

Este libro está destinado a un curso semestral de postgrado en Matemática, introductorio a la Teoría Ergódica.



La exposición está enfocada desde el punto de vista de los sistemas dinámicos, pero su lectura no requiere conocimientos previos en ese tema.

Los primeros capítulos cubren la base de la teoría: existencia de medidas invariantes y ergódicas, el teorema ergódico de Birkhoff-Khinchin, enunciados equivalentes a la ergodicidad, definición de exponentes de Lyapunov y ejemplos.

En los capítulos centrales se incluyen los teoremas ergódicos subaditivo de Kingmann y multiplicativo de Oseledets, la teoría de las medidas Sinai-Ruelle-Bowen, de atractores estadísticos de Ilyashenko, y las relaciones entre estos y las medidas físicas o pseudo-físicas.

Los capítulos finales exponen la teoría espectral, el espacio del shift, medidas de Bernoulli, y entropía métrica y topológica.

ISBN: 978-9974-0-0949-3

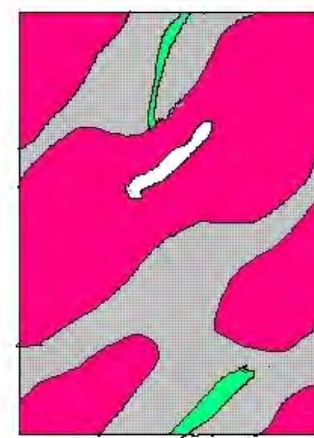
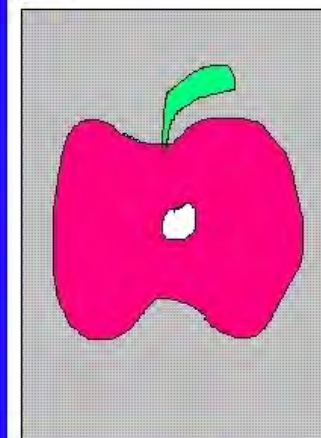


9 789974 009493

TEORÍA ERGÓDICA DE LOS SISTEMAS DINÁMICOS DISCRETOS - E. Catsigeras

TEORÍA ERGÓDICA DE LOS SISTEMAS DINÁMICOS DISCRETOS

Eleonora Catsigeras



Universidad
de la República
Uruguay

TEORÍA ERGÓDICA
DE LOS
SISTEMAS DINÁMICOS
DISCRETOS

ELEONORA CATSIGERAS

Teoría Ergódica de los Sistemas Dinámicos Discretos
Eleonora Catsigeras

© E. Catsigeras - FING - UdelaR - 2013

2010 Mathematics Subject Classification **3701**, 37A, 37D, 37C40

Palabras clave: Teoría Ergódica, Sistemas Dinámicos,
Medidas invariantes, Atractores, Medidas SRB.

Edición: Universidad de la República
Av. 18 de Julio 1968 - Montevideo - Uruguay
tel +598-2400-9201

Co-edición e impresión: M.D. Editora Reg. Edit. 2597
MirnaDominguezEditoraATTgmail.com

ISBN 978-9974-0-0949-3 (papel)

ISBN 978-9974-0-0950-9 (electrónico)

Publicación electrónica: www.fing.edu.uy
Attribution-NonCommercial-ShareAlike 3.0
CC BY-NC-SA 3.0 US

Puede reproducirse exclusivamente para fines legales no comerciales
y con atribución explícita a datos de autor y edición.

ISBN: 978-9974-0-0949-3



Impreso artesanal por reproducción digital
Mirna Domínguez - Juan Paullier 1466 - Montevideo
Depósito legal 361.992
26 de junio de 2013

A Heber

Prólogo

Este libro está destinado a un curso semestral de postgrado en Matemática, introductorio a la Teoría Ergódica de los Sistemas Dinámicos Discretos. Está organizado para que partes del libro (por ejemplo los capítulos 1 a 3, 6 y 7) puedan ser utilizadas también en cursos de las carreras de grado en Matemática y de postgrado en Física e Ingeniería Matemática.

Aunque la exposición del libro está enfocada desde el punto de vista de los sistemas dinámicos discretos y deterministas, su lectura no requiere conocimientos previos en esta materia. Las definiciones y resultados básicos en este tema están contenidos en la exposición. Requiere en cambio algunos conocimientos previos (básicos) sobre topología de espacios métricos, sobre la teoría de la medida e integración abstractas, y sobre geometría diferencial. Estos prerrequisitos están incluidos en los cursos curriculares de la carrera de grado en Matemática.

Los capítulos 1 y 2 introducen los conceptos y primeros resultados básicos de la teoría ergódica de los sistemas dinámicos discretos y de dinámica topológica: definición y existencia de medidas invariantes y ergódicas, recurrencia y transitividad, el teorema ergódico de Birkhoff-Khinchin, y las propiedades medibles que caracterizan la ergodicidad.

El capítulo 3, introduce la dinámica diferenciable. Expone, esencialmente mediante ejemplos, algunos resultados sobre sistemas dinámicos hiperbólicos.

En el capítulo 4 se enuncian y demuestran otros teoremas ergódicos, cuyas pruebas son de mayor complejidad que las de los capítulos anteriores: el teorema subaditivo de Kingmann, el teorema multiplicativo de Oseledets, y el teorema de desintegración ergódica.

Los capítulos 5, 6 y 7 tratan la teoría abstracta de atractores topológicos y estadísticos de sistemas dinámicos en variedades riemannianas. Se introduce las medidas Sinai-Ruelle-Bowen (SRB) o físicas, se demuestra su existencia para sistemas diferenciables hiperbólicos con suficiente regularidad, se define las medidas de Gibbs, y se incluye una reseña breve de la Teoría de Pesin. El capítulo 7 estudia los atractores en sistemas continuos, no necesariamente diferenciables, demostrando para ellos la existencia de

atractores topológicos de Milnor y estadísticos de Ilyashenko, y la relación entre estos y las medidas “SRB-like” o pseudo-físicas.

El capítulo 8 expone los conceptos y resultados básicos de la teoría espectral desde el punto de vista de la teoría ergódica, la jerarquía ergódica, la dinámica del shift y las medidas de Bernoulli.

El capítulo 9 trata las definiciones y propiedades básicas de la entropía métrica y topológica, incluyendo las pruebas del Teorema de Kolmogorov-Sinai y del Principio Variacional de la Entropía.

Finalmente, en el apéndice se incluye el enunciado y la prueba del Lema de Pliss, y su uso para demostrar el Teorema Ergódico Multiplicativo de Oseledets.

Eleonora Catsigeras¹

¹Instituto de Matemática y Estadística “Prof. Ing. Rafael Laguardia”, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay.
La autora agradece la financiación parcial de la Comisión Sectorial de Investigación Científica (CSIC) de la Universidad de la República, y de la Agencia Nacional de Investigación e Innovación (ANII) de Uruguay

Índice general

1. Dinámica medible y topológica	1
1.1. Dinámica de los automorfismos de medida.	1
1.2. Prueba de existencia de medidas invariantes	3
1.3. Ejemplos y puntos periódicos hiperbólicos	9
1.4. Dinámica topológica	13
1.5. Recurrencia y Lema de Poincaré	15
1.6. Ergodicidad	16
1.7. Existencia de medidas ergódicas	21
2. Teoremas Ergódicos.	27
2.1. Teorema Ergódico de Birkhoff-Khinchin Enunciado y Corolarios	27
2.2. Teorema Ergódico Maximal	32
2.3. Prueba del Teorema Ergódico de Birkhoff-Khinchin	38
2.4. Otras caracterizaciones de la ergodicidad	41
2.5. Ergodicidad Única	44
2.6. Ergodicidad de la rotación irracional	48
2.7. Transformaciones Mixing.	51
2.8. Medida de Haar	54
3. Dinámica diferenciable:	
Hiperbolicidad uniforme y no uniforme	57
3.1. Ejemplo de automorfismo lineal hiperbólico en el toro. . .	57
3.2. Ergodicidad del automorfismo lineal hiperbólico en \mathbb{T}^2 . .	63
3.3. Difeomorfismos de Anosov e hiperbolicidad uniforme . .	67
3.4. Conjuntos uniformemente hiperbólicos	70
3.5. Ejemplo: Herradura de Smale lineal	70

3.6.	Variedades invariantes de conjuntos hiperbólicos	72
3.7.	Expansividad o caos topológico.	74
3.8.	Foliaciones invariantes para dif. de Anosov.	75
3.9.	Exponentes de Lyapunov	77
3.10.	Hiperbolicidad no uniforme	81
3.11.	Región de Pesin y medidas hiperbólicas	85
3.12.	Variedades estable e inestable en la región de Pesin	87
4.	Teoremas Ergódicos:	
	Descomposición Ergódica, Subaditivo y Multiplicativo	89
4.1.	Descomposición Ergódica	89
4.2.	Teorema Ergódico Subaditivo de Kingmann	102
4.3.	Crecimiento subexponencial	116
4.4.	Generadores de exponentes de Lyapunov	118
4.5.	Primer subfibrado de Oseledets	131
4.6.	Teorema Ergódico Multiplicativo (Oseledets)	134
5.	Atractores topológicos y ergódicos	139
5.1.	Atractores topológicos	139
5.2.	Cuenca global de atracción topológica	146
5.3.	Atractores hiperbólicos caóticos	148
5.4.	Atractores ergódicos	154
5.5.	Atracción estadística y medidas SRB o físicas	160
6.	Teoría de Pesin	163
6.1.	Desintegración en medidas condicionales	163
6.2.	Medidas de Gibbs	167
6.3.	Relación entre medidas de Gibbs y SRB	171
6.4.	Sobre la Fórmula de Pesin para la entropía	175
6.5.	Demostración del Teorema 6.3.1	178
6.6.	Lema de Distorsión Acotada	181
6.7.	Demostración del Teorema 6.3.4	185
7.	Atractores estadísticos y medidas SRB-like	199
7.1.	Atractores de Milnor	199
7.2.	Atractores estadísticos o de Ilyashenko	204
7.3.	Atracción estadística de un compacto	204
7.4.	Existencia de atractores de Ilyashenko	209

7.5. Ejemplos de atractores estadísticos	212
7.6. Medidas SRB-like o pseudo-físicas	216
7.7. Relación entre atractor estadístico y medidas SRB-like	221
7.8. Ejemplos de medidas SRB-like	224
8. Espectro de Lebesgue y	
Medidas de Bernoulli	229
8.1. El operador U_T en L^2	229
8.2. Equivalencia espectral	231
8.3. Bases de Fourier en $L^2(m)$	232
8.4. Rotación irracional revisitada	233
8.5. Espectro de Lebesgue	235
8.6. Automorfismos lineales en el toro con espectro de Lebesgue.	236
8.7. El espacio del shift y su topología.	237
8.8. Espacio de medida del shift de Bernoulli	240
8.9. Ergodicidad del shift de Bernoulli.	241
8.10. Transformaciones de Bernoulli	245
9. Entropía métrica y topológica	249
9.1. Entropía de particiones	251
9.2. Entropía de partición como función de la medida	256
9.3. Refinamiento de particiones	260
9.4. Entropía de una transformación respecto a una partición	265
9.5. Entropía métrica de una transformación	269
9.6. Entropía topológica.	274
9.7. Ejemplos de cálculo de la entropía topológica	279
9.8. Prueba del Principio Variacional de la Entropía	283
10. Apéndice	
Existencia del primer subespacio de Oseledets	293
10.1. Demostración del Teorema 4.5.1: Primer Subespacio de Oseledets	293
10.2. Lema de Pliss y consecuencias	306

Capítulo 1

Dinámica medible y topológica

Usaremos definiciones, notación y algunos resultados básicos de teoría de la medida abstracta y de la topología en espacios métricos compactos, cuyos enunciados y demostraciones se pueden encontrar por ejemplo en [Fo 1984], [Rud 1979a] o [Ste 2005].

1.1. Dinámica de los automorfismos de medida.

Sea (X, \mathcal{A}) un espacio medible y sea $T : X \mapsto X$ una transformación medible, es decir $T^{-1}(A) \in \mathcal{A} \ \forall A \in \mathcal{A}$.

Definición 1.1.1. Sistema dinámico discreto

Se llama *sistema dinámico discreto* por iterados de T hacia el futuro a la aplicación que a cada $n \in \mathbb{N}$ le hace corresponder la transformación $T^n : X \mapsto X$ donde $T^n := T \circ T \dots \circ T$ n veces, si $n \geq 1$ y $T^0 := id$. Es inmediato chequear que T^n es medible para todo $n \in \mathbb{N}$.

Si además T es bi-medible (i.e. $T : X \mapsto X$ es medible, invertible y su inversa $T^{-1} : X \mapsto X$ es medible), entonces *el sistema dinámico discreto*, por iterados de T hacia el futuro y el pasado, se define como la aplicación que a cada $n \in \mathbb{Z}$ le hace corresponder T^n , donde T^{-1} es la inversa de T y $T^n := (T^{-1})^{|n|} := T^{-1} \circ T^{-1} \circ \dots \circ T^{-1}$ $|n|$ veces si $n \leq -1$.

Se observa que si $T^{n+m} = T^n \circ T^m \ \forall n, m \in \mathbb{N}$, y además si T es bi-medible se cumple esa propiedad para todos $n, m \in \mathbb{Z}$. Esta propiedad algebraica (cuando T es bi-medible) se llama *propiedad de grupo*. Significa que el sistema dinámico es una acción del grupo de los enteros en el espacio X .

Definición 1.1.2. Se llama *órbita positiva o futura* $o^+(x)$ por el punto $x \in X$, y x se llama estado inicial de la órbita, a la sucesión

$$o^+(x) := \{T^n(x)\}_{n \in \mathbb{N}}.$$

Si T es bi-medible se llama órbita negativa o pasada a la sucesión

$$o^-(x) := \{T^{-n}(x)\}_{n \in \mathbb{N}}$$

y se llama *órbita $o(x)$ bilateral* (o simplemente órbita cuando T es bi-medible) a la sucesión bi-infinita

$$o(x) := \{T^n(x)\}_{n \in \mathbb{Z}}.$$

Cuando T es medible pero no bi-medible, llamamos simplemente órbita a la órbita positiva o futura.

Se observa que $T^{n+m}(x) = T^n(T^m(x))$ y por lo tanto el iterado $T^m(x)$ es el estado inicial de la órbita por x que se obtiene corriendo el instante 0 al que antes era instante m .

Definición 1.1.3. *Punto fijo* (o periódico de período 1) es x_0 tal que $T(x_0) = x_0$. *Punto periódico* es x_0 tal que existe $p \geq 1$ tal que $T^p(x_0) = x_0$. El período es el mínimo $p \geq 1$ que cumple lo anterior. Se observa, a partir de la propiedad de grupo, que si un punto es periódico de período p entonces su órbita está formada por exactamente p puntos.

Definición 1.1.4. Sea T medible en un espacio (X, \mathcal{A}) . Una medida μ se dice que es *invariante por T* , o que *T preserva μ* , o se dice también que T es un *automorfismo del espacio de medida* (X, \mathcal{A}, μ) , si

$$\mu(T^{-1}(A)) = \mu(A) \quad \forall A \in \mathcal{A}.$$

Se observa que pueden no existir medidas de probabilidad invariantes para cierta $T : X \mapsto X$ transformación medible dada, como se muestra en el Ejemplo 1.1.7. Sin embargo:

Teorema 1.1.5. Existencia de medidas invariantes.

Sea X un espacio métrico compacto, y sea $T : X \mapsto X$ continua. Entonces existen (usualmente infinitas) medidas de probabilidad en la sigma-álgebra de Borel, que son invariantes para T .

Demostraremos este teorema en la siguiente sección 1.2.

Ejercicio 1.1.6. Sean (X, \mathcal{A}, μ) e (Y, \mathcal{B}, ν) u espacios de medida y sea $T : X \mapsto Y$ medible. Se define $T^*\mu$ como la medida en (Y, \mathcal{B}) tal que $(T^*\mu)(B) := \mu(T^{-1}(B)) = \nu(B)$ para todo $B \in \mathcal{B}$.

(a) Encontrar un ejemplo en que $T^*\mu = \nu$ pero $\mu(A) \neq \nu(T(A))$ para algún $A \in \mathcal{A}$ tal que $T(A) \in \mathcal{B}$.

(b) Encontrar un ejemplo en que T sea medible, cumpla $T^*\mu = \nu$ pero T^{-1} no sea medible.

(c) Demostrar que si T es medible, invertible y su inversa T^{-1} es medible, entonces $T^*\mu = \nu$ si y solo si $(T^{-1})^*\nu = \mu$. Cuando se cumplen todas esas condiciones, se dice que T (y por lo tanto también T^{-1}) es un isomorfismo de espacios de medida.

Ejemplo 1.1.7. Sea $T : [0, 1] \mapsto [0, 1]$ tal que $T(x) = x/2$ si $x \neq 0$ y $T(0) = 1 \neq 0$. Afirmamos que:

No existe medida de probabilidad invariante por T .

Demostración: Si existiera, llamémosla μ . Consideremos la partición de intervalo $(0, 1]$ dada por los subintervalos disjuntos dos a dos $A_n = (1/2^{n+1}, 1/2^n]$ para $n \in \mathbb{N}$. Se cumple $T^{-1}(A_n) = A_{n-1}$ para todo $n \geq 1$. Como μ es T invariante se deduce $\mu(A_n) = \mu(A_0) \forall n \geq 0$. Se tiene $\mu((0, 1]) = \sum_{n \geq 0} \mu(A_n) = \sum_{n \geq 0} \mu(A_0) \leq 1$. Luego $\mu(A_0) = 0$, de donde $\mu((0, 1]) = 0$ y $\mu(\{0\}) = 1$. Luego $\mu(T^{-1}(\{0\})) = 1$ lo cual es absurdo porque $T^{-1}(\{0\}) = \emptyset$. \square

Ejercicio 1.1.8. Probar que si $T : X \mapsto X$ es medible entonces T preserva la medida μ si y solo si para toda $f \in L^1(\mu)$ se cumple

$$\int f \circ T d\mu = \int f d\mu$$

Proposición 1.1.9. Sea $T : X \mapsto X$ una transformación medible en un espacio medible (X, \mathcal{A}) . Si μ es una medida finita o sigma-finita y \mathcal{A}_0 es un álgebra que genera a \mathcal{A} tal que $\mu(T^{-1}(A)) = \mu(A) \forall A \in \mathcal{A}_0$ entonces μ es invariante por T .

Demostración: Sea $\nu(A) = \mu(T^{-1}(A))$ definida para todo $A \in \mathcal{A}$. En la subálgebra \mathcal{A}_0 la premedida ν coincide con la premedida μ (ambas restringidas a la subálgebra son premedidas). Como existe una única extensión de una premedida dada en un álgebra a la sigma álgebra generada, entonces $\mu = \nu$ en \mathcal{A} . \square

Corolario 1.1.10. Si $T : \mathbb{R}^k \mapsto \mathbb{R}^k$ es una transformación Borel medible y μ es una medida σ -finita tal para todo conjunto A que sea unión finita de rectángulos de \mathbb{R}^k se cumple $\mu(T^{-1}(A)) = \mu(A)$ entonces μ es invariante por T en toda la sigma-álgebra de Borel.

1.2. Prueba de existencia de medidas invariantes

En esta sección X denota un espacio métrico compacto y $T : X \mapsto X$ una transformación continua, a menos que se indique lo contrario.

Introducimos algunas definiciones y resultados del Análisis Funcional:

Notación: El espacio de las funciones continuas y su dual.

Denotamos $C^0(X, \mathbb{R})$ el espacio de las funciones continuas $\psi : X \mapsto \mathbb{R}$ con la topología de la convergencia uniforme en X (inducida por la norma del supremo). Es decir

$$\text{dist}(\psi_1, \psi_2) := \|\psi_1 - \psi_2\|_0,$$

donde se denota

$$\|\psi\|_0 := \max_{x \in X} |\psi(x)| \quad \forall \psi \in C^0(X, \mathbb{R}).$$

Denotamos $C^0(X, [0, 1])$ al subespacio de funciones continuas $\psi : X \mapsto [0, 1]$ que toman valores no negativos ni mayores que 1, con la norma del supremo definida antes (o del máximo, en nuestro caso, pues X es un espacio métrico compacto).

El espacio $C^0(X, [0, 1])$ es métrico acotado y cerrado. En efecto, el límite de una sucesión uniformemente convergente de funciones continuas en $C^0(X, [0, 1])$ es continua, y pertenece a $C^0(X, [0, 1])$. Además $C^0(X, [0, 1])$ tiene una base numerable de abiertos, y por lo tanto existe un subconjunto numerable $\{\psi_i\}_{i \in \mathbb{N}}$ denso en $C^0(X, [0, 1])$ (ver por ejemplo [Fo 1984, Proposition 4.40])

Denotamos como $C^0(X, \mathbb{R})^*$ al dual de $C^0(X, \mathbb{R})$; es decir al conjunto (al que luego dotaremos de una topología adecuada) de todos los operadores lineales

$$\Lambda : C^0(X, \mathbb{R}) \mapsto \mathbb{R}.$$

Definición 1.2.1. El espacio \mathcal{M} de las probabilidades y la topología débil*

Sea X un espacio métrico compacto. Sea \mathcal{M} el espacio de todas las medidas de probabilidad.

En \mathcal{M} introducimos la siguiente topología, llamada *topología débil**:

Si $\mu_n, \mu \in \mathcal{M}$ decimos que la sucesión $\{\mu_n\}$ es convergente a μ en la topología débil*, es decir

$$\lim_{n \rightarrow +\infty} \mu_n = \mu \quad \text{en } \mathcal{M},$$

cuando:

$$\lim_{n \rightarrow +\infty} \int \psi d\mu_n = \int \psi d\mu \quad \text{en } \mathbb{R} \quad \forall \psi \in C^0(X, \mathbb{R}).$$

Observación 1.2.2. Por definición de convergencia en la topología débil* de una sucesión de medidas de probabilidad, la sucesión converge a la medida μ si y solo si para cada función continua, la sucesión de integrales converge a la integral respecto a μ . Es falso que para todo A boreliano la sucesión de medidas de A converja a $\mu(A)$. En efecto, véase el ejercicio siguiente:

Ejercicio 1.2.3. Sea $X = [0, 1]$. Para cada $n \geq 0$ sea μ_n la medida delta de Dirac concentrada en el punto $1/2^n$.

a) Probar que existe $\mu = \lim_{n \rightarrow +\infty} \mu_n$ en la topología débil* y encontrar la medida límite μ .

b) Encontrar $A \subset [0, 1]$ boreliano tal que no existe $\lim_{n \rightarrow +\infty} \mu_n(A)$.

Sugerencia: $A = \{1/2^{2^j} : j \geq 0\}$.

c) Encontrar $B \subset [0, 1]$ boreliano tal que existe $\lim_{n \rightarrow +\infty} (\mu_n(B)) \neq \mu(B)$.

Observación 1.2.4. Observemos que, para cada $\mu \in \mathcal{M}$, el operador $\Lambda_\mu : C^0(X, \mathbb{R}) \mapsto \mathbb{R}$ definido por:

$$\Lambda_\mu(\psi) := \int \psi d\mu$$

es lineal, positivo (es decir $\Lambda_\mu(\psi) \geq 0$ si $\psi \geq 0$) y acotado (es decir existe $k > 0$ tal que $|\Lambda_\mu(\psi)| \leq k$ para toda $\psi \in C^0(X, \mathbb{R})$ tal que $\|\psi\|_0 \leq 1$). Además $\Lambda_\mu(\psi) = 1$ si $\psi : X \mapsto \mathbb{R}$ es la función constante igual a 1.

El siguiente teorema establece el recíproco de la propiedad observada arriba, y es un resultado clásico de la Teoría Abstracta de la Medida y del Análisis Funcional:

Teorema de Representación de Riesz

Sea X un espacio métrico compacto.

Para todo operador lineal $\Lambda : C^0(X, \mathbb{R}) \mapsto \mathbb{R}$ que sea positivo y acotado, existe y es única una medida finita μ (de Borel y positiva) tal que

$$\Lambda(\psi) = \int \psi d\mu \quad \forall \psi \in C^0(X, \mathbb{R})$$

Además si $\Lambda(1) = 1$ entonces μ es una probabilidad. Es decir, $\mu(X) = 1$, ó, usando nuestra notación, $\mu \in \mathcal{M}$.

Una demostración del Teorema de Representación de Riesz puede encontrarse, por ejemplo, en [Rud 1979a, Teorema 2.3.1]

Debido al Teorema de Representación de Riesz, el espacio \mathcal{M} se puede identificar con el espacio de los operadores lineales de $C^0(X, \mathbb{R})$ que son positivos, acotados y que valen 1 para la función constante igual a 1. (Recordemos que el espacio de los operadores lineales de $C^0(X, \mathbb{R})$ se llama dual de $C^0(X, \mathbb{R})$ y se denota como $C^0(X, \mathbb{R})^*$). En el Análisis Funcional se definen diversas topologías en el dual de un espacio de funciones. Una de ellas es la llamada *topología débil**, que es la topología de la convergencia *punto a punto*, definida como sigue:

$$\lim_{n \rightarrow +\infty} \Lambda_n = \Lambda \text{ en } C^0(X, \mathbb{R})^*$$

si y solo si

$$\lim_{n \rightarrow +\infty} \Lambda_n \psi = \Lambda \psi \text{ en } \mathbb{R} \quad \forall \psi \in C^0(X, \mathbb{R}).$$

De las definiciones anteriores, deducimos que la topología débil estrella en \mathcal{M} es la topología inducida por la topología débil estrella (o de la convergencia punto a punto) en el dual $C^0(X, \mathbb{R})^*$ del espacio funcional de las funciones continuas. La topología débil* puede definirse como la topología producto de las definidas por la convergencia de los valores numéricos $\Lambda_n(\psi)$ que toman los operadores Λ_n para cada $\psi \in C^0(X, \mathbb{R})$ fija.

El siguiente teorema es clásico del Análisis Funcional, y es una consecuencia del teorema de Tichonov (ver por ejemplo [Fo 1984, Teorema 4.43]) que establece, bajo ciertas hipótesis, la compacidad de la topología producto:

Teorema 1.2.5. (Corolario del Teorema de Tichonov) *Si X es un espacio métrico compacto, entonces para toda constante $k > 0$ el subconjunto de los operadores lineales acotados por k es compacto en el espacio dual $C^0(X, \mathbb{R})^*$ con la topología débil estrella.*

Como caso particular, observemos que el espacio \mathcal{M} de las probabilidades de Borel en X (via el teorema de Representación de Riesz) es un subconjunto *cerrado* del espacio de los operadores lineales acotados con la topología débil estrella (es decir, si $\Lambda_n(1) = 1$ y $\Lambda_n \rightarrow \Lambda$, entonces $\Lambda(1) = 1$).

Más detalladamente:

Teorema 1.2.6. Compacidad y metrizabilidad del espacio de probabilidades

Sea X un espacio métrico compacto. Sea \mathcal{M}^1 el espacio de todas las medidas μ (de Borel, positivas y finitas) tales que $\mu(X) \leq 1$. Sea en \mathcal{M}^1 la topología débil*, definida en 1.2.1. Entonces:

(a) \mathcal{M}^1 es compacto.

(b) \mathcal{M}^1 es metrizable. (Es decir, existe una métrica dist que induce la topología débil*; esto es una distancia en $\mathcal{M}^1 \times \mathcal{M}^1$ tal que

$$\lim_n \text{dist}(\mu_n, \mu) = 0 \text{ si y solo si } \lim_n \mu_n = \mu$$

con la topología débil*).

(c) Si $\{\psi_i\}_{i \in \mathbb{N}} \subset C^0(X, [0, 1])$ es un conjunto numerable denso de funciones, entonces la siguiente métrica induce la topología débil* en \mathcal{M}^1 :

$$\text{dist}(\mu, \nu) := \sum_{i=1}^{+\infty} \frac{1}{2^i} \left| \int \psi_i d\mu - \int \psi_i d\nu \right| \quad \forall \mu, \nu \in \mathcal{M}^1. \quad (1.1)$$

(d) \mathcal{M}^1 es secuencialmente compacto. (Es decir, toda sucesión de medidas en \mathcal{M}^1 tiene alguna subsucesión convergente).

Consecuencia: Siendo $\mathcal{M} = \{\mu \in \mathcal{M}^1 : \mu(X) = 1\}$ cerrado en \mathcal{M}^1 , se deduce que

(e) El espacio de probabilidades \mathcal{M} con la topología débil* es compacto, metrizable y secuencialmente compacto.

Ejercicio 1.2.7. Demostrar el Teorema 1.2.6 como consecuencia del Corolario 1.2.5, identificando el espacio \mathcal{M}^1 con el dual de $C^0(X, \mathbb{R})$ via el Teorema de Representación de Riesz.

Proposición 1.2.8. Sea $\{\psi_i : i \geq 1\}$ un conjunto numerable denso en $C^0(X, [0, 1])$. Dos medidas μ_1 y μ_2 en $\mathcal{M}^1(X)$ coinciden si para todo $i \geq 1$ se cumple

$$\int \psi_i d\mu_1 = \int \psi_i d\mu_2$$

Demostración: Por la unicidad de la medida del teorema de Riesz alcanza probar que para toda $\psi \in C^0(X, \mathbb{R})$ se cumple:

$$\int \psi d\mu_1 = \int \psi d\mu_2 \quad (1.2)$$

La igualdad (1.2) vale obviamente para ψ idénticamente nula. Si ψ no es idénticamente nula, basta demostrar (1.2) para $\psi/\|\psi\|_0$, donde $\|\psi\|_0 = \max_{x \in X} |\psi(x)|$. Entonces supongamos que $\|\psi\|_0 = 1$. Cualquier función real ψ puede escribirse como $\psi = \psi^+ - \psi^-$, donde $\psi^+ = \max\{\psi, 0\}$, $\psi^- = -\min\{\psi, 0\}$. Observemos que $\psi^+, \psi^- \in C^0([0, 1])$. Si demostramos la igualdad (1.2) para ψ^+ y ψ^- , entonces vale también para ψ . Basta entonces probar la igualdad (1.2) para toda $\psi \in C^0(X, [0, 1])$. Por la densidad de las

funciones $\{\psi_i\}$ en $C^0(X, [0, 1])$, existe una sucesión ψ_{i_n} convergente uniformemente (es decir con la norma del supremo en $C^0(X, \mathbb{R})$), a la función ψ . Por lo tanto, converge también puntualmente y está uniformemente acotada por 1. Cada ψ_{i_n} verifica la igualdad (1.2). Luego, por el teorema de convergencia dominada, ψ también cumple (1.2). \square

Observación 1.2.9. Sea X un espacio métrico compacto no vacío, sea \mathcal{B} la sigma-álgebra de Borel, y sea \mathcal{M} el espacio de todas las medidas de probabilidad de Borel en X con la topología débil*.

El espacio \mathcal{M} es no vacío. Por ejemplo, si elegimos un punto $x \in X$ entonces $\delta_x \in \mathcal{M}$, donde δ_x es la probabilidad Delta de Dirac soportada en el punto $x \in X$. Esto es δ_x es la probabilidad que a cada boreliano $B \subset X$ le asigna $\delta_x(B) = 1$ si $x \in B$, y $\delta_x(B) = 0$ si $x \notin B$.

Ahora agreguemos una dinámica en X :

Definición 1.2.10. El pull back $T^* : \mathcal{M} \mapsto \mathcal{M}$

Sea (X, \mathcal{A}) un espacio medible, sea $T : X \mapsto X$ medible y sea \mathcal{M} el conjunto de las medidas de probabilidad en (X, \mathcal{A}) . Definimos el siguiente operador en \mathcal{M} , llamado pull back del mapa T :

$$T^* : \mathcal{M} \mapsto \mathcal{M} \quad (T^*\mu)(B) = \mu(T^{-1}(B)) \quad \forall \mu \in \mathcal{M} \quad \forall B \in \mathcal{B}.$$

Es inmediato verificar que μ es invariante por T si y solo si $T^*\mu = \mu$, es decir, las medidas invariantes por T son los puntos fijos por T^* en el espacio \mathcal{M} .

Ejercicio 1.2.11. Probar que para toda $\psi \in L^1(\mu)$ se cumple:

$$\int \psi dT^*\mu = \int \psi \circ T d\mu.$$

Sugerencia: Chequear primero para las funciones características χ_B de los borelianos, luego para las combinaciones lineales de las funciones características (funciones simples), y luego para las funciones medibles no negativas, usando el Teorema de convergencia monótona. Finalmente probar la igualdad para toda $\psi \in L^1(\mu)$ separando ψ en parte real e imaginaria, y cada ψ real en su parte positiva y negativa.

Proposición 1.2.12. Sea X espacio métrico compacto, y \mathcal{M} el espacio de las probabilidades de Borel en X con la topología débil*. Si $T : X \mapsto X$ es continuo, entonces $T^* : \mathcal{M} \mapsto \mathcal{M}$ es continuo en \mathcal{M} .

Demostración. Basta chequear que si $\lim_n \mu_n = \mu$ en \mathcal{M} entonces $\lim_n T^*\mu_n = T^*\mu$. En efecto, como $\lim_n \mu_n = \mu$, entonces, para toda $\psi \in C^0(X, \mathbb{R})$:

$$\lim_{n \rightarrow +\infty} \int \psi d\mu_n = \int \psi d\mu$$

En particular, siendo $T : X \mapsto X$ continuo, la igualdad anterior se cumple para $\psi \circ T$. Luego deducimos que

$$\lim_{n \rightarrow +\infty} \int \psi \circ T d\mu_n = \int \psi \circ T d\mu.$$

Usando el resultado del Ejercicio 1.2.11, deducimos que

$$\lim_{n \rightarrow +\infty} \int \psi dT^* \mu_n = \int \psi dT^* \mu \quad \forall \psi \in C^0(X, \mathbb{R}).$$

Por lo tanto la sucesión de medidas $T^* \mu_n$ converge en la topología débil* de \mathcal{M} , a la medida $T^* \mu$ como queríamos demostrar. \square

Ahora probaremos el Teorema 1.1.5, utilizando el llamado *procedimiento de Bogliubov-Krylov* [Bog-Kry 1937]. Este procedimiento parte de cualquier medida de probabilidad en el espacio X , toma promedios aritméticos de los iterados del operador pull back T^* de esta medida hasta tiempo n , y finalmente una subsucesión convergente en la topología débil estrella de esos promedios. Se obtienen medidas invariantes por T (bajo la hipótesis de que $T : X \mapsto X$ es continuo). El procedimiento de Bogliubov-Krylov permite “fabricar” medidas de probabilidad invariantes, usando como “semilla” cualquier medida de probabilidad.

Demostración del Teorema 1.1.5 de existencia de medidas invariantes:

Demostración. Elijamos una medida de probabilidad de Borel cualquiera $\rho \in \mathcal{M}$. Construyamos para cada $1 \leq n \in \mathbb{N}$, la siguiente probabilidad:

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} (T^*)^j \rho$$

Es inmediato probar, a partir de la definición del operador T^* (que cumple $T^* \mu(B) = \mu(T^{-1}(B))$ para todo boreliano B), que

$$T^* \mu_n = \frac{1}{n} \sum_{j=0}^{n-1} (T^*)^{j+1} \rho$$

Como el espacio \mathcal{M} de las probabilidades de Borel es secuencialmente compacto con la topología débil*, existe una subsucesión $\{\mu_{n_i}\}_{i \in \mathbb{N}}$ (con $\lim_i n_i = +\infty$), que es convergente en \mathcal{M} . Llamemos $\mu \in \mathcal{M}$ a su límite; es decir:

$$\mu = \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} (T^*)^j \rho$$

Basta demostrar ahora que $T^* \mu = \mu$, es decir μ es una probabilidad T -invariante. Usando la continuidad del operador T^* (Proposición 1.2.12), deducimos que

$$T^* \mu = \lim_{i \rightarrow +\infty} T^* \left(\frac{1}{n_i} \sum_{j=0}^{n_i-1} (T^*)^j \rho \right) = \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} (T^*)^{j+1} \rho$$

Integrando cada $\psi \in C^0(X, \mathbb{R})$ respecto de la medida $T^*\mu$, y luego respecto de la medida μ , aplicando la definición de límite de la sucesión de medidas μ_{n_i} con la topología débil*, y la continuidad del operador T^* en \mathcal{M} , obtenemos:

$$\begin{aligned} \int \psi dT^*\mu - \int \psi d\mu &= \lim_{i \rightarrow +\infty} \left(\int \psi dT^*\mu_{n_i} - \int \psi d\mu_{n_i} \right) = \\ &= \lim_{i \rightarrow +\infty} \frac{1}{n_i} \left(\int \psi d(T^*)^{n_i}\rho - \int \psi d\rho \right) \end{aligned}$$

Entonces

$$\begin{aligned} \left| \int \psi dT^*\mu - \int \psi d\mu \right| &\leq \lim_{i \rightarrow +\infty} \frac{1}{n_i} \|\psi\|_0 (\rho(X) + \rho(T^{-n_i}(X))) = \\ &= \lim_{n \rightarrow +\infty} \frac{2}{n} \|\psi\|_0 = 0. \end{aligned}$$

En la última igualdad tuvimos en cuenta que ρ es una probabilidad. Obtuvimos que

$$\left| \int \psi dT^*\mu - \int \psi d\mu \right| = 0 \quad \forall \psi \in C^0(X, \mathbb{R}).$$

Por lo tanto los operadores lineales $\psi \mapsto \int \psi dT^*\mu$ y $\psi \mapsto \int \psi d\mu$ son el mismo. Por la unicidad de la medida en el teorema de Riesz deducimos que $T^*\mu = \mu$, como queríamos demostrar. \square

Ejercicio 1.2.13. Sea (X, \mathcal{A}) un espacio medible y sea \mathcal{M} el conjunto de todas las medidas de probabilidad en (X, \mathcal{A}) . Suponga que existe $\mu \in \mathcal{M}(X)$ tal que $T^*\mu(A) \leq 2\mu(A)$ para todo boreliano $A \subset X$.

a) Probar que $2\mu - T^*\mu \in \mathcal{M}(X)$.

b) Si X es un espacio métrico compacto, \mathcal{A} es la sigma-álgebra de Borel y si T es continua, probar que dado $\mu_0 \in \mathcal{M}(X)$ existe $\mu \in \mathcal{M}(X)$ tal que $2\mu - T^*\mu = \mu_0$. Sugerencia: Para todo $\mu \in \mathcal{M}(X)$ definir $G(\mu) = 1/2 \cdot (T^*\mu + \mu_0)$, $\mu_n = 1/n \cdot \sum_{j=0}^{n-1} G^j(\mu_0)$ y tomar una subsucesión convergente en $\mathcal{M}(X)$. Probar que G es continuo en $\mathcal{M}(X)$. Observar que G conmuta con las combinaciones lineales finitas *convexas* de probabilidades. Es decir si $\mu = \sum_{i=1}^k \lambda_i \nu_i$, donde $\nu_i \in \mathcal{M}(X)$, $0 \leq \lambda_i \leq 1$ y $\sum_{i=1}^k \lambda_i = 1$, entonces $G(\mu) = \sum_{i=1}^k \lambda_i G(\nu_i)$.

1.3. Ejemplos y puntos periódicos hiperbólicos

Ejemplo 1.3.1. Sea en S^1 (el círculo) la rotación $T(e^{2\pi ix}) = e^{2\pi i(x+a)}$, donde a es una constante real. Si a es racional, T se llama *rotación racional* en el círculo, y si a es irracional, T se llama *rotación irracional*. A través de la proyección $\Pi : \mathbb{R} \mapsto S^1$ dada por $\Pi(x) = e^{2\pi ix}$, la medida de Lebesgue m en \mathbb{R} induce una medida m_\sim en S^1 dada por $m_\sim(A) = m(\Pi^{-1}(A) \cap [0, 1])$. Esta medida m_\sim se llama medida de Lebesgue en el círculo. Como m en \mathbb{R} es invariante por traslaciones, es fácil probar que m_\sim en el círculo S^1 es invariante por las rotaciones.

Ejercicio 1.3.2. Probar que para la rotación racional en el círculo todos los puntos son periódicos con el mismo período. Probar que la rotación irracional en el círculo no tiene puntos periódicos. Probar que la medida de Lebesgue en el círculo es invariante por las rotaciones.

Observación 1.3.3. Aunque no es inmediato, se puede probar que *todas las órbitas de la rotación irracional del círculo son densas*. (Lo probaremos en §2.6.2).

Ejercicio 1.3.4. Tent map Sea el intervalo $[0,1]$ dotado de la sigma álgebra de Borel. Sea $T : [0, 1] \mapsto [0, 1]$ dada por $T(x) = 2x$ si $x \in [0, 1/2]$ y $T(x) = 2 - 2x$ si $x \in [1/2, 1]$. Probar que T preserva la medida de Lebesgue en el intervalo. (Sugerencia: graficar T y probar que la preimagen de un intervalo I tiene la misma medida que I . Usar el corolario 1.1.10).

Definición 1.3.5. Sea $T : X \mapsto X$ Borel medible en un espacio métrico X . Decimos que un punto periódico x_0 de período p es *un atractor* si existe un entorno V de x_0 invariante hacia delante por T^p (es decir $T^p(V) \subset V$) y tal que

$$\text{dist}(T^n(x_0), T^n(y))_{n \rightarrow +\infty} \rightarrow 0 \quad \forall y \in V$$

Cuando T es invertible con inversa medible decimos que un punto periódico x_0 de período p es *un repulsor* si existe un entorno V de x_0 invariante hacia atrás por T^p (es decir $T^{-p}(V) \subset V$) y tal que

$$\text{dist}(T^n(x_0), T^n(y))_{n \rightarrow -\infty} \rightarrow 0 \quad \forall y \in V$$

Proposición 1.3.6. Sea $f : S^1 \mapsto S^1$ un difeomorfismo; es decir f es de clase C^1 (i.e. derivable con derivada continua), invertible (i.e. biyectiva; existe la transformación inversa $f^{-1} : S^1 \mapsto S^1$), y su inversa f^{-1} es también de clase C^1)

Supongamos que el difeomorfismo $f : S^1 \mapsto S^1$ preserva la orientación (i.e. $f' > 0$). Sea x_0 un punto periódico de período p tal que la derivada $(f^p)'(x_0)$ es menor que 1. Entonces x_0 es un atractor. Análogamente, si $(f^p)'(x_0)$ es mayor que 1 entonces x_0 es un repulsor.

Demostración: La segunda afirmación se obtiene de la primera usando f^{-p} en lugar de f^p . Demostremos la primera afirmación renombrando como f a la transformación f^p . Entonces x_0 es fijo. Grafíquese $f(x)$ para $x \in S^1 \approx [0, 1]$ del intervalo $[0, 1]$ en sí mismo, en el que se ha identificado el 0 con el 1 en el punto x_0 . La gráfica de f corta a la diagonal por lo menos en el punto $0 \sim 1 = x_0$. Gráficamente, los iterados futuros de y en un entorno de x_0 suficientemente pequeño, se obtienen trazando la vertical de abscisa y , cortándola con la gráfica de f , trazando luego la horizontal por ese punto, cortándola con la diagonal, trazando la vertical por ese punto, cortándola con la gráfica de f , y así sucesivamente (ver por ejemplo [Jo 2005, Figure on page 19]). Si la función es continua con derivada continua, y la derivada $f'(x_0) = a > 0$ en el punto fijo x_0 es menor que uno, entonces los sucesivos puntos en la gráfica de f obtenidos por la construcción anterior, tienden monótonamente al punto fijo x_0 . En efecto, por la definición de diferenciabilidad y de derivada: $\|f(y) - f(x_0)\| \leq$

$(a+(1-a)/2)\|y-x_0\|$ para todo y suficientemente cercano a x_0 , digamos $\|y-x_0\| < \delta$. Es decir $\|f(y) - x_0\| \leq b\|y - x_0\|$ donde $0 < b = a + (1 - a)/2 = (1 + a)/2 < 1$. Luego, $f(y)$ también cumple $\|f(y) - x_0\| < \delta$. Se puede aplicar, por inducción, la desigualdad anterior a todos los iterados futuros $f^n(y)$ (es decir para todo $n \in \mathbb{N}$). Obtenemos $\|f^n(y) - x_0\| \leq b^n\|y - x_0\|$. Siendo $0 < b < 1$, deducimos que $f^n(y)$ converge monótonamente a x_0 . \square

Definición 1.3.7. Un punto periódico x_0 con período p de un difeomorfismo $f : X \mapsto X$ en una variedad diferenciable X se dice *hiperbólico* si los valores propios (complejos) de la derivada $df_{x_0}^p$ de f^p en x_0 , tienen todos módulo diferente de 1. Se recuerda que la derivada $df_{x_0}^p$ es una transformación lineal de \mathbb{R}^m en \mathbb{R}^m , donde m es la dimensión de la variedad X .

Consecuencia: Si x_0 es un punto periódico hiperbólico de un difeomorfismo f de clase C^1 del círculo S^1 , entonces es un atractor si $|(f^p)'(x_0)| < 1$, y es un repulsor si $|(f^p)'(x_0)| > 1$. (Siendo x_0 hiperbólico, sabemos que $|(f^p)'(x_0)| \neq 1$ por definición, así que los dos casos anteriores son los únicos posibles). En efecto, si f preserva la orientación del círculo, aplicamos la Proposición 1.3.6, y si f invierte la orientación, aplicamos la misma Proposición a f^2 para deducir que x_0 es un punto fijo atractor (resp. repulsor) de f^2 . Es fácil probar que si x_0 es un punto fijo de f que es atractor (resp. repulsor) para f^2 , entonces también es atractor (resp. repulsor) para f .

Para un difeomorfismo f en el círculo S^1 , y más en general para un mapa de clase C^1 en una variedad de dimensión 1, un punto periódico hiperbólico x_0 se llama *pozo* si $|(f^p)'(x_0)| < 1$, y se llama *fuentes* si $|(f^p)'(x_0)| > 1$. Generalizando este resultado cuando la variedad tiene dimensión mayor que uno, adoptamos la siguiente definición:

Definición 1.3.8. Pozos, fuentes y sillars

Sea $f : X \mapsto X$ un difeomorfismo en una variedad diferenciable X . Sea x_0 un punto periódico hiperbólico para f de período p (i.e. los valores propios de $df_{x_0}^p$ tienen todos módulo diferente de 1). El punto x_0 , y también la órbita (finita) de x_0 , se llama *pozo* si los valores propios de $df_{x_0}^p$ tienen todos módulo menor que 1. Se llama *fuentes* si todos tienen módulo mayor que 1. Y se llama *silla* si alguno tiene módulo mayor que 1 y algún otro módulo menor que 1. (Obsérvese que las sillars solo pueden existir si la dimensión de la variedad es 2 o mayor).

Ejercicio 1.3.9. (a) Encontrar ejemplo de un difeomorfismo $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ con un punto fijo hiperbólico tipo silla, otro ejemplo con un pozo y otro con una fuente. (b) Encontrar un ejemplo de $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ que tenga exactamente tres puntos fijos, sean los tres hiperbólicos, uno tipo fuente, otro pozo y otro silla. (c) Demostrar que para cualquier difeomorfismo $f : M \mapsto M$, los pozos son atractores, las fuentes son repulsores, y las sillars no son atractores ni repulsores.

Ejercicio 1.3.10. Encontrar un ejemplo de difeomorfismo $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ que tenga un punto fijo atractor que no sea pozo (i.e. que no sea hiperbólico), otro que tenga un punto fijo repulsor que no sea fuente (i.e. que no sea hiperbólico) y otro que tenga un punto fijo no hiperbólico que no sea ni fuente ni pozo pero que todas las órbitas futuras en un entorno cualquiera de x_0 suficientemente pequeño, o bien tiendan a x_0 o bien se salgan del entorno.

Exponente de Lyapounov negativo significa contracción exponencial: Si $f : S^1 \mapsto S^1$ es un difeomorfismo y x_0 es un punto fijo atractor hiperbólico (i.e. $|f'(x_0)| < 1$, es decir x_0 es un pozo), entonces

$$\lim_{n \rightarrow +\infty} \frac{\log \text{dist}(f^n(x), x_0)}{n} = -\lambda < 0 \quad \forall x \text{ en algún entorno de } x_0$$

donde $-\lambda = \log |f'(x_0)| < 0$ se llama exponente de Lyapounov en x_0 .

Ejercicio 1.3.11. Demostrar la afirmación anterior. Sugerencia: Probar que

$$\frac{\text{dist}(f^{n+1}(x), x_0)}{\text{dist}(f^n(x), x_0)} \rightarrow e^{-\lambda} < 1$$

Interpretación: La distancia de $f^n(x)$ al atractor hiperbólico se contrae exponencialmente con coeficiente asintóticamente igual a e elevado al exponente de Lyapounov $-\lambda < 0$.

Exponente de Lyapounov positivo significa dilatación exponencial: Si $f : S^1 \mapsto S^1$ es un difeomorfismo y x_0 es un punto fijo repulsor hiperbólico (i.e. $|f'(x_0)| > 1$, es decir x_0 es una fuente) entonces

$$\lim_{n \rightarrow -\infty} \frac{\log \text{dist}(f^n(x), x_0)}{n} = \sigma > 0 \quad \forall x \text{ en algún entorno de } x_0$$

donde $\sigma = \log |f'(x_0)| > 0$ se llama exponente de Lyapounov en x_0 .

Ejercicio 1.3.12. Demostrar la afirmación anterior. Sugerencia: Aplicar lo ya probado a f^{-1} y la fórmula de derivada de la función inversa para deducir que

$$\lim_{n \rightarrow -\infty} \frac{\text{dist}(f^{n+1}(x), x_0)}{\text{dist}(f^n(x), x_0)} \rightarrow e^\sigma > 1$$

Interpretación: La distancia de $f^n(x)$ al repulsor hiperbólico (al crecer n y mientras $f^n(x)$ esté en un entorno pequeño del repulsor) se dilata exponencialmente con coeficiente asintóticamente igual a e elevado al exponente de Lyapounov $\sigma > 0$.

Ejercicio 1.3.13. Flujo polo norte-polo sur Llamaremos *sección de flujo polo norte-polo sur* en el círculo S^1 , a un difeomorfismo $f : S^1 \mapsto S^1$ de clase C^1 , que preserva la orientación, y tal que existen solo 2 puntos fijos N y S , son ambos hiperbólicos, N repulsor y S atractor. Graficar f en $[0, 1]_{\text{mod } 1}$ tomando $0 \sim 1 = N$. Demostrar que todas las órbitas excepto N y S son monótonas y convergen a S . (Sugerencia: ver prueba de la afirmación 1.3.6.) Demostrar que las únicas medidas de probabilidad invariantes son las combinaciones lineales convexas de δ_N y δ_S . Sugerencia: considerar una partición numerable del intervalo $(0, S)$ formado por $A_n = [x_n, x_{n+1})$ para $n \in \mathbb{Z}$ donde x_0 se elige cualquiera en el intervalo abierto $(0, S)$ y $x_n := f^n(x_0)$ para todo $n \in \mathbb{Z}$. Usando argumento similar a la prueba del ejemplo 1.1.7, probar que $\mu((0, S)) = 0$. Análogamente probar que $\mu((S, 1)) = 0$, de donde $\mu(\{N, S\}) = 1$.

Definición 1.3.14. Difeomorfismo Morse-Smale en S^1 . Un difeomorfismo $f : S^1 \mapsto S^1$ se dice *Morse-Smale* si preserva la orientación y existen exactamente una cantidad finita de puntos periódicos (todos del mismo período) y son todos ellos hiperbólicos.

Ejercicio 1.3.15. Probar que en un difeomorfismo f Morse-Smale en el círculo las únicas medidas invariantes son las combinaciones lineales convexas de las medidas

$$\frac{\delta_{x_0} + \delta_{f(x_0)} + \dots + \delta_{f^{p-1}(x_0)}}{p}$$

donde x_0 es un punto periódico de período p . Sugerencia: Graficar f^p en $S^1 = [0, 1]/\sim$ donde $0 \sim 1$ es un punto periódico de período p . Probar que los atractores y los repulsores se alternan. Probar que para toda medida invariante el arco entre un repulsor y un atractor consecutivos tiene medida cero, usando el procedimiento del ejercicio 1.3.13.

1.4. Dinámica topológica

Definición 1.4.1. Recurrencia topológica. Sea $T : X \mapsto X$ una transformación Borel medible en un espacio topológico X . Sea $x \in X$. Se llama *omega-límite de x* al conjunto

$$\omega(x) = \{y \in X : \exists n_j \rightarrow +\infty \text{ tal que } T^{n_j}(x) \mapsto y\}$$

Cuando T es bi-medible (i.e. T es medible, invertible y con inversa medible) se llama *alfa-límite de x* al conjunto

$$\alpha(x) = \{y \in X : \exists n_j \rightarrow -\infty \text{ tal que } T^{n_j}(x) \mapsto y\}$$

Un punto x se dice *recurrente* si

$$x \in \omega(x)$$

Dicho de otra forma x es recurrente si existe una subsucesión $n_j \rightarrow +\infty$ tal que $T^{n_j}(x) \rightarrow x$. Luego para todo entorno V de x existe una subsucesión $n_j \in \mathbb{N}$ tal que $T^{n_j}(V) \cap V \neq \emptyset$.

Ejercicio 1.4.2. Sea X un espacio métrico compacto y sea $T : X \mapsto X$ continua. Probar que:

- $\omega(x)$ es compacto no vacío para todo $x \in X$.
- $\omega(T^n(x)) = \omega(x)$ para todo $n \in \mathbb{N}$, es decir el conjunto $\omega(x)$ depende de la órbita por x y no de qué punto en la órbita de x se elija.
- $T(\omega(x)) = \omega(x) \subset T^{-1}(\omega(x))$ para todo $x \in X$. Es decir $\omega(x)$ es un conjunto invariante por T hacia el futuro.
- Si además T es un homeomorfismo (i.e. T es continua, invertible y con inversa T^{-1} continua) probar que:
 - $\alpha(x)$ es compacto no vacío para todo $x \in X$, $\alpha(T^n(x)) = \alpha(x)$ para todo $n \in \mathbb{Z}$ y $T(\alpha(x)) = \alpha(x) = T^{-1}(\alpha(x))$.

(ii) Si $x \in \omega(x)$ entonces $\alpha(x) \subset \omega(x)$. Sugerencia: Si $x \in \omega(x)$ entonces $T^{-n}(x) \in \omega(x)$ para todo $n \in \mathbb{Z}$, luego $\alpha(x) \subset \omega(x) = \omega(x)$.

(iii) $x \in \omega(x)$ si y solo si $\omega(x) = o(x)$, donde $o(x) = \{T^n(x) : n \in \mathbb{Z}\}$.

Definición 1.4.3. Conjunto no errante. Sea $T : X \mapsto X$ una transformación Borel medible en un espacio topológico X . Un punto $x \in X$ es *no errante* si para todo entorno V de x existe una sucesión $n_j \rightarrow +\infty$ tal que $T^{n_j}(V) \cap V \neq \emptyset$. El conjunto de los puntos no errantes de T (que puede ser vacío) se denota como $\Omega(T)$ y se llama *conjunto no errante*.

Ejercicio 1.4.4. Sea X un espacio topológico de Hausdorff y sea $T : X \mapsto X$ Borel medible.

(a) Probar que el conjunto de los puntos recurrentes está contenido en el conjunto no errante $\Omega(T)$ (la inclusión opuesta no es necesariamente cierta, como se verá más adelante).

(b) Sea μ una medida de probabilidad invariante por T . Si X tiene base numerable de abiertos probar (sin usar el enunciado del teorema de recurrencia de Poincaré que viene más adelante) que $\mu(\Omega(T)) = 1$, es decir: casi todo punto es no errante para cualquier medida de probabilidad invariante por T .

Sugerencia: Probar que para todo V de la base de abiertos que tenga medida positiva la sucesión de conjuntos medibles $T^{-n}(V)$, $n \in \mathbb{N}$ no puede ser de conjuntos disjuntos dos a dos a partir de un cierto n_0 en adelante. Deducir que existe una subsucesión $n_j \rightarrow +\infty$ tal que $T^{-n_j}(V) \cap V \neq \emptyset$ y esto implica $V \cap T^{n_j}(V) \neq \emptyset$. Un punto es errante (no es no errante) si está contenido en algún V abierto tal que no cumple lo anterior. Deducir que los puntos errantes forman un conjunto de medida nula.

Definición 1.4.5. Transitividad topológica Sea X un espacio topológico y $T : X \mapsto X$ Borel medible. Se dice que T es *transitiva* si dados dos abiertos U y V no vacíos, existe $n \geq 1$ tal que $T^n(U) \cap V \neq \emptyset$.

Supóngase que X es de Hausdorff sin puntos aislados. Es fácil probar que *si existe una órbita positiva densa entonces T es transitiva*. Y si además, T es continua y el espacio topológico tiene base numerable de abiertos y es de Baire (esto es: toda intersección numerable de abiertos densos es densa) entonces *T es transitiva si y solo si existe una órbita positiva densa*.

La transitividad significa que para cualquier abierto U , por pequeño que sea, *los iterados positivos de U transitan por todo el espacio desde el punto de vista topológico* (es decir, por todos los abiertos del espacio).

Observación 1.4.6. Se observa que para dos conjuntos cualesquiera U y V :

$$T^n(U) \cap V \neq \emptyset \text{ si y solo si } T^{-n}(V) \cap U \neq \emptyset$$

Es fácil ver que si T es continua y transitiva, entonces dados dos abiertos U y V no vacíos, existe $n_j \rightarrow +\infty$ tal que $T^{-n_j}(V) \cap U \neq \emptyset$. En particular, tomando $U = V \ni x$, se deduce que:

Si $T : X \mapsto X$ es continua y transitiva entonces $\Omega(T) = X$.

Ejercicio 1.4.7. Probar las afirmaciones de la Observación 1.4.6 y las que están inmediatamente después de la definición de transitividad en 1.4.5.

1.5. Recurrencia y Lema de Poincaré

Sea $T : X \mapsto X$ Borel medible en un espacio métrico compacto X . Recordando la Definición 1.4.1 y teniendo en cuenta la compacidad secuencial de X , un punto x es recurrente si y solo si *para todo entorno V de x existen infinitos iterados hacia el futuro de x en V* . Es decir, existe $n_j \rightarrow +\infty$ tal que $T^{n_j}(x) \in V$.

Definición 1.5.1. Sea $T : X \mapsto X$ medible en un espacio medible (X, \mathcal{A}) . Sea $E \in \mathcal{A}$. Un punto $x \in E$ *vuelve infinitas veces a E* si existen infinitos iterados hacia el futuro de x en E . Mejor dicho: existe $n_j \rightarrow +\infty$ tal que $T^{n_j}(x) \in E$ para todo $j \in \mathbb{N}$.

Los siguientes dos teoremas, llamados Lemas de Recurrencia de Poincaré, se encuentran por ejemplo en [Mañ 1983a, pag. 32-35] (ver también [Mañ 1987]). En [Wa 2000, §1.4] se encuentra también la versión medible siguiente:

Teorema 1.5.2. Lema de Recurrencia de Poincaré. Versión medible

Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Sea E un conjunto medible tal que $\mu(E) > 0$. Entonces μ -c.t.p de E vuelve infinitas veces a E .

Demostración. Sea $F_N := \bigcap_{n \geq N} T^{-n}(X \setminus E)$ y sea $F := \bigcup_{N \geq 0} F_N$. Por construcción $x \in F$ si y solo si $T^n(x) \notin E$ para todo n suficientemente grande. Esto ocurre si y solo si la órbita futura $\{T^n(x)\}_{n \geq 0}$ de x no pasa infinitas veces por E . Basta probar entonces que $\mu(F \cap E) = 0$.

Por construcción $T^{-1}(F_N) = F_{N+1}$. Como μ es T -invariante, tenemos $\mu(F_{N+1}) = \mu(F_N)$ para todo $N \geq 0$. Luego $\mu(F_N) = \mu(F_0)$ para todo $N \geq 0$. Siendo $F_{N+1} \supset F_N$ para todo $N \geq 0$, entonces $\mu(F) = \lim_{N \rightarrow +\infty} \mu(F_N) = \mu(F_0)$ y $F \supset F_0$. Luego $\mu(F \setminus F_0) = 0$. Como $E \cap F_0 = \emptyset$, deducimos que $E \cap F \subset F \setminus F_0$, de donde $\mu(E \cap F) = 0$, como queríamos demostrar. \square

Teorema 1.5.3. Lema de Recurrencia de Poincaré. Versión topológica

Sea $T : X \mapsto X$ Borel-medible en un espacio topológico X con base numerable. Si T preserva una medida de probabilidad μ , entonces μ -c.t.p. es recurrente (es decir $x \in \omega(x)$ μ -c.t.p. $x \in X$).

Demostración: Sea $\{V_j\}_{j \in \mathbb{N}}$ una base de abiertos. Por 1.4.1: $x \notin \omega(x)$ si y solo si $x \in \bigcup_{j \in \mathbb{N}} A_j$ para algún $j \in \mathbb{N}$, donde $A_j = \{x \in V_j : x \text{ no vuelve infinitas veces a } V_j\}$. Por 1.5.2 $\mu(A_j) = 0$. Luego, la unión numerable de los conjuntos A_j , que coincide con el conjunto de los puntos no recurrentes, tiene μ medida nula. \square

Ejercicio 1.5.4. Sea (X, \mathcal{A}) un espacio medible. Sea $T : X \mapsto X$ medible que preserva una probabilidad μ . Sea $E \in \mathcal{A}$ tal que $\mu(E) > 0$. Probar que

$$\sum_{n \in \mathbb{N}} \chi_E(T^n(x))$$

diverge μ - c.t.p $x \in E$

Ejercicio 1.5.5. Sea $T : X \mapsto X$ Borel medible en el espacio topológico X compacto, y preservando una medida de probabilidad μ . Sea $\text{supp}(\mu)$ el soporte compacto de μ (i.e. el mínimo compacto con medida μ igual a 1). Probar que $\emptyset \neq \text{supp}(\mu) \subset \overline{\text{Rec}(T)}$ siendo $\text{Rec}(T)$ el conjunto de los puntos recurrentes de T .

Teorema 1.5.6. Teorema de Hopf. Sea $T : \mathbb{R}^k \mapsto \mathbb{R}^k$ Borel bimedible (medible, invertible con inversa medible) que preserve la medida de Lebesgue m . Entonces casi todo punto de \mathbb{R}^n o bien es recurrente o bien tiene omega límite vacío.

Ejercicio 1.5.7. Demostrar el teorema de Hopf enunciado antes. Sugerencia: $\mathbb{R}^k = \bigcup_{i \in \mathbb{N}} X_i$ donde X_i es una bola abierta de radio $r_i \rightarrow +\infty$ creciente con i . Sea

$$\tilde{X}_i = \{x \in X_i : T^j(x) \in X_i \text{ para infinitos valores positivos de } j\}.$$

Sea $\tilde{T}_i : \tilde{X}_i \mapsto \tilde{X}_i$ la transformación que a cada $x \in \tilde{X}_i$ hace corresponder el primer retorno a X_i : $\tilde{T}_i(x) = T^j(x) \in \tilde{X}_i$ para el mínimo $j = j(x)$ natural positivo tal que $T^j(x) \in X_i$. Probar que $m(\tilde{T}_i \tilde{X}_i) = m(\tilde{X}_i)$; luego c.t.p. de \tilde{X}_i está en la imagen de \tilde{T}_i . Probar que \tilde{T}_i preserva m . Aplicar el teorema de recurrencia de Poincaré para deducir que m -c.t.p. de \tilde{X}_i es recurrente. Probar que c.t.p. de X_i o bien es recurrente o bien su omega límite no intersecta a X_i .

Observación 1.5.8. En [Fra-McC 2011], y en la bibliografía allí citada, se reseñan varios otros resultados sobre recurrencia, además de los lemas básicos de recurrencia de Poincaré. Algunos de estos resultados miden, en relación a potencias de n , la frecuencia con la que órbita futura del punto recurrente x se acerca a x , la vinculan con las medidas ergódicas y con la entropía métrica del sistema (la entropía métrica es, gruesamente hablando, una medición, ponderada según una probabilidad invariante μ , del desorden espacial que produce f al ser iterada).

1.6. Ergodicidad

Definición 1.6.1. Ergodicidad I

Sea (X, \mathcal{A}, μ) un espacio de medida de probabilidad y $T : X \mapsto X$ medible que preserve μ . Se dice que T es *ergódica* respecto a la medida μ , o que μ es una *medida ergódica* para T , si dados dos conjuntos medibles con medida positiva U y V existe $n \geq 1$ tal que $T^n(U) \cap V \neq \emptyset$.

Nota: Observar que la ergodicidad es la versión en el contexto medible de la transitividad topológica. Se resalta que por definición, si una medida es ergódica para T , entonces es T -invariante. No se define ergodicidad de medidas no invariantes.

Definición 1.6.2. Ergodicidad II

Sea (X, \mathcal{A}, μ) un espacio de medida de probabilidad y $T : X \mapsto X$ medible que preserve μ . Se dice que T es *ergódica* respecto a la medida μ , o que μ es una *medida ergódica* para T , si todo conjunto medible A que sea invariante por T (es decir $T^{-1}(A) = A$) tiene o bien medida nula o bien medida 1.

Teorema 1.6.3. T es ergódica según la definición I para la medida de probabilidad μ si y solo si es ergódica según la definición II.

Demostración: Supongamos que no se cumple la definición II. Entonces, existe un conjunto A invariante por T que tiene medida positiva distinta de 1. Luego el complemento A^c de A es también invariante por T y tiene medida positiva. Como $A = T^{-1}(A)$, toda órbita positiva con estado inicial en A está contenida en A . Deducimos que no existe $n \in \mathbb{N}$ tal que $T^n(A)$ interseca a A^c . Concluimos que T no es ergódica según la definición I.

Recíprocamente, suponemos por hipótesis que se cumple la definición II. Todo conjunto A invariante por T tiene medida o bien nula o bien 1. Supongamos por absurdo que no se cumple la definición I. Entonces existen conjuntos medibles U y V con medida positiva tales que $\bigcup_{n \in \mathbb{N}} T^{-n}(V) \cap U = \emptyset$. Entonces el conjunto

$$A = \bigcap_{N \in \mathbb{N}} \left(\bigcup_{n \geq N} T^{-n}(V) \right)$$

es medible, invariante por T (verificar que $T^{-1}(A) = A$) y tiene medida positiva (verificar que $\mu(A) \geq \mu(V)$ usando que μ es medida de probabilidad T -invariante), pero A no interseca a U que también tiene medida positiva. De lo anterior se deduce que A no puede tener medida 1, con lo que encontramos un conjunto invariante que tiene medida positiva menor que 1, contradiciendo la hipótesis. \square

Ejercicio 1.6.4. Sea $T : X \mapsto X$ que preserve una medida de probabilidad μ . (a) Probar que μ es ergódica para T si y solo si todo conjunto A medible invariante para el futuro (i.e. $A \subset T^{-1}(A)$) cumple $\mu(A) = 0$ ó $\mu(A) = 1$. (b) Probar que μ es ergódica para T si y solo si todo conjunto A medible invariante para el pasado (i.e. $T^{-1}(A) \subset A$) cumple $\mu(A) = 0$ ó $\mu(A) = 1$. Sugerencias: (a) Considerar $B = \bigcup_{n \geq 0} T^{-n}(A)$, probar que $\mu(B) = \mu(A)$ y que B es T -invariante. (b) Considerar el complemento de A .

Definición: Una medida de Borel μ en un espacio topológico se dice *positiva sobre abiertos* si $\mu(V) > 0$ para todo abierto V no vacío.

Ejercicio 1.6.5. Sea X un espacio topológico y $T : X \mapsto X$ Borel medible que preserve una medida de probabilidad μ que es positiva sobre abiertos. Probar que si T es ergódica respecto de μ entonces T es transitiva (topológicamente).

Ejercicio 1.6.6. Probar que los difeomorfismos Morse-Smale en el círculo tienen medidas ergódicas pero no son transitivos topológicamente.

Observación 1.6.7. Más adelante demostraremos que *la medida de Lebesgue es ergódica para la rotación irracional del círculo* (Teoremas 2.6.1 y 8.4.1). Luego, como la medida de Lebesgue es positiva sobre abiertos, la rotación irracional es transitiva topológicamente. Entonces existe alguna órbita densa. Es fácil ver, usando que la rotación en el círculo conserva las distancias, que al existir una órbita densa, *todas las órbitas son densas*.

También probaremos que el *tent map* T en el intervalo es ergódico respecto a la medida de Lebesgue. Luego T es topológicamente transitivo y existe órbita densa. Sin embargo no todas las órbitas en el futuro por T son densas: en efecto, existen órbitas periódicas (que, como órbitas en el futuro, son conjuntos finitos, y por lo tanto no son densos).

1.6.8. Promedios temporales asintóticos de Birkhoff

El Teorema Ergódico de Birkhoff-Khinchin, que enunciaremos completamente más adelante (Teorema 2.1.2), establece que si $T : X \mapsto X$ es una transformación medible que tiene medidas de probabilidad T -invariantes, entonces para toda probabilidad μ que sea T -invariante y para toda función $\psi \in L^1(\mu)$, existe μ -c.t.p. $x \in X$ el siguiente límite:

$$\tilde{\psi}(x) := \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x).$$

Teorema 1.6.9. Ergodicidad III. *Sea $T : X \mapsto X$ medible en un espacio medible (X, \mathcal{A}) y sea μ una probabilidad invariante por T . Las siguientes afirmaciones son equivalentes:*

- (i) μ es ergódica para T
- (ii) Toda función $\psi : X \mapsto \mathbb{R}$ que sea medible e invariante por T (es decir $\psi(x) = \psi(T(x)) \forall x \in X$), es constante μ -c.t.p.
- (iii) Para toda función medible y acotada $\psi : X \mapsto \mathbb{R}$ existe el siguiente límite μ -c.t.p. y es igual a $\int \psi d\mu$:

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x) = \int \psi d\mu \quad \mu - \text{c.t.p. } x \in X \quad (1.3)$$

Probaremos el Teorema 1.6.9 en el parágrafo 1.6.11.

Observación 1.6.10. Hipótesis de Boltzmann de la Mecánica Estadística:

Antes de demostrar el Teorema 1.6.9, interpretaremos el significado de la afirmación en la parte (iii) para fundamentar su relevancia. Ella es un caso particular del Teorema Ergódico de Birkhoff, que veremos más adelante, que establece que para toda medida μ ergódica se cumple la igualdad (1.3), no solo cuando ψ es medible y acotada, sino también para toda $\psi \in L^1(\mu)$.

La igualdad (1.3) posee un significado relevante en la teoría ergódica, pues afirma que *el promedio espacial* de cada función ψ con respecto a la probabilidad μ en el espacio X (i.e. el valor esperado $\int \psi d\mu$ de cada variable aleatoria ψ) coincide con *el promedio temporal asintótico* de los valores observados de ψ a lo largo de μ -casi toda órbita. Este promedio temporal asintótico es el límite cuando $n \rightarrow +\infty$ de los promedios temporales $(1/n) \sum_{j=0}^{n-1} \psi(T^j(x))$ de los valores de ψ , observados a lo largo del pedazo finito de órbita desde el instante 0 hasta el instante $n - 1$. Salvo casos excepcionales, es falso que el límite de los promedios temporales exista para todos los puntos $x \in M$. Además, aunque para toda medida invariante μ ese límite existe μ -c.t.p. (Teorema ergódico de Birkhoff), es falso en general (salvo cuando μ es ergódica)

que coincida con el promedio espacial de ψ respecto a la probabilidad μ . Por lo tanto las medidas ergódicas μ para T tienen un significado estadístico relevante, pues permite estimar el promedio temporal a largo plazo (esto es el promedio estadístico de las series de observaciones $\psi(T^j(x))$ a largo plazo, llamésmole por ejemplo el “clima”) en los sistemas determinísticos, para μ -casi todo estado inicial $x \in X$, calculando el valor esperado de ψ respecto a la probabilidad μ . Sin embargo, aplicar la igualdad (1.3) para hacer esa estimación, puede ser muy erróneo cuando μ no es ergódica, o cuando μ no es invariante por T , o cuando el estado inicial x no pertenece al conjunto de μ -probabilidad igual a 1.

La igualdad (1.3) de los promedios temporales asintóticos con el promedio espacial (o sea el valor esperado) es lo que en la Mecánica Estadística se llama Hipótesis de Boltzmann. Es una hipótesis importante para demostrar propiedades de la dinámica de sistemas formado por una cantidad finita pero muy grande de partículas que evolucionan determinísticamente en el tiempo (por iteraciones de un mapa T medible) preservando una medida de probabilidad dada μ en el espacio X de todos los estados posibles (llamado “espacio de fases”). Para los sistemas llamados conservativos esta medida de probabilidad μ es la medida de Lebesgue, normalizada para que $\mu(X) = 1$. Aplicando el Teorema 1.6.9, la hipótesis de Boltzmann se traduce en la hipótesis de ergodicidad de esa probabilidad μ .

