.1, $\|\mathbf{P}^{(t)}P(s)\mu - \mathbf{P}^{(t)}\mu\|_{\mathrm{TV}} \le \frac{2s}{t}$. Thus

$$\|\mathbf{P}^{(t)}P(s) - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \le \frac{2s}{t}.$$

□

The following result can be found in [73, Lemma 2]:

Lemma 6.2.13. *Let $n \in \mathbb{N}$ and $t_1, t_2, \ldots, t_n \in \mathbb{R}_+$. Then*

$$\lim_{t\to\infty} \|\mathbf{P}^{(t)}\mathbf{P}^{(t_1)}\ldots\mathbf{P}^{(t_n)} - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} = 0.$$

Proof. We prove the claim for $n = 1$. The general statement then follows by induction. Let $t_1 \in \mathbb{R}_+$ and $\mu \in \mathcal{P}(S)$. By Fubini's Theorem,

$$\mathbf{P}^{(t)}\mathbf{P}^{(t_1)}\mu - \mathbf{P}^{(t)}\mu = \frac{1}{t_1}\int_0^{t_1} P(s)\mathbf{P}^{(t)}\mu - \mathbf{P}^{(t)}\mu\, ds,$$

so by Lemma 6.2.1 and Lemma 6.2.12

$$\begin{aligned} \|\mathbf{P}^{(t)}\mathbf{P}^{(t_1)}\mu - \mathbf{P}^{(t)}\mu\|_{\mathrm{TV}} &\leq \frac{1}{t_1}\int_0^{t_1} \|P(s)\mathbf{P}^{(t)}\mu - \mathbf{P}^{(t)}\mu\|_{\mathrm{TV}}\, ds \\ &\leq \frac{1}{t_1}\int_0^{t_1} \frac{2s}{t}\, ds \\ &\leq \frac{t_1}{t}. \end{aligned}$$

Thus

$$\|\mathbf{P}^{(t)}\mathbf{P}^{(t_1)} - \mathbf{P}^{(t)}\|_{\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})} \leq \frac{t_1}{t} \to 0 \text{ as } t \to \infty.$$

□

6.3 The ergodic decomposition

Let S be a Polish space. Let d be a complete metric on S metrising the given topology and $\mathcal{S}_{\mathrm{BL}}$ the Banach space associated with (S, d).

6.3.1 Ergodic measures

Let P be a regular Markov operator with dual U. Recall that a Borel set E is a P-invariant set if $P\delta_x(E) = 1$ for every $x \in E$. If μ is an invariant probability measure, we call E μ-almost P-invariant whenever $U\mathbb{1}_E \geq \mathbb{1}_E$ μ-a.e.

Let $\mathcal{P}^P_{\mathrm{inv}}(S) \subset \mathcal{P}(S)$ denote the convex subset of invariant probability measures. There are several different, but equivalent, definitions for an invariant measure to be ergodic.

Theorem 6.3.1. *Let P be a regular Markov operator and μ be a P-invariant probability measure. Then the following are equivalent:*

(i) μ is ergodic, i.e. $\mu(E) = 0$ or 1 for every P-invariant set E.

(ii) $\mu(E) = 0$ or 1 for every μ-almost P-invariant set E.

(iii) There exists a Borel subset B of S such that $\mu(B) = 1$ and such that $U^{(n)}f(x)$ converges to $\langle \mu, f\rangle$ as $n \to \infty$ for every $x \in B$ and $f \in C_b(S)$.

(iv) For every $f \in \mathrm{BM}(S)$, $U^{(n)}f(x)$ converges to $\langle \mu, f\rangle$ for μ-a.e. $x \in S$.

(v) μ *is an extreme point of* $\mathcal{P}^P_{\text{inv}}(S)$.

Proof. *(i)*$\Leftrightarrow$*(ii)*: This follows from Lemma 5.4.2.
(i)$\Leftrightarrow$*(iii)*: This is shown in Theorem 5.4.5.
(ii)$\Rightarrow$*(iv)*: This statement is proven in [57, Proposition 2.4.2].
(iv)$\Rightarrow$*(iii)*: This can be shown with similar arguments as in the proof of Theorem 5.4.5, by using a countable convergence-determining subset of $C_b(S)$.
(ii)$\Leftrightarrow$*(v)*: This is shown in [2, Theorem 19.25]. □

Remark. Most of the equivalences in the theorem above are known and hold for more general state spaces. The proof of the equivalence of *(iii)* to ergodicity is based on Theorem 5.4.5 and needs the state space to be Polish.

Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup on S with resolvent R and dual semigroup $\mathbf{U} = (U(t))_{t\geq 0}$. Ergodicity of an invariant measure for a Markov semigroup is generally defined in the literature through ergodicity of an associated dynamical system. Following [20], we associate with $(P(t))_{t\geq 0}$ a dynamical system on the space of trajectories $\Omega = S^{\mathbb{R}}$, with σ-algebra $\mathcal{F} = \mathcal{B}^{\mathbb{R}}$ and a group of invertible, measurable transformations $(\theta_t)_{t\in\mathbb{R}}$ from Ω to Ω: $(\theta_t\omega)(s) := \omega(t+s)$, $t, s \in \mathbb{R}$. With a $(P(t))_{t\geq 0}$-invariant probability measure μ we associate, as explained in [20, Section 2.1], a probability measure $\mathbb{P}^\mu$ on $(\Omega, \mathcal{F})$, such that $\mathbb{P}^\mu(\theta_t E) = \mathbb{P}^\mu(E)$ for every $t \in \mathbb{R}$ and $E \in \mathcal{F}$. $\mathcal{S}^\mu = (\Omega, \mathcal{F}, \mathbb{P}^\mu, \theta_t)$ is called the *canonical dynamical system* associated to $(P(t))_{t\geq 0}$ and μ. As in [20, Section 2.3] a $(P(t))_{t\geq 0}$-invariant probability measure μ is $(P(t))_{t\geq 0}$*-ergodic* if the dynamical system $\mathcal{S}^\mu$ is ergodic (see [20, Chapter 1] and [97, Chapter 2] for details on ergodic dynamical systems).

It follows from [20, Proposition 3.2.7] that μ is $(P(t))_{t\geq 0}$-ergodic if and only if μ is an extreme point of the set of invariant probability measures $\mathcal{P}^{\mathbf{P}}_{\text{inv}}(S)$. In [20] the Markov semigroup is assumed to be regular and strongly stochastically continuous at zero. However, jointly measurability and regularity of the Markov semigroup suffices for the proofs of the relevant results in [20]. Working with ergodic measures either through the canonical dynamical system or the equivalent characterisation as extreme points in $\mathcal{P}^{\mathbf{P}}_{\text{inv}}(S)$ is somewhat inconvenient. Therefore we now want to discuss some equivalent characterisations of ergodicity.

A Borel set $E \subset S$ is μ*-almost* $(P(t))_{t\geq 0}$*-invariant* if it is μ-invariant with respect to $P(t)$ for every $t \in \mathbb{R}_+$. E is called a *Lebesgue-almost* $(P(t))_{t\geq 0}$*-invariant set* if for every $x \in E$, $P(t)\delta_x(E) = 1$ for *almost every* $t \in \mathbb{R}_+$. Observe that E a Lebesgue-almost $(P(t))_{t\geq 0}$-invariant set if and only if E is R-invariant.

We have the following equivalent characterisations for a $(P(t))_{t\geq 0}$-invariant probability measure to be ergodic:

Theorem 6.3.2. *Let* μ *be a* $(P(t))_{t\geq 0}$*-invariant probability measure. Then the following are equivalent:*

(i) $\mu(E) = 0$ *or* 1 *for every Lebesgue-almost* $(P(t))_{t\geq 0}$*-invariant set* E.

(ii) $\mu(E) = 0$ *or* 1 *for every* μ*-almost* $(P(t))_{t\geq 0}$ *invariant set* E*.*

(iii) There exists a Borel subset B *of* S *such that* $\mu(B) = 1$ *and such that* $\mathbf{U}^{(t)}f(x)$ *converges to* $\langle \mu, f\rangle$ *as* $t \to \infty$ *for every* $x \in B$ *and* $f \in C_b(S)$*.*

(iv) For every $f \in \mathrm{BM}(S)$*,* $\boldsymbol{U}^{(t)}f(x)$ *converges to* $\langle \mu, f\rangle$ *for* μ*-a.e.* $x \in S$*.*

(v) μ *is an extreme point of* $\mathcal{P}^{\mathbf{P}}_{\mathrm{inv}}(S)$*.*

(vi) μ *is* R*-ergodic.*

Proof. Let μ be a $(P(t))_{t\geq 0}$-invariant probability measure. By Proposition 6.2.5 μ is an R-invariant probability measure.

The equivalence between *(ii)* and *(v)* follows from [20, Theorem 3.2.4 and Proposition 3.2.7].

A Borel set E is R-invariant if and only if it is Lebesgue-almost $(P(t))_{t\geq 0}$-invariant, so μ satisfies *(i)* if and only if μ satisfies Theorem 6.3.1 *(i)*.

Theorem 6.2.7 implies that for any $f \in \mathrm{BM}(S)$ and any $x \in S$, $\mathbf{U}^{(t)}f(x) = \langle \mathbf{P}^{(t)}\delta_x, f\rangle$ converges as $t \to \infty$ if and only if $V^{(n)}f(x) = \langle R^{(n)}\delta_x, f\rangle$ converges as $n \to \infty$, where V is the dual of R. Thus *(iii)* and *(iv)* are equivalent to Theorem 6.3.1 *(iii)* and *(iv)* respectively. And by Proposition 6.2.5, $\mathcal{P}^{\mathbf{P}}_{\mathrm{inv}}(S) = \mathcal{P}^{R}_{\mathrm{inv}}(S)$ so they have the same extreme points. Thus *(v)* is equivalent to Theorem 6.3.1 *(v)*.

Now the statement follows from Theorem 6.3.1. □

Analogous to the definition of P-invariant sets for a regular Markov operator P, we define a Borel set $E \subset S$ to be $(P(t))_{t\geq 0}$*-invariant* if for every $x \in E$, $P(t)\delta_x(E) = 1$ *for all* $t \in \mathbb{R}_+$. Clearly, every $(P(t))_{t\geq 0}$-invariant set E is Lebesgue-almost $(P(t))_{t\geq 0}$-invariant, so $\mu(E) = 0$ or $\mu(E) = 1$ whenever μ is a $(P(t))_{t\geq 0}$-ergodic measure. However, the converse need not hold: one can easily construct Lebesgue-almost $(P(t))_{t\geq 0}$-invariant sets that are not $(P(t))_{t\geq 0}$-invariant.

A natural question arises: is a $(P(t))_{t\geq 0}$-invariant measure μ ergodic whenever $\mu(E) = 0$ or 1 for every $(P(t))_{t\geq 0}$-invariant set E? The following results answer this question affirmatively:

Theorem 6.3.3. *Let* E *be a Lebesgue-almost* $(P(t))_{t\geq 0}$*-invariant set. Then the set*

$$\hat{E} := \{x \in S : \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(E) = 1\}$$

contains E*, is Borel measurable and satisfies:*

(i) $\hat{E}$ *is* $(P(t))_{t\geq 0}$*-invariant,*

(ii) $\mu(E) = \mu(\hat{E})$ *for every* $(P(t))_{t\geq 0}$*-invariant probability measure* μ*.*

Consequently, any $(P(t))_{t\geq 0}$*-invariant probability measure* μ *is ergodic if and only if* $\mu(E) = 0$ *or* $\mu(E) = 1$ *for every* $(P(t))_{t\geq 0}$*-invariant Borel set* E*.*

Proof. First observe that Corollary 6.2.10 implies that

$$\hat{E} = \{x \in S : \limsup_{n\to\infty} \mathbf{P}^{(n)}\delta_x(E) = 1\}, \tag{6.1}$$

where the n ranges over $\mathbb{N}$. Second, if $x \in E$, then $P(t)\delta_x(E) = 1$ for Lebesgue almost every $t \in \mathbb{R}_+$, so $\mathbf{P}^{(t)}\delta_x(E) = 1$ for every $t \in \mathbb{R}_+$. Thus $x \in \hat{E}$.

Let $B = S\backslash\hat{E}$. Then $B = \{x \in S : \limsup_{n\to\infty} \mathbf{P}^{(n)}\delta_z(E) < 1\}$. We can write $B = \cup_{m\in\mathbb{N}} B_m$, where

$$B_m := \{x \in S : \limsup_{n\to\infty} \mathbf{P}^{(n)}\delta_x(E) \leq 1 - 1/m\}.$$

Fix $m \in \mathbb{N}$, then $B_m = \cap_{d\in\mathbb{N}} \cup_{N\in\mathbb{N}} \cap_{n\geq N} C_{d,m,n}$, where $C_{d,m,n} = \{x \in S : \mathbf{P}^{(n)}\delta_x(E) \leq 1 - 1/m + 1/d\}$. Since $\mathbf{P}^{(n)}$ is a regular Markov operator and E is a Borel set, $x \mapsto \mathbf{P}^{(n)}\delta_x(E)$ is Borel measurable by Proposition 3.3.1 and Proposition 3.2.4, thus $C_{d,m,n}$ is a Borel set for every $d, m, n \in \mathbb{N}$.

So B is Borel measurable, and consequently $\hat{E}$ as well.

The crucial part in the proof is the following:

Claim: If $\mu \in \mathcal{P}(S)$ is such that $\limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(E) = 1$, then $\mu(\hat{E}) = 1$.
Proof of claim: By assumption there is a sequence $(t_n)_n \subset \mathbb{R}_+$ such that $t_n \uparrow \infty$ and $\mathbf{P}^{(t_n)}\mu(E) \to 1$. Since $0 \leq \mathbf{U}^{(t_n)}\mathbb{1}_E(x) = \mathbf{P}^{(t_n)}\delta_x(E) \leq 1$ for every $x \in S$, we have $\mathbf{U}^{(t_n)}\mathbb{1}_E \leq \mathbb{1}_S$ and thus

$$\int_S |\mathbb{1}_S - \mathbf{U}^{(t_n)}\mathbb{1}_E|\, d\mu = \int_S (\mathbb{1}_S - \mathbf{U}^{(t_n)}\mathbb{1}_E)\, d\mu = 1 - \mathbf{P}^{(t_n)}\mu(E) \to 0$$

as $n \to \infty$. Thus $(\mathbf{U}^{(t_n)}\mathbb{1}_E)_n$ converges to $\mathbb{1}_S$ in $L^1(\mu)$. Then there is a subsequence $(\mathbf{U}^{(t_{n_k})}\mathbb{1}_E)_k$ that converges to $\mathbb{1}_S$ μ-a.e. [44, Corollary 2.32]. So there is a Borel set $D \subset S$ such that $\mu(D) = 1$ and $\mathbf{P}^{(t_{n_k})}\delta_x(E) = \mathbf{U}^{(t_{n_k})}\mathbb{1}_E(x) \to 1$ for every $x \in D$. So $D \subset \hat{E}$ and thus $\mu(\hat{E}) = 1$.

We now prove the remaining two properties:
(i) Let $x \in \hat{E}$ and $t \in \mathbb{R}_+$. There is a sequence $(t_n)_n \subset \mathbb{R}_+$ such that $t_n \uparrow \infty$ and $\mathbf{P}^{(t_n)}\delta_x(E) \to 1$. By Lemma 6.2.12, $\|\mathbf{P}^{(t_n)}P(t)\delta_x - \mathbf{P}^{(t_n)}\delta_x\|_{\mathrm{TV}} \to 0$ as $n \to \infty$, thus $\mathbf{P}^{(t_n)}P(t)\delta_x(E) \to 1$ as well. So $P(t)\delta_x(\hat{E}) = 1$, according to the claim, thus $\hat{E}$ is $(P(t))_{t\geq 0}$-invariant.

(ii) We first show that the statement holds for every R-ergodic measure. Let ν be an R-ergodic measure. E is R-invariant by assumption and by *(i)*, $\hat{E}$ is R-invariant as well. So $\nu(E) \in \{0, 1\}$ and $\nu(\hat{E}) \in \{0, 1\}$. Assume that $\nu(E) \neq \nu(\hat{E})$. Since $E \subset \hat{E}$, the only possibility is that $\nu(E) = 0$ and $\nu(\hat{E}) = 1$. Then $\mathbf{P}^{(t)}\nu(E) = \nu(E) = 0$ for every $t \in \mathbb{R}_+$. So for every $t \in \mathbb{R}_+$

$$0 = \mathbf{P}^{(t)}\nu(E) = \int_S \mathbf{P}^{(t)}\delta_x(E)\, d\nu(x) = \int_{\hat{E}} \mathbf{P}^{(t)}\delta_x(E)\, d\nu(x).$$

Thus for each $t \in \mathbb{R}_+$ there is a Borel set $F_t \subset \hat{E}$ such that $\nu(F_t) = 1$ and $\mathbf{P}^{(t)}\delta_x(E) = 0$ for every $x \in F_t$. Define $F := \cap_{n\in\mathbb{N}} F_n$. Then F is Borel measurable, $F \subset \hat{E}$ and $\nu(F) = 1$. For every $x \in F$, $\mathbf{P}^{(n)}\delta_x(E) = 0$ for all $n \in \mathbb{N}$. According to (6.1) $F \cap \hat{E} = \emptyset$, so $F = \emptyset$. This contradicts $\nu(F) = 1$ and we conclude that $\nu(E) = \nu(\hat{E})$. Now let μ be a $(P(t))_{t\geq 0}$-invariant probability measure. Then μ is R-invariant by Proposition 6.2.5. For every $x \in \Gamma^R_{cpie}$, ϵ_x is ergodic, and thus $\epsilon_x(E) = \epsilon_x(\hat{E})$. Hence

$$\mu(E) = \int_{\Gamma^R_{cpie}} \epsilon_x(E)\, d\mu(x) = \int_{\Gamma^R_{cpie}} \epsilon_x(\hat{E})\, d\mu(x) = \mu(\hat{E})$$

by Theorem 5.4.11 and Proposition 3.2.5.

The final statement is a simple consequence of Theorem 6.3.2. □

6.3.2 Preliminary Yosida-type decomposition of state space and integral decomposition of invariant measures

In this section we prove results for Markov semigroups that are similar to those obtained in Chapter 5 for Markov operators.

Recall that we assume $\mathbf{P} = (P(t))_{t\geq 0}$ to be a regular jointly measurable Markov semigroup on S. Let R be its resolvent and $\mathbf{U} = (U(t))_{t\geq 0}$ the associated dual semigroup.

We define

$$\Gamma^{\mathbf{P}}_t := \{x \in S : \{\mathbf{P}^{(t)}\delta_x : t \in \mathbb{R}_{\geq 1}\} \text{ is tight}\},$$

$$\Gamma^{\mathbf{P}}_{cp} := \{x \in S : \mathbf{P}^{(t)}\delta_x \text{ converges in } \mathcal{S}_{\mathrm{BL}} \text{ as } t \to \infty.\}$$

For $x \in \Gamma^{\mathbf{P}}_{cp}$ we define $\varepsilon_x = \lim_{t\to\infty} \mathbf{P}^{(t)}\delta_x$. Then $\varepsilon_x \in \mathcal{P}(S)$. Notice that we distinguish in notation this measure from the Markov operator analogue ϵ_x that we associate to R.

Then we set

$$\Gamma^{\mathbf{P}}_{cpi} = \{x \in S : \varepsilon_x \text{ is } (P(t))_{t\geq 0}\text{-invariant}\}.$$

It need not be true that ε_x is a $(P(t))_{t\geq 0}$-ergodic measure whenever $x \in \Gamma^{\mathbf{P}}_{cpi}$, as the following example shows.

Example. Let $S = [-1, 1]$ with Euclidean metric d and define for $x \in S$ and $t > 0$,

$$P(t)\delta_x := \begin{cases} \delta_{\max(x-t,-1)} & \text{if } x < 0 \\ \frac{1}{2}\delta_{\max(-t,-1)} + \frac{1}{2}\delta_{\min(t,1)} & \text{if } x = 0 \\ \delta_{\min(x+t,1)} & \text{if } x > 0. \end{cases}$$

It is not difficult to prove that $(P(t))_{t\geq 0}$ defines a Markov semigroup via $P(t)\mu = \int_S P(t)\delta_x\, d\mu(x)$ and that $(P(t))_{t\geq 0}$ is stochastically continuous and Markov-Feller. Clearly, δ_{-1}, δ_1 and $\varepsilon_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ are invariant with respect to $(P(t))_{t\geq 0}$.

Hence $0 \in \Gamma^{\mathbf{P}}_{cpi}$, but ε_0 is not ergodic, since ε_0 is not an extreme point of the invariant measures.

Therefore we define

$$\Gamma^{\mathbf{P}}_{cpie} := \{x \in S : \varepsilon_x \text{ is } (P(t))_{t\geq 0}\text{-ergodic}\}.$$

Clearly $\Gamma^{\mathbf{P}}_t \supset \Gamma^{\mathbf{P}}_{cp} \supset \Gamma^{\mathbf{P}}_{cpi} \supset \Gamma^{\mathbf{P}}_{cpie}$.

We define

$$\mathcal{P}^{\mathbf{P}}_t := \{\mu \in \mathcal{P}(S) : (\mathbf{P}^{(t)}\mu)_{t\geq 1} \text{ is tight } \}$$

and

$$\mathcal{P}^{\mathbf{P}}_{cp} := \{\mu \in \mathcal{P}(S) : \mathbf{P}^{(t)}\mu \text{ converges in } \mathcal{S}_{\mathrm{BL}} \text{ as } t \to \infty.\}$$

For $\mu \in \mathcal{P}^{\mathbf{P}}_{cp}$, we define $\varepsilon_\mu := \lim_{t\to\infty} \mathbf{P}^{(t)}\mu$. Then we can define

$$\mathcal{P}^{\mathbf{P}}_{cpi} := \{\mu \in \mathcal{P}^{\mathbf{P}}_{cp} : \varepsilon_\mu \text{ is } (P(t))_{t\geq 0}\text{-invariant}\}$$

and

$$\mathcal{P}^{\mathbf{P}}_{cpie} := \{\mu \in \mathcal{P}^{\mathbf{P}}_{cpi} : \varepsilon_\mu \text{ is } (P(t))_{t\geq 0}\text{-ergodic}\}.$$

Note that $x \in \Gamma^{\mathbf{P}}_\bullet$ if and only if $\delta_x \in \mathcal{P}^{\mathbf{P}}_\bullet$, for $\bullet = t, cp, cpi, cpie$. If $x \in \Gamma^{\mathbf{P}}_{cp}$, then $\varepsilon_{\delta_x} = \varepsilon_x$.

Corollary 6.2.9 implies that for every $0 < \alpha < \beta$ and $\mu \in \mathcal{P}(S)$, $\{(\mathbf{P}^{(t)}\mu : t \in [\alpha, \beta]\}$ is compact in $\mathcal{S}_{\mathrm{BL}}$, thus tight. Consequently, if $\mu \in \mathcal{P}(S)$ is such that $(\mathbf{P}^{(t)}\mu)_{t\geq\alpha}$ is tight for some $\alpha > 0$, then it is tight for every $\alpha > 0$.

Theorem 6.3.4.

$$\mathcal{P}^{\mathbf{P}}_\bullet = \mathcal{P}^R_\bullet \textit{ and } \Gamma^{\mathbf{P}}_\bullet = \Gamma^R_\bullet,$$

where $\bullet = t, cp, cpi, cpie$. Moreover, $\epsilon_\mu = \varepsilon_\mu$ for every $\mu \in \mathcal{P}^{\mathbf{P}}_{cp}$. Consequently, $\Gamma^{\mathbf{P}}_t, \Gamma^{\mathbf{P}}_{cp}, \Gamma^{\mathbf{P}}_{cpi}$ and $\Gamma^{\mathbf{P}}_{cpie}$ are Borel sets and

$$\mu(\Gamma^{\mathbf{P}}_t) = \mu(\Gamma^{\mathbf{P}}_{cp}) = \mu(\Gamma^{\mathbf{P}}_{cpi}) = \mu(\Gamma^{\mathbf{P}}_{cpie}) = 1$$

for every $(P(t))_{t\geq 0}$-invariant probability measure μ.

Proof. Let $\mu \in \mathcal{P}^{\mathbf{P}}_t$. Let $(n_k)_k \subset \mathbb{N}$ such that $n_k \uparrow \infty$. Then there is a subsequence $(n_{k_l})_l$ such that $\mathbf{P}^{(n_{k_l})}\mu$ converges in $\mathcal{S}_{\mathrm{BL}}$, so by Theorem 6.2.7 $R^{(n_{k_l})}\mu$ converges in $\mathcal{S}_{\mathrm{BL}}$. Thus $\mu \in \mathcal{P}^R_t$.

Let $\mu \in \mathcal{P}^R_t$. Let $(t_k)_k \subset [1, \infty)$. If $(t_k)_k$ is bounded, then it has a converging subsequence, so $\mathbf{P}^{(t_k)}\mu$ has a converging subsequence by Corollary 6.2.9. Else there is a subsequence of $(t_k)_k$ converging to infinity. Then there is a further subsequence $(t_{k_l})_l$, such that $R^{(\lfloor t_{k_l} \rfloor)}\mu$ converges in $\mathcal{S}_{\mathrm{BL}}$, hence $\mathbf{P}^{(t_{k_l})}\mu$ converges in $\mathcal{S}_{\mathrm{BL}}$ by Corollary 6.2.11. Thus $\mu \in \mathcal{P}^{\mathbf{P}}_t$. So $\mathcal{P}^{\mathbf{P}}_t = \mathcal{P}^R_t$ and thus $\Gamma^{\mathbf{P}}_t = \Gamma^R_t$.

Corollary 6.2.11 implies that $\mathcal{P}^{\mathbf{P}}_{cp} = \mathcal{P}^{R}_{cp}$, $\Gamma^{\mathbf{P}}_{cp} = \Gamma^{R}_{cp}$ and $\varepsilon_\mu = \epsilon_\mu$ for every $\mu \in \mathcal{P}^{\mathbf{P}}_{cp}$. By Proposition 6.2.5, $\mathcal{P}^{\mathbf{P}}_{cpi} = \mathcal{P}^{R}_{cpi}$ and $\Gamma^{\mathbf{P}}_{cpi} = \Gamma^{R}_{cpi}$. From Theorem 6.3.2, we obtain that $\mathcal{P}^{\mathbf{P}}_{cpie} = \mathcal{P}^{R}_{cpie}$ and $\Gamma^{\mathbf{P}}_{cpie} = \Gamma^{R}_{cpie}$.

The final statement follows from the results obtained in Chapter 5. □

We can give a necessary and sufficient condition for a probability measure to be in $\mathcal{P}^{\mathbf{P}}_t$.

By $\mathcal{C}_\epsilon$ we denote the family of all Borel subsets of S who are contained in a finite union of open ϵ-balls.

Lemma 6.3.5. *Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a regular jointly measurable Markov semigroup. Let $\mu \in \mathcal{P}(S)$. Then $\mu \in \mathcal{P}^{\mathbf{P}}_t$ if and only if for all $\epsilon > 0$ there is a $C \in C_\epsilon$ such that*

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\mu(C) \geq 1 - \epsilon. \tag{6.2}$$

Proof. Let $\mu \in \mathcal{P}(S)$ such that for all $\epsilon > 0$ there is a $C \in C_\epsilon$ such that (6.2) holds. Fix $\epsilon > 0$. Then there is a $C \in C_{\epsilon/2}$ and $T > 1$ such that $\mathbf{P}^{(t)}\mu(C) \geq 1 - \epsilon$ for all $t \geq T$. It follows from Corollary 6.2.9 that $\{\mathbf{P}^{(t)}\mu : 1 \leq t \leq T\}$ is tight, thus there is by Theorem 5.3.3 a $C' \in C_\epsilon$ such that $\mathbf{P}^{(t)}\mu(C') \geq 1 - \epsilon$ for all $t \in [1, T]$. Clearly $C \cup C' \in C_\epsilon$, hence Theorem 5.3.3 yields that $\mu \in \mathcal{P}^{\mathbf{P}}_t$.

The other direction follows from Theorem 5.3.3. □

On $\Gamma^{\mathbf{P}}_{cpie}$ an equivalence relation $\sim$ is defined as follows: $x \sim y$ whenever $\varepsilon_x = \varepsilon_y$. We write $[x]$ to denote the equivalence class of $x \in \Gamma^{\mathbf{P}}_{cpie}$. The following result comes from Theorem 5.4.7. It implies that we can decompose $\Gamma^{\mathbf{P}}_{cpie}$ into disjoint Borel measurable subsets, such that each ergodic measure has full measure on exactly one of these subsets.

Theorem 6.3.6. *(i) For every $x \in \Gamma^{\mathbf{P}}_{cpie}$ the set $[x]$ is Borel measurable and $\varepsilon_x([x]) = 1$.*

(ii) Any ergodic measure μ is of the form $\mu = \varepsilon_x$ for some $x \in \Gamma^{\mathbf{P}}_{cpie}$.

Using the characterisation we obtain an integral decomposition of invariant probability measures in terms of ergodic measures (Theorem 5.4.11).

Theorem 6.3.7. *Let μ be an invariant probability measure. Then the map*

$$x \mapsto \begin{cases} \varepsilon_x & \text{if } x \in \Gamma^{\mathbf{P}}_{cpie} \\ 0 & \text{if } x \notin \Gamma^{\mathbf{P}}_{cpie}. \end{cases} \tag{6.3}$$

is strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$ and

$$\mu = \int_{\Gamma^{\mathbf{P}}_{cpie}} \varepsilon_x \, d\mu(x)$$

as Bochner integral in $\mathcal{S}_{\mathrm{BL}}$.

For $f \in C_b(S)$ we define the Borel measurable function

$$f^*(x) = \begin{cases} \langle \varepsilon_x, f \rangle & \text{if } x \in \Gamma^{\mathbf{P}}_{cp} \\ 0 & \text{if } x \notin \Gamma^{\mathbf{P}}_{cp}. \end{cases}$$

By applying Theorem 6.2.7 and Theorem 6.3.4 to the results in Section 5.5, we find analogous results in the Markov semigroup setting.

Proposition 6.3.8. *Let $\mu \in \mathcal{P}(S)$.*

(i) If $\mu(\Gamma^{\mathbf{P}}_{cp}) = 1$, then $\mu \in \mathcal{P}^{\mathbf{P}}_{cp}$, $\varepsilon_\mu = \int_{\Gamma^{\mathbf{P}}_{cp}} \varepsilon_x \, d\mu(x)$ and for every $f \in C_b(S)$, $\langle \varepsilon_\mu, f \rangle = \langle \mu, f^* \rangle$.

