INVARIANT MEASURES AND A WEAK SHADOWING CONDITION

Alfredo Poirier [1]

May, 2012

Abstract

We review the concept of invariant measure and study conditions under which linear combinations of averages along periodic orbits are dense in the space of invariant measures.

MSC(2010): 37L40; 37C50.

Keywords: *Ergodic theory, invariant measures, shadowing of orbits.*

1. *Sección Matemáticas, Departamento de Ciencias, PUCP.*

1. Context

Let X be a compact Hausdorff space. Consider a continuous dynamical system $T : X \to X$. Let $\mathcal{M}(X)$ be the space of real valued measures, that is, of space of all continuous linear functionals from $C(X)$ to $\mathbb{R}$, the dual of $C(X)$.

If ν is a measure, to facilitate manipulation, we write

$$\nu(f) = \int f \, d\nu = \int f(x) \, d\nu(x),$$

where f is continuous (or more generally Borel measurable).

We say that ν is *invariant* if $\nu(f) = \nu(f \circ T)$, or alternatively if we have

$$\int f(x) \, d\nu(x) = \int f(T(x)) \, d\nu(x),$$

for all continuous $f \in C(X)$.

Clearly, the set of invariant measures $\mathcal{M}^T(X)$ is a linear subspace of $\mathcal{M}(X)$.

Example 1.1. If x_0 is a fix-point of T (i.e, a point that satisfies $T(x_0) = x_0$), then δ_{x_0}, the *delta mass located at* x_0, defined by $\delta_{x_0}(f) = f(x_0)$, is an invariant probability measure.

More generally, given a periodic orbit $x_0 \mapsto x_1 \mapsto \ldots \mapsto x_{n-1} \mapsto x_n = x_0$, *the average along the orbit*, given by $\frac{1}{n} \{\delta_{x_0} + \cdots + \delta_{x_{n-1}}\}$, is also an invariant probability measure.

Of course, $\mathcal{M}(X)$ is by design the dual space of $C(X)$, so it has a natural metric and, also, a $*$-weak topology.

Proposition 1.2. *The space $\mathcal{M}^T(X)$ of invariant measures under T is $*$-weak closed (and therefore closed).*

Proof. Let ν_n be a sequence of invariant measures that converge $*$-weakly to ν. Since ν_n annihilates all $f \circ T - f$, with $f \in C(X)$, the $*$-weak limit ν does the same. By definition ν is also invariant. □

Averages along successive points in the same orbit are important. To study them, we introduce some notation. Given f continuous, we set

$$A_n(f) = \frac{f + f \circ T + \cdots + f \circ T^{\circ n-1}}{n}.$$

In particular, with this notation we have

$$A_n(f)(x_0) = \frac{f(x_0) + f(x_1) + \cdots + f(x_{n-1})}{n}.$$

Lemma 1.3. *For $f \in C(X)$ we have $||A_n(f)|| \leq ||f||$.*

Proof. In fact, if $|f(x)| \leq M$ for all $x \in X$, then

$$|A_n(f)(x_0)| = \frac{|f(x_0)| + |f(x_1)| + \cdots + |f(x_{n-1})|}{n} \leq M.$$

The result follows. □

For other properties of these Birkhoff averages we refer the reader to any standard text in ergodic theory like [1], [3], [7].

Excluding trivial cases, the space of invariant measures is big. In fact, a way to produce a huge family of invariant probability measures is shown next.

Fix a non-principal ultrafilter $\mathcal{U}$ in $\mathbb{N}$ (i.e, one not containing finite elements). Recall that this is simply a scheme to make subsequences converge in the sense that if α_n is a sequence in a separable Hausdorff compact space, then there is a unique choice of a value $\lim_{\mathcal{U}} \alpha_n$ among the accumulation points of the sequence $\{\alpha_n\}$ (cf. [4, Theorem 4.3.5]). Recall that by definition $\lim_{\mathcal{U}} \alpha_n = \alpha$ holds if and only if for every neighborhood U of α we have $\{n : \alpha_n \in U\} \in \mathcal{U}$.