1.6.11. Proof of Theorem 1.6.9:

Demostración. (i) \Rightarrow (ii):

Sea $a \in \mathbb{R}$ y sea $A_a = \{x \in X : \psi(x) \leq a\} \subset X$. Como ψ es medible, A_a es medible. Siendo $\psi(T(x)) = \psi(x)$ para todo $x \in X$, tenemos $T^{-1}(A_a) = A_a$. Como μ es ergódica, entonces $\mu(A_a)$ vale 0 o 1. Considere la función $g : \mathbb{R} \mapsto \{0, 1\}$ definida como $g(a) := \mu(A_a)$. Por construcción $A_a \subset A_b$ si $a < b$. Luego $g(a) \leq g(b)$ si $a < b$. Entonces g es creciente y solo puede tomar valores 0 ó 1. Sea

$$k = \inf\{a \in \mathbb{R} : g(a) = 1\} = \sup\{a \in \mathbb{R} : g(a) = 0\} \in [-\infty, +\infty]. \quad (1.4)$$

Probemos que $k \in \mathbb{R}$. En efecto $\psi(x) \in \mathbb{R}$ para todo $x \in M$; entonces $\emptyset = \bigcap_{n \in \mathbb{N}} A_{-n}$, $X = \bigcup_{n \in \mathbb{N}} A_n$. Como $A_n \subset A_{n+1}$ para todo $n \in \mathbb{N}$, entonces $\lim_{n \rightarrow +\infty} \mu(A_n) = 1$ y $\lim_{n \rightarrow -\infty} \mu(A_n) = 0$. Esto implica que existe $n_0 \in \mathbb{N}$ tal que $g(n_0) = 1$ y $g(-n_0) = 0$. Por lo tanto $-n_0 \leq k \leq n_0$. Ahora probemos que $\mu(A_k) = 1$. De (1.4) deducimos que $0 = \mu(A_{k-(1/n)}) \leq \mu(A_k) \leq \mu(A_{k+(1/n)}) \forall 1 \leq n \in \mathbb{N}$ Además $A_k = \bigcap_{n=1}^{\infty} A_{k+(1/n)} = X$, y $A_{k+(1/n+1)} \subset A_{k+(1/n)}$. Luego $\mu(A_k) = \lim_{n \rightarrow +\infty} \mu(A_{k+(1/n)}) = 1$. Hemos probado que $\mu(A_k) = 1$. Por otra parte $B_k := \{x \in X : \psi(x) \geq k\} = \bigcap_{n=1}^{\infty} (X \setminus A_{k-(1/n)})$, donde $\mu(X \setminus A_{k-(1/n)}) = 1$. Como $A_{k-(1/n+1)} \supset A_{k-(1/n)}$, obtenemos $\mu(B_k) = 1$. Por lo tanto $\mu(A_k \cap B_k) = 1$, es decir para μ -c.t.p. $x \in M$ se cumple $\psi(x) = k$.

(ii) \Rightarrow (iii):

Como μ es invariante por T tenemos $\int \psi \circ T^j d\mu = \int \psi d\mu \forall j \in \mathbb{N}$. Luego $\int \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j d\mu = \int \psi d\mu \forall n \geq 1$. Tomando límite cuando $n \rightarrow +\infty$ y aplicando el Teorema

de Convergencia Dominada, deducimos que

$$\int \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j d\mu = \lim_{n \rightarrow +\infty} \int \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j d\mu = \int \psi d\mu$$

Consideremos la siguiente función definida μ -c.t.p:

$$\Phi(x) := \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x) \mu - \text{c.t.p. } x \in X.$$

Aquí, en la hipótesis de existencia μ -c.t.p. de esta función Φ , es decir en la hipótesis de existencia μ -c.t.p. del límite de los promedios temporales, estamos aplicando el Teorema Ergódico de Birkhoff (como si ya estuviera demostrado). Entonces tenemos

$$\int \Phi d\mu = \int \psi d\mu.$$

Para tener definida la función Φ en todo punto $x \in X$, tomemos

$$\Phi(x) := 0 \quad \forall x \in X \text{ tal que } \nexists \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x).$$

Afirmamos que Φ es invariante por T . Basta chequear que dado $\epsilon > 0$ existe N tal que para todo $n \geq N$ se cumple

$$\begin{aligned} & \left| \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(T(x)) - \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x) \right| = \\ & = \left| \frac{1}{n} \psi(T^{n+1}(x)) - \psi(x) \right| < \epsilon \quad \forall x \in X. \end{aligned}$$

En efecto la desigualdad anterior se verifica para todo n suficientemente grande porque, por hipótesis, $0 \leq |\psi(x)| \leq k$ para todo $x \in X$, para cierta constante $k > 0$. Entonces, para toda sucesión $n_j \rightarrow +\infty$ tal que exista el límite del promedio temporal hasta n_j de ψ , con estado inicial x , también existe ese límite con estado inicial $T(x)$, y ambos límites coinciden. Deducimos que

$$\Phi(x) = \Phi(T(x)) \quad \forall x \in X.$$

Es decir, Φ es función real invariante por T . Aplicando la hipótesis (ii) tenemos que Φ es μ -c.t.p. constante, igual a cierto número real K . Entonces

$$\int \psi d\mu = \int \Phi d\mu = \int K d\mu = K.$$

Deducimos que $K = \int \psi d\mu$, es decir $\Phi(x) = \int \psi d\mu$ para μ -c.t.p. $x \in X$. Esto implica la igualdad (1.3), probando (iii).

(iii) \Rightarrow (i)

Sea $A \subset X$ medible e invariante por T , y consideremos la función característica χ_A . Es una función medible y acotada, y como A es invariante por T , para todo $x \in X$ tenemos $\chi_A \circ T^j(x) = \chi_{T^{-j}(A)}(x) = \chi_A(x) \in \{0, 1\}$. Entonces

$$\tilde{\chi}_A = \lim_{n \rightarrow +\infty} \sum_{j=0}^{n-1} \chi_A(T^j(x)) = \chi_A(x) \in \{0, 1\} \quad \forall x \in X.$$

En particular vale la igualdad anterior para μ -c.t.p. $x \in X$. Como

$$\int \tilde{\chi}_A d\mu = \int \chi_A d\mu = \mu(A) \quad \mu - \text{c.t.p. } x \in X$$

deducimos que $\mu(A) \in \{0, 1\}$. Por lo tanto μ es ergódica, terminando de probar (i). \square

1.7. Existencia de medidas ergódicas

Teorema 1.7.1. Existencia de medidas ergódicas *Sea X un espacio métrico compacto y sea $T : X \mapsto X$ continua. Entonces existen medidas de probabilidad ergódicas para T . Además, toda medida T -invariante es el límite en la topología débil* de una sucesión de medidas que son combinaciones lineales finitas de medidas ergódicas.*

Demostraremos el teorema 1.7.1 de existencia de medidas ergódicas al final de esta sección, en el párrafo 1.7.4.

Una generalización del Teorema 1.7.1 que establece la existencia de medidas ergódicas para toda $T : X \mapsto X$ medible (no necesariamente continua) tal que preserva alguna medida de probabilidad de Borel, será demostrada más adelante, en el Corolario 4.1.3.

1.7.2. Singularidad mutua y continuidad absoluta

Recordamos que dos medidas de probabilidad μ y ν se dicen mutuamente singulares $\mu \perp \nu$ cuando existe algún conjunto medible $A \subset X$ tal que $\mu(A) = 1$ y $\nu(A) = 0$. Entonces $\mu(X \setminus A) = 0$ y $\nu(X \setminus A) = 1$ y la relación mutuamente singular es simétrica. Dadas dos medidas de probabilidad μ y ν se dice que μ es absolutamente continua respecto de ν , y se denota $\mu \ll \nu$, cuando para todo conjunto medible A tal que $\nu(A) = 0$ se cumple $\mu(A) = 0$.

Se dice que dos medidas μ y ν son equivalentes, y se denota $\mu \sim \nu$, cuando $\mu \ll \nu$ y $\nu \ll \mu$.

Se observa que si $\mu \ll \nu$ entonces $\mu \not\ll \nu$ (el recíproco es falso).

El siguiente teorema es clásico en la Teoría abstracta de la Medida (en particular en la Teoría de Probabilidades):

Teorema de Descomposición de Lebesgue-Radon-Nikodym.

Dadas dos medidas de probabilidad μ y ν existen dos probabilidades μ_1 y μ_2 , y un único real $t \in [0, 1]$, tales que

$$\mu = t\mu_1 + (1-t)\mu_2, \quad \mu_1 \ll \nu, \quad \mu_2 \perp \nu.$$

Si además $t \neq 0, 1$ entonces μ_1 y μ_2 son únicas.

El enunciado clásico de este Teorema, establece que *existen únicas las medidas finitas* $t\mu_1$ y $(1-t)\mu_2$ (posiblemente alguna de ellas es cero) tales que sumadas dan μ , siendo $t\mu_1 \ll \nu$ y $(1-t)\mu_2 \perp \nu$.

La demostración del Teorema de Descomposición de Lebesgue-Radon-Nikodym se encuentra por ejemplo en [Fo 1984, Theorem 3.8] ó en [Rud 1979a, Teorema 6.2.3].

Del teorema de Lebesgue-Radon-Nikodym se deduce que en el caso particular $\mu \ll \nu$, se cumple $t = 1$, $\mu_1 = \mu$, y μ_2 es cualquiera. Análogamente, si $\mu \perp \nu$, entonces $t = 0$, $\mu_2 = \mu$ y μ_1 es cualquiera. En el caso que $\mu \not\ll \nu$, $\mu \not\perp \nu$, es única la pareja (μ_1, μ_2) de probabilidades en la descomposición de Radon-Nikodym.

Volvamos ahora a las propiedades de las medidas de probabilidad ergódicas para una transformación T :

Teorema 1.7.3. Singularidad mutua de medidas ergódicas

Sea $T : X \mapsto X$ medible en el espacio medible (X, \mathcal{A}) .

(a) Si existen dos medidas de probabilidad diferentes μ y ν , ambas ergódicas para la transformación T , entonces $\mu \perp \nu$.

(b) Si μ y ν son medidas de probabilidad ergódicas para T y si $\mu \ll \nu$, entonces $\mu = \nu$.

Demostración. Afirmamos que, si μ y ν son ambas ergódicas para T , y si para todo conjunto A invariante por T se cumple $\mu(A) = \nu(A)$, entonces $\mu = \nu$. En efecto, sea B cualquier conjunto medible, y denotemos χ_B a la función característica de B . Aplicando la propiedad (iii) del Teorema 1.6.9 a la función $\psi = \chi_B$, sabemos que existe un conjunto A_1 , que es T -invariante, tal que $\mu(A_1) = 1$, y se cumple la igualdad (1.3) para la medida μ y para todo $x \in A_1$. Análogamente, existe un conjunto A_2 , que es T -invariante, tal que $\nu(A_2) = 1$, y se cumple la igualdad (1.3) para la medida ν y para todo $x \in A_2$. Por hipótesis $\nu(A_1) = \mu(A_1) = 1 = \nu(A_2) = \mu(A_2)$. Entonces $\nu(A_1 \cap A_2) = \mu(A_1 \cap A_2) = 1$ y para todos los puntos $x \in A_1 \cap A_2$ se cumple la igualdad

$$\mu(B) = \int \chi_B d\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_B \circ T^j(x) = \int \chi_B d\nu = \nu(B)$$

Deducimos entonces que $\mu(B) = \nu(B)$ para todo conjunto medible B , de donde $\mu = \nu$. Hemos terminado de probar la afirmación del principio.

Demostremos ahora la parte (a) del teorema 1.7.3. Debido a la afirmación recién demostrada, como $\nu \neq \mu$ y son ambas ergódicas, existe un conjunto A que es T -invariante, tal que $\mu(A) \neq \nu(A)$. Por definición de ergodicidad $\mu(A), \nu(A) \in \{0, 1\}$. Luego (eventualmente sustituyendo A por su complemente en caso necesario), tenemos $\mu(A) = 1$ y $\nu(A) = 0$, de donde $\mu \perp \nu$, demostrando (a).

La parte (b) es una consecuencia inmediata de (a), pues si $\mu \ll \nu$ entonces $\mu \not\ll \nu$. Por lo tanto, como μ y ν son ergódicas y $\mu \not\ll \nu$, aplicando la parte (a) deducimos que $\mu = \nu$, como queríamos probar. \square

1.7.4. .

Demostración del Teorema 1.7.1 (existencia de medidas ergódicas)

El teorema 1.7.1 es un corolario inmediato del teorema 1.7.5 que demostraremos a continuación. Las definiciones de *puntos extremales de un conjunto compacto y convexo* en un espacio vectorial topológico, y de *envolvente compacta convexa*, se emplean en el enunciado siguiente, y se incluyen abajo del mismo.

Teorema 1.7.5. *Sea $T : X \mapsto X$ continua en el espacio métrico compacto X . Sea \mathcal{M} el conjunto de medidas de probabilidad borelianas en X , con la topología débil*. Sea $\mathcal{M}_T \subset \mathcal{M}$ el conjunto de probabilidades invariantes por T y sea $\mathcal{E}_T \subset \mathcal{M}_T$ el conjunto de las probabilidades (invariantes) ergódicas para T .*

Entonces

- (a) \mathcal{M}_T es compacto y convexo.
- (b) \mathcal{E}_T coincide con el conjunto de puntos extremales de \mathcal{M}_T .
- (c) \mathcal{M}_T coincide con la envolvente convexa compacta de \mathcal{E}_T .

Demostraremos el Teorema 1.7.1 más adelante en esta sección. Para poder demostrarlo, necesitamos definir convexidad, envolvente convexa, puntos extremales y ver las propiedades de estos conceptos:

Combinaciones convexas y puntos extremales

Recordemos las siguientes definiciones y el teorema de Krein-Milman del Análisis Funcional:

- Sea \mathcal{A} un conjunto no vacío de un espacio vectorial topológico. Se llama *combinación convexa* de puntos de \mathcal{A} a cualquier punto del espacio que pueda escribirse como una combinación lineal finita

$$t_1 a_1 + t_2 a_2 + \dots + t_k a_k,$$

tal que $a_i \in \mathcal{A}$, $0 \leq t_i \leq 1$ para todo $1 \leq i \leq k$, y $\sum_{i=1}^k t_i = 1$. Denotamos como $ec(\mathcal{A})$ al conjunto de todas las combinaciones convexas de puntos de \mathcal{A} , llamado *envolvente convexa* de \mathcal{A} . Observar que $\mathcal{A} \subset ec(\mathcal{A})$.

- El conjunto \mathcal{A} se dice *convexo* si contiene a todas las combinaciones convexas de sus puntos, es decir $\mathcal{A} = ec(\mathcal{A})$. Es inmediato deducir, por inducción en $k \geq 2$, que un conjunto \mathcal{A} es convexo, si y solo si

$$ta + (1 - t)b \in \mathcal{A} \quad \forall t \in [0, 1], \quad \forall a, b \in \mathcal{A}.$$

- Se llama *envolvente convexa cerrada* de \mathcal{A} a $\overline{ec(\mathcal{A})}$, donde $\overline{}$ indica la clausura (o adherencia).
- Si \mathcal{A} es tal que $\overline{ec(\mathcal{A})}$ es compacto, llamaremos a este último conjunto *la envolvente convexa compacta* de \mathcal{A} .

• Si \mathcal{K} es un subconjunto no vacío de un espacio vectorial topológico, se llama *punto extremal* de \mathcal{K} (cuando existe) a un punto $a \in \mathcal{K}$ tal que las únicas combinaciones convexas $tb + (1-t)c = a$, con $0 \leq t \leq 1$ $b, c \in \mathcal{K}$, son aquellas para las cuales $b = a$ ó $c = a$.

• **Teorema de Krein-Milman**

Todo compacto no vacío y convexo en un espacio vectorial topológico contiene puntos extremales y coincide con la envolvente convexa compacta del conjunto de sus puntos extremales.

La demostración del Teorema de Krein-Milman puede encontrarse por ejemplo en [Rud 1979b, Teorema 3.21].

Ejercicio 1.7.6. Sean $x_1, x_2 \in \mathbb{R}^2$. Chequear que toda combinación convexa de $\{x_1, x_2\}$ está en el segmento con extremos x_1, x_2 y recíprocamente. Para ilustrar el teorema de Krein-Milman, considerar un polígono regular K en \mathbb{R}^2 . (K es la unión del interior del polígono con su borde). Entonces K es compacto. Chequear con argumentos geométricos que: (a) el polígono K es convexo; (b) los puntos extremales de K son los vértices del polígono; (c) todo punto de K es una combinación convexa de sus vértices. En el ejemplo del polígono, la cantidad de puntos extremales es finita. Para ilustrar el teorema de Krein-Milman cuando la cantidad de puntos extremales es infinita, considerar una circunferencia $S^1 \subset \mathbb{R}^2$ y el compacto K que es la unión de la circunferencia S^1 (borde de K) con la región acotada encerrada por S^1 (interior de K). Entonces K es compacto. Chequear con argumentos geométricos que: (d) K es convexo; (e) el conjunto de puntos extremales de K es la circunferencia S^1 ; (f) todo punto de K es combinación convexa de dos puntos extremales.

Demostración del Teorema 1.7.5

Demostración. (a) Es inmediato chequear que \mathcal{M}_T es convexo: en efecto, si $\mu_1, \mu_2 \in \mathcal{M}_T$ entonces para todo conjunto boreliano B se cumple

$$\mu_i(T^{-1}(B)) = \mu_i(B) \quad i = 1, 2.$$

Luego, cualquiera sea $0 \leq t \leq 1$ se tiene

$$[t\mu_1 + (1-t)\mu_2](T^{-1}(B)) = [t\mu_1 + (1-t)\mu_2](B),$$

de donde $t\mu_1 + (1-t)\mu_2 \in \mathcal{M}_T$, probando que \mathcal{M}_T es convexo.

Ahora probemos que \mathcal{M}_T es compacto. Recordemos que \mathcal{M} es compacto con la topología débil* (esto fue demostrado en la sección 1.2). Luego, para demostrar que \mathcal{M}_T es compacto, basta probar que \mathcal{M}_T es cerrado en \mathcal{M} . Sea $\mu_n \rightarrow \mu$ en \mathcal{M} tales que $\mu_n \in \mathcal{M}_T$. Para deducir que $\mu \in \mathcal{M}_T$ basta recordar que es continuo el operador $T^* : \mathcal{M} \rightarrow \mathcal{M}$, definido como $[T^*\mu](B) = \mu(T^{-1}(B))$ para todo boreliano B . (La continuidad de T^* fue demostrada en la sección 1.2). Entonces

$$T^*\mu = \lim_n T^*\mu_n = \lim_n \mu_n = \mu,$$

de donde $\mu \in \mathcal{M}_T$. Esto termina la prueba de que \mathcal{M}_T es cerrado en el espacio métrico compacto \mathcal{M} , y por lo tanto \mathcal{M}_T es compacto.

(b) Probemos que μ es ergódica para T si y solo si es punto extremal de \mathcal{M}_T . Primero asumamos que μ no es ergódica. Entonces existe un boreliano T -invariante $A \subset X$ tal que $0 < \mu(A) < 1$. Sea $A^c = X \setminus A$ y considérense las probabilidades definidas para todo boreliano B por:

$$\mu_1(B) := \frac{\mu(B \cap A)}{\mu(A)}, \quad \mu_2(B) := \frac{\mu(B \cap A^c)}{\mu(A^c)}.$$

Por construcción, tenemos

$$\mu = t\mu_1 + (1-t)\mu_2, \quad \text{donde } t := \mu(A).$$

Nótese que $\mu_1 \neq \mu_2$, porque $\mu_1(A) = 1$ y $\mu_2(A) = 0$. Además $0 < t < 1$, de donde

$$\mu \neq \mu_1, \quad \mu \neq \mu_2.$$

Por lo tanto μ no es punto extremal de \mathcal{M}_T . Hemos probado que si μ no es ergódica, entonces μ no es punto extremal de \mathcal{M}_T .

Ahora probemos el recíproco. Asumamos que μ es ergódica y probemos que μ es punto extremal de \mathcal{M}_T . Sea $\mu = t\mu_1 + (1-t)\mu_2$, donde $0 \leq t \leq 1$ y $\mu_1, \mu_2 \in \mathcal{M}_T$. Por un lado, si $t = 0$ ó $t = 1$ entonces $\mu = \mu_1$ ó $\mu = \mu_2$. Por otro lado, si $0 < t < 1$ entonces $\mu_1 \ll \mu$ y $\mu_2 \ll \mu$ (pues si $\mu(B) = 0$ siendo la suma de dos sumandos no negativos, cada sumando debe ser cero, de dónde $0 = \mu_1(B) = \mu_2(B)$). Sea A un conjunto T -invariante. Por hipótesis μ es ergódica, entonces $\mu(A) = 0$ ó $\mu(A^c) = 0$. Como $\mu_i \ll \mu$ para $i = 1, 2$, deducimos que $\mu_i(A) = 0$ ó $\mu_i(A) = 1$. Luego μ_i es ergódica. Por lo demostrado en el Teorema 1.7.3, $\mu_i \ll \mu$ siendo μ_i y μ ergódicas, implica que $\mu_i = \mu$. En este caso tenemos entonces $\mu = \mu_1 = \mu_2$. Hemos probado, en todos los casos, que las únicas combinaciones convexas de μ son aquellas para las cuales $\mu = \mu_1$ ó $\mu = \mu_2$. Entonces por definición, μ es extremal como queríamos demostrar.

(c) Es consecuencia directa de a) y b), y del Teorema de Krein Milman. \square

Fin de la prueba del Teorema 1.7.1: Existencia de medidas ergódicas

Demostración. Por la parte c) del Teorema 1.7.5,

$$\mathcal{M}_T = \overline{e.c.\mathcal{E}_T},$$

donde \mathcal{E}_T denota el conjunto de medidas ergódicas para T y \mathcal{M}_T el de todas las medidas T -invariantes. Por el Teorema 1.1.5, $\mathcal{M}_T \neq \emptyset$. Luego $\mathcal{E}_T \neq \emptyset$. Para probar la última afirmación del Teorema 1.1.5 aplicamos la Definición de la clausura $\overline{e.c.\mathcal{E}_T}$ de la envolvente convexa del conjunto \mathcal{E}_T de probabilidades ergódicas. Recordamos que la envolvente convexa $e.c.\mathcal{E}_T$ es el conjunto de todas las medidas que se obtienen como combinaciones lineales convexas de medidas ergódicas. Como toda medida invariante μ está en $\overline{e.c.\mathcal{E}_T}$, entonces μ se puede aproximar, tanto como uno desee en la topología débil* del espacio de probabilidades, por medidas que son por combinaciones lineales finitas y convexas de medidas ergódicas. \square

Ejercicio 1.7.7. Sea $T : X \mapsto X$ una transformación continua en un espacio métrico compacto.

(I) Asuma que existe una medida μ invariante por T y positiva sobre abiertos (no necesariamente ergódica).

(a) Probar que para cada abierto V no vacío, existe una medida ergódica ν_V tal que $\nu_V(V) > 0$.

Sugerencia: Por absurdo, si $\nu(V) = 0$ para toda medida ergódica ν , entonces $\rho(V) = 0$ para toda medida de probabilidad ρ que sea combinación convexa finita de ergódicas. Probar que $\rho(V) = 0$ para toda medida ρ en la envolvente convexa compacta de las ergódicas (con la topología débil* del espacio \mathcal{M} de probabilidades). Usar el Teorema 1.7.5 para concluir que $\mu(V) = 0$, contradiciendo la hipótesis.

(b) Usando el Lema de Recurrencia de Poincaré demostrar que el conjunto de los puntos recurrentes es denso en X .

(c) Concluir que $\Omega(T) = X$.

(II) Asuma que para todo abierto V existe una medida de probabilidad invariante μ_V (no necesariamente ergódica), tal que $\mu_V(V) > 0$.

(d) Demostrar que existe una medida de probabilidad invariante ν_V ergódica tal que $\nu_V(V) > 0$.

(e) Demostrar que existe una medida de probabilidad invariante ρ positiva sobre abiertos. Sugerencia: Tomar una base numerable de abiertos $\{V_i\}_{i \geq 1}$, demostrar que las medidas $\rho_n := \sum_{i=1}^n (1/2^i) \mu_{V_i}$ (que no son probabilidades, pero son finitas) satisfacen $0 < \rho_n(X) \leq 1$, son T -invariantes y existe $\rho = \lim_{n \rightarrow +\infty}^* \rho_n$ en la topología débil* del espacio \mathcal{M}^1 de medidas de probabilidad finitas uniformemente acotadas por 1.

Capítulo 2

Teoremas Ergódicos.

Las hipótesis generales para este capítulo son las siguientes:

(X, \mathcal{A}) es un espacio medible, $T : X \mapsto X$ es una transformación medible que preserva una medida de probabilidad μ , y $f : X \mapsto \mathbb{R}$ es una función medible.

Consideremos el siguiente resultado:

Sea $T : X \mapsto X$ medible en el espacio métrico compacto X . Sea μ una medida de probabilidad en la sigma-álgebra de Borel. Entonces μ es T -invariante si y solo si para toda $f : X \mapsto \mathbb{R}$ continua se cumple

$$\int f d\mu = \int f \circ T d\mu \quad (2.1)$$

Además, vale la igualdad (2.1) para toda f continua si y solo si vale para toda $f \in L^p(\mu)$, cualquiera sea el natural $p \geq 1$.

Ejercicio 2.0.8. Probar la afirmación anterior.

2.1. Teorema Ergódico de Birkhoff-Khinchin Enunciado y Corolarios

Definición 2.1.1. Promedios temporales Sea (X, \mathcal{A}) un espacio medible. Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ (no necesariamente ergódica). Sea $f \in L^p(\mu)$ para $1 \leq p \in \mathbb{N}$.

Se denota con \tilde{f}^+ o simplemente con \tilde{f} al límite de los llamados *promedios orbitales* (o *promedios temporales* o *promedios de Birkhoff*) hacia el futuro de f , en los puntos $x \in X$ donde exista, esto es:

$$\tilde{f}^+(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} (f(x) + f \circ T(x) + \dots + f \circ T^{n-1}(x))$$

Si además T es invertible, se denota con \tilde{f}^- al límite de los promedios orbitales (o

temporales o de Birkhoff) hacia el pasado, en los puntos x donde exista. Esto es:

$$\tilde{f}^-(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} (f(x) + f \circ T^{-1}(x) + \dots + f \circ T^{-(n-1)}(x))$$

Teorema 2.1.2. Teorema ergódico de Birkhoff-Khinchin

Sea (X, \mathcal{A}) un espacio medible. Si $T : X \mapsto X$ es medible que preserva una medida de probabilidad μ entonces:

- Para toda $f \in L^1(\mu)$ existe $\tilde{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x)$ μ -c.t.p. en $x \in X$.
- \tilde{f} es T -invariante, es decir : $\tilde{f} \circ T = \tilde{f}$ μ -c.t.p. Más precisamente, para todo $x \in X$ existe $\tilde{f}(x)$ si y solo si existe $\tilde{f}(T(x))$ y en ese caso $\tilde{f}(x) = \tilde{f}(T(x))$.
- Para todo natural $p \geq 1$, si $f \in L^p(\mu) \subset L^1(\mu)$, entonces $\tilde{f} \in L^p(\mu)$ y la convergencia es también en $L^p(\mu)$.
- $\int f d\mu = \int \tilde{f} d\mu$

Daremos la demostración del Teorema Ergódico de Birkhoff-Khinchin en la Sección 2.3 de este capítulo. La demostración de Birkhoff del Teorema 2.1.2 se encuentra en [Bi 1931]. Otra demostración, diferente de la de Birkhoff, puede encontrarse por ejemplo, en [Wa 2000, proof of Theorem 1.14, pag. 38-39], en [Ke 1998, Theorem 2.1.5] o en [Mañ 1983a, pág. 114-122] Mane (ver también [Mañ 1987]).

El Teorema Ergódico de Birkhoff-Khinchin es un caso particular del llamado **Teorema Ergódico Subaditivo de Kingmann** [Kin 1973]. Este teorema más general, establece la convergencia de la sucesión $\{f_n/n\}_n$ donde f_n , en vez de ser necesariamente una suma de Birkhoff, es una *sucesión subaditiva de funciones* en $L^1(\mu)$. Más precisamente, se asume por hipótesis, que

$$f_{n+m} \leq f_n + f_m \circ T^n \quad \forall n, m \in \mathbb{N}^+,$$

donde $f_1 \in L^1(\mu)$. En particular, $f_n = \sum_{j=0}^{n-1} f_1 \circ T^j$ es un ejemplo de sucesión subaditiva.

El Teorema Ergódico Subaditivo de Kingmann enuncia que, para toda medida μ que sea invariante por T , para toda sucesión $\{f_n\}_{n \geq 1}$ subaditiva de funciones reales en $L^1(\mu)$, la sucesión $\{f_n/n\}_n$ converge μ -c.t.p. Luego, este Teorema generaliza el Teorema Ergódico de Birkhoff-Khinchin. Demostraremos el Teorema Ergódico Subaditivo de Kingmann (Teorema 4.2.2) más adelante de este libro.

Corolario 2.1.3. del Teorema 2.1.2

Igualdad de los promedios temporales hacia el futuro y hacia el pasado.

Si T es medible, invertible, con inversa medible, y preserva una medida de probabilidad μ , entonces para toda $f \in L^1(\mu)$ se cumple \tilde{f}^+ (promedio temporal hacia el futuro) y \tilde{f}^- (promedio temporal hacia el pasado) existen μ -c.t.p. y son iguales $\tilde{f}^+ = \tilde{f}^-$ μ -c.t.p.

(Ver la definición de \tilde{f}^+ y \tilde{f}^- en 2.1.1.)

En el siguiente Ejercicio 2.1.4, se da una guía para la demostración del Corolario 2.1.3.

Ejercicio 2.1.4. Probar el corolario 2.1.3 como consecuencia del Teorema de Birkhoff-Khinchin. Sugerencias: Para demostrar que existen $\tilde{f}^+(x)$ y $\tilde{f}^-(x)$ para μ -c.t.p. $x \in X$, aplicar el teorema de Birkhoff a T^{-1} , probar que toda medida μ es invariante por T si y solo sí lo es por T^{-1} , y usar que la intersección de dos conjuntos con μ -medida igual a 1 tiene medida μ -medida 1. Para probar que $\tilde{f}^+(x) = \tilde{f}^-(x)$ para μ -c.t.p. $x \in X$, para cada natural $n \geq 1$ denote $f_n^+ := (1/n) \sum_{j=0}^{n-1} f \circ T^j$, $f_n^- := (1/n) \sum_{j=0}^{n-1} f \circ T^{-j}$. Por el Teorema de Birkhoff, $f_n^- - f_n^+$ converge en $L^1(\mu)$ a $\tilde{f}^- - \tilde{f}^+$. Entonces para todo $\epsilon > 0$ existe n suficientemente grande tal que $\|\tilde{f}^- - \tilde{f}^+\|_{L^1} \leq \|f_n^- - f_n^+\|_{L^1} + \epsilon$. Basta probar entonces que $\int |f_n^- - f_n^+| d\mu$ tiende a cero cuando $n \rightarrow +\infty$. Chequear que para todo $x \in X$ se cumple la siguiente igualdad:

$$\begin{aligned} f_{n+1}^-(x) - f_{n+1}^+(x) &= f_{n+1}^-(x) + f_{n+1}^+(x) - 2f_{n+1}^+(x) = \\ &= \frac{2n+1}{n} f_{2n+1}^+(T^{-n}(x)) + \frac{f(x)}{n+1} - 2f_{n+1}^+(x) \end{aligned}$$

Como f_{2n+1}^+ converge a \tilde{f}^+ en $L^1(\mu)$, y además la medida μ y la función \tilde{f}^+ son invariantes por T , tenemos, para todo n suficientemente grande:

$$\begin{aligned} \int |f_{2n+1}^+ - \tilde{f}^+| d\mu &< \epsilon \\ \int |f_{2n+1}^+ - \tilde{f}^+| d\mu &= \int |f_{2n+1}^+ \circ T^{-n} - \tilde{f}^+ \circ T^{-n}| d\mu = \\ &= \int |f_{2n+1}^+ \circ T^{-n} - \tilde{f}^+| d\mu < \epsilon \end{aligned}$$

Juntando todo lo anterior, deducir, para todo n suficientemente grande, que:

$$\begin{aligned} \|f_{n+1}^- - f_{n+1}^+\|_{L^1(\mu)} &\leq \\ \frac{2n+1}{n+1} \|f_{2n+1}^+ \circ T^{-n} - \tilde{f}^+\|_{L^1(\mu)} &+ \frac{|f(x)|}{n+1} + 2\|f_{n+1}^+ - \tilde{f}^+\|_{L^1(\mu)} \\ &\leq 3\epsilon + \epsilon + 2\epsilon = 6\epsilon. \end{aligned}$$

Corolario 2.1.5. del Teorema 2.1.2

Promedios de medida de transitividad. Para toda medida μ invariante por T y para todos los conjuntos A y B medibles, existe el límite siguiente:

$$\tau(A, B) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B)$$

Demostración: Denotemos χ_C a la función característica de cualquier conjunto C . Sea

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

$$\frac{1}{n} \sum_{j=0}^{n-1} \int \chi_{T^{-j}(A)} \chi_B d\mu = \int \left(\frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j \right) \chi_B d\mu,$$

El integrando a la derecha de la igualdad anterior está dominado por $1 \in L^1(\mu)$. Por el teorema de convergencia dominada $\lim I_n = \int \tilde{\chi}_A \chi_B d\mu$ \square

Ejercicio 2.1.6. Probar que el límite $\tau(A, B)$ del Corolario 2.1.5 verifica las siguientes desigualdades

$$\mu(A) - \sqrt{\mu(A)[1 - \mu(B)]} \leq \tau(A, B) \leq \sqrt{\mu(A)\mu(B)}.$$

Concluir que:

$\tau(A, B) = 0$ si $\mu(A) = 0$ ó $\mu(B) = 0$.

Si $\tau(A, B) = 0$ entonces $\mu(A) + \mu(B) \leq 1$

$\tau(A, B) = 1$ si y solo si $\mu(A) = \mu(B) = 1$.

Sugerencia para la primera parte: En la demostración del Corolario 2.1.5 se probó que $\tau(A, B) = \int \tilde{\chi}_A \chi_B d\mu$. En $L^2(\mu)$ se cumple

$$\int fg d\mu \leq \|f\|_{L^2} \|g\|_{L^2}.$$

Para probar la desigualdad de la derecha, aplicar lo anterior a $\tilde{\chi}_A$ y χ_B , y recordar que $0 \leq \tilde{\chi}_A \leq 1$, por lo cual $\tilde{\chi}_A^2 \leq \tilde{\chi}_A$. Deducir que $\|\tilde{\chi}_A\|_{L^2} \leq \sqrt{\mu(A)}$ aplicando el teorema de Birkhoff. Para probar la desigualdad de la izquierda, aplicar la desigualdad de la derecha a A y B^c y probar que $\tau(A, B) + \tau(A, B^c) = \mu(A)$.

Corolario 2.1.7. del Teorema 2.1.2

Promedios de sucesión de funciones.

Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Sea $f_n \in L^1(\mu)$ una sucesión de funciones, dominada por $f_0 \in L^1(\mu)$, que converge μ -c.t.p. y en $L^1(\mu)$ a $f \in L^1(\mu)$. Entonces

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_j \circ T^j = \tilde{f} \quad \mu\text{-c.t.p. y en } L^1(\mu).$$

Ejercicio 2.1.8. Demostrar el corolario 2.1.7. Sugerencia: Basta probarlo para $f_n \geq 0$; $f_n \rightarrow 0$ c.t.p. Sea $G_k(x) = \sup_{n \geq k} \{f_n(x)\}$. Entonces $G_k \rightarrow 0$ c.t.p., y por convergencia dominada $\|G_k\|_{L^1} \rightarrow 0$. Sea $\tilde{G}_k = \lim_{n \rightarrow \infty} (1/n) \sum_{j=0}^{n-1} G_k \circ T^j$ que existe μ -c.t.p. por el teorema de Birkhoff. La sucesión \tilde{G}_k es decreciente con k , por lo que tiene límite, y por el lema de Fatou

$$0 \leq \int \lim \tilde{G}_k d\mu \leq \lim \int \tilde{G}_k d\mu = \lim \int G_k d\mu = 0.$$

Luego $\tilde{G}_k \rightarrow 0$ c.t.p.. Finalmente, usar que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f_j \circ T^j(x) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{j=k}^{n-1} f_j \circ T^j(x) \leq \tilde{G}_k(x).$$

Definición 2.1.9. Dado A conjunto medible, se denomina *tiempo medio de estadía* $\tau_A(x)$ de un punto $x \in X$ en A , al siguiente límite, cuando existe:

$$\tau_A(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \#\{j \in \mathbb{N} : 0 \leq j \leq n-1, T^j(x) \in A\}$$

Ejercicio 2.1.10. Probar el siguiente teorema, para todo conjunto medible $A \subset X$ y para toda medida μ que sea invariante por $T : X \mapsto X$:

El tiempo medio de estadía τ_A existe μ -c.t.p. y además la convergencia es en $L^p(\mu)$ para todo $p \geq 1$ natural. Además $\int \tau_A d\mu = \mu(A)$. Sugerencia: Observar que el tiempo medio de estadía en A es la función $\tau_A = \tilde{\chi}_A$, donde χ_A es la función característica de A . Aplicar el teorema ergódico de Birkhoff-Khinchin.

Definición 2.1.11. Conjuntos de probabilidad total para T Sea $T : X \mapsto X$ una transformación medible tal que el conjunto \mathcal{M}_T de las medidas de probabilidad T -invariantes es no vacío. Un conjunto medible $\Lambda \subset X$ se dice que tiene *probabilidad total para T* si para toda $\mu \in \mathcal{M}_T$ se cumple $\mu(\Lambda) = 1$.

Entonces, la primera parte del ejercicio anterior se puede enunciar de la siguiente forma:

Para todo conjunto medible $A \subset X$, el tiempo medio de estadía $\tau_A(x)$ existe para un conjunto de puntos con probabilidad total. Si además $\mu(A) > 0$, entonces $\tau_A(x)$ no nula μ -c.t.p. (pues $\int \tau_A d\mu = \mu(A)$).

En el Corolario 4.1.4 veremos un criterio para que un conjunto tenga probabilidad total, que requiere solo el conocimiento de las medidas ergódicas.

Ejercicio 2.1.12. Conjuntos estables e inestables.

Sea $T : X \mapsto X$ Borel medible e invertible con inversa medible, en un espacio métrico compacto X , tal que el conjunto de medidas invariantes por T es no vacío. Sea $f : X \mapsto \mathbb{C}$ una función compleja continua. Sea $\Lambda \subset X$ el conjunto de probabilidad total tal que existen, y son iguales entre sí, los límites $\tilde{f}^+(x)$ y $\tilde{f}^-(x)$ de los promedios de Birkhoff hacia el futuro y hacia el pasado, respectivamente. Sea $x_0 \in \Lambda$. Se definen los conjuntos estable e inestable respectivamente por el punto x_0 (quizás se reducen solo a $\{x_0\}$):

$$W^s(x_0) := \{y \in X : \lim_{n \rightarrow +\infty} \text{dist}(T^n y, T^n x_0) = 0\}$$

$$W^u(x_0) := \{y \in X : \lim_{n \rightarrow -\infty} \text{dist}(T^n y, T^n x_0) = 0\}$$

Probar que para todo $y \in W^s(x_0)$ existe $\tilde{f}^+(y)$ y $\tilde{f}^+(y) = \tilde{f}^+(x_0)$. Probar que para todo $y \in W^u(x_0)$ existe $\tilde{f}^-(y)$ y $\tilde{f}^-(y) = \tilde{f}^-(x_0)$.

Ejercicio 2.1.13. .

Promedios de Birkhoff para funciones reales fuera de $L^1(\mu)$.

Probar la siguiente generalización del teorema de Birkhoff-Khinchin: Sea (X, \mathcal{A}) un espacio medible y sea $T : X \mapsto X$ medible que preserve la probabilidad μ . Sea $f : X \mapsto \mathbb{R}$ medible.

Entonces, μ -c.t.p. o bien $|\tilde{f}|(x) = +\infty$ o bien \tilde{f} existe y es finito.

Sugerencia: Basta probarlo para $f \geq 0$. Dado $c > 0$ sea

$$X_c = \{x \in X : \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) \leq c\}$$

X_c es T -invariante. Sea $f_m(x) = \min(f(x), m)$. Usar el teorema de Birkhoff para probar que existe \tilde{f}_m μ -c.t.p. $x \in X_c$. Sea $f|_{X_c} = \chi_{X_c} f$ donde χ_{X_c} denota la función característica de X_c . Probar que $f|_{X_c} \in L^1(\mu)$ usando el teorema de convergencia monótona y la igualdad del teorema de Birkhoff $\int f_m|_{X_c} d\mu = \int \widetilde{f_m}|_{X_c} d\mu \leq c$. Deducir que existe el límite $\tilde{f}(x)$ para μ -c.t.p. x en X_c . Tomar la unión de los X_c para todo $c \geq 1$ natural y concluir que en el complemento de esa unión se cumple $\tilde{f} = +\infty$.

Ejercicio 2.1.14. Sea $T : X \mapsto X$ una transformación medible en un espacio métrico compacto X , tal que es no vacío el conjunto \mathcal{M}_T de probabilidades invariantes por T . Probar que el conjunto siguiente tiene probabilidad total para T (i.e. tiene probabilidad 1 para toda $\mu \in \mathcal{M}_T$):

$$\{x \in X : \text{existe } \lim_{n \rightarrow +\infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} f \circ T^{pj}(x) \quad \forall f \in C^0(X, \mathbb{R}), \quad \forall p \geq 1\}$$

Sugerencia: $\mathcal{M}_{T^p}(X) \supset \mathcal{M}_T(X)$.

2.2. Teorema Ergódico Maximal

En esta sección asumiremos que $T : X \mapsto X$ es medible y que preserva una medida de probabilidad μ .

El siguiente Teorema Ergódico Maximal es un lema para demostrar el Teorema Ergódico de Birkhoff-Khinchin, pero puede tener interés en sí mismo pues (en apariencia) es más general. En la demostración del Teorema Ergódico Maximal, no asumiremos válido el Teorema de Birkhoff-Khinchin, como tampoco los Teoremas 1.7.1 y 1.6.9 de existencia y propiedades de las medidas ergódicas (que fueron probadas en el primer capítulo, asumiendo verdadero el Teorema Ergódico de Birkhoff-Khinchin.) De esa forma, en las próximas secciones podremos demostrar el Teorema de Birkhoff-Khinchin como consecuencia del Teorema Ergódico Maximal.

A diferencia del teorema de Birkhoff-Khinchin, este otro resultado, el Teorema Ergódico Maximal, no se refiere a los promedios temporales de una función f a lo largo de la órbita por un punto $x \in X$, sino al máximo o supremo de las sumas acumuladas de los valores de f a lo largo de la órbita con estado inicial x .

Teorema 2.2.1. Teorema Ergódico Maximal (Versión Particular).

Sea $f \in L^1(\mu)$ real. Sea, para $n \geq 1$, la sucesión creciente de funciones:

$$F_n = \max\{f, \quad f + f \circ T, \quad f + f \circ T + f \circ T^2, \quad \dots, \quad f + f \circ T + \dots + f \circ T^{n-1}\}$$

Sea

$$E(f) = \{x \in X : \sup_{n \geq 1} F_n(x) > 0\} = \{x \in X : \sup_{n \geq 1} \sum_{j=0}^{n-1} f \circ T^j(x) > 0\}$$

Entonces

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

Interpretación: Si la función $f : X \mapsto \mathbb{R}$ es real e integrable, aunque no sea siempre positiva, siempre sucede que es no negativo su valor esperado (con respecto a cualquier medida T -invariante) en el conjunto de puntos iniciales cuyas órbitas futuras tienen sumas acumuladas de f que a veces, aunque sea en un instante solo y muy lejano en el futuro, toman valores positivos.

Para demostrar el Teorema Ergódico Maximal 2.2.1, enunciaremos y demostraremos una versión más general, de la cual la anterior se deduce inmediatamente como aplicación. Para eso introducimos la siguiente definición:

Definición 2.2.2. Sucesión subaditiva de funciones Una sucesión $\{f_n\}_{n \geq 1}$ de funciones reales medibles $f_n : X \mapsto \mathbb{R}$ se llama *subaditiva* si

$$f_{n+m} \leq f_n + f_m \circ T^n \quad \forall n, m \geq 1.$$

La sucesión se llama *aditiva* si se cumple la igualdad en la afirmación anterior para todos $n, m \geq 1$.

Ejercicio 2.2.3. Probar que una sucesión de funciones reales medibles $\{f_n\}_{n \geq 1}$ es aditiva si y solo si

$$f_n = \sum_{j=0}^{n-1} f_1 \circ T^j \quad \forall n \geq 1.$$

Debido a lo probado en el ejercicio anterior, la sucesión de las sumas de Birkhoff de una función medible real f es una sucesión aditiva, y recíprocamente. Además como toda sucesión aditiva es subaditiva, entonces toda afirmación sobre f_n/n que se demuestre en general para las sucesiones subaditivas $\{f_n\}_{n \geq 1}$, regirá en particular para los promedios de Birkhoff.

Ejercicio 2.2.4. Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva tal que $f_1^+ \in L^1(\mu)$. Probar que $f_n^+ := \max\{0, f_n\} \in L^1(\mu)$ para todo $n \geq 1$ para toda medida de probabilidad μ invariante por T . Luego $\int (f_n/n) d\mu \in [-\infty, +\infty)$. Probar que:

$$\lim_{n \rightarrow +\infty} \int \frac{f_n}{n} d\mu = \inf_{n \geq 1} \int \frac{f_n}{n} d\mu \in [-\infty, +\infty).$$

Sugerencia para la última parte: Llame $I_n = \int f_n d\mu$ e $I = \inf_{n \geq 1} (I_n/n)$. Verificar que $I_{n+m} \leq I_n + I_m \quad \forall n, m \geq 1$. Primero asuma que $I = 0$. Fije m tal que $(I_m/m) < \epsilon$. Escriba $n = mq + r$. Pruebe que para todo q suficientemente grande $0 \leq (I_n/n) \leq (I_m/m) + \epsilon < 2\epsilon$. Concluir que $\lim I_n/n = 0$. Después, si $-\infty < I \neq 0$ aplique el

resultado anterior a $I_n - nI$. Finalmente, si $I = -\infty$ pruebe, como en el primer caso, que para todo q suficientemente grande $(I_n/n) \leq -K$, donde $K > 0$ es dado arbitrario.

Teorema 2.2.5. Teorema Ergódico Maximal (Versión General 1)

Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva de funciones reales medibles tales que $f_1 \in L^1(\mu)$. Sea

$$F_n := \max_{1 \leq j \leq n} \{f_j\}$$

Sea

$$E := \{x \in X : \sup_{n \geq 1} F_n(x) > 0\} = \{x \in X : \sup_{n \geq 1} f_n(x) > 0\}.$$

Entonces,

$$\int_E f_1 d\mu \geq 0.$$

Interpretación: Si la sucesión de funciones es subaditiva, entonces es no negativo el valor esperado de la primera de ellas en el conjunto de puntos $x \in X$ donde *alguna* de las funciones f_m es positiva (no importa si tal f_m es una sola en toda la sucesión, ni si m está muy alejado de 1).

Demostración:

Sea $E_n = \{x \in X : F_n(x) > 0\}$. Se cumple $E_n \subset E_{n+1}$, y $E = \bigcup_{n \geq 1} E_n$. Luego, denotando $f_1^+ = \max\{f_1, 0\}$, $f_1^- = -\min\{f_1, 0\}$, tenemos:

$$\begin{aligned} \int_E f_1 d\mu &= \int_E f_1^+ d\mu - \int_E f_1^- d\mu = \\ &= \lim \int_{E_n} f_1^+ d\mu - \lim \int_{E_n} f_1^- d\mu = \lim \int_{E_n} f_1 d\mu \end{aligned}$$

Entonces, para probar que $\int_E f_1 d\mu \geq 0$ basta demostrar que $\int_{E_n} f_1 d\mu \geq 0$ para todo $n \geq 1$.

En lo que sigue denotaremos $f = f_1$. Sea

$$I = \int_{E_n} f d\mu = \int_{\{F_n > 0, F_n \circ T \leq 0\}} f d\mu + \int_{\{F_n > 0, F_n \circ T > 0\}} f d\mu \quad (2.2)$$

Tenemos

$$F_n(x) = \max_{1 \leq j \leq n} \{f_j(x)\}, \quad F_n(T(x)) = \max_{1 \leq j \leq n} \{f_j(T(x))\}.$$

Estudiemos primero el caso cuando $F_n(x) > 0$ y $F_n(T(x)) \leq 0$. Esto implica $f_j(T(x)) \leq 0$ para todo $1 \leq j \leq n$. Siendo $f_{j+1}(x) \leq f(x) + f_j(T(x))$, por la subaditividad de la sucesión $\{f_n\}$, deducimos que $f_{j+1}(x) \leq f(x)$ para todo $1 \leq j \leq n$. Luego, de la construcción de $F_n(x)$ en la igualdad de arriba, deducimos que $F_n(x) = f(x)$. Hemos probado que:

$$F_n(x) = f(x) \text{ si } F_n(T(x)) \leq 0 \text{ y } F_n(x) > 0.$$

Ahora estudiemos el caso $F_n(Tx) > 0$. Tenemos

$$\begin{aligned} f(x) + F_n(T(x)) &= \max_{1 \leq j \leq n} \{f(x) + f_j(T(x))\} \geq \\ &\geq \max_{1 \leq j \leq n} \{f_{j+1}(T(x))\} \geq \max_{2 \leq h \leq n} \{f_h(x)\}. \end{aligned} \quad (2.3)$$

En la penúltima desigualdad usamos la subaditividad de la sucesión $\{f_j\}_{j \geq 1}$. En la última desigualdad se tomó el máximo en un conjunto más chico de valores del subíndice. Como estamos suponiendo que $F_n(T(x)) > 0$, entonces

$$f(x) + F_n(T(x)) > f(x). \quad (2.4)$$

Juntando las dos desigualdades (2.3) y (2.4) obtenidas cuando $F_n(Tx) > 0$, se deduce:

$$f(x) + F_n(T(x)) \geq \max_{1 \leq h \leq n} \{f_h(x)\} = F_n(x).$$

En resumen:

$$\begin{aligned} F_n(x) &= f(x) \quad \text{si } F_n(x) > 0 \text{ y } F_n(Tx) \leq 0, \\ f(x) &\geq F_n(x) - F_n(Tx) \quad \text{si } F_n(Tx) > 0. \end{aligned}$$

Sustituyendo en (2.2) resulta

$$\begin{aligned} I &\geq \int_{\{F_n > 0, F_n \circ T \leq 0\}} F_n d\mu + \int_{\{F_n > 0, F_n \circ T > 0\}} (F_n - F_n \circ T) d\mu \geq \\ &\int_{\{F_n > 0\}} F_n d\mu - \int_{\{F_n > 0, F_n \circ T > 0\}} F_n \circ T d\mu. \end{aligned} \quad (2.5)$$

Por otro lado como μ es T -invariante resulta

$$\int g d\mu = \int g \circ T d\mu \quad \forall g \in L^1(\mu)$$

De donde

$$\begin{aligned} \int_{\{F_n > 0\}} F_n d\mu &= \int \chi_{\{F_n > 0\}} F_n d\mu = \int (\chi_{\{F_n > 0\}} \circ T) \cdot (F_n \circ T) d\mu = \\ &\int (\chi_{\{F_n \circ T > 0\}}) \cdot (F_n \circ T) d\mu = \int_{\{F_n \circ T > 0\}} F_n \circ T d\mu \end{aligned}$$

Sustituyendo en (2.5) resulta:

$$\begin{aligned} I &\geq \int_{\{F_n \circ T > 0\}} F_n \circ T d\mu - \int_{\{F_n > 0, F_n \circ T > 0\}} F_n \circ T d\mu = \\ &\int_{\{F_n \leq 0, F_n \circ T > 0\}} F_n \circ T d\mu \geq 0. \end{aligned}$$

Concluimos que $I = \int_E f_1 d\mu \geq 0$, como queríamos demostrar. \square

**Teorema 2.2.6. Teorema Ergódico Maximal
(Versión General 2)**

Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva de funciones reales medibles tales que $f_1 \in L^1(\mu)$. Sea $\alpha \in \mathbb{R}$ dado. Se define el siguiente conjunto

$$E_\alpha := \{x \in X : \sup_{n \geq 1} \frac{f_n(x)}{n} > \alpha\}.$$

Sea $A \subset E_\alpha$ medible e invariante (i.e. $T^{-1}(A) = A$). Entonces:

$$\int_A f_1 d\mu \geq \alpha \mu(A) \quad \forall n \geq 1.$$

Demostración. Primero probemos el teorema cuando $\alpha = 0$. Denotaremos $E(f) := E_0$ para la sucesión $f := \{f_n\}_{n \geq 1}$. Sea la sucesión $g := \{g_n\}_{n \geq 1}$ construida por:

$$g_n(x) = \chi_A f_n(x) \quad \forall x \in X, \quad \forall n \geq 1.$$

Es inmediato verificar que g es subaditiva.

Por el Teorema 2.2.5, y teniendo en cuenta que $g_1(x) = \chi_A f_1(x)$, deducimos

$$0 \leq \int_{E(g)} g_1 d\mu = \int_{E(g)} \chi_A f_1 d\mu.$$

Afirmamos que $A \subset E(g)$. En efecto, si $x \in A \subset E(f)$, entonces $g_n(x) = f_n(x)$ para todo $n \geq 1$. Además como $x \in E(f)$, tenemos $\sup_{n \geq 1} \{f_n(x)\} > 0$. Luego $\sup_{n \geq 1} \{g_n(x)\} > 0$, es decir $x \in E(g)$.

Concluimos que:

$$0 \leq \int \chi_{E(g)} \chi_A f_1 d\mu = \int \chi_A f_1 d\mu = \int_A f_1 d\mu.$$

Ahora consideremos el caso $\alpha \neq 0$. Denotamos $E_\alpha(f) := E_\alpha$ para la sucesión $f := \{f_n\}_{n \geq 1}$. Sea la sucesión $g := \{g_n\}_{n \geq 1}$ construida por:

$$g_n := f_n - n\alpha \quad \forall n \geq 1.$$

Es inmediato chequear que g_n es subaditiva. Además $E_\alpha(f) = E_0(g)$. Por la parte b) probada en el caso $\alpha = 0$ deducimos que

$$0 \leq \int_{E_0(g)} g_1 d\mu = \int_{E_\alpha(f)} (f_1 - \alpha) d\mu,$$

de donde

$$\int_{E_\alpha(f)} f_1 d\mu \geq \alpha \mu(A),$$

terminando la prueba del Teorema 2.2.6. □

Ahora veamos que el teorema ergódico maximal permite obtener resultados sobre el supremo μ -c.t.p. de la sucesión f_n/n construida a partir de la sucesión subaditiva $\{f_n\}$ dada. En particular, permite obtener resultados sobre los promedios temporales de Birkhoff.

Corolario 2.2.7. del Teorema Ergódico Maximal. *Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva de funciones reales medibles tales que $f_1 \in L^1(\mu)$. Sea α un número real cualquiera. Se define el siguiente conjunto:*

$$\overline{G}_\alpha = \{x \in X : \limsup_{n \rightarrow +\infty} \frac{f_n}{n} > \alpha\}.$$

Sea $A \subset \overline{G}_\alpha$ medible, tal que $T^{-1}(A) = A$. Entonces

$$\int_A f_1 d\mu \geq \alpha \mu(A)$$

Corolario 2.2.8. del Teorema Ergódico Maximal.

Sea $\{f_n\}_{n \geq 1}$, con $f_1 \in L^1(\mu)$, una sucesión aditiva de funciones reales medibles (es decir se cumple la igualdad en la Definición 2.2.2). Sea β un número real cualquiera. Se define el siguiente conjunto:

$$\underline{G}_\beta = \{x \in X : \liminf_{n \rightarrow +\infty} \frac{f_n}{n} < \beta\}.$$

Sea $A \subset \underline{G}_\beta$ medible, tal que $T^{-1}(A) = A$. Entonces

$$\int_A f_1 d\mu \leq \beta \mu(A)$$

Demostración de los Corolarios 2.2.7 y 2.2.8:

Demostración del Corolario 2.2.7: Sea el conjunto

$$E_\alpha := \{x \in X : \sup_{n \geq 1} (f_n(x)/n) > \alpha\}.$$

Es inmediato que $\overline{G}_\alpha \subset E_\alpha$. Aplicando el Teorema Ergódico Maximal 2.2.6, como $T^{-1}(A) = A \subset \overline{G}_\alpha \subset E_\alpha$, tenemos $\int_A f_1 d\mu \geq \alpha \mu(A)$, como queríamos demostrar. \square

Demostración del Corolario 2.2.8: Como $\{f_n\}$ es aditiva, entonces $\{-f_n\}$ también lo es. Observemos que $\underline{G}_\beta(\{f_n\}) = \overline{G}_{-\beta}(\{-f_n\})$. Por el Corolario 2.2.7 se tiene entonces

$$\int_A (-f_1) d\mu \geq (-\beta) \mu(A), \text{ de donde } \int_A f_1 d\mu \leq \beta \mu(A). \quad \square$$

El siguiente resultado, referido a los conjuntos construidos en los Corolarios del Teorema Ergódico Maximal, será útil para probar el Teorema Ergódico de Birkhoff-Khinchin:

Lema 2.2.9. Si la sucesión de funciones es aditiva, entonces los conjuntos $A = \overline{G}_\alpha$ y $B = \underline{G}_\beta$, definidos en los Corolarios 2.2.7 y 2.2.8 respectivamente, son invariantes con T , es decir $T^{-1}(A) = A$ y $T^{-1}(B) = B$.

Demostración: $A = \{x \in X : \limsup_{n \rightarrow +\infty} f_n(x)/n > \alpha\}$, $B = \{x \in X : \liminf_{n \rightarrow +\infty} f_n(x)/n < \beta\}$.

$$T^{-1}(A) = \{x \in X : T(x) \in A\} = \{x \in X : \limsup_{n \rightarrow +\infty} \frac{f_n(T(x))}{n} > \alpha\},$$

$$T^{-1}(B) = \{x \in X : T(x) \in B\} = \{x \in X : \liminf_{n \rightarrow +\infty} \frac{f_n(T(x))}{n} < \beta\},$$

Basta demostrar que para todo $x \in X$ se cumple

$$\limsup_{n \rightarrow +\infty} \frac{f_n(T(x))}{n} = \limsup_{n \rightarrow +\infty} \frac{f_n(x)}{n}, \quad \liminf_{n \rightarrow +\infty} \frac{f_n(T(x))}{n} = \liminf_{n \rightarrow +\infty} \frac{f_n(x)}{n}. \quad (2.6)$$

En efecto

$$\frac{f_{n+1}(x)}{n+1} = \frac{f_1(x)}{n+1} + \left(\frac{n}{n+1}\right) \frac{f_n(T(x))}{n},$$

de donde, tomando límite superior e inferior respectivamente, cuando $n \rightarrow +\infty$ y $x \in X$ está fijo, se deducen las igualdades (2.6), como queríamos demostrar. \square

2.3. Prueba del Teorema Ergódico de Birkhoff-Khinchin

El Teorema 2.1.2 de Birkhoff-Khinchin es una consecuencia del Teorema Ergódico Maximal demostrado en la sección anterior. En efecto, llamemos $f_n = \sum_{j=0}^{n-1} f \circ T^j$. Cuando las funciones son reales, entonces la sucesión $\{f_n\}_{n \geq 1}$ es aditiva. Luego, podemos aplicar todos los resultados de la sección anterior.

Demostración de la Parte a) del Teorema 2.1.2 Para toda $f \in L^1(\mu)$ existe μ -c.t.p.

$$\tilde{f}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) = \lim_{n \rightarrow +\infty} \frac{f_n}{n}.$$

Demostración:

Basta demostrarlo cuando f es real en $L^1(\mu)$.

Sea $G = \{x \in X : \liminf(f_n(x))/n < \limsup(f_n(x))/n\}$. Basta probar que $\mu(G) = 0$. Se cumple

$$G = \bigcup_{\alpha > \beta \text{ racionales}} \overline{G}_\alpha(f) \cap \underline{G}_\beta(f)$$

donde los conjuntos $\overline{G}_\alpha(f)$ y $\underline{G}_\beta(f)$ son los definidos en los Corolarios 2.2.7 y 2.2.8 respectivamente.

Basta demostrar que $A = \overline{G}_\alpha(f) \cap \underline{G}_\beta(f)$ tiene μ medida cero, para todos $\alpha > \beta$ racionales fijos.

El conjunto A es invariante con T , es decir $T^{-1}(A) = A$, ya que $\overline{G}_\alpha(f)$ y $\underline{G}_\beta(f)$ lo son (ver lema 2.2.9).

Aplicando los Corolarios 2.2.7 y 2.2.8, se obtiene:

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

Luego $\alpha\mu(A) \leq \beta\mu(A)$. Siendo $\alpha > \beta$ esto implica $\mu(A) = 0$, como queríamos demostrar. \square

Demostración de la Parte b) del Teorema 2.1.2 $\tilde{f}(x)$ existe si y solo si existe $\tilde{f}(Tx)$ y en ese caso $\tilde{f}(Tx) = \tilde{f}(x)$. Luego $\tilde{f} = \tilde{f} \circ T$ μ -c.t.p.

Demostración: Basta considerar f real. Repitiendo la prueba del Lema 2.2.9 en el caso de la sucesión $\{f_n\}$ aditiva, se tiene, para todo $x \in X$:

$\limsup_{n \rightarrow +\infty} f_n(x)/n = \limsup_{n \rightarrow +\infty} f_n \circ T(x)/n$, y análogamente

$\liminf_{n \rightarrow +\infty} f_n(x)/n = \liminf_{n \rightarrow +\infty} f_n \circ T(x)/n$.

Luego, para todo $x \in X$, existe $\tilde{f}(x)$ si y solo si existe $\tilde{f}(Tx)$ y en ese caso son iguales. Por la parte a) el conjunto donde existe \tilde{f} tiene μ medida igual a uno. Luego, μ -c.t.p. $\tilde{f} = \tilde{f} \circ T$. \square

Demostración de la Parte c) del Teorema 2.1.2 Para toda $f \in L^1(\mu)$ la función \tilde{f} también pertenece a $L^1(\mu)$ y además la convergencia de f_n a \tilde{f} es también en $L^1(\mu)$.

Demostración: Basta demostrar la siguiente afirmación más general:

Afirmación c) Si $f \in L^p(\mu)$ para algún $p \geq 1$ natural entonces la función \tilde{f} también pertenece a $L^p(\mu)$ y además la convergencia de f_n/n a \tilde{f} es también en $L^p(\mu)$.

Lema 2.3.1. Si $f \in L^p(\mu)$ para algún $p \geq 1$ natural entonces las funciones \tilde{f} y $f_n/n = (1/n) \sum_{j=0}^{n-1} f \circ T^j$ también pertenecen a $L^p(\mu)$ y cumplen

$$\|f_n/n\|_p \leq \|f\|_p, \quad \|\tilde{f}\|_p \leq \|f\|_p$$

Demostración: Sea $I_n = \|f_n/n\|_p^p = \int |f_n/n|^p d\mu$. Probemos primero que $I_n \leq \|f\|_p^p$ con lo cual quedará demostrado que $f_n/n \in L^p$ y que $\|f_n/n\|_p \leq \|f\|_p$.

$$I_n = \int \left| \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j \right|^p d\mu \leq \int \left(\frac{1}{n} \sum_{j=0}^{n-1} |f| \circ T^j \right)^p d\mu =$$

$$\left\| \frac{1}{n} \sum_{j=0}^{n-1} |f| \circ T^j \right\|_p^p \leq \left(\frac{1}{n} \sum_{j=0}^{n-1} \|f \circ T^j\|_p \right)^p.$$

En la última desigualdad usamos la propiedad triangular de la norma $\|\cdot\|_p$. Observamos que $\|f \circ T^n\|_p = \|f\|_p$ para todo $n \geq 1$, pues la medida μ es T -invariante y por definición de la norma en $L^p(\mu)$ se tiene $\|f\|_p^p = \int |f|^p d\mu = \int |f \circ T|^p d\mu = \|f \circ T\|_p^p$. Hemos probado entonces que $I_n \leq \|f\|_p^p$ como afirmamos al principio.

Ahora solo resta probar que $I = \int |\tilde{f}|^p d\mu \leq \|f\|_p^p$. Por el lema de Fatou tenemos:

$$I = \int \liminf |f_n/n|^p d\mu = \int \lim |f_n/n|^p d\mu \leq \liminf \int |f_n/n|^p d\mu = \liminf I_n \leq \|f\|_p^p. \quad \square$$

Demostración de la Afirmación c) del Teorema 2.1.2:

Por el lema 2.3.1 se tiene $\tilde{f} \in L^p(\mu)$. Solo resta probar que $f_n \rightarrow \tilde{f}$ en $L^p(\mu)$.

Primer caso: $f \in L^p(\mu)$ es acotada, es decir $|f(x)| \leq K \quad \forall x \in X$. Tenemos entonces que $|f_n(x)/n| \leq K \quad \forall x \in X$ y pasando al límite μ -c.t.p. resulta $|\tilde{f}| \leq K \quad \mu$ -c.t.p., de donde $|f_n/n - \tilde{f}| \leq 2K \quad \mu$ -c.t.p.. Por el teorema de convergencia dominada se tiene:

$$\begin{aligned} \lim_{n \rightarrow +\infty} \|(f_n/n) - \tilde{f}\|_p^p &= \lim_{n \rightarrow +\infty} \int |(f_n/n) - \tilde{f}|^p d\mu = \\ &= \int \lim_{n \rightarrow +\infty} |(f_n/n) - \tilde{f}|^p d\mu = 0 \end{aligned}$$

Segundo caso: $f \in L^p(\mu)$ cualquiera. Las funciones simples son acotadas y densas en $L^p(\mu)$. Luego, dado $\epsilon > 0$, existe g simple tal que

$$\|f - g\|_p < \epsilon$$

Por la propiedad triangular de la norma $\|\cdot\|_p$, se tiene:

$$\|(f_n/n) - \tilde{f}\|_p \leq \|(f_n - g_n)/n\|_p + \|(g_n/n) - \tilde{g}\|_p + \|\tilde{g} - \tilde{f}\|_p \quad (2.7)$$

Según lo probado en el primer caso, como g es acotada se tiene:

$$\|(g_n/n) - \tilde{g}\|_p < \epsilon \quad \forall n \text{ suficientemente grande}$$

Por otra parte, el lema 2.3.1 implica:

$$\begin{aligned} \|(f_n - g_n)/n\|_p &= \|(f - g)_n\|_p \leq \|f - g\|_p < \epsilon, \\ \|\tilde{f} - \tilde{g}\|_p &= \|\lim (f_n - g_n)/n\|_p = \|\lim (f - g)_n/n\|_p = \\ &= \|(\widetilde{f - g})\|_p \leq \|f - g\|_p < \epsilon. \end{aligned}$$

Sustituyendo en (2.7), se deduce que

$$\|(f_n/n) - \tilde{f}\|_p < 3\epsilon \quad \forall n \text{ suficientemente grande,}$$

de donde $\|(f_n/n) - \tilde{f}\|_p \rightarrow 0$ como queríamos probar. \square

Demostración de la Parte d) del Teorema 2.1.2. Para toda $f \in L^1(\mu)$ se cumple

$$\int f d\mu = \int \tilde{f} d\mu$$

Demostración: Por la parte c) se tiene que $\|(f_n/n) - \tilde{f}\|_1 \rightarrow 0$. Luego:

$$\left| \int (f_n/n) d\mu - \int \tilde{f} d\mu \right| \leq \int |(f_n/n) - \tilde{f}| d\mu = \|(f_n/n) - \tilde{f}\|_1 \rightarrow 0$$

Concluimos por un lado que

$$\int (f_n/n) d\mu \rightarrow \int \tilde{f} d\mu$$

Por otro lado:

$$\int (f_n/n) d\mu = \frac{1}{n} \sum_{j=0}^{n-1} \int f \circ T^j d\mu = \frac{1}{n} \sum_{j=0}^{n-1} \int f d\mu = \int f d\mu \quad \forall n \geq 1$$

Reuniendo ambos resultados se obtiene $\int f d\mu = \int \tilde{f} d\mu$. \square

2.4. Otras caracterizaciones de la ergodicidad

En esta sección, salvo indicación en contrario, (X, \mathcal{A}) denota un espacio medible y $T : X \mapsto X$ una transformación medible que preserva alguna medida de probabilidad μ .

Se recuerda las Definiciones 1.6.1 y 1.6.2 de ergodicidad, y los Teoremas 1.6.3, 1.6.9 y 1.7.5, en los que dimos diferentes caracterizaciones de ergodicidad. Agregamos ahora las siguientes:

Teorema 2.4.1. Ergodicidad IV

Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Las siguientes propiedades son equivalentes:

- a) T es ergódica respecto de μ .
- b) Para toda $f \in L^1(\mu)$ se cumple

$$\tilde{f}(x) = \int f d\mu \quad \mu - c.t.p.$$

- c) Para toda pareja de conjuntos medibles A y B se cumple:

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

El enunciado y la prueba del Teorema 2.4.1 fueron extraídos, con leves modificaciones, de [Mañ 1983a, págs. 130-131] (ver también [Mañ 1987]).

Demostración de que a) \Rightarrow b) en el Teorema 2.4.1: Por el Teorema de Birkhoff-Khinchin $\int \tilde{f} d\mu = \int f d\mu$. Entonces basta demostrar que \tilde{f} es constante μ -c.t.p.

Por el teorema de Birkhoff-Khinchin $\tilde{f}(x)$ existe μ -c.t.p., $\tilde{f}(x)$ existe si y solo si existe $\tilde{f}(Tx)$ y en ese caso $\tilde{f}(Tx) = \tilde{f}(x)$. Dicho de otra forma, el conjunto $A = \{x \in X : \tilde{f}(x) \text{ existe}\}$ cumple $\mu(A) = 1$, $T^{-1}(A) = A$ y para todo $x \in A$: $\tilde{f}(x) = \tilde{f}(Tx)$.

Sea $g : X \mapsto \mathbb{C}$ definida como $g(x) = \tilde{f}(x)$ si $x \in A$, $g(x) = 0$ si $x \notin A$. Entonces para todo $x \in X$ se cumple $g \circ T(x) = g(x)$. Por la afirmación demostrada antes $g = cte$ μ -c.t.p. Pero por construcción $g = \tilde{f}$ μ -c.t.p., de donde se deduce que $\tilde{f} = cte$ μ -c.t.p. como queríamos. \square

Demostración de que b) \Rightarrow c) en el Teorema 2.4.1:

Obsérvese que $\mu(T^{-j}A \cap B) = \int \chi_{T^{-j}(A)} \chi_B d\mu = \int (\chi_A \circ T^j) \chi_B d\mu$, de donde:

$$I_n = \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}A \cap B) = \int \left(\frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j \right) \chi_B d\mu$$

Por convergencia dominada (el integrando está acotado por la función constante $1 \in L^1(\mu)$), resulta:

$$\lim I_n = \int \lim_{n \rightarrow +\infty} \left(\frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j \right) \chi_B d\mu = \int (\tilde{\chi}_A) \chi_B d\mu \quad (2.8)$$

Por el teorema de Birkhoff-Khinchin, observando que $\chi_A \in L^p(\mu)$ y $\tilde{\chi}_A$ es invariante con T , luego constante μ -c.t.p., se tiene que

$$\tilde{\chi}_A(x) = \int \tilde{\chi}_A d\mu = \int \chi_A d\mu = \mu(A) \quad \mu\text{-c.t.p. } x \in X$$

Sustituyendo en (2.8) resulta $\lim_n I_n = \int (\mu(A)) \chi_B d\mu = \mu(A)\mu(B)$. \square

Demostración de que c) \Rightarrow a) en el Teorema 2.4.1: Basta demostrar que si $A, B \subset X$ son medibles con μ medida positiva, entonces existe $j \geq 1$ tal que $\mu(T^{-j}(A) \cap B) > 0$ (cf. Definición 1.6.1). Por hipótesis:

$$\lim \frac{1}{n} \sum_{j=1}^{n-1} \mu(T^{-j}(A) \cap B) = \lim \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B) = \mu(A)\mu(B) > 0$$

Entonces $\sum_{j=1}^{n-1} \mu(T^{-j}(A) \cap B) > 0 \quad \forall n$ suficientemente grande, de donde $\mu(T^{-j}(A) \cap B) > 0$ para algún $j \geq 1$ como queríamos probar. \square

Ejercicio 2.4.2. Ergodicidad V.