(ii) If $\mu(\Gamma^{\mathbf{P}}_{cpi}) = 1$, then $\mu \in \mathcal{P}^{\mathbf{P}}_{cpi}$ and $\epsilon_\mu = \int_{\Gamma^{\mathbf{P}}_{cpi}} \epsilon_x \, d\mu(x)$.

(iii) If $\mu([z]) = 1$ for some $z \in \Gamma^{\mathbf{P}}_{cpie}$, then $\mu \in \mathcal{P}^{\mathbf{P}}_{cpie}$ and $\varepsilon_\mu = \varepsilon_z$.

Proposition 6.3.9. *Let μ be a finite Borel measure on S such that $|\mu|(S \backslash \Gamma^{\mathbf{P}}_{cp}) = 0$. Then there is a finite Borel measure μ^* such that the following statements holds:*

(i) $\|\mathbf{P}^{(t)}\mu - \mu^*\|^*_{\mathrm{BL}} \to 0$ as $t \to \infty$.

(ii) $\langle \mu^*, f \rangle = \langle \mu, f^* \rangle$ for every $f \in C_b(S)$

(iii) If $|\mu|(S \backslash \Gamma^{\mathbf{P}}_{cpi}) = 0$, then μ^* is invariant.

Proposition 6.3.10. *Let ν be an invariant probability measure and $\mu \in \mathcal{M}(S)$ such that $\mu \ll \nu$. Then there is an invariant probability measure μ^* such that $\|\mathbf{P}^{(t)}\mu - \mu^*\|_{\mathrm{TV}} \to 0$ and $\langle \mu^*, f \rangle = \langle \mu, f^* \rangle$ for every $f \in C_b(S)$.*

6.3.3 Full Yosida-type ergodic decomposition

Let P be a regular Markov operator and $\mathbf{P} = (P(t))_{t \geq 0}$ a regular jointly measurable Markov semigroup with resolvent R and dual semigroup $\mathbf{U} = (U(t))_{t \geq 0}$.

Let E be a Borel set in S. There is a natural bijection between $\mathcal{M}(E)$ and $\mathcal{M}_E(S) := \{\mu \in \mathcal{M}(S) : |\mu|(S \backslash E) = 0\}$: we can extend any finite Borel measure μ on E to a finite Borel measure $\overline{\mu}$ on S, by defining $\overline{\mu}(F) := \mu(F \cap E)$ for every Borel set F in S. Then clearly $|\overline{\mu}|(S \backslash E) = 0$. On the other hand, if ν is a finite Borel measure on S such that $|\nu|(S \backslash E) = 0$, then its restriction to E defines a Borel measure μ such that $\overline{\mu} = \nu$.

Let E be a $(P(t))_{t \geq 0}$-invariant Borel set. Then $(P(t))_{t \geq 0}$ leaves $\mathcal{M}_E(S)$ invariant: if $\mu \in \mathcal{M}_E(S)$, then by Proposition 3.3.1

$$P(t)|\mu|(E) = \int_S P(t)\delta_x(E) \, d|\mu|(x) \geq \int_E \mathbb{1} \, d|\mu| = |\mu|(E) = |\mu|(S).$$

Thus $|P(t)\mu|(S\backslash E) \leq P(t)|\mu|(S\backslash E) = |\mu|(S\backslash E) = 0$. So we can restrict $(P(t))_{t\geq 0}$ to $\mathcal{M}_E(S)$ This gives a 'restriction' of $(P(t))_{t\geq 0}$ to a regular Markov operator on $\mathcal{M}(E)$.

Our aim is to show an analogous result to Theorem 5.4.14 for the Markov semigroup $(P(t))_{t\geq 0}$.

Proposition 6.3.11. *Let $x \in \Gamma^{\mathbf{P}}_{cpie} = \Gamma^{R}_{cpie}$ and let $S_{[x]}$ be the R-invariant Borel set in $[x]$ given by Theorem 5.4.14. Then $S_{[x]}$ is $(P(t))_{t\geq 0}$-invariant.*

Proof. The set $S_{[x]}$ from Theorem 5.4.14 is defined by Lemma 5.4.2 as follows: $S_{[x]} = \cap_{n=1}^{\infty} B_n$, where $B_0 = [x]$ and

$$B_n = \{x \in B_{n-1} : R\delta_x(B_{n-1}) = 1\}.$$

Let E be an R-invariant Borel set in $[x]$, then $E \subset B_0$. Assume that $E \subset B_n$ for some $n \in \mathbb{N}_0$, then $R\delta_x(B_n) \geq R\delta_x(E) = 1$ for every $x \in E$, so $E \subset B_{n+1}$. Thus by induction $E \subset \cap_{n=1}^{\infty} B_n = S_{[x]}$, i.e. $S_{[x]}$ is the largest R-invariant set in $[x]$.

Let $\hat{S}_{[x]}$ be defined as in Theorem 6.3.3, i.e.

$$\hat{S}_{[x]} := \{z \in S : \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_z(S_{[x]}) = 1\}.$$

Then $\hat{S}_{[x]}$ is Borel measurable, $S_{[x]} \subset \hat{S}_{[x]}$ and $\hat{S}_{[x]}$ is $(P(t))_{t\geq 0}$-invariant by Theorem 6.3.3. We will show that $\hat{S}_{[x]} \subset [x]$. Since $\hat{S}_{[x]}$ is also R-invariant, this will imply that actually $\hat{S}_{[x]} = S_{[x]}$.

Let $z \in \hat{S}_{[x]}$. Then there is a sequence $(t_n)_n \subset \mathbb{R}_+$, such that $t_n \uparrow \infty$ and $0 < \mathbf{P}^{(t_n)}\delta_z(S_{[x]}) \to 1$ as $n \to \infty$. Since $S_{[x]} \subset [x]$, $0 < \mathbf{P}^{(t_n)}\delta_z([x]) \to 1$ as $n \to \infty$. For $E \subset S$ Borel, we define

$$\nu_n(E) := \frac{\mathbf{P}^{(t_n)}\delta_z(E \cap [x])}{\mathbf{P}^{(t_n)}\delta_z([x])}.$$

Then ν_n defines a probability measure on S, and clearly $\nu_n([x]) = 1$. Proposition 6.3.8 implies that $\nu_n \in \mathcal{P}^{\mathbf{P}}_{cpie}$ and $\varepsilon_{\nu_n} = \varepsilon_x$.

Since $\mathbf{P}^{(t_n)}\delta_z \geq \mathbf{P}^{(t_n)}\delta_z([x])\nu_n$,

$$\begin{aligned} \|\mathbf{P}^{(t_n)}\delta_z - \mathbf{P}^{(t_n)}\delta_z([x])\nu_n\|_{\mathrm{TV}} &= \mathbf{P}^{(t_n)}\delta_z(S) - \mathbf{P}^{(t_n)}\delta_z([x])\nu_n(S) \\ &= 1 - \mathbf{P}^{(t_n)}\delta_z([x]) \to 0 \end{aligned}$$

as $n \to \infty$, and

$$\|\mathbf{P}^{(t_n)}\delta_z([x])\nu_n - \nu_n\|_{\mathrm{TV}} \leq |\mathbf{P}^{(t_n)}\delta_z([x]) - 1| \to 0,$$

thus $\|\mathbf{P}^{(t_n)}\delta_z - \nu_n\|_{\mathrm{TV}} \to 0$ as $n \to \infty$.

It follows from Lemma 6.2.13 that $\lim_{t\to\infty} \|\mathbf{P}^{(t)}\rho - \mathbf{P}^{(t)}\mathbf{P}^{(s)}\rho\|_{\mathrm{TV}} = 0$ for every $\rho \in \mathcal{P}(S)$ and $s \in \mathbb{R}_+$. Thus, because $\mathbf{P}^{(t)}$ is a Markov operator,

$$\begin{aligned}\limsup_{t\to\infty} \|\mathbf{P}^{(t)}\delta_z - \mathbf{P}^{(t)}\nu_n\|_{\mathrm{TV}} &= \limsup_{t\to\infty} \|\mathbf{P}^{(t)}\mathbf{P}^{(t_n)}\delta_z - \mathbf{P}^{(t)}\nu_n\|_{\mathrm{TV}} \\ &\leq \|\mathbf{P}^{(t_n)}\delta_z - \nu_n\|_{\mathrm{TV}}.\end{aligned}$$

Now,

$$\begin{aligned}\limsup_{t\to\infty} \|\mathbf{P}^{(t)}\delta_z - \varepsilon_x\|^*_{\mathrm{BL}} &\leq \limsup_{t\to\infty} \|\mathbf{P}^{(t)}\delta_z - \mathbf{P}^{(t)}\nu_n\|_{\mathrm{TV}} \\ &\leq \|\mathbf{P}^{(t_n)}\delta_z - \nu_n\|_{\mathrm{TV}},\end{aligned}$$

which converges to zero as $n \to \infty$. Thus $z \in [x]$. □

Now we can conclude a full Yosida-type ergodic decomposition result for $(P(t))_{t\geq 0}$.

Theorem 6.3.12. *Let S be a Polish space and $\boldsymbol{P} = (P(t))_{t\geq 0}$ a regular jointly measurable Markov semigroup. If there exist invariant measures, or equivalently, if $\Gamma^{\mathbf{P}}_{cpie}$ is not empty, then for every $x \in \Gamma^{\mathbf{P}}_{cpie}$ the following statements hold:*

(i) There is a $(P(t))_{t\geq 0}$-invariant Borel set $S_{[x]} \subset [x]$ such that $\varepsilon_x(S_{[x]}) = 1$.

(ii) ε_x is the unique invariant probability measure of the restriction $(P(t)_{[x]})_{t\geq 0}$ of $(P(t))_{t\geq 0}$ to $\mathcal{M}(S_{[x]})$.

(iii) $(P(t)_{[x]})_{t\geq 0}$ is ergodic in the sense that $S_{[x]}$ cannot be written as the union of two disjoint $(P(t)_{[x]})_{t\geq 0}$-invariant sets A and B with $\varepsilon_x(A) > 0$ and $\varepsilon_x(B) > 0$.

Proof. (i) We define $S_{[x]}$ as in Proposition 6.3.11. Then the result follows from that proposition.

(ii) Since $S_{[x]}$ is $(P(t))_{t\geq 0}$-invariant, we can restrict $(P(t))_{t\geq 0}$ to a regular jointly measurable Markov semigroup $(P(t)_{[x]})_{t\geq 0}$ on $\mathcal{M}(S_{[x]})$. Let μ be a $(P(t)_{[x]})_{t\geq 0}$-invariant probability measure on $S_{[x]}$ and $\overline{\mu}$ the extension of μ to S. Then $\overline{\mu}$ is a $(P(t))_{t\geq 0}$-invariant probability measure on S such that $\overline{\mu}(S_{[x]}) = 1$, thus $\overline{\mu}(S_{[x]}) = 1$. Now, by Theorem 6.3.7 and since $\overline{\mu}(S\backslash S_{[x]}) = 0$

$$\begin{aligned}\overline{\mu} &= \int_{\Gamma^{\mathbf{P}}_{cpie}} \varepsilon_y \, d\overline{\mu}(y) = \int_{S_{[x]}} \varepsilon_y \, d\mu(y) \\ &= \int_{S_{[x]}} \varepsilon_x \, d\mu(y) = \varepsilon_x,\end{aligned}$$

thus μ is the restriction of ε_x to $S_{[x]}$.

(iii) Let A, B be disjoint $(P(t)_{[x]})_{t\geq 0}$-invariant Borel subsets of $S_{[x]}$ such that $\varepsilon_x(A) > 0$ and $\varepsilon_x(B) > 0$. Then A, B are disjoint $(P(t))_{t\geq 0}$-invariant Borel subsets of S, thus by ergodicity of ε_x, $\varepsilon_x(A) = \varepsilon_x(B) = 1$. But then $\varepsilon_x(A \cup B) = 2$, which gives a contradiction with the fact that ε_x is a probability measure. □

Remark. Theorem 6.3.12 extends [18, Proposition 5.2 and Theorem 5.1]. There, Markov semigroups associated to Borel right processes on locally compact separable metric spaces are considered (see [109, pp. 104-105] for a definition of Borel right processes). It follows from [8, Remark 5 in §29] that every locally compact separable metric space is a Polish space, so our setting is more general. Also, we do not need any continuity result on the Markov semigroup: joint measurability and regularity suffice. Finally the sets defined in [18, Theorem 5.1] depend on an *a priori* chosen invariant probability measure, while our sets do not.

From Proposition 5.5.6 and Proposition 6.3.11 we obtain the following result:

Proposition 6.3.13. *The set* $G := \bigcup_{x \in \Gamma^{\mathbf{P}}_{cpie}} S_{[x]}$ *is an invariant Borel set and* $\mu(G) = 1$ *for every invariant probability measure* μ.

Proposition 6.3.14. *Let* $\mu \in \mathcal{P}(S)$ *be such that*

$$\limsup_{t \to \infty} \mathbf{P}^{(t)} \mu([x]) = 1$$

for some $x \in \Gamma^{\mathbf{P}}_{cpie}$. *Then* $\mu([x]) = 1$, $\mu \in \mathcal{P}^{\mathbf{P}}_{cpie}$ *and* $\mathbf{P}^{(t)}\mu \to \varepsilon_x$ *in* $\mathcal{S}_{\mathrm{BL}}$ *as* $t \to \infty$.

Proof. There is a sequence $(t_n)_n \in \mathbb{R}_+$ such that $t_n \uparrow \infty$ and $\mathbf{P}^{(t_n)}\mu([x]) \to 1$ as $n \to \infty$. Proceeding as in the proof of Theorem 6.3.3,

$$\int_S |\mathbb{1}_S - \mathbf{U}^{(t_n)} \mathbb{1}_{[x]}| \, d\mu = 1 - \mathbf{P}^{(t_n)} \mu([x]) \to 0$$

as $n \to \infty$, so by [44, Corollary 2.32] there is a Borel set $D \subset S$ and subsequence $(\mathbf{U}^{(t_{n_k})} \mathbb{1}_{[x]})_k$, such that $\mu(D) = 1$ and $\mathbf{U}^{(t_{n_k})} \mathbb{1}_{[x]}(z) \to 1$ for every $z \in D$.

Re-examination of the proof of Proposition 6.3.11 shows that if $z \in S$ is such that

$$\mathbf{P}^{(t_{n_k})} \delta_z([x]) = \mathbf{U}^{(t_{n_k})} \mathbb{1}_{[x]}(z) \to 1 \text{ as } k \to \infty,$$

then $z \in [x]$. Thus $D \subset [x]$ and consequently $\mu([x]) = 1$. The final statement follows from Proposition 6.3.8. □

6.4 Notes

This chapter is based on the submitted paper [126] with some minor modifications.

Our results on the Yosida-type ergodic decomposition in Section 6.3.3 generalise results by Costa and Dufour [18]. They consider Markov semigroups on locally compact separable metric spaces, and require more regularity. Our results do not require *continuity* assumptions on the Markov semigroup: the operators $P(t)$ need not be Markov-Feller, only regular, and the orbits $t \mapsto P(t)\mu$ need not be continuous in any sense: jointly measurable is sufficient. These are natural assumptions when considering Markov semigroups associated to Markov processes.

CHAPTER

SEVEN

EQUICONTINUOUS FAMILIES OF MARKOV-FELLER OPERATORS ON POLISH SPACES WITH APPLICATIONS TO ERGODIC DECOMPOSITIONS AND EXISTENCE, UNIQUENESS AND STABILITY OF INVARIANT MEASURES

7.1 Introduction

In this chapter we consider families of Markov operators on Polish spaces with the so-called *e-property*. This property has also been considered by Szarek *et al.* [70, 73, 81, 113, 114]. If $U(t)$ is the dual of the regular Markov operator $P(t)$, then the family $(P(t))_{t\in T}$ has the e-property if and only if for each bounded Lipschitz function f on S, and any $x \in S$, the family of functions $(U(t)f)_{t\in T}$ is equicontinuous at x. Note that this definition depends on the particular metric we choose for the Polish space. The e-property is weaker than the well-studied strong Feller property [20]. We are able to obtain some stronger results than found in [73, 81, 113, 114], and even go beyond the e-property condition by requiring the weaker Cesàro e-property only. These results entail various characterisations of existence, uniqueness and stability of invariant measures for (discrete or continuous-time) semigroups of Markov operators that have the (Cesàro) e-property.

Our approach is founded on Theorem 2.3.24, which established a relationship between weak convergence and norm convergence in the Banach space $\mathcal{S}_{\mathrm{BL}}$ that we introduced in Chapter 2. We use this result to obtain equivalent conditions for families of Markov operators to be equicontinuous in Section 7.2 and are able to

conclude various interesting properties on particular sets of measures associated to such families.

In Section 7.3 we explore our general results of Section 7.2 in various settings of semigroups of Markov operators. Similar to our approach in Chapter 6 we "reduce" the continuous-time semigroup case to the discrete-time case of a single iterated Markov operator through the analysis of the *resolvent operator* associated to the continuous-time Markov semigroup (see Section 6.2 for its definition). We obtain a stronger version of a Yosida-type ergodic decomposition of the state space under the condition of the Cesàro e-property to hold. Compared to Chapters 5 and 6, the improvement in this setting consists of closedness and invariance of the sets in the decomposition, and we obtain a *continuous* surjective function from one of these sets to the ergodic invariant probability measures.

Our work on existence, uniqueness and stability of invariant measures was inspired by the interesting work of Szarek and coworkers [73, 81, 113, 114]. Our type of arguments yields various improvements, generalisations and novel results on this topic, all collected in Section 7.4. Some of these turned out to be obtained independently, with different arguments and under slightly different conditions, in [70].

Let S be a Polish space and d a complete metric on S that metrises its topology. Let $\mathcal{S}_{\mathrm{BL}}$ be the separable Banach space associated to (S, d) (see Chapter 2).

7.2 Equicontinuous families of Markov operators

Recall that a Markov operator on S is Markov-Feller if it is regular and its dual maps $C_b(S)$ into itself. We start by giving a sufficient condition for a regular Markov operator to be Markov-Feller:

Lemma 7.2.1. *Let P be a regular Markov operator with dual U such that $Uf \in C_b(S)$ for every $f \in \mathrm{BL}(S)$, then P is Markov-Feller.*

Proof. By Proposition 3.3.2, it suffices to show that $P : \mathcal{S}^+_{\mathrm{BL}} \to \mathcal{S}^+_{\mathrm{BL}}$ is continuous. Let $(\mu_n)_n, \mu$ be in $\mathcal{M}^+(S)$ such that $\|\mu_n - \mu\|^*_{\mathrm{BL}} \to 0$. Let $f \in \mathrm{BL}(S)$. By Lemma 2.3.12,

$$\langle P\mu_n, f\rangle = \langle \mu_n, Uf\rangle \to \langle \mu, Uf\rangle = \langle P\mu, f\rangle$$

as $n \to \infty$. Note that $\|P\mu_n\|_{\mathrm{TV}} = \|\mu_n\|_{\mathrm{TV}} = \|\mu_n\|^*_{\mathrm{BL}}$ is bounded as n ranges in $\mathbb{N}$, since μ_n converges in $\mathcal{S}_{\mathrm{BL}}$. Theorem 2.3.24 yields $\|P\mu_n - P\mu\|^*_{\mathrm{BL}} \to 0$. □

Let $H \subset C_b(T, X)$, where T is a topological space and (X, d) a metric space. We say that H is *equicontinuous at $t \in T$*, if for every $\epsilon > 0$ there is an open set $U \subset T$ containing t, such that $d(f(t), f(s)) < \epsilon$ for every $s \in U$. H is an *equicontinuous family* if it is equicontinuous at every $t \in T$. The following result gives several equivalent condition for a family of regular Markov operators to be equicontinuous:

Theorem 7.2.2. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of regular Markov operators on S. Let U_λ be the dual of P_λ. The following statements are equivalent:*

(i) $(P_\lambda)_{\lambda\in\Lambda}$ *is an equicontinuous family in* $C_b(\mathcal{S}^+_{\mathrm{BL}}, \mathcal{S}^+_{\mathrm{BL}})$.

(ii) For any topological space T and continuous map $\Phi : T \to \mathcal{S}^+_{\mathrm{BL}}$, $(P_\lambda \circ \Phi)_{\lambda\in\Lambda}$ is an equicontinuous family in $C_b(T, \mathcal{S}^+_{\mathrm{BL}})$.

(iii) For every $f \in \mathrm{BL}(S)$, $(U_\lambda f)_{\lambda\in\Lambda}$ is an equicontinuous family in $C_b(S)$.

Proof. (i)⇒ (ii): This is trivial.
(ii)⇒(iii): Let $T = S$ and Φ be the map $x \mapsto \delta_x : S \to \mathcal{S}^+_{\mathrm{BL}}$. This map is continuous (Lemma 2.3.6). If $f = 0$, the statement is trivial. So assume $f \neq 0$ and let $\epsilon > 0$. There exists a $\delta > 0$ such that for all $y \in S$ such that $d(x, y) < \delta$,

$$\|P_\lambda\delta_x - P_\lambda\delta_y\|^*_{\mathrm{BL}} < \frac{\epsilon}{\|f\|_{\mathrm{BL}}}, \text{ for all } \lambda \in \Lambda.$$

Then

$$|U_\lambda f(x) - U_\lambda f(y)| \quad = \quad |\langle P_\lambda\delta_x, f\rangle - \langle P_\lambda\delta_y, f\rangle| \leq \|P_\lambda\delta_x - P_\lambda\delta_y\|^*_{\mathrm{BL}}\|f\|_{\mathrm{BL}} < \epsilon.$$

(iii)⇒(i): Suppose not. Then there exists a $\mu \in \mathcal{S}^+_{\mathrm{BL}} = \mathcal{M}^+(S)$ where $(P_\lambda)_{\lambda\in\Lambda}$ is not equicontinuous. Thus there exists an $\epsilon > 0$, $\mu_k \in \mathcal{M}^+(S)$ and $\lambda_k \in \Lambda$ such that $\|\mu_k - \mu\|^*_{\mathrm{BL}} \leq \frac{1}{k}$ and

$$\|P_{\lambda_k}\mu_k - P_{\lambda_k}\mu\|^*_{\mathrm{BL}} \geq \epsilon$$

for every $k \in \mathbb{N}$. By Lemma 2.3.12 we know that μ_k converges to μ in the weak topology. This implies by [28, Theorem 7] that μ_k converges to μ uniformly on any equicontinuous and uniformly bounded family of functions on S.

Take $f \in \mathrm{BL}(S)$. By assumption, $(U_{\lambda_k} f)_k$ is an equicontinuous and uniformly bounded family of functions, thus

$$|\langle P_{\lambda_k}\mu_k - P_{\lambda_k}\mu, f\rangle| = |\langle \mu_k - \mu, U_{\lambda_k} f\rangle| \to 0.$$

So by Theorem 2.3.24 we obtain that $\|P_{\lambda_k}\mu_k - P_{\lambda_k}\mu\|^*_{\mathrm{BL}} \to 0$ which gives a contradiction. □

For a regular Markov operator P we write $P\delta$ to denote the map from S to $\mathcal{S}_{\mathrm{BL}}$ given by $P\delta(x) = P\delta_x$. Note that $P\delta$ is the composition of the continuous map $\delta : x \mapsto \delta_x$ and P, so Theorem 7.2.2 implies:

Corollary 7.2.3. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a family of regular Markov operators on $\mathcal{M}^+(S)$ with duals $(U_\lambda)_{\lambda\in\Lambda}$, such that $(U_\lambda f)_{\lambda\in\Lambda}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Then $(P_\lambda\delta)_{\lambda\in\Lambda}$ is equicontinuous as well.*

We now prove some interesting properties on certain sets associated to equicontinuous families of Markov operators.

Theorem 7.2.4. *Let $(P_\lambda)_{\lambda\in\Lambda}$ be a collection of regular Markov operators on $\mathcal{M}^+(S)$ with duals $(U_\lambda)_{\lambda\in\Lambda}$, such that $(U_\lambda f)_{\lambda\in\Lambda}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Let*

$$\mathcal{P}_t^\Lambda := \{\mu \in \mathcal{P}(S) : (P_\lambda\mu)_{\lambda\in\Lambda} \text{ is tight}\}$$

and

$$\Gamma_t^\Lambda := \{x \in S : \delta_x \in \mathcal{P}_t^\Lambda\}.$$

Then the following statements hold:

(i) $\mathcal{P}_t^\Lambda$ is closed in $\mathcal{S}_{\mathrm{BL}}$ and Γ_t^Λ is closed in S.

(ii) For every $\mu \in \mathcal{P}_t^\Lambda$,

$$\mathrm{supp}(\mu) \subset \Gamma_t^\Lambda.$$

Proof. (i) The map $\Phi : S \to \mathcal{S}_{\mathrm{BL}} : x \mapsto \delta_x$ is continuous. If $\mathcal{P}_t^\Lambda$ is closed in $\mathcal{S}_{\mathrm{BL}}$, then $\Gamma_t^\Lambda = \Phi^{-1}(\mathcal{P}_t^\Lambda)$ is closed in S.

If $\mathcal{P}_t^\Lambda = \emptyset$, the statements hold. If $\mathcal{P}_t^\Lambda \neq \emptyset$, let μ be in the closure of $\mathcal{P}_t^\Lambda$ in $\mathcal{S}_{\mathrm{BL}}$ and let $(\mu_n)_n \in \mathcal{P}_t^\Lambda$ be such that $\mu_n \to \mu$. Then $\mu \in \mathcal{P}(S)$. For $E \subset S$ and $\epsilon > 0$ recall the definition

$$f_E^\epsilon(x) = (1 - d(x, E)/\epsilon)^+$$

from Section 5.3. Let D be a countable dense subset of S, and $\mathcal{F}$ the collection of finite subsets of D, then $\mathcal{F}$ is countable and by Theorem 5.2.3 we can write

$$\mathcal{P}_t^\Lambda = \cap_{m\in\mathbb{N}} \cup_{F\in\mathcal{F}} \cap_{\lambda\in\Lambda} \{\nu \in \mathcal{P}(S) : \langle P_\lambda\nu, f_F^{1/m}\rangle \geq 1 - 1/m\}.$$

Fix $m \in \mathbb{N}$. We will show that there exists an $F \in \mathcal{F}$ such that for every $\lambda \in \Lambda$,

$$\langle P_\lambda\mu, f_F^{1/m}\rangle \geq 1 - 1/m.$$

By Theorem 7.2.2, $(P_\lambda)_{\lambda\in\Lambda}$ is equicontinuous from $\mathcal{S}_{\mathrm{BL}}^+$ to $\mathcal{S}_{\mathrm{BL}}^+$, so there exists an $n_0 \in \mathbb{N}$, such that for all $\lambda \in \Lambda$,

$$\|P_\lambda\mu_{n_0} - P_\lambda\mu\|_{\mathrm{BL}}^* < \frac{1}{2m(2m+1)}.$$

Since $\mu_{n_0} \in \mathcal{P}_t^\Lambda$, there exists an $F_0 \in \mathcal{F}$, such that

$$\langle P_\lambda\mu_{n_0}, f_{F_0}^{1/(2m)}\rangle \geq 1 - 1/(2m) \text{ for every } \lambda \in \Lambda.$$

Now, $\|f_{F_0}^{1/(2m)}\|_{\mathrm{BL}}^* \leq 1 + |f_{F_0}^{1/(2m)}|_{\mathrm{Lip}} \leq 1 + 2m$. So

$$|\langle P_\lambda\mu_{n_0} - P_\lambda\mu, f_{F_0}^{1/(2m)}\rangle| < 1/(2m) \text{ for every } \lambda \in \Lambda.$$

Thus for every $\lambda \in \Lambda$ we have

$$\begin{aligned}\langle P_\lambda\mu, f_{F_0}^{1/m}\rangle &\geq \langle P_\lambda\mu, f_{F_0}^{1/(2m)}\rangle \\ &\geq \langle P_\lambda\mu_{n_0}, f_{F_0}^{1/(2m)}\rangle - 1/(2m) \geq 1 - 2/(2m) = 1 - 1/m.\end{aligned}$$

So $\mu \in \mathcal{P}_t^\Lambda$.

(ii) Let $\mu \in \mathcal{P}_t^\Lambda$ and $x \in \operatorname{supp}(\mu)$. Suppose that $x \notin \Gamma_t^\Lambda$. Then $(P_\lambda \delta_x)_{x\in\Lambda}$ is not tight, and there exists a sequence $(\lambda_n)_n \subset \Lambda$ such that $(P_{\lambda_n}\delta_x)_n$ does not have any convergent subsequences. Then $(P_{\lambda_n}\delta_x)_n$ is not tight.

Let $D \subset S$ be countable and dense, and define $\mathcal{F}$ as before to be the countable collection of finite subsets of D. By Theorem 5.3.3 there exists an $m \in \mathbb{N}$ such that for every $F \in \mathcal{F}$ there is a strictly increasing sequence $(n_k)_k \subset \mathbb{N}$ for which

$$U_{\lambda_{n_k}} f_F^{1/m}(x) = \langle P_{\lambda_{n_k}}\delta_x, f_F^{1/m}\rangle \leq 1 - 1/m.$$

By Corollary 7.2.3 we know that $(P_\lambda \delta)_{\lambda\in\Lambda}$ is equicontinuous in x, so in particular there exists a $\delta > 0$ such that for every $z \in B_x(\delta)$ and $\lambda \in \Lambda$,

$$\|P_\lambda \delta_x - P_\lambda \delta_z\|_{\mathrm{BL}}^* \leq \frac{1}{2m(m+1)}.$$

Since $x \in \operatorname{supp}(\mu)$, $\alpha := \mu(B_x(\delta)) > 0$.

By assumption $(P_{\lambda_{n_k}}\mu)_k$ is tight, so it can be shown using Theorem 5.3.3 that there exists an $F \in \mathcal{F}$ such that

$$\langle P_{\lambda_{n_k}}\mu, f_F^{1/m}\rangle > 1 - \frac{\alpha}{2m} \text{ for every } k \in \mathbb{N}, \tag{7.1}$$

by using the fact that $f_F^{1/m} \geq f_F^{1/m'}$ whenever $m' \geq m$.

