# Notas para un curso de Teoría Ergódica

Héctor Sánchez Morgado

Instituto de Matemáticas, UNAM, Ciudad Universitaria C. P. 04510, Cd. de México, México.

# Índice general

Capítulo 1

# Transformaciones que preservan medida

## 1. Medidas invariantes

Definición 1.1. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida y $T : X \hookleftarrow$. Decimos que $T$ preserva $\mu$ ó que $\mu$ es $T$ -invariante si $\forall A \in \mathcal{A}$, $T^{-1}(A) \in \mathcal{A}$ y $\mu(A) = \mu(T^{-1}A)$.

Ejemplo 1.1. La medida de Lebesgue $\mu_0$ en $\mathbb{R}^n$ induce una medida $\mu$ en la $\sigma$-álgebra de Borel $\mathcal{B}$ del toro

$$\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1, \quad \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

La traslación $t_a(x) = x + a$ en $\mathbb{R}^n$ induce en $\mathbb{T}^n$ la rotación

$$R_\alpha(z_1, \ldots, z_n) = (\alpha_1 z_1, \ldots, \alpha_n z_n)$$

donde $\alpha_j = e^{2\pi i a_j}$. Como $t_a$ preserva $\mu_0$ tenemos que $R_\alpha$ preserva $\mu$.

Ejemplo 1.2. Sea $k \geq 2$ un entero fijo y sea $(p_0, \ldots, p_{k-1})$ un vector de probabilidad, o sea $p_i > 0$ y $p_0 + \cdots + p_{k-1} = 1$. Sean $[k] = \{0, 1, \ldots, k-1\}$ y $X = [k]^{\mathbb{Z}}$. Dados $j \in \mathbb{Z}, i_0, \ldots, i_m \in [k]$ definimos el cilindro $C(j; i_0, \ldots, i_m)$ como

$$C(j; i_0, \ldots, i_m) = \{(x_n) \in X : x_{l+j} = i_l \text{ para } 0 \leq l \leq m\}.$$

Las uniones finitas disjuntas de cilindros forman un álgebra que genera la $\sigma$-álgebra de Borel $\mathcal{A}$ de $X$, por lo tanto, para definir una medida $\mu$ en $\mathcal{A}$ basta definirla en los cilindros, lo que hacemos mediante

$$\mu(C(j; i_0, \ldots, i_m)) = p_{i_0} \ldots p_{i_m}.$$

Definimos el corrimiento bilateral $T : X \hookleftarrow$ mediante $T(x_n) = (y_n)$, donde $y_n = x_{n+1}$. Es claro que $T$ preserva $\mu$.

Una variante de este ejemplo consiste en tomar $X^+ = [k]^{\mathbb{Z}^+}$, donde $\mathbb{Z}^+ = \mathbb{N} \cup \{0\}$. Los cilindros se definen de la misma manera así como el corrimiento que ahora se llama unilateral.

Ejemplo 1.3. Consideremos la dinámica de una bola en una mesa de billar con frontera convexa $\mathcal{C}$. El movimiento esta sujeto a leyes simples: entre rebotes sucesivos la bola viaja en línea recta y en un rebote los ángulos de incidencia y reflexión son iguales. Sea $x$ la longitud de arco con respecto a un punto fijo en la frontera, que supondremos orientada en el sentido antihorario. Sea $\theta$ el ángulo de reflexión en un punto de rebote $C(x) \in \mathcal{C}$ y sea $y = -\cos\theta$. Debido a la convexidad de $\mathcal{C}$, el par $(x, y)$ determina su sucesor $(X, Y)$ y visceversa. Si incrementamos $y$ manteniendo $x$

fijo, la convexidad de $\mathcal{C}$ implica que $C(X)$ se mueve sobre $\mathcal{C}$ en la dirección positiva . O sea $\dfrac{\partial X}{\partial y} > 0$. Sea $S(x,X) = -\|C(X) - C(x)\|$ entonces, ya que $C'(x)$ es un vector tangente unitario,

$$\frac{\partial S}{\partial x} = \frac{C'(x)\cdot(C(X)-C(x))}{S(x,X)} = -y \tag{1.1}$$

$$\frac{\partial S}{\partial X} = \frac{C'(X)\cdot(C(X)-C(x))}{-S(x,X)} = Y \tag{1.2}$$

por lo que

$$Y dX - y dx = dS(x,X) \tag{1.3}$$

y así la transformación $f : (x,y) \mapsto (X,Y)$ preseva area.

Ejemplo 1.4. Definamos ahora la transformación de Gauss $T : (0,1) \hookleftarrow$ mediante

$$T(x) = \begin{cases} \left(\dfrac{1}{x}\right) = \dfrac{1}{x} - \left[\dfrac{1}{x}\right] & \text{si} \quad x > 0 \\ 0 & \text{si} \quad x = 0 \end{cases}.$$

Consideremos la medida $\mu$ en la $\sigma$-álgebra de Borel definida por

$$\mu(A) = \int_A \frac{dx}{1+x}$$

Para ver que $T$ preserva $\mu$ es suficiente probar que si $[a,b] \subset [0,1]$,

$$\mu(T^{-1}([a,b])) = \mu([a,b]) = \log\left(\frac{1+b}{1+a}\right).$$

Ahora bien

$$T^{-1}([a,b]) = \bigcup_{m=1}^{\infty} \left[\frac{1}{b+m}, \frac{1}{a+m}\right],$$

$$\begin{aligned}\mu(T^{-1}([a,b])) &= \sum_{m=1}^{\infty} \log\left(\frac{(a+m+1)(b+m)}{(b+m+1)(a+m)}\right) \\ &= \sum_{m=1}^{\infty} \log\left(\frac{a+m+1}{b+m+1}\right) - \log\left(\frac{a+m}{b+m}\right) \\ &= \lim_{m\to\infty}\left[\log\left(\frac{a+m+1}{b+m+1}\right) - \log\left(\frac{a+1}{b+1}\right)\right] = \log\left(\frac{b+1}{a+1}\right).\end{aligned}$$

Teorema de Krylov y Bogoliubov. *Sean $X$ un espacio métrico compacto, $T : X \hookleftarrow$ continua y $\mathcal{B}$ la $\sigma$-álgebra de Borel. Entonces existe $\mu$ probabilidad $T$-invariante.*

Sea $\mathcal{M}(X)$ el conjunto de las probabilidades definidas en $\mathcal{B}$. Denotaremos por $\mathcal{M}_T(X)$ el conjunto de las $\mu \in \mathcal{M}(X)$, invariantes bajo $T$.

Sea $C(X)$ el espacio de funciones continuas con la norma $\|f\| = \sup\limits_{x\in X} |f(x)|$ para $f \in C(X)$.

Como $X$ es métrico compacto $\exists \{g_n : n \in \mathbb{N}\}$ denso en la bola unitaria de $C(X)$. Consideremos en $\mathcal{M}(X)$ la métrica

$$d(\mu,\nu) = \sum_{j=1}^{\infty} \frac{1}{2^j} \left| \int_X g_j d\mu - \int_X g_j d\nu \right|$$

Lema 1.1. *Sea $\{\mu_n\}$ una sucesión en $\mathcal{M}(X)$. Las siguientes propiedades son equivalentes*

(a) $\lim\limits_{n\to\infty} d(\mu_n,\mu) = 0$

(b) $\lim\limits_{n\to\infty} \int_X g_j d\mu_n = \int_X g_j d\mu \quad \forall j \geq 1$

(c) $\lim\limits_{n\to\infty} \int_X g d\mu_n = \int_X g d\mu \quad \forall g \in C(X)$

Demostración. (a) $\Rightarrow$ (b).

$$\left| \int_X g_j d\mu_n - \int_X g_j d\mu \right| \leq 2^j d(\mu_n,\mu)$$

(b) $\Rightarrow$ (c). Es suficiente considerar $g \in C(X)$ con $\|g\| = 1$. Dado $\varepsilon > 0$, $\exists j$ tal que $\|g_j - g\| \leq \varepsilon/3$. Sea $N$ tal que $n \geq N \Rightarrow \left\| \int_X g_j d\mu_n - \int_X g_j d\mu \right\| \leq \varepsilon/3$. Si $n \geq N$,

$$\left| \int_X g d\mu_n - \int_X g d\mu \right| \leq \left| \int_X (g-g_j) d\mu_n \right| + \left| \int_X g_j d\mu_n - \int_X g_j d\mu \right| + \left| \int_X (g-g_j) d\mu \right| \leq \varepsilon$$

(c) $\Rightarrow$ (a). Sea $j_0$ tal que $\sum\limits_{j=j_0+1}^{\infty} \dfrac{1}{2^j} \leq \dfrac{\varepsilon}{4}$. Sea $N$ tal que $n \geq N \Rightarrow \left| \int_X g_j d\mu_n - \int_X g_j d\mu \right| \leq \dfrac{\varepsilon}{2}$ para $j \leq j_0$. Si $n \geq N$,

$$\begin{aligned} d(\mu_n,\mu) &\leq \sum_{j=1}^{j_0} \frac{1}{2^j} \left| \int_X g_j d\mu_n - \int_X g_j d\mu \right| + \sum_{j=j_{0+1}}^{\infty} \frac{1}{2^j} \left( \int_X |g_j|\, d\mu_n + \int_X |g_j|\, d\mu \right) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon \end{aligned}$$

□

Teorema de representación de Riesz. *Sea $\varphi : C(X) \to \mathbb{R}$ lineal tal que $\varphi(f) \geq 0$ si $f \geq 0$ y $\varphi(1) = 1$. Entonces $\exists \mu \in \mathcal{M}(X)$ tal que*

$$\int_X f d\mu = \varphi(f) \qquad \forall f \in C(X)$$

Consideremos $[-1,1]^{\mathbb{N}}$ con la métrica $\Delta(x,y) = \sum_{i=1}^{\infty} 2^{-j}|x_j - y_j|$, la cual define la topología producto. Definamos $h : \mathcal{M}(X) \to [-1,1]^{\mathbb{N}}$ $\mu \mapsto \hat{\mu}$ mediante $\hat{\mu}(j) = \int_X g_j d\mu$. Claramente $h$ es una isometría.

Proposición 1.1. *$h(\mathcal{M}(X))$ es cerrado*

Demostración. Sea $\{\mu_n\}$ una sucesión en $\mathcal{M}(X)$ tal que $\hat{\mu}_n$ converge. Así $\{\hat{\mu}_n(j)\} = \left\{\int_X g_j d\mu_n\right\}$ converge $\forall j \geq 1$ y así $\left\{\int_X g d\mu_n\right\}$ converge $\forall g \in C(X)$. En efecto, dados $g \in C(X)$ y $\varepsilon > 0$, $\exists j$ tal que $\|g_j - g/\|g\|\| \leq \varepsilon/3\|g\|$ (el caso $g = 0$ es trivial). Como $\left\{\int_X g_j d\mu_n\right\}$ es sucesion de Cauchy, existe $N$ tal que

$$k, n \geq N \Rightarrow \left\| \int_X g_j d\mu_n - \int_X g_j d\mu_k \right\| \leq \frac{\varepsilon}{3\|g\|}.$$

Si $k, n \geq N$,

$$\begin{aligned} \left| \int_X g d\mu_n - \int_X g d\mu_k \right| &\leq \|g\| \left| \int_X \left(\frac{g}{\|g\|} - g_j\right) d\mu_n \right| \\ &\quad + \|g\| \left| \int_X g_j d\mu_n - \int_X g_j d\mu_k \right| + \|g\| \left| \int_X \left(\frac{g}{\|g\|} - g_j\right) d\mu_k \right| \leq \varepsilon. \end{aligned}$$

Definamos $\varphi : C(X) \to \mathbb{R}$ mediante $\varphi(g) = \lim_{m\to\infty} \int_X g d\mu_n$.

Por el Teorema de Riesz existe $\mu \in \mathcal{M}(X)$ tal que $\varphi(g) = \int_X g d\mu$ $\forall g \in C(X)$. Entonces $\hat{\mu}$ es el límite de $\hat{\mu}_n$. □

Corolario 1.1. *$\mathcal{M}(X)$ es compacto.*

Demostración. Como $([-1,1]^{\mathbb{N}}, \Delta)$ es compacto, por la Proposición 1.1, $h(\mathcal{M}(X))$ es compacto. Pero $h$ es una isometria. □

Demostración del teorema de Krylov y Bogoliubov. Escojamos cualquier $\mu \in \mathcal{M}(X)$ por ejemplo la medida de Dirac concentrada en $x_0 \in X$.

$$\mu(A) = \begin{cases} 1 & \text{si} \quad x_0 \in A \\ 0 & \text{si} \quad x_0 \notin A \end{cases}$$

Sea $\mu_n = \frac{1}{n}(\mu + \mu \circ T^{-1} + \cdots + \mu \circ T^{-n+1})$. Por el Corolario 1.1, existe una subsucesión convergente $\{\mu_{n_m}\}$. Sea $\nu = \lim_{m\to\infty} \mu_{n_m}$, entonces

$$\begin{aligned} \nu \circ T^{-1} &= \lim_{m\to\infty} \frac{1}{n_m}(\mu \circ T^{-1} + \cdots + \mu \circ T^{-n_m}) \\ &= \lim_{m\to\infty} \frac{1}{n_m}(\mu + \cdots + \mu \circ T^{-n_m} - \mu) = \nu \end{aligned}$$

□

## 2. Formas de volumen, difeomorfismos, campos vectoriales

DEFINICIÓN 1.2. Una *forma de volumen* en un espacio vectorial $V$ de dimensión $n$ es una función multilineal antisimétrica no nula $\omega : V^n \to \mathbb{R}$

EJEMPLO 1.5. Fijando una base ordenada $(e_1, \dots, e_n)$ de $V$, definimos $\omega(e_1, \dots, e_n) = 1$. La antisimetr'ıa y multilinealidad implican que si $v_1, \dots, v_n \in V$ y escribimos $v_i = \sum_j a_{ij} e_j$, entonces

$$\omega(v_1, \dots, v_n) = \det(a_{ij}).$$

De hecho cualquier forma de volumen se obtiene de esa manera.

Dado un isomorfismo $L : V \to W$ entre espacios $n-$dimensionales y una forma de volumen $\omega$ en $W$ definimos la forma de volumen $L^*\omega$ en $V$ mediante $L^*\omega(v_1, \dots, v_n) = \omega(Lv_1, \dots, Lv_n)$. En el caso de un automorfismo $L : V \hookleftarrow$ tenemos que $L^*\omega = (\det L)\omega$.

DEFINICIÓN 1.3. Sean $M$ una variedad diferenciable y $\varphi : U \subset \mathbb{R}^n \to M$ una parametrización. Para $x \in U$, $T_{\varphi(x)}M = D\varphi_x(\mathbb{R}^n)$ es el espacio tangente a $M$ en $\varphi(x)$. Una forma de volumen $\omega$ en $\varphi(U)$ asocia, por definición, a cada $q \in \varphi(U)$ una forma de volumen $\omega_q$ en $T_qM$. Definiendo $(\varphi^*\omega)_x = D\varphi_x^*\, \omega_{\varphi(x)}$, tenemos que $\varphi^*\omega$ es una forma de volumen en $U$. Decimos que $\omega$ es suave si $\varphi^*\omega$ lo es.

Para $\omega$ suave y $A \subset M$ boreliano tal que $A \subset \varphi(U)$ definimos

$$\mu_\omega^\varphi(A) = \int_{\varphi^{-1}(A)} |\varphi^*\omega(e_1, \dots, e_n)|$$

donde $e_1, \dots, e_n$ es la base canónica de $\mathbb{R}^n$.

Si $\omega$ es una forma suave de volumen en $M$ y $\varphi, \psi$ son dos parametrizaciones tales que $A \subset \operatorname{Im}\varphi, \operatorname{Im}\psi$, el teorema del cambio de variable da

$$\begin{aligned}
\mu_\omega^\varphi(A) &= \int_{\varphi^{-1}(A)} |(\psi^{-1} \circ \varphi)^*\psi^*\omega(e_1, \dots, e_n)| \\
&= \int_{\varphi^{-1}(A)} |\det D(\psi^{-1} \circ \varphi)||\psi^*\omega(e_1, \dots, e_n)| \\
&= \int_{\psi^{-1}(A)} |\psi^*\omega(e_1, \dots, e_n)| = \mu_\omega^\psi(A).
\end{aligned}$$

Dado cualquier boreliano $A \subset M$ escogemos parametrizaciones $\{\varphi_j\}$ tales que $A \subset \bigcup_j \operatorname{Im}\varphi_j$. Sean $\{A_j \subset \operatorname{Im}\varphi_j\}$ borelianos disjuntos con $A = \bigcup_j A_j$. Definimos

$$\mu_\omega(A) = \sum_j \mu_\omega^{\varphi_j}(A_j)$$

Sean $f : M \to N$ transformación diferenciable entre variedades diferenciables de la misma dimensión y $\omega$ una forma suave de volumen en $N$.

Sea $\varphi : U \subset \mathbb{R}^n \to M$ una parametrización. Supongamos que $f|\varphi(U)$ es un difeomorfismo sobre su imagen y por lo tanto $f \circ \varphi(U)$ es una parametrización. Sea $A \subset f(\varphi(U))$, entonces

$$\begin{aligned}\mu_\omega(A) &= \int_{(f\varphi)^{-1}(A)} |(f\varphi)^*\omega(e_1, \dots, e_n)| \\ &= \int_{\varphi^{-1}f^{-1}(A)} |\varphi^* f^*\omega(e_1, \dots, e_n)| = \mu_{f^*\omega}(f^{-1}(A) \cap \varphi(U))\end{aligned}$$

Recordemos que $x \in M$ se llama *punto crítico de* $f$ si la derivada $D_x f : T_x M \to T_y N$ es singular. Un punto $y \in N$ es un *valor regular de* $f$ si ningún $x \in f^{-1}(y)$ es un punto crítico de $f$. En tal caso para cada $x \in f^{-1}(y)$ hay una vecindad $W_x$ tal que $f|W_x$ es un difeomorfismo sobre su imagen. Suponiendo además que $M$ es compacta se tiene que $f^{-1}(y)$ es finito. Sea $f^{-1}(y) = \{x_1, \dots x_k\}$ y escojamos $W_1, \dots W_k$ vecindades coordenadas disjuntas que se transforman difeomorfamente en vecindades $V_1, \dots V_k$ de $y$. Sean

$$V = V_1 \cap \dots \cap V_k \setminus f(M \setminus W_1 \dots \setminus W_k), \quad U_i = W_i \cap f^{-1}(V),$$

entonces $f^{-1}(V) = U_1 \cup \dots \cup U_k$. Para $A \subset V$ tenemos

$$\mu_{f^*\omega}(f^{-1}(A)) = \sum_{i=1}^{k} \mu_{f^*\omega}(f^{-1}(A) \cap U_i) = k\mu_\omega(A) \tag{1.4}$$

Proposición 1.2. *Sean $f : M \to N$ transformacion diferenciable entre variedades de la misma dimensión, $M$ compacta $N$ conexa. Sea $\omega$ una forma suave de volumen en $N$. Supongamos que todos los valores de $f$ son regulares. Entonces $\#f^{-1}(y)$ no depende de $y \in N$ y lo denotamos por* grado$(f)$, *$f^*\omega$ es una forma de volumen en $M$ con*

$$\mu_{f^*\omega}(M) = \text{grado}(f)\mu_\omega(N)$$

Demostración. Según la discusión previa, $f(M)$ es abierta y de hecho el número de preimagenes de los elementos de $N$ es localmente costante. Como $N$ es conexa, tenemos que $f(M) = N$ y $\#f^{-1}(y) = k \ \forall y \in N$. Descomponiendo $N$ en la unión disjunta de conuntos $A_j$ a los que podemos aplicar (1.4) tenemos

$$\mu_{f^*\omega}(M) = \sum_j \mu_{f^*\omega}(f^{-1}(A_j)) = k \sum_j \mu_\omega(A_j) = k\mu_\omega(N)$$

□

Ejemplo 1.6. Consideremos el toro

$$\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$$

con parametrizaciones $\exp : U \to \mathbb{T}^n$ donde

$$\exp(x) = x + \mathbb{Z}^n = (e^{2\pi x_1}, \dots, e^{2\pi x_n}).$$

Para cada $p = \exp(x)$, $D\exp_x : \mathbb{R}^n \to T_x\mathbb{T}^n$ define una identificación canónica ya que no depende del punto $x$ elegido puesto que si $t_k$ representa la traslación por el entero $k$

$$D\exp_x = D\exp_{x+k} \circ Dt_k = D\exp_{x+k}.$$

Así, podemos definir una forma de volumen $\omega$ en $\mathbb{T}^n$ mediante

$$\omega_{\exp(x)}(D\exp_x v_1, \dots, D\exp_x v_n) = \det(v_1, \dots, v_n)$$

o sea que $\exp^* \omega = \det$ y entonces $\mu_\omega$ es la medida del ejemplo 1.1. Sea $A$ una matriz n×n $A$ con coeficientes enteros, y consideremos la transformación $\tilde{A} : \mathbb{T}^n \hookleftarrow$ definida por $\tilde{A}(x + \mathbb{Z}^n) = Ax + \mathbb{Z}^n$. Si $\det A \neq 0$, entonces $\tilde{A}$ no tiene puntos críticos y podemos aplicar la Proposición 1.2. Por otra parte para $B \subset U$

$$\begin{aligned} \mu_{\tilde{A}^*\omega}(\exp(B)) &= \int_B |(\exp^* \tilde{A}^*\omega)(e_1, \dots, e_n)| = \int_B |(A^* \exp^* \omega)|(e_1, \dots, e_n)| \\ &= |\det A| \int_B 1 = |\det A|\, \mu_\omega(\exp(B)) \end{aligned}$$

O sea que $\mu_{\tilde{A}^*\omega} = |\det A|\, \mu_\omega$ y el número de preimagenes bajo $\tilde{A}$ de cualquier punto es $\det A$. De estas observaciones y la equación (1.4) tenemos

$$|\det A|\mu_\omega(\tilde{A}^{-1}(B)) = \mu_{\tilde{A}^*\omega}(\tilde{A}^{-1}(B)) = k\mu_\omega(B)$$

y entonces $\tilde{A}$ preserva $\mu_\omega$

Ejemplo 1.7. De acuerdo al Ejemplo 1.6 la transformación

$$\tilde{A} : \mathbb{S}^1 \to \mathbb{S}^1, \quad \tilde{A}(z) = z^m,$$

preserva la medida $\mu$ inducida por la medida de Lebesgue en $\mathbb{R}$. Defina $h : \mathbb{S}^1 \to [-1, 1]$ mediante $h(z) = \Re z$ . Sea $z = x + iy \in \mathbb{S}^1$, entonces

$$h(\tilde{A}(z)) = h(z^2) = x^2 - y^2 = 2x^2 - 1 = T(h(z)).$$

Defina $M(B) = \mu(h^{-1}(B))$ para $B \subset [-1, 1]$ boreliano.

$$M(T^{-1}(B)) = \mu(h^{-1}T^{-1}(B)) = \mu(\tilde{A}^{-1}h^{-1}(B)) = \mu(h^{-1}(B)) = M(B)$$

Explícitamente, $M(B)$ es la medida del conjunto en $\mathbb{S}^1$ que se proyecta en $[-1, 1]$

$$M(B) = \frac{1}{\pi} \int_B \frac{dx}{\sqrt{1 - x^2}}.$$

Ejemplo 1.8. Sea $\varphi : [0, 1] \hookleftarrow$ tal que existen intervalos abiertos disjuntos $I_n \subset [0, 1]$, $n \in \mathbb{N}$ satisfaciendo

- $\lambda([0, 1] \setminus \cup_n I_n) = 0$
- $\lambda(\varphi([0, 1] \setminus \cup_n I_n)) = 0$
- $\forall n\ \varphi \mid I_n$ es continuamente diferenciable con $\varphi' \neq 0$

Sean $f \in L^1([0,1])$ y $\mu$ la medida en los borelianos de $[0,1]$ definida por $\mu(A) = \int_A f(x)dx$.

Definamos $\varphi_n = \varphi|I_n$, $A_n = \varphi_n^{-1}(A)$.

Por el teorema del cambio de variable

$$\mu(\varphi^{-1}(A)) = \sum_n \int A_n f(y)dy = \sum_n \int A_n \frac{f(\varphi_n^{-1}(x))dx}{|\varphi_n'(\varphi_n^{-1}(x))|} = \int_A \sum_{y\in\varphi^{-1}(x)} \frac{f(y)}{|\varphi'(y)|}dx$$

Por lo que $\varphi$ preserva $\mu$ si y solo si para casi todo $x \in [0,1]$ se tiene

$$f(x) = \sum_{y\in\varphi^{-1}(x)} \frac{f(y)}{|\varphi'(y)|}$$

Apliquemos este criterio a la transformacion de Gauss $\varphi$ y $f(x) = \dfrac{1}{1+x}$. Tenemos $\varphi_n(y) = \dfrac{1}{y} - n$, $\varphi_n'(y) = \dfrac{1}{y^2}$, $\varphi_n^{-1}(x) = \dfrac{1}{x+n}$.

$$\begin{aligned}\sum_{y\in\varphi^{-1}(x)} \frac{f(y)}{|\varphi'(y)|} &= \sum_{y\in\varphi^{-1}(x)} \frac{y^2}{1+y} = \sum_{n=1}^{\infty} \frac{1}{(x+n)(x+n+1)} \\ &= \sum_{n=1}^{\infty} \frac{1}{x+n} - \frac{1}{x+n+1} = \frac{1}{1+x} = f(x)\end{aligned}$$

Si $f : M \hookleftarrow$ es un difeomorfismo $C^1$ y $\omega$ es una forma de volumén, entonces $\mu_\omega \circ f = \mu_{f^*\omega}$. Por lo tanto $f$ preseva $\mu_\omega$ si y solo si $f^*\omega = \omega$. Si $M$ es abierto de $\mathbb{R}^n$ $f^*\omega = (\det Df)\,\omega \circ f$.

Sea $X$ un campo vectorial $C^1$ en $M$ y sea $\{\varphi_t\}$ el flujo definido por $X$. Sea $\omega$ una forma de volumen en $M$, definimos la derivada de Lie

$$L_X\omega = \frac{d}{dt}\left(\varphi_t^*\omega\right)_{t=0}.$$

Ejemplo 1.9. Sean $M$ abierto de $\mathbb{R}^n$ y $\varphi_t^*\omega = (\det D\varphi_t)\omega \circ \varphi_t$. Definiendo $\rho(x) = \omega_x(e_1,\ldots,e_n)$, tenemos

$$L_X\omega(e_1,\ldots,e_n) = \frac{d}{dt}\left(\det D\varphi_t\right)_{t=0}\rho + \frac{d}{dt}\left(\rho\circ\varphi_t\right)_{t=0}.$$

Derivando la ecuación

$$\frac{d}{dt}\varphi_t(x) = X(\varphi_t(x))$$

tenemos la ecuación de la primera variación

$$\frac{d}{dt}D\varphi_t(x) = DX(\varphi_t(x))D\varphi_t(x).$$

Como $D\varphi_0 = I$, para $t \to 0$

$$D\varphi_t(x) = I + tDX(x) + O(t^2),$$

$$\det D\varphi_t = 1 + t \operatorname{div} X(x) + O(t^2),$$

y por lo tanto

$$L_X\omega = (\operatorname{div} X)\omega + (D\rho \cdot X)\frac{\omega}{\rho} = \frac{\operatorname{div}(\rho X)}{\rho}\omega.$$

En general definimos $\operatorname{div}_\omega X$ mediante

$$L_X\omega = (\operatorname{div}_\omega X)\omega.$$

Así, $\{\varphi_t\}$ preserva $\omega$ si y solo si $\operatorname{div}_\omega X = 0$.

PROPOSICIÓN 1.3. *Supongamos que $f : M \hookleftarrow$ preserva $\omega$ y $H : M \to \mathbb{R}$ es una primera integral de $f$, o sea $H \circ f = H$ y $H$ no es constante. Sea $C$ un valor regular de $H$, entonces $\exists\ \tilde{\omega}$ forma de volumen en $H^{-1}(C)$ preservada por $f \mid H^{-1}(C)$.*

DEMOSTRACIÓN. Sean $x \in H^{-1}(C)$, $\quad T_xH^{-1}(C) = \ker D_xH$. Sea $v$ tal que $D_xH(v) = 1$, definimos $\tilde{\omega}_x(v_2, \ldots, v_n) = \omega_x(v, v_2, \ldots, v_n)$, luego $(f^*\tilde{\omega})_x(v_2, \ldots, v_n) = \tilde{\omega}_{f(x)}(D_xfv_2, \ldots, D_xfv_n)$.

Pero $D_{f(x)}H \circ D_xf(v) = D_xH(v) = 1$ y así

$$\begin{aligned}
\tilde{\omega}_{f(x)}(D_xfv_2, \ldots, D_xfv_n) &= \omega_{f(x)}(D_xfv, D_xfv_2, \ldots, D_xfv_n)\\
&= (f^*\omega)_x(v, v_2, \ldots, v_n) = \omega_x(v, v_2, \ldots, v_n)\\
&= \tilde{\omega}_x(v_2, \ldots, v_n).
\end{aligned}$$

Así, $f^*\tilde{\omega} = \tilde{\omega}$ □

EJEMPLO 1.10. Sean $M$ abierto de $\mathbb{R}^{2n}$ y $H : M \to \mathbb{R}$ de clase $C^2$. Consideremos el campo vectorial hamiltoniano

$$X_H = \left(\frac{\partial H}{\partial y_1}, \ldots, \frac{\partial H}{\partial y_n}, -\frac{\partial H}{\partial x_1}, \ldots, -\frac{\partial H}{\partial x_n}\right).$$

$$\operatorname{div} X_H = \sum_i^n \frac{\partial^2 H}{\partial x_i \partial y_i} - \frac{\partial^2 H}{\partial y_i \partial x_i} = 0.$$

Por lo tanto el flujo hamiltoniano $\{\varphi_t\}$ preserva el volumen usual de $\mathbb{R}^{2n}$. Por otra parte

$$\frac{d}{dt}H \circ \varphi_t = DH(X_H) = \sum_i^n \frac{\partial H}{\partial x_i}\frac{\partial H}{\partial y_i} - \frac{\partial H}{\partial y_i}\frac{\partial H}{\partial X_i} = 0$$

Así, $H \circ \varphi_t = H$ para toda $t$. Luego, si $C$ es un valor regular de $H$, hay una forma de volumen $\omega_c$ en $H^{-1}(C)$, preservada por $\{\varphi_t\}$.

## 3. Recurrencia

Teorema de recurrencia de Poincaré. *Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad y $T : X \hookleftarrow$ preservadora de medida. Para cualquier $B \in \mathcal{A}$ el conjunto $B_0 = \{b \in B : \{n \in \mathbb{N} : T^n(b) \in B\}$ es infinito$\}$ pertenece a $\mathcal{A}$ y $\mu(B_0) = \mu(B)$.*

Demostración. Sea $C_n = \left\{b \in B : T^k(b) \notin B \ \ \forall k \geq n\right\}$. Entonces

$$C_n = B \setminus \bigcup_{k \geq n} T^{-k}(B) \in \mathcal{A}, \quad B_0 = B \setminus \bigcup_n C_n.$$

El teorema estará demostrado si probamos que $\mu(C_n) = 0$.

Como $C_n \subset \bigcup\limits_{k \geq 0} T^{-k}B \setminus \bigcup\limits_{k \geq n} T^{-k}B$, $\mu(C_n) \leq \mu(\bigcup\limits_{k \geq 0} T^{-k}B) - \mu(\bigcup\limits_{k \geq n} T^{-k}B)$.

Pero $\bigcup\limits_{k \geq n} T^{-k}B = T^{-n} \bigcup\limits_{k \geq 0} T^{-k}B$ y así $\mu(\bigcup\limits_{k \geq 0} T^{-k}B) = \mu(\bigcup\limits_{k \geq n} T^{-k}B)$. □

Definición 1.4. Sean $X$ espacio topológico y $T : X \hookleftarrow$ continua. El $\omega-$límite de un punto $x \in X$ es el conjunto

$$\omega(x) = \{y \in X : \forall U \text{ vecindad de } y, \exists n_k \to \infty \text{ con } T^{n_k}(x) \in U\}.$$

Si $X$ es métrico, $y \in \omega(x) \iff \exists n_k \to \infty$ tal que $d(T^{n_k}(x), y) \to 0$.

Teorema 1.2. *Sean $X$ espacio métrico separable, $\mathcal{A}$ la $\sigma-$álgebra de Borel, $T : X \hookleftarrow$ continua y $\mu \in M_T(X)$. Entonces*

$$\mu(\{x \in X : x \notin \omega(x)\}) = 0.$$

*Es decir, $\mu$-casi todo punto de $X$ es recurrente.*

Demostración. Sea $\{U_n\}_n \in \mathbb{N}$ una base de abiertos Si $x \notin \omega(x)$ hay una vecindad $U$ de $x$ tal que $\{k \in \mathbb{N} : T^k(x) \in U\}$ es finito y por lo tanto $\exists U_n$ tal que $x \in U_n$ y $\forall k \in \mathbb{N}$ $T^k(x) \notin U_n$. Reciprocamente, si el punto $x$ satisface esta propiedad entonces $x \notin \omega(x)$. Así

$$\{x \in X : x \notin \omega(x)\} = \bigcup_n (U_n \setminus \bigcup_{k \in \mathbb{N}} T^{-k}(U_n))$$

Pero en el Teorema de recurrencia de Poincaré demostramos que

$$\mu(U_n \setminus \bigcup_{k \in \mathbb{N}} T^{-k}(U_n)) = 0.$$

□

Definición 1.5. Sean $X$ espacio topológico y $T : X \hookleftarrow$ continua.

(a) $T$ se llama transitiva si $\exists x \in X$ tal que $\omega(x) = X$

(b) $T$ se llama topológicamente mezcladora si $\forall$ $U, V$ abiertos no vacios, $\exists N > 0$ talque $T^{-n}(U) \cap V \neq \phi$ $\forall n \geq N$.

EJEMPLO 1.11. La rotación $R : \mathbb{S}^1 \to \mathbb{S}^1$, $R(z) = e^{i2\pi a}z$ es transitiva si y solo si $a \notin \mathbb{Q}$ (ver Proposición 2.2).

Es sencillo demostrar que una rotación nunca es mezcladora.

EJEMPLO 1.12. La transformación $\tilde{A}(z) = z^m$ del ejemplo 1.7 es mezcladora. En efecto, para $0 < a < b < 1$

$$T^{-k}\exp(a,b) = \bigcup_{j=0}^{m^k-1} \exp(\frac{a+j}{m^k}, \frac{b+j}{m^k}).$$

Cualquier abierto $V \subset \mathbb{S}^1$ contiene un segmento $\exp(jm^{-k}, (j+1)m^{-k})$

$$\exp\Big(\frac{a+jm^n}{m^{k+n}}, \frac{b+jm^n}{m^{k+n}}\Big) \subset \exp\Big(\frac{j}{m^n}, \frac{j+1}{m^n}\Big).$$

PROPOSICIÓN 1.4. *Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua, son equivalentes*

(a) *$T$ es transitiva.*
(b) *$\{x \in X : \omega(x) = X\}$ es intersección numerable de abiertos densos.*
(c) *$\forall U$ abierto no vacio $\bigcup_{n\geq 0} T^{-n}U$ es denso en $X$.*

DEMOSTRACIÓN. (a) $\Rightarrow$ (c). Sea $x \in X \ni \omega(x) = X$. Sea $U$ abierto no vacio y sea $y \in X$. Sea $V$ vecindad de $y$, entonces $\exists n > 0$ tal que $T^n(x) \in V$. Sea $m > n$ tal que $T^m(x) \in U$. $T^m(x) = T^{m-n}(T^n(x))$ y así $T^n(x) \in T^{-(m-n)}U \cap V$. Luego $\bigcup_{j\geq 0} T^{-j}U \cap V \neq \emptyset$. Como $V$ es arbitrario $y \in \bigcup_{j\geq 0} T^{-j}U$.

(c) $\Rightarrow$ (b). Sea $\{U_n\}_{n\in\mathbb{N}}$ base para la togopolog'ıa de $X$. Por hipótesis

$$S_{i,n} \bigcup_{=} j \geq iT^{-j}U_n = \bigcup_{j\geq 0} T^{-j}(T^{-i}U_n)$$

es abierto denso. Así, $S = \bigcap_{i,n} S_{i,n}$ es intersección numerable de abiertos densos. Si $x \in S$ entonces $\omega(x) = X$. En efecto dado $U \neq \emptyset$ abierto debemos probar que $\exists m_k \to \infty$ tal que $T^{m_k}(x) \in U$. Es suficiente hacerlo para los abiertos $U_n, n \in \mathbb{N}$. Como $x \in S$, $\forall i, n$ $x \in S_{i,n}$ y así existe $m_i \geq i$ tal que $T^{m_i}(x) \in U_n$.

