Ergodic decompositions associated to regular Markov operators on Polish spaces

Daniël T. H. Worm, Sander C. Hille

Mathematical Institute, University Leiden
P.O. Box 9512, 2300 RA Leiden, The Netherlands
E-mail: dworm@math.leidenuniv.nl, shille@math.leidenuniv.nl

== MI-2009-15 ==

November 13, 2009

Abstract

For any regular Markov operator on the space of finite Borel measures on a Polish space we give a Yosida-type decomposition of the state space, which yields a parametrisation of the ergodic probability measures associated to this operator in terms of particular subsets of the state space. We use this parametrisation to prove an integral decomposition of every invariant probability measure in terms of the ergodic probability measures and give an ergodic decomposition of the state space. This extends results by Yosida (Functional analysis, Chapter XIII.4), Hernández-Lerma and Lasserre (Acta Appl. Math. **54**, 99–119) and Zaharopol (Acta. Appl. Math. **104**, 47–81), who considered the setting of locally compact separable metric spaces.

1 Introduction

Markov operators appear naturally in the context of Markov chains as transition operators. If X_n is the state of the chain at time n, i.e. a random variable that takes values in a measurable space S, and μ is the law of X_0, then the law of X_n is given by $P^n\mu$. Here P is a Markov operator: an additive and positively homogeneous map on the convex cone of positive finite measures on S, that leaves the set of probability measures invariant.Accordingly, the behaviour of the dynamical system in the set of probability measures defined by iteration of a Markov operator is of special interest, in particular the question of existence and characterisation of invariant probability measures for this action.

More structure on S is needed in order to obtain a satisfactory theory on this topic. In the literature one may encounter a line of research that focuses on a pure topological setting (e.g. [21] and references found there) and one that takes a metric perspective, in which S has the generality of a Polish space, i.e.

a topological space that is metrisable for a metric that makes it a complete separable metric space. We pursuit the latter line, driven by applications in population dynamics in biology, in which the state space S typically carries a natural metric. There is much interest lately in Markov operators on not-necessarily locally compact Polish spaces, by e.g. Szarek and coworkers [20, 24, 27, 28] and Ollivier [25]. More specific examples of Markov operators on Polish spaces are given by iterated function systems [26, 23], ARCH processes in econometrics [19] and random dynamical systems on separable Banach spaces [16, 15], 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 $\mathcal{P}_{\rm erg}(S)$. Each invariant measure μ can be represented as an integral over the extreme points, in the sense that

$$\mu(E) = \int_{\mathcal{P}_{\rm erg}(S)} \nu(E)\, d\rho_\mu(\nu) \tag{1}$$

for each Borel set E in S, see e.g. [29, Chapter 6]. Such *ergodic integral decompositions* have been considered in somewhat different formulation in the setting of Polish spaces [3], compact metric spaces [22] and standard spaces [7], using probabilistic arguments. In the pure topological setting results have been obtained for compact Hausdorff spaces using Choquet theory [18, Theorem 4.1.12] (see also M. Klünger, 'Ergodic decomposition of invariant measures', unpublished lecture notes).

There a stronger notion of ergodic decomposition is developed. That is, given an invariant probability measure μ, there exists a decomposition of the state space S into a μ-null set S^o and a disjoint union of invariant measurable subsets S_α, each carrying an ergodic invariant measure μ_α and all parametrised by a Lebesgue space A that possibly depends on μ, such that for all bounded measurable functions f on S,

$$\int_S f\, d\mu = \int_A \left(\int_{S_\alpha} f\, d\mu_\alpha \right) d\alpha. \tag{2}$$

Such a result had been obtained by Yosida in the setting of Markov operators, for which their dual operator maps the space of continuous functions with compact support into itself, on separable metric spaces in which the closed and bounded sets are compact ([31], [32, Chapter XIII.4]), while Hernández-Lerma and Lasserre later covered the more general setting of *regular* Markov operators on a locally compact separable metric space ([11], [12, Chapter 5]). Regular Markov operators are given by transition probabilities on the underlying space S. Thus the Markov operator associated with a time homogeneous Markov chain is regular. Zaharopol [33, 34] managed to extend and strengthen some of their results. In particular, he was able to obtain a partitioning S_α of the state space that does not depend on the particular invariant measure.

In this paper we are able to remove the local compactness condition and generalise Zaharopol's results to the more general setting of regular Markov operators on *Polish spaces*. We extend the definition of a suitable decreasing sequence of subsets of S, $\Gamma_t \supset \Gamma_{cp} \supset \Gamma_{cpi} \supset \Gamma_{cpie}$, which form the preliminary Yosida-type decomposition. These sets do not depend on a pre-chosen invariant measure. However, any invariant measure is concentrated on the smallest set, Γ_{cpie}. We bring the (quite technical) proofs of these results together in Section 3.

In Section 4 we show that an equivalence relation $\sim$ can be defined on Γ_{cpie} such that each equivalence class $[x]$ corresponds uniquely with an ergodic invariant measure ϵ_x. Moreover, we show that any ergodic invariant measure can be obtained in this way, which gives a bijection between $\Gamma_{cpie}/\sim$ and $\mathcal{P}_{erg}(S)$. For each $x \in \Gamma_{cpie}$ 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 the integral decomposition (2) with $A = \Gamma_{cpie}/\sim$ and $\mu_\alpha = \epsilon_{[x]}$ can actually be 'lifted' to

$$\mu = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x),$$

interpreted as Bochner integral in the Banach space $\mathcal{S}_{\mathrm{BL}}$. This space (introduced in [13], see Section 2.1) contains the positive finite measures on S. The relative topology that it induces on the positive measures equals the usual weak topology. This implies a result for μ evaluated at E similar to (1) by using results on Bochner integration in $\mathcal{S}_{\mathrm{BL}}$ (Proposition 2.5, [14]).

In Section 5 we give results on convergence of Cesàro averages of measures, based on some of the sets we define in Section 3 and Section 4.

Many of the results for Markov operators presented in this paper carry over to the setting of a regular Markov *semigroup* on a Polish space, with appropriate modifications in definitions and proofs. A follow-up publication on this case is in preparation. Moreover, in the more restrictive case of *Markov-Feller* operators and semigroups, the sets appearing in the Yosida-type decomposition have better properties, which allows to get a better characterisation of existence and stability of invariant measures. We will report on this subject in a separate paper as well.

Some notational conventions. Unless otherwise mentioned, S will denote a Polish space, viewed as a measurable space with respect to its Borel σ-algebra. We write $\mathcal{M}(S)$ for the real vector space of all signed finite Borel measures on S and $\mathcal{M}^+(S)$ for the cone of positive finite Borel measures on S. $\mathcal{P}(S)$ is the set of probability measures in $\mathcal{M}^+(S)$. $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm on $\mathcal{M}(S)$. We denote by $\mathrm{BM}(S)$ the real vector space of all bounded measurable functions from S to $\mathbb{R}$. $\mathbb{1}_E$ is the indicator function of $E \subset S$. For $f : S \to \mathbb{R}$ measurable and $\mu \in \mathcal{M}(S)$ we write $\langle \mu, f \rangle$ for $\int_S f \, d\mu$ at occasions. $C_b(S)$ is the Banach space of bounded real-valued continuous functions on S, endowed with the supremum norm $\|\cdot\|_\infty$. In the case that S is locally compact we write $C_0(S)$ for the subspace of functions f that vanish at infinity, i.e. for every $\epsilon > 0$ there is a compact $K \subset S$ such that $|f(x)| < \epsilon$ whenever $x \notin K$.

2 Preliminaries

In order to arrive at the ergodic decompositions, we need to generalise results of Zaharopol [34] 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. [2, Remark 5 in §29]. Since every locally compact separable metric space has a countable base, the following holds.

Proposition 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.

Proposition 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, i.e. if $\mu, \mu_1, \mu_2, ... \in \mathcal{P}(X)$ 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 [8, Proposition 3.4.4]: there it is shown that it suffices to check convergence for finite sums of certain bounded Lipschitz functions $f_{i,j}$, where i and j range over $\mathbb{N}$. There are countably many of such sums, which completes the proof. □

2.1 The space $\mathcal{S}_{\mathrm{BL}}$

Let (S, d) be a complete separable metric space. In this part we will recall some definitions and results from [5], [13] and [14]. The aim is to describe a Banach space $\mathcal{S}_{\mathrm{BL}}$ containing the finite signed Borel measures on S and whose topology is connected to that of weak convergence.

$\mathrm{BL}(S)$ denotes the Banach space of bounded real-valued Lipschitz functions for the metric d, endowed with the norm $\|f\|_{\mathrm{BL}} := |f|_{\mathrm{Lip}} + \|f\|_\infty$, where $|f|_{\mathrm{Lip}}$ is the global Lipschitz constant of f. The Dirac functionals $\delta_x(f) := f(x)$ for $x \in S$ are in $\mathrm{BL}(S)^*$. We denote the usual dual norm on $\mathrm{BL}(S)^*$ by $\|\cdot\|^*_{\mathrm{BL}}$.