The above explains why the two reasonable definitions of the average in the orbit of x_0 along an ultrafilter $\mathcal{U}$ agree. For $f \in C(X)$, let

$$\sigma^f_{\mathcal{U},x_0} = \lim_{\mathcal{U}} A_n(f)(x_0) = \lim_{\mathcal{U}} \frac{f(x_0) + \cdots + f(x_{n-1})}{n},$$

Here $x_0 \to x_1 \to \cdots \to x_n \to \ldots$ denotes the forward orbit of x_0 under iteration. Notice that this limit always exists and satisfies $\sigma^f_{\mathcal{U},x_0} = \sigma^f_{\mathcal{U},T(x_0)}$. Because of this last property, it is not only invariant, but does not depend in the element chosen along the full orbit of x_0.

On the other side, the probability measures $D_n(x_0) = \frac{1}{n}\{\delta_{x_0} + \cdots + \delta_{n-1}\}$ belong to the unit ball of the separable metric space $\mathcal{M}(X)$ and therefore this sequence converges along the ultrafilter $\mathcal{U}$ to a probability measure $\sigma_{\mathcal{U},x_0}$ by the Banach Alaoglu theorem.

Lemma 1.4. *For any continuous f we have $\sigma_{\mathcal{U},x_0}(f) = \sigma^f_{\mathcal{U},x_0}$.*

Proof. Key here is to notice the equality $D_n(x_0)(f) = A_n(f)(x_0)$. After this, the claim follows once we remember that $\sigma^f_{\mathcal{U},x_0} = \alpha$ is satisfied if and only if for all $\epsilon > 0$ the set $\{n : |A_n(f)(x_0) - \alpha| < \epsilon\}$ belongs to $\mathcal{U}$, while $\lim_{\mathcal{U}} D_n(x_0) = \sigma_{\mathcal{U},x_0}$ is equivalent to the fact that $\{n : |A_n(f)(x_0) - \alpha| < \epsilon\} \in \mathcal{U}$ holds for all $f \in C(X)$ and $\epsilon > 0$. □

The measure defined by any of the above quantities is *the average in the orbit of x_0 along the ultrafilter $\mathcal{U}$*. What is import here is that we are working in a bounded set of the dual of $C(X)$ with the $*$-weak topology, which is therefore metrizable (cf. [7, Theorem 6.4]). Since in compact metric spaces accumulation points are limit points, we have the following result.

Lemma 1.5. *There is a subsequence n_k so that $\sigma_{\mathcal{U},x_0} = \lim_{k\to\infty} D_{n_k}(x_0)$ (as a $*$-weak limit). In fact, for any such subsequence, given $f \in C(X)$ we have*

$$\sigma_{\mathcal{U},x_0}(f) = \lim_{k\to\infty} A_{n_k}(x_0)(f).$$

Proof. The first property is due to the fact that $\sigma_{\mathcal{U},x_0}$ is an accumulation point (and hence a limit point) of $D_n(x_0)$. The second is a consequence of Lemma 1.4 and the equality $D_n(x_0)(f) = A_n(f)(x_0)$. □

However, there is no reason to expect that the limit of the averages $A_n(x_0)$ exists. Anyway, it is well known that measures for which such averages are achieved along typical orbits play a crucial role in the study of dynamical systems. The functions for which those limits exist and vanish are easily characterized.

Proposition 1.6. *Fix $x_0 \in X$. A continuous $f \in C(X)$ satisfies $A_n(f)(x_0) \to 0$ if and only if $\sigma_{\mathcal{U},x_0}(f) = 0$ for all non-principal ultrafilters.*

Proof. This is tautological: it is the same as stating that a sequence α_n converges to 0 if and only if it converges to 0 along any non-principal ultrafilter. □

Corollary 1.7. *A continuous $f \in C(X)$ satisfies $A_n(f)(x_0) \to 0$ for all $x_0 \in K$ if and only if $\sigma_{\mathcal{U},x_0}(f) = 0$ for all non-principal ultrafilters and $x_0 \in X$.*

2. Quasicocycles

A *cocycle* is a continuous function of the form $g - g \circ T$. As a measure that annihilates all cocycles is certainly an invariant measure, we can take the reverse approach and define a *quasicocycle* f as continuous function that obeys $\nu(f) = 0$ for all $\nu \in \mathcal{M}^T(K)$. In particular, quasicocycles are invariant under iteration in the sense that f is a quasicocycle if and only $f \circ T$ is a quasicocycle.