(b) $\Rightarrow$ (a) es trivial. □

DEFINICIÓN 1.6. Sea $X = [k]^{\mathbb{Z}}$. Sea $A = (A_{ij})_{i,j\in[k]}$ matriz de ceros y unos. Sea $X_A = \big\{(x_n) \in X : A_{x_n x_{n+1}} = 1\, \forall n \in \mathbb{Z}\big\}$. Sea $\sigma$ el corrimiento, entonces $\sigma(X_A) = X_A$. $\sigma|X_A$ se llama *subcorrimiento de tipo finito.*

PROPOSICIÓN 1.5. *Para $m \geq 0$ denotemos por $A^m_{ij}$ los coeficientes de $A^m$. Entonces*

1. *$\sigma|X_A$ es transitivo $\iff$ $\forall i, j \in [k]$ $\exists m > 0$ tal que $A^m_{ij} > 0$.*

*2.* $\sigma|X_A$ *es topológicamente mezcladora* $\iff \forall i,j \in [k]\ \exists m > 0$ *tal que* $\forall l \geq m$ $A^l_{ij} > 0$.

Usaremos el siguiente lema que se demuestra facilmente por inducción.

LEMA 1.2. *Sea*

$$S^m_{ij} = \left\{a \in [k]^{[m+1]} : a_0 = i, a_m = j, A_{a_n a_{n+1}} = 1\, \forall n \in [m]\right\}.$$

*Entonces* $A^m_{ij} = \# S^m_{ij}$.

COROLARIO 1.3. $\#\{x \in X_A : \sigma^m x = x\} = Tr\, A^m$.

DEMOSTRACIÓN. En efecto $x \in X_A$ y $\sigma^m x = x$ si y solo si $x_{nm+j} = x_j$ $\forall n \in \mathbb{Z}, j \in [m]$ y $A_{x_0x_1} = A_{x_1x_2} = \ldots = A_{x_{m-1}x_0} = 1$. Así

$$\#\{x \in X_A : \sigma^m x = x\} = \sum_{i=0}^{k-1} \# S^m_{ii} = \sum_{i=0}^{k-1} A^m_{ii} = \mathrm{Tr}\, A^m.$$

$\square$

DEMOSTRACIÓN. de la Proposición 1.5. Sean $U, V$ abiertos no vac'ıos de $X_A$. Como los cilindros forman una base, $U$ contiene un abierto de la forma $C(j; i_0, \ldots, i_m) \cap X_A \neq \emptyset$, $V$ contiene un abierto de la forma $C(p; q_0, \ldots, q_l) \cap X_A \neq \emptyset$. Entonces

$$\begin{aligned} \sigma^{-n}(U) \cap V &\supset \sigma^{-n}(C(j; i_0, \ldots, i_m)) \cap C(p; q_0, \ldots, q_l) \cap X_A \\ &= C(j+n; i_0, \ldots, i_m) \cap C(p; q_0, \ldots, q_l) \cap X_A. \end{aligned}$$

Si $p+l < j+n$ este último conjunto no es vacío cuando $A^{j+n-p-l}_{q_l i_o} > 0$, ya que en tal caso $\exists a \in S^{j+n-p-l}_{q_l i_o}$ y tomando $x \in C(j; i_0, \ldots, i_n) \cap X_A$, $y \in C(p; q_0, \ldots, q_l) \cap X_A$ definimos

$$z_t = \begin{cases} y_t & t \leq p+l \\ a_{t-(p+l)} & p+l \leq t \leq j+n \\ x_{t-n} & j+n \leq t \end{cases}$$

Reciprocamente si $z \in \sigma^{-m}(C(0; j)) \cap C(0; i) \cap X_A$ entonces $z_0 = i, z_m = j$, $A_{z_n z_{n+1}} = 1\ \forall n \in [m]$. $\square$

## 4. Medidas de Markov

DEFINICIÓN 1.7. Una matriz *estocástica* $k \times k$ es una matriz $\mathbb{P} = (P_{ij})_{i,j\in[k]}$ tal que $P_{ij} \geq 0\ \forall i,j \in [k]$ y $\sum_{j\in[k]} P_{ij} = 1\ \forall i \in [k]$. Un vector $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_{k-1})$ tal que $\sum i \in [k]\pi_i = 1$ y $\pi_i > 0\ \forall_i$, se llama *vector de probabilidad* asociado a $\mathbb{P}$ si $\boldsymbol{\pi}\mathbb{P} = \boldsymbol{\pi}$.

TEOREMA 1.4. *Sean* $X = [k]^{\mathbb{Z}}$ *y* $T$ *el corrimiento. Sea* $\mathbb{P}$ *una matriz estocástica* $k \times k$ *con vector de probabilidad asociado* $\boldsymbol{\pi}$. *Entonces existe una única* $\mu \in \mathcal{M}_T(X)$ *llamada medida de Markov tal que*

(a) $\mu(C(m; i)) = \pi_i\ \forall m \in \mathbb{Z}, i \in [k]$

(b) $\mu(C(m;i) \cap C(m+1;j)) = \pi_i P_{ij}\ \forall m \in \mathbb{Z}, i, j \in [k]$
(c) $\mu(C(n;i_n) \mid C(0;i_0,\ldots,i_{n-1})) = \mu(C(n;i_n) \mid C(n-1;i_{n-1}))$

Demostración. Sea $\mathcal{A}_0$ el álgebra de las uniones finitas disjuntas de cilindros. Definimos $\mu : \mathcal{A}_0 \to \mathbb{R}$ mediante

$$\mu(C(m;i_0,\ldots,i_l)) = \pi_{i_0} P_{i_0 i_1} \ldots, P_{i_{l-1} i_l}$$

y si $A = A, \cup \cdots \cup A_r$ con $A_1, \ldots, A_r$ cilindros disjuntos definimos

$$\mu(A) = \mu(A) + \cdots + \mu(A_r).$$

$\mu$ es una medida en $\mathcal{A}_0$ que se extiende a la $\sigma$- álgebra de Borel de X. Es claro que $\mu(T^{-1}A) = \mu(A)$ si $A \in \mathcal{A}_0$ y así la extensión es invariante bajo $T$. □

Teorema de Perron y Frobenius. *Sea $\mathbb{P}$ una matriz $k \times k$ tal que alguna potencia $\mathbb{P}^N$ tiene todos sus coeficientes positivos. Entonces $\mathbb{P}$ posee un eigenvalor "dominante" $\lambda$ tal que*

(a) $\lambda > |\mu|$ *para todo eigenvalor* $\mu \neq \lambda$ *de* $\mathbb{P}$.
(b) $\dim \ker(\mathbb{P} - \lambda\mathbb{I}) = 1$.
(c) *Existe* $\mathbf{q} = (q_0, \ldots, q_{k-1})$ *tal que* $\mathbf{q}\mathbb{P} = \lambda\mathbf{q}$ *y* $q_i > 0\ \forall i$.

Corolario 1.5. *Sea $\mathbb{P}$ una matriz estocástica que satisface la hipótesis del teorema de Perron y Frobenius. Entonces*

(a) *El eigenvalor dominante es 1.*
(b) *Existe vector de probabilidad* $\boldsymbol{\pi}$ *asociado a* $\mathbb{P}$.
(c) *Si* $\mathbf{x} = (x_0, \ldots, x_{k-1})$ *satisface* $\sum\limits_{i \in [k]} x_i = 1$ *entonces*

$$\lim_{n\to\infty} \mathbf{x}\mathbb{P}^n = \boldsymbol{\pi}$$

Demostración. Sea $\lambda$ el eigenvalor dominante. Sea

$$H_a = \{\mathbf{x} = (x_0, \ldots, x_{k-1}) : \sum_i x_i = a\}.$$

Como

$$\sum_{i,j} x_i P_{ij} = \sum_i x_i \sum_j P_{ij} = \sum_i x_i \quad \forall \mathbf{x},$$

se tiene que $H_a\mathbb{P} \subset H_a$.

Por (b) y (c) del teorema se tiene que existe un único $\boldsymbol{\pi} \in H_1$ tal que $\boldsymbol{\pi}\mathbb{P} = \lambda\boldsymbol{\pi}$. Entonces $= \lambda\boldsymbol{\pi} \in H_1$ y así $\lambda = 1$.

Si $\mathbf{x} \in H_1$ entonces $\mathbf{y} = \mathbf{x} - \boldsymbol{\pi} \in H_0$. Así,

$$\mathbf{x}\mathbb{P}^n = \boldsymbol{\pi}\mathbb{P}^n + \mathbf{y}\mathbb{P}^n = \boldsymbol{\pi} + \mathbf{y}\mathbb{P}^n$$

Como $H_0\mathbb{P}^n \subset H_0$ y $\ker(\mathbb{P} - \mathbb{I}) \cap H_0 = \{0\}$ tenemos que todos los eigenvalores de $\mathbb{P} \mid H_0$ tienen módulo $< 1$ y así $\lim_{n\to\infty} \mathbf{y}\mathbb{P}^n = 0$. □

PROPOSICIÓN 1.6. *Sean $\mathbb{P}$ una matriz estocástica $k \times k$ y $\boldsymbol{\pi}$ un vector de probabilidad asociado. El soporte de la medida de Markov $\mu$ asociada es $X_A$ donde*

$$A_{ij} = \begin{cases} 1 & si \quad P_{ij} > 0 \\ 0 & si \quad P_{ij} = 0 \end{cases}$$

DEMOSTRACIÓN. Sea $x \in X_A$. Toda vecindad $V$ de $x$ en $X$ contiene un cilindro $C(i; x_i, \ldots, x_{i+m})$ Entonces

$$\mu(V) \geq \mu(C(i; x_i, \ldots, x_{i+m})) = \pi_i P_{x_i x_{i+1}} \cdot P_{x_{i+m-1} x_{i+m}} > 0$$

puesto que $\pi_i > 0$ y $P_{x_j x_{j+1}} > 0$ ya que $A_{x_j x_{j+1}} = 1$.

Reciprocamente si $x \notin X_A$ existe $i \in \mathbb{Z}$ tal que $A_{x_i x_{i+1}} = 0$. Por lo que $P_{x_i x_{i+1}} = 0$. Así $\mu(C(i; x_i, x_{i+1})) = \pi_i P_{x_i x_{i+1}} = 0$. $\square$

EJERCICIO 1.1. Sea $X$ un campo vectorial $C^r$ en un abierto $M$ de $\mathbb{R}^n$. Decimos que una variedad $N \subset M$ de dimensión $n-1$ es una sección transversal de $X$ si $\forall_p \in N$ $X(p) \notin T_p N$ y $\forall x \in M$ la órbita $\{\varphi_t(x) : t > 0\}$ intersecta $N$. Sea $N$ una sección transversal de $X$ de clase $C^r$.

a) Probar que existe $f : N \hookleftarrow$ difeomorfismo de clase $C^r$ denominado transformación de Poincaré tal que $\forall x \in N$ $\exists \tau(x) > 0$ tal que $f(x) = \varphi_{\tau(x)}(x) \in N$ y $\varphi_t(x) \notin N$ para $0 < t < \tau(x)$.

b) Defina $\tilde{\omega}$ en $N$ por $\tilde{\omega}_p(v_1, \ldots, v_{n-1}) = \det(X(p), v_1, \ldots, v_{n-1})$ donde $p \in N$ $\quad v_1, \ldots, v_{n-1} \in T_p N$. Probar que $\tilde{\omega}$ es una forma de volumen en $N$ y que si $\operatorname{div} X = 0$ entonces $f$ preseva $\tilde{\omega}$.

EJERCICIO 1.2. Hallar $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ suave para que el campo definido por

$$X(x, y) = -f(x, y)(x, y)$$

preserve la medida de Lebesgue.

EJERCICIO 1.3.

a) Probar que si $f$ es un difeomorfismo de $M$ y $H : M \to \mathbb{R}$ es una primera integral entonces $H(y) = H(x)$ para todo $x \in M$, $y \in \omega(x)$.

b) El conjunto $\omega$ límite de un difeomorfismo $f : M \hookleftarrow$ se define como el conjunto $L^+(f) = \bigcup_{x \in M} \overline{\omega(x)}$. Probar que si un difeomorfismo $f : M \hookleftarrow$ posee una primera integral $L^+(f)$ es no numerable.

EJERCICIO 1.4.

a) Probar que para toda forma de volumen $C^1 \omega$ en $\mathbb{S}^1$ y todo difeomorfismo $g : \mathbb{S}^1 \hookleftarrow$ existe un difeomorfismo $f : \mathbb{S}^1 \times \mathbb{R} \hookleftarrow$ que preserva $\omega$ y tal que $f(p, 0) = (g(p), 0) \forall p \in \mathbb{S}^1$.

b) Probar que existen difeomorfismos de variedades compactas que preservan formas de volumen para los cuales hay subvariedades invariantes compactas tal que la restricción a estas subvariedades no preservan ninguna forma de volumen.

Capítulo 2

# Ergodicidad

## 1. El teorema ergódico

TEOREMA ERGÓDICO DE BIRKHOFF. *Sean $(X, \mathcal{A}, \mu)$ un espacio de probabilidad y $T : X \hookleftarrow$ una transformación preservadora de medida. Sea $f \in L^1(X)$, entonces $\left\{\frac{1}{n}\sum_{j=0}^{n-1} f(T^j x)\right\}$ converge casi dondequiera a una $\tilde{f} \in L^1(X)$ y*

$$\int_X f d\mu = \int_X \tilde{f} d\mu.$$

$$\text{Sean} \quad f^+(x) = \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x), \quad f^-(x) = \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x).$$

De la igualdad

$$\frac{1}{n+1}\sum_{j=0}^{n} f(T^j x) = \left(\frac{n}{n+1}\right)\frac{1}{n}\sum_{j=0}^{n-1} f(T^j T x) + \frac{f(x)}{n+1}$$

tenemos que $f^+(Tx) = f^+(x)$ y $f^-(Tx) = f^-(x)$.
Utilizaremos el siguiente lema conocido como Teorema ergódico maximal.

LEMA 2.1. *Sea $f \in L^1(X)$ y definamos*

$$E(f) = \left\{x : \sup_n \sum_{j=0}^{n} f(T^j x) > 0\right\}.$$

*Entonces $\int_{E(f)} f d\mu \geq 0$*

COROLARIO 2.1. *Si $A \subset E(f), A \in \mathcal{A}$ y $T^{-1}(A) = A$, entonces $\int_A f d\mu \geq 0$.*

DEMOSTRACIÓN. Como $T^{-1}(A) = A, E(\chi_A f) = A$. Por lo tanto $\int_A f = \int_{E(\chi_A f)} \chi_A f \geq 0$. □

DEMOSTRACIÓN. del teorema ergódico. Para $\alpha \in \mathbb{R}$. Sean

$$E_\alpha^+(f) = \left\{x : f^+(x) > \alpha\right\}, \quad E_\alpha^-(f) = \left\{x : f^-(x) < \alpha\right\}.$$

Luego, $T^{-1}(E_\alpha^+(f)) = E_\alpha^+(f)$, $T^{-1}(E_\alpha^-(f)) = E_\alpha^-(f)$ y $E_\alpha^+(f) = E_0^+(f - \alpha)$, $E_\alpha^- = E_{-\alpha}^+(-f)$.

Afirmamos que si $A \subset E^+_\alpha(f)$ y $T^{-1}A = A$, entonces $\int_A f d\mu \geq \alpha\mu(A)$. En efecto

$$\int_A f d\mu = \int_A (f-\alpha) d\mu + \alpha\mu(A)$$

Como $A \subset E^+_\alpha(f) = E^+_0(f-\alpha) \subset E(f-\alpha)$, por el Corolario 2.1

$$\int_A (f-\alpha) d\mu \geq 0.$$

Similarmente, si $A \subset E^-_\beta(f)$ y $T^{-1}A = A$, entonces $\int_A f d\mu \leq \beta\mu(A)$. Por lo tanto, tomando $A = E^+_\alpha(f) \cap E^-_\beta(f)$, para $\beta < \alpha$ tenemos

$$\alpha\mu(A) \leq \int_A f d\mu \leq \beta\mu(A)$$

y así $\mu(E^+_\alpha(f) \cap E^-_\beta(f)) = 0$.

Tomando una sucesión $\{\alpha_n\}$ densa en $\mathbb{R}$ tenemos

$$\{x : f^+(x) > f^-(x)\} = \bigcup_{\alpha_n > \alpha_m} E^+_{\alpha_n}(f) \cap E^-_{\alpha_m}(f),$$

de donde $\mu(\{x : f^+(x) > f^-(x)\}) = 0$.

$\left\{\frac{1}{n}\sum\limits_{j=0}^{n-1} \mid f(T^j x) \mid\right\}$ converge casi dondequiera a $g$ y $\mid \tilde{f} \mid \leq g$.

Como $T$ preserva $\mu$,

$$\int_X \frac{1}{n}\sum_{j=0}^{n-1} \mid f \circ T^j \mid d\mu = \frac{1}{n}\sum_{j=0}^{n-1}\int_X \mid f \circ T^j \mid d\mu = \int_X \mid f \mid d\mu.$$

Por el lema de Fatou $\int_X g d\mu \leq \int_X |f|\, d\mu$ y así $\int_X |\tilde{f}| d\mu \leq \int_X |f| d\mu$. Si $f \in L^\infty(X)$, $\left\|\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j\right\|_\infty \leq \|f\|_\infty$. Así $\|\tilde{f}\|_\infty \leq \|f\|_\infty$ y $\left\|\tilde{f} - \frac{1}{n}\sum\limits_{j=0}^{n-1} f \circ T^j\right\|_\infty \leq 2\|f\|_\infty$. Por convergencia dominada

$$\lim_{n\to\infty}\int_X \left|\tilde{f} - \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j\right| d\mu = \int_X \lim_{n\to\infty}\left|\tilde{f} - \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j\right| d\mu = 0.$$

En el caso general dado $\varepsilon > 0$, $\exists\ f_0 \in L^\infty(X)$ tal que $\int_X |f - f_0|\, d\mu \leq \varepsilon$ y tomamos $N$ tal que si $n \geq N$

$$\int_X \left|\tilde{f}_0 - \frac{1}{n}\sum_{j=0}^{n-1} f_0 \circ T^j\right| d\mu \leq \varepsilon$$

Entonces

$$\int_X \left| \tilde{f} - \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \right| d\mu \leq \int_X \left| \tilde{f} - \tilde{f}_0 \right| d\mu + \int_X \left| \tilde{f}_0 - \frac{1}{n}\sum_{j=0}^{n-1} f_0 \circ T^j \right| d\mu$$
$$+ \int_X \left| \frac{1}{n}\sum_{j=0}^{n-1} (f - f_0) \circ T^j \right| d\mu$$

Pero $\int_X \left| \tilde{f} - \tilde{f}_0 \right| d\mu = \int_X \left| (f - f_0)^{\sim} \right| d\mu \leq \int_X |f - f_0| \, d\mu \leq \varepsilon$ e

$$\int_X \left| \frac{1}{n}\sum_{j=0}^{n-1} (f - f_0) \circ T^j \right| d\mu \leq \int_X |f - f_0| \, d\mu \leq \varepsilon.$$

Luego $\int_X \left| \tilde{f} - \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j \right| d\mu \leq 3\varepsilon$ si $n \geq N$. De aquí,

$$\int_X \tilde{f} d\mu = \lim_{n\to\infty} \int_X \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j = \int_X f d\mu$$

□

Definición 2.1. Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad y $T : X \to X$ transformación preservadora de medida. Decimos que $T$ es $\mu$- *ergódica* o que $\mu$ es una medida $T$-*ergódica*, si para $B \in \mathcal{A}$, $T^{-1}B = B \Rightarrow \mu(B) = 0$ ó 1. Denotamos

$$\mathcal{M}_{\text{erg}}(T) = \{\mu \in M_T(X) : \mu \text{ es } T - \text{ergódica}\}$$

Lema 2.2. *Son equivalentes*

(a) *$T$ es $\mu$-ergódica.*
(b) *$f \circ T = f$, $f \in L^1(X, \mathcal{A}, \mu) \Rightarrow f$ es constante c.d.*
(c) *$f \circ T = f$, $f \in L^2(X, \mathcal{A}, \mu) \Rightarrow f$ es constante c.d.*

Demostración. (a) $\Rightarrow$ (b). Sean $T$ $\mu$-ergódica y $f \in L^1$ tales que $f \circ T = f$. Sin pérdida de generalidad supondremos que $f$ es real.

Para $n \in \mathbb{N}$, $k \in \mathbb{Z}$ sea $X(k, n) = f^{-1}[\frac{k}{2^n}, \frac{k+1}{2^n}[$. Para $n$ fija $X = \bigcup_{k\in\mathbb{Z}} X(k, n)$.

Como $f \circ T = f$, $T^{-1}(X(k, n)) = X(k, n)$ y entonces $\mu(X(k, n)) = 0$ ó 1.

Así, $\exists! \, k_n \in \mathbb{Z}$ tal que $\mu(X(k_n, n)) = 1$. Sean $Y = \bigcap_{n\in\mathbb{N}} X(k_n, n)$ y $\{\bar{x}\} = \bigcap_{n\in\mathbb{N}} \left[\frac{k_n}{2^n}, \frac{k_n+1}{2^n}\right]$.

Entonces $\mu(Y) = 1$ y $f(x) = \bar{x}$ para todo $x \in Y$.

(b) $\Rightarrow$ (c). $L^2(X, \mathcal{A}, \mu) \subset L^1(X, \mathcal{A}, \mu)$ ya que $\mu(X) = 1$.

(c) $\Rightarrow$ (a). Si $T^{-1}A = A$, $A \in \mathcal{A}$ entonces $\chi_A \circ T = \chi_{T^{-1}A} = \chi_A$. Así, $\chi_A =$ constante c.d. O sea $\mu(A) = 0$ ó 1. □

Corolario 2.2. *Si $T$ es $\mu$-ergódica entonces*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) = \int f d\mu$$

*para toda $f \in L^1(X, \mathcal{A}, \mu)$.*

Demostración. Sea $T$ $\mu-$ergódica y sea $f \in L^1$. Sea $\tilde{f}$ dada por el teorema ergódico. Como $\tilde{f} \circ T = \tilde{f}$, $\tilde{f}$ es constante c.d. Así $\tilde{f} = \int \tilde{f} d\mu = \int f d\mu$ c. d. □

Corolario 2.3. *Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad, $T : X \hookleftarrow$ preservadora de medida. $T$ es ergódica si y solo si*

$$\forall A, B \in \mathcal{A} \quad \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap T^{-i}B) = \mu(A)\mu(B).$$

Demostración. ($\Rightarrow$) Sean $T$ ergódica y $A, B \in \mathcal{A}$, entonces

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} \chi_B \circ T^i = \mu(B)$$

casi dondequiera. Así

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap T^{-i}B) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\int \chi_B \circ T^i \chi_A d\mu \\ &= \mu(B)\int \chi_A d\mu = \mu(A)\mu(B).\end{aligned}$$

($\Leftarrow$) Si $T^{-1}A = A$,

$$\mu(A) = \frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap T^{-i}A) \to \mu(A)^2.$$

Así $\mu(A) = 0$ ó 1. □

Demostración. del Lema 2.1. Sea

$$f_n(x) = \max\left\{0, f(x), f(x) + f(Tx), \ldots, \sum_{j=0}^{n-1} f(T^j x)\right\}.$$

Como $E(f) = \bigcup_{n\in\mathbb{N}}\{x : f_n(x) > 0\}$, basta demostrar que $\forall n \in \mathbb{N}$ se tiene $\int_{\{f_n>0\}} f d\mu \geq 0$ . Para $0 \leq m < n$

$$f_n(Tx) + f(x) \geq \sum_{j=0}^{m} f(T^j Tx) + f(x) = \sum_{j=0}^{m+1} f(T^j x).$$

Así $f_n(Tx) + f(x) \geq f_n(x)$ cuando $f_n(x) > 0$. Por lo tanto

$$\begin{aligned}\int_{\{f_n>0\}} f d\mu &\geq \int_{\{f_n>0\}} f_n - \int_{\{f_n>0\}} f_n \circ T \\ &= \int_X f_n - \int_{\{f_n>0\}} f_n \circ T \geq \int_X f_n - \int_X f_n \circ T = 0\end{aligned}$$

ya que $f_n = 0$ en $X \setminus \{f_n > 0\}$ y $f_n \circ T \geq 0$. □

Definición 2.2. Sea $X$ espacio métrico compacto, $T : X \to X$ continua. Dada $\mu \in \mathcal{M}_T(X)$, su cuenca $B(\mu)$ es el conjunto de los puntos $x \in X$ tales que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu \tag{2.1}$$

para toda $\varphi \in C(X)$. Note que la cuenca es un conjunto invariante.

Proposición 2.1. *Sean $X$ espacio métrico compacto, $T : X \to X$ continua. Si $\mu \in \mathcal{M}_{erg}(T)$, entonces $\mu(B(\mu)) = 1$.*

Demostración. Sea $\{g_m : m \in \mathbb{N}\}$ denso en la bola unitaria de $C(X)$. Por el teorema ergódico el conjunto $A_m$ de los puntos $x \in X$ tales que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g_m(T^i(x)) = \int g_m d\mu,$$

tiene $\mu$-medida total. Afirmamos que $B(\mu) = \bigcap\limits_m A_m$. Probaremos la igualdad (2.1) para $x \in \bigcap\limits_m A_m$ y $\varphi \in C(X)$ con $\|\varphi\| = 1$. Dado $\varepsilon > 0$ sea $m$ tal que $\|g_m - \varphi\| \leq \varepsilon/3$. Sea $N$ tal que si $n \geq N$

$$\Big|\frac{1}{n} \sum_{i=0}^{n-1} g_m(T^i(x)) - \int_X g_m d\mu\Big| \leq \frac{\varepsilon}{3}.$$

Si $n \geq N$,

$$\begin{aligned}\Big|\frac{1}{n} \sum_{i=0}^{n-1} \varphi(T^i(x)) - \int_X \varphi d\mu\Big| &\leq \frac{1}{n} \sum_{i=0}^{n-1} |\varphi(T^i(x)) - g_m(T^i(x))| \\ &\quad + \Big|\frac{1}{n} \sum_{i=0}^{n-1} g_m(T^i(x)) - \int_X g_m d\mu\Big| + \Big|\int_X (g_m - \varphi) d\mu\Big| \leq \varepsilon\end{aligned}$$

□

## 2. Ejemplos de transformaciones ergódicas

LEMA 2.3. *Sean $X$ espacio métrico compacto, $\mathcal{A}$ la $\sigma$-álgebra de Borel y $\mu$ medida de probabilidad en $(X, \mathcal{A})$ tal que $\mu(U) > 0$ para todo abierto no vacio $U$. Sea $T : X \to X$ continua y ergódica. Entonces la órbita $\Big\{T^m(x) : m \in \mathbb{N}\Big\}$ es densa para casi todo $x \in X$.*

DEMOSTRACIÓN. Sea $\Big\{U_m\Big\}$ base numerable para la topología de $X$. $\Big\{T^m(x) : m \in \mathbb{N}\Big\}$ es densa $\iff \forall m\, \exists n_m \operatorname{con} T^{n_m} x \in U_m \iff$

$$x \in \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq 0} T^{-n} U_m.$$

$B_m := \bigcup\limits_{n\geq 0} T^{-n} U_m$ es abierto y $T^{-1} B_m \subset B_m$.

Como $T$ es ergódica, $\mu(B_m) = 0$ ó 1. En efecto sea $B_m^* = \bigcap\limits_{n\geq 0} T^{-n} B_m$, entonces $T^{-1} B_m^* = B_m^*$ y $\mu(B_m^*) = \text{lím}_{n\to\infty} T^{-n}(B_m) = \mu(B_m)$.

Como $\mu(B_m) \geq \mu(U_m) > 0$, $\mu(B_m) = 1$ y así $B_m$ es denso.

Por inducción $A_k = \bigcap\limits_{m\geq k} B_m$ es abierto no vacío y $T^{-1}(A_k) \subset A_k$. Así $\mu(A_k) = 1$ y

$$\mu\Big(\bigcap_{m\in\mathbb{N}} B_m\Big) = \lim_{k\to\infty} \mu(A_k) = 1.$$

$\square$

PROPOSICIÓN 2.2. *Consideremos la rotación $R_\alpha : \mathbb{T}^n \hookleftarrow$ definida por $R_a(x + \mathbb{Z}^n) = x + a + \mathbb{Z}^n$. Son equivalentes*

(a) *$R_a$ es ergódica*
(b) *La órbita $\{R_a^m(\mathbb{Z}^n) : m \in \mathbb{N}\}$ es densa.*
(c) *$\forall k \in \mathbb{Z}^n \setminus \{0\}\ \langle k, a\rangle \notin \mathbb{Z}$.*

DEMOSTRACIÓN. Utilizaremos series de Fourier.

Para $k \in \mathbb{Z}^n$ definimos $\varphi_k : \mathbb{T}^n \to \mathbb{S}^1$ mediante $\varphi_k(x + \mathbb{Z}^n) = \exp(2\pi i \langle k, x\rangle)$. Entonces

$$\varphi_k \circ R_a = \exp(2\pi i \,\langle k, a\rangle)\varphi_k$$

(a) $\Rightarrow$ (b) es consecuencia del lema 2.3

(b) $\Rightarrow$ (c). Si $\langle k, a\rangle \in \mathbb{Z}$, entonces $\varphi_k \circ R_a^m(\mathbb{Z}^n) = \exp(2m\pi i\langle k, a\rangle) = 1\ \forall m \in \mathbb{N}$. Como $\varphi_k$ es continua y $\{R_a^m(\mathbb{Z}^n) : m \in \mathbb{N}\}$ es densa, $\varphi_k \equiv 1$ y así $k = 0$.

(c) $\Rightarrow$ (a). Sea $f \in L^2(\mathbb{T}^n)$ entonces $f = \sum\limits_{k\in\mathbb{Z}^n} b_k \varphi_k$.

Si $f \circ R_a = f$ entonces $b_k = \exp(2\pi i\langle k, a\rangle) b_k\ \forall k \in \mathbb{Z}^n$. Si $b_k \neq 0$ entonces $\langle k, a\rangle \in \mathbb{Z}$, luego $k = 0$ y así $f$ es constante. $\square$

PROPOSICIÓN 2.3. *Sea $A \in GL(n,\mathbb{Z})$ y sea $\tilde{A} : \mathbb{T}^n \hookleftarrow$ la transformación inducida (Ejemplo 1.6). $\tilde{A}$ es ergódica $\iff$ ninguno de los eigenvalores de $A$ es raíz de la unidad*

DEMOSTRACIÓN. Sea $\varphi_k : \mathbb{T}^n \to \mathbb{S}^1$ como en la Proposición 2.2. Entonces

$$\begin{aligned}\varphi_k \circ \tilde{A}(x + \mathbb{Z}^n) &= \varphi_k(Ax + \mathbb{Z}^n) = \exp(2\pi i \langle k, Ax\rangle)\\ &= \exp(2\pi i\langle A^t k, x\rangle) = \varphi_{A^t k}(x + \mathbb{Z}^n)\end{aligned}$$

($\Rightarrow$). Supongamos que $A$ tiene un eigenvalor raíz $m$-ésima de la unidad. Entonces $A^m$ tiene a 1 como eigenvalor. Como $A^m(\mathbb{Q}^n) \subset \mathbb{Q}^n$. $\exists q \in \mathbb{Q}^n - \{0\}$ tal que $q^t A^m = q^t$ y así $\exists k \in \mathbb{Z}^n - \{0\}$ tal que $k^t A^m = k^t$ y luego $\varphi_k \circ \tilde{A}^m = \varphi_k$. Por lo tanto $f = \sum_{i=0}^{m-1} \varphi_k \circ \tilde{A}^i$ satisface $f \circ \tilde{A} = f$ y así $\tilde{A}$ no es ergódica.

($\Leftarrow$). Sea $f \in L^2(\mathbb{T}^n)$, $f = \sum\limits_{k\in\mathbb{Z}} b_k \varphi_k$ tal que $f \circ \tilde{A} = f$ Entonces $b_{A^t k} = b_k$ $\forall k \in \mathbb{Z}^n$. Como ninguno de los eigenvalores es raíz de la unidad, si $k \neq 0$, todos los elementos de la sucesión $\{(A^t)^m k : m \in \mathbb{N}\}$ son distintos y así $\|f\|^2 = \sum\limits_{r\in\mathbb{Z}^n} |b_r|^2 \geq \sum\limits_{m\in\mathbb{N}} \left|b_{(A^t)^m k}\right|^2 = \sum\limits_{m\in\mathbb{N}} |b_k|^2$ por lo tanto $b_k = 0$ y así $f$ es constante. $\square$

PROPOSICIÓN 2.4. *El corrimiento $T : X \hookleftarrow$ con vector de probabilidad $(p_0, \ldots, p_{k-1})$ es ergódico.*

DEMOSTRACIÓN. La siguiente propiedad se verifica facilmente:

Si $A_1, A_2 \subset X$ son uniones finitas de cilindros existe $m_0 \in \mathbb{N}$ tal que si $m \geq m_0$, entonces $\mu(T^{-m}(A_1) \cap A_2) = \mu(A_1)\mu(A_2)$.

Sea $A \in \mathcal{A}$ tal que $T^{-1}A = A$. Dada $\varepsilon > 0$ $\exists A_0$ unión finita disjunta de cilindros tal que $\mu(A_0 \triangle A) \leq \varepsilon$. Para algún $m$

$$\begin{aligned}\mu(T^{-m}(A_0^c) \cap A_0) &= \mu(A_0)\mu(A_0^c)\\ \mu(T^{-m}(A_0) \cap A_0^c) &= \mu(A_0)\mu(A_0^c).\end{aligned}$$

Como

$$\begin{aligned}\mu(T^{-m}(A_0)\triangle A_0) \leq \mu(T^{-m}A_0 \triangle T^{-m}A) &+ \mu(T^{-m}A\triangle A)\\ &+ \mu(A\triangle A_0) = 2\mu(A\triangle A_0) \leq 2\varepsilon\end{aligned}$$

y

$$\begin{aligned}\mu(T^{-m}(A_0)\triangle A_0) &= \mu(T^{-m}A_0 \cap A_0^c) + \mu(T^{-m}A_0^c \cap A_0)\\ &= 2\mu(A_0)\mu(A_0^c) = 2\mu(A_0)(1 - \mu(A_0)),\end{aligned}$$

se tiene que $\mu(A_0)(1 - \mu(A_0)) \leq \varepsilon$. Como $|\mu(A) - \mu(A_0)| \leq \varepsilon$,

$$\mu(A)(1 - \mu(A)) \leq (\mu(A_0) + \varepsilon)(1 - \mu(A_0) + \varepsilon) \leq 3\varepsilon$$

Como $\varepsilon > 0$ es arbitrario, $\mu(A)(1 - \mu(A)) = 0$ o sea $\mu(A) = 0$ ó 1. $\square$

LEMA 2.4. *Sea $\mathbb{P}$ una matriz estocástica con vector de probabilidad asociado $\boldsymbol{\pi}$. Entonces*

$$\mathbb{K} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{P}^i \tag{2.2}$$

*existe, $\mathbb{K}$ es estocástica y $\mathbb{K}\mathbb{P} = \mathbb{P}\mathbb{K} = \mathbb{K} = \mathbb{K}^2$.*

DEMOSTRACIÓN. Sea $\sigma : X \hookleftarrow$ el corrimiento y sea $\mu$ la medida de Markov asociada a $(\mathbb{P}, \boldsymbol{\pi})$. Sea $\chi_j$ la función característica de $C(0; j)$. Por el teorema ergódico

$$\tilde{\chi}_j = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_j \circ \sigma^i$$

existe c. d.. Multiplicando por $\chi_l$ y usando el teorema de la convergencia dominada

$$\int \chi_l \tilde{\chi}_j d\mu = \frac{1}{n} \sum_{i=0}^{n-1} \int \chi_l \, \chi_j \circ \sigma^i d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(C(0; l) \cap C(i; j)).$$

Pero

$$\mu(C(0; l) \cap C(i; j)) = \sum_{a_1, \ldots, a_{i-1}} \pi_l P_{l a_1} \cdots P_{a_{i-1} j} = \pi_l P^i_{lj}.$$

Así

$$\frac{1}{n} \sum_{i=0}^{n-1} P^i_{lj} \to K_{lj} := \frac{1}{\pi_l} \int \chi_l \, \tilde{\chi}_j d\mu.$$

Sea $\mathbb{K} = [K_{lj}]$, como cada $\mathbb{P}^i$ es estocástica se tiene que $\mathbb{K}$ es estocástica. De la definición de $\mathbb{K}$ es claro que $\mathbb{K}\mathbb{P} = \mathbb{P}\mathbb{K} = \mathbb{K}$ y así $\mathbb{K}\mathbb{P}^i = \mathbb{K}$ $\forall i \geq 0$, de donde $\mathbb{K}^2 = \mathbb{K}$. □

DEFINICIÓN 2.3. $\mathbb{P}$ matriz estocástica $k \times k$ se llama *irreducible* si

$$\forall i, j \in [k] \, \exists n \in \mathbb{N} \, \text{tal que} \, P^n_{ij} > 0.$$

TEOREMA 2.4. *Sean $\mathbb{P}, \boldsymbol{\pi}$ como en el lema 2.4 y $\mathbb{K}$ la matriz dada por* (2.2)*. Sea $\mu$ la medida de Markov asociada. Son equivalentes*

(1) *El corrimiento $\sigma$ es ergódico.*
(2) *Todos los renglones de $\mathbb{K}$ son iguales.*
(3) *$\mathbb{P}$ es irreducible.*
(4) *1 es un eigenvalor simple de $\mathbb{K}$.*

DEMOSTRACIÓN. (1) $\Rightarrow$ (2). Como en la prueba del lema 2.4

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(C(0; l) \cap C(i; j)) = \pi_l K_{lj}.$$

Como $\sigma$ es ergódico

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(C(0;l)\cap C(i;j))=\pi_l\pi_j.$$

Así $K_{lj}=\pi_j$.

