On a nonstandard approach to invariant measures for Markov operators

Abstract. We show the existence of invariant measures for Markov–Feller operators defined on completely regular topological spaces which satisfy the classical positivity condition.

1. Introduction. One of the strategies in studying Markov processes is to consider certain linear operators and to examine their properties. The theory of Markov operators is very rich and its results are very useful in many branches of pure and applied mathematics including stochastic differential equations, dynamical systems, mathematical theory of learning, population dynamics, the theory of fractals and others (see for instance [10, 20] and the references given there).

In [12], A. Lasota and J. A. Yorke gave a new sufficient condition for asymptotic stability of Markov operators defined on locally and σ-compact metric spaces. Their approach was partially based on the lower bound function technique for Markov operators acting on $L^1$ space, see [10]. T. Szarek managed in [18] to extend the Lasota–Yorke result to the case of Polish spaces. One of the difficulties in proving stability is to show the existence of an invariant measure, see [9, 11, 12, 18, 19].

In this paper we propose a nonstandard approach to the problem of an invariant measure. We use the so-called Loeb measure construction [13] to produce an invariant measure in an “ideal” space $^*X$ and then push it

2000 Mathematics Subject Classification. Primary 47A35, Secondary 28D05, 37A30.

Key words and phrases. Markov operator, invariant measure, nonstandard analysis.
down to X by using the properties of the standard part map (see [2, 8]). This approach is quite general and our main result is valid in the case of completely regular topological spaces. A related result in normal spaces was obtained in [7]. However, nonstandard approach seems to be interesting on its own for the future developments.

In Section 2 we recall definitions concerning Markov operators. Some methods of nonstandard analysis are presented in Section 3. We use these methods in Section 4 to prove our main Theorem 4.1.

2. Markov operators on measures. Let \((X, \mathcal{T})\) be a topological space. We denote by \(\mathcal{M}_{\text{fin}}\) and \(\mathcal{M}_1\) the sets of Borel measures on \(X\) such that \(\mu(X) < \infty\) and \(\mu(X) = 1\) respectively. The elements of \(\mathcal{M}_1\) are called distributions. Let \(\mathcal{B}(X)\) denote the family of Borel sets, \(\mathcal{B}(X)\) the space of all bounded Borel measurable functions \(f : X \to \mathbb{R}\) with the supremum norm and \(C(X)\) its subspace of bounded continuous functions. By \(\mathcal{M}_{\text{rad}}\) we denote the set of finite Radon measures, that is,

\[
\mu(B) = \sup \{ \mu(K) : K \subset B, \text{ K compact} \}
\]

for \(B \in \mathcal{B}(X)\) and \(\mu \in \mathcal{M}_{\text{rad}}\).

According to [10], an operator \(P : \mathcal{M}_{\text{fin}} \to \mathcal{M}_{\text{fin}}\) is called a Markov operator (on measures) if \(P(\lambda_1 \mu_1 + \lambda_2 \mu_2) = \lambda_1 P \mu_1 + \lambda_2 P \mu_2\) for \(\lambda_1, \lambda_2 \geq 0\), \(\mu_1, \mu_2 \in \mathcal{M}_{\text{fin}}\) and \(P \mu(X) = \mu(X)\) for \(\mu \in \mathcal{M}_{\text{fin}}\).

A Markov operator \(P\) is called a Feller operator if there is a linear operator \(U : \mathcal{B}(X) \to \mathcal{B}(X)\) such that

\[
\langle Uf, \mu \rangle = \langle f, P \mu \rangle \quad \text{for } f \in \mathcal{B}(X), \mu \in \mathcal{M}_{\text{fin}}
\]

and \(Uf \in C(X)\) for \(f \in C(X)\). Here \(\langle f, \nu \rangle = \int_X f(x) \nu(dx)\).

It is easy to see that \(Uf(x) = \langle f, P \delta_x \rangle\) for \(f \in \mathcal{B}(X), x \in X\), where \(\delta_x\) is the Dirac measure supported at \(x\).

A measure \(\mu \in \mathcal{M}_{\text{fin}}\) is called stationary or invariant with respect to \(P\) if \(P \mu = \mu\).

