Notas para un curso de Teoría Ergódica

Héctor Sánchez Morgado[1]

[1]Instituto de Matemáticas, UNAM Ciudad Universitaria C. P. 04510 Cd. de México, México. e-mail: hector@matem.unam.mx

Contenido

CAPíTULO 1

TRANSFORMACIONES PRESERVADORAS DE MEDIDA

1. Medidas invariantes

DEFINICIÓN 1. Sean $(X, \mathcal{A}, \mu)$ un espacio de medida y $T : X \hookleftarrow$. Decimos que T preserva μ ó que μ 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 μ_0 en $\mathbb{R}^n$ induce una medida μ en la σ-álgebra de Borel $\mathcal{B}$ del toro

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

La traslación $t_a(x) = x + a$ en $\mathbb{R}^n$ induce en K^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 μ_0 tenemos que R_α preserva μ.

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 σ-álgebra de Borel $\mathcal{A}$ de X, por lo tanto, para definir una medida μ 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 μ.

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. Definamos ahora el mapeo de Gauss $T : (0, 1) \hookleftarrow$ mediante

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

Consideremos la medida μ en la σ-álgebra de Borel definida por

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

Para ver que T preserva μ 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 σ-álgebra de Borel. Entonces existe μ 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). 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).

Sea N tal que $n \geq N \Rightarrow \left\|\int_X g_j d\mu_n - \int_X g_j d\mu\right\| \leq \varepsilon/3 \|g\|$. Si $n \geq N$,

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

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

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

$\square$

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*

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

Lema 1.2. *Toda sucesión en $\mathcal{M}(X)$ posee una subsucesión convergente.*

Demostración. Sea $\{\mu_n\}$ una sucesión en $\mathcal{M}(X)$ y sea $\hat{\mu}_n \in [-1, 1]^{\mathbb{N}}$ definida por $\hat{\mu}_n(j) = \int_X g_j d\mu_n$

Como $[-1, 1]^{\mathbb{N}}$ provisto de la topología producto es un métrico compacto existe una subsucesión $\{\mu_{n_m}\}$ tal que $\{\hat{\mu}_{n_m}(j)\}$ converge $\forall j \geq 1$ o sea que $\left\{\int_X g_j d\mu_{n_m}\right\}$ converge $\forall j \geq 1$ y así $\left\{\int_X g d\mu_{n_m}\right\}$ converge $\forall g \in C(X)$.

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

Por el Teorema de Riesz existe $\mu \in \mathcal{M}(X)$ tal que $\varphi(g) = \int_X g d\mu$ $\forall g \in C(X)$, o sea que $\{d(\mu_{n_m}, \mu)\}$ converge a 0. $\square$

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 Lema 1.2, existe una subsucesión convergente $\{\mu_{n_m}\}$. Sea $\nu = \lim\limits_{m\to\infty} \mu_{n_m}$, entonces

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

$\square$

2. Formas de volumen, difeomorfismos, campos vectoriales

Definición 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 1.4. Fijando una base ordenada $(e_1, \dots, e_n)$ de V, definimos $\omega(e_1, \dots, e_n) = 1$. La antisimetría y multilinealidad implican que si $v_1, \dots, v_n \in V$ y escribimos $v_i = \sum\limits_j a_{ij} e_j$, 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 ω 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 3. Sean M una variedad diferenciable y $\varphi : U \subset \mathbb{R}^n \to M$ una parametri- zación. 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 ω en $\varphi(U)$ asocia, por definición, a cada $q \in \varphi(U)$ una forma de volumen ω_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 ω es suave si $\varphi^*\omega$ lo es.

Si ω es 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 ω es una forma de volumen en M y φ, ψ 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 $\{A_j \subset \operatorname{Im} \varphi_j\}$ conjuntos disjuntos con $A = \bigcup_j A_j$. Definimos

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

Si $f : M \hookleftarrow$ es un difeomorfismo C^1 tenemos que $\mu_\omega \circ f = \mu_{f^*\omega}$. Por lo tanto f preseva μ_ω 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 ω una forma de volumen en M, definimos la derivada de Lie

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

EJEMPLO 1.5. Sean M abierto de $\mathbb{R}^n$ y $\varphi_t^*\omega = (\det D\varphi_t)\omega \circ \varphi_t$. Definiendo $\rho(x) = \omega_x(e_1, \dots, e_n)$, tenemos

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

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

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

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

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

PROPOSICIÓN 1.1. *Supongamos que $f : M \hookleftarrow$ preserva ω 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 $x \in H^{-1}(C)$, $\quad T_x H^{-1}(C) = \ker D_x H$. Sea v tal que $D_x H(v) = 1$, definimos $\tilde{\omega}_x(v_2, \dots, v_n) = \omega_x(v, v_2, \dots, v_n)$, luego $(f^*\tilde{\omega})_x(v_2, \dots, v_n) = \tilde{\omega}_{f(x)}(D_x f v_2, \dots, D_x f v_n)$.

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

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

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

EJEMPLO 1.6. Sean M abierto de $\mathbb{R}^{2n}$ y $H : M \to \mathbb{R}$ de clase C^2. Consideremos el campo vectorial hamiltoniano

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

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

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 ω_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 $B_0 = B \setminus \bigcup\limits_n C_n$.

Como $C_n = B \setminus \bigcup\limits_{k \geq n} T^{-k}(B) \in \mathcal{A}$, el teorema estará demostrado si probamos que y $\mu(C_n) = 0$.

Como $C_n \subset \underset{k\geq 0}{\cup} T^{-k}B \setminus \underset{k\geq n}{\cup} T^{-k}B$, $\mu(C_n) \leq \mu(\underset{k\geq 0}{\cup} T^{-k}B) - \mu(\underset{k\geq n}{\cup} T^{-k}B)$. Pero $\underset{k\geq n}{\cup} T^{-k}B = T^{-n} \underset{k\geq 0}{\cup} T^{-k}B$ y así $\mu(\underset{k\geq 0}{\cup} T^{-k}B) = \mu(\underset{k\geq n}{\cup} T^{-k}B)$.

$\square$

DEFINICIÓN 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 1.1. *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, μ-casi todo punto de X es recurrente.

Demostración.

Sea $\{U_n\}\, n \in \mathbb{N}$ una base de abiertos con $\lim\limits_{n\to\infty}$ diam $(U_n) = 0$. Sea

$$\tilde{U}_n = \{x \in U_n : \exists m_k \to \infty \,\text{con}\, T^{m_k}(x) \in U_n\}.$$

Por el teorema de Poincaré $\mu(U_n \setminus \tilde{U}_n) = 0$. Sea $\tilde{X} = \bigcap\limits_{m\geq 1} \bigcup\limits_{n\geq m} \tilde{U}_n$.

Si $x \in \tilde{X}$ entonces $x \in \omega(x)$. En efecto, sea $r > 0$ y escojamos m tal que diam $(U_n) < r$ si $n \geq m$. Como $x \in \tilde{X}$, $x \in \bigcup\limits_{n\geq m} \tilde{U}_n$ y así $\exists n \geq m$ tal que $x \in \tilde{U}_n$. Como diam$(U_n) < r$, $U_n \subset B_r(x)$, por lo que $\exists m_k \to \infty$ tal que $T^{m_k}x \in U_n \subset B_r(x)$. Así $x \in \omega(x)$.

Como $\{U_n\}_{n\in\mathbb{N}}$ es una base de abiertos $X = \bigcup\limits_{n\geq m} U_n \; \forall m$.

Luego

$$\begin{aligned}\mu(X \setminus \tilde{X}) &= \mu\left(\bigcup_{m\geq 1} (X \setminus \underset{n\geq m}{\cup} \tilde{U}_n)\right) = \mu\left(\bigcup_{m\geq 1} (\underset{n\geq m}{\cup} U_n \setminus \underset{n\geq m}{\cup} \tilde{U}_n)\right) \\ &\leq \mu\left(\bigcup_{m\geq 1} \bigcup_{n\geq m} (U_n \setminus \tilde{U}_n)\right) = 0.\end{aligned}$$

$\square$

DEFINICIÓN 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 vacíos, $\exists N > 0$ talque $T^{-n}(U) \cap V \neq \phi \; \forall n \geq N$.

Proposición 1.2. *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 togopología de X. Por hipótesis

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

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. $\square$

Definición 6. Sea $X = [k]^{\mathbb{Z}}$. Sea $A = (A_{ij})_{i,j} \in [k]$ matriz de ceros y unos. Sea $X_A = \left\{(x_n) \in X : A_{x_n x_{n+1}} = 1 \,\forall n \in \mathbb{Z}\right\}$. Sea σ el corrimiento, entonces $\sigma(X_A) = X_A$. $\sigma|X_A$ se llama *subcorrimiento de tipo finito.*

Proposición 1.3. *Denotemos por A^m_{ij} los coeficientes de A^m $m \geq 0$. 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 1.3. *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 1.2. $\#\{x \in X_A : \sigma^m x = x\} = \mathit{Tr}\, A^m$.

