3

Markov and Frobenius–Perron operators

Taking into account the concepts of the preceding chapter, we are now ready to formally introduce the Frobenius–Perron operator, which, as we saw in Chapter 1, is of considerable use in studying the evolution of densities under the operation of deterministic systems.

However, as a preliminary step, we develop the more general concept of the Markov operator and derive some of its properties. Our reasons for this approach are twofold: First, as will become clear, many concepts concerning the asymptotic behavior of densities may be equally well formulated for both deterministic and stochastic systems. Second, many of the results that we develop in later chapters concerning the behavior of densities evolving under the influence of deterministic systems are simply special cases of more general results for stochastic systems.

The theory of Markov operators is extremely rich and varied, and we have chosen an approach particularly suited to an examination of the eventual behavior of densities in dynamical systems. Foguel [1969] contains an exhaustive survey of the asymptotic properties of Markov operators.

3.1 Markov operators

We define the Markov operator as follows.

Definition 3.1.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. Any linear operator $P: L^1 \to L^1$ satisfying

$$\text{(a)} \quad Pf \geqslant 0 \qquad \text{for } f \geqslant 0, f \in L^1; \quad \text{and}$$

$$\text{(b)} \quad \|Pf\| = \|f\|, \qquad \text{for } f \geqslant 0, f \in L^1 \tag{3.1.1}$$

is called a **Markov operator.**

Remark 3.1.1. In conditions (a) and (b), the symbols f and Pf denote elements of L^1 represented by functions that can differ on a set of measure zero. Thus, for any such function, properties $f \geqslant 0$ and $Pf \geqslant 0$ hold almost everywhere. When it is clear that we are dealing with elements of L^1 (or L^p), we will drop the "almost everywhere" notation. □

Markov operators have a number of properties that we will have occasion to use. First, if f, $g \in L^1$, then

$$Pf(x) \geqslant Pg(x) \qquad \text{whenever } f(x) \geqslant g(x) \,. \tag{3.1.2}$$

Any operator P satisfying (3.1.2) is said to be **monotonic**. To show the monotonicity of P is trivial, since $(f - g) \geqslant 0$ implies $P(f - g) \geqslant 0$.

To demonstrate further inequalities that Markov operators satisfy, we offer the following proposition.

Proposition 3.1.1. If $(X, \mathcal{A}, \mu)$ is a measure space and P is a Markov operator, then, for every $f \in L^1$,

$$\text{(M1)} \quad (Pf(x))^+ \leqslant Pf^+(x) \tag{3.1.3}$$

$$\text{(M2)} \quad (Pf(x))^- \leqslant Pf^-(x) \tag{3.1.4}$$

$$\text{(M3)} \quad |Pf(x)| \leqslant P|f(x)| \tag{3.1.5}$$

and

$$\text{(M4)} \quad \|Pf\| \leqslant \|f\| \,. \tag{3.1.6}$$

Proof: These inequalities are straightforward to derive. To obtain (3.1.3), note that from the definition of f^+ and f^-, it follows that

$$\begin{aligned} (Pf)^+ = (Pf^+ - Pf^-)^+ &= \max(0, Pf^+ - Pf^-) \\ &\leqslant \max(0, Pf^+) = Pf^+ ; \end{aligned}$$

and inequality (3.1.4) is obtained in an analogous fashion. Inequality (3.1.5) follows from (M1) and (M2), namely,

$$\begin{aligned} |Pf| &= (Pf)^+ + (Pf)^- \leqslant Pf^+ + Pf^- \\ &= P(f^+ + f^-) = P|f| \,. \end{aligned}$$