3. Nonstandard preliminaries. We present here some ideas which lie behind techniques we shall use in Section 4.

Let \(\mathcal{U}\) be a free ultrafilter on \(\mathbb{N}\). For given two sequences of reals \(\langle a_n \rangle, \langle b_n \rangle\), write \(\langle a_n \rangle \equiv \langle b_n \rangle \iff \{ i \in \mathbb{N} : a_i = b_i \} \in \mathcal{U}\). It is easy to see that \(\equiv\) is an equivalence relation in the Cartesian product \(\prod_{n \in \mathbb{N}} \mathbb{R}\). The quotient space \(\prod_{n \in \mathbb{N}} \mathbb{R} / \equiv\) is called the (set theoretic) ultrapower of \(\mathbb{R}\) and is denoted by \((\mathbb{R})_{\mathcal{U}}\). Let \(\nu_1, \nu_2, \ldots\) be a sequence of probability Borel measures on \(\mathbb{R}\). It is not difficult to see that \((\mathcal{B}(\mathbb{R}))_{\mathcal{U}} = \{(A_i)_{\mathcal{U}} : A_i \in \mathcal{B}(\mathbb{R}), i \in \mathbb{N}\}\) is an algebra of subsets of \((\mathbb{R})_{\mathcal{U}}\) and

\[
\nu_0((A_i)_{\mathcal{U}}) = \lim_{\mathcal{U}} \nu_i(A_i)
\]

is a well-defined finitely additive measure defined on \((\mathcal{B}(\mathbb{R}))_{\mathcal{U}}\) (here \(\lim_{\mathcal{U}}\) denotes the limit over \(\mathcal{U}\)). It can be proved (see for instance [16, Lemma 9.1]
that \( \nu_0 \) is \( \sigma \)-additive and therefore it may be uniquely extended to the complete \( \sigma \)-algebra generated by \( (B(\mathbb{R}))_U \). The construction described above may be regarded as the ultraproduct counterpart of the famous Loeb measure construction in nonstandard analysis.

Let \( \langle x_n \rangle \) be a bounded sequence in \( \mathbb{R} \). Then there exists a unique \( x \in \mathbb{R} \) such that \( \lim_U x_n = x \). Moreover, if \( \langle y_n \rangle \equiv \langle x_n \rangle \), then \( \lim_U y_n = x \), too. A mapping \( [\langle x_n \rangle]_U \to x \) may be regarded as the counterpart of the standard part map in nonstandard analysis which allows to relate some (especially measure-theoretic) properties of \( \mathbb{R} \) and \( (\mathbb{R})_U \).

Clearly, we can consider richer structures than \( \mathbb{R} \) in a similar fashion. However, the above notions are rather difficult to handle. Nonstandard analysis offers a specific language which supports our intuition and provides us with new techniques which are not very easy to express in the ultraproduct setting.

There are several frameworks for nonstandard analysis, see [3, 5] for the surveys. In our paper we shall use the most classical superstructure approach. For a detailed account of this approach the reader is referred to [1, 4, 14].

Let \( (X, \mathcal{T}) \) be a topological Hausdorff space. We fix a \( \kappa \)-saturated nonstandard universe \((V(\Xi), V(\Xi), *)\), where \( \kappa \) is an (uncountable) cardinal number greater than \( \text{card}(\mathcal{T}) \). As usual, we assume that \( \Xi \) is some large set including \( X \) and \( \mathbb{R} \). For any \( x \in X \), \( m(x) = \bigcap \{ *T : x \in T, \ T \in \mathcal{T} \} \) is the monad of \( x \). If \( y \in *X \) and \( y \in m(x) \) for some \( x \in X \), we write \( st y = x \) and say that \( x \) is the standard part of \( y \). The set of near-standard points of \( *X \) is given by \( \text{ns}(\{X\}) = \{ y \in *X : st y \in X \} \). Thus we obtain the so-called standard part map \( st : \text{ns}(\{X\}) \to X \). We shall use the same symbol \( st \) for a mapping \( \text{ns}(\{\mathbb{R}\}) \ni y \to x \in \mathbb{R} \).

If \( \nu : *B(X) \to *[0, \infty) \) is a finite internal finitely additive function, then, by Loeb’s theorem [13], \( \nu_L(A) = st(\nu(A)) \) defines a finite \( \sigma \)-additive measure on \( *B(X) \) and therefore it may be extended to the unique complete \( \sigma \)-algebra \( L(\nu, *B(X)) \supset *B(X) \). This extension is also denoted by \( \nu_L \) and called the Loeb measure corresponding to \( \nu \). Moreover,

\[
\bar{\nu}(A) = \inf \{ \nu_L(B) : A \subset B \in *B(X) \},
\]

\[
\underline{\nu}(A) = \sup \{ \nu_L(B) : A \supset B \in *B(X) \}, \ A \subset *X,
\]

are, respectively, the outer and inner measures induced by \( \nu_L \).