Sea $T : X \mapsto X$ que preserva una probabilidad μ . Probar que:

(a) T ergódica respecto de μ si y solo si para todo conjunto A medible tal que $T^{-1}(A) \subset A$ ó $T^{-1}(A) \supset A$, se cumple $\mu(A)$ es o bien cero o bien uno.

(b) T es ergódica respecto de μ si y solo si para toda $f \in L^1(\mu)$ tal que $f \circ T \leq f$ μ -c.t.p., se cumple $f = cte$ μ -c.t.p.

(c) T es ergódica respecto de μ si y solo si para todos los conjuntos A medibles que cumplen $T^{-1}(A) \subset A$ ó $A \subset T^{-1}(A)$ se verifica $\mu(A) = 0$ ó $\mu(A) = 1$.

(d) T es ergódica respecto de μ si y solo si para todas las funciones $f \in L^1(\mu)$ tales que $f \circ T \leq f$ μ -c.t.p., ó $f \circ T \geq f$ μ -c.t.p. se cumple $f = cte$ μ -c.t.p.

Sugerencia para (a): Basta probarlo cuando $T^{-1}(A) \supset A$, pues en caso contrario, sustituimos A por su complemento. Denote χ_A a la función característica de A . Como $T^{-1}(A) \supset A$, pruebe que $\chi_{T^{-n}(A)}(x) = 1$ para todo $x \in A$. Luego $\tilde{\chi}_A(x) = 1$ para todo $x \in A$, y usando que χ_A es constante μ -c.t.p. si μ es ergódica, se deduce $\mu(A) = 1$ ó $\mu(A) = 0$.

Ejercicio 2.4.3. Sea X un espacio métrico compacto y sea $T : X \mapsto X$ Borel-medible tal que preserva una probabilidad μ . Sea $\{g_i : i \geq 1\}$ un conjunto numerable denso en $C^0(X, [0, 1])$. Probar que son equivalentes las afirmaciones siguientes:

i) μ es ergódica.

ii) $\tilde{f}(y) = \int f d\mu$ μ -c.t.p. y ; $\forall f \in C_0(X, \mathbb{R})$

iii) $\tilde{g}_i(y) = \int g_i d\mu$ μ -c.t.p. y ; $\forall i \geq 1$

Sugerencia: Recordar que μ es ergódica si y solo si para toda $h \in L^1(\mu) : \tilde{h}(y) = \int h d\mu$ μ -c.t.p. y ; y que las funciones continuas son densas en $L^1(\mu)$.

Ejercicio 2.4.4. Sea X un espacio métrico compacto y sea $T : X \mapsto X$ Borel-medible tal que preserva una probabilidad μ . Probar que T es ergódica respecto de μ si y solo si para toda función compleja $f : X \mapsto \mathbb{C}$ continua, el límite \tilde{f} de los promedios de Birkhoff de f es constante μ -c.t.p. Sugerencia: Usar lo probado en el ejercicio 2.4.3 y chequear, usando el Teorema de Birkhoff-Khinchin, que si \tilde{f} es una constante μ -c.t.p., entonces esta constante es $\int f d\mu$.

Volvamos al caso general de un espacio medible (X, \mathcal{A}) con una transformación medible $T : X \mapsto X$ que preserva una medida de probabilidad μ . Por el Lema de recurrencia de Poincaré (Teorema 1.5.2) si un conjunto medible A cumple $\mu(A) > 0$, entonces la órbita futura de μ -c.t.p. $x \in A$ vuelve infinitas veces a A . Sin embargo, ese Lema no dice nada sobre la frecuencia de visita y la duración de las estadías de la órbita de x en el conjunto A . Las medidas ergódicas dan exactamente el valor de la frecuencia asintótica en que la órbita de x pasa dentro de A .

Definición 2.4.5. Tiempo medio de estadía.

Llamamos *tiempo medio de estadía* $\tau_A(x)$ de la órbita por $x \in X$ en un conjunto medible A a:

$$\tau_A(x) = \lim_{n \rightarrow +\infty} \frac{\#\{j \in \mathbb{N} : 0 \leq j \leq n-1, T^j(x) \in A\}}{n} =$$

$$= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j(x) = \tilde{\chi}_A(x)$$

Teorema 2.4.6. Ergodicidad VI.

Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Entonces μ es ergódica para T si y solo si para todo conjunto medible A el tiempo medio de estadía $\tau_A(x)$ es constante μ -c.t.p. Además, en ese caso $\tau_A(x) = \mu(A)$ μ -c.t.p.

El enunciado y la prueba de este teorema, con leves modificaciones, fue extraído de [Mañ 1983a, pág. 133] (ver también [Mañ 1987]).

Demostración: Por el teorema de Birkhoff-Khinchin $\tilde{\chi}_A \in L^1(\mu)$ y cumple $\tilde{\chi}_A = \tilde{\chi}_A \circ T$ μ -c.t.p. Si $\tilde{\chi}_A = cte$ μ -c.t.p. entonces, aplicando nuevamente el teorema de Birkhoff-Khinchin:

$$\tilde{\chi}_A(x) = \int \tilde{\chi}_A d\mu = \int \chi_A d\mu = \mu(A) \quad \mu\text{-c.t.p.} \quad (2.9)$$

Si $\tau_A = \tilde{\chi}_A = cte$ μ -c.t.p. para todo A medible, tomemos en particular A tal que $T^{-1}(A) = A$. Para demostrar la ergodicidad de μ hay que probar que $\mu(A)$ es cero o uno.

$$\tilde{\chi}_A = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j$$

Pero $\chi_A \circ T^j = \chi_{T^{-j}(A)} = \chi_A$. Luego resulta:

$$\tilde{\chi}_A(x) = \chi_A(x) \in \{0, 1\} \quad \mu\text{-c.t.p. } x \in X \quad (2.10)$$

Por (2.9) y (2.10) se tiene $\mu(A) \in \{0, 1\}$ y μ es ergódica como queríamos probar.

Recíprocamente, si μ es ergódica, entonces por la parte b) del teorema 2.4.1. $\tau_A = \tilde{\chi}_A = cte$ μ -c.t.p. \square

Ejercicio 2.4.7. Sea $T : X \mapsto X$ es Borel medible en un espacio topológico X conexo, que preserva una medida de probabilidad μ positiva sobre abiertos. Probar que:

(a) Una función $g \in L^1(\mu)$ invariante con T (i.e. $g \circ T = g$ μ -c.t.p.) es constante μ -c.t.p. si y solo si es localmente constante c.t.p. (es decir: existe un cubrimiento de X por abiertos, tales que en cada abierto V del cubrimiento se cumple $g|_V = K_V$ constante μ -c.t.p. de V .)

(b) T es ergódica si y solo si toda función $g \in L^1(\mu)$ que sea invariante con T es localmente constante.

2.5. Ergodicidad Única

Definición 2.5.1. La transformación T es *únicamente ergódica* si existe una única medida de probabilidad que es T invariante.

Por lo visto en la sección 1.7 esta única medida es extremal en el conjunto de las probabilidades invariantes; luego es ergódica.

Teorema 2.5.2. *Sea T continua en un espacio métrico compacto. Las siguientes afirmaciones son equivalentes:*

- i) T es únicamente ergódica
- ii) Para toda $f \in C^0(X, \mathbb{R})$, para todo $x \in X$ existe el siguiente límite y es un número independiente de x :

$$\tilde{f} := \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x).$$

- iii) Para toda $f \in C^0(X, \mathbb{R})$, la sucesión de funciones continuas $\{f_n\}$ definidas por

$$f_n = \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j$$

converge uniformemente a una constante cuando $n \rightarrow \infty$.

Demostración: i) implica iii):

Sea μ la única medida de $\mathcal{M}_T(X)$. Probaremos f_n converge uniformemente a $\int f d\mu$ en X . Por absurdo, supongamos que existe $\epsilon > 0$, y una sucesión $n_j \rightarrow \infty$ tal que $\sup_{x \in X} |f_{n_j}(x) - \int f d\mu| \geq \epsilon$ para todo $j \geq 1$. Como el supremo es un máximo, existe $x_j \in X$, donde se alcanza.

Por el teorema de Riesz, existe una medida μ_j (no necesariamente T -invariante) tal que:

$$\forall g \in C^0(X, \mathbb{R}) : \int g d\mu_j = \frac{1}{n_j} \sum_{i=0}^{n_j-1} g \circ T^i(x_j)$$

Por la compacidad del espacio $\mathcal{M}(X)$, existe una subsucesión convergente de esta medidas μ_j . Por simplicidad seguiremos usando la misma notación para la subsucesión que para la sucesión original.

$$\mu_j \rightarrow \nu$$

Afirmamos que ν es T -invariante. En efecto:

$$\begin{aligned} \int g dT^* \nu &= \int g \circ T d\nu = \lim \int g \circ T d\mu_j = \lim \frac{1}{n_j} \sum_{i=0}^{n_j-1} g \circ T^{i+1}(x_j) \\ &= \lim \frac{1}{n_j} \sum_{i=0}^{n_j-1} g \circ T^i(x_j) = \lim \int g d\mu_{n_j} = \int g d\nu \end{aligned}$$

Como por hipótesis T es únicamente ergódica, se tiene $\mu = \nu$. Entonces

$$\int f d\mu = \int f d\nu = \lim f_{n_j}(x_j)$$

Pero por construcción

$$|f_{n_j}(x_j) - \int f d\mu| \geq \epsilon \quad \text{para todo } j \geq 1$$

contradiciendo la igualdad anterior.

iii) implica ii) porque la convergencia uniforme de funciones continuas en X implica la convergencia en todo punto de X .

ii) implica i): Sea Λ el funcional lineal positivo definido por

$$\Lambda(f) = \tilde{f}$$

para toda $f \in C^0(X, \mathbb{R})$.

Para toda medida μ que sea T invariante, por el teorema de Birkhoff

$$\int f d\mu = \int \tilde{f} d\mu = \tilde{f} = \Lambda(f)$$

Por el teorema de Riesz, existe una única μ que cumple

$$\int f d\mu = \Lambda(f)$$

Luego existe una única μ que es T -invariante. □

Como ejemplo, veremos en la próxima sección que la rotación irracional en el círculo es únicamente ergódica.

Definición 2.5.3. Conjuntos minimales

Sea T una transformación continua en un espacio métrico compacto X .

Un subconjunto $\Lambda \subset X$ es *minimal* (desde el punto de vista topológico) si es compacto, no vacío, invariante hacia el futuro (es decir: $\Lambda \subset T^{-1}(\Lambda)$), y no existe ningún subconjunto propio de Λ que sea compacto, no vacío e invariante hacia el futuro.

Tenemos la siguiente caracterización (ver parte (a) del Ejercicio 2.5.4):

Λ es *minimal* si y solo si Λ es compacto, no vacío y T -invariante (es decir: $T^{-1}(\Lambda) = \Lambda$) y no contiene subconjuntos propios que sean compactos, no vacíos e invariantes por T hacia el futuro.

Si además el mapa continuo T es invertible con inversa continua, entonces Λ es *minimal* si y solo si Λ es compacto, no vacío y T -invariante, y no contiene subconjuntos propios que sean también compactos, no vacíos y T -invariantes (ver parte (b) del Ejercicio 2.5.4).

Finalmente, es fácil ver que cualquiera sea T continua (no necesariamente invertible), un conjunto Λ compacto, no vacío e invariante, es *minimal* si y solo si todas sus órbitas hacia el futuro son densas en Λ . Esto es porque la clausura de cada una de ellas es un compacto no vacío, invariante hacia adelante, y contenido en Λ (ver también parte (a) del Ejercicio 2.5.4).

Ejercicio 2.5.4. (a) Probar que las siguientes afirmaciones son equivalentes (y por lo tanto cualquiera de ellas puede utilizarse como definición de Λ minimal):

(i) $\Lambda \subset X$ es compacto no vacío e invariante por T hacia el futuro, y no contiene subconjuntos propios compactos no vacíos e invariantes hacia el futuro.

(ii) $\Lambda \subset X$ es compacto no vacío y T -invariante, y no contiene subconjuntos propios compactos no vacíos e invariantes por T hacia el futuro. (Sugerencia para demostrar

(i) \Rightarrow (ii): probar que si Λ cumple (i), entonces $T^{-1}(\Lambda) \setminus \Lambda$ también.)

(iii) $\Lambda \subset X$ es compacto no vacío, y para todo $x \in \Lambda$ la clausura de $\{f^j(x)\}_{j \geq 0}$ es igual a Λ .

(b) Asumir ahora que T es un homeomorfismo (es decir, T es continua, invertible y su T^{-1} es continua). Probar que las siguientes afirmaciones son equivalentes (y por lo tanto cualquiera de ellas puede utilizarse como definición de Λ minimal):

(i) $\Lambda \subset X$ es compacto no vacío e invariante por T hacia el futuro, y no contiene subconjuntos propios compactos no vacíos e invariantes hacia el futuro.

(ii) $\Lambda \subset X$ es compacto no vacío y T -invariante, y no contiene subconjuntos propios compactos no vacíos y T -invariantes.

(iii) $\Lambda \subset X$ es compacto no vacío, y para todo $x \in \Lambda$ la clausura de $\{f^j(x)\}_{j \geq 0}$ es igual a Λ .

(iv) $\Lambda \subset X$ es compacto no vacío e invariante por T hacia el pasado (es decir $T^{-1}(\Lambda) \subset \Lambda$), y no contiene subconjuntos propios compactos no vacíos e invariantes hacia el pasado.

(v) $\Lambda \subset X$ es compacto no vacío, y para todo $x \in \Lambda$ la clausura de $\{f^{-j}(x)\}_{j \geq 0}$ es igual a Λ .

Veamos como se vincula la ergodicidad única con los conjuntos minimales:

Teorema 2.5.5. *Si T es continua en un espacio métrico compacto X , y únicamente ergódica, entonces existe un único minimal Λ , y además Λ es el soporte de la medida invariante por T .*

Ejemplo de Furstenberg: El recíproco del Teorema 2.5.5 es falso: Furstenberg en [Fu 1961] (ver también [Mañ 1983a, Capítulo 2, §7, pág. 172] o [Mañ 1987]), dio un ejemplo de transformación continua en el toro que preserva la medida de Lebesgue, para la cual todo el toro es el único minimal pero la medida de Lebesgue no es ergódica. Luego, en el Ejemplo de Furstenberg, la transformación no es únicamente ergódica, pero existe un único minimal (y además este minimal es el soporte de una medida T -invariante).

Otros ejemplos en los que se prueba la existencia de más de una medida ergódica para mapas con un único conjunto minimal, se encuentran en [Bac 1999].

Demostración del Teorema 2.5.5: Sea μ la única medida invariante de T . Sea Λ el soporte compacto de μ , definido como

$$\Lambda := \{x \in X : \forall V \text{ entorno de } x \mu(V) > 0\}.$$

Λ es cerrado en X compacto, luego es compacto.

Veamos que $\Lambda \subset T^{-1}(\Lambda)$. Sea $x \in \Lambda$ y sea $y = T(x)$. Hay que probar que $y \in \Lambda$. Para todo entorno U de y , $T^{-1}(U)$ es entorno de x , luego $0 < \mu(T^{-1}(U)) = \mu(U)$. Esto prueba que $y \in \Lambda$; luego Λ es invariante hacia adelante.

Sea Λ_0 compacto, no vacío, invariante hacia adelante. Sea $\widehat{T} = T|_{\Lambda_0} : \Lambda_0 \mapsto \Lambda_0$. Por el teorema de existencia de medidas invariantes, existe $\widehat{\nu}$ probabilidad que es \widehat{T} invariante.

Sea ν probabilidad en X , definida así:

$$\nu(A) = \widehat{\nu}(A \cap \Lambda_0)$$

El soporte de ν está contenido en Λ_0 . En efecto, si $x \notin \Lambda_0$, entonces existe V , entorno de x disjunto con Λ_0 . Luego $\nu(V) = 0$ y x no pertenece al soporte de ν .

Probemos que ν es T -invariante:

$$\nu(T^{-1}(A)) = \widehat{\nu}(T^{-1}(A) \cap \Lambda_0) = \widehat{\nu}\{x \in \Lambda_0 : T(x) \in A\} =$$

$$\widehat{\nu}(\widehat{T}^{-1}(A \cap \Lambda_0)) = \widehat{\nu}(A \cap \Lambda_0) = \nu(A).$$

Como T es únicamente ergódica, $\nu = \mu$. Luego $\Lambda = \text{sop } \nu \subset \Lambda_0$. Hemos probado que todo Λ_0 compacto, no vacío e invariante hacia adelante contiene a Λ . De ello se deducen dos resultados: Primero, si Λ_0 además está contenido en Λ , entonces coincide con Λ . Luego Λ es minimal. Segundo: si Λ_0 es minimal, entonces coincide con Λ . Luego Λ es el único minimal. \square

Ejercicio 2.5.6. Sea X un espacio métrico compacto y sea $T : X \mapsto X$ medible, tal que existe alguna medida invariante μ .

$$P = \left\{ x \in X : \text{existe } \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} h \circ T^j(x) \forall h : X \mapsto \mathbb{R} \text{ medible acotada} \right\}$$

a) Probar que P es el conjunto de los puntos periódicos de T . Sugerencia: 1) Inventar una sucesión $\{a_i\}_{i \geq 0}$ de ceros y unos tal que no exista el límite cuando $n \rightarrow \infty$ de $1/n \cdot \sum_{j=0}^{n-1} a_j$. (Por ejemplo: 1 uno, 1 cero, 10 unos, 10 ceros, 100 unos, 100 ceros, 1000 unos, 1000 ceros, etc). 2) Si x no es periódico mostrar que $\widetilde{\chi}_A(x)$ no existe para $A = \{T^i(x) : a_i = 1\}$.

b) Sea $A \subset X$ cualquier boreliano dado. Probar que tiene probabilidad total el conjunto $\{x \in X : \text{existe } \lim_{n \rightarrow \infty} (1/n) \sum_{j=0}^{n-1} \chi_A(T^j(x))\}$.

c) Probar que si T es continua, únicamente ergódica y existe un conjunto minimal con infinitos puntos, entonces $P = \emptyset$. (Un ejemplo de tal T es la rotación irracional del círculo, como veremos a continuación).

2.6. Ergodicidad de la rotación irracional

Ahora daremos una prueba de la ergodicidad de la rotación irracional en el círculo. Esta es una prueba “pedestre”, que usa argumentos topológicos, y se basa en la densidad de las órbitas y en la caracterización de la ergodicidad única. En la sección 8.4

daremos otra demostración, más breve y optimizada, usando nociones básicas de la Teoría Espectral que también caracterizan a la ergodicidad única de una transformación. La prueba breve basada en la Teoría Espectral puede encontrarse también, por ejemplo en [Wa 2000, Theorem 1.8].

Teorema 2.6.1. *La rotación irracional en el círculo es únicamente ergódica. Luego, es ergódica respecto a la medida de Lebesgue.*

Demostración. Sea $T(x) = x + x_0 \pmod{1}$, para todo $x \in S^1 = \mathbb{R}/\sim \pmod{1}$, donde $x_0 \in (0, 1)$ es irracional dado. Aplicando el Teorema 2.5.2 parte ii), para probar que T es únicamente ergódica, probaremos que para cada $f \in C^0(S^1, \mathbb{R})$ la sucesión

$$f_n = \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j$$

converge en todo punto a una constante cuando $n \rightarrow +\infty$.

Por el Teorema Ergódico de Birkhoff-Khinchin para Lebesgue-casi todo punto $x_1 \in S^1$ existe $\lim_{n \rightarrow +\infty} f_n(x_1) = \tilde{f}(x_1)$. Dado $\epsilon > 0$, como f es uniformemente continua en S^1 (porque es continua en el compacto S^1), existe $\delta > 0$ tal que

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Por la definición de la rotación T tenemos la siguiente igualdad mód. 1, para todo punto $y \in S^1$ y para todo $j \geq 0$:

$$T^j(y) = y + jx_0 = (y - x_1) + x_1 + jx_0 = (y - x_1) + T^j(x_1).$$

Luego, $|y - x_1| < \delta \Rightarrow |f(T^j(y)) - f(T^j(x_1))| < \epsilon \quad \forall j \in \mathbb{N} \Rightarrow$

$$|f_n(y) - f_n(x_1)| < \epsilon \quad \forall n \in \mathbb{N}.$$

Sea $N \geq 1$ tal que $|f_n(x_1) - \tilde{f}(x_1)| < \epsilon$ para todo $n \geq N$. Obtenemos:

$$|f_n(y) - \tilde{f}(x_1)| < 2\epsilon \quad \forall n \in \mathbb{N} \text{ si } |y - x_1| < \delta \quad (2.11)$$

Hemos probado que para todo $\epsilon > 0$ existe $\delta > 0$ (módulo de continuidad uniforme de f , y por lo tanto independiente de x_1) tal que se cumple la desigualdad (2.11) para todo punto y a distancia menor que δ de x_1 . Luego $\limsup f_n(y) - \liminf f_n(y) < 2\epsilon$ si $|y - x_1| < \delta$. Como esta desigualdad vale para Lebesgue-c.t.p. $x_1 \in S^1$, dado $\epsilon > 0$ y dado $y \in S^1$, podemos siempre elegir algún $x_1 \in S^1$ tal que $|y - x_1| < \delta$ y tal que existe $\tilde{f}(x_1)$. Entonces podemos aplicar la desigualdad (2.11) para todo $y \in S^1$ eligiendo, para cada y , algún x_1 adecuado. Deducimos que para todo $y \in S^1$ se cumple $\limsup f_n(y) - \liminf f_n(y) < 2\epsilon$. Como $\epsilon > 0$ es arbitrario, existe

$$\tilde{f}(y) = \lim f_n(y) \quad \forall y \in S^1.$$

Aplicando nuevamente la desigualdad (2.11), deducimos que la función \tilde{f} es continua. Lo anterior vale para cualquier rotación del círculo. Todavía no usamos que

la rotación es irracional, que aplicaremos ahora para probar que \tilde{f} es constante. La función \tilde{f} es invariante por T , es decir $\tilde{f} = \tilde{f} \circ T$, pues el límite de los promedios temporales de Birkhoff. Luego, \tilde{f} toma un valor constante para cada órbita. Entonces, para probar que \tilde{f} es constante (sabiendo ya que es continua) alcanza con probar que existe una órbita densa. Para probar que alguna órbita es densa alcanza con probar que existe $x_1 \in S^1$ tal que para todo $\delta > 0$ la órbita $o^+(x_1) = \{T^n(x_1) : n \geq 0\}$ es δ -densa (i.e. todo intervalo de longitud δ contiene algún punto de $o^+(x_1)$). La transformación T preserva la medida de Lebesgue. Aplicando el Lema de Recurrencia de Poincaré, Lebesgue-c.t.p. es recurrente. Elijamos un punto recurrente x_1 . Entonces existe $n_j \rightarrow +\infty$ tal que

$$\lim_j |T^{n_j}(x_1) - x_1| = 0.$$

Fijemos $m_1 \geq 1$ tal que $|T^{m_1}(x_1) - x_1| < \delta$. Observemos que $T^{m_1}(x_1) - x_1 = m_1 x_0$ (todas las igualdades son módulo 1). Como $x_0 \in (0, 1) \setminus \mathbb{Q}$, tenemos que $m_1 x_0 \neq 0$ (mód. 1). Además $T^{2m_1}(x_1) = T^{m_1}(x_1) + m_1 x_0$, de donde

$$|T^{2m_1}(x_1) - T^{m_1}(x_1)| = |T^{m_1}(x_1) - x_1| = m_1 x_0 = a \in (0, \delta) \neq 0$$

Por inducción en $k \geq 0$ obtenemos

$$|T^{(k+1)m_1}(x_1) - T^{km_1}(x_1)| = a \neq 0 \quad \forall k \in \mathbb{N}. \quad (2.12)$$

Afirmamos que el conjunto $A := \{T^{km_1}(x_1) : k \in \mathbb{N}\} \subset o^+(x_1)$ es δ -denso en S^1 . Por un lado, para valores diferentes de $k \in \mathbb{N}$, los puntos respectivos $T^{km_1}(x_1) \in A$ son diferentes. En efecto, si $T^{km_1}(x_1) = T^{hm_1}(x_1)$, entonces $km_1 x_0 = hm_1 x_0$, de donde $|k - h|m_1 x_0 = 0$ (mód. 1), donde $k, h, m_1 \in \mathbb{N}$. Como $x_0 \notin \mathbb{Q}$ deducimos que $h = k$. Por otro lado, dos puntos consecutivos de A cumplen la igualdad (2.12). Entonces $\bigcup_{k \in \mathbb{N}} (T^{km_1}(x_1) - a, T^{km_1}(x_1)) = S^1$, lo que demuestra que A es δ -denso en S^1 (pues $\delta > a$). Siendo $A \subset o^+(x_1)$, la órbita por x_1 es δ -densa, para todo $\delta > 0$, luego es densa, terminando la demostración del Teorema 2.6.1. \square

Observación 2.6.2. En la demostración del Teorema 2.6.1 probamos que existe una órbita $\{f^n(x_0)\}_{n \geq 0}$ densa en la rotación irracional del círculo $S^1 = [0, 1] \pmod{1}$ dada por $f(x) = x + a$ (mód. 1) donde a es irracional. Esto implica que

Toda órbita por la rotación irracional del círculo es densa.

Demostración. Tenemos: $f^n(y_0) = y_0 + na = x_0 + na + (y_0 - x_0) = f^n(x_0) + (y_0 - x_0)$ (mód. 1). Entonces, la órbita $\{f^n(y_0)\}_{n \geq 0}$ se obtiene de la órbita $\{f^n(x_0)\}_{n \geq 0}$ rotándola $y_0 - x_0$ (mód. 1). Como cualquier rotación en el círculo es un homeomorfismo, lleva un conjunto denso a un conjunto también denso. Por lo tanto, existe una órbita densa si y solo si todas las órbitas son densas. Ya probamos (al final de la demostración del Teorema 2.6.1), que existe una órbita densa cuando a es irracional. Concluimos que todas las órbitas son densas. \square

De la prueba del Teorema 2.6.1 deducimos que las rotaciones en el círculo son ergódicas (respecto de la medida de Lebesgue) si y solo si son únicamente ergódicas, y esto ocurre

si y solo si la rotación es topológicamente transitiva. Más en general, la transitividad topológica es equivalente a la ergodicidad única para la rotación rígida en cualquier grupo topológico compacto abeliano (ver por ejemplo [Ni-Pe 2011, page 266]).

Ejercicio 2.6.3. Considérese el toro k -dimensional

$$\mathbb{T}^k = (S^1)^k = [0, 1]^k / (\text{mód}\{1\}^k),$$

y la operación de grupo $+$ ($\text{mód}\{1\}^k$). Sea $x_0 \in \mathbb{T}^k$. Sea T la traslación $T(x) = x + x_0 \pmod{\{1\}^k}$ para todo $x \in \mathbb{T}^k$. Si $\tilde{x}_0 \in \mathbb{R}^n$ es un representante de $x_0 \in \mathbb{T}^k$ y si \langle, \rangle denota el producto interno usual en \mathbb{R}^n , asuma que $\langle x_0, m \rangle \notin \mathbb{Z} \quad \forall m \in \mathbb{Z}^k - \{0\}$.

a) Demostrar que T es únicamente ergódica.

b) Demostrar que $\forall x \in X$, la medida de Lebesgue es el límite cuando $n \rightarrow \infty$ de las medidas $1/n \cdot \sum_{j=0}^{n-1} \delta_{T^j(x)}$.

c) Probar que la medida de Lebesgue es la única medida de probabilidad en el toro invariante por todas las traslaciones, pero que no toda traslación es únicamente ergódica. Sugerencia: las traslaciones que tienen puntos periódicos no son únicamente ergódicas.

d) Deducir que para toda traslación ergódica del toro, éste es minimal y todas las órbitas son densas.

2.7. Transformaciones Mixing.

Definición 2.7.1. Sea $T : X \mapsto X$ medible que preserve una medida de probabilidad μ . Se dice que T es *mixing* respecto de μ , o que μ es *mixing* respecto de T , si para toda pareja A, B de conjuntos medibles se cumple:

$$\lim_{n \rightarrow +\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B). \quad (2.13)$$

Teorema 2.7.2. *Toda transformación mixing es ergódica.*

(El recíproco es falso como veremos en el Ejemplo 2.7.5.)

Demostración: En el teorema 2.4.1 se prueba que T es ergódica si y solo si

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}A \cap B) = \mu(A)\mu(B) \quad (2.14)$$

Si se cumple la igualdad (2.13) de la definición de transformación mixing, entonces se cumple la igualdad (2.14) de ergodicidad. En efecto si una sucesión $\{a_n\}_{n \in \mathbb{N}}$ de números reales es convergente a $a \in \mathbb{R}$, entonces (aplicando la definición de límite se puede chequear fácilmente): $\lim_{n \rightarrow +\infty} (1/n) \sum_{j=0}^{n-1} a_j = a$. \square

Para probar que el recíproco del Teorema 2.7.2 es falso, basta ver que la medida de Lebesgue para la rotación irracional en el círculo (que ya probamos que es ergódica en el Teorema 2.6.1) no es mixing. Probaremos que no es mixing en el Ejemplo 2.7.5 de esta sección.

Definición 2.7.3. Sea $T : X \mapsto X$ Borel-medible en un espacio topológico X . Se dice que T es topológicamente *mixing* si y solo si para toda pareja de abiertos U y V no vacíos existe $n_0 \in \mathbb{N}$ tal que

$$T^n(U) \cap V \neq \emptyset \quad \forall n \geq n_0$$

Ejercicio 2.7.4. Probar que si T es Borel medible en un espacio topológico X y es *mixing* respecto a una medida de probabilidad μ positiva sobre abiertos, entonces es topológicamente *mixing*.

No toda medida ergódica es mixing. En efecto:

Ejemplo 2.7.5. La medida de Lebesgue en el círculo no es *mixing* para la rotación irracional.

Demostración: Sea la rotación irracional $T : S^1 \mapsto S^1$ en el círculo S^1 , definida por $T(z) = z + \alpha$ (mód.1), donde α es un número irracional (que puede tomarse en $(0, 1)$). Para demostrar que su única medida de probabilidad invariante (la medida m de Lebesgue) no es *mixing*, basta demostrar que T no es topológicamente *mixing*. Sea $U \subset S^1$ un intervalo abierto de longitud $\epsilon : 0 < \epsilon < (1/4) \min\{\alpha, 1 - \alpha\}$. Probaremos que, si para cierto $n_0 \in \mathbb{N}$ se cumple $T^{n_0}(U) \cap U \neq \emptyset$, entonces $T^{n_0+1}(U) \cap U = \emptyset$. De lo contrario la longitud del intervalo unión $T^{n_0+1}(U) \cup U \cup T^{n_0}(U)$ sería menor que $3\epsilon < \min\{\alpha, 1 - \alpha\}$, pero contendría dos puntos x_0 y $T(x_0) = x_0 + \alpha$ que distan $\min\{\alpha, 1 - \alpha\}$. \square

Ejercicio 2.7.6. Sea T medible, invertible con inversa medible, que preserve una medida de probabilidad μ . Probar que T es *mixing* si y solo si T^{-1} lo es. (Sugerencia: para dos conjuntos A y B cualesquiera $\mu(T^{-n}(A) \cap B) = \mu(T^n(B) \cap A)$.)

Ejemplo 2.7.7. Sea en $S^1 = [0, 1]/\sim$ mód. 1 el tent map T definido por $T(x) = 2x$ si $0 \leq x \leq 1/2$, y $T(x) = 2 - 2x$ si $1/2 \leq x \leq 1$. Probaremos el siguiente resultado: La medida de Lebesgue m en S^1 es *mixing* para el tent map T ; luego es ergódica.

Demostración. Por definición de *mixing*, debemos probar que

$$\lim_{n \rightarrow +\infty} m(T^{-n}A \cap B) = m(A)m(B) \quad (2.15)$$

para toda pareja de conjuntos medibles A y B . Si $A = \emptyset$ ó $B = \emptyset$, entonces la igualdad anterior es trivialmente igual a cero. Consideremos entonces el caso en que A y B son no vacíos. Fijemos el boreliano A no vacío. Primero probaremos (2.15) cuando B es un intervalo abierto; luego cuando B es un abierto, después para B compacto, y finalmente para B boreliano cualquiera.

Sea B es un intervalo abierto, con longitud a . Para cada $n \geq 1$, la gráfica de T^n está compuesta por 2^n segmentos, con pendiente 2^n (en valor absoluto) cada uno. La imagen por T^n de cualquier intervalo de longitud 2^{n-1} cubre todo el intervalo $[0, 1]$ (croquizar la gráfica). La preimagen por T^n de cualquier boreliano A no vacío está formada por 2^n copias de A , todas homotéticas a A con razón $1/2^n$ y equidistantes en el segmento $[0, 1]$. Luego, la intersección $T^{-n}(A) \cap B$, cuando B es un intervalo de

longitud a , contiene $k = \lceil \text{parte entera}(a \cdot 2^n) \rceil - 1$ de esas copias homotéticas a A , y menos de $k + 2$ de esas copias. Luego:

$$\begin{aligned} (\text{parte entera}(a \cdot 2^n) - 1) \frac{m(A)}{2^n} &\leq m(T^{-n}(A) \cap B) \\ &\leq (\text{parte entera}(a \cdot 2^n) + 1) \frac{m(A)}{2^n}. \end{aligned}$$

De las desigualdades anteriores se deduce que $\lim_n m(T^{-n}(A) \cap B) = a \cdot m(A) = m(B)m(A)$. Hemos probado (2.15) cuando B es un intervalo.

Ahora consideremos el caso en que B es un abierto no vacío en el círculo. $B = \bigcup I_j$ donde $\{I_j\}$ es una colección finita o infinita numerable de intervalos abiertos disjuntos dos a dos. Dividimos en dos subcasos: Si B es unión finita de intervalos disjuntos dos a dos, o si es unión infinita numerable. Si $B = \bigcup_{j=1}^N I_j$, tenemos

$$\begin{aligned} \lim_{n \rightarrow +\infty} m(T^{-n}(A) \cap B) &= \lim_{n \rightarrow +\infty} \sum_{j=1}^N m(T^{-n}(A) \cap I_j) = \\ &= \sum_{j=1}^N \lim_{n \rightarrow +\infty} m(T^{-n}(A) \cap I_j) = \sum_{j=1}^N m(A)m(I_j) = m(A) \cdot m(B). \end{aligned}$$

Si $B = \bigcup_{j=1}^{+\infty} I_j$, dado $\epsilon > 0$ existe N tal que

$$0 \leq m(B) - m(B_N) = m(B \setminus B_N) < \epsilon, \quad (2.16)$$

donde $B_N = \bigcup_{j=1}^N I_j$ Luego:

$$\begin{aligned} 0 &\leq m(T^{-n}(A) \cap B) - m(T^{-n}(A) \cap B_N) = \\ &= m(T^{-n}(A) \cap (B \setminus B_N)) \leq m(B \setminus B_N) < \epsilon \quad \forall n \geq 0. \end{aligned}$$

Es decir:

$$0 \leq m(T^{-n}(A) \cap B) - m(T^{-n}(A) \cap B_N) < \epsilon \quad \forall n \geq 0 \quad (2.17)$$

Por lo probado antes $\lim_{n \rightarrow +\infty} m(T^{-n}(A) \cap B_N) = m(A)m(B_N)$. Entonces, para todo n suficientemente grande

$$|m(T^{-n}(A) \cap B_N) - m(A)m(B_N)| < \epsilon. \quad (2.18)$$

Reuniendo las desigualdades (2.16), (2.17), (2.18), obtenemos, para todo n suficientemente grande, la siguiente cadena de desigualdades:

$$\begin{aligned} |m(T^{-n}(A) \cap B) - m(A)m(B)| &\leq |m(T^{-n}(A) \cap B) - m(T^{-n}(A) \cap B_N)| + \\ &+ |m(T^{-n}(A) \cap B_N) - m(A)m(B_N)| + m(A)|m(B_N) - m(B)| < 3\epsilon. \end{aligned}$$

Lo anterior demuestra que $\lim_{n \rightarrow +\infty} m(T^{-n}(A) \cap B) = m(A)m(B)$. Hemos probado la igualdad (2.15) cuando B es abierto. Ahora demostremos que si la igualdad (2.15)

se cumple para un conjunto B , entonces también se cumple para su complemento B^c . En efecto:

$$\begin{aligned} m(T^{-n}(A) \cap B^c) &= m(T^{-n}(A)) - m(T^{-n}(A) \cap B) = \\ &= m(A) - m(T^{-n}(A) \cap B) \rightarrow_n m(A) - m(A)m(B) = m(A)m(B^c). \end{aligned}$$

Entonces, como (2.15) vale para todos los abiertos B , y es una propiedad cerrada en complementos, se cumple también para todos los compactos B . Ahora probémosla para cualquier boreliano B . Dado ϵ , existe un compacto K y un abierto V tales que $K \subset B \subset V$ y $m(V \setminus K) < \epsilon$. Luego

$$m(V) = m(B) + m(V \setminus B) \leq m(B) + m(V \setminus K) < m(B) + \epsilon.$$

Análogamente

$$m(K) = m(B) - m(B \setminus K) \geq m(B) - m(V \setminus K) > m(B) - \epsilon.$$

Entonces:

$$\begin{aligned} m(T^{-n}(A) \cap B) &\leq m(T^{-n}(A) \cap V) \rightarrow_n m(A)m(V) \leq m(A)m(B) + \epsilon, \\ m(T^{-n}(A) \cap B) &\geq m(T^{-n}(A) \cap K) \rightarrow_n m(A)m(K) \geq m(A)m(B) - \epsilon. \end{aligned}$$

Lo anterior prueba que para todo n suficientemente grande

$$|m(T^{-n}(A) \cap B) - m(A)m(B)| < 2\epsilon,$$

de donde se deduce que la igualdad (2.15), como queríamos demostrar. \square

2.8. Medida de Haar

Definición 2.8.1. Un *grupo topológico* X es un espacio topológico de Hausdorff, en el que hay definida una operación binaria \cdot que lo transforma en grupo (es decir, tiene propiedad asociativa, de neutro 1, e inverso x^{-1} para todo $x \in X$), tal que $\cdot : X \times X \mapsto X$ y $^{-1} : X \mapsto X$ son continuas.

El siguiente teorema, que solo demostraremos con hipótesis adicionales muy restrictivas, generaliza para otros grupos topológicos compactos, la propiedad de la medida de Lebesgue para las rotaciones en el círculo y en el toro n -dimensional:

Teorema 2.8.2 (Medida de Haar). *Dado un grupo topológico X de Hausdorff y localmente compacto, existe una única medida de probabilidad en los borelianos de X que es invariante por todas las traslaciones a la izquierda. Además esta medida es invariante también por todas las traslaciones a la derecha.*

Este Teorema, en la hipótesis adicional de que el grupo topológico X es compacto, fue demostrado primeramente por Von Neumann. Luego, la generalización a grupos localmente compactos fue dada por Haar. Una prueba de este Teorema, en el caso de grupos de Lie (esto es grupos topológicos compactos en los que además hay una

estructura de variedad diferenciable y la operación de grupo es diferenciable) se encuentra en [Mañ 1983a, cap. 7, pág. 58]. La demostración del Teorema 2.8.2 para grupos de Hausdorff localmente compactos cualesquiera, se encuentra en [Fo 1984, Theorem 10.5, pag. 315].

Definición 2.8.3. La medida del Teorema 2.8.2 se llama *medida de Haar*. La medida de Haar en el toro n -dimensional es la medida de Lebesgue.

Proposición 2.8.4. *Sea X un grupo topológico compacto abeliano (es decir vale la conmutativa para la operación de grupo) metrizable, y sea T una traslación que tiene una órbita densa. Entonces T es únicamente ergódica.*

Demostración: La misma prueba de que la rotación irrotacional en el círculo es únicamente ergódica (Proposición 2.6.1), demuestra esta proposición, con las siguientes modificaciones:

En vez de usar $\|y - z\| < \delta$, debe usarse $\text{dist}(y - z, 0) < \delta$. Probemos que si f es continua, dado $\epsilon > 0$, existe $\delta > 0$ tal que, para todos y, z tales que $\text{dist}(y - z, 0) < \delta$, se cumple $|f(y) - f(z)| < \epsilon$. Por absurdo, supóngase que existen $\epsilon > 0$, y sucesiones y_n e z_n , tales que

$$\text{dist}(y_n - z_n, 0) < \frac{1}{n}, \quad |f(y_n) - f(z_n)| \geq \epsilon.$$

Tomando subsucesiones de y_n y z_n convergentes a y y z respectivamente, se tiene $\text{dist}(y - z, 0) = 0$, es decir $y = z$. Por otro lado, como f es continua, $|f(y) - f(z)| \geq \epsilon$, contradiciendo que $y = z$. El resto de la prueba de 2.6.1 se aplica sin otros cambios. \square

2.8.5. Como consecuencia, se tiene una prueba del teorema 2.8.2 de la medida de Haar, con las siguientes hipótesis adicionales (que son muy restrictivas):

El grupo topológico es compacto y abeliano, su topología es metrizable y existe una traslación con alguna órbita densa:

Demostración parcial del teorema 2.8.2: Sea T_{x_0} una traslación que tiene alguna órbita densa. $T_{x_0}(x) = x + x_0$. Por la proposición 2.8.4, T_{x_0} es únicamente ergódica. Sea μ la única medida de probabilidad que es invariante por T_{x_0} . Basta demostrar que μ es también invariante por cualquier otra traslación.

Sea $x_1 \in X$ y sea $T_{x_1}(x) = x + x_1$. Basta demostrar que para toda $f \in C^0(X, \mathbb{R})$ se cumple:

$$\int f \circ T_{x_1} d\mu = \int f d\mu$$

Fijemos una función $f \in C^0(X, \mathbb{R})$. Dado $\epsilon > 0$ sea $\delta > 0$ tal que, si la diferencia de dos puntos dista de 0 menos que δ , entonces los valores de f en ellos difieren, en módulo, menos que ϵ . Por la densidad de la órbita según T_{x_0} , existe $M \geq 0$ tal que $\text{dist}(T_{x_0}^M(0) - x_1, 0) < \delta$. Luego

$$\|f \circ T_{x_1} - f \circ T_{x_0}^M\|_0 < \epsilon$$

Siendo μ una medida T_{x_0} - invariante, la integral respecto de μ de f es igual a la integral de $f \circ T_{x_0}^M$. Entonces se tiene:

$$\left| \int f \circ T_{x_1} d\mu - \int f d\mu \right| < \epsilon$$

Como lo anterior vale para cualquier $\epsilon > 0$ se tiene la igualdad de las integrales de f , como queríamos demostrar. \square

Capítulo 3

Dinámica diferenciable: Hiperbolicidad uniforme y no uniforme

El siguiente ejemplo es conocido como “Arnold’s cat map” (el mapa del gato de Arnold), porque su dinámica fue representada por el matemático ruso Vladimir Arnold en [Arn-Ave 1967], con un dibujo que se hizo famoso, y que imitamos en la Figura 3.1 de este capítulo. (Ver también otro dibujo parecido al original, por ejemplo en [Pi 2007, Figure 2].) En su dibujo, Arnold utiliza el contorno de la figura estilizada de un gato, en lugar de una manzana como hacemos nosotros en la Figura 3.1. Arnold “muestra” el efecto de la propiedad de mixing del mapa sobre los trazos del contorno del gato, el cual, en pocos iterados, se vuelve irreconocible.

3.1. Ejemplo de automorfismo lineal hiperbólico en el toro.

Sea el toro $\mathbb{T}^2 = \mathbb{R}^2 / \sim$ donde la relación de equivalencia \sim está dada por: $(a, b) \sim (c, d)$ en \mathbb{T}^2 si $c - a$ y $d - b$ son enteros. (Otra notación que se usa para \mathbb{R}^2 / \sim es $\mathbb{R}^2 /_{\text{mod } \mathbb{Z}^2} = \mathbb{R}^2 / \mathbb{Z}^2$.)

Sea $\Pi : \mathbb{R}^2 \mapsto \mathbb{Z}^2$ la proyección del espacio de recubrimiento \mathbb{R}^2 del toro definida por $\Pi(a, b) = (a, b)_{\text{mod } \mathbb{Z}^2}$ donde esto último indica la clase de equivalencia de $(a, b) \in \mathbb{R}^2$. Sea $f : \mathbb{T}^2 \mapsto \mathbb{T}^2$ dada por

$$f(x) = \Pi\left(\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} (\Pi^{-1}(x))\right) \quad \forall x \in \mathbb{T}^2.$$

Llamaremos a f automorfismo lineal hiperbólico de matriz $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro, o

simplemente “ $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro”. Abusando de la notación, a un punto $x \in \mathbb{T}^2$ lo denotaremos con cualquier representante suyo $(a, b) \in \mathbb{R}^2$: $(a, b) \in \Pi^{-1}(\{x\})$.

Ejercicio 3.1.1. Probar que $(0, 0)$ es el único punto fijo por la transformación $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro. Probar que

$$\{(1/5, 2/5), (4/5, 3/5)\} \quad \text{y} \quad \{(3/5, 1/5), (2/5, 4/5)\}$$

son dos órbitas periódicas por f y que son las únicas de período 2. Probar que

$$\{(1/2, 1/2), (1/2, 1), (1, 1/2)\}$$

es una órbita periódica de período 3.

La topología en el toro es el cociente de la topología usual en \mathbb{R}^2 . Es metrizable y la métrica está dada por

$$\text{dist}(x, y) = \min\{\sqrt{(c-a)^2 + (d-b)^2} : (a, b) \in \Pi^{-1}(x), (c, d) \in \Pi^{-1}(y)\}.$$

La medida de Lebesgue en el toro es la medida de Borel \tilde{m} definida por $\tilde{m}(B) = m(\Pi^{-1}B \cap [0, 1]^2)$ donde m es la medida de Lebesgue en \mathbb{R}^2 . Se observa que la medida de Lebesgue en el toro es una medida de probabilidad. Donde no dé lugar a confusión renombraremos como m a la medida de Lebesgue en el toro.

Proposición 3.1.2. *La medida de Lebesgue m en el toro es invariante por la transformación $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.*

Demostración:

Denotamos $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ a la matriz de coeficientes enteros que define la transformación f en el toro. Observamos que $\det(A) = 1$. Sea B un boreliano en el toro.

$$m(f^{-1}(B)) = \int \chi_{f^{-1}(B)}(x) dm(x) = \int \chi_B \circ f(x) dm(x).$$

Haciendo el cambio de variables lineal e invertible $z = f(x)$ en la integral anterior resulta

$$m(f^{-1}(B)) = \int \chi_B(z) J(z) dm(z)$$

donde $J(z) = |\det df^{-1}(z)| = |\det df(x)|^{-1}$ es el jacobiano del cambio de variables $z = f(x)$. En nuestro caso una parametrización local de la superficie del toro \mathbb{T}^2 está dada por $\Pi|_{B_\delta(\Pi^{-1}(x))}$, donde $B_\delta(\Pi^{-1}(x))$ es la bola abierta en \mathbb{R}^2 de radio $\delta > 0$ (suficientemente pequeño) y centro en un punto denotado como $\Pi^{-1}(x) \in \mathbb{R}^2$ que se

proyecta por $\Pi : \mathbb{R}^2 \mapsto \mathbb{T}^2$ en el punto $x \in \mathbb{T}^2$. Calculando la derivada $df(x)$ y el Jacobiano con las coordenadas en esa carta local, resulta $T_x\mathbb{T}^2 \sim \mathbb{R}^2$, y

$$J(z) = (\det A)^{-1} = 1 \quad \forall z \in \mathbb{T}^2.$$

Entonces $m(f^{-1}(B)) = \int \chi_B(z) dm(z) = m(B)$. □

Dinámica del ejemplo de automorfismo lineal hiperbólico en el toro.

Estudiamos la dinámica por iterados de la transformación lineal $F : \mathbb{R}^2 \mapsto \mathbb{R}^2$ que tiene como matriz $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

F se llama levantado de f a \mathbb{R}^2 y f se llama proyección de F en el toro. La dinámica de F está relacionada fuertemente con la dinámica de su proyección f en el toro.

En efecto la proyección $\Pi : \mathbb{R}^2 \mapsto \mathbb{T}^2$ cuando restringida a un entorno suficientemente pequeño del origen $(0,0)$ es un homeomorfismo sobre su imagen.

Π transforma la dinámica de F en la de f . Más precisamente

$$\Pi \circ F = f \circ \Pi$$

Luego, aplicando $\Pi : \mathbb{R}^2 \mapsto \mathbb{T}^2$ a una órbita de F se obtiene una órbita de f . Y la preimagen por Π de una órbita por f es una infinidad numerable de órbitas por F tal que una se obtiene de otra trasladándola en \mathbb{R}^2 según un vector de coordenadas enteras.

El comportamiento topológico local de las órbitas de F en un entorno suficientemente pequeño V del origen es el mismo (a menos del homeomorfismo $\Pi|_V$) que el de las órbitas de f en el abierto $\Pi(V)$.

Variedades invariantes.

Los valores propios de la matriz A son

$$\sigma := (3 + \sqrt{5})/2 > 1, \quad 0 < \lambda := (3 - \sqrt{5})/2 < 1.$$

Las direcciones propias respectivas tienen pendientes irracionales, la primera positiva y la segunda negativa. Se deduce que las dos rectas que pasan por el origen y tienen direcciones según los vectores propios de la matriz A son invariantes por F en el plano. Entonces sus proyecciones en el toro son curvas invariantes por f y se cortan transversalmente en el origen (y también se cortan transversalmente en todos sus otros puntos de intersección en el toro).

F restringida a la recta r_1 que tiene dirección propia de valor propio mayor que 1, expande las distancias exponencialmente con tasa $\log(3 + \sqrt{5})/2 > 0$. Es decir, contrae distancias hacia el pasado como $\sigma^{-n} = e^{-n \log \sigma}$:

$$\frac{\text{dist}(F^{-n}(a, b), (0, 0))}{\text{dist}((a, b), (0, 0))} = e^{-n \log \sigma} \quad \forall n \in \mathbb{N}, \quad \forall (a, b) \in r_1 \setminus (0, 0).$$

Proyectando r_1 en el toro se obtiene una curva $W^u(0, 0) = \Pi(r_1)$ inmersa en el toro, que se llama *variedad inestable por* $(0, 0)$.

$W^u(0,0)$ pasa por el origen, *es densa en el toro* (esto se puede demostrar usando que la pendiente de r_1 en el plano es irracional y usando que la rotación irracional en el círculo es densa en el círculo), *es invariante por f y cumple:*

$$W^u(0,0) = \{y \in \mathbb{T}^2 : \lim_{n \rightarrow -\infty} f^n(y) = (0,0)\}$$

Esto último se debe a que $r_1 = \{(a,b) \in \mathbb{R}^2 : \lim_{n \rightarrow -\infty} F^n(a,b) = (0,0)\}$. Observamos que la subvariedad $W^u(0,0) = \Pi(r_1)$ es inmersa y densa en \mathbb{T}^2 , pero *no es subvariedad encajada en \mathbb{T}^2* . Es decir, la topología que se define a lo largo de la subvariedad $W^u(0,0)$ no es la inducida por su inclusión en \mathbb{R}^2 . Los abiertos en $W^u(0,0)$ están generados por los arcos abiertos (homeomorfos a intervalos abiertos en la recta real). Estos no se obtienen como intersección de un abierto en \mathbb{T}^2 con $W^u(0,0)$ pues cualquier abierto en \mathbb{T}^2 cortado con $W^u(0,0)$ contiene una infinidad de arcos conexos, debido a la densidad de $W^u(0,0)$.

Variedad inestable local: En este ejemplo, existe $\epsilon > 0$ suficientemente pequeño, tal que, denotando $W_{loc}^u(0,0)$ (variedad inestable local) a la componente conexa de $W^u(0,0)$ intersecada con la bola de centro $(0,0)$ y radio ϵ en el toro \mathbb{T}^2 , se tiene:

$$\frac{\text{dist}(f^{-n}(y), (0,0))}{\text{dist}(y, (0,0))} = e^{-n \log \sigma} \quad \forall n \in \mathbb{N}, \quad \forall y \in W_{loc}^u(0,0).$$

Esta igualdad se obtiene porque la bola de centro $(0,0)$ y radio $\epsilon > 0$ en el toro \mathbb{T}^2 es difeomorfa por un preimagen de Π , con la bola de centro $(0,0)$ y radio $\epsilon > 0$ en \mathbb{R}^2 , si $\epsilon > 0$ es suficientemente pequeño.

Análogamente F restringida a la recta r_2 que tiene dirección propia de valor propio menor que 1, contrae las distancias exponencialmente con tasa $\log(3 - \sqrt{5})/2 < 0$. Es decir, contrae distancias hacia el futuro como $\lambda^n = e^{n \log \lambda}$:

$$\frac{\text{dist}(F^n(a,b), (0,0))}{\text{dist}((a,b), (0,0))} = e^{n \log \lambda} \quad \forall n \in \mathbb{N}, \quad \forall (a,b) \in r_2 \setminus (0,0).$$

Proyectando r_2 en el toro se obtiene una curva $W^s(0,0) = \Pi(r_2)$ inmersa en el toro, que se llama *variedad estable por* $(0,0)$.

$W^s(0,0)$ pasa por el origen, *es densa en el toro* (porque la pendiente de r_2 en el plano es irracional), *es invariante por f , no encajada en \mathbb{R}^2 y cumple:*

$$W^s(0,0) = \{y \in \mathbb{T}^2 : \lim_{n \rightarrow +\infty} f^n(y) = (0,0)\}$$

Esto último se debe a que $r_2 = \{(a,b) \in \mathbb{R}^2 : \lim_{n \rightarrow +\infty} F^n(a,b) = (0,0)\}$.

Denotando $W_{loc}^s(0,0)$ (variedad estable local) a la componente conexa de $W^s(0,0)$ intersecada con la bola de centro $(0,0)$ y radio ϵ en el toro \mathbb{T}^2 , se tiene:

$$\frac{\text{dist}(f^n(y), (0,0))}{\text{dist}(y, (0,0))} = e^{n \log \lambda} \quad \forall n \in \mathbb{N}, \quad \forall y \in W_{loc}^s(0,0).$$

Variedades invariantes por cualquier punto: El argumento anterior puede aplicarse a cualquier punto periódico x . Deducimos que todos los puntos periódicos son

hiperbólicos tipo silla (tienen un valor propio mayor que uno y otro positivo menor que uno).

En general, para cualquier punto $x \in \mathbb{T}^2$, aunque no sea periódico, la variedad estable $W^s(x)$ y la variedad inestable $W^u(x)$ se definen como la proyecciones sobre el toro de las rectas en \mathbb{R}^2 que pasan por $\Pi^{-1}(x)$, según las direcciones de los vectores propios de la matriz A (que son las mismas que las direcciones de las rectas r_1 y r_2 en \mathbb{R}^2 que pasan por el origen). Argumentando como más arriba se tiene

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(y), f^n(x)) = 0 \quad \forall y \in W^s(x)$$

$$\lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(y), f^{-n}(x)) = 0 \quad \forall y \in W^u(x),$$

y el acercamiento a cero de esas distancias se realiza como λ^n o σ^{-n} respectivamente. Las variedad inestable por un punto x no es invariante si el punto x no es fijo por f , pero su imagen por f es la variedad inestable por el punto $f(x)$ (esto se chequea inmediatamente de la construcción de $W^u(x)$ y $W^u(f(x))$ como las imágenes por Π de las rectas paralelas a r_1 por $\Pi^{-1}(x)$ y $\Pi^{-1}(f(x)) = F(\Pi^{-1}(x))$ respectivamente).

Foliaciones invariantes:

La familia de todas las variedades inestables, forman en el toro \mathbb{T}^2 lo que se llama una *foliación* invariante: pues cada subvariedad de la foliación (llamada hoja), al aplicarle f se transforma en otra hoja de la foliación. Análogamente, la foliación formada por las variedades estables, es invariante.

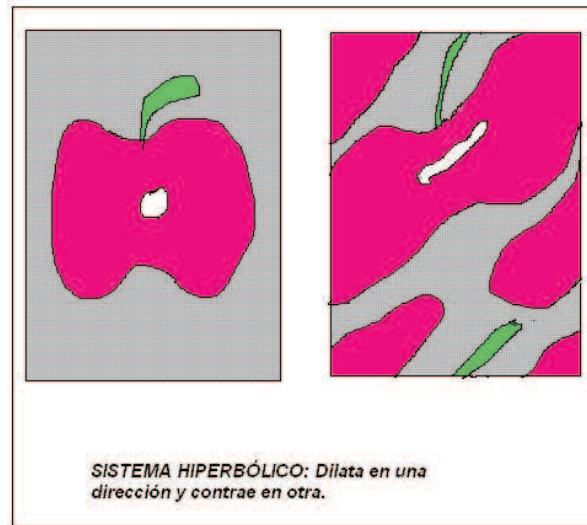


Figura 3.1: Deformación producida por el automorfismo lineal en el toro.

Interpretación gráfica del automorfismo lineal hiperbólico: La deformación que produce f en este ejemplo de automorfismo lineal hiperbólico en el toro \mathbb{T}^2 , está representado en la figura 3.1. Esa figura es una modificación de la conocida llamada “gato de Arnold” (que en vez de una manzana, deforma la imagen de un gato, figura creada por Arnold, en la década de 1960 para ilustrar la deformación

hiperbólica de un automorfismo lineal en el toro \mathbb{T}^2). Al iterar sucesivas veces f , la figura representada, se estira a lo largo de la foliación inestable y se contrae a lo largo de la foliación estable. Como f es invertible, los pedazos que se obtienen de identificar 0 con 1 en vertical y horizontal, no se intersecan (proviene de subconjuntos disjuntos antes de aplicar f).

Definición 3.1.3. Exponentes de Lyapunov en puntos fijos hiperbólicos.

Los logaritmos de los módulos de los valores propios de $df^p(x_0)$ dividido p , cuando x_0 es un punto periódico de período p , se llaman *exponentes de Lyapunov* en x_0 . En el capítulo 3 definiremos con generalidad los exponentes de Lyapunov de cualquier sistema dinámico diferenciable, para casi todas sus órbitas (no necesariamente puntos fijos ni órbitas periódicas).

En el ejemplo del $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ f es lineal (mejor dicho F , dada por la matriz A , es lineal), por lo tanto la matriz A es, en una base adecuada de $T_{(0,0)}\mathbb{T}^2$, la derivada $df(0,0) : T_{(0,0)}\mathbb{T}^2 \sim \mathbb{R}^2 \mapsto T_{(0,0)}\mathbb{T}^2 \sim \mathbb{R}^2$. Los logaritmos de los valores propios de $A = df(0,0)$ son dos, uno positivo $\chi^+ = \log \sigma$ y el otro negativo $\chi^- = \log \lambda$. (El origen es un punto fijo hiperbólico tipo silla, pues $0 < \lambda < 1 < \sigma$).

El exponente de Lyapunov $\chi^+ := \log \mu = \log(3 + \sqrt{5})/2 > 0$ es la *tasa exponencial de dilatación* al aplicar $df_{(0,0)}$, de las normas de los vectores en el subespacio propio U , que corresponde al valor propio μ de $df_{(0,0)}$. Precisamente:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_{(0,0)}^n u\|/\|u\|)}{n} = \log \mu = \chi^+ > 0 \quad \forall u \in U \setminus \{0\} \subset T_x M.$$

Observar que para $n < 0$ también se cumple la misma igualdad anterior; es decir:

$$\lim_{n \rightarrow -\infty} \frac{\log(\|df_{(0,0)}^n u\|/\|u\|)}{n} = \log \mu = \chi^+ \quad \forall u \in U \setminus \{0\} \subset T_x M.$$

Además, el exponente de Lyapunov positivo χ^+ es la tasa exponencial de dilatación de las distancias a lo largo de la variedad inestable local por el punto fijo. Más precisamente

$$\lim_{n \rightarrow -\infty} \frac{\log \text{dist}(f^n(y), (0,0))}{n} = \chi^+ = \log \mu > 0 \quad \forall y \in W^u(0,0)$$

Análogamente, el exponente de Lyapunov $\chi^- := \log \lambda = \log(3 - \sqrt{5})/2 < 0$ negativo, se interpreta como la *tasa exponencial de contracción por f* en un entorno suficientemente pequeño del punto fijo a lo largo de la variedad estable por ese punto. Más precisamente

$$\lim_{n \rightarrow +\infty} \frac{\log \text{dist}(f^n(y), (0,0))}{n} = \chi^- = \log \lambda < 0 \quad \forall y \in W^s(0,0).$$

Ejercicio 3.1.4. Sea f la transformación $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro. Probar que m -c.t.p es recurrente pero que hay puntos no recurrentes (Sugerencia: $y \in W^s(0,0) \setminus \{(0,0)\}$ no

es recurrente). Probar que dados dos abiertos U y V cualesquiera en el toro, existe una subsecuencia n_j de naturales tales que $f^{n_j}U \cap V \neq \emptyset$. (Sugerencia: Usar que $W^u((0,0))$ es invariante por f y pruébese que para cualquier arco compacto $K \subset W^u(0,0)$ la unión de sus iterados hacia el futuro $\bigcup_{n \in \mathbb{N}} f^n(K)$ es densa en el toro.)

Deducir que f es transitiva.

Probar que todo punto es no errante. Concluir que si bien todo punto recurrente es no errante, no necesariamente todo punto no errante es recurrente.

Observación: En la próxima sección probaremos que la transformación

$$f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

en el toro es *ergódica respecto a la medida de Lebesgue*. La ergodicidad en este ejemplo, es un caso particular de la ergodicidad de mapas llamados de Anosov (Definición 3.3.1), de clase C^2 y que preservan la medida de Lebesgue m en una variedad compacta riemanniana M de dimensión finita. Demostraremos el caso general, para los difeomorfismos de Anosov, en el Capítulo 6, Corolario 6.3.5. De todas formas, en la próxima sección y con argumentos diferentes a los que utilizaremos después en el Capítulo 6, demostraremos la ergodicidad en el caso particular del ejemplo 2, 1, 1, 1 en el toro.

Ejercicio 3.1.5. Sea el automorfismo lineal hiperbólico $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ del toro \mathbb{T}^2 .

Sea P el conjunto de puntos periódicos por f . Probar que $P \neq \emptyset$, pero $\mu(P) = 0$ para toda μ invariante y ergódica para f que sea positiva sobre abiertos.

3.2. Ergodicidad del automorfismo lineal hiperbólico en \mathbb{T}^2

Vamos a probar que el automorfismo lineal hiperbólico f en el toro \mathbb{T}^2 , dado por la matriz $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, es *ergódico respecto a la medida de Lebesgue m* . Como m es positiva sobre abiertos, su ergodicidad implica la transitividad topológica de f . Por lo tanto f tiene órbitas densas.

Más adelante en los próximos capítulos probaremos nuevamente la ergodicidad de $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en \mathbb{T}^2 , demostrando que cumple una condición más fuerte (tiene “espectro de Lebesgue con base numerable”), la cual implica además, que T sea mixing.

Teorema 3.2.1. *La transformación $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro es ergódica respecto a la medida de Lebesgue m .*

Demostración: Sea $\psi : \mathbb{T}^2 \mapsto \mathbb{C}$ una función continua compleja. Usando los resultados de los ejercicios 2.4.4 y 2.4.7, basta probar que el límite $\tilde{\psi}$ de los promedios de Birkhoff de ψ es localmente constante c.t.p.

Llamaremos Λ al conjunto de los puntos $p \in \mathbb{T}^2$ donde existen y son iguales los límites de los promedios de Birkhoff hacia el futuro $\psi^+(p)$ y hacia el pasado $\psi^-(p)$. Por el corolario 2.1.3, y teniendo en cuenta que la medida de Lebesgue m es invariante por T , sabemos que $m(\Lambda) = 1$, es decir m -casi todo punto p pertenece a Λ .

Elijamos un punto $p_0 \in \mathbb{T}^2$. Dibujemos un rectángulo pequeño $R \subset V \subset \mathbb{T}^2$ con un vértice en p_0 , y cuyos lados tienen direcciones u y s , siendo u (resp. s) la dirección propia con autovalor mayor que 1 (resp. menor que 1) de la matriz $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. En la Figura 3.2 dibujamos el rectángulo R y tomamos coordenadas locales tales que $p_0 = (0, 0)$ y $p = (x_u, x_s)$ para cualquier $p \in R$.

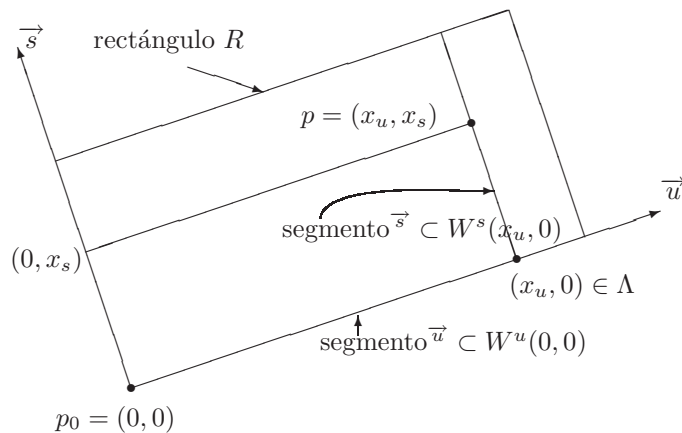


Figura 3.2: Estructura de producto local en el rectángulo R .

Consideremos la variedad estable $W^s(x_u, 0)$ por el punto $(x_u, 0)$, definida por

$$W^s(x_u, 0) := \{p \in \mathbb{T}^2 : \lim_{n \rightarrow +\infty} \text{dist}(T^n(p), T^n(x_u, 0)) = 0\}.$$

Afirmamos que $W^s(x_u, 0)$ contiene, en el rectángulo R , al segmento de dirección s que pasa por el punto $(x_u, 0)$. En efecto, al aplicar T a puntos de este segmento, las distancias se contraen con factor de contracción igual al valor propio $\lambda < 1$ de la matriz A . En particular:

$$(x_u, x_s) \in W^s(x_u, 0).$$

Asumamos hipotéticamente que el punto $(x_u, 0)$ pertenece a Λ .

Afirmamos que, bajo la hipótesis anterior, existe el límite $\psi^+(x_u, x_s)$ de los promedios de Birkhoff hacia el futuro de ψ , en el punto $(x_u, x_s) \in W^s(x_u, 0)$, y además se cumple:

$$\psi^+(x_u, x_s) = \psi^+(x_u, 0) = \psi^-(x_u, 0). \quad (3.1)$$

En efecto, la igualdad de la derecha es debida a que estamos asumiendo, hipotéticamente, que $(x_u, 0) \in \Lambda$, y entonces, por definición del conjunto Λ , los promedios de

Birkhoff hacia el futuro y hacia el pasado en la órbita de $(x_u, 0)$, existen y son iguales. La igualdad de la izquierda es debida a que existe $\psi^+(x_u, 0)$ porque $(x_u, 0) \in \Lambda$ y a que $(x_u, x_s) \in W^s(x_u, 0)$, es decir, $\lim_{n \rightarrow +\infty} (f^n(x_u, x_s), f^n(0, x_s)) = 0$. Por lo probado en el Ejercicio 2.1.12, estas últimas condiciones implican que $\psi^+(x_u, x_s) = \psi^+(x_u, 0)$ para toda función continua $\psi : \mathbb{T}^2 \mapsto \mathbb{C}$. Hemos probado la igualdad (3.1) bajo la hipótesis $(x_u, 0) \in \Lambda$.

Ahora analicemos el límite de los promedios de Birkhoff hacia el pasado por el punto $(x_u, 0)$. La variedad inestable $W^u(0, 0)$ por el punto $(0, 0)$ está definida por

$$W^u(0, 0) := \{q \in \mathbb{T}^2 : \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(q), f^{-n}(0, 0)) = 0\}.$$

Análogamente a lo probado arriba, pero aplicado a f^{-1} en vez de f , deducimos que $W^u(0, 0)$ contiene, en el rectángulo R , al segmento de dirección u que pasa por el punto $(0, 0)$. En efecto, esta dirección u corresponde a la dirección propia con valor propio $\sigma > 1$ de la matriz A . Entonces f dilata distancias (localmente) según la dirección u , con factor σ . Luego f^{-1} contrae distancias según la dirección u con factor de dilatación $\sigma^{-1} < 1$. En particular

$$(x_u, 0) \in W^u(0, 0).$$

Usando nuevamente lo probado en el Ejercicio 2.1.12, afirmamos que:

$$\psi^-(x_u, 0) = \psi^-(0, 0) \tag{3.2}$$

En efecto, la igualdad proviene de la hipótesis $(x_u, 0) \in \Lambda$, por lo cual existe $\psi^-(x_u, 0)$, y aplicando lo probado en el Ejercicio 2.1.12, deducimos que existe también $\psi^-(0, 0)$ y es igual a $\psi^+(x_u, 0)$. Entonces hemos probado la igualdad (3.2), bajo la hipótesis $(x_u, 0) \in \Lambda$.

Reuniendo las igualdades (3.1) y (3.2) deducimos lo siguiente:

Para todo punto $(x_u, x_s) \in R$, si el punto proyección $(x_u, 0)$ está en Λ , entonces $\psi^+(x_u, x_s) = \psi^-(0, 0)$. (Ver Figura 3.2)

Luego, para terminar que probar que ψ^+ es localmente constante, basta probar la siguiente afirmación:

Afirmación (A) a probar: *Para Lebesgue-casi todo punto (x_u, x_s) en el rectángulo R , su proyección $(x_u, 0)$ pertenece a Λ .*

Para probar la afirmación (A), admitiremos, si fuera necesario, re-elegir el punto de referencia, que llamamos $(0, 0)$, sustituyéndolo por algún otro punto del rectángulo R , que tomaremos como nuevo origen de coordenadas para hacer todo el desarrollo anterior.

Probemos la afirmación (A): La medida de probabilidad de Lebesgue m es invariante por f . Aplicando el teorema de Birkhoff y su Corolario 2.1.3, sabemos que $m(\Lambda) = 1$. Además $m(R) > 0$, pues el rectángulo R tiene interior no vacío, y todo abierto tiene medida de Lebesgue positiva. Luego, m -c.t.p. de R cumple $x \in R \cap \Lambda$. Es decir $\chi_\Lambda(x) = 1$ para m -casi todo punto $x \in R$.

La medida de Lebesgue m en el rectángulo R (bidimensional) es la medida producto de la medida de Lebesgue m^u , unidimensional a lo largo de segmentos con dirección u , por la medida de Lebesgue m^s , también unidimensional, a lo largo de segmentos con dirección s . Es decir

$$m = m^u \times m^s.$$

Llamemos $dx_s = dm^s$ y $dx_u = dm^u$. Aplicando el teorema de Fubini, la integral en el rectángulo R de cualquier función integrable respecto a la medida de Lebesgue, se puede escribir como una integral doble, integrando primero en “bastones” paralelos a la dirección u , con x_s constante, e integrando luego en x_s . (Ver Figura 3.3).

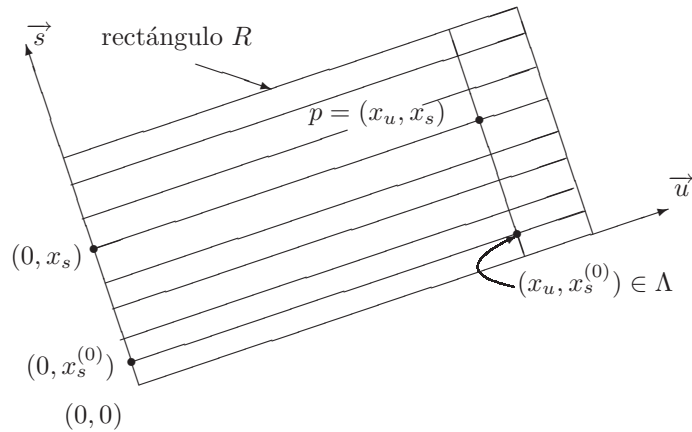


Figura 3.3: Integración por Fubini en el rectángulo R y elección del nuevo punto de referencia $(0, x_s^{(0)})$.

Obtenemos:

$$m(R) = m(R \cap \Lambda) = \int_R \chi_\Lambda dm = \int_0^a dx_s \int_0^b \chi_\Lambda(x_s, x_u) dx_u.$$

Sabemos que $\chi_\Lambda(x_s, x_u)$ solo toma valores 0 ó 1 en cualquier punto (x_s, x_u) , pero para m^s -casi todo punto (x_s, x_u) , vale 1. Esto solo se puede cumplir, usando la igualdad de la integral doble de arriba, si para m^s -casi todo valor *fijo* de x_s , la integral $\int_0^b \chi_\Lambda(x_s, x_u) dx_u = 1$. Esto último solo se puede cumplir si la función $x_u \mapsto \chi_\Lambda(x_u, x_s)$, con x_s constante, es igual a 1 para m^u -casi todo valor de x_u .