Finally, by integrating (3.1.5) over X, we have

$$\begin{aligned} \|Pf\| &= \int_X |Pf(x)|\mu(dx) \leqslant \int_X P|f(x)|\mu(dx) \\ &= \int_X |f(x)|\mu(dx) = \|f\| , \end{aligned}$$

which confirms (3.1.6). ■

Inequality (3.1.6) is extremely important, and any operator P that satisfies it is called a **contraction**. The actual inequality (3.1.6) is known as the **contractive property** of P. To illustrate its power note that for any $f \in L^1$ we have

$$\|P^n f\| = \|P(P^{n-1}f)\| \leq \|P^{n-1}f\|$$

and, thus, for any two $f_1, f_2 \in L^1$, $f_1 \neq f_2$,

$$\begin{aligned}\|P^n f_1 - P^n f_2\| &= \|P^n(f_1 - f_2)\| \\ &\leq \|P^{n-1}(f_1 - f_2)\| = \|P^{n-1}f_1 - P^{n-1}f_2\|. \end{aligned} \tag{3.1.7}$$

Inequality (3.1.7) simply states that during the process of iteration of two individual functions the distance between them can decrease, but it can never increase. This is referred to as the **stability property** of iterates of Markov operators.

By the **support** of a function g we simply mean the set of all x such that $g(x) \neq 0$, that is,

$$\operatorname{supp} g = \{x: g(x) \neq 0\}. \tag{3.1.8}$$

Remark 3.1.2. This is not the customary definition of the support of a function, which is usually defined by

$$\operatorname{supp} g = \operatorname{closure} \{x: g(x) \neq 0\}. \tag{3.1.9}$$

But, because the customary definition (3.1.9) requires the introduction of topological notions not used elsewhere, we have presented the slightly unusual definition (3.1.8). □

Remark 3.1.3. If g is an element of L^p, then the set (3.1.8) is not defined in a completely unique manner, since g may be represented by functions that differ on a set of measure zero. This inaccuracy never leads to any difficulties in calculating measures and integrals. Thus, it is customary to simplify the terminology and to talk about the supports of elements from L^p as if we were speaking of sets. However, if we want to emphasize that a relation between sets does not hold precisely but may be violated on a set of measure zero, we say that it holds **modulo zero.** Thus $A = B$ modulo zero means that the set of x in A that does not belong to B, or vice versa, has measure zero. □

One might wonder under what conditions the contractive property (3.1.6) is a strong inequality. The answer is quite simple.

Proposition 3.1.2. $\|Pf\| = \|f\|$ if and only if Pf^+ and Pf^- have disjoint supports.

Proof: We start from the inequality

$$|Pf^+(x) - Pf^-(x)| \leq |Pf^+(x)| + |Pf^-(x)|.$$

Clearly the inequality will be strong if both $Pf^+(x) > 0$ and $Pf^-(x) > 0$, while the equality holds if $Pf^+(x) = 0$ or $Pf^-(x) = 0$. Thus, by integrating over the space X, we have

$$\int_X |Pf^+(x) - Pf^-(x)|\mu(dx) = \int_X |Pf^+(x)|\mu(dx) + \int_X |Pf^-(x)|\mu(dx)$$

if and only if there is no set $A \in \mathcal{A}$, $\mu(A) > 0$, such that $Pf^+(x) > 0$ and $Pf^-(x) > 0$ for $x \in A$, that is, $Pf^+(x)$ and $Pf^-(x)$ have disjoint support. Since $f = f^+ - f^-$, the left-hand integral is simply $\|Pf\|$. Further, the right-hand side is $\|Pf^+\| + \|Pf^-\| = \|f^+\| + \|f^-\| = \|f\|$, so the proposition is proved. ■

Having developed some of the more important elementary properties of Markov operators, we now introduce the concept of a fixed point of P.

Definition 3.1.2. If P is a Markov operator and, for some $f \in L^1$, $Pf = f$ then f is called a **fixed point** of P.

From Proposition 3.1.1 it is easy to show the following.

Proposition 3.1.3. If $Pf = f$ then $Pf^+ = f^+$ and $Pf^- = f^-$.