Now we have for every $z \in B_x(\delta), k \in \mathbb{N}$, that

$$\begin{aligned} U_{\lambda_{n_k}} f_F^{1/m}(z) &\leq U_{\lambda_{n_k}} f_F^{1/m}(x) + |U_{\lambda_{n_k}} f_F^{1/m}(z) - U_{\lambda_{n_k}} f_F^{1/m}(x)| \\ &\leq 1 - 1/m + \|P_{\lambda_{n_k}}\delta_z - P_{\lambda_{n_k}}\delta_x\|_{\mathrm{BL}}^* \|f_F^{1/m}\|_{\mathrm{BL}} \\ &\leq 1 - 1/m + 1/(2m) = 1 - 1/(2m), \end{aligned}$$

since $\|f_F^{1/m}\|_{\mathrm{BL}}^* \leq m+1$. So for every $k \in \mathbb{N}$ we obtain

$$U_{\lambda_{n_k}} f_F^{1/m} \leq (1 - 1/(2m))\mathbb{1}_{B_x(\delta)} + \mathbb{1}_{S\setminus B_x(\delta)}.$$

Therefore,

$$\langle \mu, U_{\lambda_{n_k}} f_F^{1/m}\rangle \leq \alpha(1 - 1/(2m)) + 1 - \alpha = 1 - \frac{\alpha}{2m}$$

for every $k \in \mathbb{N}$, which contradicts (7.1). □

Let $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ be a regular Markov operator. Following [73], we say that P has the *e-property* if $(U^n f)_{n\in\mathbb{N}_0}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Every such Markov operator is automatically Markov-Feller by Lemma 7.2.1. P has the *Cesàro e-property* if $(U^{(n)} f)_{n\in\mathbb{N}}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Again

Lemma 7.2.1 implies that such a Markov operator is Markov-Feller. Note that every Markov operator with the e-property has the Cesàro e-property, though the converse statement need not be true in general. By Theorem 7.2.2, P has the e-property if and only if $(P^n)_{n\in\mathbb{N}_0}$ is an equicontinuous family in $C_b(\mathcal{S}^+_{\mathrm{BL}}, \mathcal{S}^+_{\mathrm{BL}})$, and the Cesàro e-property if and only if $(P^{(n)})_{n\in\mathbb{N}}$ is an equicontinuous family in $C_b(\mathcal{S}^+_{\mathrm{BL}}, \mathcal{S}^+_{\mathrm{BL}})$.

Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a regular jointly measurable Markov-Feller semigroup with dual semigroup $\mathbf{U} = (U(t))_{t\geq 0}$.

We say $\mathbf{P}$ has the *eventual e-property* if there is a $\tau \geq 0$ such that $(U(t)f)_{t\geq\tau}$ is equicontinuous for every $f \in \mathrm{BL}(S)$ and it has the *e-property* if τ can be chosen to be zero. Note that the corresponding definition of eventual e-property for Markov-Feller operators coincides with the definition of e-property, because a finite union of equicontinuous families of functions is equicontinuous.

$\mathbf{P}$ has the *Cesàro e-property* if there is a $\tau > 0$ such that $(\mathbf{U}^{(t)}f)_{t\geq\tau}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Since $t \mapsto \mathbf{P}^{(t)}$ is continuous from $(0,\infty)$ to $\mathcal{L}(\mathcal{M}(S)_{\mathrm{TV}})$ (endowed with the operator norm) according to Corollary 6.2.9, the map $t \mapsto \mathbf{U}^{(t)}f$ is continuous from $(0,\infty)$ to $(\mathrm{BM}(S), \|\cdot\|_\infty)$. Using this it can easily be shown that equicontinuity of $(\mathbf{U}^{(t)}f)_{t\geq\tau}$ for some $\tau > 0$ implies equicontinuity of $(\mathbf{U}^{(t)}f)_{t\geq\tau'}$ for any $\tau' > 0$.

Lemma 7.2.5. *If $\mathbf{P}$ has the eventual e-property, then it has the Cesàro e-property.*

Proof. Let $\tau > 0$ be such that $(U(t)f)_{t\geq\tau}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Let $t \geq \tau$. Then for every $x \in S$,

$$\mathbf{U}^{(t)}f(x) = \frac{1}{t}\int_0^\tau U(s)f(x)\,ds + \frac{1}{t}\int_\tau^t U(s)f(x)\,ds.$$

Note that $x \mapsto \int_0^\tau U(s)f(x)\,ds$ is a bounded continuous function.

Fix $x \in S$ and $\epsilon > 0$. Then there is a $\delta > 0$ such that $d(x,y) < \delta$ implies $|U(s)f(x) - U(s)f(y)| < \epsilon$ for every $s \geq \tau$ and

$$\left|\int_0^\tau U(s)f(x)\,ds - \int_0^\tau U(s)f(y)\,ds\right| < \epsilon.$$

Then for every $t \geq \tau$ and $y \in B_x(\delta)$,

$$|\mathbf{U}^{(t)}f(x) - \mathbf{U}^{(t)}f(y)| \leq \frac{1}{t}\epsilon + \frac{t-\tau}{t}\epsilon < (\frac{1}{\tau}+1)\epsilon.$$

This implies that $(\mathbf{U}^{(t)}f)_{t\geq\tau}$ is equicontinuous. □

Proposition 7.2.6. *(i) If $\mathbf{P}$ has the Cesàro e-property, then the resolvent R has the Cesàro e-property.*

(ii) If $\mathbf{P}$ has the eventual e-property, then R has the e-property.

Proof. Let V be the dual operator of the regular Markov operator R (cf. Proposition 6.2.2).

(i) From Theorem 6.2.7 it follows that

$$\lim_{n\to\infty} \|V^{(n)}f - \mathbf{U}^{(n)}f\|_\infty = 0.$$

Suppose $\mathbf{P}$ has the Cesàro e-property and let $f \in \mathrm{BL}(S)$ and $x \in S$. Fix $\epsilon > 0$. Then there is a $\delta > 0$ such that $|\mathbf{U}^{(n)}f(x) - \mathbf{U}^{(n)}f(y)| < \epsilon$ for every $y \in B_x(\delta)$ and $n \in \mathbb{N}$. There is also an $N \in \mathbb{N}$ such that $\|V^{(n)}f - \mathbf{U}^{(n)}f\|_\infty < \epsilon$ for every $n \geq N$. This implies that $|V^{(n)}f(x) - V^{(n)}f(y)| < 3\epsilon$ for every $y \in B_x(\delta)$ and $n \geq N$. Since $\{V^{(n)}f : 1 \leq n \leq N-1\}$ is equicontinuous, there is a $0 < \delta' < \delta$ such that $|V^{(n)}f(x) - V^{(n)}f(y)| < 3\epsilon$ for every $y \in B_x(\delta')$. This implies that $(V^{(n)}f)_{n\in\mathbb{N}}$ is equicontinuous at x.

(ii): In Lemma 6.2.6 we obtained that

$$R^n\mu = \int_0^\infty \frac{t^{n-1}}{(n-1)!} e^{-t} P(t)\mu \, dt.$$

Consequently, for every $f \in \mathrm{BM}(S)$ and $x \in S$,

$$V^n f(x) = \int_0^\infty \frac{t^{n-1}}{(n-1)!} e^{-t} U(t) f(x) \, dt.$$

Suppose $\mathbf{P}$ has the eventual e-property, then there is a $\tau > 0$ such that $(U(t)f)_{t\geq\tau}$ is equicontinuous for every $f \in \mathrm{BL}(S)$. Fix $f \in \mathrm{BL}(S)$, $x \in S$ and $\epsilon > 0$. Then there is a $\delta > 0$ such that $|U(t)f(x) - U(t)f(y)| < \epsilon$ for every $y \in B_x(\delta)$ and $t \geq \tau$. Now, for every $y \in B_x(\delta)$ we have

$$\begin{aligned}
|V^n f(x) - V^n f(y)| &\leq \int_0^\tau \frac{t^{n-1}}{(n-1)!} e^{-t} |U(t)f(x) - U(t)f(y)| \, dt \\
&\quad + \int_\tau^\infty \frac{t^{n-1}}{(n-1)!} e^{-t} |U(t)f(x) - U(t)f(y)| \, dt \\
&\leq 2\|f\|_\infty \int_0^\tau \frac{t^{n-1}}{(n-1)!} e^{-t} \, dt + \epsilon \int_\tau^\infty \frac{t^{n-1}}{(n-1)!} e^{-t} \, dt \\
&= 2\|f\|_\infty e^{-\tau} \sum_{k=n}^\infty \frac{\tau^k}{k!} + \epsilon e^{-\tau} \sum_{k=0}^{n-1} \frac{\tau^k}{k!}.
\end{aligned}$$

The first term goes to zero as $n \to \infty$ and the second term is bounded from above by ϵ, so there is an $N \in \mathbb{N}$ such that $|V^n f(x) - V^n f(y)| < 2\epsilon$ for every $n \geq N$ and $y \in B_x(\delta)$. Since $\{V^n f : 0 \leq n \leq N-1\}$ is equicontinuous, there is a $0 < \delta' < \delta$ such that $|V^n f(x) - V^n f(y)| < 2\epsilon$ for every $y \in B_x(\delta')$. This implies that $(V^n f)_{n\in\mathbb{N}}$ is equicontinuous at x. □

The strong Feller property is often used in combination with irreducibility ([20, p. 42]) to show uniqueness of invariant measures for Markov operators and Markov semigroups, by applying Doob's Theorem [20, Theorem 4.2.1]. A regular Markov operator P is called *strong Feller* if $U(\mathrm{BM}(S)) \subset C_b(S)$ and *eventually strong Feller* if there is an $N \in \mathbb{N}$ such that $U^N(\mathrm{BM}(S)) \subset C_b(S)$. Note that $U^N(\mathrm{BM}(S)) \subset C_b(S)$ implies that $U^n(\mathrm{BM}(S)) \subset C_b(S)$ for every $n \geq N$. P is called *ultra Feller* if $x \mapsto P\delta_x$ is continuous from S to $\mathcal{M}(S)_{\mathrm{TV}}$. Obviously, when P is ultra Feller, it is also strong Feller, though not necessary the other way around. However, it is a remarkable fact that if P and Q are both strong Feller operators, then PQ is ultra Feller [27, Théorème IX.18] (see also [107]).

A Markov semigroup $(P(t))_{t\geq 0}$ is *strong Feller* if $P(t)$ is strong Feller for every $t > 0$ and *eventually strong Feller* if there is a $\tau > 0$ such that $P(\tau)$ is strong Feller, which implies that $P(t)$ is strong Feller for every $t \geq \tau$. Since $P(t) = P(t/2)P(t/2)$, in these cases $P(t)$ is ultra Feller for every $t > 0$ respectively $t \geq 2\tau$.

Proposition 7.2.7. *If the Markov-Feller operator P is eventually strong Feller, then it has the e-property. If the Markov-Feller semigroup $(P(t))_{t\geq 0}$ is eventually strong Feller, then it has the eventual e-property.*

Proof. There is an $N \in \mathbb{N}$ such that P^N is strong Feller. Then P^{2N} is ultra Feller. For every $n \in \mathbb{N}$ we have

$$\|P^{n+2N}\delta_x - P^{n+2N}\delta_y\|_{\mathrm{TV}} \leq \|P^{2N}\delta_x - P^{2N}\delta_y\|_{\mathrm{TV}}.$$

This implies that for every $f \in \mathrm{BM}(S)$, $(U^n f)_{n\geq 2N}$ is equicontinuous. Hence $(U^n f)_{n\in\mathbb{N}_0}$ is equicontinuous.

The proof for the Markov semigroup case proceeds in an analogous manner. □

Note that the (eventual) strong Feller property is a rather restrictive condition on the Markov operator or semigroup. For instance, if $S = \mathbb{R}$ and $\Phi_t(x) := x + t$, then the induced Markov semigroup $(\Phi_t)_{t\geq 0}$ does not satisfy the eventual strong Feller property, but it *does* satisfy the e-property when S is endowed with Euclidean metric:

Example. Let $(\Phi_t)_{t\geq 0}$ be a jointly measurable semigroup of continuous maps on a complete separable metric space (S, d), and let $(P_\Phi(t))_{t\geq 0}$ be the associated Markov semigroup. Then one can show that $(P_\Phi(t))_{t\geq 0}$ satisfies the e-property if and only if the family of maps $(\Phi_t)_{t\geq 0}$ is equicontinuous. The analogous statement holds for a Markov operator associated to a continuous map $\Phi : S \to S$. A particular example is given by $S = \mathbb{R}$ with Euclidean metric and $\Phi_t(x) := x + t$.

Unlike some of the concepts we defined in Chapters 5 and 6, the e-property depends on the metric d we put on the Polish space S. If we have a Markov-Feller operator or semigroup that satisfies the e-property on (S, d), then it need not satisfy the e-property on (S, d'), where d' is a different metric on S metrising its topology:

Example. Let $S = \mathbb{R}_+$ with Euclidean metric d and define $\Phi : S \to S$ by $\Phi(x) := x + 1$. Then the associated Markov operator P_Φ satisfies the e-property on (S, d). Define $d'(x, y) := d(e^x, e^y)$, then d' is topologically equivalent to d, and (S, d') is again a complete separable metric space. Now, for all $n \in \mathbb{N}$ and $x > 0$ we have

$$d'(\Phi^n(0), \Phi^n(x)) = e^{n+x} - e^n \geq xe^n,$$

thus the family of maps $(\Phi^n)_{n\in\mathbb{N}}$ is not equicontinuous at zero, which implies that P_Φ does not satisfy the e-property on (S, d').

It would be interesting to be able to give topological criteria on the Polish space S and the Markov operator or semigroup, such that there exists a complete metric d' on S that metrises its topology, and such that the Markov operator or semigroup satisfies the e-property on (S, d').

We call a regular Markov operator P on (S, d) *non-expansive* (with respect to $\|\cdot\|^*_{\mathrm{BL}}$) if $\|P\mu\|^*_{\mathrm{BL}} \leq \|\mu\|^*_{\mathrm{BL}}$ for all $\mu \in \mathcal{M}(S)$. A regular jointly measurable Markov semigroup $(P(t))_{t\geq 0}$ on S is non-expansive if $P(t)$ is non-expansive for all $t \in \mathbb{R}_+$. Clearly non-expansive Markov operators and semigroups satisfy the e-property. Non-expansive Markov operators can be extended to bounded linear operators on $\mathcal{S}_{\mathrm{BL}}$. In [80, 110, 111, 112], non-expansivity of Markov operators and semigroups with respect to the equivalent norm

$$\|\mu\|^*_{\mathrm{BL,max}} = \{\langle \mu, f\rangle : f \in \mathrm{BL}(S) : \max(|f|_{\mathrm{Lip}}, \|f\|_\infty) \leq 1.\}$$

has been used to give conditions for existence, uniqueness and stability of invariant measures. Note that a non-expansive Markov operator P with respect to $\|\cdot\|^*_{\mathrm{BL}}$ need not be non-expansive with respect to $\|\cdot\|^*_{\mathrm{BL,max}}$ in general. However, we *can* conclude that for this P, the family $\{U^n f : n \in \mathbb{N}_0, f \in \mathrm{BL}(S), \|f\|_{\mathrm{BL,max}} \leq 1\}$ is uniformly equicontinuous, which seems to be the important ingredient for many of the results from the above mentioned papers.

One can easily construct examples of Markov operators and semigroups with the e-property that are not non-expansive. However, non-expansivity (as well as e-property) depends on the metric d. So this prompts the following question: If a regular Markov operator P on S satisfies the e-property, is there a topologically equivalent metric d' such that (S, d') is complete and P is non-expansive on (S, d')?

And of course one can ask the same question for regular jointly measurable Markov semigroups with the e-property.

If these questions are answered positively, then the various results for non-expansive Markov operators and semigroups can be applied to Markov operators and semigroups as well. We do not believe this is the case in general, though it seems difficult to construct counter examples. However, in certain specific cases the questions can be answered positively, as the following example shows:

Example. Let $(\Phi_t)_{t\geq 0}$ be a jointly measurable semigroup of continuous maps on a complete separable metric space (S, d), and let $(P_\Phi(t))_{t\geq 0}$ be the associated Markov

semigroup. Assume that $(P_\Phi(t))_{t\geq 0}$ satisfies the e-property. We define the metric

$$d'(x,y) := \sup_{t\in\mathbb{R}_+} \|P_\Phi(t)\delta_x - P_\Phi\delta_y\|^*_{\mathrm{BL}} = \sup_{t\in\mathbb{R}_+} \|\delta_{\Phi_t(x)} - \delta_{\Phi_t(y)}\|^*_{\mathrm{BL}}. \tag{7.2}$$

From Lemma 2.3.6, the e-property and Corollary 7.2.3 it follows that the metrics d and d' are topologically equivalent, and moreover, that every Cauchy sequence in S with respect to d' is a Cauchy sequence with respect to d. Thus (S,d') is a complete separable metric space. Now let $f : S \to 1$ be such that $\|f\|_{\mathrm{BL},d'} := |f|^{d'}_{\mathrm{Lip}} + \|f\|_\infty \leq 1$, where $|\cdot|^{d'}_{\mathrm{Lip}}$ is the Lipschitz seminorm with respect to the metric d'. Then

$$|U_\Phi(s)f(x) - U_\Phi(s)f(y)| = |f(\Phi_s(x)) - f(\Phi_s(y))| \leq d'(\Phi_s(x),\Phi_s(y)) \leq d'(x,y),$$

thus $\|U_\Phi(t)f\|_{\mathrm{BL},d'} \leq \|f\|_{\mathrm{BL},d'}$ for all bounded Lipschitz (with respect to d') f. And this implies that $(P_\Phi(t))_{t\geq 0}$ is non-expansive on (S,d').

7.3 Ergodic decomposition of Markov operators and semigroups with the Cesàro e-property

In this section we show that the ergodic decomposition associated to Markov operators and semigroups that we obtained in Chapter 5 and 6 has nice properties when we assume the Cesàro e-property. We start by considering a regular Markov operator P with the Cesàro e-property.

Theorem 7.3.1. *Let P be a regular Markov operator that satisfies the Cesàro e-property. Then $\mathcal{P}^P_t = \mathcal{P}^P_{cp}$ and consequently also $\Gamma^P_t = \Gamma^P_{cp}$.*

Proof. It is clear that $\mathcal{P}^P_{cp} \subset \mathcal{P}^P_t$. Let $\mu \in \mathcal{P}^P_t$. Define

$$D_\mu := \{\nu \in \mathcal{P}(S) : \lim_{n\to\infty} \|P^{(n)}\mu - P^{(n)}\nu\|^*_{\mathrm{BL}} \to 0\}.$$

Step 1. D_μ is closed in $\mathcal{S}_{\mathrm{BL}}$.

We can write

$$D_\mu = \bigcap_{m\in\mathbb{N}} \bigcup_{N\in\mathbb{N}} \{\rho \in \mathcal{P}(S) : \|P^{(n)}\mu - P^{(n)}\rho\|^*_{\mathrm{BL}} \leq 1/m \text{ for all } n \geq N\}.$$

Let $\nu_n \in D_\mu$ such that $\|\nu_n - \nu\|^*_{\mathrm{BL}} \to 0$ for some $\nu \in \mathcal{S}_{\mathrm{BL}}$, then $\nu \in \mathcal{P}(S)$, since $\mathcal{P}(S)$ is closed in $\mathcal{S}_{\mathrm{BL}}$.

Fix $m \in \mathbb{N}$. By Theorem 7.2.2, $(P^{(n)})_{n\in\mathbb{N}}$ is an equicontinuous family in $C_b(\mathcal{S}^+_{\mathrm{BL}}, \mathcal{S}^+_{\mathrm{BL}})$. Thus there exists a $k \in \mathbb{N}$ such that

$$\|P^{(n)}\nu_k - P^{(n)}\nu\|^*_{\mathrm{BL}} \leq 1/(2m) \text{ for every } n \in \mathbb{N}.$$

Since $\nu_k \in D_\mu$, there exists $N_0 \in \mathbb{N}$ such that

$$\|P^{(n)}\nu_k - P^{(n)}\mu\|^*_{\mathrm{BL}} \leq 1/(2m) \text{ for every } n \geq N_0.$$

Thus for every $m \in \mathbb{N}$,

$$\nu \in \bigcup_{N\in\mathbb{N}} \{\rho \in \mathcal{P}(S) : \|P^{(n)}\mu - P^{(n)}\rho\|^*_{\mathrm{BL}} \leq 1/m \text{ for all } n \geq N\}.$$

So $\nu \in D_\mu$.

Step 2. $P^{(n)}\mu \in \mathcal{D}_\mu$ for every $n \in \mathbb{N}$.

Let $\rho = \nu - P\nu$ for some $\nu \in \mathcal{M}^+(S)$. It is easy to see that $P^{(n)}\rho = \frac{1}{n}\rho - \frac{1}{n}P^n\rho$ for every $n \in \mathbb{N}$. Thus $\lim_{n\to\infty} \|P^{(n)}\rho\|^*_{\mathrm{BL}} = 0$.

Fix $n \in \mathbb{N}$. Then we can write $\mu - P^{(n)}\mu = \nu - P\nu$ where

$$\nu = \frac{n-1}{n}\mu + \frac{n-2}{n}P\mu + ... + \frac{1}{n}P^{n-2}\mu.$$

Now, $\nu \in \mathcal{M}^+(S)$, thus we know that

$$\lim_{m\to\infty} \|P^{(m)}\mu - P^{(m)}P^{(n)}\mu\|^*_{\mathrm{BL}} = \lim_{m\to\infty} \|P^{(m)}(\nu - P\nu)\|^*_{\mathrm{BL}} = 0$$

so $P^{(n)}\mu \in D_\mu$.

Step 3. $\mu \in \mathcal{P}^P_{cp}$.

Since $\mu \in \mathcal{P}^P_t$, there is a subsequence $(P^{(n_k)}\mu)_k$ that converges to a $\mu^* \in \mathcal{P}(S)$. Now, Step 1 and Step 2 imply that $\mu^* \in D_\mu$. Since P is Markov-Feller, μ^* is invariant, so $P^{(n)}\mu^* = \mu^*$ for every $n \in \mathbb{N}$. Thus

$$\|P^{(n)}\mu - \mu^*\|^*_{\mathrm{BL}} = \|P^{(n)}\mu - P^{(n)}\mu^*\|^*_{\mathrm{BL}} \to 0$$

as $n \to \infty$. □

By exploiting the relationship between the ergodic decomposition of a Markov semigroup and that of its resolvent (Theorem 6.2.7), we easily obtain the Markov semigroup version of Theorem 7.3.1 (see Theorem 7.3.13), which is also proven in [70, Lemma 1] (where the e-property instead of the Cesàro e-property is assumed to hold).

Because a regular Markov operator with the Cesàro e-property is Markov-Feller, $\Gamma^P_{cp} = \Gamma^P_{cpi}$ and $\mathcal{P}^P_{cp} = \mathcal{P}^P_{cpi}$. A direct consequence of Theorem 7.2.4, Theorem 7.3.1 and the statement above is as follows:

Corollary 7.3.2. *$\mathcal{P}^P_t = \mathcal{P}^P_{cp} = \mathcal{P}^P_{cpi}$ is closed in $\mathcal{S}_{\mathrm{BL}}$ and $\Gamma^P_t = \Gamma^P_{cp} = \mathcal{P}^P_{cpi}$ is closed in S.*

For $f \in C_b(S)$ we define

$$f^*(x) := \begin{cases} \langle \epsilon_x, f \rangle & \text{if } x \in \Gamma^P_{cp} \\ 0 & \text{if } x \notin \Gamma^P_{cp}. \end{cases}$$

In general, $f^* \in \mathrm{BM}(S)$. However, in our setting, when P is a regular Markov operator with the Cesàro e-property, one has:

Proposition 7.3.3. *The map $x \mapsto \epsilon_x$ is continuous from Γ^P_{cp} to $\mathcal{S}_{\mathrm{BL}}$ and for every $f \in C_b(S)$, $x \mapsto f^*(x)$ is continuous from Γ^P_{cp} to $\mathbb{R}$.*

Proof. Let $f \in \mathrm{BL}(S)$ and $x \in \Gamma^P_{cp}$. Let $\epsilon > 0$. Since P has the Cesàro e-property, there exists a $\delta > 0$, such that

$$|\langle P^{(n)}\delta_x - P^{(n)}\delta_y, f\rangle| < \epsilon$$

for every $n \in \mathbb{N}$ and $y \in B_x(\delta)$. Let $y \in B_x(\delta) \cap \Gamma^P_{cp}$, then we have

$$|\langle \epsilon_x - \epsilon_y, f\rangle| = \lim_{n\to\infty} |\langle P^{(n)}\delta_x - P^{(n)}\delta_y, f\rangle| \leq \epsilon.$$

Thus $x \mapsto \langle \epsilon_x, f\rangle$ is continuous. Now let $x_n \in \Gamma^P_{cp}$ such that $x_n \to x$ in S. Then $x \in \Gamma^P_{cp}$ according to Corollary 7.3.2, and thus $\langle \epsilon_{x_n}, f\rangle \to \langle \epsilon_x, f\rangle$ for every $f \in \mathrm{BL}(S)$. So $\epsilon_{x_n} \to \epsilon_x$ by Theorem 2.3.24, hence $x \mapsto \epsilon_x$ is continuous from Γ^P_{cp} to $\mathcal{S}_{\mathrm{BL}}$. The last statement follows from Lemma 2.3.12 and the fact that $f^*(x) = \langle \epsilon_x, f\rangle$. □

Remark. The conclusion of Corollary 7.3.2 need not hold for Markov-Feller operators that do not satisfy the Cesàro e-property. In [113] an example is given of a Markov-Feller operator P on a closed subset S of $\mathbb{R}^2$ (endowed with Euclidean metric) that has a unique invariant probability measure μ^* with $\mathrm{supp}(\mu^*) = S$. But there exists an $x \in S$ for which $(P^{(n)}\delta_x)_n$ is not tight, thus $x \notin \Gamma^P_t$. Hence we can conclude that Γ^P_t is not closed, because otherwise $\mathrm{supp}(\mu^*)$ would be contained in Γ^P_t, since $\mu^*(\Gamma^P_t) = 1$.

Theorem 7.3.4. *Γ^P_{cpie} is closed.*

Proof. By definition

$$\Gamma^P_{cpie} = \left\{ x \in \Gamma^P_{cpi} : \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) = 0 \text{ for every } f \in C_b(S) \right\}.$$

Let $x_n \in \Gamma^P_{cpie}$ such that $x_n \to x$ for some $x \in S$. Then $x \in \Gamma^P_{cp} = \Gamma^P_{cpi}$ by Corollary 7.3.2 and $\epsilon_{x_n} \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$ by Proposition 7.3.3.

Let $f \in C_b(S)$. We need to show that

$$\int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) = 0.$$

Since $\|f^*\|_\infty \leq \|f\|_\infty$, we have for every $y \in \Gamma^P_{cpi}$ and $n \in \mathbb{N}$

$$\begin{aligned} |(f^*(y) - f^*(x))^2 &- (f^*(y) - f^*(x_n))^2| \\ &= |[2f^*(y) - f^*(x) - f^*(x_n)] \cdot [f^*(x_n) - f^*(x)]| \\ &\leq 4\|f\|_\infty |f^*(x) - f^*(x_n)|. \end{aligned}$$

So

$$\begin{aligned} &\left| \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 - (f^*(y) - f^*(x_n))^2 \, d\epsilon_{x_n}(y) \right| \\ &\qquad \leq 4\|f\|_\infty \int_{\Gamma^P_{cpi}} |f^*(x) - f^*(x_n)| \, d\epsilon_{x_n}(y) \\ &\qquad = 4\|f\|_\infty |f^*(x) - f^*(x_n)| \to 0 \end{aligned}$$

as $n \to \infty$.

For every $n \in \mathbb{N}$

$$\begin{aligned} &\left| \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) - \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x_n))^2 \, d\epsilon_{x_n}(y) \right| \leq \\ &\left| \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d[\epsilon_x(y) - \epsilon_{x_n}(y)] \right| \\ &+ \left| \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 - (f^*(y) - f^*(x_n))^2 \, d\epsilon_{x_n}(y) \right|. \end{aligned} \tag{7.3}$$

The final term in inequality (7.3) above goes to zero as $n \to \infty$.

By Proposition 7.3.3 $y \mapsto (f^*(y) - f^*(x))^2$ is bounded and continuous from $\Gamma^P_{cp} = \Gamma^P_{cpi}$ to $\mathbb{R}$. By Corollary 7.3.2 Γ^P_{cpi} is closed, thus we can apply the Tietze Extension Theorem, and so there exists a $g \in C_b(S)$, such that $g(y) = (f^*(y) - f^*(x))^2$ for every $y \in \Gamma^P_{cpi}$. Since $\epsilon_x(\Gamma^P_{cpi}) = \epsilon_{x_n}(\Gamma^P_{cpi}) = 1$ for every $n \in \mathbb{N}$, we have

$$\left| \int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d[\epsilon_x(y) - \epsilon_{x_n}(y)] \right| = |\langle \epsilon_x, g \rangle - \langle \epsilon_{x_n}, g \rangle| \to 0$$

as $n \to \infty$ since $\epsilon_{x_n} \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$, $g \in C_b(S)$ and norm convergence in $\mathcal{S}^+_{\mathrm{BL}}$ coincides with weak convergence of measures.