Hemos probado lo siguiente: *Para m^s -casi todo valor fijo de x_s , el segmento de dirección u que pasa por $(0, x_s)$ tiene m^u -casi todo punto en Λ .* Luego, existe algún punto $(0, x_s^{\{0\}})$ tal que $(x_u, x_s^{\{0\}}) \in \Lambda$ para m^u -casi todo valor de x_u . Para todo punto (x_u, x_s) consideramos su proyección $(x_u, x_s^{\{0\}})$. (Ver Figura 3.3). Deducimos que para Lebesgue-casi todo punto (x_u, x_s) su proyección $(x_u, x_s^{\{0\}})$ pertenece a Λ . Luego, si se toma el punto $(0, x_s^{\{0\}})$ como nuevo origen de coordenadas, se cumple la afirmación (A) como queríamos demostrar. \square

3.3. Difeomorfismos de Anosov e hiperbolicidad uniforme

En esta sección asumiremos que M es una variedad diferenciable, compacta y conexa, provista de una estructura riemanniana (i.e. existe un producto interno $\langle \cdot, \cdot \rangle: TM \times TM \mapsto \mathbb{R}$ diferenciable, definido en el fibrado tangente TM). Por lo tanto, para todo $x \in M$ y para todo $v \in T_x M$, está definida la norma $\|v\| := \sqrt{\langle v, v \rangle_x}$, determinada por la métrica riemanniana en M).

Sea $f: M \mapsto M$ un difeomorfismo, lo que se denota $f \in \text{Diff}^1(M)$ y significa que f es de clase C^1 , invertible, y su inversa $f^{-1}: M \mapsto M$ también es de clase C^1 . Si además f y f^{-1} fueran de clase C^r para algún natural $r > 1$, se denota $f \in \text{Diff}^r(M)$ (para lo cual se requiere que M sea una variedad de clase C^r por lo menos). En esta sección, asumiremos que $f \in \text{Diff}^1(M)$ e indicaremos expresamente cuando además $f \in \text{Diff}^r(M)$ para algún $r > 1$.

Definición 3.3.1. Difeomorfismos de Anosov [An 1967]

$f \in \text{Diff}^1(M)$ se llama *difeomorfismo de Anosov* si existe una descomposición (llamada *splitting*) $S \oplus U = TM$ del fibrado tangente TM , no trivial (i.e. $0 < \dim(S) < \dim(M)$), que es invariante por df (i.e. $df(x)S_x = S_{f(x)}$, $df(x)U_x = U_{f(x)}$), y existen constantes $C > 0$ y $0 < \lambda < 1 < \sigma$, tales que, para todo $x \in M$:

$$\|df^n(x)s\| \leq C\lambda^n \|s\| \quad \forall n \geq 0, \quad \forall s \in S_x, \quad (3.3)$$

$$\|df^n(x)u\| \geq C^{-1}\sigma^n \|u\| \quad \forall n \geq 0, \quad \forall u \in U_x. \quad (3.4)$$

Para cada punto $x \in M$, los subespacios S_x y U_x se llaman *estable e inestable* respectivamente, en x . Los subfibrados S y U (formados por los subespacios S_x y U_x al variar $x \in M$), se llaman fibrados estable e inestable respectivamente. La constante $0 < \lambda < 1$ se llama tasa o coeficiente de contracción (uniforme) en el futuro a lo largo del fibrado estable, y la constante $\sigma > 1$, tasa o coeficiente de dilatación (uniforme) en el futuro a lo largo del fibrado inestable.

Nota: En la definición de difeomorfismo de Anosov, la constante C no depende de x , así como tampoco dependen de x los coeficientes λ y σ de contracción y dilatación respectivamente. Siendo la variedad M compacta, si se cambia la métrica Riemanniana, la norma de los vectores en cada subespacio tangente se cambia por una equivalente. Luego, se modifica la constante C por otra constante C' (que tampoco depende de x) manteniéndose las mismas tasas λ y σ de contracción y dilatación respectivamente. Esto permite que la definición de difeomorfismo de Anosov sea intrínseca al difeomorfismo, y no dependa de la métrica riemanniana elegida. Se puede probar que para todo f de Anosov existe una métrica riemanniana (llamada *métrica adaptada a f*) para la cual la constante C puede tomarse igual a 1).

Ejercicio 3.3.2. Probar que el autormorfismo lineal $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro \mathbb{T}^2 es un difeomorfismo de Anosov. Sugerencia: Tomar S_x y U_x las direcciones propias de la matriz A que define a f , λ y μ los valores propios respectivos y $C = 1$.

Ejercicio 3.3.3. Sea f un difeomorfismo de Anosov. Probar que la inversa $0 < \sigma^{-1} < 1$ de la tasa σ de dilatación hacia el futuro a lo largo del fibrado inestable, es una tasa de contracción hacia el pasado a lo largo del mismo fibrado. Precisamente:

$$\|df^{-n}(x)u\| \leq C\sigma^{-n}\|u\| \quad \forall n \geq 0, \quad \forall u \in U_x, \quad \forall x \in M. \quad (3.5)$$

Análogamente, probar que la inversa $\lambda^{-1} > 1$ de la tasa de contracción λ hacia el futuro a lo largo del fibrado estable, es una tasa de dilatación hacia el futuro a lo largo del mismo fibrado estable.

Deducir que $f : M \mapsto M$ es un difeomorfismo de Anosov, si y solo si f^{-1} también lo es.

Ejercicio 3.3.4. Mostrar que el splitting $S_x \oplus U_x$ es único. Sugerencia: Probar, para todas las direcciones $[s] \subset S_x$ y $[v] \subset T_x M \setminus S_x$, las siguientes desigualdades:
 $\limsup_{n \rightarrow +\infty} \frac{1}{n} \log \|df^n s\| < 0 < \limsup_{n \rightarrow +\infty} \frac{1}{n} \log \|df^n v\|$.

Ejercicio 3.3.5. Probar que las aplicaciones $x \mapsto S_x$ y $x \mapsto U_x$ son continuas y deducir que si M es conexa, entonces $\dim S_x$ y $\dim U_x$ son constantes.

Sugerencia: Considerar una sucesión $\{x_n\}$ convergente a x y tomar una subsucesión de ella tales que S_{x_n} y U_{x_n} tengan dimensiones constantes con n y sean convergentes. Mostrar que los subespacios $\lim_n S_{x_n}$ y $\lim_n U_{x_n}$ satisfacen las desigualdades de hiperbolicidad uniforme, y forman un splitting de $T_x M$. Finalmente, usar la unicidad del splitting hiperbólico en el punto $x \in M$.

Ejercicio 3.3.6. Probar que en la Definición 3.3.1, la condición $0 < \dim(S_x) < \dim(M)$ es redundante.

Sugerencia: Por absurdo, si $\dim(S_x) = \dim(M)$, elegir $N \geq 1$ tal que $C\lambda^N < 1/4$. Usando la definición de diferenciabilidad de f en el punto x , probar que existe $\delta_x > 0$ tal que:

$$\text{dist}(x, y) < \delta_x \Rightarrow \text{dist}(f(x), f(y)) \leq \frac{\text{dist}(x, y)}{2}.$$

Cubrir M con una cantidad k finita de bolas abiertas $\{B_{\delta_i}(x_i)\}_{1 \leq i \leq k}$ de centros $x_1, \dots, x_i, \dots, x_k$ y radios $\delta_i := \delta_{x_i}$. Probar que para todo $y \in M$, y para todo $n \in \mathbb{N}$ existe x_i (que depende de y y de n) tal que $\text{dist}(f^{nN}(x_i), f^{nN}(y)) < \epsilon$. Deducir que $f^{nN}(M)$ se puede cubrir con k bolas de radio ϵ . Siendo f^{nN} un difeomorfismo, toda la variedad M se puede cubrir con k bolas de radio $\epsilon > 0$, siendo k constante y $\epsilon > 0$ arbitrario. Deducir que $\text{diam}(M) < k\epsilon$ para todo $\epsilon > 0$, lo cual es una contradicción pues el diámetro de M es positivo.

Ejercicio 3.3.7. Sea $f : M \mapsto M$ un difeomorfismo de Anosov con tasa $\lambda < 1$ de contracción a lo largo del fibrado estable S , y tasa $\sigma > 1$ de dilatación a lo largo del fibrado inestable U . Probar que para todo $x \in M$:

$$\limsup_{n \rightarrow +\infty} \frac{\log \|df^n(x)s\|}{n} \leq \log \lambda < 0 \quad \forall s \in S_x$$

$$\limsup_{n \rightarrow -\infty} \frac{\log \|df^n(x)s\|}{-n} \leq \log \lambda < 0 \quad \forall s \in S_x$$

$$\liminf_{n \rightarrow +\infty} \frac{\log \|df^n(x)u\|}{n} \geq \log \sigma > 0 \quad \forall u \in U_x$$

$$\liminf_{n \rightarrow -\infty} \frac{\log \|df^n(x)u\|}{-n} \geq \log \sigma > 0 \quad \forall u \in U_x$$

Exponentes de Lyapunov para difeomorfismos de Anosov:

En las próximas secciones enunciaremos el teorema de Oseledets que establece que los límites superior e inferior de las desigualdades de arriba, son límites que existen para μ -casi todo punto $x \in M$ (donde μ es cualquier medida de probabilidad invariante por f). Esos límites se llaman *exponentes de Lyapunov* (en el futuro) para la órbita de $x \in M$.

Por lo tanto, admitiendo el teorema de Oseledets, de las desigualdades de arriba se deducen los siguientes enunciados, que satisfacen todo difeomorfismo de Anosov f para toda medida de probabilidad μ que sea f -invariante:

(a) *Los exponentes de Lyapunov de la órbita por μ - casi todo punto $x \in M$, correspondientes a las direcciones u del subespacio inestable, son positivos y mayores o iguales que el logaritmo del coeficiente de dilatación $\sigma > 1$ del difeomorfismo de Anosov.*

(b) *Los exponentes de Lyapunov de la órbita por μ - casi todo punto $x \in M$, correspondientes a las direcciones s del subespacio estable, son negativos y menores o iguales que el logaritmo del coeficiente de contracción $\lambda < 1$ del difeomorfismo de Anosov.*

De los enunciados (a) y (b), teniendo en cuenta que por definición de difeomorfismo de Anosov, el espacio tangente $T_x M$ es la suma directa de los subespacios estable S_x e inestable U_x , se deduce el siguiente enunciado (para una demostración del mismo usar el teorema de Oseledets, que probaremos en el próximo capítulo, y ver Ejercicio 3.9.5):

(c) *Los exponentes de Lyapunov para un difeomorfismo de Anosov no son cero, y están "bounded away from zero" (i.e. están fuera de un entorno de cero).*

Observación 3.3.8. .

Sobre la transitividad de los difeomorfismos de Anosov. Una conjetura cuya demostración es aún un problema abierto es la siguiente:

Conjetura: Los difeomorfismos de Anosov en variedades compactas y conexas son transitivos.

Se conocen pruebas parciales: si la variedad M donde actúa el difeomorfismo de Anosov es un toro \mathbb{T}^n , entonces f es transitivo. Otro caso conocido: si la dimensión del fibrado inestable o la del estable es uno (de un difeomorfismo f de Anosov), entonces f es transitivo. En general, no se conocen ejemplos de difeomorfismos f de Anosov que no sean transitivos.

Observación 3.3.9. .

Sobre la ergodicidad de los difeomorfismos de Anosov.

Sea M una variedad compacta y conexas de dimensión $n \geq 2$. Si $f \in \text{Diff}^2(M)$ es un difeomorfismo de Anosov (de clase C^2), si f es transitivo y si f preserva la medida de Lebesgue m , entonces m es ergódica. Demostraremos este resultado en el Corolario 6.3.5, al definir y estudiar las medidas invariantes de probabilidad llamadas SRB (Sinai-Ruelle-Bowen) para los difeomorfismos de Anosov.

3.4. Conjuntos uniformemente hiperbólicos

Definición 3.4.1. Sea $f : M \mapsto M$ y sea $\Lambda \subset M$ un subconjunto compacto e invariante (i.e. $f^{-1}(\Lambda) = \Lambda$). El conjunto Λ se llama *uniformemente hiperbólico* (en breve, hiperbólico) si para todo $x \in \Lambda$ existe un splitting $S_x \oplus U_x = T_x M$ del espacio tangente $T_x M$ a M en x , que dependen continuamente de $x \in \Lambda$, y constantes $C > 0$ y $0 < \lambda < 1 < \sigma$, que verifican las desigualdades (3.3) y (3.4) para todo $x \in \Lambda$.

Nota: La condición de dependencia continua de S_x y U_x al variar $x \in \Lambda$ es redundante. Se puede demostrar esta continuidad a partir de la invariancia de esos subespacios, de las desigualdades de hiperbolicidad uniforme y de la compacidad del conjunto Λ (ver Ejercicios 3.3.4 y 3.3.5 y generalizarlos sustituyendo M por Λ).

Sobre las dimensiones de S_x y U_x : Son complementarias (su suma es $\dim M$), pero no necesariamente son ambas mayores que cero. Es decir, alguno de los dos subespacios puede tener dimensión cero, y el otro coincidir con $T_x M$. Las dimensiones pueden depender de $x \in \Lambda$. La dependencia continua de los subespacios invariantes implica que las dimensiones de S_x y U_x sean localmente constantes. Como Λ es compacto, si existen varios subconjuntos de Λ para los cuales S_x y U_x tienen dimensiones diferentes, entonces son dos a dos aislados.

De las definiciones anteriores, es inmediato deducir que $f : M \mapsto M$ es un difeomorfismo de Anosov, si y solo si la variedad M es un conjunto uniformemente hiperbólico (con dimensiones estable e inestable no nulas).

Ejercicio 3.4.2. Sea $f : M \mapsto M$. Probar que un punto periódico x de período p (i.e. $f^p(x) = x$, $f^j(x) \neq x \forall 1 \leq j < p$) es hiperbólico (i.e. los valores propios de df^p tienen módulo diferente de 1), si y solo si su órbita (finita) $\{x, f(x), \dots, f^{p-1}(x)\}$ es un conjunto uniformemente hiperbólico. Probar que si x es punto silla (es decir, $df^p(x)$ tiene valores propios con módulo menor que uno y también con módulo mayor que uno), entonces los subespacios propios de $df^p(x)$ son los subespacios estable S_x e inestable U_x , respectivamente. Probar que si x es un pozo hiperbólico (i.e. $df^p(x)$ tiene todos sus valores propios con módulo menor que uno), entonces $S_x = T_x M$ y $U_x = \mathbf{0}$. Enunciar y probar resultado dual si x es una fuente hiperbólica (i.e. $df^p(x)$ tiene todos sus valores propios con módulo mayor que uno).

3.5. Ejemplo: Herradura de Smale lineal

Un ejemplo paradigmático de conjunto hiperbólico (transitivo) que no es toda la variedad, es el siguiente, debido a Smale [Sm 1967] (ver también, por ejemplo, [Jo 2005, pages 97-98]):

Definición 3.5.1. Herradura de Smale lineal.

Se llama *herradura de Smale lineal* (en dimensión 2 y con 2 patas) a un difeomorfismo $T : Q = [0, 1]^2 \subset \mathbb{R}^2 \mapsto T(Q) \subset \mathbb{R}^2$ que satisface las siguientes condiciones (ver Figura 3.4):

- a) $T(Q) \cap Q = Q_0 \cup Q_1$ donde $Q = [0, 1]^2$,
 $Q_0 = [1/5, 2/5] \times [0, 1]$, $Q_1 = [3/5, 4/5] \times [0, 1]$.

- b) $T^{-1}(Q_0) = [0, 1] \times [1/5, 2/5]$, $T^{-1}(Q_1) = [0, 1] \times [3/5, 4/5]$.
- c) Para $j = 0, 1$, la restricción $T|_{T^{-1}(Q_j)}(x, y)$ es una transformación afín en (x, y) con direcciones propias $(1, 0)$ y $(0, 1)$ y valores propios λ y μ reales tales que $|\lambda| = 1/5$ y $|\mu| = 5$ respectivamente. Por ejemplo:

$$T|_{T^{-1}(Q_0)}(x, y) = ((1/5)(x + 1), 5y - 1),$$

$$T|_{T^{-1}(Q_1)}(x, y) = ((-1/5)(x - 4), -5y + 4)$$

Para comprender cómo es la transformación T , ver la figura 3.4 e imaginar T como la composición de dos transformaciones: Primero considerar una transformación afín que lleva el cuadrado $Q = [0, 1]^2$ a un rectángulo 5 veces más alto y 5 veces menos ancho que el cuadrado Q (contrae en horizontal y dilata en vertical). Después aplicar una transformación que "dobla" en forma biyectiva al rectángulo alto y flaco, dándole forma de herradura, superponiendo esta sobre el cuadrado Q en Q_0 y Q_1 (sin deformar Q_0 ni Q_1).

Se observa que hemos restringido la definición eligiendo valores numéricos fijos para $|\lambda|$ y $|\mu|$, iguales a $1/5$ y 5 respectivamente. Sin embargo, si se toman otros valores numéricos, $0 < |\lambda| < 1/2$ y $2 < |\mu|$, y se definen (coherentemente con esos nuevos valores numéricos) los rectángulos compactos Q_0 y Q_1 disjuntos, como en la Figura 3.4, y tales $T(Q) \cap Q = Q_0 \cup Q_1$; y si las transformaciones $T|_{T^{-1}(Q_0)}$ y $T|_{T^{-1}(Q_1)}$ son afines con valores propios $\pm\lambda$ y $\pm\mu$, entonces T también se llama *herradura de Smale lineal*.

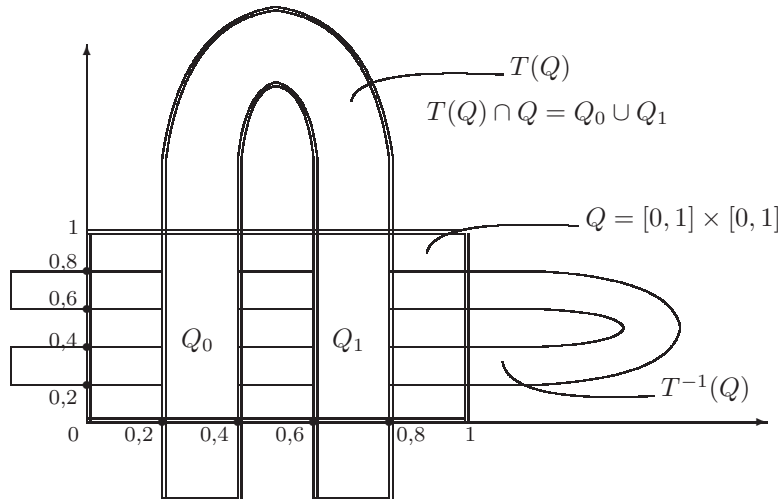


Figura 3.4: Herradura de Smale

Ejercicio 3.5.2. Sea $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ una herradura de Smale lineal. Sea $Q = [0, 1]^2$.
 (a) Dibujar esquemáticamente $T(T(Q) \cap Q) \subset T(Q)$, $T^2(Q)$ y $\cap_{n=0}^N T^n(Q)$ para $N = 2, 3$.

- (b) Dibujar esquemáticamente $T^{-1}(Q) \cap Q$, y $\cap_{n=0}^N T^{-n}Q$ para $N = 2, 3$.
- (c) Dibujar esquemáticamente el conjunto “estable” $W^s \cap Q$ de todos los puntos de Q cuyas órbitas futuras permanecen en Q para todos los iterados $n \geq 0$. (Sugerencia: Ver que $W^s \cap Q = \cap_{n=0}^{+\infty} T^{-n}(Q)$.)
- (d) Calcular el **exponente de Lyapunov negativo (tasa exponencial de contracción hacia el futuro)** de la herradura de Smale: $\lim_{n \rightarrow +\infty} (1/n) \log \|DT_x^n(1, 0)\|$ para todo punto $x \in \cap_{n \in \mathbb{N}} T^{-n}(Q)$.
- (d) Definir el conjunto “inestable” $W^u \cap Q$ y el exponente de Lyapunov positivo (tasa exponencial de dilatación hacia el futuro.) Calcularlo.

Definición 3.5.3. Sea $T : Q = [0, 1]^2 \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$ es la herradura de Smale lineal. Se llama *conjunto invariante maximal de T en Q* a

$$\Lambda := \bigcap_{n \in \mathbb{Z}} T^{-n}Q.$$

Obsérvese que por la propiedad de intersecciones finitas no vacías de compactos, el conjunto Λ es compacto no vacío.

Ejercicio 3.5.4. Probar que el maximal invariante Λ de una herradura de Smale lineal, es un conjunto hiperbólico. Probar que Λ es el producto cartesiano de dos conjuntos de Cantor en el intervalo $[0, 1]$.

3.6. Variedades invariantes de conjuntos hiperbólicos

A continuación enunciamos algunos teoremas sobre dinámica diferenciable, que generalizan propiedades de los difeomorfismos de Anosov en una variedad compacta M , y en particular algunos de los resultados vistos en el ejemplo del automorfismo lineal hiperbólico $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro \mathbb{T}^2 .

Teorema 3.6.1. Variedades invariantes para Λ unif. hiperbólico Sea $f : M \mapsto M$ un difeomorfismo y $\Lambda \subset M$ un conjunto invariante, compacto y uniformemente hiperbólico. Para $x \in M$ denotemos U_x, S_x los subespacios inestable y estable respectivamente. Entonces, para cada $x \in \Lambda$ existen y son únicas:

A) Una subvariedad conexa $W^s(x)$, C^1 -inmersa en M (pero no necesariamente “embedded”, i.e. no encajada en M , ni necesariamente compacta), llamada *variedad estable por x* , tal que:

$$x \in W^s(x), \quad f(W^s(x)) = W^s(f(x)), \quad T_x(W^s(x)) = S_x. \quad (3.6)$$

B) Una subvariedad C^1 conexa $W^u(x)$, C^1 -inmersa en M (pero no necesariamente encajada en M ni compacta), llamada *variedad inestable por x* , tal que:

$$x \in W^u(x), \quad f(W^u(x)) = W^u(f(x)), \quad T_x(W^u(x)) = U_x. \quad (3.7)$$

tales que:

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(y), f^n(x)) = 0 \Leftrightarrow y \in W^s(x), \quad (3.8)$$

$$\lim_{n \rightarrow -\infty} \text{dist}(f^n(y), f^n(x)) = 0 \Leftrightarrow y \in W^u(x). \quad (3.9)$$

Si además f es un difeomorfismo C^r para algún $r > 1$ entonces W_x^s y W_x^u son subvariedades de clase C^r .

Una prueba del teorema de existencia de variedades estable e inestable puede encontrarse en [Hi-Pu-Sh 1977].

Observación 3.6.2. Sea $f : M \mapsto M$ un difeomorfismo de Anosov con coeficiente $\lambda < 1$ de contracción en el fibrado estable S , y coeficiente $\sigma > 1$ de dilatación en el fibrado inestable U . Admitiendo la existencia de C^1 -subvariedades $W^s(x)$ y $W^u(x)$ conexas, que cumplen las condiciones (3.6) y (3.7) respectivamente, y sabiendo que los subespacios S_x y U_x dependen continuamente de x , se puede demostrar que

$$\limsup_{n \rightarrow +\infty} \frac{\log \text{dist}(f^n(y), f^n(x))}{n} \leq \log \lambda < 0 \quad \forall y \in W^s(x); \quad (3.10)$$

$$\liminf_{n \rightarrow -\infty} \frac{\log \text{dist}(f^n(y), f^n(x))}{n} \geq \log \sigma > 0 \quad \forall y \in W^u(x). \quad (3.11)$$

Definición 3.6.3. Sistemas dinámicos conjugados Sean X e Y espacios topológicos y sean $f : X \mapsto X$ y $g : Y \mapsto Y$ mapas Borel-medibles. Se dice que f y g son *conjugados o topológicamente equivalentes* si existe un homeomorfismo $h : X \mapsto Y$ llamado *conjugación* tal que $g \circ h = h \circ f$. La conjugación implica que cada una de las propiedades de la dinámica topológica de f (transitividad, conjunto no errante, omega y alfa límites, recurrencia) se satisface para f si y solo si se satisface para g .

Observación 3.6.4. Volvamos a los difeomorfismos de Anosov, como caso particular de sistema uniformemente hiperbólico:

El Teorema o Lema de Franks [Fr 1969] (ver también [Ma 1974] para difeomorfismos de Anosov en el toro n -dimensional, para todo $n \geq 2$) establece que:

Teorema o Lema de Franks *Las únicas superficies compactas, conexas, sin borde y orientables que soportan un difeomorfismo de Anosov f son homeomorfas al toro \mathbb{T}^2 , y f es conjugado a un automorfismo lineal hiperbólico.*

La demostración del Lema de Franks se encuentra en [Fr 1969]. En [Ma 1974], Manning generaliza la última parte del Lema de Franks, probando que los difeomorfismos de Anosov en el toro de dimensión n son conjugados a automorfismos lineales. Recientemente, en [Ha 2013] se obtiene el mismo resultado que Manning, pero para difeomorfismos de Anosov en nilmanifolds de dimensión 3, y no solo para el toro.

Debido al teorema de Franks los automorfismos lineales en el toro \mathbb{T}^2 son el paradigma de los difeomorfismos de Anosov en superficies, y entre ellos en particular el $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro \mathbb{T}^2 .

En cambio, conjuntos hiperbólicos como la herradura de Smale, o topológicamente conjugadas a ella, se pueden construir en un abierto homeomorfo a un cuadrado de \mathbb{R}^2 , en cualquier superficie (variedad de dimensión dos).

Se puede generalizar la herradura de Smale a dimensiones mayores que dos, tomando un cubo n -dimensional en el rol de cuadrado Q , y eligiendo dimensiones complementarias de contracción uniforme y de dilatación uniforme.

3.7. Expansividad o caos topológico.

Se han dado varias definiciones precisas de *caos* en la literatura matemática. Pero estas definiciones no son equivalentes entre sí (ver por ejemplo [Bu 2011]). Entonces cuando uno se refiere a *sistema caótico* debería siempre incluir la definición que está utilizando, con precisión, y observar que los resultados que se obtengan con esa definición, no son necesariamente ciertos si se hubiese adoptado otra. Según sea el objetivo de investigación de quien estudia el sistema dinámico (por ejemplo, su objetivo puede ser estudiar la dinámica topológica de un conjunto abierto y denso de órbitas, o solo estudiar la de μ -casi toda órbita cuando μ es invariante, o la de Lebesgue-casi toda órbita, cuando la medida de Lebesgue no es invariante, etc), adopta una u otra definición.

El estudio de las propiedades topológicas de los difeomorfismos lineales en el toro \mathbb{T}^2 es paradigmático para estudiar más en general, los sistemas caóticos desde el punto de vista topológico. En efecto:

Definición 3.7.1. Sea $T : X \mapsto X$ un homeomorfismo en un espacio métrico compacto X . Se dice que T es *caótico* (topológicamente) o *expansivo*, si existe una constante $\alpha > 0$ (llamada *constante de expansividad*) tal que

$$\sup_{n \in \mathbb{Z}} \text{dist}(T^n(x), T^n(z)) > \alpha \quad \forall x \neq y \in X$$

Interpretación: El caos topológico o expansividad es una versión de “hiperbolicidad topológica.” Significa que la *dinámica es α -sensible a las condiciones iniciales*: dos órbitas con estados iniciales $x \neq y$ diferentes se separan más que la distancia α , hacia el futuro ó hacia el pasado. Podrá haber parejas de órbitas, con estados iniciales próximos $x \neq y$, que se separan solo para el futuro pero no para el pasado, o al revés. Generalmente (por ejemplo en un difeomorfismo de Anosov), para la mayor parte de las parejas de estados iniciales $x \neq y$, sus órbitas se separan en el futuro y en el pasado.

La separación hacia el futuro o hacia el pasado, más que una constante uniforme α , de dos órbitas con estados iniciales $x \neq y$ (por más que x e y estén tan cerca entre sí como se desee) implica lo siguiente:

Si uno aproxima con error $\epsilon > 0$, aunque sea arbitrariamente pequeño, el estado inicial x sustituyéndolo por y tal que $0 < \text{dist}(x, y) \leq \epsilon$, entonces el estado del sistema $T^n(y)$ en algún instante $n \in \mathbb{Z}$, se modificará más que $\alpha > 0$ respecto del estado $T^n(x)$ que tendría si no se hubiese cometido el error en el estado inicial. La sensibilidad a las condiciones iniciales, o expansividad, o caos topológico, se llama

también “efecto mariposa”: la leve modificación del estado inicial producido por el aleteo de una mariposa, por más leve que esta sea (es decir por más pequeño que sea la diferencia $\epsilon > 0$ provocada en ese estado inicial) produce una modificación “drástica” (es decir mayor que una constante uniforme $\alpha > 0$) en el estado del sistema en otro instante).

Ejercicio 3.7.2. Demostrar que la rotación del círculo (racional o irracional) no es expansiva (i.e. no es topológicamente caótica).

Ejercicio 3.7.3. Demostrar que el tent map $f : [0, 1] \mapsto [0, 1]$ es expansivo para el futuro, esto es, existe una constante $\alpha > 0$ tal que

$$\text{Si } \text{dist}(f^j(x), f^j(y)) < \alpha \quad \forall j \in \mathbb{N} \quad \text{entonces } x = y.$$

Sugerencia: Probar que la derivada $|(f^n)'|$ tiende uniformemente a $+\infty$ con n y usar $\alpha = 1/4$.

Ejercicio 3.7.4. Demostrar que la transformación $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro \mathbb{T}^2 es expansiva.

Observación 3.7.5. En [Lew 1989], Lewowicz demostró los siguientes resultados:

Teorema de Lewowicz *Las únicas superficies (variedades de dimensión dos) compactas y conexas donde existen homeomorfismos expansivos, son homeomorfas al toro \mathbb{T}^2 .*

Todos los homeomorfismos expansivos en el toro \mathbb{T}^2 son conjugados a un Anosov (y por el teorema de Franks son conjugados a un difeomorfismo de Anosov lineal).

Por eso resulta paradigmático estudiar los difeomorfismos lineales, y el $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en particular.

3.8. Foliaciones invariantes para dif. de Anosov.

La unicidad de las variedades invariantes para los difeomorfismos de Anosov, implica que la siguiente colección no numerable de subconjuntos (subvariedades) $\{W_x^s\}_{x \in M}$, sea una partición de M . En efecto:

Proposición 3.8.1. *Sea $f : M \mapsto M$ un difeomorfismo de Anosov, y sean $x, y \in M$. Entonces o bien $W_x^s \cap W_y^s = \emptyset$, o bien $W_x^s = W_y^s$. Análogamente para W_x^u y W_y^u .*

Demostración. Debido a (3.8) y (3.9), $y \in W_x^s$ si y solo si $x \in W_y^s$. Supongamos que existe $z \in W_x^s \cap W_y^s$. Sea $x' \in W_x^s$. Usando nuevamente (3.8) y (3.9) y la propiedad triangular, deducimos que

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(y), f^n(x')) = 0.$$

En efecto:

$$\begin{aligned} \text{dist}(f^n(y), f^n(x')) &\leq \\ \text{dist}(f^n(y), f^n(z)) + \text{dist}(f^n(z), f^n(x)) + \text{dist}(f^n(x), f^n(x')), \end{aligned}$$

y estos tres sumandos tienden a cero cuando $n \rightarrow +\infty$. Luego $x' \in W_y^s$ para todo $x' \in W_x^s$, probando que $W_x^s \subset W_y^s$. Simétricamente, $W_y^s \subset W_x^s$. Hemos demostrado que si existe $z \in W_x^s \cap W_y^s$ entonces $W_y^s = W_x^s$. \square

Definición 3.8.2. Foliaciones estable e inestable.

Se llama *foliación estable* a la partición $\{W_x^s\}_{x \in M}$ de M en variedades estables. Análogamente, se llama *foliación inestable* a la partición $\{W_x^u\}_{x \in M}$ de M en variedades inestables.

Nota sobre la definición geométrica de “foliación”: Por definición, una foliación tiene una estructura geométrica precisa que trasciende a la mera partición del espacio M en subvariedades inmersas disjuntas dos a dos todas de la misma dimensión $1 \leq k_1 < \dim M$. Estas variedades inmersas se llaman hojas de la foliación. La estructura geométrica de foliación consiste en la existencia de un atlas de cartas locales ξ de M que son “trivializadoras” de las hojas de la foliación: i.e. la imagen de cada ξ es el producto cartesiano $B_{k_1} \times B_{k_2}$, donde B_{k_i} es la bola unitaria de R^{k_i} ($k_1 + k_2 = \dim M$), tal que transforma las componentes conexas locales de cada hoja, en las secciones $B_{k_1} \times \{v_2\}$ (con $v_2 \in B_{k_2}$ fijo).

En general, para las foliaciones dinámicamente definidas (por ejemplo, las foliaciones estable e inestable de un f de Anosov), las cartas locales trivializadoras ξ existen, pero son solo homeomorfismos sobre sus imágenes (no son siquiera C^1). Como las hojas de la foliación son subvariedades inmersas de clase C^1 , entonces las *restricciones* de las cartas trivializadoras a las hojas locales, son C^1 .

Definición 3.8.3. Regularidad de foliaciones invariantes.

Un partición $\mathcal{W} := \{W_x\}_{x \in M}$ del espacio M en subvariedades diferenciables inmersas en M y disjuntas dos a dos, todas de la misma dimensión, se dice que es *una foliación invariante* C^0 (desde el punto de vista dinámico) si:

- (1) $f(W_x) = W_{f(x)} \quad \forall x \in M$ (invariancia),
- (2) cada hoja W_x es de clase C^1 , y
- (3) la aplicación $x \in M \mapsto T_x W_x$ que lleva cada punto x en el subespacio tangente $T_x W_x$ en x a la hoja W_x , es continua (es decir, el subespacio tangente a la hoja varía en forma C^0 con $x \in M$).

Nota: La condición (2) de que cada hoja sea de clase C^1 , implica que su subespacio tangente varíe en forma C^0 con x *variando a lo largo de la hoja respectiva*. Pero no implica que ese subespacio tangente *varíe además en forma C^0 con x en todas las direcciones transversales a la hoja*. Por lo tanto la condición (3) no es redundante: es más fuerte que la condición (2).

Análogamente, para todo $r \in \mathbb{N}$, la partición \mathcal{W} es *una foliación invariante C^r* (desde el punto de vista dinámico) si cada hoja $W_x \in \mathcal{W}$ es de clase C^{r+1} y la aplicación $x \in M \mapsto T_x W_x$ es de clase C^r . Para que esta definición sea aplicable, hay que asumir que la variedad ambiente M es de clase C^{r+1} por lo menos.

Sea $0 < \alpha < 1$. Una foliación invariante \mathcal{W} es α -Hölder continua y se denota de clase C^α , si cada hoja $W_x \in \mathcal{W}$ es de clase C^1 y la aplicación $x \in M \mapsto T_x W_x$ satisface:

$$\text{dist}(W_x, W_y) \leq K [\text{dist}(x, y)]^\alpha \quad (3.12)$$

para cierta constante K . Se dice que la foliación es Lipschitz y se denota como C^{Lip} , si es 1-Hölder continua (es decir vale la desigualdad (3.12) para $\alpha = 1$).

Debido al Teorema 3.6.1, tenemos el siguiente resultado:

Corolario 3.8.4. *Si f es difeomorfismo de Anosov de clase C^1 , entonces las foliaciones estable e inestable son de clase C^0 .*

Demostración. $T_x W_x^s = S_x$, $T_x W_x^u = U_x$ y los subespacios S_x y U_x dependen continuamente de x . \square

Uno podría esperar que si f es difeomorfismo de Anosov de clase C^{r+1} con $r \geq 0$, entonces las foliaciones invariantes estable e inestables sean foliaciones de clase C^r . Este último resultado es FALSO. En general, cada hoja es C^{r+1} . Pero no se puede mejorar casi nada la regularidad C^0 de las foliaciones invariantes simplemente aumentando la regularidad de f .

Teorema 3.8.5. Hölder continuidad de foliaciones invariantes para difeomorfismos de Anosov.

Si $f : M \mapsto M$ es un difeomorfismo de Anosov de clase C^k con $k > 2$, entonces las variedades estables e inestables de f son subvariedades C^k , y las foliaciones estables e inestables que forman son α -Hölder continuas para cierto $0 < \alpha < 1$.

Una prueba de este teorema se puede encontrar en [Ba-Pe 2001, Theorem 2.3.1, pag. 48].

3.9. Exponentes de Lyapunov

A lo largo de las próximas secciones asumiremos las siguientes hipótesis:

- M es una variedad diferenciable, compacta, conexa y provista de una estructura riemanniana.
- $f : M \mapsto M$ es de clase C^1 (es decir f es diferenciable y su derivada df es continua). El mapa f no es necesariamente invertible.

Notación 3.9.1.

- $f : M \mapsto M$ es un difeomorfismo si de clase C^1 , invertible y además su inversa también es de clase C^1 . Denotamos $f \in \text{Diff}^1(M)$.
- Si además f y su inversa son de clase C^r para algún $r \geq 2$ (para lo cual la variedad M también debe ser de clase C^r por lo menos), indicaremos $f \in \text{Diff}^r(M)$.
- Denotaremos $f \in \text{Diff}^{1+\alpha}$ cuando $f \in \text{Diff}^1(M)$ y además $df : TM \mapsto TM$ es α -Hölder continua para cierta constante $0 < \alpha < 1$; i.e. existen constantes $\delta > 0$ y $K > 0$ tales que

$$\|df_x - df_y\| \leq K(\text{dist}(x, y))^\alpha \quad \forall x, y \in M \text{ such that } \text{dist}(x, y) < \delta.$$

En el primer miembro de la desigualdad anterior, $\|A\|$ denota la norma de la transformación lineal $A \in L(\mathbb{R}^k)$ donde $k = \dim(M)$, es decir $\|A\| = \max\{\|Av\| : v \in \mathbb{R}^k, \|v\| = 1\}$.

• Denotaremos $f \in \text{Diff}^{1+Lip}(M)$ si $f \in \text{Diff}^1(M)$ y además df es Lipschitz, i.e. existen constantes $\delta > 0$ y $K > 0$ que satisfacen la desigualdad de α -Hölder continuidad con $\alpha = 1$.

Definición 3.9.2. Exponentes de Lyapunov

Sea $f \in \text{Diff}^1(M)$. Un punto $x \in M$ se llama *débilmente regular* si existen los siguiente límites para todo $v \in T_x M \setminus \{0\}$:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^n(v)\|)}{n}; \quad \lim_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}(v)\|)}{-n}.$$

Estos dos límites (números reales), se llaman *exponentes de Lyapunov en el futuro y en el pasado respectivamente, de la órbita por x correspondientes a la dirección $[v]$* .

No se definen los exponentes de Lyapunov en los puntos no regulares.

Se puede definir también puntos débilmente regulares y exponentes de Lyapunov (en el futuro) para $f \in C^1(M)$ aunque f no sea un difeomorfismo.

Más adelante veremos el Teorema de Oseledets, que demuestra, entre otros resultados, lo siguiente:

Para toda medida de probabilidad μ que sea f -invariante, μ -casi todo punto $x \in M$ es regular.

Dicho de otra forma, el conjunto de los puntos no regulares tiene medida nula para toda medida de probabilidad μ invariante por f .

Nota sobre el concepto de regularidad: En la literatura, suele llamarse punto regular a aquellos puntos x que cumplen una condición más fuerte que la que hemos adoptado nosotros en la Definición 3.9.2. En efecto, se definen condiciones adicionales de existencia de subespacios invariantes tales que para todo vector en ellos el exponente de Lyapunov en el futuro coincide con el exponente de Lyapunov en el pasado. A los puntos regulares que cumplen esa condición más fuerte, los llamaremos Lyapunov-regulares (ver Definición 3.9.6). El teorema de Oseledets establece que μ -casi todo punto $x \in M$ no solo es débilmente regular según nuestra definición 3.3.7, sino también Lyapunov-regular según la Definición 3.9.6.

Los exponentes de Lyapunov dependen de la órbita de x , pero no dependen de cuál punto se tome en la misma órbita. En efecto:

Ejercicio 3.9.3. Probar que si x es débilmente regular, entonces para todo $k \in \mathbb{Z}$ fijo, el punto $y = f^k(x)$ (es decir, cualquier punto en la órbita de x) es también débilmente regular y que el conjunto de los exponentes de Lyapunov en y coincide con el de los de x .

El siguiente ejercicio tiene como objetivo adelantarse al enunciado del teorema de Oseledets (que enunciaremos al final de esta sección).

Ejercicio 3.9.4. Sea x un punto débilmente regular.

(a) Sean $0 \neq v \in T_x M$, χ el exponente de Lyapunov en el futuro (o en el pasado) de la órbita de x correspondiente a una dirección $u \neq 0$. Sea $[u] \subset T_x M$ el subespacio de dimensión 1 generado por u . Demostrar que para todo $0 \neq u' = [u]$ el exponente de Lyapunov en el futuro (o en el pasado respectivamente) correspondiente a u es el mismo que el de u' .

(b) Sean χ_u^+ y χ_v^+ los exponentes de Lyapunov en el futuro de dos direcciones l.i. $0 \neq u, v \in T_x M$. Sean χ_u^- y χ_v^- los exponentes de Lyapunov en el pasado de esas dos mismas direcciones u y v . Asuma $\chi_u^+ \neq \chi_v^+$, $\chi_u^- \neq \chi_v^-$. Sea $w = u + v$. Probar que $\chi_w^+ = \max\{\chi_u^+, \chi_v^+\}$ y $\chi_w^- = \min\{\chi_u^-, \chi_v^-\}$.

Sugerencia: Asuma $\chi_u^+ > \chi_v^+$. Use el primer límite de 3.9.2 con los vectores u y v , para probar que para todo $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que:

$$\begin{aligned} \|df^n(w)\| &\geq \|df^n(u)\| - \|df^n(v)\| \geq \\ e^{n(\chi_u^+ - \epsilon)}(\|u\| - e^{n(\chi_v^+ - \chi_u^+ + 2\epsilon)}\|v\|) &\quad \forall n \geq N. \end{aligned}$$

Fije $\epsilon < (\chi_u^+ - \chi_v^+)/2$. Tome logaritmo, divida entre n y luego $n \rightarrow +\infty$.

(c) Como en la parte (b) asuma ahora $\chi_u^+ = \chi_v^+$ y $\chi_u^- = \chi_v^-$. Pruebe que $\chi_w^+ \leq \chi_u^+ = \chi_v^+$, $\chi_w^- \geq \chi_u^- = \chi_v^-$.

(d) Sea una base u_1, \dots, u_k de $T_x M$. Asuma que los exponentes de Lyapunov en el futuro χ_i^+ y en el pasado χ_i^- en las direcciones u_i , cumplen $\chi_i^+ \neq \chi_j^+$ para todo $i \neq j$.

Probar que para todo $0 \neq u = \sum_{i=1}^k b_i u_i \in T_x M$, se cumple:

$$\chi_u^+ = \max_{1 \leq i \leq k; b_i \neq 0} \chi_i^+.$$

$$\chi_u^- = \min_{1 \leq i \leq k; b_i \neq 0} \chi_i^-.$$

Sugerencia: usar inducción en la cantidad de coeficientes $b_i \neq 0$, y la parte (b).

Ejercicio 3.9.5. Sea x un punto débilmente regular.

Para cada $k \geq 0$ sean $E_{f^k(x)}^1$ y $E_{f^k(x)}^2$ dos subespacios L.I. (de dimensiones no nulas) de $T_{f^k(x)} M$, invariantes por f (es decir $E_{f^{k+1}(x)}^i = df_{f^k(x)} E_{f^k(x)}^i$ para todo $k \geq 0$ y para $i = 1, 2$). Asuma la siguiente hipótesis:

Hipótesis I Para cada i coinciden los exponentes de Lyapunov en el futuro y en el pasado entre sí y en todas las direcciones del subespacio E_x^i . Llámelo χ^i . Además para todo $0 \neq v \in E_x^1 \oplus E_x^2$, el exponente de Lyapunov en el futuro de v es mayor o igual que el exponente de Lyapunov en el pasado de v .

(a) Probar que

(i) Para todo $k \geq 0$ y para cada $i = 1, 2$, los exponentes de Lyapunov en el futuro y en el pasado en $f^k(x)$, correspondientes a la dirección $E_{f^k(x)}^i$, coinciden entre sí y con χ^i .

(ii) Probar que para todo $0 \neq v \in E_{f^k(x)}^1 \oplus E_{f^k(x)}^2$ tal que $v \notin E_{f^k(x)}^1, E_{f^k(x)}^2$, el exponente de Lyapunov $\chi_v^+(x)$ de $[v]$ en el futuro es igual a $\max\{\chi^1, \chi^2\}$, y el exponente de Lyapunov $\chi_v^-(x)$ de $[v]$ en el pasado, es igual a $\min\{\chi^1, \chi^2\}$.

(iii) Deducir que si $\chi_1 = \chi_2 = \chi$, entonces el subespacio $E_{f^k(x)} = E_{f^k(x)}^1 \oplus E_{f^k(x)}^2$ es invariante, de dimensión mayor que los subespacios $E_{f^k(x)}^1$ y $E_{f^k(x)}^2$ que lo generaron, y los exponentes de Lyapunov en el futuro y en el pasado para todos los vectores $\mathbf{0} \neq v \in E$ coinciden con χ .

(b) Extender (enunciar y demostrar) los resultados de la parte a), asumiendo que existe un splitting $T_x M = E_x^1 \oplus E_x^2 \oplus \dots \oplus E_x^h$ que verifica la Hipótesis I.

(c) Deducir, asumiendo (b), que el conjunto de exponentes de Lyapunov diferentes en cualquier punto regular que cumpla la Hipótesis I es finito, y menor o igual que $\dim(M)$.

El ejercicio anterior motiva preguntarse cuándo se cumple la hipótesis asumida sobre la existencia de los subespacios E_x^i invariantes para los cuales los exponentes de Lyapunov en el futuro y en el pasado existen y son iguales entre sí, y diferentes para diferentes valores de i . Esta pregunta motiva la siguiente definición de regularidad:

Definición 3.9.6. Puntos Lyapunov-regulares

Un punto $x \in M$ se llama *Lyapunov regular* si existe un splitting del espacio tangente

$$T_x M = E_x^1 \oplus E_x^2 \oplus E_x^{k(x)}$$

(que puede reducirse como caso particular a $k(x) = 1$, $T_x M = E_x^1$), tal que existen y coinciden entre sí *los exponentes de Lyapunov en el futuro y en el pasado* $\chi_i(x)$ en toda dirección $[v] \subset E_x^i$ y además

$$\chi_1(x) < \chi_2(x) < \dots < \chi_h(x)(x).$$

En otras palabras:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^n(v)\|)}{n} = \lim_{n \rightarrow -\infty} \frac{\log(\|df_x^n(v)\|)}{n} = \chi_i(x) \quad \forall \{\mathbf{0}\} \neq [v] \subset E_x^i.$$

Los subespacios E_x^i en un punto Lyapunov regular, se llaman *subespacios de Oseledets*.

Nota: El Ejercicio 3.9.5 muestra que todo punto Lyapunov-regular es regular en el sentido de la Definición 3.9.2. El recíproco es falso (ver el ejemplo del Ejercicio 1.3.13).

En el enunciado del siguiente ejercicio se establece que la regularidad Lyapunov, el splitting y los exponentes de Lyapunov dependen de la órbita y no de qué punto x en cada órbita se elija.

Ejercicio 3.9.7. Sea x un punto Lyapunov-regular.

(a) Demostrar que el splitting $T_x M = \bigoplus_{i=1}^{h(x)} E^i(x)$ y el conjunto de exponentes de Lyapunov, son únicos.

(b) Demostrar que para todo $k \in \mathbb{Z}$ el punto $f^k(x)$ también es Lyapunov regular, y además $E_{f^k(x)}^i = df^k E^i(x) \forall 1 \leq i \leq h(f^k(x)) = h(x)$, y $\chi_i(x) = \chi_i(f^k(x))$.

Pregunta: ¿Existen abundantes puntos Lyapunov regulares?

El siguiente resultado es un teorema fundamental en la Teoría Ergódica Diferenciable y responde afirmativamente a la pregunta anterior, por lo menos desde el punto de vista medible, y con respecto a las medidas invariantes:

Teorema 3.9.8. de Oseledets

Sea M una variedad compacta riemanniana de dimensión finita. Sea $f \in \text{Diff}^1(M)$. Entonces

- (a) El conjunto R de puntos Lyapunov regulares para f es medible.
- (b) Las funciones que a cada punto $x \in R$ asignan los exponentes de Lyapunov son medibles.
- (c) El conjunto R tiene probabilidad total (para toda medida de probabilidad f -invariante μ , se cumple $\mu(R) = 1$).

Demostraremos el teorema de Oseledets en la Sección 4.6. La demostración de V.I. Oseledets del Teorema 3.9.8 se encuentra en [Os 1966] y en [Os 1968]. Otras demostraciones del teorema de Oseledets pueden encontrarse por ejemplo en [Ba-Pe 2001], en [Mañ 1983a, Cap. IV §10] (ver también [Mañ 1987], en [Via 2012a]) y en [Ra 1979]. Generalizaciones del Teorema de Oseledets, llamados Teoremas Ergódicos Multiplicativos, enuncian resultados en los cuales df_x es sustituido por una aplicación lineal que depende de x en un espacio vectorial finito dimensional, o incluso por un cociclo en ciertos espacios de Banach infinito dimensionales. Por ejemplo, en [Kar-Mar 1999], se prueba un Teorema Ergódico Multiplicativo que generaliza al Teorema de Oseledets a ciertos cociclos en espacios de Banach uniformemente convexos.

3.10. Hiperbolicidad no uniforme

Interpretación de los exponentes de Lyapunov no nulos:

Veremos por qué los exponentes de Lyapunov, cuando no son nulos, se interpretan como *la tasa exponencial asintótica de crecimiento (dilatación) o decrecimiento (contracción)* en el futuro de la norma de los vectores en el subespacio tangente, por iteración del diferencial, es decir al aplicar df^n . En efecto, supongamos que un exponente de Lyapunov en el futuro y en el pasado, para la dirección $[s] \subset T_x M$ de la órbita por x , es $\chi(x) < 0$. Entonces, de la definición del límite en 3.9.2, para todo $\epsilon > 0$ existe $N \geq 1$:

$$\|df_x^n(s)\| \leq e^{n(\chi+\epsilon)} \|s\| \quad \forall n \geq N$$

Luego, para ciertos números reales $C(x) > 0$, $\lambda(x) = e^{\chi(x)+\epsilon} < 1$ (si se fija $0 < \epsilon < -\chi(x)$), se cumple

$$\|df^n(s)\| \leq C(x)[\lambda(x)]^n \|s\| \quad \forall n \geq 0, \quad \text{donde } 0 < \lambda(x) < 1, \quad (3.13)$$

(Demostramos con detalle la existencia de $C(x) > 0$ y la desigualdad anterior en el Lema 3.10.3.) La desigualdad anterior significa que, en la dirección $[s]$, el diferencial n -ésimo contrae exponencialmente las normas de los vectores hacia el futuro con coeficiente $0 < \lambda = e^{\chi+\epsilon} < 1$. Este coeficiente tiende a e^χ cuando $\epsilon \rightarrow 0^+$ (y por lo tanto

cuando $n \rightarrow +\infty$). Entonces un exponente de Lyapunov negativo χ es asintóticamente igual a $\log \lambda$. Decimos así que *un exponente de Lyapunov negativo $\chi(x)$ es la tasa exponencial asintótica de contracción en el futuro (o de dilatación hacia el pasado) por la derivada de f^n en la dirección $[s]$.*

Análogamente, si para la misma órbita de x existe alguna dirección $[u] \subset T_x M$, para la cual el exponente de Lyapunov en el pasado y en el futuro es $\chi(x) > 0$, entonces existen un número real $C(x) > 0$ y un valor real $\sigma(x) = e^{\chi(x)-\epsilon} > 1$ (si se fija $0 < \epsilon < \chi(x)$) tales que:

$$\|df^{-n}(u)\| \leq C(x)[\sigma(x)]^{-n}\|u\| \quad \forall n \geq 0, \quad \text{donde } \sigma(x) > 1, \quad (3.14)$$

(Demostraremos la existencia de la constante $C(x) > 0$ y la desigualdad anterior en el Lema 3.10.3.) Siendo σ asintóticamente igual a e^χ , decimos que *un exponente de Lyapunov positivo $\chi(x)$ es la tasa exponencial asintótica de constracción hacia el pasado (o de dilatación hacia el futuro) por la derivada de f^{-n} en la dirección $[u]$.*

Observemos las similitudes y diferencias entre las desigualdades (3.13) y (3.14) y las de la Definición 3.4.1 de hiperbolicidad uniforme (desigualdades (3.3) y (3.4)) Las similitudes justifican la siguiente definición:

Definición 3.10.1. Hiperbolicidad no uniforme Sea $\Lambda \subset M$ un conjunto medible f -invariante: $f^{-1}(\Lambda) = \Lambda$ (Λ no es necesariamente compacto).

Decimos que f es *no uniformemente hiperbólico en Λ* , (o que Λ es un conjunto no uniformemente hiperbólico para f), si para todo punto $x \in \Lambda$ existe un splitting $T_x M = S_x \oplus U_x$ que depende mediblemente de $x \in \Lambda$ y que es df -invariante, i.e.

- $df S_x = S_{f(x)}$, $df U_x = U_{f(x)} \quad \forall x \in \Lambda$,
- y existen números reales $C(x) > 0$ y $0 \leq \lambda(x) < 1 < \sigma(x)$, que dependen mediblemente de $x \in \Lambda$, tales que
- $\lambda(f(x)) = \lambda(x)$, $\sigma(f(x)) = \sigma(x)$,
- se verifican las dos desigualdades (3.13) y (3.14); es decir, para todo $n \geq 0$ y para todos $u \in U_x$ y $s \in S_x$:

$$\|df^n s\| \leq C(x)\lambda(x)^n \|s\|, \quad \|df^{-n} u\| \leq C(x)\sigma(x)^{-n} \|u\|$$

Observación 3.10.2. Hiperbolicidad No Uniforme. A diferencia de la Definición 3.4.1 de hiperbolicidad uniforme en compactos, la hiperbolicidad no uniforme no implica la continuidad del splitting $S_x \oplus U_x$ al variar $x \in \Lambda$. Pero sí exige, por definición, que el splitting sea medible.

Nota: Si $C(x)$, $\lambda(x)$ y $\sigma(x)$ son constantes independientes de $x \in \Lambda$, y además Λ es compacto, se dice que f es uniformemente hiperbólico en Λ , o que Λ es un conjunto uniformemente hiperbólico para f , de acuerdo con la definición 3.4.1.

Lema 3.10.3. *Sea Λ un conjunto f -invariante, medible tal que todo $x \in \Lambda$ es un punto Lyapunov regular cuyo splitting $E_x^1 \oplus E_x^2 \oplus \dots \oplus E_x^r = T_x M$ depende mediblemente de x ($r = r(x)$ también), y cuyos exponentes de Lyapunov respectivos $\chi_1(x) < \chi_2(x) < \dots < \chi_i(x) < \dots < \chi_r(x)$ son **todos diferentes de cero** y dependen mediblemente de x . Entonces Λ es un conjunto no uniformemente hiperbólico.*

En extenso, existe un splitting medible $T_x M = S_x \oplus U_x$ invariante por df y funciones medibles $C(x) > 0$ y $0 < \lambda(x) < 1 < \sigma(x)$, tales que $\lambda(f(x)) = \lambda(x)$ y $\sigma(f(x)) = \sigma(x)$ y tales que se verifican las desigualdades (3.13) y (3.14) para todo $n \geq 0$, para todo $u \in U_x$ y todo $s \in S_x$, y para todo $x \in \Lambda$.

Probaremos el Lema 3.10.3 más abajo. En relación a su enunciado, en [Al-Ar 2003] se demuestra que para ciertos conjuntos no uniformemente hiperbólicos, si μ -c.t.p. $x \in M$ cumple las hipótesis del Lema 3.10.3 para *toda* medida invariante μ , entonces el conjunto es localmente uniformemente hiperbólico.

Demostración. del Lema 3.10.3: Sean: $\alpha := \max\{\chi_i(x) : 1 \leq i \leq r, \chi_i(x) < 0\} < 0$,
 $\beta := \min\{\chi_i(x) : 1 \leq i \leq r, \chi_i(x) > 0\} > 0$.

Como α y β son máximo y mínimo de funciones medibles, son medibles. Fijemos un valor constante $\epsilon > 0$ suficientemente pequeño tal que $\alpha + \epsilon < 0$, $\beta - \epsilon > 0$. Denotamos

$$0 < \lambda = \lambda(x) := e^{\alpha + \epsilon} < 1, \quad \sigma = \sigma(x) := e^{\beta - \epsilon} > 1.$$

Como $\lambda(x)$ y $\sigma(x)$ son composición de funciones continuas con funciones medibles, son medibles. Por construcción, como los exponentes de Lyapunov son los mismos para x que para $f(x)$, tenemos $\lambda(x) = \lambda(f(x))$, $\sigma(x) = \sigma(f(x))$. Sean

$$S_x := \oplus\{E_x^i(x) : 1 \leq i \leq r, \chi_i(x) < 0\},$$

$$U_x := \oplus\{E_x^i(x) : 1 \leq i \leq r, \chi_i(x) > 0\},$$

donde $\oplus_{i=1}^r E_x^i(x) = T_x M$ es el splitting en los subespacios de Oselecs del espacio tangente en el punto x . Estos subespacios existen por la Definición 3.9.6 de punto Lyapunov regular; y por hipótesis dependen mediblemente de x . Las funciones $\chi_i(x)$ son medibles. Luego las preimágenes por $\chi_i : T\Lambda \mapsto \mathbb{R}$ de \mathbb{R}^+ y de \mathbb{R}^- son conjuntos medibles. Finalmente, la suma directa de un conjunto finito de subfibrados medibles, es un subfibrado medible. Concluimos que S_x y U_x son subfibrados medibles de $T\Lambda$. Por construcción, como los subespacios de Oseledets son invariantes por df , tenemos

$$df S_x = S_{f(x)}, \quad df U_x = U_{f(x)}.$$

Aplicando el resultado del Ejercicio 3.9.5, tenemos:

Para todo $s \in S_x$ el exponente de Lyapunov en el futuro (y también en el pasado) es menor o igual que $\alpha < 0$. Para todo $u \in U_x$ el exponente de Lyapunov en el pasado (y también en el futuro) es mayor o igual que $\beta > 0$. En extenso:

$$\chi_i^+(x, s) = \lim_{n \rightarrow +\infty} \frac{\log \|df^n(s)\|}{n} \leq \alpha < \alpha + \epsilon = \log \lambda < 0 \quad \forall 0 \leq s \in S_x, \quad (3.15)$$

$$\chi_i^-(x, u) = \lim_{m \rightarrow -\infty} \frac{\log \|df^m(u)\|}{m} \geq \beta > \beta - \epsilon = \log \sigma > 0 \quad \forall 0 \leq u \in U_x. \quad (3.16)$$

Probemos que para cada $0 \leq s \in S_x$ y para cada $0 \leq u \in U_x$, existen los siguientes números reales $H(x, s)$, $K(x, s) > 0$:

$$H(x, s) := \sup_{n \geq 0} \frac{\|df_x^n(s)\|}{\lambda^n \|s\|}, \quad K(x, u) := \sup_{n \geq 0} \frac{\|df_x^{-n}(u)\|}{\sigma^{-n} \|u\|}. \quad (3.17)$$

En efecto, fijemos x, u, s . En las igualdades (3.15) y (3.16), aplicamos la definición de límite, multiplicamos por n (con $|n|$ suficientemente grande), y aplicamos la exponencial. (Hay que cuidar que cuando n es negativo, al multiplicar por n se invierten el sentido de las desigualdades). Concluimos que existe $N = N(x, u, s) \geq 0$ tal que

$$\|df^n(s)\| / \lambda^n \|s\| < 1, \quad \|df^{-n}(u)\| / \sigma^{-n} \|u\| < 1,$$

para todo $n \geq N \geq 0$. Entonces, el supremo que define $H(x, s)$, así como el supremo que define $K(x, s)$, existe y es un número real no negativo (porque para $|n| \geq N$, todos los cocientes cuyo supremo buscamos son menores que 1; y para $0 \leq |n| \leq N$ los cocientes a maximizar son positivos y una cantidad finita). Afirmamos que existen los números reales $K(x), H(x) > 0$, definidos por:

$$H(x) := \sup \{H(x, s) : s \in S_x, \|s\| = 1\}, \quad (3.18)$$

$$K(x) := \sup \{K(x, u) : u \in U_x, \|u\| = 1\} \quad (3.19)$$

Probaremos que existe $K(x)$ real (la prueba de que existe $H(x)$ real es similar). Tomemos una base u_1, \dots, u_k de U_x , donde $k = \dim(U_x)$, formada por vectores u_i de norma 1, que se encuentran todos en los subespacios de Oseledets, según la definición 3.9.6 de punto Lyapunov regular. Entonces, dado $u \in U_x$ se puede escribir: $u = \sum_{i=1}^{k_2} b_i u_i$. Si $\|u\| = 1$, entonces existe $M(x)$ tal que $0 \leq |b_i| \leq M(x) \forall 1 \leq i \leq k$. En efecto, fijada la base, cada $|b_i|$ es una función real continua del vector $u \in U_x$ (pues es la norma de la proyección ortogonal del vector u sobre el subespacio generado por u_i). Por el teorema de Weierstrass, la función continua $|b_i|$ tiene un máximo M_i en el subconjunto compacto $\{u \in U_x : \|u\| = 1\} \subset T_x M$. Entonces basta tomar $M(x) := \max_{i=1}^k M_i$.

Tenemos:

$$\frac{\|df^{-n}(u)\|}{\sigma^{-n}} \leq \sum_{i=1}^k |b_i| \frac{\|df^n(u_i)\|}{\sigma^{-n}} \leq M(x) \sum_{i=1}^k K(x, u_i) =: K_1(x) \quad \forall n \geq 0. \quad (3.20)$$

Para la primera desigualdad usamos la propiedad triangular de la norma. Para la última desigualdad, usamos que $|b_i| \leq M(x)$ y la definición de $K(x, u_i)$. Esto prueba que existe $K_1(x) < +\infty$ definido por la igualdad (3.20). Como la desigualdad a la izquierda en (3.20) vale para todo $u \in U_x$ con $\|u\| = 1$, entonces el supremo $K(x)$ definido en (3.19) cumple $K(x) \leq K_1(x) < +\infty$. Análogamente se prueba que existe $H(x) < +\infty$ definido por (3.18).

De las definiciones de los números $H(x, s), K(x, u), H(x), K(x)$ en las igualdades (3.17), (3.18) y (3.19), definiendo $C(x) = \max\{K(x), H(x)\}$, deducimos:

$$\|df^n(s)\| \leq C(x) \lambda(x)^n \|s\| \quad \forall n \geq 0, \quad \forall s \in S_x,$$

$$\|df^{-n}(u)\| \leq C(x) \sigma(x)^{-n} \|u\| \quad \forall n \geq 0, \quad \forall u \in U_x.$$

En efecto, estas desigualdades valen cuando $\|s\| = 1$ y $\|u\| = 1$ por la definición de $H(x)$ y $K(x)$. Entonces valen para todo $s \in S_x$ y para todo $u \in U_x$ por la linealidad de df^n . Esto termina de probar el Lema 3.10.3. \square

3.11. Región de Pesin y medidas hiperbólicas

En esta sección $f \in \text{Diff}^1(M)$ y M es una variedad compacta y riemanniana.

Definición 3.11.1. La *región de Pesin* $P_f \subset M$ es el conjunto de los puntos $x \in M$ Lyapunov-regulares tales que los exponentes de Lyapunov $\chi_x^1 < \chi_x^2 \dots \chi_x^{h(x)}$ son todos diferentes de cero.

Observar que por el Teorema 3.9.8 de Oseledets, la región de Pesin es medible. Para algunos difeomorfismos la región de Pesin puede ser vacía. Por ejemplo, trivialmente, si f es la identidad $P_f = \emptyset$. Otro ejemplo: las rotaciones de la esfera S^2 (f es la rotación de la esfera, de ángulo constante alrededor de un diámetro de S^2 , llamado eje polo norte-polo sur): $P_f = \emptyset$.

Ejercicio 3.11.2. (a) Construir un difeomorfismo $f : S^1 \mapsto S^1$ en el círculo S^1 que preserve la orientación, tal que $P_f = \emptyset$ y tal que en todo abierto $V \subset S^1$ la derivada f' no sea idénticamente igual a 1.

(b) Idem en la esfera S^2 con la condición de que en todo abierto $V \subset S^2$ la derivada df no es idénticamente igual a la identidad Id , ni a $-Id$.

(c) Construir un ejemplo en el círculo S^1 que cumplan las condiciones de la parte (a) y además tal que el conjunto de puntos Lyapunov regulares sea finito.

(d) ¿Existen ejemplos en el círculo que cumplan las condiciones de la parte (a) y además tal que el conjunto de los puntos Lyapunov-regulares sea infinito?

(e) Construir un difeomorfismo $f : S^1 \mapsto S^1$ tal que la región de Pesin $P_f \neq \emptyset$ pero que no coincida con el conjunto de todos los puntos Lyapunov-regulares.

(f) Probar que para todo difeomorfismo f del círculo, o bien la región de Pesin P_f es vacía o bien es finita o bien es infinita numerable, y encontrar ejemplos de los tres casos (sugerencia: probar que todo $x \in P_f$ es aislado en P_f).

(g) Demostrar que existen difeomorfismos $f : S^2 \mapsto S^2$ en la esfera tal que la región de Pesin P_f es no numerable (sugerencia: la Herradura de Smale definida en 3.5.1).

Definición 3.11.3. (Medida hiperbólica)

Una medida de probabilidad f -invariante μ se dice *hiperbólica* si $\mu(P_f) = 1$, donde P_f es la región de Pesin. En otras palabras, μ es hiperbólica si y solo si los exponentes de Lyapunov son diferentes de cero μ -c.t.p.

Para demostrar el siguiente resultado, utilizaremos el Teorema 3.9.8 de Oseledets.

Teorema 3.11.4. Medidas hiperbólicas y conjuntos no uniformemente hiperbólicos

(a) Si μ es medida de probabilidad hiperbólica entonces existe un conjunto invariante Λ (no necesariamente compacto) tal que $\mu(\Lambda) = 1$ y f es hiperbólica (unif. o no unif.) en Λ .

(b) Si Λ es un conjunto invariante tal que f es (unif. o no unif.) hiperbólica en Λ , y si μ es una probabilidad f -invariante tal que $\mu(\Lambda) = 1$, entonces μ es medida hiperbólica.

(c) Si Λ es un conjunto invariante y compacto tal que f es (unif. o no unif.) hiperbólica en Λ , entonces existen medidas de probabilidad μ tales que $\mu(\Lambda) = 1$. Luego, por

la parte (b), todas estas medidas son hiperbólicas. En particular existen medidas de probabilidad hiperbólicas y ergódicas soportadas en Λ .

Demostración. (a) Sea Λ la región de Pesin. Por definición de puntos regulares, Λ es f -invariante, y por el Teorema 3.9.8 de Oseledets, Λ es medible. Por definición de medida hiperbólica $\mu(\Lambda) = 1$. Por el Teorema 3.9.8 de Oseledets, los exponentes de Lyapunov y el splitting en subespacios correspondientes, son funciones medibles. Debido al Lema 3.10.3, como los exponentes de Lyapunov de todo punto $x \in \Lambda$ son no nulos, Λ es (unif. o no unif.) hiperbólico.

(b) De la desigualdad (3.13), tomando logaritmo, dividiendo entre n y haciendo $n \rightarrow +\infty$, se deduce que para μ -c.t.p. $x \in \Lambda$, y para toda dirección $[s] \in S_x$, los exponentes de Lyapunov *hacia el futuro* (según Definición 3.9.2) son menores que $\log \lambda(x) < 0$. Análogamente, de la desigualdad (3.14), tomando logaritmo, dividiendo entre $-n < 0$ (se invierte el sentido de la desigualdad), y haciendo $n \rightarrow +\infty$, deducimos que para toda dirección $[u] \in U_x$ los exponentes de Lyapunov *hacia el pasado* son mayores que $\log \sigma(x) > 0$.

Por el teorema de Oseledets, μ -casi todo punto es Lyapunov regular. Entonces para μ -casi todo punto existen los subespacios de Oseledets para los cuales los exponentes de Lyapunov hacia el pasado son iguales a los exponentes de Lyapunov hacia el futuro. A priori, hay tres casos: el subespacio de Oseledets E_x^i está contenido en S_x , o está contenido en U_x , o ninguna de las dos cosas. En el primer caso, el exponente de Lyapunov hacia el futuro y hacia el pasado en E_x^i es negativo (porque es negativo hacia el futuro por estar contenido en S_x y coincide con el exponente hacia el pasado por ser un subespacio de Oseledets). En el segundo caso, el Lyapunov hacia el futuro y hacia el pasado en E_x^i es positivo (porque es positivo hacia el pasado por estar contenido en U_x y porque coincide con el exponente hacia el futuro por ser un subespacio de Oseledets). Probemos que el tercer caso es vacío; es decir todo subespacio de Oseledets está contenido en S_x o en U_x . Por absurdo, sea una dirección $[v] \in E_x^i$, $[v] \notin S_x, U_x$. Como $T_x M = U_x \oplus S_x$, entonces $v = u + s$ con $0 \neq u \in U_x$, $0 \neq s \in S_x$. Como $\chi_s^+ < 0$, aplicando lo probado en el Ejercicio 3.9.5, el exponente de Lyapunov χ_i hacia el futuro en E_x^i es menor o igual que $\chi_s^+ < 0$. Entonces es negativo. Análogamente, como $\chi_u^- > 0$ (porque $u \in U_x$ y en U_x los exponentes de Lyapunov hacia el pasado son positivos), aplicamos lo probado en el Ejercicio 3.9.5 y deducimos que el exponente de Lyapunov χ_i hacia el pasado en E_x^i es mayor o igual que $\chi_u^- > 0$. Entonces es positivo. Pero en un subespacio de Oseledets, el exponente de Lyapunov hacia el futuro y hacia el pasado coincide. No puede ser negativo y positivo a la vez. Entonces no hay subespacio de Oseledets que no esté incluido en S_x o en U_x .

Concluimos que en todos los subespacios de Oseledets, los exponentes de Lyapunov son no nulos. Por lo probado en el ejercicio 3.9.5, todos los exponentes de Lyapunov en cualquier dirección, hacia el futuro o hacia el pasado, son iguales a algún exponente de Lyapunov en los subespacios de Oseledets. Entonces son no nulos. Esto vale para μ -c.t.p. $x \in M$. Entonces μ -c.t.p. no tiene exponentes de Lyapunov iguales a cero; es decir, μ es una medida de probabilidad hiperbólica.

(c) Sea $\tilde{f} = f|_\Lambda : \Lambda \mapsto \Lambda$. Siendo Λ un espacio métrico compacto y \tilde{f} continua en Λ , el Teorema 1.7.1 asegura que existen medidas de probabilidad μ soportadas

en Λ invariantes y ergódicas para \tilde{f} . Es inmediato chequear que μ , como medida de probabilidad en M , es invariante y ergódica para f . Por la parte b) toda tal medida es hiperbólica. \square

Observación 3.11.5. Medidas hiperbólicas ergódicas: Si una medida es ergódica, entonces los exponentes de Lyapunov son constantes μ -c.t.p. (pues son funciones medibles invariantes c.t.p.). Por el mismo motivo, las dimensiones de los subespacios del splitting en la definición de Lyapunov regularidad son constantes μ -c.t.p. Por lo tanto, para las medidas hiperbólicas ergódicas, la definición del conjunto no uniformemente hiperbólico Λ , tal que $\mu(\Lambda) = 1$, adquiere las particularidades siguientes:

- (i) En las desigualdades (3.13) y (3.14), las tasas de contracción y dilatación $0 < \lambda < 1 < \sigma$ son constantes independientes de $x \in \Lambda$ (mientras que en general el coeficiente $C(x) > 0$ varía con x).
- (ii) Las dimensiones de los subespacios estable S_x e inestable U_x , son constantes, independientes de $x \in \Lambda$.

