Perron-Frobenius Theorem

1 Hilbert Distance

Let $(V, \|\cdot\|)$ be a normed vector space. A subset $C \subset V$ is called a *convex cone* if $C+C \subset C$ and $\lambda C \subset C$ for all $\lambda \geq 0$. If furthermore $C \cap (-C) = \{0\}$ then C is said to be a *pointed convex cone*. When C is closed in the topology of V we say that C is a *closed convex cone*. Given a convex cone $C \subset V$, the set $E = C + (-C)$ is a linear subspace of V, referred to as the *linear span* of C. The interior of a convex cone, denoted by C°, is by definition the topological interior of C in its linear span. Similarly, we denote by ∂C the topological boundary of C in V.

Given a pointed closed convex cone $C \subset V$, we now introduce a pseudo distance $\theta_C : C^\circ \times C^\circ \to [0, +\infty)$ with the property that $\theta_C(x, y) = 0$ if there exists $\lambda > 0$ such that $y = \lambda x$. This means that θ_C is really a distance between rays contained in C and for that is called a *projective distance* in C.

Given a pointed closed convex cone $C \subset V$ with linear span V we define the following order relation in V: $x \preceq_C y$ if $y - x \in C$. This relation partially orders V.

(1) The order $\preceq_C$ is reflexive, i.e., $x \preceq_C x$, because $0 \in C$.

(2) The order $\preceq_C$ is antisymmetric, i.e., $x \preceq_C y$ and $y \preceq_C x \;\Rightarrow\; x = y$, because $C \cap (-C) = \{0\}$.

(3) Finally the order $\preceq_C$ is transitive, i.e., $x \preceq_C y$ and $y \preceq_C z \;\Rightarrow\; x \preceq_C z$, because $C + C \subset C$.

Moreover the convexity of C implies that the partial order $\preceq_C$ is compatible with the linear structure of V in the sense that for all $x, y, z \in V$

(4) $x \preceq_C y \;\Rightarrow\; x + z \preceq_C y + z$;

(5) $x \preceq_C y$ and $\lambda \geq 0 \;\Rightarrow\; \lambda x \preceq_C \lambda y$;

Finally, because C is closed, the partial order $\preceq_C$ is also closed.

(6) If $x_n \preceq_C y_n$, $x = \lim_{n\to\infty} x_n$ and $y = \lim_{n\to\infty} y_n \;\Rightarrow\; x \preceq_C y$.

Remark 1. *Given points $x, y \in C^\circ$, since they are interior points there exists $\delta > 0$ small enough such that $y - \delta x \in C$ and $x - \delta y \in C$. These two relations imply that*

$$\delta y \preceq_C x \preceq_C \delta^{-1} y.$$

Define now the functions $\alpha, \beta : C^\circ \times C^\circ \to [0, +\infty)$

$$\alpha(x,y) := \inf\{\lambda > 0 : x \preceq_C \lambda\, y\},$$
$$\beta(x,y) := \sup\{\mu > 0 : \mu\, y \preceq_C x\}.$$

By the previous remark, the sets where these infimum and supremum are taken are always non-empty. Hence the functions α and β are well-defined with positive and finite values.

Given a closed pointed convex cone $C \subset V$, with linear span V, we define the *Hilbert distance* $\theta_C : C^\circ \times C^\circ \to [0, +\infty)$

$$\theta_C(x,y) := \log \frac{\alpha(x,y)}{\beta(x,y)}.$$

Proposition 1. *The Hilbert distance θ_C is a projective pseudo-metric on C°.*

(1) $\theta_C(x,y) \geq 0$,

(2) $\theta_C(x,y) = 0 \quad \Leftrightarrow \; x = \lambda\, y$ *for some* $\lambda > 0$,

(3) $\theta_C(x,y) = \theta_C(y,x)$,

(4) $\theta_C(x,z) \leq \theta_C(x,y) + \theta_C(y,z)$.

Proof. For item (1) notice that

$$\beta\, y \preceq_C x \preceq_C \alpha\, y \tag{1}$$

with $\alpha = \alpha(x,y)$ and $\beta = \beta(x,y)$. Hence $\alpha \geq \beta$ and $\theta_C(x,y) = \log \frac{\alpha}{\beta} \geq 0$.

If $\theta_C(x,y) = 0$ then $\alpha = \beta$ and by (1) $\alpha\, y \preceq_C x \preceq_C \alpha\, y$, which implies $x = \alpha\, y$. Conversely, if $x = \lambda\, y$ then $\alpha(x,y) = \beta(x,y) = \lambda$, which implies $\theta_C(x,y) = 0$.

Relation (1) is equivalent to

$$\alpha^{-1}\, x \preceq_C y \preceq_C \beta^{-1}\, x,$$

from which

$$\alpha(y,x) = \beta(x,y)^{-1} \;\text{ and }\; \beta(y,x) = \alpha(x,y)^{-1}.$$

Hence

$$\theta_C(y,x) = \log \frac{\beta^{-1}}{\alpha^{-1}} = \log \frac{\alpha}{\beta} = \theta_C(x,y).$$

For the triangle inequality let us write $\alpha = \alpha(x,y)$, $\beta = \beta(x,y)$, $\alpha' = \alpha(y,z)$ and $\beta' = \beta(y,z)$. Then by (1)

$$\beta\, y \preceq_C x \preceq_C \alpha\, y \;\text{ and }\; \beta'\, z \preceq_C y \preceq_C \alpha'\, z.$$

These relations imply that

$$\beta\, \beta'\, z \preceq_C x \preceq_C \alpha\, \alpha'\, z.$$

Therefore $\beta\, \beta' \leq \beta(x,z) \leq \alpha(x,z) \leq \alpha\, \alpha'$ and

$$\theta_C(x,z) = \log \frac{\alpha(x,z)}{\beta(x,z)} \leq \log \frac{\alpha\, \alpha'}{\beta\, \beta'} = \log \frac{\alpha}{\beta} + \log \frac{\alpha'}{\beta'} = \theta_C(x,y) + \theta_C(y,z).$$

$\square$

Proposition 2. *Let $T : V_1 \to V_2$ be a bounded linear map between normed spaces V_1 and V_2 and for $i = 1, 2$ let $C_i \subset V_i$ be pointed closed convex cones with linear span V_i. If $T(C_1^\circ) \subset C_2^\circ$ then for all $x, y \in C_1^\circ$,*

$$\theta_{C_2}(Tx, Ty) \leq \theta_{C_1}(x, y).$$

Moreover, if T is an isomorphism such that $T(C_1^\circ) = C_2^\circ$ then for all $x, y \in C_1^\circ$,

$$\theta_{C_2}(Tx, Ty) = \theta_{C_1}(x, y).$$

Proof. Since the Hilbert distances θ_{C_i} are defined in terms of the partial orders $\preceq_{C_i}$, it is enough to remark that if $T(C_1) \subset C_2$ then

$$x \preceq_{C_1} y \quad \Rightarrow \quad Tx \preceq_{C_2} Ty,$$

and if $T(C_1) = C_2$ then

$$x \preceq_{C_1} y \quad \Leftrightarrow \quad Tx \preceq_{C_2} Ty.,$$

□

The Hilbert distance between two points is a 2d measurement. Given two vectors $x, y \in C^\circ$ consider the subspace E generated by $\{x, y\}$ and apply the following proposition.