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 $A_{x_0 x_1} = A_{x_1 x_2} = \ldots = A_{x_{m-1} x_0} = 1$. Así

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

□

Demostración de la Proposición 1.3. Sean U, V abiertos no vacíos 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

$$\begin{aligned}\sigma^{-n}(U) \cap V &\supset \sigma^{-n}(C(1; i_0, \ldots, i_m)) \cap C(p; q_0, \ldots, q_l) \cap X_A \\ &= C(j+n; i_0, \ldots, i_m) \cap C(p; q_0, \ldots, q_l) \cap X_A.\end{aligned}$$

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

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

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

4. Medidas de Markov

DEFINICIÓN 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 1.3. *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

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

μ es una medida en $\mathcal{A}_0$ que se extiende a la σ- 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" λ 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 1.4. *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* $\sum_{i \in [k]} x_i = 1$ *entonces*

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

Demostración. Sea λ el eigenvalor dominante. Sea

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

Por (b) y (c) del teorema se tiene que existe un único $\boldsymbol{\pi} \in H_1$ tal que $\boldsymbol{\pi}\mathbb{P} = \lambda\mathbb{P}$. Entonces $\sum_i \pi_i P_{ij} = \lambda \pi_j$ y sumando $\sum_{i,j} \pi_i P_{ij} = \lambda$. Pero

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

Luego $\lambda = 1$. Si $\mathbf{x} \in H_1$ entonces $\mathbf{y} = \mathbf{x} - \boldsymbol{\pi} \in H_0$. Así

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

$H_0\mathbb{P}^n \subset H_0$ y como $\ker(\mathbb{P} - \mathbb{I}) \cap H_0 = \{0\}$ tenemos que todos los eigenvalores de $\mathbb{P} \mid H_0$ tienen modulo < 1 y

así $\lim_{n \to \infty} \mathbf{y}\mathbb{P}^n = 0$. □

PROPOSICIÓN 1.4. *Sean $\mathbb{P}$ una matriz estocástica $k \times k$ y $\boldsymbol{\pi}$ un vector de probabilidad asociado. El soporte de la medida de Markov μ 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 $C(i; x_i, \dots, x_{i+m})$ Entonces

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

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 $\mu(C(i; x_i, x_{i+1})) = \pi_i P_{x_i x_{i+1}} = 0$. □

Ejercicio 1.1. Una matriz n×n A con coeficientes enteros, induce una transformación continua $\tilde{A} : \mathbb{K}^n \hookleftarrow$ mediante $\tilde{A}(x + \mathbb{Z}) = Ax + \mathbb{Z}$. Demostrar que si $\det A \neq 0$, entonces $\tilde{A}$ preserva la medida μ del Ejemplo 1.1.

Ejercicio 1.2. De acuerdo al Ejercicio 1.1 la transformación $\tilde{A} : K^1 \hookleftarrow$ dada por $\tilde{A}(z) = z^m$, preserva la medida inducida por la medida de Lebesgue en $\mathbb{R}$. Deducir que la transformación $T : [-1, 1] \hookleftarrow$ dada por $T(x) = 2x^2 - 1$ preserva una medida equivalente a la medida de Lebesgue.

Sugerencia. Defina $h : K^1 \to [-1, 1]$ mediante $h(z) = \Re z$ y pruebe que $h(z^2) = T(h(z))$. Definase $M(A) = \mu(h^{-1}(A))$ para $A \subset [-1, 1]$ boreliano.

Ejercicio 1.3. Sea $\varphi : [0, 1] \hookleftarrow$ tal que existen intervalos abiertos disjuntos $I_n \subset [0, 1]$, $n \in \mathbb{N}$ satisfaciendo

1) $\lambda([0, 1] \setminus \underset{n}{\cup} I_n) = 0$

2) $\lambda(\varphi([0, 1] \setminus \underset{n}{\cup} I_n)) = 0$

3)$\forall\ n \varphi \mid I_n$ es continuamente diferenciable con $\varphi' \neq 0$

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

Probar que φ preserva μ si y solo si

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

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

Ejercicio 1.4. 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 $p \in N \quad v_1, \ldots, v_{n-1} \in T_pN$. Probar que $\tilde{\omega}$ es una forma de volumen en N y que si div $X = 0$ entonces f preseva $\tilde{\omega}$.

EJERCICIO 1.5. 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 1.6. 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 ω- límite de un difeomorfismo $f : M \hookleftarrow$ se define como el conjunto $L^+(f) = \overline{\bigcup_{x \in M} \omega(x)}$. Probar que si un difeomorfismo $f : M \hookleftarrow$ posee una primera integral $L^+(f)$ es no numerable.

EJERCICIO 1.7. a) Probar que para toda forma de volumen $C^1\omega$ en K^1x y todo difeomorfismo $g : K^1 \hookleftarrow$ existe un difeomorfismo $f : K^1 \times \mathbb{R} \hookleftarrow$ que preserva ω y tal que $f(p, 0) = (g(p), 0) \forall p \in K^1$.

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 teorena 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*

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

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 $T^{-1}(A) = A$, $E(X_A f) = A$. Por lo tanto $\int_A f = \int_{E(X_A f)} X_A f \geq 0$. $\square$

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 $\int_{E_\alpha^+(f)} f d\mu \geq \alpha\mu(E_\alpha^+(f))$. En efecto

$$\begin{aligned} \int_{E_\alpha^+(f)} f d\mu &= \int_{E_\alpha^+(f)} (f-\alpha) d\mu + \alpha\mu(E_\alpha^+(f)) \\ &= \int_{E_0^+(f-\alpha)} (f-\alpha) d\mu + \alpha\mu(E_\alpha^+(f)) \end{aligned} \tag{2.1}$$

Pero $E_0^+(f-\alpha) \subset E(f-\alpha)$ y así

$$\int_{E_0^+(f-\alpha)} (f-\alpha) d\mu \geq 0.$$

Si $A \subset E_\alpha^+(f)$ y $T^{-1}A = A$, entonces

$$\int_A f d\mu = \int_A \chi_A f d\mu = \int_{E_\alpha^+(\chi_A f)} \chi_A f d\mu \geq \alpha\mu(E_\alpha^+(\chi_A f)) = \alpha\mu(A).$$

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

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

Como T preserva μ,

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

Por el lema de Fatou $\int_X g d\mu \leq \int_X |f| \, d\mu$ y así $\int_X \left|\tilde{f}\right| d\mu \leq \int_X |f| \, d\mu$. 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 $\left\|\tilde{f}\right\|_\infty \leq \|f\|_\infty$ y $\left\|\tilde{f} - \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j\right\|_\infty \leq 2\|f\|_\infty$. 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 $\varepsilon > 0$ existe $f_0 \in L^\infty(X)$ tal que $\int_X |f - f_0|\, d\mu \leq \varepsilon$ y tomamos N tal que $\int_X \left|\tilde{f}_0 - \frac{1}{n}\sum_{j=0}^{n-1} f_0 \circ T^j\right| d\mu \leq \varepsilon$ si $n \geq N$. Entonces

(2.2)
$$\int_X \left| f - \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \right| d\mu \leq \int_X \left|\tilde{f} - \tilde{f}_0\right| d\mu + \int_X \left| f_0 - \frac{1}{n}\sum_{j=0}^{n-1} f_0 \circ T^j \right| d\mu$$
$$+ \int_X \left| \frac{1}{n}\sum_{j=0}^{n-1} (f - f_0) \circ T^j \right| d\mu$$

Pero $\int_X \left|\tilde{f} - \tilde{f}_0\right| d\mu = \int_X |(f - f_0)^\sim|\, d\mu \leq \int_X |f - f_0|\, d\mu \leq \varepsilon$ 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 8. Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad y $T : X \to X$ transformación preservadora de medida. Decimos que T es μ- *ergódica* o que μ es una medida T-*ergódica*, si $T^{-1}B = B$ para $B \in \mathcal{A} \Rightarrow \mu(B) = 0$ ó 1. Denotamos

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

Lema 2.2. *Son equivalentes*

(a) T *es* μ-*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 μ-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 $X(k, n) = f^{-1}[\frac{k}{2^n}, \frac{k+1}{2^n}[$. Como $f \circ T = f$, $T^{-1}(X(k, n)) = X(k, n)$. Así $\mu(X(k, n)) = 0$ ó 1. Para n fija $X = \underset{k\in\mathbb{Z}}{\cup} X(k, n)$. Así, $\exists!\, k_n \in \mathbb{Z}$ tal que $\mu(X(k, n)) = 1$. Sean $Y = \underset{n\in\mathbb{Z}}{\cap} X(k_n, n)$ 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. $\square$

COROLARIO 2.2. *Si T es μ-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. $\square$

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. $\square$

Demostración de 2.1. Sea

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

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 $f_n \circ T \geq 0$.