3.12. Variedades estable e inestable en la región de Pesin

En esta sección, M es una variedad compacta y riemanniana y $f \in \text{Diff}^{1+\alpha}(M)$, es decir f es difeomorfismo C^1 más Hölder. Recordamos que esto significa que $f \in \text{Diff}^1(M)$ y tanto df_x como df_x^{-1} son funciones Hölder-continuas del punto $x \in M$, i.e. existen constantes $\alpha, K > 0$ tales que

$$\|df_x - df_y\| \leq K [\text{dist}(x, y)]^\alpha,$$

y análogamente para df_x^{-1} .

El siguiente teorema, generaliza al caso no uniformemente hiperbólico, el Teorema 3.6.1 de existencia de variedades invariantes para conjuntos uniformemente hiperbólicos (en particular para difeomorfismos de Anosov). Sin embargo, la validez de la siguiente generalización, así como la de los resultados que fundamentan la Teoría de Pesin, está restringida a difeomorfismos de clase $C^{1+\alpha}$.

Teorema 3.12.1. Variedades Estable e Inestable locales (Pesin)

Sea $f : M \mapsto M$ un difeomorfismo C^1 más Hölder en una variedad riemanniana compacta M tal que la región de Pesin $P(f)$ es no vacía. Para $x \in P(f)$ denotamos S_x y U_x los subespacios estable e inestable, respectivamente, correspondientes a los exponentes de Lyapunov negativos y positivos de $x \in \Lambda$. (Notar que $S_x \oplus U_x = T_x M$). Entonces, para todo $x \in P(f)$ existen subvariedades locales $W_{\text{loc}}^s(x)$ y $W_{\text{loc}}^u(x)$, C^1 -encajadas en M , a las que llamamos variedad estable e inestable local respectivamente, tales que:

(a)

$$T_x W_{\text{loc}}^s(x) = S_x, \quad T_x W_{\text{loc}}^u(x) = U_x.$$

(b)

$$f(W_{\text{loc}}^s(x)) \subset W_{\text{loc}}^s(f(x)), \quad f(W_{\text{loc}}^u(x)) \supset W_{\text{loc}}^u(f(x)).$$

(c) Para todo $x \in P(f)$ y para todo $y \in M$ se cumple:

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), f^n(y)) = 0 \quad \text{si } y \in W_{\text{loc}}^s(x),$$

$$\lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(x), f^{-n}(y)) = 0 \quad \text{si } y \in W_{\text{loc}}^u(x).$$

El teorema anterior y su demostración se encuentran en Pesin [Pe 1976]. La demostración puede encontrarse también en [Ba-Pe 2001, Theorem 4.1.1, pag.81] o en [Hi-Pu-Sh 1977].

Se observa, en el Teorema 3.12.1, que la hipótesis f de clase C^1 más Hölder es necesaria. En efecto, Pugh en [Pug 1984] construyó un ejemplo $f \in \text{Diff}^1(M)$ cuya derivada df es continua pero no es Hölder continua, con región de Pesin no vacía, para el que no vale el Teorema 3.12.1 de existencia de variedades invariantes.

Capítulo 4

Teoremas Ergódicos: Descomposición Ergódica, Subaditivo y Multiplicativo

4.1. Descomposición Ergódica

El propósito de esta sección es enunciar y demostrar el Teorema 4.1.2, de Descomposición o Desintegración Ergódica. El mismo extiende el resultado de existencia de medidas ergódicas para transformaciones continuas en espacios métricos compactos, a transformaciones medibles que preservan alguna medida de probabilidad. Además, el Teorema de Descomposición Ergódica muestra cómo se puede descomponer o desintegrar una medida invariante μ en función de las que son ergódicas (sus componentes ergódicas).

Definición 4.1.1. Descomposición o Desintegración Ergódica

Sea (X, \mathcal{A}) un espacio medible y $T : X \mapsto X$ una transformación medible tal que existe alguna medida de probabilidad μ (definida en \mathcal{A}) invariante por T .

(i) Sea $A \in \mathcal{A}$. Decimos que μ tiene *descomposición o desintegración ergódica* para el conjunto A si para μ -c.t.p. $x \in X$ existe una medida ergódica μ_x tal que:

(a) La función real definida μ -c.t.p. por $x \mapsto \mu_x(A)$ es medible para μ -c.t.p. $x \in X$. (Entonces está en $L^1(\mu)$ pues está acotada por 1)

(b)

$$\mu(A) = \int_{x \in X} (\mu_x(A)) d\mu.$$

(ii) Sea $h \in L^1(\mu)$. Decimos que μ tiene *descomposición o desintegración ergódica* para la función h si para μ -c.t.p. $x \in X$ existe una medida ergódica μ_x tal que $h \in L^1(\mu_x)$ y tal que:

(c) La función real definida μ -c.t.p. por $x \mapsto \int h d\mu_x$ es medible para μ -c.t.p. $x \in X$ y está en $L^1(\mu)$.

(d)

$$\int h d\mu = \int_{x \in X} \left(\int h d\mu_x \right) d\mu.$$

Teorema 4.1.2.**Descomposición Ergódica en espacios métricos compactos**

Sea X un espacio métrico compacto provisto de la sigma-álgebra de Borel. Sea $T : X \mapsto X$ una transformación medible que preserva una medida de probabilidad μ .

Entonces, para todo $A \in \mathcal{A}$ y para toda $h \in L^1(\mu)$ existe descomposición ergódica de μ .

Más aún, para μ -c.t.p. $x \in X$ existe y es única una medida ergódica μ_x (llamada **componente ergódica** de μ a la que pertenece x) tal que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j(x)} = \mu_x,$$

donde el límite es en la topología débil estrella del espacio \mathcal{M} de probabilidades.

Daremos la demostración del Teorema 4.1.2 en espacios métricos compactos, más adelante en esta sección. La demostración sigue los mismos argumentos de la que se encuentra en [Mañ 1983a, Cap.II §] (ver también [Mañ 1987]). Una generalización para transformaciones medibles que preservan una medida de probabilidad en espacios de medida no necesariamente métricos compactos, se encuentra en [Ke 1998, Theorem 2.3.3]. Ahora veamos algunas consecuencias del Teorema 4.1.2 de Descomposición Ergódica:

Corolario 4.1.3. *Sea X un espacio métrico compacto y sea $T : X \mapsto X$ medible tal que existe alguna medida de probabilidad μ invariante por T . Entonces existen medidas invariantes ergódicas para T .*

Demostración. Es consecuencia inmediata del Teorema 4.1.2 ya que este establece que toda medida invariante μ tiene componentes ergódicas. \square

Recordamos la Definición 2.1.11: Un conjunto medible $A \subset X$ se dice que tiene *probabilidad total* si $\mu(A) = 1$ para toda medida de probabilidad μ en X que sea invariante por T (bajo la hipótesis que existen medidas de probabilidad invariantes por T).

Corolario 4.1.4. *Sea X un espacio métrico compacto, y sea $T : X \mapsto X$ continua que preserva alguna medida de probabilidad. Entonces un conjunto medible $A \subset X$ tiene probabilidad total si y solo si $\nu(A) = 1$ para toda medida de probabilidad ν ergódica para T .*

Demostración. Usando el Teorema 4.1.2, y la definición 4.1.1, para cualquier conjunto medible A , y cualquiera sea la medida invariante μ , obtenemos la siguiente igualdad para el complemento A^c de A :

$$\mu(A^c) = \int \left(\int \chi_{A^c} d\mu_x \right) d\mu,$$

donde μ_x es una medida ergódica, que depende del punto x y está definida para μ -c.t.p. $x \in X$. Como $\chi_{A^c} \geq 0$ entonces la función $x \mapsto \int \chi_{A^c} d\mu_x = \mu_x(A^c)$ es no negativa. Las medidas μ_x son ergódicas según enuncia el Teorema 4.1.2 y la Definición 4.1.1. Concluimos que $\mu(A^c) = 0$ para toda medida invariante μ , si y solo si $\nu(A^c) = 0$ para toda medida ergódica ν . \square

4.1.5. .

Reducción de probabilidad total a probabilidad total ergódica

Sea (X, \mathcal{A}) un espacio medible cualquiera y sea $T : X \mapsto X$ medible tal que es no vacío el conjunto \mathcal{M}_T de probabilidades T -invariantes y tal que, por hipótesis, existe descomposición ergódica en (X, \mathcal{A}) de cualquier medida $\mu \in \mathcal{M}_T$ para todo $A \in \mathcal{A}$ y toda $h \in L^1(\mu)$, c.f. Definición 4.1.1.

Entonces, argumentando de la misma forma que en el Corolario 4.1.4, tenemos la siguiente propiedad:

Un conjunto $A \in \mathcal{A}$ tiene probabilidad total si y solo si $\nu(A) = 1$ para toda medida ergódica ν en (X, \mathcal{A}) . En otras palabras, para demostrar que $\mu(A) = 1 \quad \forall \mu \in \mathcal{M}$, no es restrictivo asumir que μ es ergódica, y por lo tanto que se cumplen para μ todas las propiedades que caracterizan la ergodicidad.

En el caso particular en que X es compacto, si \mathcal{A} es la sigma-álgebra de Borel, entonces, para cualquier $T : X \mapsto X$ medible que tenga medidas invariantes, la propiedad de descomposición ergódica se cumple en virtud del Teorema 4.1.2.

Ahora pasaremos a la demostración del Teorema 4.1.2 de Descomposición Ergódica en espacios métricos compactos.

4.1.6. Demostración del Teorema 4.1.2.

Dividiremos la prueba en 4 pasos:

Paso 1 de la Prueba del Teorema 4.1.2:

Construcción de las medidas μ_x de desintegración

Probaremos el siguiente:

Lema 4.1.7. *Existe un conjunto $\Sigma_0 \subset X$ de probabilidad total, tal que para todo $x \in \Sigma_0$, existe una medida de probabilidad μ_x (no necesariamente invariante ni ergódica) que satisface las siguientes afirmaciones (4.1), (4.2) y (4.3), para toda medida μ -invariante y para toda función continua $h : X \mapsto \mathbb{R}$:*

$$\int h d\mu_x = \tilde{h}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} h(T^j(x)) \quad \text{para } \mu\text{-c.t.p. } x \in \Sigma_0 \quad (4.1)$$

$$x \mapsto \int h d\mu_x \text{ es medible para } \mu\text{-c.t.p. } x \in \Sigma_0 \text{ y está en } L^1(\mu) \quad (4.2)$$

$$\int h d\mu = \int \left(\int h d\mu_x \right) d\mu. \quad (4.3)$$

Demostración. Observamos que si probamos (4.1), entonces las afirmaciones (4.2) y (4.3) siguen del teorema ergódico de Birkhoff que dice que \tilde{h} es medible, está en $L^1(\mu)$ y satisface $\int h d\mu = \int \tilde{h} d\mu$. Entonces construyamos un conjunto Σ_0 de probabilidad total, tal que toda función continua h satisface (4.1) para cualquier medida invariante μ .

Sea $\{f_i\}_{i \geq 1}$ una familia de funciones continuas densa en $C^0(X, [0, 1])$. En la sección 1.2 probamos que la topología débil estrella del espacio \mathcal{M} , definida por la siguiente métrica, le dota una estructura de espacio métrico compacto:

$$\text{dist}^*(\mu, \nu) := \sum_{i=1}^{+\infty} \left| \int f_i d\mu - \int f_i d\nu \right| \quad \forall \mu, \nu \in \mathcal{M}. \quad (4.4)$$

Además probamos que las siguientes tres afirmaciones son equivalentes:

$$\lim_{n \rightarrow +\infty}^* \mu_n = \mu \text{ en } \mathcal{M}, \text{ es decir } \lim_{n \rightarrow +\infty} \text{dist}^*(\mu_n, \mu) = 0 \quad (4.5)$$

$$\lim_{n \rightarrow +\infty} \int f_i d\mu_n = \int f_i d\mu \quad \forall i \geq 1, \quad (4.6)$$

$$\lim_{n \rightarrow +\infty} \int f d\mu_n = \int f d\mu \quad \forall f \in C^0(X, \mathbb{R}). \quad (4.7)$$

Por el teorema ergódico de Birkhoff aplicado a cada función f_i , existe un conjunto Σ_i con probabilidad total (es decir $\mu(\Sigma_i) = 1 \quad \forall \mu \in \mathcal{M}_T$) tal que existe:

$$\tilde{f}_i(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j(x)) \quad \forall x \in \Sigma_i. \quad (4.8)$$

Luego para todo punto x en el conjunto

$$\Sigma_0 := \bigcap_{i \geq 1} \Sigma_i$$

existe $\tilde{f}_i(x)$ para todo $i \geq 1$. El conjunto Σ_0 tiene probabilidad total porque es intersección numerable de conjuntos con probabilidad total.

Para cada $x \in \Sigma_0$ y para cada $n \geq 1$ definamos la siguiente medida de probabilidad (en general no es invariante):

$$\mu_{x,n} := \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j(x)}, \quad (4.9)$$

donde $\delta_{T^j(x)}$ es la delta de Dirac soportada en el punto $T^j(x)$. Por construcción, para cada $x \in \Sigma$ fijo, tenemos

$$\lim_{n \rightarrow +\infty} \int f_i d\mu_{n,x} = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j(x)) = \tilde{f}_i(x) \quad \forall i \geq 1. \quad (4.10)$$

Como el espacio \mathcal{M} es secuencialmente compacto con la topología débil estrella, existe (para cada elección fija del punto $x \in \Sigma_0$) una subsucesión convergente de $\{\mu_{n,x}\}_{n \geq 1}$. Es decir, existe $n_h \rightarrow +\infty$ y una medida de probabilidad μ_x tales que:

$$\mu_x = \lim_{h \rightarrow +\infty}^* \mu_{n_h, x}. \quad (4.11)$$

Reuniendo las afirmaciones (4.5), (4.6), (4.10) y (4.11) deducimos:

$$\int f_i d\mu_x = \tilde{f}_i(x) \quad \forall x \in \Sigma_0, \quad \forall i \geq 1. \quad (4.12)$$

Como $\lim_h^* \mu_{n_h, x} = \mu_x$, por la igualdad (4.7) deducimos, para toda función continua $f : X \mapsto \mathbb{R}$:

$$\int f d\mu_x = \lim_{h \rightarrow +\infty} \int f d\mu_{n_h, x} = \lim_h \frac{1}{n_h} \sum_{j=1}^{n_h-1} f(T^j(x)) = \tilde{f}(x)$$

para μ -c.t.p. $x \in \Sigma_0$

Concluimos que las afirmaciones (4.1), (4.2) y (4.3) se satisfacen para toda función $h = f$ continua, para toda medida T -invariante μ , terminando la demostración del Lema 4.1.7. \square

Antes de demostrar que μ_x es invariante y ergódica, probaremos que las afirmaciones (4.2) y (4.3) valen también para toda $h \in L^1(\mu)$, para cualquier $\mu \in \mathcal{M}_T$. Además probaremos que la afirmación (4.1) vale también para toda h medible y acotada.

Paso 2 de la Prueba del Teorema 4.1.2:

Demostración de (4.1), (4.2) y (4.3) para h medible acotada.

Para probar (4.1) para h medible acotada, basta demostrar (4.1) para las funciones h medibles acotadas no negativas. En efecto, usando la linealidad de las integrales y del operador $h \mapsto \tilde{h}$, (4.1) se extiende a toda función compleja $h : X \mapsto \mathbb{C}$ medible y acotada.

Lema 4.1.8. *Para cada función medible h no negativa y acotada, existe un conjunto $\Sigma_h \subset \Sigma_0$ de probabilidad total tal que:*

$$\int h d\mu_x = \tilde{h}(x) = \lim_n \frac{1}{n} \sum_{j=1}^{n-1} h(T^j(x)) \quad \forall x \in \Sigma_h.$$

Más aún, fijada la medida T -invariante, entonces para toda $h \in L^1(\mu)$ (aunque no sea acotada), también valen las afirmaciones (4.2) y (4.3).

Demostración. En el Paso 1, probamos la afirmación (4.1) para funciones continuas. Es decir, el Lema 4.1.8 ya está demostrado cuando h es continua.

Vamos a demostrar (4.1) para toda h medible, no negativa y acotada, en el siguiente orden: Primero para funciones características χ_V de abiertos V , luego para funciones

características χ_A de conjuntos medibles A cualesquiera. Después para funciones simples, y finalmente para funciones medibles no negativas acotadas h cualesquiera. Sea V abierto. Sea $\{\psi_n\}_{n \geq 1}$ una sucesión de funciones continuas $\psi_n : X \mapsto [0, 1]$ tales que $\lim_{n \rightarrow +\infty} \psi_n(y) = \chi_V(y)$ para todo $y \in X$. Tales funciones continuas existen por el lema de Urysohn. Tenemos, para cada $x \in \Sigma_0$:

$$\left| \int \chi_V d\mu_x - \tilde{\chi}_V(x) \right| \leq \int |\chi_V - \psi_n| d\mu_x + \left| \int \psi_n d\mu_x - \tilde{\psi}_n(x) \right| + |\tilde{\psi}_n(x) - \tilde{\chi}_V(x)|.$$

El primer sumando es menor que $\epsilon > 0$ para todo n suficientemente grande, por el teorema de convergencia dominada y porque $\lim \psi_n(y) = \chi_V(y)$ para todo $y \in X$. El sumando del medio es cero para μ -c.t.p. $x \in \Sigma_0$ porque ya probamos en el paso 1 que cuando una función ψ es continua entonces $\int \psi d\mu_x = \psi(x)$ para μ -c.t.p. $x \in \Sigma_0$, c.f. igualdad (4.1). Entonces tenemos para todo n suficientemente grande:

$$\left| \int \chi_V d\mu_x - \tilde{\chi}_V(x) \right| \leq \epsilon + |\tilde{\psi}_n(x) - \tilde{\chi}_V(x)| \quad \text{para } \mu - \text{c.t.p. } x \in \Sigma_0. \quad (4.13)$$

Ahora veamos el tercer sumando, integrándolo con respecto a la medida T -invariante μ , aplicando la propiedad triangular del valor absoluto, y el teorema de Birkhoff:

$$\int |\tilde{\psi}_n(x) - \tilde{\chi}_V(x)| d\mu \leq \int |\psi_n - \chi_V| d\mu = \int |\psi_n - \chi_V| d\mu \rightarrow_n 0.$$

El límite al final es debido al teorema de convergencia dominada y a que $\lim_n \psi_n(x) = \chi_V(x)$ para todo punto x . Por lo tanto $\tilde{\psi}_n$ converge en $L^1(\mu)$ a $\tilde{\chi}_V$. Luego, existe una subsucesión de $\{\tilde{\psi}_n\}$ que converge μ -c.t.p. $\tilde{\chi}_V$. Entonces, eligiendo un n suficientemente grande de esa subsucesión, obtenemos

$$|\tilde{\psi}_n(x) - \tilde{\chi}_V(x)| < \epsilon \quad \text{para } \mu - \text{c.t.p. } x \in \Sigma_0. \quad (4.14)$$

Reuniendo las desigualdades (4.13) y (4.14), deducimos

$$\left| \int \chi_V d\mu_x - \tilde{\chi}_V(x) \right| < 2\epsilon \quad \text{para } \mu - \text{c.t.p. } x \in \Sigma_0.$$

Como $\epsilon > 0$ es arbitrario y el término de la izquierda no depende de ϵ obtenemos

$$\int \chi_V d\mu_x = \tilde{\chi}_V(x) \quad \text{para } \mu - \text{c.t.p. } x \in \Sigma_0.$$

Por lo tanto el conjunto Σ_V de puntos $x \in \Sigma_0$ que satisface la igualdad anterior, para cada abierto fijo V , tiene probabilidad total. Hemos probado el Lema 4.1.8 cuando $h = \chi_V$ y V es abierto.

Ahora probemos el Lema cuando $h = \chi_A$ y A es cualquier boreliano. Fijada una medida μ , existe una sucesión decreciente de abiertos V_n tales que $V_n \supset A$ y $\mu(V_n \setminus A) \rightarrow_n 0$ y una sucesión creciente de compactos K_n tales que $K_n \subset A$ y $\mu(A \setminus K_n) \rightarrow_n 0$. Luego $\mu(V_n \setminus K_n) = \mu(V_n \setminus A) + \mu(A \setminus K_n) \rightarrow_n 0$

$$\begin{aligned} \int \left(\left| \int \chi_A d\mu_x - \tilde{\chi}_A(x) \right| \right) d\mu &\leq \int \left| \int \chi_A - \chi_{V_n} d\mu_x \right| d\mu + \\ &\int \left(\left| \int \chi_{V_n} d\mu_x - \tilde{\chi}_{V_n}(x) \right| \right) d\mu + \int |\tilde{\chi}_{V_n}(x) - \tilde{\chi}_A(x)| d\mu. \end{aligned} \quad (4.15)$$

El primer sumando de (4.15):

$$\begin{aligned} \int \left| \int \chi_{V_n} - \chi_A d\mu_x \right| d\mu &\leq \int \left(\int \chi_{V_n \setminus K_n} d\mu_x \right) d\mu = \\ &\int \tilde{\chi}_{V_n \setminus K_n} d\mu = \int \chi_{V_n \setminus K_n} d\mu = \mu(V_n \setminus K_n) < \epsilon \end{aligned}$$

En la primera igualdad de arriba usamos que $V_n \setminus K_n$ es abierto. Entonces, por lo probado antes para los abiertos, tenemos $\int \chi_{V_n \setminus K_n} d\mu_x = \tilde{\chi}_{V_n \setminus K_n}$ para μ -c.t.p. $x \in \Sigma_0$. En la segunda igualdad usamos el teorema de Birkhoff. Por el mismo motivo, el segundo sumando de (4.15) es cero, pues V_n es abierto. Queda:

$$\int \left| \int \chi_A d\mu_x - \tilde{\chi}_A(x) \right| d\mu \leq \epsilon + \int |\tilde{\chi}_{V_n}(x) - \tilde{\chi}_A(x)| d\mu. \quad (4.16)$$

Además, usando el teorema ergódico de Birkhoff, y la propiedad $V_n \supset A$, tenemos

$$\begin{aligned} \int |\tilde{\chi}_{V_n} - \tilde{\chi}_A| d\mu &= \int (\tilde{\chi}_{V_n} - \tilde{\chi}_A) d\mu = \\ &\int (\chi_{V_n} - \chi_A)^\sim d\mu = \int (\chi_{V_n} - \chi_A) d\mu = \mu(V_n \setminus A) < \epsilon \end{aligned}$$

Sustituyendo en la desigualdad (4.16) resulta:

$$\int \left| \int \chi_A d\mu_x - \tilde{\chi}_A(x) \right| d\mu < 2\epsilon.$$

Lo anterior vale para todo $\epsilon > 0$ y el término de la izquierda no depende de ϵ . Entonces concluimos que

$$\int \chi_A d\mu_x = \tilde{\chi}_A(x) \quad \text{para } \mu\text{-c.t.p. } x \in \Sigma_0.$$

Por lo tanto para el boreliano A fijo, el conjunto Σ_A de puntos $x \in \Sigma_0$ que verifica la igualdad de arriba, tiene probabilidad total. Hemos probado el Lema 4.1.8 cuando la función $h = \chi_A$ para cualquier boreliano A .

Por linealidad de la integral y del operador $h \mapsto \tilde{h}$, deducimos que el Lema 4.1.8 vale también para cualquier función simple.

Ahora vamos a probarlo para cualquier función h medible, acotada y no negativa. Sea una medida T -invariante μ . Sea s_n una sucesión monótona creciente de funciones simples no negativas tales que $s_n \leq h$ y que converge μ -c.t.p. a h . Tenemos:

$$\begin{aligned} \int \left| \int h d\mu_x - \tilde{h}(x) \right| d\mu &\leq \int \left| \int (h - s_n) d\mu_x \right| d\mu + \\ &+ \int \left| \int s_n d\mu_x - \tilde{s}_n(x) \right| d\mu + \int |\tilde{s}_n(x) - \tilde{h}(x)| d\mu. \end{aligned} \quad (4.17)$$

Analicemos el primer sumando $\int \left(\int (h - s_n) d\mu_x \right) d\mu$. Tenemos $0 \leq h(y) - s_n(y) \leq h(y)$ para todo $y \in X$. Fijemos $\epsilon > 0$:

$$0 \leq \int (h - s_n) d\mu_x \leq \epsilon \mu_x(A^c) + k \mu_x(A) \leq \epsilon + k \mu_x(A),$$

donde

$$A := \{y \in X : h(y) - s_n(y) > \epsilon\}, \quad k := \sup_{y \in X} h(y).$$

Luego:

$$\begin{aligned} \int \left| \int (h - s_n) d\mu_x \right| d\mu &= \int \left(\int (h - s_n) d\mu_x \right) d\mu \leq \\ &\int (\epsilon + k \mu_x(A)) d\mu < \epsilon + k \int \left(\int \chi_A d\mu_x \right) d\mu = \\ &\epsilon + k \int \tilde{\chi}_A(x) d\mu = \epsilon + k \int \chi_A d\mu = \epsilon + k \mu(A). \end{aligned} \quad (4.18)$$

En las tres últimas igualdades hemos usado la afirmación (4.1), ya probada para todo conjunto medible A , y luego el teorema ergódico de Birkhoff. Como $\lim_{n \rightarrow +\infty} s_n(x) = h(x)$ para μ -c.t.p. $x \in X$, tenemos que $\mu(A) < \epsilon$ para todo n suficientemente grande. En efecto, por convergencia dominada, $0 \leq \int (h - s_n) d\mu < \epsilon^2$ para todo n suficientemente grande. Si fuera $\mu(A) > \epsilon$, entonces, como para todo $y \in A$ se cumple $h(y) - s_n(y) > \epsilon$, tendríamos $\int (h - s_n) d\mu > \epsilon \cdot \mu(A) > \epsilon^2$.

Juntando con la ecuación (4.18), obtenemos:

$$\int \left| \int (h - s_n) d\mu_x \right| d\mu \leq (1 + k)\epsilon. \quad (4.19)$$

Por lo tanto hemos acotado el primer sumando de la igualdad (4.17). El segundo sumando es cero porque ya probamos para las funciones simples s_n que $\int s_n d\mu_x = \tilde{s}_n(x)$ para μ -c.t.p. $x \in \Sigma_0$. Entonces tenemos para todo n suficientemente grande:

$$\int \left| \int h d\mu_x - \tilde{h} \right| d\mu \leq (1 + k)\epsilon + \int |\tilde{s}_n(x) - \tilde{h}(x)| d\mu \quad (4.20)$$

Ahora acotemos el último sumando, aplicando la propiedad triangular del valor absoluto, y el teorema de Birkhoff:

$$\int |\tilde{s}_n(x) - \tilde{h}(x)| d\mu \leq \int |s_n - h|^\sim d\mu = \int |s_n - h| d\mu \rightarrow_n 0. \quad (4.21)$$

El límite al final es debido al teorema de convergencia dominada ya que, por construcción de la sucesión de funciones simples tenemos $\lim_n s_n(x) = h(x)$ para μ casi todo punto x . Reuniendo las desigualdades (4.20) y (4.21), deducimos

$$\int \left| \int h d\mu_x - \tilde{h}(x) \right| d\mu < (1+k)\epsilon + \epsilon$$

Como $\epsilon > 0$ es arbitrario y el término de la izquierda no depende de ϵ obtenemos

$$\int h d\mu_x = \tilde{h}(x) \quad \text{para } \mu - \text{c.t.p. } x \in \Sigma_0.$$

Como la medida μ es cualquiera, para cada h fija (medible y acotada) el conjunto Σ_h de puntos $x \in \Sigma_0$ que satisface la igualdad anterior tiene probabilidad total. Hemos probado el Lema 4.1.8 para todas las funciones medibles acotadas no negativas.

Finalmente probemos la última afirmación del Lema 4.1.8, es decir, para una medida T -invariante fija μ probemos (4.2) y (4.3) para toda $h \in L^1(\mu)$.

Por la linealidad de las integrales basta probar (4.2) y (4.3) para $h \in L^1(\mu)$ no negativa. Consideremos la sucesión monótona creciente de funciones acotadas $h_n(x) = \max\{f(x), n\}$. Tenemos $\lim_n h_n(x) = h(x)$ para todo $x \in X$. Luego, por convergencia monótona

$$\lim_{n \rightarrow +\infty} \int h_n d\mu_x = \int h d\mu_x$$

Ya sabemos que $\int h_n d\mu_x = \tilde{h}_n(x)$ (lo probamos para toda función medible acotada), y la función \tilde{h}_n , por el Teorema de Birkhoff es medible y está en $L^1(\mu)$. Entonces la aplicación $x \mapsto \int h d\mu_x$ es medible (porque es el límite de una sucesión de funciones medibles). Además como h_n es no negativa y es una sucesión creciente con n , entonces $\tilde{h}_n d\mu_x$, para cada x fijo, es no negativa y es creciente con n . Luego, por el teorema de convergencia monótona:

$$\begin{aligned} \int \left(\int h d\mu_x \right) d\mu &= \lim_{n \rightarrow +\infty} \int \left(\int h_n d\mu_x \right) d\mu = \\ &= \lim_{n \rightarrow +\infty} \int \tilde{h}_n d\mu = \lim_{n \rightarrow +\infty} \int h_n d\mu = \int h d\mu. \end{aligned}$$

En la penúltima igualdad usamos el teorema de Birkhoff. En la última igualdad usamos nuevamente el teorema de convergencia monótona. Concluimos que:

$$\int \left(\int h d\mu_x \right) d\mu = \int h d\mu.$$

Como $h \in L^1(\mu)$, entonces $h \in L^1(\mu_x)$ para μ -c.t.p. x y además la aplicación $x \mapsto \int h d\mu_x$ (que ya probamos que es medible) está en $L^1(\mu)$. \square

Paso 3 de la Prueba del Teorema 4.1.2:

Demostración de que μ_x es T -invariante.

Sea Σ_0 el conjunto de probabilidad total construido en el Paso 1. Hacemos notar que no necesariamente μ_x es T -invariante para todo punto $x \in \Sigma_0$. El enunciado preciso de la T -invariancia de μ_x es el siguiente:

Lema 4.1.9. *Existe un conjunto $\Sigma_1 \subset \Sigma_0$ con probabilidad total tal que μ_x es T -invariante para cada punto $x \in \Sigma_1$.*

Demostración. Consideremos el conjunto numerable $\{f_i\}_{i \geq 1}$ de funciones continuas denso en $C^0(X, [0, 1])$ como en la demostración del Lema 4.1.7. Tenemos en la igualdad (4.12):

$$\int f_i d\mu_x = \widetilde{f}_i(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j(x)) \quad \forall x \in \Sigma_0.$$

Además por el Lema 4.1.8, para la función medible, no negativa y acotada $f_i \circ T$ existe un conjunto $\Sigma'_i \subset \Sigma_0$, con probabilidad total tal que

$$\int f_i \circ T d\mu_x = \widetilde{f_i \circ T}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_i(T^j(T(x))) \quad \forall x \in \Sigma'_i.$$

Como además $\widetilde{f_i \circ T}(x) = \widetilde{f}_i(x)$ para todo $x \in \Sigma'_i \cap \Sigma_0 = \Sigma'_i$, deducimos que:

$$\int f_i \circ T d\mu_x = \int f_i d\mu_x \quad \forall x \in \Sigma'_i. \quad (4.22)$$

Recordamos el operador $T^* : \mathcal{M} \mapsto \mathcal{M}$ definido por

$$(T^*\nu)(B) = \nu(T^{-1}(B)) \quad \forall B \in \mathcal{B}.$$

De la definición de invariancia de una medida, deducimos que ν es T -invariante si y solo si $T^*\nu = \nu$. Además, para toda función medible acotada $h : X \mapsto \mathbb{R}$ (en particular para $h = f_i$), y para cualquier medida de probabilidad ν (no necesariamente T invariante):

$$\int h \circ T d\nu = \int h d(T^*\nu)$$

En efecto, de la definición del operador T^* la igualdad de arriba se prueba inmediatamente para las funciones características de los borelianos B . Por la linealidad de la integral, es cierta también para las funciones simples, y por convergencia monótona, vale también para las funciones medibles no negativas. Luego por linealidad, es cierta también para toda h medible acotada (descomponiendo esta en su partes positiva y negativa).

Concluimos que la igualdad (4.22) puede escribirse como:

$$\int f_i d(T^*\mu_x) = \int f_i d\mu_x \quad \forall x \in \Sigma'_i.$$

Definimos:

$$\Sigma_1 := \bigcap_{i=1}^{+\infty} \Sigma'_i.$$

Como cada Σ'_i tiene probabilidad total, y son una cantidad numerable, entonces Σ_1 tiene probabilidad total. Luego, dejando fija la elección de un punto $x \in \Sigma_1$, se cumple:

$$\int f_i d(T^* \mu_x) = \int f_i d\mu_x \quad \forall i \geq 1.$$

Recordando la definición de la distancia débil estrella dist^* en el espacio de probabilidades, según la igualdad (4.4), deducimos

$$\text{dist}^*(T^* \mu_x, \mu_x) = 0 \quad \forall x \in \Sigma_1.$$

Luego $T^* \mu_x = \mu_x$, es decir μ_x es T -invariante. \square

Paso 4 de la Prueba del Teorema 4.1.2:

Prueba de que μ_x es ergódica.

Sea Σ_1 el conjunto de probabilidad total que existe por el Lema 4.1.9. Probaremos el siguiente:

Lema 4.1.10. *Existe un conjunto $\Sigma_2 \subset \Sigma_1$ con probabilidad total tal que μ_x es ergódica para T para cada punto $x \in \Sigma_2$.*

Demostración. Primero probaremos la siguiente afirmación:

Afirmación (A): *Si para un punto $x \in \Sigma_1$ (elegido fijo) se cumple*

$$\int_{y \in X} (\tilde{f}_i(y) - \tilde{f}_i(x))^2 d\mu_x = 0 \quad \forall i \geq 1, \quad (4.23)$$

entonces μ_x es ergódica.

En efecto, si se cumple la igualdad de arriba, entonces $\tilde{f}_i(y) = \tilde{f}_i(x)$ para μ_x -c.t.p. $y \in X$, para cada $i \geq 1$. Luego $\tilde{f}_i(y)$ es constante μ_x -c.t.p. Entonces, como μ_x es invariante, por el teorema de Birkhoff, tenemos

$$\tilde{f}_i(y) = \int \tilde{f}_i d\mu_x = \int f_i d\mu_x \quad \text{para } \mu_x - \text{c.t.p. } y \in X. \quad (4.24)$$

Si demostramos que para cualquier conjunto medible A se cumple

$$\tilde{\chi}_A(y) = \int \chi_A d\mu_x \quad \text{para } \mu_x - \text{c.t.p. } y \in X, \quad (4.25)$$

entonces en el caso particular que $T^{-1}(A) = A$, tendremos

$\tilde{\chi}_A(y) = \lim_{n \rightarrow +\infty} (1/n) \sum_{j=0}^{n-1} \chi_A(T^j y) = \chi_A(y) \in \{0, 1\}$ para todo $y \in X$ (porque $\chi_A \circ T^j = \chi_{T^{-j}(A)} = \chi_A$). Luego, $\int \chi_A d\mu_x \in \{0, 1\}$, lo cual implica $\mu_x(A) \in \{0, 1\}$ y μ_x es ergódica.

Demostremos (4.25): por hipótesis, se cumple para f_i en lugar de χ_A , cualquiera sea $i \geq 1$, c.f. Igualdad (4.24). Veamos que se cumple para cualquier función continua f .

Basta probarlo para funciones continuas $f \in C^0(X, [0, 1])$. Como $\{f_i\}_{i \geq 1}$ es denso en $C^0(X, [0, 1])$ dado $\epsilon > 0$ existe f_i tal que

$$\|f - f_i\|_{C^0} = \max_{x \in X} |f(x) - f_i(x)| < \epsilon.$$

Entonces

$$\begin{aligned} \left| \tilde{f}(y) - \int f d\mu_x \right| &\leq |\tilde{f}(y) - \tilde{f}_i(y)| + \left| \tilde{f}_i(y) - \int f_i d\mu_x \right| + \\ &+ \left| \int f_i d\mu_x - \int f d\mu_x \right| \leq |f - f_i|^\sim(y) + \int |f_i - f| d\mu_x \leq 2\epsilon, \end{aligned}$$

para μ_x -c.t.p. $y \in X$ (donde $\tilde{f}_i(y) = \int f_i d\mu_x$). Concluimos que $|\tilde{f}(y) - \int f d\mu_x| < 2\epsilon$ para todo $\epsilon > 0$. Luego $\tilde{f}(y) = \int f d\mu_x$ para μ_x -c.t.p. $y \in X$. Hemos probado la igualdad (4.25) para f continua en lugar de χ_A . Ahora probemos la igualdad (4.25) para χ_A con A medible cualquiera. Recordemos que la medida μ_x está fija, pues el punto $x \in \Sigma_1$ está fijo por ahora. Sean una sucesión de compactos K_n y una de abiertos V_n tales que $K_n \subset A \subset V_n$ con $\lim_n \mu_x(V_n \setminus A) = 0$ y $\lim_n \mu_x(A \setminus K_n) = 0$. Tales abiertos y compactos existen por la regularidad de cualquier medida de probabilidad en la sigma-álgebra de Borel. Sea $\psi_n \in C^0(X, [0, 1])$ tal que $\psi_n|_{K_n} = 1$ y $\psi_n|_{V_n^c} = 0$. Tal función continua existe por el Lema de Urysohn. Tenemos

$$\begin{aligned} \left| \tilde{\chi}_A(y) - \int \chi_A(y) d\mu_x \right| &\leq |\tilde{\chi}_A(y) - \tilde{\psi}_n(y)| + \\ &+ \left| \tilde{\psi}_n(y) - \int \psi_n d\mu_x \right| + \left| \int \psi_n d\mu_x - \int \chi_A d\mu_x \right| \leq \\ &\leq |\tilde{\chi}_A(y) - \tilde{\psi}_n(y)| + \int |\psi_n - \chi_A| d\mu_x, \end{aligned} \quad (4.26)$$

para μ_x -c.t.p. $y \in X$ (donde $\tilde{\psi}_n(y) = \int \psi_n d\mu_x$). Probemos que el primer sumando $|\tilde{\chi}_A(y) - \tilde{\psi}_n(y)|$ en la desigualdad (4.26) es menor que ϵ para μ_x -c.t.p. $y \in X$. En efecto, observamos que $\chi_{K_n} \leq \chi_A \leq \chi_{V_n}$ porque $K_n \subset A \subset V_n$. Además $\chi_{K_n} \leq \psi_n \leq \chi_{V_n}$ porque $\psi_n|_{K_n} = 1$, $\psi_n|_{V_n^c} = 0$ y $0 \leq \psi_n(y) \leq 1$ para todo y . Entonces deducimos que

$$|\chi_A(y) - \psi_n(y)| \leq \chi_{V_n}(y) - \chi_{K_n}(y) \quad \forall y \in X.$$

Luego $|\chi_A - \psi_n|^\sim \leq (\chi_{V_n} - \chi_{K_n})^\sim$. Por la propiedad triangular de la norma y por el teorema de Birkhoff:

$$\begin{aligned} \int |\tilde{\chi}_A(y) - \tilde{\psi}_n(y)| d\mu_x &\leq \int |\chi_A - \psi_n|^\sim d\mu_x \leq \int (\chi_{V_n} - \chi_{K_n})^\sim d\mu_x = \\ &= \int (\chi_{V_n} - \chi_{K_n}) d\mu_x = \mu_x(V_n \setminus K_n) = \mu_x(V_n \setminus A) + \mu_x(A \setminus K_n) \rightarrow_n 0. \end{aligned}$$

Hemos probado $\tilde{\psi}_n$ converge a $\tilde{\chi}_A$ en la norma $L^1(\mu_x)$. Luego, existe una subsucesión de $\tilde{\psi}_n(y)$ que converge a $\tilde{\chi}_A(y)$ para μ_x -c.t.p. $y \in X$. Entonces, eligiendo un n suficientemente grande de esa subsucesión, tenemos que

$$|\tilde{\chi}_A(y) - \tilde{\psi}_n(y)| < \epsilon \quad \text{para } \mu_x - \text{c.t.p. } y \in X.$$

Sustituyendo en la desigualdad (4.26):

$$\left| \tilde{\chi}_A(y) - \int \chi_A(y) d\mu_x \right| \leq \epsilon + \int |\psi_n - \chi_A| d\mu_x, \quad (4.27)$$

Acotemos ahora el término a la derecha de (4.27).

$$\int |\psi_n - \chi_A| d\mu_x \leq \int |\chi_{V_n} - \chi_{K_n}| d\mu_x = \mu_x(V_n \setminus K_n) \rightarrow_n 0.$$

Luego es $\int |\psi_n - \chi_A| d\mu_x < \epsilon$ si n es suficientemente grande. Sustituyendo en (4.27):

$$\left| \tilde{\chi}_A(y) - \int \chi_A(y) d\mu_x \right| < 2\epsilon, \quad \text{para } \mu_x - \text{c.t.p. } y \in X. \quad (4.28)$$

para todo $\epsilon > 0$. Como el término a la izquierda es independiente de ϵ , deducimos que:

$$\tilde{\chi}_A(y) = \int \chi_A d\mu_x \quad \text{para } \mu_x - \text{c.t.p. } y \in X,$$

probando la igualdad (4.25) como queríamos demostrar.

Hemos demostrado la Afirmación (A) del principio:

Para cada $x \in \Sigma_1$ fijo, la medida μ_x (que ya sabemos que es invariante), es ergódica si para todo $i \geq 1$ se cumple la igualdad (4.23).

Entonces, ahora solo falta probar que existe un conjunto $\Sigma_2 \subset \Sigma_1$ con probabilidad total, tal que para todo $x \in \Sigma_2$ se cumple la igualdad (4.23) para todo $i \geq 1$. Fijemos $x \in \Sigma_1$ e $i \geq 1$. Calculemos la siguiente integral:

$$I_i(x) = \int_{y \in X} (\tilde{f}_i(y) - \tilde{f}_i(x))^2 d\mu_x.$$

$$I_i(x) = \int_{y \in X} (\tilde{f}_i(y))^2 d\mu_x - 2\tilde{f}_i(x) \int_{y \in X} \tilde{f}_i(y) d\mu_x + (\tilde{f}_i(x))^2$$

Observamos que $I_i(x) \geq 0$. Queremos probar que $I_i(x) = 0$ para un conjunto $\Sigma''_i \subset \Sigma_1$ de puntos x que tiene probabilidad total. Integramos $I_i(x)$ respecto de una medida μ cualquiera T -invariante:

$$\begin{aligned} 0 &\leq \int I_i(x) d\mu = \\ &= \int \left(\int (\tilde{f}_i)^2 d\mu_x \right) d\mu - 2 \int \tilde{f}_i(x) \left(\int \tilde{f}_i d\mu_x \right) d\mu + \int (\tilde{f}_i(x))^2 d\mu \end{aligned} \quad (4.29)$$

Sabemos que μ_x es invariante con T . Entonces $\int \tilde{f}_i d\mu_x = \int f_i d\mu_x$. Recordemos la igualdad (4.12): $\int f_i d\mu_x = \tilde{f}_i(x)$ para todo $x \in \Sigma_0$, en particular para todo $x \in \Sigma_1 \subset \Sigma_0$. Sustituyendo en (4.29):

$$\begin{aligned} 0 &\leq \int I_i(x) d\mu = \\ &= \int \left(\int (\tilde{f}_i)^2 d\mu_x \right) d\mu - 2 \int (\tilde{f}_i(x))^2 d\mu + \int (\tilde{f}_i(x))^2 d\mu. \end{aligned} \quad (4.30)$$

Por el lema 4.1.8, para toda h medible y acotada, en particular para $h = (\tilde{f}_i)^2$, vale la igualdad (4.3). Es decir: $\int \left(\int h d\mu_x \right) d\mu = \int h d\mu$. Entonces sustituyendo en (4.30), queda:

$$\int I_i(x) d\mu = \int (\tilde{f}_i(x))^2 d\mu - 2 \int (\tilde{f}_i(x))^2 d\mu + \int (\tilde{f}_i(x))^2 d\mu = 0. \quad (4.31)$$

Como $I_i(x) \geq 0$ para todo x , concluimos que $I_i(x) = 0$ para μ -c.t.p. $x \in \Sigma_1$. Es decir, se cumple la igualdad (4.23), para cada $i \geq 1$ fijo, para un conjunto $\Sigma_i'' \subset \Sigma_1$ de puntos x que tiene probabilidad total. Entonces construimos

$$\Sigma_2 = \bigcap_{i=1}^{+\infty} \Sigma_i'',$$

que tiene probabilidad total, pues es intersección numerable de conjuntos con probabilidad total. Concluimos que para cada punto $x \in \Sigma_2$ se cumple la igualdad (4.23) para todo $i \geq 1$. Luego, por la afirmación probada al principio μ_x es ergódica para todo $x \in \Sigma_2$, como queríamos probar, terminando de demostrar el Teorema 4.1.2. \square

4.2. Teorema Ergódico Subaditivo de Kingmann

En esta sección solo asumimos que $T : X \mapsto X$ es medible en un espacio medible X . Se recuerda la Definición 2.2.2 de sucesión subaditiva de funciones medibles reales $\{f_n\}_{n \geq 1}$:

$$f_{n+m} \leq f_n + f_m \circ T^n \quad \forall m, n \geq 1.$$

Ejercicio 4.2.1. Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva de funciones medibles reales.

Sea $a(x) := \limsup_{n \rightarrow +\infty} \frac{f_n(x)}{n} \in [-\infty, +\infty)$ y $b(x) := \liminf_{n \rightarrow +\infty} \frac{f_n(x)}{n} \in [-\infty, +\infty)$.

(a) Probar que $a(x) \leq a(T(x))$, $b(x) \leq b(T(x))$ para todo $x \in X$. Sugerencia: $(a_{n+1}(x))/n \leq a_1(x)/n + a_n(T(x))/n$.

(b) Probar que para toda medida invariante

$$b(x) = b(T(x)) \quad a(x) = a(T(x)) \quad \mu - \text{c.t.p.} \quad x \in X.$$

Sugerencia: Dada una constante $\lambda \in [-\infty, +\infty)$, considerar el conjunto $B(\lambda) = \{x \in X : b(x) \geq \lambda\}$. Chequear $B(\lambda) \subset T^{-1}(B(\lambda))$. Como μ es invariante por T , deducir que $\mu(T^{-1}B(\lambda) \setminus B(\lambda)) = 0$.

El Teorema Ergódico de Birkhoff-Khinchin es un caso particular del siguiente teorema:

Teorema 4.2.2. Teorema Ergódico Subaditivo (Kingmann)

Sea (X, \mathcal{A}) un espacio medible. Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Sea $\{f_n\}_{n \geq 1}$ una sucesión subaditiva de funciones reales medibles tales que $f_1^+ \in L^1(\mu)$, donde $f_1^+ := \max\{0, f_1\}$. Entonces existe el siguiente límite μ -c.t.p.

$$\tilde{f}(x) = \lim_{n \rightarrow +\infty} \frac{f_n(x)}{n} \in [-\infty, +\infty) \quad \mu - \text{c.t.p. } x \in X.$$

Además $\tilde{f}^+ := \max\{0, \tilde{f}\} \in L^1(\mu)$ y se satisface la siguiente igualdad:

$$\int \tilde{f} d\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \int f_n d\mu = \inf_{n \geq 1} \frac{1}{n} \int f_n d\mu \in [-\infty, +\infty). \quad (4.32)$$

A continuación daremos la prueba del Teorema Ergódico Subaditivo 4.2.2, usando el Teorema Ergódico 2.1.2 de Birkhoff-Khinchin. La demostración de Kingmann se encuentra en [Kin 1973]. Otra demostración, que no usa el Teorema Ergódico de Birkhoff-Khinchin, y por lo tanto puede sustituir la demostración de este último teorema, se encuentra en [Boc 2012].

Observación: Separando f_n en parte positiva $f_n^+ := \max\{0, f_n\}$ y negativa $f_n^- := -\min\{0, f_n\}$, resulta $f_n = f_n^+ - f_n^-$. Es fácil ver que $\{f_n^+\}$ es subaditiva, luego $0 \leq \int f_n^+ \leq n \int f_1^+ < +\infty$. Entonces $f_n^+ \in L^1(\mu)$ para todo $n \geq 1$. Sin embargo f_n^- puede no ser integrable. Por lo tanto $\int f_n/n d\mu \in [-\infty, +\infty)$ para todo $n \geq 0$, y el valor de $\tilde{f}(x)$, así como las igualdades (4.32) de la tesis en el Teorema 4.2.2, pueden ser $-\infty$.

Demostración del Teorema 4.2.2:

Paso 1. Primero veremos que el conjunto A de los puntos x donde existe (en $[-\infty, +\infty)$) $\tilde{f}(x) = \lim_{n \rightarrow +\infty} (f_n(x))/n$, es medible.

Demostración. Las funciones

$$a(x) := \limsup_{n \rightarrow +\infty} f_n(x)/n \in [-\infty, +\infty),$$

$$b(x) := \liminf_{n \rightarrow +\infty} f_n(x)/n \in [-\infty, +\infty),$$

son medibles por ser límite superior e inferior respectivamente, de una sucesión de funciones medibles. Luego, si dos funciones a y $b : X \mapsto [-\infty, +\infty)$ son medibles, el conjunto de puntos $x \in X$ tales que $a(x) = b(x)$ es medible: Es la unión del conjunto $\{a(x) > -\infty, b(x) > -\infty, (a - b)(x) = 0\}$ con el conjunto $\{a(x) = -\infty\} \cap \{b(x) = -\infty\}$. \square

Para demostrar el Teorema 4.2.2 hay que probar que A tiene probabilidad total y que además se cumplen las igualdades (4.32) con las integrales para cualquier medida invariante μ .

Paso 2. Veamos que para μ -casi todo punto $x \in X$, las funciones $(f_n^+(x)/n)$ tienen parte positiva en $L^1(\mu)$ y acotemos superiormente sus integrales.

Demostración. Usando la subaditividad de $\{f_n\}$, es fácil ver que $\{f_n^+\}$ también es subaditiva. Luego, tenemos

$$\frac{f_n(x)}{n} \leq \frac{f_n^+(x)}{n} \leq \frac{1}{n} \sum_{j=0}^{n-1} f_1^+(T^j(x)) \in L^1(\mu).$$

Entonces para todo $n \geq 1$ están definidas las integrales (extendidas) $\int f_n/n d\mu \in [-\infty, +\infty)$. Además se cumple:

$$\int \frac{f_n}{n} d\mu \leq \int \frac{f_n^+}{n} d\mu \leq \int f_1^+ d\mu \quad \forall n \geq 1.$$

□

Paso 3. Ahora probaremos la igualdad de la derecha en (4.32) para toda medida μ que sea invariante por T .

Demostración. Denotemos, para todo $n \geq 1$:

$$I_n := \int f_n d\mu \in [-\infty, +\infty), \quad I := \inf\{I_n/n : n \geq 1\} \in [-\infty, +\infty).$$

Probaremos que

$$I = \lim_{n \rightarrow +\infty} I_n/n.$$

Es inmediato chequear que $I_{n+m} \leq I_n + I_m$ para todos $n, m \geq 1$ (porque μ es una medida T -invariante, y entonces $\int f_n \circ T^m d\mu = \int f_n d\mu$). Entonces, si $I_{n_0} = -\infty$ para un primer natural $n_0 \geq 1$, deducimos que la sucesión $\{I_n\}_{n \geq 1}$ es finalmente constante igual a $-\infty$ para todo $n \geq n_0$. En este caso trivialmente se verifica $\lim_n I_n = \inf I_n = -\infty$. Ahora supongamos que $I_n \in \mathbb{R}$ para todo $n \geq 1$. Dados $n, m \geq 1$, hacemos la división entera $n = qm + r$, $0 \leq r \leq m - 1$. De la subaditividad de la sucesión de números I_n (denotando $I_0 = 0$), obtenemos:

$$I_n = I_{qm+r} \leq qI_m + I_r \quad \forall n = mq + r \geq 1, \quad m \geq 1, q \geq 0, 0 \leq r \leq m - 1.$$

Luego, dividiendo entre $n \geq 1$:

$$\frac{I_n}{n} \leq \frac{qm}{qm+r} \cdot \frac{I_m}{m} + \frac{I_r}{qm+r}.$$

Dejando fijo m , si se toma $qm + r = n \rightarrow +\infty$, como r está acotado por m , resulta $q \rightarrow +\infty$. Siendo $|I_r| \leq \max_{0 \leq j \leq m-1} |I_j|$, queda:

$$[-\infty, +\infty) \ni \limsup_{n \rightarrow +\infty} \frac{I_n}{n} \leq \frac{I_m}{m} \quad \forall m \geq 1.$$

Luego

$$\limsup_{n \rightarrow +\infty} \frac{I_n}{n} \leq \inf_{m \geq 1} \frac{I_m}{m} \leq \liminf_{n \rightarrow +\infty} \frac{I_n}{n} \in [-\infty, +\infty),$$

de donde se deduce que existe $\lim_{n \rightarrow +\infty} I_n/n \in [-\infty, +\infty)$ y es igual a $I = \inf_{n \geq 1} I_n/n$. Hemos probado la igualdad de la derecha en (4.32). \square

Para demostrar el Teorema 4.2.2, ahora resta probar que para cualquier medida de probabilidad μ invariante por T , para μ -c.t.p. $x \in X$ existe $\tilde{f}(x) := a(x) = b(x)$, que $\tilde{f}^+ \in L^1(\mu)$ y que $\int \tilde{f} d\mu = I$ para toda medida invariante μ .

Afirmación a demostrar: *Bajo la hipótesis adicional de que existe una constante $-K < 0$ tal que*

$$\frac{f_n(x)}{n} \geq -K \quad \forall n \geq 1, \quad \forall x \in X, \quad (4.33)$$

se cumplen, para toda medida invariante μ , las siguientes desigualdades:

$$\int a(x) d\mu \leq I \leq \int b(x) d\mu, \quad (4.34)$$

donde $a(x) := \limsup_n f_n(x)/n$, $b(x) := \liminf f_n(x)/n$.

Más adelante probaremos esta afirmación.

Paso 4. Ahora probaremos que alcanza demostrar la afirmación anterior, es decir, las desigualdades (4.34) bajo la hipótesis adicional (4.33), para terminar la demostración del Teorema 4.2.2 en general (sin la hipótesis adicional).

Demostración. Discutimos dos casos: o bien existe una constante $-K < 0$ tal que $f_n(x)/n \geq -K$ para todo $n \geq 1$ y todo $x \in X$, o bien no existe tal $-K$.

En el primer caso, $-K \leq b(x) \leq a(x)$ porque a es el límite superior y b el inferior de $f_n/n \geq -K$, y ninguno de ellos es $-\infty$. Además $I = \inf \int (f_n/n) d\mu \geq -K$ (también como consecuencia de que $f_n/n \geq -K$). Debido a (4.34): $\int a d\mu \leq I \leq \int b d\mu$, donde μ es una medida de probabilidad, I es un número real y por lo tanto $b(x) \leq a(x)$ son reales μ -c.t.p. $x \in X$. Restando $\int b d\mu$ queda $\int (a - b) d\mu \leq 0$. Como la integral de la función real $a - b \geq 0$ solo puede dar ≤ 0 (en realidad = 0) si $a(x) = b(x)$ μ -c.t.p., deducimos que ambas funciones a y b coinciden μ -c.t.p. y que las desigualdades (4.34) son en realidad igualdades. Por lo tanto existe $\lim f_n/n = \tilde{f}$ y su integral es igual a I . Esto termina, una vez probadas las desigualdades (4.34), la demostración del Teorema 4.2.2 en el primer caso en que rige la hipótesis adicional (4.33).

En el segundo caso, cuando no vale la hipótesis adicional (4.33), fijamos una constante entera $-K < 0$ cualquiera (que luego haremos tender a $-\infty$) y construimos la sucesión auxiliar de funciones

$$g_n(x) := \max\{f_n(x), -nK\}. \quad (4.35)$$

Es inmediato chequear que g_n es subaditiva. Entonces, $\{g_n\}$ está en el primer caso. Una vez demostradas las desigualdades (4.34), sabremos que existe

$$\lim_{n \rightarrow +\infty} \frac{g_n}{n}(x) = \tilde{g}(x) \quad \mu - \text{c.t.p.}$$

y además

$$\int \tilde{g} d\mu = \lim_n \int \frac{g_n}{n} d\mu = \inf_n \int \frac{g_n}{n} d\mu.$$

Entonces, para μ -c.t.p. $x \in X$, tenemos:

$$\begin{aligned} \max \left\{ \limsup_n \frac{f_n(x)}{n}, -K \right\} &= \limsup_n \max \left\{ \frac{f_n(x)}{n}, -K \right\} = \\ &= \limsup_n \frac{g_n(x)}{n} = \tilde{g}(x) = \liminf_n \frac{g_n(x)}{n} = \\ &= \liminf_n \max \left\{ \frac{f_n(x)}{n}, -K \right\} = \max \left\{ \liminf_n \frac{f_n(x)}{n}, -K \right\} \end{aligned}$$

para todo $-K < 0$ entero fijo. Haciendo $-K \rightarrow -\infty$ en la igualdad anterior, el término en el extremo de la izquierda tiende monótonamente a $a(x) = \limsup_n f_n/n \in [-\infty, +\infty)$, mientras que el término en el extremo de la derecha lo hace a $b(x) = \liminf_n f_n/n \in [-\infty, +\infty)$. Resulta la siguiente igualdad μ -c.t.p.:

$$a(x) = \limsup_n \frac{f_n(x)}{n} = \lim_{-K \rightarrow -\infty} \tilde{g}(x) = \liminf_n \frac{f_n(x)}{n} = b(x) \quad \forall x \in X.$$

Esto prueba que existe $\tilde{f}(x) = a(x) = b(x) \in [-\infty, +\infty)$ para μ -c.t.p. $x \in X$. Además

$$\lim_{n \rightarrow +\infty} \frac{f_n(x)}{n} = \tilde{f}(x) = \lim_{-K \rightarrow +\infty} \tilde{g}(x) \quad (4.36)$$

(recordar que $\{g_n\}$ depende de $-K$ y por lo tanto $\tilde{g} := \lim_{n \rightarrow +\infty} g_n/n$ depende de $-K$). Ahora probemos que $\int \tilde{f} d\mu = I = \inf \int (f_n/n) d\mu$:

Por convergencia monótona (de las partes negativas, pues las partes positivas son independientes de K ; es decir $f_n^+ = g_n^+$), queda:

$$\begin{aligned} \int \frac{f_n}{n} d\mu &= \int \lim_{-K \rightarrow -\infty} \max \left\{ \frac{f_n}{n}, -K \right\} d\mu = \\ &= \lim_{-K \rightarrow -\infty} \int \max \left\{ \frac{f_n}{n}, -K \right\} d\mu. \end{aligned}$$

Tomando el ínfimo en $n \geq 1$:

$$I = \inf_{n \geq 1} \int \frac{f_n}{n} d\mu = \inf_{n \geq 1} \lim_{-K \rightarrow -\infty} \int \max \left\{ \frac{f_n}{n}, -K \right\} d\mu \quad (4.37)$$

Afirmamos que

$$\inf_{n \geq 1} \lim_{-K \rightarrow -\infty} \int \max \left\{ \frac{f_n}{n}, -K \right\} d\mu = \lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \max \left\{ \frac{f_n}{n}, -K \right\} d\mu. \quad (4.38)$$

En efecto, como para cada n fijo, la sucesión $\left\{ \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \right\}_{K \in \mathbb{N}}$ es decreciente con K , se tiene:

$$\lim_{-K \rightarrow -\infty} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \quad \forall n \geq 1, \forall K.$$

Tomando $\inf_{n \geq 1}$ en ambos lados, tenemos:

$$\inf_{n \geq 1} \lim_{-K \rightarrow -\infty} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \quad \forall K$$

Ahora tomando límite cuando $-K \rightarrow -\infty$:

$$\inf_{n \geq 1} \lim_{-K \rightarrow -\infty} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu.$$

Ahora probemos la desigualdad opuesta. Como

$$\inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \quad \forall n \geq 1 \forall K,$$

tomando límite cuando $-K \rightarrow -\infty$ en ambos miembros, obtenemos

$$\lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \lim_{-K \rightarrow -\infty} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \quad \forall n \geq 1,$$

y tomando ínfimo en $n \geq 1$, resulta

$$\lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu \leq \inf_{n \geq 1} \lim_{-K \rightarrow -\infty} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu.$$

Hemos probado (4.38). Sustituyendo en (4.37), obtenemos:

$$\begin{aligned} I &= \lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \max\left\{\frac{f_n}{n}, -K\right\} d\mu = \\ &= \lim_{-K \rightarrow -\infty} \inf_{n \geq 1} \int \frac{g_n}{n} d\mu = \lim_{K \rightarrow -\infty} \int \tilde{g} d\mu. \end{aligned} \quad (4.39)$$

En la última igualdad usamos la propiedad de existencia μ -c.t.p. de $\tilde{g} = \lim g_n/n$, porque $\{g_n\}$ cumple la hipótesis adicional (4.33) y está en el primer caso. Además, por lo visto en el primer caso, para $\{g_n\}$ las desigualdades (4.34) son igualdades (escribiendo $I_g := \lim_n \int g_n/n d\mu = \inf_n \int g_n/n d\mu$, en lugar de I). Entonces

$$\inf_{n \geq 1} \int \frac{g_n}{n} d\mu = \int \tilde{g} d\mu,$$

lo que justifica la última igualdad en (4.39).

Ahora, en el último término de (4.39), por convergencia monótona cuando $-K \rightarrow -\infty$ en las partes negativas de las funciones \tilde{g} que dependen de K (las partes positivas son independientes de K), deducimos:

$$I = \lim_{K \rightarrow -\infty} \int \tilde{g} d\mu = \int \lim_{K \rightarrow -\infty} \tilde{g} d\mu.$$

Finalmente, en (4.36) probamos que $\lim_{K \rightarrow -\infty} \tilde{g} = \tilde{f}$. Concluimos

$$I = \int \tilde{f} d\mu.$$

Esto termina de demostrar el Teorema 4.2.2, una vez probada la afirmación con las desigualdades (4.34). \square

Ahora solo falta demostrar las desigualdades (4.34) para toda medida invariante μ , y bajo la hipótesis adicional (4.33). Dividiremos la demostración en varias partes, la primera para demostrar la desigualdad de la izquierda en (4.34), y las restantes para demostrar la desigualdad de la derecha.

Paso 5: Prueba de la primera desigualdad (4.34)

Demostración. Recordamos que $I = \lim_{m \geq 1} \int (f_m/m) d\mu$. Entonces, dado $\epsilon > 0$ fijemos $m \geq 1$ tal que

$$\int \frac{f_m}{m} d\mu \leq I + \epsilon.$$

Para cada $n \geq 1$ escribimos $n = mq + r$, $0 \leq r < m$. Por la subaditividad de $\{f_n\}_n$, en cada punto $x \in X$ (denotando $f_0 = 0$), se cumple:

$$\begin{aligned} \frac{f_n}{n} &= \frac{f_{mq+r}}{mq+r} \leq \frac{1}{mq+r} \left(f_m + f_m \circ T^m + \dots + f_m \circ T^{(q-1)m} + f_r \circ T^{mq} \right) \\ \frac{f_n}{n} &\leq \left(\frac{q}{qm+r} \right) \left(\frac{1}{q} \sum_{j=0}^{q-1} f_m \circ (T^m)^j \right) + \left(\frac{mq}{mq+r} \right) \left(\frac{f_r \circ T^{mq}}{mq} \right). \end{aligned} \quad (4.40)$$

Para cada r fijo, $f_r \in L^1(\mu)$ pues $f_r^+ \in L^1(\mu)$ (esto fue demostrado en el paso 2), y además $f_r^- : -\min\{f_r, 0\} \in L^1(\mu)$ porque $-(f_r^-)/r \geq -K$ (c.f. la hipótesis adicional (4.33)). Entonces, por el teorema ergódico de Birkhoff, tomando la transformación T^m con $m \geq 1$ fijo, en vez de T , existe μ -c.t.p. el siguiente límite $\tilde{f}_r(x) = \tilde{f}_r(T^m(x))$:

$$\begin{aligned} \tilde{f}_r(x) &= \lim_{q \rightarrow +\infty} \frac{1}{q} \sum_{j=0}^{q-1} f_r \circ T^{mj}(x), \\ \tilde{f}_r(T^m(x)) &= \lim_{q \rightarrow +\infty} \frac{1}{q} \sum_{j=1}^q f_r \circ T^{mj}(x). \end{aligned}$$

Como $\tilde{f}_r(T^m(x)) = \tilde{f}_r(x)$, resulta:

$$\lim_{q \rightarrow +\infty} \frac{f_r(T^{mq}(x))}{q} - \frac{f_r(x)}{q} = 0 \quad \mu - \text{c.t.p. } x,$$

de donde

$$\lim_{q \rightarrow +\infty} \frac{f_r(T^{mq}(x))}{qm} = 0 \quad \mu - \text{c.t.p. } x, \quad (4.41)$$

para cada $m \geq 1$ y cada $1 \leq r \leq m-1$ fijos que se tomen. Entonces, el último sumando a la derecha en la desigualdad (4.40) tiende a cero cuando $q \rightarrow +\infty$, con m fijo. Más precisamente:

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \frac{f_n(x)}{n} &\leq \limsup_{q \rightarrow +\infty} \left(\frac{q}{qm+r} \right) \left(\frac{1}{q} \sum_{j=0}^{q-1} f_m \circ (T^m)^j \right) + \\ &+ \limsup_{q \rightarrow +\infty} \left(\frac{mq}{mq+r} \right) \left(\frac{f_r \circ T^{mq}}{mq} \right) \leq \\ &\frac{1}{m} \lim_{q \rightarrow +\infty} \left(\frac{1}{q} \sum_{j=0}^{q-1} f_m \circ (T^m)^j \right) + \\ &+ \max_{1 \leq r \leq m-1} \limsup_{q \rightarrow +\infty} \left(\frac{mq}{mq+r} \right) \left(\frac{f_r \circ T^{mq}}{mq} \right), \end{aligned}$$

donde el límite del sumando de la izquierda existe μ -c.t.p. por el teorema ergódico de Birkhoff, y es igual a \tilde{f}_m , mientras que el máximo de los límites superiores del sumando de la derecha es cero debido a la igualdad (4.41). En definitiva, obtenemos:

$$\limsup_{n \rightarrow +\infty} \frac{f_n(x)}{n} \leq \frac{\tilde{f}_m(x)}{m} \quad \mu - c.t.p., \forall \quad m \geq 1,$$

donde \tilde{f}_m es el límite de los promedios de Birkhoff, y es una función en $L^1(\mu)$. Integrando respecto de μ , aplicando el Teorema de Birkhoff (para la transformación T^m): $\int \tilde{f}_m d\mu = \int f_m d\mu$, y recordando la construcción del natural fijo m , se obtiene

$$\int \limsup_{n \rightarrow +\infty} \frac{f_n}{n} d\mu \leq \int \frac{f_m}{m} d\mu \leq I + \epsilon.$$

Como la acotación superior de la integral de la izquierda con $I + \epsilon$, vale para todo $\epsilon > 0$, deducimos

$$\int \limsup_{n \rightarrow +\infty} \frac{f_n}{n} d\mu \leq I,$$

como queríamos demostrar. \square

La segunda desigualdad (4.34):

Recordamos que $I = \inf_{n \geq 1} \int (f_n/n) d\mu$.

Hay que probar que $I \leq \int \liminf_{n \rightarrow +\infty} (f_n/n) d\mu$, bajo la hipótesis adicional $f_n/n \geq -K$ para todo $n \geq 1$, donde $K > 0$ es una constante. Debido a esta hipótesis tenemos $I \geq -K$.

Denotamos:

$$b(x) = \liminf_{n \rightarrow +\infty} \frac{f_n(x)}{n}.$$

Por lo visto en el ejercicio 4.2.1,

$$b(x) = b(T^n(x)) \quad \forall n \geq 1, \quad \mu - c.t.p.$$

Paso 6. Construcción de conjunto y funciones auxiliares. Fijemos $\epsilon > 0$. Para todo natural $k \geq 1$ definimos el siguiente conjunto A_k y las siguientes funciones ψ_k y φ_k :

$$A_k := \{x \in X : \exists j \in \{1, \dots, k-1\} \text{ tal que } \frac{f_j(x)}{j} < b(x) + \epsilon\},$$

$$\psi_k(x) := \begin{cases} b(x) + \epsilon & \text{si } x \in A_k \\ \text{máx}\{b(x) + \epsilon, f_1(x)\} & \text{si } x \notin A_k. \end{cases}$$

$$\varphi_k(x) := (f_1 - \psi_k)^+(x) = \text{máx}\{0, f_1(x) - \psi_k(x)\}.$$

Entonces

$$\lim_{k \rightarrow +\infty} \chi_{A_k}(x) = 1 \quad \forall x \in X. \quad (4.42)$$

y $\lim_{k \rightarrow +\infty} \mu(A_k) = 1$.

Demostración. $A_{k+1} \supset A_k$; entonces:

$$\chi_{A_{k+1}} \geq \chi_{A_k} \quad \mu(A_k) \leq \mu(A_{k+1}) \quad \forall k \geq 1,$$

donde χ_A denota la función característica de un conjunto A . Además

$$\bigcup_{k \geq 1} A_k = X,$$

debido a la definición de límite inferior: para cada punto $x \in X$, existen infinitos valores de $j \geq 1$ tales que $f_j(x)/j < b(x) + \epsilon$. \square

Paso 7. Fijemos cualquier valor de $k \geq 1$, tal que $\mu(A_k) > 0$. Afirmamos que

Para todo $n \geq 1$ y para μ -casi todo punto $x \in X$ existe $m = m(n, x) \in \mathbb{N}$ tal que:

$$\text{máx}\{0, n - k\} \leq m = m(n, x) \leq n - 1, \quad y$$

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^m(x))). \quad (4.43)$$

Demostración. Probaremos esta afirmación por inducción completa en $n \geq 1$. Si $n = 1$, para todo $x \in X$ se cumple

$$f_1(x) = \psi_k(x) + f_1(x) - \psi_k(x) \leq \psi_k(x) + (f_1 - \psi_k)^+(x) = \psi_k(x) + \varphi_k(x).$$

Como $\varphi(x) \geq 0$, tomando $m(1, x) = 0$, queda:

$$f_1(x) \leq \psi_k(x) + \sum_{j=0}^{k-1} \varphi_k(T^j(x)).$$

Concluimos que se verifica la desigualdad (4.43) cuando $n = 1$ para todo $x \in X$. Hemos probado el paso base de la demostración por inducción completa. Ahora demostraremos el paso inductivo.

Fijemos $n \geq 2$. Asumimos por hipótesis de inducción que vale la afirmación (4.43) para todo $x \in X$ y para todo natural positivo menor que n . Probaremos que también vale para n .

Fijemos el punto $x \in X$. Discutamos dos casos: o bien no existe $n_1 \in \{0, 1, \dots, n-1\}$ tal que $T^{n_1}(x) \in A_k$, o bien existe.

Primer caso: No existe $n_1 \in \{0, 1, \dots, n-1\}$ tal que $T^{n_1}(x) \in A_k$. Tenemos, por la subaditividad de la sucesión $\{f_n(x)\}_{n \geq 1}$, la siguiente desigualdad:

$$f_n(x) \leq f_1(x) + f_1(T(x)) + f_1(T^2(x)) + \dots + f_1(T^{n-1}(x)).$$

Como $T^j(x) \notin A_k$, entonces

$$\psi_k(T^j(x)) = \max\{b(T^j(x)) + \epsilon, f_1(T^j(x))\} \geq f_1(T^j(x)).$$

Por lo tanto:

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)).$$

Como la función φ_k es no negativa, si elegimos $m(n, x) = n-1$, resulta:

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^m(x))).$$

Hemos probado la afirmación (4.43) para todos los puntos x que estén en el primer caso.

Segundo caso Existe un menor natural $n_1 \in \{0, 1, \dots, n-1\}$, tal que $T^{n_1}(x) \in A_k$. Primero veamos el subcaso $n_1 > 0$ y luego $n_1 = 0$.