Proposition 3. *Let V be a normed space. Given a pointed closed convex cone $C \subset V$ with linear span V and any subspace $E \subset V$, then $(C \cap U)^\circ = C^\circ \cap E$ and for all $x, y \in C^\circ \cap E$,*

$$\theta_C(x, y) = \theta_{C \cap E}(x, y).$$

Proof. Just notice that for all $x, y \in E$,

$$x \preceq_C y \quad \Leftrightarrow \quad x \preceq_{C \cap E} y.$$

□

Given four (ordered) points A, X, Y, B on a straight line ℓ in some Euclidean space, their *cross-ratio* is the quotient

$$(A, X, Y, B) := \frac{|AY|\ |XB|}{|AX|\ |YB|}.$$

This number equals 1 when $X = Y$ and otherwise it is greater than 1.

Proposition 4. *Let V be a normed space. Given a pointed closed convex cone $C \subset V$ with linear span V and two points $x, y \in C^\circ$ let $a, b \in \partial C$ be such that a, x, y, b form an ordered sequence of points in the line joining x to y. Then*

$$\theta_C(x, y) = \log(a, x, y, b).$$

Proof. By Proposition 4 we just need to check this formula for two-dimensional cones. By Proposition 2 we can pick any such cone, say $C = \mathbb{R}^2_+ \subset \mathbb{R}^2$. Given $x = (x_1, x_2), y = (y_1, y_2) \in C^\circ = (0, +\infty)^2$, rescaling these vectors we can assume that $x_1+x_2 = 1 = y_1+y_2$. Assume for instance that $\frac{x_1}{x_2} \geq \frac{y_1}{y_2}$ so that the points $a = (1, 0)$, x, y and $b = (0, 1)$ are collinear and ordered with $a, b \in \partial C$. A simple calculation shows that

$$\alpha(x, y) = \max\{\frac{x_1}{y_1}, \frac{x_2}{y_2}\} = \frac{x_1}{y_1},$$
$$\beta(x, y) = \min\{\frac{x_1}{y_1}, \frac{x_2}{y_2}\} = \frac{x_2}{y_2}.$$

Hence, because $\|x-a\| = \sqrt{2}\,(1-x_1)$, $\|x-b\| = \sqrt{2}\,x_1$, $\|y-a\| = \sqrt{2}\,(1-y_1)$ and $\|x-b\| = \sqrt{2}\,y_1$,

$$\begin{aligned}\theta_C(x, y) &= \log \frac{x_1}{y_1} \cdot \frac{y_2}{x_2} = \log \frac{x_1}{y_1} \cdot \frac{1-y_1}{1-x_1} \\ &= \log \frac{\|y-a\|\,\|x-b\|}{\|x-a\|\,\|y-b\|} = \log(a, x, y, b).\end{aligned}$$

□

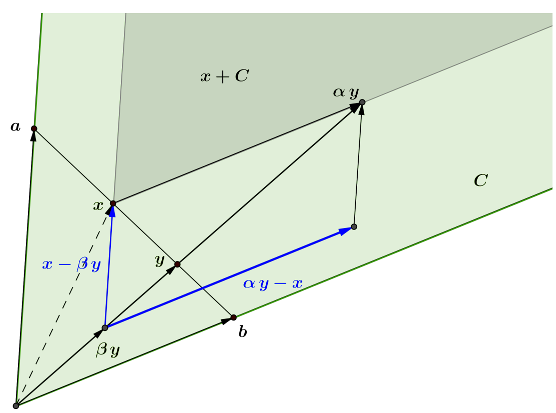

Figure 1: The vectors $x - \beta\, y$ and $\alpha\, y - x$ lie on the boundary of C.

Theorem 1 (G. Birkhoff). *Let $T : V_1 \to V_2$ be a bounded linear map between normed spaces V_1 and V_2 and for each $i = 1, 2$ let $C_i \subset V_i$ be pointed closed convex cones with linear span V_i. If $T(C_1) \subset C_2^\circ$ and the diameter of $T(C_1)$ in C_2° w.r.t. θ_{C_2} is $\leq \Delta$ then for all $x, y \in C_1^\circ$,*

$$\theta_{C_2}(Tx, Ty) \leq (1 - e^{-\Delta/2})\, \theta_C(x, y).$$

In particular the map $T : C_1 \to C_2^\circ$ is a strict contraction w.r.t. θ_{C_1} and θ_{C_2}.

Proof. Take two points $x, y \in C_1^\circ$ and let $\alpha = \alpha(x,y)$ and $\beta = \beta(x,y)$ so that relations (1) hold. Then, because C_1 is closed the vectors $x - \beta\, y$ and $\alpha\, y - x$ lie on the boundary of the cone C_1 (see Figure 1). Hence $\theta_{C_2}(T(\alpha\, y - x), T(x - \beta\, y)) \leq \Delta$. This means there are positive numbers $\lambda \geq \mu$ such that $\log(\lambda/\mu) \leq \Delta$ and

$$\mu\, T(x - \beta\, x) \preceq_{C_2} T(\alpha y - x) \preceq_{C_2} \lambda\, T(x - \beta\, x).$$

These relations imply that

$$\frac{\alpha + \beta\,\lambda}{1+\lambda}\, T(y) \preceq_{C_2} T(x) \preceq_{C_2} \frac{\alpha + \beta\,\mu}{1+\mu}\, T(y).$$