Since an invariant measure is by definition a measure that annihilates all cocycles, as a consequence of the Hahn Banach theorem we know that a quasicocycle can be uniformly approximated by cocycles. There are several other characterizations.

Theorem 2.1. *For a continuous function $f \in C(X)$ the following are equivalent:*

a) f is a quasicocycle;

b) for all $\epsilon > 0$ there is a continuous g such that $||f - g + g \circ T|| < \epsilon$;

c) $A_n(f)$ converges uniformly to 0;

d) for any $x_0 \in K$ we have $A_n(f)(x_0) \to 0$.

Proof. We go around the circle of claimed properties.

If f is a quasicocycle, then f belongs to the uniform closure of the linear span of functions of type $g - g \circ T$.

Fix $\epsilon > 0$ and a continuous g such that $||f-g+g\circ T|| < \epsilon/2$. Notice the telescopic relation

$$|A_n(g - g \circ T)(x_0)| = \left| \frac{g(x_0) - g(x_n)}{n} \right| \leq \frac{2||g||}{n} \to 0.$$

Therefore, if N is such that $2||g|| \leq N\epsilon/2$, then for $n \geq N$ we get

$$\begin{aligned} |A_n(f)| = |A_n(f - g + g \circ T) + A_n(g - g \circ T)| \leq \\ |A_n(f - g + g \circ T)| + |A_n(g - g \circ T)| \leq \epsilon/2 + \epsilon/2, \end{aligned}$$

where we have used Lemma 1.3. Uniform convergence follows at once.

Uniform convergence implies pointwise convergence.

Finally, suppose that $A_n(f)(x_0)$ converges to 0 for all x_0. Let ν be an invariant measure. If $\epsilon > 0$, Egoroff's theorem asserts that there is a measurable set E with $|\nu|(E) < \epsilon$, outside which the convergence is uniform: there is N so that $n \geq N$ implies $|A_n(f)(x)| \leq \epsilon$ for $x \notin E$. Also, by invariance we get

$$\int f(x)\, d\nu(x) = \frac{1}{n} \int f(x) + \cdots + f(T^{\circ n-1}(x))\, d\nu(x) = \int A_n(f)(x)\, d\nu(x),$$

and splitting the integration domain into $X - E$ and E we arrive to

$$\int f(x)\,d\nu(x) = \int_{X-E} A_n(f)(x)\,d\nu(x) + \int_E A_n(f)(x)\,d\nu(x).$$

Thus, we get

$$\left| \int f(x)\,d\nu(x) \right| \leq ||\nu||\epsilon + \frac{n+1}{n+1}||f||\,|\nu|(E) \leq ||\nu||\epsilon + ||f||\epsilon.$$

Letting ϵ go to 0 we obtain $\displaystyle\int f\,d\nu = 0$, as desired. □

Theorem 2.2. *Linear combinations of measures of the form $\sigma_{\mathcal{U},x_0}$ are $*$-weak dense in the space of invariant measures.*

Proof. This is a consequence of Corollary 1.7, Theorem 2.1 and the inclusions

$$\bigcap_{\nu \text{ invariant}} Ker(\nu) \subset \bigcap_{\nu \text{ linear combination of } \mathcal{U}-\text{type}} Ker(\nu) \subset \bigcap_{\nu \text{ of } \mathcal{U}-\text{type}} Ker(\nu).$$

□

3. A Shadowing Criterion for Density of Periodic Averages

Recall that the ambient space X is assumed to be a compact metric space. The idea now is to approximate a general average by a periodic one.

For this, fix $n < m$, two positive integers. In the following estimate think of x_0 as paying the price of x_n shadowing y_0.