$\mathrm{BL}(S)$ is in fact isometrically isomorphic to the dual of a separable Banach space $\mathcal{S}_{\mathrm{BL}}$, which can be defined as the closure of the finite linear span of the δ_x, $x \in S$, in $\mathrm{BL}(S)^*$. Then, as shown in [5, Lemma 6], each $\mu \in \mathcal{M}(S)$ defines a unique element in $\mathrm{BL}(S)^*$, which we will also denote by μ, by sending $f \in \mathrm{BL}(S)$ to $\langle \mu, f\rangle = \int_S f\, d\mu$. A function $f \in \mathrm{BL}(S)$ defines a bounded linear functional on $\mathcal{S}_{\mathrm{BL}}$ by sending ϕ to $\phi(f)$. Using [13, Lemma 3.5] one can show that the map $x \mapsto \delta_x$ is a continuous embedding from S into $\mathcal{S}_{\mathrm{BL}}$.

By [13, Theorem 3.9 and Corollary 3.10], $\mathcal{M}^+(S)$ is a closed convex cone of $\mathcal{S}_{\text{BL}}$, and $\mathcal{M}(S)$ is a $\|\cdot\|_{\text{BL}}^*$-dense subspace of $\mathcal{S}_{\text{BL}}$.

The restriction of the weak-star topology on $C_b(S)^*$ to $\mathcal{M}^+(S)$, also called the topology of weak convergence on $\mathcal{M}^+(S)$, *equals* the restriction of the norm topology on $\mathcal{S}_{\text{BL}}$ to $\mathcal{M}^+(S)$ by [5, Theorem 18]. In particular the following lemma holds:

Lemma 2.3. *Let $\mu_n, \mu \in \mathcal{M}^+(S)$. Then $\|\mu_n - \mu\|_{\text{BL}}^* \to 0$ if and only if $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f \in C_b(S)$.*

Let
$$\mathcal{S}_{\text{BL}}^+ := \{\phi \in \mathcal{S}_{\text{BL}} : \phi(f) \geq 0 \text{ for all } f \in \text{BL}(S), f \geq 0\}.$$
Then $\mathcal{S}_{\text{BL}}^+ = \mathcal{M}^+(S)$ by [13, Corollary 4.2].

The following results come from [14, Proposition 2.5, Proposition 2.6 and Corollary 2.7]. Let (Ω, Σ) be a measurable space.

Proposition 2.4. *Let $p : \Omega \to \mathcal{S}_{\text{BL}}^+$. The following conditions are equivalent:*

(i) p is strongly measurable.

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

Proposition 2.5. *Let $\mu \in \mathcal{M}^+(\Omega)$ and let $p : \Omega \to \mathcal{S}_{\text{BL}}^+$ be Bochner integrable with respect to μ. 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 $f \in \text{BM}(S)$. In particular, for any Borel set $E \subset S$,
$$\left[\int_\Omega p(\omega) d\mu(\omega)\right](E) = \int_\Omega p(\omega)(E) d\mu(\omega).$$

Corollary 2.6. *For any $\mu \in \mathcal{M}^+(S)$, $\int_S \delta_x\, d\mu(x) = \mu$, as a Bochner integral in $\mathcal{S}_{\text{BL}}$.*

Proof. Since $x \mapsto \delta_x$ is bounded and continuous from S to $\mathcal{S}_{\text{BL}}$, and $\mathcal{S}_{\text{BL}}$ is separable, this map is Bochner integrable with respect to μ. For every $f \in \mathcal{S}_{\text{BL}}^* \cong \text{BL}(S)$ we have, according to Proposition 2.5,
$$\langle \int_S \delta_x\, d\mu(x), f\rangle = \int_S \langle \delta_x, f\rangle\, d\mu(x) = \int_S f(x)\, d\mu(x) = \langle \mu, f\rangle.$$
Thus $\int_S \delta_x\, d\mu(x) = \mu$. □

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 $\epsilon > 0$.

For $E \subset S$ and $\epsilon > 0$ we define $E_\epsilon := \{x \in S : d(x, E) < \epsilon\}$. Then the following holds:

Theorem 2.7. *Let $M \subset \mathcal{P}(S)$. Then the following 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 [8, Theorem 3.2.2]. And the equivalence between *(i)* and *(iii)* follows from Prokhorov's Theorem and Lemma 2.3.

Remark 2.8. While we looked at a specific metric on S, we will show that most of the concepts above do not depend on the particular metric. Suppose that S is a Polish space and let d be a complete metric on S metrising the given topology. 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-star topology of $C_b(S)^*$ 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 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, 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 $\mu \in \mathcal{M}^+(\Omega)$ does not depend on the specific metric. Furthermore, it follows from Proposition 2.5 that the value of the Bochner integral also does not depend on d.

2.2 Regular Markov operators

Let S be a Polish space. Let d be a complete metric on S that metrises the given topology and $\mathcal{S}_{\mathrm{BL}}$ the Banach space associated with (S, d).

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

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

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

Since $(\mathcal{M}(S), \|\cdot\|_{\mathrm{TV}})$ is a Banach lattice, condition *(MO1)* ensures that a Markov operator P extends to a positive *bounded* linear operator on $(\mathcal{M}(S), \|\cdot\|_{\mathrm{TV}})$ given by $P\mu := P(\mu^+) - P(\mu^-)$. The operator norm of this extension is

$$\|P\| = \sup\{\|P\mu\|_{\mathrm{TV}} : \mu \in \mathcal{M}^+(S),\ \|\mu\|_{\mathrm{TV}} \leq 1\} = 1$$

according to *(MO2)*.

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

Following [10, 14, 23], we will call a Markov operator P *regular* if there exists a map $U : \mathrm{BM}(S) \to \mathrm{BM}(S)$, the *dual* of P, such that $\langle P\mu, h\rangle = \langle \mu, Uh\rangle$ for every $\mu \in \mathcal{M}(S), h \in \mathrm{BM}(S)$.

For $n \in \mathbb{N}$ we define the *Cesàro averages*

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

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

The following important ergodic theorem was proven by Kakutani [17, Theorem 1].

Theorem 2.9. *(Kakutani's Ergodic Theorem) Let P be a regular Markov operator 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 $(U^{(n)}f)_n$ converges pointwise μ-a.e. to g and $\langle \mu, f \rangle = \langle \mu, g \rangle$.*

If P is regular, we can bring it inside integrals:

Proposition 2.10. *Let (Ω, Σ) be a measurable space and $h : \Omega \mapsto \mathcal{S}_{\mathrm{BL}}$ Bochner integrable with respect to $\mu \in \mathcal{M}^+(\Omega)$. Then, for any regular Markov operator P the following holds:*

$$P \int_\Omega h(\omega)\, d\mu(\omega) = \int_\Omega Ph(\omega)\, d\mu(\omega).$$

Proof. By Proposition 2.5 we obtain for $f \in \mathrm{BM}(S)$

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

□

Corollary 2.11. *Let P be a regular Markov operator and $\mu \in \mathcal{M}^+(S)$. Then $x \mapsto P\delta_x$ is a strongly measurable map from S to $\mathcal{S}_{\mathrm{BL}}$ and*

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

Moreover, for every Borel set $E \subset S$

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

Proof. By Corollary 2.6, $\int_S \delta_x\, d\mu(x) = \mu$. For every $f \in \mathrm{BM}(S)$, $x \mapsto \langle P\delta_x, f \rangle = Uf(x)$ is Borel measurable. Thus by Proposition 2.4, $x \mapsto P\delta_x$ is strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$, hence Bochner integrable with respect to μ. We can now apply Proposition 2.10 and Proposition 2.5. □

We shall write $\mathcal{B}(S)$ to denote the σ-algebra of all Borel sets of S. A map $p : S \times \mathcal{B}(S) \to \mathbb{R}$ is a *transition probability* if:

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

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

It follows from conditions $(T1)$ and $(T2)$ that every transition probability p *generates* a Markov operator P, by $P\mu(E) := \int_S p(x, E)\, d\mu(x)$. Note that then $P\mu = \int_S p_x\, d\mu(x)$ in Bochner sense by Proposition 2.4. P is regular: let $f \in \mathrm{BM}(S)$, then by Proposition 2.5

$$\langle P\mu, f\rangle = \langle \int_S p_x\, d\mu(x), f\rangle = \int_S \langle p_x, f\rangle\, d\mu(x) = \langle \mu, Uf\rangle,$$

where $Uf(x) = \langle p_x, f\rangle = \int_S f\, dp_x$. The converse also holds:

Proposition 2.12. *Any regular Markov operator is generated by a transition probability.*

Proof. Let $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ be a regular Markov operator. Define $p(x, E) := P\delta_x(E)$ for all $x \in S, E \in \mathcal{B}(S)$. Then p defines a transition function and $P\mu(E) = \int_S p(x, E)\, d\mu(x)$ for every $E \in \mathcal{B}(S)$ by Corollary 2.11. □

An important class of examples of regular Markov operators is given by measurable maps from S to S: Let $\Phi : S \to S$ be measurable. Then it is straightforward to show that $P_\Phi \mu := \mu \circ \Phi^{-1}$ defines a regular Markov operator.