Now note that $\int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x_n))^2 \, d\epsilon_{x_n}(y) = 0$ for every $n \in \mathbb{N}$, since $x_n \in \Gamma^P_{cpie}$, thus $\int_{\Gamma^P_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) = 0$ as well. So $x \in \Gamma^P_{cpie}$. □

Corollary 7.3.5. *For every $x \in \Gamma^P_{cpie}$, $[x]$ is closed.*

Proof. Let $x \in \Gamma^P_{cpie}$. Then $[x] = \{z \in \Gamma^P_{cpie} : \epsilon_z = \epsilon_x\}$. The map $h : \Gamma^P_{cpie} \to \mathcal{S}_{\mathrm{BL}}, z \mapsto \epsilon_z$ is continuous by Proposition 7.3.3, thus $[x]$ is closed in Γ^P_{cpie}, and hence in S as well by Theorem 7.3.4. □

Now we have shown that all the sets in our ergodic decomposition are closed. Our aim now is to show that the sets are also P-invariant. Theorem 7.3.1 and Theorem 7.2.4 imply:

Corollary 7.3.6. *For every $\mu \in \mathcal{P}^P_{cp} = \mathcal{P}^P_t$, $\mathrm{supp}(\mu) \subset \Gamma^P_{cp} = \Gamma^P_t$, so in particular $\mu(\Gamma^P_{cp}) = 1$.*

The following results follow directly from Proposition 5.5.1 and Corollary 7.3.6:

Corollary 7.3.7. *Let $\mu \in \mathcal{P}(S)$. Then $\mu \in \mathcal{P}^P_{cp}$ if and only if $\mu(\Gamma^P_{cp}) = 1$.*

Corollary 7.3.8. *Let $\mu \in \mathcal{P}^P_{cp}$, then*

$$\epsilon_\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x).$$

Corollary 7.3.9. *Γ^P_{cp} is a P-invariant set, i.e.*

$$P\delta_x(\Gamma^P_{cp}) = 1 \text{ for every } x \in \Gamma^P_{cp}.$$

Proof. Let $x \in \Gamma^P_{cp}$. Since

$$\lim_{n\to\infty} \|P^{(n)}P\delta_x - P^{(n)}\delta_x\|^*_{\mathrm{BL}} \to 0,$$

$P^{(n)}P\delta_x \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$. Thus $P\delta_x \in \mathcal{P}^P_{cp}$ and by Corollary 7.3.7, $P\delta_x(\Gamma^P_{cp}) = 1$. □

We define $\mathcal{P}^P_{cpie} := \{\mu \in \mathcal{P}^P_{cp} : \epsilon_\mu \text{ is ergodic.}\}$

Lemma 7.3.10. *Let μ be an invariant probability measure and $z \in \Gamma^P_{cpie}$ such that $\mu([z]) = 1$. Then $\mu = \epsilon_z$.*

Proof. By the integral decomposition of invariant measures into ergodic measures, we have

$$\mu = \int_{\Gamma^P_{cpie}} \epsilon_x \, d\mu(x) = \int_{[z]} \epsilon_x \, d\mu(x),$$

since $\mu([z]) = 1$. If $x \in [z]$, then $\epsilon_x = \epsilon_z$, so $\mu = \int_{[z]} \epsilon_z \, d\mu(x) = \epsilon_z$. □

Proposition 7.3.11. *Let $\mu \in \mathcal{P}(S)$. Then $\mu \in \mathcal{P}^P_{cpie}$ if and only if $\mu([z]) = 1$ for some $z \in \Gamma^P_{cpie}$. In that case $\epsilon_\mu = \epsilon_z$.*

Proof. Suppose that $\mu([z]) = 1$ for some $z \in \Gamma^P_{cpie}$, then $\mu(\Gamma^P_{cp}) = 1$, so $\mu \in \mathcal{P}^P_{cp}$ by Corollary 7.3.7, and $\epsilon_\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x)$ by Corollary 7.3.8. Since $\mu([z]) = 1$ we have

$$\epsilon_\mu = \int_{\Gamma^P_{cp}} \epsilon_x \, d\mu(x) = \int_{[z]} \epsilon_x \, d\mu(x) = \int_{[z]} \epsilon_z \, d\mu(x) = \epsilon_z,$$

so ϵ_μ is ergodic, thus $\mu \in \mathcal{P}^P_{cpie}$.

On the other hand, let $\mu \in \mathcal{P}^P_{cpie}$. Then ϵ_μ is ergodic, so there is a $z \in \Gamma^P_{cpie}$ such that $\epsilon_\mu = \epsilon_z$. Since $\mu \in \mathcal{P}^P_{cp}$, Corollary 7.3.8 implies

$$1 = \epsilon_z([z]) = \int_{\Gamma^P_{cp}} \epsilon_x([z]) \, d\mu(x),$$

thus $\epsilon_x([z]) = 1$ for μ-a.e. $x \in \Gamma^P_{cp}$. Since ϵ_x is an invariant probability measure for every $x \in \Gamma^P_{cp}$, Lemma 7.3.10 implies that $\epsilon_x = \epsilon_z$ for μ-a.e. $x \in \Gamma^P_{cp}$. Thus $x \in [z]$ for μ-a.e. $x \in \Gamma^P_{cp}$. Since $\mu(\Gamma^P_{cp}) = 1$, this implies that $\mu([z]) = 1$. □

Corollary 7.3.12. *For every $z \in \Gamma^P_{cpie}$, $[z]$ is a P-invariant set. Consequently Γ^P_{cpie} is P-invariant.*

Proof. Let $z \in \Gamma^P_{cpie}$ and $x \in [z]$. Because

$$\lim_{n\to\infty} \|P^{(n)} P\delta_x - P^{(n)}\delta_x\|^*_{\mathrm{BL}} \to 0,$$

$P^{(n)} P\delta_x \to \epsilon_x = \epsilon_z$, so $\epsilon_{P\delta_x} = \epsilon_z$ and thus Proposition 7.3.11 implies that $P\delta_x([z]) = 1$. Since Γ^P_{cpie} is a union of P-invariant sets, Γ^P_{cpie} is also P-invariant. □

Now let $(P(t))_{t\geq 0}$ be a jointly measurable Markov-Feller semigroup with the Cesàro e-property. By Proposition 7.2.6 the resolvent R also has the Cesàro e-property. By Theorem 6.2.7, Corollary 6.2.11 and the previous results we obtain:

Theorem 7.3.13. *Let $\mathbf{P} = (P(t))_{t\geq 0}$ be a jointly measurable Markov-Feller semigroup with the Cesàro e-property. Then the following holds:*

(i) $\mathcal{P}^{\mathbf{P}}_t = \mathcal{P}^{\mathbf{P}}_{cp} = \mathcal{P}^{\mathbf{P}}_{cpi}$ *is closed in* $\mathcal{S}_{\mathrm{BL}}$ *and* $\Gamma^{\mathbf{P}}_t = \Gamma^{\mathbf{P}}_{cp} = \Gamma^{\mathbf{P}}_{cpi}$ *is closed.*

(ii) $\Gamma^{\mathbf{P}}_{cpie}$ *is closed and* $[z]$ *is closed for every* $z \in \Gamma^{\mathbf{P}}_{cpie}$.

(iii) The map $x \mapsto \varepsilon_x$ *is continuous from* $\Gamma^{\mathbf{P}}_{cp}$ *to* $\mathcal{S}_{\mathrm{BL}}$.

(iv) For $\mu \in \mathcal{P}(S)$, $\mu \in \mathcal{P}^{\mathbf{P}}_{cp}$ *if and only if* $\mu(\Gamma^{\mathbf{P}}_{cp}) = 1$. *In this case* $\varepsilon_\mu = \int_{\Gamma^{\mathbf{P}}_{cp}} \varepsilon_x \, d\mu(x)$

(v) For $\mu \in \mathcal{P}(S)$, $\mu \in \mathcal{P}^{\mathbf{P}}_{cpie}$ *with* $\varepsilon_\mu = \varepsilon_z$ *if and only if* $\mu([z]) = 1$.

Corollary 7.3.14. $\Gamma^{\mathbf{P}}_{cp}$ *is a* $(P(t))_{t\geq 0}$*-invariant set, i.e.*

$$P(t)\delta_x(\Gamma^{\mathbf{P}}_{cp}) = 1 \text{ for every } x \in \Gamma^{\mathbf{P}}_{cp} \text{ and } t \in \mathbb{R}_+.$$

Proof. Let $x \in \Gamma_{cp}^{\mathbf{P}}$. Since

$$\lim_{s\to\infty} \|\mathbf{P}^{(s)}P(t)\delta_x - \mathbf{P}^{(s)}\delta_x\|_{\mathrm{BL}}^* \to 0$$

by Lemma 6.2.12, $\mathbf{P}^{(s)}P(t)\delta_x \to \varepsilon_x$. Thus $P(t)\delta_x \in \mathcal{P}_{cp}^{\mathbf{P}}$ and by Theorem 7.3.13 $P(t)\delta_x(\Gamma_{cp}^{\mathbf{P}}) = 1$. □

Since every ergodic measure is of the form ε_z for some $z \in \Gamma_{cpie}^{\mathbf{P}}$, we know that for every $\mu \in \mathcal{P}_{cpie}^{\mathbf{P}}$, $\varepsilon_\mu = \varepsilon_z$ for some $z \in \Gamma_{cpie}^{\mathbf{P}}$.

Corollary 7.3.15. *For every $z \in \Gamma_{cpie}^{\mathbf{P}}$, $[z]$ is a $(P(t))_{t\geq 0}$-invariant set. Consequently $\Gamma_{cpie}^{\mathbf{P}}$ is $(P(t))_{t\geq 0}$-invariant.*

Proof. Let $z \in \Gamma_{cpie}^{\mathbf{P}}$ and $x \in [z]$. Then since

$$\lim_{s\to\infty} \|\mathbf{P}^{(s)}P(t)\delta_x - P^{(s)}\delta_x\|_{\mathrm{BL}}^* \to 0$$

by Lemma 6.2.12, $\mathbf{P}^{(s)}P(t)\delta_x \to \varepsilon_x = \varepsilon_z$, so $\varepsilon_{P(t)\delta_x} = \varepsilon_z$ and thus Theorem 7.3.13 implies that $P(t)\delta_x([z]) = 1$. Since $\Gamma_{cpie}^{\mathbf{P}}$ is a union of $(P(t))_{t\geq 0}$-invariant sets, $\Gamma_{cpie}^{\mathbf{P}}$ is also $(P(t))_{t\geq 0}$-invariant. □

7.4 Existence, uniqueness and stability of invariant measures

Every Markov semigroup we consider in this section will be Markov-Feller and jointly measurable. We will find conditions on existence, uniqueness and stability of invariant measures of such semigroups with a (Cesàro) e-property. Our proofs can easily be adapted such that similar results hold for regular Markov operators with the (Cesàro) e-property as well.

In this section as well as in the next chapter we will occasionally make use of the Portmanteau Theorem (also known as the Alexandrov Theorem) on weak convergence of measures:

Theorem 7.4.1. [11, Theorem 2.1] *Let (X, d) be a metric space and $(\mu_n)_n$ and μ in $\mathcal{P}(X)$. Then the following statements are equivalent:*

(i) $\mu_n \to \mu$ *as* $n \to \infty$ *in the weak topology* $\sigma(\mathcal{M}(X), C_b(X))$.

(ii) $\liminf_{n\to\infty} \mu_n(U) \geq \mu(U)$ *for all open* $U \subset X$.

(iii) $\limsup_{n\to\infty} \mu_n(C) \leq \mu(C)$ *for all closed* $C \subset X$.

7.4.1 Existence of invariant measures

The proof of the following proposition is inspired by that of [81, Theorem 3.1], though simplified by exploiting our equicontinuity results. It will help us in giving equivalent conditions for the existence of invariant measures.

Proposition 7.4.2. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the Cesàro e-property. If $z \notin \Gamma_t^{\boldsymbol{P}}$, then there is a $\delta > 0$ such that*

$$\lim_{t\to\infty} \mathbf{P}^{(t)}\nu(B_z(\delta)) = 0 \textit{ for all } \nu \in \mathcal{P}(S). \tag{7.4}$$

Proof. Let $z \notin \Gamma_t^{\mathbf{P}}$. We will first show that there exist $\epsilon > 0$, compact sets $K_n \subset S$ and $t_n \in \mathbb{R}_+$ for $n \in \mathbb{N}$, such that $t_n \geq n$,

$$\mathbf{P}^{(t_n)}\delta_z(K_n) > 2\epsilon \text{ for every } n \in \mathbb{N}$$

and

$$\min\{d(x,y) : x \in K_m, y \in K_n\} \geq \epsilon \text{ whenever } m \neq n.$$

Since $z \notin \Gamma_t^{\mathbf{P}}$, $\{\mathbf{P}^{(t)}\delta_z : t \geq N\}$ is not tight for every $N \in \mathbb{N}$. So there exists an $0 < \epsilon < \frac{1}{2}$ such that for every $N \in \mathbb{N}$ and every compact $K \subset S$, there is a $t \geq N$, such that

$$\mathbf{P}^{(t)}\delta_z(K^\epsilon) \leq \mathbf{P}^{(t)}\delta_z(K^{2\epsilon}) < 1 - 2\epsilon.$$

Let $t_1 = 1$ and $K_1 \subset S$ be compact such that $\mathbf{P}^{(t_1)}\delta_z(K_1) > 2\epsilon$. Then there is a $t_2 \geq 2$ such that $\mathbf{P}^{(t_2)}\delta_z(K_1^\epsilon) < 1 - 2\epsilon$. So there is a compact $K_2 \subset S\backslash K_1^\epsilon$ such that $\mathbf{P}^{(t_2)}\delta_z(K_2) > 2\epsilon$. And since $K_1 \cup K_2$ is compact, there is a $t_3 \geq 3$ such that $\mathbf{P}^{(t_3)}\delta_z((K_1 \cup K_2)^\epsilon) < 1 - 2\epsilon$. Continuing inductively we find the required K_n and t_n.

By the Cesàro e-property and Theorem 7.2.2, $(\mathbf{P}^{(t)}\delta)_{t\geq 1}$ is equicontinuous in $z \in S$, so it follows that there exists a $\delta > 0$ such that

$$\|\mathbf{P}^{(t)}\delta_z - \mathbf{P}^{(t)}\delta_y\|_{\mathrm{BL}}^* < \frac{\epsilon}{1 + 3/\epsilon}$$

for every $t \geq 1$ and $y \in B_z(\delta)$. Let $f_n := f_{K_n}^{\epsilon/3}$. Then $f_n \in \mathrm{BL}(S)$ with $\|f_n\|_{\mathrm{BL}} \leq 1 + 3/\epsilon$. Therefore

$$|\langle \mathbf{P}^{(t_n)}\delta_z - \mathbf{P}^{(t_n)}\delta_y, f_n\rangle| < \epsilon \text{ for every } n \in \mathbb{N}, y \in B_z(\delta).$$

Since $\mathbb{1}_{K_n} \leq f_n \leq \mathbb{1}_{(K_n)^{\epsilon/3}}$, we have for every $n \in \mathbb{N}$ and $y \in B_z(\delta)$:

$$\begin{aligned}
\mathbf{P}^{(t_n)}\delta_y((K_n)^{\epsilon/3}) &\geq \langle \mathbf{P}^{(t_n)}\delta_y, f_n\rangle \\
&\geq \langle \mathbf{P}^{(t_n)}\delta_z, f_n\rangle - |\langle \mathbf{P}^{(t_n)}\delta_z - \mathbf{P}^{(t_n)}\delta_y, f_n\rangle| \\
&\geq \mathbf{P}^{(t_n)}\delta_z(K_n) - |\langle \mathbf{P}^{(t_n)}\delta_z - \mathbf{P}^{(t_n)}\delta_y, f_n\rangle| > \epsilon.
\end{aligned}$$

We proceed by contradiction: assume that $\nu \in \mathcal{P}(S)$ is such that

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(B_z(\delta)) =: 2\alpha > 0.$$

Then there is a sequence $(r_k)_k \subset \mathbb{R}_+$ such that $r_k \to \infty$ and

$$\lim_{k\to\infty} \mathbf{P}^{(r_k)}\nu(B_z(\delta)) = 2\alpha.$$

Now, for all $t \geq 0$ and $n \in \mathbb{N}$ we have

$$\begin{aligned} P(t)\mathbf{P}^{(t_n)}\nu(K_n^{\epsilon/3}) &= \int_S \mathbf{P}^{(t_n)}\delta_y(K_n^{\epsilon/3})\, dP(t)\nu(y) \\ &\geq \int_{B_z(\delta)} \mathbf{P}^{(t_n)}\delta_y(K_n^{\epsilon/3})\, dP(t)\nu(y) \geq \epsilon P(t)\nu(B_z(\delta)). \end{aligned}$$

By Lemma 6.2.13,

$$\liminf_{k\to\infty} \mathbf{P}^{(r_k)}\nu(K_n^{\epsilon/3}) = \liminf_{k\to\infty} \mathbf{P}^{(r_k)}\mathbf{P}^{(t_n)}\nu(K_n^{\epsilon/3})$$

for all $n \in \mathbb{N}$. Now let $n \in \mathbb{N}$. Then

$$\begin{aligned} \liminf_{k\to\infty} \mathbf{P}^{(r_k)}\mathbf{P}^{(t_n)}\nu(K_n^{\epsilon/3}) &= \liminf_{k\to\infty} \frac{1}{r_k}\int_0^{r_k} P(t)\mathbf{P}^{(t_n)}\nu(K_n^{\epsilon/3})\, dt \\ &\geq \epsilon \liminf_{k\to\infty} \mathbf{P}^{(r_k)}\nu(B_z(\delta)) = 2\alpha\epsilon. \end{aligned}$$

Since $K_n^{\epsilon/3} \cap K_m^{\epsilon/3} = \emptyset$ whenever $n \neq m$,

$$\liminf_{k\to\infty} \mathbf{P}^{(r_k)}\nu(\bigcup_{n=1}^{N} K_n^{\epsilon/3}) \geq 2N\alpha\epsilon$$

for every $N \in \mathbb{N}$, which is not possible. So

$$\lim_{t\to\infty} \boldsymbol{P}^{(t)}\nu(B_z(\delta)) = 0$$

for all $\nu \in \mathcal{P}(S)$. □

Proposition 7.4.2 implies that if there is a $z \in S$ such that for all $\delta > 0$ there is a $\nu \in \mathcal{P}(S)$ for which

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(B_z(\delta)) > 0,$$

then $z \in \Gamma_t^{\mathbf{P}}$. Note that this is not a necessary condition for z to be in $\Gamma_t^{\mathbf{P}}$:

Example. Let $S = [0,1]$ with Euclidean metric and define

$$P(t)\delta_x = \delta_{(x-t)\vee 0} \text{ for all } x \in S, t \in \mathbb{R}_+.$$

Then $\mathbf{P} = (P(t))_{t\geq 0}$ defines a Markov semigroup with the e-property, and clearly $\Gamma_t^{\mathbf{P}} = S$. Since $\lim_{t\to\infty} P(t)\nu = \delta_0$ for all $\nu \in \mathcal{P}(S)$, the Portmanteau Theorem implies that for any $0 < \delta < z$, $\limsup_{t\to\infty} P(t)\nu(\overline{B_z(\delta)}) \leq \delta_0(\overline{B_z(\delta)}) = 0$.

We also obtain the following result from Proposition 7.4.2.

Corollary 7.4.3. *Let $K \subset S$ be non-empty and compact, and such that $K \cap \Gamma_t^{\mathbf{P}} = \emptyset$. Then there is a $\delta > 0$ such that*

$$\lim_{t\to\infty} \mathbf{P}^{(t)}\nu(K^\delta) = 0 \text{ for all } \nu \in \mathcal{P}(S).$$

Proof. For each $z \in K$ let $r_z > 0$ be such that (7.4) holds. Then there exist $z_1, ..., z_n \in K$ such that $K \subset \cup_{i=1}^n B_{z_i}(r_{z_i}/2)$. Let $\delta = \min_{1\le i\le n} r_{z_i}/2$, then $\delta > 0$ and $K^\delta \subset \cup_{i=1}^n B_{z_i}(r_{z_i})$. Thus

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(K^\delta) \le \sum_{i=1}^n \limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(B_{z_i}(r_{z_i})) = 0.$$

□

Using these results we obtain necessary and sufficient conditions for the existence of invariant measures in the upcoming theorem. The statement *(v)⇒(i)* is slightly more general than that of [81, Theorem 3.1], since there ν has to be of the form δ_x for some $x \in S$. We also do not require the Markov semigroup to be strongly stochastically continuous at zero: joint measurability is sufficient. Finally, we only require the Cesàro e-property, instead of the e-property.

Theorem 7.4.4. *Let $(P(t))_{t\ge 0}$ be a Markov semigroup with the Cesàro e-property. The following results are equivalent:*

(i) $\Gamma_t^{\mathbf{P}}$ *is not empty.*

(ii) There exist invariant measures.

(iii) There exist $x, z \in S$ such that for every $\delta > 0$,

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\delta)) > 0.$$

(iv) There exists $z \in S$ such that for every $\delta > 0$ there is a $\nu \in \mathcal{P}(S)$ such that

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(B_z(\delta)) > 0.$$

(v) There exists a non-empty compact $K \subset S$ such that for every open set U with $K \subset U$ there is a $\nu \in \mathcal{P}(S)$ such that

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\nu(U) > 0.$$

Proof. (i) ⇒ (ii): By Theorem 7.3.13 $\Gamma_t^{\mathbf{P}} = \Gamma_{cp}^{\mathbf{P}} = \Gamma_{cpi}^{\mathbf{P}}$. So there exist invariant measures.
(ii) ⇒ (iii): If there exist invariant measures, then there exist ergodic measures.

Let μ be an ergodic measure and $z \in \operatorname{supp}(\mu)$. Now, $\mu = \varepsilon_x$ for some $x \in \Gamma^{\mathbf{P}}_{cpie}$, so $\mathbf{P}^{(t)}\delta_x \to \varepsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$ as $t \to \infty$. This implies, by Lemma 2.3.12 and the Portmanteau Theorem, that for every $\delta > 0$,

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\delta)) \geq \varepsilon_x(B_z(\delta)) > 0.$$

(iii)$\Rightarrow$ *(iv)*: Take $\nu = \delta_x$.
(iv)$\Rightarrow$ *(v)*: Take $K = \{z\}$.
(v)$\Rightarrow$ *(i)*: Corollary 7.4.3 implies that $K \cap \Gamma^{\mathbf{P}}_t \neq \emptyset$. In particular, $\Gamma^{\mathbf{P}}_t \neq \emptyset$. □

7.4.2 Uniqueness of invariant measures

The following notion is defined by Szarek in [113]: A Markov operator P *overlaps support* if for every $x, y \in S$ there is an $n_0 \in \mathbb{N}$, such that

$$\operatorname{supp}(P^{n_0}\delta_x) \cap \operatorname{supp}(P^{n_0}\delta_y) \neq \emptyset.$$

The main result in [113] is the following:

Theorem 7.4.5. *Let P be a regular Markov operator satisfying the e-property and such that P overlaps support. Then P has at most one invariant probability measure.*

The condition of overlapping support is far from necessary for uniqueness of invariant measures. We give a simple example to show this:

Example. Let $S = \{0, 1\}$ with $d(0, 1) = 1$, and $\Phi : S \to S$ defined by $\Phi(0) := 1$ and $\Phi(1) := 0$. Then $P_\Phi\mu := \mu \circ \Phi^{-1}$ defines a regular Markov operator satisfying the e-property. P_Φ has a unique invariant probability measure $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$, but clearly does not overlap supports, since $\operatorname{supp}(P^n_\Phi\delta_0) \cap \operatorname{supp}(P^n_\Phi\delta_1) = \emptyset$ for every $n \in \mathbb{N}$.

We show that a more general condition suffices for uniqueness of invariant probability measures, and the previous example satisfies *this* condition:

Theorem 7.4.6. *Let P be a regular Markov operator with the Cesàro e-property. Suppose that for every $x, y \in S$ there exists a $z \in S$, such that for every $\delta > 0$ there are $n_1, n_2 \in \mathbb{N}$ such that*

$$P^{n_1}\delta_x(B_z(\delta)) > 0 \text{ and } P^{n_2}\delta_y(B_z(\delta)) > 0. \tag{7.5}$$

Then there is at most one invariant probability measure.

Proof. Suppose there are at least two invariant probability measures. Theorem 5.4.11 implies that there are at least two distinct ergodic measures ϵ_x and ϵ_y, where $x, y \in \Gamma^P_{cpie}$ such that $[x] \neq [y]$. Let $z \in S$ be given by the assumption. Since $[x]$ and $[y]$ are disjoint, z cannot be in both. Say $z \notin [x]$. Since $[x]$ is closed by Corollary 7.3.5 there is a $\delta > 0$ such that $B_z(\delta) \cap [x] = \emptyset$. By Corollary 7.3.15 $[x]$ is an invariant set, so $P^n\delta_x([x]) = 1$ for every $n \in \mathbb{N}$, which implies that $P^n\delta_x(B_z(\delta)) = 0$ for every $n \in \mathbb{N}$, contradicting (7.5). □

By a similar proof, but now using Theorem 7.3.13 and Corollary 7.3.15, we obtain the Markov semigroup-variant of Theorem 7.4.6:

Theorem 7.4.7. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the Cesàro e-property. Suppose that for every $x, y \in S$ there exists a $z \in S$, such that for every $\delta > 0$ there are $t_1, t_2 > 0$ such that*

$$P(t_1)\delta_x(B_z(\delta)) > 0 \text{ and } P(t_2)\delta_y(B_z(\delta)) > 0. \tag{7.6}$$

Then there is at most one invariant probability measure.

We recently discovered that, using different techniques, Theorem 7.4.7 has also been obtained in [70, Theorem 2]. However, in [70] strong stochastic continuity at zero and the e-property are required, while joint measurability and the Cesàro e-property suffice for our proof. The condition required in Theorem 7.4.7 is far more general than the notion of irreducibility, as defined in e.g. [20], where it is needed to prove uniqueness of invariant measures.

The following result comes from [73, Theorem 1] (translated into our notation).

Theorem 7.4.8. *Let$(P(t))_{t\geq 0}$ be a Markov semigroup with the e-property. If there is a $z \in S$ such that for every $x \in S$ and every $\delta > 0$*

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\delta)) > 0,$$

then there exists a unique invariant probability measure μ^* and $\mathbf{P}^{(t)}\nu \to \mu^*$ for every $\nu \in \mathcal{P}(S)$ such that $\operatorname{supp}(\nu) \subset \Gamma_t^{\mathbf{P}}$.

The second part of Theorem 7.4.8, that $\mathbf{P}^{(t)}\nu \to \mu^*$ for every $\nu \in \mathcal{P}(S)$ such that $\operatorname{supp}(\nu) \subset \Gamma_t^{\mathbf{P}}$, *always* holds whenever the Markov semigroup has the Cesàro e-property and a unique invariant measure:

Proposition 7.4.9. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the Cesàro e-property, and with a unique invariant probability measure μ^*. Then the following are equivalent for $\nu \in \mathcal{P}(S)$:*

(i) $\mathbf{P}^{(t)}\nu \to \mu^*$ *in* $\mathcal{S}_{\mathrm{BL}}$ *as* $t \to \infty$.

(ii) $\nu(\Gamma_t^{\mathbf{P}}) = 1$.

(iii) $\operatorname{supp}(\nu) \subset \Gamma_t^{\mathbf{P}}$

Proof. Let $\nu \in \mathcal{P}(S)$. Since μ^* is the only invariant probability measure, $\lim_{t\to\infty} \mathbf{P}^{(t)}\nu = \mu^*$ in $\mathcal{S}_{\mathrm{BL}}$ if and only if $\nu \in \mathcal{P}_{cp}^{\mathbf{P}}$. By Theorem 7.3.13 this holds if and only if $\nu(\Gamma_t^{\mathbf{P}}) = 1$ if and only if $\operatorname{supp}(\nu) \subset \Gamma_t^{\mathbf{P}}$. □

We generalise Theorem 7.4.8 by showing that we can replace *limes inferior* with *limes superior* and the e-property with the Cesàro e-property:

Theorem 7.4.10. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the Cesàro e-property. If there is a $z \in S$ such that for every $x \in S$ and every $\delta > 0$*

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\delta)) > 0, \tag{7.7}$$

then there exists a unique invariant probability measure μ^*. Moreover, $z \in \operatorname{supp}(\mu^*)$ and $\mathbf{P}^{(t)}\nu \to \mu^*$ for every $\nu \in \mathcal{P}(S)$ such that $\nu(\Gamma_t^{\mathbf{P}}) = 1$.

Proof. By Theorem 7.4.4 and Theorem 7.4.7 there exists a unique invariant measure μ^*.

Suppose $z \notin \operatorname{supp}(\mu^*)$, then there is a $\delta > 0$ such that $\mu^*(B_z(\delta)) = 0$. Let $f := f^{\delta/2}_{B_z(\delta/2)}$, then $\mathbb{1}_{B_z(\delta/2)} \leq f \leq \mathbb{1}_{B_z(\delta)}$. We showed above that $\mathcal{P}_t^{\mathbf{P}} \neq \emptyset$. Note that $\Gamma_t^{\mathbf{P}} = \Gamma_{cp}^{\mathbf{P}}$ is non-empty, because there exists an invariant measure. Let $x \in \Gamma_{cp}^{\mathbf{P}}$. Then $\mathbf{P}^{(t)}\delta_x \to \mu^*$ in $\mathcal{S}_{\mathrm{BL}}$, thus by the Portmanteau Theorem

$$\begin{aligned}\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\delta/2)) &\leq \limsup_{t\to\infty}\langle \mathbf{P}^{(t)}\delta_x, f\rangle \\ &= \langle \mu^*, f\rangle \leq \mu^*(B_z(\delta)) = 0,\end{aligned}$$

which contradicts (7.7). The final statement follows from Proposition 7.4.9. □

7.4.3 Stability of invariant measures

The proof of our next theorem is based on that of [114, Theorem 2]. We will show further on that our statement is actually a generalisation of [114, Theorem 2].

Theorem 7.4.11. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the eventual e-property. Assume that $(P(t))_{t\geq 0}$ has an ergodic measure μ^* and let $w \in \Gamma_{cpie}^{\mathbf{P}}$ be such that $\mu^* = \varepsilon_w$. Then the following statements are equivalent:*

(i) There is a $z \in \operatorname{supp}(\mu^*)$ and an $x \in S$, such that

$$\liminf_{t\to\infty} P(t)\delta_x(B_z(\delta)) > 0 \text{ for every } \delta > 0.$$

(ii) $\lim_{t\to\infty} P(t)\nu = \mu^*$ in $\mathcal{S}_{\mathrm{BL}}$ for every $\nu \in \mathcal{P}(S)$ such that $\nu([w]) = 1$.

(iii) For every $z \in \operatorname{supp}(\mu^*)$ and every $\nu \in \mathcal{P}(S)$ such that $\nu([w]) = 1$,

$$\liminf_{t\to\infty} P(t)\nu(B_z(\delta)) > 0 \text{ for every } \delta > 0.$$

Proof. (i) ⇒ (ii):

Step 1: For every $\delta > 0$ there is an $\alpha > 0$ such that

$$\liminf_{t\to\infty} P(t)\nu(B_z(\delta)) > \alpha$$

for every $\nu \in \mathcal{P}(S)$ *such that* $\nu([w]) = 1$.