Proof: Note that from $Pf = f$ we have

$$f^+ = (Pf)^+ \leq Pf^+ \quad \text{and} \quad f^- = (Pf)^- \leq Pf^-,$$

hence

$$\begin{aligned}
&\int_X [Pf^+(x) - f^+(x)]\mu(dx) + \int_X [Pf^-(x) - f^-(x)]\mu(dx) \\
&= \int_X [Pf^+(x) + Pf^-(x)]\mu(dx) - \int_X [f^+(x) + f^-(x)]\mu(dx) \\
&= \int_X P|f(x)|\mu(dx) - \int_X |f(x)|\mu(dx) \\
&= \|P|f|\| - \||f|\|.
\end{aligned}$$

However, by the contractive property of P we know that

$$\|P|f|\| - \||f|\| \leq 0.$$

Since both of the integrands $(Pf^+ - f^+)$ and $(Pf^- - f^-)$ are nonnegative, this last inequality is possible only if $Pf^+ = f^+$ and $Pf^- = f^-$. ■

Definition 3.1.3. Let $(X, \mathcal{A}, \mu)$ be a measure space and the set $D(X, \mathcal{A}, \mu)$ be defined by $D(X, \mathcal{A}, \mu) = \{f \in L^1(X, \mathcal{A}, \mu): f \geqslant 0 \text{ and } \|f\| = 1\}$. Any function $f \in D(X, \mathcal{A}, \mu)$ is called a **density**.

Definition 3.1.4. If $f \in D(X, \mathcal{A}, \mu)$, then the normalized measure

$$\mu_f(A) = \int_A f(x)\mu(dx) \qquad \text{for } A \in \mathcal{A}$$

is said to be **absolutely continuous** with respect to μ, and f is called the **density** of $\mu_f(A)$.

From Corollary 2.2.1 it follows that a normalized measure ν is absolutely continuous with respect to μ if $\nu(A) = 0$ whenever $\mu(A) = 0$. This property is often used as the definition of an absolutely continuous measure.

Using the notion of densities we may extend the concept of a fixed point of a Markov operator with the following definition.

Definition 3.1.5. Let $(X, \mathcal{A}, \mu)$ be a measure space and P be a Markov operator. Any $f \in D$ that satisfies $Pf = f$ is called a **stationary density** of P.

The concept of a stationary density of an operator is extremely important and plays a central role in many of the sections that follow.

3.2 The Frobenius–Perron operator

Having developed the concept of Markov operators and some of their properties, we are in a position to examine a special class of Markov operators, the Frobenius–Perron operator, which we introduced intuitively in Chapter 1.

We start with the following definitions.

Definition 3.2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. A transformation $S: X \to X$ is **measurable** if

$$S^{-1}(A) \in \mathcal{A} \qquad \text{for all } A \in \mathcal{A}.$$

Definition 3.2.2. A measurable transformation $S: X \to X$ on a measure space $(X, \mathcal{A}, \mu)$ is **nonsingular** if $\mu(S^{-1}(A)) = 0$ for all $A \in \mathcal{A}$ such that $\mu(A) = 0$.

Before stating a precise definition of the Frobenius–Perron operator, consider the following. Assume that a nonsingular transformation $S: X \to X$ on a measure space is given. We define an operator $P: L^1 \to L^1$ in two steps.

1. Let $f \in L^1$ and $f \geqslant 0$. Write

$$\int_{S^{-1}(A)} f(x)\mu(dx). \tag{3.2.1}$$

Because

$$S^{-1}(\bigcup_i A_i) = \bigcup_i S^{-1}(A_i),$$

it follows from property (L5) of the Lebesgue integral that the integral (3.2.1) defines a finite measure. Thus, by Corollary 2.2.1, there is a unique element in L^1, which we denote by Pf, such that

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx) \qquad \text{for A} \in \mathcal{A}.$$

2. Now let $f \in L^1$ be arbitrary, that is, not necessarily nonnegative. Write $f = f^+ - f^-$ and define

$$Pf = Pf^+ - Pf^-.$$

From this definition we have

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f^+(x)\mu(dx) - \int_{S^{-1}(A)} f^-(x)\mu(dx)$$

or, more completely,

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx), \qquad \text{for } A \in \mathcal{A}. \tag{3.2.2}$$

From Proposition 2.2.1 and the nonsingularity of S, it follows that equation (3.2.2) uniquely defines P.