(2) $\Rightarrow$ (3). Supongamos que todos los renglones de $\mathbb{K}$ son iguales a $(q_0,\ldots,q_{k-1})$. Como $\pi\mathbb{K}=\pi$ se tiene que $\sum\limits_{j\in[k]}\pi_j q_i=\pi_i\ \forall_i\in[k]$ o sea $q_i=\pi_i>0$. Si existieran $i,j\in[k]$ tales que $P_{ij}^n=0\ \forall n\in\mathbb{N}$ entonces tendríamos $q_j=0$.

(3) $\Rightarrow$ (2). Sea $A_i=\{j\in[k]:K_{ij}>0\}$. Como $\mathbb{K}$ es estocástica, $A_i\neq\emptyset$. Sean $l\in A_i$, $j\in[k]$ y sea $n\in\mathbb{N}$ tal que $P_{lj}^n>0$. Como $\mathbb{K}=\mathbb{K}\mathbb{P}^n$,

$$K_{ij}=\sum_{r\in[k]}K_{ir}P_{rj}^n\geq K_{il}P_{lj}^n>0,$$

o sea $j\in A_i$. Así $K_{ij}>0\ \forall i,j\in[k]$.

Fijo $j\in[k]$ sea $q_j=\max\{K_{ij}:i\in[k]\}$. Como $\mathbb{K}^2=\mathbb{K}$, $\forall l\in[k]$.

$$K_{lj}=\sum_{i\in[k]}K_{li}K_{ij}.$$

Si $K_{ij}<q_j$ para algún $i\in[k]$ entonces

$$K_{lj}<\sum_{i\in[k]}K_{li}q_j=q_j\quad\forall l\in[k].$$

Por lo tanto $K_{ij}=q_j\ \forall i\in[k]$.

(2) $\Rightarrow$ (1). Para demostrar que $\sigma$ es ergódica es suficiente demostrar que si $A=C(j;i_0,\ldots,i_s)\ \ B=C(l;r_0,\ldots,r_m)$ se tiene

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap\sigma^{-i}B)=\mu(A)\mu(B).$$

Para $i>j+s-l$,

$$\mu(A\cap\sigma^{-i}B)=\pi_{i_0}P_{i_0i_1}\cdots P_{i_{s-1}i_s}P_{i_sr_0}^{i+l-(j+s)}P_{r_0r_1}\cdots P_{r_{m-1}r_m}.$$

como (2) $\Rightarrow K_{ij}=\pi_j$, tenemos que

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mu(A\cap\sigma^{-i}B)$$
$$=\pi_{i_0}P_{i_0i_1}\cdots p_{i_{s-1}i_s}\pi_{r_0}P_{r_0r_1}\cdots P_{r_{m-1}r_m}=\mu(A)\mu(B)$$

(2) $\Rightarrow$ (4). $K_{ij}=\pi_j$. Si $\mathbb{K}\alpha=\alpha$ entonces

$$\alpha_i=\sum_j K_{ij}\alpha_j=\sum_j\pi_j\alpha_j\quad\forall i\in[k]$$

o sea que $\alpha$ es multiplo del vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

(4) $\Rightarrow$ (2). Ya que $\mathbb{K}\mathbb{K} = \mathbb{K}$ cada columna de $\mathbb{K}$ es un eigenvector para el eigenvalor 1. Como además $\mathbb{K}$ es estocástica, cada columna es un multiplo del eigenvector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. O sea que todos los renglones de $\mathbb{K}$ son iguales. $\square$

## 3. Aplicaciones del teorema ergódico

Tomando $f = \chi_B$ para $B \in \mathcal{A}$ en el corolario 2.2 tenemos que si $T : X \hookleftarrow$ es ergódica entonces para casi todo $x \in X$

$$\tau_B(x) = \lim_{n\to\infty} \frac{1}{n} \# \left\{ 0 \le i \le n-1 : T^i x \in B \right\} = \mu(B)$$

Ejemplo 2.1. Considere el primer dígito de los números de la sucesión $2^n$ : $1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, \ldots$. Aparece el dígito 7 ? Que dígito aparece más frecuentemente, 7 u 8? Con cuánta más frecuencia?

El primer dígito de $2^n$ es $m$, precisamente cuando

$$m\,10^k \le 2^n < (m+1)10^k,\ k \in \mathbb{N}.$$

Tomando logaritmo de base 10

$$k + \log m \le n \log 2 < k + \log(m+1),$$

o sea

$$n \log 2 \in (\log m, \log(m+1)) \mod 1.$$

Por la proposición 2.2 la rotación $R : \mathbb{S}^1 \hookleftarrow$ por un ángulo $2\pi \log 2$, es ergódica y así

$$\lim_{n\to\infty} \frac{1}{n} \#\{0 \le k < n : x + k \log 2 \in (a,b) \mod 1\} = b - a$$

para casi todo $x \in [0,1]$

Ejemplo 2.2. Para casi todo $x \in [0,1]$ (con respecto a la medida de Lebesgue) el número promedio de ceros en la expansión decimal $x = 0.x_1 x_2 \cdots$ es $\frac{1}{10}$.

En efecto, consideremos $T : \mathbb{S}^1 \hookleftarrow$, $T(z) = z^{10}$. De acuerdo a la Proposición 2.3, $T$ es ergódica. Sea $B = \left\{ e^{2\pi i x} : x \in [0, \frac{1}{10}) \right\}$, entonces $x_j = 0 \iff T^j(e^{2\pi i x}) \in B$ y $\mu(B) = \frac{1}{10}$.

EJEMPLO 2.3. Sea $T : [0,1) \hookleftarrow$ la transformación de Gauss

$$T(x) = \begin{cases} 0 & \text{si } \ x = 0 \\ \dfrac{1}{x} - \left[\dfrac{1}{x}\right] & \text{si } \ x \neq 0 \end{cases}$$

Si $x \neq 0$, $n \in \mathbb{N}$, definamos $a_n(x) = \left[\dfrac{1}{T^{n-1}(x)}\right]$ y así $a_n(x) = k \iff T^{n-1}(x) \in I_k = \left]\frac{1}{k+1}, \frac{1}{k}\right]$. Tenemos que $\dfrac{1}{T^{n-1}(x)} = a_n(x) + T^n(x)$ y

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{\ddots \, \cfrac{}{a_n(x) + T^n(x)}}}.$$

Sea

$$\frac{p_n(x)}{q_n(x)} := [a_1(x), \ldots, a_n(x)] := \cfrac{1}{a_1(x) + \cfrac{1}{\ddots \, \cfrac{}{a_{n-1}(x) + \cfrac{1}{a_n(x)}}}}$$

con $p_n(x), q_n(x)$ primos relativos.

$$\begin{aligned} \frac{p_n(x)}{q_n(x)} = [a_1(x), \ldots, a_n(x)] &= \frac{1}{a_1(x) + [a_2(x), \ldots, a_n(x)]} \\ &= \frac{1}{a_1(x) + p_{n-1}(Tx)/q_{n-1}(Tx)} = \frac{q_{n-1}(Tx)}{a_1(x) q_{n-1}(Tx) + p_{n-1}(Tx)}. \end{aligned}$$

Las fracciones de la izquierda y la derecha no se pueden simplificar, así $p_n(x) = q_{n-1}(Tx)$. Entonces

$$\frac{p_n(x)}{q_n(x)} \cdot \frac{p_{n-1}(Tx)}{q_{n-1}(Tx)} \cdots \frac{p_1(T^{n-1}x)}{q_1(T^{n-1}x)} = \frac{1}{q_n(x)},$$

de donde

$$-\frac{1}{n} \log q_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} \log \left( \frac{p_{n-i}(T^i x)}{q_{n-i}(T^i x)} \right).$$

Poniendo $p_0(x) = 0$, $q_0(x) = 1$, se demuestra por inducción que para $n \geq 2$

$$p_n(x) = a_n(x) p_{n-1}(x) + p_{n-2}(x) \tag{2.3}$$

$$q_n(x) = a_n(x) q_{n-1}(x) + q_{n-2}(x), \tag{2.4}$$

lo que implica que las sucesiones $\{p_n\}$ y $\{q_n\}$ son crecientes. Hagamos el paso inductivo

$$\begin{aligned} p_{n+1}(x) &= q_n(Tx) = a_n(Tx)q_{n-1}(Tx) + q_{n-2}(Tx) \\ &= a_{n+1}(x)p_n(x) + p_{n-1}(x) \\ q_{n+1}(x) &= a_1(x)q_n(Tx) + p_n(Tx) = a_1(x)(a_{n+1}(x)q_{n-1}(Tx) + q_{n-2}(Tx)) \\ &\quad + a_{n+1}(x)p_{n-1}(Tx) + p_{n-2}(Tx) = a_{n-1}(x)q_n(x) + q_{n-1}(x). \end{aligned}$$

Tenemos que

$$x = \frac{p_n(x) + T^n(x)p_{n-1}(x)}{q_n(x) + T^n(x)q_{n-1}(x)}.$$

En efecto, para $n = 1$ esto es claro y para el paso inductivo tenemos

$$\begin{aligned} x =& \frac{p_{n-1}(x) + T^{n-1}(x)p_{n-2}(x)}{q_{n-1}(x) + T^{n-1}(x)q_{n-2}(x)} = \frac{p_{n-1}(x) + \frac{p_{n-2}(x)}{a_n(x)+T^n(x)}}{q_{n-1}(x) + \frac{q_{n-2}(x)}{a_n(x)+T^n(x)}} \\ =& \frac{p_{n-1}(x)(a_n(x) + T^n(x)) + p_{n-2}(x)}{q_{n-1}(x)(a_n(x) + T^n(x)) + q_{n-2}(x)} = \frac{p_n(x) + T^n(x)p_{n-1}(x)}{q_n(x) + T^n(x)q_{n-1}(x)}. \end{aligned}$$

Utilizando (2.3) se demuestra por inducción que $p_n \geq 2^{(n-2)/2}$ y $q_n \geq 2^{(n-1)/2}$ para $n \geq 2$ y que

$$p_{n-1}(x)q_n(x) - p_n(x)q_{n-1}(x) = (-1)^n \quad n \geq 1 \tag{2.5}$$

Como

$$\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{q_n p_{n-1} - p_n q_{n-1}}{q_{n+1}q_n} = \frac{q_{n-1}}{q_{n+1}}\left(\frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n}\right), \tag{2.6}$$

$q_{n+1} > q_{n-1}$ y $p_1/q_1 > p_2/q_2$ se tiene que $\{p_{2n}/q_{2n}\}$ es creciente y $\{p_{2n-1}/q_{2n-1}\}$ es decreciente.

Proposición 2.5. *La transformación de Gauss $T$ es $\mu$-ergódica para $\mu$ definida en los borelianos por*

$$\mu(A) = \frac{1}{\log 2}\int_A \frac{dx}{1+x}.$$

Demostración. Denotando por $\lambda$ a la medida de Lebesgue tenemos

$$\frac{1}{2\log 2}\lambda(A) \leq \mu(A) \leq \frac{1}{\log 2}\lambda(A).$$

Para $n \in \mathbb{N}$ definamos la partición formada por los intervalos

$$J_{k_1,\ldots,k_n} := \bigcap_{i=1}^{n} T^{-i}I_{k_i} = \{x \in (0,1] : a_i(x) = k_i, i = 1, \ldots, n\},\ k_1, \ldots, k_n \in \mathbb{N}.$$

Para $x \in J_{k_1,\dots,k_n}$ tenemos que $p_i, q_i,\ i = 1, \dots, n$ son constantes y una biyección $\psi_{k_1,\dots,k_n} : [0,1[\to J_{k_1,\dots,k_n}$ dada por

$$\psi_{k_1,\dots,k_n}(t) = \frac{p_n + tp_{n-1}}{q_n + tq_{n-1}}.$$

Por lo tanto

$$J_{k_1,\dots,k_n} = \left[\frac{p_n + p_{n-1}}{q_n + q_{n-1}}, \frac{p_n}{q_n}\right[ \quad n \text{ par},$$

$$J_{k_1,\dots,k_n} = \left]\frac{p_n}{q_n}, \frac{p_n + p_{n-1}}{q_n + q_{n-1}}\right] \quad n \text{ impar}$$

$$\lambda(J_{k_1,\dots,k_n}) = 1/(q_n(q_n + q_{n-1}))$$

Para una medida $\nu$ consideremos la medida condicional

$$\nu(A|B) = \frac{\nu(A \cap B)}{\nu(B)}$$

Tomemos $J_n$ un intervalo de la forma $J_{k_1,\dots,k_n}$

$$\lambda(\{x : y < T^n(x) < z\}|J_n) = \frac{\psi(z) - \psi(y)}{\psi(1) - \psi(0)}$$

Como

$$\psi(z) - \psi(y) = (y - z)\frac{p_{n-1}q_n - p_n q_{n-1}}{(q_n + yq_{n-1})(q_n + zq_{n-1})} = \frac{(-1)^n(y - z)}{(q_n + yq_{n-1})(q_n + zq_{n-1})},$$

tenemos

$$\lambda(T^{-n}[y, z[|J_n) = \frac{(z - y)q_n(q_n + q_{n-1})}{(q_n + yq_{n-1})(q_n + zq_{n-1})},$$

y así

$$\frac{1}{2}(z - y) \le \lambda(T^{-n}[y, z[|J_n) \le 2(z - y).$$

Ya que los intervalos $[y, z[$ generan la $\sigma$-álgebra de Borel, tenemos que

$$\frac{1}{2}\lambda(A) \le \lambda(T^{-n}A|J_n) \le 2\lambda(A)$$

para cualquier boreliano $A$ y por consiguiente

$$\frac{1}{4}\mu(A) \le \mu(T^{-n}A|J_n) \le 2\mu(A).$$

Sea $A$ boreliano tal que $T^{-1}A = A$ con $\mu(A) > 0$, entonces $\frac{1}{4}\mu(A) \le \mu(A|J_n)$, ó equivalentemente, $\frac{1}{4}\mu(J_n) \le \mu(J_n|A)$. Como los intervalos $J_n$ generan la $\sigma$-algebra de Borel, tenemos que $\frac{1}{4}\mu(B) \le \mu(B|A)$ para todo boreliano $B$. Escogiendo $B = [0,1] - A$ obtenemos que $\mu(A) = 1$. $\square$

Corolario 2.5. *Para casi todo $x$ (respecto a la medida de Lebesgue)*

$$\lim_{n\to\infty} \sqrt[n]{a_1(x)\cdots a_n(x)} = \prod_{k=1}^{\infty}\left(1+\frac{1}{k^2+2k}\right)^{\log k/\log 2} \tag{2.7}$$

$$\lim_{n\to\infty} \frac{\log q_n(x)}{n} = \frac{\pi^2}{12\log 2} \tag{2.8}$$

Demostración. 2.7 Sea $f : ]0,1[ \to \mathbb{R}$ definida por $f(x) = \log k$ para $x \in I_k$

$$\frac{1}{n}\sum_{i=1}^{n}\log a_i = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$$

converge a

$$\frac{1}{\log 2}\int_0^1 \frac{f(x)dx}{1+x} = \sum_{k=1}^{\infty}\frac{\log k}{\log 2}\log\left(\frac{1+\frac{1}{k}}{1+\frac{1}{k+1}}\right).$$

2.8. Utilizando (2.6) se sigue fácilmente por inducción que

$$\left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| \leq \frac{1}{q_{n-1}q_n}.$$

Similarmente se prueba por inducción que

$$\frac{p_{2n}(x)}{q_{2n}(x)} < x < \frac{p_{2n-1}(x)}{q_{2n-1}(x)}$$

y luego

$$\left|\frac{xq_n(x)}{p_n(x)} - 1\right| = \frac{q_n(x)}{p_n(x)}\left|x - \frac{p_n(x)}{q_n(x)}\right| \leq \frac{1}{p_n(x)q_{n+1}(x)} \leq 2^{1-n},$$

$$\left|\frac{p_n(x)}{xq_n(x)} - 1\right| = \frac{p_n(x)}{q_n(x)}\left|\frac{1}{x} - \frac{q_n(x)}{p_n(x)}\right| \leq \frac{1}{q_n(x)p_{n+1}(x)} \leq 2^{1-n},$$

de donde

$$\begin{aligned}\sum_{i=0}^{n-1}\left|\log\left(\frac{p_{n-i}(T^i x)}{q_{n-1}(T^i x)}\right) - \log\left(T^i x\right)\right| &= \sum_{i=0}^{n-1}\left|\log\frac{(T^i x)q_{n-i}(T^i x)}{p_{n-i}(T^i x)}\right| \\ \leq \sum_{n-i \text{ par}}\left|\frac{(T^i x)q_{n-i}(T^i x)}{p_{n-i}(T^i x)} - 1\right| &+ \sum_{n-i \text{ impar}}\left|\frac{p_{n-i}(T^i x)}{(T^i x)q_{n-i}(T^i x)} - 1\right| \\ &\leq \sum_{i=0}^{n-1} 2^{1+i-n} \leq 2.\end{aligned}$$

Por lo tanto

$$\lim_{n\to\infty}\frac{\log q_n(x)}{n} = \lim_{n\to\infty}\frac{-1}{n}\sum_{i=0}^{n-1}\log(T^i x) = \frac{-1}{\log 2}\int_0^1 \frac{\log x}{1+x}dx = \frac{\pi^2}{12\log 2}.$$

□

## 4. Descomposición en medidas ergódicas

Dado un subconjunto convexo $C$ de un espacio vectorial topológico $V$ definimos el conjunto de puntos extremos de $C$ como

$\mathrm{Ext}(C) = \{u \in V : u = (1-\alpha)x + \alpha y,\ \alpha \in [0,1],\ x \neq y \Rightarrow \alpha = 0, 1\}$

Si $C \neq \emptyset$, $\mathrm{Ext}(C) \neq \emptyset$.

LEMA 2.5. *Los puntos extremos de $\mathcal{M}_T(X)$ son las medidas $T$-ergódicas. Dos medidas ergódicas $T$-distintas son mutuamente singulares.*

DEMOSTRACIÓN. Sea $\mu$ punto extremo de $\mathcal{M}_T(X)$ y supongamos que $\exists A \in \mathcal{A}$ tal que $T^{-1}A = A$ y $\mu(A) \neq 0, 1$. Definamos $\mu_A, \mu_{A^c}$ mediante

$$\mu_A(B) = \frac{\mu(B \cap A)}{\mu(A)}, \quad \mu_{A^c}(B) = \frac{\mu(B \setminus A)}{1 - \mu(A)},$$

entonces $\mu_A, \mu_{A^c} \in M_T(X)$ y $\mu = \mu(A)\mu_A + (1 - \mu(A))\mu_{A^c}$. Contradiciendo el hecho que $\mu$ es punto extremo.

Reciprocamente supongamos que $\mu$ es ergódica y $\mu = (1-\alpha)\mu_1 + \alpha\mu_2$ con $0 < \alpha < 1$ $\mu_i \in M_T(X)$. Entonces $\mu(A) = 0 \Rightarrow \mu_i(A) = 0$, es decir $\mu_i \ll \mu$ y así $\exists f_i \in L^1(X, \mathcal{A}, \mu)$ tal que $f_i \geq 0$ y $\mu_i(A) = \int_A f_i d\mu$ como $T$ preserva $\mu$ y $\mu_i$ se tiene que $f_i \circ T = f_i$. Como $\mu$ es ergódica $f_i$ es constante y así $\mu_i = \mu$.

Sean $\mu_1, \mu_2$ medidas $T$-ergodicas distintas y sea $A \in \mathcal{A}$ tal que $\mu_1(A) \neq \mu_2(A)$. Definiendo

$$\tau_A(x) = \lim_{n\to\infty} \frac{1}{n} \# \left\{0 \leq i \leq n-1 : T^i x \in A\right\},$$

$$A_i = \{x \in X : \tau_A(x) = \mu_i(A)\},$$

tenemos que $A_1, A_2$ son ajenos y $\mu_1(A_2) = \mu_2(A_1) = 0$. □

PROPOSICIÓN 2.6. *Dadas $\mathcal{J} \subset \mathcal{A}$ $\sigma$-álgebra y $f \in L^1(X, \mathcal{A}, m)$, existe $E_m(f \mid \mathcal{J}) \in L^1(X, \mathcal{J}, m \mid \mathcal{J})$ llamada* esperanza condicional *tal que*

(i) $\int_A f dm = \int_A E_m(f \mid \mathcal{J}) dm\, \forall A \in \mathcal{J}$
(ii) $E_m(1 \mid \mathcal{J}) = 1$
(iii) $E_m(\chi_A \mid \mathcal{J}) = \chi_A\, \forall A \in \mathcal{J}$

DEMOSTRACIÓN. Consideremos la medida definida en $\mathcal{J}$ por

$$\mu_f(A) = \int_A f dm.$$

Entonces $\mu_f \ll m \mid \mathcal{J}$ y así por el Teorema de Radon-Nicodym existe $E_m(f \mid \mathcal{J}) \in L^1(X, \mathcal{J}, m \mid \mathcal{J})$ que satisface (i). □

Cuando $\mathcal{J}$ es una suálgebra finita, sus átomos definen una partición finita $\mathcal{Q}$ y tenemos
$$E_\mu(f \mid \mathcal{J}) = \sum_{q\in\mathcal{Q}} \frac{1}{\mu(Q)}(\int_Q f d\mu)\chi_Q$$

TEOREMA 2.6. *Sea $m \in \mathcal{M}_T(X)$, existe una probabilidad $\mu_m$ en $\mathcal{M}_T(X)$ tal que*

$\mu_m(\mathcal{M}_{erg}(T)) = 1.$

$$\int_X f dm = \int_{\mathcal{M}_{erg}(T)} \left(\int_X f d\nu\right) d\mu_m(\nu).$$

DEMOSTRACIÓN. Sea $\mathcal{J} = \{A \in X \mid T^{-1}A = A\}$ la $\sigma$-álgebra $T$ -invariante. Para $m$ -casi todo $x \in X$ definimos $m_x$ en $(X, \mathcal{A})$ mediante $m_x(A) = E_m(\chi_A \mid \mathcal{J})(x)$. $m_x$ es una probabilidad ergódica, en efecto sea $A \in \mathcal{J}$, entonces
$$m_x(A) = E_m(\chi_A \mid \mathcal{J})(x) = \chi_A(x) = \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A \end{cases}.$$
Además
$$\int_X f dm = \int_X E_m(f \mid \mathcal{J})(x) dm = \int_X \left(\int_X f dm_x\right) dm.$$
Resta interpretar $m$ como una medida en $\mathcal{M}_T(X)$ soportada en $\mathcal{M}_{\text{erg}}$. De hecho podemos definir un encaje $h : X \to M_T(X)$ mediante $h(x) = m_x$ y definir $\mu_m$ en la $\sigma$-álgebra $\mathcal{B} := \{\mathcal{N} \subset \mathcal{M}_T(X) : h^{-1}(\mathcal{N}) \in \mathcal{A}\}$ mediante $\mu_m(\mathcal{N}) = m(h^{-1}(\mathcal{N}))$. □

## 5. Transformaciones mezcladoras

DEFINICIÓN 2.4. Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad, $T : X \hookleftarrow$ preservadora de medida. Decimos que $T$ es *mezcladora* si $\forall A, B \in \mathcal{A}$
$$\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

OBSERVACIÓN 2.1. Las transformaciones mezcladoras son ergódicas porque si $T^{-1}A = A$, entonces $\mu(A^c)\mu(A) = \lim_{n\to\infty} \mu(A^c \cap T^{-n}A) = \mu(A^c \cap A) = 0$.

PROPOSICIÓN 2.7. *Sean $X$ espacio topológico y $\mu$ probabilidad sobre los borelianos tal que $\mu(U) > 0$ $\forall U$ abierto no vacío. Supongamos que $T : X \hookleftarrow$ preserva $\mu$ y es mezcladora. Entonces $T$ es topológicamente mezcladora.*

DEMOSTRACIÓN. Sean $U, V$ abiertos no vacíos, entonces
$$\lim_{n\to\infty} \mu(T^{-n}U \cap V) = \mu(U)\mu(V) > 0.$$
Así que existe $N$ tal que $\mu(T^{-n}U \cap V) > 0$ para $n > N$. Por lo tanto $T^{-n}U \cap V \neq \emptyset$ para $n > N$. □

OBSERVACIÓN 2.2. Si $(X, \mathcal{A}, \mu)$ es espacio de probabilidad y $T : X \hookleftarrow$ preserva $\mu$ tenemos un operador $U_T : L^2(X, \mathcal{A}, \mu) \hookleftarrow$ definido por $U_T(f) = f \circ T$.

PROPOSICIÓN 2.8. *Sea $(X, \mathcal{A}, \mu)$ espacio de probabilidad, la transformación preservadora de medida $T : X \hookleftarrow$ es mezcladora si y solo si $\forall f, g \in L^2(X, \mathcal{A}, \mu)$*

$$\lim_{n\to\infty} \langle U_T^n f, g\rangle = \langle f, 1\rangle \langle 1, g\rangle \tag{2.9}$$

DEMOSTRACIÓN. Si $U_T$ satisface la igualdad (2.9),

$$\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \lim_{n\to\infty} \langle U_T^n \chi_B, \chi_A\rangle = \langle \chi_B, 1\rangle = \mu(B)\mu(A)$$

Reciprocamente si $T$ es mezcladora se tiene que (2.9) se satisface cuando $f, g$ son funciones caracter'ısticas. De la bilinealidad de la igualdad (2.9) se tiene que vale para funciones simples Dadas $f, g \in L^2$ y $\varepsilon > 0$ $\exists$ $f_0, g_0$ funciones simples tales que $\|f - f_0\|_2, \|g - g_0\|_2 < \varepsilon$ y así $\|g_0\|_2 < \|g\|_2 + \varepsilon$

$$\langle U_T^n f, g\rangle = \langle U_T^n (f - f_0), g_0\rangle + \left\langle U_n^T f_0, g_0\right\rangle + \left\langle U_n^T f, g - g_0\right\rangle$$
$$\langle f, 1\rangle \langle 1, g\rangle = \langle f - f_0, 1\rangle \langle 1, g_0\rangle + \langle f_0, 1\rangle \langle 1, g_0\rangle + \langle f, 1\rangle \langle 1, g - g_0\rangle .$$

Así

$$\begin{aligned} &|\langle U_T^n f, g\rangle - \langle f, 1\rangle \langle 1, g\rangle| \\ &\quad \le \left|\left\langle U_n^T f_0, g_0\right\rangle - \langle f_0, 1\rangle \langle 1, g_0\rangle\right| + 2\|f - f_0\|_2 \|g_0\|_2 + 2\|f\|_2 \|g - g_0\|_2 \\ &\qquad\qquad < \varepsilon(1 + 2\|g\|_2 + 2\varepsilon + 2\|f\|_2) \end{aligned}$$

si $n$ es suficientemente grande. □

PROPOSICIÓN 2.9. *Sea $\mathbb{P}$ matriz estocástica $k \times k$ con vector de probabilidad $\boldsymbol{\pi}$ y sea $\mu$ la medida de Markov asociada. Son equivalentes.*

(1) *El corrimiento $\sigma$ es mezclador.*
(2) *$\exists n > 0$ tal que $P_{ij}^n > 0$ $\forall i, j \in [k]$*
(3) *$\lim_{n\to\infty} P_{ij}^n = \pi_j$*

DEMOSTRACIÓN. (1) $\Rightarrow$ (3).

$$\mu(C(0; i) \cap \sigma^{-n} C(0; j)) = \pi_i P_{ij}^n \to \mu(C(0; i))\mu(C(0; j)) = \pi_i \pi_j .$$

(3) $\Rightarrow$ (2). Como $\pi_j > 0\,\forall j \in [k]$, $\exists N \ni P_{ij}^n > 0\,\forall i, j \in [k], n > N$.

(3) $\Rightarrow$ (1). Para demostrar que $\sigma$ es mezcladora es suficiente demostrar que si $A = C(j; i_0, \cdots, i_s)$, $B = C(l; r_0, \ldots, r_m)$ entonces

$$\lim_{n\to\infty} \mu(A \cap \sigma^{-n} B) = \mu(A)\mu(B).$$

Para $n > j + s - l$ se tiene

$$\mu(A \cap \sigma^{-n} B) = \mu(A)\, P_{i_s r_0}^{n+l-(j+s)} \frac{\mu(B)}{\pi_{r_0}}$$

que por (3) converge a $\mu(A)\mu(B)$.

(2) $\Rightarrow$ (3). Por el Corolario 1.5, si $\{\mathbf{e}_j : j \in [k]\}$ es la base canónica de $\mathbb{R}^k$, $\lim_{n\to\infty} \mathbf{e}_j \mathbb{P}^n = \boldsymbol{\pi}$. Pero $\mathbf{e}_j \mathbb{P}^n$ es el $j$-ésimo renglon de $\mathbb{P}^n$. $\square$

PROPOSICIÓN 2.10. *Sean $A$ matriz $n \times n$ con coeficientes enteros y $\tilde{A} : \mathbb{T}^n \hookleftarrow$ la transformación inducida. Entonces $\tilde{A}$ es ergódica si y solo si es mezcladora.*

DEMOSTRACIÓN. De acuerdo a la observación 2.1, toda transformación mezcladora es ergódica.

Si $\tilde{A}$ es ergódica y $k \in \mathbb{Z} \setminus \{0\}$, todos los elementos de la sucesión $\{k^t A^m\}_{m \in \mathbb{N}}$ son distintos entre si. Por lo tanto si $l \in \mathbb{Z}^n$, y $m$ es suficientemente grande $k^t A^m \neq l$ y así $\left\langle U_{\tilde{A}}^m \varphi_k, \varphi_l \right\rangle = \left\langle \varphi_{(A^m)^t k}, \varphi_l \right\rangle = 0$ y $\langle \varphi_k, 1 \rangle \langle 1, \varphi_l \rangle = 0$.

También $\left\langle U_{\tilde{A}}^m 1, \varphi_l \right\rangle = \langle 1, \varphi_l \rangle = \langle 1, 1 \rangle \langle 1, \varphi_l \rangle$.

Por lo tanto, si $f$ y $g$ son combinaciones lineales (finitas) de $\{\varphi_k : k \in \mathbb{Z}^n\}$, $\left\langle U_{\tilde{A}}^m f, g \right\rangle = \langle f, 1 \rangle \langle 1, g \rangle$ si $m$ es suficientemente grande. Como dichas combinaciones lineales forman un conjunto denso en $L^2(\mathbb{T}^n)$, tenemos que $\forall f, g \in L^2(\mathbb{T}^n)$,

$$\lim_{m\to\infty} \left\langle U_{\tilde{A}}^m f, g \right\rangle = \langle f, 1 \rangle \langle 1, g \rangle ,$$

y así $\tilde{A}$ es mezcladora. $\square$

EJERCICIO 2.1. Sea $(X, \mathcal{A}, \mu)$ un espacio de probabilidad y sea $T : X \hookleftarrow$ una transformación preservadora de medida.

a) Probar que $\forall A \in \mathcal{A}$ con $\mu(A) > 0$, $\exists A_0 \in \mathcal{A}$ tal que $A_0 \subset A$, $\mu(A_0) > 0$ y $\forall x \in A_0$, $\tau_A(x) \geq \mu(A)$.

**Sugerencia**: Sea $A_1 = \{x : \tau_A(x) \geq \mu(A)\}$. Probar que $\mu(A_1) > 0$ y que $A_1 = \bigcup_{n\geq 0} T^{-n}(A_1 \cap A)$.

b) Probar que $\tau_A(x) > 0$ para casi todo $x \in A$.

c) Si $X$ es un espacio métrico separable y $\mathcal{A}$ es la $\sigma$ -álgebra de Borel probar que para casi todo $x \in X$ se tiene que $\tau_U(x) > 0$ para toda vecindad $U$ de x.

EJERCICIO 2.2. Sea $(X, \mathcal{A}, \mu)$ un espacio de probabilidad y sea $T : X \hookleftarrow$ una transformación preservadora de medida.

a) Probar que si $f : X \to (0, \infty)$ es medible entonces para casi todo $x \in X$

$$\liminf_{n\to\infty} \frac{1}{n} f(T^n(x)) = 0.$$

**Sugerencia**: Si la propiedad es falsa existen $A \in \mathcal{A}, K_1, K_2 > 0$ tales que $\mu(A) > 0$ y si $x \in A$

$$\liminf_{n\to\infty} \frac{1}{n} f(T^n(x)) \geq K_1, \quad f(x) \leq K_2.$$

Aplicando el teorema de Poincaré obtener una contradicción entre las dos desigualdades.

b) Si $f : X \to (0, \infty)$ es medible y $f \circ T - f$ es integrable probar que

$$\lim_{n\to\infty} \frac{1}{n} f(T^n(x)) = 0$$

**Sugerencia**: Aplicar el teorema ergódico a $f \circ T - f$.

Ejercicio 2.3. Sea $(X, \mathcal{A}, \mu)$ un espacio de probabilidad y sea $T : X \hookleftarrow$ una transformación preservadora de medida. Decimos que $T$ es debilmente mezcladora si para todo $A, B \in \mathcal{A}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} |\mu(A \cap T^{-j}B) - \mu(A)\mu(B)| = 0$$

a) Probar que si $T$ es debilmente mezcladora, entonces es ergódica y que si $T$ es mezcladora entonces es debilmente mezcladora.

b) Probar que si $T$ es debilmente mezcladora entonces $T^n$ también lo es.

c) Probar que el producto cartesiano de una transformación ergódica y una debilmente mezcladora es ergódica.

d) Sea $R_\alpha$ la rotación del c'ırculo por el ángulo $\alpha$. Demostrar que $R_\alpha \times R_\alpha$ no es ergódica y por lo tanto $R_\alpha$ no es debilmente mezcladora.

c) Probar que si $T \times T$ es ergódica entonces $T$ es debilmente mezcladora.

e) Probar que el producto cartesiano de dos transformaciones debilmente mezcladoras es debilmente mezcladora.

f) Probar que $T$ es debilmente mezcladora si y solo si para toda $f \in L^2$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\langle U_T^i f, f\rangle - |\langle f, 1\rangle|^2| = 0.$$

g) Probar que $T$ es debilmente mezcladora si y solo si el único eigenvalor de $U_T$ es 1.

Ejercicio 2.4. Sea $A$ matriz $n \times n$ con coeficientes enteros y $\det A \neq 0$. Sea $\tilde{A}$ la transformación inducida en $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Sea $\pi : \mathbb{R}^n \to \mathbb{T}^n$ la proyección natural. Considere la rotación $R_{\pi(a)}, a \in \mathbb{R}^n$. Sea $T = R_{\pi(a)} \circ \tilde{A}$.

a) Probar que $T$ es ergódica si y solo si

$$\forall k \in \mathbb{Z}^n - \{0\}, m \in \mathbb{N} \quad kA^m = k \Rightarrow m\langle k, a\rangle \notin \mathbb{Z}$$

b) Probar que son equivalentes

(i) $T$ es mezcladora

(ii) $T$ es debilmente mezcladora

(iii) $\tilde{A}$ es ergódica.

**Sugerencias**: Recuerde la prueba de (iii) $\Rightarrow$ $\tilde{A}$ es mezcladora. Para (ii) $\Rightarrow$ (iii) suponga que $\tilde{A}$ no es ergódica y utilice el ejercicio 2.3f) para probar que $T$ no es debilmente mezcladora.

Capítulo 3

# Entropía

## 1. Entropía con respecto a una medida

Definiciones 3.1. Sea $(X, \mathcal{A}, \mu)$ un espacio de probabilidad.

(a) Decimos que $A \subset B$ mod (0), si $\mu(A \setminus B) = 0$.

(b) Una *partición* de $(X, \mathcal{A}, \mu)$ es una familia $\mathcal{P} \subset \mathcal{A}$ de conjuntos de medida positiva tales que $X = \bigcup \mathcal{P}$ mod(0) y $\mu(A \cap B) = 0$ si $A, B \in \mathcal{P}, A \neq B$. En tal caso $\mathcal{P}$ es una familia contable. Los elementos de $\mathcal{P}$ se llaman *átomos*.

(c) Si $\mathcal{P}$ y $\mathcal{Q}$ son particiones decimos que $\mathcal{Q}$ es un *refinamiento* de $\mathcal{P}$ lo que denotamos $\mathcal{P} \prec \mathcal{Q}$ si todo átomo de $\mathcal{Q}$ esta contenido en un átomo de $\mathcal{P}$ mod (0). Esto implica que todo átomo de $\mathcal{P}$ es la unión de átomos de $\mathcal{Q}$ mod (0).

Definiciones 3.2. Sean $\mathcal{P}, \mathcal{Q}$ particiones finitas del espacio de probabilidad $(X, \mathcal{A}, \mu)$. Sea $\phi(x) = x \log x$ si $0 < x \leq 1$, $\phi(0) = 0$.

(a) Definimos la *entropí* a de $\mathcal{P}$ como

$$H(\mathcal{P}) = H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \phi(\mu(P))$$

(b) Definimos la *entropí a condicional* de $\mathcal{P}$ dado $\mathcal{Q}$ como

$$H_\mu(\mathcal{P} \mid \mathcal{Q}) = -\sum_{P \in \mathcal{P}, Q \in \mathcal{Q}} \mu(Q)\, \phi\left(\frac{\mu(P \cap Q)}{\mu(Q)}\right).$$

Observaciones 3.1.