Let $\delta > 0$. Define

$$\gamma := \frac{1}{2} \liminf_{t\to\infty} P(t)\delta_x(B_z(\delta/2)).$$

By assumption $\gamma > 0$. Let $f := f^{\delta/2}_{B_z(\delta/2)}$. Then $1\!\!1_{B_z(\delta/2)} \leq f \leq 1\!\!1_{B_z(\delta)}$ and $f \in \mathrm{BL}(S)$. By the eventual e-property, there exists an $\eta > 0$ and a $T > 0$ such that

$$|\langle P(t)\delta_y - P(t)\delta_x, f\rangle| < \gamma$$

for every $y \in B_x(\eta)$ and $t \geq T$. Thus for all $y \in B_x(\eta)$ and $t \geq T$,

$$\begin{aligned} P(t)\delta_y(B_z(\delta)) &\geq \langle P(t)\delta_y, f\rangle \\ &\geq \langle P(t)\delta_x, f\rangle - |\langle P(t)\delta_y - P(t)\delta_x, f\rangle| \\ &> P(t)\delta_x(B_z(\delta/2)) - \gamma. \end{aligned}$$

Hence

$$\liminf_{t\to\infty} P(t)\delta_y(B_z(\delta)) \geq \liminf_{s\to\infty} P(s)\delta_x(B_z(\delta/2)) - \gamma = \gamma$$

for every $y \in B_x(\eta)$. Set $\theta := \frac{\mu^*(B_z(\eta))}{2}$, then $\theta > 0$ since $z \in \mathrm{supp}(\mu^*)$.

Fix $\nu \in \mathcal{P}(S)$ such that $\nu([w]) = 1$. Then Theorem 7.3.13 implies that

$$\lim_{t\to\infty} \mathbf{P}^{(t)}\nu \to \varepsilon_w = \mu^*$$

in $\mathcal{S}_{\mathrm{BL}}$. Thus by Lemma 2.3.12 and the Portmanteau Theorem

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\nu(B_z(\eta)) \geq \mu^*(B_z(\eta)).$$

Then there exists a $t_0 > 0$ such that $P(t_0)\nu(B_z(\eta)) > \theta$. Thus by Fatou's Lemma

$$\begin{aligned} \liminf_{t\to\infty} P(t)\nu(B_z(\delta)) &= \liminf_{t\to\infty} P(t+t_0)\nu(B_z(\delta)) \\ &\geq \int_S \liminf_{t\to\infty} P(t)\delta_y(B_z(\delta))d[P(t_0)\nu](y) \\ &\geq \int_{B_z(\eta)} \liminf_{t\to\infty} P(t)\delta_y(B_z(\delta))d[P(t_0)\nu](y) > \gamma\theta, \end{aligned}$$

so we can take $\alpha = \gamma\theta$.

Step 2: For every $\nu_1, \nu_2 \in \mathcal{P}(S)$ *with* $\nu_1([w]) = \nu_2([w]) = 1$,

$$\lim_{t\to\infty} \|P(t)\nu_1 - P(t)\nu_2\|^*_{\mathrm{BL}} = 0.$$

Since $\mu^*([w]) = 1$, this will imply that $\lim_{t\to\infty} \|P(t)\nu - \mu^*\|^*_{\mathrm{BL}} = 0$, which proves this part of the theorem.

Fix $\epsilon > 0$. For $\delta > 0$ we define $\mathcal{P}^\delta(S) := \{\mu \in \mathcal{P}(S) : \mathrm{supp}(\mu) \subset B_z(\delta)\}$. Let $\mu_1, \mu_2 \in \mathcal{P}^\delta(S)$. If $f \in \mathrm{BL}(S)$ with $|f|_{\mathrm{Lip}} \leq \|f\|^*_{\mathrm{BL}} \leq 1$, then $|f_{\sup} - f_{\inf}| \leq 2\delta$, where

$$f_{\inf} := \inf\{f(y) : y \in B_z(\delta)\} \text{ and } f_{\sup} := \sup\{f(y) : y \in B_z(\delta)\}.$$

Now, $f_{\inf} \leq \langle \mu_i, f\rangle \leq f_{\sup}$ for $i \in \{1,2\}$, thus $|\langle \mu_1 - \mu_2, f\rangle| \leq 2\delta$. So $\|\mu_1 - \mu_2\|^*_{\mathrm{BL}} \leq 2\delta$. Since $(P(t))_{t\geq 0}$ satisfies the eventual e-property, there exists by Theorem 7.2.2 a $\delta > 0$ and a $T > 0$ such that $\|P(t)\mu_1 - P(t)\mu_2\|^*_{\mathrm{BL}} < \epsilon/2$ whenever $\mu_1, \mu_2 \in \mathcal{P}^\delta(S)$ and $t \geq T$.

By Step 1 there exists an $0 < \alpha < 1$, such that

$$\liminf_{t\to\infty} P(t)\nu(B_z(\delta)) > \alpha$$

for every $\nu \in \mathcal{P}^{\mathbf{P}}_t$. By induction we will define a sequence of $(t_n)_{n\geq 0} \subset \mathbb{R}_+$ and four sequences of probability measures $(\nu^n_1)_{n\geq 0}$, $(\nu^n_2)_{n\geq 0}$, $(\mu^n_1)_{n\geq 0}$, $(\mu^n_2)_{n\geq 0}$, such that $\nu^n_i \in \mathcal{P}^\delta(S), \mu^n_i([w]) = 1$ for every $i \in \{1,2\}, n \in \mathbb{N}$, and

$$P(t_n)\mu^{n-1}_i = \alpha\nu^n_i + (1-\alpha)\mu^n_i \text{ for every } i \in \{1,2\}, n \in \mathbb{N}. \tag{7.8}$$

Define $t_0 = 0$, $\nu^0_1 = \mu^0_1 = \nu_1$ and $\nu^0_2 = \mu^0_2 = \nu_2$. If $n \in \mathbb{N}$ and $t_{n-1}, \nu^{n-1}_1, \nu^{n-1}_2, \mu^{n-1}_1$, μ^{n-1}_2 are given, then since $\mu^{n-1}_i \in \mathcal{P}^{\mathbf{P}}_t$ for $i \in \{1,2\}$, we may choose $t_n > 0$ such that

$$P(t_n)\mu^{n-1}_i(B_z(\delta)) > \alpha \text{ for } i \in \{1,2\}.$$

For $E \subset S$ Borel, define

$$\nu^n_i(E) := \frac{P(t_n)\mu^{n-1}_i(E \cap B_z(\delta))}{P(t_n)\mu^{n-1}_i(B_z(\delta))} \text{ for } i \in \{1,2\},$$

then $\nu^n_i \in \mathcal{P}^\delta(S)$ for $i \in \{1,2\}$. Note that

$$\nu^n_i \leq \frac{P(t_n)\mu^{n-1}_i}{P(t_n)\mu^{n-1}_i(B_z(\delta))} < \frac{P(t_n)\mu^{n-1}_i}{\alpha}.$$

We also define

$$\mu^n_i(E) := \frac{1}{1-\alpha}(P(t_n)\mu^{n-1}_i(E) - \alpha\nu^n_i(E)) \text{ for } i \in \{1,2\}.$$

Then $\mu^n_i \in \mathcal{P}(S)$. Also, for $i \in \{1,2\}$, $\mu^n_i \leq \frac{1}{1-\alpha}P(t_n)\mu^{n-1}_i$. Since $\mu^{n-1}_i([w]) = 1$, $P(t_n)\mu^{n-1}_i([w]) = 1$ as well, since $[w]$ is $(P(t))_{t\geq 0}$-invariant by Corollary 7.3.15. So $P(t_n)\mu^{n-1}_i(S\backslash[w]) = 0$, and thus $\mu^n_i(S\backslash[w]) = 0$ as well. Hence $\mu^n_i([w]) = 1$ for $i \in \{1,2\}$.

Now choose $N \in \mathbb{N}$ such that $(1-\alpha)^N < \epsilon/4$. From (7.8) we obtain for every $i \in \{1,2\}$ and $t \in \mathbb{R}_+$:

$$\begin{aligned} P(t_1 + t_2 + ... + t_N)\nu_i &= \alpha P(t_2 + ... + t_N + t)\nu^1_i + \alpha(1-\alpha)P(t_3 + ... + t_N + t)\nu^2_i \\ &+ ... + \alpha(1-\alpha)^{N-1}P(t)\nu^k_i + (1-\alpha)^N P(t)\mu^N_i. \end{aligned}$$

Since $\nu_i^n \in \mathcal{P}^\delta(S)$ for every $i \in \{1,2\}, n \in \mathbb{N}$, we have $\|P(t)\nu_1^n - P(t)\nu_2^n\|_{\mathrm{BL}}^* \leq \epsilon/2$ for every $t \geq T$. Also,

$$\|(1-\alpha)^N(P(t)\mu_1^N - P(t)\mu_2^N)\|_{\mathrm{BL}}^* \leq 2(1-\alpha)^N < \epsilon/2.$$

Thus for every $t \geq t_1 + ... + t_N + T$ we have $\|P(t)\nu_1 - P(t)\nu_2\|_{\mathrm{BL}}^* \leq \epsilon$. Since $\epsilon > 0$ was arbitrary, we have proven Step 2, and thus *(ii)*.

(ii) $\Rightarrow$ *(iii)*: Let $z \in \mathrm{supp}(\mu^*)$ and $\nu([w]) = 1$. $P(t)\nu \to \mu^*$ as $t \to \infty$, so by the Portmanteau Theorem, $\liminf_{t\to\infty} P(t)\nu(B_z(\delta)) \geq \mu^*(B_z(\delta)) > 0$ for any $\delta > 0$.

(iii) $\Rightarrow$ *(i)*: By assumption, $[w]$ is non-empty, so take $x = w$, then $\delta_x([w]) = 1$ and the statement holds. □

If $(P(t))_{t\geq 0}$ has the eventual e-property and a unique invariant probability measure μ^*, then μ^* is ergodic and $\Gamma_t^{\mathbf{P}} = \Gamma_{cpie}^{\mathbf{P}} = [w]$, where $w \in \Gamma_{cpie}^{\mathbf{P}}$. So Theorem 7.4.11 and Theorem 7.3.13 imply the following:

Corollary 7.4.12. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the eventual e-property. Assume $(P(t))_{t\geq 0}$ has a unique invariant probability measure μ^*, then the following are equivalent:*

(i) There is a $z \in \mathrm{supp}(\mu^*)$ and an $x \in S$, such that

$$\liminf_{t\to\infty} P(t)\delta_x(B_z(\delta)) > 0 \text{ for every } \delta > 0.$$

(ii) $\lim_{t\to\infty} P(t)\nu = \mu^*$ in $\mathcal{S}_{\mathrm{BL}}$ for every $\nu \in \mathcal{P}_t^{\mathbf{P}}$.

(iii) For every $z \in \mathrm{supp}(\mu^*)$ and every $\nu \in \mathcal{P}_t^{\mathbf{P}}$

$$\liminf_{t\to\infty} P(t)\nu(B_z(\delta)) > 0 \text{ for every } \delta > 0.$$

Let $(P(t))_{t\geq 0}$ be a semigroup with the e-property. In [114, Theorem 2] it is shown that if $(P(t))_{t\geq 0}$ satisfies the condition of Theorem 7.4.8 and if additionally $\liminf_{t\to\infty} P(t)\delta_z(B_z(\delta)) > 0$ with z such as in Theorem 7.4.8, then $\lim_{t\to\infty} P(t)\nu = \mu^*$ in $\mathcal{S}_{\mathrm{BL}}$ for every $\nu \in \mathcal{P}(S)$ with $\nu(\Gamma_t^{\mathbf{P}}) = 1$, where μ^* is the unique invariant probability measure. As we have shown in our generalisation of Theorem 7.4.8, Theorem 7.4.10, the z from that theorem is in the support of the invariant measure. In particular, Corollary 7.4.12 is a generalisation of [114, Theorem 2], since it implies that $(P(t))_{t\geq 0}$ need not satisfy the condition of Theorem 7.4.8 (or the condition of its generalisation), as long as it has a unique invariant measure.

We also obtain the following result:

Corollary 7.4.13. *Let $(P(t))_{t\geq 0}$ be a Markov semigroup with the eventual e-property. Assume that for every ergodic measure μ^* there is a $z \in \mathrm{supp}(\mu^*)$ and an $x \in S$ such that*

$$\liminf_{t\to\infty} P(t)\delta_x(B_z(\delta)) > 0 \text{ for every } \delta > 0.$$

If $\nu \in \mathcal{P}(S)$ is such that $\nu(\Gamma_{cpie}^{\mathbf{P}}) = 1$, then $\lim_{t\to\infty} P(t)\nu = \varepsilon_\nu$ in $\mathcal{S}_{\mathrm{BL}}$.

Proof. Let $\nu \in \mathcal{P}(S)$ be such that $\nu(\Gamma^{\mathbf{P}}_{cpie}) = 1$. Then

$$P(t)\nu = \int_S P(t)\delta_x \, d\nu(x) = \int_{\Gamma^{\mathbf{P}}_{cpie}} P(t)\delta_x \, d\nu(x).$$

By Theorem 7.4.11, $\lim_{t\to\infty} P(t)\delta_x = \varepsilon_x$ for every $x \in \Gamma^{\mathbf{P}}_{cpie}$, so by the Dominated Convergence Theorem we get

$$\begin{aligned} \lim_{t\to\infty} P(t)\nu &= \int_{\Gamma^{\mathbf{P}}_{cpie}} \lim_{t\to\infty} P(t)\delta_x \, d\nu(x) \\ &= \int_{\Gamma^{\mathbf{P}}_{cp}} \varepsilon_x \, d\nu(x) = \varepsilon_\nu. \end{aligned}$$

□

7.5 Notes

This chapter is based on the submitted paper [127], with some minor modifications.

The notion of equicontinuity of Markov operators and semigroups and its implications on existence, uniqueness and stability of invariant measures has been well-studied in various forms, where equicontinuity of iterates of the dual operator is demanded on various subsets of the space of bounded continuous functions. Among one of the first is Rosenblatt's paper on equicontinuous Markov operators on compact Hausdorff spaces [105]. See Meyn and Tweedie's monograph [86], where, among other things, so-called e-chains on locally compact separable metric spaces are studied. These are also considered in [132]. The e-property of Markov operators and semigroups on Polish spaces that we consider in this chapter has also been studied by Szarek and coworkers [70, 73, 81, 113, 114]. Our results on existence, uniqueness and stability of invariant measures extend some of the results from [73, 81, 113, 114].

One area of possible applications of the results in this chapter (and Chapter 8) lies in the realm of stochastic differential equations on separable Hilbert spaces or Banach spaces. Under certain conditions, one can associate regular jointly measurable Markov semigroups to solutions of stochastic differential equations, and investigate existence, uniqueness and stability of invariant measures by studying this semigroup. See e.g. [20] for an introduction into this topic. In [37, 70, 73, 81, 114] stochastic differential equations are studied for which the associated Markov semigroup has the e-property. Previously mentioned results by Szarek and coworkers on such semigroups are exploited to obtain results for existence, uniqueness and stability of invariant measures.

CHAPTER

EIGHT

SET OF ERGODIC MEASURES OF MARKOV SEMIGROUPS WITH THE CESÀRO E-PROPERTY

8.1 Introduction

In this chapter we are interested in determining ergodic measures as limits of Cesàro averages starting from some compact set. A concentrating condition related to the one introduced in [81] (see also Theorem 7.4.4*(v)*) appears to be perfectly fitted to our task. Namely, we prove that then the number of ergodic measures is closely related to the behaviour of the Markov semigroup on this concentrating compact set. This allows us to provide a condition for the existence of finitely many ergodic measures. Similar problems for infinite dimensional systems were studied in [84]. We also find a sufficient condition for the existence of at most countably many ergodic invariant measures.

The chapter is organised as follows. In Section 8.2 we first show some new consequences of the (Cesàro) e-property on the set of ergodic measures. We show for instance that this set is closed in the weak topology, hence in $\mathcal{S}_{\mathrm{BL}}$. The remainder of the section is devoted to the study of conclusions we are able to draw from a weak concentrating condition (at some compact set). The main result, Theorem 8.2.9, says that the set of all ergodic measures is obtained as weak limits of Cesàro averages starting at some points from the given compact set. In fact, we find in this way a Borel subset of the compact set that maps bijectively to the set of ergodic measures. The condition assuring the existence of finitely or countably many ergodic measures is provided in Section 8.3. In Section 8.4, in turn, we show (Theorem 8.4.3) that two different conditions (related to our weak concentrating condition) on a Markov semigroup with the Cesàro e-property both ensure that for every probability mea-

sure the Cesàro weak limit exists and is an invariant measure. This theorem implies corollaries that give necessary and sufficient conditions for a Markov semigroup to be weakly mean ergodic and asymptotically stable.

Let (S, d) be a complete separable metric space and $\mathcal{S}_{\mathrm{BL}}$ the associated Banach space (see Chapter 2). In the sequel we will assume that $\mathbf{P} = (P(t))_{t\geq 0}$ is a jointly measurable Markov-Feller semigroup on S.

Remark. While all the results in this chapter are formulated for Markov semigroups, the proofs can easily be adapted to show the analogous results for Markov operators hold as well.

8.2 Weak concentrating condition

Recall the results from Theorem 7.3.13 on properties of the various sets in the Yosida-type ergodic decomposition and the map $x \mapsto \varepsilon_x$ for a Markov semigroup with the Cesàro e-property. In both this section as well as in Section 8.4 we will often use these results implicitly, without referring each time to Theorem 7.3.13.

As before, we will write $\mathcal{P}_{\mathrm{erg}}(S)$ and $\mathcal{P}_{\mathrm{inv}}(S)$ to denote the set of ergodic measures and invariant probability measures (with respect to $(P(t))_{t\geq 0}$) respectively.

Furthermore, we define

$$\Phi : \Gamma^{\mathbf{P}}_{cp} \to \mathcal{S}_{\mathrm{BL}} : x \mapsto \varepsilon_x$$

and recall that Φ is continuous and that $\Gamma^{\mathbf{P}}_{cp}$ and $\Gamma^{\mathbf{P}}_{cpie}$ are closed in S by Theorem 7.3.13, whenever $(P(t))_{t\geq 0}$ satisfies the Cesàro e-property.

We begin this section by giving some properties of the set of ergodic measures when the Cesàro e-property holds.

Proposition 8.2.1. *Suppose $(P(t))_{t\geq 0}$ has the Cesàro e-property. Then $\mathcal{P}_{\mathrm{erg}}(S)$ is closed in $\mathcal{S}_{\mathrm{BL}}$.*

Proof. An invariant probability measure μ is ergodic if and only if for every $f \in C_b(S)$, $\lim_{t\to\infty} \mathbf{U}^{(t)} f(x) = \langle \mu, f\rangle$ for μ-a.e. $x \in S$ (Theorem 6.3.2).

Let us define for $f \in C_b(S)$

$$f^*(x) := \begin{cases} \langle \varepsilon_x, f\rangle & \text{if } x \in \Gamma^{\mathbf{P}}_{cp} \\ 0 & \text{if } x \notin \Gamma^{\mathbf{P}}_{cp}. \end{cases}$$

For every $x \in \Gamma^{\mathbf{P}}_{cp}$, $U^{(t)} f(x) \to \langle \varepsilon_x, f\rangle = f^*(x)$ as $t \to \infty$. Since $\mu(\Gamma^{\mathbf{P}}_{cp}) = 1$ we obtain that μ is ergodic if and only if $f^* = \langle \mu, f\rangle$ μ-a.e. for every $f \in C_b(S)$, or equivalently

$$\int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 \, d\mu(x) = 0. \tag{8.1}$$

Now assume that $(\mu_n)_n$ is a sequence of ergodic measures such that $\mu_n \to \mu$ in $\mathcal{S}_{\mathrm{BL}}$. Then μ is invariant, since $(P(t))_{t\geq 0}$ is Markov-Feller. Fix $f \in C_b(S)$. We need to show that (8.1) holds. Since $\|f^*\|_\infty \leq \|f\|_\infty$, we have for every $x \in \Gamma^{\mathbf{P}}_{cp}$ and $n \in \mathbb{N}$

$$\begin{aligned} |(f^*(x) - \langle \mu, f\rangle)^2 &- (f^*(x) - \langle \mu_n, f\rangle)^2| \\ &= |2f^*(x) - \langle \mu, f\rangle - \langle \mu_n, f\rangle| \cdot |\langle \mu_n, f\rangle - \langle \mu, f\rangle| \\ &\leq 4\|f\|_\infty |\langle \mu - \mu_n, f\rangle|. \end{aligned}$$

So

$$\left| \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 - (f^*(x) - \langle \mu_n, f\rangle)^2 \, d\mu_n(x) \right| \leq 4\|f\|_\infty |\langle \mu - \mu_n, f\rangle| \to 0$$

as $n \to \infty$.

For every $n \in \mathbb{N}$,

$$\begin{aligned} &\left| \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 \, d\mu(x) - \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu_n, f\rangle)^2 \, d\mu_n(x) \right| \leq \\ &\left| \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 \, d[\mu(x) - \mu_n(x)] \right| \\ &+ \left| \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle))^2 - (f^*(y) - \langle \mu_n, f\rangle)^2 \, d\mu_n(x) \right|. \end{aligned} \tag{8.2}$$

We showed above that the final term in inequality (8.2) goes to zero as $n \to \infty$.

Since $x \mapsto \varepsilon_x$ is continuous from $\Gamma^{\mathbf{P}}_{cp}$ to $\mathcal{S}_{\mathrm{BL}}$, we also know that $x \mapsto (f^*(x) - \langle \mu, f\rangle)^2$ is bounded and continuous from $\Gamma^{\mathbf{P}}_{cp}$ to $\mathbb{R}$. $\Gamma^{\mathbf{P}}_{cp}$ is closed, thus we can apply the Tietze Extension Theorem, and so there exists a $g \in C_b(S)$, such that $g(x) = (f^*(x) - \langle \mu, f\rangle)^2$ for every $x \in \Gamma^{\mathbf{P}}_{cp}$. Since $\mu(\Gamma^{\mathbf{P}}_{cp}) = \mu_n(\Gamma^{\mathbf{P}}_{cp}) = 1$ for every $n \in \mathbb{N}$, we have

$$\left| \int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 \, d[\mu(x) - \mu_n(x)] \right| = |\langle \mu, g\rangle - \langle \mu_n, g\rangle| \to 0$$

as $n \to \infty$ since $\mu_n \to \mu$ in $\mathcal{S}_{\mathrm{BL}}$ and $g \in C_b(S)$.

Now note that $\int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu_n, f\rangle)^2 \, d\mu_n(x) = 0$ for every $n \in \mathbb{N}$, since the measures μ_n are ergodic, thus $\int_{\Gamma^{\mathbf{P}}_{cp}} (f^*(x) - \langle \mu, f\rangle)^2 \, d\mu(x) = 0$ as well. Thus μ is ergodic. $\square$

As mentioned before, the proof can easily be adapted to show a similar statement holds for Markov operators. If a Markov operator or semigroup is merely Markov-Feller, then the set of invariant probability measures is closed in $\mathcal{S}_{\mathrm{BL}}$. However, the Markov-Feller property is not sufficient to obtain closedness of the set of ergodic measures, as the following example shows. This indicates that the Cesàro e-property is a crucial assumption in Proposition 8.2.1.

Example 8.2.2. Let $S = \{0,1\}^{\mathbb{N}}$ be endowed with the product topology. Then S is compact and separable, and can be metrised using the following metric:

$$d(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} |x_k - y_k|.$$

Since S is compact, it is complete. We define the map $\Phi : S \to S$ as follows: $\Phi(x_1, x_2, \cdots) = (x_2, x_3, \ldots)$. Then Φ is continuous, and thus defines a Markov-Feller operator P_Φ on S. We will show that the set of ergodic measures for this operator is not closed. For $n \in \mathbb{N}$ we define $z_n \in S$ consisting of n consecutive ones followed by n consecutive zeroes, and this pattern repeated indefinitely. Then P_Φ acts periodically on δ_{z_n}, with period $2n$. From this we define an invariant probability measure by

$$\mu_n := \frac{1}{2n} \sum_{k=0}^{2n-1} P_\Phi^k \delta_{z_n}.$$

Because μ_n is an invariant measure supported on a single periodic orbit, it is ergodic. Furthermore, $\mu_n \to \mu$ in weak topology, where

$$\mu = \frac{1}{2}\delta_{(0,0,0,\cdots)} + \frac{1}{2}\delta_{(1,1,1,\cdots)}.$$

μ is the convex combination of two different ergodic measures, hence not ergodic itself.

Proposition 8.2.3. *If $\mathcal{P}_{\text{inv}}(S)$ is non-empty, then there exists an invariant probability measure μ_0 with*

$$\operatorname{supp}(\mu_0) = \overline{\bigcup_{\mu \in \mathcal{P}_{\text{inv}}(S)} \operatorname{supp}(\mu)}.$$

Hence $\bigcup_{\mu \in \mathcal{P}_{\text{inv}}(S)} \operatorname{supp}(\mu)$ is closed.

Proof. Let $D = \bigcup_{\mu \in \mathcal{P}_{\text{inv}}(S)} \operatorname{supp}(\mu)$. Then D is separable, so there exist $(x_n)_n \subset D$ such that $D \subset \overline{\{x_n : n \in \mathbb{N}\}}$. Let $\mu_n \in \mathcal{P}_{\text{inv}}(S)$ be such that $x_n \in \operatorname{supp}(\mu_n)$ and define

$$\mu_0 = \sum_{n=1}^{\infty} 2^{-n} \mu_n,$$

then $\mu_0 \in \mathcal{P}_{\text{inv}}(S)$ and

$$\bigcup_{n=1}^{\infty} \operatorname{supp}(\mu_n) \subset \operatorname{supp}(\mu_0),$$

thus

$$D = \overline{\bigcup_{n=1}^{\infty} \operatorname{supp}(\mu_n)} = \operatorname{supp}(\mu_0).$$

□

We can also ask ourselves what we can say about the union of the supports of all ergodic measures. Even with the e-property, this set need not be closed, as the following example shows:

Example. Let $S = [0,1]$ with Euclidean metric. For $x \in S$ and $t \in \mathbb{R}_+$ define

$$P(t)\delta_x = [x + e^{-t}(1-x)]\delta_x + [(1-x) - e^{-t}(1-x)]\delta_{1-x}.$$

Then $P(0)\delta_x = \delta_x$, and easy calculation shows that $P(t)P(s)\delta_x = P(t+s)\delta_x$ for all $x \in S$ and $s, t \in \mathbb{R}_+$. For every $E \subset S$ Borel,

$$P(t)\delta_x(E) = [x + e^{-t}(1-x)]\mathbb{1}_E(x) + [(1-x) - e^{-t}(1-x)]\mathbb{1}_E(1-x)$$

so $(t,x) \mapsto P(t)\delta_x(E)$ is jointly measurable. Thus we can define a jointly measurable Markov semigroup on S as follows:

$$P(t)\mu = \int_S P(t)\delta_x \, d\mu(x) \text{ for all } \mu \in \mathcal{M}^+(S).$$

It can be shown that $(P(t))_{t\geq 0}$ is Markov-Feller and satisfies the e-property. Now, for all $x \in S$, $P(t)\delta_x \to x\delta_x + (1-x)\delta_{1-x} = \varepsilon_x$ as $t \to \infty$. Because these measures cannot be written as the convex combination of different invariant probability measures, these are ergodic measures. So $\Gamma^{\mathbf{P}}_{cpie} = S$, and thus each ergodic measure equals $x\delta_x + (1-x)\delta_{1-x}$ for some $x \in S$. Now, for all $0 < x < 1$, $\operatorname{supp}(\varepsilon_x) = \{x, 1-x\}$, and $\operatorname{supp}(\varepsilon_0) = \operatorname{supp}(\varepsilon_1) = \{1\}$. Thus

$$\bigcup_{x\in S} \operatorname{supp}(\varepsilon_x) = (0,1],$$

which is open but not closed in S.

We show that in general the Cesàro e-property implies that union of the supports of ergodic measures is a G_δ subset of S, i.e. a countable intersection of open sets.

Theorem 8.2.4. *Let $(P(t))_{t\geq 0}$ be a Markov-Feller semigroup with the Cesàro e-property. Then*

$$D := \bigcup_{\mu \in \mathcal{P}_{\mathrm{erg}}(S)} \operatorname{supp}(\mu)$$

is a G_δ set. In particular, D is a Polish space in its relative topology.

Proof. If $x \in \operatorname{supp}(\mu)$ for an ergodic measure μ, then $x \in \Gamma^{\mathbf{P}}_{cpie}$, and $\operatorname{supp}(\mu) \subset [x]$, so $\mu = \varepsilon_x$. So we can write

$$D = \{x \in \Gamma^{\mathbf{P}}_{cpie} : x \in \operatorname{supp}(\varepsilon_x)\} = \bigcap_{k\in\mathbb{N}} D_k,$$

where

$$D_k = \{x \in \Gamma^{\mathbf{P}}_{cpie} : \varepsilon_x(B_x(1/k)) > 0\}.$$

Let $E_k := \Gamma^{\mathbf{P}}_{cpie} \backslash D_k = \{x \in \Gamma^{\mathbf{P}}_{cpie} : \varepsilon_x(B_x(1/k)) = 0\}$. We will show that E_k is closed. Let $x_n \in E_k$ such that $x_n \to x$ in S. Then $x \in \Gamma^{\mathbf{P}}_{cpie}$ and $\varepsilon_{x_n} \to \varepsilon_x$.

For $N \in \mathbb{N}$ define $V_N = B_x(1/k) \cap (\cap_{n \geq N} B_{x_n}(1/k))$. Let $y \in V_N$ and define $r := \sup\{d(y, x_n) : n \geq N\}$. Since $d(y, x) < 1/k$ and $x_n \to x$, $r < 1/k$, which implies that V_N is open in S. Now, for all $N \in \mathbb{N}$,

$$\varepsilon_x(V_N) \leq \liminf_{n\to\infty} \varepsilon_{x_n}(V_N) \leq \liminf_{n\to\infty} \varepsilon_{x_n}(B_{x_n}(1/k)) = 0,$$

by the Portmanteau Theorem. Since $V_N \subset V_{N+1}$ for all N and $\cup_N V_N = B_x(1/k)$, $\varepsilon_x(B_x(1/k)) = 0$ and E_k is closed.

We can write:

$$D = \Gamma^{\mathbf{P}}_{cpie} \cap \bigcap_{k\in\mathbb{N}} (S \backslash E_k).$$

Since $\Gamma^{\mathbf{P}}_{cpie}$ is a closed subset of a metric space, it is a G_δ set, thus D is a G_δ set.

The final statement follows from [6, Theorem 3.1.2], which states that every G_δ subset of a Polish space is again a Polish space. □

We introduce the *weak concentrating condition*:

($\mathcal{C}$) There exists a compact $K \subset S$ such that for every $\epsilon > 0$ and every $x \in S$

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(K^\epsilon) > 0.$$

It turns out that we can obtain every ergodic measure from K (note that we do not require the (Cesàro) e-property for this result):

Lemma 8.2.5. *Suppose* ($\mathcal{C}$) *is satisfied. For every* $x \in \Gamma^{\mathbf{P}}_{cpie}, K \cap \operatorname{supp}(\varepsilon_x) \neq \emptyset$.