The following theorem will be used in the next section. Write

\[
\nu \circ st^{-1}(B) = \nu(st^{-1}(B)), \ B \in B(X).
\]

**Theorem 3.1** (see [8, Theorem 4]). Let \( (X, \mathcal{T}) \) be a regular topological space and let \( \nu : *B(X) \to *[0, \infty) \) be a finite internal finitely additive function. Then

\[
st^{-1}(B) \in \{ C \cap \text{ns}(\{X\}) : C \in L(\nu, *B(X)) \}\]
for $B \in \mathcal{B}(X)$ and
\[ \nu \circ \text{st}^{-1} : \mathcal{B}(X) \to [0, \infty) \]
is a Radon measure.

Let
\[ \mathcal{T}_0 = \{ f^{-1}(O) : O \subset \mathbb{R} \text{ open, } f \text{ continuous} \} \]
denote the family of exact open sets. Note that in the case of completely regular topological spaces $\mathcal{T}_0$ is a base for the topology $\mathcal{T}$. The following theorem follows easily from [1, Proposition 3.4.5], (see also [8, Theorem 1]).

**Theorem 3.2.** Let $(X, \mathcal{T})$ be a completely regular topological space and let $\nu : *\mathcal{B}(X) \to *[0, \infty)$ be a finite internal finitely additive function. Then
\[ \text{st}^{-1}(K) \in L(\nu, *\mathcal{B}(X)) \]
for all compact sets $K$ and
\[ \nu_L\left(\text{st}^{-1}(K)\right) = \inf\{\nu_L(*U) : K \subset U, U \in \mathcal{T}_0\} \]

4. **Invariant measures.** In this section we prove a general theorem about the existence of an invariant distribution for Markov operators satisfying a classical positivity condition. A related result in normal spaces was obtained by Foguel in [7].

**Theorem 4.1.** Let $X$ be a completely regular topological space and let $P : \mathcal{M}_{\text{fin}} \to \mathcal{M}_{\text{fin}}$ be a Feller operator with $P(\mathcal{M}_{\text{rad}}) \subset \mathcal{M}_{\text{rad}}$. Assume that there is a compact set $Y_0 \subset X$ such that
\[ \limsup_{n \to \infty} \sup_{\mu \in \mathcal{M}_1} \left( \frac{1}{n} \sum_{k=1}^{n} P^k \mu(Y_0) \right) > 0. \]
Then $P$ has an invariant Radon distribution.

**Proof.** It follows from (4.1) that there exist $\varepsilon > 0$, a sequence of distributions $\langle \mu_n \rangle$ and an increasing sequence of integers $\langle q_n \rangle$ such that
\[ \frac{1}{q_n} \sum_{k=1}^{q_n} P^k \mu_n(Y_0) \geq \varepsilon, \quad \text{for } n = 1, 2, \ldots \]
Put
\[ m_n = \frac{1}{q_n} \sum_{k=1}^{q_n} P^k \mu_n \]
and notice that
\[ |Pm_n(A) - m_n(A)| \leq \frac{2}{q_n} \]
for every $A \in \mathcal{B}(X)$ and $n \in \mathbb{N}$. By transfer,
\[ |^*PM_N(B) - M_N(B)| \leq \frac{2}{q_N} \]
for every $B \in ^*\mathcal{B}(X)$, $N \in ^*\mathbb{N}$ (note that, since $\langle m_n \rangle \in (M_1)^N$, $^*\langle m_n \rangle = \langle M_N \rangle \in ^*((M_1)^N) \subset (^*\mathcal{M}_1)^N$). Fix a hyperinteger $\omega \in ^*\mathbb{N} \setminus \mathbb{N}$. Since $\lim_{n \to \infty} g_n = \infty$, we have $\frac{1}{g_n} \approx 0$ and hence $^*PM_\omega(B) \approx M_\omega(B)$ for every $B \in ^*\mathcal{B}(X)$. Put $\nu = M_\omega$ and denote by

$$
(^*P\nu)_L, \nu_L : L(\nu, ^*\mathcal{B}(X)) \to [0, 1]
$$

the Loeb measures associated with $^*P\nu$ and $\nu$, respectively. Clearly, $(^*P\nu)_L = \nu_L$. Let $\nu$ be the inner measure induced by $\nu_L$, that is,

$$
\nu(B) = \sup \{ \nu_L(A) : B \supset A \in ^*\mathcal{B}(X) \}.
$$

We show that $\nu \circ \text{st}^{-1}$ is an invariant, nontrivial measure for $P$. The proof will be given in four steps.