Lemma 3.1. *Take x_0 and y_0 in X. Write $M = \max_{0 \leq i \leq m-n-1} |f(x_{n+i}) - f(y_i))|$. Then*

$$|A_m(x_0)(f) - A_{m-n}(y_0)(f)| \leq \frac{2n}{m}||f|| + M.$$

Proof. This is a matter of simple bookkeeping. If we write

$$A_m(x_0)(f) = \frac{1}{m}\sum_{i=0}^{n-1} f(x_i) - \frac{n}{m(m-n)}\sum_{i=n}^{m-1} f(x_i) + \frac{1}{m-n}\sum_{i=n}^{m-1} f(x_i),$$

we immediately obtain

$$|A_m(x_0)(f) - A_{m-n}(y_0)(f)| \ \leq$$
$$\frac{n}{m}||f|| + \frac{n}{m}||f|| + \frac{1}{m-n}\sum_{i=0}^{m-n-1} |f(x_{n+i}) - f(y_i))|$$

and the result follows. □

We next introduce an important technical concept.

Start with $x_0 \in X$ and a convergent subsequence x_{n_α} of its forward orbit. We say that x_{n_α} is *weakly shadowable* when for all $\epsilon > 0$ there are $k < l$ and a periodic point p of period $n_l - n_k$ such that

$$d(T^{\circ i}(x_{n_k}), T^{\circ i}(p)) < \epsilon,$$

for $i = 0, \ldots, n_l - n_k - 1$. We insist in two points. First, the shadowing does not necessarily starts at $x_{n_0} = x_0$ but at any x_{n_k}. Second, the shadowing not necessarily goes on for a single approximation to the cluster point of the sequence but for as many as necessary ($l - k$ in the notation).

We say that T *enjoys the weak shadowing property for* x_0 if every convergent subsequence of x_n is weakly shadowable. More generally, T *enjoys the weak shadowing property* if it enjoys this property for all x_0.

Proposition 3.2. *Suppose that T enjoys the weak shadowing property for $x_0 \in X$. Then any measure $\sigma_{\mathcal{U},x_0}$ can be $*$-weakly approximated by averages along periodic orbits.*

Proof. Start with a representing subsequence $n_0 = 0 < n_1 < n_2 < \ldots$ of the convergence of $D_n(x_0)$ to $\sigma_{\mathcal{U},x_0}$ (compare Lemma 1.5). As X is compact, we can suppose x_{n_α} converges. By taking a further subsequence, we can assume that $10^\alpha n_\alpha < n_{\alpha+1}$ holds. Note that this property will still be satisfied even if we keep taking refinements.

Let $k_0 < l_0$ be such that there is a periodic point p_0 of period $n_{l_0} - n_{k_0}$ such that

$$d(T^{\circ i}(x_{n_{k_0}}), T^{\circ i}(p_0)) < 1$$

for $i = 0, \dots, n_{l_0} - n_{k_0} - 1$.

We continue inductively as follows. Momentarily erase $n_1, n_2, \dots n_{l_m}$ from the list and use the shadowing property to get $n_{l_{m+1}} > n_{k_{m+1}}$ (notice that $n_{l_{m+1}}$ is bigger than n_{l_m}, hence also bigger than m by induction) and a periodic point p_{m+1} of period $n_{l_{m+1}} - n_{k_{m+1}}$ such that

$$d(T^{\circ i}(x_{n_{k_{m+1}}}), T^{\circ i}(p_{m+1})) < 1/2^{m+1}$$

for $i = 0, \dots, n_{l_{m+1}} - n_{k_{m+1}} - 1$.

For them we get

$$|A_{n_{l_m}}(x_0)(f) - A_{n_{l_m} - n_{k_m}}(p_m)(f)| \leq \frac{2}{10^m}||f|| + M_m,$$

where $M_m = \max_{0 \leq i \leq n_{l_m} - n_{k_m} - 1} |f(x_{n_{k_m}+i}) - f(T^{\circ i}(p_m)))|$. Here we are applying Lemma 3.1 with the trivial estimate $n_{k_m}/n_{l_m} \leq 1/10^m$.