2. Ejemplos de transformaciones ergódicas

LEMA 2.3. *Sean X espacio métrico compacto, $\mathcal{A}$ la σ-álgebra de Borel y μ 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 $\{T^m(x) : m \in \mathbb{N}\}$ es densa para casi todo $x \in X$.*

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

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

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

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

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

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

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

□

PROPOSICIÓN 2.1. *Consideremos la rotación $R_\alpha : K^n \hookleftarrow$ 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 $\varphi_k : K^n \to K^1$ 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 φ_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 $f \in L^2(K^n)$ 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. □

Proposición 2.2. *Sea $A \in GL(n,\mathbb{Z})$ y sea $\tilde{A} : K^n \hookleftarrow$ la transformación inducida (Ejercicio 1.1). $\tilde{A}$ es ergódica $\iff$ ninguno de los eigenvalores de A es raíz de la unidad*

Demostración. Sea $\varphi_k : K^n \to K^1$ como en la Proposición 2.1. Entonces y así

$$\begin{aligned}\varphi_k \circ \tilde{A}(x + \mathbb{Z}^n) &= \varphi_k(Ax + \mathbb{Z}^n) = \exp(2\pi i\langle k, Ax\rangle) \\ &= \exp(2\pi i\langle A^t k, x\rangle) = \varphi_{A^t k}(x + \mathbb{Z}^n)\end{aligned} \tag{2.3}$$

($\Rightarrow$). 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 $f \in L^2(K^n)$, $f = \sum\limits_{k\in\mathbb{Z}} b_k\varphi_k$ 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. □

Proposición 2.3. *El corrimiento $T : X \hookleftarrow$ con vector de probabilidad $(p_0, \dots, p_{k-1})$ 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 $\mu(T^{-m}(A_1) \cap A_2)$.

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

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

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 $\varepsilon > 0$ 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*

$$\mathbb{K} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{P}^i \tag{2.4}$$

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 μ la medida de Markov asociada a $(\mathbb{P}, \boldsymbol{\pi})$. Sea χ_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 χ_l y usando el teorema de la convergencia dominada

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

Pero

$$\mu(C(0;l) \cap C(i;j)) = \sum_{a_1,\dots,a_{i-1}} \pi_l P_{la_1}\cdots P_{a_{i-1}j} = \pi_l P^i_{lj}.$$

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

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

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

(1) *El corrimiento σ 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 σ 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 tendriamos $q_j = 0$.

(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 $\mathbb{K} = \mathbb{K}\mathbb{P}^n$,

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

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 σ 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 $i > j + s - l$,

$$\mu(A \cap \sigma^{-i} B) = \pi_{i_0} P_{i_0 i_1} \cdots P_{i_{s-1} i_s} P_{i_s r_0}^{i+l-(j+s)} P_{r_0 r_1} \cdots P_{r_{m-1} r_m}.$$

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

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(A \cap \sigma^{-i} B)$$
$$= \pi_{i_0} P_{i_0 i_1} \cdots p_{i_{s-1} i_s} \pi_{r_0} P_{r_0 r_1} \cdots P_{r_{m-1} r_m} = \mu(A)\mu(B)$$

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

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

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

(4) $\Rightarrow$ (2). Ya que $\mathbb{P}\mathbb{K} = \mathbb{K}$ cada columna de $\mathbb{K}$ es un eigenvector. Así, 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. □

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. 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 $T : K^1 \hookleftarrow$, $T(z) = z^{10}$. De acuerdo a la Proposición 2.2, T es ergódica. Sea $B = \left\{ e^{2\pi i x} \mid x \in [0, \frac{1}{10}) \right\}$ $x_j = 0 \iff T^j(e^{2\pi i x}) \in B$ y $\mu(B) = \frac{1}{10}$.

EJEMPLO 2.2. Sea $T : [0, 1) \hookleftarrow$ la transformación de Gauss

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

Si $x \neq 0$,

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

Sea

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

Para casi todo x (respecto a la medida de Lebesgue) se tiene que

$$\lim_{n\to\infty} \sqrt[n]{a_1 \cdots a_n} = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k^2 + 2k}\right)^{\log k / \log 2} \tag{2.5a}$$

$$\lim_{n\to\infty} \frac{\log q_n}{n} = \frac{\pi^2}{12 \log 2} \tag{2.5b}$$

Para demostrar estas ecuaciones, utilizaremos el hecho de que T es μ- ergódica donde μ esta definida en los borelianos por

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

(2.5a). Observemos que $a_i(x) = k \iff T^{i-1}(x) \in \left[\frac{1}{k+1}, \frac{1}{k}\right[$. Sea $f :]0, 1[\to \mathbb{R}$ definida por $f(x) = \log k$ para $x \in \left[\frac{1}{k+1}, \frac{1}{k}\right[$.

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

converge a

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

(2.5b).

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

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

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

Se demuestra por inducción que para $n \geq 3$

$$\begin{aligned} p_n(x) &= a_n(x) p_{n-1}(x) + p_{n-2}(x) \\ q_n(x) &= a_n(x) q_{n-1}(x) + q_{n-2}(x), \end{aligned} \tag{2.6}$$

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

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

Utilizando (2.6) se demuestra facilmente por inducción que $p_k \geq 2^{(k-2)/2}$ y $q_k \geq 2^{(k-1)/2}$ para $k \geq 1$. Como

$$\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} - \frac{p_n}{q_n} = \frac{q_n p_{n-1} - p_n q_{n-1}}{q_{n+1} q_n} = \frac{q_{n-1}}{q_{n+1}} \left(\frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n}\right), \tag{2.7}$$

$q_{n+1} > q_{n-1}$ 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.

Utilizando (2.7) 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

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

de donde

$$\sum_{i=0}^{n-1}\left|\log\left(\frac{p_{n-i}(T^ix)}{q_{n-1}(T^ix)}\right)-\log\left(T^ix\right)\right|=\sum_{i=0}^{n-1}\left|\log\frac{(T^ix)q_{n-i}(T^ix)}{p_{n-i}(T^ix)}\right|$$
$$\leq\sum_{i=0}^{n-1}\left|\frac{(T^ix)q_{n-i}(T^ix)}{p_{n-i}(T^ix)}-1\right|\leq\sum_{i=0}^{n-1}2^{1+i-n}\leq 2.$$

Por lo tanto

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

□

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.*

Demostración. Sea μ 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 μ_A,μ_{A^c} mediante

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

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 μ es punto extremo.

Reciprocamente supongamos que μ 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 μ y μ_i se tiene que $f_i\circ T=f_i$. Como μ es ergódica f_i es constante y así $\mu_i=\mu$. □

TEOREMA 2.5. *Sea $m\in\mathcal{M}_T(X)$, existe una medida de probabilidad μ_m en $\mathcal{M}_T(X)$ tal que*

(a) $\mu_m(M_{erg}(T))=1$.
(b) $\int_X f dm=\int_{\mathcal{M}_{erg}(T)}(\int_X f d\nu)d\mu_m(\nu)$.

Demostración. Sea $\mathcal{J} = \{A \in X \mid T^{-1}A = A\}$ la σ -álgebra T -invariante. Para cada $f \in L^1(X, \mathcal{A}, m)$, $\exists E_m(f \mid \mathcal{J}) \in L^1(X, \mathcal{J}, m \mid \mathcal{J})$ 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) $E_m(\chi_A \mid \mathcal{J}) = \chi_A \, \forall A \in \mathcal{J}$

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 $\mathcal{M}_{\text{erg}}$. De hecho podemos encajar X en M_{erg} mediante $x \to m_x$ y definir μ_m mediante $\mu_m(\mathcal{N}) = m(\mathcal{N} \cap X)$. □

5. Transformaciones mezcladoras

DEFINICIÓN 10. Sean $(X, \mathcal{A}, \mu)$ espacio de probabilidad, $T : X \hookleftarrow$ preservadora de medidad. 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 2.4. *Sean X espacio topológico y μ probabilidad sobre los borelianos tal que $\mu(U) > 0$ $\forall U$ abierto no vacío. Supongamos que $T : X \hookleftarrow$ preserva μ 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 μ tenemos un operador $U_T : L^2(X, \mathcal{A}, \mu) \hookleftarrow$ definido por $U_T(f) = f \circ T$.

Proposición 2.5. *Sea $(X, \mathcal{A}, \mu)$ espacio de probabilidad, la transformación presrvadora de medida $T : X \hookleftarrow$ es mezcladora si y solo si $\forall f, g \in L^2(X, \mathcal{A}, \mu)$*

$$\lim_{n\to\infty} \langle U_T^n f, g\rangle = \langle f, 1\rangle \langle 1, g\rangle \tag{2.8}$$

Demostración. Si U_T satisface la igualdad (2.8),

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

Reciprocamente si T es mezcladora se tiene que (2.8) se satisface cuando f, g son funciones características. De la bilinealidad de la igualdad (2.8) 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

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

si n es suficientemente grande. ☐

Proposición 2.6. *Sea $\mathbb{P}$ matriz estocástica $k \times k$ con vector de probabilidad $\boldsymbol{\pi}$ y sea μ la medida de Markov asociada. Son equivalentes.*

(1) *El corrimiento σ 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).

