Teorema Central de Límite y Estabilidad Estocástica para transformaciones Expansoras $C^{1+\nu_0}$

Angel E. Terán González

16 de septiembre de 2014

CONTENIDO

CAPÍTULO 1

Introducción

La evolución en el tiempo de diversos fenómenos naturales es descrito por transformaciones $f : M \to M$ sobre una variedad M, la cual modela los posibles estados del fenómeno y a la cual denominamos *espacio de fase*. Físicamente se estudian observaciones numéricas relativas a los estados, las cuales corresponden a funciones φ reales o complejas definidas sobre el espacio fase M. De este modo, los datos experimentales sobre el sistema vienen dados, usualmente, en la forma de sucesiones de "mediciones" $\varphi(f^j(x))$, donde $x \in M$ y $j \geqslant 0$. De tales datos se intenta extraer las principales propiedades intrínsecas del proceso dinámico fundamental.

Muy frecuentemente, estas *observaciones temporales* $\varphi(f^j(x))$ tienen un comportamiento bastante complicado y "errático" cuando el tiempo j varía, e incluso para leyes de evolución simples (representada por f).

Tal comportamiento es interpretado como "caótico", de manera que la dinámica puede ser difícil de entender en términos determinísticos, y un analisis estocástico de las observaciones temporales puede ser un enfoque alternativo para el estudio de tales dinámicas. Esto es, uno considera las observaciones temporales esencialmente como una sucesión aleatória, y se centra en determinar sus propiedades estadísticas. Son de particular interés aquellas propiedades las cuales son intrínsecas de la dinámica del sistema, esto es, independientes de la elección de un estado inicial x, e incluso la persistencia (robustez) de estas propiedades intrínsecas bajo pequeñas modificaciones aleatorias en el sistema original.

El presente trabajo está dedicado al estudio de las propiedades estadísticas de una clase particular de sistemas determinísticos con fuertes propiedades dinámicas, conocidas como transformaciones uniformemente expansoras. El estudio de estos sistemas y los métodos aplicados en ellos son el punto de partida para el abordaje de sistemas más generales y complejos.

1.1. Medidas Físicas

El primer objetivo del estudio de propiedades estadísticas es la existencia de los *promedios temporales asintóticos*

$$E_x(\varphi) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x))$$

para "muchos" puntos $x \in M$. Claramente, $E_x(\varphi)$ existe siempre que x es un punto periódico de f, esto es, cuando $f^k(x) = x$ para algún $k \geqslant 1$. Más generalmente, el Teorema Ergódico de Birkhoff afirma que los promedios temporales asintóticos existen para casi todo punto, con respecto de cualquier medida de probabilidad invariante por f, esto es, que el límite de los promedios temporales asintóticos existe para todo punto x en un conjunto $S \subset M$ con $\mu(S) = 1$ si la medida de probabilidad μ sobre M satisface la propiedad de *invarianza*: $\mu(f^{-1}B) = \mu(B)$ para todo boreliano $B \subset M$. Pero las medidas invariantes arbitrarias pueden carecer de significado físico. En general, tomamos "muchos" arriba en sentido de un "conjunto con medida de Lebesgue positiva" debido a que la medida de Lebesgue m es la que generaliza las nociones físicas de longitud, área, volumen, etc.

También es de interés conocer si los promedios temporales son independientes del punto inicial. Suponga que para toda función continua $\varphi : M \to \mathbb{R}$, el promedio asintótico $E_x(\varphi)$ existe y es independiente del punto x tomado en algún conjunto $B \subset M$ con medida de Lebesgue positiva. Entonces $\varphi \mapsto E(\varphi) = E_x(\varphi)$ (cualquier $x \in B$) define un operador lineal no-negativo sobre el espacio $C^0(M, \mathbb{R})$ de funciones reales continuas el cual, por el Teorema de representación de Riesz-Fischer, puede ser considerado como una medida de Borel μ en M:

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

(cualquier $x \in B$). Observe que tal medida μ puede ser "observada físicamente" calculando promedios temporales de funciones continuas para puntos $x \in M$ escogidos aleatoriamente (teniendo probabilidad positiva de tomar $x \in B$).

Esto motiva a la siguiente definición. Una medida de probabilidad μ invariante por f es una *medida física o SRB* (Sinai-Ruelle-Bower) para f si existe un conjunto de puntos $x \in M$ con medida de Lebesgue positiva tal que

$$\int \varphi d\mu = E(\varphi) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x)) \qquad \text{para toda } \varphi \in C^0(M, \mathbb{R}).$$

El conjunto de puntos $x \in M$ satisfaciendo esta propiedad es llamado *cuenca ergódica* de μ y es denotado $B(\mu)$. Una situación particular, se presenta cuando se tiene una transformación $f : M \to M$ (M variedad diferenciable) con una medida invariante μ_0 que es ergódica y absolutamente continua con respecto de la medida de Lebesgue m en M.

1.2. Propiedades Estadísticas

1.2.1 Funciones de Correlación

De ahora en adelante, sea $U \subset M$ algún conjunto abierto con $f(U) \subset U$. Supongamos que μ es la única medida SRB para f con soporte contenido en U, y analicemos el sistema $(f|_U\ ,\mu)$.

El próximo paso es entender la pérdida de memoria en el pasado por el sistema como evolución del tiempo. En terminos más precisos, uno desea conocer cómo el límite de las observaciones temporales $\varphi(f^n(x))$ en un instante $n \gg 1$ es afectado por valores iniciales $\psi(x)$ de alguna observación ψ (posiblemente $\psi = \varphi$). Esto es, expresado por medio de las *funciones de correlación*

$$C_n(\varphi,\psi) = \int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \quad \text{para toda } \varphi,\psi \in L^1(\mu).$$

Note que $C_n(\varphi,\psi) = 0$ corresponde, en términos probabilísticos, que $\varphi \circ f^n$ y ψ son variables aleatórias independientes. Decimos que (f,μ) es *mixing* si $C_n(\varphi,\psi) \to 0$ para todo par de funciones φ y ψ: los valores de $\varphi \circ f^n$ llegan a ser menos y menos dependientes de los valores de ψ cuando el tiempo tiende a infinito. Diremos que (f,μ) es *exponencialmente mixing* o tiene *decaimiento exponencial de correlaciones* si su "pérdida de memoria" ocurre con rapidez exponencial: existe $\tau < 1$ y para cada φ y ψ existe $c > 0$ tal que

$$|C_n(\varphi,\psi)| \leqslant c\tau^n \quad \text{para todo } n \geqslant 1.$$

1.2.2 Aleatoriedad

Otra caracterización importante de casi independencia (más precisamente, correlación débil) de observaciones sucesivas se puede dar a través del teorema central de límite y el teorema de las grandes desviaciones. Ambos resultados describen el comportamiento de los promedio temporales

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi(f^j(x))$$

alrededor de su valor esperado $\int \varphi d\mu$. Empezamos mencionando los siguientes teoremas (las demostraciones pueden ser vistas en [V1]), los cuales son resultados clásicos para variables aleatorias independientes idénticamente distribuidas.

Teorema 1 (Teorema Central de Límite para v.a.i.i.d.) *Sean $X_0, X_1, \cdots, X_n, \cdots$ variable aleatorias independientes idénticamente distribuidas que toman valores en $\mathbb{R}$, con promedio $\overline{X} = E(X_n) < +\infty$ y varianza $\sigma^2 = E((X_n - \overline{X})^2) \in (0, \infty^+)$. Entonces, dado cualquier intervalo abierto $A \subset \mathbb{R}$, la probabilidad de*

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (X_j - \overline{X}) \in A \quad \textit{converge a} \quad \frac{1}{\sigma\sqrt{2\pi}} \int_A \exp\left(\frac{-t^2}{2\sigma^2}\right) dt$$

cuando $n \to \infty$.

Teorema 2 (Teorema de las Grandes Desviaciones para v.a.i.i.d.) *Sean $X_0, X_1, \cdots, X_n, \cdots, \overline{X}$,y σ como en el teorema 1. Suponer, además, que $E(\exp(tX_n)) < \infty$ para todo $t \in \mathbb{R}$. Entonces, dado cualquier $\varepsilon > 0$, la probabilidad $\mathcal{P}(n, \varepsilon)$ de*

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} (X_j - \overline{X}) \right| > \varepsilon$$

converge a cero exponencialmente rápido cuando $n \to \infty$, en el sentido de que

$$\limsup \frac{1}{n} \log \mathcal{P}(n, \varepsilon) < 0.$$

Cada unos de los antes mencionados son parte de una familia completa de teoremas relacionados que incluyen resultados mucho más sofisticados. En los procesos estocásticos se tiene una versión de Teorema Central de Límite para Martingalas que mencionaremos sin demostrar (ver la demostración en [W]). Antes de establecer el teorema recordemos algunas nociones de la teoria de probabilidad que vamos a utilizar. Dada una sucesión $(X_n)_{n\geqslant 0}$ de variables aleatorias reales en un espacio de probabilidad $(M, \mathcal{B}, \mu)$, considere la función shift

$$\begin{aligned} \tau : \mathbb{R}^{\mathbb{N}} &\longrightarrow \mathbb{R}^{\mathbb{N}} \\ (\omega_n)_{n\geqslant 0} &\longmapsto (\omega_{n+1})_{n\geqslant 0} \end{aligned}$$

y la medida de probabilidad ν sobre $\mathbb{R}^{\mathbb{N}}$ dada por:

$$\nu(\{(\omega_n) \ : \ \omega_k \in A_0, \cdots, \omega_{k+m} \in A_m\}) = \mu(X_k^{-1}(A_0), \cdots, X_{k+m}^{-1}(A_m))$$

para cada $k \geqslant 0$, $m \geqslant 0$ y $A_0, \cdots, A_0$ intervalos en $\mathbb{R}$. Entonces, la sucesión $(X_n)_{n\geqslant 0}$ es llamada *estacionaria* si la medida ν es τ-invariante, y *ergódica* si ν es τ-ergódica. Un caso particular de relevante interés es cuando se tiene $X_n = X_0 \circ f^n$ para todo $n \geqslant 0$, donde X_0 es una función medible: la sucesión es estacionaria si μ es f-invariante, y es ergódica si μ es f-ergódica.
Ahora, sea $(\mathcal{B}_n)_{n\geqslant 0}$ una sucesión de σ-álgebras no creciente sobre M. Entonces $(X_n)_{n\geqslant 0}$ es un *Martingala de diferencia inversa* para $(\mathcal{B}_n)_{n\geqslant 0}$ si cumple las siguientes condiciones:

$$\begin{aligned} &(a) \quad X_n \text{ es } \mathcal{B}_n\text{-medible para todo } n \geqslant 0 \\ &(b+) \quad E(X_n|\mathcal{B}_{n+1}) = 0 \text{ para todo } n \geqslant 0. \end{aligned}$$

La condición $(b+)$ solo significa que $\int_A X_n d\mu = 0$ para todo $A \in \mathcal{B}_{n+1}$. La *Martingala de diferencia directa* son definida similarmente, reemplazando "no creciente" por "no decreciente" y $(b+)$ por

$$(b-) \quad E(X_n|\mathcal{B}_{n+1}) = 0 \quad \text{para todo} \quad n \geqslant 1, \quad E(X_0) = 0.$$

Una situación especial corresponde cuando X_n es independiente e idénticamente distribuida con esperanza igual a cero: Entonces la sucesión $(X_n)_{n\geqslant 0}$ es una martingala de diferencia directa para $\mathcal{B}_n = \sigma$-álgebra generada por $\{X_j \ : \ j \leqslant n\}$, y una martingala de diferencia inversa para $\mathcal{B}_n = \sigma$-álgebra generada por $\{X_j \ : \ j \geqslant n\}$ (la σ-álgebra generada por un conjunto de funciones es la σ-álgebra más pequeña con respecto a la que todas las funciones son medibles).

Teorema 3 (Teorema Central de Límite para Martingala) *Sea $(X_n)_{n\geqslant 0}$ una sucesión estacionaria, ergódica, martingala de diferencia inversa o directa, tal que $\sigma^2 = E(X_0^2)$ es estrictamente positiva y finita. Entonces, para todo $z \in \mathbb{R}$*

$$\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} X_j(x) < z\right\}\right) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}}\int_A \exp\left(\frac{-t^2}{2\sigma^2}\right)dt,$$

cuando $n \to \infty$.

Volviendo a nuestro contexto dinámico, decimos que un observable φ satisface el teorema central de límite para (f, μ) si existe $\sigma > 0$ tal que, para todo intervalo $A \subset \mathbb{R}$

$$\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\left[\varphi(f^j(x)) - \int \varphi d\mu\right] \in A\right\}\right) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}}\int_A \exp\left(\frac{-t^2}{2\sigma^2}\right)dt,$$

cuando $n \to \infty$. Por otro lado, se dice que φ satisface el teorema de las grandes desviaciones para (f, μ) si dado cualquier $\varepsilon > 0$ existe $h(\varepsilon) > 0$ tal que

$$\mu\left(\left\{x \in M \ : \ \left|\frac{1}{n}\sum_{j=0}^{n-1}\left[\varphi(f^j(x)) - \int \varphi d\mu\right]\right| > \varepsilon\right\}\right) \leqslant \exp(-nh(\varepsilon))$$

$n \geqslant 1$ suficientemente grande. Como veremos, las propiedades de este tipo cuando está φ tiene decaimiento de correlación suficientemente rápido. En cierto sentido, esto significa que los promedios temporales se comporta de una manera esencialmente aleatoria en grandes rangos de tiempo.

1.2.3 Estabilidad Estocástica

Muy a menudo, los modelos matemáticos $f : M \to M$ de un proceso físico dado involucran simplificaciones, donde una parte principal del proceso es aislado e influencias externas son descartadas ya que son demasiado complejas para ser tomadas en consideración y, con esperanza, demasiado pequeñas para ser pertinentes. Claramente estos procedimientos requieren una justificación, especialmente si estas ocurren con frecuencia, el sistema simplificado f puede ser estructuralmente inestable (las transformaciones g próximas de f pueden tener un comportamiento dinámico muy diferente).

En muchos ejemplos donde tales influencias externas no son completamente entendidas, o son demasiado complejas para ser expresadas efectivamente en términos determinísticos, podemos pensar de ellas como una especie de perturbaciones aleatorias. Uno entonces habla de *estabilidad estocástica* si la presencia de pequeñas perturbaciones con carácter aleatorio tienen sólo un pequeño efecto sobre el comportamiento asintótico de f. En términos más precisos, para cada $\varepsilon > 0$ pequeño uno considera iterados

$$x_j^\varepsilon = f_j \circ f_{j-1} \circ \cdots \circ f_1(x) \quad \text{para } x \in U \text{ y } j \geqslant 0,$$

donde las f_i son escogidas aleatoriamente e independientes cada una de la otra en la ε-vecindad de f, para alguna ley de distribución dada. Es conveniente la suposición que $f(\overline{U}) \subset U$ para asegurar que $f_i(U) \subset U$ para cada i siempre que ε sea suficientemente pequeño. Entonces, bajo condiciones generales, existen medidas de probabilidad μ_ε con $\mathrm{supp}(\mu_\varepsilon) \subset U$ y

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1}\varphi(x_j^\varepsilon) = \int \varphi d\mu_\varepsilon$$

para toda $\varphi : M \to \mathbb{R}$ continua y muchas (probabilidad positiva) trayectorias aleatorias $(x_j^\varepsilon)_{j\geqslant 0}$ con $x_0 \in U$. Diremos que $(f|_U\ , \mu)$ es *estable estocásticamente* si μ_ε converge a μ

en el sentido débil-$*$, esto es,

$$\int \varphi d\mu_\varepsilon \to \int \varphi d\mu \qquad \text{para toda } \varphi \in C^0(M, \mathbb{R})$$

cuando $\varepsilon \to 0$ (si μ_ε no es única entonces requeriremos convergencia a μ para todas estas medidas estacionarias con soporte contenido en U).

Observamos que mientras la estabilidad estructural requiere restricciones muy rígidas sobre la dinámica del sistema, la estabilidad estocástica parece ser mucho más factible para clases amplias de sistemas dinámicos determinísticos.

CAPÍTULO 2

Métricas Proyectivas y el Operador de Perron-Frobenius

En este capítulo establecemos las definiciones y algunas propiedades de las transformaciones expasoras, el operador de Perron-Frobenius (o transfer), conos convexos y sus métricas proyectivas. Además, en este capítulo se mostrará la construcción de una medida invariante absolutamente continua con respecto a la medida de Lebesgue.

2.1. Transformaciones Expansoras y Operador de Perron-Frobenius

Comenzamos esta sección introduciendo la definición de transformación expansora y el operador de Perron-Frobenius. Estableceremos algunas propiedades elementales de estas transformaciones.

Sean M una variedad compacta conexa y $f : M \to M$ una transformación de clase $C^{1+\nu_0}$. Diremos que f es *expansora*, si existe $\sigma > 1$ tal que

$$\|Df(x) \cdot v\| \geqslant \sigma \|v\| \quad \text{para todo } x \in M \text{ y todo } v \in T_x M. \tag{2.1}$$

para alguna métrica Riemanniana $\| \cdot \|$ (la elección en particular de tal métrica no es pertinente en este trabajo).

Una clase de ejemplos estándar proviene de la siguiente construcción. Sea $F : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal tal que $F(\mathbb{Z}^n) \subset \mathbb{Z}^n$. Entonces existe una única transformación f sobre el toro n-dimensional $T^n = \mathbb{R}^n/\mathbb{Z}^n$ tal que $f \circ \pi = \pi \circ F$, donde π es la proyección canonica sobre el toro. Además, si todos los autovalores de F tienen norma mayor que 1 entonces f es una transformación expansora. Claramente, la existencia de una constante $\sigma > 1$ como en (2.1) es una condición abierta sobre las transformaciones de clase C^1. De este modo, cualquier transformación suave de M bastante próxima de f, en el sentido C^1, también es una transformación expansora. Supondremos en lo sucesivo que el diámetro de M no es mayor que 1. También sea m el volumen Riemanniano normalizado en M.

La condición de expansión de vectores (2.1) implica, que la derivada Df es un isomorfismo en todo punto, en particular, f es un difeomorfismo local. Como una consecuencia, todo punto $y \in M$ tiene el mismo número $k \geqslant 1$ de preimágenes. Demostremos esto, procediendo por reducción al absurdo. Suponga que existen $y_1, y_2 \in M, y_1 \neq y_2$ tal que

$$\#f^{-1}(y_1) \neq \#f^{-1}(y_2).$$

Consideremos los siguientes conjuntos

$$A = \{y \in M : \#f^{-1}(y) = \#f^{-1}(y_1)\} \quad \text{y} \quad B = \{y \in M : \#f^{-1}(y) \neq \#f^{-1}(y_1)\}.$$

Tenemos entonces que $A \neq \varnothing$ (porque, $y_1 \in A$) y $B \neq \varnothing$ (porque, $y_2 \in B$). Además, A y B son conjuntos abiertos, dísjuntos y $M = A \cup B$. Esto establece que M tiene una separación no trivial, lo cual contradice el hecho de que M es conexa. Por tanto,

$$\#f^{-1}(y_1) = \#f^{-1}(y_2).$$

Así, todo punto $y \in M$ tiene el mismo número de preimágenes. Además, $\#f^{-1}(y) < +\infty$ porque de lo contrario existe una sucesión $(x_n)_n$ en M con $x_n \neq x_k$ $(n \neq k)$, tal que $x_n \in f^{-1}(y)$ y $(x_n)_n$ converge a un punto $x_0 \in M$. Entonces $f(x_0) = \lim_{n\to+\infty} f(x_n) = y$. Como consecuencia se tendría que f no es localmente inyectiva lo cual es contradictorio. Luego, $\#f^{-1}(y) = k < +\infty$.