Si $n_1 \geq 1$, por la subaditividad, escribimos:

$$f_n(x) \leq f_{n_1}(x) + f_{n-n_1}(T^{n_1}(x)). \quad (4.44)$$

Como $1 \leq n_1 \leq n-1$, entonces $1 \leq n-n_1 \leq n-1$. Por la hipótesis de inducción, (y la T -invariancia de μ) el sumando de la derecha satisface la afirmación (4.43) para μ -c.t.p. x . Es decir, existe m , que cumple

$$\max\{0, n-n_1-k\} \leq m = m(n-n_1, T^{n_1}(x)) \leq n-n_1-1$$

y tal que

$$f_{n-n_1}(T^{n_1}(x)) \leq \sum_{i=0}^{n-n_1-1} \psi_k(T^i(T^{n_1}(x))) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^m(T^{n_1}(x)))).$$

Como $\max\{n_1, n-k\} \leq m+n_1 \leq n-1$, si elegimos $m' = m'(n, x) = m+n_1$, se satisface

$$\max\{0, n-k\} \leq \max\{n_1, n_k\} \leq m' = m'(n, x) \leq n-1,$$

y además:

$$f_{n-n_1}(T^{n_1}(x)) \leq \sum_{j=n_1}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi(T^j(T^{m'}(x))). \quad (4.45)$$

Por otro lado, el primer sumando del término a la derecha en la desigualdad (4.44) satisface:

$$f_{n_1}(x) \leq \sum_{j=0}^{n_1-1} f_1(T^j(x)). \quad (4.46)$$

Como n_1 es el primer natural tal que $T^{n_1}(x) \in A_k$, deducimos que $T^j(x) \notin A_k$ para todo $0 \leq j \leq n_1 - 1$. Entonces, por la definición de la función ψ_k , tenemos:

$$\psi_k(T^j(x)) = \max\{f_1(T^j(x)), b(T^j(x)) + \epsilon\} \geq f_1(T^j(x)) \quad \forall 0 \leq j \leq n_1 - 1.$$

Sustituyendo en (4.46), queda

$$f_{n_1}(x) \leq \sum_{j=0}^{n_1-1} \psi_k(T^j(x)). \quad (4.47)$$

Juntando las desigualdades (4.44), (4.45) y (4.47) resulta:

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=1}^{k-1} \varphi_k(T^j(T^{m'}(x))),$$

donde $\max\{0, n - k\} \leq m' = m'(n, x) \leq n - 1$. Esto demuestra la afirmación (4.43) en el primer sub-caso $n_1 > 0$.

Tercer caso: Es $n_1 = 0$ el primer natural tal que $T^{n_1}(x) \in A_k$. Por construcción de n_1 sabemos que $T^{n_1}(x) = x \in A_k$. Por definición del conjunto A_k , existe un menor natural $m_1 \in \{1, \dots, k\}$ tal que

$$f_{m_1}(x) < m_1(b(x) + \epsilon). \quad (4.48)$$

Volvemos a discutir nuevamente: o bien $1 \leq m_1 \leq n - 1$, o bien $m_1 = n$, o bien $m_1 > n$.

Si $n_1 = 0$ y $1 \leq m_1 \leq n - 1$, entonces $1 \leq n - m_1 \leq n - 1$. Por la subaditividad y por la hipótesis de inducción, tenemos:

$$\begin{aligned} f_n(x) &\leq f_{m_1}(x) + f_{n-m_1}(T^{m_1}(x)) \leq \\ &\leq f_{m_1}(x) + \sum_{i=0}^{n-m_1-1} \psi_k(T^i(T^{m_1}(x))) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^{m+m_1}(x))), \end{aligned} \quad (4.49)$$

donde $\max\{0, n - m_1 - k\} \leq m = m(n - m_1, T^{m_1}(x)) \leq n - m_1 - 1$. Luego,

$$\max\{0, n - k\} \leq \max\{m_1, n - k\} \leq m + m_1 \leq n - 1.$$

Si definimos $m'(n, x) = m + m_1$, resulta:

$$f_n(x) \leq f_{m_1}(x) + \sum_{j=m_1}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^{m'}(x))), \quad (4.50)$$

donde

$$\max\{0, n - k\} \leq m' = m'(n, x) \leq n - 1.$$

Ahora acotemos por arriba el término $f_{m_1}(x)$ en la desigualdad (4.50). Por construcción de m_1 , este es el primer natural en $\{1, \dots, k\}$ que cumple: $f_{m_1}(x) \leq m_1(b(x) + \epsilon)$. Es decir:

$$f_{m_1}(x) \leq \sum_{j=0}^{m_1-1} (b(x) + \epsilon). \quad (4.51)$$

Recordando lo probado en el ejercicio 4.2.1, $b(x) = b(T^j(x))$ μ -c.t.p., para todo $j \geq 0$. Entonces

$$f_{m_1}(x) \leq \sum_{j=0}^{m_1-1} (b(T^j(x)) + \epsilon) \quad \mu - \text{c.t.p.} \quad (4.52)$$

Por construcción de la función ψ_k , tenemos: $\psi_k(T^j(x)) \geq b(T^j(x)) + \epsilon$ para todo $j \geq 0$, independientemente de si el punto $T^j(x)$ pertenece o no pertenece a A_k . Sustituyendo en (4.52), resulta:

$$f_{m_1}(x) \leq \sum_{j=0}^{m_1-1} \psi_k(T^j(x)). \quad (4.53)$$

Reuniendo las desigualdades (4.50) y (4.53) resulta:

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=1}^k \varphi_k(T^j(T^m(x))),$$

donde $m = m(n, x)$ es tal que $\max\{0, n - k\} \leq m \leq n - 1$. Por lo tanto se satisface la afirmación (4.43) como queríamos demostrar.

Cuarto caso: $n_1 = 0$ y $m_1 = n$. Observamos que las desigualdades (4.51), (4.52) y (4.53), valen cuando $n_1 = 0$ (es decir $x \in A_k$, cualquiera sea el primer natural $m_1 \in \{1, \dots, k\}$ para el cual $f_{m_1}(x) < m_1(b(x) + \epsilon)$). En particular valen cuando $m_1 = n$. En este caso particular, la desigualdad (4.53) afirma que:

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)).$$

Entonces, si elegimos $m = m(n, x) = n - 1$, como $\varphi_k \geq 0$, resulta

$$f_n(x) \leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi_k(T^j(T^m(x))),$$

probando la afirmación (4.43). Finalmente:

Quinto (y último) caso: $n_1 = 0$ y $m_1 > n \geq 1$. Por construcción $m_1 \leq k$. Luego

$$1 < n < k,$$

de donde

$$\max\{0, n - k\} = 0 \leq n - 1.$$

Por la subaditividad, escribimos:

$$f_n(x) \leq f_1(x) + f_1(T(x)) + \dots + f_1(T^{n-1}(x)).$$

Luego

$$\begin{aligned} f_n(x) &\leq \sum_{j=0}^{n-1} f_1(T^j(x)) = \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{n-1} \left(f_1(T^j(x)) - \psi_k(T^j(x)) \right) \leq \\ &\leq \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{n-1} (f_1 - \psi_k)^+(T^j(x)) = \\ &= \sum_{j=0}^{n-1} \psi_k(T^j(x)) + \sum_{j=0}^{k-1} \varphi_k(T^j(x)), \end{aligned}$$

porque $1 < n < k$ y $\varphi_k \geq 0$. Entonces, si elegimos $m = m(n, x) = 0$ tenemos $\max\{0, n - k\} = 0 = m \leq n - 1$, terminando de probar la afirmación (4.43) como queríamos. \square

Paso 8 (último paso). Ahora que terminamos de probar (4.43), vamos a usar esa afirmación para demostrar que

$$\lim_n \int \frac{f_n}{n} d\mu \leq \int (b(x) + \epsilon) d\mu. \quad (4.54)$$

para cualquier medida invariante μ .

Demostración. De la definición de la función ψ_k , observamos que $\varphi_k(x) = \max\{0, f_1(x) - \psi_k(x)\}$ y que $\psi_k(x) \geq b(x) + \epsilon \geq b(x)$ para todo x . Además por la hipótesis adicional de acotación por abajo, sabemos que $f_n(x)/n \geq -K$. Luego, $b(x) = \liminf_n f_n(x)/n \geq -K$, de donde

$$\begin{aligned} 0 \leq \varphi_k(x) &\leq \max\{0, f_1(x) - b(x)\} \leq \\ &\leq \max\{0, f_1(x) + K\} \leq f_1^+(x) + K \quad \forall x \in X. \end{aligned}$$

Concluimos que

$$\varphi_k \in L^1(\mu),$$

pues por hipótesis $f_1^+ \in L^1(\mu)$.

De la afirmación (4.43) con $n, k \geq 1$ fijos, observamos que

$$\text{máx}\{0, n - k\} \leq m(n, x) \leq n - 1.$$

Como $\varphi_k(x) \geq 0$ tenemos:

$$\varphi_k(T^j(T^{m(n,x)}(x))) \leq \sum_{m=\text{máx}\{0, n-k\}}^{n-1} \varphi_k \circ T^j \circ T^m(x) \quad \forall x \in X, \forall j \geq 0.$$

(En efecto, para cada x fijo existe un solo sumando a la derecha que es igual al término de la izquierda: es el sumando que corresponde a $m = m(n, x)$. Luego, como los demás sumandos son no negativos, el término de la izquierda es menor o igual que la suma de la derecha).

Entonces, como $\varphi_k \in L^1(\mu)$, deducimos que, para todos n, k, j fijos:

$$\varphi_k \circ T^j \circ T^{m(n, \cdot)} \in L^1(\mu)$$

y además

$$\int \varphi_k(T^j(T^{m(n,x)}(x))) d\mu \leq \sum_{m=\text{máx}\{0, n-k\}}^{n-1} \int \varphi_k \circ T^j \circ T^m d\mu.$$

Como μ es T invariante, la integral de la derecha es constante al variar j , n y m (con k fijo), siempre igual a $\int \varphi_k d\mu$. Deducimos que

$$\int \varphi_k T^j(T^{m(n,x)}(x)) d\mu \leq k \int \varphi_k d\mu.$$

Integrando la desigualdad (4.43) con $n, k \geq 1$ fijos, usando la desigualdad de arriba y recordando que $\int \psi_k \circ T^j d\mu = \int \psi_k d\mu$, resulta:

$$\begin{aligned} \int \frac{f_n}{n} d\mu &\leq \int \psi_k d\mu + \frac{1}{n} \sum_{j=0}^{k-1} \int \varphi_k T^j(T^{m(n,x)}(x)) d\mu \leq \\ &\leq \int \psi_k d\mu + \frac{k}{n} \sum_{j=0}^{k-1} \int \varphi_k d\mu = \int \psi_k d\mu + \frac{k^2}{n} \int \varphi_k d\mu. \end{aligned} \quad (4.55)$$

Ahora hacemos $n \rightarrow +\infty$ con k fijo, en la desigualdad (4.55), recordando que denotamos $I := \lim_{n \rightarrow +\infty} \int f_n/n d\mu = \inf_{n \geq 1} \int f_n/n d\mu$:

$$I \leq \int \psi_k d\mu = \int \chi_{A_k} \psi_k d\mu + \int (1 - \chi_{A_k}) \psi_k d\mu \quad \forall k \geq 1. \quad (4.56)$$

Recordemos la definición de la función ψ_k , igual a $b(x) + \epsilon$ para todo $x \in A_k$. Deducimos

$$I \leq \int \chi_{A_k} (b(x) + \epsilon) d\mu + \int (1 - \chi_{A_k}) \psi_k d\mu \quad \forall k \geq 1. \quad (4.57)$$

Veamos que $|\psi_k|$ está dominada μ -c.t.p., para todo $k \geq 1$ por una función en $L^1(\mu)$. Como por hipótesis $f_n/n \geq -K$, tenemos $b(x) = \liminf f_n(x)/n \geq -K$. Luego $b^-(x) = -\min\{0, b(x)\} \leq K$. Además, $f_1^- = -\min\{0, f_1\} \leq K$. Por otra parte, $f_1^+ \in L^1(\mu)$ por hipótesis, y

$$b^+(x) \leq \limsup_{n \rightarrow +\infty} \max\{0, f_n(x)/n\} = \limsup_n \frac{f_n^+(x)}{n}.$$

Como la sucesión $\{f_n^+\}_n$ también es subaditiva, tenemos

$$f_n^+ \leq f_1^+ + f_1^+ \circ T + \dots + f_1^+ \circ T^{n-1},$$

de donde deducimos que

$$b^+(x) \leq \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_1^+(T^j(x)) = \tilde{f}_1^+ \quad \mu - \text{c.t.p.},$$

donde $\tilde{f}_1^+ \in L^1(\mu)$ es el límite de los promedios de Birkhoff de f_1^+ . Hemos probado que

$$|b|, |f_1| \in L^1(\mu).$$

Por construcción, la función ψ_k coincide en cada punto x , o bien con $f_1(x)$, o bien con $b(x) + \epsilon$. Ambas funciones f_1 y b están en $L^1(\mu)$. Entonces:

$$|\psi_k| \leq \max\{|f_1|, |b + \epsilon|\} \in L^1(\mu) \quad \forall k \geq 1.$$

Por convergencia dominada, haciendo $k \rightarrow +\infty$ en la desigualdad (4.57), recordando la convergencia puntual $\chi_{A_k}(x) \rightarrow 1$ cuando $k \rightarrow +\infty$ dada por la igualdad (4.42), resulta:

$$I \leq \int (b(x) + \epsilon) d\mu = \epsilon + \int b(x) d\mu,$$

probando la afirmación (4.54) como queríamos. Como $\epsilon > 0$ es arbitrario, deducimos

$$I \leq \int b(x) d\mu,$$

que es la desigualdad de la derecha en (4.34), terminando la demostración del teorema 4.2.2. \square

4.3. Crecimiento subexponencial

Antes de enunciar y demostrar el Teorema Ergódico Multiplicativo en la próxima sección (que incluye como caso particular al Teorema 3.9.8 de Oseledets), veremos una serie de lemas y resultados que son útiles para demostrar, tanto el teorema ergódico multiplicativo, como otros resultados de la teoría ergódica.

En la primera parte de esta sección $T : X \mapsto X$ denota una transformación bimedible en un espacio métrico compacto X , tal que existe alguna medida μ invariante por T .

Definición 4.3.1. Crecimiento subexponencial

Sea $H : X \mapsto \mathbb{R}^+$ medible y sea μ una medida invariante por T . Decimos que ψ tiene crecimiento subexponencial μ -c.t.p., si cumple

$$\lim_{n \rightarrow +\infty} \frac{\log^+ H(T^n(x))}{n} = 0 \quad \text{y} \quad \lim_{n \rightarrow +\infty} \frac{\log^+ H(T^{-n}(x))}{n} = 0,$$

para μ -c.t.p. $x \in X$, donde $\log^+ := \max\{0, \log\}$.

El significado del crecimiento subexponencial es el siguiente: Si se cumplen las igualdades de la definición, entonces, para cualquier $\epsilon > 0$ arbitrariamente pequeño $H(T^n(x)), H(T^{-n}(x)) \leq e^{\epsilon n}$ si n es suficientemente grande (dependiendo de x). Esto significa que, si bien la función H no tiene porqué estar uniformemente acotada a lo largo de las órbitas, si crece, lo hace menos que cualquier exponencial e^{an} que tenga coeficiente positivo a en el exponente, donde $n \geq 1$ es la cantidad de iterados (para el futuro o para el pasado).

Proposición 4.3.2. Criterio de crecimiento subexponencial

Sea $H : X \mapsto \mathbb{R}^+$ medible y sea μ una medida invariante por T . Si

$$\log^+ H \circ T - \log^+ H \in L^1(\mu)$$

entonces H tiene crecimiento subexponencial μ -c.t.p.

Nota: Vamos a usar este criterio de crecimiento subexponencial cuando $\log^+ H \notin L^1(\mu)$. Es decir asumimos que $\int |\log^+ H \circ T - \log^+ H| d\mu < +\infty$, pero podría suceder (aún siendo μ una medida T -invariante), que $\int \log^+ H d\mu = +\infty$.

Demostración. Sea $\psi := \log^+ H \circ T - \log^+ H$. Por la hipótesis, $\psi \in L^1(\mu)$. Luego, por el Teorema Ergódico de Birkhoff, para μ -c.t.p. $x \in X$ existe el límite $\tilde{\psi}(x)$ de los promedios temporales de ψ hacia el futuro:

$$\tilde{\psi}(x) := \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ T^j(x) = \lim_{n \rightarrow +\infty} \left(\frac{\log^+ H \circ T^n(x)}{n} - \frac{\log^+ H(x)}{n} \right),$$

Como para cada x fijo $\lim_n (\log^+ H(x))/n = 0$ tenemos

$$\exists \tilde{\psi}(x) = \lim_{n \rightarrow +\infty} \frac{\log^+ H \circ T^n(x)}{n} \quad \mu - \text{c.t.p. } x \in X. \quad (4.58)$$

Sea $\epsilon > 0$ fijo, arbitrario y sea, para cada $n \geq 1$ el siguiente conjunto:

$$A_n := \{x \in X : \log^+ H(x) \geq n\epsilon\}.$$

La sucesión de conjuntos A_n es decreciente con n ; es decir $A_{n+1} \subset A_n$ para todo $n \geq 1$. Entonces

$$\mu\left(\bigcap_{n \geq 1} A_n\right) = \lim_{n \rightarrow +\infty} \mu(A_n).$$

Como para todo punto x la función $\log^+ H(x)$ toma un valor real, entonces $\bigcap_{n \geq 1} A_n = \emptyset$. Luego

$$\lim_{n \rightarrow +\infty} \mu(A_n) = 0.$$

Luego, existe $N \geq 1$ tal que

$$\mu(A_n) < \epsilon \quad \forall n \geq N.$$

Como μ es invariante por T , tenemos $\mu(T^{-n}A_n) < \epsilon \quad \forall n \geq N$. Escrito de otra forma:

$$\begin{aligned} \mu(\{x \in X : T^n(x) \in A_n\}) &< \epsilon \quad \forall n \geq N, \\ \mu\left(\left\{x \in X : \frac{\log H^+(T^n(x))}{n} \geq \epsilon\right\}\right) &< \epsilon \quad \forall n \geq N. \end{aligned}$$

La desigualdad anterior, por definición, significa que la sucesión de funciones medibles $\log H^+ \circ T^n/n$ converge a cero en medida, para la medida de probabilidad μ . La convergencia en medida implica que existe una subsucesión $n_k \rightarrow +\infty$ para la cual las funciones convergen μ -c.t.p. (ver por ejemplo [Fo 1984, Theorem 2.30]). Entonces, obtenemos

$$\lim_{k \rightarrow +\infty} \frac{\log^+ H \circ T^{n_k}(x)}{n_k} = 0 \quad \mu - \text{c.t.p. } x \in X.$$

Juntando esta afirmación con (4.58) concluimos que

$$\lim_{n \rightarrow +\infty} \frac{\log^+ H \circ T^n(x)}{n} = 0 \quad \mu - \text{c.t.p. } x \in X.$$

Observemos que la hipótesis $\log^+ H \circ T - \log^+ H \in L^1(\mu)$, implica $\log^+ H - \log^+ H \circ T^{-1} \in L^1(\mu)$, porque μ es T invariante. Entonces, $\log^+ H \circ T^{-1} - \log^+ H \in L^1(\mu)$. Aplicando lo demostrado antes a T^{-1} en el lugar de T , se obtiene:

$$\lim_{n \rightarrow +\infty} \frac{\log^+ H \circ T^{-n}(x)}{n} = 0 \quad \mu - \text{c.t.p. } x \in X.$$

Hemos probado que H tiene crecimiento subexponencial μ -c.t.p., terminando la demostración de la Proposición 4.3.2. \square

4.4. Generadores de exponentes de Lyapunov

Desde ahora y hasta el final del capítulo, el espacio donde transcurre la dinámica es una variedad compacta y riemanniana M de dimensión finita, y la transformación, que denotaremos $f \in \text{Diff}^1(M)$, es un difeomorfismo de clase C^1 .

En esta sección definiremos, para los subfibrados medibles del fibrado tangente, las sucesiones de funciones auxiliares ψ_n, φ_n , que son subaditivas y aproximan a los exponentes de Lyapunov. Luego demostraremos que ciertas funciones obtenidas como supremo de estas sucesiones, tienen crecimiento subexponencial (aplicando el criterio de la Proposición 4.3.2). Usaremos el crecimiento subexponencial de los supremos para

demostrar la Proposición 4.4.7 sobre la existencia y mutua coincidencia de los límites de las sucesiones subaditivas de funciones ψ_n y φ_n que aproximan a los exponentes de Lyapunov. Esta propiedad da una condición suficiente para la existencia de exponentes de Lyapunov y subespacios de Oseledets, en el Teorema 4.4.8 y sus Corolarios 4.4.9 y 4.4.10, probados al final de esta sección. El Teorema 4.4.8 y sus Corolarios constituyen una herramienta clave que nos permitirá demostrar el Teorema 3.9.8 de Oseledets en las próximas secciones.

Definición 4.4.1. Fibrados

Sea V un espacio vectorial de dimensión finita. Denotamos con $\mathcal{G}(V)$ el Grassmaniano de V : esto es la familia de todos los subespacios F de V . Se observa que $0 \leq \dim(F) \leq \dim(V)$. En particular $\{0\}, V \in \mathcal{G}(V)$

En $\mathcal{G}(V)$ definimos la métrica:

$$\text{dist}(F^1, F^2) := \max \left\{ \max \{ \|v_1 - \pi_2 v_1\| : v_1 \in F^1, \|v_1\| = 1 \}, \right. \\ \left. \max \{ \|v_2 - \pi_1 v_2\| : v_2 \in F^2, \|v_2\| = 1 \} \right\},$$

donde, para $i = 1, 2$, la aplicación $\pi_i : V \mapsto F^i$ es la proyección ortogonal sobre F^i . Esta métrica en $\mathcal{G}(V)$ define su topología, y esta define la sigma-álgebra de Borel en el Grassmaniano $\mathcal{G}(V)$.

Sea M una variedad compacta y riemanniana de dimensión finita m . Un *fibrado o subfibrado* (del fibrado tangente TM) es una aplicación $F : M \mapsto \bigcup_{x \in M} \mathcal{G}(T_x M)$ tal que $F(x) \in \mathcal{G}(T_x M)$ para todo $x \in M$.

Para cada punto $x \in M$ denotamos con F_x y llamamos fibra en el punto x , al subespacio $F(x) \subset T_x M$. Dado un atlas de M , en cada carta local $\xi : U \subset M \mapsto \xi(U) \subset \mathbb{R}^m$, la fibra F_x se identifica (vía $d\xi$) con un subespacio de \mathbb{R}^m .

Un fibrado F es medible, si para toda carta local $\xi : U \subset M \mapsto \xi(U) \subset \mathbb{R}^m$ la aplicación que a cada $x \in U$ le hace corresponder el subespacio $d\xi F_x \in \mathcal{G}(\mathbb{R}^m)$ es medible.

Punto singular de un subfibrado F es un punto $x \in M$ tal que $F_x = \{0\}$. Denotamos como A_F el conjunto de puntos no singulares:

$$A_F := \{x \in M : F_x \neq \{0\}\}.$$

Sea $f \in \text{Diff}^1(M)$. Un subfibrado F es *df*-invariante si

$$df_x F_x = F_{f(x)} \quad \forall x \in M.$$

Observación: Si F es un fibrado medible, entonces el conjunto de sus puntos singulares, definido por

$$A_F := \{x \in M : F_x = \{0\}\},$$

es medible en M . En efecto, la función $\|\cdot\| : TM \mapsto \mathbb{R}$ definida por $\|(x, v)\| = \|v\|$ es continua. Entonces la preimagen en M del 0 por $\|\cdot\| \circ F$, es un conjunto medible.

Definición 4.4.2. Sea F un subfibrado medible. Definimos las siguientes funciones reales $\psi_n, \varphi_n : M \mapsto \mathbb{R}$, para cada $n \in \mathbb{N}$, llamadas *generadoras de exponentes de Lyapunov*:

$$\psi_n(x) := \begin{cases} \log \sup \left\{ \frac{\|df_x^n(v)\|}{\|v\|} : 0 \neq v \in F_x \right\} & \text{si } F_x \neq \{0\} \\ 0 & \text{si } F_x = \{0\} \end{cases}$$

$$\varphi_n(x) := \begin{cases} -\log \inf \left\{ \frac{\|df_x^{-n}(v)\|}{\|v\|} : 0 \neq v \in F_x \right\} & \text{si } F_x \neq \{0\} \\ 0 & \text{si } F_x = \{0\} \end{cases}$$

Observaciones: El supremo y el ínfimo en las igualdades anteriores son máximo y mínimo respectivamente, pues, siendo $\dim(T_x M) = m$ finito, todo subespacio $F_x \subset T_x M$ es cerrado.

Se observa que $\psi_0 = \varphi_0 = 0$.

Si F es medible, entonces $\psi_n, \varphi_n : M \mapsto \mathbb{R}$ son medibles. En efecto, $df^n : TM \mapsto TM$ es continua. Entonces la preimagen A por $\log(\|df^n \cdot \|/\| \cdot \|)$ de la semirrecta $(-\infty, a]$ (para todo $a \in \mathbb{R}$) es un conjunto medible de TM . La preimagen por F de A es medible en M , porque F es un fibrado medible. Análogamente, la preimagen por F de A^c es medible en M . El conjunto de puntos $x \in M$ tal que $\psi_n(x) \leq a$ (i.e. la preimagen por ψ_n de la semirrecta $(-\infty, a]$) es $F^{-1}(A) \setminus F^{-1}(A^c)$; luego es medible. En forma similar se prueba que φ_n es medible.

Proposición 4.4.3. Sea F un fibrado medible y df -invariante. Sean las funciones ψ_n y φ_n , generadoras de los exponentes de Lyapunov, construidas en 4.4.2. Entonces:

- (a) $\varphi_n(x) = \psi_n(f^{-n}(x))$, $\psi_n(x) = \varphi_n(f^n(x)) \quad \forall n \in \mathbb{N}, \quad \forall x \in X$.
(b) Sea $k \geq 1$ la constante definida por $k = \max_{z \in M} \{1, \|df_z\|, \|df_z^{-1}\|\}$. Para todo $n \in \mathbb{N}$ y para todo $x \in X$, se cumple:

$$|\psi_n(f(x)) - \psi_{n+1}(x)|, \quad |\varphi_n(f^{-1}(x)) - \varphi_{n+1}(x)| \leq \log k. \quad (4.59)$$

$$|\psi_n(x)|, \quad |\varphi_n(x)| \leq n \log k, \quad (4.60)$$

- (c) La sucesión $\{\psi_n\}_{n \geq 1}$ es subaditiva respecto de f y la sucesión $\{\varphi_n\}_{n \geq 1}$ es subaditiva respecto de f^{-1} .

Demostración. Cuando $F_x = \{0\}$, tomando cualquier constante $k \geq 1$, las desigualdades son triviales, pues $\psi_n = \varphi_n = 0$ a lo largo de toda la órbita futura y pasada de x , para todo $n \in \mathbb{Z}$. Cuando $F_x \neq \{0\}$:

- (a) Por la regla de la cadena la transformación lineal $L = df_x^{-n}$ es la inversa de la transformación lineal $L^{-1} = df_{f^{-n}(x)}^n$. Luego, $\inf \|Lv\|/\|v\| = 1 / \sup \|L^{-1}v\|/\|v\|$. Tomando logaritmo y multiplicando por -1 , resulta $\varphi_n(x) = \psi_n(f^{-n}(x))$, probando la primera de las desigualdades de (a). La segunda desigualdad de (a) es consecuencia de la primera, que vale para todo $x \in M$, escribiendo $f^n(x)$ en lugar de x .

- (b) Por la regla de la cadena, tenemos

$$df_{f(x)}^n = df_x^{n+1} df_x^{-1}, \quad df_x^{n+1} = df_{f(x)}^n df_x,$$

$$e^{\psi_n(f(x))} = \sup_{0 \neq v \in F_{f(x)}} \frac{\|df_{f(x)}^n v\|}{\|v\|} \leq \|df_x^{-1}\| \cdot \sup_{0 \neq u \in F_x} \frac{\|df_x^{n+1} u\|}{\|u\|} \leq \leq k e^{\psi_{n+1}(x)},$$

donde $k = \max_{z \in M} \{1, \|df_z\|, \|df_z^{-1}\|\}$. Análogamente:

$$\begin{aligned} e^{\psi_{n+1}(x)} &= \sup_{0 \neq u \in F_x} \frac{\|df_x^{n+1} u\|}{\|u\|} \leq \\ &\leq \|df_x\| \sup_{0 \neq v \in F_{f(x)}} \frac{\|df_x^n v\|}{\|v\|} \leq k e^{\psi_n(f(x))}. \end{aligned}$$

Tomando logaritmo en las dos últimas desigualdades, obtenemos:

$$\psi_n(f(x)) - \psi_{n+1}(x) \leq \log k, \quad \psi_{n+1}(x) - \psi_n(f(x)) \leq \log k,$$

de donde

$$|\psi_n(f(x)) - \psi_{n+1}(x)| \leq \log k \quad \forall n \in \mathbb{Z} \quad \forall x \in M.$$

Hemos probado la primera desigualdad (4.59). Ahora probaremos la segunda: Por la regla de la cadena tenemos

$$df_{f^{-1}(x)}^{-n} v = df_x^{-(n+1)} df_{f^{-1}(x)} v, \quad df_x^{-(n+1)} u = df_{f^{-1}(x)}^{-n} df_x^{-1} u.$$

Tomando norma, denotando $u := df_{f^{-1}(x)} v$, y recordando la definición de la función φ_n , resulta:

$$\begin{aligned} e^{-\varphi_n(f^{-1}(x))} &= \inf\{\|df_{f^{-1}(x)}^{-n} v\|/\|v\| : 0 \neq v \in F_{f^{-1}(x)}\} \geq \\ &(\inf_{0 \neq u \in F_x} \|df_x^{-(n+1)} u\|/\|u\|) (\inf_{0 \neq v \in F_{f^{-1}(x)}} \|df_{f^{-1}(x)} v\|/\|v\|) = \\ &= e^{-\varphi_{n+1}(x)} \left(1 / \sup_{0 \neq u \in F_x} \|df_x^{-1} u\|/\|u\|\right) \geq (1/k) e^{-\varphi_{n+1}(x)}. \end{aligned}$$

Entonces,

$$-\varphi_n(f^{-1}(x)) + \varphi_{n+1}(x) \geq -\log k.$$

Ahora acotemos por el otro lado:

$$\begin{aligned} e^{-\varphi_{n+1}(x)} &= \inf\{\|df_x^{-(n+1)} u\|/\|u\| : 0 \neq u \in F_x\} \geq \\ &(\inf_{0 \neq v \in F_{f^{-1}(x)}} \|df_{f^{-1}(x)}^{-n} v\|/\|v\|) (\inf_{0 \neq u \in F_x} \|df_x^{-1} u\|/\|u\|) = \\ &= e^{-\varphi_n(f^{-1}(x))} \left(1 / \sup_{0 \neq v \in F_{f^{-1}(x)}} \|df_{f^{-1}(x)} v\|/\|v\|\right) \geq (1/k) e^{-\varphi_n(f^{-1}(x))}. \end{aligned}$$

Obtenemos:

$$-\varphi_{n+1}(x) + \varphi_n(f^{-1}(x)) \geq -\log k.$$

Concluimos que

$$|\varphi_{n+1}(x) - \varphi_n(f^{-1}(x))| \leq \log k,$$

terminando de probar las desigualdades (4.59). Ahora probemos (4.60). Argumentando por inducción completa en $n \geq 0$, tenemos $\psi_0 = 0 \leq 0 \cdot \log k$. Aplicando (4.59) para $n \geq 0$, obtenemos

$$\begin{aligned} |\psi_n(x)| &\leq |\psi_{n-1}(f^{-1}(x))| + |\psi_{n-1}(f^{-1}(x)) - \psi_n(x)| \\ &\leq |\psi_{n-1}(f^{-1}(x))| + \log k \leq (n-1) \log k + \log k = n \log k, \end{aligned}$$

(el primer sumando fue acotado usando la hipótesis de inducción). Finalmente, la misma prueba funciona, escribiendo φ_n en lugar de ψ_n , terminando de probar la parte (b).

(c) El siguiente ejercicio guiado se prueba la subaditividad. \square

Ejercicio 4.4.4. Probar la subaditividad de $\{\psi_n\}_{n \geq 1}$ con respecto de f y la subaditividad de $\{\varphi_n\}_{n \geq 1}$ con respecto de f^{-1} . Sugerencia: para la subaditividad de ψ_n usar la regla de la cadena $df_x^{n+m} = df_{f^n(x)}^m df_x^n$, y que el supremo de un producto de dos funciones positivas, es menor o igual que el producto de los supremos. Para la subaditividad de $\{\varphi_n\}_{n \geq 1}$ usar la parte (a) de la Proposición 4.4.3 y la subaditividad de $\{\psi_n\}_{n \geq 1}$.

Límites de las sucesiones ψ_n/n y φ_n/n

Sea F un subfibrado medible df -invariante. Sean ψ_n, φ_n las funciones generadoras de los exponentes de Lyapunov, según la Definición 4.4.2.

Como consecuencia de la subaditividad de las sucesiones $\{\psi_n\}_{n \geq 1}$ (respecto de f) y $\{\varphi_n\}_{n \geq 1}$ (respecto de f^{-1}), y a los resultados de acotación uniforme de $|\varphi_n|/n$, $|\psi_n|/n$, según la Fórmula (4.60) de la Proposición 4.4.3, podemos aplicar el Teorema Ergódico Subaditivo de Kingmann (Teorema 4.2.2). Deducimos que para toda medida f -invariante μ (la cual también es invariante por f^{-1}), existen los siguientes límites μ -c.t.p.:

$$\tilde{\psi}(x) := \lim_{n \rightarrow +\infty} \frac{\psi_n(x)}{n}, \quad \tilde{\varphi}(x) := \lim_{n \rightarrow +\infty} \frac{\varphi_n(x)}{n} \quad \mu - \text{c.t.p.} \quad (4.61)$$

Es decir, el conjunto donde existen $\tilde{\psi}$ y $\tilde{\varphi}$ tiene probabilidad total. En los demás puntos (con probabilidad nula para toda medida invariante), convenimos en definir $\tilde{\psi} = 0$; $\tilde{\varphi} = 0$.

Por lo probado en el Ejercicio 4.2.1, se cumple:

$$\tilde{\psi}(f(x)) = \tilde{\psi}(x), \quad \tilde{\varphi}(f(x)) = \tilde{\varphi}(x) \quad \forall x \in M. \quad (4.62)$$

Denotamos

$$\begin{aligned} \sup \|df|_{F_x}\| &:= \sup\{\|df_x v\|/\|v\| : 0 \neq v \in F_x\}, \\ \inf \|df|_{F_x}\| &:= \inf\{\|df_x v\|/\|v\| : 0 \neq v \in F_x\}. \end{aligned}$$

En resumen, para todo subfibrado F que sea df -invariante y medible, existen los siguientes límites:

$$\begin{aligned}\tilde{\psi}(x) &:= \lim_{n \rightarrow +\infty} \frac{\sup \|df^n|_{F_x}\|}{n} \quad \mu - \text{c.t.p.} \\ \tilde{\varphi}(x) &:= \lim_{n \rightarrow +\infty} \frac{\inf \|df^{-n}|_{F_x}\|}{-n} \quad \mu - \text{c.t.p.}\end{aligned}\tag{4.63}$$

El primer objetivo de esta sección, es demostrar en la Proposición 4.4.7, que $\tilde{\varphi} = \tilde{\psi}$ para μ -c.t.p. $x \in M$, para cualquier medida μ que sea f -invariante.

Definición 4.4.5. Funciones supremo

Sea F un fibrado df -invariante y medible. Sea $\epsilon > 0$. Definimos μ -c.t.p. las siguientes funciones supremo, que dependen de F y de ϵ (además de depender del difeomorfismo f):

$$\begin{aligned}\log H(x) &:= \sup\{\psi_n(x) - n\tilde{\psi}(x) - n\epsilon : n \geq 0\}, \\ \log K(x) &:= \sup\{\varphi_n(x) - n\tilde{\varphi}(x) - n\epsilon : n \geq 0\},\end{aligned}$$

donde

$$\psi_n := \log \sup \|df^n|_{F_x}\|, \quad \varphi_n := -\log \inf \|df^n|_{F_x}\|$$

(en estas últimas fórmulas de ψ_n y φ_n convenimos en que debe tomarse $\sup \|df^n|_{F_x}\|$ cuando $F_x \neq \{0\}$; y que $\psi_n(x) = \varphi_n(x) = 0$ si $F_x = \{0\}$. Luego, para todo punto x tal que $F_x = \{0\}$, se cumple por convención que $H(x) = K(x) = 1$).

Observamos que, como para $n = 0$ son nulos los términos cuyo supremo se buscan, entonces

$$H(x), K(x) \geq 1 \quad \forall x \in M.$$

Además, las funciones supremo son medibles porque son supremo de una sucesión de funciones medibles.

Proposición 4.4.6. Crecimiento subexponencial de supremos

Las funciones supremo H y K tienen crecimiento subexponencial μ -c.t.p. para toda medida de probabilidad f -invariante μ .

Demostración. Aplicaremos el criterio de crecimiento subexponencial de la Proposición 4.3.2. De la fórmula (4.60) sabemos que existe $k \geq 1$ tal que

$$|\tilde{\psi}|, |\tilde{\varphi}| \leq \log k$$

están uniformemente acotadas en valor absoluto por $\log k$, donde $k \geq 1$ es una constante.

Acotemos $|H(f(x)) - H(x)|$:

$$\log H(f(x)) = \sup\{\psi_n(f(x)) - n\tilde{\psi}(f(x)) - n\epsilon : n \geq 0\}.$$

Recordando que $\tilde{\psi}(f(x)) = \tilde{\psi}(x)$ (c.f. Fórmula (4.62)), denotamos simplemente $\tilde{\psi}$ (pues $x \in M$ está fijo y solo nos movemos a lo largo de la órbita de x). Usando la desigualdad (4.59), obtenemos $\psi_n(f(x)) \leq \psi_{n+1}(x) + \log k$. Entonces:

$$\begin{aligned} \log H(f(x)) &\leq \sup\{\psi_{n+1}(x) - (n+1)\tilde{\psi} - (n+1)\epsilon + \tilde{\psi} + \epsilon : n \geq 0\} = \\ &= \log k + \tilde{\psi} + \epsilon + \sup\{\psi_m(x) - m\tilde{\psi} - m\epsilon : m \geq 1\} \leq \\ &\leq 2 \log k + \epsilon + \log H(x). \end{aligned}$$

Hemos probado que

$$\log H(f(x)) - \log H(x) \leq 2 \log k + \epsilon. \quad (4.64)$$

Ahora acotemos esa diferencia por abajo, para probar que está en $L^1(\mu)$:

$$\begin{aligned} \log H(x) &= \sup\{\psi_n(x) - n\tilde{\psi} - n\epsilon : n \geq 0\} = \\ &= \max\left\{0, \sup\{\psi_n(x) - n\tilde{\psi} - n\epsilon : n \geq 1\}\right\} \end{aligned}$$

Para obtener la última igualdad hemos separado el caso $n = 0$ del caso $n \geq 1$, observando que la función cuyo supremo se busca, es nula cuando $n = 0$. Usando la fórmula (4.59), tenemos que $\psi_n(x) \leq \psi_{n-1}(f(x)) + \log k$, para todo $n \geq 1$. Entonces:

$$\begin{aligned} \log H(x) &\leq \max\left\{0, \sup\{\psi_{n-1}(f(x)) + \log k - (n-1)\tilde{\psi} - (n-1)\epsilon + \right. \\ &\quad \left. (-\tilde{\psi} - \epsilon) : n \geq 1\}\right\} = \\ &= \max\left\{0, \log k - \tilde{\psi} - \epsilon + \sup\{\psi_m(f(x)) - m\tilde{\psi} - m\epsilon : m \geq 0\}\right\} \leq \\ &= \max\left\{0, 2 \log k + \epsilon + \log H(f(x))\right\} = 2 \log k + \epsilon + \log H(f(x)), \end{aligned}$$

pues $k, H \geq 1$ y $\epsilon > 0$.

Hemos probado que

$$\log H(x) - \log H(f(x)) \leq 2 \log k + \epsilon. \quad (4.65)$$

Reuniendo las desigualdades (4.64) y (4.65), deducimos que:

$$|\log H(f(x)) - \log H(x)| \leq 2 \log k + \epsilon; \quad \log H \circ f - \log H \in L^1(\mu).$$

Observamos que como $H \geq 1$, entonces $\log H = \log^+ H$. Aplicando el criterio de la Proposición 4.3.2, concluimos que H tiene crecimiento subexponencial. La demostración de que K tiene crecimiento subexponencial es análoga, usando la desigualdad de la derecha en (4.60), y f^{-1} en lugar de f , para concluir que

$$|\log K - \log K \circ f^{-1}| \leq 2 \log k + \epsilon; \quad \log K \circ f^{-1} - \log K \in L^1(\mu),$$

terminando de probar la Proposición 4.4.6. \square

Proposición 4.4.7. Igualdad de las funciones límite

Sean $\tilde{\psi}$, $\tilde{\varphi}$ las funciones definidas en (4.63) para un fibrado F dado, medible y df -invariante. Entonces, para toda medida de probabilidad f -invariante μ , se satisface la siguiente igualdad:

$$\tilde{\psi} = \tilde{\varphi} \quad \mu - \text{c.t.p.}$$

Demostración. Aplicando la regla de la cadena, tenemos, para todo $0 \neq u \in F_x$:

$$u = df_{f^n(x)}^{-n} df_x^n u; \quad 1 = \frac{\|df_{f^n(x)}^{-n} v\|}{\|v\|} \frac{\|df_x^n u\|}{\|u\|},$$

donde $v = df_x^n u$. Tomando logaritmo y dividiendo entre $n \geq 1$:

$$0 = \frac{\log(\|df_{f^n(x)}^{-n} v\|/\|v\|)}{n} + \frac{\log(\|df_x^n u\|/\|u\|)}{n}.$$

Entonces, para todo $n \geq 1$ y para todo $0 \neq u \in F_x$, si denotamos

$$v = df_x^n u \in F_{f^n(x)},$$

se cumple:

$$\frac{\log(\|df_{f^n(x)}^{-n} v\|/\|v\|)}{n} = \frac{-\log(\|df_x^n u\|/\|u\|)}{n}.$$

Luego

$$-\varphi_n(f^n(x)) = \inf \|df_{f^n(x)}^{-n}|_{F_{f^n(x)}}\| \leq \frac{-\log(\|df_x^n u\|/\|u\|)}{n},$$

de donde

$$\frac{\log(\|df_x^n u\|/\|u\|)}{n} \leq \varphi_n(f^n(x)) \quad (4.66)$$

Por la definición de la función supremo $K(f^n(x))$, y por la f -invariancia de la función $\tilde{\varphi}$, tenemos:

$$\varphi_n(f^n(x)) \leq \log K(f^n(x)) + n\tilde{\varphi}(x) + n\epsilon \quad \forall n \geq 0.$$

Sustituyendo en la igualdad (4.66), resulta:

$$\frac{\log(\|df_x^n u\|/\|u\|)}{n} \leq \frac{\log K(f^n(x))}{n} + \tilde{\varphi}(x) + \epsilon \quad \forall u \in F_x \quad \forall n \geq 0.$$

Luego, tomando supremo en el miembro de la izquierda, para $0 \neq u \in F_x$, y recordando la definición de la sucesión subaditiva de funciones ψ_n , resulta:

$$\frac{\psi_n(x)}{n} \leq \frac{\log K(f^n(x))}{n} + \tilde{\varphi}(x) + \epsilon \quad \forall n \geq 0.$$

Debido a que la función K tiene crecimiento subexponencial μ -c.t.p., el primer sumando del segundo miembro tiende a cero cuando $n \rightarrow +\infty$ para μ -c.t.p. $x \in M$. Obtenemos:

$$\tilde{\psi}(x) = \lim_{n \rightarrow +\infty} \frac{\psi_n(x)}{n} \leq \tilde{\varphi}(x) + \epsilon \quad \mu - \text{c.t.p.}$$

Como esta desigualdad vale para todo $\epsilon > 0$ arbitrario, entonces:

$$\tilde{\psi} \leq \tilde{\varphi} \quad \mu - \text{c.t.p.},$$

probando la primera de las desigualdades a demostrar. Ahora probemos la desigualdad contraria. Por la regla de la cadena, para todo $0 \neq u \in F_x$:

$$u = df_{f^{-n}(x)}^n df_x^{-n} u, \quad 1 = \frac{\|df_{f^{-n}(x)}^n v\|}{\|v\|} \frac{\|df_x^{-n} u\|}{\|u\|},$$

donde $v = df_x^{-n} v$. Tomando logaritmo y dividiendo entre $n \geq 1$:

$$0 = \frac{\log(\|df_{f^{-n}(x)}^n v\|/\|v\|)}{n} + \frac{\log(\|df_x^{-n} u\|/\|u\|)}{n}.$$

Entonces, para todo $n \geq 1$ y para todo $0 \neq u \in F_x$, si denotamos

$$v = df_x^{-n} u \in F_{f^{-n}(x)},$$

se cumple:

$$\frac{\log(\|df_{f^{-n}(x)}^n v\|/\|v\|)}{n} = \frac{-\log(\|df_x^{-n} u\|/\|u\|)}{n}. \quad (4.67)$$

Por la definición de la función supremo $H(f^{-n}(x))$, y por la f -invariancia de la función $\tilde{\psi}$, tenemos:

$$\frac{\log(\|df_{f^{-n}(x)}^n v\|/\|v\|)}{n} \leq \log H(f^{-n}(x)) + n\tilde{\psi}(x) + n\epsilon \quad \forall n \geq 0.$$

Sustituyendo en la igualdad (4.67), resulta:

$$\frac{-\log(\|df_x^{-n} u\|/\|u\|)}{n} \leq \frac{\log H(f^{-n}(x))}{n} + \tilde{\psi}(x) + \epsilon \quad \forall u \in F_x \quad \forall n \geq 0.$$

$$\frac{\log(\|df_x^{-n} u\|/\|u\|)}{n} \geq \frac{-\log H(f^{-n}(x))}{n} - \tilde{\psi}(x) - \epsilon \quad \forall u \in F_x \quad \forall n \geq 0.$$

Luego, tomando ínfimo en el miembro de la izquierda, para $0 \neq u \in F_x$, y recordando la definición de la sucesión subaditiva de funciones φ_n , resulta:

$$-\frac{\varphi_n(x)}{n} \geq \frac{-\log H(f^{-n}(x))}{n} - \tilde{\psi}(x) - \epsilon \quad \forall n \geq 0,$$

$$\frac{\varphi_n(x)}{n} \leq \frac{\log H(f^{-n}(x))}{n} + \tilde{\psi}(x) + \epsilon \quad \forall n \geq 0.$$

Debido a que la función H tiene crecimiento subexponencial μ -c.t.p., el primer sumando del segundo miembro tiende a cero cuando $n \rightarrow +\infty$ para μ -c.t.p. $x \in M$. Obtenemos:

$$\tilde{\varphi}(x) = \lim_{n \rightarrow +\infty} \frac{\varphi_n(x)}{n} \leq \tilde{\psi}(x) + \epsilon \quad \mu - \text{c.t.p.}$$

Como esta desigualdad vale para todo $\epsilon > 0$ arbitrario, entonces:

$$\tilde{\varphi} \leq \tilde{\psi} \quad \mu - \text{c.t.p.},$$

probando la segunda de las desigualdades y terminando la demostración de la Proposición 4.4.7. \square

Teorema 4.4.8.

Cotas de exponentes de Lyapunov

Sea $f \in \text{Diff}^1(M)$, sea μ una medida de probabilidad f -invariante y sea F un subfibrado medible y df -invariante. Sea

$$A_F := \{x \in M : F_x \neq \{0\}\}.$$

Asumimos que $\mu(A_F) > 0$.

Entonces, existen los siguientes límites para μ -c.t.p. $x \in A_F$, son funciones medibles del punto x , uniformemente acotadas en valor absoluto por $\log k$, donde $k = \max_{z \in M} \{\|df_z\|, \|df_z^{-1}\|\}$; y satisfacen las siguientes igualdades (a), (b), (d), y la desigualdad (c):

$$(a) \quad \chi_F^+(x) := \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in F_x} \|df_x^n v\| / \|v\|}{n} =$$

$$= \lim_{n \rightarrow +\infty} \frac{\log \inf_{0 \neq v \in F_x} \|df_x^{-n} v\| / \|v\|}{-n}$$

$$(b) \quad \chi_F^-(x) := \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in F_x} \|df_x^{-n} v\| / \|v\|}{-n} =$$

$$= \lim_{n \rightarrow +\infty} \frac{\log \inf_{0 \neq v \in F_x} \|df_x^n v\| / \|v\|}{n}$$

$$(c) \quad \chi_F^-(x) \leq \chi_F^+(x) \quad \mu - \text{c.t.p. } x \in A_F$$

$$(d) \quad \chi_F^+(f(x)) = \chi_F^+(x), \quad \chi_F^-(f(x)) = \chi_F^-(x) \quad \mu - \text{c.t.p. } x \in A_F.$$

Nota: El significado del Teorema 4.4.8 está dado, por un lado, por los Corolarios 4.4.9 y 4.4.10 que dan condiciones suficientes para la existencia y coincidencia de los exponentes de Lyapunov hacia el futuro y hacia el pasado. Por otro lado, el Teorema 4.4.8 será un resultado intermedio esencial para probar, en la próxima sección, la existencia de un primer subespacio de Oseledets.

Como se verá en su prueba, el Teorema 4.4.8 es una simple reformulación de la Proposición 4.4.7.

Demostración. De las definiciones de las funciones $\tilde{\psi}, \tilde{\varphi}$ en las igualdades (4.63), y de la Proposición 4.4.7, se obtiene:

$$\chi_F^+ = \chi_{F,f}^+ = \tilde{\psi} = \tilde{\varphi}, \quad (4.68)$$

probando las dos igualdades de la parte (a). Además la función $\tilde{\psi}$ es medible, porque es el límite μ -c.t.p., y el límite superior en todos los puntos de M , de la sucesión de funciones medibles ψ_n , que fueron construidas en la Definición 4.4.2 para el subfibrado medible F . Entonces χ_F^+ es una función medible. Por la desigualdad (4.60), $|\tilde{\psi}| \leq \log k$. La parte (b) resulta de aplicar la parte (a) a f^{-1} en el rol de f , y multiplicar por -1 , donde:

$$\chi_F^- = \chi_{F,f}^- := \chi_{F,f^{-1}}^+ \quad (4.69)$$

La parte (c) es consecuencia de la siguiente desigualdad:

$$\inf_{0 \neq v \in F_x} \frac{\|df_x^n v\|}{\|v\|} \leq \sup_{0 \neq v \in F_x} \frac{\|df_x^n v\|}{\|v\|} \quad \forall n \geq 0, \forall x \in M.$$

Tomando logaritmo, dividiendo entre $n \geq 1$, haciendo $n \rightarrow +\infty$, y usando una de las igualdades de (a), y otra de (b), se obtiene $\chi_F^- \leq \chi_F^+$, probando (c).

La parte (d) resulta de la igualdad $\chi_F^+(x) = \tilde{\psi}(x)$, y de la igualdad (4.62). Entonces χ_F^+ es una función f -invariante. Observando las igualdades (a) y (b) notamos que $-\chi_{F,f}^-$ para el difeomorfismo f , coincide con $\chi_{F,f^{-1}}^+$ para el difeomorfismo f^{-1} . Entonces $\chi_{F,f}^-$ es también una función f -invariante, terminando de demostrar el Teorema 4.4.8. \square

Corolario 4.4.9. . En las hipótesis del Teorema 4.4.8, si además se cumple

$$\chi_F^+(x) = \chi_F^-(x) \quad \mu\text{-c.t.p. } x \in A_F,$$

entonces para μ -c.t.p. $x \in A_F$ y para todo $0 \neq v \in F_x$ existen los siguientes límites, son funciones medibles del punto $x \in M$, invariantes por f , iguales entre sí, y a $\chi_F^+(x) = \chi_F^-(x)$:

$$\lim_{n \rightarrow +\infty} \frac{\log \|df_x^n v\| / \|v\|}{n} = \lim_{n \rightarrow +\infty} \frac{\log \|df_x^{-n} v\| / \|v\|}{-n}.$$

Nota: Debido al Corolario 4.4.9, cuando $\chi_F^+(x) = \chi_F^-(x)$, este número se denota simplemente como $\chi_F(x)$, y se llama *exponente de Lyapunov para el futuro y para el pasado* en las direcciones del subespacio F_x .

Por lo tanto este Corolario, así como el siguiente, dan condiciones suficientes para la existencia de exponente de Lyapunov, coincidente para el futuro y para el pasado e independiente de la dirección que se tome en el subfibrado F .

Demostración.

$$\frac{\log(\inf \|df_x^n|_{F_x}\|)}{n} \leq \frac{\log(\sup \|df_x^n v\| / \|v\|)}{n} \quad \forall 0 \neq v \in F_x.$$

Luego:

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{\log(\inf \|df_x^n|_{F_x}\|)}{n} &\leq \liminf_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\| / \|v\|)}{n} \quad \forall 0 \neq v \in F_x. \\ \lim_{n \rightarrow +\infty} \frac{\log(\sup \|df_x^n|_{F_x}\|)}{n} &\geq \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\| / \|v\|)}{n} \quad \forall 0 \neq v \in F_x. \end{aligned}$$

Aplicando las partes (a) y (b) del Teorema 4.4.8, las funciones medibles χ_F^+ y χ_F^- cumplen, para todo $0 \neq v \in F_x$, las siguientes desigualdades:

$$\begin{aligned} \chi_F^-(x) &= \lim_{n \rightarrow +\infty} \frac{\log(\inf \|df_x^n|_{F_x}\|)}{n} \leq \liminf_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\|/\|v\|)}{n} \leq \\ &\leq \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\|/\|v\|)}{n} \leq \lim_{n \rightarrow +\infty} \frac{\log(\sup \|df_x^n|_{F_x}\|)}{n} = \chi_F^+(x). \end{aligned} \quad (4.70)$$

Como por hipótesis $\chi_F^- = \chi_F^+$, entonces el \liminf coincide con el \limsup . Es decir, existe el exponente de Lyapunov para el futuro en cualquier dirección $0 \neq v \in F_x$, y coincide con $\chi_F^- = \chi_F^+$. Ahora veamos el límite de Lyapunov para el pasado:

$$\begin{aligned} -\chi_F^+(x) &= \lim_{n \rightarrow +\infty} \frac{\log \inf \|df_x^{-n}|_{F_x}\|}{n} \leq \liminf_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n} v\|/\|v\|)}{n} \leq \\ &\leq \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n} v\|/\|v\|)}{n} \leq \lim_{n \rightarrow +\infty} \frac{\log \sup \|df_x^{-n}|_{F_x}\|}{n} = -\chi_F^-(x). \end{aligned} \quad (4.71)$$

Como por hipótesis $\chi_F^- = \chi_F^+$, entonces el \liminf coincide con el \limsup . Es decir, existe el exponente de Lyapunov para el futuro en cualquier dirección $0 \neq v \in F_x$, es igual al del para el pasado, y coincide con $\chi_F^- = \chi_F^+$. \square

Corolario 4.4.10. *En las hipótesis del Teorema 4.4.8, asumimos además lo siguiente, para μ -c.t.p. $x \in A_F$:*

Hipótesis adicional: *Para todo $0 \neq v \in F_x$, existe el siguiente límite (exponente de Lyapunov hacia el pasado), y es igual a $\chi_F^+(x)$:*

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n} v\|/\|v\|)}{-n} = \chi_F^+(x). \quad (4.72)$$

Asumimos además que la convergencia por abajo al límite (4.72) es uniforme para todo $v \in F_x \setminus \{0\}$; es decir:

Para todo $\epsilon > 0$ existe $N = N(\epsilon, x)$ tal que

$$\frac{\log(\|df_x^{-n} v\|/\|v\|)}{-n} \geq \chi_F^+(x) - \epsilon \quad \forall n \geq N \quad \forall 0 \neq v \in F_x.$$

Entonces, también existe el siguiente límite (exponente de Lyapunov hacia el futuro), y es igual a $\chi_F^+(x)$:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\|/\|v\|)}{n} = \chi_F^+(x),$$

y además

$$\chi_F^-(x) = \chi_F^+(x).$$

Demostración. Por la hipótesis de uniformidad en el convergencia por abajo al límite χ_F^+ , para todo $\epsilon > 0$ existe $N = N(\epsilon, x)$ tal que

$$\frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} \leq -\chi_F^+(x) + \epsilon \quad \forall n \geq N \quad \forall 0 \neq v \in F_x.$$

Entonces tomando supremo en $v \in F_x \setminus \{0\}$ (y observando que el logaritmo del supremo es el supremo de los logaritmos), tenemos:

$$\frac{\log \sup \|df_x^{-n}|_{F_x}\|}{n} \leq -\chi_F^+(x) + \epsilon \quad \forall n \geq N$$

Luego, tomando límite cuando $n \rightarrow +\infty$ y usando la parte (b) del Teorema 4.4.8, deducimos:

$$-\chi_F^-(x) = \lim_{n \rightarrow +\infty} \frac{\log \sup \|df_x^{-n}|_{F_x}\|}{n} \leq -\chi_F^+(x) + \epsilon$$

Deducimos la desigualdad $-\chi_F^-(x) \leq -\chi_F^+(x) + \epsilon$ para todo $\epsilon > 0$. Entonces

$$\chi_F^-(x) \geq \chi_F^+(x),$$

lo cual, junto con la desigualdad de la parte (c) del Teorema 4.4.8, implica

$$\chi_F^-(x) = \chi_F^+(x).$$

Aplicando el Corolario 4.4.9, deducimos que existe el límite de la tesis (el exponente de Lyapunov hacia el futuro en todas las direcciones de F_x), y es igual a $\chi_F^-(x) = \chi_F^+(x)$, como queríamos demostrar. \square

En el siguiente enunciado generalizamos el Teorema 4.4.8 a otras transformaciones lineales T que no son necesariamente el diferencial df .

Definición 4.4.11. Transformaciones lineales de subfibrados

Sea F un subfibrado medible del fibrado tangente (no necesariamente df -invariante) tal que $\dim(F_x) = \dim(F_{f(x)})$ para todo $x \in M$.

Decimos que una aplicación $L : F \mapsto F$ es *lineal invertible* si

$$L_x : F_x \mapsto F_{f(x)}$$

es una aplicación lineal invertible para todo $x \in M$. Denotamos como $L^{-1} : F \mapsto F$ a la aplicación lineal tal que $(L^{-1})_x : F_x \mapsto F_{f^{-1}(x)}$ es la inversa de $L_{f^{-1}(x)} : F_{f^{-1}(x)} \mapsto F_x$ para todo $x \in M$.

Decimos que una aplicación lineal $L : F \mapsto F$ es medible si la aplicación tal que a cada x hace corresponder $L_x \in \mathcal{L}(F_x, F_{f(x)})$ es medible. Esto significa que para todo atlas de cartas locales $\xi_x : U_x \mapsto \xi(U_x) \subset \mathbb{R}^m$ de la variedad diferenciable m -dimensional M , la transformación $x \mapsto d\xi_{f(x)}^{-1} L_x d\xi_x \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ es medible, donde $d = \dim(F_x)$. (Se recuerda que dado un espacio vectorial V de dimensión finita, el espacio $\mathcal{L}(V, V)$ de todas las transformaciones lineales invertibles, es un espacio vectorial normado, y por lo tanto un espacio topológico, donde la norma $\|L\|$ de una transformación lineal

$L \in \mathcal{L}(V, V)$ se define por $\|L\| := \max_{0 \neq v \in V} \|Lv\|/\|v\|$. Luego $\mathcal{L}(V, V)$ es un espacio medible, con la sigma-álgebra de Borel definida a partir de su topología.)

Decimos que la aplicación lineal invertible L es bimedible si L y L^{-1} son medibles. Se puede chequear que si $L : F \mapsto F$ es lineal, invertible y medible, entonces es bimedible.

Teorema 4.4.12. Generalización del Teorema 4.4.8

Sea $f \in \text{Diff}^1(M)$ y sea μ una medida de probabilidad f -invariante. Sea F un subfibrado medible (no necesariamente df -invariante). Sea $L : F \mapsto F$ una aplicación lineal invertible medible.

Sea

$$A_F := \{x \in M : F_x \neq \{0\}\}.$$

Asumimos que $\mu(A_F) > 0$.

Entonces, existen los siguientes límites para μ -c.t.p. $x \in A_F$, son funciones medibles del punto x , uniformemente acotadas en valor absoluto por $\log k$, donde $k = \max_{z \in M} \{\|L_z\|, \|L_z^{-1}\|\}$; y satisfacen las siguientes igualdades (a), (b), (d), y la desigualdad (c):

$$(a) \quad \chi_{F,L}^+(x) := \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in F_x} \|L_x^n v\|/\|v\|}{n} =$$

$$= \lim_{n \rightarrow +\infty} \frac{\log \inf_{0 \neq v \in F_x} \|L_x^{-n} v\|/\|v\|}{-n}$$

$$(b) \quad \chi_{F,L}^-(x) := \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in F_x} \|L_x^{-n} v\|/\|v\|}{-n} =$$

$$= \lim_{n \rightarrow +\infty} \frac{\log \inf_{0 \neq v \in F_x} \|L_x^n v\|/\|v\|}{n}$$

$$(c) \quad \chi_{F,L}^-(x) \leq \chi_{F,L}^+(x) \quad \mu - \text{ c.t.p. } x \in A_F$$

$$(d) \quad \chi_{F,L}^+(f(x)) = \chi_{F,L}^+(x), \quad \chi_{F,L}^-(f(x)) = \chi_{F,L}^-(x) \quad \mu - \text{ c.t.p. } x \in A_F.$$

Demostración. Basta repetir todas las definiciones y demostraciones de esta sección, en particular la del Teorema 4.4.8, reemplazando $df_x|_{F_x} : F_x \mapsto F_{f(x)}$ por la transformación lineal $L_x : F_x \mapsto F_{f(x)}$. \square

4.5. Primer subfibrado de Oseledets

El propósito de esta sección es enunciar el Teorema 4.5.1, que construye un primer subfibrado vectorial E^1 de dimensión no nula del fibrado tangente TM , para el cual los exponentes de Lyapunov en el futuro y en el pasado son iguales (cf. Definición 3.9.2). Este fibrado define entonces un subespacio no nulo $E_x^1 \subset T_x M$ para casi todo punto $x \in M$ que constituye un primer subespacio de Oseledets. El Teorema 4.5.1 es la esencia para demostrar en forma recursiva el Teorema 3.9.8 de Oseledets.

Teorema 4.5.1.**Primer subespacio de Oseledets.**

Sea $f \in \text{Diff}^1(M)$. Sea μ una medida f -invariante. Sea E un subfibrado vectorial medible de TM , invariante con df , y tal que

$$\dim(E_x) \geq 1$$

para un conjunto de puntos $x \in M$ con medida μ positiva.

Entonces, existe un subfibrado vectorial $E^1 \subset E$ medible y df -invariante (llamado primer subespacio de Oseledets) tal que:

(a) **No nulidad del primer subfibrado de Oseledets E^1 :**

$$\dim(E_x^1) \geq 1 \text{ para } \mu\text{-c.t.p. tal que } E_x \neq \{0\}.$$

(b) **Existencia del exponente de Lyapunov χ_1 en E^1 :**

Para μ -c.t.p. tal que $E_x \neq \{0\}$, para todo $0 \neq v \in E_x^1$, existen los siguientes límites (exponentes de Lyapunov $\chi_1^+(x)$ y $\chi_1^-(x)$, para el futuro y para el pasado respectivamente, en todas las direcciones de E_x^1), son independientes de la dirección v , e iguales entre sí:

$$\chi_1^+(x) = \lim_{n \rightarrow +\infty} \frac{\log(\|df^n(v)\|/\|v\|)}{n} = \chi_1^-(x) = \lim_{n \rightarrow +\infty} \frac{\log(\|df^{-n}(v)\|/\|v\|)}{-n}$$

(c) **Maximalidad de χ_1 :**

El exponente de Lyapunov $\chi_1^+(x) = \chi_1^-(x)$ en las direcciones de E_x^1 (que denotamos como $\chi_1(x)$), satisface:

$$\chi_1(x) := \chi_E^+(x) = \chi_{E^1}^-(x) = \chi_{E^1}^+(x),$$

y es una función medible y f -invariante.

Nota: Esta propiedad es de maximalidad (respecto al fibrado dado E) del exponente de Lyapunov χ_1 en el primer subfibrado de Oseledets E^1 , porque por la parte (a) del Teorema 4.4.8, cuando existen exponentes de Lyapunov para el futuro o para el pasado para vectores no nulos de E (es decir existen $\lim_{n \rightarrow +\infty} \frac{\log \|df^n v\|}{n}$ y $\lim_{n \rightarrow -\infty} \frac{\log \|df^n v\|}{n}$), estos deben ser siempre menores o iguales a χ_E^+ - c.f. Desigualdades (4.70) y (4.71).

(d) **Subespacio complementario de E^1 :**

Sea G el complemento de E^1 en E , definido por

$$G_x = \{0\} \cup (E_x \setminus E_x^1)$$

para μ -c.t.p $x \in M$ tal que $E_x \neq \{0\}$ y definido por $G_x = \{0\}$ para los demás puntos $x \in M$. Entonces G es invariante por df y para todo subfibrado vectorial medible $F^1 \subset G$ invariante con df , y para μ -c.t.p. x tal que $F_x^1 \neq \{0\}$, se verifica la siguiente desigualdad:

$$\sup_{0 \neq v \in F_x^1} \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\|/\|v\|)}{n} \leq \chi_{F^1}^+(x) < \chi_1(x) = \chi_E^+(x).$$

Nota: La afirmación (d) significa que la propiedad de maximalidad del exponente de Lyapunov $\chi_1(x) = \chi_E^+(x)$ en el primer subespacio E_x^1 de Oseledets, es *exclusiva* de E_x^1 . Es decir, si existiera algún otro subfibrado F^1 invariante con dimensión no nula, contenido en el complemento de E^1 , entonces $\chi_{F^1}^+$ debe ser estrictamente menor que $\chi_1(x)$.

(e) Subfibrado complementario de E^1 :

Existe un subfibrado $F^1 \subset G$ medible y df -invariante (llamado subfibrado complementario del primer subfibrado de Oseledets), tal que μ -c.t.p. se cumple:

$$E_x^1 \subsetneq E_x \Rightarrow F_x^1 \neq \{0\}, \quad E_x^1 \oplus F_x^1 = E_x.$$

Nota: La propiedad (e) es la que permite aplicar un proceso inductivo para demostrar el Teorema 3.9.8 de Oseledets. Al subfibrado complementario F^1 podemos aplicar este mismo Teorema 4.5.1: existirá un subfibrado de Oseledets $E^2 \subset F^1$, con exponente de Lyapunov $\chi_2 = \chi_{F^1}^+ < \chi_1$, etc, y así sucesivamente, hasta agotar la dimensión finita de la variedad M . Por lo tanto, la demostración de este Teorema 4.5.1 es en esencia, la del Teorema de Oseledets.

La demostración del Teorema 4.5.1 se encuentra en el Apéndice del Capítulo 10 y está basada en el Lema 10.2.2 de Pliss [Pl 1972].

Corolario 4.5.2. Sea M una variedad compacta riemanniana con dimensión finita. Sea $f \in \text{Diff}^1(M)$. Para μ -casi todo punto débilmente regular $x \in M$ existe el siguiente límite y es igual al mayor exponente de Lyapunov $\chi^+(x)$ en las direcciones del espacio tangente $T_x M$:

$$\chi^+(x) = \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in T_x M} \|df^n(v)\|/\|v\|}{n}.$$

Demostración. Tomando como fibrado E medible y df -invariante a todo el fibrado tangente $E = TM$, y aplicando la parte (A) del Teorema 4.4.8, existe el límite siguiente y es igual a $\chi_{TM}^+(x)$ para μ -c.t.p. $x \in M$:

$$\chi_{TM}^+(x) = \lim_{n \rightarrow +\infty} \frac{\log \sup_{0 \neq v \in T_x M} \|df^n(v)\|/\|v\|}{n}. \tag{4.73}$$

La función logaritmo es creciente, y por la definición de supremo, tenemos:

$$\frac{\log(\|df^n(v)\|/\|v\|)}{n} \leq \frac{\log \sup_{0 \neq v \in T_x M} \|df^n(v)\|/\|v\|}{n} \quad \forall 0 \neq v \in T_x M. \tag{4.74}$$

Por hipótesis el punto x es débilmente regular. Aplicando la Definición 3.9.2 de punto débilmente regular, para todo $0 \neq v \in T_x M$ existe el siguiente límite, llamado exponente de Lyapunov hacia el futuro en la dirección de v :

$$\chi^+(x, v) = \lim_{n \rightarrow +\infty} \frac{\log(\|df^n(v)\|/\|v\|)}{n}$$

Juntando esta igualdad con la desigualdad (4.74) y la igualdad (4.73), resulta:

$$\begin{aligned}\chi^+(x, v) &= \lim_{n \rightarrow +\infty} \frac{\log(\|df^n(v)\|/\|v\|)}{n} \leq \\ &\leq \frac{\log \sup_{0 \neq v \in T_x M} \|df^n(v)\|/\|v\|}{n} = \chi_{TM}^+(x).\end{aligned}$$

Entonces, el mayor exponente de Lyapunov a lo largo de la órbita de x es menor o igual que $\chi_{TM}^+(x)$. Pero por la partes (a), (b), (c) del Teorema 4.5.1, existen direcciones $0 \neq v \in E_x^1$ (en el primer subespacio de Oseledets), tales que el exponente de Lyapunov en la dirección v coincide con $\chi_{TM}^+(x)$. Concluimos que $\chi_{TM}^+(x)$ es igual al mayor exponente de Lyapunov en las direcciones de $T_x M$. \square

4.6. Teorema Ergódico Multiplicativo (Oseledets)

Recordamos la definición 3.9.6 de puntos Lyapunov regulares y el enunciado del Teorema 3.9.8 de Oseledets:

El conjunto R de puntos Lyapunov regulares para f es medible y tiene probabilidad total. Además, los exponentes de Lyapunov son funciones medibles del punto $x \in R$. Vamos a generalizar este resultado introduciendo la siguiente definición:

Definición 4.6.1. .

Puntos Lyapunov regulares para un fibrado

Sea E un subfibrado vectorial medible y df -invariante de TM ; es decir:

$$df_x(E_x) = E_{f(x)} \quad \forall x \in M,$$

tal que $\dim(E_x) \geq 1$ para un conjunto de puntos $x \in M$ con probabilidad total.

Decimos que $x \in M$ es un *punto Lyapunov regular para el fibrado E* si existe un splitting

$$E_x = \bigoplus_{i=1}^{k(x)} E_x^i$$

en $k(x) \geq 1$ subfibrados E_x^i tales que para todo $0 \neq v \in E_x^i$ existen, reales, los siguientes límites para el futuro y para el pasado, y coinciden entre sí:

$$\lim_{n \rightarrow \pm\infty} \frac{\log \|df_x^n v\|}{n} = \chi_i(x)$$

Además, $\chi_i(x) \neq \chi_j(x)$ para todo $i \neq j \in \{1, \dots, k(x)\}$. Entonces siempre se pueden ordenar los subfibrados de forma que

$$\chi_1(x) > \chi_2(x) \dots > \chi_{k(x)}(x).$$

Los números $\chi_i(x)$ se llaman *exponentes de Lyapunov* para el subfibrado E , de (la órbita del) punto Lyapunov regular x . Los subespacios E_x^i se llaman *subespacios de Oseledets* de E_x .

Teorema 4.6.2.**Teorema Ergódico Multiplicativo (versión diferenciable)**

Sea $f \in \text{Diff}^1(M)$ un difeomorfismo en una variedad compacta riemanniana M . Sea E un subfibrado vectorial df -invariante, medible y tal que $\dim(E) \geq 1$ para un conjunto de puntos con probabilidad total.

Entonces:

- (a) El conjunto R de puntos Lyapunov regulares para el fibrado E es medible, y los exponentes de Lyapunov son funciones medibles del punto $x \in R$
- (b) El conjunto R de puntos Lyapunov regulares para el fibrado E tiene probabilidad total: es decir $\mu(R) = 1$ para toda medida de probabilidad μ que sea f -invariante.

Antes de demostrar el Teorema Ergódico Multiplicativo 4.6.2, veamos su Corolario inmediato:

Corolario 4.6.3. Teorema de Oseledets.

Sea $f \in \text{Diff}^1(M)$. Entonces, el conjunto de puntos Lyapunov regulares (para el fibrado tangente TM) es medible, tiene probabilidad total, y las funciones exponentes de Lyapunov son medibles.