Therefore

$$\begin{aligned}
\theta_{C_2}(T(x), T(y)) &\leq \log\left(\frac{\alpha + \beta\,\mu}{1+\mu} \cdot \frac{1+\lambda}{\alpha + \beta\,\lambda}\right) = \log \frac{(\frac{\alpha}{\beta} + \mu)\,(1+\lambda)}{(1+\mu)\,(\frac{\alpha}{\beta} + \lambda)} \\
&= \log\left(\frac{\alpha}{\beta} + \mu\right) - \log(1+\mu) - \log\left(\frac{\alpha}{\beta} + \lambda\right) + \log(1+\lambda) \\
&= \int_0^{\log(\alpha/\beta)} \left[\frac{e^x}{e^x + \mu} - \frac{e^x}{e^x + \lambda}\right] dx = \int_0^{\log(\alpha/\beta)} \frac{(\lambda - \mu)\, e^x}{(e^x + \lambda)\,(e^x + \lambda)}\, dx \\
&\leq \log\left(\frac{\alpha}{\beta}\right) \max_{x \geq 0} \frac{(\lambda - \mu)\, e^x}{(e^x + \lambda)\,(e^x + \lambda)} = \log\left(\frac{\alpha}{\beta}\right) \frac{(\lambda - \mu)\,\sqrt{\lambda\,\mu}}{(\sqrt{\lambda\,\mu} + \lambda)\,(\sqrt{\lambda\,\mu} + \lambda)} \\
&= \log\left(\frac{\alpha}{\beta}\right) \frac{\sqrt{\lambda} - \sqrt{\mu}}{\sqrt{\lambda} + \sqrt{\mu}} = \log\left(\frac{\alpha}{\beta}\right) \left(1 - \frac{2\,\sqrt{\mu}}{\sqrt{\lambda} + \sqrt{\mu}}\right) \\
&\leq \theta_{C_1}(x,y)\,\left(1 - \sqrt{\frac{\mu}{\lambda}}\right) \leq \theta_{C_1}(x,y)\,\left(1 - e^{-\Delta/2}\right)
\end{aligned}$$

which concludes the proof. Above we have made use of the integral relation

$$\log(K + \lambda) - \log(1+\lambda) = \int_0^{\log K} \frac{e^x}{e^x + \lambda}\, dx.$$

Notice also that the function $\varphi(x) := \frac{e^x}{(e^x+\lambda)\,(e^x+\mu)}$ has derivative

$$\varphi'(x) = \frac{e^x\,(\mu\lambda\, - e^{2x})}{(e^x+\lambda)^2\,(e^x+\mu)^2}$$

which explains that the global maximum of $\varphi(x)$ is attained when $e^x = \sqrt{\lambda\,\mu}$. $\square$

Given any subset $A \subset V$ of a normed space V let

$$\mathrm{dist}(x, A) := \inf\{\|x - a\| \colon a \in A\}.$$

Given subsets $A, B \subset V$ let

$$\mathrm{dist}(A, B) := \inf\{\|x - y\| \colon\ x \in A,\ y \in B\}.$$

Proposition 5. *Let V be a normed space and $C \subset V$ a pointed closed convex cone with linear span V. Let $H \subset V$ be a subspace such that $C \cap H$ is bounded in norm*[1]*, i.e., $\|x\| \leq L$ for all $x \in C \cap H$. Then for all $x, y \in C^\circ \cap H$ with $\operatorname{dist}(x, \partial C) \geq \varepsilon$ and $\operatorname{dist}(y, \partial C) \geq \varepsilon \geq 0$*

$$\frac{1}{L} \|x - y\| \leq \theta_C(x, y) \leq \frac{2}{\varepsilon} \|x - y\|.$$

In particular convergence in the Hilbert distance θ_C implies norm convergence.

Proof. Given vectors $x, y \in C^\circ \cap H$ find vectors $a, b \in \partial C$ such that a, x, y, b are ordered points in some line $\ell \subset H$. Because these four points are in $C \cap H$, $\|x - a\| \leq L$ and $\|y - b\| \leq L$. Therefore

$$\begin{aligned}
\theta_C(x, y) = \log(a, x, y, b) &= \log \frac{\|y - a\| \, \|x - b\|}{\|x - a\| \, \|y - b\|} \\
&= \log \frac{(\|x - a\| + \|x - y\|) \, (\|x - y\| + \|y - b\|)}{\|x - a\| \, \|y - b\|} \\
&= \log \left(1 + \frac{\|x - y\|}{\|x - a\|}\right) + \log \left(1 + \frac{\|x - y\|}{\|y - b\|}\right) \\
&\geq 2 \, \log \left(1 + \frac{\|x - y\|}{L}\right) \geq \frac{\|x - y\|}{L}.
\end{aligned}$$

In the last step we have used that $\log(1 + x) \geq \frac{x}{2}$ for all $0 \leq x \leq 1$. For the converse inequality we use instead that $\log(1 + x) \leq x$ for all $x > 0$. Since a and b lie in the boundary of C, $\|x - a\| \geq \operatorname{dist}(x, \partial C) > \varepsilon$ and $\|y - b\| \geq \operatorname{dist}(y, \partial C) > \varepsilon$. Hence

$$\begin{aligned}
\theta_C(x, y) &= \log \left(1 + \frac{\|x - y\|}{\|x - a\|}\right) + \log \left(1 + \frac{\|x - y\|}{\|y - b\|}\right) \\
&\leq 2 \, \log \left(1 + \frac{\|x - y\|}{\varepsilon}\right) \leq \frac{2 \, \|x - y\|}{\varepsilon}.
\end{aligned}$$

□

Corollary 6. *If V is a Banach space and H is closed in V then $C^\circ \cap H$ is complete with respect to the metric θ_C.*

Proof. To prove that $(C^\circ \cap H, \theta_C)$ is complete take a Cauchy sequence $\{x_n\} \subset C^\circ \cap H$. By Proposition 5, $\{x_n\}$ is a Cauchy sequence in H w.r.t. the norm distance. Because V is a Banach space and H is closed in V, the limit $x = \lim_{n \to \infty} x_n$ exists with $x \in C \cap H$.

Given $x, y \in C^\circ \cap H$ we can not have $x \preceq_C y$, nor $y \preceq_C x$, for otherwise $C \cap H$ would contain a ray. Hence $\beta(x, y) \leq 1 \leq \alpha(x, y)$ for all $x, y \in C^\circ \cap H$.

Given $\epsilon > 0$, since $\{x_n\} \subset C^\circ \cap H$ is a Cauchy sequence, there exists $n_0 \in \mathbb{N}$ such that for all $n, m \geq n_0$,

$$\max\{\log \alpha(x_n, x_m), - \log \beta(x_n, x_m)\} \leq \theta_C(x_n, x_m) < \epsilon.$$

[1] This condition expresses the transversality between H and C because it prevents rays in C from being contained in H.