Step I. Let $f \in C(X)$. Then

$$
\int_X f(x) P(\nu \circ \text{st}^{-1})(dx) = \int_X (Uf)(x) (\nu \circ \text{st}^{-1})(dx)
$$

$$
= \int_{\text{st}^{-1}(X)} (Uf)(st y) \nu(dy) = \int_{\text{st}^{-1}(X)} \text{st}(Uf)(y) \nu(dy)
$$

$$
\leq \int_{\text{st}^{-1}(X)} \text{st}(U^*f)(y) \nu_L(dy) = \int_{\text{st}^{-1}(X)} \text{st}(U^*f)(y) \nu(dy)
$$

$$
= \text{st} \int_{\text{st}^{-1}(X)} \text{st}(U^*f)(y) \nu_L(dy) = \text{st} \int_{\text{st}^{-1}(X)} \text{st}(U^*f)(y) \nu_L(dy)
$$

$$
= \int_X \text{st}(f(y)) \nu_L(dy).
$$

Step II. We show that $P(\nu \circ \text{st}^{-1})(U) \leq \nu_L(U)$ for every $U \in T_0$, where

$$
T_0 = \{ f^{-1}(O) : O \subset \mathbb{R} \text{ open, } f \text{ continuous} \}
$$

is the family of exact open sets. If $U \in T_0$, then there exists an open $O \subset \mathbb{R}$ and a continuous function $f : X \to \mathbb{R}$ such that $U = f^{-1}(O)$. Put $F = \mathbb{R} \setminus O$ and let $g_n(x) = \min \{ 1, n \text{ dist}(f(x), F) \}$. It is easy to see that $g_n \uparrow I_U$. Therefore

$$
P(\nu \circ \text{st}^{-1})(U) = \int_X I_U(x) P(\nu \circ \text{st}^{-1})(dx)
$$

$$
= \int_X \lim_{n \to \infty} g_n(x) P(\nu \circ \text{st}^{-1})(dx) = \lim_{n \to \infty} \int_X g_n(x) P(\nu \circ \text{st}^{-1})(dx)
$$

$$
\leq \lim_{n \to \infty} \int_X \text{st}(g_n(y)) \nu_L(dy)
$$

by Step I. But $^*g_n(y) \leq ^*I_U(y)$ by transfer and hence $\text{st}(^*g_n(y)) \leq \text{st}(^*I_U(y)) = I_U(y)$. Consequently

$$
\lim_{n \to \infty} \int_X \text{st}(^*g_n(y)) \nu_L(dy) \leq \int_X I_U(y) \nu_L(dy) = \nu_L(U).
$$
Step III. Let $K$ be a compact subset of $X$. Then, by Theorem 3.2, 
$$
\nu_L(\text{st}^{-1}(K)) = \inf \{ \nu_L(U) : K \subset U, U \in \mathcal{T}_0 \}.
$$

Hence
$$
\nu \circ \text{st}^{-1}(K) = \nu_L(\text{st}^{-1}(K)) = \inf \{ \nu_L(U) : K \subset U, U \in \mathcal{T}_0 \}
\geq \inf \{ P(\nu \circ \text{st}^{-1})(U) : K \subset U, U \in \mathcal{T}_0 \}
\geq P(\nu \circ \text{st}^{-1})(K).
$$

But $\nu \circ \text{st}^{-1}$ is a Radon measure by Theorem 3.1, so $P(\nu \circ \text{st}^{-1})$ is Radon, too, by assumption. This gives $P(\nu \circ \text{st}^{-1})(A) \leq \nu \circ \text{st}^{-1}(A)$ for every $A \in \mathcal{B}(X)$ and consequently $P(\nu \circ \text{st}^{-1})(A) = \nu \circ \text{st}^{-1}(A)$.

Step IV. It remains to prove non triviality. To this aim, note that, by (4.2), $m_n(Y_0) \geq \varepsilon$ for every $n \in \mathbb{N}$. By transfer, we have $\nu(\star Y_0) \geq \varepsilon$ and thus $\nu \circ \text{st}^{-1}(Y_0) \geq \nu(\star Y_0) = \nu_L(\star Y_0) \geq \varepsilon$.