We summarize these comments as follows.

Definition 3.2.3. Let $(X, \mathcal{A}, \mu)$ be a measure space. If $S: X \to X$ is a nonsingular transformation the unique operator $P: L^1 \to L^1$ defined by equation (3.2.2) is called the **Frobenius–Perron operator** corresponding to S.

It is straightforward to show from (3.2.2) that P has the following properties:

(FP1) $$P(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 Pf_1 + \lambda_2 Pf_2 \tag{3.2.3}$$

for all $f_1, f_2 \in L^1$, λ_1, $\lambda_1 \in R$, so P is a linear operator;

(FP2) $$Pf \geqslant 0 \quad \text{if} \quad f \geqslant 0; \qquad \text{and} \tag{3.2.4}$$

(FP3) $$\int_X Pf(x)\mu(dx) = \int_X f(x)\mu(dx) \tag{3.2.5}$$

(FP4) If $S_n = S \circ \overset{n}{\cdots} \circ S$ and P_n is the Frobenius–Perron operator corresponding to S_n, then $P_n = P^n$, where P is the Frobenius–Perron operator corresponding to S.

Remark 3.2.1. Although the definition of the Frobenius–Perron operator P by (3.2.2) is given by a quite abstract mathematical theorem of Radon–Nikodym, it should be realized that it describes the evolution of f by a transformation S. Properties (3.2.4–5) of the transformed distribution $Pf(x)$ are exactly what one would expect on intuitive grounds. □

Remark 3.2.2. From the preceding section the Frobenius–Perron operator is also a Markov operator.

As we wish to emphasize the close connection between the behavior of stochastic systems and the chaotic behavior of deterministic systems, we will formulate concepts and results for Markov operators wherever possible. The Frobenius–Perron operator is a particular Markov operator, and thus any property of Markov operators is immediately applicable to the Frobenius–Perron operator. □

In some special cases equation (3.2.2) allows us to obtain an explicit form for Pf. As we showed in Chapter 1, if $X = [a,b]$ is an interval on the real line R, and $A = [a,x]$, then (3.2.2) becomes

$$\int_a^x Pf(s)\,ds = \int_{S^{-1}([a,x])} f(s)\,\mathrm{ds}\,,$$

and by differentiating

$$Pf(x) = \frac{d}{dx}\int_{S^{-1}([a,x])} f(s)\,\mathrm{ds}\,. \tag{3.2.6}$$

It is important to note that in the special case that the transformation S is differentiable and invertible, an explicit form for Pf is available. If S is differentiable and invertible, then S must be monotone. Suppose S is an increasing function and S^{-1} has a continuous derivative. Then

$$S^{-1}([a,x]) = [S^{-1}(a), S^{-1}(x)]\,,$$

and from (3.2.6)

$$Pf(x) = \frac{d}{dx}\int_{S^{-1}(a)}^{S^{-1}(x)} f(s)\,ds = f(S^{-1}(x))\frac{d}{dx}[S^{-1}(x)]\,.$$

If S is decreasing, then the sign of the right-hand side is reversed. Thus, in the general one-dimensional case, for S differentiable and invertible with continuous dS^{-1}/dx,