Hence

$$\beta_{n,m}\, x_m \preceq_C x_n \preceq_C \alpha_{n,m}\, x_m$$

for some $e^{-\epsilon} \leq \beta_{n,m} \leq 1 \leq \alpha_{n,m} \leq e^{\epsilon}$. Fix now $n \geq n_0$ and let m tend to $+\infty$. Then x_m converges to x and we can choose a subsequence $m_k \to +\infty$ such that $\alpha_{n,m_k} \to \alpha_n$ and $\beta n, m_k \to \beta_n$. Taking the limit, because the cone C is closed we get

$$\beta_n\, x \preceq_C x_n \preceq_C \alpha_n\, x$$

with $e^{-\epsilon} \leq \beta_n \leq 1 \leq \alpha_n \leq e^{\epsilon}$, which implies that $\theta_C(x_n, x) \leq 2\,\epsilon$ for all $n \geq n_0$. Hence $x \in C^\circ$ and the sequence $\{x_n\}$ converges to x w.r.t. θ_C. □

Corollary 7. *Given a norm compact subspace $K \subset C^\circ \cap H$, the metric θ_C induces the norm topology in K. In particular K has finite diameter w.r.t. θ_C.*

Proof. Because K is compact and $K \subset C^\circ$ we have $\operatorname{dist}(K, \partial C) > 0$. Then by Proposition 5 the metric θ_C and the norm distance are equivalent over K. □

2 Perron-Frobenius Theorem

A matrix $A \in \mathrm{Mat}_k(\mathbb{R})$ with non-negative entries is called *primitive* if there exists $m \geq 1$ such that all entries of A^m are strictly positive.

Theorem 2 (Perron-Frobenius)**.** *Given a matrix $A \in \mathrm{Mat}_k(\mathbb{R})$ with non-negative entries, if A is primitive then*

(a) there exists $\lambda > 0$ and $v \in \mathbb{R}^k$ with strictly positive entries such that $A\,v = \lambda\, v$;

(b) λ is a simple eigenvalue, i.e., $\dim \mathrm{Ker}(A - \lambda\, I) = 1$;

(c) λ is a dominant eigenvalue, i.e., $|\alpha| < \lambda$ for every other eigenvalue α of A;

(d) The limit $Q = \lim_{n\to+\infty} \lambda^{-n}\, A^n$ exists and every column of Q is a λ-eigenvector.

Proof. The cone $C = \mathbb{R}^k_+$ is pointed, closed and convex. Any matrix A with non-negative entries leaves the cone C invariant, i.e., $A\,C \subset C$. If A is primitive and $m \geq 1$ is such that all entries of A^m are strictly positive then $A^m C \subset C^\circ$. By Theorem 1 there exists $0 < \kappa < 1$ such that for all $x, y \in C^\circ$ and $q \geq 0$, $\theta_C(A^{qm}x, A^{qm}y) \leq \kappa^q\, \theta_C(x, y)$. On the other hand, by Proposition 2, for all $x, y \in C^\circ$ and $i \geq 0$, $\theta_C(A^i x, A^i y) \leq \theta_C(x, y)$. Thus dividing any given integer $n \geq 1$ by m, $n = qm + i$ for some $0 \leq i < m$ and

$$\begin{aligned} \theta_C(A^n x, A^n y) &= \theta_C(A^{qn+i}x, A^{qn+i}y) \leq \theta_C(A^{qn}x, A^{qn}y) \\ &\leq \kappa^n\, \theta_C(x, y) \leq L\,(\kappa^{1/m})^n\, \theta_C(x, y) \end{aligned}$$

with $L = \kappa^{-m}$, i.e., the projective action of A on C° is contractive.

By the Banach fixed point theorem there exists $v \in C^\circ$ such that for all $x \in C^\circ$, $\lim_{n\to+\infty} \theta_C(A^n x, v) = 0$. Thus $\theta_C(A\,v, v) = 0$ and there exists $\lambda > 0$ such that $A\,v = \lambda\, v$, which proves (a).

Item (b) follows because there exists a unique ray in $\{0\} \cup C^\circ$ fixed by the projective action of A on C.

Let H denote the direct sum of all eigenspaces associated with eigenvalues of A distinct from λ. Since $A\,H \subset H$ and $\lambda^{-1} A\, v = v$, the linear map $\hat{A}(x) = \lambda^{-1} A x$ leaves invariant H as well as the affine subspace $v + H$. The convex set $K = C \cap (v + H)$ is compact (prove it). The set K is also $\hat{A}$-invariant. By Proposition 5, for all $x \in K^\circ = C^\circ \cap (v + H)$,

$$\|\lambda^{-n} A^n x - v\| \lesssim \theta_C(\hat{A}^n(x), v) = \theta_C(A^n(x), v) \leq L\, \kappa^{\frac{n}{m}}.$$

This implies (prove it) that for all $n \geq 1$ and $x \in H$

$$\|\lambda^{-n} A^n x\| \lesssim \kappa^{\frac{n}{m}}. \tag{2}$$

Hence, for any vector $x \in \mathbb{R}^k$, if $x = \mu\, v + h$, with $h \in H$, then $\lim_{n\to\infty} \lambda^{-n} A^n x = \mu\, v$, which proves (d). Finally item (c) holds because the bound (2) for vectors in H implies that $|\alpha| \leq \lambda\, \kappa^{1/m} < \lambda$ for every other eigenvalue α of A. □

3 Mixing property of the Gauss Map

In this section we prove the mixing property of the Gauss map $T : [0, 1) \to [0, 1)$, defined by $Tx := \frac{1}{x} - \lfloor \frac{1}{x} \rfloor$. This map preserves the measure $\mu = \frac{1}{\log 2} \frac{dx}{1+x}$.

Theorem 3. *The Gauss map (T, μ) is mixing.*

We can conjugate the Gauss map (T, μ) to a piecewise expanding map $f : [0, 1) \to [0, 1)$ preserving the Lebesgue measure $m = dx$ on the interval $[0, 1)$. For this consider the orientation preserving diffeomorphism $h : [0, 1] \to [0, 1]$, $h(x) := \frac{\log(1+x)}{\log 2}$. Because

$$\mu([0, x)) = \frac{1}{\log 2} \int_0^x \frac{dt}{1+t} = \frac{\log(1+x)}{\log 2} = h(x) = m([0, h(x))) = m(h[0, x))$$

the map $h : [0, 1) \to [0, 1)$ is a measure preserving transformation from $([0, 1), \mu)$ onto $([0, 1), m)$. Hence, the map $f := h \circ T \circ h^{-1}$ is a piecewise expanding map which preserves the the Lebesgue measure, and (f, m) is conjugated to (T, μ).