Proof. Suppose $\Gamma^{\mathbf{P}}_{cpie}$ is non-empty and let $x \in \Gamma^{\mathbf{P}}_{cpie}$ be such that $K \cap \operatorname{supp}(\varepsilon_x) = \emptyset$. Since $\operatorname{supp}(\varepsilon_x)$ is closed and K is compact, there exists an $\epsilon > 0$ such that $K^\epsilon \cap \operatorname{supp}(\varepsilon_x) = \emptyset$. Thus $\varepsilon_x(K^\epsilon) = 0$. In particular, $\varepsilon_x(\overline{K^{\epsilon/2}}) = 0$.
Let $y \in \operatorname{supp}(\varepsilon_x) \cap [x]$. The latter set is non-empty since $\varepsilon_x([x]) = 1$. Then, using the Portmanteau Theorem,

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_y(\overline{K^{\epsilon/2}}) \leq \varepsilon_x(\overline{K^{\epsilon/2}}) = 0,$$

which contradicts ($\mathcal{C}$). □

Now, if $(P(t))_{t\geq 0}$ satisfies the Cesàro e-property and ($\mathcal{C}$), then Theorem 7.4.4 implies that there exists an invariant measure. Then there must also exist an ergodic measure, and thus $\Gamma^{\mathbf{P}}_{cpie}$ and $\Gamma^{\mathbf{P}}_{cp}$ are non-empty. The following result implies that we can assume, without loss of generality, that the K from ($\mathcal{C}$) is contained in $\Gamma^{\mathbf{P}}_{cp}$.

Proposition 8.2.6. *Let $(P(t))_{t\geq 0}$ satisfy the Cesàro e-property. Let $K \subset S$ be compact. Then there is a $K_0 \subset K \cap \Gamma^{\mathbf{P}}_{cp}$ compact and for all $\epsilon > 0$ there is a $\delta > 0$ such that for every $\mu \in \mathcal{P}(S)$*

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(K_0^\epsilon) \geq \limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(K^\delta).$$

Proof. We define by $K_0 \subset K$ the set of all $x \in K$ such that for any $\epsilon > 0$

$$\limsup_{t\to\infty} \boldsymbol{P}^{(t)}\mu(B_x(\epsilon)) > 0 \tag{8.3}$$

for some $\mu \in \mathcal{P}(S)$. Observe that K_0 is closed, hence compact. Fix $\epsilon > 0$. Then for each $x \in K_0$ we define $r_x = \epsilon/2$. For $x \in K\backslash K_0$ there exists an $r_x > 0$ such that $\lim_{t\to\infty} \mathbf{P}^{(t)}\mu(B_x(2r_x)) = 0$ for all $\mu \in \mathcal{P}(S)$. By compactness of K there exist $m, n \in \mathbb{N}_0$, $x_1, \cdots, x_m \in K_0$ and $x_{m+1}, \cdots, x_n \in K\backslash K_0$ such that $K \subset \cup_{i=1}^n B_{x_i}(r_{x_i})$. Let $\delta := \min\{r_{x_i} : 1 \leq i \leq n\}$, then $\delta > 0$ and

$$K^\delta \subset \cup_{i=1}^m B_{x_i}(\epsilon) \cup (\cup_{i=1}^m B_{x_i}(2r_{x_i})) \subset K_0^\epsilon \cup \left(\cup_{i=n+1}^m B_{x_i}(2r_{x_i})\right).$$

Now,

$$\begin{aligned}\limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(K_0^\epsilon) &= \limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(K_0^\epsilon) + \sum_{k=m+1}^{n} \limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(B_{x_i}(2r_{x_i})) \\ &\geq \limsup_{t\to\infty} \mathbf{P}^{(t)}\mu(K^\delta).\end{aligned}$$

Proposition 7.4.2 yields that $K_0 \subset \Gamma^{\mathbf{P}}_t = \Gamma^{\mathbf{P}}_{cp}$. □

Note that if we apply Proposition 8.2.6 to the K from $(\mathcal{C})$, then the obtained K_0 must be non-empty. We shall write $\hat{K} := K \cap \Gamma^{\mathbf{P}}_{cpie}$. Since $\Gamma^{\mathbf{P}}_{cpie}$ is closed, $\hat{K}$ is compact. By Lemma 8.2.5, $\Phi(\hat{K}) = \mathcal{P}_{\text{erg}}(S)$, and by continuity of Φ we can conclude:

Corollary 8.2.7. *If the Cesàro e-property and $(\mathcal{C})$ hold for $(P(t))_{t\geq 0}$, then $\mathcal{P}_{\text{erg}}(S)$ is compact in $\mathcal{S}_{\text{BL}}$.*

Remark. From Example 8.2.2 we see that the statement in Corollary 8.2.7 need not hold if we replace the Cesàro e-property by the weaker Markov-Feller property.

The following result can be found in [13, Corollary 6.9.18]:

Proposition 8.2.8. *Let X be a Polish space and R an equivalence relation on X with closed equivalence classes. If $R(Z) \subset X$ is a Borel set for every closed $Z \subset X$, then R admits a Borel section, i.e. there is a Borel set $B \subset X$ such that B contains exactly one element of every equivalent class.*

Theorem 8.2.9. *If $(P(t))_{t\geq 0}$ satisfies the Cesàro e-property and $(\mathcal{C})$, then there exists a Borel set $K_0 \subset K$ such that*

(i) $x \in \text{supp}(\varepsilon_x)$ for all $x \in K_0$. In particular $K_0 \subset \Gamma^{\mathbf{P}}_{cpie}$.

(ii) If $x, y \in K_0$ with $x \neq y$, then $\varepsilon_x \neq \varepsilon_y$.

(iii) For every ergodic measure μ there is an $x \in K_0$ such that $\mu = \varepsilon_x$.

Proof. Let

$$X := \bigcup_{\mu \in \mathcal{P}_{\mathrm{erg}}(S)} \mathrm{supp}(\mu) \cap K,$$

then X is a G_δ set by Theorem 8.2.4, hence a Polish space in its relative topology by [6, Theorem 3.1.2]. Also, $X \subset \Gamma^{\mathbf{P}}_{cpie}$.

Let us define an equivalence relation R on X as follows: xRy if and only if x and y are in the support of the same ergodic measure, so if and only if $\varepsilon_x = \varepsilon_y$. Note that xRy if and only if $x \in \mathrm{supp}(\varepsilon_y)$ if and only if $y \in \mathrm{supp}(\varepsilon_x)$. Note that R is the restriction to X of the equivalence class $\sim$ on $\Gamma^{\mathbf{P}}_{cpie}$ we introduced earlier. For every $x \in X$, $R(x) = \mathrm{supp}(\varepsilon_x) \cap K = \mathrm{supp}(\varepsilon_x) \cap X$, thus $R(x)$ is closed in X. In order to apply Proposition 8.2.8, we need to show that $R(Z)$ is Borel for all closed subsets Z in X.

Let Z be closed in X. Let $\bar{Z}$ be its closure in S, then $Z = \bar{Z} \cap X$. Furthermore, $\bar{Z}$ is a closed subset of $K \cap \Gamma^{\mathbf{P}}_{cpie}$, hence compact. We claim that

$$R(Z) = \{x \in \Gamma^{\mathbf{P}}_{cpie} \cap K : \varepsilon_x(\bar{Z}^{1/n}) > 0 \text{ for all } n \in \mathbb{N}\} =: W_Z.$$

Let $x \in R(Z)$, then there is a $z \in Z$ such that $x \in \mathrm{supp}(\varepsilon_z)$, thus also $z \in \mathrm{supp}(\varepsilon_x)$. Hence, for all $n \in \mathbb{N}$, $\varepsilon_x(\bar{Z}^{1/n}) \geq \varepsilon_x(B_z(1/n)) > 0$, so $x \in W_Z$.

Now let $x \in W_Z$. Suppose that $\mathrm{supp}(\varepsilon_x) \cap \bar{Z} = \emptyset$, then compactness of $\bar{Z}$ implies that there is an $n \in \mathbb{N}$ such that $\varepsilon_x(\bar{Z}^{1/n}) = 0$, which is a contradiction. So there is a $z \in \bar{Z}$ such that $z \in \mathrm{supp}(\varepsilon_x)$. Since $\bar{Z} \subset K$, $z \in \bar{Z} \cap X = Z$. Since $z \in \mathrm{supp}(\varepsilon_x)$, $x \in \mathrm{supp}(\varepsilon_z)$, so $x \in R(Z)$.

Now it remains to show that W_Z is Borel in X. We can write

$$W_Z = \bigcap_{n \in \mathbb{N}} W_Z^n,$$

where $W_Z^n = \{x \in \Gamma^{\mathbf{P}}_{cpie} \cap K : \varepsilon_x(\bar{Z}^{1/n}) > 0\}$. Let

$$V_Z^n := (\Gamma^{\mathbf{P}}_{cpie} \cap K) \backslash W_Z^n = \{x \in \Gamma^{\mathbf{P}}_{cpie} \cap K : \varepsilon_x(\bar{Z}^{1/n}) = 0\},$$

and let $x_k \in V_Z^n$ such that $x_k \to x \in S$. Then $x \in \Gamma^{\mathbf{P}}_{cpie} \cap K$ and by the Portmanteau Theorem

$$\varepsilon_x(\bar{Z}^{1/n}) \leq \liminf_{k \to \infty} \varepsilon_{x_k}(\bar{Z}^{1/n}) = 0,$$

thus $x \in V_Z^n$. So V_Z^n is closed, and thus W_Z^n is open (in the relative topology on $\Gamma^{\mathbf{P}}_{cpie} \cap K$). Then $W_Z^n \cap X$ is open in X, so

$$R(Z) = W_Z = X \cap W_Z = \bigcap_{n \in \mathbb{N}} W_Z^n \cap X$$

is a G_δ subset of X, thus Borel. Application of Proposition 8.2.8 yields the existence of a Borel set $K_0 \subset X \subset \Gamma^{\mathbf{P}}_{cpie} \cap K$ such that for every $x \in X$ there is exactly one $y \in K_0$ such that $y \in R(x)$.

Thus *(i)* and *(ii)* are satisfied. Now let μ be an ergodic measure, then there is an $x \in \Gamma^{\mathbf{P}}_{cpie}$ with $\mu = \varepsilon_x$. Lemma 8.2.5 implies that there is a $z \in \operatorname{supp}(\varepsilon_x) \cap K \subset X$ and thus there is exactly one $y \in K_0$ such that $y \in R(z)$. Consequently $\varepsilon_y = \varepsilon_z = \varepsilon_x = \mu$. This concludes the proof. □

Note that the set K_0 from Theorem 8.2.9 need not be unique. For instance, if we let S be the unit circle and $P(t)\delta_x := \delta_{e^{2\pi i t}x}$, then $(P(t))_{t\geq 0}$ defines a Markov-Feller semigroup with the e-property with a unique ergodic measure given by the normalised Lebesgue measure on S. Obviously we can choose $K_0 = \{z\}$ for any $z \in S$.

Theorem 8.2.9 raises the interesting open question if for Markov-Feller semigroups with the (Cesàro) e-property, but without $(\mathcal{C})$, a result analogous to Theorem 8.2.9 holds.

We say that $(P(t))_{t\geq 0}$ is *sweeping* from some family $\mathcal{A}$ of Borel subsets of S if

$$\lim_{t\to\infty} P(t)\mu(A) = 0,$$

for all $\mu \in \mathcal{P}(S)$ and $A \in \mathcal{A}$.

The following result generalises [114, Proposition 3] (where $(\mathcal{C})$ is assumed to hold with K equal to a single element):

Proposition 8.2.10. *Suppose $(P(t))_{t\geq 0}$ satisfies the eventual e-property and $(\mathcal{C})$. Then $(P(t))_{t\geq 0}$ is sweeping from compact sets disjoint from $\Gamma^{\mathbf{P}}_t = \Gamma^{\mathbf{P}}_{cp}$.*

Proof. By Proposition 8.2.6 we can assume that the K from $(\mathcal{C})$ is contained in $\Gamma^{\mathbf{P}}_{cp}$. Thus $\Gamma^{\mathbf{P}}_t = \Gamma^{\mathbf{P}}_{cp}$ is non-empty.

Suppose there is a compact L such that $L \cap \Gamma^{\mathbf{P}}_t = \emptyset$ and an $\alpha > 0$ and $\mu \in \mathcal{P}(S)$ such that $\limsup_{t\to\infty} P(t)\mu(L) > \alpha$. Since $\Gamma^{\mathbf{P}}_t$ is closed, there is an $\eta > 0$ such that $L^\eta \cap \Gamma^{\mathbf{P}}_t = \emptyset$.

We define

$$M := \{\nu \in \mathcal{P}(S) : \text{there exists } \gamma < \eta \text{ such that } \liminf_{t\to\infty} P(t)\nu((\Gamma^{\mathbf{P}}_t)^\gamma) > 1 - \alpha/2\}.$$

Note that $\Gamma^{\mathbf{P}}_t$ is $(P(t))_{t\geq 0}$-invariant (see Corollary 7.3.14), thus $\{\delta_x : x \in \Gamma^{\mathbf{P}}_t\} \subset M$, and in particular M is non-empty as well. Also, M is convex and $P(t)M \subset M$ for all $t \in \mathbb{R}_+$. From [114, Lemma 3] it follows that M is an open subset of $\mathcal{P}(S)$ endowed with the weak topology. Thus, since $K \subset \Gamma^{\mathbf{P}}_t$ and K is compact, there is a $\sigma > 0$ such that if $\nu \in \mathcal{P}(S)$ with $\nu(K^\sigma) = 1$, then $\nu \in M$.

Let $x \in L$, then $(\mathcal{C})$ implies that there is a $t_x > 0$ such that $\alpha_x := P(t_x)\delta_x(K^{\sigma/2}) > 0$. Since $P(t_x)$ is Markov-Feller, there is an $r_x > 0$ such that $P(t_x)\delta_y(K^{\sigma/2}) > \alpha_x/2$ for all $y \in B_x(r_x)$. By compactness of L there exist $x_1, \cdots, x_m \in L$ such that $L \subset \cup_{i=1}^m B_{x_i}(r_{x_i})$. Define $\Theta = \min_{1\le i\le m} \alpha_{x_i}/2$ and

$$\gamma := \sup\{\beta \ge 0 : P(t_0)\mu \ge \beta\nu \text{ for some } \nu \in M, t_0 > 0\}.$$

Clearly $\gamma \le 1$. Now choose $\nu \in M$ and $t_0 > 0$ such that $P(t_0)\mu \ge \beta\nu$ holds with $\beta > \gamma - \Theta\alpha/(2m)$. Then for all $t \ge 0$, $P(t+t_0)\mu \ge \beta P(t)\nu$ and $P(t)\nu \subset M$, thus we can choose $\nu \in M$ and t_0 in such a way that $P(t_0)\mu(L) > \alpha$ and $\nu(L) < \alpha/2$. Then, since $\beta \le 1$,

$$(P(t_0)\mu - \beta\nu)(L) > \alpha - \alpha/2 = \alpha/2.$$

So there exist $j \in \{1, \cdots, m\}$ such that $(P(t_0)\mu - \beta\nu)(B_{x_j}(r_{x_j})) \ge \alpha/(2m)$. Now

$$\begin{aligned}
\langle P(t_{x_j})(P(t_0)\mu - \beta\nu), 1\!\!1_{K^\sigma}\rangle &= \langle P(t_0)\mu - \beta\nu, U(t_{x_j})1\!\!1_{K^\sigma}\rangle \\
&= \int_S P(t_{x_j})\delta_x(K^\sigma)d[P(t_0)\mu - \beta\nu](x) \\
&\ge \int_{B_{x_j}(r_{x_j})} P(t_{x_j})\delta_x(K^\sigma)d[P(t_0)\mu - \beta\nu](x) \\
&\ge \Theta\alpha/(2m).
\end{aligned}$$

Set

$$\tilde{\nu} = \frac{(P(t_{x_j} + t_0)\mu - \beta P(t_{x_j})\nu)(\cdot \cap K^\sigma)}{(P(t_{x_j} + t_0)\mu - \beta P(t_{x_j})\nu)(K^\sigma)},$$

then $\tilde{\nu} \in M$, since $\nu(K^\sigma) = 1$. Let

$$\hat{\nu} = \beta(\beta + \Theta\alpha/(2m))^{-1}P(t_{x_j})\nu + \Theta\alpha/(2m)(\beta + \Theta\alpha/(2m))^{-1}\tilde{\nu}.$$

Since $P(t_{x_j})\nu$ and $\tilde{\nu}$ are in M, $\hat{\nu}$ is in M as well, since M is convex. Furthermore,

$$P(t_{x_j} + t_0)\mu \ge (\beta + \Theta\alpha/(2m))\hat{\nu},$$

which contradicts the fact that $\gamma < \beta + \Theta\alpha/(2m)$. This completes the proof. □

8.3 Countably many ergodic measures

Let $\mathbf{P} = (P(t))_{t\ge0}$ be a jointly measurable Markov-Feller semigroup with dual $\mathbf{U} = (U(t))_{t\ge0}$. In this section we give some sufficient conditions for the set of ergodic measures to be countable or finite.

First we note that the e-property, even when combined with the weak concentrating condition $(\mathcal{C})$, does not guarantee that the set of ergodic measures is countable. A trivial example is given by any uncountable Polish space S and $(P(t))_{t\ge0}$ the identity semigroup on S. Then δ_x is an ergodic measure for all $x \in S$.

Hairer and Mattingly introduce in [52] the asymptotic strong Feller property, which generalises the strong Feller property (see Section 7.2), and use it in combination with other conditions to show uniqueness of invariant measures of certain Markov semigroups associated to stochastic differential equations which need not be strong Feller. The definition of the asymptotic strong Feller property is quite involved, so we omit it here. A sufficient condition for a Markov semigroup to be asymptotic strong Feller is given in [52, Proposition 3.12]. This sufficient condition is easier to state. However, it makes sense only for Hilbert spaces. We give a more general condition that works for Polish spaces as well. We will show that this implies that there are at most countably many ergodic measures, and when combined with $(\mathcal{C})$, the number of ergodic measures is finite. The condition is as follows:

Fix $x_0 \in S$. For $f : S \to \mathbb{R}$ and $\theta > 0$ we define the local Lipschitz constant

$$|f|_{\mathrm{Lip},\theta} := \sup\left\{\frac{|f(x) - f(y)|}{d(x,y)} : x, y \in B_{x_0}(\theta)\right\}.$$

We assume there are sequences $t_n \in \mathbb{R}_+$ and $\delta_n \downarrow 0$ and a non-decreasing function $C : \mathbb{R}_+ \to \mathbb{R}$, such that for all $f \in \mathrm{BL}(S)$ and $\theta \geq 0$

$$|U(t_n)f|_{\mathrm{Lip},\theta} \leq C(\theta)[\|f\|_\infty + \delta_n |f|_{\mathrm{Lip}}]. \tag{8.4}$$

Whenever a Markov semigroup on a Hilbert space satisfies the condition from [52, Proposition 3.12], it also satisfies (8.4). See also [52] for certain examples of Markov semigroups satisfying this condition.

Our next result gives lower bounds on distances between points in the supports of different ergodic measures. It generalises [53, Theorem 2.1] and its proof is based on the proof of that theorem. We include it here for completeness.

Proposition 8.3.1. *Let μ and ν be ergodic measures and $x \in \mathrm{supp}(\mu)$, $y \in \mathrm{supp}(\nu)$. Then* (8.4) *implies that*

$$d(x,y) \geq \frac{1}{C(d(x,x_0) \vee d(y,x_0))}.$$

Proof. We define for $n \in \mathbb{N}$ the following metric on S: $d_n(x,y) = 1 \wedge (\frac{1}{\sqrt{\delta_n}} d(x,y))$. These metrics induce metrics on $\mathcal{P}(S)$ in the following way:

$$d_n(\mu,\nu) = \sup\{|\langle \mu - \nu, f\rangle| : f \in \mathrm{Lip}^1_{d_n}(S)\},$$

where

$$\mathrm{Lip}^1_{d_n}(S) = \{f : S \to \mathbb{R} : |f(x) - f(y)| \leq d_n(x,y) \text{ for all } x, y \in S\}.$$

Then $d_n(\mu,\nu) \leq 1$ and $\lim_{n\to\infty} d_n(\mu,\nu) = \frac{1}{2}\|\mu - \nu\|_{\mathrm{TV}}$, by [52, Lemma 3.4]. Note that it suffices to only consider those $f \in \mathrm{Lip}^1_{d_n}(S)$ for which $f(x_0) = 0$. For such

f, $|f(x)| = |f(x) - f(x_0)| \leq d_n(x, x_0) \leq 1$, so $\|f\|_\infty \leq 1$. Moreover,

$$|f(x) - f(y)| \leq \frac{1}{\sqrt{\delta_n}} d(x, y) \text{ for all } x, y \in S,$$

so $|f|_{\mathrm{Lip}} \leq \frac{1}{\sqrt{\delta_n}}$. Now we apply (8.4):

$$\begin{aligned} d_n(P(t_n)\delta_x, P(t_n)\delta_y) &\leq \sup\{|U(t_n)f(x) - U(t_n)f(y)| : f \in \mathrm{Lip}^1_{d_n}(S), f(x_0) = 0\} \\ &\leq d(x, y)C(d(x, x_0) \vee d(y, x_0))(1 + \sqrt{\delta_n}). \end{aligned}$$

Let μ_1 and μ_2 be two distinct ergodic measures, then they are mutually singular, so $\|\mu_1 - \mu_2\|_{\mathrm{TV}} = 2$. Suppose that there are $x \in \mathrm{supp}(\mu_1)$ and $y \in \mathrm{supp}(\mu_2)$ such that $d(x, y) < (C(d(x, x_0) \vee d(y, x_0)))^{-1}$. Then we will show that $\|\mu_1 - \mu_2\|_{\mathrm{TV}} < 2$, which gives a contradiction. By assumption there is a Borel set E containing x and y such that $\alpha := \min(\mu_1(E), \mu_2(E)) > 0$ and $\beta := \mathrm{diam}(E)C(d(x, x_0) \vee d(y, x_0)) < 1$. We can write $\mu_i = \alpha\nu_i + (1 - \alpha)\rho_i$, with $\nu_i, \rho_i \in \mathcal{P}(S)$ and $\nu_i(E) = 1$. Then

$$\begin{aligned} d_n(\mu_1, \mu_2) &= d_n(P(t_n)\mu_1, P(t_n)\mu_2) \leq \alpha d_n(P(t_n)\nu_1, P(t_n)\nu_2) + 1 - \alpha \\ &\leq \alpha \int_S \int_S d_n(P(t_n)\delta_w, P(t_n)\delta_z)\, d\nu_1(w)\, d\nu_2(z) + 1 - \alpha \\ &\leq \alpha\beta(1 + \sqrt{\delta_n}) + 1 - \alpha, \end{aligned}$$

thus

$$\tfrac{1}{2}\|\mu_1 - \mu_2\|_{\mathrm{TV}} \leq 1 - \alpha(1 - \beta) < 1,$$

which is a contradiction. □

Corollary 8.3.2. *Assume that* (8.4) *holds for some non-decreasing* $C : \mathbb{R}_+ \to \mathbb{R}$. *Then there exist at most countably many ergodic measures.*

Proof. We will show that for every bounded set B in S, there exist at most countably many ergodic measures whose support intersects B. Since we can cover S with countably many bounded sets, this proves that there are at most countably many ergodic measures.

Let $B \subset S$ bounded, and define $R := \sup\{d(x, x_0) : x \in B\} < \infty$. Let μ be an ergodic measure with $x \in \mathrm{supp}(\mu) \cap B$. Then Proposition 8.3.1 implies that for any ergodic measure ν with $\mu \neq \nu$ and $y \in \mathrm{supp}(\nu) \cap B$ we have

$$d(x, y) \geq 1/C(R). \tag{8.5}$$

Now we choose for every ergodic measure μ with $\mathrm{supp}(\mu) \cap B \neq \emptyset$ an $x \in \mathrm{supp}(\mu) \cap B$ and consider the open ball $B_x(1/(2C(R)))$. By (8.5) these balls are mutually disjoint. Separability of S implies that there can be only countably many of such balls, which concludes the proof. □

Corollary 8.3.3. *Assume that* (8.4) *holds for some non-decreasing* $C : \mathbb{R}_+ \to \mathbb{R}$ *and that condition* $(\mathcal{C})$ *holds. Then there are only finitely many ergodic measures.*

Proof. Lemma 8.2.5 implies that the support of every ergodic measure has non-empty intersection with K. Since K is compact, it is bounded, so by the proof of Corollary 8.3.2 there is an $R > 0$ such that whenever μ and ν are two distinct ergodic measures, then $B_x(1/(2C(R)) \cap B_y(1/(2C(R))) = \emptyset$ for every $x \in \operatorname{supp}(\mu) \cap K$ and $y \in \operatorname{supp}(\nu) \cap K$. Since any subset of K is totally bounded, there can be only a finite number of mutually disjoint balls with radius $(2C(R))^{-1}$ and center in K, so the number of ergodic measures is finite as well. □

Notice that the conditions in Corollary 8.3.3 not necessarily imply the existence of invariant measures. However, when combined with the Cesàro e-property there do exist invariant measures (cf. Theorem 7.4.4). There exist examples of Markov-Feller semigroups with the e-property that do not satisfy (8.4) or even the asymptotic strong Feller property. At the beginning of this section we gave a trivial example of such a semigroup. See also [73, Remark 6]. As for now we do not know any Markov-Feller semigroups that satisfy the asymptotic strong Feller property but not the e-property. We now give a condition to ensure both properties:

Proposition 8.3.4. *Suppose there exists a non-decreasing* $C : \mathbb{R}_+ \to \mathbb{R}$ *such that for all* $t \in \mathbb{R}_+$ *and* $f \in \mathrm{BL}(S)$

$$|U(t)f|_{\mathrm{Lip},\theta} \leq C(\theta)[\|f\|_\infty + |f|_{\mathrm{Lip}}] = C(\theta)\|f\|_{\mathrm{BL}}.$$

Then $(P(t))_{t\geq 0}$ *satisfies the e-property. If in addition there is a function* $h : \mathbb{R}_+ \to \mathbb{R}_+$ *such that* $\lim_{t\to\infty} h(t) = 0$ *and*

$$|U(t)f|_{\mathrm{Lip},\theta} \leq C(\theta)[\|f\|_\infty + h(t)|f|_{\mathrm{Lip}}] \text{ for all } t \in \mathbb{R}_+,$$

then $(P(t))$ *satisfies* (8.4) *as well.*

Proof. If $x_n \to x \in S$, then for all $f \in \mathrm{BL}(S)$,

$$\sup_{t\geq 0} |U(t)f(x_n) - U(t)f(x)| \leq C(d(x, x_0) + 1)\|f\|_{\mathrm{BL}} d(x_n, x)$$

for n large enough. This proves the e-property. It is clear that under the extra assumption, $(P(t))_{t\geq 0}$ satisfies (8.4) as well. □

8.4 Convergence of Cesàro averages

In this section we will formulate a condition on Markov-Feller semigroups with the Cesàro e-property such that $\Gamma_{cp}^{\mathbf{P}} = S$, i.e. such that the Cesàro averages of each probability measure will converge weakly to an invariant probability measure.

We would like to remark at this point that $(\mathcal{C})$ is not sufficient for this to happen. See [73, Remark 1] for an example of a Markov-Feller semigroup $(P(t))_{t\geq 0}$ having the e-property that satisfies an even stronger condition than $(\mathcal{C})$, i.e. there is a $z \in S$ such that $\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\epsilon)) > 0$ for all $\epsilon > 0$ and all $x \in S$. However, as shown in [73, Remark 1], the set $\Gamma_t^{\mathbf{P}}$ for this semigroup does not equal the whole space, and since $\Gamma_{cp}^{\mathbf{P}} = \Gamma_t^{\mathbf{P}}$, there exist probability measures for which the Cesàro averages do not converge.

It turns out that strengthening $(\mathcal{C})$ by demanding a uniform lower bound depending on ϵ (precisely stated in Theorem 8.4.3 *(i)*) will give the result. We first prove some preliminary results:

Lemma 8.4.1. *Let $(P(t))_{t\geq 0}$ be a Markov-Feller semigroup that satisfies the Cesàro e-property. Let $K \subset S$ be compact. Then for all $\epsilon > 0$ there exists a $\delta > 0$ such that for every $y \in K$ and $x \in B_y(\delta)$*

$$\|\mathbf{P}^{(t)}\delta_x - \mathbf{P}^{(t)}\delta_y\|_{\mathrm{BL}}^* < \epsilon \textit{ for all } t \geq 1.$$

Proof. Suppose that the statement does not hold. Then there exists an $\epsilon > 0$, $y_n \in K, x_n \in B_{y_n}(1/n)$ and $t_n \in [1, \infty)$ such that

$$\|\mathbf{P}^{(t_n)}\delta_{x_n} - \mathbf{P}^{(t_n)}\delta_{y_n}\|_{\mathrm{BL}}^* \geq \epsilon. \tag{8.6}$$

There is a subsequence y_{n_k} such that $y_{n_k} \to y \in K$ and thus $x_{n_k} \to y$.

The Cesàro e-property and Theorem 7.2.2 imply that the family of maps $(\mathbf{P}^{(t)}\delta_\cdot)_{t\geq 1}$ from S to $\mathcal{S}_{\mathrm{BL}}$ is equicontinuous. So, as $k \to \infty$,

$$\|\mathbf{P}^{(t_{n_k})}\delta_{y_{n_k}} - \mathbf{P}^{(t_{n_k})}\delta_y\|_{\mathrm{BL}}^* \to 0$$

and

$$\|\mathbf{P}^{(t_{n_k})}\delta_{x_{n_k}} - \mathbf{P}^{(t_{n_k})}\delta_y\|_{\mathrm{BL}}^* \to 0.$$

This contradicts (8.6). □

Lemma 8.4.1 allows us to prove the following important result, showing that in certain cases we can replace a *limes superior* condition by a *limes inferior* condition.