$$Pf(x) = f(S^{-1}(x))\left|\frac{d}{dx}[S^{-1}(x)]\right|\,. \tag{3.2.7}$$

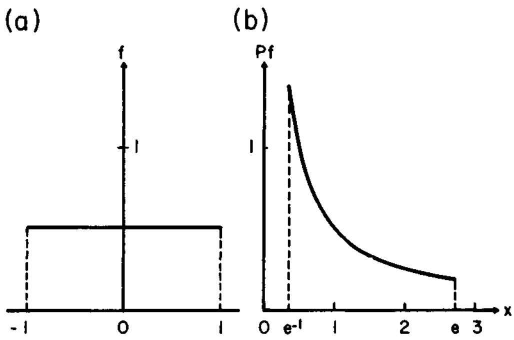

Figure 3.2.1. Operation of the Frobenius–Perron operator corresponding to $S(x) = e^x$, $x \in R$. (a) An initial density $f(x) = \frac{1}{2}1_{[-1,1]}(x)$ is transformed into the density $Pf(x) = (2x)^{-1}1_{[e^{-1},e]}(x)$ by S as shown in (b).

Example 3.2.1. To see how the Frobenius–Perron operator works, pick $S(x) = \exp(x)$. Thus from (3.2.7), we have

$$Pf(x) = (1/x)f(\ln x).$$

Consider what happens to an initial f given by

$$f(x) = \tfrac{1}{2}1_{[-1,1]}(x),$$

and shown in Figure 3.2.1a. Under the action of P, the function f is carried into

$$Pf(x) = (1/2x)1_{[e^{-1},e]}(x)$$

as shown in Figure 3.2.1b. □

Two important points are illustrated by this example. The first is that for an initial f supported on a set $[a, b]$, Pf will be supported on $[S(a), S(b)]$. Second, Pf is small where (dS/dx) is large and vice versa.

We generalize the first observation as follows.

Proposition 3.2.1. Let $S: X \to X$ be a nonsingular transformation and P the associated Frobenius–Perron operator. Assume that an $f \geq 0$, $f \in L^1$, is given. Then

$$\operatorname{supp} f \subset S^{-1}(\operatorname{supp} Pf) \tag{3.2.8}$$

and, more generally, for every set $A \in \mathcal{A}$ the following equivalence holds: $Pf(x) = 0$ for $x \in A$ if and only if $f(x) = 0$ for $x \in S^{-1}(A)$.

Proof: The proof is straightforward. By the definition of the Frobenius–Perron operator, we have

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx)$$

or

$$\int_X 1_A(x)Pf(x)\mu(dx) = \int_X 1_{S^{-1}(A)}(x)f(x)\mu(dx).$$

Thus $Pf(x) = 0$ on A implies, by property (L2) of the Lebesgue integral, that f(x) = 0 for $x \in S^{-1}(A)$ and vice versa. Now setting $A = X\backslash \mathrm{supp}(Pf)$, we have $Pf(x) = 0$ for $x \in A$ and, consequently, $f(x) = 0$ for $x \in S^{-1}(A)$, which means that $\mathrm{supp}\, f \subset X\backslash S^{-1}(A)$. Since $S^{-1}(A) = X\backslash S^{-1}(\mathrm{supp}(Pf))$ this completes the proof. ■

Remark 3.2.3. In the case of arbitrary $f \in L^1$, then, in Proposition 3.2.1 we only have: If $f(x) = 0$ for all $x \in S^{-1}(A)$ then $Pf(x) = 0$ for all $x \in A$. That the converse is not true can be seen by the following example. Take $S(x) = 2x$ (mod 1) and let

$$f(x) = \begin{cases} 1 & 0 \leqslant x < \frac{1}{2} \\ -1 & \frac{1}{2} \leqslant x \leqslant 1. \end{cases}$$

Then from (1.2.13) $Pf(x) = 0$ for all $x \in [0,1)$ but $f(x) \neq 0$ for any $x \in [0,1]$. □

For a second important case consider the rectangle $X = [a,b] \times [c,d]$ in the plane R^2. Set $A = [a,x] \times [b,y]$ so that (3.2.2) now becomes

$$\int_a^x ds \int_b^y Pf(s,t)\, dt = \iint\limits_{S^{-1}([a,x]\times[b,y])} f(s,t)\, ds\, dt.$$

Differentiating first with respect to x and then with respect to y, we have immediately that

$$Pf(x,y) = \frac{\partial^2}{\partial y\, \partial x} \iint\limits_{S^{-1}([a,x]\times[b,y])} f(s,t)\, ds\, dt.$$

Analogous formulas can be derived in the case of $X \subset R^d$.