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

(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 σ 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 1.4, 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 2.7. *Sean A matriz $n \times n$ con coeficientes enteros y $\tilde{A} : K^n \hookleftarrow$ 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 $\{k^t A^n\}\, m \in \mathbb{N}$ 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 $L^2(K^n)$, tenemos que $\forall f, g \in L^2(K^n)$,

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

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 σ -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_α la rotación del círculo por el ángulo α. Demostrar que $R_\alpha \times R_\alpha$ no es ergódica y por lo tanto R_α 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 $K^n = \mathbb{R}^n/\mathbb{Z}^n$. Sea $\pi : \mathbb{R}^n \to K^n$ 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

ENTROPIA

1. Entropía con respecto a una medida

DEFINICIONES 11. 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 $\mathcal{P} \leq \mathcal{Q}$ 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 12. 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 *entropía condicional* de $\mathcal{P}$ dado $\mathcal{Q}$ como

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

OBSERVACIONES 3.1.

(a) ϕ 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 σ -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 *entropía condicional* de $\mathcal{P}$ dada $\mathcal{J}$ como

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

Recuperamos nuestra definición tomando $\mathcal{J} = \mathcal{J}(\mathcal{Q})$ la σ -álgebra generada por $\mathcal{Q}$.

PROPOSICIÓN 3.1. *Sean $\mathcal{P}, \mathcal{Q}, M$ 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* $\mathcal{P} \leq \mathcal{Q}$,$H(\mathcal{M} \mid \mathcal{P}) \geq H(\mathcal{M} \mid \mathcal{Q})$ *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* μ *entonces* $H(T^{-1}\mathcal{P} \mid T^{-1}\mathcal{Q}) = H(\mathcal{P} \mid \mathcal{Q})$
(f) $H(\mathcal{P} \mid \mathcal{Q}) = 0 \Leftrightarrow \mathcal{P} \leq \mathcal{Q}$

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 $\mathcal{P} \leq \mathcal{Q}$ entonces $\mathcal{P} \vee \mathcal{Q} = \mathcal{Q}$. Así

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

Como ϕ es convexa, $\phi\left(\sum_i \alpha_i x_i\right) \leq \sum_i \alpha_i \phi(x_i)$ si $\sum_i \alpha_i = 1$, $x_i \in [0,1]$. Luego

$$H(\mathcal{M} \mid \mathcal{Q}) = -\sum_{M,Q} \mu(M \cap Q) \log \frac{\mu(M \cap Q)}{\mu(Q)}$$

$$= -\sum_{M,P} \mu(P) \sum_{Q \subset P} \frac{\mu(Q)}{\mu(P)} \phi\left(\frac{\mu(M \cap Q)}{\mu(Q)}\right) \geq -\sum_{M,P} \mu(P)\, \phi\left(\sum_{Q \subset P} \frac{\mu(M \cap Q)}{\mu(P)}\right)$$

$$= -\sum_{M,P} \mu(P)\, \phi\left(\frac{\mu(M \cap P)}{\mu(P)}\right) = H(\mathcal{M} \mid \mathcal{P})$$

(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). $H(\mathcal{P} \mid \mathcal{Q}) = -\sum\limits_{P,Q} \mu(Q)\phi\left(\frac{\mu(P\cap Q)}{\mu(Q)}\right) = 0$ si y solo si $\forall P, Q$ se tiene $\mu(P \cap Q) = 0$ ó $\mu(Q)$; es decir $\mathcal{P} \leq \mathcal{Q}$.

Definición 13. 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 $a_{n+m} \leq a_n + a_m$. 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 $a_N/N \leq C + \varepsilon$. Si $n > N$ escribimos $n = pN + q$ con $q = 1, \ldots, N$.

$$\frac{a_n}{n} = \frac{a_{pN+q}}{pN+q} \leq \frac{pa_{N+}a_q}{pN+q} \leq a_N + \frac{a_q}{n} \leq C + \varepsilon + \frac{1}{n}\sup\{a_1, \ldots, a_N\}.$$

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

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

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

Proposición 3.2.
(a) $h(T, \mathcal{P}) - h(T, \mathcal{Q}) \leq H(\mathcal{P} \mid \mathcal{Q})$.
(b) *Si $\mathcal{P} \leq \mathcal{Q}$ entonces $h(T, \mathcal{P}) \leq h(T, \mathcal{Q})$.*
(c) $h(T, T^{-1}\mathcal{P}) = h(T, \mathcal{P})$.
(d) $\forall n \geq 0$, $h(T, \mathcal{P} \vee T^{-1}\mathcal{P} \vee \cdots \vee T^{-n}\mathcal{P}) = h(T, \mathcal{P})$.
(e) $h(T, P) = \lim\limits_{n\to\infty} H(\mathcal{P} \mid T^{-1}\mathcal{P} \vee \cdots \vee T^{-n}P)$.
(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

$$(3.1)\quad H(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{P})-H(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{Q})\leq H(\bigvee_{i=0}^{n-1}T^{-i}\mathcal{P}\mid\bigvee_{i=0}^{n-1}T^{-i}\mathcal{Q})$$

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

(b). Si $\mathcal{P}\leq\mathcal{Q}$, $H(\mathcal{P}\mid\mathcal{Q})=0$ y entonces $h(T,\mathcal{P})-H(T,\mathcal{Q})\leq 0$.
(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 14. 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 3.2. $h(T) = \sup\{h(T,\mathcal{P}) : H(\mathcal{P}) < \infty\}$.

En efecto, sea $\mathcal{P} = \{P_1, P_2, \cdots\}$ con $H(\mathcal{P}) < \infty$.
Consideremos $P^{(n)} = \{P_1, \cdots, P_n, \underset{j>n}{\cup} P_j\}$. Entonces

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

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 15. Si $\mathcal{A}_n$ es una familia de subconjuntos de X $\forall n \in \mathbb{N}$, $\bigvee_{n=1}^{\infty} \mathcal{A}_n$ es la σ-álgebra generada por $\cup_{n=1}^{\infty}\mathcal{A}_n$.

Teorema 3.1. *Sean $\mathcal{P}_1 \leq \mathcal{P}_2 \leq \cdots$ particiones y sea $\mathcal{P}$ una partición con $H(\mathcal{P}) < \infty$. Entonces $\mathcal{P} \subset \bigvee_{i=1}^{\infty} \mathcal{P}_i$(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 $\mathcal{P}_1 \leq \mathcal{P}_2 \leq \cdots$ es una sucesión de particiones con entropía finita tales que $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$, entonces $h(T) = \sup\limits_n h(T, \mathcal{P}_n)$*

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

Definición 16. Sea $(X, \mathcal{A}, \mu)$ espacio de probabilidad. Supongamos que $T : X \hookleftarrow$ es invertible y preserva μ. Una partición $\mathcal{P}$ con $H(\mathcal{P}) < \infty$ se llama T -generador si $\bigvee_{n=1}^{\infty} T^n\mathcal{P} = \mathcal{A}$.

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

Demostración.

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

□

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

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

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

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

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

con $\mu(I_j^m) = 1/n^m$. Entonces $\bigvee_{j=0}^{\infty} T^{-j}\mathcal{P}$ es la σ -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 $T : K^1 \hookleftarrow$ la rotación $T(z) = \alpha z$.

(1) $\arg(\alpha)/2\pi$ es irracional. 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, $\bigvee_{j=0}^{\infty} T^{-j}\mathcal{P}$ es la σ -álgebra de Borel de K^1. Así $h(T) = h(T, \mathcal{P})$.

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

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

(2) Si $\arg(\alpha)/2\pi$ es racional, existe m tal que $T^m = I$ y así $h(T^m) = 0$. Se sigue de la siguiente proposición que $h(T) = 0$.

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

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

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

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

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

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

2. Entropía Topológica

Definición 17. Sean X espacio métrico compacto y $T : X \hookleftarrow$ continua, decimos que $S \subset X$ es un (n,ε) *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 3.3. Sea $\{U_1,\dots,U_n\}$ una cubierta de X por conjuntos de diametro $\leq \varepsilon$ y en cada conjunto no vacío de la forma $\bigcap_{j=0}^{n} T^{-j}(U_{i_j})$ escojamos un punto. Con esto formamos un conjunto S (n,ε) generador con a lo más m^n elementos.

Definición 18. Sea $r_T(n,\varepsilon) = \min\{\#S : S \text{ es } (n,\varepsilon)\,\text{generador}\}$. Por la observación 3.3, $r_T(n,\varepsilon) \leq m^n$, así que

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

Definimos la *entropía topológica* de T como

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

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

Definición 19. Decimos que $S \subset X$ es (n,ε) *separado respecto a* T si $\forall x,y \in S, x \neq y$ $\exists j \in \{0,\dots,n\}$ 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,ε) 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 así $x = y$. Por lo tanto $\#E \leq \#F$ y así

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

Sea S conjunto (n,ε) separado de cardinalidad máxima. Si $y \notin S$ entonces $S \cup \{y\}$ no es (n,ε) 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,ε) 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.