We will also need the following result on the dual U of a regular Markov operator:

Lemma 2.13. *Let μ 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, so

$$\begin{aligned} \int_S |Uf - Ug|\, d\mu &= \langle \mu, |U(f - g)|\rangle \leq \langle \mu, U|f - g|\rangle \\ &= \langle P\mu, |f - g|\rangle = \langle \mu, |f - g|\rangle = 0. \end{aligned}$$

Thus $Uf = Ug$ μ-a.e.

(ii) Let $f \in \mathrm{BM}(S)$, then

$$\begin{aligned} \|Uf\|_1 &= \int_S |Uf|\, d\mu \leq \langle \mu, U|f|\rangle \\ &= \langle P\mu, f^+\rangle + \langle P\mu, f^-\rangle = \langle \mu, |f|\rangle = \int_S |f|\, d\mu. \end{aligned}$$

□

2.3 Markov-Feller operators

A regular Markov operator P is a *Markov-Feller operator* if the dual U leaves $C_b(S)$ invariant.

Proposition 2.14. *A Markov operator P is a Markov-Feller operator if and only if $P : \mathcal{S}_{\mathrm{BL}}^+ \to \mathcal{S}_{\mathrm{BL}}^+$ is continuous.*

Proof. Suppose P is regular such that U leaves $C_b(S)$ invariant. Let $\mu_n, \mu \in \mathcal{M}^+(S)$ such that $\|\mu_n - \mu\|_{\mathrm{BL}}^* \to 0$, then by Lemma 2.3 $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. Thus

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

for every $f \in C_b(S)$. Thus, again by Lemma 2.3, $\|P\mu_n - P\mu\|_{\mathrm{BL}}^* \to 0$.

The other direction follows from [14, Lemma 3.3]. □

Let $\Phi : S \to S$ be measurable. Then the regular Markov operator P_Φ is a Markov-Feller operator if and only if Φ is continuous.

3 A preliminary Yosida-type decomposition

From now on, we let S be a Polish space and $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$ a regular Markov operator with dual U. 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 $P^{(n)}\delta_x$. These are interesting in their own right, and will prove to be an important ingredient in proving the ergodic decompositions in Section 4. Some of these sets are generalisations in the setting of Polish spaces of those formulated by Yosida in [32, Chapter XIII, section 4] and [31] in order to obtain an ergodic decomposition of state space, which is why we call the decomposition of S into these sets a preliminary Yosida-type decomposition. Zaharopol [34] extended this decomposition to the setting of regular Markov operators of locally compact separable metric spaces, and 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 given by

$$\Gamma_t := \{x \in S : (P^{(n)}\delta_x)_n \text{ is tight}\}.$$

So Γ_t consists of those $x \in S$ such that there is a subsequence n_k for which $P^{(n_k)}\delta_x$ converges in $\mathcal{S}_{\mathrm{BL}}$ as $k \to \infty$. Thus it is natural to consider

$$\Gamma_{cp} := \{x \in S : (P^{(n)}\delta_x)_n \text{ converges in } \mathcal{S}_{\mathrm{BL}}\}.$$

For $x \in \Gamma_{cp}$ we write ϵ_x to denote the limit of $(P^{(n)}\delta_x)_n$, which is a probability measure.

If P is Markov-Feller, then ϵ_x is an invariant probability measure for every $x \in \Gamma_{cp}$. This follows from the fact that by Proposition 2.14 for every $x \in \Gamma_{cp}$

$$\begin{aligned} P\epsilon_x &= P \lim_{n\to\infty} P^{(n)}\delta_x = \lim_{n\to\infty} PP^{(n)}\delta_x = \lim_{n\to\infty} \left[P^{(n)} + \frac{1}{n} - \frac{P^{n+1}}{n}\right]\delta_x \\ &= \lim_{n\to\infty} P^{(n)}\delta_x = \epsilon_x. \end{aligned}$$

If P is not a Markov-Feller operator, then the measures ϵ_x may not be invariant for any $x \in \Gamma_{cp}$. An example is given in [12, Example 5.2.5]. Another set of relevance therefore is

$$\Gamma_{cpi} := \{x \in \Gamma_{cp} : \epsilon_x \text{ is invariant}\}.$$

Finally we will consider a certain subset $\Gamma_{cpie} \subset \Gamma_{cpi}$. We postpone its definition here (see (8)), because it requires some concepts that will be defined later on. In Section 4.1 it will be shown that Γ_{cpie} consists of exactly those $x \in \Gamma_{cpi}$ for which ϵ_x is ergodic. Obviously

$$\Gamma_{cpie} \subset \Gamma_{cpi} \subset \Gamma_{cp} \subset \Gamma_t. \tag{3}$$

We will show in this section that these sets are all Borel measurable and – more importantly – that for *every* invariant probability measure μ,

$$\mu(\Gamma_t) = \mu(\Gamma_{cp}) = \mu(\Gamma_{cpi}) = \mu(\Gamma_{cpie}) = 1. \tag{4}$$

Of course, there need not be any invariant probability measure. In that case Γ_{cpi} and Γ_{cpie} are empty. There exist examples of regular Markov operators however, for which there are no invariant probability measures, while Γ_{cp} and Γ_t are non-empty.

In order to prove (4) it would suffice to show that $\mu(\Gamma_{cpie}) = 1$. However, technically we cannot achieve this directly. Instead we proceed stepwise in the chain (3) 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 2.9).

In order to deal with Γ_t, 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

$$f_E^\epsilon(x) := \left(1 - \frac{d(x,E)}{\epsilon}\right)^+.$$

This function is in $\mathrm{BL}(S)$. In particular, $0 \leq f_E^\epsilon \leq 1$ and $|f_E^\epsilon|_{\mathrm{Lip}} = \frac{1}{\epsilon}$. Also, $f(x) = 0$ for every $x \notin E_\epsilon$. If $E \subset F \subset S$ then $d(x,F) \leq d(x,E)$, so $f_E^\epsilon \leq f_F^\epsilon$. Moreover, if $\epsilon \leq \epsilon'$ then $f_F^\epsilon \leq f_F^{\epsilon'}$. Recall the notation used in Theorem 2.7.

Theorem 3.1. *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

$$\mu(F_\epsilon) \geq 1 - \epsilon \text{ for every } \mu \in M.$$

(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

$$\langle \mu, f_F^{\frac{1}{m}} \rangle > 1 - \frac{1}{m} \text{ for every } \mu \in M.$$

Proof. *(ii)⇒(i)*: Follows from Theorem 2.7 since every finite set is compact.

(i)⇒(ii): Let $\epsilon > 0$. Then there is, by Theorem 2.7 a compact $K \subset S$, such that $\mu(K_{\frac{\epsilon}{2}}) \geq 1 - \frac{\epsilon}{2}$. Since K is compact and $D \subset S$ is dense, there exists a finite $F \subset D$, such that $K \subset \cup_{x \in F} B_x(\frac{\epsilon}{2})$. Then $K_{\frac{\epsilon}{2}} \subset F_\epsilon$, hence for every $\mu \in M$

$$\mu(F_\epsilon) \geq \mu(K_{\frac{\epsilon}{2}}) \geq 1 - \frac{\epsilon}{2} > 1 - \epsilon.$$

(ii)⇒(iii): One can easily verify that, for any $\epsilon > 0$, one can choose $0 < \delta < \epsilon$ such that

$$\frac{\delta}{\epsilon} + \delta - \frac{\delta^2}{\epsilon} < \epsilon. \tag{5}$$

By *(ii)* there exists a finite $F \subset D$, such that $\mu(F_\delta) \geq 1 - \delta$ for all $\mu \in M$. If $x \in F_\delta$, then

$$f_F^\epsilon(x) = 1 - \frac{d(x, F)}{\epsilon} \geq 1 - \frac{\delta}{\epsilon} > 0.$$

Thus, by (5) we obtain for every $\mu \in M$

$$\begin{aligned} \langle \mu, f_F^\epsilon \rangle &\geq (1 - \frac{\delta}{\epsilon})\mu(F_\delta) \geq (1 - \frac{\delta}{\epsilon})(1 - \delta) \\ &= 1 - \frac{\delta}{\epsilon} - \delta + \frac{\delta^2}{\epsilon} > 1 - \epsilon. \end{aligned}$$

(iii)⇒(ii): Let $\epsilon > 0$. Then there is a finite subset $F \subset D$ such that $\langle \mu, f_F^\epsilon \rangle \geq 1 - \epsilon$. Then $\mu(F_\epsilon) \geq \langle \mu, f_F^\epsilon \rangle > 1 - \epsilon$.