(a) $\phi$ es convexa y $\phi(x) < 0$ si $0 < x < 1$.

(b) Si $\mathbb{1}$ es la partición trivial $\{X\}$ entonces $H(\mathcal{P} \mid \mathbb{1}) = H(\mathcal{P})$

(c) Una definición más general se obtiene considerando una sub $\sigma$-álgebra $\mathcal{J}$ de $\mathcal{A}$. Se introduce la *información condicional* de $\mathcal{P}$ dada $\mathcal{J}$

$$I_\mu(\mathcal{P} \mid \mathcal{J}) = -\sum_{P \in \mathcal{P}} \chi_P \log E_\mu(\chi_P \mid \mathcal{J})$$

y se define la *entrop'ıa condicional* de $\mathcal{P}$ dada $\mathcal{J}$ como

$$H_\mu(\mathcal{P} \mid \mathcal{J}) = \int I_\mu(P \mid \mathcal{J}) d\mu = \int -\sum_{P \in \mathcal{P}} \phi \circ E_\mu(\chi_P \mid \mathcal{J}) d\mu.$$

Para verificar la segunda igualdad observamos primero que para cualquier $A \in \mathcal{J}$ se tiene

$$\int \chi_P \chi_A d\mu = \int E_\mu(\chi_P \mid \mathcal{J})\chi_A d\mu,$$

de donde se sigue de que para cualquier $f \in L^1(X, \mathcal{J}, \mu)$

$$\int \chi_P f = \int E_\mu(\chi_P \mid \mathcal{J}) f d\mu.$$

Tomando $\mathcal{J} = \mathcal{J}(\mathcal{Q})$ la $\sigma$-álgebra generada por $\mathcal{Q}$, tenemos

$$\begin{aligned} E_\mu(\chi_P \mid \mathcal{J}) &= \sum_{q\in\mathcal{Q}} \frac{\mu(P \cap Q)}{\mu(Q)} \chi_Q, \\ \phi \circ E_\mu(\chi_P \mid \mathcal{J}) &= \sum_{q\in\mathcal{Q}} \phi\Big(\frac{\mu(P \cap Q)}{\mu(Q)}\Big) \chi_Q, \\ H_\mu(\mathcal{P} \mid \mathcal{J}) &= H_\mu(\mathcal{P} \mid \mathcal{Q}) \end{aligned}$$

Observación 3.2. De la convexidad de $\phi$ se demuestra por inducción que

$$\sum_i \alpha_i = 1,\ x_i \in [0,1] \Rightarrow \phi\Big(\sum_i \alpha_i x_i\Big) \leq \sum_i \alpha_i \phi(x_i) \tag{3.1}$$

Sea $\mathcal{P} = \{P_1, \ldots, P_k\}$ una partición, poniendo $\alpha_i = 1/k$ tenemos

$$-\frac{1}{k}\log k = \phi(\frac{1}{k}) = \phi\Big(\sum_{i=1}^k \frac{1}{k}\mu(P_i)\Big) \leq \frac{1}{k}\sum_{i=1} \phi(\mu(P_i)) = \frac{1}{k} H_\mu(\mathcal{P})$$

y así $H_\mu(\mathcal{P}) \leq k$.

Proposición 3.1. *Sean $\mathcal{P}, \mathcal{Q}, \mathcal{M}$ particiones de $(X, \mathcal{A}, \mu)$. Denotamos por $\mathcal{P} \vee \mathcal{Q}$ la partición cuyos átomos son los conjuntos $P \cap Q$ con medida positiva y $P \in \mathcal{P}, Q \in \mathcal{Q}$. Entonces*

(a) $H(\mathcal{P} \vee \mathcal{Q} \mid \mathcal{M}) = H(\mathcal{P} \mid \mathcal{M}) + H(\mathcal{Q} \mid \mathcal{M} \vee \mathcal{P})$
(b) *Si* $\mathcal{P} \prec \mathcal{Q}$ ,$H(\mathcal{M} \mid \mathcal{P}) \geq H(\mathcal{M} \mid \mathcal{Q})$ *y* $H(\mathcal{P} \mid \mathcal{M}) \leq H(\mathcal{Q} \mid \mathcal{M})$
(c) $H(\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q} \mid \mathcal{P})$
(d) $H(\mathcal{P} \vee \mathcal{Q} \mid \mathcal{M}) \leq H(\mathcal{P} \mid \mathcal{M}) + H(\mathcal{Q} \mid \mathcal{M})$
(e) *Si* $T : X \hookleftarrow$ *preserva* $\mu$ *entonces* $H(T^{-1}\mathcal{P} \mid T^{-1}\mathcal{Q}) = H(\mathcal{P} \mid \mathcal{Q})$
(f) $H(\mathcal{P} \mid \mathcal{Q}) = 0 \Leftrightarrow \mathcal{P} \prec \mathcal{Q}$

Demostración. (a)

$$\begin{aligned}
H(\mathcal{P}\vee\mathcal{Q}\mid\mathcal{M}) &= -\sum_{P,Q,M}\mu(P\cap Q\cap M)\,\log\frac{\mu(P\cap Q\cap M)}{\mu(M)}\\
&= -\sum_{P,Q,M}\mu(P\cap Q\cap M)\,\log\frac{\mu(P\cap Q\cap M)}{\mu(P\cap M)}\\
&\quad -\sum_{P,Q,M}\mu(P\cap Q\cap M)\,\log\frac{\mu(P\cap M)}{\mu(M)}\\
&=H(\mathcal{Q}\mid\mathcal{P}\vee\mathcal{M})-\sum_{P,M}\mu(P\cap M)\,\log\frac{\mu(P\cap M)}{\mu(M)}\\
&=H(\mathcal{Q}\mid\mathcal{P}\vee\mathcal{M})+H(\mathcal{P}\mid\mathcal{M})
\end{aligned}$$

(b). Si $\mathcal{P}\prec\mathcal{Q}$ entonces $\mathcal{P}\vee\mathcal{Q}=\mathcal{Q}$. Así

$$H(\mathcal{Q}\mid\mathcal{M})=H(\mathcal{P}\mid\mathcal{M})+H(\mathcal{Q}\mid\mathcal{P}\vee\mathcal{M})\geq H(\mathcal{P}\mid\mathcal{M}).$$

Por (3.1) tenemos

$$\begin{aligned}
H(\mathcal{M}\mid\mathcal{Q}) &= -\sum_{M,Q}\mu(M\cap Q)\,\log\frac{\mu(M\cap Q)}{\mu(Q)}\\
&= -\sum_{M,P}\mu(P)\sum_{Q\subset P}\frac{\mu(Q)}{\mu(P)}\,\phi\left(\frac{\mu(M\cap Q)}{\mu(Q)}\right)\leq -\sum_{M,P}\mu(P)\,\phi\left(\sum_{Q\subset P}\frac{\mu(M\cap Q)}{\mu(P)}\right)\\
&= -\sum_{M,P}\mu(P)\,\phi\left(\frac{\mu(M\cap P)}{\mu(P)}\right)=H(\mathcal{M}\mid\mathcal{P})
\end{aligned}$$

(c). Poniendo $\mathcal{M}=\mathbb{1}$ tenemos por (a) y (b)

$$H(\mathcal{Q})\leq H(\mathcal{Q}\vee\mathcal{P})=H(\mathcal{P})+H(\mathcal{Q}\mid\mathcal{P}).$$

(d).

$$\begin{aligned}
H(\mathcal{P}\vee\mathcal{Q}\mid\mathcal{M}) &= H(\mathcal{P}\mid\mathcal{M})+H(\mathcal{Q}\mid\mathcal{P}\vee\mathcal{M})\\
&\leq H(\mathcal{P}\mid\mathcal{M})+H(\mathcal{Q}\mid\mathcal{M})
\end{aligned}$$

(e). Inmediato

(f). $H(\mathcal{P}\mid\mathcal{Q})=-\sum_{P,Q}\mu(Q)\phi\left(\dfrac{\mu(P\cap Q)}{\mu(Q)}\right)=0$ si y solo si $\forall P,Q$ se tiene $\mu(P\cap Q)=\mu(Q)$ ó 0; es decir $\mathcal{P}\prec\mathcal{Q}$. □

Definición 3.3. Supongamos que $T:X\hookleftarrow$ preserva medida y sea $\mathcal{P}$ una partición, definimos la entropía $h(T,\mathcal{P})$ de $T$ respecto a $\mathcal{P}$ como

$$h(T,\mathcal{P})=\lim_{n\to\infty}\frac{1}{n}H(\mathcal{P}\vee T^{-1}\mathcal{P}\vee\cdots\vee T^{-(n-1)}\mathcal{P}).$$

Para ver que este límite existe utilizaremos el siguiente lema.

LEMA 3.1. *Sea $\{a_n\}$ sucesión de reales positivos tales que $a_{n+m} \le a_n + a_m$. Entonces*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_n \left(\frac{a_n}{n}\right).$$

DEMOSTRACIÓN. Sea $C = \inf (a_n/n)$. Dado $\varepsilon > 0$ $\exists N$ tal que $a_N/N \le C + \varepsilon$. Si $n > N$ escribimos $n = pN + q$ con $q = 1, \dots, N$.

$$\frac{a_n}{n} = \frac{a_{pN+q}}{pN+q} \le \frac{pa_N + a_q}{pN+q} \le \frac{a_N}{N} + \frac{a_q}{n} \le C + \varepsilon + \frac{1}{n}\sup\{a_1, \dots, a_N\}.$$

Si $n$ es suficientemente grande, $C \le a_n/n \le C + 2\varepsilon$. □

Pongamos ahora $a_n = H(\mathcal{P} \vee T^{-1}\mathcal{P} \vee \cdots \vee T^{-(n-1)}\mathcal{P})$, entonces

$$\begin{aligned} a_{n+m} &= H(\mathcal{P} \vee T^{-1}\mathcal{P} \vee \cdots \vee T^{-(n-1)}\mathcal{P} \vee T^{-n}\mathcal{P} \vee \cdots \vee T^{-(n+m-1)}\mathcal{P}) \\ &\le H(\mathcal{P} \vee \cdots \vee T^{-(n-1)}\mathcal{P}) + H(T^{-n}(\mathcal{P} \vee \cdots \vee T^{-(m-1)}P)) \\ &= a_n + a_m. \end{aligned}$$

PROPOSICIÓN 3.2.
(a) $h(T,\mathcal{P}) - h(T,\mathcal{Q}) \le H(\mathcal{P} \mid \mathcal{Q})$.
(b) *Si* $\mathcal{P} \prec \mathcal{Q}$ *entonces* $h(T,\mathcal{P}) \le h(T,\mathcal{Q})$.
(c) $h(T,T^{-1}\mathcal{P}) = h(T,\mathcal{P})$.
(d) $\forall n \ge 0$, $h(T,\mathcal{P} \vee T^{-1}\mathcal{P} \vee \cdots \vee T^{-n}\mathcal{P}) = h(T,\mathcal{P})$.
(e) $h(T,\mathcal{P}) = \lim_{n\to\infty} H(\mathcal{P} \mid T^{-1}\mathcal{P} \vee \cdots \vee T^{-n}P)$.
(f) *La sucesión* $\left\{\frac{1}{n}H(\mathcal{P} \vee \cdots \vee T^{-(n-1)}\mathcal{P})\right\}$ *es decreciente.*

DEMOSTRACIÓN. (a).

$$h(T,\mathcal{P}) - h(T,\mathcal{Q}) = \lim_{n\to\infty} \frac{1}{n}(H(\bigvee_{i=0}^{n-1} T^{-i}P) - H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q})).$$

Pero

$$\begin{aligned} H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) - H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}) &\le H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \mid \bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}) \\ &\le \sum_{i=0}^{n-1} H(T^{-i}\mathcal{P} \mid \bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) \le \sum_{i=0}^{n-1} H(T^{-i}\mathcal{P} \mid T^{-i}\mathcal{Q}) = nH(\mathcal{P} \mid \mathcal{Q}) \end{aligned} \tag{3.2}$$

(b). Si $\mathcal{P} \prec \mathcal{Q}$, $H(\mathcal{P} \mid \mathcal{Q}) = 0$ y entonces $h(T,\mathcal{P}) - h(T,\mathcal{Q}) \le 0$.
(c). Inmediato.

$$\text{(d).}\quad h(T, \bigvee_{i=0}^{n} T^{-i}\mathcal{P}) = \lim_{m\to\infty} \frac{1}{m} H(\bigvee_{j=0}^{m-1} T^{-j}\big(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}\big))$$

$$= \lim_{m\to\infty} \frac{1}{m} H(\bigvee_{j=0}^{m+n-1} T^{-j}\mathcal{P}) = \lim_{m\to\infty} \frac{m+n-1}{m} \frac{1}{m+n-1} H(\bigvee_{j=0}^{m+n-1} T^{-j}\mathcal{P})$$

$$= h(T, P)$$

(e). La sucesión $\left\{H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P})\right\}$ es decreciente. Sea $c$ su límite.

$$\begin{aligned}
H(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}) &= H(\bigvee_{i=1}^{n} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P}) \\
&= H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P}) \\
&= H(\bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P}) \\
&= H(\bigvee_{i=0}^{n-2} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}) + H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P}) \\
\cdots &= H(\mathcal{P}) + \sum_{k=1}^{n} H(\mathcal{P} \mid \bigvee_{i=1}^{k} T^{-i}\mathcal{P})
\end{aligned}$$

Por el teorema de Cesaro

$$h(T, \mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} H(\mathcal{P} \mid \bigvee_{i=1}^{k} T^{-i}\mathcal{P}) = c$$

(f). De la prueba de (e) tenemos

$$(n+1)H(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}) = (n+1)H(\bigvee_{i=1}^{n} T^{-i}\mathcal{P}) + (n+1)H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P})$$

y

$$\begin{aligned}
H(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}) &= H(\mathcal{P}) + \sum_{k=1}^{n} H(\mathcal{P} \mid \bigvee_{i=1}^{k} T^{-i}\mathcal{P}) \\
&\geq (n+1)H(\mathcal{P} \mid \bigvee_{i=1}^{n} T^{-i}\mathcal{P})
\end{aligned}$$

Así

$$(n+1)H(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}) \leq (n+1)H(\bigvee_{i=1}^{n} T^{-i}\mathcal{P}) + H(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}),$$

y entonces

$$nH(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}) \leq (n+1)H(\bigvee_{i=1}^{n} T^{-i}\mathcal{P}) = (n+1)H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P})$$

□

Definición 3.4. La entropía de una transformación $T : (X, \mathcal{A}, \mu) \hookleftarrow$ que preserva medida se define como

$$h(T) = h_\mu(T) = \sup\{h(T, \mathcal{P}) : \mathcal{P} \text{ es partición finita}\}$$

Observación 3.3. $h(T) = \sup\{h(T, \mathcal{P}) : H(\mathcal{P}) < \infty\}$.

En efecto, sea $\mathcal{P} = \{P_1, P_2, \cdots\}$ con $H(\mathcal{P}) < \infty$.

Consideremos $\mathcal{P}^{(n)} = \{P_1, \cdots, P_n, \underset{j>n}{\cup} P_j\}$. Observemos que para $1 \leq i, j \leq n$, $\dfrac{\mu(P_i \cap P_j)}{\mu(P_i)} = 0, 1$. Entonces

$$\begin{aligned}
0 &\leq h(T, \mathcal{P}) - h(T, \mathcal{P}^{(n)}) \leq H(\mathcal{P} \mid \mathcal{P}^{(n)}) \\
&= -\sum_{i,j=1}^{n} \mu(p_j) \log\left(\frac{\mu(P_i \cap P_j)}{\mu(P_i)}\right) - \sum_{j>n} \mu(P_j)\ \log \frac{\mu(P_j)}{\mu(\underset{k>n}{\cup} P_k)} \\
&= -\sum_{j>n} \phi(\mu(P_j)) + (\sum_{j>n} \mu(P_j))\ \log \mu(\underset{k>n}{\cup} P_k) \\
&= -\sum_{j>n} \phi(\mu(P_j)) + \phi(\mu(\underset{k>n}{\cup} P_k))
\end{aligned}$$

Como $H(\mathcal{P}) = -\sum_j \phi(\mu(P_j)) < \infty$, tenemos $\lim\limits_{n\to\infty} \sum\limits_{j>n} \phi(\mu(P_j)) = 0$. Como $\lim\limits_{n\to\infty} \mu(\underset{j>n}{\cup} P_j) = 0$, también $\lim\limits_{n\to\infty} \phi(\mu \underset{j>n}{\cup} P_j)) = 0$. Y resulta que $\lim\limits_{n\to\infty} h(T, \mathcal{P}^{(n)}) = h(T, \mathcal{P})$.

Definición 3.5. Si $\forall n \in \mathbb{N}$ $\mathcal{A}_n$ es una familia de subconjuntos de $X$ , $\bigvee\limits_{n=1}^{\infty} \mathcal{A}_n$ denota la $\sigma$-álgebra generada por $\bigcup_{n=1}^{\infty} \mathcal{A}_n$.

Teorema 3.1. *Sean $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots$ particiones y sea $\mathcal{P}$ una partición con $H(\mathcal{P}) < \infty$. Entonces $\mathcal{P} \subset \bigvee\limits_{i=1}^{\infty} \mathcal{P}_i$(mod 0) si y solo si $\lim\limits_{n\to\infty} H(\mathcal{P} \mid \mathcal{P}_n) = 0$.*

No demostraremos este teorema.

Corolario 3.2. *Si* $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots$ *es una sucesión de particiones con entropía finita tales que* $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$*, entonces* $h(T) = \sup_n h(T, \mathcal{P}_n)$

Demostración. Sea $\mathcal{P}$ una partición finita. Por el teorema y la hipótesis $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$, tenemos que $\lim_{n\to\infty} H(\mathcal{P} \mid \mathcal{P}_n) = 0$. Como $h(T, \mathcal{P}) \leq h(T, \mathcal{P}_n) + H(\mathcal{P} \mid \mathcal{P}_n)$, se tiene $h(T, \mathcal{P}) \leq \sup_n h(T, \mathcal{P}_n)$. □

Definición 3.6. Sea $(X, \mathcal{A}, \mu)$ espacio de probabilidad. Supongamos que $T : X \hookleftarrow$ es invertible y preserva $\mu$. Una partición $\mathcal{P}$ con $H(\mathcal{P}) < \infty$ se llama $T$-generador si $\bigvee_{n=-\infty}^{\infty} T^n\mathcal{P} = \mathcal{A}$.

Corolario (Kolmogorov - Sinai). *Si* $\mathcal{P}$ *es un* $T-$*generador entonces* $h(T, \mathcal{P}) = h(T)$.

Demostración.

$$h(T) = \sup_n h(T, \bigvee_{j=-n}^{n} T^{-j}\mathcal{P}) = \sup_n h(T, T^n(\bigvee_{j=0}^{2n} T^{-j}\mathcal{P}))$$
$$= \sup_n h(T, \bigvee_{j=0}^{2n} T^{-j}\mathcal{P}) = h(T, \mathcal{P}).$$

□

Para $T$ no invertible se tiene

Corolario 3.3. *Si* $\mathcal{P}$ *es una partición tal que* $H(\mathcal{P}) < \infty$ *y* $\bigvee_{n=0}^{\infty} T^{-n}\mathcal{P} = \mathcal{A}$*, entonces* $h(T, \mathcal{P}) = h(T)$.

Demostración. $h(T) = \sup_{n\geq 0} h(T, \bigvee_{j=0}^{n} T^{-j}\mathcal{P}) = h(T, \mathcal{P})$. □

Corolario 3.4. *Supongamos existe una partición* $\mathcal{P}$ *tal que* $H(\mathcal{P}) < \infty$ *y* $\bigvee_{n=1}^{\infty} T^{-n}\mathcal{P} = \mathcal{A}$*, entonces* $h(T) = 0$.

Demostración. Como $\bigvee_{n=0}^{\infty} T^{-n}\mathcal{P} = \mathcal{A}$, tenemos que $h(T, \mathcal{P}) = h(T)$.

$$h(T, \mathcal{P}) = \lim_{m\to\infty} H(\mathcal{P} \mid T^{-1}\mathcal{P} \vee \cdots \vee T^{-n}\mathcal{P}).$$

Como $\mathcal{P} \subset \bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}$, tenemos que $h(T) = h(T, \mathcal{P}) = 0$. □

Denotaremos por $\mathcal{P}_n(x)$ al elemento de $\mathcal{P}_n$ que contiene un punto dado $x \in X$.

Corolario 3.5. *Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua. Si $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots$ es una sucesión de particiones con entropía finita tales que* $\operatorname{diam} \mathcal{P}_n(x)$ *converge a cero para $\mu$-casi todo $x \in X$. Entonces*

$$h(f) = \lim_{n\to\infty} h(f, \mathcal{P}_n).$$

Demostración. Sea $U$ un abierto de $X$. La hip'tesis garantiza que para cada $x$ existe $n(x)$ tal que el conjunto $P_x = \mathcal{P}_{n(x)}(x)$ está contenido en $U$. Es claro que $P_x$ pertence al álgebra $\mathcal{A}$ generada por $\bigcup_{n=1}^{\infty} \mathcal{P}_n$. Observe también que esta álgebra es numerable, ya que está formada por las uniones finitas de elementos de las particiones $\mathcal{P}_n$. En particular, el conjunto de los valores tomados por $P_x$ es numerable. Se sigue que $U = \bigcup_{x\in U} P_x$ pertenece a$\mathcal{A}$. Esto prueba que la $\sigma$-álgebra de Borel esta contenida en $\bigvee_{n=1}^{\infty} \mathcal{P}_n$. Aplicamos ahora el Corolario 3.2 $\square$

Proposición 3.3. *$h(T^m) = mh(T)$ $\forall m \geq 0$, y si $T$ es invertible*
*$h(T^m) = |m|\, h(T)$ $\forall m \in \mathbb{Z}$*

Demostración. Sea $\mathcal{P}$ una partición de $X$.

$$\begin{aligned} h(T, \mathcal{P}) &= \lim_{n\to\infty} \frac{1}{nm} H(\bigvee_{j=0}^{nm-1} T^{-j}\mathcal{P}) \\ &\geq \lim_{n\to\infty} \frac{1}{nm} H(\bigvee_{j=0}^{n-1} T^{-jm}\mathcal{P}) = \frac{1}{m} h(T^m, \mathcal{P}). \end{aligned}$$

Así $h(T^m) \leq mh(T)$. También

$$\begin{aligned} h(T^m, \bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}) &= \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-im}(\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P})) \\ &= \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{j=0}^{nm-1} T^{-j}\mathcal{P}) = mh(T, P). \end{aligned}$$

Si $T$ es invertible

$$\begin{aligned} h(T, \mathcal{P}) &= \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} H(T^{-(n-1)} \bigvee_{i=0}^{n-1} T^{i}\mathcal{P}) \\ &= \lim_{n\to\infty} H(\bigvee_{i=0}^{n-1} T^{i}\mathcal{P}) = h(T^{-1}, \mathcal{P}). \end{aligned}$$

$\square$

Ejemplo 3.1. Sea $T : \mathbb{S}^1 \hookleftarrow$ definido por $T(z) = z^n$. Tomemos la partición $\mathcal{P} = \{P_1, \dots, P_n\}$ donde

$$P_j = \left\{ z \in \mathbb{S}^1 : \frac{2\pi(j-1)}{n} < \arg(z) < \frac{2\pi j}{n} \right\}.$$

Los atomos de $\mathcal{P} \vee \cdots \vee T^{-(m-1)}P$ son los intervalos.

$$I_j^m = \left\{ z \in \mathbb{S}^1 : \frac{2\pi(j-1)}{n^m} < \arg(z) < \frac{2\pi j}{n^m} \right\},$$

con $\mu(I_j^m) = 1/n^m$. Entonces $\bigvee_{j=0}^{\infty} T^{-j}\mathcal{P}$ es la $\sigma$-álgebra de Borel. Así $h(T) = h(T, \mathcal{P})$.

$$h(T,\mathcal{P}) = \lim_{m\to\infty} \frac{1}{m} H(\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}) = \lim_{m\to\infty} \frac{-1}{m} \log(\frac{1}{n^m}) = \log n.$$

Ejemplo 3.2. Sea $T : \mathbb{S}^1 \hookleftarrow$ la rotación $T(z) = \alpha z$.

(1) Si $\arg(\alpha) \notin \pi\mathbb{Q}$, consideremos una partición formada por 2 intervalos $(a_1, a_2)(a_2, a_1)$. Entonces los atomos de $\bigvee_{j=0}^{n} T^{-j}\mathcal{P}$ son los intervalos con extremos $T^{-j}a_1, T^{-i}a_2$. Como las orbitas $\{T^{-j}a_1\}$, $\{T^{-j}a_2\}$ son densas, $\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}$ es la $\sigma$-álgebra de Borel de $\mathbb{S}^1$. Por el Corolario 3.4 $h(T) = 0$.

(2) Si $\arg(\alpha) \in \pi\mathbb{Q}$, existe $m$ tal que $T^m = I$ y así $mh(T) = h(T^m) = 0$.

Ejemplo 3.3. Sea $\sigma : X \hookleftarrow$ el corrimiento con vector de probabilidad $(p_0, \dots, p_{k-1})$. Sea $\mathcal{P} = \{C(0;0), \dots, C(0;k-1)\}$ la partición al tiempo cero. Los átomos de $\bigvee_{i=j}^{j+m} \sigma^{-i}\mathcal{P}$ son los cilindros $C(j; i_0, \dots, i_n)$ $i_0, \dots, i_n \in [k]$. Por lo tanto $\mathcal{P}$ es un generador.

$$\begin{aligned}
h(\sigma,\mathcal{P}) &= \lim_{m\to\infty} \frac{1}{m} H(\mathcal{P} \vee \cdots \vee \sigma^{-(m-1)}\mathcal{P}) \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{i_0,\dots,i_{m-1}} p_{i_0} \dots p_{i_{m-1}} \log(p_{i_0} \cdots p_{i_{m-1}}) \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{i_0,\dots,i_{m-1}} p_{i_0} \cdots p_{i_{m-1}} \log p_{i_0} + \cdots + p_{i_0} \cdots p_{i_{m-1}} \log p_{i_{m-1}} \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{j=0}^{m-1} \sum_{i_j} p_{i_j} \log p_{i_j} = -\sum_{i\in[k]} p_i \log p_i
\end{aligned}$$

EJEMPLO 3.4. Sea $\mathbb{P}$ matriz estocástica $k \times k$ con vector de probabilidad $\boldsymbol{\pi}$ y sea $\mu$ la medida de Markov asociada. Sean $\sigma$ y $\mathcal{P}$ como en el ejemplo anterior

$$\begin{aligned}
h(\sigma, \mathcal{P}) &= \lim_{m\to\infty} \frac{1}{m} H(\mathcal{P} \vee \cdots \vee \sigma^{-(m-1)}\mathcal{P}) \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{i_1,\ldots,i_m} \pi_{i_1} P_{i_1 i_2} \cdots P_{i_{m-1} i_m} \log(\pi_{i_1} P_{i_1 i_2} \cdots P_{i_{m-1} i_m}) \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{i_1} \pi_{i_1} \log \pi_i \\
&\quad -\lim_{m\to\infty} \frac{1}{m} \sum_{j=1}^{m-1} \sum_{i_j i_{j+1}} \sum_{i_1,\ldots,\hat{i}_j,\hat{i}_{j+1},\ldots,i_{m-1}} \pi_{i_1} P_{i_1 i_2} \cdots P_{i_{m-1} i_m} \log P_{i_j i_{j+1}} \\
&= -\lim_{m\to\infty} \frac{1}{m} \sum_{j=1}^{m-1} \sum_{i_j i_{j+1}} \pi_{i_j} P_{i_j i_{j+1}} \log P_{i_j i_{j+1}} \\
&= -\lim_{m\to\infty} \frac{1}{m} m \sum_{i,j} \pi_i P_{ij} \log P_{ij} = -\sum_{i,j\in[k]} \pi_i P_{ij} \log P_{ij}
\end{aligned}$$

## 2. Entropía Topológica

DEFINICIÓN 3.7. Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua, decimos que $S \subset X$ es un $(n, \varepsilon)$ *generador para* $T$ si $\forall x \in X\ \exists y \in S$ tal que $d(T^j x, T^j y) \leq \varepsilon$ para $0 \leq j \leq n$.

OBSERVACIÓN 3.4. Sea $\{U_1, \ldots, U_m\}$ una cubierta de $X$ por conjuntos de diametro $\leq \varepsilon$ y en cada conjunto no vac'ıo de la forma $\bigcap_{j=0}^{n} T^{-j}(U_{i_j})$ escojamos un punto. Con esto formamos un conjunto $S$ $(n, \varepsilon)$ generador con a lo más $m^{n+1}$ elementos.

DEFINICIÓN 3.8. Sea $r_T(n, \varepsilon) = \min\{\#S : S \text{ es } (n, \varepsilon) \text{ generador}\}$. Por la observación 3.4, $r_T(n, \varepsilon) \leq m^{n+1}$, así que

$$r_T(\varepsilon) = \limsup_{n\to\infty} \frac{1}{n} \log r_T(n, \varepsilon) < \infty.$$

Definimos la *entrop'ıa topológica* de $T$ como

$$h_{\text{top}}(T) = \lim_{\varepsilon\to 0} r_T(\varepsilon)$$

Como $r_T(\varepsilon)$ es decreciente, el límite anterior existe aunque puede ser infinito.

DEFINICIÓN 3.9. Decimos que $S \subset X$ es $(n, \varepsilon)$ *separado respecto a* $T$ si $\forall x, y \in S, x \neq y\ \exists j \in \{0, \ldots, n\}$ tal que $d(T^j x, T^j y) > \varepsilon$.

PROPOSICIÓN 3.4. *Sea* $s_T(n, \varepsilon) = \max\{\#S : S \text{ es } (n, \varepsilon) \text{ separado}\}$. *Entonces* $r_T(n, \varepsilon) \leq s_T(n, \varepsilon) \leq r_T(n, \varepsilon/2)$.

Demostración. Sea $E$ un conjunto $(n, \varepsilon)$ separado y sea $F$ un conjunto $(n, \varepsilon/2)$ generador definimos $\phi : E \to F$ de la siguiente manera, para $x \in E$ escogemos $\phi(x)$ tal que $d(T^j x, T^j \phi(x)) \leq \varepsilon/2$ para $0 \leq j \leq n$. Si $\phi(x) = \phi(y)$ entonces $d(T^j x, T^j y) \leq \varepsilon$ para $0 \leq j \leq n$ y as'ı $x = y$. Por lo tanto $\#E \leq \#F$ y así

$$s_T(n, \varepsilon) \leq r_T(n, \varepsilon/2).$$

Sea $S$ conjunto $(n, \varepsilon)$ separado de cardinalidad máxima. Si $y \notin S$ entonces $S \cup \{y\}$ no es $(n, \varepsilon)$ separado y así $\exists x \in S$ tal que $\forall j \in \{0, \dots, n\}$ $d(T^j x, T^j y) \leq \varepsilon$, o sea que $S$ es $(n, \varepsilon)$ generador y entonces

$$r_T(n, \varepsilon) \leq s_T(n, \varepsilon).$$

□

Se sigue de la Proposición 3.4 que

$$h_{\text{top}}(T) = \lim_{\varepsilon \to 0} s_T(\varepsilon) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s_T(n, \varepsilon).$$

Proposición 3.5. *Para cualquier* $\varepsilon > 0$

$$r_T(2\varepsilon) \leq \inf_n \frac{1}{n} \log r_T(n, \varepsilon), \tag{3.3}$$

*y por consiguiente*

$$h_{top}(T) = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log r_T(n, \varepsilon).$$

Utilizaremos el siguiente lema en la demostración.

Lema 3.2. *Sean* $n_1, \dots, n_N \in \mathbb{N}$ *y* $\varepsilon > 0$. *Entonces*

$$r_T(n_1 + \cdots + n_N, 2\varepsilon) \leq r_T(n_1, \varepsilon) \cdots r_T(n_N, \varepsilon).$$

Demostración. Sea $S_i$ un conjunto $(n_i, \varepsilon)$ generador de cardinalidad mínima. Sea $S$ el conjunto de $N$-adas $\alpha = (x_1, \dots, x_N)$ con $x_i \in S_i$ tales que $\exists z(\alpha) \in X$ tal que

$$d(T^t(T^{m_i} z(\alpha)), T^t(x_i)) \leq \varepsilon \,\text{para}\, 0 \leq t \leq n_i,$$

donde $m_1 = 0$, $m_i = n_1 + \cdots + n_{i-1}$, $i = 1, \dots, N$. Luego

$$\#S \leq r_T(n_1, \varepsilon) \cdots r_T(n_N, \varepsilon).$$

$\{z(\alpha) : \alpha \in S\}$ es un $(n_1 + \cdots + n_N, 2\varepsilon)$ generador ya que si $x \in X$ $\exists x_i \in S_i$ tal que

$$d(T^t(T^{m_i} x), T^t(x_i)) \leq \varepsilon \;\text{ para }\; 0 \leq t \leq n_i,$$

por lo que $\alpha = (x_1, \dots, x_N) \in S$ y

$$d(T^t(T^{m_i} x), T^t(T^{m_i} z(\alpha))) \leq 2\varepsilon,\, 0 \leq t \leq n_i,\, i = 1, \dots, N$$

O sea

$$d(T^j(x), T^j(z(\alpha))) \leq 2\varepsilon \,\text{para}\, 0 \leq j \leq n_1 + \cdots + n_N.$$

Así

$$r_T(n_1 + \cdots + n_N, 2\varepsilon) \leq \#S \leq r_T(n_1, \varepsilon) \cdots r_T(n_N, \varepsilon).$$

□

Demostración. de la Proposición 3.5. Si $n_0 > 0$ todo $n \geq n_0$ se escribe $n = k\, n_0 + m$ con $0 \leq m < n_0$. Entonces

$$\log r_T(n, 2\varepsilon) = \log r_T(n_0 + \cdots + n_0 + m, 2\varepsilon) \leq k \log r_T(n_0, \varepsilon) + \log r_T(m, \varepsilon),$$

$$\frac{1}{n} \log r_T(n, 2\varepsilon) \leq \frac{kn_0}{nn_0} \log r_T(n_0, \varepsilon) + \frac{1}{n} \sup\{\log r_T(m, \varepsilon) : 1 \leq m < n_0\}.$$

Así

$$\limsup_{n\to\infty} \frac{1}{n} \log r_T(n, 2\varepsilon) \leq \frac{1}{n_0} \log r_T(n_0, \varepsilon)$$

y entonces

$$r_T(2\varepsilon) \leq \inf_n \frac{1}{n} \log r_T(n, \varepsilon).$$

□

Definición 3.10. Un homeomorfismo $T$ de un espacio métrico $(X, d)$ se llama *expansivo* si $\exists \varepsilon_0 > 0$ llamada constante de expansividad tal que si $d(T^n x, T^n y) \leq \varepsilon_0\, \forall n \in \mathbb{Z}$, entonces $x = y$.

Proposición 3.6. *Sea $T$ un homeomorfismo expansivo de un espacio métrico compacto $X$ con constante de expansividad $\varepsilon_0$. Si $0 < \varepsilon < \varepsilon_0$, se tiene*

$$h_{top}(T) = r_T(\varepsilon) = \inf_n \frac{1}{n} \log r(n, \varepsilon/2)$$

Usaremos el siguiente lema

Lema 3.3. *Con las hipótesis de la Proposición 3.6, si $< \varepsilon < \varepsilon_0$, existen $k > 0, N > 2k$ tales que si $x, y \in X$, $n \geq N$ satisfacen $d(T^t(x), T^t(y)) \leq \varepsilon\ \forall 0 \leq t \leq n$, entonces $d(T^t(x), T^t(y)) \leq \varepsilon/2\ \forall k \leq t \leq n - k$*

Demostración. Supongamos que no, entonces existe existe $0 < \varepsilon < \varepsilon_0$ talque $\forall k > 0$ existen $x_k, y_k \in X, t_k, n_k \in \mathbb{N}$ con $n_k > 2k$ y $k \leq t_k \leq n_k - k$ tales que

$$d(T^t(x_k), T^t(y_k)) \leq \varepsilon \quad \text{si}\, 0 \leq t \leq n_k$$

$$d(T^{t_k}(x_k), T^{t_k}(y_k)) > \varepsilon/2.$$

Así $d(T^t T^{t_k}(x_n), T^t T^{t_k}(y_k)) \leq \varepsilon$ para $-t_k \leq t \leq n_k - t_k$.

Sean $x, y$ puntos de acumulación de $\{T^{t_k}(x_k)\}$ y $\{T^{t_k}(y_k\}$.