Lemma 8.4.2. *Let $(P(t))_{t\geq 0}$ be a Markov-Feller semigroup that satisfies the Cesàro e-property and $(\mathcal{C})$. Let $L \subset S$ be compact. Then there is a compact set $\hat{L} \subset \Gamma_{cp}^{\mathbf{P}}$ such that if $A \subset S$ satisfies*

$$\inf_{x\in A} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(L^\epsilon) > 0 \textit{ for all } \epsilon > 0, \tag{8.7}$$

then

$$\inf_{x\in A} \liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(\hat{L}^\epsilon) > 0 \textit{ for all } \epsilon > 0. \tag{8.8}$$

Proof. By Proposition 8.2.6 we can assume, without loss of generality, that the compact set L is contained in $\Gamma^{\mathbf{P}}_{cp}$.

By Corollary 8.2.7 the set of ergodic measures is compact in $\mathcal{S}_{\mathrm{BL}}$, hence tight by Theorem 5.2.3. So there exists a compact set L_1 such that $\nu(L_1) \geq 1/2$ for every ergodic measure ν. Now for an arbitrary invariant probability measure μ, we obtain

$$\mu(L_1) = \int_{\Gamma^{\mathbf{P}}_{cpie}} \varepsilon_x(L_1)\, d\mu(x) \geq \mu(\Gamma^{\mathbf{P}}_{cpie})/2 = 1/2.$$

Since every invariant measure is concentrated on the closed set $\Gamma^{\mathbf{P}}_{cp}$, we may assume that $L_1 \subset \Gamma^{\mathbf{P}}_{cp}$. Now we define the compact set $\hat{L} := L \cup L_1 \subset \Gamma^{\mathbf{P}}_{cp}$.

Fix $0 < \epsilon < 1/8$. By Lemma 8.4.1 there exists a $\delta > 0$ such that for all $y \in \hat{L}$ and $x \in B_y(\delta)$, $\|\mathbf{P}^{(t)}\delta_x - \mathbf{P}^{(t)}\delta_y\|^*_{\mathrm{BL}} < \epsilon^2$ for all $t \geq 1$.

Define $g := (1 - d(\cdot, \hat{L})/\epsilon)^+$, then $|g|_{\mathrm{Lip}} \leq 1/\epsilon$ and $\|g\|_\infty \leq 1$, so $g \in \mathrm{BL}(S)$ with $\|g\|_{\mathrm{BL}} \leq 1/\epsilon + 1$. Moreover, $\frac{1}{2}\mathbb{1}_{\hat{L}^{\epsilon/2}} \leq g \leq \mathbb{1}_{\hat{L}^{\epsilon}}$. Fix $x \in \hat{L}^\delta$ and let $y \in \hat{L}$ be such that $d(x, y) < \delta$. Then

$$\begin{aligned}
\mathbf{P}^{(t)}\delta_x(\hat{L}^\epsilon) &\geq \langle \mathbf{P}^{(t)}\delta_x, g\rangle \geq \langle \mathbf{P}^{(t)}\delta_y, g\rangle - \|\mathbf{P}^{(t)}\delta_x - \mathbf{P}^{(t)}\delta_y\|^*_{\mathrm{BL}}(1/\epsilon + 1) \\
&\geq \tfrac{1}{2}\mathbf{P}^{(t)}\delta_y(\hat{L}^{\epsilon/2}) - (\epsilon + \epsilon^2).
\end{aligned}$$

Since $y \in \Gamma^{\mathbf{P}}_{cp}$, $\mathbf{P}^{(t)}\delta_y$ converges to the invariant probability measure ε_y, so we obtain by the Portmanteau Theorem

$$\begin{aligned}
\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(\hat{L}^\epsilon) &\geq \tfrac{1}{2}\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_y(\hat{L}^{\epsilon/2}) - (\epsilon + \epsilon^2) \\
&\geq \tfrac{1}{2}\varepsilon_y(\hat{L}^{\epsilon/2}) - (\epsilon + \epsilon^2) \\
&\geq 1/4 - \epsilon - \epsilon^2 > 1/4 - 2\epsilon > 0.
\end{aligned}$$

Now let $A \subset S$ be such that (8.7) is satisfied. Then

$$\alpha := \inf_{x\in A} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(\hat{L}^\delta) > 0$$

since $L \subset \hat{L}$.

Fix $x \in A$. Then there is a $T > 0$ such that $\mathbf{P}^{(T)}\delta_x(\hat{L}^\delta) \geq \alpha/2$. Define

$$\rho := \frac{\mathbf{P}^{(T)}\delta_x(\hat{L}^\delta \cap \cdot)}{\mathbf{P}^{(T)}\delta_x(\hat{L}^\delta)}.$$

Then $\rho \in \mathcal{P}(S)$ and $\mathbf{P}^{(T)}\delta_x \geq \mathbf{P}^{(T)}\delta_x(\hat{L}^\delta)\rho \geq (\alpha/2)\rho$.

By Fatou's Lemma

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\rho(\hat{L}^\epsilon) \geq \int_{\hat{L}^\delta} \liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_y(\hat{L}^\epsilon)\, d\rho(y) \geq 1/4 - 2\epsilon.$$

Now,

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(\hat{L}^\epsilon) = \liminf_{t\to\infty} \mathbf{P}^{(t)}\mathbf{P}^{(T)}\delta_x(\hat{L}^\epsilon) \geq \tfrac{\alpha}{2}(1/4 - 2\epsilon) > 0,$$

where the first equality follows from Lemma 6.2.13. Thus (8.8) is satisfied. □

Theorem 8.4.3. *Let $(P(t))_{t\geq 0}$ be a Markov-Feller semigroup that satisfies the Cesàro e-property. Then the following three statements are equivalent:*

(i) There exists a compact set $K \subset S$ such that for every $\epsilon > 0$

$$\inf_{x\in S} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(K^\epsilon) > 0. \tag{8.9}$$

(ii) There exists a compact set $K \subset S$ such that for every bounded set A and every $\epsilon > 0$

$$\inf_{x\in A} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(K^\epsilon) > 0 \tag{8.10}$$

and for all $\eta > 0$ and $x \in S$ there exists a bounded Borel set D such that

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(D) > 1 - \eta. \tag{8.11}$$

(iii) The set of ergodic measures is compact and $(\mathbf{P}^{(t)}\mu)_{t\geq 0}$ converges to an invariant measure for every $\mu \in \mathcal{P}(S)$.

Proof. *(iii)⇒(i)* and *(iii)⇒(ii)*: By compactness the set of ergodic measures is tight. So for every $\eta > 0$ there exists a compact set K_η such that $\mu(K_\eta) \geq 1 - \eta$ for all ergodic μ. Let ν be an invariant probability measure, then

$$\nu(K_\eta) = \int_{\Gamma^{\mathbf{P}}_{cpie}} \varepsilon_x(K_\eta)\, d\nu(x) \geq \nu(\Gamma^{\mathbf{P}}_{cpie})(1-\eta) = 1 - \eta.$$

Let $x \in S$. By assumption $\mathbf{P}^{(t)}\delta_x$ converges to an invariant probability measure ν, thus the Portmanteau Theorem implies that for all $\epsilon > 0$ and $\eta > 0$,

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_x(K^\epsilon_\eta) \geq \nu(K_\eta) \geq 1 - \eta.$$

From this *(i)* and *(ii)* both follow.

(i)⇒(iii) and *(ii)⇒(iii)*: Either of the statements *(i)* and *(ii)* imply that $(\mathcal{C})$ is satisfied, so by Corollary 8.2.7 the set of ergodic measures is compact in $\mathcal{S}_{\mathrm{BL}}$.

It follows from Theorem 7.3.13 that it is sufficient to show that $\Gamma^{\mathbf{P}}_t = S$, i.e. $(\mathbf{P}^{(t)}\delta_z)_{t\geq 1}$ is tight for any $z \in S$. This follows from Lemma 6.3.5 if we can show that for every $\epsilon > 0$ there is a $C \in \mathcal{C}_\epsilon$ such that $\liminf_{t\to\infty} \mathbf{P}^{(t)}\delta_z(C) \geq 1 - \epsilon$.

By Lemma 8.4.2 there exists a compact $\hat{K} \subset \Gamma_{cp}^{\mathbf{P}}$ such that if *(i)* is satisfied, then

$$\inf_{x \in S} \liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_x(\hat{K}^\epsilon) > 0 \text{ for all } \epsilon > 0, \tag{8.12}$$

and if *(ii)* is satisfied, then

$$\inf_{x \in A} \liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_x(\hat{K}^\epsilon) > 0 \text{ for all bounded } A \subset S \text{ and } \epsilon > 0. \tag{8.13}$$

Step 1. For every $\epsilon > 0$ there exists an open set U with $\hat{K} \subset U$ and a $C \in \mathcal{C}_\epsilon$ such that for all $\mu \in \mathcal{P}(S)$ with $\mu(U) = 1$ the following holds:

$$\liminf_{t \to \infty} \mathbf{P}^{(t)} \mu(C) \geq 1 - \epsilon.$$

Fix $\epsilon > 0$ and let $x \in \hat{K}$. Since $(\mathbf{P}^{(t)} \delta_x)_{t \geq 1}$ is tight, we may find a $C_x \in \mathcal{C}_\epsilon$ and $r_x > 0$ such that

$$\liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_y(C_x) \geq 1 - \epsilon \qquad \text{for all } y \in B_x(r_x).$$

Indeed, Lemma 6.3.5 yields the existence of a $\tilde{C}_x \in \mathcal{C}_{\epsilon/2}$ for which

$$\liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_x(\tilde{C}_x) \geq 1 - \epsilon/2.$$

Choose an arbitrary function $f \in \mathrm{BL}(S)$ such that $\mathbb{1}_{\tilde{C}_x} \leq f \leq \mathbb{1}_{C_x}$, where $C_x = \tilde{C}_x^{\epsilon/2}$. Obviously, $C_x \in \mathcal{C}_\epsilon$. By the Cesàro e-property there exists $r_x > 0$ such that $|\mathbf{U}^{(t)} f(x) - \mathbf{U}^{(t)} f(y)| < \epsilon/2$ for $t \geq 1$ and $y \in B_x(r_x)$. Then

$$\begin{aligned} \liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_y(C_x) &\geq \liminf_{t \to \infty} \mathbf{U}^{(t)} f(y) \geq \liminf_{t \to \infty} \mathbf{U}^{(t)} f(x) - \epsilon/2 \\ &\geq \liminf_{t \to \infty} \mathbf{U}^{(t)} \mathbb{1}_{\tilde{C}_x}(x) - \epsilon/2 \geq 1 - \epsilon. \end{aligned} \tag{8.14}$$

Let $\{x_1, \cdots x_N\} \subset \hat{K}$ be such that

$$\hat{K} \subset \bigcup_{i=1}^{N} B_{x_i}(r_{x_i}) := U.$$

Set $C = \bigcup_{i=1}^{N} C_{x_i}$ and observe that $C \in \mathcal{C}_\epsilon$. For $\mu \in \mathcal{P}(S)$ with $\mu(U) = 1$ we have by Fatou's Lemma

$$\liminf_{t \to \infty} \mathbf{P}^{(t)} \mu(C) \geq \int_S \liminf_{t \to \infty} \mathbf{P}^{(t)} \delta_x(C) \, d\mu(x) \geq 1 - \epsilon. \tag{8.15}$$

The following step is technical. It is essential though for the proof of the subsequent step, that concludes the proof of the theorem.

Step 2. For all $z \in S$, $\epsilon > 0$ and $0 \leq \gamma < 1$ there exists a $\beta > 0$ such that for all $N \in \mathbb{N}, t_1, \cdots, t_N \in \mathbb{R}_+$, $0 \leq \alpha \leq \gamma$ and $\nu \in \mathcal{P}(S)$ such that $\mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z \geq \alpha\nu$, the probability measure

$$\hat{\mu} := \frac{\mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z - \alpha\nu}{1-\alpha} \tag{8.16}$$

satisfies

$$\liminf_{t\to\infty} \mathbf{P}^{(t)} \hat{\mu}(\hat{K}^\epsilon) \geq \beta. \tag{8.17}$$

If *(i)* is satisfied, we can take

$$\beta = \inf_{x\in S} \liminf_{t\to\infty} \mathbf{P}^{(t)} \delta_x(\hat{K}^\epsilon).$$

Then $\beta > 0$ by (8.12), and by Fatou's Lemma

$$\liminf_{t\to\infty} \mathbf{P}^{(t)} \nu(\hat{K}^\epsilon) \geq \beta$$

for every $\nu \in \mathcal{P}(S)$.

Now suppose *(ii)* is satisfied. Fix $z \in S$, $\epsilon > 0$ and $0 \leq \gamma < 1$. Let $0 < \delta < (1-\gamma)/2$. Let $D \subset S$ be bounded and Borel such that $\limsup_{t\to\infty} \mathbf{P}^{(t)} \delta_z(D) \geq 1-\delta$ and define

$$\beta_0 := \inf_{x\in D} \liminf_{t\to\infty} \mathbf{P}^{(t)} \delta_x(\hat{K}^\epsilon),$$

then $\beta_0 > 0$ by (8.13). Define

$$\beta := \left(1 - \frac{2\delta}{1-\gamma}\right) \beta_0.$$

Let $N \in \mathbb{N}, t_1, \cdots, t_N \in \mathbb{R}_+$, $0 \leq \alpha \leq \gamma$ and $\nu \in \mathcal{P}(S)$ be such that

$$\mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z \geq \alpha\nu,$$

and define the probability measure $\hat{\mu}$ as in (8.16). According to Lemma 6.2.13,

$$\limsup_{t\to\infty} \mathbf{P}^{(t)} \mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z(D) = \limsup_{t\to\infty} \mathbf{P}^{(t)} \delta_z(D) \geq 1-\delta,$$

so there is a $T \in \mathbb{R}_+$ such that $\mathbf{P}^{(T)} \mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z(D) \geq 1 - 2\delta$. Then

$$\mathbf{P}^{(T)} \hat{\mu}(S \backslash D) \leq \frac{\mathbf{P}^{(T)} \mathbf{P}^{(t_1)} \cdots \mathbf{P}^{(t_N)} \delta_z(S \backslash D)}{1-\alpha} \leq \frac{2\delta}{1-\alpha} \leq \frac{2\delta}{1-\gamma}.$$

So $\mathbf{P}^{(T)} \hat{\mu}(D) \geq 1 - \frac{2\delta}{1-\gamma} > 0$. Define

$$\tilde{\mu} := \frac{\mathbf{P}^{(T)} \hat{\mu}(\cdot \cap D)}{\mathbf{P}^{(T)} \hat{\mu}(D)},$$

then $\tilde{\mu} \in \mathcal{P}(S)$, $\tilde{\mu}(D) = 1$ and

$$\mathbf{P}^{(T)}\hat{\mu} \geq \mathbf{P}^{(T)}\hat{\mu}(D)\tilde{\mu} \geq \left(1 - \frac{2\delta}{1-\gamma}\right)\tilde{\mu}.$$

By Fatou's Lemma $\liminf_{t\to\infty} \mathbf{P}^{(t)}\tilde{\mu}(\hat{K}^{\epsilon}) \geq \beta_0$, so by Lemma 6.2.13

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\hat{\mu}(\hat{K}^{\epsilon}) = \liminf_{t\to\infty} \mathbf{P}^{(t)}\mathbf{P}^{(T)}\hat{\mu}(\hat{K}^{\epsilon}) \geq \beta_0\left(1 - \frac{2\delta}{1-\gamma}\right) = \beta.$$

Step 3. For every $\epsilon > 0$ and every $z \in S$ there is a $C \in \mathcal{C}_\epsilon$ for which

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\mu(C) \geq 1 - \epsilon.$$

Let $\epsilon > 0$ and $z \in S$. Let U and $C \in \mathcal{C}_\epsilon$ be given by Step 1. Define

$$\gamma = \sup\{\alpha \geq 0 : \exists N, t_1, \cdots, t_N \geq 0 \text{ such that } \mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z \geq \alpha\nu \\ \text{for some } \nu \in \mathcal{P}(S) \text{ satisfying } \liminf_{t\to\infty} \mathbf{P}^{(t)}\nu(C) \geq 1 - \epsilon\}.$$

We prove that $\gamma = 1$. Note that $\gamma \leq 1$. So assume, contrary to our claim, that $\gamma < 1$.

Let $\tilde{\epsilon} > 0$ be such that $\hat{K}^{\tilde{\epsilon}} \subset U$. Let $\beta \in (0,1)$ be given by Step 2, such that condition (8.17) holds with ϵ replaced with $\tilde{\epsilon}$. If $\gamma > 0$, choose $\alpha \in ((\gamma-\beta)(1-\beta)^{-1}, \gamma) \cap [0,1)$ and else choose $\alpha = 0$. Then there exist $N \in \mathbb{N}$, $t_1, \cdots, t_N \geq 0$, and $\nu \in \mathcal{P}(S)$ such that

$$\mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z \geq \alpha\nu$$

and

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\nu(C) \geq 1 - \epsilon.$$

Set

$$\hat{\mu} = (1-\alpha)^{-1}(\mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z - \alpha\nu)$$

and observe that $\hat{\mu} \in \mathcal{P}(S)$. Further, by (8.17),

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\hat{\mu}(U) \geq \beta,$$

so there is a $T > 0$ such that

$$\mathbf{P}^{(T)}\hat{\mu}(U) \geq \beta/2.$$

Define

$$\nu_1 = \frac{\mathbf{P}^{(T)}\hat{\mu}(\cdot \cap U)}{\mathbf{P}^{(T)}\hat{\mu}(U)}.$$

Then

$$\mathbf{P}^{(T)}\hat{\mu} = (1-\alpha)^{-1}(\mathbf{P}^{(T)}\mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z - \alpha\mathbf{P}^{(T)}\nu) \geq (\beta/2)\nu_1$$

and hence

$$\begin{aligned}&\mathbf{P}^{(T)}\mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z \geq \alpha\mathbf{P}^{(T)}\nu + \tfrac{\beta}{2}(1-\alpha)\nu_1 \\ &= (\alpha + \tfrac{\beta}{2}(1-\alpha))[(\alpha + \tfrac{\beta}{2}(1-\alpha))^{-1}(\alpha\mathbf{P}^{(T)}\nu + \tfrac{\beta}{2}(1-\alpha)\nu_1)].\end{aligned} \tag{8.18}$$

Set $\nu_2 := (\alpha + \frac{\beta}{2}(1-\alpha))^{-1}(\alpha\mathbf{P}^{(T)}\nu + \frac{\beta}{2}(1-\alpha)\nu_1)$. Observe that Step 1 implies that

$$\begin{aligned}\liminf_{t\to\infty}\mathbf{P}^{(t)}\nu_2(C) &\geq (\alpha + \tfrac{\beta}{2}(1-\alpha))^{-1}[\alpha\liminf_{t\to\infty}\mathbf{P}^{(t)}\mathbf{P}^{(T)}\nu(C) \\ &\qquad + \tfrac{\beta}{2}(1-\alpha)\liminf_{t\to\infty}\mathbf{P}^{(t)}\nu_1(C)] \\ &\geq 1-\epsilon.\end{aligned} \tag{8.19}$$

Here we also used Lemma 6.2.13 in order to show that

$$\liminf_{t\to\infty}\mathbf{P}^{(t)}\mathbf{P}^{(T)}\nu(C) = \liminf_{t\to\infty}\mathbf{P}^{(t)}\nu(C) \geq 1-\epsilon.$$

(8.18) and (8.19) yield that $\gamma \geq \alpha + \frac{\beta}{2}(1-\alpha)$. However, α was chosen such that $\alpha + \frac{\beta}{2}(1-\alpha) > \gamma$, so we arrive at a contradiction. Hence $\gamma = 1$. To continue the main line of proof, for any $\alpha < 1$ we find $N \in \mathbb{N}$, $t_1, \cdots, t_N > 0$ and $\nu \in \mathcal{P}(S)$, such that

$$\begin{aligned}\liminf_{t\to\infty}\mathbf{P}^{(t)}\delta_z(C) &= \liminf_{t\to\infty}\mathbf{P}^{(t)}\mathbf{P}^{(t_1)}\mathbf{P}^{(t_2)}\cdots\mathbf{P}^{(t_N)}\delta_z(C) \\ &\geq \alpha\liminf_{t\to\infty}\mathbf{P}^{(t)}\nu(C) \geq \alpha(1-\epsilon),\end{aligned}$$

where the first equality follows from Lemma 6.2.13. Hence $\liminf_{t\to\infty}\mathbf{P}^{(t)}\delta_z(C) \geq 1-\epsilon$, which completes the proof of implication *(i)⇒(iii)* and *(ii)⇒(iii)*.

□

We call a Markov semigroup $(P(t))_{t\geq 0}$ *weakly mean ergodic* if there exists a $\mu_* \in \mathcal{P}(S)$ such that

$$\lim_{t\to\infty}\mathbf{P}^{(t)}\mu = \mu_* \text{ for all } \mu \in \mathcal{P}(S).$$

In [73, Theorem 2] there are given sufficient conditions for weak mean ergodicity. We can use Theorem 8.4.3 to give necessary and sufficient conditions for a Markov semigroup to be weakly mean ergodic that generalise [73, Theorem 2].

Corollary 8.4.4. *Suppose $(P(t))_{t\geq 0}$ satisfies the Cesàro e-property. Then the following are equivalent:*

(i) $(P(t))_{t\geq 0}$ *is weakly mean ergodic.*

(ii) *There exists a* $z \in S$ *such that for every* $\epsilon > 0$

$$\inf_{x\in S} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\epsilon)) > 0. \tag{8.20}$$

(iii) *There exists a* $z \in S$ *such that for every bounded set* A *and every* $\epsilon > 0$

$$\inf_{x\in A} \limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(B_z(\epsilon)) > 0$$

and for all $\eta > 0$ *and* $x \in S$ *there exists a bounded Borel set* D *such that*

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(D) > 1 - \eta.$$

Proof. *(ii)⇒(i)* and *(iii)⇒(i)*: This follows from Theorem 8.4.3 and the observation that if *(ii)* or *(iii)* is satisfied, then condition $(\mathcal{C})$ holds with K replaced with $\{z\}$, so Theorem 8.2.9 implies that there is exactly one ergodic measure.

(i)⇒(ii) and *(i)⇒(iii)*: Let $\mu_* \in \mathcal{P}(S)$ be the invariant probability measure of $(P(t))_{t\geq 0}$. Let $z \in \operatorname{supp}(\mu^*)$. Then for all $\epsilon > 0$ and $\mu \in \mathcal{P}(S)$, by the Portmanteau Theorem,

$$\liminf_{t\to\infty} \mathbf{P}^{(t)}\mu(B_z(\epsilon)) \geq \mu^*(B_z(\epsilon)) > 0.$$

Let $\eta > 0$, then there is a D compact such that $\mu^*(D) > 1-\eta$. Thus for all $\mu \in \mathcal{P}(S)$

$$\liminf_{t\to\infty} \boldsymbol{P}^{(t)}\mu(D) > 1 - \eta.$$

□

We define $(P(t))_{t\geq 0}$ to be *asymptotically stable* if there exists a probability measure μ_* such that $P(t)\mu \to \mu_*$ for all $\mu \in \mathcal{P}(S)$. Then μ_* is a unique invariant probability measure.

Corollary 8.4.5. *Suppose* $(P(t))_{t\geq 0}$ *has the eventual e-property. Then the following are equivalent:*

(i) $(P(t))_{t\geq 0}$ *is asymptotically stable.*

(ii) *There exists a* $z \in S$ *such that for every* $\epsilon > 0$

$$\inf_{x\in S} \liminf_{t\to\infty} P(t)\delta_x(B_z(\epsilon)) > 0.$$

(iii) *There exists a* $z \in S$ *such that for every bounded set* A *and* $\epsilon > 0$

$$\inf_{x\in A} \liminf_{t\to\infty} P(t)\delta_x(B_z(\epsilon)) > 0$$

and for all $\eta > 0$ *and* $x \in S$ *there exists a bounded Borel set* D *such that*

$$\limsup_{t\to\infty} \mathbf{P}^{(t)}\delta_x(D) > 1 - \eta.$$

Proof. *(i)⇒(ii)* and *(i)⇒(iii)*: Let $(P(t))_{t\geq 0}$ be asymptotically stable with invariant probability measure μ_*. Let $z \in \operatorname{supp}(\mu_*)$, then for all $\mu \in \mathcal{P}(S)$, by the Portmanteau Theorem,

$$\liminf_{t\to\infty} P(t)\delta_x(B_z(\epsilon)) \geq \mu_*(B_z(\epsilon)) > 0.$$

Let $\eta > 0$, then there is a D compact such that $\mu^*(D) > 1-\eta$. Thus for all $\mu \in \mathcal{P}(S)$

$$\liminf_{t\to\infty} \boldsymbol{P}^{(t)}\mu(D) > 1 - \eta.$$

(ii)⇒(i) and *(iii)⇒(i)*: Corollary 8.4.4 yields that $(P(t))_{t\geq 0}$ is weakly mean ergodic, so $\Gamma^{\mathbf{P}}_{cp} = S$. Thus by Corollary 7.4.12, $P(t)\mu \to \mu_*$ for all $\mu \in \mathcal{P}(S)$. This completes the proof. □

Example. [73, Theorem 3] shows for a certain class of stochastic differential equations that the associated Markov semigroup satisfies the conditions of [73, Theorem 2], yielding the existence of a unique invariant probability measure and the weak mean ergodicity of the semigroup. We do not want to go into the specifics of these results here, but we do want to mention that if we leave out Assumption (2.14) in [73, Theorem 3], then it follows from the first three steps of the proof of that theorem that the associated Markov semigroup satisfies (8.10) for a certain compact set K. Thus Theorem 8.4.3 implies then that the set of ergodic measures is compact and that the Cesàro averages of every measure converges to an invariant measure.

8.5 Notes

This chapter is based on the submitted paper [115], with some generalisations and new results: several results in [115] are based on the e-property; in this chapter we show that the Cesàro e-property in many cases suffices. Example 8.2.2 is also new. It emerged from a discussion between Tomasz Szarek and Anna Zdunik and was suggested to us by Tomasz Szarek. Finally, the equivalence of Theorem 8.4.3*(ii)* with *(i)* and *(iii)* is new compared to the paper.

BIBLIOGRAPHY

[1] Albiac, F. and N.J. Kalton (2006), *Topics in Banach Space Theory*, Graduate Texts in Mathematics, Vol. 233, Springer, New York. (Pages 33, 62).

[2] Aliprantis, C. and K. Border (2006), *Infinite dimensional analysis. A hitchhiker's guide*, Third edition, Springer, Berlin. (Pages 102, 103, 127).

[3] Aliprantis, C. and O. Burkinshaw (1985), *Positive Operators*, Academic Press, New York. (Pages 60, 61).

[4] Arendt, W., C. Batty, N. Hieber and F. Neubrander (2001), *Vector-valued Laplace transforms and Cauchy problems*, Birkhäuser Verlag, Basel. (Page 121).

[5] Arens, R.F. and J. Eels, Jr. (1956). On embedding uniform and topological spaces, *Pacific J. Math.* **6**, 397-403. (Page 29).

[6] Arveson, W. (1976), *An invitation to C^*-algebras*, Springer-Verlag, New York-Heidelberg. (Pages 168, 170).

[7] Barnsley, M.F. (1988), *Fractals Everywhere*, Academic Press, Boston. (Pages 3, 51).

[8] Bauer, H. (2001), *Measure and Integration Theory*, Walter de Gruyter, Berlin. (Pages 89, 95, 136).

[9] Beboutov, M. (1942), Markov chains with a compact state space, *Rec. Math. (Mat. Sb.)* **10(52)**, 213–238. (Pages 58, 88, 114).

[10] Bierkens, J., O. van Gaans and S. Verduyn Lunel (2009), Existence of an invariant measure for stochastic evolutions driven by an eventually compact semigroup, *J. Evol. Equ.* **9**, 771–786. (Page 117).

[11] Billingsley, P. (1999), *Convergence of probability measures*, Second edition, Wiley, New York. (Pages 78, 152).

[12] Blömker, D. (200), *Amplitude equations for stochastic partial differential equations*, World Scientific Publishing, Singapore. (Page 18).

[13] Bogachev, V.I. (2007), *Measure Theory; Volume II*, Springer, Berlin. (Pages 56, 57, 169).

[14] Chalub, F.A.C.C., Markowich, P.A., Perthame, B. and Ch. Schmeiser (2004). Kinetic Models for Chemotaxis and their Drift-Diffusion Limits, *Monatsh. Math.* **142**, 123–141. (Pages 1, 6, 6).

[15] Chung, K.L. (1967), *Markov Chains with Stationary Transition Probabilities*, Second edition, Springer-Verlag, New York. (Pages 58, 92, 107).

[16] Connors, K.A. (1990), *Chemical Kinetics–The Study of Reaction Rates in Solution*, VCH Publishers, New York. (Page 2).

[17] Conway, J.B. (1990), *A Course in Functional Analysis*, second edition, Springer. (Pages 33, 37).

[18] Costa, O. and F. Dufour (2006), Ergodic properties and ergodic decompositions of continuous-time Markov processes, *J. Appl. Prob.* **43**, 767–781. (Pages 11, 118, 136, 136, 136).

[19] Da Prato, G. and J. Zabczyk (1992), *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, Cambridge. (Pages 4, 117).

[20] Da Prato, G. and J. Zabczyk (1996), *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Note Series **229**, Cambridge University Press, Cambridge. (Pages 4, 5, 11, 43, 58, 117, 127, 127, 127, 127, 127, 127, 127, 128, 137, 144, 144, 157, 162).

[21] Diekmann, O. and Ph. Getto (2005), Boundedness, global existence and continuous dependence for nonlinear dynamical systems describing physiologically structured populations, *J. Diff. Eq.* **215**, 268–319. (Page 6).

[22] Diekmann, O., S.A. van Gils, S.M. Verduyn Lunel and H.-O. Walther (1995), *Delay equations. Functional, complex, and nonlinear analysis*, Springer-Verlag, New York. (Page 2).

[23] Diekmann, O., M. Gyllenberg, H.R. Thieme and S.M. Verduyn Lunel (1993), *A cell-cycle model revisited*, Report AM-R9305, Centrum voor Wiskunde en Informatica, Amsterdam. (Pages 3, 3, 5).

[24] Diekmann, O., M. Gyllenberg, J.A.J. Metz and H.R. Thieme (1998), On the formulation and analysis of general deterministic structured population models; I. Linear Theory, *J. Math. Biol.* **36**, 349–388. (Page 6, 6).

[25] Diekmann, O., M. Gyllenberg, H. Huang, M. Kirkilionis, J.A.J. Metz and H.R. Thieme (2001), On the formulation and analysis of general deterministic structured population models; II. Nonlinear Theory, *J. Math. Biol.* **43**, 157–189. (Page 6, 6).