(iii)⇒(iv): This is trivial.

(iv)⇒(iii): Let $\epsilon > 0$. Then there is an $m \in \mathbb{N}$ such that $\frac{1}{m} < \epsilon$. By assumption there is a finite subset $F \subset D$ such that $\langle \mu, f_F^{\frac{1}{m}} \rangle > 1 - \frac{1}{m} > 1 - \epsilon$. Now, $f_F^\epsilon \geq f_F^{\frac{1}{m}}$ thus $\langle \mu, f_F^\epsilon \rangle > 1 - \epsilon$. □

Proposition 3.2. *Γ_t is a Borel set and $\mu(\Gamma_t) = 1$ 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

$$\begin{aligned} K_{F,m,n} &:= \{x \in S : \langle P^{(n)}\delta_x, f_F^{\frac{1}{m}} \rangle > 1 - \frac{1}{m}\} \\ &= \{x \in S : U^{(n)} f_F^{\frac{1}{m}}(x) > 1 - \frac{1}{m}\}. \end{aligned}$$

By Theorem 3.1 we have $\Gamma_t = \cap_{m\in\mathbb{N}} \cup_{F\in\mathcal{F}} \cap_{n\in\mathbb{N}} K_{F,m,n}$.

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 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 $m \in \mathbb{N}$. This implies that $\mu(\Gamma_t) = 1$.

Fix $m \in \mathbb{N}$ and $0 < \delta < 1$. Because (S, d) is a complete separable metric space, $\{\mu\}$ is tight by Theorem 2.7. Thus there exists an $F_0 \in \mathcal{F}$, depending on δ, m and μ, such that

$$\langle \mu, f_{F_0}^{\frac{1}{m}} \rangle \geq \langle \mu, f_{F_0}^{\frac{\delta}{m}} \rangle \geq 1 - \frac{\delta}{m}.$$

For convenience, put $f := f_{F_0}^{\frac{1}{m}}$.

By Kakutani's Ergodic Theorem (Theorem 2.9) 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 $x \in S$

$$g(x) \leq (1 - \frac{1}{2m}) \mathbb{1}_A(x) + \mathbb{1}_{S\setminus A}.$$

Therefore

$$1 - \frac{\delta}{m} \leq \langle \mu, f \rangle = \langle \mu, g \rangle \leq \mu(A)(1 - \frac{1}{2m}) + (1 - \mu(A)).$$

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 $U^{(n)} f(x) > 1 - \frac{1}{m}$ for every $n > N$. The finite set of measures $\{\delta_x, P^{(1)}\delta_x, ..., P^{(N)}\delta_x\}$ is tight, so by Theorem 3.1*(iv)* there exists an $F_1 \in \mathcal{F}$ such that $U^{(n)} f_{F_1}^{\frac{1}{m}}(x) > 1 - \frac{1}{m}$ for $1 \leq n \leq N$. Now put $F := F_0 \cup F_1$. Then $F \in \mathcal{F}$ and $f_F^{\frac{1}{m}} \geq f_{F_i}^{\frac{1}{m}}$ for $i = 0, 1$. Thus $U^{(n)} f_F^{\frac{1}{m}}(x) > 1 - \frac{1}{m}$ 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 $\mu(\Gamma_t) = 1$. □

We consider the set $\Gamma_{cp} = \{x \in S : P^{(n)}\delta_x \text{ converges in } \mathcal{S}_{\mathrm{BL}}\}$.

Lemma 2.3 implies that $x \in \Gamma_{cp}$ 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} does not depend on the choice of the metric.

If S is a locally compact separable metric space, then the definition in [33, 34] of the set Γ_{cp} can be written as follows: $x \in \Gamma_{cp}$ 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 [2, 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}.

Our aim is to show that Γ_{cp} is a Borel set and that $\mu(\Gamma_{cp}) = 1$ for every invariant probability measure μ, which will extend [34, Proposition 5.4 and Proposition 5.9]. We first need some preliminary results.

Let $\{f_k : k \in \mathbb{N}\}$ be the countable subset of $\mathrm{BL}(S)$ from Proposition 2.2. Define

$$\Gamma_c = \{x \in S : \langle P^{(n)}\delta_x, f_k\rangle \text{ converges for every } k \in \mathbb{N}\}.$$

Proposition 3.3. *Γ_c is a Borel set and $\mu(\Gamma_c) = 1$ for every invariant probability measure μ.*

Proof. We can write

$$\Gamma_c = \{x \in S : (U^{(n)}f_k(x))_n \text{ converges for every } k \in \mathbb{N},\} = \bigcap_{m\in\mathbb{N}} \bigcap_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} H_{m,n,k}$$

where

$$H_{m,n,k} := \{x \in S : |U^{(n_1)}f_k(x) - U^{(n_2)}f_k(x)| \leq \frac{1}{m} \text{ whenever } n_1, n_2 \geq n.\}$$

Now,

$$H_{m,n,k} = \bigcap_{n_1 \geq n} \bigcap_{n_2 \geq n} \{x \in S : |U^{(n_1)}f_k(x) - U^{(n_2)}f_k(x)| \leq \frac{1}{m}|\}.$$

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 is Borel measurable as well.

Let μ be an invariant probability measure. By Kakutani's Ergodic Theorem (Theorem 2.9) 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 $C \subset \Gamma_c$, so $\mu(\Gamma_c) = 1$. □

Note that $\Gamma_{cp} \subset \Gamma_c$, but they need not be equal. However, we do have the following:

Proposition 3.4. *$\Gamma_{cp} = \Gamma_t \cap \Gamma_c$. Consequently Γ_{cp} is Borel measurable and $\mu(\Gamma_{cp}) = 1$ for every invariant probability measure μ.*

Proof. It is obvious that $\Gamma_{cp} \subset \Gamma_t \cap \Gamma_c$.

Now take $x \in \Gamma_t \cap \Gamma_c$. Since $x \in \Gamma_t$, 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 $\langle P^{(n_m)}\delta_x, f\rangle \to \langle \mu, f\rangle$ for every $f \in C_b(S)$. Because $x \in \Gamma_c$ 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 2.2 we know that $\|P^{(n)}\delta_x - \mu\|^*_{\mathrm{BL}} \to 0$ as $n \to \infty$, so $x \in \Gamma_{cp}$. So Proposition 3.2 and Proposition 3.3 imply that Γ_{cp} is Borel measurable and $\mu(\Gamma_{cp}) = 1$ for every invariant probability measure μ. □

For $f \in \mathrm{BM}(S)$, let

$$A_f := \{x \in S : (U^{(n)}f(x))_n \text{ converges.}\} \tag{6}$$

Then it can easily be shown that A_f is Borel measurable. Kakutani's Ergodic Theorem (Theorem 2.9) implies that $\mu(A_f) = 1$ for every invariant probability measure μ. We can define

$$f^*(x) := \begin{cases} \lim_{n\to\infty} U^{(n)}f(x) & \text{if } x \in A_f \cap \Gamma_{cp} \\ 0 & \text{if } x \notin A_f \cap \Gamma_{cp}. \end{cases}$$

Then f^* is measurable as the pointwise limit of a sequence of measurable functions. Observe that by Proposition 3.4, $U^{(n)}f \to f^*$ μ-a.e. for every invariant probability measure μ, thus f^* plays the role of the g in Kakutani's Ergodic Theorem and does not depend on the specific invariant probability measure.

If $f \in C_b(S)$, then $\Gamma_{cp} \subset A_f$, since $U^{(n)}f(x) = \langle P^{(n)}\delta_x, f\rangle \to \langle \epsilon_x, f\rangle$, so in this case

$$f^*(x) = \begin{cases} \lim_{n\to\infty} U^{(n)}f(x) & \text{if } x \in \Gamma_{cp} \\ 0 & \text{if } x \notin \Gamma_{cp}. \end{cases}$$

For $f \in \mathrm{BM}(S)$, we can define the function

$$f^\square(x) := \begin{cases} \langle \epsilon_x, f\rangle & \text{if } x \in \Gamma_{cp} \\ 0 & \text{if } x \notin \Gamma_{cp}. \end{cases}$$

Clearly if $f \in C_b(S)$ then $f^\square = f^*$.

In [34], f^* and $f^\square$ are also defined, though their definitions differs slightly from ours. As in [34], the functions f^* and $f^\square$ play an important role in the proof of the Borel measurability of Γ_{cpi}. We first show some properties of these functions, analogous to those shown in [34], 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 $f^\square$ also plays the role of the g in Kakutani's Ergodic Theorem for any invariant probability measure μ.

Lemma 3.5. *For every $f \in \mathrm{BM}(S)$, $f^\square$ is in* $\mathrm{BM}(S)$.