Now, the invariant Radon distribution is given by $\mu_{\nu \circ \text{st}^{-1}}(X)$.

**Remark.** In the case of complete metric spaces, locally compact spaces or regular $\sigma$-compact spaces we could simplify the proof by considering a measure $\nu_L \circ \text{st}^{-1}$ instead of $\nu \circ \text{st}^{-1}$ (see [8, Corollary 3]).

The following example shows that some assumptions of the kind $P(\mathcal{M}_{\text{rad}}) \subset \mathcal{M}_{\text{rad}}$ are needed to obtain an invariant Radon measure.

**Example 4.2.** Let $X$ be a nonmeasurable set for Lebesgue measure $\lambda$ on $[0,1]$ with outer measure $\lambda(X) = 1$ and inner measure $\lambda(X) = 0$. Then $\mu(B) = \lambda(B)$ is a probability measure on Borel subsets of $X$ which is not Radon (see [6, Problem 7.1.9, p. 177]). Fix $x \in X$ and put $\mu_0 = \frac{1}{2} \mu + \frac{1}{2} \delta_x$. Let $P \nu = \nu(\cdot) \mu_0$ for every finite Borel measure on $X$. Then $P$ is a Feller operator which satisfies (4.1) but the only invariant distribution is $\mu_0$ which is not Radon.

**Corollary 4.3** (see [9, Theorem 6.1]). Let $X$ be a complete and separable metric space (Polish space). Assume that $P$ is a Feller operator and that there is a compact set $Y \subset X$ and a measure $\mu_0 \in \mathcal{M}_1$ such that
$$
\limsup_{n \to \infty} \frac{1}{n} \sum_{k=1}^n P^k \mu_0(Y) > 0.
$$

Then $P$ has an invariant distribution.

**Proof.** It is well known that in Polish spaces every finite Borel measure is Radon (the Ulam theorem [15]).
Corollary 4.4 (see [17, Corollary 2]). Assume $A$ is a compact subset of a Polish space $X$. Then either

$$\sup_{x \in X} \left| \frac{1}{n} \sum_{k=1}^{n} U^k I_A (x) \right| \to 0$$

or there exists an invariant distribution for a Feller operator $P$.

Proof. It is enough to notice that $U^k I_A (x) = P^k \delta_x (A)$ for each $k \in \mathbb{N}$. □

There are many natural examples of Markov operators which satisfy the condition $P (\mathcal{M}_{\text{rad}}) \subset \mathcal{M}_{\text{rad}}$.

Example 4.5. Let $X$ be a Hausdorff topological space and put

$$P \mu (A) = \sum_{i=1}^{n} \int_{S_i^{-1} (A)} p_i (x) \mu (dx) , A \in \mathcal{B} (X) ,$$

where $S_i : X \to X$, $p_i : X \to [0,1]$, $i = 1, \ldots, n$ are continuous functions with $\sum_{i=1}^{n} p_i (x) = 1$ for $x \in X$. The pair of sequences $(S_1, \ldots, S_n; p_1, \ldots, p_n)$ is called an iterated function system. Notice that

$$P \mu \left( \bigcup_{i=1}^{n} S_i (A) \right) \geq \mu (A)$$

for every $A \in \mathcal{B} (X)$ and hence $P$ transforms tight measures into tight measures. In perfectly normal spaces (metric spaces, in particular) Borel sets = Baire sets and hence all tight measures are Radon (see [6, Theorem 7.1.3]). In a general case, a similar conclusion holds if we restrict measures to Baire sets.

Example 4.6. Let $X$ be a Hausdorff topological space and put

$$P \mu (A) = \int_{X} \left( \int_{0}^{T} I_A (S (x,t)) p (x,t) dt \right) \mu (dx) , A \in \mathcal{B} (X) ,$$

where $S : X \times [0,T] \to X$ is a continuous function and $p : X \times [0,T] \to \mathbb{R}_+$ is Borel measurable and normalized. Operators of this kind appear in studying stochastically perturbed dynamical systems (see [10, 20]). Notice that $P \mu (S (A \times [0,T])) \geq \mu (A)$ for every $A \in \mathcal{B} (X)$ and, as in the previous example, $P$ transforms tight measures into tight measures.

References


Andrzej Wiśnicki
Institute of Mathematics
Maria Curie-Skłodowska University
pl. Marii Curie-Skłodowskiej 1
20-031 Lublin
Poland
e-mail: awisnic@hektor.umcs.lublin.pl

Received September 9, 2009