The map f is given by

$$f(x) = \frac{1}{\log 2} \log\left(\frac{1}{2^x - 1} - (n-1)\right) \qquad \text{for all } \; a_{n+1} \leq x < a_n$$

where $a_0 = 1$ and $a_n = h(1/n)$ if $n \geq 1$. The map f is piecewise expanding, in fact $|f'(x)| \geq 2$ for all $x \in [0, 1)$.

Theorem 3 reduces to prove that:

Proposition 8. *The system (f, m) is mixing.*

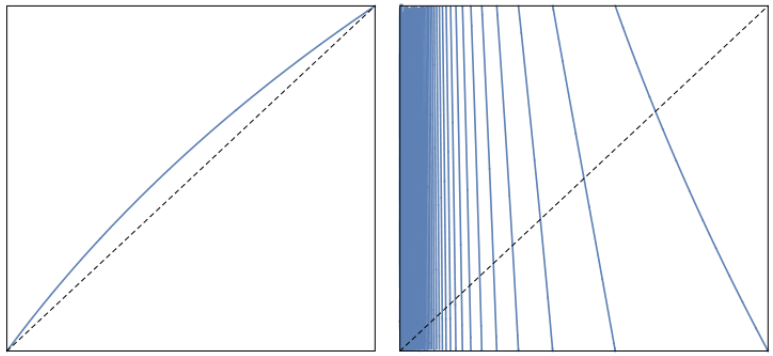

Figure 2: Graphs of h (left) and f (right).

We analyze the adjoint U_f^* of the Koopman operator $U_f : L^2([0,1),m) \to L^2([0,1),m)$, $U_f(\varphi) := \varphi \circ f$. To make U_f^* explicit we need the *inverse branches* of $f : [0,1) \to [0,1)$, which are the diffeomorphisms $g_n : [0,1] \to [a_{n-1}, a_n]$ given by

$$g_n(x) := \frac{\log(2^x + n)}{\log 2} - \frac{\log(2^x + n - 1)}{\log 2}.$$

Proposition 9. *The inverse branches of f satisfy for all $n \geq 1$ and all $x \in [0,1]$,*

(1) $g_n'(x) = \dfrac{2^x}{2^x + n} - \dfrac{2^x}{2^x + n - 1} < 0$;

(2) $|g_1'(x)| \leq \frac{1}{2}$;

(3) $|g_n'(x)| \leq (\sqrt{n} + \sqrt{n-1})^{-2} \leq \frac{1}{2}$ *when* $n \geq 2$;

(4) $\sum_{n=1}^{\infty} |g_n'(x)| = 1$;

(5) $g_n''(x) = \dfrac{2^x\,(n^2 - n - 4^x)\,\log 2}{(2^x+n)^2\,(2^x+n-1)^2}$;

(6) $|(\log |g_n'(x)|)'| = \frac{|g_n''(x)|}{|g_n'(x)|} \leq 1$.

Proof. The proof is left as an exercise. □

Proposition 10. *The adjoint U_f^* of the Koopman operator is given by*

$$U_f^*(\psi)(x) = \sum_{n=1}^{\infty} \psi(g_n(x))\;|g_n'(x)| = \sum_{y \in f^{-1}(x)} \frac{\psi(y)}{|f'(y)|}$$

for all $\psi \in L^2([0,1),m)$.

Proof. Take L^2 observables φ and ψ on $[0,1]$, decompose the integral on $[0,1]$ over the subintervals $[a_{n+1}, a_n]$ and perform the changes of variables $x = g_n(y)$ with $0 \le x \le 1$.

$$\begin{aligned}
\langle U_f(\varphi), \psi\rangle &= \int_0^1 \varphi(f(x))\,\psi(x)\,dx = \sum_{n=1}^{\infty} \int_{a_{n+1}}^{a_n} \varphi(f(x))\,\psi(x)\,dx \\
&= \sum_{n=1}^{\infty} \int_1^0 \varphi(y)\,\psi(g_n(y))\,g_n'(y)\,dy \\
&= \int_0^1 \varphi(y) \left(\sum_{n=1}^{\infty} \psi(g_n(y))\ |g_n'(y)| \right) dy.
\end{aligned}$$

□

Given $a = (a_1, \dots, a_n) \in \mathbb{N}^n$, with the convention that $\mathbb{N} = \{1, 2, \dots\}$, the map $g_a := g_{a_1} \circ \cdots \circ g_{a_n}$ is a diffeomorphism which maps $[0,1]$ onto the interval bounded between the h images of the values of the finite continued fractions with coefficients $[a_1, \dots, a_{n-1}, a_n]$ and $[a_1, \dots, a_{n-1}, a_n + 1]$, respectively. The images of these maps, for $a \in \mathbb{N}^n$, partition the interval $[0,1]$ mod 0.

Proposition 11. *The iterates of U_f^* are given by*

$$(U_f^*)^n(\psi)(x) = \sum_{a \in \mathbb{N}^n} \psi(g_a(x))\ |g_a'(x)|\,.$$

Proof. The proof is left as an exercise. □

Proposition 12. *For all $a \in \mathbb{N}^n$ and $x, y \in [0,1]$,*

$$e^{-2\,|x-y|} \le \frac{|g_a'(x)|}{|g_a'(y)|} \le e^{2\,|x-y|}.$$

Proof. Given $x, y \in [0,1]$, define $x_0 = x$, $y_0 = y$, $x_j = (g_{a_{n-j+1}} \circ \ldots \circ g_{a_n})(x)$ and $y_j = (g_{a_{n-j+1}} \circ \ldots \circ g_{a_n})(y)$ for all $1 \le j \le n$, so that $|x_j - y_j| \le 2^{-j}\ |x - y|$. Then

$$\begin{aligned}
\left| \log \frac{g_a'(x)}{g_a'(y)} \right| &= |\log|g_a'(x)| - \log|g_a'(y)|| \le \sum_{j=0}^{n-1} \left| \log \left| g_{a_{j+1}}'(x_j) \right| - \log \left| g_{a_{j+1}}'(y_j) \right| \right| \\
&\le \sum_{j=0}^{n-1} |x_j - y_j| \le \sum_{j=0}^{n-1} 2^{-j}\ |x - y| \le 2\ |x - y|
\end{aligned}$$