Proof. We define the following map from S to $\mathcal{S}_{\mathrm{BL}}$:

$$\eta(x) := \begin{cases} \epsilon_x & \text{if } x \in \Gamma_{cp} \\ 0 & \text{if } x \notin \Gamma_{cp}. \end{cases}$$

We will show that η is strongly measurable. Then it follows from Proposition 2.4 that $x \mapsto \langle \eta(x), f\rangle$ is measurable for every $f \in \mathrm{BM}(S)$, and $\langle \eta(x), f\rangle = f^{\square}(x)$ for every $x \in S$.

First we define the map $\eta_n : S \to \mathcal{S}_{\mathrm{BL}}$ by

$$\eta_n(x) := \begin{cases} P^{(n)}\delta_x & \text{if } x \in \Gamma_{cp} \\ 0 & \text{if } x \notin \Gamma_{cp}. \end{cases}$$

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 $\langle \eta_n(x), g\rangle = \mathbb{1}_{\Gamma_{cp}}(x)U^{(n)}g(x)$, so from measurability of Γ_{cp} (Proposition 3.4) 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. $\square$

The proof of the following lemma is based on that of the locally compact version [34, Lemma 5.10].

Lemma 3.6. *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 $|f_n(x)| \le M$ for every $n \in \mathbb{N}, x \in S$.

(ii) The sequence $(f_n)_n$ converges pointwise to f

(iii) $f_n^* = f_n^{\square}$ μ-a.e. for every $n \in \mathbb{N}$.

Then $f^* = f^{\square}$ μ-a.e.

Proof. Clearly $|f(x)| \le M$ for every $x \in S$, thus $|f_n(x) - f(x)| \le 2M$. Therefore $\int_S |f_n - f|\, d\mu \to 0$ by the Dominated Convergence Theorem. By Lemma 2.13

$$\int_S |U^{(m)} f_n - U^{(m)} f|\, d\mu \le \int_S |f_n - f|\, d\mu \text{ for every } m, n \in \mathbb{N}. \tag{7}$$

Let $h \in \mathrm{BM}(S)$, then $\mu(A_h) = 1$, where A_h is the Borel set defined as in (6). So Proposition 3.4 implies that $A := \Gamma_{cp} \cap A_f \cap (\cap_{n\in\mathbb{N}} A_{f_n})$ 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 (7) it follows that $\int_S |f_n^* - f^*|\, d\mu \le \int_S |f_n - f|\, d\mu \to 0$ 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 $x \in \Gamma_{cp}$. Thus $f_n^{\square} \to f^{\square}$ pointwise. Now,

$$|f_n^{\square}(x) - f^{\square}(x)| = |\langle \epsilon_x, f_n - f\rangle| \le 2M\|\epsilon_x\|_{\mathrm{TV}} = 2M,$$

thus again by the Dominated Convergence Theorem $\int_S |f_n^{\square} - f^{\square}|\, d\mu \to 0$.

Note that $\int_S |f_n^* - f_n^\square| \, d\mu = 0$ for every $n \in \mathbb{N}$ by assumption. Thus

$$\int_S |f^* - f^\square| \, d\mu \leq \int_S |f^* - f_n^*| \, d\mu + \int_S |f_n^\square - f^\square| \, d\mu \to 0$$

as $n \to \infty$. Therefore $f^* = f^\square$ μ-a.e. □

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 [34, Proposition 5.11], since that approach is based on a result, [34, Proposition 2.2], that uses the local compactness of the state space. Instead, we use the Monotone Class Theorem for functions, which we state here for convenience (see e.g. [30, Theorem II.4]).

Theorem 3.7. *Let $\mathcal{E}$ be a π-system for S and let $\mathcal{H}$ be a vector space of functions from S 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 $\mathbb{1}_S$,

(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 $\mathcal{E}$.

Proposition 3.8. *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 3.6, 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 3.6. 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 $\mathcal{H} = \mathrm{BM}(S)$. □

Proposition 3.9. *Γ_{cpi} is Borel measurable and $\mu(\Gamma_{cpi}) = 1$ for every invariant probability measure μ.*

Proof. Let $\{f_n : n \in \mathbb{N}\} \subset \mathrm{BL}(S)$ be the countable subset given by Proposition 2.2, and let $B_n := \{x \in \Gamma_{cp} : f_n^\square(x) = (Uf_n)^\square(x)\}$, then B_n is measurable, since $f_n^\square$ and $(Uf_n)^\square$ are measurable by Lemma 3.5. Let $x \in \Gamma_{cp}$. According to Proposition 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 $x \in \Gamma_{cpi}$ if and only if $(Uf_n)^{\Box}(x) = f_n^{\Box}(x)$ for every $n \in \mathbb{N}$. So $\Gamma_{cpi} = \bigcap_{n=1}^{\infty} B_n$ 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 $n \to \infty$, thus $f_n^* = (Uf_n)^*$. Since $f_n \in C_b(S)$, $f_n^* = f_n^{\Box}$, so $(Uf_n)^* = f_n^{\Box}$ for every $n \in \mathbb{N}$. Thus $B_n = \{x \in S : (Uf_n)^*(x) = (Uf_n)^{\Box}(x)\} \cap \Gamma_{cp}$. By Proposition 3.4 and Proposition 3.8 we can conclude that $\mu(B_n) = 1$ for every $n \in \mathbb{N}$. Hence $\mu(\Gamma_{cpi}) = 1$. □

Now we consider the subset

$$\Gamma_{cpie} := \{x \in S : \int_S (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) = 0 \text{ for every } f \in C_b(S).\} \tag{8}$$

Notice that Γ_{cpie} is well-defined, since f^* is a Borel measurable function and Γ_{cpi} is Borel measurable. Also, Γ_{cpi}, f^* and ϵ_x are all independent of the choice of the metric d, thus Γ_{cpie} also does not depend on the choice of the metric.

This set is similar to the set defined in [34, 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 $\int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) = 0$ for every $f \in C_0(S)$, then $\int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) = 0$ for every $f \in C_b(S)$, so our formulation of Γ_{cpie} generalises the one in [34] to the setting of Polish spaces.

We can reformulate Γ_{cpie}:

Lemma 3.10. *Let $\{f_n : n \in \mathbb{N}\} \subset \mathrm{BL}(S)$ be the countable subset from Proposition 2.2. Then*

$$\Gamma_{cpie} = \{x \in \Gamma_{cpi} : \int_{\Gamma_{cpi}} (f_n^*(y) - f_n^*(x))^2 d\epsilon_x(y) = 0 \textit{ for every } n \in \mathbb{N}.\}$$

Proof. Clearly '$\subseteq$' holds. Now let $x \in \Gamma_{cpi}$ be such that $\int_{\Gamma_{cpi}} (f_n^*(y) - f_n^*(x))^2 d\epsilon_x(y) = 0$ for every $n \in \mathbb{N}$. Then there are Borel sets $B_n \subset \Gamma_{cpi}$ 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 $y \in \Gamma_{cp}$, 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 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 $\epsilon_x(B) = 1$, $x \in \Gamma_{cpie}$. □

Lemma 3.11. *Let μ be an invariant probability measure. Then*

$$\int_{\Gamma_{cpi}} \left(\int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \right) d\mu(x) = 0$$

for every $f \in C_b(S)$.

Proof. Let $f \in C_b(S)$ and let $x \in \Gamma_{cpi}$, then $\epsilon_x(\Gamma_{cpi}) = 1$ by Proposition 3.9, so

$$\begin{aligned}\int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) &= \int_S (f^*(y) - f^*(x))^2 d\epsilon_x(y) \\ &= \langle \epsilon_x, (f^*)^2 \rangle - 2f^*(x)\langle \epsilon_x, f^* \rangle + (f^*(x))^2 \epsilon_x(\Gamma_{cpi}) \\ &= ((f^*)^2)^{\square}(x) - 2f^*(x) f^{\square}(x) + (f^*)^2(x).\end{aligned}$$

So the map from Γ_{cpi} to $\mathbb{R}$ given by $x \mapsto \int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y)$ is Borel measurable since Γ_{cpi} is Borel measurable by Proposition 3.9, f^* is Borel measurable and $((f^*)^2)^{\square}$ and $f^{\square}$ are Borel measurable by Lemma 3.5.

By Proposition 3.8 $((f^*)^2)^{\square} = ((f^*)^2)^*$ μ-a.e. and $f^{\square} = f^*$ since $f \in C_b(S)$. Also note that by Kakutani's Ergodic Theorem (Theorem 2.9)

$$\langle \mu, ((f^*)^2)^* \rangle = \langle \mu, (f^*)^2 \rangle.$$

Thus, using the fact that $\mu(\Gamma_{cpi}) = 1$,

$$\begin{aligned}\int_{\Gamma_{cpi}} \int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \, d\mu(x) &= \langle \mu, ((f^*)^2)^{\square} - 2f^* f^{\square} + (f^*)^2 \rangle \\ &= \langle \mu, ((f^*)^2)^* \rangle - 2\langle \mu, f^* f^* \rangle + \langle \mu, (f^*)^2 \rangle = 0.\end{aligned}$$

□

The following theorem generalises [34, Theorem 6.1 and Theorem 6.3].