In the general case, where $X = R^d$ and $S: X \to X$ is invertible, we can derive an interesting and useful generalization of equation (3.2.7). To do this we first state and prove a change of variables theorem based on the Radon–Nikodym theorem.

Theorem 3.2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, $f \in L^1 \cap L^\infty$, namely, a bounded integrable function, and $S: X \to X$ a nonsingular transformation. Then for every $A \in \mathcal{A}$,

$$\int_{S^{-1}(A)} f(S(x))\mu(dx) = \int_A f(x)\mu S^{-1}(dx) = \int_A f(x)J^{-1}(x)\mu(dx)$$

where μS^{-1} denotes the measure

$$\mu S^{-1}(B) = \mu(S^{-1}(B)), \qquad \text{for } B \in \mathcal{A},$$

and J^{-1} is the density of μS^{-1} with respect to μ, that is,

$$\mu(S^{-1}(B)) = \int_B J^{-1}(x)\mu(dx) \qquad \text{for B} \in \mathcal{A}.$$

Remark 3.2.4. We use the notation $J^{-1}(x)$ to draw the connection with differentiable invertible transformations on R^d, in which case $J(x)$ is the determinant of the Jacobian matrix:

$$J(x) = \left|\frac{dS(x)}{dx}\right| \quad \text{or} \quad J^{-1}(x) = \left|\frac{dS^{-1}(x)}{dx}\right|. \quad \square$$

Proof of Theorem 3.2.1: To prove this change of variables theorem, we recall Remark 2.2.6 and first take $f(x) = 1_B(x)$ so that $f(S(x)) = 1_B(S(x)) = 1_{S^{-1}(B)}(x)$ and, hence,

$$\begin{aligned}\int_{S^{-1}(A)} f(S(x))\mu(dx) &= \int_X 1_{S^{-1}(A)}(x)f(S(x))\mu(dx)\\ &= \int_X 1_{S^{-1}(A)}(x)1_{S^{-1}(B)}(x)\mu(dx)\\ &= \mu(S^{-1}(A) \cap S^{-1}(B)) = \mu(S^{-1}(A \cap B)).\end{aligned}$$

The second integral of the theorem may be written as

$$\int_A f(x)\mu S^{-1}(dx) = \int_X 1_A(x)1_B(x)\mu S^{-1}(dx) = \mu(S^{-1}(A \cap B))$$

whereas the third and last integral has the form

$$\begin{aligned}\int_A f(x)J^{-1}(x)\mu(dx) &= \int_A 1_B(x)J^{-1}(x)\mu(dx)\\ &= \int_{A\cap B} J^{-1}(x)\mu(dx) = \mu(S^{-1}(A \cap B)).\end{aligned}$$

Thus we have shown that the theorem is true for functions of the form $f(x) = 1_B(x)$. To complete the proof we need only to repeat it for simple functions

$f(x)$, which will certainly be true by linearity [property (L3)] of the Lebesgue integral. Finally, we may pass to the limit for arbitrary bounded and integrable functions f. [Note that f bounded is required for the integrability of $f(x)J^{-1}(x)$.] ■

With this change of variables theorem it is easy to prove the following extension of equation (3.2.7).

Corollary 3.2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, $S: X \to X$ an invertible nonsingular transformation, and P the associated Frobenius–Perron operator. Then for every $f \in L^1 \cap L^\infty$ (f is bounded and integrable),

$$Pf(x) = f(S^{-1}(x))J^{-1}(x). \tag{3.2.10}$$

Proof: By the definition of P, for $A \in \mathcal{A}$ we have

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx).$$

Change the variables in the right-hand integral with $y = S(x)$, so that

$$\int_{S^{-1}(A)} f(x)\mu(dx) = \int_A f(S^{-1}(y))J^{-1}(y)\mu(dy)$$

by Theorem 3.2.1. Thus we have

$$\int_A Pf(x)\mu(dx) = \int_A f(S^{-1}(x))J^{-1}(x)\mu(dx)$$

with the result that, by Proposition 2.2.1,

$$Pf(x) = f(S^{-1}(x))J^{-1}(x) \quad ■$$

3.3 The Koopman operator

To close this chapter, we define a third type of operator closely related to the Frobenius–Perron operator.