Entonces $d(x, y) \geq \varepsilon/2$ y como $t_k, n_k - t_k \to \infty$, $d(T^t(x), T^t(y)) \leq \varepsilon\ \forall t \in \mathbb{Z}$, lo que contradice el hecho de que $\varepsilon_0 > \varepsilon$ es una constante de expansividad. □

Corolario 3.6. *Con las hipótesis de la Proposición 3.6, si $0 < \varepsilon < \varepsilon_0$, existen $k > 0\, N > 2k$ tales que $r_T(n, \varepsilon) \geq r_T(n - 2k, \varepsilon/2)\ \forall n \geq N$.*

Demostración. Dado $0 < \varepsilon < \varepsilon_0$ sean $k > 0, N > 2k$ dados por el Lema 3.3. Sea $n \geq N$ y sea $S$ un conjunto $(n, \varepsilon)$ generador de cardinalidad mínima. Se sigue del Lema 3.3 que $T^k(S)$ es un conjunto $(n-2k, \varepsilon/2)$ generador. Así $r_T(n,\varepsilon) \geq r_T(n-2k, \varepsilon/2)$. □

Demostración. de la Proposición 3.6. Por el Corolario 3.6 y (3.3)

$$\liminf_{n\to\infty} \frac{1}{n} \log r_T(n,\varepsilon) \geq \liminf_{n\to\infty} \frac{1}{n} \log r_T(n-2k,\varepsilon/2)$$
$$= \liminf_{n\to\infty} \frac{1}{n} \log r_T(n,\varepsilon/2) \geq \inf_n \frac{1}{n} \log r_T(n,\varepsilon/2) \geq r_T(\varepsilon)$$

□

Proposición 3.7. *Sean $X = [k]^{\mathbb{Z}}$ y $\sigma : X \hookleftarrow$ el corrimiento. Sea $\Lambda \subset X$ compacto e invariante es decir $\sigma(\Lambda) = \Lambda$. Entonces*

$$h_{top}(\sigma \mid \Lambda) = \lim_{n\to\infty} \frac{1}{n} \log r(n),$$

*donde $r(n) = \#\{\theta \mid \{0,\dots,n\} : \theta \in \Lambda\}$.*

**Demostación**. Sea $d$ la métrica en $X$ definida por

$$d(x,y) = \begin{cases} 3^{-N} & \text{si} \quad N = \min\{|n| : x_n \neq y_n\} \\ 0 & \text{si} \quad x = y \end{cases}.$$

Nótese que todo $0 < \varepsilon_0 < 1$ es una constante de expansividad para $\sigma$.

Sean $N \in \mathbb{N}$ y $\varepsilon = 3^{-(1+N)}$. Sea $S_n$ un conjunto tal que para todo $\theta \in \Lambda$ existe un único $\alpha \in S_n$ tal que

$$\alpha|\{-N,\dots,N+n\} = \theta|\{-N,\dots,N+n\}$$

y por consiguiente $d(\sigma^j\alpha, \sigma^j\theta) \leq \varepsilon$ para todo $0 \leq j \leq n$. Entonces $S_n$ es un $(n,\varepsilon)$ generador para $\sigma$ y así $r_\sigma(n,\varepsilon) \leq \#S_n = r(n+2N)$. Si $S$ fuera un $(n,\varepsilon)$ generador con $\#S < \#S_n$, existirían $\alpha_1, \alpha_2 \in S_n, \theta \in S$ tales que $d(\sigma^j\alpha_i, \sigma^j\theta) \leq \varepsilon$ para $0 \leq j \leq n, i = 1,2$. Entonces $d(\sigma^j\alpha_1, \sigma^j\alpha_2) \leq 2\varepsilon < 3^{-N}$ para $0 \leq j \leq n$ y así $\alpha_1|\{-N,\dots,N+n\} = \alpha_2|\{-N,\dots,N+n\}$, lo que contradice la definición de $S_n$.

Por lo tanto $r_\sigma(n,\varepsilon) = r(n+2N)$ y luego

$$h_{\text{top}}(\sigma \mid \Lambda) = \lim_{n\to\infty} \frac{1}{n} \log r(n+2N) = \lim_{n\to\infty} \frac{1}{n} \log r(n).$$

Apliquemos la proposición al subcorrimiento de tipo finito $X_A$ donde $A = (a_{ij})_{i,j\in[k]}$ es una matriz de ceros y unos tal que $A^n > 0$ para algún $n \in \mathbb{N}$. Por el Teorema de Perron $A$ tiene un eigenvalor dominante $\lambda$ con $\dim\ker(A-\lambda I) = 1$

La sucesión $A^n/\lambda^n$ converge porque si $JAJ^{-1}$ es la forma canónica de Jordan con $\lambda$ en el primer renglón y columna entonces

$$\lim_{n\to\infty} \frac{JA^n J^{-1}}{\lambda^n} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}.$$

y así $\lambda^{-n}A^n$ converge a una matriz $\mathbb{K} \neq 0$.

$$\text{Como}\, r(n) = \sum_{i,j} a_{ij}^n,\ \lim_{n\to\infty} \lambda^{-n} r(n) = \sum_{i,j} K_{ij} > 0.$$

$$\frac{1}{n}\log r(n) = \lim_{n\to\infty} \frac{1}{n}\log\frac{r(n)}{\lambda^n} + \log\lambda = \log\lambda.$$

así $h_{\mathrm{top}}(\sigma \mid X_A) = \log\lambda$.

Proposición 3.8. *Sean $X$ compacto y $T : X \hookleftarrow$ continua. Entonces*

$$h_{top}(T^m) = m\, h_{top}(T).$$

Demostración. Si $E$ es un $(nm, \varepsilon)$ generador para $T$ entonces $E$ es un $(n, \varepsilon)$ generador pasa $T^m$. Así $r_{T^m}(n, \varepsilon) \leq r_T(nm, \varepsilon)$ y luego

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n}\log r_{T^m}(n,\varepsilon) &\leq \limsup_{n\to\infty} \frac{m}{nm}\log r_T(nm,\varepsilon) \leq \limsup_{k\to\infty} \frac{m}{k}\log r_T(k,\varepsilon), \\ h_{\mathrm{top}}(T^m) &\leq m\, h_{\mathrm{top}}(T). \end{aligned}$$

Como $T^j$ es uniformemente continua,

$$\forall\varepsilon > 0\, \exists\delta = \delta(\varepsilon) : d(x,y) \leq \delta \Rightarrow d(T^j x, T^j y) \leq \varepsilon,\ 0 \leq j < m$$

Sea $F$ un $(n, \delta)$ generador para $T^m$ entonces $F$ es un $(n, \varepsilon)$ generador para $T$. Así $r_T(nm, \varepsilon) \leq r_{T^m}(n, \delta)$ y luego

$$\begin{aligned} \liminf_{k\to\infty} \frac{m}{k}\log r_T(k,\varepsilon) &\leq \liminf_{n\to\infty} \frac{m}{nm}\log r_T(nm,\varepsilon) \leq \liminf_{n\to\infty} \frac{1}{n}\log r_{T^m}(n,\delta), \\ m h_{\mathrm{top}}(T) &\leq h_{\mathrm{top}}(T^m). \end{aligned}$$

□

Proposición 3.9. *Sean $(X_1, d_1)(X_2, d_2)$ compactos. Consideremos la métrica $d((x_1, x_2), (y_1, y_2)) = \max\{d(x_1, y_1), d(x_2, y_2)\}$ para $X_1 \times X_2$. Si $T_i : X_1 \hookleftarrow i = 1, 2$ son continuas, entonces*

$$h_{top}(T_1 \times T_2) = h_{top}(T_1) + h_{top}(T_2).$$

Demostración. Sea $F_i$ un $(n, \varepsilon)$ generador para $T_i$. Entonces $F_1 \times F_2$ es un $(n, \varepsilon)$ generador para $T_1 \times T_2$. Así

$$\begin{aligned} r_{T_1\times T_2}(n,\varepsilon) &\leq r_{T_1}(n,\varepsilon)\cdot r_{T_2}(n,\varepsilon), \\ r_{T_1\times T_2}(\varepsilon) &\leq r_{T_1}(\varepsilon) + r_{T_2}(\varepsilon), \\ h_{\mathrm{top}}(T_1\times T_2) &\leq h_{\mathrm{top}}(T_1) + h_{\mathrm{top}}(T_2). \end{aligned}$$

Si $E_i$ es $(n,\varepsilon)$ separado para $T_i$, $E_1 \times E_2$ es $(n,\varepsilon)$ separado para $T_1 \times T_2$.

$$\begin{aligned} \text{Así} \qquad s_{T_1\times T_2}(n,\varepsilon) &\geq s_{T_1}(n,\varepsilon)\cdot s_{T_2}(n,\varepsilon), \\ s_{T_1\times T_2}(\varepsilon) &\geq s_{T_1}(\varepsilon)+s_{T_2}(\varepsilon), \\ h_{\mathrm{top}}(T_1\times T_2) &\geq h_{\mathrm{top}}(T_1)+h_{\mathrm{top}}(T_2). \end{aligned}$$

□

PROPOSICIÓN 3.10. *Sea $T : X \hookleftarrow$ un homeomorfismo expansivo. Entonces $\#Fix(T^n) < \infty$ $\forall n \in \mathbb{N}$ y*

$$h_{top}(T) \geq \limsup_{n\to\infty} \frac{1}{n} \log \#Fix(T^n).$$

DEMOSTRACIÓN. Sea $\varepsilon_0$ una constante de expansividad de $T$. Si $x, y \in \mathrm{Fix}(T^n)$ y $d(T^j x, T^j y) \leq \varepsilon_0$ para $0 \leq j \leq n-1$, entonces $d(T^j x, T^j y) \leq \varepsilon_0$ $\forall j \in \mathbb{Z}$ y así $x = y$. Por lo tanto $\mathrm{Fix}(T^n)$ es un conjunto $(n,\varepsilon_0)$ separado y $\#\mathrm{Fix}(T^n) \leq s_T(n,\varepsilon_0) < \infty$. Así

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\mathrm{Fix}(T^n) \leq \limsup_{n\to\infty} \frac{1}{n} \log s_T(n,\varepsilon_0) \leq h_{\mathrm{top}}(T)$$

□

## 3. El principio variacional de la entropía

PRINCIPIO VARIACIONAL. *Sean $X$ espacio métrico compacto, $T : X \hookleftarrow$ continua. Entonces*

$$h_{top}(T) = \sup\{h_\mu(T) : \mu \in \mathcal{M}_T(X)\}$$

Para demostrar el Principio variacional, introducimos otra definición de entropía topológica y demostramos que coincide con la anterior.

DEFINICIONES 3.11. Sea X un espacio topológico compacto. Sean $\mathcal{U}$, $\mathcal{V}$ cubiertas abiertas, y $T : X \hookleftarrow$ continua.

(a) Decimos que $\mathcal{V}$ es un refinamiento de $\mathcal{U}$, y escribimos $\mathcal{U} \leq \mathcal{V}$, si $\forall V \in \mathcal{V}$ $\exists U \in \mathcal{U}$ tal que $V \subset U$.
(b) Definimos $\mathcal{U} \vee \mathcal{V} = \{U \cap V : U \in \mathcal{U}, V \in V\}$.
(c) Definimos $T^{-1}\mathcal{U} = \{T^{-1}U : U \in \mathcal{U}\}$.
(d) Denotamos por $N(\mathcal{U})$ la mínima cardinalidad de una subcubierta finita de $\mathcal{U}$ y definimos la entropía de $\mathcal{U}$ por $H(\mathcal{U}) = \log N(\mathcal{U})$.
(e) Definimos la entropía de $T$ respecto a $\mathcal{U}$ como

$$H(T,\mathcal{U}) = \lim_{n\to\infty} \frac{1}{n} H(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U})$$

PROPOSICIÓN 3.11.

(a) *$H(\mathcal{U}) \geq 0$ y $H(\mathcal{U}) = 0 \iff N(\mathcal{U}) = 1 \iff X \in \mathcal{U}$.*
(b) *Si $\mathcal{U} \leq \mathcal{V}$ entonces $H(\mathcal{U}) \leq H(\mathcal{V})$.*

(c) $H(\mathcal{U} \vee \mathcal{V}) \leq H(\mathcal{U}) + H(\mathcal{V})$.
(d) $H(T^{-1}\mathcal{U}) \leq H(\mathcal{U})$. *Si* $T$ *es sobre entonces* $H(T^{-1}\mathcal{U}) = H(\mathcal{U})$.
(e) *El límite en la Definición 3.11(e) existe.*
(f) *Si* $\mathcal{U} \leq \mathcal{V}$ *entonces* $H(T,\mathcal{U}) \leq H(T,\mathcal{V})$

Demostración. (a) y (b) son inmediatos.

(c). Sean $\mathcal{U}_0$ y $\mathcal{V}_0$ subcubiertas de $\mathcal{U}$ y $\mathcal{V}$ respectivamente, de cardinalidades $N(\mathcal{U})$ y $N(\mathcal{V})$ respectivamente. Entonces $\mathcal{U}_0 \vee \mathcal{V}_0$ es una subcubierta de $\mathcal{U} \vee \mathcal{V}$. Así $N(\mathcal{U} \vee \mathcal{V}) \leq N(\mathcal{U})\, N(\mathcal{V})$.

(d). Sea $\mathcal{U}_0$ subcubierta de $\mathcal{U}$ de cardinalidad $N(\mathcal{U})$ entonces $T^{-1}\mathcal{U}_0$ es una subcubierta de $T^{-1}\mathcal{U}$ y luego $N(T^{-1}\mathcal{U}) \leq N(\mathcal{U})$. Si $T$ es sobre y $T^{-1}\mathcal{U}_0$ es una subcubierta de $T^{-1}\mathcal{U}$ de cardinalidad $N(T^{-1}\mathcal{U})$, entonces $\mathcal{U}_0$ cubre $X$ y luego $N(\mathcal{U}) \leq N(T^{-1}\mathcal{U})$.

(e) Sea $a_n = H(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U})$, entonces

$$\begin{aligned} a_{n+m} &= H(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U} \vee T^{-n}(\mathcal{U} \vee \cdots \vee T^{1-m}\mathcal{U})) \\ &\leq H(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U}) + H(T^{-n}(\mathcal{U} \vee \cdots \vee T^{1-m}\mathcal{U})) \\ &\leq a_n + a_m. \end{aligned}$$

Por el Lema 3.1, $\lim a_n/n$ existe.

(f) Se sigue de (b) y la definición de entropía de $T$ respecto a una cubierta. □

Definición 3.12. Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua. Definimos

$$H_{\text{top}}(T) = \sup\{H(T,\mathcal{U}) : \mathcal{U} \quad \text{es cubierta abierta de} \quad X\}.$$

Si $U$ es cubierta abierta de $X$, definimos el *diámetro* de $\mathcal{U}$ como
$\operatorname{diam}\mathcal{U} = sup\{\operatorname{diam} U : U \in \mathcal{U}\}$

Lema 3.4. *Si* $\{\mathcal{U}_n\}$ *es una sucesión de particiones tal que* $\lim_{n\to\infty} \operatorname{diam}\mathcal{U}_n = 0$*, entonces*

$$H_{top}(T) = \lim_{n\to\infty} H(T,\mathcal{U}_n).$$

Demostración. Dada una cubierta abierta $\mathcal{U}$ de $X$, para $n$ suficientemente grande tenemos que $\operatorname{diam}\mathcal{U}_n$ es menor que el número de Lebesgue de $\mathcal{U}$. Entonces $\mathcal{U}_n$ es un refinamiento de $\mathcal{U}$ y así $H(T,\mathcal{U}) \leq H(T,\mathcal{U}_n)$. □

Proposición 3.12.

$$h_{top}(T) = H_{top}(T)$$

Demostración. Sea $\varepsilon > 0$. Sea $\mathcal{U}$ una cubierta con $\operatorname{diam}\mathcal{U} < \varepsilon$. Dos puntos distintos de un conjunto $(n,\varepsilon)$- separado no pueden estar en el mismo elemento de la cubierta $\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U}$. Por lo tanto

$$s_T(n,\varepsilon) \leq N(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U}).$$

Así, $s_T(\varepsilon) \leq H(T,\mathcal{U}) \leq H_{\text{top}}(T)$ y entonces

$$h_{\text{top}}(T) \leq H_{\text{top}}(T).$$

Sea $\mathcal{U}$ una cubierta abierta de $X$ y sea $\lambda$ su número de Lebesgue. Sea $2\varepsilon < \lambda$ y sea $F$ un $(n, \varepsilon)$- generador de cardinalidad mínima. Para cada $y \in F$, $0 \leq i < n$, escojamos $U_{y,i} \in \mathcal{U}$ conteniendo la bola de radio $\varepsilon$ centrada en $T^i(y)$. El conjunto

$$\{U_{y,0} \cap \cdots \cap T^{-}U_{y,n-1} : y \in F\}$$

es una subcubierta de $\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U}$ y por lo tanto

$$N(\mathcal{U} \vee \cdots \vee T^{1-n}\mathcal{U}) \leq r_T(n, \varepsilon).$$

Así, $H(T,\mathcal{U}) \leq r_T(\varepsilon) \leq h_{\text{top}}(T)$ y entonces

$$H_{\text{top}}(T) \leq h_{\text{top}}(T).$$

$\square$

Proposición 3.13. *Sean $X$ espacio métrico compactoy $\mu \in \mathcal{M}(X)$. Para todo $\varepsilon > 0$ existe una partición finita $\mathcal{P}$ tal que* $\operatorname{diam} \mathcal{P} < 2\varepsilon$ *y* $\mu(\partial P) = 0 \ \forall P \in \mathcal{P}$.

Demostración. Sea $\{B_{\varepsilon/2}(x_i) : i = 1, \ldots k\}$ una cubierta de $X$. Para cada $i = 1, \ldots k$ los conjuntos $\partial B_r(x_i)$, $\varepsilon/2 < r < \varepsilon$ son disjuntos y por lo tanto hay un $\varepsilon/2 < r_i < \varepsilon$ tal que $\mu(\partial B_{r_i}(x_i)) = 0$. Sea $P_1 = B_{r_1}(x_1)$ y defina inductivamente $P_i = B_{r_i}(x_i) \setminus \bigcup\limits_{j<i} P_i$. $\square$

Proposición 3.14. *Sea $X$ espacio métrico compacto. Si $\mu_n \to \mu$ en $M(X)$ entonces*

- $\forall F \subset X$ *cerrado* $\limsup \mu_n(F) \leq \mu(F)$
- $\forall U \subset X$ *abierto* $\liminf \mu_n(U) \geq \mu(U)$
- *Si $A \subset X$ es un boreliano con $\mu(\partial A) = 0$, $\mu_n(A) \to \mu(A)$*

Demostración. Sea $F \subset X$ cerrado y sea

$$V_m = \{x \in X : d(x, F) < \frac{1}{m}\},$$

entonces $\mu(V_m) \to \mu(F)$. Por el lema de Uryson existe $f_m : X \to [0, 1]$ continua tal que $f_m = 1$ en $F$ y $f_m = 0$ en $X \setminus V_m$. Entonces

$$\limsup \mu_n(F) \leq \limsup \int f_m d\mu_n \leq \int f_m d\mu \leq \mu(V_m)$$

y así $\limsup \mu_n(F) \leq \mu(F)$. Si $U \subset X$ es abierto

$$\liminf \mu_n(U) = 1 - \limsup \mu_n(X \setminus U) \geq 1 - \mu(X \setminus U) = \mu(U).$$

Si $A \subset X$ es boreliano con $\mu(\partial A) = 0$, $\mu(\overline{A}) = \mu(A) = \mu(\text{int}A)$ y así

$$\limsup \mu_n(A) \leq \limsup \mu_n(\overline{A}) \leq \mu(A) \leq \liminf \mu_n(\text{int}A) \leq \liminf \mu_n(A)$$

$\square$

LEMA 3.5. *Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua. Dados $\varepsilon > 0$, $n \in \mathbb{N}$, sea $E_n$ un conjunto $(n, \varepsilon)$ separado. Consideremos las medidas*

$$\mu_n = \frac{1}{\# E_n} \sum_{x \in E_n} \delta_x \quad y \quad \nu_n = \frac{1}{n} \sum_{k=0}^{n-1} \mu_n \circ T^{-k}.$$

*Entonces existe un punto de acumulación $\mu \in \mathcal{M}_T(X)$ de $\{\nu_n\}$ tal que*

$$h_\mu(T) \geq \limsup_{n \to \infty} \frac{1}{n} \log(\# E_n).$$

DEMOSTRACIÓN. Por la convexidad de $\phi$, para cualquier particion $\mathcal{Q}$

$$H_{\nu_n}(\mathcal{Q}) = -\sum_Q \phi(\nu_n(Q)) \geq -\frac{1}{n} \sum_Q \sum_{k=0}^{n-1} \phi(\mu_n(T^{-k}Q)) = \frac{1}{n} \sum_{k=0}^{n-1} H_{\mu_n}(T^{-k}\mathcal{Q}).$$

Sea $\{n_j\}$ tal que $\lim_{j \to \infty} \frac{1}{n_j} \log \#(E_{n_j}) = \limsup \frac{1}{n} \log(\# E_n)$ y la subsucesión $\{\nu_{n_j}\}$ converge a una $\mu \in \mathcal{M}(X)$ que por construcción es $T$-invariante.

Sea $\mathcal{P}$ la partición dada por la Proposición 3.13 . Entonces cada átomo $P$ de $\mathcal{P}_n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$ contiene a lo más un punto de $E_n$ y $\mu(\partial P) = 0$. Así

$$H_{\mu_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) = -\sum_{P \in \mathcal{P}_n} \phi(\mu_n(P)) = -\sum_{P \in \mathcal{P}_n} \phi\big(\#(E_n \cap P)/\# E_n\big) \tag{3.4}$$

$$= -\# E_n \phi\big(1/\# E_n\big) = \log(\# E_n) \tag{3.5}$$

Fijemos $m \in \mathbb{N}$. $\forall n \in \mathbb{N}\ \exists d \in \mathbb{N}$, $r \in \{0, \ldots, m-1\}$ tales que $n = dm + r$. Sea $k \leq m$

$$\begin{aligned} \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} &= \left(\bigvee_{i=0}^{d-1} T^{-im}\big(\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}\big)\right) \vee \bigvee_{i=0}^{r-1} T^{-dm-i}\mathcal{P} \\ &= \left(\bigvee_{i=0}^{d-1} T^{-im-k}\big(\bigvee_{j=0}^{m-1} T^{-i}\mathcal{P}\big)\right) \vee \mathcal{P}_{d,r,k} \end{aligned}$$

donde

$$\mathcal{P}_{d,r,k} = \bigvee_{i=0}^{r-1} T^{-dm-i}\mathcal{P} \vee \bigvee_{i=0}^{k-1} T^{-i}\mathcal{P}.$$

es una partición con a lo más $(\#\mathcal{P})^{2m}$ átomos. Por lo tanto

$$\begin{aligned} H_{\mu_n}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) &\leq \sum_{i=0}^{d-1} H_{\mu_n}\big(T^{-im-k}\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}\big) + H_{\mu_n}(\mathcal{P}_{d,k,r}) \\ &\leq \sum_{i=0}^{d-1} H_{\mu_n}\big(T^{-im-k}\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}\big) + 2m\log(\#\mathcal{P}) \end{aligned}$$

y luego

$$H_{\nu_n}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq \frac{1}{n}\sum_{l=0}^{n-1} H_{\mu_n}\big(T^{-l}\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}\big) \tag{3.6}$$

$$\geq \frac{1}{n}\sum_{k=0}^{m-1}\sum_{i=0}^{d-1} H_{\mu_n}\big(T^{-im-k}\bigvee_{j=0}^{m-1} T^{-j}\mathcal{P}\big) \tag{3.7}$$

$$\geq \frac{1}{n}\sum_{k=0}^{m-1}\left(H_{\mu_n}\big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\big) - 2m\log(\#\mathcal{P})\right) \tag{3.8}$$

$$\geq \frac{m}{n} H_{\mu_n}\big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\big) - \frac{2m^2}{n}\log(\#\mathcal{P}). \tag{3.9}$$

Por (3.4) y (3.6)

$$\frac{1}{m} H_{\nu_{n_j}}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq \frac{1}{n_j}\log(\#E_{n_j}) - \frac{2m}{n_j}\log(\#\mathcal{P})$$

Tomando límite cuando $j \to \infty$

$$\frac{1}{m} H_{\mu}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq \limsup_{n\to\infty} \frac{1}{n}\log(\#E_n)$$

y así

$$h_\mu(T,\mathcal{P}) \geq \limsup_{n\to\infty}\frac{1}{n}\log(\#E_n).$$

□

Demostración. del Principio Variacional

I) Sean $\mu \in \mathcal{M}_T(X)$ y $\delta > 0$. Sea $\mathcal{P} = \{P_1, \ldots, P_k\}$ una partición de $(X, \mathcal{A}, \mu)$ tal que $h_\mu(T,\mathcal{P}) \geq h_\mu(T) - \delta$. Dado $\varepsilon \in (0, 1/k\log k)$ escojamos compactos $B_i \subset P_i$, $i = 1, \ldots, k$ tales que $\mu(P_i \setminus B_i) < \varepsilon$. Sean

$$B_0 = X \setminus \bigcup_{i=1}^{k} B_i, \quad \mathcal{B} = \{B_0, B_1, \ldots, B_k\}.$$

Entonces

$$\begin{aligned} H_\mu(\mathcal{P} \mid \mathcal{B}) &= -\mu(B_0)\sum_{i=1}^{k}\phi\Big(\frac{\mu(P_i\cap B_0)}{\mu(B_0)}\Big) \\ &\le \mu(B_0)\log k \le \varepsilon\, k\log k < 1 \end{aligned}$$

Por la Proposición 3.2(a)

$$h_\mu(T,\mathcal{P}) \le h_\mu(T,\mathcal{B}) + H_\mu(\mathcal{P}\mid\mathcal{B}) < h_\mu(T,\mathcal{B}) + 1. \tag{3.10}$$

Sea $U_i = B_0 \cup B_i$, y por lo tanto $X \setminus U_i = \bigcup_{j\neq 0,i} B_j$. O sea que cada elemento de la cubierta

$$\mathcal{U} = \{U_1,\dots,U_k\}$$

intersecta a dos elementos de la partición $\mathcal{B}$. Como $B_i \subset U_i$ para todo $i$, cada elemento de $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}$ está contenido en un elemento de $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}$ y cada uno de estos últimos contiene a lo más $2^n$ elementos de $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}$. Luego

$$\# \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B} \le 2^n N\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}\right),$$

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right) \le n\log 2 + H\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{U}\right),$$

$$h_\mu(T,\mathcal{B}) \le \log 2 + H(T,\mathcal{U}).$$

Por (3.10)

$$h_\mu(T,\mathcal{P}) \le 1 + \log 2 + H(T,\mathcal{U}),$$

luego

$$h_\mu(T) - \delta \le 1 + \log 2 + h_{\text{top}}(T),$$

y como $\delta > 0$ es arbitraria

$$h_\mu(T) \le 1 + \log 2 + h_{\text{top}}(T).$$

Aplicando esta desigualdad a $T^m$ y usando las Proposiciones 3.3 y 3.8, tenemos

$$mh_\mu(T) = h_\mu(T^m) \le 1 + \log 2 + h_{\text{top}}(T^m) = 1 + \log 2 + mh_{\text{top}}(T)$$

Luego, para toda $m \in \mathbb{N}$ tenemos

$$h_\mu(T) \le \frac{1}{m}(1+\log 2) + h_{\text{top}}(T)$$

Y así $h_\mu(T) \le h_{\text{top}}(T)$.

II) Sean $\varepsilon > 0$, $n \in \mathbb{N}$. Sea $E_n$ un conjunto $(n,\varepsilon)$ separado de cardinalidad máxima. Por el Lema 3.5, hay una medida $\mu \in \mathcal{M}_T(X)$ tal que $h_\mu(T) \ge s_T(\varepsilon)$. Así,

$$\sup\{h_\mu(T) : \mu \in \mathcal{M}_T(X)\} \ge h_{\text{top}}(T).$$

$\square$

COROLARIO 3.7. *Si $T : X \hookleftarrow$ es un homeomorfismo expansivo, entonces existe $\mu \in \mathcal{M}_T(X)$ tal que $h_\mu(T) = h_{top}(T)$*

DEMOSTRACIÓN. Sea $\varepsilon_0$ una constante de expansividad para $T$. Por la Proposición 3.6, tenemos que para $\varepsilon < \varepsilon_0$ $h_{\text{top}}(T) = s_T(\varepsilon)$. Tomando para $\varepsilon < \varepsilon_0$ $E_n$ un conjunto $(n, \varepsilon)$ separado de cardinalidad máxima, la medida $\mu \in \mathcal{M}_T(X)$ dada por el Lema 3.5 satisface $h_\mu(T) = h_{\text{top}}(T)$. □

## 4. Entropía de los automorfismos torales

TEOREMA 3.8. *Sean $A \in GL(N, \mathbb{Z})$ y $\tilde{A} : \mathbb{T}^N \hookleftarrow$ la transformación inducida. Sea $R_a : \mathbb{T}^N \hookleftarrow$ una rotación y sea $T = R_a \circ \tilde{A}$. Sea $m$ la medida inducida por la medida de Lebesgue en $\mathbb{R}^N$. Entonces*

$$h_m(T) = h_{top}(T) = h_m(\tilde{A}) = h_{top}(\tilde{A}).$$

DEMOSTRACIÓN. Sean $\pi : \mathbb{R}^N \to \mathbb{T}^N$ la proyección natural y $d$ la distancia en $\mathbb{T}^N$ inducida por la euclidiana. Sea

$$B_n(x, \varepsilon, T) = \left\{y \in \mathbb{T}^n : d(T^i y, T^i x) < \varepsilon, 0 \leq i \leq n-1\right\}.$$

Demostraremos por inducción que $T^{-k}(B_\varepsilon(T^k(x))) = x \cdot \left[\tilde{A}^{-k} B_\varepsilon \pi(0)\right]$ la cual es válida para $k = 0$ por la invariancia de $d$ bajo rotaciones. Si es válida para $k$ entonces

$$\begin{aligned} T^{-(k+1)} B_\varepsilon(T^{k+1}x) = T^{-1}(T^{-k} B_\varepsilon(T^k(Tx))) = T^{-1}\left(Tx \cdot \left[\tilde{A}^{-k} B_\varepsilon(\pi(0))\right]\right) \\ = \tilde{A}^{-1}(a^{-1}Tx \cdot \tilde{A}^{-k} B_\varepsilon(\pi(0))) = x \cdot \tilde{A}^{-(k+1)}(B_\varepsilon(\pi(0))). \end{aligned}$$

Así

$$B_n(x, \varepsilon, T) = x \cdot \bigcap_{k=0}^{n-1} \tilde{A}^{-k}(B_\varepsilon(\pi(0))) = x \cdot B_n(\pi(0), \varepsilon, \tilde{A})$$

Luego $m(B_n(x, \varepsilon, T)) = m(B_n(\pi(0), \varepsilon, \tilde{A}))$. Sea $\varepsilon > 0$ y sea $\mathcal{P} = \{P_1, \ldots, P_n\}$ partición de diametro $< \varepsilon$.

$$x \in \bigcap_{k=0}^{n-1} T^{-k} P_{i_k} \Rightarrow \bigcap_{k=0}^{n-1} T^{-k} P_{i_k} \subset x \cdot B_n(\pi(0), \varepsilon, \tilde{A}),$$

porque si $y \in \bigcap_{k=0}^{n-1} T^{-k}P_{i_k}$ entonces $T^kx, T^ky \in P_{i_k}$ y así $d(T^kx, T^ky) < \varepsilon$ para $0 \leq k \leq n-1$, o sea $y \in B_n(x, \varepsilon, T) = x \cdot B_n(\pi(0), \varepsilon, \tilde{A})$. Por lo tanto

$$\begin{aligned} m\left(\bigcap_{k=0}^{n-1} T^{-k}P_{i_k}\right) &\leq m(B_n(\pi(0), \varepsilon, \tilde{A}), \\ \sum_{i_0,\ldots,i_{n-1}} \phi(m(\bigcap_{k=0}^{n-1} T^{-k}P_{i_k})) &\leq \sum_{i_0,\ldots,i_{n-1}} m(\bigcap_{k=0}^{n-1} T^{-k}P_{i_k}) \log m(B_n(\pi(0), \varepsilon, \tilde{A})) \\ &= \log m(B_n(\pi(0), \varepsilon, \tilde{A})). \end{aligned}$$

Entonces

$$h_m(T) \geq h_m(T, \mathcal{P}) \geq \limsup_{n\to\infty} \left[ -\frac{1}{n} \log m(B_n(\pi(0), \varepsilon, \tilde{A})) \right].$$

Como $\varepsilon > 0$ fue arbitraria

$$h_m(T) \geq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[ -\frac{1}{n} \log m(B_n(\pi(0), \varepsilon, \tilde{A})) \right].$$

Sea $E$ un conjunto $(n, \varepsilon)$ separado respecto a $T$ con cardinalidad máxima. Entonces

$$\bigcup_{x\in E} B_n(x, \varepsilon/2, T) = \bigcup_{x\in E} x \cdot B_n(\pi(0), \varepsilon/2, \tilde{A})$$

es unión ajena y así

$$s_T(n, \varepsilon) m(B_n(\pi(0), \varepsilon/2, \tilde{A})) \leq 1.$$

Luego

$$\limsup_{n\to\infty} \frac{1}{n} \log s_T(n, \varepsilon) \leq \limsup_{n\to\infty} \left[ -\frac{1}{n} \log m(B_n(\pi(0), \varepsilon/2, \tilde{A})) \right],$$

y entonces

$$h_{\text{top}}(T) \leq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[ -\frac{1}{n} \log m(B_n(\pi(0), \varepsilon/2, \tilde{A})) \right] \leq h_m(T).$$

Como este límite no depende de $a$ , $h_{top}(T) = h_{top}(\tilde{A})$. □

Extenderemos ahora la definición de entropía topológica a espacios métricos no necesariamente compactos para incluir $A : \mathbb{R}^N \hookleftarrow$ lineal.

Definiciones 3.13. Sean $(X, d)$ espacio métrico y $T : X \hookleftarrow$ uniformemente continua. Sea $K \subset X$ compacto.

(a) Decimos que $F \subset X$ es un $(n, \varepsilon)$ generador para $K$ si $\forall y \in K\ \exists x \in F$ tal que $d(T^ix, T^iy) \leq \varepsilon$ para $0 \leq i \leq n-1$. Sea $r(K, n, \varepsilon)$ la m'ınima cardinalidad de un $(n, \varepsilon)$ generador.

(b) Decimos que $E \subset K$ está $(n, \varepsilon)$ separado si $\forall x, y \in E,\ x \neq y\ \exists i \in \{0, \ldots, n-1\}$ tal que $d(T^ix, T^i, y) > \varepsilon$. Sea $s(K, n, \varepsilon)$ la máxima cardinalidad de un $(n, \varepsilon)$ separado.

Se tiene como antes $r(K,n,\varepsilon) \le s(K,n,\varepsilon) \le r(K,n,\varepsilon/2)$.

(c)

$$h(K,T) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log r(K,n,\varepsilon) = \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log s(K,n,\varepsilon)$$

(d)

$$h_{top}(T) := \sup \{h(K,A) : K \subset X \text{ es compacto}\}.$$

Lema 3.6. *Sea $A$ como en el Teorema 3.8. Sean $\mu$ la medida de Lebesgue y $\|\|$ una norma en $\mathbb{R}^N$. Sean*

$$B_{\|\|}(0;\varepsilon) = \left\{x \in \mathbb{R}^N : \|x\| < \varepsilon\right\}, \quad B_n(0,\varepsilon,A) = \bigcap_{i=0}^{n-1} A^{-i} B_{\|\|}(0;\varepsilon).$$

*Entonces*

$$h_{top}(A) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n} \log \mu(B_n(0,\varepsilon,A))\right].$$

Demostración. Si $F$ es un $(n,\varepsilon)$ generador para $K \subset \mathbb{R}^N$ compacto,

$$K \subset \bigcup_{x\in F} B_n(x,2\varepsilon,A) = \bigcup_{x\in F} x + B_n(0,2\varepsilon,A).$$

Así $\mu(K) \le (\#F)\mu(B_n(0,2\varepsilon,A))$, y entonces

$$r(K,n,\varepsilon) \ge \mu(K)/\mu(B_n(0,2\varepsilon,A).$$

Por lo tanto, cuando $\mu(K) > 0$

$$h(K,A) \ge \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n} \log \mu(B_n(0,\varepsilon,A))\right].$$

Si $q$ es suficientemente grande $K \subset K_q := [-q,q]^N$. Si $E \subset K$ es $(n,\varepsilon)$ separado entonces $\bigcup_{x\in E} B_n(x,\varepsilon/2,A)$ es unión disjunta y está contenida en $K_{q+\varepsilon}$. Así

$$s(K,n,\varepsilon)\mu(B_n(0,\varepsilon/2,A)) \le 2^N(q+\varepsilon)^N,$$

y luego

$$\limsup_{n\to\infty} \frac{1}{n} \log s(K,n,\varepsilon) \le \limsup_{n\to\infty} \left[-\frac{1}{n} \log \mu(B_n(0,\varepsilon/2,A))\right],$$

$$h(K,A) \le \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n} \log \mu(B_n(0,\varepsilon,A))\right].$$

Así $h_{top}(A) = h(K,A)$ para cualquier compacto con $\mu(K) > 0$. □

Lema 3.7. *Sean $A, \tilde{A}$ como en el Teorema 3.8. Entonces*

$$h_{top}(A) = h_{top}(\tilde{A})$$

Demostración. Sea $q < \frac{1}{2}$ Sea $0 < \varepsilon < \text{mín}(1, \|A\|^{-1})$. Supongamos que $E \subset K_q$ es $(n, \varepsilon)$ separado. Entonces $\pi(E)$ es $(n, \varepsilon)$ separado para $\tilde{A}$. En efecto si $x, y \in E$ y $\pi(x) \neq \pi(y)$ entonces $x \neq y$. Sea $i_0$ tal que $\|A^i(x-y)\| \leq \varepsilon$ si $i \leq i_0$ y $\|A^{i_0+1}(x-y)\| > \varepsilon$. Entonces

$$\|A^{i_0+1}(x-y)\| \leq \|A\|\|A^{i_0}(x-y\| \leq \|A\|\varepsilon < 1.$$

y así $d(\tilde{A}^{i_0+1}\pi(x), \tilde{A}^{i_0+1}\pi(y) = \|A^{i_0+1}(x-y)\| > \varepsilon$. Luego $s(K_q, n, \varepsilon) \leq s_{\tilde{A}}(n, \varepsilon)$, de donde $h_{K_q}(A) \leq h_{top}(\tilde{A})$.