Theorem 3.12. *Γ_{cpie} is Borel measurable in S and $\mu(\Gamma_{cpie}) = 1$ for every invariant probability measure μ.*

Proof. Let $x \in \Gamma_{cpi}$. Let $\{f_n : n \in \mathbb{N}\}$ be the countable set from Proposition 2.2. For every $n \in \mathbb{N}$ we define $g_n : \Gamma_{cpi} \to \mathbb{R}$ as follows:

$$g_n(x) := \int_{\Gamma_{cpi}} (f_n^*(y) - f_n^*(x))^2 d\epsilon_x(y).$$

We have shown in the proof of Lemma 3.11 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 3.10 $\Gamma_{cpie} = \bigcap_{n=1}^{\infty} g_n^{-1}(\{0\})$, hence Γ_{cpie} is a Borel subset of Γ_{cpi}, hence Borel measurable.

Let μ be an invariant probability measure. By Lemma 3.11 $\int_{\Gamma_{cpi}} g_n \, d\mu = 0$ and since the g_n are positive, $g_n(x) = 0$ for μ-a.e. $x \in \Gamma_{cpi}$. So $\mu(g_n^{-1}(\{0\})) = \mu(\Gamma_{cpi}) = 1$ by Proposition 3.9. Thus $\mu(\Gamma_{cpie}) = 1$ as well. □

4 The ergodic decompositions

We start in Section 4.1 with the definition and some properties of ergodic measures. In Section 4.2 we show that the set Γ_{cpie} defined in Section 3 consists of exactly those $x \in \Gamma_{cpi}$ for which ϵ_x is ergodic. Moreover, we show that *every* ergodic probability measure is of the form ϵ_x for some $x \in \Gamma_{cpie}$. Using these results we prove an integral decomposition of invariant probability measures in terms of ergodic probability measures. In Section 4.3 we complete the Yosida-type ergodic decomposition of the state space S.

4.1 Ergodic measures

A Borel measurable subset A of S is an *invariant set* (with respect to P) if $P\delta_x(E) = 1$ for every $x \in E$. Following [11, 12, 33] 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 $X_n^x = P^n\delta_x$. A is an invariant set if, for every $x \in E$, $Pr(X_n^x \in E) = 1$ for all $n \in \mathbb{N}$. I.e. the associated Markov chain starting in $x \in E$ will remain in E with probability 1.

We shall write $\mathcal{P}_{inv}(S)$ to denote the convex set of invariant probability measures and $\mathcal{P}_{erg}(S)$ to denote the subset of ergodic probability measures, both with respect to P.

Let μ be an invariant probability measure. A Borel measurable subset E is a μ*-invariant set* (with respect to P) if $U\mathbb{1}_E = \mathbb{1}_E$ μ-a.e.

In some places in the literature, e.g. in [1, 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-invariant set A. We first show that this alternative definition is equivalent to our definition for ergodic measures.

Note that if E is an invariant set, then $U\mathbb{1}_E(x) = P\delta_x(E) = 1$ for every $x \in E$, thus $U\mathbb{1}_E \geq \mathbb{1}_E$. Since $\|U\mathbb{1}_E\|_1 \leq \|\mathbb{1}_E\|_1$ by Lemma 2.13, $U\mathbb{1}_E = \mathbb{1}_E$ μ-a.e. So every invariant set is μ-invariant.

Lemma 4.1. *Let μ be an invariant probability measure and let $B \subset S$ be μ-measurable. 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 2.13 $U\mathbb{1}_{B_{n-1}} = U\mathbb{1}_B$ 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,

$$\mu(B_n) = \mu(B_{n-1} \cap \{x \in B : U\mathbb{1}_{B_{n-1}}(x) = 1\}) = \mu(B_{n-1} \cap B) = \mu(B_{n-1}) = \mu(B),$$

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 4.2. *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 2.13 $U\mathbb{1}_B = U\mathbb{1}_S$ 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 4.1 concludes the proof. □

A proof of the following result, for a more general state space, i.e. just metrisable, can be found in [1, Theorem 19.25].

Theorem 4.3. *The extreme points of $\mathcal{P}_{inv}(S)$ are exactly the ergodic measures.*

The following theorem gives an equivalent condition for invariant measures to be ergodic.

Theorem 4.4. *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 $(U^{(n)}f(x))_n$ converges to $\langle \mu, f\rangle$ for every $x \in B$ and $f \in C_b(S)$.

An analogous result has been proven by Zaharopol [33, 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 [33, 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 2.2 will be exactly what we need in this situation. We first need some preliminary results.

Lemma 4.5. *Let μ be an ergodic probability 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 [12, Proposition 2.4.2]. □

Proof. (Theorem 4.4)

(i) ⇒ (ii): Let $\{f_n : n \in \mathbb{N}\}$ be the countable subset in $C_b(S)$ given by Proposition 2.2. Since μ is ergodic, Lemma 4.5 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 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 (Theorem 2.9) 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 $\langle \mu_1, f\rangle = \langle \mu_2, f\rangle$ 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 4.3. □

4.2 An integral decomposition of invariant measures

It need not be true that ϵ_x is an ergodic measure whenever $x \in \Gamma_{cpi}$. A very simple example is given in [33, 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_{cpie}, defined in Section 3, consists of exacty those $x \in \Gamma_{cpi}$ 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_{cpie} as follows: $x \sim y$ if and only if $\epsilon_x = \epsilon_y$. Let $[x]$ be the equivalence class of $x \in \Gamma_{cpie}$ defined by $\sim$.

The following theorem extends [34, Lemma 6.4 and Theorem 6.5]:

Theorem 4.6. *(i) For every $x \in \Gamma_{cpie}$ the set $[x]$ is Borel measurable, ϵ_x is an ergodic probability measure and $\epsilon_x([x]) = 1$.*

(ii) Conversely, any ergodic measure μ is of the form $\mu = \epsilon_x$ for some $x \in \Gamma_{cpie}$.

Proof. *(i)* Let $x \in \Gamma_{cpie}$. Let $\{f_n : n \in \mathbb{N}\}$ be the countable subset of $\mathrm{BL}(S)$ from Proposition 2.2. We define

$$F_n = \{y \in \Gamma_{cpi} : f_n^*(y) = f_n^*(x)\} = (f_n^*)^{-1}(f_n^*(x))$$

and $E_n = \Gamma_{cpie} \cap F_n$. Since f_n^* is a Borel measurable function and Γ_{cpie} is a Borel set, E_n and F_n are also Borel sets. From Proposition 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 $\Gamma_{cpie} \subset \Gamma_{cpi}$, 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 4.4. Since $x \in \Gamma_{cpie}$ we know that $\int_{\Gamma_{cpi}} (f_n^*(y) - f_n^*(x))^2 \, d\epsilon_x(y) = 0$ for every $n \in \mathbb{N}$. This implies that $\epsilon_x(F_n) = 1$, since $\epsilon_x(\Gamma_{cpi}) = 1$ by Proposition 3.9. So $\epsilon_x(E_n) = 1$ as well, since $\epsilon_x(\Gamma_{cpie}) = 1$ by Theorem 3.12, thus $\epsilon_x([x]) = 1$.

(ii) Let μ be an ergodic probability measure. Then by Theorem 4.4 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 $A \subset \Gamma_{cpi}$. 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 $\epsilon_x(\Gamma_{cpi} \backslash A) = 0$

$$\begin{aligned}\int_{\Gamma_{cpi}} (f^*(y) - f^*(x))^2 d\epsilon_x(y) &= \int_A (f^*(y) - f^*(x))^2 \, d\epsilon_x(y) \\ &= \int_A (\langle \epsilon_x, f \rangle - \langle \epsilon_x, f \rangle)^2 \, d\epsilon_x(y) = 0,\end{aligned}$$

for every $f \in C_b(S)$. So $x \in \Gamma_{cpie}$. □

Corollary 4.7. *Let μ, ν be ergodic probability measures. Then either $\mu = \nu$ or μ and ν are mutually singular.*

Proof. By Theorem 4.6 there are $x, y \in \Gamma_{cpie}$ 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 3.12 and Theorem 4.6 the following holds:

Corollary 4.8. *The following statements are equivalent:*

(i) There exists an invariant probability measure.

(ii) Γ_{cpie} is not empty.

(iii) There exists an ergodic probability measure.

This implies

Corollary 4.9. *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 probability measures.