Más aún, dada cualquier preimágen x de y, existen una vecindad V de y y una transformación $g : V \to M$ tales que $f \circ g = \mathrm{Id}$ y $g(y) = x$. Esta g es una inversa local de f y, por ser f expansora, la función g es contractiva:

$$d(g(y), g(z)) \leqslant \sigma^{-1} d(y, z) \qquad \text{para todo } y, z \in V$$

donde $\sigma > 1$ es la tasa de expansión de f. Probemos esta afirmación. Como f es localmente inyectiva, existe $r > 0$ tal que si $x_1 \neq x_2$ y $d(x_1, x_2) \leqslant r$ entonces $f(x_1) \neq f(x_2)$. Por otro lado, sea $r_1 > 0$ tal que para $y, z \in M$ con $d(y, z) < r_1$ existe una geodésica $\gamma : [0, 1] \to M$ con $\gamma(0) = y, \gamma(1) = z$ y $d(y, z) = \ell(\gamma)$, donde $\ell(\gamma)$ es la longitud de γ según la métrica Riemanniana. Tomemos $\rho = \min\{r, r_1\}$. Entonces para todo $y \in M$ existe un conjunto abierto $V = B(y; \rho_y)$, donde $B(y; \rho_y)$ es la bola abierta centrada en y y radio ρ_y, tal que la función $g = (f|_V)^{-1}$ satisface

- $(f \circ g)(z) = z \quad$ para todo $z \in V$,
- $g(y) = f^{-1}(y) = x$.

Con lo cual obtenemos $f \circ g = \mathrm{Id}$ y $g(y) = x$. Además,

$$\begin{aligned} d(g(y), g(z)) \leqslant \ell(g \circ \gamma) &= \int_0^1 \|D(g \circ \gamma)(t)\| dt \\ &= \int_0^1 \|D(g(\gamma(t)))\gamma'(t)\| dt \leqslant \sigma^{-1} \int_0^1 \|\gamma'(t)\| dt \\ &= \sigma^{-1}\ell(\gamma) = \sigma^{-1} d(y, z), \end{aligned}$$

y así queda demostrada la afirmación.
Usando la compacidad de M, dado $\bigcup_{y \in M} B(y; \rho_y)$ un cubrimiento abierto de M podemos extraer un subcubrimiento abierto finito $B(y_1; \rho_1), \ldots, B(y_n; \rho_n)$. Luego, tomando $\rho_0 = \min\{\rho_1, \ldots, \rho_n\}$ tenemos una constante uniforme. Resumiendo:

(i) Existe $k \geqslant 1$ tal que todo punto $y \in M$ tiene exactamente k preimágenes bajo f.

(ii) Existe $\rho_0 > 0$ tal que, dados $y_1, y_2 \in M$ con $d(y_1, y_2) \leqslant \rho_0$ podemos escribir $f^{-1}(y_j) = \{x_1^j, \ldots, x_k^j\}$, $j = 1, 2$ con

$$d(x_i^1, x_i^2) \leqslant \sigma^{-1} d(y_1, y_2) \quad \text{para cada } i = 1, \ldots, k.$$

El objetivo principal de este trabajo es estudiar algunas de las propiedades estadísticas, el teorema central de límite y la estabilidad estocástica para transformaciones expansoras de clase $C^{1+\nu_0}$. Primero demostremos la existencia y unicidad de medidas físicas para está clase particular de transformaciones, esto se obtiene basandose en dos propiedades fuertes: ergodicidad y continuidad absoluta con relación de la medida de Lebesgue. A continuación el Teorema principal del trabajo.

Teorema 4 *Sea f una transformación expansora de clase $C^{1+\nu_0}$, para algún $\nu_0 \in (0, 1]$. Entonces*

1) *f admite una única medida invariante μ_0 la cual es absolutamente continua con respecto a m. Además, μ_0 es exacta (así ergódica) y $d\mu_0/dm$ es estrictamente positiva, en particular, μ_0 es la única medida SRB de f.*
2) *(f, μ_0) es exponencialmente mixing y satisface el teorema central de límite en el espacio de Banach de funciones ν-Hölder continuas, para cualquier $\nu \in (0, \nu_0]$.*
3) *(f, μ_0) es estocásticamente estable bajo pequeñas perturbaciones aleatorias.*

La demostración de existencia de medidas invariantes absolutamente continuas consiste en encontrar una función de densidad φ_0 tal que la medida $\mu_0(A) := \int_A \varphi_0 dm$ sea invariante por f y desde luego, $\mu_0 \ll m$. Esto nos conduce naturalmente a considerar los operadores $\mathcal{L}$ (operador de *Perron-Frobenius*) y U definidos sobre $E = C^0(M, \mathbb{R})$, por

$$(\mathcal{L}\varphi)(y) = \sum_{i=1}^{k} \varphi(x_i) |\det Df(x_i)|^{-1} \quad \text{y} \quad (U\varphi)(x) = \varphi(f(x)),$$

donde $f^{-1}(y) = \{x_1, \ldots, x_k\}$ (ver arriba (ii)). De hecho U es el adjunto de $\mathcal{L}$ y así, tenemos la siguiente propiedad

$$\int (\mathcal{L}\varphi)\psi dm = \int \varphi(U\psi) dm. \tag{2.2}$$

Esta implica, en particular, que los puntos fijos de $\mathcal{L}$ estan directamente relacionados con las medidas invariantes absolutamente continuas, esto es, $\mu_0 = \varphi_0 m$ es una medida de probabilidad de Borel invariante por f si y sólo si φ_0 es una función no negativa con $\mathcal{L}(\varphi_0) = \varphi_0$ y $\int \varphi_0 dm = 1$.

Se demostrara en la sección 2.4 la existencia de un punto fijo para el operador $\mathcal{L}$ usando métrica proyectiva asociada a un cono convexo invariante del espacio de funciones continuas con logaritmo localmente Hölder continuo. La siguiente sección contiene la notación y algunas propiedades de las métricas proyectivas y conos que usaremos en este trabajo.

2.2. Conos y Métricas Proyectivas

El propósito de está sección es introducir las definiciones básicas de conos y las métricas proyectivas. Además estudiaremos algunas de sus propiedades las cuales utilizaremos en las próximas secciones.

Sea E un espacio vectorial. Un *cono* en E es cualquier subconjunto $C \subset E \setminus \{0\}$ tal que $tv \in C$ para todo $v \in C$ y todo $t > 0$. Diremos que C es un cono *convexo* si además $t_1 v_1 + t_2 v_2 \in C$ para todo $v_1, v_2 \in C$ y todo $t_1, t_2 > 0$. Definimos la *clausura* $\overline{C}$ del cono C como el conjunto de vectores $w \in \overline{C}$ para los cuales existen $v \in C$ y $(t_n)_{n \geqslant 1} \searrow 0$ tales que $(w + t_n v) \in C$ para todo $n \geqslant 1$.

En lo que sigue supondremos que los conos son *no degenerados* en el sentido que

$$\overline{C} \cap (-\overline{C}) = \{0\}. \tag{2.3}$$

Dados $v_1, v_2 \in C$ definimos

$$\begin{aligned}\alpha(v_1,v_2) &= \sup\{t>0 : v_2 - tv_1 \in C\} \quad \text{ y}\\ \beta(v_1,v_2) &= \text{ínf}\{s>0 : sv_1 - v_2 \in C\},\end{aligned}$$

con los usuales convenios $\sup\varnothing = 0$ e $\text{ínf}\,\varnothing = +\infty$. Note que $\alpha(v_1,v_2) \leqslant \beta(v_1,v_2)$ para todo $v_1, v_2 \in C$. En efecto, sean $t_0 = \alpha(v_1,v_2)$ y $s_0 = \beta(v_1,v_2)$. Entonces $v_2 - t_0v_1 \in C$ y $s_0v_1 - v_2 \in C$. Sumando y usando la convexidad de C obtenemos $(s_0 - t_0)v_1 \in C$. Como $-v_1 \notin C$ (por (2.3)) tenemos $s_0 \geqslant t_0$ y por tanto $\alpha(v_1,v_2) \leqslant \beta(v_1,v_2)$.

Además, $\alpha(v_1,v_2) < +\infty$ y $\beta(v_1,v_2) > 0$ para todo $v_1, v_2 \in C$. En efecto, procediendo por reducción al absurdo, suponga que $\alpha(v_1,v_2) = +\infty$. Entonces existe $(t_n)_{n\geqslant 1}$ con $t_n \nearrow +\infty$ y $v_2 - t_nv_1 \in C$ para cada $n \geqslant 1$. Haciendo $r_n = \frac{1}{t_n}$ tenemos que $r_n \searrow 0$ y $r_nv_2 - v_1 \in C$ para cada $n \geqslant 1$ y, por definición de la clausura de conos, resulta que $-v_1 \in \overline{C}$ contradiciendo (2.3). Por tanto, $\alpha(v_1,v_2) < +\infty$.

Por otro lado, si $\beta(v_1,v_2) = 0$ entonces existe $(s_n)_{n\geqslant 1}$ tal que $s_n \to 0$ y $s_nv_1 - v_2 \in C$ para todo $n \geqslant 1$. Usando la definición de clausura de conos se tiene $-v_2 \in \overline{C}$ lo que contradice (2.3). Por tanto, $\beta(v_1,v_2) > 0$.

Ahora, sea $\theta(v_1,v_2) = \log\left[\dfrac{\beta(v_1,v_2)}{\alpha(v_1,v_2)}\right]$ (con $\theta = +\infty$, si $\alpha = 0$ ó $\beta = +\infty$). En vista de la observación anterior $\theta(v_1,v_2)$ está bien definida y toma valores en $[0,+\infty]$ para $v_1, v_2 \in C$.

Proposición 5 *$\theta(\cdot,\cdot)$ es una métrica en el cociente proyectivo de C, esto es,*

(a) $\theta(v_1,v_2) = \theta(v_2,v_1)$ *para todo* $v_1, v_2 \in C$.

(b) $\theta(v_1,v_2) + \theta(v_2,v_3) \geqslant \theta(v_1,v_3)$ *para todo* $v_1, v_2, v_3 \in C$.

(c) $\theta(v_1,v_2) = 0$ *si y solo si existe* $t > 0$ *tal que* $v_2 = tv_1$.

Demostración. Afirmamos que $\alpha(v_2,v_1) = [\beta(v_1,v_2)]^{-1}$. Supongamos primero que $\alpha(v_2,v_1) > 0$. Por la definición de α se tiene

$$\begin{aligned}\alpha(v_2,v_1) &= \sup\{t>0 : v_1 - tv_2 \in C\}\\ &= \text{ínf}\{s>0 : sv_1 - v_2 \in C\}^{-1}\\ &= \beta(v_1,v_2).^{-1}\end{aligned}$$

Por otra parte, si $\alpha(v_2,v_1) = 0$ entonces $v_1 - tv_2 \notin C$ para todo $t > 0$. Usando la convexidad de C tenemos que $sv_1 - v_2 \notin C$ para todo $s > 0$ y por tanto, $\beta(v_1,v_2) = +\infty$. Luego, la afirmación se cumple también en este caso. En forma análoga se muestra que

$\beta(v_2, v_1) = [\alpha(v_1, v_2)]^{-1}$. Con esto la parte (a) de la Proposición se obtiene inmediatamente de la siguiente manera. Para todo $v_1, v_2 \in C$

$$\begin{aligned}\theta(v_1, v_2) &= \log\left[\frac{\beta(v_1, v_2)}{\alpha(v_1, v_2)}\right] = \log\left[\frac{1}{\alpha(v_1, v_2)\alpha(v_2, v_1)}\right]\\ &= \log\left[\frac{\beta(v_2, v_1)}{\alpha(v_2, v_1)}\right] = \theta(v_2, v_1).\end{aligned}$$

Para demostrar la parte (b), afirmamos que $\alpha(v_1, v_2)\alpha(v_2, v_3) \leqslant \alpha(v_1, v_3)$ para todo $v_1, v_2, v_3 \in C$. Esto es obvio si $\alpha(v_1, v_2) = 0$ o $\alpha(v_2, v_3) = 0$. Considere $\alpha(v_1, v_2) > 0$ y $\alpha(v_2, v_3) > 0$. Entonces existen sucesiones $(\gamma_n)_n$ y $(s_n)_n$ convergiendo crecientemente a $\alpha(v_1, v_2)$ y $\alpha(v_2, v_3)$ respectivamente, con $v_2 - \gamma_n v_1 \in C$ y $v_3 - s_n v_2 \in C$. Por convexidad y la definición de α se tiene que $s_n\gamma_n \leqslant \alpha(v_1, v_3)$ para todo $n \geqslant 1$. Ahora, tomando límite obtenemos la afirmación.

Por otro lado, probaremos que $\beta(v_1, v_2)\beta(v_2, v_3) \geqslant \beta(v_1, v_3)$ para todo $v_1, v_2, v_3 \in C$. Esto se cumple si $\beta(v_1, v_2) = +\infty$ o $\beta(v_2, v_3) = +\infty$. Así que supongamos $\beta(v_1, v_2) < +\infty$ y $\beta(v_2, v_3) < +\infty$. Entonces, existen sucesiones $(t_n)_n$ y $(s_n)_n$ convergiendo decrecientemente a $\beta(v_1, v_2)$ y $\beta(v_2, v_3)$ respectivamente tales que $t_n v_1 - v_2 \in C$ y $s_n v_2 - v_3 \in C$. Ahora usamos el hecho que el cono C es convexo y la definición de β para obtener $s_n t_n \geqslant \beta(v_1, v_3)$. Luego tomamos límite cuando $n \to +\infty$ para completar la prueba de $\beta(v_1, v_2)\beta(v_2, v_3) \geqslant \beta(v_1, v_3)$.

De las afirmaciones anteriores se obtiene

$$\frac{\beta(v_1, v_2)\beta(v_2, v_3)}{\alpha(v_1, v_2)\alpha(v_2, v_3)} \geqslant \frac{\beta(v_1, v_3)}{\alpha(v_1, v_3)}.$$

La parte (b) es ahora una consecuencia de propiedades del logaritmo y la definición de θ.

Pasemos a la demostración de la parte (c). Para la suficiencia notemos que la condición $\log\left[\dfrac{\beta(v_1, v_2)}{\alpha(v_1, v_2)}\right] = \theta(v_1, v_2) = 0$ implica que $\alpha(v_1, v_2) = \beta(v_1, v_2) = t$ para algún $t \in (0, +\infty)$. De modo que existen sucesiones $(t_n)_n$ convergiendo crecientemente a t y $(s_n)_n$ convergiendo decrecientemente a t, tales que $v_2 - t_n v_1 \in C$ para todo $n \geqslant 1$ y $s_n v_1 - v_2 \in C$ para todo $n \geqslant 1$. Usando la definición de clausura y (2.3) se obtiene $v_2 - tv_1 = 0$. Por tanto, existe $t > 0$ tal que $v_2 = tv_1$.

Para la necesidad supongamos $v_2 = tv_1$ para algún $t \in (0, +\infty)$. Entonces por (2.3), se tiene $v_2 - tv_1 \in \overline{C} \cap (-\overline{C})$. Aplicando la definición de clausura de conos, existen sucesiones convergentes $(t_n)_n$ crecientemente a t y $(s_n)_n$ decrecientemente a t tales que $v_2 - t_n v_1 \in C$ para todo $n \geqslant 1$ y $s_n v_1 - v_2 \in C$ para todo $n \geqslant 1$. Así, $\alpha(v_1, v_2) = t$ y $\beta(v_1, v_2) = t$ y por tanto $\theta(v_1, v_2) = \log\left[\dfrac{\beta(v_1, v_2)}{\alpha(v_1, v_2)}\right] = 0$, lo cual demuestra la parte (c) de la Proposición.

■

Llamaremos a $\theta(\cdot,\cdot)$ la *métrica proyectiva* asociada al cono convexo C. Note que la métrica proyectiva depende de una manera monótona con respecto al cono. En efecto, sean $C_1 \subset C_2$ dos conos convexos en E y $\alpha_i(\cdot,\cdot)$, $\beta_i(\cdot,\cdot)$, $\theta_i(\cdot,\cdot)$ los objetos correspondientes, $i = 1, 2$ como fueron definidos arriba. Claramente $\alpha_1(v_1, v_2) \leqslant \alpha_2(v_1, v_2)$ y $\beta_1(v_1, v_2) \geqslant \beta_2(v_1, v_2)$ porque

$$\alpha_1(v_1, v_2) = \sup\{t > 0 : v_2 - tv_1 \in C_1\} \leqslant \sup\{t > 0 : v_2 - tv_1 \in C_2\} = \alpha_2(v_1, v_2);$$

y también,

$$\beta_1(v_1, v_2) = \inf\{s > 0 : sv_1 - v_2 \in C_1\} \geqslant \inf\{s > 0 : sv_1 - v_2 \in C_2\} = \beta_2(v_1, v_2).$$

Con esto se tiene

$$\frac{\beta_1(v_1, v_2)}{\alpha_1(v_1, v_2)} \geqslant \frac{\beta_2(v_1, v_2)}{\alpha_2(v_1, v_2)},$$

y por tanto,

$$\theta_1(v_1, v_2) \geqslant \theta_2(v_1, v_2) \quad \text{para todo } v_1, v_2 \in C_1 \subset C_2.$$

Más generalmente, sean E_1, E_2 dos espacios vectoriales y $C_i \subset E_i$, $i = 1, 2$ conos convexos. Sea $L : E_1 \to E_2$ un operador lineal y suponga que $L(C_1) \subset C_2$. Entonces,

$$\alpha_1(v_1, v_2) = \sup\{t > 0 : v_2 - tv_1 \in C_1\} \leqslant \sup\{t > 0 : L(v_2 - tv_1) \in C_2\} = \alpha_2(L(v_1), L(v_2)).$$

Análogamente, $\beta_1(v_1, v_2) \geqslant \beta_2(L(v_1), L(v_2))$. Por tanto,

$$\theta_1(v_1, v_2) \geqslant \theta_2(L(v_1), L(v_2)) \quad \text{para todo } v_1, v_2 \in C_1.$$

Denotemos $\operatorname{diam}_\theta$ al diámetro de un conjunto con respecto a la métrica proyectiva θ. En general, L no es necesariamente una contracción (en el sentido estricto) con respecto a θ_1 y θ_2, pero el resultado siguiente afirma que éste es el caso si $L(C_1)$ tiene $\operatorname{diam}_{\theta_2}$ finito.