Demostración. Basta aplicar el Teorema 4.6.2 al caso particular en que el fibrado medible E invariante dado es TM , es decir el subespacio E_x es todo espacio tangente T_xM para todo $x \in M$. \square

4.6.4. Demostración del Teorema Ergódico Multiplicativo 4.6.2

Demostración. En realidad la verdadera dificultad en la demostración del Teorema de Ergódico Multiplicativo fue resuelta en la prueba del Teorema 4.5.1 de la sección anterior, en que probamos la existencia de un primer subespacio de Oseledets con exponente de Lyapunov máximo, y la de su subespacio invariante complementario. Ahora solo resta mostrar cómo se pega el Teorema 4.5.1 con el enunciado del Teorema Ergódico Multiplicativo:

Paso 1. Aplicando el Teorema 4.5.1, existe un primer subfibrado medible $E^1 \subset E$ definido para todo $x \in M$ (primer subfibrado de Oseledets), que es no nulo μ -c.t.p. para cualquier $\mu \in \mathcal{M}_f$, que es df -invariante y tal que para todo $0 \neq v \in E_x^1$ (para μ -c.t.p. $x \in M$) existe, y no depende de la dirección v , el exponente de Lyapunov $\chi_1(x)$. Además el Teorema 4.5.1 afirma que existe un subfibrado medible F^1 , complementario E^1 en el fibrado dado E ($E_x^1 \oplus F_x^1 = E_x$ para todo $x \in M$), df -invariante y tal que para todo $0 \leq v \in F_x^1$ (para μ -c.t.p. $x \in M$ para toda $\mu \in \mathcal{M}_f$), si existe exponente de Lyapunov, entonces este es estrictamente menor que $\chi_1(x)$.

Como E^1 es subfibrado medible, el conjunto A^1 de puntos x donde $E_x^1 \neq \{0\}$, es medible.

Además el teorema 4.5.1 demuestra que el primer exponente de Lyapunov χ_1 es una función medible.

En todo punto $x \in M$ sucede uno de los siguientes dos casos: o bien $E_x^1 = E_x$, o bien $E_x^1 \neq E_x$. En el primer caso, el punto x satisface la definición de ser Lyapunov regular para el fibrado dado E , con un splitting de Oseledets trivial $E_x^1 = E_x$ y un único exponente de Lyapunov. El conjunto de estos puntos es medible, pues es aquel

donde el subfibrado medible complementario F^1 es nulo (ya que $E_x^1 \oplus F_x^1 = E_x$ para todo $x \in M$). Denotando $B_1 = \{x \in M : F_x^1 \neq \{0\}\}$, en resumen tenemos:

$$\forall x \in A_1 \setminus B_1, \quad x \text{ es Lyapunov regular}$$

con un splitting

$$E_x = E_x^1, \quad \dim(E_x^1) \geq 1,$$

y un único exponente de Lyapunov χ_1 .

Además:

$$A_1, B_1 \text{ son medibles y } \mu(A_1) = 1 \quad \forall \mu \in \mathcal{M}_f.$$

Paso 2. En el segundo caso, si el conjunto de puntos $x \in B_1 \subset A_1$ donde $F_x^1 \neq \{0\}$ tiene probabilidad nula (es decir $\mu(B_1) = 0 \quad \forall \mu \in \mathcal{M}_f$), entonces no hay más nada que demostrar ya que el conjunto R de puntos Lyapunov regulares es A_1 .

Si existe alguna medida de probabilidad $\mu \in \mathcal{M}_f$ tal que $\mu(B_1) > 0$, tomamos ahora el subfibrado F^1 en el rol de E . Aplicando de nuevo el Teorema 4.5.1 a F^1 , en vez de E , existe, para todo $x \in B^1$, un subfibrado medible $E^2 \subset F^1$ y df -invariante (segundo subfibrado de Oseledets), tal que para μ -c.t.p. $x \in B^1$ es $E^2 \neq \{0\}$ para toda $\mu \in \mathcal{M}$ para la cual $\mu(B_1) > 0$. Además, el Teorema 4.5.1 afirma existe el exponente de Lyapunov χ_2 en toda dirección no nula de E^2 μ -c.t.p. y que χ_2 es una función medible.

El conjunto

$$A_2 \subset B_1$$

de puntos $x \in B_1$ donde $E^2 \neq \{0\}$ es medible, pues el subfibrado E^2 es medible y tiene probabilidad 1 para toda $\mu \in \mathcal{M}_f$ para la cual $\mu(B_1) > 0$. Por la propiedad demostrada en el Teorema 4.5.1 para el complementario F^1 del primer subfibrado de Oseledets:

$$\chi_1 > \chi_2$$

(porque $E^2 \subset F^1$). Además, el Teorema 4.5.1 afirma la existencia del subfibrado medible F^2 , complementario de E^2 en F^1 (i.e. $F^1 = E^2 \oplus F^2$, definido para todo punto $x \in B_1$, tal que si para alguna dirección no nula de F^2 existe el exponente de Lyapunov, entonces éste exponente debe ser estrictamente menor que χ_2).

Distinguimos dos casos para todo punto $x \in B_1$: o bien $E^2 = F^1$, o bien $E^2 \neq F^1$. En el primer caso $F_x^2 = \{0\}$ ya que $E_x^2 \oplus F_x^2 = F_x^1$. El subconjunto de los puntos $x \in B_1$ en este primer caso, es medible, ya que es el complemento en el conjunto medible B_1 del conjunto $B_2 := \{x \in B_1 : F_x^2 \neq \{0\}\}$, y F^2 , por el Teorema 4.5.1, es un subfibrado medible. En resumen, en este primer caso del segundo paso, tenemos:

$$\forall x \in (A_2 \setminus B_2) \subset B_1, \quad x \text{ es Lyapunov regular}$$

con splitting

$$E_x = E_x^1 \oplus E_x^2 \quad \dim(E_x^{1,2}) \geq 1$$

y dos exponentes de Lyapunov respectivos

$$\chi_1 > \chi_2.$$

Además:

$$A_2, B_2 \text{ son medibles y } \mu(A_2) = \mu(B_1) \quad \forall \mu \in \mathcal{M}_f.$$

Paso $(i + 1)$ -ésimo En el segundo caso del paso anterior i -ésimo, si el conjunto medible $B_i = \{x \in M : F_x^i \neq \{0\}\}$ tiene medida μ nula para toda $\mu \in \mathcal{M}_f$, no hay más nada que probar ya que el conjunto R de puntos Lyapunov regulares es

$$(A_1 \setminus B_1) \cup (A_2 \setminus B_2) \cup \dots \cup (A_i \setminus B_i)$$

Si $\mu(B_i) > 0$ para alguna medida $\mu \in \mathcal{M}_f$ tomamos el subfibrado F^i , complementario del subespacio de Oseledets E^i construido en el paso anterior i -ésimo. Usando F^i en el rol de E para aplicar el Teorema 4.5.1, deducimos que existe, para todo $x \in B^i$, un subfibrado medible $E^{i+1} \subset F^i$, df -invariante ($i + 1$ -ésimo subfibrado de Oseledets), tal que para μ -c.t.p. $x \in B^i$ es $E^{i+1} \neq \{0\}$ para toda $\mu \in \mathcal{M}$ para la cual $\mu(B_i) > 0$. Además, el Teorema 4.5.1 afirma existe el exponente de Lyapunov χ_{i+1} en toda dirección no nula de E^{i+1} μ -c.t.p. y que χ_{i+1} es una función medible.

El conjunto

$$A_{i+1} \subset B_i$$

de puntos $x \in B_i$ donde $E^{i+1} \neq \{0\}$ es medible, pues el subfibrado E^{i+1} es medible y tiene probabilidad 1 para toda $\mu \in \mathcal{M}_f$ para la cual $\mu(B_i) > 0$. Por la propiedad demostrada en el Teorema 4.5.1 para el complementario F^i del subfibrado de Oseledets anterior:

$$\chi_i > \chi_{i+1}$$

Además, el Teorema 4.5.1 afirma la existencia del subfibrado medible F^{i+1} , complementario de E^{i+1} en F^i (i.e. $F^i = E^{i+1} \oplus F^{i+1}$, definido para todo punto $x \in B_i$). El conjunto de puntos $x \in B_i$ tal que $E^{i+1} = F^i$, es medible, ya que es el complemento en el conjunto medible B_i de $B_{i+1} := \{x \in B_i : F_x^{i+1} \neq \{0\}\}$, y F^{i+1} , por el Teorema 4.5.1, es un subfibrado medible. En resumen, en este paso $i + 1$, tenemos:

$$\forall x \in (A_{i+1} \setminus B_{i+1}) \subset B_i, \quad x \text{ es Lyapunov regular}$$

con splitting

$$E_x = E_x^1 \oplus E_x^2 \dots \oplus E_x^i \oplus E_x^{i+1},$$

$$\dim E_x^{1, \dots, i, i+1} \geq 1,$$

e $i + 1$ exponentes de Lyapunov respectivos, todos diferentes:

$$\chi_1 > \chi_2 > \dots > \chi_i > \chi_{i+1}.$$

Además:

$$A_{i+1}, B_{i+1} \text{ son medibles y } \mu(A_{i+1}) = \mu(B_i) \quad \forall \mu \in \mathcal{M}_f.$$

Paso final Como en cada paso se construye un subespacio de Oseledets nuevo, L.I. con los anteriores, y siempre de dimensión mayor o igual que 1, pero la variedad M tiene dimensión finita, el proceso anterior debe terminarse después un número finito k de pasos. Concluimos que:

El conjunto R de puntos Lyapunov regulares es la siguiente unión de conjuntos disjuntos dos a dos:

$$R = (A_1 \setminus B_1) \cup (A_2 \setminus B_2) \cup \dots \cup (A_i \setminus B_i) \cup \dots \cup A_k.$$

R es medible pues cada A_i y cada B_i lo es. R tiene probabilidad total, pues $A_{i+1} \subset B_i$, y para toda μ -invariante $\mu(A_{i+1} \setminus B_{i+1}) = 1$. (Más precisamente, se observa que $R = A_1$ y ya desde primer paso teníamos que $\mu(A_1) = 1$).

Las funciones exponentes de Lyapunov χ_i obtenidas en cada paso, son medibles por aplicación del Teorema 4.5.1. Con esto termina la demostración del Teorema Ergódico Multiplicativo 4.6.2. \square

Capítulo 5

Atractores topológicos y ergódicos

5.1. Atractores topológicos

A lo largo de esta sección $f : X \mapsto X$ es una transformación continua en un espacio métrico compacto X .

Definición 5.1.1. Estabilidad Lyapunov y orbital.

Sea $K \subset X$ un conjunto compacto no vacío invariante por f (es decir $f^{-1}(K) = K$.) K se dice *orbitalmente estable* (hacia el futuro) si para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$\text{dist}(p, K) < \delta \Rightarrow \text{dist}(f^n(p), K) < \epsilon \quad \forall n \geq 0.$$

K se dice *Lyapunov estable* (hacia el futuro) si para todo $\epsilon > 0$ existe $\delta > 0$ tal que

$$\text{dist}(p, K) < \delta \Rightarrow \exists q \in K \text{ tal que } \text{dist}(f^n(p), f^n(q)) < \epsilon \quad \forall n \geq 0.$$

De las definiciones anteriores se deduce que si K es Lyapunov estable, entonces es orbitalmente estable. Sin embargo el recíproco no es cierto en general.

Definición 5.1.2. Atractor topológico I

Un *atractor topológico* es un conjunto K compacto y no vacío tal que

1) $K = f^{-1}(K) = f(K)$.

2) Existe un abierto $V \supset K$, llamado *cuenca local de atracción topológica* de K , tal que

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0 \quad \forall x \in V. \tag{5.1}$$

Consideramos, de particular importancia, los atractores topológicos *que sean orbitalmente estables* (para los cuales daremos una definición equivalente en 5.1.10).

Un atractor topológico K se llama *minimal* (como atractor topológico) si el único compacto no vacío $K' \subset K$ que cumple 1) y 2) es $K' = K$.

Observación 5.1.3. Muchos autores requieren además de las condiciones 1) y 2), que el compacto no vacío K sea orbitalmente estable o Lyapunov estable, para llamarse atractor topológico. La condición de estabilidad orbital en la Definición 5.1.2 no es redundante con las otras dos condiciones 1) y 2). En el ejemplo 5.1.8 parte (B), se muestra un conjunto compacto $\{p_0\}$ formado por un solo punto fijo p_0 que cumple las condiciones 1) y 2) de la definición 5.1.2, pero que no es orbitalmente estable.

Ejemplo: En el capítulo 1 vimos que toda órbita periódica hiperbólica tipo pozo (de un difeomorfismo $f : M \mapsto M$ en una variedad M), es un atractor topológico.

Ejercicio 5.1.4. Probar que una órbita periódica hiperbólica tipo pozo es un atractor topológico Lyapunov estable.

Ejercicio 5.1.5. (a) Sea $f : M \mapsto M$ un homeomorfismo transitivo en una variedad compacta M . Probar que M es un atractor topológico y que es minimal (como atractor topológico). Deducir que M es el único atractor topológico. Sugerencia: De la transitividad deducir que existe alguna órbita futura densa en M . Para probar que M es minimal como atractor topológico y el único atractor asumir que $K \subset M$ es un atractor topológico. Su cuenca local de atracción topológica V contiene algún punto x de una órbita futura densa. Probar que la órbita futura de x es densa. Usando la definición de atractor, probar que el omega-límite de x para cualquier $x \in V$ debe estar en K . Deducir que $K = M$.

(b) Probar que el único atractor topológico del autormorfismo lineal $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro \mathbb{T}^2 es todo el toro.

Ejercicio 5.1.6. Sea $K \subset X$ un atractor topológico. Sea en K la topología inducida por su inclusión en el espacio métrico X . Suponga que existe $x \in K$ tal que la órbita futura de x es densa en K . Demostrar que K es minimal como atractor topológico.

Ejercicio 5.1.7. Sea $X \subset \mathbb{R}^2 \sim \mathbb{C}$ un disco compacto de centro en el origen y radio $1 < r < 2$. Consideremos en X coordenadas polares $p = \rho e^{i\varphi} \in X : (\rho, \varphi) : 0 \leq \rho \leq r, 0 \leq \varphi < 2\pi$. Sea $f : X \mapsto X$ el homeomorfismo dado por las siguientes ecuaciones

$$f(p) = f(\rho e^{i\varphi}) = \rho * e^{i\varphi*}, \text{ donde}$$

$$\rho^* = \frac{\rho(4 - \rho)}{3}, \quad \varphi^* = \varphi + a$$

donde a es una constante $0 \leq a < 2\pi$.

(a) Dibujar algunas de las órbitas y probar que la circunferencia K de centro en el origen y radio 1 es un atractor topológico Lyapunov estable. (Sugerencia: ver la demostración de que K es atractor topológico en el ejemplo 5.1.8).

(b) Probar que si $a/(2\pi)$ es irracional entonces K es minimal como atractor topológico. (Sugerencia: usar el ejercicio 5.1.6)

(c) Si $a/(2\pi)$ es racional, ¿es K minimal como atractor topológico?

Ejemplo 5.1.8.

Sea $X \subset \mathbb{R}^2 \sim \mathbb{C}$ el disco compacto de centro en el origen y radio $1 < r < 2$, como en el ejercicio 5.1.7.

(A) Sea, en coordenadas polares, el siguiente homeomorfismo f :

$$f(p) = f(\rho e^{i\varphi}) = \rho^* e^{i\varphi^*}, \text{ donde}$$

$$\rho^* = \frac{\rho(4-\rho)}{3}, \quad \varphi^* = \varphi + (\rho - 1)$$

Tenemos $\rho^* - 1 = (3 - \rho)(\rho - 1)/3$. Luego $|\rho^* - 1| \leq \lambda|\rho - 1|$ si $\rho > 3 \cdot (1 - \lambda)$ donde $0 < \lambda < 1$. Entonces, por inducción en n se deduce que $\lim_{n \rightarrow +\infty} |\rho^{(n)} - 1| = 0$ donde $\rho^{(n)}$ es la distancia al origen de $f^n(p)$ para cualquier $p \neq (0, 0)$. Deducimos que la distancia de $f^n(p)$ a la circunferencia K de centro 0 y radio 1 tiende a cero con $n \rightarrow +\infty$, para cualquier punto inicial $0 \neq p \in X$. Además $\rho^* = 1$ si $\rho = 1$, de donde $\rho^{(n)} = 1$ si $\rho^{(0)} = 1$. Entonces la circunferencia K de centro en el origen y radio 1 es invariante por T y por lo probado antes $\text{dist}(f^n(p), K) \rightarrow 0$. Luego K es un atractor topológico de acuerdo con la definición 5.1.2. Además K es orbitalmente estable porque $|\rho^* - 1| \leq |\rho - 1|$. Entonces $|\rho^{(n)} - 1|$ es decreciente con n si $p \neq 0$. Luego, la distancia de $f^n(p)$ a la circunferencia K decrece con n . Por lo tanto, es menor que $\epsilon > 0$ si inicialmente es menor que $\delta = \epsilon$. Esto prueba la estabilidad orbital de K .

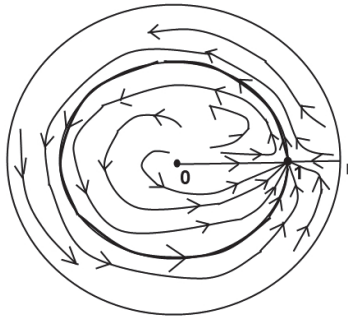


Figura 5.1: Algunas órbitas del Ejemplo g en 5.1.8 (B). El punto 1 es fijo, es un atractor topológico no orbitalmente estable

(B) Sea ahora $g : X \mapsto X$ dado por las ecuaciones

$$g(p) = f(\rho e^{i\varphi}) = \rho^* e^{i\varphi^*}, \text{ donde}$$

$$\rho^* = \frac{\rho(4-\rho)}{3}, \quad \varphi^* = \varphi \left(2 - \frac{\varphi}{2\pi} \right) \text{ para } \varphi \in [0, 2\pi].$$

En la figura 5.1 se representan algunas de las órbitas. Como en la parte (A), se prueba que la circunferencia K de centro en el origen y radio 1, es un atractor topológico orbitalmente estable. Además, si graficamos φ^* en función de φ para todo $\varphi \in [0, 2\pi]$, observamos que es creciente, con puntos fijos en 0 y 2π , siendo 0 repulsor (a la derecha de 0) y 2π atractor (a la izquierda de 2π). Entonces, si $\varphi^{(n)}$ es el ángulo en coordenadas

polares del punto $f^n(p)$, para $p = \rho e^{i\varphi}$, se tiene que $\lim_{n \rightarrow +\infty} \varphi^{(n)} = 2\pi \sim 0$. Luego, en este ejemplo, $1 = 1e^{2\pi i}$ es un punto fijo, y $\lim_{n \rightarrow +\infty} f^n(p) = 1$ para todo $p \neq 0$. Entonces $\{1\} \subset K$ también es atractor topológico según la definición 5.1.2. Pero $\{1\}$ no es orbitalmente estable: Por un lado 1 es un punto fijo, por lo tanto, si $\{1\}$ fuera orbitalmente estable, toda órbita futura con punto inicial p suficientemente cercano a 1, debería mantenerse arbitrariamente cerca de 1. Consideremos la órbita futura $o := \{f^n(p)\}_{n \geq 0}$ por $0 \neq p = \rho e^{i\varphi} \notin K$ con $0 < \varphi < \delta$ (para $\delta > 0$ suficientemente pequeño). Aunque o se acerca a K cuando n crece, y además $\lim_n f^n(p) = 1$, el ángulo $\varphi^{(n)}$ (esto es la coordenada angular de $f^n(p)$ en polares) no se mantiene a distancia menor que $\epsilon > 0$ de 0. Dicho de otra forma, la órbita $\{f^n(p)\}_n$, para algunos $n \geq 1$, se aleja del punto fijo 1 más que una constante positiva (digamos $1/6$, si $\delta > 0$ es suficientemente pequeño), aunque finalmente tiende a 1 cuando $n \rightarrow +\infty$ (ver Figura 5.1).

El siguiente resultado da una caracterización de los atractores topológicos orbitalmente estables. Debido a este resultado algunos autores definen atractor topológico agregando, a las condiciones 1) y 2) de la Definición 5.1.2, la condición de que sea orbitalmente estable.

Proposición 5.1.9. .

Caracterización de atractores topológicos orbitalm. estables

Sea $f : X \mapsto X$ un homeomorfismo en un espacio métrico compacto X .

(a) Si V es un abierto no vacío tal que $\overline{f(V)} \subset V$ y si

$$K := \bigcap_{n=0}^{+\infty} f^n(V), \quad (5.2)$$

entonces K es compacto no vacío, es un atractor topológico orbitalmente estable y V es cuenca local de atracción de K .

(b) Recíprocamente, si K es un atractor topológico orbitalmente estable, entonces existe un abierto V que es cuenca local de atracción de K , tal que $\overline{f(V)} \subset V$ y $K = \bigcap_{n=0}^{+\infty} f^n(V)$.

Demostremos la Proposición 5.1.9 más abajo en esta sección. Antes veamos sus consecuencias:

Definición de conjunto maximal invariante: Si K es compacto invariante no vacío que satisface la igualdad (5.2) para un abierto $V \supset K$, entonces K se llama *conjunto maximal invariante (hacia el futuro)* de V .

La Proposición 5.1.9 justifica la siguiente definición:

Definición 5.1.10. Atractor topológico II

Sea $f : X \mapsto X$ es un homeomorfismo en un espacio métrico compacto X . Un subconjunto no vacío, compacto e invariante $K \subset X$ se llama *atractor topológico orbitalmente estable* o también *atractor topológico maximal invariante* si existe un entorno abierto

$V \supset K$, llamado cuenca local de atracción topológica de K , tal que

$$\overline{f(V)} \subset V, \quad y \quad K := \bigcap_{n=0}^{+\infty} f^n(V).$$

Demostración de la Proposición 5.1.9

Demostración. (a) Como $\overline{f(V)} \subset V$ y f es un homeomorfismo, tenemos, por inducción:

$$\begin{aligned} f^{n+1}(V) &\subset \overline{f^{n+1}(V)} \subset f^n(V) \quad \forall n \in \mathbb{N}. \\ K &:= \bigcap_{n=0}^{+\infty} f^n(V) = V \cap \left(\bigcap_{n=1}^{+\infty} f^n(V) \right) \subset V \cap \left(\bigcap_{n=1}^{+\infty} \overline{f^n(V)} \right) = \\ &= \bigcap_{n=1}^{+\infty} \overline{f^n(V)} \subset \bigcap_{n=0}^{+\infty} f^n(V) = K. \end{aligned}$$

Luego, todas las inclusiones anteriores son igualdades y tenemos

$$K = \bigcap_{n=1}^{+\infty} \overline{f^n(V)}, \quad \bigcap_{n=1}^{+N} f^n(V) = f^N(V), \quad \bigcap_{n=1}^{+N} \overline{f^n(V)} = \overline{f^N(V)} \quad \forall N \geq 1.$$

Por la propiedad de intersecciones finitas no vacías de compactos, deducimos que K es compacto no vacío. Además

$$f^{-1}(K) = \bigcap_{n=1}^{+\infty} f^{-1}(\overline{f^n(V)}) = \bigcap_{n=0}^{+\infty} \overline{f^n(V)} = \overline{V} \cap K = K,$$

pues, por construcción $K \subset V$. Hemos probado que K es compacto no vacío e invariante con f .

Para $\epsilon > 0$, denotamos $B_\epsilon(K) := \{x \in X : \text{dist}(x, K) < \epsilon\}$.

Afirmación A: Para todo $\epsilon > 0$ existe $N \geq 1$ tal que $\overline{f^N(V)} \subset B_\epsilon(K)$.

Por absurdo, si existe $\epsilon > 0$ y para todo $n \geq 1$ existe $x_n \in \overline{f^n(V)}$ tal que $\text{dist}(x_n, K) \geq \epsilon$, entonces tomando una subsucesión de $\{x_n\}$ convergente a un punto x , tenemos $\text{dist}(x, K) \geq \epsilon$. Además $x \in \overline{f^N(V)} \quad \forall N \geq 1$ (pues $x_N \in \overline{f^N(V)} = \bigcap_{n=1}^N \overline{f^n(V)}$, de donde $x_m \in \overline{f^N(V)}$ para todo $m \geq N$). Entonces $x \in \bigcap_{N \geq 1} \overline{f^N(V)} = K$, lo cual contradice que $\text{dist}(x, K) \geq \epsilon$. Hemos probado la afirmación A.

Para probar que K es un atractor topológico, y que V es cuenca local de atracción de K , basta probar que si $y = \lim_{j \rightarrow +\infty} f^{n_j}(x)$ para algún $x \in V$ y para alguna subsucesión $n_j \rightarrow +\infty$, entonces $y \in K$. En efecto para todo $\epsilon > 0$, tenemos $\text{dist}(y, \overline{f^{n_j}(V)}) < \epsilon$ para todo j suficientemente grande. Luego, usando la afirmación (A), como $\overline{f^{n_j}(V)} \subset B_\epsilon(K)$ para todo n_j suficientemente grande, deducimos: $\text{dist}(y, K) < 2\epsilon$. La desigualdad anterior vale para todo $\epsilon > 0$. Entonces $\text{dist}(y, K) = 0$; es decir $y \in K$, como queríamos demostrar.

Ahora resta probar que K es orbitalmente estable. Dado $\epsilon > 0$ sea $N \geq 1$ como en la afirmación (A). Siendo f un homeomorfismo, $f^N(V)$ es abierto. Además $K \subset f^N(V)$ porque $K := \bigcap_{N \geq 0} f^N(V)$. Como K es compacto, existe $\delta > 0$ tal que $B_\delta(K) \subset f^N(V)$. Si $x \in B_\delta(K)$ entonces, para todo $m \geq 0$ se cumple $f^m(x) \in f^m(B_\delta(K)) \subset$

$f^m(f^N(V)) \subset \overline{f^{m+N}(V)} \subset \overline{f^N(V)} \subset B_\epsilon(K)$. Hemos probado que si $\text{dist}(x, K) < \delta$ entonces $\text{dist}(f^m(x), K) < \epsilon$ para todo $m \geq 0$. Luego K es orbitalmente estable, terminando la demostración de la parte (a).

(b) Sea K un atractor topológico orbitalmente estable y sea V' un abierto que es cuenca local de atracción de K , según la condición 2) de la Definición 5.1.2. Sea $\epsilon > 0$ tal que

$$\overline{B_{2\epsilon}(K)} \subset V',$$

donde $B_\epsilon(K) := \{x \in X : \text{dist}(x, K) < \epsilon\}$. Por la definición de estabilidad orbital de K existe $0 < \delta < \epsilon$ tal que $f^m(B_{2\delta}(K)) \subset B_\epsilon(K)$ para todo $m \geq 0$. Como $\overline{B_\delta(K)} \subset B_{2\delta}(K)$, tenemos

$$f^m(\overline{B_\delta(K)}) \subset B_\epsilon(K) \quad \forall m \geq 0.$$

Afirmación (B): Existe $N \geq 1$ tal que

$$f^m(x) \in B_\epsilon(K) \quad \forall m \geq N, \quad \forall x \in \overline{B_{2\epsilon}(K)}.$$

En efecto, como $\overline{B_{2\epsilon}(K)} \subset V'$ y V' satisface la condición 2) de la Definición 5.1.2, para cada $x \in \overline{B_{2\epsilon}(K)}$ existe $N_x \geq 1$ tal que $\text{dist}(f^{N_x}(x), K) < \delta$. Como f es continua, existe un entorno abierto U_x de x tal que $\text{dist}(f^{N_x}(y), K) < \delta$ para todo $y \in U_x$. Siendo $\overline{B_{2\epsilon}(K)}$ compacto, existe un subcubrimiento finito $\{U_{x_1}, U_{x_2}, \dots, U_{x_h}\}$. Deducimos que para todo $y \in \overline{B_{2\epsilon}(K)}$ existe $N_{x_i} \geq 1$ tal que $\text{dist}(f^{N_{x_i}}(y), K) < \delta$. Por la construcción de δ deducimos que $\text{dist}(f^m(y), K) < \epsilon$ para todo $m \geq N_{x_i}$. Luego, si $N = \max\{N_{x_i} : 1 \leq i \leq h\}$, se cumple $f^m(y) \in B_\epsilon(K)$ para todo $m \geq N$ y para todo $y \in \overline{B_{2\epsilon}(K)}$, terminando la prueba de la Afirmación (B).

Definimos el abierto V de la siguiente manera:

$$\begin{aligned} V := & B_{\epsilon_0}(K) \cup f(B_{\epsilon_1}(K)) \cup f^2(B_{\epsilon_2}(K)) \cup \\ & \cup f^3(B_{\epsilon_3}(K)) \cup \dots \cup f^N(B_{\epsilon_N}(K)), \\ \text{donde } & \epsilon_i := \epsilon \left(1 + \frac{i}{2^N}\right) < 2\epsilon \quad \forall 0 \leq i \leq N. \end{aligned}$$

Probemos que $\overline{f(V)} \subset V$. En efecto, por un lado si $0 \leq i \leq N-1$ entonces

$$\overline{f(f^i(B_{\epsilon_i}(K)))} \subset \overline{f^{i+1}(B_{\epsilon_i}(K))} \subset f^{i+1}(B_{\epsilon_{i+1}}(K)) \subset V.$$

Por otro lado, si $i = N$ entonces

$$\overline{f(f^N(B_{\epsilon_N}(K)))} \subset \overline{f(f^N(B_{2\epsilon}(K)))} = f^{N+1}(\overline{B_{2\epsilon}(K)}) \subset B_\epsilon(K) \subset V,$$

debido a la construcción del natural N según la afirmación (B).

Ahora probemos que V es cuenca local de atracción de K . Sea $x \in V$. Entonces $x \in f^i(B_{2\epsilon}(K))$ para algún $0 \leq i \leq N$, es decir $x = f^i(y)$ para algún $y \in B_{2\epsilon}(K) \subset V'$,

donde V' es por hipótesis una cuenca local de atracción de K . Entonces, por la propiedad 2) de la Definición 5.1.2, tenemos $\lim_{n \rightarrow +\infty} \text{dist}(f^n(y), K) = 0$. Luego

$$\lim_{m \rightarrow +\infty} \text{dist}(f^m(x), K) = \lim_{m \rightarrow +\infty} \text{dist}(f^{m-i}(y), K) = 0 \quad \forall x \in V.$$

Hemos probado que V satisface la condición 2) de la Definición 5.1.2. Entonces, por definición V es cuenca local de atracción topológica de K .

Para terminar la demostración solo falta probar que $K = \bigcap_{n=0}^{+\infty} f^n(V)$. Por un lado como $f(K) = K$ (porque por hipótesis K es invariante con f), y como $K \subset V$ por construcción, entonces $K = f^n(K) \subset f^n(V)$ para todo $n \geq 0$. Luego $K \subset \bigcap_{n=0}^{+\infty} f^n(V)$. Ahora probemos la inclusión opuesta. Sea $y \in \bigcap_{n=0}^{+\infty} f^n(V)$. Entonces

$$\forall n \geq 1 \quad \exists x_n \in f(V) \text{ tal que } y = f^{n-1}(x_n).$$

Sea x el límite de una subsucesión convergente de $\{x_n\} \subset f(V)$, es decir

$$x = \lim_{j \rightarrow +\infty} x_{n_j}, \quad n_j \rightarrow +\infty.$$

Tenemos $x \in \overline{f(V)} \subset V$.

Dado $\epsilon' > 0$, por la estabilidad orbital de K , existe $\delta' > 0$ tal que $f^n(B_{\delta'}(K)) \subset B_{\epsilon'}(K)$ para todo $n \geq 0$. Como V es cuenca local de atracción de K , y $x \in V$, deducimos que existe N tal que $\text{dist}(f^m(x), K) < \delta' \quad \forall m \geq N$, en particular

$$\text{dist}(f^N(x), K) < \delta'.$$

Por la continuidad de f^N , existe un entorno abierto U_x de x , tal que

$$\text{dist}(f^N(z), K) < \delta' \quad \forall z \in U_x.$$

Por construcción del punto x , $x_{n_j} \in U_x$ para todo j suficientemente grande. Luego:

$$\text{dist}(f^N(x_{n_j}), K) < \delta' \quad \forall j \text{ suficientemente grande.}$$

Por la construcción del número δ' a partir de la estabilidad orbital de K , deducimos que

$$\text{dist}(f^m(x_{n_j}), K) < \epsilon' \quad \forall m \geq N.$$

Luego, en particular, para $m = n_j - 1$, si elegimos j suficientemente grande tal que

$$n_j \geq N + 1,$$

tenemos

$$\text{dist}(f^{n_j-1}(x_{n_j}), K) < \epsilon'.$$

Por construcción de la sucesión $\{x_n\}$ tenemos $y = f^{n_j-1}(x_{n_j})$. Deducimos que $\text{dist}(y, K) < \epsilon'$. Como $\epsilon' > 0$ es arbitrario, concluimos que $y \in K$ como queríamos demostrar. \square

5.2. Cuenca global de atracción topológica

Definición 5.2.1. Cuenca global de atracción topológica

Sea $f : X \mapsto X$ continua en un espacio métrico compacto X . Sea $K \subset X$ (no vacío, compacto e invariante) un atractor topológico según la Definición 5.1.2. Se llama *cuenca de atracción topológica (global) de K* al siguiente conjunto:

$$C(K) = \{x \in X : \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0\}. \quad (5.3)$$

Nota: Dado un compacto K no vacío e invariante cualquiera, aunque el conjunto $C(K)$ construido mediante la igualdad (5.3) resulte no vacío, en general no se llama a este conjunto “cuenca de atracción topológica” de K , si K no es un atractor topológico. Es decir, cuando se usa el nombre “cuenca de atracción topológica”, previamente sabemos que K satisface todas las condiciones de la Definición 5.1.2, en particular la condición 2) de existencia de un *entorno abierto* de K contenido en la cuenca global de atracción topológica.

Proposición 5.2.2. .

Propiedades de la cuenca global de atracción topológica.

Sea $f : X \mapsto X$ continua en un espacio métrico compacto X . Si $K \subset X$ es un atractor topológico, entonces su cuenca (global) de atracción topológica $C(K)$, tiene las siguientes propiedades:

- (a) $C(K)$ es abierto y no vacío.
- (b) $C(K)$ es invariante con f , es decir $f^{-1}(C(K)) = C(K)$.
- (c) $K \subset C(K)$.

Como consecuencia de (a) y (b) si X es conexo, entonces o bien $K = X$ (en cuyo caso $C(K) = K = X$), o bien la cuenca de atracción $C(K)$ contiene propiamente a K (no coincide K con su cuenca).

Demostración. (b) Si $x \in C(K)$ entonces, por la igualdad (5.3):
 $\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0$. Luego,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \text{dist}(f^n(f(x)), K) &= \lim_{n \rightarrow +\infty} \text{dist}(f^{n-1}(f(x)), K) = \\ &= \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0. \end{aligned}$$

Esto muestra que $f(x) \in C(K)$ para todo $x \in C(K)$. Hemos probado que $f(C(K)) \subset C(K)$, o lo que es lo mismo, $C(K) \subset f^{-1}(C(K))$.

Sea ahora $y \in f^{-1}(C(K))$. Entonces $f(y) = x \in C(K)$. Luego, para todo $n \geq 0$ tenemos $f^n(x) = f^{n+1}(y)$. Entonces:

$$\begin{aligned} \lim_{n \rightarrow +\infty} \text{dist}(f^n(y), K) &= \lim_{n \rightarrow +\infty} \text{dist}(f^{n+1}(y), K) = \\ &= \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0. \end{aligned}$$

Se deduce que $y \in C(K)$ para todo $y \in f^{-1}(C(K))$, probando que $f^{-1}(C(K)) \subset C(K)$.

(a) y (c) Por hipótesis K es un atractor topológico. Entonces existe $V \supset K$ abierto, cuenca local de atracción topológica de K , de acuerdo a la condición 2) de la Definición 5.1.2. De la definición de $C(K)$ en (5.3), junto con la igualdad (5.1), tenemos $V \subset C(K)$, probando que $K \subset C(K)$ y, por lo tanto, $C(K)$ es no vacío. Si $x \in C(K)$, entonces $\lim_n \text{dist}(f^n(x), K) = 0$. Como V es un entorno abierto de K , deducimos que existe $N \geq 1$ (que depende de $x \in C(K)$), tal que

$$f^n(x) \in V \quad \forall n \geq N.$$

En particular, la afirmación de arriba vale para $n = N$. Luego, por la continuidad de f y la apertura de V , existe un entorno abierto U_x de x tal que

$$f^N(x) \in V \quad \forall x \in U_x.$$

Como $V \subset C(X)$, deducimos que $f^N(U_x) \subset C(X)$, o dicho de otra forma $U_x \subset f^{-N}(C(X)) = C(X)$. Luego $C(X)$ es abierto. \square

Proposición 5.2.3. *Sea $f : X \mapsto X$ continua en un espacio métrico compacto X . Sea $K \subset X$ un atractor topológico con cuenca global de atracción topológica $C(K)$. Entonces:*

- (a) *Existen medidas de probabilidad invariantes para $f|_{C(K)}$.*
- (b) *Existen medidas de probabilidad ergódicas para $f|_{C(K)}$.*
- (c) *Toda medida de probabilidad invariante para $f|_{C(K)}$ está soportada en el atractor K (es decir $\mu(K) = 1$).*

Demostración. (a) y (b) Consideremos $f|_K : K \mapsto K$. Es la restricción al atractor $K \subset C(K)$ de $f|_{C(K)} : C(K) \mapsto C(K)$. Como f es continua y K es compacto, entonces, por lo demostrado en el capítulo 1, existen medidas de probabilidad invariantes para $f|_K$, y también existen medidas de probabilidad ergódicas para $f|_K$. Tomemos una de estas medidas ν invariante para $f|_K$. Definamos la medida de probabilidad μ en los borelianos B de $C(K)$ de la siguiente forma:

$$\mu(B) := \nu(B \cap K) \quad \forall B \subset C(K) \text{ boreliano}.$$

(Observar que $\mu(C(K)) = \nu(C(K) \cap K) = \nu(K) = 1$, porque $K \subset C(K)$.) Veamos que la probabilidad μ es invariante para $f|_{C(K)}$. Sea $B \subset C(K)$ un boreliano cualquiera:

$$\begin{aligned} \mu(f^{-1}(B)) &= \nu(f^{-1}(B) \cap K) = \nu(f^{-1}(B) \cap f^{-1}(K)) = \\ &= \nu(f^{-1}(B \cap K)) = \nu(B \cap K) = \mu(B), \end{aligned}$$

pues ν es invariante con $f|_K$. Ahora veamos que si ν es ergódica para $f|_K$ entonces μ es ergódica para $f|_{C(K)}$. Sea $B \subset C(K)$ boreliano tal que $f^{-1}(B) = B$. Tenemos que $f^{-1}(B \cap K) = f^{-1}(B) \cap f^{-1}(K) = B \cap K$. Luego, como ν es ergódica para $f|_K$, se cumple $\nu(B \cap K) \in \{0, 1\}$. Entonces:

$\mu(B) = \nu(B \cap K) \in \{0, 1\}$, lo que prueba la ergodicidad de μ .

(c) Sea μ una medida de probabilidad invariante para $f|_{C(K)}$. Entonces:

$\mu(C(K)) = 1$. Debemos probar que $\mu(K) = 1$ (lo cual implica que $\mu(C(K) \setminus K) = 0$). Por el Lema de Recurrencia de Poincaré, μ -c.t.p. es recurrente; es decir $x \in \omega(x)$ para μ -c.t.p. $x \in C(K)$. Para probar que $\mu(K) = 1$ basta probar que $x \in K$ para cualquier $x \in C(K)$ tal que $x \in \omega(x)$. Para esto, alcanza demostrar que $\omega(x) \subset K$ para todo $x \in C(K)$. En efecto, por la definición de omega-límite, si $y \in \omega(x)$ entonces

$$y = \lim_{j \rightarrow +\infty} f^{n_j}(x), \quad n_j \rightarrow +\infty.$$

Luego:

$$\text{dist}(y, K) = \lim_{j \rightarrow +\infty} \text{dist}(f^{n_j}(x), K) \leq \limsup_{n \rightarrow +\infty} \text{dist}(f^n(x), K).$$

Como $x \in C(K)$, por la definición de cuenca de atracción topológica dada en (5.3), tenemos:

$$\limsup_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0.$$

Luego $\text{dist}(y, K) = 0$, o lo que es lo mismo $y \in K$ para todo $y \in \omega(x)$, como queríamos demostrar. \square

5.3. Atractores hiperbólicos caóticos

En esta sección consideramos el caso particular en que $X = M$ es una variedad diferenciable compacta y riemanniana, y $f \in \text{Diff}^1(M)$.

Se llaman *atractores hiperbólicos*, a aquellos atractores topológicos K que están soportados en las variedades inestables de un conjunto compacto no vacío (unif. o no unif.) hiperbólico $\Lambda \subset K$. Más precisamente: K es un atractor hiperbólico si es un atractor topológico y existe un conjunto invariante compacto y (unif. o no unif.) hiperbólico $\Lambda \subset K$ tal que $K \subset \bigcup_{p \in \Lambda} W^u(p)$.

El caso no trivial es cuando la dimensión de estas variedades inestables es mayor o igual que 1, para μ - casi todo punto $x \in K$ para alguna medida invariante μ , soportada en el atractor K . Como son variedades inestables de puntos de un conjunto hiperbólico, los exponentes de Lyapunov en las direcciones tangentes a estas variedades (tangentes al atractor) son positivos. Llamaremos a tales atractores topológicos, *atractores hiperbólicos caóticos*.

La búsqueda de variedades inestables que soporten el atractor, es una de las motivaciones más relevantes de la teoría de sistemas dinámicos diferenciables. Significa que la dinámica dentro de un atractor topológico K (es decir la dinámica de $f|_K$) puede ser expansiva en el futuro, o en otras palabras, caótica; por tener exponentes de Lyapunov positivos. Dicho de otra forma, veamos el caso en que el atractor K sea orbitalmente estable (ó Lyapunov estable, respectivamente). La estabilidad orbital (de Lyapunov, resp.) rige en la cuenca de atracción $C(K)$. En efecto, por definición, la estabilidad orbital (de Lyapunov resp.) tiene significado no trivial solo para las órbitas de $C(K)$ fuera de K . Sin embargo, no es contradictoria con una dinámica inestable y expansiva, dada por los exponentes de Lyapunov positivos, dentro del atractor K (ver Ejercicio 5.3.1).

El siguiente ejercicio muestra un ejemplo en el que la dinámica dentro del atractor topológico es expansiva o caótica: el atractor está formado por las variedades inestables de puntos con exponentes de Lyapunov positivos:

Ejercicio 5.3.1. Atractor Solenoide (Smale-Williams)

Sea $S \subset \mathbb{R}^3$ el toro sólido compacto, definido como la imagen en \mathbb{R}^3 de la siguiente parametrización con tres parámetros reales (en coordenadas cilíndricas)

$$(r, \varphi, \theta) \in [0, a] \times [0, 2\pi] \times [0, 2\pi],$$

donde $0 < a < 1/2$ es una constante real:

$$\begin{aligned} x &= (r - 1/2) \cos \theta \cos \varphi \\ y &= r \cos \theta \sin \varphi \\ z &= r \sin \theta \end{aligned}$$

(a) Interpretar el significado geométrico de los parámetros (r, φ, θ) en \mathbb{R}^3 y dibujar el toro sólido S . (Sugerencia, dibujar primero el disco circular de radio a en el plano xz (es decir el plano $\{\varphi = 0\}$). Luego, observar que S es el sólido de revolución que se obtiene haciendo girar ese disco alrededor del eje de las z).

(b) Sea $f : S \mapsto \text{int}(S)$ continua, tal que lleva cada disco circular D_φ que se obtiene cortando S con el plano φ constante, en un disco circular $f(D_\varphi)$ de radio $a/4$, contenido en la sección $D_{2\varphi}$, de modo que $f|_{D_\varphi} = f_3 \circ f_2 \circ f_1$ donde:

- f_1 es una rotación con eje z de ángulo φ (esto implica que $f_1(D_\varphi) = D_{2\varphi}$; cuando escribimos 2φ nos referimos a este ángulo módulo 2π).
- f_2 es una homotecia de razón $1/4$ que lleva el centro del disco $D_{2\varphi}$ al punto, en el interior de $D_{2\varphi}$, con parámetros $(a/2, 2\varphi, 0)$
- f_3 es una rotación de ángulo φ alrededor del eje perpendicular al plano del disco $D_{2\varphi}$ que pasa por el centro del círculo $D_{2\varphi}$.

Verificar que $f : S \mapsto f(S)$ es inyectiva, y que $f(S)$ es un compacto contenido en el interior de S . Dibujar $f(S)$. Sugerencia: $f(S)$ da dos vueltas alrededor del eje de las z . Ver por ejemplo [Pi 2007, Figure 1].

(c) Sea $K = \bigcap_{n \geq 0} f^n(S)$ el atractor topológico con cuenca local de atracción S . Sea, para cada $N \geq 0$ el compacto $K_N = \bigcap_{n=0}^N f^n(S)$. Sea $A_N = K_N \cap \{(r, \varphi, \theta) \in S : \varphi = 0\}$. Dibujar A_1, A_2, K_1, K_2 , y probar que $K_{N+1} \subset \text{int}(K_N)$ para todo $N \geq 0$. (Sugerencia: inducción en N).

(d) Para cada punto $p \in K$ se define la variedad inestable de p del siguiente modo:

$$\begin{aligned} W^u(p) &= \\ &= \{q \in S : \exists f^{-n}(q) \in S \forall n \geq 0 \text{ y } \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(q), f^{-n}(p)) = 0\}. \end{aligned}$$

Probar que $W^u(p)$ es una variedad inmersa en \mathbb{R}^3 de dimensión 1. Sugerencia: Fijar un punto cualquiera $q \in W^u(p)$. No es restrictivo asumir que la coordenada φ de q es cero. Usando las secciones A_N definidas en la parte (c) probar que la componente

conexa de la intersección de $W^u(p)$ con un entorno pequeño de punto q en \mathbb{R}^3 , es un arco diferenciable.

(e) Probar que el atractor K es:

$$K = \bigcup_{p \in K} W^u(p).$$

(f) Probar que K es un atractor topológico caótico. Esto significa que para alguna medida invariante μ soportada en K , para μ -c.t.p. $p \in K$ y para toda dirección tangente inestable $0 \neq v \in T_p(W^u(p)) = T_p(K)$ el exponente de Lyapunov de v hacia el futuro es positivo. Alcanza con probar que:

$$\liminf_{n \rightarrow +\infty} \frac{\log \|df^n(v)\|}{n} \geq \log 2 > 0 \quad \forall p \in K, \quad \forall 0 \neq v \in T_p(W^u(p)).$$

(Sugerencia: probar que $\|df_p(v)\|/\|v\| \geq 2$ para todo $p \in K$ y para todo $0 \neq v \in T_p(W^u(p))$.)

El siguiente ejemplo muestra, además del atractor solenoide expuesto en el Ejercicio 5.3.1, otro atractor topológico formado por la clausura de la variedad inestable de un conjunto hiperbólico.

Ejemplo 5.3.2. Herradura de Smale en el disco

Sea $T : [0, 1]^2 \subset \mathbb{R}^2 \mapsto T([0, 1]^2) \subset \mathbb{R}^2$ la herradura de Smale lineal, definida en 3.5.1. Sea el “disco” rectangular $D = [0, 1] \times [-3, 3] \subset \mathbb{R}^2$. D es homeomorfo a un disco compacto. Ver Figura 5.2. Tenemos

$$[0, 1]^2 \subset D.$$

Denotamos

$$D = D_1 \cup Q \cup D_2, \text{ donde}$$

$$D_1 = \{(x, y) \in D : y \leq -2\}, \quad D_2 = \{(x, y) \in D : y \geq 2\}, \text{ y}$$

$$Q = \{(x, y) \in D : -2 < y < 2\}. \text{ Luego: } [0, 1]^2 \subset Q, \quad \text{int}(Q) \setminus [0, 1]^2 \neq \emptyset.$$

Sea $f : D \mapsto f(D) \subset \text{int}(D)$ un difeomorfismo construido como se describe a continuación (ver Figura 5.2):

$$f = f_2 \circ f_1, \text{ donde}$$

$$f_1 : D \mapsto f_1(D) \subset \mathbb{R}^2 \text{ tal que } f_1(x, y) = \left(\frac{1}{5}x, 5y\right) \quad \forall (x, y) \in D,$$

$$f_2 : f_1(D) \mapsto f(D) \subset \text{int}(D) \text{ tal que}$$

- $f_2|_{f_1(D_1)}$ es una traslación compuesta con la transformación lineal $(x, y) \mapsto (x, y/10)$, tal que

$$f(D_1) = f_2 \circ f_1(D_1) \subset \text{int}(D_1).$$

- $f_2|_{f_1(D_2)}$ es una traslación compuesta con la transformación lineal $(x, y) \mapsto (-x, -y/10)$, tal que

$$f(D_2) = f_2 \circ f_1(D_2) \subset \text{int}(D_1), \quad f(D_1) \cap f(D_2) = \emptyset.$$

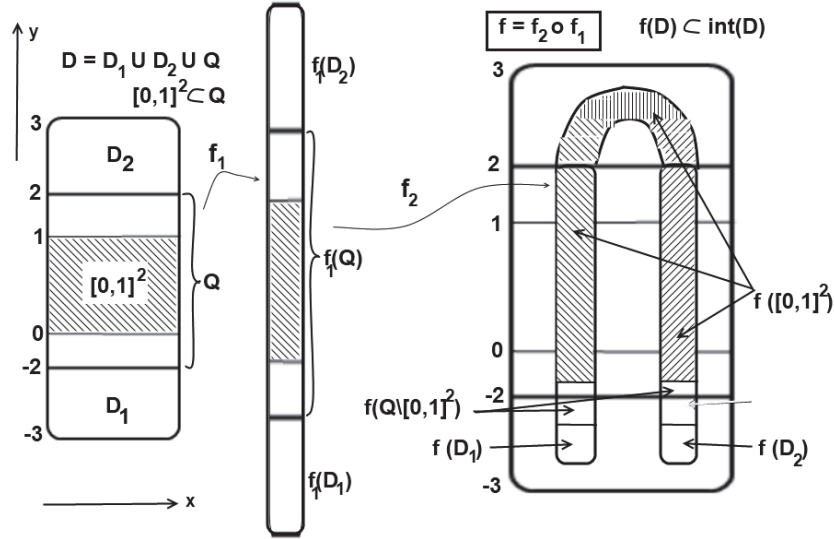


Figura 5.2: Herradura de Smale en el disco D (las escalas están deformadas).

- $f(Q) = f_2 \circ f_1(Q)$, tal que f restringida al cuadrado $[0, 1]^2 \subset \text{int}(Q)$ es la transformación T definida en la Sección 3.5:

$$f|_{[0,1]^2} = f_2 \circ f_1|_{[0,1]^2} = T : [0, 1]^2 \mapsto T([0, 1]^2) \subset \text{int}(Q \cup D_2).$$

- $f|(Q \setminus [0, 1]^2)$ es tal que

$$\{p \in Q \setminus [0, 1]^2 : \exists f^{-n}(p) \in Q \setminus [0, 1]^2 \forall n \geq 0\} = \emptyset.$$

(a) La construcción anterior puede realizarse de modo que f sea un difeomorfismo sobre su imagen, a pesar que las fórmulas de construcción de $f|_{D_1}$, $f|_{D_2}$ y $f|_{[0,1]^2}$ están fijas de antemano. Esto es porque $D_1, D_2, [0, 1]^2$ son compactos disjuntos dos a dos, cuyas imágenes son compactas disjuntas dos a dos.

(b) De la definición de $f|_{D_1}$ se deduce que es una transformación afín (lineal compuesta con una traslación) cuya matriz asociada es diagonal con valores propios $1/5$ y $1/2$. Además $f|_{D_1}$ es continua y $f(D_1) \subset \text{int}D_1$. Entonces $f|_{D_1}$ es una contracción uniforme del conjunto compacto D_1 en su interior. Por el teorema del punto fijo de Banach, existe $p_0 = (x_0, y_0) \in \text{int}(D_1)$ tal que $f(p_0) = p_0$. Además, como los valores propios de $df(p_0)$ son $1/5$ y $1/2$, p_0 es un pozo hiperbólico. Consideremos su variedad inestable $W^u(p_0)$ definida por

$$W^u(p_0) := \{q \in D : \exists f^{-n}(q) \forall n \in \mathbb{N} \text{ y } \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(q), p_0) = 0\}.$$

Siendo p_0 un pozo, deducimos que $W^u(p_0)$ tiene dimensión cero. Precisamente

$$W^u(p_0) = \{p_0\} = \bigcap_{n \geq 0} f^n(D_1).$$

(c) Sea $\Lambda \subset \text{int}([0, 1]^2)$ el conjunto compacto maximal invariante de la Herradura de Smale lineal $T : [0, 1]^2 \mapsto \text{int}(D)$ (según Definición 3.5.3). Por construcción de $\Lambda = \bigcap_{n \in \mathbb{Z}} f^n([0, 1])$. Entonces, para todo $p \in \Lambda$ existe $f^{-n}(p)$ y está en $[0, 1]^2$ cualquiera sea $n \in \mathbb{N}$.

Para cada $p \in \Lambda$ consideramos su variedad inestable $W^u(p)$, definida por:

$$W^u(p) := \{q \in D : \exists f^{-n}(q) \forall n \in \mathbb{N} \text{ y } \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(q), f^{-n}(p)) = 0\}.$$

Definimos $W^u(\Lambda)$, que llamamos variedad inestable del conjunto Λ por:

$$W^u(\Lambda) := \bigcup_{p \in \Lambda} W^u(p).$$

Sea $W_{loc}^u(\Lambda)$, que llamamos variedad inestable local en el cuadrado $[0, 1]^2$ del conjunto Λ , definida por:

$$W_{loc}^u(\Lambda) := \{q \in W^u(\Lambda) : f^{-n}(q) \in [0, 1]^2 \forall n \geq 0\}.$$

Por construcción

$$\Lambda \subset W_{loc}^u(\Lambda) \subset W^u(\Lambda) \quad \text{y} \quad f^n(p) \in [0, 1]^2 \forall n \in \mathbb{Z} \text{ si y solo si } p \in \Lambda.$$

Ejercicio 5.3.3. En el ejemplo 5.3.2 probar que

$$\begin{aligned} W_{loc}^u(\Lambda) &= \bigcap_{n \geq 0} f^n([0, 1]^2) = \bigcap_{n \geq 0} T^n([0, 1]^2) \quad \text{y} \\ \bigcap_{n=0} T^n([0, 1]^2) &= [0, 1]^2 \bigcap \left(\bigcap_{n \geq N} T^n([0, 1]^2) \right) \quad \forall N \geq 0. \end{aligned} \quad (5.4)$$

Deducir que para todo $p \in \Lambda$, $W_{loc}^u(p)$ es una subvariedad de dimensión uno.

Probar que $W^u(\Lambda) = \bigcup_{n \geq 0} f^n(W_{loc}^u(\Lambda))$.

Probar que $f(W^u(\Lambda)) = W^u(\Lambda)$

Del Ejercicio 5.3.3 y teniendo en cuenta que f es un difeomorfismo, deducimos que $W^u(p)$ es una variedad inmersa de dimensión uno para todo $p \in \Lambda$. En la Figura 5.3 ilustramos gráficamente un arco de la variedad inestable global $W^u(p)$ para un punto $p \in \Lambda$. Esta subvariedad global se arrolla infinitamente sobre sí misma, de tal forma que resulta imposible dibujarla completamente, pero sí puede ser imaginada mentalmente en su totalidad.

En el Ejemplo 5.3.2 sea Λ^* el conjunto hiperbólico definido por

$$\Lambda^* := \Lambda \cup \{p_0\}. \quad (5.5)$$

Sea $W^u(\Lambda^*)$ la variedad inestable de Λ^* definida por

$$W^u(\Lambda^*) := \bigcup_{p^* \in \Lambda^*} W^u(p^*) = \{p_0\} \cup W^u(\Lambda),$$

donde p_0 es el pozo construido en $D_1 \subset D$. En efecto, se que como p_0 es un pozo, entonces

$$W_{p_0}^u = \{q \in \Lambda^* : \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(q), f^{-n}(p_0)) = 0\} = \{p_0\}.$$

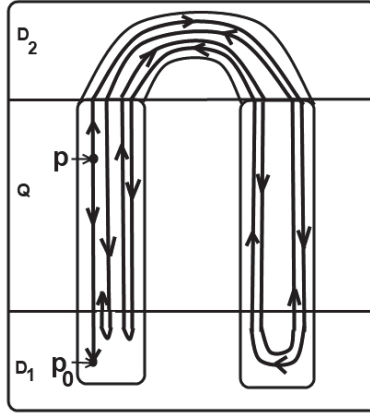


Figura 5.3: Arco de la variedad inestable $W^u(p)$ por el punto p en el atractor de la Herradura de Smale.

Proposición 5.3.4. *Sea Λ^* el conjunto hiperbólico construido por la igualdad (5.5) en el Ejemplo 5.3.2. Entonces $W^u(\Lambda^*)$ es un atractor topológico e $\text{int}(D)$ es cuenca local de atracción de $W^u(\Lambda^*)$.*

Demostración. Como por construcción $\overline{f(\text{int}(D))} \subset \text{int}(D)$, aplicando la Proposición 5.1.9, basta demostrar que

$$W^u(\Lambda^*) = K := \bigcap_{n \geq 0} f^n(D).$$

$W^u(\Lambda^*) \subset D$ y $f^n(W^u(\Lambda^*)) = W^u(\Lambda^*)$ para todo $n \geq 0$. Luego $W^u(\Lambda^*) \in f^n(D)$ para todo $n \geq 0$. Luego $W^u(\Lambda^*) \subset K$. Ahora demostramos la inclusión opuesta. Sea $p \in \bigcap_{n \geq 0} f^n(D)$. Entonces, teniendo en cuenta que $f : D \mapsto f(D)$ es inyectiva, para todo $n \geq 0$ existe único $p_n \in D$ tal que

$$p = f^n(p_n), \quad p_n = f(p_{n+1}).$$

Tenemos $D = D_1 \cup D_2 \cup Q$, siendo estos tres conjuntos disjuntos dos a dos. Por construcción de f se cumple $f(D_1) \subset \text{int}(D_1)$, $f(D_2) \subset \text{int}(D_1)$, $\{p : f^{-n}(p) \in Q \setminus [0, 1]^2 \forall n \geq 0\} = \emptyset$. Luego, existen solo dos posibilidades para la sucesión de puntos $\{p_n\}_{n \geq 0}$:

Caso 1: Existe $N \geq 0$ tal que $p_n \in [0, 1]^2$ para todo $n \geq N$.

Caso 2: $p_n \in D_1$ para todo $n \geq 0$.

En el Caso 1:

$$p \in f^N \left(\bigcap_{m=0}^{+\infty} f^m([0, 1]^2) \right) = f^N(W_{loc}^u(\Lambda)) \subset W^u(\Lambda).$$

En el Caso 2:

$$p \in \bigcap_{m=0}^{+\infty} f^m(D_1) = \{p_0\}.$$

Hemos probado que $p \in \{p_0\} \cup W^u(\Lambda) = W^u(\Lambda^*) \subset W^u(\Lambda^*)$, como queríamos. Para terminar la demostración de la Proposición 5.3.4, resta probar que $K = W^u(\Lambda^*)$ es un atractor hiperbólico caótico: es decir existe una medida invariante μ soportada en K tal que para μ -c.t.p. $p \in K$ la dimensión de la variedad inestable es uno. En efecto, ya probamos que $W^u(p)$ tiene dimensión uno para todo $p \in \Lambda = \Lambda^* \setminus \{p_0\}$. Entonces basta demostrar que existe una medida invariante μ tal que $\mu(\Lambda) = 1$. En efecto, como Λ es el conjunto maximal invariante de la Herradura de Smale lineal definida en la sección 3.5, se tiene Λ compacto y $f(\Lambda) = \Lambda$. La transformación $f|_\Lambda$ es continua de un espacio métrico compacto Λ en sí mismo. Por el teorema de existencia de medidas invariantes, existe alguna probabilidad ν invariante para $f|_\Lambda$. Definimos la siguiente medida μ , en los borelianos $B \subset D$:

$$\mu(B) = \nu(B \cap \Lambda).$$

Es inmediato chequear que μ es una probabilidad en D , invariante por $f : D \mapsto D$ y que $\mu(\Lambda) = 1$, terminando la demostración de la Proposición 5.3.4. \square

5.4. Atractores ergódicos

En esta sección consideraremos $f : M \mapsto M$ continua en una variedad compacta y riemanniana M . Por el teorema ergódico de Birkhoff-Khinchin, cualquiera sea la medida invariante μ , para μ -c.t.p. $x \in M$ y para toda función $\psi \in L^1(\mu)$ (en particular para toda función continua ψ) existe el promedio temporal asintótico $\tilde{\psi}$ (en el futuro) definido por:

$$\tilde{\psi}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)).$$

Es decir, para μ -c.t.p. x fijo en la variedad M , está definido el funcional lineal

$$\psi \in C^0(M, \mathbb{R}) \mapsto \tilde{\psi}(x).$$

Por el teorema de Representación de Riesz, existe una medida de probabilidad μ_x tal que

$$\tilde{\psi}(x) = \int \psi d\mu_x \quad \text{para } \mu\text{-c.t.p. } x \in M, \quad \forall \mu \in \mathcal{M}_f,$$

donde \mathcal{M}_f denota el espacio de todas las medidas de probabilidad en la sigma-álgebra de Borel de M que son f -invariantes. En los capítulos anteriores probamos que $\mu_x \in \mathcal{M}_f$. Además, dotando el espacio \mathcal{M} de probabilidades (no necesariamente f -invariantes) de la topología débil-estrella, tenemos:

$$\mu_x = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} \quad \text{para } \mu\text{-c.t.p. } x \in M, \quad \forall \mu \in \mathcal{M}_f,$$

donde el límite es tomado en \mathcal{M} con la topología débil* y δ_y denota la probabilidad Delta de Dirac soportada en el punto $y \in M$.

En las secciones anteriores vimos además que μ_x es ergódica para μ -c.t.p. $x \in M$, para toda $\mu \in \mathcal{M}_f$. Luego, el promedio temporal asintótico $\tilde{\psi}(x)$ coincide con el valor esperado de ψ con respecto a la probabilidad μ_x .

Por un lado, notamos que la teoría ergódica desarrollada hasta ahora es válida para μ -c.t.p. $x \in M$, y en general, no para cualquier punto $x \in M$. Dicho de otra forma, si el criterio de selección de los puntos iniciales $x \in M$ no es μ -c.t.p. para alguna medida $\mu \in \mathcal{M}_f$, entonces no necesariamente existen los promedios asintóticos de Birkhoff $\tilde{\psi}(x)$. Y si existen, en general, no coinciden con el promedio espacial que resulta de integrar ψ respecto a una medida ergódica.

Por otro lado, cuando se tiene un atractor topológico K con cuenca local V (abierto), el criterio de selección de los estados iniciales, por la propia definición de atractor topológico, reside en tomar los puntos $x \in V$ para algún abierto V . El criterio topológico de relevancia u “observabilidad” de conjuntos de órbitas, es que estos conjuntos sean abiertos, o más en general, con interior no vacío.

Sin embargo, en la mayoría de los ejemplos de atractores topológicos que vimos en la sección anterior, el atractor K (que está contenido en su cuenca local V) tiene interior vacío. Además vimos (Proposición 5.2.3), que las medidas μ invariantes por $f|_V$ (que siempre existen) están soportadas en K . Entonces μ -c.t.p. x de la cuenca V , está en K . Luego, el teorema ergódico de Birkhoff, y el teorema de existencia de medidas invariantes y ergódicas, *no aseguran la existencia de los promedios temporales asintóticos*, para las órbitas con punto inicial en un conjunto de estados iniciales relevante u “observable”, desde el punto de vista topológico.

Notamos que, en general, no hay esperanza que los promedios temporales asintóticos existan para las órbitas con punto inicial en un conjunto de estados iniciales con interior no vacío (es decir, relevante u “observable”, desde el punto de vista topológico). En efecto, si K es un atractor topológico con cuenca de atracción $C(K)$ (abierto), se puede demostrar que si $f|_{C(K)}$ no es únicamente ergódica (y en general no lo es), entonces el conjunto de estados iniciales $x \in C(K)$ para los cuales no existe el promedio asintótico de Birkhoff, es denso. Luego, el conjunto de estados iniciales para los cuales existe ese promedio temporal asintótico, tiene interior vacío.

Por estos motivos, entre otras razones, se adopta otro criterio de “observabilidad” de las órbitas o de selección de los estados iniciales. Es un criterio medible en vez de topológico, pero que considera a la mayoría de los puntos de la cuenca $C(K)$.

Definición 5.4.1. Criterio de observabilidad medible

Cuando el espacio es una variedad riemanniana M , el criterio medible de “observabilidad” de las órbitas, es que formen un conjunto con medida de Lebesgue positiva.

Lo anterior justifica las siguientes definiciones:

Definición 5.4.2. Atractor ergódico

Sea $f : M \mapsto M$ continua en una variedad compacta y Riemanniana M de dimensión finita. Denotamos con m a la medida de Lebesgue en M . Notamos que m no es necesariamente f -invariante.

Un conjunto compacto no vacío K se llama *atractor ergódico* si:

- $f^{-1}(K) = K = f(K)$

- Existe un abierto $V \supset K$ tal que

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0 \quad \text{para Lebesgue-c.t.p. } x \in V. \quad (5.6)$$

-

$\exists \mu$ ergódica tal que: $\mu(K) = 1$,

$\mu(K') < 1$ para todo compacto $K' \subset K$, $K' \neq K$, y

$$\exists \tilde{\psi}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)) = \int \psi d\mu \quad (5.7)$$

para Lebesgue-c.t.p. $x \in V$ y $\forall \psi \in C^0(M, \mathbb{R})$.

Un abierto $V \supset K$ que satisface las condiciones anteriores, se llama *cuenca local de atracción* del atractor ergódico K .

Nota: No todos adoptan la Definición 5.4.2. Algunos autores exigen que K sea, por definición, un atractor topológico que satisface además la condición (5.7), para llamarlo atractor ergódico (ver por ejemplo [Pug-Shu 1989]).

Definición: Medida SRB o física soportada en un atractor ergódico

Se observa que cuando existe una medida μ ergódica, soportada en el atractor ergódico K , que satisface la igualdad (5.7) para Lebesgue-c.t.p. $x \in V$, donde V es un entorno de K , entonces esta medida es *única*. En efecto, dos medidas diferentes no pueden satisfacer la igualdad (5.7), ambas para Lebesgue c.t.p. $x \in V$, pues cuando un punto satisface la igualdad (5.7) para una medida, no puede satisfacerla para otra medida diferente.

Tal medida ergódica única μ se llama *medida SRB ergódica* o también *medida física ergódica*, del atractor ergódico K dado.

En 5.5.6 definiremos medida SRB o medida física μ , en un contexto más general, aunque μ no esté soportada en un atractor ergódico.

Observación: Dado cualquier atractor topológico K con cuenca local V , debido a la Proposición 5.2.3 siempre existen medidas invariantes y ergódicas para $f|_V$, y están soportadas en K . Si alguna de estas medidas ergódicas μ satisface la condición (5.7) para Lebesgue-casi todo punto $x \in V$, entonces K es un atractor ergódico.

Ejercicio 5.4.3. Sea K el atractor topológico del Ejemplo 5.3.2 de la Herradura de Smale f en el disco D bidimensional.

(a) Probar que para todo $k \geq 2$ natural, existe en K órbita periódica o de período k . Probar que la medida invariante soportada en esa órbita periódica es ergódica, soportada en el atractor, pero no es medida SRB.

Sugerencia: probar que el conjunto de puntos $x \in D$ tales que $\text{dist}(f^n(x), o) \rightarrow 0$ es la variedad estable global $W^s(o)$ de la órbita periódica o . Probar que $W^s(o)$ es localmente, alrededor de o , el producto cartesiano de un conjunto de Cantor en el intervalo, por un segmento de recta. Demostrar que ese conjunto de Cantor tiene

medida de Lebesgue (unidimensional) igual a 0 y concluir que la medida de Lebesgue (bidimensional) de $W^s(o)$ es nula.

(b) Probar que existe un único punto fijo $p_0 \in D$ por f en el disco D tal que Lebesgue-c.t.p. $x \in D$ cumple $\text{dist}(f^n(x), p_0) \rightarrow 0$. Concluir que $\{p_0\}$ es un atractor ergódico.

(c) Probar que el atractor ergódico $\{p_0\}$ **no** es caótico de acuerdo a la siguiente definición:

Definición (atractor ergódico caótico): Un atractor ergódico se llama *caótico*, si para μ -c.t.p. existen exponentes de Lyapunov estrictamente positivos, donde μ es la (única) medida SRB ergódica del atractor.

(d) Concluir que el atractor K del Ejemplo 5.3.2 es atractor topológico hiperbólico caótico (cf. definición en la Sección 5.3), pero no es atractor ergódico caótico.

Observación 5.4.4. (Observabilidad Lebesgue-medible de la cuenca)

A diferencia de los atractores topológicos, para los atractores ergódicos K la atracción a K de las órbitas en su cuenca V dada por la igualdad (5.6) es solo para Lebesgue c.t.p. $x \in V$, y no necesariamente para todo punto $x \in V$. Es un *criterio de observabilidad medible Lebesgue-c.t.p.* de la cuenca. Entonces un atractor ergódico no es necesariamente un atractor topológico. En el Ejercicio 5.4.6 se muestra un ejemplo de atractor ergódico que no es topológico.

Por otra parte, un atractor topológico satisface la igualdad (5.1) para todo $x \in V$. Luego satisface (5.6). Pero no necesariamente satisface la condición de existencia de una medida SRB ergódica para la cual valga la igualdad (5.7). Entonces un atractor topológico no es necesariamente un atractor ergódico. En el ejercicio 5.4.5 se muestra un ejemplo de atractor topológico que no es ergódico.

Ejercicio 5.4.5. Sea en $Q = [0, 1]^2$ la aplicación $T(x, y) = ((1/2)x, y)$. Probar que el segmento $K = \{0\} \times [0, 1]$ es un atractor topológico pero no es un atractor ergódico. Sugerencia: para probar que K no es atractor ergódico, demostrar que toda medida μ ergódica es delta de Dirac en un punto fijo y que el conjunto de puntos $x \in Q$ para los cuales vale la igualdad (5.6) tiene medida de Lebesgue cero.

Ejercicio 5.4.6. Sea en el disco $D = \{z \in \mathbb{C} : |z| \leq 1\}$ la transformación $f : D \mapsto D$ que deja fijo el origen y tal que para todo $z \neq 0$ expresado en polares, $f(z)$ está dado por la siguiente fórmula:

$$f(\rho e^{i\varphi}) = \widehat{\rho} e^{i\widehat{\varphi}},$$

donde

$$\begin{aligned} \widehat{\varphi} &= \frac{3}{2}\varphi \text{ si } 0 \leq \varphi < \pi(\text{mód}2\pi), \\ \widehat{\varphi} &= \pi + \frac{\varphi}{2} \text{ si } \pi \leq \varphi < 2\pi(\text{mód}2\pi), \\ \widehat{\rho} &= \left(1 - \frac{\widehat{\varphi}}{2\pi}\right)\rho. \end{aligned}$$

(a) Bosquejar las órbitas (Sugerencia: los puntos $\varphi = 0$ son fijos, y las demás órbitas son tales que el argumento tiende a 2π por abajo y el módulo tiende a cero).

(b) Probar que toda órbita con punto inicial z que no se encuentre en el segmento $\varphi = 0$ es tal que la distancia al origen tiende a cero y la sucesión de promedios de Birkhoff de las funciones continuas ψ tiende a $\int \psi d\delta_0$, donde δ_0 es la Delta de Dirac soportada en el origen.

(c) Concluir que $K = \{0\}$ es un atractor ergódico pero no es atractor topológico.

Nota: En este ejemplo f es discontinuo en el semieje real positivo. Sin embargo, puede construirse un ejemplo continuo con bosquejo similar de órbitas.

Observación 5.4.7. Por un lado tenemos el criterio de observabilidad de la cuenca local V del atractor. O bien la atracción se produce para todo estado inicial x en el abierto V (criterio de observabilidad topológica) o bien la atracción se produce solo para Lebesgue c.t.p. $x \in V$ (criterio de observabilidad Lebesgue-medible). El criterio de observabilidad de la cuenca es entonces el *criterio con el cual se eligen los estados iniciales* para observar a dónde son atraídas las órbitas.

Definición 5.4.8. Atracción topológica En forma independiente al criterio de observabilidad de los estados iniciales en la cuenca local, la atracción en sí misma, definida por la igualdad (5.1) para los atractores topológicos, y por la igualdad (5.6) para los atractores ergódicos se llama *atracción topológica*. Esta significa, por definición, que el límite de la distancia al atractor K existe y es cero, desde los puntos iniciales elegidos según el criterio de observabilidad que corresponda.

Definición 5.4.9. Atracción estadística. Esta significa, por definición, que los promedios de Birkhoff de las funciones continuas convergen (o por lo menos, en un contexto más general, tienen subsucesiones convergentes) al valor esperado respecto a alguna medida invariante μ soportada en el atractor K , desde los puntos iniciales elegidos según el criterio de observabilidad que corresponda. En un contexto más general, μ no es necesariamente ergódica.