Finally let us prove that the periodic averages $A_{n_{l_m} - n_{k_m}}(p_m)$ converge $*$-weakly to $\sigma_{\mathcal{U},x_0}$. By Lemma 1.5, for $f \in C(X)$ we have

$$\sigma_{\mathcal{U},x_0}(f) = \lim_{m \to \infty} A_{n_{k_m}}(x_0)(f).$$

Because f is uniformly continuous in X, given $\epsilon > 0$ there is m so that $|x - y| < \frac{1}{2^m}$ implies $|f(x) - f(y)| < \epsilon$. In particular we get $M_m \leq \epsilon$ as $d(T^{\circ i}(x_{n_{k_m}}), T^{\circ i}(p_m)) < 1/2^m$ for $i = 0, \dots, n_{l_m} - n_{k_m} - 1$. Therefore we get $M_m \to 0$ and with this

$$\sigma_{\mathcal{U},x_0}(f) = \lim A_{n_{k_m}}(x_0)(f) = \lim A_{n_{l_m} - n_{k_m}}(p_m)(f),$$

as claimed. □

Theorem 3.3. *If T enjoys the weak showing property, then linear combinations of averages along periodic orbits are $*$-weak dense in $\mathcal{M}^T(X)$.*

Proof. In view of Theorem 2.2, this a corollary to last proposition. $\square$

It is important to notice that the weak shadowing property is hereditary in the sense that it passes through to factor systems.

Proposition 3.4. *Let X and Y be compact metric spaces. Suppose $S : Y \to Y$ is a continuous semiconjugacy of $T : X \to X$ in the sense that there is a continuous surjection $\varphi : X \to Y$ such that the diagram*

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \Big\downarrow{\scriptstyle\varphi} & & \Big\downarrow{\scriptstyle\varphi} \\ Y & \xrightarrow{S} & Y \end{array}$$

commutes. If T enjoys the weak shadowing property, then so does S.

Proof. Fix $y_0 \in Y$ and a convergent subsequence y_{n_k}. As φ is surjective, we can choose x_0 such that $\varphi(x_0) = y_0$. Since X is compact, by taking a refinement we can suppose that x_{n_k} is convergent. But $\varphi : X \to Y$ is uniformly continuous, so given $\epsilon > 0$, there is $\eta > 0$ such that $|x - \hat{x}| < \eta$ implies $|\varphi(x) - \varphi(\hat{x})| < \epsilon$. Now, by the shadowing property of T, there are indices $l > k$ and a periodic p of period $n_l - n_m$ such that $d(T^{\circ i}(x_{n_k}), T^{\circ i}(p)) < \eta$, for $i = 0, \ldots, n_l - n_k - 1$. Now since $\varphi(p)$ has period (dividing) $n_l - n_m$, the result follows. $\square$

4. Examples

Proposition 4.1. *Multiplication by $d \geq 2$ modulo 1 enjoys the weak shadowing property. Hence averages along periodic orbits generate the invariant measures.*

Proof. Suppose $d_{n_k}\theta_0$ converges to $\hat{\theta}$. Fix $\epsilon > 0$ and pick $n_k < n_l$ so that $|d^j\theta_0 - \hat{\theta}|_{S^1} \leq \epsilon/4$ for $j = n_k, n_l$. In the circle we have $|d^{n_k}\theta_0 - d^{n_l}\theta_0|_{S^1} \leq \epsilon/2$ so there is an integer m such that $|m - \theta_0(d^{n_k} - d^{n_l})| \leq \epsilon/2$. Since $2d^{n_k} \leq dd^{n_k} \leq d^{n_k+1} \leq d^{n_l}$ implies $d^{n_l} \leq 2(d^{n_l} - d^{n_k})$, we get

$$\left|\frac{m}{d^{n_l} - d^{n_k}} - \theta_0\right| \leq \frac{\epsilon}{2(d^{n_l} - d^{n_k})} \leq \frac{\epsilon}{d^{n_l}}.$$

Therefore, θ_0 and $\dfrac{m}{d^{n_l} - d^{n_k}}$ stay ϵ-close for at least n_l iterates. However, the n_k-th iterate of $\dfrac{m}{d^{n_l} - d^{n_k}}$, that is, the angle $\dfrac{md^{n_k}}{d^{n_l} - d^{n_k}} = \dfrac{m}{d^{n_l - n_k} - 1}$, is periodic of period $n_l - n_k$. □