Proposición 6 *Sea* $D = \sup\{\theta_2(L(v_1), L(v_2)) : v_1, v_2 \in C_1\}$ *el diámetro del cono convexo* C_1. *Si* $D < \infty$ *entonces*

$$\theta_2(L(v_1), L(v_2)) \leqslant (1 - e^{-D})\theta_1(v_1, v_2) \quad \textit{para todo } v_1, v_2 \in C_1.$$

Demostración. Podemos suponer que $\alpha_1(v_1, v_2) > 0$ y $\beta_1(v_1, v_2) < +\infty$, pues de lo contrario no tenemos nada que demostrar porque $\theta_1(v_1, v_2) = +\infty$. Entonces existen sucesiones $(t_n)_n$ convergiendo crecientemente a $\alpha_1(v_1, v_2)$ y $(s_n)_n$ convergiendo decrecientemente a $\beta_1(v_1, v_2)$, tales que $v_2 - t_nv_1 \in C_1$ para todo $n \geqslant 1$ y $s_nv_1 - v_2 \in C_1$ para todo $n \geqslant 1$. Usando la definición de D se obtiene

$$\theta_2(L(v_2 - t_nv_1), L(s_nv_1 - v_2)) \leqslant D \quad \text{para todo } n \geqslant 1.$$

Esto implica, por definición de θ_2, que existen sucesiones $(T_n)_n$ y $(S_n)_n$ (de manera que $T_n = \alpha_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2))$ y $S_n = \beta_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2)))$ tales que $\lim_{n\to+\infty}\left[\log\left(\dfrac{S_n}{T_n}\right)\right] \leqslant D$ y

$$\begin{aligned} L(s_n v_1 - v_2) - T_n L(v_2 - t_n v_1) \in C_2 &\Leftrightarrow (s_n + T_n t_n)L(v_1) - (1 + T_n)L(v_2) \in C_2 \\ &\Rightarrow \beta_2(L(v_1), L(v_2)) \leqslant \frac{s_n + T_n t_n}{1 + T_n} \end{aligned}$$

$$\begin{aligned} S_n L(v_2 - t_n v_1) - L(s_n v_1 - v_2) \in C_2 &\Leftrightarrow (1 + S_n)L(v_2) - (s_n + t_n S_n)L(v_1) \in C_2 \\ &\Rightarrow \alpha_2(L(v_1), L(v_2)) \geqslant \frac{s_n + t_n S_n}{1 + S_n}. \end{aligned}$$

Entonces,

$$\begin{aligned} \theta_2(L(v_1), L(v_2)) &\leqslant \log\left[\left(\frac{s_n + T_n t_n}{1 + T_n}\right)\cdot\left(\frac{1 + S_n}{s_n + t_n S_n}\right)\right] \\ &= \log\left[\frac{s_n}{t_n} + T_n\right] - \log\left[1 + T_n\right] + \log\left[1 + S_n\right] - \log\left[\frac{s_n}{t_n} + S_n\right] \\ &= \int_0^{\log(s_n/t_n)}\left(\frac{e^x}{e^x + T_n} - \frac{e^x}{e^x + S_n}\right)dx \\ &\leqslant \log\left(\frac{s_n}{t_n}\right)\cdot\sup_{x>0}\frac{e^x(S_n - T_n)}{(e^x + T_n)(e^x + S_n)} \\ &= \log\left(\frac{s_n}{t_n}\right)\cdot\sup_{x>0}\left[\left(\frac{e^x}{e^x + T_n}\right)\cdot\left(\frac{1 - \frac{T_n}{S_n}}{\frac{e^x}{S_n} + 1}\right)\right] \\ &= \log\left(\frac{s_n}{t_n}\right)\cdot\left[1 - \frac{T_n}{S_n}\right]. \end{aligned}$$

Ahora, tomando límite cuando $n \to +\infty$ se obtiene

$$\begin{aligned} \theta_2(L(v_1), L(v_2)) &\leqslant \lim_{n\to+\infty}\left[\log\left(\frac{s_n}{t_n}\right)\cdot\left[1 - \frac{T_n}{S_n}\right]\right] \\ &= (1 - e^{-D})\theta_1(v_1, v_2). \end{aligned}$$

Esto completa la demostración. ■

En los siguientes ejemplos se muestra cómo se calcula la métrica proyectiva asociada a un cono para diferentes espacios.

Ejemplo 2.1. Sean $E = \mathbb{R}^2$ y $C = \{(x, y) : y > |x|\}$. El cociente proyectivo de C puede ser identificado naturalmente con el segmento $(-1, 1) \times \{1\}$, luego con el intervalo $(-1, 1)$.

Dados $-1 < x_1 \leqslant x_2 < 1$ y por las respectivas definiciones, se tiene

$$\alpha(x_1, x_2) = \sup\{t > 0 : x_2 - tx_1 < 1 - t\} = \frac{1 - x_2}{1 - x_1}.$$

Análogamente, $\beta(x_1, x_2) = \dfrac{x_2 + 1}{x_1 + 1}$. Luego,

$$\begin{aligned}\theta(x_1, x_2) &= \log\left[\frac{\beta(x_1, x_2)}{\alpha(x_1, x_2)}\right] = \log\left[\left(\frac{x_2 + 1}{x_1 + 1}\right) \cdot \left(\frac{1 - x_1}{1 - x_2}\right)\right] \\ &= \log R(-1, x_1, x_2, 1).\end{aligned}$$

Por tanto, $\theta(x_1, x_2) = \log R(-1, x_1, x_2, 1)$ donde R denota la razón transversal de cuatro puntos $a < b \leqslant c < d$ en la recta real, $R(a, b, c, d) = \dfrac{c - a}{b - a} \cdot \dfrac{d - b}{d - c}$.

En este sentido, las métricas proyectivas generalizan la métrica hiperbólica usual o métrica de Poincaré en el intervalo o el disco (ver [DB], pág 136).

Ejemplo 2.2. Sean X un espacio métrico compacto y $E = C^0(X, \mathbb{R})$ el espacio de las funciones continuas a valores reales definidas en X. Denotemos

$$C = C_+ = \{\varphi \in E : \varphi(x) > 0 \quad \text{para todo } x \in X\}.$$

Entonces, para cualesquiera $\varphi_1, \varphi_2 \in C$ tenemos

$$\begin{aligned}\alpha(\varphi_1, \varphi_2) &= \sup\{t > 0 : \varphi_2 - t\varphi_1 \in C\} \\ &= \sup\{t > 0 : (\varphi_2 - t\varphi_1)(x) > 0 \text{ para todo } x \in X\} \\ &= \text{ínf}\left\{\frac{\varphi_2(x)}{\varphi_1(x)} : x \in X\right\}.\end{aligned}$$

De igual forma se obtiene

$$\beta(\varphi_1, \varphi_2) = \sup\left\{\frac{\varphi_2(x)}{\varphi_1(x)} : x \in X\right\}.$$

Por tanto,

$$\theta(\varphi_1, \varphi_2) = \log\left[\frac{\sup(\varphi_2/\varphi_1)}{\text{ínf}(\varphi_2/\varphi_1)}\right] = \log\sup\left\{\frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} : x, y \in X\right\}.$$

Ejemplo 2.3. Sean X y E como antes y tomemos $C = C(a, \nu)$ el conjunto de funciones $\varphi \in E$ tales que $\varphi(x) > 0$ para todo $x \in X$ y $\log(\varphi)$ es (a, ν)-Hölder continuo. Esta condición significa lo siguiente

$$\exp(-ad(x, y)^\nu) \leqslant \frac{\varphi(x)}{\varphi(y)} \leqslant \exp(ad(x, y)^\nu) \quad \text{para todo } x, y \in X.$$

Observemos que esto equivale a la definición usual de condición de Hölder, porque la condición de (a,ν)-Hölder para $\log(\varphi)$, $|\log(\varphi(x)) - \log(\varphi(y))| \leqslant ad(x,y)^\nu$ implica

$$-ad(x,y)^\nu \leqslant \log\left(\frac{\varphi(x)}{\varphi(y)}\right) \leqslant ad(x,y)^\nu,$$

de lo cual se obtiene

$$\exp(-ad(x,y)^\nu) \leqslant \frac{\varphi(x)}{\varphi(y)} \leqslant \exp(ad(x,y)^\nu) \quad \text{para todo } x,y \in X.$$

Dados $\varphi_1, \varphi_2 \in C$ y $t_1, t_2 > 0$ se tiene

$$\exp(-ad(x,y)^\nu) \leqslant \frac{t_1\varphi_1(x) + t_2\varphi_2(x)}{t_1\varphi_1(y) + t_2\varphi_2(y)} \leqslant \exp(ad(x,y)^\nu) \quad \text{para todo } x,y \in X.$$

Usemos esta para mostrar que C es un cono convexo. En efecto, primero veamos que C es un cono. Sean $\varphi \in C$ y $t > 0$. De la definición de C se tiene $\varphi(x) > 0$ para todo $x \in X$. Entonces $t\varphi(x) > 0$ para todo $x \in X$ y

$$|\log(t\varphi(x)) - \log(t\varphi(y))| = \left|\log\left(\frac{\varphi(x)}{\varphi(y)}\right)\right| \leqslant ad(x,y)^\nu \quad \text{para todo } x,y \in X,$$

con lo cual C es un cono. Resta ver que C es convexo. Esto se verifica fácilmente mediante argumentos algebraicos. Por definición, $\alpha(\varphi_1,\varphi_2) = \sup\{t > 0 : \varphi_2 - t\varphi_1 \in C\}$. Estudiemos las condiciones que debe cumplir la función $\varphi_2 - t\varphi_1$ para pertenecer al cono C. Ella debe de ser positiva sobre X, es decir, $(\varphi_2 - t\varphi_1)(x) > 0$ para todo $x \in X$. Lo que es equivalente a $t < \frac{\varphi_2}{\varphi_1}(x)$ para todo $x \in X$. Además $\log(\varphi_2 - t\varphi_1)$ tiene que ser (a,ν)-Hölder continuo lo cual significa que

$$\exp(-ad(x,y)^\nu) \leqslant \frac{(\varphi_2 - t\varphi_1)(x)}{(\varphi_2 - t\varphi_1)(y)} \leqslant \exp(ad(x,y)^\nu) \quad \text{para todo } x,y \in X. \tag{2.4}$$

La desigualdad de la derecha en (2.4) implica

$$t(\exp(ad(x,y)^\nu)\varphi_1(y) - \varphi_1(x)) \leqslant \exp(ad(x,y)^\nu)\varphi_2(y) - \varphi_2(x).$$

Luego,

$$t \leqslant \frac{\exp(ad(x,y)^\nu)\varphi_2(y) - \varphi_2(x)}{\exp(ad(x,y)^\nu)\varphi_1(y) - \varphi_1(x)} \quad \text{para todo } x,y \in X.$$

Procedemos de igual manera con la desigualdad de la izquierda en (2.4) para obtener

$$t \leqslant \frac{\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y)} \quad \text{para todo } x,y \in X.$$

Con lo anterior se deduce que

$$\alpha(\varphi_1,\varphi_2) = \text{ínf}\left\{ \frac{\varphi_2}{\varphi_1}(x), \frac{\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y)} : x,y \in X, x \neq y \right\},$$

y $\beta(\varphi_1,\varphi_2)$ es dada por una expresión similar, precisamente con supremo en lugar de ínfimo.

El siguiente ejemplo sencillo muestra que, para conos convexos en general, la proyección del conjunto $\{\varphi : \theta(\varphi,1) < \infty\}$ no necesariamente es un espacio metrico completo.

Ejemplo 2.4. Sean X una variedad diferenciable compacta y $E = C^1(X,\mathbb{R})$ el espacio de Banach de funciones reales de clase C^1 sobre X dotado con la topologia C^1. Considere el cono convexo $C = \{\varphi \in E : \varphi(x) > 0 \text{ para todo } x \in X\}$. Sea φ una función estrictamente positiva sobre X la cual es continua pero no diferenciable y sea $(\varphi_n)_n$ una sucesión de funciones estrictamente positivas de C^1 convergiendo uniformemente a φ. Claramente la métrica proyectiva θ del cono C está dada por la siguiente expresión como en el ejemplo 2.2

$$\theta(\varphi_1,\varphi_2) = \log\left[\frac{\sup(\varphi_2/\varphi_1)}{\text{ínf}(\varphi_2/\varphi_1)}\right] = \log\sup\left\{\frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} : x,y \in X\right\}.$$

Usando esta observación, veamos que $(\varphi_n)_n$ es una sucesión θ-Cauchy en el cono C. En efecto, la convergencia uniforme de $(\varphi_n)_n$ implica que $(\varphi_n)_n$ es una sucesión de Cauchy en E. Dado $\varepsilon > 0$ existe un número natural N_0 tal que si $k,l \geqslant N_0$ entonces $\varphi_l(x)$ está próximo de $\varphi_k(x)$ y $\varphi_k(y)$ está próximo de $\varphi_l(y)$. De esto se tiene

$$e^{-\varepsilon} < \sup\frac{\varphi_k(x)\varphi_l(y)}{\varphi_l(x)\varphi_k(y)} < e^{\varepsilon}.$$

Luego, por la definición de la métrica θ obtenemos que $\theta(\varphi_k,\varphi_l) < \varepsilon$. Con esto hemos mostrado que $(\varphi_n)_n$ es una sucesión θ-Cauchy. Por otro parte, esta sucesión no puede ser θ-convergente en C, ya que el θ-límite es un límite uniforme y éste tendría que coincidir con φ, lo que es absurdo por que φ no pertenece a C.

Con lo anterior se tiene un ejemplo de un cono que no es completo con respecto de la métrica proyectiva.

2.3. Conos Invariantes

Retornando al contexto de la sección 2.1. Comenzamos está sección considerando el cono convexo $C(a,\nu)$ estudiado en el ejemplo 2.3 de la sección anterior, pero para funciones continuas definidas sobre la variedad M. Llamaremos una *ρ-vecindad* a la clausura de un conjunto abierto con diámetro ρ. Si W es una ρ-vecindad entonces $d(x,y)\leqslant\rho$ para todo $x,y\in W$. Fijemos $\rho_0>0$ como la condición (ii) en la página 10 y para constantes positivas a, $\nu<\nu_0$. Definimos $C(a,\nu)$ el cono convexo de funciones $\varphi\in E=C^0(M,\mathbb{R})$ satisfaciendo

1) $\varphi(x)>0$ para todo $x\in M$.

2) $\log(\varphi)$ es (a,ν)-Hölder continuo sobre ρ-vecindades, esto es,

$$\exp(-ad(y_1,y_2)^\nu)\leqslant\frac{\varphi(y_1)}{\varphi(y_2)}\leqslant\exp(ad(y_1,y_2)^\nu)\quad\text{para todo } y_1,y_2 \text{ con } d(y_1,y_2)\leqslant\rho_0.$$

El objetivo principal en está sección es mostrar que el operador de Perron-Frobenius (o transfer) $\mathcal{L}$ es una contracción con respecto de la métrica proyectiva $\theta=\theta_{a,\nu}$ del cono convexo $C(a,\nu)$. El siguiente resultado muestra que el cono $C(a,\nu)$ es invariante por el operador $\mathcal{L}$.

Proposición 7 (Invarianza) *Existe una constante positiva $\lambda_1<1$ tal que $\mathcal{L}(C(a,\nu))\subset C(\lambda_1 a,\nu)$ para todo $a>0$ suficientemente grande.*

Demostración. Si $\varphi>0$, entonces $\mathcal{L}(\varphi)>0$ y así sólo tenemos que tratar con la condición (2) de la definición del cono $C(a,\nu)$. Sean $y_1,y_2\in M$ con $d(y_1,y_2)\leqslant\rho_0$ y escribamos $f^{-1}(y_j)=\{x_1^j,\dots,x_k^j\}$, $j=1,2$, como la condición (ii) en la página 10. Afirmamos que $\log|\det Df|$ es (a_0,ν_0)-Hölder continua para algún $a_0>0$. Demostremos esta afirmación. Como f es $C^{1+\nu_0}$ se tiene que Df es (b_1,ν_0)-Hölder continua para algún $b_1>0$ y las funciones determinante y logaritmo son Lipschitzianas, con constantes de Lipschitz l_1 y l_2 respectivamente. Haciendo la compuesta de las funciones Df, det y log se obtiene

$$\begin{aligned}|\log|\det Df(x)|-\log|\det Df(y)||&\leqslant l_2|\det Df(x)-\det Df(y)|\\&\leqslant l_2l_1\|Df(x)-Df(y)\|\\&\leqslant l_2l_1b_1d(x,y)^{\nu_0}.\end{aligned}$$

Tomando $a_0=l_1l_2b_1>0$ queda demostrada la afirmación. Por otro lado, fijemos arbitrariamente λ_1 con $\sigma^{-\nu}<\lambda_1<1$ y escojemos $a>0$ suficientemente grande de modo que $\sigma^{-\nu}+\frac{a_0}{a}\leqslant\lambda_1$, lo que implica que $a\sigma^{-\nu}+a_0\leqslant\lambda_1 a$. Por (ii) tenemos que

$d(x_i^1, x_i^2)^{\nu_0} \leqslant d(y_1, y_2)^{\nu_0} \leqslant d(y_1, y_2)^{\nu}$. Usando lo anterior se obtiene, para toda $\varphi \in C(a, \nu)$,

$$\begin{aligned}
(\mathcal{L}\varphi)(y_1) &= \sum_{i=1}^{k} \varphi(x_i^1) |\det Df(x_i^1)|^{-1} \\
&\leqslant \sum_{i=1}^{k} \varphi(x_i^2) \exp(ad(x_i^1, x_i^2)^{\nu}) |\det Df(x_i^2)|^{-1} \exp(a_0 d(x_i^1, x_i^2)^{\nu_0}) \\
&\leqslant \sum_{i=1}^{k} \varphi(x_i^2) \exp(a\sigma^{-\nu} d(y_1, y_2)^{\nu}) |\det Df(x_i^2)|^{-1} \exp(a_0 d(y_1, y_2)^{\nu}) \\
&= \exp([a\sigma^{-\nu} + a_0] d(y_1, y_2)^{\nu}) \sum_{i=1}^{k} \varphi(x_i^2) |\det Df(x_i^2)|^{-1} \\
&\leqslant \exp(\lambda_1 a d(y_1, y_2)^{\nu}) (\mathcal{L}\varphi)(y_2).
\end{aligned}$$

Así, $\mathcal{L}(\varphi) > 0$ y $\log(\mathcal{L}\varphi)$ es $(\lambda_1 a, \nu)$-Hölder continuo. De modo que $\mathcal{L}\varphi \in C(\lambda_1 a, \nu)$. ■

Denotemos $\theta = \theta_{a,\nu}$ a la métrica proyectiva asociada al cono convexo $C(a, \nu)$ y recordemos algunos hechos que fueron establecidos previamente. Entonces por el ejemplo 2.3 tenemos $\theta(\varphi_1, \varphi_2) = \log \left[\dfrac{\beta(\varphi_1, \varphi_2)}{\alpha(\varphi_1, \varphi_2)} \right]$, donde $\alpha(\varphi_1, \varphi_2)$ está dado por

$$\inf \left\{ \frac{\varphi_2}{\varphi_1}(x), \frac{\exp(ad(x,y)^{\nu})\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^{\nu})\varphi_1(x) - \varphi_1(y)} :\ x, y \in X,\ x \neq y,\ d(x,y) \leqslant \rho_0 \right\}, \tag{2.5}$$

mientras que $\beta(\varphi_1, \varphi_2)$ es

$$\sup \left\{ \frac{\varphi_2}{\varphi_1}(x), \frac{\exp(ad(x,y)^{\nu})\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^{\nu})\varphi_1(x) - \varphi_1(y)} :\ x, y \in X,\ x \neq y,\ d(x,y) \leqslant \rho_0 \right\}.$$

También consideremos el cono

$$C_+ = \{\varphi \in E : \varphi(x) > 0 \text{ para todo } x \in M\}.$$

La métrica proyectiva θ_+ asociada a C_+ fue calculada en el ejemplo 2.2 y está dada por $\theta_+(\varphi_1, \varphi_2) = \log \left[\dfrac{\beta_+(\varphi_1, \varphi_2)}{\alpha_+(\varphi_1, \varphi_2)} \right]$ donde

$$\alpha_+(\varphi_1, \varphi_2) = \inf \left\{ \frac{\varphi_2(x)}{\varphi_1(x)} : x \in M \right\} \quad \text{y} \quad \beta_+(\varphi_1, \varphi_2) = \sup \left\{ \frac{\varphi_2(x)}{\varphi_1(x)} : x \in M \right\}.$$

Proposición 8 (Diámetro Finito) *El diámetro* $D_1 = \sup\{\theta(\varphi_1, \varphi_2) : \varphi_1, \varphi_2 \in C(\lambda_1 a, \nu)\}$ *del cono* $C(\lambda_1 a, \nu)$ *es finito, dadas las constantes positivas* a, ν *y* $\lambda_1 < 1$.