Sea $\tilde{E}$ un $(n, \varepsilon)$ separado para $\tilde{A}$. Sea $E = \pi^{-1}(E) \cap [-\frac{1}{2}, \frac{1}{2}[^N$. Entonces $E$ es $(n, \varepsilon)$ separado para $K_{\frac{1}{2}}$ porque si $\|A^i(x) - A^i y\| \leq \varepsilon$ $x, y \in E$, entonces $d(\tilde{A}^i\pi(x), \tilde{A}^i(\pi(y)) \leq \varepsilon$. Así $s_{\tilde{A}}(n, \varepsilon) \leq s(K_{\frac{1}{2}}, n, \varepsilon)$ y entonces $h_{\text{top}}(\tilde{A}) \leq h(K_{\frac{1}{2}}, A)$. □

Teorema 3.9. *Sean $A, \tilde{A}$ como en el Teorema 3.8. Sean $\lambda_1, \ldots, \lambda_N$ los eigenvalores de $A$. Entonces*

$$h_{top}(\tilde{A}) = \sum_{|\lambda_i|>1} \log|\lambda_i|.$$

Demostración. Descompongamos $\mathbb{R}^N = F_1 \oplus F_2$ tal que $A(F_i) \subset F_i$, los eigenvalores de $A_1 = A \mid F_1$ tienen módulo $> 1$ y los eigenvalores de $A_2 = A \mid F_2$ tienen módulo $\leq 1$ Eligiendo una base en $F_i$ podemos suponer $F_i = R^{p_i}$ Sea $m_i$ la medida de Lebesgue en $F_i$ y sea $m = m_1 \times m_2$ la medida de Lebesgue en $\mathbb{R}^N$. Sea $\|\|_i$ norma en $F^i$, entonces

$$\|(x_1, x_2)\| = \text{máx}(\|x_1\|, \|x_2\|)$$

define una norma en $\mathbb{R}^N$. Por el Lema 3.6

$$h_{top}(A) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n}\left[-\log m_1(B_n(0, \varepsilon, A_1)) - \log m_2(B_n(0, \varepsilon, A_2))\right].$$

Se tiene que

$$\begin{aligned} m_1(B_n(0, \varepsilon, A_1)) &\leq m_1(A_1^{1-n} B_{\|\|_1}(0, \varepsilon)) = |\det A_1|^{1-n} m_1(B_{\|\|_1}(0, \varepsilon)) \\ m_2(B_n(0, \varepsilon, A_2)) &\leq m_2(B_{\|\|_2}(0, \varepsilon)). \end{aligned}$$

Por lo tanto

$$\frac{1}{n}\left[-\log m_1(B_n(0, \varepsilon, A_1)) - \log m_2(B_n(0, \varepsilon, A_2))\right]$$
$$\geq \frac{(n-1)}{n} \log|\det A_1| - \frac{1}{n} \log m(B_{\|\|}(0, \varepsilon)),$$

$$h_{\text{top}}(A) \geq \log|\det A_1| = \sum_{|\lambda_i > 1|} \log|\lambda_i|.$$

Para probar la desigualdad opuesta escribamos $\mathbb{R}^N = E_1 \oplus \cdots \oplus E_k$ tal que $A(E_j) \subset E_j$ y todos los eigenvalores de $A_j = A \mid E_j$ tienen el mismo módulo $\mu_j$. Así $A = A_1 \oplus \cdots \oplus A_k$ y entonces

$$h_{\text{top}}(A) \leq h_{\text{top}}(A_1) + \cdots + h_{\text{top}}(A_k).$$

Demostrando $h_{top}(A_j) \leq \text{máx}\,\{0, (\dim E_j)\log \mu_j\}$, tendremos

$$h_{top}(A) \leq \sum_{\mu_j > 1} (\dim E_j) \log \mu_j = \sum_{|\lambda_i > 1|} \log |\lambda_i| \,.$$

Resta pues, demostrar el siguiente

LEMA 3.8. *Sea $A : \mathbb{R}^P \hookleftarrow$ lineal tal que todos sus eigenvalores tienen el mismo módulo $\mu$, entonces*

$$h_{top}(A) \leq \text{máx}\,\{0, P \log \mu\}\,.$$

DEMOSTRACIÓN. Si $\|A\| \leq 1$, un $(1, \varepsilon)$ generador para un compacto $K$ es también un $(n, \varepsilon)$ generador. Así $r(K, n, \varepsilon) \leq r(K, 1, \varepsilon)$, luego $h(K, A) = 0$ y así

$$h_{\text{top}}(A) = 0 = \text{máx}\,\{0, P \log \|A\|\}\,.$$

Si $\|A\| > 1$, sean $K = K_1$ y $c = \text{máx}\,\{\|x\| : x \in K\}$.

Para $0 < \delta < 1$, sea $F(\delta) = \{\delta(n_1, \ldots, n_P) \in K : n_i \in \mathbb{Z}\}$.

Si $N = [2/\delta]$, $\#F(\delta) \leq (N+1)^P < (3/\delta)^P$.

Como $\forall y \in K\ \exists x \in F(\delta)$ tal que $\|x - y\| < c\delta$, se tiene que $F(\delta)$ es un $(n, \|A\|^n c\delta)$ generador para $K$ ya que

$$\left\|A^i x - A^i y\right\| \leq \left\|A^i\right\| \|x - y\| \leq \|A\|^n \|x - y\|$$

si $0 \leq i \leq n$. Sea $\varepsilon > 0$, tomemos $\delta = \varepsilon \|A\|^{-n}/c < 1$ para $n$ suficientemente grande. Entonces

$$r(K, n, \varepsilon) \leq \#F(\delta) \leq (3\|A\|^n c/\varepsilon)^P,$$

$$\limsup_{n\to\infty} \frac{1}{n} \log r(K, n, \varepsilon) \leq \limsup_{n\to\infty} P \left[\log \|A\| + \frac{1}{n}\log(3c/\varepsilon)\right] = P \log \|A\|,$$

$$h_{\text{top}}(A) \leq P \log \|A\| = \text{máx}\,\{0, P \log \|A\|\}\,.$$

Aplicando esta desigualdad a $A^n$ y utilizando la Proposición 3.8

$$h_{\text{top}}(A) = \frac{1}{n} h_{\text{top}}(A^n) \leq \text{máx}\left\{0, P \log \|A^n\|^{\frac{1}{n}}\right\}.$$

Como $\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} =$ radio espectral de $A = \mu$, tenemos

$$h_{\text{top}}(A) \leq \text{máx}\,\{0, P \log \mu\}$$

$\square$

$\square$

## 5. Presión topológica

Sean $X$ espacio métrico compacto, $T : X \hookleftarrow$ continua y $\varphi \in C(X)$. Definimos $S_n\varphi(x) = \sum_{i=0}^{n-1} \varphi(T^i x)$

$$r_T(\varphi, n, \varepsilon) := \inf \left\{ \sum_{x \in E} e^{S_n\varphi(x)} : E \text{ es } (n, \varepsilon)\text{-generador} \right\} \tag{3.11}$$

$$s_T(\varphi, n, \varepsilon) := \sup \left\{ \sum_{x \in E} e^{S_n\varphi(x)} : E \text{ es } (n, \varepsilon)\text{-separado} \right\} \tag{3.12}$$

Se tiene

$$r_T(\varphi, n, \varepsilon) \leq \exp(\|S_n(\varphi)\|) r_T(n, \varepsilon)$$

y así

$$r_T(\varphi, \varepsilon) = \limsup_{n \to \infty} \frac{1}{n} \log r_T(\varphi, n, \varepsilon) \leq \|\varphi\| + r_T(\varepsilon) < \infty$$

Definimos la *presión topológica* de $\varphi$ como

$$P(\varphi) = P_T(\varphi) = \lim_{\varepsilon \to 0} r_T(\varphi, \varepsilon).$$

Sea $s_T(\varphi, \varepsilon) = \limsup_n \frac{1}{n} \log s_T(\varphi, n, \varepsilon)$.

Dado $C \subset X$ no vacío denotamos $S_n\varphi(C) = \sup_{x \in C} S_n\varphi(x)$. Dada una cubierta $\mathcal{U}$ de $X$ definimos la sucesión

$$p_T(\varphi, n, \mathcal{U}) := \inf \left\{ \sum_{V \in \mathcal{V}} e^{S_n\varphi(V)} : \mathcal{V} \text{ es subcubierta finita de } \bigvee_{i=1}^{n-1} T^{-i}\mathcal{U} \right\} \tag{3.13}$$

Proposición 3.15. *La sucesión* $\log p_T(\varphi, n, \mathcal{U})$ *es subaditiva.*

Demostración. Si $\mathcal{V}_n, \mathcal{V}_m$ son subcubiertas de $\bigvee_{i=1}^{n-1} T^{-i}\mathcal{U}$ y $\bigvee_{i=1}^{m-1} T^{-i}\mathcal{U}$ respectivamente, entonces $\mathcal{V}_n \vee T^{-n}\mathcal{V}_m$ es subcubierta de $\bigvee_{i=1}^{n+m-1} T^{-i}\mathcal{U}$ y

$$\sum_{V \in \mathcal{V}_n, W \in \mathcal{V}_m} e^{S_{n+m}\varphi(V \cap T^{-n}W)} \leq \sum_{V \in \mathcal{V}_n} e^{S_n\varphi(V)} \sum_{W \in \mathcal{V}_m} e^{S_n\varphi(V)}.$$

Por lo tanto $\quad p_T(\varphi, n+m, \mathcal{U}) \leq p_T(\varphi, n, \mathcal{U})\, p_T(\varphi, m, \mathcal{U}).$ □

Definimos

$$p_T(\varphi, \mathcal{U}) := \lim_{n \to \infty} \frac{1}{n} \log p_T(\varphi, n, \mathcal{U}).$$

LEMA 3.9. *El límite* $\lim_{\operatorname{diam}\mathcal{U}\to 0} p_T(\varphi,\mathcal{U})$ *existe, o sea que hay un* $p_T(\varphi)\in[0,\infty]$ *tal que para cualquier sucesión de cubiertas* $(\mathcal{U}_k)$ *con* $\operatorname{diam}\mathcal{U}_k\to 0$ *se tiene*

$$p_T(\varphi)=\lim_k p_T(\varphi,\mathcal{U}_k).$$

DEMOSTRACIÓN. Sean $(\mathcal{U}_k)_k$, $(\mathcal{V}_k)_k$ sucesiones con diámetros convergiendo a cero. Dado $\varepsilon>0$ fijemos $\delta>0$ tal que si $d(x,y)<\delta$ entonces $|\varphi(x)-\varphi(y)|<\varepsilon$. Fijemos $k$ suficientemente grande tal que $\operatorname{diam}\mathcal{U}_k<\delta$ y sea $\rho>0$ el número de Lebesgue de $\mathcal{U}_k$. Para $l$ suficientemente grande $\operatorname{diam}\mathcal{V}_l<\rho$, y así todo $V\in\mathcal{V}_l$ esta contenido en algún $U\in\mathcal{U}_k$. Por consiguiente, para cualquier $n\in\mathbb{N}$ todo $V\in\bigvee_{i=1}^{n-1}T^{-i}\mathcal{V}_l$ esta contenido en algún $U\in\bigvee_{i=1}^{n-1}T^{-i}\mathcal{U}_k$ y como $\operatorname{diam}\mathcal{U}_k<\delta$ tenemos que $S_n\varphi(U)\le n\varepsilon+S_n\varphi(V)$. Luego

$$p_T(\varphi,n,\mathcal{U}_k)\le e^{n\varepsilon}p_T(\varphi,n,\mathcal{V}_l)$$

y por lo tanto $p_T(\varphi,\mathcal{U}_k)\le\varepsilon+p_T(\varphi,\mathcal{V}_l)$. Haciendo tender primero $l\to\infty$ y luego $k\to\infty$ obtenemos

$$\limsup_k p_T(\varphi,\mathcal{U}_k)\le\varepsilon+\liminf_l p_T(\varphi,\mathcal{V}_l)$$

Como $\varepsilon>0$ es arbitrario, se sigue que

$$\limsup_k p_T(\varphi,\mathcal{U}_k)\le\liminf_l p_T(\varphi,\mathcal{V}_l)$$

Ahora intercambiemos los papeles de $(\mathcal{U}_k)_k$ y $(\mathcal{V}_k)_k$. □

PROPOSICIÓN 3.16.

1. *Sea* $\mathcal{U}$ *cubierta con* $\operatorname{diam}\mathcal{U}<\varepsilon$ *entonces*

$$r_T(\varphi,n,\varepsilon)\le s_T(\varphi,n,\varepsilon)\le p_T(\varphi,n,\mathcal{U})$$

   *de donde* $\quad P_T(\varphi)\le\lim_{\varepsilon\to 0}s_T(\varphi,\varepsilon)\le p_T(\varphi)$
2. $P_T(\varphi)=\lim_{\varepsilon\to 0}s_T(\varphi,\varepsilon)=p_T(\varphi)=\lim_{\varepsilon\to 0}\liminf_n\frac{1}{n}\log r_T(\varphi,n,\varepsilon)$
3. *Para* $m\in\mathbb{N}$, $P_{T^m}(S_m\varphi)=mP_T(\varphi)$

DEMOSTRACIÓN. 1. Si $E$ es un $(n,\varepsilon)$-separado de cardinalidad máxima, entonces es un $(n,\varepsilon)$-generador y así

$$r_T(\varphi,n,\varepsilon)\le\sum_{x\in E}e^{S_n\varphi(x)}\le s_T(\varphi,n,\varepsilon).$$

Si $E$ es un $(n,\varepsilon)$-separado y $\mathcal{V}$ es una subcubierta de $\bigvee_{i=1}^{n-1}T^{-i}\mathcal{U}$, definimos una función inyectiva $\phi:E\to\mathcal{V}$ tal que $x\in\phi(x)$. Por lo tanto

$$\sum_{x\in E}e^{S_n\varphi(x)}\le\sum_{V\in\mathcal{V}}e^{S_n\varphi(V)}$$

Tomando el supremo sobre los $E$ y el ínfimo sobre las $\mathcal{V}$, obtenemos

$$s_T(\varphi,n,\varepsilon) \le p_T(\varphi,n,\mathcal{U}).$$

2. Para todo $\delta < 0$ existe $\varepsilon_0 < 0$ tal que si $0 < \varepsilon < \varepsilon_0$ y $d(x,y) < \varepsilon$ entonces $|\varphi(x) - \varphi(y)| < \delta$.

Sea $\mathcal{U}$ una cubierta con $\operatorname{diam}\mathcal{U} < \varepsilon$ y sea $\rho$ un número de Lebesgue de $\mathcal{U}$. Sea $E$ es un $(n,\rho)$-generador. Para cada $x \in E$, $0 \le i < n$ existe $U_{x,i} \in \mathcal{U}$ tal que $B_\rho(T^i x) \subset U_{x,i}$. Denotamos $V(x) = \bigcap_{i=0}^{n-1} U_{x,i}$. Como $E$ es un $(n,\rho)$-generador, $\mathcal{V} = \{V(x) : x \in E\}$ es una sucubierta de $\bigvee_{i=1}^{n-1} T^{-i}\mathcal{U}$. Además, ya que $\operatorname{diam} U_{i,x} < \varepsilon$, para todo $x \in E$ tenemos

$$S_n\varphi(V(x)) \le n\delta + S_n\varphi(x)$$
$$\sum_{V\in\mathcal{V}} e^{S_n\varphi(V)} \le e^{n\delta} \sum_{x\in E} e^{S_n\varphi(x)}$$

De donde $p_T(\varphi,n,\mathcal{U}) \le e^{n\delta} r_T(\varphi,n,\rho)$ para todo $n \in \mathbb{N}$ y por consiguiente

$$p_T(\varphi,\mathcal{U}) \le \delta + \liminf_n \frac{1}{n} \log r_T(\varphi,n,\rho).$$

Haciendo tender primero $\rho \to 0$ y luego $\varepsilon \to 0$ tenemos

$$p_T(\varphi) \le \delta + \lim_{\rho\to 0} \liminf_n \frac{1}{n} \log r_T(\varphi,n,\rho)$$

para cualquie $\delta > 0$ y así

$$p_T(\varphi) \le \lim_{\rho\to 0} \liminf_n \frac{1}{n} \log r_T(\varphi,n,\rho)$$

3. Si $F$ es un $(nm,\varepsilon)$ generador de cardinalidad mínima para $T$, entonces $F$ es un $(n,\varepsilon)$ generador para $T^m$. Así $r_{T^m}(S_m\varphi,n,\varepsilon) \le r_T(\varphi,nm,\varepsilon)$ y luego

$$P_{T^m}(S_m\varphi) \le mP_T(\varphi).$$

Dado $\varepsilon > 0$ escogemos $\delta > 0$ tal que $d(T^i x, T^i y) < \varepsilon$ si $i < m$ y $d(x,y) < \delta$. Si $F$ es $(n,\delta)$ generador para $T^m$, entonces $F$ es $(nm,\varepsilon)$ generador para $T$. Así $r_{T^m}(S_m\varphi,n,\delta) \ge r_T(\varphi,nm,\varepsilon)$ y luego

$$P_{T^m}(S_m\varphi) \ge r_{T^m}(S_m\varphi,\delta) \ge m \liminf_n \frac{1}{n} \log r_T(\varphi,n,\varepsilon).$$

Haciendo tender $\varepsilon$ a cero tenemos $P_{T^m}(S_m\varphi) \ge mP_T(\varphi)$ $\square$

Tenemos el Lema análogo al Lema 3.5

Lema 3.10. *Sean $X$ espacio métrico compacto y $T : X \hookleftarrow$ continua, $\varphi \in C(X)$. Dados $\varepsilon > 0$, $n \in \mathbb{N}$, sea $E_n$ un conjunto $(n,\varepsilon)$ separado. Consideremos las medidas*

$$\mu_n = \left(\sum_{y\in E_n} e^{S_n\varphi(y)}\right)^{-1} \sum_{x\in E_n} e^{S_n\varphi(x)}\delta_x \quad y \quad \nu_n = \frac{1}{n}\sum_{k=0}^{n-1} \mu_n \circ T^{-k}.$$

*Entonces existe un punto de acumulación $\mu \in \mathcal{M}_T(X)$ de $\{\nu_n\}$ tal que*

$$h_\mu(T) + \int \varphi \, d\mu \geq \limsup_{n\to\infty} \frac{1}{n} \log \sum_{x\in E_n} e^{S_n\varphi(x)}. \tag{3.14}$$

Demostración. Sea $n_j$ tal que

$$\lim_{j\to\infty} \frac{1}{n_j} \log \sum_{x\in E_{n_j}} \exp(S_{n_j}\varphi(x)) = \limsup_{n\to\infty} \frac{1}{n} \log \sum_{x\in E_n} \exp(S_n\varphi(x))$$

y la subsucesión $\{\nu_{n_j}\}$ converge a una $\mu \in \mathcal{M}(X)$ que por construcción es $T$-invariante. Sea $\mathcal{P}$ la partición dada por la Proposición 3.13 . Entonces cada átomo $P$ de $\mathcal{P}_n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$ contiene a lo más un punto de $E_n$ y $\mu(\partial P) = 0$. Así

$$\begin{aligned} H_{\mu_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) + n\int \varphi \, d\nu_n &= H_{\mu_n}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) + \int S_n(\varphi)\, d\mu_n \\ &= \sum_{x\in E_n} -\mu_n(\{x\})\log \mu_n(\{x\}) + \mu_n(\{x\})S_n\varphi(x) \\ = \sum_{x\in E_n} -\frac{e^{S_n\varphi(x)}}{\sum_{y\in E_n} e^{S_n\varphi(y)}}\left[S_n\varphi(x) - \log \sum_{y\in E_n}^{n} e^{S_n\varphi(y)}\right] &+ \frac{e^{S_n\varphi(x)}}{\sum_{y\in E_n} e^{S_n\varphi(y)}} S_n\varphi(x) \\ &= \log \sum_{x\in E_n} e^{S_n\varphi(x)} \end{aligned} \tag{3.15}$$

Por (3.6) en la prueba del Lemma 3.5 y (3.15)

$$\begin{aligned} H_{\nu_n}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) + m\int \varphi\, d\nu_n &\geq \frac{m}{n}(H_{\mu_n}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) + n\int \varphi\, d\nu_n) - \frac{2m^2}{n}\log(\#\mathcal{P}) \\ &\geq \frac{m}{n}\log \sum_{x\in E_n} e^{S_n\varphi(x)} - \frac{2m^2}{n}\log(\#\mathcal{P}) \end{aligned} \tag{3.16}$$

Poniendo $n = n_j$ y tomando límite cuando $j \to \infty$ tenemos

$$\frac{1}{m} H_\mu(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) + \int \varphi\, d\mu \geq \limsup_{n\to\infty} \frac{1}{n}\log \sum_{x\in E_n} e^{S_n\varphi(x)}$$

y así

$$h_\mu(T,\mathcal{P}) + \int \varphi\, d\mu \geq \limsup_{n\to\infty} \frac{1}{n}\log \sum_{x\in E_n} e^{S_n\varphi(x)}.$$

□

Tenemos la siguiente versión del Princio Variacional para la presión

Teorema 3.10. *Sean $X$ espacio métrico compacto, $T : X \hookleftarrow$ continua, $\varphi \in C(X)$. Entonces*

$$P_T(\varphi) = \sup\{h_\mu(T) + \int \varphi \, d\mu : \mu \in \mathcal{M}_T(X)\}$$

Demostración. I) Sea $\mu \in \mathcal{M}_T(X)$. Sea $\mathcal{P} = \{P_1, \ldots, P_k\}$ una partición de $(X, \mathcal{A}, \mu)$. Dado $\epsilon \in (0, 1/k \log k)$ escojamos compactos $B_i \subset P_i$, $i = 1, \ldots, k$ tales que $\mu(P_i \setminus B_i) < \epsilon$. Sean

$$B_0 = X \setminus \bigcup_{i=1}^{k} B_i, \quad \mathcal{B} = \{B_0, B_1, \ldots, B_k\}.$$

Entonces tenemos (3.10)

$$h_\mu(T, \mathcal{P}) \leq h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{P} \mid \mathcal{B}) < h_\mu(T, \mathcal{B}) + 1.$$

Sea $d = \min_{1 \leq i,j \leq k} d(B_i, B_j)$ y sea $\delta \in (0, d/2)$ tal que $|\varphi(x) - \varphi(y)| \leq 1$ si $d(x, y) < \delta$. Para cada $B \in \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}$, existe $x_B \in \overline{B}$ tal que $S_n\varphi(x_B) = S_n(B)$. Para $E$ un conjunto $(n, \delta)$-generador cualquiera sea $y_B \in E$ tal que $d(T^i(x_B), T^i(y_B)) < \delta$ para $0 \leq i \leq n$ y entonces $|S_n\varphi(x_B) - S_n\varphi(y_B)| \leq n$. Sean $A = \sum_B e^{S_n\varphi(x_B)}$, $b_B = e^{S_n\varphi(x_B)}/A$. Por la convexidad de $\phi$, tenemos

$$\begin{aligned}
&H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right) + \int S_n\varphi \, d\mu \leq \sum_B \mu(B)(-\log \mu(B) + S_n\varphi(x_B)) \\
&\qquad = -A \sum_B b_B \phi(\mu(B) e^{-S_n\varphi(x_B)}) \leq -A\phi(\sum_B b_B \mu(B) e^{-S_n\varphi(x_B)}) \\
&= -A\,\phi(A^{-1}) = \log A \leq \log \sum_B e^{S_n\varphi(y_B)+n} = n + \log \sum_B e^{S_n\varphi(y_B)}.
\end{aligned} \tag{3.17}$$

Ya que $\delta < d/2$ tenemos que $\#\{B \in \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B} : y_B = x\} \leq 2^n$ para todo $x \in E$. Así, se sigue de (3.17) que

$$\frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right) + \int \varphi \, d\mu \leq 1 + \log 2 + \frac{1}{n} \log \sum_{x \in E} e^{-S_n\varphi(x)}.$$

Entonces

$$h_\mu(T, \mathcal{P}) + \int \varphi \, d\mu \leq h_\mu(T, \mathcal{B}) + 1 + \int \varphi \, d\mu \leq 2 + \log 2 + P_T(\varphi)$$

como $\mathcal{P}$ es arbitrario

$$h_\mu(T) + \int \varphi \, d\mu \leq 2 + \log 2 + P_T(\varphi) \tag{3.18}$$

para todo $T, \varphi$. Aplicando (3.18) a $T^n, S_n\varphi$ tenemos

$$n\left(h_\mu(T) + \int \varphi\, d\mu\right) \le 2 + \log 2 + P_{T^n}(S_n\varphi) = 2 + \log 2 + nP_T(\varphi).$$

de donde se sigue que

$$\sup\{h_\mu(T) + \int \varphi\, d\mu : \mu \in \mathcal{M}_T(X)\} \le P_T(\varphi).$$

II) Sea $\delta > 0$. Sean $\varepsilon > 0$, $n \in \mathbb{N}$. Sea $E_n$ un conjunto $(n, \varepsilon)$ separado tal que $\sum\limits_{x \in E} e^{-S_n\varphi(x)} > s_T(\varphi, n, \varepsilon)e^{-\delta}$. Por el Lema 3.10, hay una medida $\mu \in \mathcal{M}_T(X)$ tal que

$$h_\mu(T) + \int \varphi\, d\mu \ge \limsup_{n\to\infty} \frac{1}{n} \log \sum_{x \in E_n} e^{S_n\varphi(x)} = \limsup_{n\to\infty} \frac{1}{n} \log s_T(\varphi, n, \varepsilon).$$

Así,

$$\sup\{h_\mu(T) + \int \varphi\, d\mu : \mu \in \mathcal{M}_T(X)\} \ge P_T(\varphi).$$

□

## 6. Entropía y descomposición ergódica

PROPOSICIÓN 3.17. *Sean $X$ espacio métrico compacto, $T : X \hookleftarrow$ continua. Sean $\mu \in \mathcal{M}_T(X)$ y $\mathcal{P}$ partición finita tales que $\mu(\partial P) = 0\ \forall P \in \mathcal{P}$. Entonces la funcional $\mathcal{M}_T(X) \to \mathbb{R}$, $\nu \mapsto h_\nu(T, \mathcal{P})$ es semicontinua superior en $\mu$.*

DEMOSTRACIÓN. Para todo átomo de $Q$ de $\bigvee\limits_{i=0}^{n-1} T^{-i}\mathcal{P}$ tenemos $\mu(\partial Q) = 0$. Si $\nu_m \to \mu$ en $\mathcal{M}_T(x)$ entonces

$$\lim_{m\to\infty} H_{\nu_m}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) = H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right).$$

Así $\nu \to H_\nu\left(\bigvee\limits_{i=0}^{n-1} T^{-i}\mathcal{P}\right)$ es continua en $\mu$. Como

$$h_\nu(T, \mathcal{P}) = \inf_n \frac{1}{n} H_\nu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right),$$

tenemos que $\nu \mapsto h_\nu(T, \mathcal{P})$ es semicontinua superior en $\mu$. □

COROLARIO 3.11. *Supongamos que hay una partición finita $\mathcal{P}$ tal que $\mu(\partial P) = 0\ \forall P \in \mathcal{P}$ y $\bigvee_{n=0}^{\infty} T^n\mathcal{P} = \mathcal{A}$, entonces la funcional $\nu \to h_\nu(T)$ es semicontinua superior en $\mu$.*

DEMOSTRACIÓN. Por el corolario 3.3, $h_\nu(T) = h_\nu(T, \mathcal{P})$. □

PROPOSICIÓN 3.18. *La funcional* $\nu \to h_\nu(T)$ *es afín, o sea que* $\forall \mu, \nu \in \mathcal{M}_T(X)$, $t \in [0,1]$, *tenemos*

$$h_{t\mu+(1-t)\nu}(T) = th_\mu(T) + (1-t)h_\nu(T)$$

DEMOSTRACIÓN. Por la convexidad de $\phi$ para cualquier $B \in \mathcal{A}$

$$\phi(t\mu(B) + (1-t)\nu(B)) \leq t\phi(\mu(B)) + (1-t)\phi(\nu(B))$$

También

$$\begin{aligned} t\phi(\mu(B)) + (1-t)\phi(\nu(B)) - \phi(t\mu(B) + (1-t)\nu(B)) \\ = t\mu(B)\log\mu(B) + (1-t)\nu(B)\log\nu(B) \\ - [t\mu(B) + (1-t)\nu(B)]\log(t\mu(B) + (1-t)\nu(B)) = \\ t\mu(B)\log\left[\frac{\mu(B)}{t\mu(B) + (1-t)\nu(B)}\right] + (1-t)\nu(B)\log\left[\frac{\nu(B)}{t\mu(B) + (1-t)\nu(B)}\right] \\ \leq -t\mu(B)\log t - (1-t)\nu(B)\log(1-t) \end{aligned}$$

Así, para cualquier particion $\mathcal{P}$ tenemos

$$\begin{aligned} H_{t\mu+(1-t)\nu}(T) &\geq tH_\mu(T) + (1-t)H_\nu(T) \\ H_{t\mu+(1-t)\nu}(T) &\geq tH_\mu(T) + (1-t)H_\nu(T) - \phi(t) - \phi(1-t) \end{aligned}$$

lo que implica

$$h_{t\mu+(1-t)\nu}(T, \mathcal{P}) = th_\mu(T, \mathcal{P}) + (1-t)h_\nu(T, \mathcal{P}) \tag{3.19}$$

y luego

$$h_{t\mu+(1-t)\nu}(T) \leq th_\mu(T) + (1-t)h_\nu(T)$$

Si $\mathcal{Q}$ es otra partición, se sigue de (3.19) que

$$h_{t\mu+(1-t)\nu}(T, \mathcal{P} \vee \mathcal{Q}) = th_\mu(T, \mathcal{P}) + (1-t)h_\nu(T, \mathcal{Q})$$

y entonces

$$h_{t\mu+(1-t)\nu}(T) \geq th_\mu(T) + (1-t)h_\nu(T).$$

□

TEOREMA 3.12. *Sea* $X$ *espacio métrico compacto. Sean* $m \in \mathcal{M}_T(X)$ *y* $\mu_m$ *su decomposición ergódica. Entonces*

$$h_m(T) = \int_{\mathcal{M}_{erg}(T)} h_\nu(T) d\mu_m(\nu)$$

Deduciremos este Teorema de un resultado general sobre transformaciones afines en el espacio $\mathcal{M}_T(X)$. Sea $W$ una probabilidad en la $\sigma$-álgebra de Borel de $\mathcal{M}_T(X)$. El baricentro de $W$ es la probabilidad $\operatorname{bar} W$ en $X$ dada por

$$\int_X \varphi \, d \operatorname{bar} W = \int_{\mathcal{M}_T(X)} \left( \int_X \varphi d\nu \right) dW(\nu)$$

TEOREMA 3.13. *Sean $H : \mathcal{M}_T(X) \to \mathbb{R}$ un funcional afín, no negativo y semicontnuo superiormente y $W$ una probabilidad en la $\sigma$-álgebra de Borel de $\mathcal{M}_T(X)$. Entonces,*

$$H(\operatorname{bar} W) = \int_{\mathcal{M}_T(X)} H(\nu) dW(\nu)$$

PRUEBA DEL TEOREMA 3.12 USANDO EL TEOREMA 3.13.

LEMA 3.11. *Para toda partición finita $\mathcal{P}$ se tiene*

$$h_m(T, \mathcal{P}) = \int_{\mathcal{M}_T(X)} h_\nu(T, \mathcal{P}) d\mu_m(\nu)$$

DEMOSTRACIÓN. El Teorema 2.6 dice que $m$ es el baricentro de su descomposición ergódica $\mu_m$. El funcional $H : \mathcal{M}_T(X) \to \mathbb{R}$ dado por $H(\nu) = h_\nu(T, \mathcal{P})$ es afín y no negativo. Para aplicar el Teorema 3.13 resta verificar que $H$ es semicontinuo superiormente.

Primero suponemos que $X = [k]^{\mathbb{N}}$, que $T$ es el corrimiento $\sigma$ y que $\mathcal{P}$ es la partición al tiempo cero $\{C(0; 0), \ldots, C(0; k-1)\}$ cuyos elementos son abiertos y cerrados y por lo tanto $\partial Q = \emptyset$ para todo $Q \in \mathcal{P}$. Por la Proposición 3.17 tenemos que $H$ es semicontinuo superior y así aplicamos el Teorema 3.13 para obtener

$$h_m(T, \mathcal{P}) = H(\operatorname{bar} \mu_m) = \int\limits_{\mathcal{M}_T(X)} H(\nu) d\mu_m(\nu) = \int\limits_{\mathcal{M}_T(X)} h_\nu(T, \mathcal{P}) d\mu_m(\nu)$$

Ahora tratamos el caso general reduciendolo al caso anterior. Dada la partición finita $\mathcal{P}$, consideremos $\Sigma = \mathcal{P}^{\mathbb{N}}$ definamos $f : X \to \Sigma$ mediante $f(x) = (\mathcal{P}(T^n x))_{n \in \mathbb{N}}$ de tal forma que $f \circ T = \sigma \circ f$ donde $\sigma : \Sigma \hookleftarrow$ es el corrimiento. Por lo tanto, si $\nu \in \mathcal{M}_T(X)$ entonces $\nu' = f_* \nu \in \mathcal{M}_\sigma(\Sigma)$. Además si $\nu$ es ergódica entonces $\nu'$ es ergódica pues si $B' = \sigma(B')$ y $B = f^{-1}(B')$, entonces $B = T^{-1}(B)$ y así $\nu'(B') = \nu(B) = 0, 1$.