Theorem 4.10. *Let μ be an invariant probability measure. Then the map*

$$x \mapsto \begin{cases} \epsilon_x & \text{if } x \in \Gamma_{cpie} \\ 0 & \text{if } x \notin \Gamma_{cpie}. \end{cases} \tag{9}$$

is strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$ and

$$\mu = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu.$$

Proof. Corollary 2.6 and Theorem 3.12 imply that

$$\mu = \int_{\Gamma_{cpie}} \delta_x \, d\mu(x).$$

By Corollary 2.11 and the invariance of μ we obtain for every $n \in \mathbb{N}$

$$\mu = P^{(n)}\mu = \int_{\Gamma_{cpie}} P^{(n)} \delta_x \, d\mu(x).$$

Now, $P^{(n)}\delta_x \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$ for every $x \in \Gamma_{cpie}$. So the measurability of Γ_{cpie} implies that the map defined in (9) is strongly measurable from S to $\mathcal{S}_{\mathrm{BL}}$, and by the Dominated Convergence Theorem we can conclude that

$$\mu = \lim_{n\to\infty} P^{(n)}\mu = \int_{\Gamma_{cpie}} \lim_{n\to\infty} P^{(n)}\delta_x \, d\mu(x) = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x).$$

□

Corollary 4.11. *Suppose P has a unique ergodic probability measure μ^*. Then μ^* is the only invariant probability measure.*

Proof. Let μ be an invariant probability measure. Then by Theorem 4.10 $\mu = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x)$. ϵ_x is ergodic, so $\epsilon_x = \mu^*$ for every $x \in \Gamma_{cpie}$. The result now follows from $\mu(\Gamma_{cpie}) = 1$ (Theorem 3.12. □

Theorem 4.12. *The following statements hold:*

(i) $\mathcal{P}_{inv}(S)$ is dense in the closed convex hull of $\mathcal{P}_{erg}(S)$ in $\mathcal{S}_{\mathrm{BL}}$.

(ii) If P is a Markov-Feller operator then $\mathcal{P}_{inv}(S)$ equals the closed convex hull of $\mathcal{P}_{erg}(S)$ in $\mathcal{S}_{\mathrm{BL}}$.

Proof. If there exists no invariant probability measure then both $\mathcal{P}_{inv}(S)$ and $\mathcal{P}_{erg}(S)$ 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 4.10 $\mu = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x)$. By [4, Corollary 8]

$$\frac{1}{\mu(\Gamma_{cpie})} \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x) \in \overline{\mathrm{conv}}(\{\epsilon_x : x \in \Gamma_{cpie}\}).$$

By Theorem 3.12 $\mu(\Gamma_{cpie}) = 1$ and by Theorem 4.6 $\epsilon_x \in \mathcal{P}_{erg}(S)$ for every $x \in \Gamma_{cpie}$, thus

$$\mu = \int_{\Gamma_{cpie}} \epsilon_x \, d\mu(x) \in \overline{\mathrm{conv}}(\mathcal{P}_{erg}(S)).$$

(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 $\mathcal{P}_{inv}(S)$ is closed and convex in $\mathcal{S}_{\mathrm{BL}}$. So the closed convex hull of $\mathcal{P}_{erg}(S)$ is contained in $\mathcal{P}_{inv}(S)$. □

If P is not a Markov-Feller operator, $\mathcal{P}_{inv}(S)$ need not be closed in $\mathcal{S}_{\mathrm{BL}}$. 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 2.2, $P\mu := \mu \circ \Phi^{-1}$ defines a regular Markov operator $P : \mathcal{M}^+(S) \to \mathcal{M}^+(S)$. Clearly $\delta_x \in \mathcal{P}_{inv}(S)$ 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 $\mathcal{P}_{inv}(S)$ is not closed in $\mathcal{S}_{\mathrm{BL}}$.

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 $\mathcal{M}_R(S) := \{\mu \in \mathcal{M}(S) : |\mu|(S \backslash R) = 0\}$: we 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 Corollary 2.11

$$P|\mu|(R) = \int_S P\delta_x(R)\, d|\mu|(x) \geq \int_R d|\mu| = |\mu|(R) = |\mu|(S).$$

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 [11, Proposition 4.5] for regular Markov operators on locally compact separable metric spaces.

Theorem 4.13. *Let S be a Polish space and P a regular Markov operator. If there exist ergodic measures or equivalently if Γ_{cpie} is not empty, then for every $[x] \in \Gamma_{cpie}/\sim$ 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 $x \in \Gamma_{cpie}$.

(i) By Theorem 4.6 $\epsilon_x([x]) = 1$, so by Corollary 4.2 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 4.10 and since $\overline{\mu}(S \backslash S_{[x]}) = 0$

$$\begin{aligned} \overline{\mu} &= \int_{\Gamma_{cpie}} \epsilon_y\, d\overline{\mu}(y) = \int_{S_{[x]}} \epsilon_y\, d\mu(y) \\ &= \int_{S_{[x]}} \epsilon_x\, d\mu(y) = \epsilon_x, \end{aligned}$$

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. □

5 Application to convergence of Cesàro averages

Section 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_{cp}, Γ_{cpi} or $[z]$ for some $z \in \Gamma_{cpie}$, 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.

Let

$$\mathcal{P}_{cp} := \{\mu \in \mathcal{P}(S) : (P^{(n)}\mu)_n \text{ converges in } \mathcal{S}_{\mathrm{BL}} \text{ as } n \to \infty\},$$

and define

$$\epsilon_\mu := \lim_{n\to\infty} P^{(n)}\mu$$

for every $\mu \in \mathcal{P}_{cp}$. Note that by definition $\delta_x \in \mathcal{P}_{cp}$ if and only if $x \in \Gamma_{cp}$, and $\epsilon_{\delta_x} = \epsilon_x$. Analogous to Γ_{cpi} and Γ_{cpie} we can define

$$\mathcal{P}_{cpi} := \{\mu \in \mathcal{P}_{cp} : \epsilon_\mu \text{ is invariant}\}$$

and

$$\mathcal{P}_{cpie} := \{\mu \in \mathcal{P}_{cpi} : \epsilon_\mu \text{ is ergodic.}\}$$

Proposition 5.1. *Let $\mu \in \mathcal{P}(S)$ such that $\mu(\Gamma_{cp}) = 1$. Then $\mu \in \mathcal{P}_{cp}$, $\epsilon_\mu = \int_{\Gamma_{cp}} \epsilon_x \, d\mu(x)$ and for every $f \in C_b(S)$, $\langle \epsilon_\mu, f\rangle = \langle \mu, f^*\rangle$.*

Proof. By Corollary 2.11 and $\mu(\Gamma_{cp}) = 1$, we have

$$P^{(n)}\mu = \int_{\Gamma_{cp}} P^{(n)}\delta_x \, d\mu(x).$$

Now, for every $x \in \Gamma_{cp}$, $P^{(n)}\delta_x \to \epsilon_x$ in $\mathcal{S}_{\mathrm{BL}}$, so by the Dominated Convergence Theorem we obtain that $P^{(n)}\mu$ converges and

$$\lim_{n\to\infty} P^{(n)}\mu = \int_{\Gamma_{cp}} \epsilon_x \, d\mu(x),$$

thus $\mu \in \mathcal{P}_{cp}$. According to Proposition 2.5, for every $f \in C_b(S)$,

$$\langle \epsilon_\mu, f\rangle = \int_{\Gamma_{cp}} \langle \epsilon_x, f\rangle \, d\mu(x) = \int_S f^*(x) \, d\mu(x).$$

□

Proposition 5.2. *Suppose $\mu \in \mathcal{P}(S)$ is such that $\mu(\Gamma_{cpi}) = 1$. Then $\mu \in \mathcal{P}_{cpi}$ and $\epsilon_\mu = \int_{\Gamma_{cpi}} \epsilon_x \, d\mu(x)$.*

Proof. Since $\mu(\Gamma_{cp}) = 1$, Proposition 5.1 implies that $\mu \in \mathcal{P}_{cp}$ and $\epsilon_\mu = \int_{\Gamma_{cp}} \epsilon_x \, d\mu(x)$. Because $\mu(\Gamma_{cpi}) = 1$ we obtain

$$\epsilon_\mu = \int_{\Gamma_{cpi}} \epsilon_x \, d\mu(x).$$

According to Proposition 2.10,

$$P\epsilon_\mu = P\epsilon_x \, d\mu(x) = \int_{\Gamma_{cpi}} \epsilon_x \, d\mu(x) = \epsilon_\mu.$$

So $\mu \in \mathcal{P}_{cpi}$. □

We now state some results on convergence of Cesàro averages of (possibly signed) finite Borel measures.

Corollary 5.3. *Let μ be a finite Borel measure on S such that $|\mu|(S \backslash \Gamma_{cp}) = 0$. 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 $|\mu|(S \backslash \Gamma_{cpi}) = 0$, *then* μ^* *is invariant.*

Proof. These results follow by writing $\mu = \mu^+ - \mu^-$ and applying Proposition 5.1 and Proposition 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 [11, Theorem 3.1(g)] from the locally compact setting to the Polish setting.