Demostración. Denotemos diam_θ, diam_{θ_+} a los diámetros según las métricas θ y θ_+ respectivamente. La demostración tiene dos pasos: primero demostremos que

$$\mathrm{diam}_\theta(C(\lambda_1 a,\nu)) \leqslant \mathrm{diam}_{\theta_+}(C(\lambda_1 a,\nu)) + K'(\lambda_1),$$

entonces se obtiene $\mathrm{diam}_{\theta_+}(C(a,\nu)) \leqslant K''(a)$, donde $K'(\cdot), K''(\cdot) < +\infty$. Dadas cualesquiera $\varphi_1, \varphi_2 \in C(\lambda_1 a,\nu)$, por la definición del cono $C(\lambda_1 a,\nu)$, se tiene

- $\varphi_1(x) > 0$ para todo $x \in M$.
- $\log(\varphi_1)$ es $(\lambda_1 a,\nu)$-Hölder continuo sobre ρ_0-vecindades, esto es,

$$\exp(-a\lambda_1 d(y_1,y_2)^\nu) \leqslant \frac{\varphi_1(y_1)}{\varphi_1(y_2)} \leqslant \exp(a\lambda_1 d(y_1,y_2)^\nu), \tag{2.6}$$

para todo y_1, y_2 con $d(y_1,y_2) \leqslant \rho_0$.

Sean x, y en una ρ_0-vecindad, es decir, $d(x,y) \leqslant \rho_0$. De la desigualdad de la izquierda en (2.6) se obtiene $-\exp(-a\lambda_1 d(x,y)^\nu)\varphi_1(x) \geqslant -\varphi_1(y)$. Entonces

$$[\exp(ad(x,y)^\nu) - \exp(-a\lambda_1 d(x,y)^\nu)]\varphi_1(x) \geqslant \exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y).$$

Por otro lado, en (2.6) para la función $\log(\varphi_2)$ tomamos la desigualdad de la derecha, para obtener $-\varphi_2(y) \geqslant -\exp(a\lambda_1 d(x,y)^\nu)\varphi_2(x)$. Luego,

$$\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y) \geqslant [\exp(ad(x,y)^\nu) - \exp(a\lambda_1 d(x,y)^\nu)]\varphi_2(x).$$

Con todo esto, estimamos

$$\begin{aligned}
\frac{\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y)} &\geqslant \frac{[\exp(ad(x,y)^\nu) - \exp(a\lambda_1 d(x,y)^\nu)]\varphi_2(x)}{[\exp(ad(x,y)^\nu) - \exp(-a\lambda_1 d(x,y)^\nu)]\varphi_1(x)} \\
&\geqslant K_1 \frac{\varphi_2(x)}{\varphi_1(x)} \\
&\geqslant K_1 \inf\left\{\frac{\varphi_2(x)}{\varphi_1(x)} : x \in M\right\} \\
&= K_1 \alpha_+(\varphi_1,\varphi_2),
\end{aligned}$$

donde

$$K_1 = \inf\left\{\frac{z - z^{\lambda_1}}{z - z^{-\lambda_1}} : z > 1\right\}.$$

Note que $K_1 \in (0,1)$, porque

$$\lim_{z\to+\infty} \frac{z - z^{\lambda_1}}{z - z^{-\lambda_1}} = 1 \quad \text{y} \quad \lim_{z\to 1^+} \frac{z - z^{\lambda_1}}{z - z^{-\lambda_1}} = \frac{1-\lambda_1}{1+\lambda_1} < 1.$$

Luego usando (2.5) se sigue que, $\alpha(\varphi_1,\varphi_2) \geqslant K_1\alpha_+(\varphi_1,\varphi_2)$. Para β aplicamos un procedimiento análogo al desarrollado para α y obtenemos

$$\begin{aligned}\frac{\exp(ad(x,y)^\nu)\varphi_2(x)-\varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x)-\varphi_1(y)} &\leqslant K_2\frac{\varphi_2(x)}{\varphi_1(x)}\\ &\leqslant K_2\sup\left\{\frac{\varphi_2(x)}{\varphi_1(x)} : x\in M\right\}\\ &= K_2\beta_+(\varphi_1,\varphi_2),\end{aligned}$$

donde

$$K_2=\sup\left\{\frac{z-z^{-\lambda_1}}{z-z^{\lambda_1}} : z>1\right\}\in(1,+\infty).$$

De lo anterior y (2.5) tenemos $K_2\beta_+(\varphi_1,\varphi_2) \geqslant \beta(\varphi_1,\varphi_2)$. Con lo ya demostrado

$$\frac{K_2\beta_+(\varphi_1,\varphi_2)}{K_1\alpha_+(\varphi_1,\varphi_2)} \geqslant \frac{\beta(\varphi_1,\varphi_2)}{\alpha(\varphi_1,\varphi_2)}.$$

Entonces $\theta(\varphi_1,\varphi_2) \leqslant \theta_+(\varphi_1,\varphi_2)+\log\left[\frac{K_2}{K_1}\right]$, esto concluye el primer paso de la demostración.

Para el segundo paso de la demostración, observe que

$$\theta_+(\varphi_1,\varphi_2)=\log\left[\frac{\beta_+(\varphi_1,\varphi_2)}{\alpha_+(\varphi_1,\varphi_2)}\right]=\log\sup\left\{\frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} : x,y\in X\right\}. \tag{2.7}$$

Afirmamos que si $\varphi\in C(a,\nu)$ entonces $\log(\varphi)$ es (b,ν)-Hölder continuo sobre toda la variedad M (no precisamente sobre ρ_0-vecindades), para algún $b>0$. Demostremos está afirmación. Por conexidad de M cualquier par de puntos x y y pueden ser unidos por medio de una curva. Empleando la compacidad de M, existe $N\geqslant 1$ tal que dados $x,y\in M$ podemos hallar puntos z_i con $d(z_{i-1},z_i)\leqslant\rho_0$ para cada $i=1,\dots,N$ y $x=z_0,z_1,\dots,z_N=y$ se tiene que

$$\sum_{i=1}^{N} d(z_{i-1},z_i)\leqslant 2d(x,y).$$

Para demostrar está desigualdad se procede por inducción. Para $k=1$, tomamos $x=z_0$, $z_1=y$ usando la conexidad de M existe una geodésica $\gamma:[0,1]\to M$ tal que $\gamma(0)=z_0$ y $\gamma(1)=z_1$. Así $d(z_0,z_1)\leqslant 2d(z_0,z_1)=2d(x,y)$. Para $k=2$, tenemos $z_0=x$, $z_2=y$ con $d(z_0,z_2)\leqslant\rho_0$. Por ser M conexa podemos unir estos dos puntos por medio de una geodésica $\gamma:[0,1]\to M$ con $\gamma(0)=z_0$ y $\gamma(1)=z_2$. Consideremos la bola B_0 centrada en z_0 y radio $\frac{\rho_0}{2}$ para encontrar un punto z_1' que pertenece a la intersección de ∂B_0 con γ. Ahora tomemos la bola B_1 centrada en z_1' y radio $\frac{\rho_0}{2}$ conteniendo a z_2 (porque $d(z_0,z_2)\leqslant\rho_0$).

Además denotemos $z_0' = z_0$ y $z_2' = z_2$. Sea z_1 cualquier punto en $B_0 \cap B_1$ ($z_1 \neq z_1'$) tal que cumpla con la siguientes desigualdades $d(z_1', z_1) \leqslant d(z_0, z_1')$ y $d(z_1, z_1') \leqslant d(z_1', z_2)$ (porque de lo contrario estaremos en el caso anterior). Ahora,

$$\begin{aligned}
\sum_{i=1}^{2} d(z_{i-1}, z_i) &\leqslant d(z_0, z_1') + d(z_1', z_1) + d(z_1, z_1') + d(z_1', z_2) \\
&\leqslant d(z_0, z_1') + d(z_0, z_1') + d(z_1', z_2) + d(z_1', z_2) \\
&= 2\left[\sum_{i=1}^{2} d(z_{i-1}', z_i')\right] = 2d(x, y).
\end{aligned}$$

Se concluye que,

$$\sum_{i=1}^{2} d(z_{i-1}, z_i) \leqslant 2d(x, y).$$

Hipotesis inductiva:

$$\sum_{i=1}^{N-1} d(z_{i-1}, z_i) \leqslant 2d(x, y)$$

donde $x = z_0, \ldots, z_{N-1} = y$ con $d(z_{i-1}, z_i) \leqslant \rho_0$. Demostremos para $k = N$ tomamos $x = z_0$, $z_N = y$. Nuevamente por la conexidad de M existe una geodésica $\gamma : [0, 1] \to M$ tal que $\gamma(0) = z_0$, $\gamma(1) = z_N$. Esto implica que existen $z_1' \in \partial B_0 \cap \gamma$ y $z_{i+1}' \in \gamma \cap (\partial B_i - B_{i-1})$ donde $B_i = B(z_i', \frac{\rho_0}{2})$ para todo $i = 0, \ldots, N-2$ además B_{N-1} contiene a z_N y denotemos $z_0 = z_0'$, $z_N = z_N'$. Con esto tenemos $N - 1$ bolas cubriendo a γ. Sean $z_i \in B_{i-1} \cap B_i$ tal que $d(z_i, z_i') \leqslant d(z_i', z_{i+1}')$ para todo $i = 1, \ldots, N-1$ (de lo contrario estaremos en el caso $k = N - 1$). Notamos que:

$$\begin{aligned}
\sum_{i=1}^{N} d(z_{i-1}, z_i) &= \sum_{i=1}^{N-1} d(z_{i-1}, z_i) + d(z_{N-1}, z_N) \\
&\leqslant \sum_{i=1}^{N-1} d(z_{i-1}, z_i) + d(z_{N-1}, z_{N-1}') + d(z_{N-1}', z_N) \\
&\leqslant \sum_{i=1}^{N-1} d(z_{i-1}', z_i') + 2d(z_{N-1}', z_N') \\
&= 2\sum_{i=1}^{N} d(z_{i-1}', z_i') = 2d(x, y)
\end{aligned}$$

Con este hecho demostramos la afirmación. Aplicamos el resultado anterior en la siguiente estimación

$$\frac{\varphi(y)}{\varphi(x)} = \prod_{i=1}^{N} \frac{\varphi(z_i)}{\varphi(z_{i-1})} \leqslant \exp\left(\sum_{i=1}^{N} ad(z_{i-1}, z_i)^{\nu}\right)$$
$$\leqslant \exp\left(Na\left(\sum_{i=1}^{N} d(z_{i-1}, z_i)\right)^{\nu}\right) \leqslant \exp(Na2^{\nu} d(x,y)^{\nu}))$$

y podemos tomar $b = Na2^{\nu}$. Luego, $\log\varphi$ es (b,ν)-Hölder continuo. Se obtiene, dados $\varphi_1, \varphi_2 \in C(a,\nu)$ vale lo siguiente

$$\log\left[\frac{\varphi_1(x)}{\varphi_1(y)}\right] \leqslant bd(x,y)^{\nu} \leqslant b \quad \text{y} \quad \log\left[\frac{\varphi_2(y)}{\varphi_2(x)}\right] \leqslant b$$

para todo $x, y \in M$. De lo cual obtenemos

$$\sup\left\{\frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} : x, y \in M\right\} \leqslant e^{2b}$$

y con esto concluimos, $\theta_+(\varphi_1, \varphi_2) \leqslant 2b$. Así, $\operatorname{diam}_{\theta_+}(C(a,\nu)) \leqslant K''(a)$ donde $K''(a) = 2^{1+\nu}Na$, esto termina la demostración de la Proposición. ∎

Por la Proposición 7 en la página 19 sabemos que $\mathcal{L}(C(a,\nu)) \subset C(\lambda_1 a, \nu)$. Luego, los diámetros de estos conos con respecto de la métrica proyectiva $\theta = \theta_{a,\nu}$ cumplen $\operatorname{diam}_\theta(\mathcal{L}(C(a,\nu))) \leqslant \operatorname{diam}_\theta(C(\lambda_1 a, \nu))$. En la Proposición 8 de la página 20 demostramos que D_1 es finito, a lo cual aplicamos la Propoción 6 de la página 14 para obtener

$$\theta(\mathcal{L}(\varphi_1), \mathcal{L}(\varphi_2)) \leqslant (1 - e^{-D_1})\theta(\varphi_1, \varphi_2) = \Lambda_1 \theta(\varphi_1, \varphi_2)$$

donde $\Lambda_1 = 1 - e^{-D_1}$. Por tanto, el operador $\mathcal{L} : C(a,\nu) \to C(a,\nu)$ es una Λ_1-contracción con respecto de la métrica proyectiva $\theta = \theta_{a,\nu}$.

2.4. Medidas Invariantes Absolutamente Continuas

Para demostrar que $\mathcal{L}$ tiene un único punto fijo nos apoyaremos en el hecho que el cono C_+ es completo para la correspondiente métrica proyectiva. Más precisamente, tenemos la siguiente Proposición.

PROPOSICIÓN 9 (Completitud) *Cualquier sucesión θ_+-Cauchy $(\varphi_n)_n$ en C_+ es θ_+-convergente en C_+. Además si la sucesión está normalizada, en el sentido que $\int \varphi_n dm = 1$ para todo $n \geqslant 1$, entonces $(\varphi_n)_n$ también es convergente uniformemente.*

Demostración. Sea $(\varphi_n)_n$ una sucesión θ_+-Cauchy, normalizada por $\int \varphi_n dm = 1$ para todo $n \geqslant 1$. En particular, $(\varphi_n)_n$ es θ_+-acotada y, recordando la definición de θ_+ de (2.7), existe $R_1 > 0$ tal que

$$\frac{1}{R_1} \leqslant \frac{\varphi_n(x)\varphi_1(y)}{\varphi_n(y)\varphi_1(x)} \leqslant R_1 \quad \text{para todo } x, y \in M \text{ y todo } n \geqslant 1.$$

En particular,

$$\frac{1}{R_2} \leqslant \frac{\varphi_n(x)}{\varphi_n(y)} \leqslant R_2$$

para todo $x, y \in M$ y todo $n \geqslant 1$ donde $R_2 = R_1 \sup \left\{ \frac{\varphi_1(s)}{\varphi_1(t)} : s, t \in M \right\}$. Por otro lado, la normalización $\int \varphi_n dm = 1$ implica que $\inf \varphi_n \leqslant 1 \leqslant \sup \varphi_n$ y con esto tenemos que

$$\frac{1}{R_2} \leqslant \varphi_n(x) \leqslant R_2 \text{ para todo } x \in M \text{ y } n \geqslant 1.$$

Ahora, la condición de Cauchy establece que dado cualquier $\varepsilon > 0$ existe $N \geqslant 1$ tal que $\dfrac{\sup(\varphi_k/\varphi_l)}{\inf(\varphi_k/\varphi_l)} \leqslant e^{\varepsilon}$ para todo $k, l \geqslant N$. Como la sucesión $(\varphi_n)_n$ está normalizada $\int \varphi_k dm = \int \varphi_l dm = 1$ implica $\inf \varphi_k \leqslant 1 \leqslant \sup \varphi_k$ y $\inf \varphi_l \leqslant 1 \leqslant \sup \varphi_l$ con lo cual se obtiene $\inf(\varphi_k/\varphi_l) \leqslant 1 \leqslant \sup(\varphi_k/\varphi_l)$. Luego, por la condición de Cauchy tenemos

$$e^{-\varepsilon} \leqslant \inf\left(\frac{\varphi_k}{\varphi_l}\right) \leqslant 1 \leqslant \sup\left(\frac{\varphi_k}{\varphi_l}\right) \leqslant e^{\varepsilon}$$

lo que nos dice que $e^{-\varepsilon} \leqslant \dfrac{\varphi_k}{\varphi_l} \leqslant e^{\varepsilon}$ con lo cual se obtiene

$$\sup|\varphi_k - \varphi_l| \leqslant \sup|\varphi_l| \cdot \sup\left|\frac{\varphi_k}{\varphi_l} - 1\right| \leqslant R_2(e^{\varepsilon} - 1).$$

Por tanto, $(\varphi_n)_n$ es una sucesión de Cauchy con respecto de la métrica uniforme y como consecuencia, $(\varphi_n)_n$ es convergente uniformemente. Sea φ_0 el límite uniforme de $(\varphi_n)_n$. Observe que $\varphi_0 \geqslant R_2^{-1}$ de modo que $\varphi_0 \in C_+$. Pasando el límite cuando $l \to +\infty$ se obtiene

$$e^{-\varepsilon} \leqslant \inf\left(\frac{\varphi_k}{\varphi_0}\right) \leqslant 1 \leqslant \sup\left(\frac{\varphi_k}{\varphi_0}\right) \leqslant e^{\varepsilon},$$

para todo $k \geqslant N$. Esto demuestra que tanto $\sup(\varphi_k/\varphi_0)$ como $\inf(\varphi_k/\varphi_0)$ ambos convergen a 1 y, por definición de la métrica θ_+, se tiene $\theta_+(\varphi_n, \varphi_0)$ converge a cero cuando $n \to +\infty$. En efecto, de lo anteriormente demostrado tenemos $\dfrac{\sup(\varphi_k/\varphi_0)}{\inf(\varphi_k/\varphi_0)} \leqslant e^{2\varepsilon}$ y aplicando logaritmo a ambos lados de la desigualdad se tiene $\theta_+(\varphi_k, \varphi_0) \leqslant 2\varepsilon$ para todo $k \geqslant N$. ■

Observación 2.1. La forma de la condición de normalización en la proposición anterior es algo arbitraria que puede ser reemplazada por $\sup(\varphi_n) = 1$. Esto se puede hacer en el sentido más general del contexto de los ejemplos 2.2 y 2.3.