which proves the proposition. □

The previous statement is usually referred to as the *bounded distortion* of f, which can be reformulated as follows: Given two points $x, y \in [0,1)$, if $f^j(x)$ and $f^j(y)$ lie on the same branch of f for all $j = 0, 1, \dots, n-1$, then

$$e^{-2\,|x-y|} \le \frac{|(f^n)'(x)|}{|(f^n)'(y)|} \le e^{2\,|x-y|}.$$

Lemma 1. *Given $\psi \in C^1([0,1])$ and an interval $J \subset [0,1]$,*

$$\max_J |\psi| \leq \int_J |\psi'| \ dm + \frac{1}{m(J)} \int_J |\psi| \ dm$$

Proof. The maximum value of $|\psi(x)|$ on J is bounded by its mean value over J plus the total length of the curve traced by the $\psi(x)$. □

Proposition 13. *The adjoint $U_f^* : L^2([0,1), m) \to L^2([0,1), m)$ of the Koopman operator satisfies the following properties:*

(a) $\|U_f^*(\psi)\|_\infty \leq \|\psi\|_\infty$ *for all* $\psi \in C^0([0,1])$;

(b) It is a positive operator, i.e., $\psi \geq 0 \Rightarrow U_f^*(\psi) \geq 0$;

(c) $\int_0^1 U_f^*(\psi) \, dm = \int_0^1 \psi \, dm$;

(d) It fixes the constant functions, i.e., $U_f^*(\mathbf{1}) = \mathbf{1}$;

(e) It preserves continuity, i.e., it maps $C^0([0,1])$ *to* $C^0([0,1])$;

(f) It preserves C^1 *regularity, i.e., it maps* $C^1([0,1])$ *to* $C^1([0,1])$;

(g) $\|((U_f^*)^n \psi)'\|_\infty \leq \frac{1+e^2}{2^n} \|\psi'\|_\infty + e^4 \int_0^1 |\psi| \ dm$ *for all* $\psi \in C^1([0,1])$.

Proof. Having the explicit formula for the adjoint operator U_f^*, items (a)-(d) become the matter of a simple verification. These properties are valid for the adjoint of the Koopman operator of any measure preserving transformation, whose proof we leave as an exercise.

Let us prove items (e), (f) and (g). Given $\psi \in C^0([0,1])$, the compositions $|g_n'(y)| \ \psi(g_n(y))$ are continuous functions. Thus, by Weierstrass M-test and property (4) of Proposition 9, $U_f^*(\psi) \in C^0([0,1])$, which proves (e).

For each $a \in \mathbb{N}^n$, by the mean value theorem the interval $J_a = g_a([0,1])$ has length $m(J_a) = |g_a'(\xi_a)|$ for some $\xi_a \in (0,1)$. By Proposition 12, for all $x, y \in [0,1]$,

$$|\log |g_a'(x)| - \log |g_a'(y)|| = \left| \log \frac{|g_a'(x)|}{|g_a'(y)|} \right| \leq 2 \ |x-y|$$

which implies that for all $x \in [0,1]$,

$$\frac{|g_a''(x)|}{|g_a'(x)|} = |(\log |g_a'(x)|)'| \leq e^2.$$

Given $\psi \in C^1([0,1])$, the function $|g_a'(y)| \ \psi(g_a(y))$ is of class C^1 and by Lemma 1 we obtain the following bound on its derivative

$$\begin{aligned}
&\left|\frac{d}{dy}\ |g_a'(y)|\ \psi(g_a(y))\right| = \psi'(g_a(x))\ |g_a'(x)|^2 + |\psi(g_a(x))|\ |g_a''(x)| \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty\ |g_a'(x)| + \sup_{J_a}|\psi|\ \frac{|g_a''(x)|}{|g_a'(x)|}\ |g_a'(x)| \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty\ |g_n'(x)| + e^2\ \sup_{J_a}|\psi|\ |g_a'(x)| \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty\ |g_n'(x)| + e^2\ |g_a'(x)|\ \left(\int_{J_a} |\psi'|\ dm + \frac{1}{m(J_a)}\ \int_{J_a} |\psi|\ dm\right) \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty\ |g_n'(x)| + \frac{e^2}{2^n}\ \int_{J_a} |\psi'|\ dm + e^2\ \frac{|g_a'(x)|}{|g_a'(\xi_a)|}\ \int_{J_a} |\psi|\ dm \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty\ |g_n'(x)| + \frac{e^2}{2^n}\ \int_{J_a} |\psi'|\ dm + e^4\ \int_{J_a} |\psi|\ dm.
\end{aligned}$$

By Weierstrass M-test, the function $(U_f^*)^n\psi$ is of class C^1 with a derivative satisfying

$$\begin{aligned}
\left|((U_f^*)^n\psi)'(x)\right| &\leq \sum_{a\in\mathbb{N}^n} \left|\frac{d}{dy}\ |g_a'(y)|\ \psi(g_a(y))\right| \\
&\leq \frac{1}{2^n}\ \|\psi'\|_\infty + \frac{e^2}{2^n}\ \int_0^1 |\psi'|\ dm + e^4\ \int_0^1 |\psi|\ dm \\
&\leq \frac{1+e^2}{2^n}\ \|\psi'\|_\infty + e^4\ \int_0^1 |\psi|\ dm.
\end{aligned}$$

This proves item (g). □

Lemma 2. *For every $\psi \in C^0([0,1])$, if $\int_0^1 \psi\, dm = 1$ and $\psi \geq 0$ then there exist $N \in \mathbb{N}$ and $\delta > 0$ such that $(U_f^*)^n\psi \geq \delta\, \mathbf{1}$.*

Proof. Since $\psi \geq 0$ and $\int_0^1 \psi\, dm = 1$, there exists an open interval $I \subset [0,1]$, say with length $m(I) \geq \varepsilon > 0$, where $\psi(y) \geq \delta_0$ for all $x \in I$.