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

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,ε) 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

$$d(T^t(T^{m_i} x), T^t(x_1)) \leq \varepsilon \;\text{ para }\; 0 \leq t \leq n_i,$$

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

$$\overline{r_T}(2\varepsilon) \leq \underline{n}\mathrm{inf} \frac{1}{n} \log r_T(n, \varepsilon) \tag{3.2}$$

□

Definición 20. Un homeomorfismo T de un espacio métrico (X, d) se llama *expansivo* si $\exists \varepsilon_0 > 0$ llamada constante de expansividad tal que

$$d(T^n x, T^n y) \leq \varepsilon_0 \, \forall n \in \mathbb{Z} \Rightarrow x = y$$

Proposición 3.6. *Sea T un homeomorfismo expansivo de un espacio métrico compacto X con constante de expansividad ε_0. Si $0 < \varepsilon < \varepsilon_0$, se tiene*

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

Demostración. 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 3.4. *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,ε) generador de cardinalidad mínima. Se sigue del Lema 3.3 que $T^k(S)$ es un conjunto $(n-2k, \varepsilon/2)$ generador. Así $r_T(n,\varepsilon) \geq r_T(n-2k,\varepsilon/2)$. □

Utilizando el corolario y (3.2)

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

□

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

Demostación. Sea d la métrica en X definida por

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

Es claro que cualquier $0 < \varepsilon_0 < 1$ es una constante de expansividad para σ. 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 $\alpha|\{-N, \ldots, N+n\} = \theta|\{-N, \ldots, N+n\}$.

Así, $d(\sigma^j\alpha, \sigma^j\theta) \leq \varepsilon$ para todo $0 \leq j \leq n$. Entonces S_n es un (n,ε) generador para σ y así $r_\sigma(n,\varepsilon) \leq \#S_n = r(n+2N)$. Si S fuera un (n,ε) 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 $\alpha_1|\{-N, \ldots, N+n\} = \alpha_2|\{-N, \ldots, N+n\}$, 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 λ con $\dim\ker(A - \lambda I) = 1$

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

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

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

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

$$\limsup_{n\to\infty} \frac{1}{n} \log r_{T^m}(n,\varepsilon) \leq \limsup_{n\to\infty} \frac{m}{nm} \log r_T(nm,\varepsilon) \leq \limsup_{k\to\infty} \frac{m}{k} \log r_T(k,\varepsilon).$$

Por lo tanto $h_{\text{top}}(T^m) \leq m\, h_{\text{top}}(T)$. $\forall\varepsilon > 0$, $\exists\delta = \delta(\varepsilon)$ tal que $d(T^jx, T^jy) \leq \varepsilon$ si $d(x,y) \leq \delta$, $0 \leq j \leq m-1$. Sea F un (n,δ) 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

$$\liminf_{k\to\infty} \frac{m}{k} \log r_T(k,\varepsilon) \leq \liminf_{n\to\infty} \frac{m}{nm} \log r_T(nm,\varepsilon) \leq \liminf_{n\to\infty} \frac{1}{n} \log r_{T^m}(n,\delta).$$

Así $m h_{\text{top}}(T) \leq h_{\text{top}}(T^m)$. □

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*

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

Demostración. Sea F_i un (n,ε) generador para T_i. Entonces F_1 x F_2 es un (n,ε) generador para $T_1 \times T_2$. Así $r_{T_1\times T_2}(n,\varepsilon) \leq r_{T_1}(n,\varepsilon)\cdot r_{T_2}(n,\varepsilon)$ y luego $\overline{r}_{T_1\times T_2}(\varepsilon) \leq \overline{r}_{T_1}(\varepsilon) + \overline{r}_{T_2}(\varepsilon)$ y entonces

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

□

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 ε_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,ε_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

En estsa sección demostraremos el siguiente Teorema

PRINCIPIO VARIACIONAL. *Sean X espacio métrico compacto, $T : X \hookleftarrow$ continua y $\mathcal{A}$ la σ -álgebra de Borel. Entonces*

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

Para comenzar daremos otra definición de entropía topológica y demostraremos que coincide con la definición anterior.

DEFINICIONES 21. 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

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

OBSERVACIONES 3.4.

(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 el inciso (e) de la Definición 21 existe.

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

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

By Lemma 3.1, $\lim a_n/n$ exists. □

Definición 22. 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 diam($\mathcal{U}$)=supdiam($\mathcal{U}$) : $U \in \mathcal{U}$

Lema 3.4. *Si $\{\mathcal{U}_n\}$ es una sucesión de particiones tal que* $\lim_{n\to\infty}$ *$diam(\mathcal{U}_n) = 0$, 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 diam($\mathcal{U}_n$) es menor que el número de Lebesgue de $\mathcal{U}$. Entonces $\mathcal{U}_n$ es un refinamiento de U y así $H(T, \mathcal{U}) \leq H(T, \mathcal{U}_n)$. □

Proposición 3.11.

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

Demostración.

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

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

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 λ su número de Lebesgue. Sea $\varepsilon < \lambda$ y sea F un (n,ε)- 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 ε centrada en $T^i(y)$. El conjunto

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

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

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

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

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

□

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 $\epsilon \in (0, 1/k\log k)$ 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

$$\begin{aligned} H_\mu(\mathcal{P} \mid \mathcal{B}) &= \mu(B_0) \sum_{i=1}^{k} \mu(P_i \setminus B_0) \log \mu(P_i \setminus B_0) \\ &\leq \mu(B_0) \log k \leq \epsilon\, k \log k < 1 \end{aligned}$$

Como en (3.1), tenemos que

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \mid \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right) \leq nH_\mu(\mathcal{P} \mid \mathcal{Q})$$

y así

(3.3)
$$\begin{aligned}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) &\leq H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \vee \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right)\\ &\leq H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \mid \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right) + H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right)\\ &\leq n + H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right)\end{aligned}$$

Sea $U_i = B_0 \cap B_i$, o sea $X \setminus U_i = \bigcap_{j\neq 0,i} B_j$. Por lo tanto, cada elemento de la cubierta

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

intersecta a dos elementos de la cubierta $\mathcal{B}$. 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

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

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

Por (3.3)

$$\begin{aligned}h_\mu(\mathcal{P}) = \lim_{n\to\infty}\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) &\leq 1 + \lim_{n\to\infty}\frac{1}{n}H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}\right)\\ &\leq 1 + \log 2 + H(T,\mathcal{U})\end{aligned} \tag{3.4}$$

Así

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

y como $\delta > 0$ es arbitraria

$$h_\mu(T) \leq 1 + \log 2 + h_{\mathrm{top}}(T).$$

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

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

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

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

Y luego $h_\mu(T) \leq h_{\text{top}}(T)$

II) Sea $\varepsilon > 0$. Para cada $n \in \mathbb{N}$ sea E_n un conjunto (n, ε) separado de cardinalidad máxima. Consideremos la medida

$$\mu_{E_n} = \frac{1}{s_T(n,\varepsilon)} \sum_{x \in E_n} \delta_x$$

y definamos

$$\mu_n \sum_{k=0}^{n-1} \mu_{E_n} \circ T^{-k}.$$

Por la convexidad de ϕ para cualquier particion $\mathcal{Q}$

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

Escojamos una subsucesión $\{n_j\}$ tal que $\{\mu_{n_j}\}$ converge a $\mu \in \mathcal{M}_T(X)$ y

$$\lim_{j \to \infty} \frac{1}{n_j} \log s_T(n_j, \varepsilon) = \overline{s}_T(\varepsilon).$$

Demostraremos que $h_\mu(T) \geq s_T(\varepsilon)$, para lo cual utilizaremos el siguiente

Lema 3.5. *Sea $\mu \in \mathcal{M}(X)$. Para todo $\varepsilon > 0$ existe una partición finita $\mathcal{P}$ tal que diam$(\mathcal{P}) < \varepsilon$ y $\mu(\partial P) = 0$ $\forall P \in \mathcal{P}$.*

Sea $\mathcal{P}$ la partición dada por el Lema 3.5 . Entonces cada átomo de $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$ contiene a lo más un punto de E_n. Así

$$\begin{aligned}
H_{\mu_{E_n}}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) &= -\sum_P \phi(\mu_{E_n}(P)) = -\sum_P \phi\big(\frac{1}{s_T(n,\varepsilon)} \#(E_n \cap P)\big) \\
&= \sum_{x \in E_n} \phi\big(\frac{1}{s_T(n,\varepsilon)}\big) = \log s_T(n, \varepsilon)
\end{aligned}$$

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

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

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. Po lo tanto

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

y luego

$$\begin{aligned}H_{\mu_n}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq H_{\mu_{E_n}}(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) &= \frac{1}{n}\sum_{l=0}^{n-1} H_{\mu_{E_n}}\big(T^{-l}\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}\big) \\ &\geq \frac{1}{n}\sum_{k=0}^{m-1} H_{\mu_{E_n}}\big(T^{-jm-k}\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}\big) \\ &\geq \frac{1}{n}\sum_{k=0}^{m-1}\left(H_{\mu_{E_n}}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) - 2m\log(\#\mathcal{P})\right) \\ &\geq \frac{m}{n} H_{\mu_{E_n}}(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) - \frac{2m^2}{n}\log(\#\mathcal{P}) \\ &= \frac{m}{n}\log s_T(n,\varepsilon) - \frac{2m^2}{n}\log(\#\mathcal{P}).\end{aligned}$$

Poniendo $n = n_j$ y tomando límite cuando $j \to \infty$

$$H_\mu(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq m s_T(\varepsilon)$$

y así

$$h(T,\mathcal{P}) = \lim_{m\to\infty} \frac{1}{m} H_\mu(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}) \geq s_T(\varepsilon).$$

$\square$

TEOREMA 3.5. *Sean $A \in GL(n,\mathbb{Z})$ y $\tilde{A} : \mathbb{K}^N \hookleftarrow$ la transformación inducida. Sea $R_a : \mathbb{K}^N \hookleftarrow$ 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 $\pi : \mathbb{R}^N \to \mathbb{K}^N$ la proyección natural y d la distancia en $\mathbb{K}^n$ inducida por la euclidiana. Sea

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

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 valida 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

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

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

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

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

$$m\left(\bigcap_{k=0}^{n-1} T^{-k} P_{i_k}\right) \leq m(D_n(\pi(0),\varepsilon,\tilde{A}),$$

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

Entonces

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

Como $\varepsilon > 0$ fue arbitraria

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

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

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

es unión ajena y así

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

Luego

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

y entonces

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

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

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

DEFINICIONES 23. 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,ε) generador para K si $\forall y \in K$ $\exists x \in F$ tal que $d(T^i x, T^i y) \leq \varepsilon$ para $0 \leq i \leq n-1$. Sea $r(K,n,\varepsilon)$ la mínima cardinalidad de un (n,ε) generador.