Aplicando el resultado previo a la sucesión $\varphi_n = \mathcal{L}^n 1$. Como $\mathcal{L}$ es una θ-contracción, $(\varphi_n)_n$ es una sucesión θ-Cauchy y es también θ_+-Cauchy: recordemos que $\theta_+ \leqslant \theta$ porque $C(a, \nu) \subset C_+$. Por otro lado,

$$\int \varphi_n dm = \int (\mathcal{L}^n 1) dm = 1 \quad \text{para todo } n \geqslant 1.$$

Luego la sucesión φ_n converge uniformemente a alguna $\varphi_0 \in C_+$. De hecho $\varphi_0 \in C(\lambda_1 a, \nu)$ porque la condición de Hölder en la definición de este cono es cerrada bajo límite uniforme. Demostremos esto. Claramente $\varphi_n \in C(\lambda_1 a, \nu)$ para todo $n \geqslant 1$. De la condición de Hölder para esta sucesión se tiene

$$\exp\left(-a\lambda_1 d(x,y)^\nu\right) \leqslant \frac{\varphi_n(x)}{\varphi_n(y)} \leqslant \exp\left(a\lambda_1 d(x,y)^\nu\right) \quad \text{con } d(x,y) \leqslant \rho_0 \text{ y todo } n \geqslant 1.$$

Ahora tomando límite cuando $n \to +\infty$ tenemos

$$\exp\left(-a\lambda_1 d(x,y)^\nu\right) \leqslant \frac{\varphi_0(x)}{\varphi_0(y)} \leqslant \exp\left(a\lambda_1 d(x,y)^\nu\right), \text{ con } d(x,y) \leqslant \rho_0$$

y así $\varphi_0 \in C(a\lambda_1, \nu)$. Puesto que $\mathcal{L}$ es un operador acotado con respecto de la norma $\|\cdot\|_0$ de convergencia uniforme (observe que $\|\mathcal{L}\varphi - \mathcal{L}\psi\|_0 \leqslant k \sup |\det Df|^{-1} \|\varphi - \psi\|_0$ donde $k = \# f^{-1}(y)$ para cualquier $y \in M$) se tiene que φ_0 es punto fijo de $\mathcal{L}$. Como una consecuencia, $\mu_0 = \varphi_0 m$ es una medida de probabilidad invariante por f:

$$\int \varphi d\mu_0 = \int \varphi \varphi_0 dm = \int \varphi (\mathcal{L}\varphi_0) dm = \int (\varphi \circ f) \varphi_0 dm = \int (\varphi \circ f) d\mu_0$$

para toda $\varphi \in C^0(M, \mathbb{R})$. Además, $d\mu_0/dm = \varphi_0 \geqslant R_2^{-1} > 0$ y, por tanto, μ_0 es equivalente a la medida de Lebesgue m.

CAPÍTULO 3

Propiedades Estadísticas

En el presente capítulo nos concentraremos en estudiar algunas de las propiedades estadísticas como son ergodicidad, exactitud y exponencialmente mixing. Además la relación entre estas propiedades.

3.1. Mixing Exponencial

Iniciaremos esta sección estableciendo la definición de mixing, ergodicidad y recordemos la definición de exponencialmente mixing. Considere una transformación que preserva medida T de un espacio de probabilidad $(X, \mathcal{B}, \mu)$. Recuerde que un conjunto $A \in \mathcal{B}$ es T-*invariante* si $T^{-1}A = A$

Definición 1 Diremos que T es *ergódica* si cada conjunto T-invariante tiene medida igual 0 o 1.

Definición 2 T se denomina *mixing* si para cualquier par de conjuntos $A, B \in \mathcal{B}$

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

Observación 3.2. Note que toda transformación mixing es ergódica, pero el reciproco no se cumple (ver [RM], pág 181).

Recordemos la definición de exponencialmente mixing. Considere las funciones de correlacion

$$C_n(\varphi, \psi) = \int (\varphi \circ T^n)\psi d\nu - \int \varphi d\mu \int \psi d\mu \quad \text{para toda } \varphi, \psi \in L^1(\mu).$$

Decimos que (T, μ) es *exponencialmente mixing* o tiene *decaimiento exponencial de correlación*si existe $\tau < 1$ y para cada $\varphi, \psi \in L^1(\mu)$ existe una constante $c > 0$ tal que

$$|C_n(\varphi, \psi)| \leqslant c\tau^n \quad \text{para todo } n \geqslant 1.$$

Observe que exponencialmente mixing implica mixing.

Sea $\varphi \in C(a\lambda_1, \nu)$. El argumento de la sección anterior nos da

$$\theta_+(\mathcal{L}^n\varphi, \varphi_0) \leqslant \theta(\mathcal{L}^n\varphi, \varphi_0) \leqslant \theta(\varphi, \varphi_0)\Lambda_1^n \leqslant D_1\Lambda_1^n$$

donde la primera desigualdad se obtiene de $\theta_+ \leqslant \theta$, la segunda se obtiene por inducción usando que $\mathcal{L}$ es una Λ_1-contracción, φ_0 es punto fijo; y la última resulta de la definición de D_1. Procediendo como en la demostración de la Proposición 9, y como $\Lambda_1 \leqslant 1$, se tiene $D_1\Lambda_1^n \leqslant D_1\Lambda_1$. Aplicando el hecho que $\exp(t) - 1 \leqslant 2t$ para todo $0 < t \leqslant D_1\Lambda_1$ tenemos que

$$\sup|\mathcal{L}^n\varphi - \varphi_0| \leqslant R_2(e^{D_1\Lambda_1^n} - 1) \leqslant R_3\Lambda_1^n$$

para $R_3 = 2R_2D_1 > 0$ y todo $n \geqslant 1$. Está puede ser vista como una afirmación de pérdida de memoria con velocidad exponencial en el sistema: los iterados $(\mathcal{L}^n\varphi)m$ de una distribución de masa inicial φm convergen exponencialmente rápido a la distribución de equilibrio $\varphi_0 m$. El siguiente resultado nos proporciona una estimación la cual usaremos luego para demostrar que (f, μ_0) es exponencialmente mixing.

Proposición 10 *Dadas una función ν-Hölder φ y una función m-integrable ψ sobre M, existe $K_0 = K_0(\varphi, \psi) > 0$ tal que*

$$\left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0 \int \varphi dm\right| \leqslant K_0\Lambda_1^n \quad \text{para todo } n \geqslant 0.$$

Demostración. Suponga primero que $\varphi \in C(a\lambda_1, \nu)$. Podemos suponer además que $\int \varphi dm = 1$. Entonces, denotando $\|\psi\|_1 = \int |\psi| d\mu_0$, con las estimaciones hechas al principio de esta sección se tiene

$$\begin{aligned}
\left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0 \int \varphi dm\right| &= \left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0\right| \\
&= \left|\int (\mathcal{L}^n\varphi)\psi dm - \int \psi d\mu_0\right| = \left|\int \psi \left(\frac{\mathcal{L}^n\varphi}{\varphi_0} - 1\right) d\mu_0\right| \\
&\leqslant \left\|\frac{\mathcal{L}^n\varphi}{\varphi_0} - 1\right\|_0 \cdot \|\psi\|_1 \leqslant (e^{D_1\Lambda_1^n} - 1)\|\psi\|_1 \\
&\leqslant R_3\|\psi\|_1\Lambda_1^n,
\end{aligned}$$

donde la igualdad $\int \varphi(\psi \circ f^n) dm = \int (\mathcal{L}^n\varphi)\psi dm$ es obtenida de aplicar inducción en la ecuación (2.2). Ahora sean φ una función ν-Hölder cualquiera y $A > 0$ tal que φ es (A, ν)-Hölder. Para $B > 0$, se escribe

$$\varphi = \varphi_B^+ - \varphi_B^- \quad \text{donde } \varphi_B^\pm = \frac{1}{2}(|\varphi| \pm \varphi) + B.$$

De está suposición veamos que $\varphi_B^{\pm}$ es (A,ν)-Hölder continua. Demostremos primero para φ_B^{+}, si $d(x,y) \leqslant \rho_0$ entonces

$$\begin{aligned}
|\varphi_B^+(x) - \varphi_B^+(y)| &= \left| \left(\frac{1}{2}(|\varphi| + \varphi)(x) + B \right) - \left(\frac{1}{2}(|\varphi| + \varphi)(y) + B \right) \right| \\
&\leqslant \frac{1}{2}||\varphi|(x) + |\varphi|(y)| + \frac{1}{2}|\varphi(x) - \varphi(y)| \\
&\leqslant \frac{1}{2}Ad(x,y)^\nu + \frac{1}{2}Ad(x,y)^\nu = Ad(x,y)^\nu.
\end{aligned}$$

Se demuestra una desigualdad similar para φ_B^-. Fácilmente se ve que $\varphi_B^{\pm} \geqslant B$. Tomando $B = (\lambda_1 a/A)$ obtenemos $\varphi_B^{\pm} \in C(\lambda_1 a, \nu)$ y así la Proposición se cumple para $\varphi_B^{\pm}$ con $\int \varphi_B^{\pm} = 1$.

- $\left|\int (\psi \circ f^n)\varphi_B^+ dm - \int \psi d\mu_0 \int \varphi_B^+ dm\right| \leqslant K_0(\varphi_B^+, \psi)\Lambda_1^n$ y
- $\left|\int (\psi \circ f^n)\varphi_B^- dm - \int \psi d\mu_0 \int \varphi_B^- dm\right| \leqslant K_0(\varphi_B^-, \psi)\Lambda_1^n$,

donde $K_0(\varphi_B^{\pm}, \psi) = R_3\|\psi\|_1$. Aplicando linealidad demostremos que la Proposición también se cumple para φ,

$$\begin{aligned}
\left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0 \int \varphi dm\right| &= \left|\int (\psi \circ f^n)(\varphi_B^+ - \varphi_B^-) dm - \int \psi d\mu_0 \int (\varphi_B^+ - \varphi_B^-) dm\right| \\
&\leqslant \left|\int (\psi \circ f^n)\varphi_B^+ dm - \int \psi d\mu_0 \int \varphi_B^+ dm\right| \\
&\quad + \left|\int (\psi \circ f^n)\varphi_B^- dm - \int \psi d\mu_0 \int \varphi_B^- dm\right| \\
&\leqslant K_0(\varphi, \psi)\Lambda_1^n.
\end{aligned}$$

donde $K_0(\varphi, \psi) = K_0(\varphi_B^+, \psi) + K_0(\varphi_B^-, \psi)$. Esto completa la demostración de la Proposición. ■

Observación 3.3. La constante $K_0 = K_0(\varphi, \psi)$ que aparece en la demostración de la Proposición 7 tiene la forma

$$K_0(\varphi, \psi) \leqslant R\|\psi\|_1(\|\varphi\|_1 + H_\nu(\varphi)) \tag{3.1}$$

donde $H_\nu(\varphi)$ denota algún número que depende de A tal que φ es (A,ν)-Hölder, y $R > 0$ es independiente de φ y ψ. En efecto, la primera parte de la demostración nos da

$$K_0(\varphi_B^{\pm}, \psi) = R_3\|\psi\|_1 \int \varphi_B^{\pm} dm \leqslant R_3\|\psi\|_1\Big(\int |\varphi| dm + B\Big) \leqslant R_3\|\psi\|_1\Big(R_2\|\varphi\|_1 + \frac{a\lambda_1}{A}\Big)$$

con esto y la parte final de la demostración obtenemos

$$\begin{aligned} K_0(\varphi,\psi) &= K_0(\varphi_B^+,\psi)+K_0(\varphi_B^-,\psi) \\ &\leqslant R_3\|\psi\|_1\Big(R_2\|\varphi\|_1+\frac{a\lambda_1}{A}\Big)+R_3\|\psi\|_1\Big(R_2\|\varphi\|_1+\frac{a\lambda_1}{A}\Big) \\ &\leqslant 2R_3R_2\|\psi\|_1\Big(\|\varphi\|_1+\frac{a\lambda_1}{R_2A}\Big) \end{aligned}$$

tomando $R = 2R_2R_3 > 0$ y $H_\nu(\varphi) = (a\lambda_1/AR_2)$ se concluye nuestra afirmación.

Corolario 1 (Decaimiento Exponencial de Correlaciones) *Dadas φ_1 y φ_2 funciones ν-Hölder continuas, existe $K_0 = K_0(\varphi_1,\varphi_2)$ tal que*

$$\left|\int(\varphi_1\circ f^n)\varphi_2 d\mu_0-\int\varphi_1 d\mu_0\int\varphi_2 d\mu_0\right|\leqslant K_0\Lambda_1^n \text{ para todo } n\geqslant 0.$$

Demostración. Tomando $\psi = \varphi_1$ y $\varphi = \varphi_2\varphi_0$ en la proposición anterior se obtiene

$$\begin{aligned} \left|\int(\varphi_1\circ f^n)\varphi_2 d\mu_0-\int\varphi_1 d\mu_0\int\varphi_2 d\mu_0\right| &= \left|\int(\varphi_1\circ f^n)\varphi_2\varphi_0 dm-\int\varphi_1 d\mu_0\int\varphi_2\varphi_0 dm\right| \\ &= \left|\int(\psi\circ f^n)\varphi dm-\int\psi d\mu_0\int\varphi dm\right| \\ &\leqslant K_0\Lambda_1^n \text{ para todo } n\geqslant 0. \end{aligned}$$

Usando la definición de mixing exponencial obtenemos que (f,μ_0) tiene decaimiento exponencial de correlaciones. ■

Sea $\mathcal{B}_n = f^{-n}(\mathcal{B})$, para $n\geqslant 0$, donde $\mathcal{B}$ es la σ-algebra de Borel de M. Una función $\xi : M\to\mathbb{R}$ es $\mathcal{B}_n$-*medible* si y sólo si $\xi = \xi_n\circ f^n$ para alguna función medible ξ_n. En efecto, supongamos primero que ξ es una función $\mathcal{B}_n$-medible. Como $\mathcal{B}_n = f^{-n}(\mathcal{B})$ tenemos que para $y\in\mathbb{R}$ el conjunto $\xi^{-1}(y)$ tiene la forma $f^{-n}(A_y)$ para algún conjunto $A_y\in\mathcal{B}$. Tomando ξ_n una función $\mathcal{B}$-medible con $\xi_n\big|_{A_y}\equiv y$. Por otro lado, para todo $x\in f^{-n}(A_y)$ existe $w_x\in A_y$ tal que $f^n(x) = w_x$. Así $(\xi_n\circ f^n)(x) = \xi_n(w_x) = y = \xi(x)$, ya que $x\in\xi^{-1}(y)$. Para el reciproco, suponga que $\xi = \xi_n\circ f^n$ para alguna función ξ_n medible. Del hecho que ξ_n es una función medible, para $y\in\mathbb{R}$ se tiene $\xi_n^{-1}(y)$ es un conjunto $\mathcal{B}$-medible. Entonces $\xi^{-1}(y) = f^{-n}(\xi_n^{-1}(y))$ es un conjunto $\mathcal{B}_n$-medible. Por tanto ξ es una función $\mathcal{B}_n$-medible, con esto se completa la demostración de la afirmación. Además, $\mathcal{B}_0\supset\mathcal{B}_1\supset\cdots\supset\mathcal{B}_n\supset\cdots$. Una medida f-invariante μ es llamada *exacta* si la σ-algebra $\mathcal{B}_\infty = \bigcap_{n\geqslant 0}\mathcal{B}_n$ es μ-trivial, en el sentido que todas las funciones $\mathcal{B}_\infty$-medibles son constante en μ-casi todas partes. Note que medidas exactas son ergódicas: si $A\subset M$ es f-invariante entonces $\chi_A\in\mathcal{B}_\infty$.

Corolario 2 (Unicidad y Exactitud) *La medida μ_0 es exacta y es la única medida f-invariante que es absolutamente continua con respecto a m*

Demostración. Ya hemos demostrado que las medidas μ_0 y m son equivalentes. Sea $\psi \in L^1(\mu_0)$ una función $\mathcal{B}_\infty$-medible. Entonces para cada $n \geqslant 0$ existe una función medible ψ_n tal que $\psi = \psi_n \circ f^n$. Note que $\|\psi_n\|_1 = \|\psi\|_1 < +\infty$. Por la Proposición 10 y (3.1), dada cualquier función ν-Hölder continua φ, existe $K_0(\varphi) = R(\|\varphi\|_1 + H_\nu(\varphi)) > 0$ tal que

$$\left|\int \left(\psi - \int \psi d\mu_0\right) \varphi dm\right| = \left|\int (\psi_n \circ f)\varphi dm - \int \psi d\mu_0 \int \varphi dm\right|$$
$$\leqslant K_0(\varphi)\|\psi_n\|_1 \Lambda_1^n = K_0(\varphi)\|\psi\|_1 \Lambda_1^n$$

y esta última expresión tiende a cero cuando $n \to \infty$. Con todo esto obtenemos

$$\int \left(\psi - \int \psi d\mu_0\right)\varphi dm = 0.$$

Luego, $\psi = \int \psi d\mu_0$ en casi todas partes con respecto a m y μ_0, y como $\psi \in L^1(\mu_0)$ es arbitraria esto demuestra que μ_0 es una medida exacta. Finalmente, si μ es otra medida f-invariante con $\mu \ll m$ entonces $\mu \ll \mu_0$ (porque m y μ_0 son equivalentes) y luego $\mu = \mu_0$ ya que μ_0 es ergódica (ver [RM], pág 133). ∎

3.2. Teorema Central de Límite

En está sección procederemos a mostrar que el sistema (f, μ_0) satisface el Teorema Central de Límite, para eso primero haremos un estudio sobre las funciones en el espacio $L^2(\mu_0)$, luego introducimos la definición del operador proyección como la esperanza condicional de una función dada una σ-álgebra de Borel de M.

Comencemos está sección recordando $\mathcal{B}_n = f^{-n}(\mathcal{B})$, para $n \geqslant 0$, donde $\mathcal{B}$ es la σ-álgebra de Borel de M, y escribamos $L^2(\mathcal{B}_n) = \{\xi \in L^2(\mu_0) : \xi \text{ es } \mathcal{B}_n\text{-medible}\}$, para cada valor de $n \geqslant 0$. Observe que: $L^2(\mu_0) = L^2(\mathcal{B}_0) \supset L^2(\mathcal{B}_2) \supset \cdots \supset L^2(\mathcal{B}_n) \supset \cdots$ y

μ_0 es exacta si y sólo si $\bigcap_{n\geqslant 0} L^2(\mathcal{B}_n) = \{ \text{ constante } \}$

Ahora definamos el operador proyección ortogonal de $L^2(\mu_0)$ sobre $L^2(\mathcal{B}_n)$ y lo denotaremos como $E(\varphi \mid \mathcal{B}_n)$ con $\varphi \in L^2(\mu_0)$ y $n \geqslant 0$. Esta definición en la Teoría de Probabilidad es conocida como la esperanza condicional.(ver [A], pág 215)

Corolario 3 *Para cualquier función ν-Hölder continua φ con $\int \varphi d\mu_0 = 0$, existe $R_0 = R_0(\varphi)$ tal que*

$$\|E(\varphi \mid \mathcal{B}_n)\|_2 \leqslant R_0 \Lambda_1^n \quad \textit{para todo } n \geqslant 0.$$

Demostración. Procedamos a realizar la acotación:

$$\begin{aligned}\|E(\varphi \mid \mathcal{B}_n)\|_2 &= sup\left\{\int \xi\varphi d\mu_0 \ :\ \xi \in L^2(\mathcal{B}_n) \text{ y } \|\xi\|_2\right\}\\ &= sup\left\{\int (\psi\circ f^n)\varphi d\mu_0 \ :\ \psi \in L^2(\mu_0) \text{ y } \|\psi\|_2\right\}\\ &= sup\left\{\int (\psi\circ f^n)\varphi\varphi_0 dm \ :\ \psi \in L^2(\mu_0) \text{ y } \|\psi\|_2\right\}\\ &\leqslant K_0(\varphi\varphi_0)\|\psi\|_1\Lambda_1^n\end{aligned}$$

por el corolario 2 debido a que μ_0 es una medida exacta y tenemos como hipótesis $\int \varphi\varphi_0 dm = \int \varphi d\mu_0 = 0$; además, $\|\psi\|_1 \leqslant \|\psi\|_2 = 1$ porque $L^2(\mu_0) \subset L^1(\mu_0)$. Por ultimo tomando $R_0(\varphi) = K_0(\varphi\varphi_0)$ se completa la demostración. ■

Observación 3.4. La siguiente descripción de la acción del operador $\mathcal{L}$ en L^2 está contenida en lo que hemos hecho hasta el momento. Sea φ_0 como antes y definimos el siguiente conjunto $H = \{\varphi \in L^2(m) \ :\ \int \varphi dm = 0\}$. Claramente podemos ver que $\mathcal{L}(\varphi_0) = \varphi_0$ (punto fijo) y $\mathcal{L}(H) \subset H$ (2.2). En efecto, sea $\varphi \in H$, veamos que ocurre con $\mathcal{L}(\varphi)$,

$$\int \mathcal{L}(\varphi)dm = \int \mathcal{L}(\varphi)1dm = \int \varphi(1\circ f)dm = \int \varphi dm = 0.$$

Consideremos la isometría $h : L^2(m) \longrightarrow L^2(\mu_0)$ dada por:

$$h(\varphi) = \frac{\varphi}{\varphi_0},$$

y además introducimos el operador $\mathcal{P} : L^2(\mu_0) \longrightarrow L^2(\mu_0)$ definido como

$$\mathcal{P} = (h\circ\mathcal{L}\circ h^{-1})(\varphi).$$

Veamos la imagen de conjunto H a traves de la isometría h y obtenemos el siguiente conjunto:

$$N = h(H) = \left\{\psi \in L^2(m) \ :\ \int \psi d\mu_0 = 0\right\}.$$

Resulta que $\mathcal{P}(1) = 1$ y $\mathcal{P}(N) \subset N$, y por (2.2) se puede afirmar que $\mathcal{P}$ es el operador adjunto del operador de Koopman en $L^2(\mu_0)$, es decir, $U : L^2(\mu_0) \longrightarrow L^2(\mu_0)$ dado por $U(\psi) = \psi\circ f$. Denotemos por $L_0^2(\mathcal{B}_n)$ a la intersección del conjunto N con el espacio $L^2(\mathcal{B}_n)$, entonces U y $\mathcal{P}$ son operadores unitario con $U(L_0^2(\mathcal{B}_n)) = L_0^2(\mathcal{B}_{n+1})$ y $\mathcal{P}(L_0^2(\mathcal{B}_{n+1})) = L_0^2(\mathcal{B}_n)$.