Corollary 4.2. *Let P be a degree $d \geq 2$ polynomial whose Julia set J is connected and locally connected. Then P enjoys the weak shadowing property and therefore averages along periodic orbits generate the invariant measures.*

Proof. In fact, in this case the dynamics in the Julia set is semiconjugated to multiplication by d in the unit circle (cf. [5, Theorem 18.3]) □

We provide $\Sigma_n = \{(a_0, a_1, \dots) : a_i \in \{0, \dots, d-1\}\}$, *the space of infinite words on an alphabet of d symbols*, with a distance

$$\mathcal{D}((a_0, a_1, \dots), (b_0, b_1, \dots)) = \sum \frac{\rho(a_i, b_i)}{3^i},$$

where $\rho(a_i, b_i)$ equals 1 if $a_i = b_i$ but equals 0 otherwise. With this metric Σ_n becomes a compact metric space. Here, to claim that two point are close really means that their first, say r, entries agree. Therefore neighborhoods are given by sequences with the same initial blocks. The *shift* $\sigma : \Sigma_n \to \Sigma_n$ is the continuous map given by $\sigma((a_0, a_1, a_2, \dots)) = (a_1, a_2, \dots)$. For those and further properties of the shift we refer the reader to [6].

Proposition 4.3. *The shift σ in the space Σ_n of n symbols enjoys the weak shadowing property.*

Proof. Given $x_0 = (a_0, \dots)$ suppose $x_{n_k} = (a_{n_k}, a_{n_k+1}, \dots)$ converges to $y = (b_0, b_1, \dots)$. For $\epsilon > 0$ let r be such that for two points to be ϵ-close is equivalent for them to agree at least for the first r spots. By refining into a further subsequence we assume that the first r entries of all x_{n_k} agree. Define $p_i = x_j$, where $j = i \mod n_1$, in order to get a period $n_1 = n_1 - n_0$ symbol sequence. By definition p and x_0 agree for the first n_1 symbols. If we apply $\sigma^{\circ n_1}$ to p and x_0 we get p (because of the period) and x_{n_1}, respectively. But both of these words agree in the first r terms with those of x_0. In brief, we have just proved that p and x_0 agree for at least $n_1 + r$ terms. In other words, we have that $\sigma^i(p)$ and $\sigma^i(x_0)$ have in common their first r terms for $i = 0, \dots, n_1 - 1$. □

Corollary 4.4. *Let P be a degree $d \geq 2$ polynomial all of its critical points escape to infinity. Then P restricted to its Julia sets enjoys the weak shadowing property and therefore averages along periodic orbits generate the invariant measures.*

Proof. In fact, that in this case the dynamics is conjugated to the shift in d symbols is classical, and was already known to Fatou and Julia (compare [2, Theorem 2.1]). □

References

[1] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms,* Second Revised Edition, Lecture Notes in Mathematics 470, Springer Verlag 2008.

[2] B. Branner, *Cubic Polynomials: Turning Around the Connectedness locus*, in Topological Methods in Modern Mathematics: A Sympo-

sium in Honor of John Milnor's Sixtieth Birthday, 391-427. Publish or Perish, Inc. 1993.

[3] Cornfeld, Fomin, Sinai, *Ergodic Theory*, Grundlehren der mathematischen Wissenshafften 245, Springer Verlag 1982.

[4] J. Dixmier, *General Topology*, Undergraduate Texts in Mathematics, Springer Verlag 1984.

[5] J. Milnor, *Dynamics in One Complex Variable, Introductory Lectures,* Vieweg, 1999.

[6] C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, Second Edition, CRC Press 1999.

[7] Walters, *An introduction to Ergodic Theory,* Graduate Texts in Mathematics 79, Springer Verlag 1982.

Resumen

Revisamos el concepto de medida invariante y estudiamos condiciones bajo las cuales combinaciones lineales de promedios a lo largo de órbitas periódicas son densas en el espacio de medidas invariantes

Palabras clave: Teoría ergódica, medidas invariantes, persecución de órbitas.

Alfredo Poirier
Departmento de Ciencias,
Sección Matemáticas,
Pontificia Universidad Católica del Perú,
Apartado 1761, Lima 100, Perú
apoirie@pucp.edu.pe