(b) Decimos que $E \subset K$ está (n,ε) separado si $\forall x, y \in E$, $x \neq y$ $\exists i \in \{0, \dots, n-1\}$ tal que $d(T^i x, T^i, y) > \varepsilon$. Sea $s(K,n,\varepsilon)$ la máxima cardinalidad de un (n,ε) separado.
Se tiene como antes $r(K,n,\varepsilon) \leq s(K,n,\varepsilon) \leq r(K,n,\varepsilon/2)$.

(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 3.5. Sean μ la medida de Lebesgue y $\|\|$ una norma en $\mathbb{R}^N$. Sean*

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

Entonces

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

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

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

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

$$r(K,n,\varepsilon) \geq \mu(K)/\mu(D_n(0,2\varepsilon,A).$$

Por lo tanto, cuando $\mu(K) > 0$

$$h(K,A) \geq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n}\log \mu(D_n(0,\varepsilon,A))\right].$$

Si q es suficientemente grande $K \subset K_q := [-q,q]^N$. Si $E \subset K$ es (n,ε) separado entonces $\bigcup_{x\in E} D_n(x,\varepsilon/2,A)$ es unión adjunta y está contenida en $K_{q+\varepsilon}$. Así

$$s(K,n,\varepsilon)\mu(D_n(0,\varepsilon/2,A)) \leq 2^N(q+\varepsilon)^N,$$

y luego

$$\limsup_{n\to\infty} \frac{1}{n}\log s(K,n,\varepsilon) \leq \limsup_{n\to\infty} \left[-\frac{1}{n}\log \mu(D_n(0,\varepsilon/2,A))\right],$$

$$h(K,A) \leq \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n}\log \mu(D_n(0,\varepsilon,A))\right].$$

Así $h_{top}(A) = h(K,A)$ para cualquier compacto con $\mu(K) > 0$. $\square$

LEMA 3.7. *Sean $A, \tilde{A}$ como en el Teorema 3.5. Entonces*

$$h_{top}(A) = h_{top}(\tilde{A})$$

Demostración. Sea $q < \frac{1}{2}$ Sea $0 < \varepsilon < \min(1, \|A\|^{-1})$. Supongamos que $E \subset K_q$ es (n,ε) separado. Entonces $\pi(E)$ es (n,ε) 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, ε) separado para $\tilde{A}$. Sea $E = \pi^{-1}(E) \cap [-\frac{1}{2}, \frac{1}{2}[^N$. Entonces E es (n, ε) 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 $h_{\text{top}}(\tilde{A}) \leq h(K_{\frac{1}{2}}, A)$. $\square$

Teorema 3.6. *Sean $A, \tilde{A}$ como en el Teorema 3.5. 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 ≤ 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

$$\|(x_1, x_2)\| = \max(\|x_1\|, \|x_2\|)$$

define una norma en $\mathbb{R}^N$. Por el Lema 3.6

$$h_{top}(A) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\left[-\log m_1(D_n(0,\varepsilon,A_1)) - \log m_2(D_n(0,\varepsilon,A_2))\right].$$

Se tiene que

$$m_1(D_n(0,\varepsilon,A_1)) \leq m_1(A_1^{-(n-1)}B_{\|\|_1}(0,\varepsilon)) = |\det A_1|^{1-n}\, m_1(B_{\|\|_1}(0,\varepsilon))$$
$$m_2(D_n(0,\varepsilon,A_2)) \leq m_2(B_{\|\|_2}(0,\varepsilon)).$$

Por lo tanto

$$\frac{1}{n}\left[-\log m_1(D_n(0,\varepsilon,A_1)) - \log m_2(D_n(0,\varepsilon,A_2))\right]$$
$$\geq \frac{(n-1)}{n}\log|\det A_1| - \frac{1}{n}\log m(B_{\|\|}(0,\varepsilon)),$$

$$h_{\text{top}}(A) \geq \log|\det A_1| = \sum_{|\lambda_i>1|}\log|\lambda_i|.$$

Para probar la desigualdad opuesta escribamos $\mathbb{R}^N = E_1 \oplus E_k$ tal que $A(E_j) \subset E_j$ y todos los eigenvalores de $A_j = A \mid E_j$ tienen el mismo módulo μ_j. Así $A = A_1 \oplus \quad \oplus A_k$ y entonces

$$h_{\text{top}}(A) \leq h_{\text{top}}(A_1) + \cdots + h_{\text{top}}(A_k).$$

Demostrando $h_{top}(A_j) \leq \max\{0, (demE_j)\log\mu_j\}$, 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 μ, entonces*

$$h_{top}(A) \leq \max\{0, P\log\mu\}\,.$$

Demostración.

Si $\|A\| \leq 1$, un $(1,\varepsilon)$ generador para un compacto K es también un (n,ε) generador. Así $r(K,n,\varepsilon) \leq r(K,1,\varepsilon)$, luego $h(K,A) = 0$ y así

$$h_{\mathrm{top}}(A) = 0 = \max\{0, P\log\|A\|\}\,.$$

Si $\|A\| > 1$, sean $K = K_1$ y $c = \max\{\|x\| : x \in K\}$.

Para $0 < \delta < 1$, sea $F(\delta) = \{\delta(n_1, \dots, n_P) \in K : n_i \in \mathbb{Z}\}$.

Si $N = [2/\delta]$, $\#F(\delta) \leq (N+1)^P < (2/\delta)^P$.

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

$$r(K,n,\varepsilon) \leq \#F(\delta) \leq (2\|A\|^n c/\varepsilon)^P,$$

$$\limsup_{n\to\infty} \frac{1}{n}\log r(K,n,\varepsilon) \leq \limsup_{n\to\infty} P\left[\log\|A\| + \frac{1}{n}\log(2c/\varepsilon)\right] = P\log\|A\|,$$

$$h_{\mathrm{top}}(A) \leq P\log\|A\| = \max\{0, P\log\|A\|\}\,.$$

Aplicando esta desigualdad a A^n y utilizando la Proposición 3.8

$$h_{\mathrm{top}}(A) = \frac{1}{n} h_{\mathrm{top}}(A^n) \leq \max\left\{0, P\log\|A^n\|^{\frac{1}{n}}\right\}.$$

Como $\lim_\infty \|A^n\|^{\frac{1}{n}} =$ radio espectral de $A = \mu$, tenemos

$$h_{\mathrm{top}}(A) \leq \max\{0, P\log\mu\}$$

$\square$

4. Presión topológica y estados de equilibrio

Consideremos el subcorrimiento de tipo finito (X_A, σ). Dado $0 < \theta < 1$ definimos en X_A la métrica

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

Sea $F_\theta = \{f : (X_A, d_\theta) \to \mathbb{C} \mid f \quad \text{es Lipschitz}\}$. Nuestro propósito es demostrar la siguiente versión del principio variacional: Si $f \in F_\theta$, $f(X_A) \subset \mathbb{R}$, entonces existe una única $m \in \mathcal{M}_\sigma(X_A)$ llamada *estado de equilibrio* tal que

$$h_m(\sigma) + \int f dm = P(f) := \sup\left\{h_\mu(\sigma) + \int f d\mu : \mu \in m_\sigma(X_A)\right\}$$

$P(f)$ se llama *presión topológica*. En particular tenemos que $P(0) = h_{\text{top}}(\sigma)$. Si $f : X_A \to \mathbb{C}$ definimos

$$\text{Var}_n f = \sup\{|f(x) - f(y)| : d_\theta(x,y) \leq \theta^n\}.$$

Es fácil ver que $|f(x) - f(y)| \leq c\, d_\theta(x,y) \iff \forall n \geq 0 \ \text{Var}_n f \leq c\,\theta^n$. Para $f \in F_\theta$ definimos

$$\|f\|_\infty = \sup\{|f(x)| : x \in X_A\},\ |f|_\theta = \sup\{\text{Var}_n f/\theta^n : n \geq 0\}$$
$$\|f\|_\theta = \|f\|_\infty + |f|_\theta$$

PROPOSICIÓN 3.12. $(F_\theta, \|\quad\|_\theta)$ *es un espacio de Banach*

DEFINICIÓN 24. Dos funciones $f, g \in F_\theta$ se llaman *cohomologas* y escribimos $f \sim g$, si $\exists\ h \in C(X_A)$ tal que $f = g + h \circ \sigma - h$.