La propiedad de exactitud de la medida μ_0 significa que el espacio $L^2(\mu_0)$ lo podemos escribir como la suma directa de dos conjuntos ortogonales

$$L^2(\mu_0) = \{\text{ constante }\} \oplus N.$$

Más aun, usemos la siguiente notación $L_0^2(\mathcal{B}_n) \perp L_0^2(\mathcal{B}_{n+1})$ para representar el complemento ortogonal de $L_0^2(\mathcal{B}_{n+1})$ con respecto al subespacio $L_0^2(\mathcal{B}_n)$, de esta manera podemos escribir

$$L^2(\mu_0) = \{\text{ constante }\} \oplus (\oplus_{n=0}^{\infty}[L_0^2(\mathcal{B}_n) \perp L_0^2(\mathcal{B}_{n+1})]).$$

Con está división y el corolario anterior implica que las componentes de cualquier función $\psi \in L^2(\mu_0)$, disminuyen exponencialmente rapído conforme n crece. Este hecho lo usaremos para la demostración del Teorema Central de Límite.

Basandomos en el análisis anterior, ahora podemos mostrar que las oscilaciones de las sumas de Birkhoff de una función observable, ν-Hölder continua, alrededor de su valor esperado convergen, en distribución, a un proceso gausianno (distribución normal). En primer lugar, enunciaremos y probaremos una versión abstracta del Teorema Central de Límite para funciones medibles (no invertible)

Teorema 11 *Sea $(M, \mathcal{B}, \mu)$ un espacio de probabilidad, $f : M \longrightarrow M$ una función medible tal que μ es f-invariante y f-ergódica, y $\phi \in L^2(\mu)$ tal que $\int \phi d\mu = 0$. Sea $(\mathcal{B}_n)_{n\geqslant 0}$ la sucesión no creciente de σ-álgebras $\mathcal{B}_n = f^{-n}(\mathcal{B})$, $n \geqslant 0$. Supongamos que*

$$\sum_{n=0}^{\infty} \|E(\phi|\mathcal{B}_n)\|_2 < +\infty$$

Entonces $\sigma \geqslant 0$ dado por

$$\sigma^2 = \int \phi^2 d\mu + 2\sum_{n=1}^{\infty}\int \phi(\phi \circ f^n) d\mu \tag{3.2}$$

es finito; y $\sigma = 0$ si y sólo si $\phi = u \circ f - u$ para alguna $u \in L^2(\mu)$. Por otra parte, si $\sigma > 0$, entonces dado cualquier intervalo $A \subset \mathbb{R}$,

$$\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) \in A\right\}\right) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}}\int \exp\left(\frac{-t^2}{2\sigma^2}\right) dt$$

cuando $n \to \infty$.

Antes de realizar la demostración este teorema, mostraremos que como consecuencia de él y el corolario 3 se obtiene el siguiente resultado que es la versión del Teorema Central de Límite para transformaciones expansoras.

Proposición 12 (Teorema Central de Límite) *Sea φ una función ν-Hölder continua y*

$$\sigma^2 = \int \phi^2 d\mu_0 + 2\sum_{j=1}^{\infty}\int \phi(\phi\circ f^j)d\mu \ , \qquad \textit{donde } \phi = \varphi - \int \varphi d\mu_0.$$

Entonces σ está bien definida; y $\sigma = 0$ si y sólo si $\phi = u\circ f - u$ para alguna $u \in L^2(\mu_0)$. Si $\sigma > 0$, entonces para todo intervalo $A \subset \mathbb{R}$ se cumple

$$\mu_0\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\left(\varphi(f^j(x)) - \int\varphi d\mu_0\right) \in A\right\}\right) \longrightarrow \frac{1}{\sigma\sqrt{2\pi}}\int_A \exp\left(\frac{-t^2}{2\sigma^2}\right) dt$$

cuando $n \to \infty$.

En lo que resta de sección nos encargamos de hacer la demostración del Teorema 11.

Demostración. La idea general para la demostración es tratar de escribir $\phi = \eta + \zeta\circ f - \zeta$ para funciones $\eta, \zeta \in L^2(\mu)$ tal que

(i) $\frac{1}{\sqrt{n}}(\zeta\circ f^n - \zeta)$ converge a cero en medida, cuando $n \to \infty$ (también probamos un hecho más fuerte: que converge a cero en $L^2(\mu)$).

(ii) Las variables aleatorias $\eta\circ f^n$ (identicamente distribuidas) satisfacen una condición de independencia, que se establece en (3.3)

El primer item significa que, dado cualquier $\varepsilon > 0$,

$$\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}|\zeta(f^n(x)) - \zeta(x)| > \varepsilon\right\}\right) \longrightarrow 0$$

lo cual implica que,

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\phi\circ f^j) &= \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}[\eta + \zeta\circ f - \zeta](f^j) \\
&= \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}[\eta(f^j) + (\zeta\circ f)(f^j) - \zeta(f^j)] \\
&= \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\eta\circ f^j) + \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}[(\zeta\circ f^{j+1}) - (\zeta\circ f^j)]
\end{aligned}$$

como el ultimo termino es una suma telescópica al desarrollarla se obtiene

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\phi\circ f^j) = \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\eta\circ f^j) + \frac{1}{\sqrt{n}}(\zeta\circ f^n - \zeta)$$

De esta manera podemos afirmar que $\dfrac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\phi\circ f^j)$ tiene la misma distribución límite que $\dfrac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\eta\circ f^j)$. No parece posible escoger las variables aleatorias $(\eta\circ f^n)$ para que sean independientes, en cuyo caso (la mayoria de) el teorema seguiría directamente del teorema 1 de la página 4. En lugar de ello, construimos η de tal forma que

$$E(\eta\circ f^n|\mathcal{B}_{n+1})=0 \qquad \text{para todo } n\geqslant 0 \tag{3.3}$$

sea suficiente para tener de la conclusión del teorema 1. En el lenguaje de la teoría de Probabilidad, (3.3) significa que la sucesión $(\eta\circ f^n)_{n\geqslant 0}$ es una martingala de diferencia inversa, y este resultado se conoce con el nombre de Teorema Central de Límite para Martingalas (ver Teorema 3, página 5).

Procedamos a desarrollar los detalles de la demostración. Introducimos el operador $U:L^2(\mu)\longrightarrow L^2(\mu)$ dado por $U\varphi=\varphi\circ f$, y sea $\mathcal{P}$ el operador adjunto con la siguiente propiedad

$$\int(\mathcal{P}\phi)\varphi d\mu=\int\phi(U\varphi)d\mu.$$

Note que U es una isometría sobre $L^2(\mathcal{B}_1)\subset L^2(\mathcal{B}_0)=L^2(\mu)$, y así, $\mathcal{P}:L^2(\mathcal{B}_1)\longrightarrow L^2(\mathcal{B}_0)$ también lo es (demostrar esto no es parte del trabajo, pero el interesado puede usar [A], pág 127). Por otra parte,

$$U(L^2(\mathcal{B}_n))=L^2(\mathcal{B}_{n+1}) \quad \text{y} \quad \mathcal{P}(L^2(\mathcal{B}_{n+1}))=L^2(\mathcal{B}_n),$$

para cada $n\geqslant 1$. Veamos esto ultimo, para $U(L^2(\mathcal{B}_n))=L^2(\mathcal{B}_{n+1})$. Sea $\phi\in L^2(\mathcal{B}_n)$, por definición del espacio $L^2(\mathcal{B}_n)$ se tiene que $\phi\in L^2(\mu)$ tal que ϕ es $\mathcal{B}_n$-medible, por este hecho podemos escribir $\phi=\xi\circ f^n$ para alguna ξ una función $\mathcal{B}$-medible. Por otro lado,

$$\begin{aligned}(U\phi)(x)&=(\phi\circ f)(x)=\phi(f(x))=(\xi\circ f^n)(f(x))\\&=\xi[f^n(f(x))]=\xi[f^{n+1}(x)]=(\xi\circ f^{n+1})(x)\end{aligned}$$

de lo cual se concluye que $(U\phi)\in L^2(\mu)$ y es una función $\mathcal{B}_{n+1}$-medible. Con un procedimiento análogo se puede probar que $\mathcal{P}(L^2(\mathcal{B}_{n+1}))=L^2(\mathcal{B}_n)$. Ahora definamos las siguientes funciones

$$\zeta=\sum_{j=1}^{\infty}\mathcal{P}^j(E(\phi|\mathcal{B}_j)) \quad \text{y} \quad \eta=\sum_{j=0}^{\infty}\mathcal{P}^j(E(\phi|\mathcal{B}_j)-E(\phi|\mathcal{B}_{j+1})).$$

Dado que $\mathcal{P}$ es una isometría, se tiene $\|\mathcal{P}^j(E(\phi|\mathcal{B}_j))\|_2 = \|E(\phi|\mathcal{B}_j)\|_2$, esto unido a la hipótesis del teorema se asegura que la serie en la definición de ζ converge en $L^2(\mu)$. Este hecho lo usaremos para probar la primera propiedad de como deberian de ser las funciones al comienzo de la demostración

$$\left\|\frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)\right\|_2 \leqslant \frac{1}{\sqrt{n}}(\|\zeta \circ f^n\|_2 + \|\zeta\|_2) = \frac{2}{\sqrt{n}}\|\zeta\|_2 \longrightarrow 0,$$

cuando $n \to \infty$. Como consecuencia de lo anterior, $\dfrac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)$ converge a cero en medida.

Para el segundo item del comienzo de la demostración, que tiene que ver con la función η. Como $E(\phi|\mathcal{B}_j)$ es una función en $L^2(\mathcal{B}_j)$ y este espacio lo podemos escribir como la suma directa de $L^2(\mathcal{B}_{j+1})$ y $(L^2(\mathcal{B}_{j+1}))^{\perp}$ de la siguiente manera:

$$E(\phi|\mathcal{B}_j) = v + w \quad con \quad v \in L^2(\mathcal{B}_{j+1}),\ w \in (L^2(\mathcal{B}_{j+1}))^{\perp}$$

por propiedades de la esperanza condicional ([A], página 220; o [W], página 83), $v = E(E(\phi|\mathcal{B}_j)|\mathcal{B}_{j+1}) = E(\phi|\mathcal{B}_{j+1})$ (por que $\mathcal{B}_{j+1} \subset \mathcal{B}_j$). Luego,

$$E(\phi|\mathcal{B}_j) = E(\phi|\mathcal{B}_{j+1}) + w$$

lo que implica que la proyección ortogonal de $E(\phi|\mathcal{B}_j)$ sobre $(L^2(\mathcal{B}_{j+1}))^{\perp}$, la podemos escribir como $(E(\phi|\mathcal{B}_j) - E(\phi|\mathcal{B}_{j+1}))$. Nuevamente, como $\mathcal{P}$ es una isometría, se obtiene

$$\|\mathcal{P}^j(E(\phi|\mathcal{B}_j) - E(\phi|\mathcal{B}_{j+1}))\|_2 = \|E(\phi|\mathcal{B}_j) - E(\phi|\mathcal{B}_{j+1})\|_2 \leqslant \|E(\phi|\mathcal{B}_j)\|_2$$

y con esto podemos asegurar que la serie en la definición η es también convergente en $L^2(\mu)$. Ahora desarrollemos la serie en la definición de la función η para hacer un reordenamiento de los terminos

$$\begin{aligned}
\eta =& [E(\phi|\mathcal{B}_0) - E(\phi|\mathcal{B}_1)] + [\mathcal{P}(E(\phi|\mathcal{B}_1)) - \mathcal{P}(E(\phi|\mathcal{B}_2))] \\
& + [\mathcal{P}^2(E(\phi|\mathcal{B}_2)) - \mathcal{P}^2(E(\phi|\mathcal{B}_3))] + \cdots \\
=& E(\phi|\mathcal{B}_0) + \mathcal{P}(E(\phi|\mathcal{B}_1)) + \mathcal{P}^2(E(\phi|\mathcal{B}_2)) + \cdots \\
& - E(\phi|\mathcal{B}_1) - \mathcal{P}(E(\phi|\mathcal{B}_2)) - \mathcal{P}^2(E(\phi|\mathcal{B}_3)) - \cdots \\
=& E(\phi|\mathcal{B}_0) + \sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi|\mathcal{B}_j)) - \sum_{j=1}^{\infty} \mathcal{P}^{j-1}(E(\phi|\mathcal{B}_j))
\end{aligned}$$

Es evidente que $E(\phi|\mathcal{B}_0) = \phi$, el segundo término es la función ζ, y procedamos a estudiar el tercer término para poder tener una expresión más clara de la función η.

Por otra parte, sabemos que $E(\phi|\mathcal{B}_j) \in L^2(\mathcal{B}_j)$ lo cual implica que $\mathcal{P}^{j-1}(E(\phi|\mathcal{B}_j)) \in L^2(\mathcal{B}_1)$, y $U\mathcal{P} = \mathrm{id}|_{L^2(\mathcal{B}_1)}$ por el hecho que el operador U y su operador adjunto $\mathcal{P}$ son isometrías sobre sus imagenes (son operadores unitarios). Con esto tenemos la siguiente propiedad, para cada $j \geqslant 1$

$$\mathcal{P}^{j-1}(E(\phi|\mathcal{B}_j)) = U\mathcal{P}[\mathcal{P}^{j-1}(E(\phi|\mathcal{B}_j))] = U\mathcal{P}^j(E(\phi|\mathcal{B}_j)) = \mathcal{P}^j(E(\phi|\mathcal{B}_j)) \circ f$$

Con lo antes mostrado, reescribimos la función η de tal forma que nos quede

$$\begin{aligned}\eta &= \phi + \sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi|\mathcal{B}_j)) - \sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi|\mathcal{B}_j)) \circ f \\ &= \phi + \zeta - \zeta \circ f\end{aligned}$$

Ya hemos visto que $E(\phi|\mathcal{B}_j) \in L^2(\mathcal{B}_j)$ y su proyección ortogonal sobre $(L^2(\mathcal{B}_{j+1}))^{\perp}$ la escribimos como $(E(\phi|\mathcal{B}_j) - E(\phi|\mathcal{B}_{j+1}))$, entonces $\mathcal{P}^j(E(\phi|\mathcal{B}_j) - E(\phi|\mathcal{B}_{j+1})) \in (L^2(\mathcal{B}_1))^{\perp}$ para todo $j \geqslant 0$. Luego,

$$\eta \in (L^2(\mathcal{B}_1))^{\perp}, \quad \text{es decir,} \quad E(\eta|\mathcal{B}_1) = 0$$

Esto implica que la condición en (3.3) es verificada:

$$E((\eta \circ f^n)|\mathcal{B}_{n+1}) = E(\eta|\mathcal{B}_1) \circ f^n = 0, \quad \text{para todo} \quad n \geqslant 0.$$

En particular, las variables $(\eta \circ f^n)$ son ortogonales dos a dos:

$$\int (\eta \circ f^k)(\eta \circ f^n) d\mu = 0 \quad \text{para todo } k > n \geqslant 0,$$

porque $(\eta \circ f^k)$ es una función de $L^2(\mathcal{B}_{n+1})$. Lo cual nos ayuda a tener el siguiente resultado,

$$\|\eta\|_2^2 = \frac{1}{n} \sum_{j=0}^{n-1} \|\eta \circ f^j\|_2^2 = \frac{1}{n} \left\| \sum_{j=0}^{n-1} (\eta \circ f^j) \right\|_2^2 = \left\| \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\eta \circ f^j) \right\|_2^2$$

anteriormente vimos que

$$\left\| \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\phi \circ f^j) - \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\eta \circ f^j) \right\|_2 = \left\| \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta) \right\|_2 \longrightarrow 0,$$

podemos concluir

$$\begin{aligned}\|\eta\|_2^2 &= \lim \left\| \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\phi \circ f^j) \right\|_2^2 \\ &= \lim \frac{1}{n} \left(\sum_{k=0}^{n-1} \int (\phi \circ f^k)^2 d\mu + 2 \sum_{0 \leqslant k < l \leqslant n-1} \int (\phi \circ f^k)(\phi \circ f^l) d\mu \right)\end{aligned}$$

veamos como podemos escribir el ultimo término, primero fijamos el índice l para poder ver la suma respecto del otro índice (k), luego desarrollar las sumas del índice k para contar los elementos comunes,

$$\begin{aligned}\sum_{0\leqslant k<l\leqslant n-1}\int(\phi\circ f^k)(\phi\circ f^l)d\mu =&\sum_{k=0}^{n-2}\int(\phi\circ f^k)(\phi\circ f^{n-1})d\mu+\sum_{k=0}^{n-3}\int(\phi\circ f^k)(\phi\circ f^{n-2})d\mu\\&+\sum_{k=0}^{n-4}\int(\phi\circ f^k)(\phi\circ f^{n-3})d\mu+\cdots+\\&+\sum_{k=0}^{1}\int(\phi\circ f^k)(\phi\circ f^{2})d\mu+\int\phi(\phi\circ f)d\mu\\=&(n-1)\int\phi(\phi\circ f)d\mu+(n-2)\int\phi(\phi\circ f^2)d\mu\\&+(n-3)\int\phi(\phi\circ f^3)d\mu+\cdots+\\&+2\int\phi(\phi\circ f^{n-2})d\mu+\int\phi(\phi\circ f^{n-1})d\mu\\=&\sum_{j=1}^{n-1}(n-j)\int\phi(\phi\circ f^j)d\mu\end{aligned}$$

sustituyendo en la expresión de arriba,

$$\begin{aligned}\|\eta\|_2^2&=\text{lím}\left(\int\phi^2d\mu+\frac{2}{n}\sum_{j=1}^{n-1}(n-j)\int\phi(\phi\circ f^j)d\mu\right)\\&=\text{lím}\left(\int\phi^2d\mu+2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\int\phi(\phi\circ f^j)d\mu\right)\\&=\int\phi^2d\mu+2\sum_{j=1}^{\infty}\int\phi(\phi\circ f^j)d\mu\end{aligned}$$

para la ultima igualdad note que $\phi\circ f^j\in L^2(\mathcal{B}_j)$, entonces $E(\phi|\mathcal{B}_j)=\phi$, y tenemos

$$\int\phi(\phi\circ f^j)d\mu=\int E(\phi|\mathcal{B}_j)(\phi\circ f^j)d\mu\leqslant\|E(\phi|\mathcal{B}_j)\|_2\|\phi\circ f^j\|_2=\|E(\phi|\mathcal{B}_j)\|_2\|\phi\|_2$$

y así

$$\begin{aligned}\left\|2\sum_{j=1}^{n-1}\frac{j}{n}\int\phi(\phi\circ f^j)d\mu+2\sum_{j=n}^{\infty}\int\phi(\phi\circ f^j)d\mu\right\|_2&\\
\leqslant 2\|\phi\|_2\left(\sum_{j=1}^{n-1}\frac{j}{n}\|E(\phi|\mathcal{B}_j)\|_2+\sum_{j=n}^{\infty}\|E(\phi|\mathcal{B}_j)\|_2\right)&\\
\leqslant 2\|\phi\|_2\left(\varepsilon\sum_{j=1}^{\infty}\|E(\phi|\mathcal{B}_j)\|_2+\sum_{j=n\varepsilon}\|E(\phi|\mathcal{B}_j)\|_2\right)&\end{aligned}$$

donde la ultima expresión puede hacerse arbitrariamente pequeña cuando se fija $\varepsilon > 0$ próximo de cero y tomando $n \gg \varepsilon^{-1}$. De esta manera, hemos demostrado que

$$\|\eta\|_2^2 = \sigma^2$$

recordar que σ fue definido en (3.2). En particular, se obtuvo que σ es finita. Por otra parte, $\sigma = 0$ implica $\eta = 0$ y $\phi = \zeta\circ f - \zeta$. Reciprocamente, si ϕ tiene la forma $\phi = u\circ f - u$ entonces podemos tomar $\zeta = u$, lo cual nos lleva a tener $\eta = 0$ y por los argumentos anteriores, $\sigma = \|\eta\|_2 = 0$. Con esto hemos realizado la primera parte de la demostración del teorema.