Por construcción $\mathcal{P} = f^{-1}(\mathcal{P}')$ donde $\mathcal{P}'$ es la partición al tiempo cero. Más generalmente, $\bigvee\limits_{i=0}^{n-1} T^{-i}\mathcal{P} = f^{-1}\left(\bigvee\limits_{i=0}^{n-1} T^{-i}\mathcal{P}\right)$ y así

$$H_\nu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) = H_{\nu'}\left(\bigvee_{i=0}^{n-1} \sigma^{-i}\mathcal{P}'\right)$$

$$h_\nu(T, \mathcal{P}) = h_{\nu'}(\sigma, \mathcal{P}') \tag{3.20}$$

Sean $m' = f_* m$, $\mu'_m = (f_*)_* \mu_m$. Para $\psi : \Sigma \to \mathbb{R}$ Borel medible tenemos

$$\begin{aligned} \int_\Sigma \psi dm' &= \int_X \psi \circ f dm = \int_{\mathcal{M}_T(X)} \Big( \int_X \psi \circ f d\nu \Big) d\mu_m(\nu) \\ &= \int_{\mathcal{M}_T(X)} \Big( \int_\Sigma \psi d(f_*\nu) \Big) d\mu_m(\nu) = \int_{\mathcal{M}_\sigma(\Sigma)} \Big( \int_\Sigma \psi d\nu' \Big) d\mu'_m(\nu') \end{aligned}$$

Lo que significa que $\mu'_m$ es la decomposición ergódica de $m'$. Entonces, de acuerdo con el caso anterior y (3.20)

$$h_m(T,\mathcal{P}) = h_{m'}(\sigma,\mathcal{P}') = \int\limits_{\mathcal{M}_\sigma(\Sigma)} h_{\nu'}(\sigma,\mathcal{P}')d\mu'_m(\nu') = \int\limits_{\mathcal{M}_T(X)} h_\nu(T,\mathcal{P})d\mu_m(\nu)$$

□

Prosiguiendo con la prueba del teorema 3.12 consideremos una sucesión creciente de particiones finitas $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots$ de $X$ tal que el diámetro de $\mathcal{P}_n(x)$ converge a cero para todo $x \in X$. Tal sucesión puede construirse a partir de bolas centradas en puntos de un subconjunto denso numerable, con radios convergiendo a cero. Por el Lema 3.11

$$h_\mu(T,\mathcal{P}_n) = \int_{\mathcal{M}_T(X)} h_\nu(T,\mathcal{P}_n)d\nu_m(\nu). \tag{3.21}$$

Para cualquier $\nu \in \mathcal{M}_T(x)$ la sucesión $h_\nu(T,\mathcal{P}_n)$ no decrece y por el Corolario 3.5 su límite es $h_\nu(T)$. Usando el teorema de la convergencia monótona podemos pasar al límite en (3.21) para obtener

$$h_\mu(T) = \int_{\mathcal{M}_T(X)} h_\nu(T)d\nu_m(\nu).$$

□

Lema 3.12. *Sea* $: \mathcal{M}(X) \to \mathbb{R}$ *funcional afín no negativo y sean* $\nu_i \in \mathcal{M}(x)$, $t_i \geq 0$, $i \in \mathbb{N}$ *tales que* $\sum\limits_{i=1}^{\infty} t_i = 1$. *Entonces*

$$H(\sum_{i=1}^{\infty} t_i\nu_i) \geq \sum_{i=1}^{\infty} t_i H(\nu_i)$$

Demostración. Sea $s_n = \sum\limits_{i=1}^{n} t_i$; si $s_n < 1$ tome $\eta_n = \sum\limits_{i>n} t_i\nu_i/(1-s_n)$ y si no tome $\eta_n$ cualquier probablidad. Entonces

$$\sum_{i=1}^{\infty} t_i\nu_i = \sum_{i=1}^{n} t_i\nu_i + (1-s_n)\eta_n.$$

Como $H$ es afín, para cualquier $n$

$$H(\sum_{i=1}^{\infty} t_i\nu_i) = \sum_{i=1}^{n} t_i H(\nu_i) + (1-s_n)H(\eta_n) \geq \sum_{i=1}^{n} t_i H(\nu_i).$$

Ahora hacemos tender $n$ a infinito □

Corolario 3.14. *Si* $H : \mathcal{M}_T(X) \to \mathbb{R}$ *es un funcional afín no negativo entonces está acotado.*

DEMOSTRACIÓN. Si $H$ no está acotado existen $\nu_i \in \mathcal{M}_T(X)$ tales que $H(\nu_i) \geq 2^i$ para todo $i \in \mathbb{N}$. Entonces

$$H(\sum_{i=1}^{\infty} 2^i \nu_i) \geq \sum_{i=1}^{\infty} 2^i H(\nu_i) = \infty$$

$\square$

PRUEBA DEL TEOREMA 3.13. Sea $\mu = \operatorname{bar} W$. Como $H$ se semicontinuo superiormente para todo $\varepsilon > 0$, hay una $\delta$ tal que $H(\eta) < N(\nu) + \varepsilon$ si $d(\eta, \nu) < \delta$. También para cada $\nu \in \mathcal{M}_T(X)$ hay una vecindad $V(\nu)$ contenida en $B_\delta(\nu) = \{\eta \in \mathcal{M}_T(X) : d(\eta, \nu) < \delta\}$ tal que si $\eta \in V(\nu)$ entonces $H(\eta) < N(\nu) + \varepsilon$. La colección de tales vecindades es una cubierta abierta de $\mathcal{M}_T(X)$, por lo tanto hay una subcubierta finita $V(\nu_i)$, $i = 1, \ldots, n$. Consideremos la particion de $\mathcal{M}_T(X)$ definida inductivamente por $P_1 = V(\nu_1)$, $P_i = V(\nu_i) \setminus \bigcup_{j<i} P_i$.

$$\begin{aligned}
d(\sum_{i=1}^{n} W(P_i)\nu_i, \mu) &= \sum_{j=1}^{\infty} 2^{-j} \Big| \int g_j d\mu - \sum_{i=1}^{n} W(P_i) \int g_j d\nu_i \Big| \\
&= \sum_{j=1}^{\infty} 2^{-j} \Big| \sum_{i=1}^{n} \int_{P_i} \Big( \int g_j d\eta - \int g_j d\nu_i \Big) dW(\eta) \Big| \\
&\leq \sum_{i=1}^{n} \int_{P_i} \sum_{j=1}^{\infty} 2^{-j} \Big| \int g_j d\eta - \int g_j d\nu_i \Big| dW(\eta) \\
&= \sum_{i=1}^{n} \int_{P_i} d(\eta, \nu_i) dW(\eta) < \delta \sum_{i=1}^{n} W(P_i) = \delta
\end{aligned}$$

Por lo tanto

$$\sum_{i=1}^{n} W(P_i) H(\nu_i) = H(\sum_{i=1}^{n} W(P_i)\nu_i) < H(\mu) + \varepsilon$$

Por otro lado

$$\begin{aligned}
\int H(\eta) dW(\eta) - \sum_{i=1}^{n} W(P_i) H(\nu_i) &= \sum_{i=1}^{n} \int_{P_i} (H(\eta) - H(\nu_i)) dW(\eta) \\
&< \sum_{i=1}^{n} \varepsilon W(P_i) = \varepsilon
\end{aligned}$$

Sumando las dos últimas desigualdades $\int H(\eta) dW(\eta) < H(\mu) + 2\varepsilon$. Como $\varepsilon > 0$ es arbitrario, $\int H(\eta) dW(\eta) \leq H(\mu)$

Para probar la otra desigualdad considere cualquier sucesión $(\mathcal{P}_n)_n$ de particiones finitas de $\mathcal{M}_T(X)$ tal que el diámetro de $\mathcal{P}_n(\nu)$ converge a cero, cualquiera que sea

$\nu \in \mathcal{M}_T(X)$. Para $n$ fijo consideremos la probabilidad condicional $W_P$ para cada $P \in \mathcal{P}_n$,

$$W_P(A) = \frac{W(A \cap P)}{W(A)},$$

siendo claro que $W = \sum_{P \in \mathcal{P}_n} W(P) W_P$. Como el baricentro es una función afín

$$\mu = \operatorname{bar} W = \sum_{P \in \mathcal{P}_n} W(P) \operatorname{bar} W_P$$

y por lo tanto

$$H(\mu) = \sum_{P \in \mathcal{P}_n} W(P) H(\operatorname{bar} W_P)$$

la cual, definiendo $H_n(\nu) = H(\operatorname{bar} W_{\mathcal{P}_n}(\nu)))$, puede ser reescrita en la forma:

$$H(\mu) = \int H_n(\nu) dW(\nu). \tag{3.22}$$

Por el Corolario 3.14, $H_n(\nu) \leq \sup H < \infty$.

Para $\nu$ fijo, como $H$ es semicontinuo superiormente, dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $d(\eta, \nu) < \delta$ entonces $H(\eta) < H(\nu) + \varepsilon$. Ya que $\operatorname{diam} \mathcal{P}_n(\nu)$ converge a cero, para $n$ suficientemente grande tenemos que $\mathcal{P}_n(\nu) \subset B_\delta(\nu)$. Suponiendo que $W(\mathcal{P}_n(\nu)) > 0$ tenemos que $W_{\mathcal{P}_n(\nu)}(B_\delta(\nu)) = 1$ para $n$ suficientemente grande. Entonces

$$\begin{aligned}
d(\operatorname{bar} W_{\mathcal{P}_n(\nu)}, \nu) &= \sum_{j=1}^{\infty} 2^{-j} \Big| \int g_j d \operatorname{bar} W_{\mathcal{P}_n(\nu)} - \int g_j d\nu \Big| \\
&= \sum_{j=1}^{\infty} 2^{-j} \Big| \int \Big( \int g_j d\eta - \int g_j d\nu \Big) dW_{\mathcal{P}_n(\nu)}(\eta) \Big| \\
&\leq \int_{B_\delta(\nu)} \sum_{j=1}^{\infty} 2^{-j} \Big| \int g_j d\eta - \int g_j d\nu \Big| dW_{\mathcal{P}_n(\nu)}(\eta) \\
&\leq \int_{B_\delta(\nu)} d(\eta, \nu) dW_{\mathcal{P}_n(\nu)}(\eta) < \delta W_{\mathcal{P}_n(\nu)}(B_\delta(\nu)) = \delta
\end{aligned}$$

Por lo tanto $H_n(\nu) \leq H(\nu) + \varepsilon$ para $n$ suficientemente grande. Así

$$\limsup_n H_n(\nu) \leq H(\nu).$$

Aplicando el lema de Fatou a $\sup H - H_n$ tenemos

$$\limsup_n \int H_n(\nu) dW(\nu) \leq \int \limsup_n H_n(\nu) dW(\nu) \leq \int H(\nu) dW(\nu)$$

que junto con (3.22) implica

$$H(\mu) \leq \int H(\nu) dW(\nu).$$

$\square$

## 7. Jacobianos

Sean $(X, \mathcal{A})$ espacio medible y $T : X \hookleftarrow$ medible. Diremos que $T$ es localmente invertible si existe una cubierta numerable $\{U_k : k \geq 1\}$ de $X$ por conjuntos en $\mathcal{A}$ tal que $T|U_k$ es inyectiva con $T(U_k) \in \mathcal{A}$ y su inversa $S_k : T(U_k) \to U_k$ es medible. Los $U_k$ se llaman dominios de inyectividad. Note que si $A \in \mathcal{A}$ es cualquier dominio de inyectividad entonces $T(A) \in \mathcal{A}$ y $\forall y \in X$, $T^{-1}(y)$ es numerable.

Sea $\eta$ una probabilidad no necesariamente $T$-invariante. Una función medible $g : X \to [0, \infty)$ es un *jacobiano* de $T$ respecto a $\eta$ si su restrición a cualquier dominio de inyectividad $A \in \mathcal{A}$ es $\eta$-integrable y

$$\eta(T(A)) = \int_A g d\eta.$$

Decimos que la medida $\eta$ no es singular para la transformación $T$ si para cualquier dominio de inyectividad $A$ con $\eta(A) = 0$ se tiene $\eta(T(A)) = 0$.

Proposición 3.19. *Sea $T : X \hookleftarrow$ una transformación localmente invertible y sea $\eta$ una medida en $(X\mathcal{A})$, no singular para $T$. Entonces, existe un jacobiano de $T$ con respecto a $\eta$ y dos jacobianos cualesquiera coinciden $\eta$-c.t.p.*

Demostración. Para una cubierta $\{U_k : k \geq 1\}$ de $X$ por dominios de inyectividad de $T$, definimos una partición de $X$por dominios de inyectividad mediante $P_1 = U_1$, $P_k = U_k \setminus (U_1 \cup \cdots \cup U_{k-1})$, $k > 1$. En cada $P_k$ definimos la medida $\eta_k$ por $\eta_k(A) = \eta(T(A))$. Como $\eta$ no es singular para $T$, $\eta_k << \eta|P_k$. Sea $g_k$ la derivada de Radon-Nykodim de $\eta_k$ respecto a $\eta|P_k$, por lo que

$$\eta(T(A)) = \int_A g_k \, d\eta$$

para cualquier $A \subset P_k$ medible. Definamos ahora $g : X \to [0, \infty)$ por $g|P_k = g_k$. Es claro que $g$ es un jacobiano de $T$ respecto a $\eta$.

Sea $h$ otro jacobiano de $T$ respecto a $\eta$ y supongamos que $\eta(\{x : g(x) \neq h(x)\}) > 0$. Permutando los papeles de $h$ y $g$ si es necesario, para algún $k \geq 1$ tenemos que si $B_k = \{x \in P_k : g(x) < h(x)\}$, $\eta(B_k) > 0$. Entonces

$$\eta(T(B_k)) = \int_{B_k} g_k \, d\eta < \int_{B_k} h_k \, d\eta = \eta(T(B_k))$$

$\square$

Denotaremos por $J_\eta T$ al jacobiano de $T$ com respecto a $\eta$.

Si $\#T^{-1}(y) \leq l$ entonces

$$\int J_\eta T\, d\eta = \sum_k \int_{P_k} J_\eta T\, d\eta = \sum_k \eta(T(P_k)) \leq l$$

TEOREMA 3.15. *Sea $T: X \to X$ una transformación localmente invertible y sea $\mu$ probabilidad $T$-invariante. Suponga que existe alguna partición finita o numerable por dominios de inyectividad tal que $\mathcal{A} = \bigvee_{i=0}^{\infty} T^{-i}\mathcal{P}$. Entonces*

$$h_\mu(T) = \int \log J_\mu T\, d\mu.$$

DEMOSTRACIÓN. Denotando $\mathcal{Q}_n = \bigvee\limits_{i=1}^{n} T^{-i}\mathcal{P}$, por la Proposición 3.2 (e) y el Corolario 3.3 tenemos

$$h_\mu(T) = h_\mu(T, \mathcal{P}) = \lim_n H_\mu(\mathcal{P} \mid \mathcal{Q}_n).$$

Por la Observación 3.1 (c)

$$H_\mu(P \mid \mathcal{Q}_n) = -\sum_{P\in\mathcal{P}} \int \phi \circ E_\mu(\chi_P \mid \mathcal{Q}_n)\, d\mu. \tag{3.23}$$

Para terminar la demostración utilizaremos los siguientes Lemas que demostraremos después.

LEMA 3.13. *Sea $\mathcal{Q}_1 \prec \mathcal{Q}_2 \prec \cdots$ sucesión creciente de particiones numerables. Dada $\psi: X \to \mathbb{R}$ medible y acotada, el límite $e[\psi] = \lim_n E_\mu(\psi \mid \mathcal{Q}_n)$ existe $\mu$-c.t.p.*

LEMA 3.14. *Sean $\psi: X \to \mathbb{R}$ medible, acotada y $\eta$ una probabilidad $T$-invariante. Entonces*

$$e[\psi] = \hat{\psi} \circ T, \eta - c.t.p., \text{ donde } \hat{\psi}(y) = \sum_{z\in T^{-1}(y)} \frac{\psi}{J_\eta T}(z)$$

De (3.23) obtenemos por el Teorema de la Convergencia dominada

$$h_\mu(T) = -\sum_{p\in\mathcal{P}} \int \phi \circ e[\chi_P]\, d\mu$$

Como $T|P$ es inyectiva $\forall P \in \mathcal{P}$, $\#(P \cap T^{-1}(y)) = 0, 1$. Por el Lema 3.14, $e[\chi_P] = \hat{\chi}_P \circ T$ donde

$$\hat{\chi}_P(y) = \begin{cases} 1/J_\mu T((T|P)^{-1}(y)) & y \in T(P) \\ 0 & y \notin T(P) \end{cases}$$

Entonces

$$\int \phi \circ e[\chi_P]\, d\mu = \int \phi \circ \hat{\chi}_P\, d\mu = -\int_{T(P)} \frac{\log J_\mu}{J_\mu} \circ (T|P)^{-1}\, d\mu = -\int_P \log J_\mu\, d\mu$$

y por lo tanto

$$h_\mu(T) = \sum_{P\in\mathcal{P}} \int_P \log J_\mu \, d\mu = \int \log J_\mu \, d\mu$$

□

Demostración del Lema 3.13. Primero suponga que $\psi \geq 0$. Para $\alpha < \beta$ sea

$$S(\alpha,\beta) = \{x \in X \liminf_n E_\mu(\psi, \mid \mathcal{Q}_n)(x) < \alpha < \beta < \limsup E_\mu(\psi \mid \mathcal{Q}_n)(x).$$

El conjunto de puntos donde límite $e[\psi]$ existe es $X_\psi = X \setminus \bigcup_{\alpha,\beta\in\mathbb{Q}} S(\alpha,\beta)$.

Para probar que $\mu(X_\psi) = 1$ basta probar que $\mu(S(\alpha,\beta)) = 0$ para todo $\alpha < \beta$.

Para $\alpha, \beta$ fijos escribimos $S = S(\alpha,\beta)$. Dado $x \in S$, tome una sucesión de enteros $1 \leq n_1^x < m_1^x < \cdots < n_i^x < m_i^x < \cdots$ tales que

$$E_\mu(\psi \mid \mathcal{Q}_{n_i^x})(x) < \alpha, E_\mu(\psi \mid \mathcal{Q}_{m_i^x})(x) > \beta \quad \forall i \geq 1$$

Defina $A_i(x) = \mathcal{Q}_{n_i^x}(x)$, $B_i(x) = \mathcal{Q}_{n_i^x}(x)$, $A_i = \bigcup_{x\in S} A_i(x)$, $B_i = \bigcup_{x\in S} B_i(x)$.

Por construcción $S \subset A_{i+1} \subset B_i \subset A_i$ para todo $i \geq 1$. En particular

$$S \subset \tilde{S} = \bigcap_{i=1}^\infty A_i = \bigcap_{i=1}^\infty B_i$$

Como la sucesión $\mathcal{Q}_n$, $n \geq 1$, es creciente, dados dos cualesquiera de los conjuntos $A_i(x)$ que formam $A_i$, ó son disjuntos o uno de ellis está contenido en el otro. Entonces los conjuntos $A_i(x)$ maximales son disjuntos por pares y constityuen una partición de $A_i$. Luego, sumando solo sobre estos conjuntos maximales con medida positiva,

$$\int_{A_i} \psi \, d\mu = \sum_{A_i(x)} \int_{A_i(x)} \psi \, d\mu \leq \sum_{A_i(x)} \alpha\mu(A_i(x)) = \alpha\mu(A_i),$$

para cualquier $i \geq 1$. Analogamente,

$$\int_{B_i} \psi \, d\mu = \sum_{B_i(x)} \int_{B_i(x)} \psi \, d\mu \geq \sum_{B_i(x)} \beta\mu(B_i(x)) = \beta\mu(B_i).$$

Como $A_i \supset B_i$ y estamos suponiendo que $\psi \geq 0$, se sigue que

$$\alpha\mu(A_i) \geq \int_{Ai} \psi \, d\mu \geq \int_{B_i} \psi \, d\mu \geq \beta\mu(B_i),$$

para todo $i \geq 1$. Tomando el límite cuando $i \to \infty$, obtenemos que $\alpha\mu(\tilde{S}) \geq \beta\mu(\tilde{S})$. Esto implica que $\mu(\tilde{S}) = 0$ y así $\mu(S) = 0$, lo que prueba la afirmación cuando $\psi$ es no-negativa. En el caso general, escribiendo $\psi = \psi^+ - \psi^-$ con $\psi^\pm$ medibles, no-negativas y acotadas, tenemos $E_\mu(\psi \mid \mathcal{Q}_n) = E_\mu(\psi^+ \mid \mathcal{Q}_n) - E_\mu(\psi^- \mid \mathcal{Q}_n)$ para todo $n \geq 1$. □

Demostración del Lema 3.14. Definiendo $\mathcal{P}_n = \bigvee_{i=0}^{n} T^{-i}\mathcal{P}$ tenemos que $\mathcal{Q}_n(x) = T^{-1}(\mathcal{P}_{n-1}(T(x)))$ y $\mathcal{P}_n(x) = \mathcal{P}(x) \cap \mathcal{Q}_n(x)$. Por lo tanto

$$\begin{aligned}\int_{\mathcal{P}_{n-1}(T(x))} \hat{\psi}\, d\eta &= \sum_{P\in\mathcal{P}} \int_{T(P)\cap\mathcal{P}_{n-1}(T(x))} \frac{\psi}{J_\eta T} \circ (T|P)^{-1}\, d\eta \\ &= \sum_{P\in\mathcal{P}} \int_{P\cap\mathcal{Q}_n(x)} \psi\, d\eta = \int_{\mathcal{Q}_n(x)} \psi\, d\eta\end{aligned}$$

Como $\eta$ es $T$-invariante, $\eta(\mathcal{Q}_n(x)) = \eta(\mathcal{P}_{n-1}(T(x)))$ y así

$$\begin{aligned}E_\mu(\psi \mid \mathcal{Q}_n) &= \sum_{Q\in\mathcal{Q}_n} \frac{1}{\eta(Q)}\Big(\int_Q \psi d\eta\Big)\chi_Q \\ &= \sum_{P\in\mathcal{P}_{n-1}} \frac{1}{\eta(P)}\Big(\int_P \hat{\psi} d\eta\Big)\chi_P \circ T = E_\mu(\hat{\psi} \mid \mathcal{P}_{n-1}) \circ T\end{aligned}$$

Como $\bigcup_n \mathcal{P}_n$ genera $\mathcal{A}$ tenemos que $\lim_n E_\mu(\hat{\psi} \mid \mathcal{P}_n) = \hat{\psi}$ $\eta$-c.t.p. □

Ejercicio 3.1. Sea $T : X \to X$ una transformación localmente invertible y sea $\eta$ una probabilidad boreliana en $X$ no singular con respecto a $T$. Muestre que para toda función medible acotada $\psi : X \to R$,

$$\int \psi d\eta = \int \sum_{x\in T^{-1}(y)} \frac{\psi}{J_\eta T}(x) d\eta(y).$$

Capítulo 4

# Transformaciones expansoras

## 1. Transformaciones expansoras en variedades

Sea $M$ una variedad compacta y sea $f : M \to M$ de clase $C^1$. Decimos que $f$ es expansora si existe $\sigma > 1$ y una métrica Riemanniana en $M$ tal que

$$\|Df(x)v\| \geq \sigma \|v\| \quad \forall x \in M, v \in T_x M$$

En particular, $f$ es un difeomorfismo local. Llamaremos medida de Lebesgue en $M$ a la medida de volume $m$ inducida por la métrica Riemanniana. La elección de la métrica no es importante. Recordemos que si $\mu$ es ergódica, entonces su cuenca $B(\mu)$ tiene $\mu$-medida total. El resultado principal de esta sección es el siguiente

Teorema 4.1. *Sean $M$ una variedad compacta conexa y $f : M \to M$ transformación expansora con jacobiano* $\det Df$ *Hölder. Entonces $f$ admite una única probabilidad invariante $\mu$ absolutamente continua respecto a la medida de Lebesgue $m$. Además $\mu$ es ergódica, su soporte es todo $M$ y su cuenca tiene medida de Lebesgue total.*

En general decimos que una probabilidad $\mu$ invariante por un difeomorfismo local $f$ es una *medida física* si su cuenca tiene medida de Lebesgue positiva. El Teorema 4.1 afirma que en el presente contexto existe una única medida física, que es absolutamente continua y cuya cuenca tiene medida de Lebesgue total. Este último hecho puede escribirse en la siguiente forma:

$$\mu = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} \quad m- \text{ c.t.p. } x$$

Comenzamos la demostración de Teorema 4.1 con el siguiente

Lema 4.1. *Sea $f : M \to M$ un difeomorfismo local de una variedad Riemanniana compacta y supongamos que $\|Df^{-1}\| \leq 1/\sigma$. Entonces existe $\rho > 0$ tal que, para cualquier preimagen $x$ de un punto $y \in M$, existe una transformación $h : B_\rho(y) \to M$ de clase $C^1$ tal que $f \circ h = id$, $h(y) = x$ y*

$$d(h(y_1), h(y_2)) \leq \sigma^{-1} d(y_1, y_2) \quad \forall y_1, y_2 \in B_\rho(y). \tag{4.1}$$

Transformaciones $h$ como en este enunciado se llaman ramas inversas de $f$ y (4.1) dice que estas ramas son contracciones con una taza uniforme de contracción. En particular, podemos definir ramas inversas $h^n$ de cualquier iterado $f^n$, $n \geq 1$, de la siguiente forma. Dado $y \in M$ y $x \in f^{-n}(y)$, sean $h_1, ..., h_n$ ramas inversas de $f$ con

$$h_j(f^{n-j+1}(x)) = f^{n-j}(x)$$

para todo $1 \leq j \leq n$. Como cada $h_j$ es una contración, su imagen está contenida en una bola de radio menor que $\rho$ alrededor de $f^{n-j}(x)$. Entonces $h^n = h_n \circ \cdots \circ h_1$ está bien definida en la bola de radio $\rho$ alrededor de $y$. Es claro que $f^n \circ h^n =$id y $h^n(y) = x$. Ademas,

$$d(h^n(y_1), h^n(y_2)) \leq \sigma^{-n} d(y_1, y_2) \quad \forall y_1, y_2 \in B_\rho(y).$$

Lema 4.2. *Si $f : M \to M$ es una transformación expansora $C^1$ en una variedad compacta entonces $f$ es expansiva.*

Demostración. Por la compacidad de $M$ existe $\rho > 0$ tal que, para cualquier preimagen $x$ de un punto $y \in M$, existe uma transformación $h : B_\rho(y) \to M$ de classe $C^1$ tal que $f \circ h =$ id y $h(y) = x$. Además

$$d(h(y_1), h(y_2)) \leq \sigma^{-1} d(y_1, y_2) \quad y_1, y_2 \in B_\rho(y).$$

Así, si $d(f^n(x), f^n(y)) \leq \rho^n$ para todo $n \geq 0$

$$d(x, y) \leq \sigma^{-n} d(f_n(x), f_n(y)) \leq \sigma^{-n}\rho,$$

lo que implica imediatamente que $x = y$. $\square$

Siendo $f$ difeomorfismo local y $M$ compacta, $C^{-1} \leq |\det Df| \leq C$. La hipótesis de que $\det Df$ es Hölder implica que también $\log|\det Df|$ lo es: existen $C_0, \nu > 0$ tales que

$$\big|\log|\det Df(x)| - \log|\det Df(y)|\big| \leq C_0 d(x, y)^\nu, \quad \forall x, y \in M$$

Lema 4.3 (Lema de distorsión). *Existe $C_1 > 0$ tal que para todo $n \geq 1$, todo $y \in M$ y toda rama inversa $h^n : B_\rho(y) \to M$ de $f^n$ se tiene*

$$\log \frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} \leq C_1 d(y_1, y_2)^\nu \leq C_1(2\rho)^\nu \quad \forall y_1, y_2 \in B_\rho(y).$$

Demostración. Escribimos $h^n$ como composición de ramas inversas de $f$, $h^n = h_n \circ \cdots \circ h_1$. Sean $h^i = h_i \circ \cdots \circ h_1$ para $1 \leq i < n$, y $h_0 =$ id. Entonces

$$\log \frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} = \sum_{i=1}^{n} \log|\det Dh_i(h^{i-1}(y_1)| - \log|\det Dh_i(h^{i-1}(y_2)|.$$

Note que $\log|\det Dh_i| = -\log|\det Df| \circ h_i$ y que por el Lema 4.1 cada $h_j$ es una contracccción con taza $\sigma^{-1}$. Luego,

$$\log \frac{|\det Dh^n(y_1)|}{|\det Dh^n(y_2)|} \leq \sum_{i=1}^{n} C_0 d(h^i(y_1), h^i(y_2))^\nu \leq \sum_{i=1}^{n} C_0 \sigma^{-i\nu} d(y_1, y_2)^\nu$$

Tomando $C_1 = C_0 \sum_{i=1}^{\infty} \sigma^{-i\nu}$ obtenemos el Lema. $\square$

Corolario 4.2. *Existe $C_2 > 0$ tal que, para todo $y \in M$ y cualesquiera conjuntos medibles $B_1, B_2 \subset B_\rho(y)$ tenemos*

$$\frac{1}{C_2}\frac{m(B_1)}{m(B_2)} \leq \frac{m(h^n(B_1))}{m(h^n(B_2))} \leq C_2 \frac{m(B_1)}{m(B_2)}$$

Demostración. Por el Lema de distorsión

$$m(h^n(B_1)) = \int_{B_1} |\det Dh^n| dm \leq \exp(C_1(2\rho)^\nu)|\det Dh^n(y)|m(B_1)$$

$$m(h^n(B_1)) = \int_{B_1} |\det Dh^n| dm \geq \exp(-C_1(2\rho)^\nu)|\det Dh^n(y)|m(B_2)$$

Tomando $C_2 = \exp(2C_1(2\rho)\nu)$ y dividiendo las dos desigualdades

$$\frac{m(h^n(B_1))}{m(h^n((B_2))} \leq C_2 \frac{m(B_1)}{m(B_2)}.$$

Invirtiendo los papeles de $B_1$ y $B_2$ obtenemos la otra desigualdad. □

Proposición 4.1. *Existe $C_3 > 0$ tal que $(f_*^n m)(B) \leq C_3 m(B)$ para todo conjunto medible $B \subset M$ y todo $n \geq 1$.*

Demostración. Sin perdida de generalidad suponemos que $B$ está contenido em alguna bola $B_0 = B_\rho(z)$. Usando el Corolario 4.2 vemos que para toda ramas inversa de $f^n$ alrededor de $z$

$$\frac{m(h^n(B))}{m(h^n((B_0))} \leq C_2 \frac{m(B)}{m(B_0)}.$$

Como $(f_*^n m)(B) = m(f^{-n}(B))$ es la suma sobre todas las ramas inversas de $m(h^n(B))$ y lo mismo ocurre para $B_0$, tenemos

$$\frac{m(f_*^n(B))}{m(f_*^n((B_0))} \leq C_2 \frac{m(B)}{m(B_0)}.$$

con $f_*^n m(B_0) \leq f_*^n m(M) \leq 1$. Además la medida de Lebesgue de las bolas con radio fijo $\rho$ esta separada de 0 por una constante $a(\rho)$. Basta entonces tomar $C_3 = \exp(C_1(2\rho)^\nu/a(\rho)$ para concluir la demostración. □

Dada uma función $\varphi$ y una medida $\nu$, representamos por $\varphi\nu$ a la medida definida por $(\varphi\nu)(B) = \int_B \varphi d\nu$.

Lema 4.4. *Sea $\nu$ una probabilidad en el espacio métrico compacto $X$, y sea $\varphi \in L^1(\nu)$. Supongamos que $\mu_n \to \mu$ en $\mathcal{M}_T(X)$ y $\mu_n \leq \varphi\nu$ para todo $n \geq 1$. Entonces $\mu \leq \varphi\nu$.*

Demostración. Por la Proposición 3.14, para cualquier abierto $U$ tenemos que $\mu(U) \leq \liminf_n \mu_n(U) \leq \varphi\nu(U)$. Luego, para cualquier cerrado $F$

$$\mu(F) = \inf\{\mu(U) : U \text{ abierto} \supset F\} \leq \inf\{\varphi\nu(U) : U \text{ abierto} \supset F\} = \varphi\nu(F)$$

y finalmente, para cualquier boreliano $B$

$$\mu(B) = \sup\{\mu(F) : F \text{ cerrado} \subset B\} \leq \sup\{\varphi\nu(F) : F \text{ cerrado} \subset B\} = \varphi\nu(B)$$

□

Corolario 4.3. *Todo punto de acumulación $\mu$ de la sucesión $\frac{1}{n}\sum_{j=0}^{n-1} f_*^j m$ es una probabilidad invariante para $f$, absolutamente continua respecto a la medida de Lebesgue.*

Demostración. Tomemos $\varphi$ constante igual a $C_3$ y $\nu = m$. Tomemos también $\mu_i = \frac{1}{n_i}\sum_{j=0}^{n_i-1} f_*^j m$, para cualquier subsucesión $(n_i)_i$ tal que $(\mu_i)_i$ converge a una medida $\mu$. La Proposición 4.1 garantiza que $\mu_i \leq \varphi\nu$. Entonces $\mu \leq \varphi\nu = C_3 m$, por el Lema 4.4. Esto implica que $\mu << m$, con densidad acotada por $C_3$ . □

Ahora demostraremos que la medida $\mu$ que acabamos de construir es la única probabilidad invariante absolutamente continua respecto a la medida de Lebesgue y que además es $f$-ergódica.

Comezamos por fijar una partición $\mathcal{P}_0 = \{U_1, \ldots, U_s\}$ de $M$ en regiones con interior no vacio y diámetro menor que $\rho$. Entonces, definimos a $\mathcal{P}_n$ como la partición de $M$ que consiste de las imagenes de cada uno de los $U_i, 1 \leq i \leq s$, por las respectivas ramas inversas de $f^n$. El diámetro de la partición $\mathcal{P}_n$ es menor que $\rho\sigma^{-n}$.

Lema 4.5. *Sea $\mathcal{P}_n, n \geq 1$, una sucesión de particiones de un espacio métrico compacto $X$, con diámetros convergiendo a cero cuando $n \to \infty$. Sea $\nu \in \mathcal{M}(X)$ y sea $B$ boreliano con $\nu(B) > 0$. Entonces existen $V_n \in \mathcal{P}_n$, para $n \geq 1$, tales que*

$$\nu(V_n) > 0, \quad \lim_{n\to\infty} \frac{\nu(B \cap V_n)}{\nu(V_n)} = 1$$

Demostración. Dado $0 < \varepsilon < \nu(B)$, sea $K_\varepsilon$ un subconjunto compacto de $B$ con $\nu(B \setminus K_\varepsilon) < \varepsilon$. Como el diámetro de las particiones converge a cero, para $n$ suficientemente la unión $K_{\varepsilon,n}$ de todos los elementos de $\mathcal{P}_n$ que intersectan $K_\varepsilon$ satisface $\nu(K_{\varepsilon,n} \setminus K_\varepsilon) < \varepsilon$. Suponga que

$$\nu(K_\varepsilon \cap V_n) \leq \frac{\nu(B) - \varepsilon}{\nu(B) + \varepsilon}\nu(V_n)$$

para todo $V_n \in \mathcal{P}_n$ que intersecta $K_\varepsilon$. Se tendria que

$$\nu(K_\varepsilon) \le \sum_{V_n} \nu(K_\varepsilon \cap V_n) \le \sum_{V_n} \frac{\nu(B)-\varepsilon}{\nu(B)+\varepsilon}\nu(V_n) = \frac{\nu(B)-\varepsilon}{\nu(B)+\varepsilon}\nu(K_{\varepsilon,n})$$
$$\le \frac{\nu(B)-\varepsilon}{\nu(B)+\varepsilon}(\nu(K_\varepsilon)+\varepsilon) \le \nu(B)-\varepsilon < \nu(K_\varepsilon).$$

Esta contradición muestra que debe existir algun $V_n \in \mathcal{P}_n$ tal que

$$\nu(V_n) \ge \nu(B \cap V_n) \ge \nu(K_\varepsilon \cap V_n) > \frac{\nu(B)-\varepsilon}{\nu(B)+\varepsilon}\nu(V_n)$$

lo que implica que $\nu(V_n) > 0$. Haciendo tender $\varepsilon$ a cero obtenemos el Lema. □

En lo que sigue, la igualdad mód 0 entre conjuntos se refiere a la medida de Lebesgue. Usaremos el hecho de que para un difeomorfismo local de una variedad $f : M \to M$ se tiene que si $f^{-1}(A) = A \mod 0$ entonces $f(A) = A \mod 0$.

Lema 4.6. *Supongamos que $f^{-1}(A) = A \mod 0$ y $m(A) > 0$, entonces existe $1 \le i \le s$ tal que $m(U_i \setminus A) = 0$.*

Demostración. Por el Lema 4.5, podemos escoger $V_n \in \mathcal{P}_n$ de tal forma que $m(V_n \setminus A)/m(V_n)$ converge a cero. Sea $U_{i(n)} = f^n(V_n)$ . Por el Corolario 4.2 aplicado a la rama inversa de $f^n$ que manda $U_{i(n)}$ a $V_n$ tenemos que

$$\frac{m(U_{i(n)} \setminus A)}{m(U_{i(n)})} \le \frac{m(f^n(V_n \setminus A))}{m(f^n(V_n))} \le \exp(C_1(2\rho)^\nu)\frac{m(V_n \setminus A)}{m(V_n)}$$

también converge a cero. Como $\mathcal{P}_0$ es finita, debe existir $1 \le i \le s$ tal que $i(n) = i$ para una infinidad de valores de $n$. Entonces, $m(U_i \setminus A) = 0$. □

Corolario 4.4. *La transformación $f : M \to M$ admite alguna probabilidad invariante ergódica absolutamente continua respecto a la medida de Lebesgue.*

Demostración. Se sigue del Lema anterior que existen a lo más $s = \#\mathcal{P}_0$ conjuntos invariantes con medida de Lebesgue positiva disjuntos por parejas. Por lo tanto, se puede partir $M$ en un número finito de conjuntos invariantes $A_1, \dots, Ar, r \le s$ con medida de Lebesgue positiva y que son minimales, en el sentido de que no existe subconjunto invariante $B_i \subset A_i$ con $0 < m(B_i) < m(A_i)$. Dada cualquier medida invariante absolutamente continua $\mu$, existe algun $i$ tal que $\mu(A_i) > 0$. Entonces la restrición normalizada

$$\mu_i(B) = \frac{\mu(B \cap A_i)}{\mu(A_i)}$$

de es invariante y absolutamente continua. Además, como $A_i$ es minimal, $\mu_i$ es ergódica. □

Observación 4.1. El argumento anterior muestra que existe a lo más un número finito de probabilidades ergódicas y absolutamente continuas.