PROPOSICIÓN 3.13. *Si $f \in F_\theta$ entonces existen $g, h \in F_{\theta^{\frac{1}{2}}}$ tales que $f = g + h - h \circ \sigma$ y $g(x) = g(y)$ siempre que $x_n = y_n$ para $n \geq 0$. O sea que g depende solo de las coordenadas en el "futuro".*

Demostración. Para cada $j \in [k]$ escojamos una sucesión del "pasado" $z(j) = (z_n(j))_{n \leq 0}$ con $A_{z_n(1) z_{n+1}(j)} = 1$, $z_0(j) = j$. Para cada $x \in X_A$ definamos $\varphi(x)$ mediante

$$\varphi(x)_n = \begin{cases} x_n & \text{si} \quad n \geq 0 \\ z_n(x_0) & \text{si} \quad n \leq 0 \end{cases}.$$

Definamos $h(x) = \sum_{n=0}^{\infty} f(\sigma^n(x)) - f(\sigma^n \varphi(x))$. La serie converge ya que

$$|f(\sigma^n(x)) - f(\sigma^n \varphi(x)| \leq \text{Var}_n f \leq \|f\|_\theta\ \theta^n.$$

$$\begin{aligned} h(x) - h(\sigma x) &= \sum_{n\geq 0} f(\sigma^n(x)) - f(\sigma^n\varphi(x)) - f(\sigma^{n+1}x) - f(\sigma^n\varphi\sigma(x)) \\ &= f(x) - \left[f(\varphi(x)) + \sum_{n\geq 0} f(\sigma^{n+1}\varphi(x)) - f(\sigma^n\varphi\sigma(x)) \right] \\ &:= f(x) - g(x) \end{aligned}$$

obviamente g depende solo del futuro. Probando que $h \in F_{\theta^{\frac{1}{2}}}$, tenemos que $g \in F_{\theta^{\frac{1}{2}}}$. Basta probar que $\exists K > 0$ tal que $\mathrm{Var}_{2N} h \leq K\theta^N$, ya que entonces $\mathrm{Var}_{2N+1} h \leq (K\theta^{-\frac{1}{2}})\theta^{N+\frac{1}{2}}$.

Sean $x, y \in X_A$ tales que $x_i = y_i$ para $|i| < 2N$, entonces

$$|f(\sigma^n x) - f(\sigma^n y)|, |f(\sigma^n\varphi x) - f(\sigma^n\varphi y)| \leq |f|_\theta \, \theta^{2N-n}$$

para $0 \leq n \leq N$. Si $n \geq 0$

$$|f(\sigma^n x) - f(\sigma^n\varphi x)|, |f(\sigma^n y) - f(\sigma^n\varphi y)| \leq |f|_\theta \theta^n$$

Así

$$\begin{aligned} |h(x) - h(y)| &\leq 2|f|_\theta \sum_{h=0}^{N} \theta^{2N-n} + 2|f|_\theta \sum_{n\geq N+1} \theta^n \\ &= 2|f|_\theta \theta^N \left[\frac{1-\theta^{N+1}}{1-\theta} + \frac{\theta}{1-\theta} \right] \leq 4|f|_\theta \frac{\theta^N}{1-\theta} \end{aligned}$$

La Proposición anterior define una transformación continua $W : F_\theta \to F_{\theta^{\frac{1}{2}}}$ dada por $W(f) = g$. Más aún, si d_θ^+ denota la métrica en X_θ^+ análoga a d_θ, $W(f)$ puede identificarse con un elemento de

$$F^+_{\theta^{\frac{1}{2}}} = \left\{ g : (X_A^+, d_{\theta^{\frac{1}{2}}} \right\}^+) \to \mathbb{C} : g \text{es Lipschitz}.$$

Cuando $f \in F_\theta$ depende unicamente del futuro $W(f) = f$. El corrimiento $\sigma : X_A \hookleftarrow$ ($\sigma : X_A^+ \hookleftarrow$) induce un operador $\sigma^* : F_\theta \to F_\theta$ ($\sigma^* : F_\theta^+ \to F_\theta^+$) definido por $\sigma^* f = f \circ \sigma$.

Para $f \in F_\theta^+$ definimos el operador $L_f : F_\theta^+ \to F_\theta^+$ mediante

$$L_f W(x) = \sum_{\sigma y = x} \exp \circ f(y) \, W(y).$$

L_f es un operador lineal acotado. Cuando f es real y $L_f 1 = 1$ decimos que L_f está normalizado

Proposición (Desigualdad básica). *Sea $f = u + iv \in F_\theta^+$. Si L_u está normalizado entonces $\exists C > 0$ tal que*

$$|L_f^n W|_\theta \leq C \, \|W\|_\infty + \theta^n \, |W|_\theta$$

para todo $W \in F_\theta^+$, $n \geq 0$.

Demostración Primero demostraremos que $\exists C_0 > 0$ tal que

$$|L_f W|_\theta \leq C_0 \|W\|_\infty + \theta |W|_\theta$$

Para $x \in X_A^+$, $i \in [k]$ tal que $A_{ix_0} = 1$ definimos $ix \in X_A^+$ mediante $(ix)_0 = i$, $(ix)_{n+1} = x_n$ para $n \geq 0$. Si $x, y \in X_A^+$, $x_0 = y_0$, entonces $d_\theta^+(ix, iy) \leq \theta d_\theta^+(x, y)$ y así

$$\begin{aligned}
|L_f W(x) - L_f W(y)| &\leq \sum_{Aix_0=1} |e^{f(ix)} W(ix) - e^{f(iy)} W(iy)| \\
&\leq \sum_{Aix_0=1} |e^{f(ix)} - e^{f(iy)}||W(ix)| + \sum_{Aix_0=1} e^{u(iy)} |W(ix) - W(iy)| \\
&\leq k |\exp \circ f|_\theta \, \theta \, d_\theta^+(x, y) \, \|W\|_\infty + \theta \, d_\theta^+(x, y)
\end{aligned}$$

Procediendo por inducción supongamos $|L_f^n W|_\theta \leq C_{n-1} \|W\|_\infty + \theta^n |W|_\theta$. Entonces

$$\begin{aligned}
|L_f^{n+1} W|_\theta &\leq C_{n-1} \|L_f W\|_\infty + \theta^n |L_f W|_\theta \\
&\leq C_{n-1} \|W\|_\infty + \theta^n (C_0 \|W\|_\infty \, \theta |W|_\theta) \\
&\leq (C_{n-1} + \theta^n C_0) \|W\|_\infty + \theta^{n+1} |W|_\theta
\end{aligned}$$

$$C_{n-1} + \theta^n C_0 = \sum_{k=1}^{n} \theta^k C_0 = \frac{1 - \theta^{n-1}}{1 - \theta} C_0 \leq \frac{C_0}{1 - \theta}$$

Teorema de Perron y Frobenius. *Supongamos que $\sigma : X_A^+ \hookleftarrow$ es topológicamente mezcladora. Sea $f \in F_\theta^+$, $f(X_A^+) \subset \mathbb{R}$, entonces*

1. *$\exists\ \lambda$ eigenvalor maximal positivo simple de $L_f : C(X_A^+) \to C(X_A^+$, con eigenfunción estrictamente positiva $h \in F_\theta^+$*
2. *El resto del espectro de $L_f : F_\theta^+ \to F_\theta^+$ está contenido en un disco de radio $R < \lambda$.*
3. *$\exists!\ \mu \in \mathcal{M}(X_A^+)$ tal que $L_f^* \mu = \lambda \mu$. O sea*
$$\int L_f v \, d\mu = \lambda \int v d\mu$$
para toda $v \in C(X_A^+)$.
4. *Si h es como en (1) y tal que $\int h d\mu = 1$ entonces*
$$\lim_{n \to \infty} \lambda^{-n} L_f^n(v) = h \int v d\mu$$
para toda $v \in C(X_A^+)$
5. *Si h es como en (4), $h\mu$ es invariante bajo σ.*

Demostración Sea

$$\Lambda = \left\{ g \in C(X_A^+, [0,1]) : g(x) \le g(y) \exp\big(\frac{|f|_\theta\, \theta^n}{1-\theta}\big) \,\text{si}\, d_\theta(x,y) < \theta^n \right\}.$$

Es fácil ver que Λ es convexo y cerrado bajo límites uniformes. Si $x, y \in X_A^+$ $x_i = y_i$ $|i| \le n$,

$$|g(x) - g(y)| \le g(y)(\exp\left(\frac{|f|_\theta\, \theta^n}{1-\theta}\right) - 1) \le \|g\|_\infty \frac{|f|_\theta\, \theta^n}{1-\theta} \;\exp\left(\frac{|f|_\theta\, \theta^n}{1-\theta}\right).$$

Así, $\Lambda \subset F_\theta^+$ y Λ es una familia equicontinua. Por el teorema de Ascoli, Λ es compacto con la norma uniforme.