Para la parte que falta de la prueba vamos a suponer $\sigma > 0$. Sea $a < b$ fijos. Definamos la función $\Phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ por la siguiente expresión

$$\Phi(r,s) = \frac{1}{\sigma\sqrt{2\pi}}\int_r^s \exp\left(\frac{-t^2}{2\sigma^2}\right)dt.$$

Dado $\delta > 0$ tomamos $\varepsilon > 0$ tal que

$$\begin{aligned}\Phi(a-\varepsilon,b+\varepsilon) =&\frac{1}{\sigma\sqrt{2\pi}}\int_{a-\varepsilon}^{b+\varepsilon}\exp\left(\frac{-t^2}{2\sigma^2}\right)dt\\
=&\frac{1}{\sigma\sqrt{2\pi}}\int_{a-\varepsilon}^{a}\exp\left(\frac{-t^2}{2\sigma^2}\right)dt+\frac{1}{\sigma\sqrt{2\pi}}\int_{a}^{b}\exp\left(\frac{-t^2}{2\sigma^2}\right)dt\\
&+\frac{1}{\sigma\sqrt{2\pi}}\int_{b}^{b+\varepsilon}\exp\left(\frac{-t^2}{2\sigma^2}\right)dt\end{aligned}$$

lo que implica

$$\Phi(a-\varepsilon,b+\varepsilon)\leqslant\Phi(a,b)+\delta$$

Luego, tomando $n_1 \geqslant 1$ tal que

$$\mu\left(\left\{x\in M\ :\ \left|\frac{1}{\sqrt{n}}(\zeta\circ f^n-\zeta)(x)\right|>\varepsilon\right\}\right)\leqslant\delta\quad\text{para todo } n\geqslant n_1$$

Por el Teorema Central de Límite para Martingalas, ver en la introducción, existe un $n_2 \geqslant n_1$ tal que

$$\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\eta(f^j(x)) \in (a-\varepsilon, b+\varepsilon)\right\}\right) \leqslant \Phi(a-\varepsilon, b+\varepsilon) + \delta.$$

Del siguiente hecho

$$a < \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) < b$$

podemos tener

$$a-\varepsilon < \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\eta \circ f^j) < b+\varepsilon \quad \text{o} \quad \left|\frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)(x)\right| > \varepsilon$$

con esto concluimos que

$$\begin{aligned}
&\mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) \in (a,b)\right\}\right) \\
&\qquad\qquad \leqslant \mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\eta(f^j(x)) \in (a-\varepsilon, b+\varepsilon)\right\}\right) \\
&\qquad\qquad\quad + \mu\left(\left\{x \in M \ : \ \left|\frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)(x)\right| > \varepsilon\right\}\right) \\
&\qquad\qquad \leqslant \Phi(a-\varepsilon, b+\varepsilon) + \delta + \delta = \Phi(a-\varepsilon, b+\varepsilon) + 2\delta \\
&\qquad\qquad \leqslant \Phi(a,b) + 3\delta
\end{aligned}$$

lo anterior es valido para cualquier $\delta > 0$,

$$\limsup \mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) \in (a,b)\right\}\right) \leqslant \Phi(a,b),$$

de manera análoga, podemos obtener

$$\liminf \mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) \in (a,b)\right\}\right) \geqslant \Phi(a,b).$$

Por tanto,

$$\lim \mu\left(\left\{x \in M \ : \ \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x)) \in (a,b)\right\}\right) = \Phi(a,b) = \frac{1}{\sigma\sqrt{2\pi}}\int_a^b \exp\left(\frac{-t^2}{2\sigma^2}\right)dt.$$

el teorema queda demostrado. ■

3.3. Estabilidad Estocástica

A continuación mostraremos que las transformaciones expansoras son estables bajo perturbaciones aleatorias. Sea T un espacio métrico con su σ-álgebra de Borel $\mathcal{B}(T)$. Sea $p : T \to C^{1+\nu_0}(M)$ es una función continua en un punto τ con $p(\tau) = f$. Denotamos $p(t) = f_t$ para simplificar. De esta manera si t está próximo de τ entonces f_t es uniformemente próxima de f y Df_t es próxima de Df con respecto de la norma ν_0-Hölder:

$$|||G|||_{\nu_0} = \sup\{\|G(x)\| : x \in M\} + \sup\left\{\frac{\|G(x) - G(y)\|}{d(x, y)^{\nu_0}} : x, y \in M, 0 < d(x, y) \leqslant \rho_0\right\}.$$

(recordemos que ρ_0 proviene de la condición de Hölder en la definición del cono $C(a, \nu)$). Así tenemos una familia parametrizada $f_t : M \to M$ de transformaciones $C^{1+\nu_0}$ donde $t \in T$. Un caso particular importante es seguido cuando T es alguna vecindad de f en $C^{1+\nu_0}$ y $p(t) = t$.
Ahora consideremos una familia $(\theta_\varepsilon)_{\varepsilon>0}$ de medidas de probabilidad de Borel regulares en T tal que $\operatorname{supp}(\theta_\varepsilon) \to \{\tau\}$, cuando $\varepsilon \to 0$ en la métrica de Hausdorff para subconjuntos cerrados de T y denotamos $\operatorname{supp}(\theta_\varepsilon)$ como el soporte de la θ_ε. Fijando un $\varepsilon > 0$ pequeño considere una familia cualquiera $(t_i) = (t_i^\varepsilon)_{i\geqslant 1}$ de variables aleatorias $t_i : (\Omega, \mathcal{A}, P) \to (T, \mathcal{B}(T), \theta_\varepsilon)$ independientes y todas con distribución θ_ε esto es $(t_i)_* P = \theta_\varepsilon$ para todo $i \geqslant 1$.
Nuestro intéres es comparar el comportamiento asintótico de las trayectorias aleatorias $X_j^\varepsilon = p(t_j(\omega)) \circ \cdots \circ p(t_1(\omega))$ con $\omega \in \Omega$ del proceso $X_j^\varepsilon : \Omega_\varepsilon \to C^{1+\nu_0}(M)$ con el comportamiento asintótico de las trayectorias determinísticas $f^j(x)$ en M con $\varepsilon > 0$ pequeño y, las comparaciones se efectúan desde el punto de vista probabilístico.

Para tales efectos se introducen las versiones perturbadas de los operadores $\mathcal{L}$ y U usados antes con relación a la transformación f,

$$(U_t\varphi)(x) = \varphi(f_t(x)), \qquad (\mathcal{L}_t\varphi)(y) = \sum_{f_t(x)=y} \varphi(x)|\det Df_t(x)|^{-1}$$

y también

$$(\widehat{U}_\varepsilon\varphi)(x) = \int (U_t\varphi)(x)d\theta_\varepsilon(t), \qquad (\widehat{\mathcal{L}}_\varepsilon\varphi)(y) = \int (\mathcal{L}_t\varphi)(y)d\theta_\varepsilon(t)$$

para toda $\varphi \in E = C^0(M, \mathbb{R})$.

De acuerdo con [K2] diremos que una medida de probabilidad μ en M es *estacionaria* para el proceso de Markov (X_n) si

$$\int P(x, B)d\mu(x) = \mu(B), \quad \text{para todo } B \subset M \text{ medible}.$$

Denotamos $P(x,B) = P\{X_{n+1} \in B | X_n = x\}$ a las probabilidades de transición de (X_n).

Con esto podemos dar condiciones que caracterizan las medidas estacionarias para el proceso $(X_n^\varepsilon)_{n\geqslant 1}$.

Lema 1 *Sea φ_ε una función no negativa en $L^1(m)$ con $\widehat{\mathcal{L}}_\varepsilon \varphi_\varepsilon = \varphi_\varepsilon$ y $\int \varphi_\varepsilon dm = 1$. Entonces $\mu_\varepsilon = \varphi_\varepsilon m$ es una medida de probabilidad estacionaria para (X_n^ε).*

Demostración. Afirmamos que la medida μ_ε cumple con la siguiente igualdad

$$\int (\widehat{U}_\varepsilon \psi) d\mu_\varepsilon = \int \psi d\mu_\varepsilon \quad \text{para toda } \psi \text{ continua.} \tag{3.4}$$

En efecto, por el Teorema de Fubini y la versión de (2.2) para f_t, tenemos

$$\begin{aligned}
\int (\widehat{\mathcal{L}}_\varepsilon \varphi)(y)\psi(y)dm(y) &= \int\int (\mathcal{L}_t \varphi)(y) d\theta_\varepsilon(t)\psi(y)dm(y) \\
&= \int\int (\mathcal{L}_t \varphi)(y)\psi(y)dm(y)d\theta_\varepsilon(t) \\
&= \int\int \varphi(x)(U_t\psi)(x)dm(x)d\theta_\varepsilon(t) \\
&= \int \varphi(x)\left(\int (U_t\psi)(x)d\theta_\varepsilon(t)\right)dm(x) \\
&= \int \varphi(x)(\widehat{U}_\varepsilon \varphi(x))dm(x).
\end{aligned}$$

Ahora tomando $\mu_\varepsilon = \varphi_\varepsilon m$ y $\varphi = \varphi_\varepsilon$ en lo anterior se obtiene

$$\int (\widehat{U}_\varepsilon \psi)d\mu_\varepsilon = \int (\widehat{U}_\varepsilon \psi)\varphi_\varepsilon dm = \int (\widehat{\mathcal{L}}_\varepsilon \varphi_\varepsilon)\psi dm = \int \varphi_\varepsilon \psi dm = \int \psi d\mu_\varepsilon.$$

Por tanto

$$\int (\widehat{U}_\varepsilon \psi)d\mu_\varepsilon = \int \psi d\mu_\varepsilon,$$

para toda ψ continua y también para $\psi \in L^1(m)$. Esto completa la demostración de nuestra afirmación. Por otro lado, consideremos $Y_j^\varepsilon(w) = ev_x \circ X_j^\varepsilon$ las trayectorias aleatorias en M, denotemos $ev_x : C^{1+\nu_0}(M) \to M$ a la función evaluación dada por $ev_x(g) = g(x)$ y B un subconjunto de M medible.

$$\begin{aligned}
P^\varepsilon(x,B) &= P\{Y_1^\varepsilon \in B | Y_0^\varepsilon = x\} \\
&= P\{w \in \Omega_\varepsilon : (ev_x \circ X_1^\varepsilon)(w) \in B | Y_0^\varepsilon = x\} \\
&= P\{w \in \Omega_\varepsilon : (ev_x \circ p \circ t_1)(w) \in B\} \\
&= \big[(ev_x)_* \circ p_* \circ (t_1)_*\big] P(B) \\
&= \big[(ev_x)_* \circ p_*\big](\theta_\varepsilon(B)) \\
&= \theta_\varepsilon\{t \in T : f_t(x) \in B\}
\end{aligned}$$

Llamemos $A_x = \{t \in T : f_t(x) \in B\}$ y es fácil ver que $\chi_{f_t^{-1}(B)}(x) = \chi_{A_x}(t)$ para todo $x \in M$ y todo $t \in T$. Luego,

$$P^\varepsilon(x,B) = \theta_\varepsilon\{t \in T : f_t(x) \in B\} = \theta_\varepsilon(A_x) = \int \chi_{A_x}(t) d\theta_\varepsilon(t) = \int \chi_{f_t^{-1}(B)}(x) d\theta_\varepsilon(t)$$

integrando $P^\varepsilon(x,B)$ con respecto de μ_ε, tenemos

$$\begin{aligned}\int P^\varepsilon(x,B) d\mu_\varepsilon(x) &= \int \left(\int \chi_{f_t^{-1}(B)}(x) d\theta_\varepsilon(t)\right) d\mu_\varepsilon(x) \\ &= \int \left(\int (U_t \chi_B)(x) d\theta_\varepsilon(t)\right) d\mu_\varepsilon(x) \\ &= \int (\widehat{U}_\varepsilon \chi_B)(x) d\mu_\varepsilon(x)\end{aligned}$$

Ahora, tomando $\psi = \mathcal{X}_B$ en (3.4) se tiene

$$\int (\widehat{U}_\varepsilon \chi_B) d\mu_\varepsilon = \int \chi_B d\mu_\varepsilon = \mu_\varepsilon(B)$$

con lo cual obtenemos

$$\int P^\varepsilon(x,B) d\mu_\varepsilon(x) = \mu_\varepsilon(B).$$

Por tanto μ_ε es una medida estacionaria para el proceso (X_j^ε). ■

Procedamos a demostrar la existencia de la función φ_ε. Veremos además que está es única y próxima de φ_0 si $\varepsilon > 0$ es pequeño.

Nuestra suposición implica que si $\varepsilon > 0$ suficientemente pequeño entonces toda f_t con $t \in \operatorname{supp}(\theta_\varepsilon)$ es una transformación expansora $C^{1+\nu_0}$, con cotas uniformes $\sigma_1 > 1$ y $a_1 > 0$ iguales a la taza de expansión de f_t y la constante de Hölder de $\log|\det Df_t|$ respectivamente. De esta manera la estimación en la demostración de la Proposición 4 se aplica uniformemente para obtener

$$\mathcal{L}_t(C(a,\nu)) \subset C(a\lambda_1,\nu) \quad \text{para todo } t \in \operatorname{supp}(\theta_\varepsilon),$$

con $\sigma_1^{-\nu} < \lambda_1 < 1$ y $a \geqslant a_0/(\lambda_1 - \sigma_1^{-\nu})$. Por tanto, por convexidad del cono $C(a\lambda_1,\nu)$, resulta $\widehat{\mathcal{L}}_\varepsilon(C(a,\nu)) \subset C(\lambda_1 a,\nu)$ para todo $\varepsilon > 0$ pequeño.
Argumentado como antes para $\varphi_n = \mathcal{L}^n\mathbf{1}$, concluimos que la sucesión de funciones $(\varphi_{\varepsilon,n})$ definida por $\varphi_{\varepsilon,n} = \widehat{\mathcal{L}}_\varepsilon^n\mathbf{1}$ para $n \geqslant 1$, converge uniformemente a alguna $\varphi_\varepsilon \in C(\lambda_1 a,\nu)$ cuando $n \to +\infty$, la cual es punto fijo de $\widehat{\mathcal{L}}_\varepsilon$. Tomemos $\mu_\varepsilon = \varphi_\varepsilon m$. Note que está medida de probabilidad es equivalente a m.

Primero establecemos una consecuencia del argumento previo, el cual tiene importancia por si mismo.

COROLARIO 4 (Mixing Exponencial para Perturbaciones Aleatorias) *Sean φ una función ν-Hölder y ψ una función $L^1(m)$. Entonces existe $K_0 = K_0(\varphi, \psi) > 0$ tal que*

$$\left| \int (\widehat{U}_\varepsilon^n \psi)\varphi dm - \int \psi d\mu_\varepsilon \int \varphi dm \right| \leqslant K_0 \Lambda_1^n$$

para todo $n \geqslant 0$ y todo $\varepsilon > 0$ suficientemente pequeño.

Demostración. Se demuestra de manera ánaloga a la Proposición 10 en la página 28, reemplazando U, $\mathcal{L}$, μ_0, φ_0 por $\widehat{U}_\varepsilon$, $\widehat{\mathcal{L}}_\varepsilon$, μ_ε, φ_ε respectivamente. ∎

A continuación, mostraremos que μ_ε determina el comportamiento asintótico de casi todas las trayectorias aleatorias $(X_j^\varepsilon)_{j \geqslant 0}$, en el sentido que

$$\frac{1}{n} \sum_{j=0}^{n-1} \varphi(X_j^\varepsilon) \to \int \varphi d\mu_\varepsilon,$$

para m-casi toda elección del punto inicial x_0 y θ_ε-casi toda elección de suceciones $(t_n)_{n \geqslant 1}$.

Ahora considere la medida de probabilidad producto $\nu_\varepsilon = \mu_\varepsilon \times \theta_\varepsilon^{\mathbb{N}}$ caracterizada en la σ-álgebra producto de $M \times T^{\mathbb{N}}$ por

$$\nu_\varepsilon(A \times B_1 \times \cdots \times B_k) = \mu_\varepsilon(A) \cdot \theta_\varepsilon(B_1) \cdots \theta_\varepsilon(B_k)$$

para conjuntos de Borel $A \subset M, B_1, \ldots, B_k \subset T$ y $k \geqslant 1$ (ver [RD], pág 255). Se introduce la transformación shift $\sigma : M \times T^{\mathbb{N}} \to M \times T^{\mathbb{N}}$ dada por $\sigma(x, t_1, t_2, \ldots) = (f_{t_1}(x), t_2, \ldots)$.