Proposition 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 3.9, $\nu(S \backslash \Gamma_{cpi}) = 0$, thus $|\mu|(S \backslash \Gamma_{cpi}) = 0$ as well. So Corollary 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 [14, Lemma 4.2] 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. [9, Theorem VII.A]) there exists, for every $f \in L^1(\nu)$, an $\hat{f} \in L^1(\nu)$ such that $\|\frac{1}{n}\sum_{k=0}^{n-1} T^k f - \hat{f}\|_1 \to 0$.

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 $\mu \in \mathcal{P}_{cpie}$ whenever $\mu(\Gamma_{cpie}) = 1$. However, this is generally not true: $\mu(\Gamma_{cpie}) = 1$ for every invariant probability measure μ according to Theorem 3.12, and $\epsilon_\mu = \mu$ for these measures, while they obviously need not be ergodic. The following result does hold, however:

Proposition 5.5. *Let $\mu \in \mathcal{P}(S)$. If $\mu([z]) = 1$ for some $z \in \Gamma_{cpie}$, then $\mu \in \mathcal{P}_{cpie}$ and $\epsilon_\mu = \epsilon_z$.*

Proof. Suppose that $\mu([z]) = 1$ for some $z \in \Gamma_{cpie}$, then $\mu(\Gamma_{cpi}) = 1$, so $\mu \in \mathcal{P}_{cpi}$ by Proposition 5.2, and $\epsilon_\mu = \int_{\Gamma_{cpi}} \epsilon_x \, d\mu(x)$. Since $\mu([z]) = 1$ we have

$$\epsilon_\mu = \int_{\Gamma_{cpi}} \epsilon_x \, d\mu(x) = \int_{[z]} \epsilon_x \, d\mu(x) = \int_{[z]} \epsilon_z \, d\mu(x) = \epsilon_z,$$

so $\epsilon_\mu = \epsilon_z$ is ergodic, thus $\mu \in \mathcal{P}_{cpie}$. □

The conditions in Proposition 5.1, Proposition 5.2 and Proposition 5.5 are not *necessary* in general for a measure μ to be in $\mathcal{P}_{cp}$, $\mathcal{P}_{cpi}$ and $\mathcal{P}_{cpie}$ 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 $\mu(\Gamma_{cp}) = 0$.

Example 5.6. Let $S = [0,1]$ with Euclidean metric d. Let $(a_n)_n$ be a sequence in $\{0,1\}$ such that $\frac{1}{n}\sum_{k=1}^n a_k$ does not converge. Define $b_n := a_n + \frac{(-1)^{a_n}}{2^{n+1}}$ and $c_n = 1 - a_n$. 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 be the regular Markov operator associated to Φ, i.e. $P\mu(E) := \mu(\Phi^{-1}(E))$ for every $\mu \in \mathcal{M}^+(S)$ and $E \subset S$ Borel. Then $P\delta_{b_n} = \delta_{b_{n+1}}$ and $P\delta_{c_n} = \delta_{c_{n+1}}$ for every $n \in \mathbb{N}$.

We will show that $b_1 \notin \Gamma_{cp}$. To that end, let $f(x) := x$. Then

$$\langle P^{(n)}\delta_{b_1}, f\rangle = \frac{1}{n}\sum_{k=1}^n \langle \delta_{b_k}, f\rangle = \frac{1}{n}\sum_{k=1}^n b_k.$$

Now $|\frac{1}{n}\sum_{k=1}^n b_k - \frac{1}{n}\sum_{k=1}^n a_k| \leq \frac{1}{n}\sum_{k=1}^n 2^k \to 0$, so $\frac{1}{n}\sum_{k=1}^n b_k$ does not converge, thus $b_1 \notin \Gamma_{cp}$. By similar reasonings, $c_1 \notin \Gamma_{cp}$.

Let $\mu = \frac{1}{2}\delta_{b_1} + \frac{1}{2}\delta_{c_1}$. Then

$$P^{(n)}\mu = \frac{1}{2n}\sum_{k=1}^n \delta_{\frac{1}{2^{k+1}}} + \frac{1}{2n}\sum_{k=1}^n \delta_{1-\frac{1}{2^{k+1}}} \to \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$$

in $\mathcal{S}_{\mathrm{BL}}$ as $n \to \infty$. Note that $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ is an ergodic invariant probability measure, thus $\mu \in \mathcal{P}_{cpie}$, but $\mu(\Gamma_{cp}) = 0$.

References

[1] Aliprantis, C. and K. Border (2006), *Infinite dimensional analysis. A hitchhiker's guide*, Third edition, Springer, Berlin.

[2] Bauer, H. (2001), *Measure and Integration Theory*, Walter de Gruyter, Berlin.

[3] Deuschel, J.-D. and D. Stroock (1989), *Large Deviations*, Academic Press, New York.

[4] Diestel, J. and J.J. Uhl jr. (1977), *Vector Measures*, Math. Surveys, Nr. 15, Amer. Math. Soc.

[5] Dudley, R.M. (1966), Convergence of Baire measures, *Stud. Math.* **27**, 251–268.

[6] Dudley, R.M. (1974), Correction to: "Convergence of Baire measures", *Stud. Math.* **51**, 275.

[7] Dynkin, E.B. (1978), Sufficient statistics and extreme points, *Annals of Probability* **6**, 705–730.

[8] Ethier, S.N. and T.G. Kurtz (1986), *Markov Processes; Characterization and Convergence*, Wiley, New York.

[9] Foguel, S.R. (1969), *The ergodic theory of Markov processes*, Van Nostrand Mathematical Studies, Vol. 21, New York, London, Toronto, Melbourne.

[10] Gacki, H. (2007), Applications of the Kantorovich-Rubinstein maximum principle in the theory of Markov semigroups, *Dissertationes Math* **448**.

[11] Hernández-Lerma, O. and J.B. Lasserre (1998), Ergodic theorems and ergodic decomposition for Markov chains, *Acta Appl. Math.* **54**, 99–119.

[12] Hernández-Lerma, O. and J.B. Lasserre (2003), *Markov chains and invariant probabilities*, Birkhäuser, Basel.

[13] Hille, S.C. and D.T.H. Worm (2009), Embedding of semigroups of Lipschitz maps into positive linear semigroups on ordered Banach spaces generated by measures, *Integr. Equ. Oper. Theory* **63**, 351–371.

[14] 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.

[15] Horbacz, K., Myjak, J. and T. Szarek (2005), On stability of some general random dynamical system, *J. Stat. Phys.* **119**, 35–60.

[16] Horbacz, K. (2004), Random dynamical systems with jumps, *J. Appl. Prob.* **41**, 890–910.

[17] Kakutani, S. (1940), Ergodic theorems and the Markoff process with a stable distribution, *Proc. Imp. Acad. Tokyo* **16**, 49–54.

[18] Katok, A. and B. Hasselblatt (1995), *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge.

[19] 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

[20] Lasota, A., J. Myjak and T. Szarek (2002), Markov operators with a unique invariant measure, *J. Math. Anal. Appl.* **276**, 343–356.

[21] Lant, T. and H.R. Thieme (2007), Markov transition functions and semigroups of measures, *Semigroup Forum* **74**, 337–369.

[22] Mañé, R. (1987), *Ergodic theory and differentiable dynamics*, Springer, Berlin.

[23] Myjak, J. and T. Szarek (2003), Attractors of iterated function systems and Markov operators, *Abstr. Appl. Anal.* **8**, 479–502.

[24] Myjak, J. and T. Szarek (2004), On the existence of an invariant measure for Markov-Feller operators, *J. Math. Anal. Appl.* **294**, 215–222.

[25] Ollivier, Y. (2009), Ricci curvature of Markov chains on metric spaces, *J. Funct. Anal.* **256**, 810–864.

[26] Szarek, T. (2003), Invariant measures for Markov operators with application to function systems, *Studia Math.* **154**, 207–222.

[27] Szarek, T. (2003), Invariant measures for nonexpansive Markov operators on Polish spaces, *Dissertationes Math.* **415**.

[28] Szarek, T. (2008), The uniqueness of invariant measures for Markov operators, *Studia Math.* **189** (3), 225–233.

[29] Varadhan, S.R.S. (2001), *Probability theory*, Courant Lecture Notes in Mathematics 7. Providence, RI: American Mathematical Society.

[30] Williams, D. (1979), *Diffusions, Markov Processes and Martingales. Volume 1: Foundations*, Wiley, Chichester.

[31] Yosida, K. (1948), Simple Markoff process with a locally compact phase space, *Math. Japonicae* **1**, 99-103.

[32] Yosida, K. (1980), *Functional analysis*, Sixth edition, Springer-Verlag, Berlin.

[33] Zaharopol, R. (2005), *Invariant probabilities of Markov-Feller operators and their supports*, Birkhäuser, Basel.

[34] Zaharopol, R. (2008), An ergodic decomposition defined by transition probabilities, Acta. Appl. Math. **104**, 47–81.