[26] Diestel, J. and J.J. Uhl jr. (1977), *Vector Measures*, Math. Surveys, Nr. 15, Amer. Math. Soc. (Pages 25, 25, 25, 72, 107).

[27] Dellacherie, C. and P.-A. Meyer (1983), *Probabilités et Potentiel. Chapitres IX à XI: Théorie discrète du potentiel*, Hermann, Paris. (Page 144).

[28] Dudley, R.M. (1966), Convergence of Baire measures, *Stud. Math.* **27**, 251–268. (Pages 8, 14, 14, 16, 20, 20, 24, 24, 25, 32, 32, 34, 42, 139).

[29] Dudley, R.M. (1974), Correction to: "Convergence of Baire measures", *Stud. Math.* **51**, 275. (Page 8).

[30] Dudley, R.M. (2002), *Real Analysis and Probability*, Cambridge University Press, Cambridge. (Page 29).

[31] Dunford, N. and J.T. Schwartz (1958), *Linear operators, Part I*, Wiley-Interscience. (Page 65).

[32] Dynkin, E.B. (1978), Sufficient statistics and extreme points, *Annals of Probability* **6**, 705–730. (Page 87).

[33] Edelstein-Keshet, L. (2005), *Mathematical models in biology*, Society for Industrial and Applied Mathematics, Philadelphia. (Page 2).

[34] Emel'yanov, E.Yu. (2007), *Non-spectral asymptotic analysis of one-parameter operator semigroups*, Operator Theory: Advances and Applications, 173, Birkhuser Verlag, Basel. (Pages 9, 58, 58, 59).

[35] Engel, K.-J. and R. Nagel (2000), *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics, Vol.194, Springer. (Pages 2, 2, 2, 2, 40, 62, 63).

[36] Es-Sarhir, A. (2009), Existence and uniqueness of invariant measures for a class of transition semigroups on Hilbert spaces, *J. Math. Anal. Appl.* **353**, 497–507. (Page 117).

[37] Es-Sarhir, A. and M.-K. von Renesse (2010), Ergodicity of stochastic curve shortening flow in the plane, submitted. (Pages 11, 162).

[38] Ethier, S.N. and T.G. Kurtz (1986), *Markov Processes; Characterization and Convergence*, Wiley, New York. (Pages 3, 4, 7, 43, 50, 54, 55, 58, 59, 89, 90, 90).

[39] Farkas, B., A note on semigroups on spaces of continuous functions and of measures, arXiv:0805.4280. (Page 7).

[40] Feller, W. (1965), *An Introduction to Probability Theory and its Applications*, Vol. II, Wiley, New York. (Pages 4, 59).

[41] Filbet, F., Ph. Laurençot and B. Perthame (2005), Derivation of hyperbolic models for chemosensitive movement, *J. Math. Biol.* **50**, No. 2, 189–207. (Page 6).

[42] Fisher, Y. (1995), *Fractal Image Compression; Theory and Application*, Springer-Verlag, New York. (Page 3).

[43] Foguel, S.R. (1969), *The ergodic theory of Markov processes*, Van Nostrand Mathematical Studies, Vol. 21, New York, London, Toronto, Melbourne. (Pages 58, 111).

[44] Folland, G.B. (1999), *Real analysis: modern techniques and their applications*, New York: John Wiley & Sons Inc. (Pages 57, 67, 129, 136).

[45] Fortet, R. and B. Mourier (1953), Convergence de la répartition empirique vers la répartition théorétique, *Ann. Sci. École Norm. Sup.* **70**, 267–285. (Page 18).

[46] Frasson, M.V.S. (2005), *Large time behaviour of neutral delay systems*, PhD thesis, Leiden University. (Page 3).

[47] Gacki, H. (2007), Applications of the Kantorovich-Rubinstein maximum principle in the theory of Markov semigroups, *Dissertationes Math.* **448**. (Page 50).

[48] Goldstein, J.A. (1985), *Semigroups of operators and applications*, Oxford University Press, New York. (Page 2).

[49] Goldys, B. and J.M.A.M. van Neerven (2003), Transition semigroups of Banach space-valued Ornstein-Uhlenbeck processes, *Acta Appl. Math.* **76**, 283–330. (Page 117).

[50] Greschonig, G. and K. Schmidt (2000), Ergodic decomposition of quasi-invariant probability measures, *Colloq. Math.* **84/85**, 495–514. (Page 114).

[51] Gulick, S.L., T.S. Liu and A.C.M. van Rooij (1967), *Group Algebra Modules II*, Canad. J. Math. **19**, 151-173. (Pages 60, 69, 69, 69, 69, 69, 69, 69, 85).

[52] Hairer, M. and J.C. Mattingly (2006), Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, *Ann. of Math. (2)* **164**, 993–1032. (Page 173, 173, 173, 173, 173).

[53] Hairer, M. and J.C. Mattingly (2008), A Theory of Hypoellipticity and Unique Ergodicity for Semilinear Stochastic PDEs, preprint. (Page 173).

[54] Hale, J.K. (1988), *Asymptotic behavior of dissipative systems*, American Mathematical Society, Providence. (Page 1).

[55] Halmos, P.H. (1950), *Measure Theory*, Van Nostrand, Princeton. (Page 68).

[56] Hernández-Lerma, O. and J.B. Lasserre (1998), Ergodic theorems and ergodic decomposition for Markov chains, *Acta Appl. Math.* **54**, 99–119. (Pages 10, 88, 102, 108, 111, 114).

[57] Hernández-Lerma, O. and J.B. Lasserre (2003), *Markov chains and invariant probabilities*, Birkhäuser, Basel. (Pages 58, 88, 91, 102, 104, 114, 127).

[58] Hille, E. (1948), *Functional Analysis and Semi-Groups*, American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, New York. (Page 2).

[59] Hille, E. and R.S. Phillips (1957), *Functional Analysis and Semi-groups*, revised ed., American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, New York. (Pages 2, 68).

[60] Hille, S.C. (2008), Local well-posedness of kinetic chemotaxis models, *J. Evol. Equ.* **8**, 423–448. (Pages 1, 3, 6).

[61] Hille, S.C. and D.T.H. Worm (2007), Global existence of positive mild solutions to a class of kinetic chemotaxis equations, preprint. (Pages 1, 3, 6).

[62] Hille, S.C. and D.T.H. Worm (2009), Embedding of semigroups of Lipschitz maps into positive linear semigroups on Banach spaces generated by measures, *Integr. Equ. Oper. Theory* **63**, 351–371. (Pages 12, 42, 42, 42).

[63] Hille, S.C. and D.T.H. Worm (2009), Continuity properties of Markov semigroups and their restrictions to invariant L^1-spaces, *Semigroup Forum* **79**, 575–600. (Pages 12, 45, 58, 84, 84, 85).

[64] Horbacz, K. (2004), Random dynamical systems with jumps, *J. Appl. Prob.* **41**, 890–910. (Pages 4, 87).

[65] Horbacz, K., J. Myjak and T. Szarek (2005), On stability of some general random dynamical system, *J. Stat. Phys.* **119**, 35–60. (Pages 4, 87).

[66] Hutchinson, J.E. (1981), Fractals and self similarity, *Indiana Univ. Math. J.* **30**, 713-747. (Pages 3, 3, 51).

[67] Hwang, H.J., K. Kang and A. Stevens (2005), Global solutions of nonlinear transport equations for chemosensitive movement, *Siam J. Math. Anal.* **36**, 1177–1199. (Pages 1, 6).

[68] Johnson, J.A. (1970), Banach spaces of Lipschitz functions and vector-valued Lipschitz functions, *Trans. Amer. Math. Soc.* **148**, 147-169. (Pages 29, 42).

[69] Kakutani, S. (1940), Ergodic theorems and the Markoff process with a stable distribution, *Proc. Imp. Acad. Tokyo* **16**, 49–54. (Pages 58, 90).

[70] Kapica, R., T. Szarek and M. Ślęczka, On a unique ergodicity of some Markov processes, submitted. (Pages 11, 137, 138, 147, 157, 157, 162, 162).

[71] Kazakevičius, V. and R. Leipus (2003), A new theorem on the existence of invariant distributions with applications to ARCH processes, *J. Appl. Probab.* **40**, no. 1, 147–162 (Page 87).

[72] Klünger M., *Ergodic decomposition of invariant measures*, unpublished lecture notes. (Page 87).

[73] Komorowski, T., S. Peszat and T. Szarek (2010), On ergodicity of some Markov processes, *Ann. Probab.* **38**, 1401–1443. (Pages 11, 117, 125, 137, 137, 138, 141, 157, 162, 162, 162, 175, 176, 176, 182, 182, 184, 184, 184).

[74] Kravchenko, A.S. (2006), Completeness of the space of separable measures in the Kantorovich-Rubinstein metric, *Sib. Math. Journal* **47**, 68–76. (Page 29).

[75] Krylov, N. and N. Bogolioubov (1937), La théorie générale de la mesure dans son application à l'étude des systémes de la mécanique non linéaire, *Ann. Math.* **38**, 65–113. (Pages 58, 88, 114).

[76] Kunze, M. (2009), Continuity and equicontinuity of semigroups on norming dual pairs, *Semigroup Forum* **79**, 540–560. (Page 7).

[77] Lant, T. and H.R. Thieme (2007), Markov transition functions and semigroups of measures, *Semigroup Forum* **74**, 337–369. (Pages 5, 7, 43, 54, 58, 66).

[78] Lasota, A. and M.C. Mackey (1994), *Chaos, fractals, and noise. Stochastic aspects of dynamics. Second edition*, Springer Verlag, New York. (Pages 5, 9, 49, 51, 58, 58, 59).

[79] Lasota, A., J. Myjak and T. Szarek (2002), Markov operators with a unique invariant measure, *J. Math. Anal. Appl.* **276**, 343–356. (Page 87).

[80] Lasota, A. and J.A. Yorke (1994), Lower bound technique for Markov operators and iterated function systems, *Random Comput. Dynam.* **2**, 41–77. (Page 145).

[81] Lasota, A. and T. Szarek (2006), Lower bound technique in the theory of a stochastical differential equation, *J. Diff. Eq.* **231**, 513–533. (Pages 11, 117, 137, 137, 138, 153, 155, 162, 162, 162, 163).

[82] Liu, T.S. and A.C.M. van Rooij (1968), Transformation groups and absolutely continuous measures, *Indag. Math.* **30**, 225–231. (Pages 69, 69, 79, 80, 85).

[83] Manca, L. (2008), Kolmogorov equations for measures, *J. Evol. Equ.* **8**, 231–262. (Page 117).

[84] Masmoudi, N. and L.S. Young (2002), Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs, *Commun. Math. Phys.* **227**, 461–481. (Page 163).

[85] Meyer-Nieberg, P. (1991), *Banach lattices*, Springer-Verlag, Berlin. (Pages 35, 60, 60, 61, 61, 61, 79, 80).

[86] Meyn, S.P. and L.R. Tweedie (1993), *Markov Chains and Stochastic Stability*, Springer, London. (Pages 3, 4, 43, 50, 58, 58, 58, 162).

[87] Myjak, J. and T. Szarek (2003), Attractors of iterated function systems and Markov operators, *Abstr. Appl. Anal.* **8**, 479–502. (Pages 50, 51, 87).

[88] Myjak, J. and T. Szarek (2004), On the existence of an invariant measure for Markov-Feller operators, *J. Math. Anal. Appl.* **294**, 215–222. (Page 87).

[89] Nekvinda, A. and L. Zajíček (1988), A simple proof of the Rademacher theorem, *Casopis Pest. Mat.* **113**, no. 4, 337–341. (Page 33).

[90] van Neerven, J.M.A.M. and B. de Pagter (1994), The adjoint of a positive semigroup, *Comp. Math.* **90**, 99–118. (Pages 84, 85).

[91] Ollivier, Y. (2009), Ricci curvature of Markov chains on metric spaces, *J. Funct. Anal.* **256**, 810–864. (Page 87).

[92] Pachl, J. (1979), Measures as functionals on uniformly continuous functions, *Pacific J. Math.* **82**, 515–521. (Pages 14, 31, 31, 42).

[93] de Pagter, B. (1992), A Wiener-Young type theorem for dual semigroups, *Acta Appl. Math.* **27**, 101–109. (Pages 84, 85).

[94] Pazy, A. (1983), *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York. (Page 2).

[95] Peng, J. and Z. Xu (2005), A novel dual approach to nonlinear semigroups of Lipschitz operators, *Trans. Amer. Math. Soc.* **357**, no. 1, 409-424. (Pages 20, 40, 40, 40, 42).

[96] Perthame, B. (2007), *Transport equations in biology*, Birkhäuser, Basel. (Page 2).

[97] Petersen, K. (1983), *Ergodic theory*, Cambridge University Press. (Page 127).

[98] Phelps, R.R. (2001), *Lectures on Choquet's theorem*, second edition, Springer-Verlag, Berlin. (Page 9).

[99] Pichór, K. and R. Rudnicki (1997), Asymptotic behaviour of Markov semigroups and applications to transport equations, *Bull. Polish Acad. Math.* **45**, 379–397. (Pages 9, 58, 59).

[100] Pichór , K. and R. Rudnicki (2000), Continuous Markov semigroups and stability of transport equations, *Journal of Math. An. and Appl.* **249**, 668–685. (Pages 9, 58, 59).

[101] Plessner, A. (1929), Eine Kennzeichnung der totalstetigen Funktionen, *J. fr Reine und Angew. Math* **60**, 26–32. (Pages 9, 79).

[102] Priola, E. (1999), On a class of Markov type semigroups in spaces of uniformly continuous and bounded functions, *Studia Math.* **136**, 271-295. (Page 7).

[103] Raugi, A. (2009), A probabilistic ergodic decomposition result, *Ann. Inst. H. Poincaré Probab. Statist.* **45**, 932–942. (Page 114).

[104] Revuz, D. (1984), *Markov chains, Second edition*, North-Holland Publishing Co., Amsterdam. (Pages 3, 58).

[105] Rosenblatt, M. (1964), Equicontinuous Markov Operators, *Theor. Prob. Appl.* **9**, 205-222 (Pages 58, 162).

[106] Rosenblatt, M. (1967), Transition probability operators, *Proc. Fifth Berkely Symposium Math. Stat. Prob.* **2**, 473–483. (Page 102).

[107] Seidler, J. (2001), *A note on the strong Feller property*, unpublished lecture notes. (Page 144).

[108] Sell, G.R. and You, Y. (2002), *Dynamics of evolutionary equations*, Springer-Verlag, New York. (Pages 1, 2).

[109] Sharpe, M. (1988), *General theory of Markov processes*, Academic Press, London. (Page 136).

[110] Szarek, T. (2000), The stability of Markov operators on Polish spaces, *Studia Math.* **143**, 145–152. (Pages 87, 145).

[111] Szarek, T. (2003), Invariant measures for Markov operators with application to function systems, *Studia Math.* **154**, 207–222. (Pages 51, 87, 145).

[112] Szarek, T. (2003), Invariant measures for nonexpansive Markov operators on Polish spaces, *Dissertationes Math.* **415**. (Pages 51, 58, 87, 145).

[113] Szarek, T. (2008), The uniqueness of invariant measures for Markov operators, *Studia Math.* **189** (3), 225–233. (Pages 87, 137, 137, 138, 148, 156, 156, 162, 162).

[114] Szarek, T., M. Ślęczka and M. Urbański, On stability of velocity vectors for some passive tracer models, accepted for publication in *Bull. London Math. Soc.* (Pages 11, 117, 137, 137, 138, 158, 158, 161, 161, 162, 162, 162, 171, 171).

[115] Szarek, T. and D.T.H. Worm, Ergodic measures of Markov semigroups with the e-property, submitted. (Page 184, 184).

[116] Temam, R. (1997), *Infinite-dimensional dynamical systems in mechanics and physics*, Second edition, Springer-Verlag, New York. (Pages 1, 2).

[117] Thieme, H.R. (2003), *Mathematics in population biology*, Princeton University Press, Princeton. (Page 2).

[118] Thieme, H.R. and J. Voigt (2006), *Stochastic semigroups: their construction by perturbation and approximation* "Proc: Positivity IV–Theory and Applications", (M.R. Weber, J. Voigt, eds.), Dresden, 135–146. (Pages 59, 77).

[119] Varadarajan, V.S., Measures on topological spaces, *Mat. Sb.* **55**, 35-100 (1961); English translation: *Amer. Math. Soc. Transl.* **2**, 161-228 (1965). (Pages 8, 14, 14, 42).

[120] Varadarajan, V.S. (1963), Groups of automorphisms of Borel spaces, *Trans. Amer. Math. Soc.* **109**, 191–220. (Page 114).

[121] Varadhan, S.R.S. (2001), *Probability theory*, Courant Lecture Notes in Mathematics 7. Providence, RI: American Mathematical Society. (Pages 9, 87, 115).

[122] Weaver, N. (1999), *Lipschitz Algebras*, World Scientific Publishing Co. Pte. Ltd. (Pages 15, 18, 20, 29, 42).

[123] Wiener, N. and R.C. Young (1935), *The total variation of* $g(x+h)-g(x)$, Trans. Amer. Math. Soc. **33**, 327–340. (Pages 60, 83).

[124] Williams, D. (1979), *Diffusions, Markov Processes and Martingales. Volume 1: Foundations*, Wiley, Chichester. (Pages 43, 44).

[125] Worm, D.T.H. and S.C. Hille (2009), Ergodic decompositions associated to regular Markov operators on Polish spaces, accepted for publication in *Ergodic Theory and Dynamical Systems*, appeared online: http://dx.doi.org/10.1017/S0143385710000039. (Pages 12, 114).

[126] Worm, D.T.H. and S.C. Hille (2010), An ergodic decomposition defined by regular jointly measurable Markov semigroups on Polish spaces, submitted. (Page 136).

[127] Worm, D.T.H. and S.C. Hille (2010), Equicontinuous families of Markov operators on complete separable metric spaces with applications to ergodic decompositions and existence, uniqueness and stability of invariant measures, submitted. (Pages 42, 162).

[128] Yosida, K. (1948), On the differentiability and the representation of one-parameter semi-group of linear operators, *J. Math. Soc. Japan* **1**, 15-21. (Page 2).

[129] Yosida, K. (1948), Simple Markoff process with a locally compact phase space, *Math. Japonicae* **1**, 99–103. (Pages 88, 88, 114).

[130] Yosida, K. (1980), *Functional analysis*, Sixth edition, Springer-Verlag, Berlin. (Pages 58, 58, 88, 88, 91, 91, 114).

[131] Zaanen, A.C. (1997), *Introduction to operator theory in Riesz spaces*, Springer, Berlin. (Page 60).

[132] Zaharopol, R. (2005), *Invariant probabilities of Markov-Feller operators and their supports*, Birkhäuser, Basel. (Pages 58, 88, 88, 92, 95, 102, 103, 103, 104, 107, 114, 162).

[133] Zaharopol, R. (2007), Equicontinuity and existence of attractive probability measures for some iterated function systems, *Rev. Roumaine Math. Pures Appl.* **52**, 259-286. (Page 51).

[134] Zaharopol, R. (2008), An ergodic decomposition defined by transition probabilities, *Acta. Appl. Math.* **104**, 47–81. (Pages 10, 88, 88, 89, 91, 95, 95, 97, 97, 97, 98, 99, 99, 100, 100, 101, 104, 114).

[135] Zaharopol, R. (2008), Vector integrals and a pointwise mean ergodic theorem for transition functions, *Rev. Roumaine Math. Pures Appl.* **53**, 63–78. (Page 58).

INDEX OF NOTATION

INDEX

SAMENVATTING

Het onderwerp van dit proefschrift, *halfgroepen op ruimten van maten*, bevindt zich in een gedeelte van de wiskunde waar abstracte analyse en kansrekening elkaar ontmoeten. We bekijken het vanuit een analytisch perspectief. Hier volgt een korte toelichting van het onderwerp voor niet-wiskundigen, met aan het eind een korte samenvatting van de inhoud van het proefschrift.

Het onderwerp behoort tot het gebied van dynamische systemen, dat een grote rol speelt in allerlei wetenschappen buiten de wiskunde, zoals biologie, natuurkunde en scheikunde, vooral voor het modelleren van bepaalde processen die in de tijd veranderen. Het idee is dat we een verzameling Ω hebben van mogelijke toestanden, de *toestandsruimte*, en dat we willen beschrijven hoe deze toestanden door de tijd heen bewegen. Als voorbeeld van een biologisch model zou je Ω kunnen kiezen als de mogelijke gewichten van een bepaald organisme, en beschrijven hoe deze in verloop van tijd veranderen, of je kan Ω laten bestaan uit posities, en beschrijven hoe het organisme zich voortbeweegt door Ω. Het uiteindelijk doel is dan om van de beweging van een individu over te gaan naar de beweging van een gehele populatie. Een populatie kan dan niet langer door een element in Ω beschreven worden, maar door een *maat* μ, die aan bepaalde deelverzamelingen E in Ω een getal $\mu(E)$ toekent, dat aangeeft welk gedeelte van de populatie zich in E bevindt. Vaak eisen we meer structuur op Ω, bijvoorbeeld het bestaan van een afstand d, dat voor twee elementen x en y van Ω aangeeft wat de afstand $d(x, y)$ tussen x en y is. In dit geval heet Ω een *metrische ruimte*.

We kijken eerst naar *deterministische* dynamische systemen. Dit zijn dynamische systemen waarbij toeval geen rol speelt. De toestand van het systeem op ieder tijdstip wordt gerepresenteerd door een element x van de toestandsruimte Ω. En er is een bijbehorende beschrijving van de beweging van zo'n element door de toestandsruimte als we de tijd laten oplopen. Dit doen we met behulp van afbeeldingen Φ_t van Ω naar Ω, waarbij $\Phi_t(x)$ aangeeft waar in de toestandsruimte het element x zich na tijd t bevindt. We beschouwen *discrete tijd* (t zit in de verzameling van niet-negatieve gehele getallen $\mathbb{N}_0 = \{0, 1, 2, \cdots\}$) of *continue tijd* ($t$ zit in de verzameling

van alle niet-negatieve getallen $\mathbb{R}_+ = \{x \text{ zodat } x \geq 0\}$). We gaan ervan uit dat de afbeeldingen Φ_t aan de *halfgroep* eigenschap voldoen: Φ_0 houdt alles op zijn plek, en eerst Φ_t en dan Φ_s toepassen is gelijk aan Φ_{t+s} toepassen.

Zodra toeval wel een rol speelt, spreken we van niet-deterministische of *stochastische* dynamische systemen. Als voorbeeld bekijken we een eenvoudig discreet stochastisch dynamisch systeem: we nemen als toestandsruimte Ω de verzameling van alle gehele getallen $\{\cdots, -2, -1, 0, 1, 2, \cdots\}$, die we kunnen zien als een verzameling punten op een horizontale getallenlijn. Stel dat we beginnen met toestand x in Ω. We gooien na één tijdstap een munt op, en bij kop gaat de toestand één stap naar links, en bij munt één stap naar rechts. We kunnen de toestand op een tijdstip in de toekomst dus niet langer beschrijven door een element van de toestandsruimte Ω, maar we moeten het over kansen hebben: de nieuwe toestand is met kans $1/2$ gelijk aan $x - 1$, en met kans $1/2$ gelijk aan $x + 1$. Beginnend bij een element x in Ω, wat is dan de kans dat we na n tijdstappen in een deelverzameling E van Ω zitten? Dit kun je beschrijven met behulp van *kansmaten*. Een kansmaat μ geeft aan bepaalde deelverzamelingen E van Ω een positieve waarde, die we noteren met $\mu(E)$, en die gelijk is aan de kans om in de verzameling E te zitten. Dus $\mu(\Omega) = 1$, want het is zeker dat je ergens in de toestandsruimte zit. Als we zeker in een punt x zitten, dan gebruiken we de notatie δ_x, de zogeheten Dirac-maat op x (dus $\delta_x(E) = 1$ als x in E zit, en 0 als x niet in E zit). Als de begintoestand x is, dan is de kans dat de toestand na één tijdstap in E zit als volgt:

$$\tfrac{1}{2}\delta_{x-1}(E) + \tfrac{1}{2}\delta_{x+1}(E).$$

En als we beginnen met een kansmaat μ, dan kunnen we de kans dat we na één tijdstap in E zitten als volgt beschrijven: met kans $1/2$ wordt er één stap naar links gesprongen. Deze kans moet vermenigvuldigd worden met de kans dat de nieuwe toestand na een sprong naar links in E terechtkomt. Die kans is $\mu(E + 1)$, waarbij $E + 1$ de verzameling van gehele getallen is die je krijgt door de getallen in E één op te hogen. Iets soortgelijks geldt bij het naar rechts springen. Dus de kans dat we na één tijdstap in E zitten, is te schrijven als

$$P\mu(E) := \tfrac{1}{2}\mu(E + 1) + \tfrac{1}{2}\mu(E - 1).$$

$P\mu$ is een nieuwe kansmaat. Nu hebben we een afbeelding P gedefinieerd van de ruimte van kansmaten naar zichzelf. P is een voorbeeld van een *Markov-operator*, oftewel een zogeheten lineaire afbeelding van de ruimte van maten naar zichzelf, die kansmaten afbeeldt op kansmaten. Zo'n Markov-operator geeft middels iteratie een discreet deterministisch dynamisch systeem op de ruimte van maten. Als we P n keer toepassen op de Dirac-maat δ_x, dan krijgen we een kansmaat die we noteren met $P^n\delta_x$, zodat $P^n\delta_x(E)$ de kans geeft dat we na n tijdstappen in Ω zitten als we in x beginnen.

Dit proefschrift gaat over algemene Markov-operatoren en Markov-halfgroepen. Een Markov-halfgroep is een halfgroep van Markov-operatoren $P(t)$ met t in $\mathbb{R}_+$, zodat $P(0)$ alles op zijn plek houdt, en eerst $P(t)$ en dan $P(s)$ toepassen hetzelfde

is als $P(t+s)$ toepassen. Er zijn allerlei toepassingen van Markov-operatoren en Markov-halfgroepen, bijvoorbeeld in de biologie, maar ook binnen de wiskunde, zoals stochastische differentiaalvergelijkingen, maar de focus in dit proefschrift ligt bij de theorie. In de meeste hoofdstukken gaan we ervan uit dat Ω een (bepaald soort) metrische ruimte is.

Nu geven we een kort overzicht van de inhoud:

Hoofdstuk 1 is de introductie, waarin we een uitgebreide samenvatting van dit proefschrift geven. Hoofdstuk 2 gaat over verschillende topologieën op de ruimte van maten. Een topologie geeft een manier om te zeggen wanneer twee elementen dicht bij elkaar zitten. Op de ruimte van maten hoort een natuurlijke topologie, gegeven door de totalevariatie-norm. Deze heeft echter niet bepaalde eigenschappen die we zouden willen hebben, dus bekijken we zwakkere topologieën.

In Hoofdstuk 3 geven we definities en eigenschappen voor Markov-operatoren en Markov-halfgroepen. Hoofdstuk 4 gaat over continuïteit van Markov-halfgroepen: als we beginnen met een maat μ, dan zouden we willen dat $P(t)\mu$ dicht bij μ is als t dicht bij nul zit, bijvoorbeeld in de topologie gegeven door de totalevariatie-norm, omdat dan de Markov-halfgroep een zogeheten C_0-halfgroep is, waar veel theorie voor is ontwikkeld. Vaak is dit niet het geval, maar we kunnen wel kijken naar de deelruimte van maten waarvoor dit wel geldt, en hier eigenschappen over bewijzen.

Een belangrijk concept bij het bestuderen van langetermijngedrag van Markov-operatoren en Markov-halfgroepen wordt gegeven door *invariante kansmaten*, oftewel kansmaten die door de Markov-operator of Markov-halfgroep op zichzelf worden afgebeeld (dus $P\mu = \mu$ of $P(t)\mu = \mu$ voor alle t). Een belangrijk deel binnen de verzameling van invariante kansmaten wordt gegeven door de *ergodische maten*, die je zou kunnen beschrijven als de 'atomen' van de invariante maten: deze kunnen niet worden opgedeeld in twee verschillende invariante maten. Hoofdstuk 5 gaat over ergodische maten voor Markov-operatoren, en geeft onder andere aan hoe we invariante maten kunnen opdelen in ergodische maten. In Hoofdstuk 6 bereiken we soortgelijke resultaten voor Markov-halfgroepen.

Het is van belang voorwaarden te vinden waaronder er invariante maten bestaan. In Hoofdstuk 7 geven we bepaalde condities op Markov-operatoren en Markov-halfgroepen die we hiervoor kunnen gebruiken. Ook vinden we voorwaarden waaronder een invariante kansmaat uniek is. En als er een unieke invariante kansmaat μ bestaat, kunnen we voorwaarden formuleren waaronder μ *stabiel* is: dat wil zeggen dat andere kansmaten uiteindelijk in de tijd naar μ zullen toebewegen.

Uiteindelijk gebruiken we in Hoofdstuk 8 resultaten uit eerdere hoofdstukken om allerlei informatie over de structuur van de verzameling van ergodische maten te verkrijgen. Zo vinden we bijvoorbeeld voorwaarden die impliceren dat er slechts eindig veel ergodische maten bestaan.

Er worden door dit proefschrift ook interessante vragen opgeroepen, die aanleiding kunnen geven voor vervolgonderzoek. Het is bijvoorbeeld van belang om te kijken

in hoeverre de theoretische resultaten in dit proefschrift kunnen worden gebruikt in specifieke toepassingen.

CURRICULUM VITAE

Daniël Worm is geboren te Warnsveld op 10 augustus 1983. Hij heeft tot zijn achttiende gewoond in het kleine dorpje Baak in de Achterhoek. In 2001 behaalde hij zijn gymnasium diploma aan het Baudartius College in Zutphen. In september van dat jaar begon hij zijn studie wiskunde aan de Universiteit Leiden. Hij behaalde in 2002 een dubbele propedeusediploma voor zowel wiskunde als informatica. Onder begeleiding van dr. M.F.E. de Jeu schreef hij zijn afstudeerscriptie *The interplay between flows and C^*-algebras*, waarmee hij in 2006 cum laude afstudeerde. Hij begon in oktober 2006 als Assistent in Opleiding aan het Mathematisch Instituut in Leiden. Het promotieonderzoek, waarvan dit proefschrift het resultaat is, werd begeleid door dr. S.C. Hille, die zijn dagelijkse begeleider was, en prof. dr. S.M. Verduyn Lunel.