Para concluir la demostración del Teorema 4.1 resta probar que tal probabilidad es única. Para eso usamos el hecho de que $f$ es *topológicamente exacta*:

LEMA 4.7. *Dado cualquier abierto no vacio $U$ , existe $N \geq 1$ tal que $f^N(U) = M$ .*

DEMOSTRACIÓN. Sean $x \in U$ y $r > 0$ tales que $B_r(x) \subset U$. Dado $n \geq 1$, suponga que $f^n(U)$ no cubre toda la variedad. Entonces existe alguna curva $\gamma$ uniendo $f^n(x)$ a un punto $y \in M \setminus f^n(U)$, la cual puede tomarse de longitud menor que $\operatorname{diam} M + 1$. Levantando $\gamma$ por el difeomorfismo local $f^n$, obtenemos una curva $\gamma_n$ uniendo $x$ a un punto $y_n \in M \setminus U$. Entonce $r \leq \operatorname{long} \gamma_n \leq \sigma^{-n}(\operatorname{diam} M + 1)$. Esto da una cota superior para el valor posible de $n$. Luego, $f^n(U) = M$ para $n$ suficientemente grande. $\square$

COROLARIO 4.5. *Si $f^{-1}(A) = A \mod 0$ y $m(A) > 0$ , entonces $m(A) = m(M)$.*

DEMOSTRACIÓN. Sea $U$ el interior de um conjunto $U_i$ como en el Lema 4.6, y sea $N \geq 1$ tal que $f^N(U) = M$. Entonces $m(U \setminus A) = 0$, y por lo tanto

$$M \setminus A = f^N(U) \setminus f^N(A) \subset f^N(U \setminus A) \mod 0$$

también tiene medida de Lebesgue cero, pues $f$ es difeomorfismo local. $\square$

COROLARIO 4.6. *Sea $\mu$ probabilidad invariante absolutamente continua. Entonces $\mu$ es ergódica y su cuenca $B(\mu)$ tiene medida de Lebesgue total en $M$. Consecuentemente $\mu$ es única y además su soporte es toda la variedad $M$.*

DEMOSTRACIÓN. Si A es un subconjunto invariante cualquiera entonces, por el Corolario 4.5, $A$ tiene medida de Lebesgue cero o $A^c$ tiene medida de Lebesgue cero. Como $\mu$ es absolutamente continua, se sigue que $\mu(A) = 0$ o $\mu(A^c) = 0$. Esto prueba que $\mu$ es ergódica, luego $\mu(B(\mu)) = 1$. Entonces $B(\mu)$ es un conjunto invariante com medida de Lebesgue positiva y, consecuentemente, debe tener medida de Lebesgue total. Analogamente, como el soporte de $\mu$ es un conjunto compacto invariante, tien que coincidir com $M$. Finalmente, sean $\mu$ y $\nu$ dos probabilidades invariantes absolutamente continuas. Se sigue que ellas son ergódicas y que sus cuencas se intersectan. Para cualquier punto $x \in B(\mu) \cap B(\nu)$,

$$\mu = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} = \nu$$

Por unicidad del límite, se sigue que $\mu = \nu$. $\square$

En la siguiente sección reobtendremos esta probabilidad invariante absolutamente continua $\mu$ en una forma que demuestra que la densidad $h = d\mu/dm$ es Hölder y está separada de cero. En particular, $\mu$ es equivalente a la medida de Lebesgue $m$, no solamente absolutamente continua.

Además, su jacobiano está dado por $J_\mu f = |\det Df|(h \circ f)/h$. Entonces, el Teorema 3.15 da que

$$h_\mu(f) = \int \log J_\mu f d\mu = \int \log|\det Df| d\mu + \int \log h \circ f d\mu - \int \log h$$

y como $\mu$ es $f$-invariante, $h_\mu(f) = \int \log|\det Df| d\mu$

## 2. El formalismo termodinámico

Decimos que una transformación continua $T : X \to X$ del espacio métrico compacto $X$ es *expansora* si existen $\sigma > 1$, $\rho > 0$ tales que para todo $p \in X$, $T(B_\rho(p))$ contiene una vecindad de $\overline{B_\rho(T(p))}$ y

$$d(T(x), T(y)) \geq \sigma d(x, y) \quad \forall x, y \in B_\rho(p). \tag{4.2}$$

Así toda transformación expansora $T$ es inyectiva en cada bola $B_\rho(p)$. Luego, la restrición de $T$ a cualquer compacto $K \subset B_\rho(p)$ es un homeomorfismo sobre su imagen. Tome como $K = T^{-1}(\overline{B_\rho(T(p))}) \cap B_\rho(p)$. Llamamos rama inversa de $T$ en $p$ a la inversa $h_p : \overline{B_\rho(T(p))} \to B_\rho(p)$ de la restrición de $T$ a $K$. Es claro que $h_p(T(p)) = p$ y que $T \circ h_p = \mathrm{id}$. La condición (4.2) implica que

$$d(h_p(z), h_p(w)) \leq \sigma^{-1} d(z, w) \quad \forall z, w \in B_\rho(T(p)). \tag{4.3}$$

LEMA 4.8. *Si $T : X \to X$ es expansora entonces, para todo $y \in X$,*

$$T^{-1}(B_\rho(y)) = \bigcup_{x \in T^{-1}(y)} h_x(B_\rho(y)).$$

DEMOSTRACIÓN. La relación $T \circ h_x = \mathrm{id}$ implica que $h_x(B_\rho(y))$ está contenido en la preimagen de $B_\rho(y)$ para todo $x \in T^{-1}(y)$. Para probar la otra contención, sea $z$ un punto tal que $T(z) \in B_\rho(y)$. Por la definición de transformación expansora, $T(B_\rho(z))$ contiene $B_\rho(T(z))$ y por lo tanto, contiene a $y$. Sea $h_z : B_\rho(T(z)) \to X$ la rama inversa de $T$ que envia $T(z)$ en $z$ y sea $x = h_z(y)$. Tanto $z$ como $h_x(T(z))$ estan en $B_\rho(x) \cap T^{-1}(z)$. Como $T$ es inyectiva en cada bola de radio $\rho$, se sigue que $z = h_x(T(z)$ lo que completa la demonstración. □

DEFINICIÓN 4.1 (Operador de transferencia o de Perron-Frobenius-Ruelle).
Sea $T : X \to X$ transformación expansora. Para $\varphi : X \to \mathbb{R}$ continua definimos $\mathcal{L} = \mathcal{L}_\varphi : C(X) \to C(X)$ mediante

$$\mathcal{L}_\varphi g(y) = \sum_{x \in T^{-1} y} e^{\varphi(y)} g(x).$$

Veamos que efectivamente $\mathcal{L}_\varphi g \in C(X)$ siempre que $g \in C(X)$. Por el Lema 4.8, para cada $y \in X$ existen ramas inversas $h_i : B_\rho(y) \to X$, $i = 1, \ldots, k$ de la transformación

$T$ tales que $\bigcup_{i=1}^{k} h_i(B_\rho(y))$ coincide con la preimagen de la bola $B_\rho(y)$. Entonces

$$\mathcal{L}_\varphi g|B_\rho(y) = \sum_{i=1}^{k} (e^\varphi g) \circ h_i \tag{4.4}$$

que claramente define una función continua. $\mathcal{L}_\varphi$ es un operador lineal acotado. Cuando $\varphi$ es real y $\mathcal{L}_\varphi 1 = 1$ decimos que $\mathcal{L}_\varphi$ está normalizado.

El espacio dual de $C(X)$ se identica de manera natural con el espacio de medidas complejas $\mathcal{M}_\mathbb{C}(X)$. La transformación dual $\mathcal{L}^*$ esta definida por

$$\int g\, d(\mathcal{L}^*\nu) = \int \mathcal{L}g\, d\nu \quad \nu \in \mathcal{M}_\mathbb{C}(X), g \in C(X)$$

TEOREMA 4.7 (Ruelle). *Sean $T : X \to X$ transformación expansora topologicamente exacta y $\varphi : X \to \mathbb{R}$ Hölder. Entonces*

1. *$\exists\lambda$ eigenvalor maximal positivo simple de $\mathcal{L}_\varphi^* : \mathcal{M}_\mathbb{C}(X) \to \mathcal{M}_\mathbb{C}(X)$ con eigenvector $\nu \in \mathcal{M}(X)$.*
2. *$\lambda$ es eigenvalor de $\mathcal{L}_\varphi$ con eigenfunción estrictamente positiva Hölder $h$.*
3. *Si $h$ es como en (2) y tal que $\int h d\nu = 1$ entonces*
$$\lim_{n\to\infty} \lambda^{-n}\mathcal{L}_\varphi^n(g) = h \int g\, d\nu \quad \forall g \in C(X).$$
4. *Si $h$ es como en (3), $\mu = h\nu$ es $T$-invariante y además*
$$h_\mu(T) + \int \varphi\, d\mu = P_T(\varphi)$$

Para comenzar la demostración, utilizaremos un resultado del Análisis Funcional.

DEFINICIONES 4.2. Sea $E$ un espacio de Banach.

1. Un subconjunto cerrado y convexo $C$ se llama *cono* de $E$ si
$$\forall\lambda \geq 0\, \lambda C \subset C,\, C \cap (-C) = \{0\}$$
2. Decimos que el cono $C$ es *normal* cuando
$$\inf\{\|x + y\| : x, y \in C, \|x\| = \|y\| = 1\} > 0.$$
Fijemos un cono $C$ de $E$.
3. Dado un operador linear continuo $L : E \to E$, diremos que $L$ es un *operador positivo sobre* $C$ si $L(C) \subset C$.
4. Dado un funcional lineal continuo $\varphi : E \to \mathbb{R}$, diremos que $\varphi$ es un *funcional positivo sobre* $C$ si $\varphi(v) \geq 0$ paratodo $v \in C$.
5. El *cono dual* $C^*$ es el cono en $E^*$ formado por todos los funcionales positivos sobre $C$.

EJEMPLO 4.1. $C_+(X) = \{\varphi \in C(X) : \varphi \geq 0\}$ es un cono normal de $C(X)$. Por el teorema de representación de Riesz, el cono dual se identifica naturalmente con el espacio de medidas (positivas) finitas en $X$.

TEOREMA 4.8. *Sea $C$ un cono normal en un espacio de Banach $E$ y sea $L : E \to E$ un operador linear positivo sobre $C$. Entonces el radio espectral $r(T^*)$ es eigenvalor del operador dual $L^* : E^* \to E^*$ con algun eigenvector $v^* \in C^*$.*

LEMA 4.9. *Considere el radio espectral $\lambda = r(\mathcal{L}^*) = r(\mathcal{L})$. Entonces existe una probabilidad $\nu$ en $X$ tal que $\mathcal{L}^*\nu = \lambda\nu$.*

DEMOSTRACIÓN. Aplicando el Teorema 4.8 tenemos que $\mathcal{L}^*$ posee algun eigenvector $\nu \in C_+(X)^*$ correspondiente al eigenvalor $\lambda$. Como acabamos de decir, $\nu$ se identifica com una medida positiva finita. Podemos normalizar $\nu$, para que sea una probabilidad. $\square$

EJEMPLO 4.2. Sea $f : M \to M$ un difeomorfismo local en la variedad Riemanniana compacta M. Considere el operador de transferencia $\mathcal{L}$ asociado al potencial $\varphi = -\log|\det Df|$. La medida de Lebesgue $m$ en $M$ es una eigenmedida del operador dual, correspondente al eigenvalor $\lambda = 1$: $\mathcal{L}^*m = m$. Para verificar ese hecho, basta mostrar que $\mathcal{L}^*m(E) = m(E)$ para todo boreliano $E$ contenido en la imagen de una bola $B_\rho(y)$ por alguna rama inversa $h_j : B_\rho(y) \to M$ (pues, por la compacidade de $M$, todo boreliano es la union finita disjunta de subconjuntos $E$ de este tipo). De la definición de $\mathcal{L} = \mathcal{L}_\varphi$

$$\mathcal{L}^*m(E) = \int \chi_E \, d(\mathcal{L}^*m) = \int \mathcal{L}\chi_E \; dm = \int \sum_{i=1}^{k} \frac{\chi_E}{|\det Df|} \circ h_i \, dm.$$

Entonces, por la elección de $E$ y la fórnula de cambio de variables,

$$\mathcal{L}^*m(E) = \int \frac{\chi_E}{|\det Df|} \circ h_j \, dm = \int \chi_E dm = m(E).$$

Esto prueba que $m$ es un punto fijo de $\mathcal{L}^*$.

Una conclusión análoga vale para las medidas de Markov.

A partir de ahora supondremos que $\nu$ es una medida de referencia, o sea, una probabilidad satisfaciendo $\mathcal{L}^*\nu = \lambda\nu$ para algún $\lambda > 0$. Una de las conclusiones al final de la demostración del teorema de Ruelle será que $\lambda$ está univocamente determinado (en vista del Lema 4.9, eso quiere decir que $\lambda$ es necesariamente igual al radio espectral de $\mathcal{L}$ y $\mathcal{L}^*$) y que la propia medida $\nu$ también es única.

LEMA 4.10. *La transformación $T : X \to X$ admite jacobiano relativo a $\nu$, dado por $J_\nu T = \lambda e^{-\varphi}$.*

DEMOSTRACIÓN. Sea $A$ cualquier dominio de inyectividad de $T$. Sea $(g_n)_n$ una sucesión de funcionoes continuas que converge en $\nu$-casi todo punto a la $\chi_A$ y tal que $\sup |g_n| \leq 1$ para todo $n$. Observe que

$$\mathcal{L}(e^{-\varphi} g_n)(y) = \sum_{x \in T^{-1}(y)} g_n(x).$$

La expresión del lado derecho esta acotada por el grado($T$) y converge a $\chi_{T(A)}(y)$ en $\nu$-casi todo punto. Por el teorema de la convergencia dominada, el lado derecho de

$$\int \lambda e^{-\varphi} g_n \, d\nu = \int e^{-\varphi} g_n \, d(\mathcal{L}^* \nu) = \int \mathcal{L}(e^{-\varphi} g_n) \, d\nu$$

converge a $\nu(T(A))$, pero el lado izquierdo converge a $\int_A \lambda e^{-\varphi} d\nu$. Concluimos que

$$\nu(T(A)) = \int_A \lambda e^{-\varphi} \, d\nu$$

$\square$

Lema 4.11. *Sea $T : X \to X$ transformación expansora topologicamente exacta y sea $\eta$ cualquier probabilidad en los borelianos tal que existe jacobiano de $T$ relativo a $\eta$. Entonces el soporte de $\eta$ es todo $X$.*

Demostración. Suponga que por el contrario existe algun abierto $U \subset X$ tal que $\eta(U) = 0$. Ya que $T$ es un homeomorfismo local, es una aplicación abierta y entonces $T(U)$ es abierto. Podemos cubrir $U$ con una unión finita de dominios de inyectividad $A$. Para cada uno de ellos,

$$\eta(T(A)) = \int_A J_\eta T d\eta = 0.$$

Por lo tanto, $\eta(T(U)) = 0$. Por inducion, se sgue que $\eta(T^n(U)) = 0$ para todo $n \geq 0$. Como $T$ es topologicamente exacta, existe $k \geq 1$ tal que $T^k(U) = X$. Como $\eta(X) = 1$, esto genera uma contradición. $\square$

Fijemos constantes $K_0 > 0$, $\alpha > 0$ tales que $|\varphi(z) - \varphi(w)| \leq K_0 d(z,w)^\alpha$ para cuaquier $z, w \in X$

Lema 4.12. *Existe $K_1 > 0$ tal que para todo $n \geq 1$, todo $x \in X$ y todo $y \in B_{n+1}(x, \rho, T)$,*

$$|S_n\varphi(x) - S_n\varphi(y)| \leq K_1 d(T^n(x), T^n(y))^\alpha$$

Demostración. Por hipótesis, $d(T^i(x), T^i(y)) < \rho$ para todo $0 \leq i \leq n$. Entonces, para cada $j = 1, \ldots, n$, la rama contractiva $h_j : B_\rho(T^n(x)) \to X$ de $T^j$ que envia $T^n(x)$ en $T^{n-j}(x)$ también envia $T^n(y)$ en $T^{n-j}(y)$. Por el Lema 4.8 tenemos que $d(T^{n-j}(x), T^{n-j}(y)) \leq \sigma^{-j} d(T^n(x), T^n(y))$ para todo $j = 1, \ldots, n$.

$$|S_n\varphi(x) - S_n\varphi(y)| \leq \sum_{j=1}^{n} |\varphi(T^{n-j}(x)) - \varphi(T^{n-j}(y))| \ \leq K_0 \sum_{j=1}^{n} \sigma^{-j\alpha} d(T^n(x), T^n(y))^\alpha.$$

Por lo tanto, basta tomar $K_1 \geq K_0 \sum_{j=0}^{\infty} \sigma^{-j\alpha}$. □

COROLARIO 4.9. *Existe $K_2 > 0$ tal que para todo $n \geq 1$, todo $x \in X$ y todo $y \in D_{n+1}(x, \rho, T)$,*

$$K_2^{-1} \leq \frac{J_\nu T^n(x)}{J_\nu T^n(y)} \leq K_2.$$

DEMOSTRACIÓN. Del Lema 4.10 se sigue que

$$J_\nu T^n(z) = \lambda^n e^{-S_n\varphi}(z) \quad \forall z \in X, n \geq 1.$$

Entonces el Lema 4.12 da que

$$\big| \log \frac{J_\nu T^n(x)}{J_\nu T^n(y)} \big| = |S_n\varphi(x) - S_n\varphi(y)| \leq K_1 d(T^n(x), T^n(y))^\alpha \leq K_1\rho^\alpha.$$

Así, basta escoger $K_2 = \exp(K_1\rho^\alpha)$. □

Ahora probaremos que $\nu$ es un estado de Gibbs:

LEMA 4.13. *Escribamos $P = \log \lambda$. Para todo $\varepsilon > 0$ suficientemente pequeño existe $K_3 = K_3(\varepsilon) > 0$ tal que*

$$K_3^{-1} \leq \frac{\nu(B_n(x, \varepsilon, T))}{\exp(S_n\varphi(x) - nP)} \leq K_3 \quad \forall x \in X, n \geq 1.$$

DEMOSTRACIÓN. Considere $\varepsilon < \rho$. Entonces $T|B_\varepsilon(y)$ es inyectiva para todo $y \in X$ y así , $T^n|B_n(x, \varepsilon, T)$ es inyectiva para todo $x \in X$ y todo $n$. Entonces,

$$\nu(T^n(B_n(x, \varepsilon, T))) = \int_{B_n(x,\varepsilon,T)} J_\nu T^n(y) d\nu(y).$$

Por el Corolario 4.9, el valor de $J_\nu T^n$ en un punto cualquiera $y \in B_n(x, \varepsilon, T)$ difiere de $J_\nu T^n(x)$ por un factor acotado por la constante $K_2$. Se sigue que

$$K_2^{-1}\nu(T^n(B_n(x, \varepsilon, T))) \leq J_\nu T^n(x)\nu(B_n(x, \varepsilon, T)) \leq K_2\nu(T^n(B_n(x, \varepsilon, T))). \tag{4.5}$$

Como vimos en 4.10 $J_\nu T^n(x) = \lambda^n e^{-S_n\varphi(x)} = \exp(nP - S_n\varphi(x))$. También tenemos que $T^n(B_n(x, \varepsilon, T)) = T(B_\varepsilon(T^{n-1}(x)))$ y por lo tanto,

$$\nu(T^n(B_n(x, \varepsilon, T))) = \int_{B_\varepsilon(T^{n-1}(x))} J_\nu T \, d\nu \quad \forall x \in X, n. \tag{4.6}$$

El lado izquierdo de (4.6) está mayorado por 1. Además, $J_\nu T = \lambda e^{-\varphi}$ esta separado de cero, lo mismo que $\{\nu(B_\varepsilon(y)) : y \in X\}$. Por lo tanto el lado derecho de (4.6) esta minorado por algun número $a > 0$. Usando estas observaciones en (4.5), obtenemos

$$K_2^{-1}a \leq \nu(B_n(x, \varepsilon, T)) \leq K_2.$$

Basta tomar $K_3 = K_2 \max(1, a^{-1})$. □

En seguida vamos a demostrar que $\lambda$ es eigenvalor de $\mathcal{L}$ con eigenfunción estrictamente positiva Hölder, la cual será construida como punto de acumulación de Cesaro de la sucesión de funciones $\lambda^{-n}\mathcal{L}^n 1$.

Lema 4.14. *Sea $K_1 > 0$ como en el Lema (4.12), entonces para todo $n \geq 1$, $y_1, y_2 \in X$ con $d(y_1, y_2) < \rho$*

$$-K_1 d(y_1, y_2)^\alpha \leq \log \frac{\mathcal{L}^n 1(y_1)}{\mathcal{L}^n 1(y_2)} \leq K_1 d(y_1, y_2)^\alpha$$

Demostración. Se sigue de (4.4) que

$$\mathcal{L}^n g | B_\rho(y) = \sum_i (e^{S_n\varphi} g) \circ h_i^n$$

donde la suma es sobre las ramas inversas $h_i^n : B_\rho(y) \to X$ del iterado $T^n$. En particular,

$$\frac{\mathcal{L}^n 1(y_1)}{\mathcal{L}^n 1(y_2)} = \frac{\sum_i e^{S_n\varphi(h_i^n(y_1))}}{\sum_i e^{S_n\varphi(h_i^n(y_2))}}$$

Por el Lema (4.12), para cada una de esas ramas inversas

$$|S_n\varphi(h_i^n(y_1)) - S_n\varphi(h_i^n(y_2))| \leq K_1 d(y_1 y_2)^\alpha$$

y así

$$e^{-K_1 d(y_1,y_2)^\alpha} \sum_i e^{S_n\varphi(h_i^n(y_2))} \leq \sum_i e^{S_n\varphi(h_i^n(y_1))} \leq e^{K_1 d(y_1,y_2)^\alpha} \sum_i e^{S_n\varphi(h_i^n(y_2))}.$$

□

Corolario 4.10. *Existe $K_5 > 0$ tal que $K_5^{-1} \leq \lambda^{-n}\mathcal{L}^n 1(x) \leq K_5$ para todo $n \geq 1$ y $x \in X$ .*

Demostración. Comecemos observando que, para todo $n \geq 1$,

$$\int \mathcal{L}^n 1 \, d\nu = \int 1 d(\mathcal{L}^{*n}\nu) = \int \lambda^n d\nu = \lambda^n.$$

En particular, para todo $n \geq 1$,

$$\min_{y\in X} \lambda^{-n}\mathcal{L}^n 1(y) \leq 1 \leq \max_{y\in X} \lambda^{-n}\mathcal{L}^n 1(y). \tag{4.7}$$

Como $T$ es topologicamente exacta, existe $N \geq 1$ tal que $T^N B_\rho(x) = X$ para todo $x \in X$. Dados $x, y \in X$, podemos encontrar $x' \in B_\rho(x)$ tal que $T^N(x') = y$. Tomando $c = \sup |\varphi|$, por un lado tenemos

$$\mathcal{L}^{n+N} 1(y) = \sum_{z\in T^{-N}(y)} e^{S_N\varphi(z)} \mathcal{L}^n 1(z) \geq e^{S_N\varphi(x')} \mathcal{L}^n 1(x') \geq e^{cN} \mathcal{L}^n 1(x').$$

Por otro lado, el Lema 4.14 da que $\mathcal{L}^n 1(x') \geq \mathcal{L}^n 1(x)\exp(-K_1\rho\alpha)$. Tome $K_4 \geq \exp(K_1\rho\alpha + cN)\lambda^N$, combinando las desigualdades anteriores vemos que

$$\mathcal{L}^{n+N} 1(y) \geq \exp(-K_1\rho\alpha - cN)\mathcal{L}^n 1(x) \geq K_4^{-1}\lambda^N \mathcal{L}^n 1(x)$$

para todo $x, y \in X$. Por lo tanto, para todo $n \geq 1$,

$$\min \lambda^{-(n+N)}\mathcal{L}^{n+N} 1 \geq K_4^{-1} \max \lambda^{-n}\mathcal{L}^n 1. \tag{4.8}$$

Combinando (4.7) y (4.8) obtenemos

$$\max \lambda^{-n}\mathcal{L}^n 1 \leq K_4 \min \lambda^{-(n+N)}\mathcal{L}^{n+N} 1 \leq K_4 \quad \forall n \geq 1$$
$$\min \lambda^{-n}\mathcal{L}^n 1 \geq K_4^{-1} \max \lambda^{-n+N}\mathcal{L}^{n-N} 1 \geq K_4^{-1} \quad \forall n > N.$$

Falta extender la última estimación para $1 \leq n \leq N$. Para eso, observe que cada $\mathcal{L}^n 1$ es una función continua y positiva. Luego, por la compacidad de $X$, el mínimo de $\mathcal{L}^n 1$ es positivo para todo $n$. Entonces, podemos tomar $K_5 \geq K_4$ tal que $\min \lambda^{-n}\mathcal{L}^n 1 \geq K_5^{-1}$ para todo $n = 1, \dots, N$. □

Se sigue del Corolario que el eigenvalor $\lambda$ esta unicamente determinado.

LEMA 4.15. *Existe $K_6 > 0$ tal que para todo $n \geq 1$ y cualesquiera $x, y \in X$*

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq K_6 d(x,y)^\alpha$$

*En particular, la sucesión $\lambda^{-n}\mathcal{L}^n 1$ es equicontinua.*

DEMOSTRACIÓN. Primero supongamos que $d(x, y) < \rho$. Por el Lema 4.14

$$\lambda^{-n}\mathcal{L}^n 1(x) \leq \lambda^{-n}\mathcal{L}^n 1(y)\exp(K_1 d(x,y)^\alpha)$$

y por lo tanto

$$\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y) \leq (\exp(K_1 d(x,y))^\alpha - 1)\lambda^{-n}\mathcal{L}^n 1(y).$$

Tome $K > 0$ tal que $|\exp(K_1 t) - 1| \leq K|t|$ siempre que $|t| \leq \rho^\alpha$. Entonces, usando el Corolario

$$\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y) \leq KK_5 d(x,y)^\alpha.$$

Invirtiendo los papeles de $x$ e $y$ concluimos que si $d(x, y) < \rho$ entonces

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq KK5d(x,y)^\alpha.$$

Cuando $d(x, y) \geq \rho$ el Corolario 4.10 da que

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq 2K_5 \leq 2K_5\rho^{-\alpha}d(x,y)^\alpha.$$

Luego, basta tomar $K_6 \geq \max(KK_5, 2K_5\rho^{-\alpha})$. □

El Corolario 4.10 y el Lema 4.15 implican que la sucesión

$$h_n = \frac{1}{n}\sum_{i=0}^{n-1} \lambda^{-i}\mathcal{L}^i 1$$

es acotada y equicontinua y asi por el Teorema de Arzela-Ascoli tiene una subsucesión $(h_{n_i})$ que converge uniformemente a una función continua $h$.

LEMA 4.16. *La función $h$ satisface $\mathcal{L}h = \lambda h$, $\int h\, d\nu = 1$ y además*

$$K_5^{-1} \leq h(x) \leq K_5,\ |h(x) - h(y)| \leq K_6 d(x,y)^\alpha \quad \forall x, y \in X.$$

DEMOSTRACIÓN. Como $\mathcal{L}$ es continuo

$$\begin{aligned}\mathcal{L}h &= \lim_i \mathcal{L}h_{n_i} = \lim_i \frac{1}{n_i}\sum_{k=0}^{n_i-1} \lambda^{-k}\mathcal{L}^{k+1}1 = \lim_i \frac{\lambda}{n_i}\sum_{k=1}^{n_i}\lambda^{-k}\mathcal{L}^k 1 \\ &= \lim_i \frac{\lambda}{n_i}\sum_{k=0}^{n_i-1}\lambda^{-k}\mathcal{L}^k 1 + \frac{\lambda}{n_i}(\lambda^{-n_i}\mathcal{L}^{n_i}1 - 1) = \lambda h\end{aligned}$$

Por la definición de $\nu$, tenemos $\int \lambda^{-n}\mathcal{L}^n 1 d\nu = \int \lambda^{-n} d(\mathcal{L}^{*n}\nu) = \int \nu = 1$ para todo $n \in \mathbb{N}$, de donde $\int h_n d\nu = 1$ y por el teorema de la convergencia dominada $\int h d\nu = 1$. Las otras afirmaciones se siguen del Corolario 4.10 y el Lema 4.15. □

LEMA 4.17. *La probabilidad $\mu = h\nu$ es $T$-invariante y $T$ admite jacobiano respecto a $\mu$, dado por $J_\mu T = \lambda e^{-\varphi}(h \circ T)/h$.*

DEMOSTRACIÓN. Notemos que $\mathcal{L}(g_1 \circ T)g_2) = g_1\mathcal{L}g_2$, cualquiera que sean las funciones continuas $g_1, g_2 : X \to \mathbb{R}$ ya que para todo $y \in X$,

$$\begin{aligned}\mathcal{L}((g_1 \circ T)g_2)(y) &= \sum_{x\in T^{-1}(y)} e^{\varphi(x)} g_1(T(x))g_2(x) \\ &= g_1(y)\sum_{x\in T^{-1}(y)} e^{\varphi(x)} g_2(x) = g_1(y)\mathcal{L}g_2(y).\end{aligned}$$

Entonces, para toda función continua $g : X \to \mathbb{R}$,

$$\begin{aligned}\int (g\circ T)d\mu &= \int (g \circ T)hd\nu = \lambda^{-1}\int (g\circ T)hd(\mathcal{L}^*\nu) = \lambda^{-1}\int \mathcal{L}((g\circ T)h)d\nu \\ &= \lambda^{-1}\int g\mathcal{L}hd\nu = \int ghd\nu = \int gd\mu,\end{aligned}$$

lo que prueba que la probabilidad $\mu$ es $T$-invariante. Para probar la segunda afirmación, considere cualquier dominio de inyectividad $A$ de $T$.

$$\mu(T(A)) = \int_{T(A)} hd\nu = \int_A J_\nu T(h \circ T)d\nu = \int_A J_\nu T\frac{h\circ T}{h}d\mu,$$

lo que significa que

$$J_\mu T = J_\nu T\frac{h\circ T}{h} = \lambda e^{-\varphi}\frac{h \circ T}{h}.$$

□

COROLARIO 4.11. *La probabilidad $\mu = h\nu$ satisface $h_\mu(T) + \int \varphi d\mu = P$.*

DEMOSTRACIÓN. Por el Teorema 3.15 y el Lema 4.17

$$h_\mu(T) = \int \log J_\mu T d\mu = \log\lambda - \int \varphi d\mu + \int (\log h\circ T - \log h) d\mu = P - \int \varphi d\mu$$

□

Sea $\eta \in \mathcal{M}_T(X)$ satisfaciendo

$$h_\eta(T) + \int \varphi d\eta \geq P \tag{4.9}$$

Sean $g_\eta = 1/J_\eta T$ y $g = 1/J_\mu T$. Observe que para toda $y \in X$.

$$\sum_{x\in T^{-1}(y)} g(x) = \frac{1}{\lambda h(y)} \sum_{x\in T^{-1}(y)} e^{\varphi}(x)h(x) = \frac{\mathcal{L}h(y)}{\lambda h(y)} = 1. \tag{4.10}$$

Como $\eta$ es $T$-invariante, para $\eta$-casi todo $y \in X$.

$$\sum_{x\in T^{-1}(y)} g_\eta(x) = 1. \tag{4.11}$$

Usando (4.9) y el Teorema 3.15, la definición de $g$ y la invariancia de $\eta$

$$0 \leq h_\eta(T) + \int \varphi d\eta - P = \int (\log J_\eta T + \varphi - \log\lambda) d\eta \tag{4.12}$$

$$= \int (-\log g_\eta + \log g + \log h\circ T - \log h) d\eta = \int \log\frac{g}{g_\eta} d\eta. \tag{4.13}$$

Por la definición de $g_\eta$ y el ejercicio 3.1 tenemos

$$\int \log\frac{g}{g_\eta} d\eta = \int \Big( \sum_{x\in T^{-1}(y)} g_\eta(x) \log\frac{g}{g_\eta}(x)\Big) d\eta(y) \tag{4.14}$$

De la concavidad estricta del logaritmo se sigue:

LEMA 4.18. *Sean $p_i, b_i, i = 1, \dots, k$ nmeros reales positivos tales que $\sum_{i=1}^k p_i = 1$. Entonces $\sum_{i=1}^k p_i \log b_i \leq \log\Big(\sum_{i=1}^k p_i b_i\Big)$ y la igualdad se da si y sólo si los números $b_j$ son todos iguales a $\sum_{i=1}^k p_i b_i$.*

Para $y \in X$, sean $x_i$ las pre-imagenes de $y$. Tome $p_i = g_\eta(x_i)$ y $b_i = \log(g(x_i)/g_\eta(x_i))$. La igualdad (4.11) significa que $\sum_i p_i = 1$ para $\eta$-casi todo $y$. Aplicando el Lema 4.18

$$\sum_{x\in T^{-1}(y)} g_\eta(x) \log\frac{g}{g_\eta}(x) \leq \log \sum_{x\in T^{-1}(y)} g_\eta(x)\frac{g}{g_\eta}(x) = \log \sum_{x\in T^{-1}y)} g(x) = 0 \tag{4.15}$$

Combinando las relaciones (4.12)-(4.15) obtenemos

$$h_\eta(T) + \int \varphi d\eta - P = \int \log \frac{g}{g_\eta} d\eta = 0. \tag{4.16}$$

COROLARIO 4.12. $P_T(\varphi) = P = \log r(\mathcal{L})$

COROLARIO 4.13. *Si $\eta$ es un estado de equilibrio para $\varphi$ entonces el soporte de $\eta$ es $X$, $J_\eta T = \lambda e^{-\varphi}(h \circ T)/h$ y $\mathcal{L}^*(\eta/h) = \lambda(\eta/h)$.*

DEMOSTRACIÓN. La primera afirmación del enunciado es consecuencia de la segunda y el Lema 4.11 La igualdad (4.16) también implica que vale la igualdade en (4.15) para casi todo $y \in X$. De acuerdo al Lema 4.18, eso ocurre si y sólo si, los números $b_i = \log(g(x_i)/g_\eta(x_i))$ son todos iguales. En otras palabras, para $\eta$-casi todo $y \in X$ existe un número $c(y)$ tal que

$$\frac{g(x)}{g_\eta(x)} = c(y) \text{ para todo } x \in T^{-1}(y).$$

Además por las igualdades (4.10) y (4.11),

$$c(y) = \sum_{x \in T^{-1}(y)} c(y) g_\eta(x) = \sum_{x \in T^{-1}(y)} g(x) = 1$$

para $\eta$-casi todo $y$. Se sigue que $g_\eta = g$ en $\eta$ casi todo punto, o sea, la función $1/g = \lambda e^{-\varphi}(h \circ T)/h$ es un jacobiano para $T$ relativo a $\eta$. Esto prueba la segunda afirmación. Para probar la tercera afirmación, sea $\xi : X \to R$ una función continua cualquiera. Por un lado, usando a definición del operador de transferencia

$$\int \xi d\mathcal{L}^*\Big(\frac{\eta}{h}\Big) = \int \frac{1}{h} L\xi d\eta = \int \frac{1}{h(y)} \Big( \sum_{x \in T^{-1}(y)} e^{\varphi(x)} \xi(x) \Big) d\eta(y). \tag{4.17}$$

Por la definición de la función $g$, para $x \in T^{-1}(y)$

$$\frac{e^{\varphi(x)}}{h(y)} = \frac{\lambda g(x)}{h(x)}.$$

Substituyendo esta igualdad en (4.17) y usando el ejercicio 3.1, obtenemos

$$\int \xi d\mathcal{L}^*\Big(\frac{\eta}{h}\Big) = \int \Big( \sum_{x \in T^{-1}(y)} \frac{\lambda g \xi}{h}(x) \Big) d\eta(y) = \int \frac{\lambda g}{h} d\eta. \tag{4.18}$$

□

**2.1. Medidas absolutamente continuas.** Discutiremos el caso particular en que $f : M \to M$ es un difeomorfismo local de una variedade compacta y $\varphi = -\log|\det Df|$. Supondremos que el potencial $\varphi$ es Hölder. Primero vamos a comparar las conclusiones del teorema de Ruelle com los resultados de la Sección 1.

Proposición 4.2. *La probabilidad absolutamente continua invariante bajo $f$ coincide con el estado de equilibrio $\mu$ del potencial $\varphi = -\log|\det Df|$. Consecuentemente, es equivalente a la medida de Lebesgue $m$, con densidad $d\mu/dm$ Hölder y alejada de cero e infinito.*

Demostración. Vimos en el Ejemplo 4.2 que la medida de Lebesgue $m$ es eigenvector para el eigenvalor 1 del dual $\mathcal{L}^*$ del operador de transferencia correspondiente al potencial $\varphi = -\log|\det Df|$. Aplicando la teoría del Lema 4.10 en adelante com $\lambda = 1$ y $\nu = m$, encontramos una función Hölder $h : M \to R$, alejada de cero e infinito, tal que $\mathcal{L}h = h$ y la medida $\mu = hm$ es el estado de equilibrio del potencial $\varphi$. Del Corolario 4.6, se sigue que $\mu$ es también la única probabilidad $f$-invariante absolutamente continua respecto a $m$. Ya que $h$ es positiva tenemos que $\mu$ y $m$ son equivalentes. □

También se tiene que

$$h\mu(f) - \int \log|\det Df| d\mu = P(f,\varphi) = \log 1 = 0. \tag{4.19}$$

Sea $\tilde{\varphi}$ la media temporal de la función $\varphi$, dada por el teorema ergódico de Birkhoff.

$$\int \log|\det Df| d\mu = -\int \varphi d\mu = -\int \tilde{\varphi} d\mu. \tag{4.20}$$

Además

$$-\tilde{\varphi}(x) = \lim_n \frac{1}{n}\sum_{j=0}^{n-1} \log|\det Df(f^j(x))| = \lim_n \frac{1}{n}\log|\det Df^n(x)| \tag{4.21}$$

en $\mu$-casi todo punto.

# Bibliografía


[B-S] M. Brin, G. Stuck. *Introduction to Dynamical Systems*. Cambridge University Press 2004.

[K-H] A. Katok, B. Hasselblatt. *Modern introduction to the theory of Dynamical Systems*. Cambridge University Press 1995.

[Ma] R. Mañe. *Teoria Ergódica*. Projeto Euclides. IMPA, Rio de Janeiro 1983.

[O-V] K. Oliveira, M. Viana. *Fundamentos da Teoria Ergódica*

[Pe] K. Petersen. *Ergodic Theory*. Cambridge University Press 1983.

[Po] M. Pollicott. *Lectures on Ergodic Theory and Pesin Theory on Compact Manifolds.* London Math. Soc. LNS 180, Cambridge University Press 1994.

[Wa] P. Walters. *An introduction to Ergodic Theory*. GTM 79, Springer 1982.