Nuestro siguiente paso es aplicar el Teorema Ergódico de Birkhoff, pero antes necesitamos mostrar que la medida producto ν_ε es σ-invariante. Esto lo haremos en el siguiente Lema empleando el hecho de que la medida μ_ε es estacionaria. Denotaremos $\theta_\varepsilon^{(k)}$ a la medida producto $\bigotimes_{i \geqslant k} \theta_\varepsilon$

LEMA 2 *Si μ_ε es estacionaria, entonces ν_ε es σ-invariante.*

Demostración. Para cualquier función g en $C^0(M \times T^{\mathbb{N}}, \mathbb{R})$ definimos

$$\widehat{g}(x) = \int g(x, w) d\theta_\varepsilon^{(1)}(w)$$

donde $x \in M$ y $w \in T^{\mathbb{N}}$ con $w = (t_1, t_2, \ldots)$. Estudiemos la parte derecha de (3.4) para $\widehat{g}$

$$\begin{aligned} \int \widehat{g} d\nu_\varepsilon &= \int \int g(x, w) d\theta_\varepsilon^{(1)}(w) d\mu_\varepsilon(x) \\ &= \int g(x, w) d(\mu_\varepsilon \times \theta_\varepsilon^{(1)})(x, w) \\ &= \int g(x, w) d\nu_\varepsilon(x, w). \end{aligned}$$

Por otro lado, veamos la parte izquierda de la igualdad (3.4) para la función $\widehat{g}$ y tomemos $\tilde{w} = (t_2, t_3, \ldots) \in T^{\mathbb{N}}$

$$\begin{aligned}
\int (\widehat{U}_\varepsilon \widehat{g})(x) d\mu_\varepsilon(x) &= \int \left(\int (U_{t_1}\widehat{g})(x) d\theta_\varepsilon(t_1) \right) d\mu_\varepsilon(x) \\
&= \int \int (\widehat{g} \circ f_{t_1})(x) d\theta_\varepsilon(t_1) d\mu_\varepsilon(x) \\
&= \int \int \int g(f_{t_1}(x), \tilde{w}) d\theta_\varepsilon^{(2)}(\tilde{w}) d\theta_\varepsilon(t_1) d\mu_\varepsilon(x) \\
&= \int \int g(f_{t_1}(x), \tilde{w}) d\theta_\varepsilon^{(1)}(t_1, t_2, \ldots) d\mu_\varepsilon(x) \\
&= \int \int g(\sigma(x, w)) d\theta_\varepsilon^{(1)}(w) d\mu_\varepsilon(x) \\
&= \int (g \circ \sigma)(x, w) d\nu_\varepsilon(x, w).
\end{aligned}$$

Como la medida estacionaria μ_ε cumple con (3.4) obtenemos

$$\int (g \circ \sigma) d\nu_\varepsilon = \int g d\nu_\varepsilon \quad \text{para toda } g \in C^0(M \times T^{\mathbb{N}}, \mathbb{R}).$$

Por tanto, la medida producto ν_ε es σ-invariante. ∎

Con el resultado previo aplicamos el Teorema Ergódico de Birkhoff, para obtener que el límite

$$\widetilde{\varphi}(x, t_1, t_2, \ldots) = \lim_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ \pi_0)(\sigma^j(x, t_1, t_2, \ldots))$$

existe para ν_ε-casi todo $(x, t_1, t_2, \ldots) \in M \times T^{\mathbb{N}}$, donde π_0 es la proyección canonica $\pi_0 : M \times T^{\mathbb{N}} \to M$ dada por $\pi_0(x, t_1, t_2, \ldots) = x$. Note que $\pi_0(\sigma^j(x, t_1, t_2, \ldots))$ es precisamente lo que hemos denotado X_j^ε. También por el Teorema de Birkhoff tenemos $\int \widetilde{\varphi} d\nu_\varepsilon = \int (\varphi \circ \pi_0) d\nu_\varepsilon$. Nos resta demostrar

$$\widetilde{\varphi}(x, t_1, t_2, \ldots) = \int (\varphi \circ \pi_0) d\nu_\varepsilon = \int \varphi d\mu_\varepsilon \tag{3.5}$$

para completar la demostración de la afirmación enunciada antes del Lema 2. La segunda igualdad es fácil de ver haciendo uso de las definiciones de la medida producto ν_ε y la proyección π_0. Para la igualdad de la izquierda. Definimos para cada $k \geqslant 0$

$$\widetilde{\varphi}_k(x, t_1, t_2, \ldots, t_k) = \int \widetilde{\varphi}(x, t_1, \ldots, t_k, t_{k+1}, \ldots) d\theta_\varepsilon^{(k+1)}(t_{k+1}, t_{k+2}, \ldots).$$

Por el Teorema de Birkhoff $\widetilde{\varphi}$ es σ-invariante, es decir, $\widetilde{\varphi}\circ\sigma=\widetilde{\varphi}$ de lo cual obtenemos

$$\begin{aligned}\widetilde{\varphi}_0(x) &= \int \widetilde{\varphi}(x,t_1,t_2,\dots)d\theta_\varepsilon^{(1)}(t_1,t_2,\dots)\\ &= \int (\widetilde{\varphi}\circ\sigma)(x,t_1,t_2,\dots)d\theta_\varepsilon^{(1)}(t_1,t_2,\dots)\\ &= \int \widetilde{\varphi}(f_{t_1}(x),t_2,\dots)d\theta_\varepsilon^{(2)}(t_2,t_3,\dots)d\theta_\varepsilon(t_1)\\ &= \int \widetilde{\varphi}_0(f_{t_1}(x))d\theta_\varepsilon(t_1) = (\widehat{U}_\varepsilon\widetilde{\varphi}_0)(x)\end{aligned}$$

para μ_ε-casi todo $x\in M$. Entonces

$$\int\left(\widetilde{\varphi}_0-\int\widetilde{\varphi}_0 d\mu_\varepsilon\right)\varphi dm = \int(\widehat{U}_\varepsilon^n\widetilde{\varphi}_0)\varphi dm-\int\widetilde{\varphi}_0 d\mu_\varepsilon\int\varphi dm$$

y por el Corolario 3 la última expresión tiende a cero cuando $n\to\infty$ para toda φ ν-Hölder. Por tanto

$$\widetilde{\varphi}_0=\int\widetilde{\varphi}_0 d\mu_\varepsilon=\int\widetilde{\varphi}d\nu_\varepsilon=\int(\varphi\circ\pi_0)d\nu_\varepsilon=\int\varphi d\mu_\varepsilon$$

en m-casi todas partes, y así también en μ_ε-casi todas partes. La segunda igualdad proviene de la definición de $\widetilde{\varphi}_0$, la tercera del Teorema de Birkhoff y la última es precisamente la igualdad de la derecha en (3.5).

Más generalmente, para $k\geqslant 1$,

$$\begin{aligned}\widetilde{\varphi}_k(x,t_1,\dots,t_k) &= \int\widetilde{\varphi}(x,t_1,\dots,t_k,t_{k+1},\dots)d\theta_\varepsilon^{(k+1)}(t_{k+1},t_{k+2},\dots)\\ &= \int(\widetilde{\varphi}\circ\sigma)(x,t_1,\dots,t_k,t_{k+1},\dots)d\theta_\varepsilon^{(k+1)}(t_{k+1},t_{k+2},\dots)\\ &= \int\widetilde{\varphi}(f_{t_1}(x),t_2,\dots,t_k,t_{k+1},\dots)d\theta_\varepsilon^{(k+1)}(t_{k+1},t_{k+2},\dots)\\ &= \widetilde{\varphi}_{k-1}(f_{t_1}(x),t_2,\dots,t_k).\end{aligned}$$

Se sigue, por inducción, que para todo $k\geqslant 0$,

$$\widetilde{\varphi}_k=\int\varphi d\mu_\varepsilon \quad \text{en } (\mu_\varepsilon\times\theta_\varepsilon^k)-\text{casi todas partes.}$$

Esto nos da, $\widetilde{\varphi}=\int\varphi d\mu_\varepsilon$ en $(\mu_\varepsilon\times\theta_\varepsilon^k)$-casi todas partes (ver [RD]). Resumiendo tenemos

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi(X_j^\varepsilon)=\frac{1}{n}\sum_{j=0}^{n-1}(\varphi\circ\pi_0)(\sigma^j(x,t_1,t_2,\dots))\to\widetilde{\varphi}=\int\varphi d\mu_\varepsilon,$$

y así se completa la demostración de nuestra afirmación.

A continuación mostraremos el Teorema principal de esta sección. En este Teorema se demostrara que la función φ_ε converge uniformemente a la función φ_0 cuando $\varepsilon \to 0$. Para eso mostraremos que la sucesión de funciones $\mathcal{L}_\varepsilon^n 1$ es próxima de la sucesión de funciones $\mathcal{L}^n 1$ si $\varepsilon > 0$ es suficientemente pequeño.

Teorema 13 *φ_ε converge uniformemente a φ_0, cuando $\varepsilon \to 0$. En particular, $\mu_\varepsilon \to \mu_0$ debil$\star$ cuando $\varepsilon \to 0$.*

Demostración. Por la Proposición 10 y la ecuación (3.1), existe $K_0 = K_0(\varphi)$ tal que

$$\left| \int \psi(\mathcal{L}^n 1) dm - \int \psi d\mu_0 \right| \leqslant K_0 \|\psi\|_1 \Lambda_1^n$$

y un resultado similar para $\widehat{\mathcal{L}}_\varepsilon$ se deduce usando el Corolario 4 y (3.1),

$$\left| \int \psi(\widehat{\mathcal{L}}_\varepsilon^n 1) dm - \int \psi d\mu_\varepsilon \right| \leqslant K_0 \|\psi\|_1 \Lambda_1^n.$$

Dado $y \in M$ y $t \in \operatorname{supp}(\theta_\varepsilon)$ con $\varepsilon > 0$ suficientemente pequeño, porque f_t es expansora ver la página 43, podemos escribimos

$$f^{-1}(y) = \{x_1, \ldots, x_k\} \quad \text{y} \quad f_t^{-1}(y) = \{x_1^t, \ldots, x_k^t\}, \tag{3.6}$$

con $\sup\{d(x_i^t, x_i) : \ 1 \leqslant i \leqslant k, \ t \in \operatorname{supp}(\theta_\varepsilon), \ y \in M\}$ convergiendo a cero cuando $\varepsilon \to 0$. Así tenemos

$$(\mathcal{L}\varphi)(y) = \sum_{i=1}^{k} \varphi(x_i) |\det Df(x_i)|^{-1} \quad \text{y} \quad (\mathcal{L}_t \varphi)(y) = \sum_{i=1}^{k} \varphi(x_i^t) |\det Df_t(x_i^t)|^{-1}.$$

Como Df esta próxima de Df_t se obtiene que $|\det Df|$ esta próximo de $|\det Df_t|$ y el cociente de estás expresiones es próximo de 1. Además dado $\varphi \in C(a, \nu)$ se tiene

$$\frac{\varphi(x_i^t)}{\varphi(x_i)} \leqslant \exp\left(a d(x_i^t, x_i)^\nu\right)$$

donde x_i y x_i^t como en (3.6) con $d(x_i^t, x_i) \leqslant \varepsilon$. Aplicando esto al cociente de los operadores

$$\begin{aligned} \frac{(\mathcal{L}_t \varphi)(y)}{(\mathcal{L}\varphi)(y)} &= \frac{\sum_{i=1}^k \varphi(x_i^t) |\det Df_t(x_i^t)|^{-1}}{\sum_{i=1}^k \varphi(x_i) |\det Df(x_i)|^{-1}} \\ &\leqslant \max_{1 \leqslant i \leqslant k} \left\{ \frac{\varphi(x_i^t) |\det Df_t(x_i^t)|^{-1}}{\varphi(x_i) |\det Df(x_i)|^{-1}} \right\} \\ &\leqslant \max_{1 \leqslant i \leqslant k} \left\{ \exp\left(a d(x_i^t, x_i)^\nu\right) \frac{|\det Df_t(x_i^t)|^{-1}}{|\det Df(x_i)|^{-1}} \right\} \\ &\leqslant \exp\left(a \varepsilon^\nu\right) \max_{1 \leqslant i \leqslant k} \left\{ \frac{|\det Df_t(x_i^t)|^{-1}}{|\det Df(x_i)|^{-1}} \right\} \end{aligned}$$

y como la útima expresión es próximo de 1 cuando $\varepsilon > 0$ es suficientemente pequeño. Esto lo podemos escribir

$$1 - \xi(\varepsilon) \leqslant \frac{(\mathcal{L}_t\varphi)(y)}{(\mathcal{L}\varphi)(y)} \leqslant 1 + \xi(\varepsilon) \quad \text{para todo } y \in M, t \in \operatorname{supp}(\theta_\varepsilon), \varphi \in C(a,\nu)$$

donde $\xi(\varepsilon)$ es independiente de y, t, φ; además converge a 0 cuando $\varepsilon \to 0$. Como una consecuencia, dado cualquier $\varphi \in C(a,\nu)$,

$$\left| \frac{(\widehat{\mathcal{L}}_\varepsilon\varphi)(y)}{(\mathcal{L}\varphi)(y)} - 1 \right| \leqslant \xi(\varepsilon), \quad \text{para todo } y \in M.$$

Luego, $\|\widehat{\mathcal{L}}_\varepsilon\varphi - \mathcal{L}\varphi\|_0 \leqslant \xi(\varepsilon)\|\mathcal{L}\varphi\|_0$. Empleando lo anterior para cada $\varphi = \mathcal{L}^i 1, 0 \leqslant i < n$ obtenemos

$$\begin{aligned}
\left| \int (\widehat{\mathcal{L}}_\varepsilon^n 1 - \mathcal{L}^n 1)\psi dm \right| &= \left| \sum_{i=0}^{n-1} \int \widehat{\mathcal{L}}_\varepsilon^{n-i-1}(\widehat{\mathcal{L}}_\varepsilon - \mathcal{L})(\mathcal{L}^i 1) \cdot \psi dm \right| \\
&= \left| \sum_{i=0}^{n-1} \int (\widehat{\mathcal{L}}_\varepsilon - \mathcal{L})(\mathcal{L}^i 1)(\widehat{U}_\varepsilon^{n-i-1}\psi) dm \right| \\
&\leqslant \sum_{i=0}^{n-1} \|(\widehat{\mathcal{L}}_\varepsilon - \mathcal{L})(\mathcal{L}^i 1)\|_0 \|\widehat{U}_\varepsilon^{n-i-1}\psi\|_1 \sup|\varphi_\varepsilon|^{-1} \\
&\leqslant \sum_{i=0}^{n-1} \xi(\varepsilon) \cdot \sup|\mathcal{L}^{i+1} 1| \cdot \|\psi\|_1 \cdot \sup|\varphi_\varepsilon|^{-1} \\
&\leqslant Kn\xi(\varepsilon)\|\psi\|_1,
\end{aligned}$$

donde K depende sólo de f, $\|\widehat{U}_\varepsilon\psi\|_1 = \|\psi\|_1$ por (3.4) y en la última desigualdad usamos el hecho que $\mathcal{L}^i 1 \in C(a\lambda_1, \nu)$ y $\int \mathcal{L}^i 1 dm = 1$. De modo que $\sup|\mathcal{L}^{i+1}|$ admite una cota superior independiente de i (vea la prueba de la Proposición 9). De todas las estimaciones previas se obtiene

$$\begin{aligned}
\left| \int (\varphi_\varepsilon - \varphi_0)\psi dm \right| &= \left| \int \psi d\mu_\varepsilon - \int \psi d\mu_0 \right| \\
&\leqslant \left| \int \psi d\mu_\varepsilon - \int \psi(\widehat{\mathcal{L}}_\varepsilon^n 1) dm \right| + \left| \int \psi(\mathcal{L}^n 1) dm - \int \psi d\mu_0 \right| \\
&\quad + \left| \int (\widehat{\mathcal{L}}_\varepsilon^n 1)\psi dm - \int (\mathcal{L}^n 1)\psi dm \right| \\
&\leqslant (2R\Lambda_1^n + Kn\xi(\varepsilon))\|\psi\|_1
\end{aligned}$$

para todo $n \geqslant 0, \varepsilon > 0$ pequeño y $\psi \in L^1(m)$. Por tanto

$$\|\varphi_\varepsilon - \varphi_0\|_0 \leqslant (2R\Lambda_1^n + Kn\xi(\varepsilon))$$

para todo $n \geqslant 0$. Finalmente fijamos $n \geqslant 0$ con $\Lambda_1^{n+1} \leqslant \xi(\varepsilon) \leqslant \Lambda_1^n$ para obtener la estimación

$$\|\varphi_\varepsilon - \varphi_0\|_0 \leqslant K'\xi(\varepsilon)\log(\xi(\varepsilon))$$

donde $K' > 0$ sólo depende de f y $\xi(\varepsilon)\log(\xi(\varepsilon))$ converge a cero cuando $\varepsilon \to 0$. El Teorema está demostrado. ■

Bibliografía

[A] R. Ash. *Probability Measure Theory.* Second edition. Academic Press. Boston (2000).

[DB] D. Branhan. *Geometry.* Cambridge University Press (2000).

[RD] R.M. Dudley. *Real Analysis and Probability.* Cambridge University Press, Boston (2002).

[E] R. Ellis. *Entropy, Large Deviations and Statistical Mechanics.* Reprint of the 1985 edition. Springer (2000).

[FS] P. Ferrero & B. Schmitt. *Ruelle's Perron-Frobenius theorem and projective metrics.* Coll. Math. Soc. János Bolyai, **27** (1979) pp. 333-336.

[Ke] G. Keller. *The transfer operator approach to interval maps.* Preprint en:

`www.mi.uni-erlangen.de/%7EKeller/gklit.html`

[K1] Yu. Kifer. *Ergodic theory of random transformations.* Birkhauser. Boston, 1986.

[K2] Yu. Kifer. *Random perturbations of dynamical systems.* Progress in Probability and Statistics, Vol 16, Birkhauser. Boston, 1988.

[K3] Yu. Kifer. *General random perturbations of hyperbolic and expanding transformations.* Journal D'Analyse Math. **47** (1986), 111–150.

[KS] K. Kryzewski & Z. Slenk. *On invariant measures for expanding differentiable mappings.* Studia Math. **33** (1969), 83–92.

[LM] A. Lasota M. C. Mackay. *Chaos Fractals and Noise, Stochastic aspects of dynamics.* Springer-Verlag. New York, 1995.

[L1] C. Liverani. *Decay of correlations.* Annals of Math. **142** (1995), 239–301.

[L2] C. Liverani. *Decay of correlations for piecewise expanding maps.* J. Stat. Phys. **78** (1995). 1111–1129.

[RM] R. Mañe. *Introducão á teoria ergódica.* IMPA, Rio de Janeiro 1983.

[R] D. Ruelle. *The thermodynamical formalism for expanding maps.* Comm. Math. Phys. **125** (1989). 239–262.

[V1] M. Viana. *Stochastic dynamics of deterministic systems.* Brazillian Coll. of Math. (1998). IMPA, Rio de Janeiro.

[W] D. Williams. *Probability with Martingales.* Cambridge University Press. Melbourne (1991).

[Y1] L.-S. Young. *Stochastic stability of hyperbolic attractors.* Ergodic Thoery & Dynamical Systems. **6** (1986), 311–319.

[Y2] L.-S. Young. *On the ergodic theory of dynamical systems.* in: From Topology to Computation: Proceedings of the Smalefest. M. W. Hirsch, J. E. Marsden and M. Shub eds. Springer Verlag. New York. 1990.

[Y3] L.-S. Young. *What are SRB measures, and which dynamical systems have them?* Preprint en: `www.cims.nyu.edu/~lsy/papres/srbSurvey.pdf`

Angel E. Terán González

Producido por el autor usando Texmaker, Latex, Ubuntu.

Barquisimeto, Mayo 2014