Definition 3.3.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, $S: X \to X$ a nonsingular transformation, and $f \in L^\infty$. The operator $U: L^\infty \to L^\infty$ defined by

$$Uf(x) = f(S(x)) \tag{3.3.1}$$

is called the **Koopman operator** with respect to S.

This operator was first introduced by Koopman [1931]. Due to the nonsingularity of S, U is well defined since $f_1(x) = f_2(x)$ a.e. implies $f_1(S(x)) = f_2(S(x))$ a.e. Operator U has some important properties:

(K1) $U(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 Uf_1 + \lambda_2 Uf_2$ (3.3.2)

for all $f_1, f_2 \in L^\infty$, $\lambda_1, \lambda_2 \in R$;

(K2) For every $f \in L^\infty$,

$$\|Uf\|_{L^\infty} \leq \|f\|_{L^\infty}, \tag{3.3.3}$$

that is, U is a contraction on L^∞;

(K3) For every $f \in L^1$, $g \in L^\infty$,

$$\langle Pf, g \rangle = \langle f, Ug \rangle \tag{3.3.4}$$

so that U is adjoint to the Frobenius–Perron operator P.

Property (K1) is trivial to check. Further, property (K2) follows immediately from the definition of the norm since $|f(x)| \leq \|f\|_{L^\infty}$ a.e. implies $|f(S(x))| \leq \|f\|_{L^\infty}$ a.e. The latter inequality gives equation (3.3.3) since, by (3.3.1), $Uf(x) = f(S(x))$.

Finally, to obtain (K3) we first check it with $g = 1_A$. Then the left-hand side of (3.3.4) becomes

$$\langle Pf, g \rangle = \int_X Pf(x) 1_A(x) \mu(dx) = \int_A Pf(x) \mu(dx),$$

while the right-hand side becomes

$$\begin{aligned}\langle f, Ug \rangle &= \int_X f(x) U 1_A(x) \mu(dx) \\ &= \int_X f(x) 1_A(S(x)) \mu(dx) = \int_{S^{-1}(A)} f(x) \mu(dx).\end{aligned}$$

Thus (K3) is equivalent to

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx)$$

which is the equation defining Pf. Because (K3) is true for $g(x) = 1_A(x)$ it is true for any simple function $g(x)$. Thus, by Remark 2.2.6, property (K3) must be true for all $g \in L^\infty$.

With the Koopman operator it is easy to prove that the Frobenius–Perron operator is weakly continuous. Precisely, this means that for every sequence $\{f_n\} \subset L^1$ the condition

$$f_n \to f \text{ weakly} \tag{3.3.5}$$

implies

$$Pf_n \to Pf \text{ weakly}. \tag{3.3.6}$$

To show this note that by property (K3) we have

$$\langle Pf_n, g\rangle = \langle f_n, Ug\rangle \qquad \text{for } g \in L^\infty.$$

Furthermore, from (3.3.5) it follows that $\langle f_n, Ug\rangle$ converges to $\langle f, Ug\rangle = \langle Pf, g\rangle$, which means that Pf_n converges weakly to Pf.

The same proof can be carried out for an arbitrary Markov operator P (or even more generally for every bounded linear operator). In this case we must use the fact that for every Markov operator there exists a unique adjoint operator $P^*: L^\infty \to L^\infty$ that satisfies

$$\langle Pf, g\rangle = \langle f, P^*g\rangle \qquad \text{for } f \in L^1,\ g \in L^\infty.$$