Since the diffeomorphism $h : [0,1] \to [0,1]$ has derivative $|h'(x)| \leq 1/\log 2 < 2$, if a finite set $D \subset [0,1]$ is $\frac{\varepsilon}{2}$-dense then $h(D)$ is ε-dense in $[0,1]$. Take $m \in \mathbb{N}$ such that $\frac{1}{m} < \frac{\varepsilon}{4}$ and $n \in \mathbb{N}$ such that $\frac{1}{2^{n-1}} < \frac{\varepsilon}{4}$. Given $a = (a_1, \ldots, a_n) \in \{1, \ldots, m\}^n$, let $[a] = [a_1, \ldots, a_n]$ denote the finite continued fraction with coefficients a_j. The set $D := \{[a] \colon a \in \{1, \ldots, m\}^n\}$ is $\frac{\varepsilon}{4}$-dense in $[0,1]$ (check this fact), and hence $h(D) = \{h([a]) \colon a \in \{1, \ldots, m\}^n\}$ is $\frac{\varepsilon}{2}$-dense in $[0,1]$.

Notice that $h([a]) = g_a(0)$ for each $a \in \{1, \ldots, m\}^n$, and pick $a \in \{1, \ldots, m\}^n$ such that the distance between $h([a])$ and the middle point of I is $< \frac{\varepsilon}{2}$. Since the range $g_a([0,1])$ has length less than $2^{-n} < \frac{\varepsilon}{2}$ it follows that $g_a([0,1]) \subset I$.

Finally, because $\psi \geq 0$, using Proposition 11 we have

$$((U_f^*)^n\psi)(x) \geq |g_a'(x)|\ \psi(g_a(x)) \geq \delta_0 \min_{a\in\{1,\ldots,m\}^n} \min_{t\in[0,1]} |g_a'(t)| > 0$$

for all $x \in [0,1]$. □

Proposition 14. *Given $\psi \in C^1([0,1])$, $\psi \geq 0$,*

$$\lim_{n\to+\infty} (U_f^*)^n(\psi) = \int_0^1 \psi\, dm$$

with uniform convergence.

Proof. $C^1([0,1])$ is dense in the space of Lipschitz functions. Denote by $\mathrm{Lip}(\psi)$ the Lipschitz seminorm of a function $\psi : [0,1] \to \mathbb{R}$

$$\mathrm{Lip}(\psi) := \sup_{x,y\in[0,1]} |\psi(x) - \psi(y)| \,.$$

A function $\psi : [0,1] \to \mathbb{R}$ is Lipschitz iff $\mathrm{Lip}(\psi) < +\infty$. By the mean value theorem, if $\psi : [0,1] \to \mathbb{R}$ is of class C^1 then ψ is Lipschitz with $\mathrm{Lip}(\psi) = \|\psi'\|_\infty$.

Consider now the family of sets

$$\mathcal{C}_a := \{\psi \in C^0([0,1])\colon \psi \geq 0,\ \mathrm{Lip}(\psi) \leq a \int_0^1 \psi\, dm\}$$

with of $a > 0$. Each $\mathcal{C}_a$ is a pointed, closed and convex cone in the Banach space $(C^0([0,1]), \|\cdot\|_\infty)$. We claim that $(U_f^*)^N(\mathcal{C}_a) \subset \mathcal{C}_a^\circ$ for all large enough $a > 0$ and some some $N = N(a) \geq 1$.

From item (g) of Proposition 13, we get for all $\psi \in C^1([0,1]) \cap \mathcal{C}_a$

$$\begin{aligned}
\|(U_f^*)^n(\psi)'\|_\infty &\leq \frac{1+e^2}{2^n}\,\|\psi'\|_\infty + e^4 \int_0^1 |\psi|\; dm \\
&= \frac{1+e^2}{2^n}\,\mathrm{Lip}(\psi) + e^4 \int_0^1 |\psi|\; dm \\
&\leq \left(\frac{(1+e^2)\,a}{2^n} + e^4\right) \int_0^1 |\psi|\; dm \\
&= \left(\frac{(1+e^2)\,a}{2^n} + e^4\right) \int_0^1 \left|(U_f^*)^n\psi\right|\; dm.
\end{aligned}$$

Thus if we take $a > 150$ and $n \geq 6$, given $\psi \in \mathcal{C}_a$ of class C^1

$$\|(U_f^*)^n(\psi)'\|_\infty \leq \left(\frac{1+e^2}{2^n}\,a + e^4\right) \int_0^1 \left|(U_f^*)^n\psi\right|\; dm \leq \frac{a}{2} \int_0^1 \left|(U_f^*)^n\psi\right|\; dm.$$

Because class C^1 functions are dense in the space of Lipschitz functions and the functionals $\psi \mapsto \mathrm{Lip}(\psi)$ and $\psi \mapsto \int_0^1 |\psi|\; d\mu$ are continuous on $C^0([0,1])$, it follows that for any $\psi \in \mathcal{C}_a$,

$$\mathrm{Lip}((U_f^*)^n\psi) \leq \frac{a}{2} \int_0^1 \left|(U_f^*)^n\psi\right|\; dm.$$

This proves that $(U_f^*)^n\psi \in \mathcal{C}_{\frac{a}{2}} \subset \mathcal{C}_a$.

Consider now the hyperplane $H := \{\psi \in C^0([0,1])\colon \int_0^1 \psi\, dm = 1\}$ and the convex set $\mathcal{K}_a := \mathcal{C}_a \cap H$. By Arzelà-Ascoli's theorem, since $\mathrm{Lip}(\psi) \leq a$ for all $\psi \in \mathcal{K}_a$, the closed set

$\mathcal{K}_a$ is compact in $C^0([0,1])$. By Lemma 2, for every $\psi \in \mathcal{K}_a$ there exists $N = N(\psi) \in \mathbb{N}$ and $\delta = \delta(\psi) > 0$ such that $(U_f^*)^N(\psi) \geq 2\,\delta\,\mathbf{1}$. Denoting by $U(\psi)$ the ball of radius $\delta/2$ around ψ in $C^0([0,1])$, we have $(U_f^*)^N(\varphi) \geq \delta\,\mathbf{1}$ for all $\varphi \in U(\psi)$. By property (b) of Proposition 13, we also have $(U_f^*)^n(\varphi) \geq \delta\,\mathbf{1}$ for all $\varphi \in U(\psi)$ and every $n \geq N(\psi)$. By compactness of $\mathcal{K}_a$, there are functions $\psi_1, \ldots, \psi_m \in \mathcal{K}_a$ such that $\mathcal{K}_a \subset U(\psi_1) \cup \ldots \cup U(\psi_m)$. Hence, defining $N = \max_{1 \leq j \leq m} N(\psi_j) \in \mathbb{N}$ and $\delta = \min_{1 \leq j \leq m} \delta(\psi_j) > 0$, we have $(U_f^*)^n(\psi) \geq \delta\,\mathbf{1}$ for all $n \geq N$ and $\psi \in \mathcal{K}_a$. Thus, given $\psi \in \mathcal{C}_a$, $\psi \neq 0$, $(U_f^*)^N(\psi) \geq \delta\,(\int_0^1 \psi\,dm)\,\mathbf{1}$ but also $\mathrm{Lip}((U_f^*)^N\psi) \leq \frac{a}{2}\int_0^1 |\psi|\;dm = \frac{a}{2}\int_0^1 \left|(U_f^*)^N\psi\right|\,dm$, which proves that $(U_f^*)^N\psi \in \mathcal{C}_a^\circ$. Therefore $(U_f^*)^N(\mathcal{C}_a) \subset \mathcal{C}_a^\circ$.