Para cada $N \in \mathbb{N}$ definimos $L_N(g) = L_f(g{+}1/N)/\left\|L_f(g+1/N)\right\|_\infty$, $g \in \Lambda$. Si $x, y \in X_A^+$, $x_i = y_i$ para $|i| \le N$, entonces

$$L_f\left(g + \frac{1}{N}\right)(x) \le L_f\left(g + \frac{1}{N}\right)(y)\;\exp(\frac{|f_\theta|\theta^N}{1-\theta})$$

y por lo tanto $L_N(\Lambda) \subset \Lambda$.

Como Λ es convexo y compacto (con la norma uniforme) podemos aplicar el teorema de punto fijo de Schauder-Tijonov a cada $L_N : \Lambda \to \Lambda$ para ver que $\exists h_N \in \Lambda$ tal que $L_N(h_N + 1/N) = \lambda_N h_N$, donde $\lambda_N = \left\|L_f(h_N + 1/N)\right\|_\infty$

Por la compacidad de Λ existe una subsucesión $h_{N_m} \to h \in \Lambda$, y por continuidad, $L_f h = \lambda h$ donde $\lambda = \left\|L_f(h)\right\|_\infty$. Para ver que λ es positivo notemos que

$$\lambda_N h_N(x) = \sum_{\sigma y = x} e^{f(y)}(h_N(y) + 1/N) \ge (\inf|h_N| + 1/N)\exp(-lVfrV_\infty),$$

luego $\lambda_N(\inf h_N) \ge (\inf h_N + 1/N)\exp(-\left\|f\right\|_\infty)$.

Así $\lambda_N \ge \exp(-\left\|f\right\|_\infty)$ y entonces $\lambda \ge \exp(-\left\|f\right\|_\infty)$.

Supongamos que h no es estrictamente positiva, entonces $\exists x \in X_A^+$ tal que $h(x) = 0$. Luego

$$\sum_{\sigma^n y = x} \exp \circ f^n(y) h(y) = \lambda^n h(x) = 0$$

para $n \ge 1$, donde $f^n(y) = f(y) + \cdots + f(\sigma^{n-1}y)$. En particular $h(y) = 0$ si $\sigma^n y = x$ para algún $n \ge 0$. Como σ es transitivo, el conjunto de tales y es denso en X_A^+, así $h \equiv 0$. Pero $\lambda = \left\|L_f h\right\|_\infty > 0$, lo cual da una contradicción.

Para demostrar que λ es simple, sea $g \in C(X_A^+)$ otra eigenfunción correspondiente a λ y sea

$$t = \inf\left\{\frac{g(x)}{h(x)} : x \in X_A^+\right\} = \frac{g(z)}{h(z)}.$$

Entonces $g(x)-th(x) \geq 0 = g(z)-th(z)$ para todo $x \in X_A^+$. Repitiendo el razonamiento previo concluímos que $g - th \equiv 0$.

Con h y λ como arriba, sea $g = f - \log h \circ \sigma + \log h - \log \lambda$. Entonces $L_g = \lambda^{-1}\Delta(h)^{-1}L_f\Delta(h)$ donde $\Delta(h)$ es el operador multiplicación por h. Más aún $L_g 1 = 1$ y así L_g está normalizado. Observemos que esto nos permite deducir las propiedades restantes de L_f de las propiedades correspondientes de L_g.

Como $\mathcal{M}(X_A^+)$ es convexo y compacto, el operador $L_g^* : \mathcal{M}(X_A^+) \to \mathcal{M}(X_A^+)$ tiene un punto fijo m, por el teorema de Schauder-Tijonov. Demostraremos la unicidad de m y (4) simultaneamente.

Sea $v \in F_\theta^+$, entonces $\left\{L_g^n v\right\}$ es equicontinua ya que

$$\mathrm{Var}_k(L_g^n v) \leq |L_g^n v\theta|_\theta \theta^k \leq C\theta^k \left\|v\right\|_\infty + \theta^{k+n}|v|_\theta.$$

Por lo tanto hay una subsucesión convergente $L_g^{n_m} v \to v^*$. Ya que $\sup L_g^{n+1} v \leq \sup L_g^n v$, para todo $N \in \mathbb{N}$

$$\sup v^* = \inf_n (\sup L_g^n \theta) = \sup L_g^N v^*.$$

Sea $v^*(x_0) = \sup v^* = L^N g v^*(x_N)$. Así

$$v^*(x_0) = \sum_{\sigma^N y = x_N} \exp \circ g^N(y) v^*(y)$$

Como L_g está normalizado, ésta es una combinación convexa y así $v^*(y) = v^*(x_0)$ cuando $\sigma^N y = x_N$. Como σ es topológicamente mezcladora, v^* es constante. Como $L_g^* m = m, \int L_g^n v \; dm = \int v dm$ y así $v^* = \int v dm$.

Como F_θ^+ es uniformemente denso en $\subset (X_A^+)$ podemos suponer que $v \in \subset (X_A^+)$. El argumento anterior puede repetirse a cualquier subsucesión $\left\{L_g^{k_n} v\right\}$ y así $\lim_{n\to\infty} L_g^n v = \int v dm$. Esto completa la demostración de (3) y (4).

Para demostrar (5) Sea $v \in C(X_A^+)$, entonces $L_g(v \circ \sigma) = v$ y así $\int v \circ \sigma dm = \int L_g(v \circ \sigma) dm = \int v dm$.

Para probar (2) es suficiente demostrar que $L_g \mid C^\perp$ tiene radio espectral < 1 donde

$$C^\perp = \left\{ w \in F_\theta^+ : \int w dm = 0 \right\}.$$

$$|L_g^{n+r} w|_\theta \leq C \left\|L_g^r w\right\|_\infty + \theta^n |L_g^r w|_\theta.$$

$$\lim_{r\to\infty} \left\|L_g^r w\right\|_\infty = \int w dm = 0.$$

Como el conjunto $\left\{w \in C^{\perp} : \|w\|_{\theta} \leq 1\right\}$ es compacto (con la norma uniforme), para n, r grandes tenemos $\varepsilon > 0$ tal que $\left\|L_g^{n+r} w\right\|_{\theta} \leq \varepsilon < 1$ si $w \in C^{\perp}$, $\|w\|_{\theta} \leq 1$.

OBSERVACIÓN 3.5. **Teorema de Perron para matrices**

En el caso en que f depende de dos coordenadas solamente o sea $f(x) = f(x_0, x_1)$ se tiene la matriz k x k $M_{ij} = A_{ij}\ \exp f(i,j)$. Sea λ el eigenvalor dominante positivo. En este caso $h(x) = h(x_0)$ donde $\sum\limits_i h(i) Mij = h(j)$.

Definiendo $g(i,j) = \log h(i) - \log h(j) - \log \lambda + f(i,j)$ la matriz correspondiente a L_g es

$$P_{ij} = A_{ij}\ \exp g(i,j) = \frac{M_{ij} h(i)}{\lambda h(j)},$$

su transpuesta es la matriz correspondiente a L_g^* y es estocástica ya que

$$\sum_i P_{ij} = \frac{\sum_i M_{ij} h(i)}{\lambda h(j)} = \frac{\lambda h(j)}{\lambda h(j)} = 1$$

Sea π tal que $\sum_j P_{ij} \pi(j) = \pi(i)$ y $\sum_i \pi(i) = 1$. Sea $C = C(0; i_0, \ldots, i_n)$ un cilindro,

$$L_g \chi_C(X) = \sum_i P_{ix_0} \chi_C(ix) = P_{i_0 i_1}\ \chi_C(1; i_1, \ldots, i_n)(x).$$

Así

$$m(C(j; i_0, \ldots, i_n)) = \int \chi_C\, dm = \int L_g \chi_C\, dm = P_{i_0 i_1} m(C(1; i_1, \ldots, i_n)).$$

También

$$L_g \chi_{C(0;i)}(X) = P_{i_0 x_0} = \sum_l P_{i_0 l} \chi_{C(0;l)}(x).$$

Así

$$m(C(0;i)) = \int L_g X_{c(0;i)} = \sum_j P_{ij} m(C(0;j)).$$

Luego $m(C(0;i)) = \pi(i)$ y entonces

$$m(C(j; i_0, \ldots, i_n)) = P_{i_0 i_1} \cdots P_{i_{n-1} i_n} \pi(i_n).$$

Bibliografía

[Ma] R. Mañe. *Teoria Ergódica.* Projeto Euclides. IMPA, Rio de Janeiro 1983.
[Pa] W. Parry, M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics.* Asterisque 187-188, 1990.
[Pe] K. Petersen. *Ergodic Theory.* Cambridge University Press 1983.
[Po] M. Pollicott. *Lectures on Ergodic Theory and Pesin Theory.* London Math. Soc. LNS 180, Cambridge University Press 1994.
[Wa] P. Walters. *An introduction to Ergodic Theory.* GTM 79, Springer 1982.