Since $\mathcal{K}_a$ is compact then so is $(U_f^*)^N(\mathcal{K}_a)$. Because this image is contained in $\mathcal{C}_a^\circ \cap H$, by Corollary 7, $(U_f^*)^N(\mathcal{K}_a)$ has finite diameter w.r.t. the Hilbert metric $\theta_{\mathcal{C}_a}$. Hence, by Theorem 1, the bounded linear operator $(U_f^*)^N$ acts as a strict contraction on $\mathcal{K}_a$ w.r.t. $\theta_{\mathcal{C}_a}$. Since $U_f^*\mathbf{1} = \mathbf{1}$ (property (d) of Proposition 13) by Proposition 5,

$$\|(U_f^*)^n(\psi) - \mathbf{1}\|_\infty \lesssim \theta_{\mathcal{C}_a}((U_f^*)^n(\psi), \mathbf{1})$$

converges to 0 at some geometric rate.

To finish the proof, consider any non-zero observable $\psi \in C^1([0,1])$, $\psi \geq 0$. Then $\psi \in \mathcal{C}_a$ for some large enough a. By the previous argument the iterates $(U_f^*)^n(\bar{\psi})$ of the normalized observable $\bar{\psi} = (\int_0^1 \psi\,dm)^{-1}\,\psi$ converge uniformly to the constant function $\mathbf{1}$. Therefore

$$\lim_{n \to +\infty} (U_f^*)^n(\psi) = \int_0^1 \psi\,dm$$

with uniform (and geometric) rate of convergence. $\square$

Proof of Proposition 8. Given $\varphi, \psi \in L^2(X, m)$, and $\epsilon > 0$, consider non-negative functions $\psi_0^+, \psi_0^- \in C^1([0,1])$ such that

$$\|\psi^+ - \psi_0^+\|_2 < \frac{\epsilon}{4\,\|\varphi\|_2} \quad \text{and} \quad \|\psi^- - \psi_0^-\|_2 < \frac{\epsilon}{4\,\|\varphi\|_2}.$$

Notice that

$$\begin{aligned}
|\langle (U_f)^n\varphi,\, \psi\rangle - \langle \varphi, \mathbf{1}\rangle\,\langle \psi, \mathbf{1}\rangle| &= \left|\langle \varphi, (U_f^*)^n\psi\rangle - \langle \varphi, \mathbf{1}\rangle\,\langle \psi, \mathbf{1}\rangle\right| \\
&= \left|\langle \varphi, (U_f^*)^n\psi\rangle - \langle \varphi, \mathbf{1}\rangle\,\langle \psi, \mathbf{1}\rangle\right| \\
&\leq \left|\langle \varphi, (U_f^*)^n\psi^+\rangle - \langle \varphi, \mathbf{1}\rangle\,\langle \psi^+, \mathbf{1}\rangle\right| + \left|\langle \varphi, (U_f^*)^n\psi^-\rangle - \langle \varphi, \mathbf{1}\rangle\,\langle \psi^-, \mathbf{1}\rangle\right|.
\end{aligned}$$

On the first summand we have

$$\begin{aligned}
\left|\langle\varphi,(U_f^*)^n\psi^+\rangle - \langle\varphi,\mathbf{1}\rangle\langle\psi^+,\mathbf{1}\rangle\right| &\le \left|\langle\varphi,(U_f^*)^n\psi^+\rangle - \langle\varphi,(U_f^*)^n\psi_0^+\rangle\right| \\
&\quad + \left|\langle\varphi,(U_f^*)^n\psi_0^+\rangle - \langle\varphi,\mathbf{1}\rangle\langle\psi_0^+,\mathbf{1}\rangle\right| \\
&\quad + \left|\langle\varphi,\mathbf{1}\rangle\langle\psi_0^+,\mathbf{1}\rangle - \langle\varphi,\mathbf{1}\rangle\langle\psi^+,\mathbf{1}\rangle\right| \\
&\le \left|\langle U_f^n\langle\varphi),\psi^+ - \psi_0^+\rangle\right| + \|\varphi\|_1\,\|(U_f^*)^n\psi_0^+ - \langle\psi_0^+,\mathbf{1}\rangle\|_\infty \\
&\qquad + |\langle\varphi,\mathbf{1}\rangle|\,\left|\langle\psi^+ - \psi_0^+,\mathbf{1}\rangle\right| \\
&\le 2\,\|\varphi\|_2\,\|\psi^+ - \psi_0^+\|_2 + \|\varphi\|_1\,\|(U_f^*)^n\psi_0^+ - \int_0^1 \psi_0^+\,dm\|_\infty \\
&\le \frac{\epsilon}{2} + \|\varphi\|_1\,\|(U_f^*)^n\psi_0^+ - \int_0^1 \psi_0^+\,dm\|_\infty.
\end{aligned}$$

By Proposition 14 this summand is $< \epsilon$ for all large enough n. A similar bound shows that $\left|\langle\varphi,(U_f^*)^n(\psi^-)\rangle - \langle\varphi,\mathbf{1}\rangle\langle\psi^-,\mathbf{1}\rangle\right| < \epsilon$ for all sufficiently large n. Together these two bounds prove that (f,m) is mixing. □

References

[1] Garrett Birkhoff, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc. **85** (1957), 219–227. MR 0087058

[2] M Viana, *Stochastic dynamics of deterministic systems*, Publicações Matemáticas, 21° Colóquio Brasileiro de Matemática, IMPA, 1997.

[3] Marcelo Viana and Krerley Oliveira, *Foundations of ergodic theory*, Cambridge Studies in Advanced Mathematics, vol. 151, Cambridge University Press, Cambridge, 2016. MR 3558990