Lo usual es que cuando se estudia la atracción estadística, el criterio de selección de puntos iniciales sea el de observabilidad medible Lebesgue c.t.p.

En los próximos ejemplos veremos casos particulares de existencia y de no existencia de atractor ergódico.

Ejemplo 5.4.10. En el tent map del intervalo, todo el intervalo es un atractor topológico y ergódico a la vez, cuya medida ergódica SRB (o física) es la medida de Lebesgue (ver por ejemplo [Bu 2011]).

Ejemplo 5.4.11. En el flujo Polo Norte-Polo Sur, el Polo Sur es atractor ergódico y topológico a la vez, cuya media ergódica SRB (o física) es la delta de Dirac soportada en el Polo Sur.

Ejemplo 5.4.12. .

Contraejemplo: Rotación irracional de la esfera.

En este ejemplo no existen atractores ergódicos. Sea $f : S^2 \mapsto S^2$ continua, definida en la superficie esférica S^2 por la siguiente parametrización en coordenadas “esféricas”:
 $S^2 : \{(x, y, z) \in \mathbb{R}^3 : x = \cos \varphi \sin \theta, y = \sin \varphi \sin \theta, z = \cos \theta, 0 \leq \varphi \leq 2\pi, 0 \leq \theta \leq \pi\}$

f está definida por las siguientes ecuaciones en coordenadas esféricas:

$$f(\varphi, \theta) = (\widehat{\varphi}, \widehat{\theta}) \text{ donde:}$$

$$\begin{aligned}\widehat{\theta} &= \theta \\ \widehat{\varphi} &= \varphi + a\end{aligned}$$

siendo $a \in (0, 2\pi)$ una constante tal que $a/2\pi$ es irracional.

Es inmediato chequear que los polos $C_0 = \{\theta = 0\}$ y $C_\pi = \{\theta = \pi\}$ son puntos fijos por f , y que cada circunferencia C_θ , con θ constante diferente de 0 y de π , es invariante por f . Además, $f|_{C_\theta}$ es una rotación irracional si $\theta \neq 0, \pi$. Por lo tanto toda órbita por f es densa en la sección C_θ donde está contenida. Deducimos que no hay órbitas densas en S^2 . Luego, $f : S^2 \mapsto S^2$ no es transitivo.

Sea m la medida de Lebesgue bidimensional en la esfera S^2 (normalizada para que sea una probabilidad, es decir, dividimos la medida de Lebesgue en la esfera, entre el área de toda la esfera, para que $m(S^2) = 1$). Esta medida es invariante con f , pues el Jacobiano $|\det df|$ es idénticamente igual a 1. Sin embargo m no es ergódica, pues m es positiva sobre abiertos pero f no es transitivo.

Afirmación: *No existen atractores ergódicos para la rotación irracional f en la esfera S^2 .*

Demostración. Por absurdo, supongamos que existe un atractor ergódico K y llamemos V a su cuenca local de atracción. Entonces K es compacto no vacío e invariante por f y V es abierto que contiene a K . Sea $p \in V$. Vimos que $\omega(p) = C_{\theta_p}$ donde C_{θ_p} es la sección horizontal de la esfera que contiene al punto p . De la igualdad (5.6) deducimos $\omega(p) \subset K$ para Lebesgue c.t.p. $p \in V$. Entonces $C_{\theta_p} \subset K$, y como $p \in C_{\theta_p}$ deducimos que $p \in K$. Hemos probado que, bajo la hipótesis de absurdo, Lebesgue c.t.p. $p \in V$ está contenido en K . Como K es compacto, tenemos $V \subset K$. Pero por definición de atractor ergódico $K \subset V$. Entonces $K = V$ es compacto y abierto a la vez, y es no vacío. Como S^2 es conexo concluimos que, si existiera un atractor ergódico K , éste sería toda la esfera $K = S^2$.

Por (5.7), el promedio temporal asintótico $\widetilde{\psi}(p)$ debería ser constante para m -c.t.p. $p \in V = S^2$, para cualquier función continua $\psi : S^2 \mapsto \mathbb{R}$. Como en este ejemplo m es invariante, entonces m sería una medida ergódica, contradiciendo que m es positiva sobre abiertos y f no es transitivo. \square

Nota: En este ejemplo 5.4.12 toda la esfera es un atractor topológico (en realidad es el único atractor topológico). Luego, este ejemplo prueba que no todo atractor topológico es un atractor ergódico.

Ejemplo 5.4.13. Sea $f = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}_{\text{mód}(1,1)}$ en el toro \mathbb{T}^2 . Es ergódico respecto a

la medida de Lebesgue m . Entonces todo el toro \mathbb{T}^2 es un atractor ergódico y m es la medida ergódica SRB o física.

5.5. Atracción estadística y medidas SRB o físicas

Independientemente de si existe o no un atractor ergódico, dada una medida de probabilidad μ (no necesariamente ergódica ni invariante) definimos el siguiente conjunto $B(\mu)$ en el espacio X donde actúa f :

Definición 5.5.1. Cuenca de atracción estadística

Sea $f : X \mapsto X$ una transformación continua en un espacio métrico compacto. Sea $\mu \in \mathcal{M}$ una probabilidad.

Se llama *cuenca de atracción estadística* de μ al siguiente conjunto:

$$B(\mu) := \left\{ x \in X : \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)) = \int \psi d\mu \quad \forall \psi \in C^0(X, \mathbb{R}) \right\}.$$

Usando la caracterización de la topología débil* en el espacio \mathcal{M} obtenemos:

$$B(\mu) = \left\{ x \in X : \lim_{n \rightarrow +\infty} \sigma_{n,x} = \mu \right\}, \quad \text{dónde} \quad (5.8)$$

$$\sigma_{n,x} := \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)}. \quad (5.9)$$

Observación: La cuenca de atracción estadística de cualquier medida de probabilidad μ es un conjunto medible. En efecto, tomando una familia numerable $\{\psi_i\}_{i \in \mathbb{N}}$ de funciones reales continuas $\psi : X \mapsto [0, 1]$ que sea denso en el espacio $C^0(X, [0, 1])$, la igualdad dentro de la definición de la cuenca $B(\mu)$, se verifica para toda $\psi \in C^0(X, \mathbb{R})$ si y solo si se verifica para ψ_i para todo $i \in \mathbb{N}$. Para cada i fijo, la igualdad se satisface para un conjunto medible, pues el límite puntual de una sucesión de funciones continuas es medible. Luego, $B(\mu)$ es la intersección numerable de conjuntos medibles; es decir, es medible.

Definición 5.5.2. Probabilidades empíricas

Se observa que $\sigma_{n,x}$, definida por la igualdad (5.9), es una medida de probabilidad, para todo $n \geq 1$ y para todo $x \in X$ (en general $\sigma_{n,x}$ no es f invariante). Es la probabilidad promedio concentrada en los puntos de tramos finitos de la órbita futura por x .

Las probabilidades $\sigma_{n,x}$ (para cualquier $n \geq 1$, con $x \in X$ fijo) se llaman *probabilidades empíricas* de la órbita futura por x . Este nombre proviene de que un experimentador no puede observar los promedios temporales asintóticos (en tiempo infinito), sino que observa los promedios hasta tiempo n finito. Estos promedios se pueden calcular como el valor esperado integrando respecto a las probabilidades empíricas. Más precisamente, debido a la igualdad (5.9) tenemos:

$$\frac{1}{n} \sum_{j=0}^{n-1} \psi(f^j(x)) = \int \psi d\sigma_{n,x} \quad \forall \psi \in C^0(X, \mathbb{R}).$$

Observación 5.5.3. Si μ es una probabilidad tal que su cuenca de atracción estadística $B(\mu)$ es no vacía, entonces μ es invariante con T (Ejercicio 5.5.5 a)). Se puede caracterizar a una medida ergódica por la siguiente afirmación (Ejercicio 5.5.5 b)):

$$\mu \text{ es invariante y ergódica si y solo si } \mu(B(\mu)) = 1.$$

Para cualquier medida de probabilidad μ no invariante, y para cualquier medida de probabilidad invariante no ergódica, se cumple $\mu(B(\mu)) = 0$ (Ejercicio 5.5.5 c)), aunque $B(\mu)$ puede ser no vacío.

Observación 5.5.4. La cuenca de atracción estadística $B(\mu)$ de una medida invariante μ no ergódica, puede ser no vacía, y además puede cubrir Lebesgue c.t.p.

En efecto, un ejemplo de homeomorfismo no diferenciable en un disco bidimensional compacto D (que es adaptación de un difeomorfismo en el disco compacto, atribuido a Bowen) exhibe una medida invariante y no ergódica μ tal que $B(\mu) \neq \emptyset$ (ver [Cat 2012, Example 7.2, case B]). Más aún, en este ejemplo, $B(\mu)$ contiene Lebesgue casi todo punto $x \in D$ a pesar que $\mu(B(\mu)) = 0$, es decir, el soporte de μ tiene medida de Lebesgue nula. Luego, en este ejemplo, existe una medida SRB μ no ergódica, la medida de Lebesgue m no es invariante, existe un atractor topológico no ergódico cuya cuenca de atracción topológica es abierta y cubre Lebesgue casi todo punto, y no existen atractores ergódicos.

En la versión C^1 del ejemplo atribuido a Bowen (ver [Go-K1 2007]), se cumple $m(B(\mu)) = 0$ para toda medida de probabilidad μ invariante soportada en el atractor topológico. Luego, no existen medidas SRB. Además en este ejemplo, toda medida invariante es hiperbólica (tiene exponentes de Lyapunov no nulos) y tiene exponentes de Lyapunov positivos. Como no existen medidas SRB, no existen atractores ergódicos. Sin embargo, existe un atractor topológico cuya cuenca es abierta y cubre Lebesgue casi todo punto.

Ejercicio 5.5.5. Sea $f : X \mapsto X$ continua en un espacio métrico X . Denotamos \mathcal{M} el espacio de todas las probabilidades de Borel (no necesariamente f -invariantes), dotado de la topología débil*. Sea $\mu \in \mathcal{M}$.

- a) Probar que si la cuenca de atracción estadística $B(\mu)$ es no vacía, entonces μ es f -invariante.
- b) Probar que μ es invariante y ergódica si y solo si $\mu(B(\mu)) = 1$.
- c) Probar que si μ no es invariante, o si μ es invariante pero no ergódica, entonces

$$\mu(B(\mu)) = 0.$$

(Sugerencias: $B(\mu)$ es siempre un conjunto f -invariante. En el caso μ invariante, aplicar el teorema de descomposición ergódica).

De la parte c) del Ejercicio 5.5.5 deducimos:

Una medida de probabilidad μ es f -invariante y ergódica si y solo si

$$\mu(B(\mu)) = 1,$$

donde $B(\mu)$ denota la cuenca de atracción estadística de μ .

Definición 5.5.6. Medidas SRB o físicas

(Sinai [Si 1972]- Ruelle [Rue 1976]-Bowen [Bow 1971, Bow-Rue 1975])

Sea $f : M \mapsto M$ continua en una variedad compacta y riemanniana M . Sea m la medida de Lebesgue en M , normalizada para que sea una probabilidad: $m(M) = 1$. (En general, m no es necesariamente f -invariante.) Sea $\mu \in \mathcal{M}$.

Decimos que la medida de probabilidad μ es *SRB* (*Sinai-Ruelle-Bowen*) o, indistintamente, que es *física*, si

$$m(B(\mu)) > 0,$$

donde $B(\mu)$ es la cuenca de atracción estadística de μ , definida en 5.5.1.

Nota: Si μ es SRB, entonces $B(\mu) \neq \emptyset$, y por lo observado en la parte (a) del Ejercicio 5.5.5, $\mu \in \mathcal{M}_f$ (es decir las medidas SRB son invariantes con f).

De acuerdo a la Definición 5.5.6, una medida SRB puede no ser ergódica. En efecto, en el ejemplo mencionado en 5.5.4, que es adaptación C^0 de un ejemplo atribuido a Bowen, existe medida SRB no ergódica.

Sobre la nomenclatura “SRB” y “física”. La Definición 5.5.6 de medida SRB no es adoptada por todos los autores. En general se utiliza esta definición que solo requiere $m(B(\mu)) > 0$, solo para llamar *física* a la medida μ . Pero, para una parte importante de matemáticos (por ejemplo [Bu 2011], [You 2002]), medida SRB no es sinónimo de medida física. Llaman física a cualquier probabilidad μ cuya cuenca de atracción estadística $B(\mu)$ tenga medida de Lebesgue positiva. Pero para llamar SRB a μ , requieren que la probabilidad μ sea ergódica, tenga exponentes de Lyapunov positivos, y tenga medidas condicionales inestables absolutamente continuas (veremos más adelante qué significa esta propiedad adicional, al introducir las medidas de Gibbs, mediante las Definiciones 6.1.9, 6.2.1 y 6.2.3). Esto es debido a que, entre otros motivos, en el contexto restringido de los difeomorfismos de clase $C^{1+\alpha}$ uniformemente hiperbólicos, ambas definiciones son equivalentes (ver Teorema 6.3.4).

Nosotros adoptamos la definición de, por ejemplo, [Bon-Dia-Via 2005, Definition 1.9] o [Jo 2005, Definition 22]. Aquí, en la Definición 5.5.6, *no estamos asumiendo ninguna condición adicional a la fisicalidad de la medida para llamarla SRB o física (ni la ergodicidad, ni la positividad de exponentes de Lyapunov, ni la continuidad absoluta condicionada inestable). En resumen, usamos ambas palabras “SRB” ó “física”, como sinónimos.*

Capítulo 6

Teoría de Pesin

6.1. Desintegración en medidas condicionales

En esta sección asumimos que X es un espacio métrico compacto, provisto de la sigma-álgebra de Borel \mathcal{B} , y que μ es una medida de probabilidad en (X, \mathcal{B}) . Expondremos el enunciado de un resultado de la teoría de la medida (el Teorema de Rohlin), válido aunque no exista una dinámica definida en el espacio X . En la sección siguiente veremos el uso del Teorema de Rohlin, junto con la Teoría de Pesin, en sistemas dinámicos de clase C^1 -más Hölder (cuando X tiene además, una estructura de variedad).

Definición 6.1.1. Partición medible

Se llama *partición* en X a una colección (puede ser finita, infinita numerable o infinita no numerable)

$$\mathcal{P} := \{W(x)\}_{x \in X}$$

de subconjuntos $W(x) \subset X$ (pueden ser medibles o no medibles) tales que:

- (a) $x \in W(x)$ para todo $x \in X$ (esto implica que $\bigcup_{x \in X} W(x) = X$)
- (b) Los conjuntos $W(x)$ son dos a dos disjuntos; es decir, para toda pareja de puntos $x \neq y$ en X , ó bien $W(x) = W(y)$ ó bien $W_x \cap W_y = \emptyset$.

La partición \mathcal{P} se dice *medible*, si sus piezas $W(x)$ son todas medibles y \mathcal{P} está generada por una colección numerable de particiones finitas con piezas medibles. Esto es:

- (c) Existe una colección numerable $\{\mathcal{P}_n\}_{n \geq 1}$ donde:

$$(c1) \quad \mathcal{P}_n := \{E_{n,j}\}_{1 \leq j \leq k_n}$$

es una partición finita de X para todo $n \geq 1$, con exactamente k_n piezas $E_{n,j} \subset X$ medibles (disjuntas dos a dos al cambiar j con n fijo, y cuya unión en j es X para todo n fijo).

$$(c2) \quad \mathcal{P}_{n+1} \prec \mathcal{P}_n \quad \forall n \geq 1.$$

Esto significa que cada pieza de la partición \mathcal{P}_{n+1} está contenida en alguna pieza de la partición \mathcal{P}_n (Se dice que \mathcal{P}_{n+1} es *más fina* que \mathcal{P}_n y se denota $\mathcal{P}_{n+1} \prec \mathcal{P}_n$).

$$(c3) \quad \forall x \in X : W(x) = \bigcap_{n \geq 1} E_{n, j_n(x)} \in \mathcal{P}$$

donde $1 \leq j_n(x) \leq k_n$ es el único índice, para cada n fijo, tal que $x \in E_{n, j_n(x)} \quad \forall n \geq 1$. Observar que esta última condición implica que para todo $x \in X$ y para todos $n \geq 1$ y $1 \leq j \leq k_n$, o bien $x \notin E_{n, j}$, o bien $W(x) \subset E_{n, j}$. En otras palabras, cada conjunto medible $E_{n, j}$ están formado por piezas enteras de la partición \mathcal{P} . Dicho de otra forma: $\mathcal{P} \prec \mathcal{P}_n$ para todo $n \geq 1$.

La condición (3) establece una condición más fuerte que $\mathcal{P} \prec \mathcal{P}_n$ para todo $n \geq 1$: La “intersección” decreciente de las particiones \mathcal{P}_n es \mathcal{P} . Esto se denota como

$$\mathcal{P} = \bigvee_{n=1}^{+\infty} \mathcal{P}_n,$$

donde para cualquier pareja \mathcal{R}, \mathcal{S} de particiones finitas se define

$$\mathcal{R} \vee \mathcal{S} := \{R \cap S : R \in \mathcal{R}, S \in \mathcal{S}\}.$$

La condición (c3) (junto con la propiedad de medibilidad de las piezas de cada partición \mathcal{P}_n) implica, en particular, la medibilidad de las piezas de \mathcal{P} . Sin embargo el recíproco es falso, como veremos en el ejemplo del Ejercicio 6.1.5: existen particiones \mathcal{P} cuyas piezas son todas medibles y que no cumplen la condición (c). Estas particiones, no son particiones medibles, de acuerdo a esta definición, a pesar que sus piezas son todas medibles, porque no está generada por ninguna colección numerable de particiones con piezas medibles.

Ejercicio 6.1.2. Probar que si \mathcal{P} es una partición con piezas medibles, y si \mathcal{P} es finita o infinita numerable, entonces \mathcal{P} es una partición medible.

En los siguientes ejercicios veremos ejemplos de particiones medibles y no medibles, con una cantidad no numerable de piezas:

Ejercicio 6.1.3. Sea $X = [0, 1]^2$. Para cada $(x_0, y_0) \in X$ se denota $S(x_0, y_0) := \{(x, y) \in X : x = x_0\}$ al segmento vertical de X que se obtiene seccionando X con la recta $x = x_0$ constante.

(a) Probar que la foliación $\mathcal{F} := \{S(x, y)\}_{(x, y) \in X}$ es una partición medible de X . (Sugerencia: considerar la colección numerable $\{E_{n, i}\}_{n \geq 1, 1 \leq i \leq n}$ de borelianos $E_{n, i} := [(i-1)/n, i/n) \times [0, 1]$ si $i < n$, $E_{n, n} := [(n-1)/n, 1] \times [0, 1]$.)

(b) En el intervalo $[0, 1]$ sea $K = \bigcap_{n=0}^{+\infty} K_n$ el conjunto de Cantor de los “tercios mitad”, i.e. $K_0 = [0, 1]$ y K_{n+1} se obtiene de K_n retirando de cada intervalo cerrado I que forma a K_n el subintervalo abierto con punto medio en el punto medio de I y con longitud $(1/3)\text{long}(I)$. Sea en X , la partición $\{\mathcal{P}\} = \{P(x, y)\}_{(x, y) \in X}$ definida por $P(x_0, y_0) = S(x_0, y_0)$ si $x_0 \in K$, y $P(x_0, y_0) := \{(x, y) \in X : x \notin K\}$ si $x_0 \notin K$. Probar que \mathcal{P} es una partición medible en X .

(c) Sea $\xi : X \mapsto Y$ un homeomorfismo. Sea en Y la siguiente foliación \mathcal{G} :

$$\mathcal{G} := \{\xi(S(\xi^{-1}(x, y)))\}_{(x, y) \in Y}.$$

(Se dice que ξ^{-1} es una trivialización C^0 de la foliación \mathcal{G}). Probar que \mathcal{G} es una partición medible.

(d) Sean X e Y espacios métricos compactos, y sea en $X \times Y$ la partición $\mathcal{F} = \{S(x, y)\}_{(x, y) \in X \times Y}$ por secciones verticales $S(x_0, y_0) := \{(x, y) \in X \times Y : x = x_0\}$. Probar que \mathcal{F} es una partición medible.

Ejercicio 6.1.4. Sean (X, \mathcal{A}) e (Y, \mathcal{B}) dos espacios medibles. Sea $\xi : X \mapsto Y$ una transformación bimedible. Sea en X una partición $\mathcal{P} = \{P(x)\}_{x \in X}$ cualquiera (no necesariamente medible).

Sea en Y la partición $\mathcal{Q} := \{\xi(S(\xi^{-1}(y)))\}_{y \in Y}$. Probar que \mathcal{P} es partición medible si y solo si \mathcal{Q} lo es.

Ejercicio 6.1.5. Sea $f : \text{Diff}^\infty(\mathbb{T}^2)$ en el toro \mathbb{T}^2 el autormorfismo lineal $f = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ (mód. \mathbb{Z}^2). Sea $W^u(p)$ la variedad inestable (global) por cada punto $p \in \mathbb{T}^2$, i.e.:

$$W^u(p) = \{q \in \mathbb{T}^2 : \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(p), f^{-n}(q)) = 0\}.$$

(a) Probar que $W^u(p)$ es un conjunto medible para todo $p \in \mathbb{T}^2$.

(b) Probar que la partición $\mathcal{P} := \{W^u(p)\}_{p \in \mathbb{T}^2}$ no es medible.

Sugerencia: Considerar la medida de Lebesgue m en el toro. Chequear que $m(W^u(p)) = 0$ para todo p . Sea $E \subset \mathbb{T}^2$ medible con $m(E) > 0$ y tal que si $p \in E$ entonces $W(p) \subset E$. Probar que $m(E) = 1$ (Recordar que $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Tomar un levantado \widehat{E} a \mathbb{R}^2 del conjunto E : si $\widehat{p} \in \widehat{E}$ entonces toda la recta con dirección inestable en \mathbb{R}^2 que pasa por \widehat{p} está contenida en \widehat{E} . Mirar la intersección de \widehat{E} con cada recta horizontal de altura entera en \mathbb{R}^2 . Proyectar todas esas intersecciones, siguiendo las verticales en \mathbb{R}^2 , sobre una sola recta horizontal. Observar que esa proyección (mód. 1) en $[0, 1] = \mathbb{R}/\mathbb{Z}$, es invariante con la rotación irracional y tiene medida de Lebesgue positiva en el intervalo.) De la propiedad $m(E) = 0$ ó $m(E) = 1$, deducir que para m.c.t.p. $p \in \mathbb{T}^2$ no se puede verificar la condición (c3) de la Definición 6.1.1) de partición medible.

Ejercicio 6.1.6. (a) Probar que si \mathcal{P} y \mathcal{Q} son dos particiones medibles, entonces $\mathcal{P} \vee \mathcal{Q}$ es una partición medible. (Sugerencia, construir $\mathcal{P}_n \vee \mathcal{Q}_n$ donde \mathcal{P}_n y \mathcal{Q}_n son las particiones finitas que satisfacen, para \mathcal{P} y \mathcal{Q} respectivamente, la condición (c) de la Definición 6.1.1.)

(b) Probar que si \mathcal{Q} es medible, entonces \mathcal{P} es medible si y solo si $\mathcal{P} \vee \mathcal{Q}$ es medible.

Espacio cociente por una partición medible.

Dada una partición medible $\mathcal{P} = \{W(x)\}_{x \in X}$ de (X, \mathcal{B}) consideramos el conjunto cociente X/\sim por la relación de equivalencia $x \sim y$ si y solo si $W(x) = W(y)$. Denotando $[x]$ a la clase de equivalencia que contiene a x , tenemos

$$X/\sim = \{[x]\}_{x \in X} \text{ donde } [x] = W(x)$$

Dotamos el conjunto X/\sim de la estructura medible cociente, definiendo la sigma-álgebra:

$$(X/\sim, \mathcal{A}/\sim) : \hat{A} \in \mathcal{A}/\sim \quad \text{si y solo si} \\ A := \{x \in X : W(x) \in \hat{A}\} \in \mathcal{A}.$$

Para cada $W(x) \in X/\sim$, es decir, en cada pieza $W(x)$ de la partición medible \mathcal{P} , definimos la sigma-álgebra $\mathcal{A}_{W(x)}$ que resulta de restringir la sigma-álgebra \mathcal{A} a $W(x)$. Esto es:

$$(W(x), \mathcal{A}_{W(x)}) : \forall A \subset W(x), \quad A \in \mathcal{A}_{W(x)} \text{ si y solo si } A \in \mathcal{A}.$$

Teorema 6.1.7. Desintegración Medible (Teorema de Rohlin)

Sea X un espacio métrico compacto y sea \mathcal{A} la sigma-álgebra de Borel. Sea $\{W(x)\}_{x \in X}$ una partición medible. Sea μ una medida de probabilidad en (X, \mathcal{A}) . Entonces:

(i) Existe una medida de probabilidad $\hat{\mu}$ en el espacio medible cociente X/\sim , tal que para $\hat{\mu}$ -c.t.p. $W(x) \in X/\sim$, existe una medida de probabilidad $\mu^{W(x)}$ en el espacio medible $(W(x), \mathcal{A}_{W(x)})$, tal que:

- Para todo $A \in \mathcal{A}$:

$$\int_X \chi_A d\mu = \int_{[x] \in X/\sim} \left(\int_{y \in [x]=W(x)} \chi_A(y) d\mu^{W(x)} \right) d\hat{\mu}. \quad (6.1)$$

- Más en general, para toda $\psi \in L^1(\mu)$:

$$\int_X \psi d\mu = \int_{[x] \in X/\sim} d\hat{\mu} \left(\int_{y \in [x]=W(x)} \psi(y) d\mu^{W(x)} \right). \quad (6.2)$$

(ii) La medida de probabilidad $\hat{\mu}$, y para $\hat{\mu}$ -casi todas las piezas $W(x) \in X/\sim$, las medidas de probabilidad $\mu^{W(x)}$ que verifican las igualdades (6.1) y (6.2), son únicas.

La prueba original del Teorema de Rohlin se encuentra en [Roh 1966], o también en [Roh 1962]. La demostración se encuentra reformulada también, por ejemplo, en [Led 1984] o en [Via 2012b].

El teorema de Rohlin generaliza a un contexto medible, y para cualquier medida de probabilidad, el teorema de Fubini que vale para la medida de Lebesgue en un rectángulo de $\mathbb{R}^a \times \mathbb{R}^b$. En efecto, por el Teorema de Fubini en un rectángulo, la integral con respecto a la medida de Lebesgue se obtiene como la integral doble en una sección “horizontal” del rectángulo, de las integrales sobre las secciones “verticales,” con respecto a las medidas de Lebesgue a lo largo de dichas secciones, respectivamente.

Definición 6.1.8. Medidas condicionales

En las hipótesis y conclusiones del Teorema 6.1.7 de Rohlin, las medidas μ^W se llaman *medidas condicionales de μ a lo largo de las piezas W de la partición \mathcal{P}* .

También las llamamos, en breve *medidas \mathcal{P} -condicionales de μ* , o si la partición está clara en el contexto, simplemente *medidas condicionales de μ* .

Observar que cada medida condicionada μ^W está soportada en una pieza W , es una probabilidad (es decir $\mu^W(W) = 1$). Además, μ^W está definida para $\hat{\mu}$ -casi todas las piezas W de la partición \mathcal{P} , y no necesariamente para todas las piezas.

Continuidad absoluta de las medidas condicionales.

Cuando el espacio métrico X tiene una estructura de variedad compacta y riemanniana M (es decir $X = M$), y si las piezas W de la partición medible \mathcal{P} son subvariedades de M , se considera, como medida privilegiada a lo largo de cada una de estas subvariedades, la medida de Lebesgue en W , que denotamos m^W .

En un contexto general, las medidas condicionales μ^W de μ a lo largo de las piezas W de la partición \mathcal{P} , podrían no tener relación con las medidas de Lebesgue m^W a lo largo de estas piezas. Sin embargo, dentro de la Teoría de Pesin que veremos en la próxima sección, tienen especial importancia las medidas μ que cumplen la siguiente definición:

Definición 6.1.9. Se dice que μ tiene medidas \mathcal{P} -condicionales absolutamente continuas cuando para $\hat{\mu}$ - casi toda pieza W de la partición \mathcal{P} , se cumple:

$$\mu^W \ll m^W, \quad \text{i.e. } \mu^W(A) = 0 \text{ si } m^W(A) = 0,$$

donde $A \subset W$ es medible, W es una subvariedad u -dimensional de M , y m^W es la medida de Lebesgue u -dimensional a lo largo de la subvariedad W .

6.2. Medidas de Gibbs

Definición 6.2.1.**Medidas condicionales inestables**

Sea M una variedad compacta y riemanniana y sea $f : \text{Diff}^1(M)$ tal que el siguiente conjunto

$$W^u(x) := \{y \in M : \lim_{n \rightarrow +\infty} \text{dist}(f^{-n}(x), f^{-n}(y)) = 0\} \quad (6.3)$$

es, por hipótesis, una subvariedad C^1 inmersa en M para μ -casi todo punto $x \in M$, para cierta medida de probabilidad μ invariante con f . Bajo esta hipótesis $W^u(x)$ se llama *subvariedad inestable global* por el punto x . Debido a la continuidad de f , es inmediato chequear que $f^{-1}(W^u(x)) = W^u(f^{-1}(x))$ para todo $x \in M$.

Sea B_δ una bola (compacta) de radio suficientemente pequeño con $\mu(B_\delta) > 0$, y sea en ella la partición $\mathcal{F}_\delta^u = \{W_\delta^u(x)\}_{x \in M}$ definida como sigue:

- $W_\delta^u(x) = c.c.x(W^u(x) \cap B_\delta(x))$ es la componente conexa que contiene al punto x de la intersección $W^u(x) \cap B_\delta(x)$, para todo punto x tal que $W^u(x)$ es una subvariedad C^1 -inmersa en M (es decir para μ -c.t.p $x \in M$). Por hipótesis, esta subvariedad $W_\delta^u(x)$ está C^1 -encajada en M para μ -c.t.p $x \in B_\delta$. Se llama *subvariedad inestable local* por el punto x .

- Abusando de la notación, para los restantes puntos $x \in B_\delta$, denotamos $W_\delta^u(x) := \{y \in B_\delta : \text{no existe subvariedad inestable por el punto } y\}$. Por construcción, esta única pieza de la partición \mathcal{F}^u (que no es necesariamente una variedad) tiene μ -medida cero. Rescalando μ para que $\mu(B_\delta) = 1$, tenemos lo siguiente:

Si la partición $\mathcal{F}^u = \{W_\delta^u(p)\}_{p \in B_\delta}$ de subvariedades inestables locales así construida, fuera una partición μ -medible, llamamos *medidas condicionales inestables* de μ , a las

medidas condicionales $\mu^{W_\delta^u(p)}$ de μ a lo largo de las piezas $W_\delta^u(p)$ de esa partición, es decir a lo largo de las variedades inestables locales.

Por simplicidad en la notación escribiremos $\mu^u = \mu^{W_\delta^u}$ a las medidas condicionales inestables, y $m^u = m^{W_\delta^u}$ a las medidas de Lebesgue a lo largo de las subvariedades inestables W_δ^u .

Definición 6.2.2. Medidas condicionales inestables absolutamente continuas. En el contexto de la Definición 6.2.1, una medida μ se dice que tiene *medidas condicionales inestables absolutamente continuas* cuando

$$\mu^u \ll m^u$$

para $\widehat{\mu}$ - casi toda variedad inestable local W_δ^u de la partición \mathcal{F}^u , donde $\widehat{\mu}$ y μ^u son las medidas del Teorema 6.1.7 de desintegración de Rohlin en la partición de variedades inestables locales.

Se recuerda que, por definición, dadas dos medidas ν_1 y ν_2 , se dice que ν_1 es *absolutamente con respecto de ν_2* , y se denota $\nu_1 \ll \nu_2$ cuando para todo boreliano A se cumple

$$\nu_2(A) = 0 \Rightarrow \nu_1(A) = 0.$$

Dos medidas finitas ν_1 y ν_2 cumplen $\nu_1 \ll \nu_2$ si y solo si existe una función $h \in L^1(\nu_2)$, llamada *derivada de Radon-Nikodym* de ν_1 con respecto de ν_2 , tal que para todo conjunto medible A se cumple

$$\nu_1(A) = \int_A h d\nu_2.$$

(Ver por ejemplo [Fo 1984, page 85] o [Rud 1979a, page 113].)

Luego, si μ tiene medidas condicionales inestables absolutamente continuas, entonces para μ -c.t.p. $x \in M$ existe una función $h_x \in L^1(m_x^u)$, donde m_x^u es la medida de Lebesgue a lo largo de la variedad inestable local $W_\delta^u(x)$, tal que la medida condicional inestable μ_x^u a lo largo de $W_\delta^u(x)$ cumple:

$$\mu_x^u(A \cap W_\delta^u(x)) = \int_{y \in W_\delta^u(x)} \chi_A(y) h_x(y) dm_x^u(y)$$

para todo boreliano $A \subset M$.

Definición 6.2.3. Medidas de Gibbs

Sea M una variedad compacta y riemanniana y sea $f \in \text{Diff}^1(M)$. Una medida de probabilidad μ f -invariante se dice que es *medida de Gibbs* cuando cumple:

- (i) Para μ -c.t.p. $x \in M$ el conjunto definido por la igualdad (6.3) es una subvariedad C^1 -inmersa en M (subvariedad inestable global por el punto x).
- (ii) Para $\delta > 0$ suficientemente pequeño, si $B_\delta \subset M$ es una bola compacta de radio δ tal que $\mu(B_\delta) > 0$, entonces la siguiente familia de variedades inestables locales, es una partición μ -medible

$$\mathcal{F}^u := \{W_\delta^u(x)\}_{x \in B_\delta}, \quad \text{donde } W_\delta^u(x) := c.c.x(W^u(x) \cap B_\delta).$$

(c.c. $_x$ denota la componente conexa que contiene al punto x).

(iii) μ tiene medidas condicionales inestables absolutamente continuas, de acuerdo a la definición dada en el último párrafo de 6.2.1

Veamos ahora que la propiedad de tener medidas condicionales absolutamente continuas a lo largo de las piezas de una partición, se transmite de una medida invariante μ a sus componentes ergódicas, y recíprocamente. Por lo tanto una medida invariante μ es de Gibbs, si y solo si sus componentes ergódicas son medidas de Gibbs.

Corolario 6.2.4. del Teorema de Rohlin

Sea $f: X \mapsto X$ continua en el espacio métrico compacto, sea μ una medida de probabilidad f -invariante y sea $\mathcal{P} = \{W(x)\}_{x \in X}$ una partición medible y f -invariante (i.e. $f^{-1}(W(x)) = W(f^{-1}(x))$ para todo $x \in X$).

Entonces:

(a) Las medidas condicionales μ^W de μ a lo largo de las piezas W de \mathcal{P} , y la medida inducida $\hat{\mu}$ por μ en el espacio cociente X/\sim de la partición, son medidas invariantes con f .

(b) $\mu_x^{W(x)} = \mu^{W(x)}$ para μ -c.t.p. $x \in M$, donde $W(x)$ denota la pieza de la partición \mathcal{P} que contiene al punto x , y μ_x denota la componente ergódica de la medida μ a la que pertenece el punto x según el Teorema 4.1.2 (es decir, $\lim_{n \rightarrow +\infty} (1/n) \sum_{j=0}^{n-1} \delta_{f^j(x)} = \mu_x$ en la topología débil estrella).

(c) Si además $X = M$ es una variedad compacta y riemanniana, $f \in \text{Diff}^1(M)$, y $\mathcal{P} = \{W_\delta(x)\}_{x \in M}$ es una partición medible de la bola $B_\delta(x)$ formada por subvariedades $W_\delta(x) \subset C^1$ inmersas en M para μ -c.t.p. $x \in X$, entonces:

- Las medidas condicionales μ^W de μ son absolutamente continuas a lo largo de las subvariedades $W \in \mathcal{P}$ si y solo si para μ -c.t.p. $x \in X$ las medidas condicionales μ_x^W de las componentes ergódicas μ_x de μ , son absolutamente continuas.

- μ es una medida de Gibbs si y solo si las componentes ergódicas μ_x de μ son medidas de Gibbs para μ -c.t.p. $x \in M$.

Demostración. (a) Por el Teorema 6.1.7 de Rohlin y la f -invariancia de μ tenemos, para todo conjunto medible A :

$$\begin{aligned} \mu(A) &= \int_{[x] \in X/\sim} d\hat{\mu} \int_{x \in [x]} \chi_A(x) d\mu^{W(x)} = \\ &= \int_{[x] \in X/\sim} \left(\mu^{W(x)}(W(x) \cap A) \right) d\hat{\mu} = \\ &= \mu(f^{-1}(A)) = \int_{[x] \in X/\sim} d\hat{\mu} \int_{x \in [x]} \chi_{f^{-1}(A)}(x) d\mu^{W(x)} = \\ &= \int_{[x] \in X/\sim} \left(\mu^{W(x)}(W(x) \cap f^{-1}(A)) \right) d\hat{\mu}. \end{aligned} \quad (6.4)$$

Las medidas de probabilidad $\mu^{W(x)}$ están definidas para $\hat{\mu}$ -casi toda pieza de la partición \mathcal{P} . Para las otras piezas tomamos por convención cualquier medida de probabilidad soportada en ellas. De esta forma tenemos definidas $\mu^{W(x)}$ para todo $x \in M$.

Debido a la f -invariancia de las piezas $W \in \mathcal{P}$, se cumple:

$$W(x) \cap f^{-1}(A) = f^{-1}(W(f(x)) \cap A)$$

Definimos la siguiente medida $(\mu^{W(y)})^*$ a lo largo de la pieza $W(y)$ para todo $y \in M$:

$$(\mu^{W(y)})^*(B) := \mu^{W(f^{-1}(y))}(f^{-1}(B)) \quad \forall B \in \mathcal{A},$$

donde \mathcal{A} denota la sigma-álgebra de Borel en X . Entonces:

$$\mu^{W(x)}(W(x) \cap f^{-1}(A)) = (\mu^{W(f(x))})^*(W(f(x)) \cap A),$$

y sustituyendo en (6.4) obtenemos

$$\begin{aligned} \mu(A) &= \int_{[x] \in X/\sim} \left((\mu^{W(f(x))})^*(W(f(x)) \cap A) \right) d\hat{\mu} = \\ &= \int_{[y] \in X/\sim} \left((\mu^{W(y)})^*(W(y) \cap A) \right) d(\hat{\mu})^*, \end{aligned} \quad (6.5)$$

donde $(\hat{\mu})^*$ es la medida de probabilidad en X/\sim definida por

$$\int \varphi d\hat{\mu}^* := \int \varphi \circ f d\hat{\mu} \quad \forall \varphi \in L^1(\hat{\mu}).$$

(En la última igualdad $f : X/\sim \rightarrow X/\sim$ denota la aplicación que lleva la pieza $W \in \mathcal{P}$ en la pieza $f^{-1}(W)$). De (6.5) deducimos

$$\begin{aligned} \mu(A) &= \int_{[y] \in X/\sim} \left((\mu^{W(y)})^*(W(y) \cap A) \right) d(\hat{\mu})^* = \\ &= \int_{[y] \in X/\sim} d\hat{\mu}^* \int_{y \in [y]} \chi_A(y) d(\mu^{W(y)})^* \quad \forall A \in \mathcal{A}. \end{aligned}$$

Luego hemos encontrado otra desintegración de Rohlin de la medida μ con respecto a la partición \mathcal{P} . Por la unicidad de las medidas de probabilidad $\hat{\mu}$ y μ^W de la desintegración de Rohlin, se cumple

$$(\hat{\mu})^* = \hat{\mu}, \quad \mu^{W(x)} = (\mu^{W(f(x))})^*,$$

demostrando la f -invariancia de $\hat{\mu}$ y de μ^W .

(b) Para todo conjunto $A \in \mathcal{A}$, por el Teorema de Descomposición Ergódica (Teorema 4.1.2) tenemos:

$$\mu(A) = \int \mu_x(A) d\mu.$$

Por el Teorema de Descomposición de Rohlin, aplicado a cada medida ergódica μ_x , se cumple:

$$\mu(A) = \int \mu_x(A) d\mu = \int_{x \in X} d\mu \int_{[y] \in X/\sim} d\hat{\mu}_x \int_{y \in [y]} \chi_A d\mu_x^{W(y)} =$$

$$= \int_{[y] \in X/\sim} d\hat{\nu} \int_{y \in [y]} \chi_A d\mu_x^{W(y)}, \quad (6.6)$$

donde la medida de probabilidad $\hat{\nu}$ en el espacio cociente X/\sim , \mathcal{A} está definida por:

$$\int_{X/\sim} \varphi d\hat{\nu} := \int_{x \in X} \left(\int_{[y] \in X/\sim} \varphi([y]) d\hat{\mu}_x \right) d\mu \quad \forall \varphi \in L^1(\hat{\mu}).$$

Luego, la igualdad (6.6) es una descomposición de Rohlin de la medida μ con respecto a la partición \mathcal{P} . Por la unicidad de las probabilidades de la descomposición de Rohlin, concluimos que $\hat{\mu} = \hat{\nu}$ y $\mu_x^W = \mu^W$.

(c) Como consecuencia de la parte (b) $\mu^W \ll m^W$ si y solo si $\mu_x^W \ll m^W$. En el caso particular en que la partición medible \mathcal{P} es la partición de una bola $B \subset M$ en variedades inestables locales, obtenemos que μ es medida de Gibbs si y solo si μ_x lo es para μ -c.t.p. $x \in M$. \square

6.3. Relación entre medidas de Gibbs y SRB

En esta sección asumimos que M es una variedad compacta y riemanniana y que $f \in \text{Diff}^{1+\alpha}(M)$. Algunos de los teoremas que veremos en esta sección se obtienen de resultados de la Teoría de Pesin, bajo la hipótesis de que f es de clase $C^{1+\alpha}$. Pero son falsos si f es solo de clase C^1 . En particular las condiciones de continuidad absoluta de las medidas condicionales inestables (existencia de medida ergódica de Gibbs que es también SRB) no rige para todo $f \in \text{Diff}^1(M)$.

Más adelante, en las secciones posteriores de este capítulo, veremos algunas formas de reformular los resultados de esta sección, generalizando algunas definiciones (mediante la introducción de las medidas SRB-like) y modificando adecuadamente los enunciados, para que sean aplicables a todo difeomorfismo de clase C^1 en la variedad M (y más aún, algunos de ellos son aplicables también a transformaciones continuas $f : M \mapsto M$).

Teorema 6.3.1. de Gibbs ergódica implica SRB o física

Sea $f \in \text{Diff}^{1+\alpha}(M)$ y sea μ una medida invariante hiperbólica.

Si μ es una medida de Gibbs, entonces las componentes ergódicas de μ son medidas SRB o físicas.

En el párrafo 6.5.3 veremos la demostración de este Teorema, reduciéndola a resultados de la Teoría de Pesin. También se puede encontrar la demostración del teorema 6.3.1 en [Pug-Shu 1989] o en [Bon-Dia-Via 2005, Proposition 11.24].

Es cierto también el recíproco del Teorema 6.3.1, bajo hipótesis de $C^{1+\alpha}$ -hiperbolicidad uniforme: si las componentes ergódicas de una medida invariante μ son medidas SRB o físicas, entonces μ es una medida de Gibbs. Más precisamente:

Teorema 6.3.2. SRB ergódica implica de Gibbs

Sea $f \in \text{Diff}^{1+\alpha}(M)$, sea $\Lambda \subset M$ un atractor ergódico uniformemente hiperbólico, y sea μ la medida ergódica SRB o física soportada en Λ (de acuerdo a la Definición 5.4.2 de atractor ergódico).

Entonces μ es medida de Gibbs.

El Teorema 6.3.2 es consecuencia del siguiente Teorema de Pesin-Sinai, que demuestra la existencia de medida SRB ergódica, bajo hipótesis de hiperbolicidad uniforme en el contexto $C^{1+\alpha}$

Teorema 6.3.3. Pesin-Sinai [Pe-Si 1982]

Sea $f \in \text{Diff}^{1+\alpha}(M)$. Sea Λ un atractor topológico uniformemente hiperbólico. Entonces:

Existen medidas SRB-ergódicas soportadas en Λ . Toda medida SRB ergódica soportada en Λ es de Gibbs. Recíprocamente, toda medida ergódica de Gibbs soportada en Λ es SRB. (cf. Teorema 6.3.1).

En el caso particular que f sea Anosov, demostraremos este resultado más adelante en este capítulo (Teorema 6.3.4).

Generalización del Teorema 6.3.3 para difeomorfismos $C^{1+\alpha}$ no uniformemente hiperbólicos

La implicación SRB-ergódica \Rightarrow Gibbs, como en el Teorema 6.3.3 de Pesin-Sinai, rige aún en hipótesis más generales que la hiperbolicidad uniforme que nosotros enunciarnos en ese Teorema, asumiendo siempre que $f \in \text{Diff}^{1+\alpha}(M)$. En efecto, el atractor topológico Λ puede no ser hiperbólico, sino que alcanza que sea *parcialmente hiperbólico* (ver Definición, por ejemplo en [Bon-Dia-Via 2005, Definition B.3, page 289]). Puede ser no uniformemente hiperbólico con singularidades, como el atractor de Lorenz [Pe 1992], por ejemplo.

Para los difeomorfismos parcialmente hiperbólicos, las medidas invariantes no son necesariamente hiperbólicas, sino que puede existir un subespacio de Oseledets con exponente de Lyapunov igual a cero. Sin embargo, puede existir una separación acotada uniformemente lejos de cero, entre los exponentes de Lyapunov positivos y los no positivos.

La demostración de Pesin-Sinai del Teorema 6.3.3 se encuentra en [Pe-Si 1982] para difeomorfismos $C^{1+\alpha}$ -parcialmente hiperbólicos; en particular para los que son uniformemente hiperbólicos ó Anosov. También puede encontrarse la demostración general para difeomorfismos $C^{1+\alpha}$ parcialmente hiperbólicos, en [Bon-Dia-Via 2005, Theorem 11.16].

Observemos que la hiperbolicidad parcial no implica que las medidas invariantes sean hiperbólicas, y por lo tanto, aunque existan medidas de Gibbs ergódicas, no podremos aplicar el Teorema 6.3.1 para deducir que estas son SRB. En general, el problema de existencia de medidas SRB para difeomorfismos parcialmente hiperbólicos, está esencialmente abierto. Recientemente, en [Via-You 2013] se demuestra que los difeomorfismos parcialmente hiperbólicos con dimensión central 1 (1 es la dimensión del

subespacio de Oseledets con exponente de Lyapunov igual a cero), $C^{1+\alpha}$ -genéricamente existe una, y a lo sumo una cantidad finita, de medidas ergódicas SRB cuyas cuencas de atracción estadística cubren Lebesgue c.t.p.

Los difeomorfismos hiperbólicos son un caso particular de los llamados difeomorfismos con *splitting dominado*. La definición, los enunciados y las demostraciones de propiedades dinámicas topológicas de los difeomorfismos con *splitting dominado* se encuentran en [Puj-Sam 2009]. El problema de existencia de medidas SRB para los difeomorfismos con *splitting dominado*, salvo en casos particulares, está esencialmente abierto.

Otro caso en el que la existencia de medidas SRB se ha estudiado, es el de los llamados difeomorfismos *derivados de Anosov*. La demostración de la misma tesis del Teorema 6.3.3 para difeomorfismos derivados de Anosov transitivos con exponente de Lyapunov positivo y de clase $C^{1+\alpha}$, y además la unicidad de su medida SRB, fueron dadas por M.F. Carvalho en [Car 1992] (o también, más detalladamente explicadas, en [Car 1991]).

También existen otras generalizaciones del Teorema 6.3.3 para ciertas clases de difeomorfismos no uniformemente ni parcialmente hiperbólicos ni derivados de Anosov y que no tienen *splitting dominado* (por ejemplo en [Hu 2000], [Enr 1998] y [Cat-Enr 2001]). Estas clases de difeomorfismos tienen C^r regularidad para valores de $r \geq 2$ suficientemente grande y se componen de ciertos difeomorfismos f_1 llamados *casi-Anosov*, que están en el borde de los de Anosov en el espacio $\text{Diff}^r(M)$ para cierto $r > 1$ suficientemente grande. Por definición, un difeomorfismo f_1 es casi Anosov, o *almost Anosov*, si f_1 se obtiene por medio de una isotopía (es decir, por medio de una deformación continua $f_t \in \text{Diff}^r(M)$ para $t \in [0, 1] \subset \mathbb{R}$), tal que f_t es de Anosov transitivo para todo $0 \leq t < 1$. La isotopía en un casi-Anosov debe cumplir la siguiente condición: el *splitting* $T_x M = S_x^t \oplus U_x^t$ es un *splitting* uniformemente hiperbólico de f_t para todo $t \in [0, 1)$ fijo, y para todo $x \in M$. Además, por hipótesis, existe una órbita $o(x_0)$ tal que para todo $x \notin o(x_0)$ existe un *splitting* $S_x^1 \oplus U_x^1$ invariante por f_1 , obtenido como el límite cuando $t \rightarrow 1$ del *splitting* hiperbólico de $S_x^t \oplus U_x^t$. Finalmente, el difeomorfismo f_1 , o bien posee también un *splitting* (no hiperbólico) definido como $S^1 \oplus U^1$, donde $S^1 = \lim_{t \rightarrow 1^-} S^t$, $U^1 = \lim_{t \rightarrow 1^-} U^t$ (cuando existen dichos límites en la órbita de x_0 y son mutuamente transversales); o bien existen dichos límites pero son tangentes; o bien no existen.

La existencia de una única medida de Gibbs ergódica que es SRB, en ejemplos del primer caso de difeomorfismos C^2 casi-Anosov, es demostrada en [Car 1992], donde x_0 es un punto fijo por f_t para todo $0 \leq t \leq 1$, y los exponentes de Lyapunov en $U_{x_0}^1$ (para $t = 1$) son estrictamente positivos. En cambio, en el ejemplo de casi-Anosov también del primer caso, estudiado en [Hu-You 1995], se prueba que no existe ninguna medida de probabilidad de Gibbs, si se toma como foliación inestable, la formada por las C^1 -subvariedades que son tangentes a E_x^1 en todo punto $x \in M$. En este ejemplo, el difeomorfismo f_1 , de clase C^2 casi-Anosov, se construye tomando x_0 en un punto fijo, pero tal que los exponentes de Lyapunov en $U_{x_0}^1$ son nulos. Sin embargo, aún en este ejemplo, en [Hu-You 1995] se demuestra que existe una única medida SRB ergódica y que su cuenca de atracción estadística cubre Lebesgue casi todo punto.

En el segundo caso de difeomorfismos casi-Anosov, en el toro \mathbb{T}^2 , [Enr 1998] prueba

la existencia de la medida de Gibbs ergódica que es SRB, cuando $o(x_0)$ es una órbita no periódica (más precisamente una órbita de tangencia heteroclínica). También en el segundo caso (cuando $S_{x_0}^1$ y $U_{x_0}^1$ son tangentes) y en \mathbb{T}^2 , la existencia y unicidad de una medida SRB ergódica con cuenca de atracción que cubre Lebesgue-c.t.p., es demostrada en [Cat-Enr 2001], cuando x_0 es un punto fijo (no hiperbólico) para f_1 de clase C^3 , y además una de las derivadas segundas parciales de f_1 en x_0 se anula. Este caso incluye los ejemplos de difeomorfismos casi-Anosov en el toro bidimensional introducidos por Lewowicz en [Lew 1980], en los que las direcciones estable e inestable en un punto fijo x_0 para $t < 1$, se hacen tangentes cuando el parámetro t de la isotopía alcanza el valor 1 (sin saber a priori si los exponentes de Lyapunov en casi todo punto diferente de x_0 , son o no son, diferentes de cero).

El tercer caso de difeomorfismos casi-Anosov, en que los límites $S_{x_0}^1$ y $U_{x_0}^1$ no existen, es estudiado en [Hu 2000] cuando x_0 es un punto fijo, f es de clase C^r para $r \geq 2$ suficientemente grande, y se cumplen ciertas hipótesis sobre las derivadas de orden mayor que 1 en x_0 .

Estos resultados parciales, fueron obtenidos en subclases muy particulares de difeomorfismos casi-Anosov. La dificultad grande para demostrar la existencia de medidas SRB, cuando no se tienen a priori hipótesis de existencia de una constante uniforme que separe los exponentes de Lyapunov positivos de los no positivos, provoca que el problema general de existencia de medidas SRB para los casi-Anosov, permanezca aún mayormente abierto.

Ahora re-enunciamos el Teorema 6.3.3 en el caso particular de $f \in \text{Diff}^{1+\alpha}(M)$ difeomorfismo de Anosov. Más precisamente:

Teorema 6.3.4. Sinai

Sea $f \in \text{Diff}^{1+\alpha}(M)$ un difeomorfismo de Anosov. Entonces:

- (a) Existen medidas ergódicas SRB (o físicas).
- (b) Toda medida SRB (o física) es de Gibbs ergódica, y recíprocamente
- (c) La unión de las cuencas de atracción estadística de las medidas SRB cubren Lebesgue c.t.p. $x \in M$.
- (d) Si además f es topológicamente transitivo, entonces existe una única medida SRB, es ergódica y de Gibbs, y su cuenca de atracción estadística cubre Lebesgue c.t.p. $x \in M$.

Demostraremos el Teorema 6.3.4 más adelante en este capítulo, en el párrafo 6.7.2.

Corolario 6.3.5. Sea $f \in \text{Diff}^{1+\alpha}(M)$ de Anosov transitivo que preserva la medida de Lebesgue m . Entonces, m es ergódica, es la única medida SRB de f y es de Gibbs.

Demostración. Por la parte (d) del Teorema 6.3.4, Lebesgue c.t.p. $x \in M$ cumple

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} = \mu,$$

donde μ es la única medida SRB de f , es ergódica y de Gibbs.

Por el Teorema 4.1.2, como la medida de Lebesgue m es invariante, entonces para m -c.t.p. $x \in M$

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} = m_x,$$

donde m_x es la componente ergódica de m a la que pertenece f .

Cuando existe, el límite de una sucesión de medidas de probabilidad en la topología débil estrella, es único (esta es una propiedad de todo espacio métrico). Concluimos que $m_x = \mu$ para m -c.t.p. $x \in M$. Dicho de otra forma, la descomposición ergódica de m tiene una única componente ergódica que es μ . Entonces $m = \mu$, como queríamos demostrar. \square

Observación 6.3.6. En la demostración del Teorema 6.3.4, probaremos además los siguientes resultados:

Las medidas condicionales inestables μ_x^u de cualquier medida μ que sea SRB (para un difeomorfismo de Anosov de clase $C^{1+\alpha}$) no solo son absolutamente continuas respecto a las medidas de Lebesgue inestables m_x^u (a lo largo de las respectivas variedades inestables locales $W_\delta^u(x)$ para μ -c.t.p. $x \in M$), sino que son además equivalentes a estas; es decir:

$$\mu_x^u \ll m_x^u, \quad \mu_x^u \ll m_x^u \quad \mu - \text{c.t.p. } x \in M.$$

Además, la derivada de Radon-Nikodym $d\mu_x^u/dm_x^u$ (es decir la densidad de las medidas condicionales inestables) es una función continua $h_x(y)$ y positiva, y está definida por:

$$\frac{d\mu_x^u}{dm_x^u}(y) = h_x(y) := \prod_{j=0}^{+\infty} \frac{J^u(f^j(x))}{J^u(f^j(y))} \in C_0(M, \mathbb{R}^+) \quad \mu - \text{c.t.p. } x \in M,$$

donde $J^u(x) := |\det(df_x|_{E_x^u})|$ se llama *Jacobiano inestable de f en el punto x* .

Cuando intentamos generalizar el Corolario 6.3.5 a difeomorfismos no uniformemente hiperbólicos que preservan la medida de Lebesgue, el método de demostración que usamos para los Anosov transitivos no funciona si uno no sabe a priori, o demuestra primero, la existencia de una medida SRB:

Conjetura 6.3.7. (Viana [Via 1998]) Sea f un difeomorfismo que preserva la medida de Lebesgue m . Si m es una medida hiperbólica (cf. Definición 3.11.3) entonces existe alguna medida SRB.

La demostración de la conjetura de Viana, o el hallazgo de un contraejemplo, es un problema abierto.

6.4. Sobre la Fórmula de Pesin para la entropía

Como consecuencia de los teoremas 6.3.1 y 6.3.2, la búsqueda de medidas SRB o físicas en el caso de difeomorfismos de clase C^1 más Hölder, se centra en la búsqueda

de medidas de Gibbs, o sea, de medidas invariantes cuyas medidas condicionales inestables sean absolutamente continuas (ver Definiciones 6.2.1 y 6.2.3). Por ese motivo, la caracterización de las medidas invariantes que son medidas de Gibbs, adquiere en el contexto $C^{1+\alpha}$ especial relevancia. Una tal caracterización está dada por la igualdad del siguiente Teorema 6.4.1, llamada *Fórmula de Pesin para la entropía*.

Entropía métrica Para poder enunciar la Fórmula de Pesin, introducimos brevemente el concepto de *entropía métrica* $h_\mu(f)$ de una transformación continua $f : M \mapsto M$ con respecto a una medida de probabilidad μ invariante con f . La definición de entropía métrica, y el estudio de las primeras propiedades que dan su forma de cálculo para difeomorfismos de Anosov, son debidos a Kolmogorov [Ko 1958] y a Sinai [Si 1959]. La definición precisa y sus propiedades, puede encontrarse además, por ejemplo en [Wa 2000, §4.4], [Mañ 1983a, Capítulo 4], [Kat-Has 1995, pag. 168-170], [Si 2007], [Si 1994, pag. 55-76], [Jo 2005, §4.1-4.2], [Ke 1998, Chapter 3], o [Ki 2011], entre muchos otros textos que tratan matemáticamente el concepto de la entropía métrica para los sistemas dinámicos.

Para comprender los enunciados siguientes, admítase que tenemos definido un número real no negativo $h_\mu(f) \geq 0$, llamado *entropía métrica de f* con respecto de la medida de probabilidad f -invariante μ , que depende solo de f y de μ y tal que $h_\mu(f)$ es invariante por isomorfismos de espacios de medida. El número $h_\mu(f)$, por la forma en que se define, mide con respecto a la probabilidad μ , la tasa exponencial maximal en que *los iterados futuros de f desordenan los pedazos de cualquier partición finita \mathcal{P} del espacio M* , ponderados con la probabilidad μ .

Más precisamente, consideremos la partición $\mathcal{P}_n := \bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})$, definida por la siguiente condición: $x, y \in A \in \mathcal{P}_n$ si y solo si para todo $0 \leq j \leq n-1$ existe $P_j \in \mathcal{P}$ tal que $x, y \in P_j$. Cada pedazo diferente de esta partición \mathcal{P}_n , es el conjunto de puntos x que mutuamente se acompañan dentro de un mismo pedazo de \mathcal{P} , al ser iterados hacia el futuro. Cuando mayor es la cantidad de pedazos diferentes de la partición \mathcal{P}_n al crecer n , más desordena el iterado f^n a las órbitas en el espacio, en relación a la partición inicial \mathcal{P} .

Se define

$$h_\mu(f, \mathcal{P}) := \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{A \in \mathcal{P}_n} \mu(A) \log \mu(A).$$

$\sum_{A \in \mathcal{P}_n} \mu(A) \log \mu(A)$ es un promedio ponderado de logaritmos. Luego, se interpreta como un exponente (ponderado). Entonces, al dividirlo entre n , da una *tasa o coeficiente de crecimiento* exponencial con n . Por este motivo, ese cociente se interpreta como la tasa o velocidad de crecimiento exponencial con n (hasta el instante n), del “desorden espacial” que producen los iterados de f , ponderado con la medida de probabilidad μ . Por lo tanto, intuitivamente hablando, $h_\mu(f, \mathcal{P})$ es la tasa exponencial, μ -ponderada y asintótica, en que los iterados de f (mejor dicho μ -casi toda órbita de f) desordenan a la partición inicial \mathcal{P} con la que mirábamos, como referencia, la distribución de puntos iniciales.

Finalmente, se define la entropía métrica

$$h_\mu(f) := \sup_{\mathcal{P} \in \mathbb{P}} h_\mu(f, \mathcal{P}),$$

donde \mathbb{P} denota el conjunto de todas las particiones finitas de M en piezas medibles. Luego $h_\mu(f) \geq 0$ puede interpretarse como la tasa exponencial asintótica maximal de crecimiento del desorden espacial de f^n al hacer $n \rightarrow +\infty$, ponderada con la probabilidad μ . Si $h_\mu(f) > 0$, el sistema restringido al soporte de μ se llama *caótico* en sentido medible. Cuanto mayor es $h_\mu(f)$, más rápidamente se desordena el espacio al iterar f .

Luego, al observador que quiera cuantificar el caos, interesan, si existen, aquellas medidas de probabilidad invariantes μ que maximicen la entropía métrica $h_\mu(f)$ en relación al valor esperado de cierta función real, llamada *potencial* (el cual, gruesamente hablando, cuantifica el significado de “optimizar la observación del caos”). Estas medidas que maximizan la diferencia entre la entropía métrica y el valor esperado de una función potencial, se llaman *estados de equilibrio*. Su estudio se constituye en la sub-teoría, dentro de la teoría ergódica de los sistemas dinámicos, llamada *formalismo termodinámico* (ver por ejemplo [Ke 1998, Chapter 4]).

Teorema 6.4.1. Fórmula de Pesin para la entropía

Sea $f \in \text{Diff}^{1+\alpha}(M)$ que preserva una medida μ hiperbólica y de Gibbs. Entonces

$$h_\mu(f) = \int \sum_{i=1}^{\dim(M)} \chi_i^+(x) d\mu, \quad (6.7)$$

donde $h_\mu(f)$ es la entropía métrica de f con respecto de μ , y

$$\chi_i^+(x) := \max\{0, \chi_i(x)\},$$

siendo $\{\chi_i\}_{1 \leq i \leq \dim(M)}$ los exponentes de Lyapunov de f en el punto x , repetido cada uno de ellos tantas veces como la dimensión del espacio de Oseledets E_x^i correspondiente.

Se recuerda que, por el Teorema de Oseledets, μ -c.t.p. es regular. Por lo tanto existe la función real no negativa $\sum_{i=1}^{\dim(M)} \chi_i^+(x)$ y es medible. Si ninguno de los exponentes de Lyapunov $\chi_i(x)$ positivo, resulta $\sum_{i=1}^{\dim(M)} \chi_i^+(x) = 0$.

La demostración de la Fórmula de Pesin se encuentra en [Pe 1977]. Una prueba diferente que prescinde de la propiedad de continuidad absoluta de la foliación estable se encuentra en [Mañ 1983b]. La demostración también se encuentra, por ejemplo, en [Ba-Pe 2001, Theorem 5.4.5]. Finalmente, en [Ta 2002] se generaliza el Teorema 6.4.1 para difeomorfismos C^1 -genéricos.

Si una medida μ satisface la fórmula de Pesin, entonces ella mide óptimamente el caos medible del sistema en relación al valor esperado de la suma de los exponentes de Lyapunov positivos. En efecto, la igualdad (6.7) da el máximo posible de la entropía métrica $h_\mu(f)$ en relación a ese valor esperado, ya que para toda aplicación de clase C^1 rige la siguiente cota superior de $h_\mu(f)$:

Teorema 6.4.2.**Desigualdad de Margulis-Ruelle [Mar 1966], [Rue 1978]**

Sea $f : M \mapsto M$ de clase C^1 . Sea μ una medida de probabilidad invariante por f . Entonces

$$h_\mu(f) \leq \int \sum_{i=1}^{\dim(M)} \chi_i^+(x) d\mu, \quad (6.8)$$

donde $h_\mu(f)$ es la entropía métrica de f con respecto de μ , y

$$\chi_i^+(x) := \text{máx}\{0, \chi_i(x)\},$$

donde $\{\chi_i\}_{1 \leq i \leq \dim(M)}$ son los exponentes de Lyapunov de f en el punto x , repetido cada uno de ellos tantas veces como la dimensión del espacio de Oseledets E_x^i correspondiente.

La demostración de la Desigualdad (6.8) de Margulis-Ruelle se puede encontrar en [Rue 1978], o también en [Ba-Pe 2001, Theorem 5.4.1].

La fórmula de Pesin, caracteriza a las medidas de Gibbs, debido al siguiente recíproco del Teorema 6.4.1:

Teorema 6.4.3. [Ledrappier-Young, [Led-You 1985]]

Sea $f \in \text{Diff}^2(M)$ y sea μ una medida invariante tal que

$$\sum_{i=1}^{\dim(M)} \chi_i^+(x) > 0 \quad \mu - \text{c.t.p. } x \in M.$$

Si se verifica la Fórmula de Pesin (6.7) de la entropía métrica $h_\mu(f)$, entonces la medida μ es de Gibbs, es decir, μ tiene medidas condicionales inestables absolutamente continuas.

La demostración del Teorema 6.4.3 se encuentra en [Led-You 1985], y también en [Led 1984] con la hipótesis adicional de μ hiperbólica.

Corolario 6.4.4. Sea Λ un atractor topológico uniformemente hiperbólico de $f \in \text{Diff}^2(M)$. Entonces las siguientes afirmaciones son equivalentes:

- (i) μ es una medida ergódica SRB (o física).
- (ii) μ es una medida ergódica de Gibbs.
- (iii) μ es una medida ergódica que satisface la Fórmula (6.7) de Pesin para la entropía.

Demostración. Se obtiene inmediatamente de reunir los Teoremas 6.3.3, 6.4.1 y 6.4.3. \square

6.5. Demostración del Teorema 6.3.1

Para poder demostrar los Teoremas 6.3.1 y 6.3.4, necesitamos introducir algunos resultados relevantes de la llamada *Teoría de Pesin*: Esta teoría estudia el comportamiento de μ -casi toda órbita, para $f \in \text{Diff}^{1+\alpha}(M)$ donde μ es una medida de

probabilidad f -invariante e *hiperbólica*. Es decir, la región de Pesin Σ (i.e. el conjunto de los puntos regulares cuyos exponentes de Lyapunov son todos diferentes de cero) tiene μ -probabilidad igual a 1.

Definición 6.5.1. Holonomía

Sea μ una medida de probabilidad invariante hiperbólica de $f \in \text{Diff}^{1+\alpha}(M)$. Es decir, la región de Pesin Σ cumple $\mu(\Sigma) = 1$. Tomemos $x_0 \in \Sigma$ tal que para todo $\delta > 0$ la bola $B_\delta(x_0) \subset M$ de centro x_0 y radio $\delta > 0$ cumple $\mu(B_\delta(x_0)) > 0$. Por el Teorema 3.12.1, existen para μ -c.t.p. $x \in B_\delta(x_0)$ las variedades estables e inestables locales por el punto x , que denotamos $W_\delta^s(x)$, $W_\delta^u(x) \subset B_\delta(x_0)$ (Si fuera necesario tomamos la componente conexa que contiene al punto x de la intersección de la variedad estable o inestable local por el punto x con la bola $B_\delta(x_0)$). Definimos la holonomía

$$h_s : B_\delta^*(x_0) \mapsto W_\delta^u(x_0),$$

donde

$$B_\delta^*(x_0) = \{y \in B_\delta(x_0) \cap \Sigma : \#(W_\delta^s(y) \cap W_\delta^u(x_0)) = 1\}$$

como la transformación que a cada punto $y \in B_\delta^*(x_0)$ hace corresponder el único punto $h_s(y) \in W_\delta^s(y) \cap W_\delta^u(x_0)$. Debido al Teorema 3.12.1, si $\delta > 0$ es suficientemente pequeño, entonces h_s existe, pues el conjunto $B_\delta^*(x_0)$ contiene por lo menos a la variedad estable local $W_\delta^s(x_0)$ que interseca transversalmente a $W_\delta^u(x_0)$ en el punto x_0 .

La transformación h_s se llama *holonomía a lo largo de las variedades estables locales* en la bola $B_\delta(x_0)$ sobre la variedad inestable local del punto x_0 , o en breve, *holonomía local estable*.

Decimos que *la holonomía estable h_s es absolutamente continua* si para μ -c.t.p. $x_0 \in \Sigma$ y para todo boreliano $A \subset B_\delta(x_0)$ se cumple:

$$m(h_s^{-1}(A)) = 0 \Leftrightarrow m^u(A \cap W_\delta^u(x_0)) = 0, \quad (6.9)$$

donde m es la medida de Lebesgue en la variedad M y m^u es la medida de Lebesgue a lo largo de la subvariedad inestable local $W_\delta^u(x_0)$.

Análogamente se define *holonomía inestable* y *continuidad absoluta de la holonomía inestable*, intercambiando entre sí los roles de las variedades estables e inestables locales en la definición anterior.

Teorema 6.5.2. Pesin [Pe 1976]

(Teorema fundamental de la Teoría de Pesin)

Sea $f \in \text{Diff}^{1+\alpha}(M)$ y sea μ una medida ergódica e hiperbólica. Entonces existe $\delta > 0$ suficientemente pequeño, tal que la holonomía estable y la holonomía inestable en las bolas $B_\delta(x_0)$ para μ -c.t.p. $x_0 \in M$ son absolutamente continuas.

La prueba del Teorema 6.5.2 se encuentra en [Pe 1976], y también, por ejemplo, en [Theorem 4.3.1][Ba-Pe 2001].

Observamos que el Teorema 6.5.2 es falso en la topología C^1 . En efecto, en [Bow 1975a] y [Rob-You 1980] (y también en [Sch-Gor 1989] para endomorfismos) se construyen

ejemplos de atractores hiperbólicos para los cuales la holonomía a lo largo de las variedades estables locales no es absolutamente continua.

Por ese motivo la demostración del Teorema 6.3.1 no funciona para difeomorfismos de clase C^1 que no sean $C^{1+\alpha}$. En particular, la demostración que daremos del Teorema 6.3.4 no funciona para difeomorfismos de clase C^1 que no sean de clase $C^{1+\alpha}$. Más adelante expondremos algunos resultados utilizando las medidas “SRB-like” (ver Definición 7.6.4), que no son necesariamente medidas de Gibbs, pero que existen para difeomorfismos y endomorfismos sin más regularidad que C^1 manteniendo propiedades de atracción estadística similares a las medidas SRB.

Ahora, veremos que la demostración del Teorema 6.3.1, que relaciona para difeomorfismos de clase $C^{1+\alpha}$ las medidas de Gibbs hiperbólicas y ergódicas con las medidas SRB o físicas, se reduce al Teorema fundamental de la Teoría de Pesin que establece la continuidad absoluta de la holonomía estable local.

6.5.3. Demostración del Teorema 6.3.1

Demostración. Por la parte (c) del Corolario 6.2.4 si μ es medida de Gibbs, entonces μ -casi toda componente ergódica μ_x de μ es medida de Gibbs. Además como μ es hiperbólica por hipótesis, entonces $\mu(\Sigma) = 1$, donde Σ es la región de Pesin. Luego, por el Teorema 4.1.2 de descomposición ergódica μ -casi toda componente ergódica μ_x de μ cumple $\mu_x(\Sigma) = 1$; es decir μ_x es hiperbólica y de Gibbs, además de ser ergódica. Tomemos una tal μ_x y renombremosla como μ .

Para probar el Teorema 6.3.1, aplicando la Definición 5.5.6 de medida SRB o física, debemos probar que la cuenca B de atracción estadística de μ , definida por:

$$B = \left\{ x \in M : \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} = \mu \right\}$$

tiene medida de Lebesgue $m(B)$ positiva.

Como μ es ergódica, $\mu(B) = 1$. Por el Teorema 6.1.7 de Descomposición de Rohlin en una bola $B_\delta(x_0)$ con μ -medida positiva, se cumple:

$$\mu^u(B \cap W_\delta^u(x_0)) = 1$$

para μ -c.t.p. x_0 , donde μ^u es la medida condicionada inestable de μ . Como μ es medida de Gibbs, por la Definición 6.2.3, se cumple $\mu^u \ll m^u$, para μ -c.t.p. x_0 , donde m^u es la medida de Lebesgue a lo largo de la variedad $W_\delta^u(x_0)$. Luego deducimos

$$m^u(B \cap W_\delta^u(x_0)) > 0$$

Aplicando el Teorema 6.5.2 de la Teoría de Pesin, obtenemos:

$$m(h_s^{-1}(B \cap W_\delta^u(x_0))) > 0.$$

Entonces, para terminar de demostrar que μ es una medida SRB, es decir para terminar de probar que $m(B) > 0$, basta demostrar ahora que $h_s^{-1}(B \cap W_\delta^u(x_0)) \subset B$. En

efecto, sea $y \in h_s^{-1}(B \cap W_\delta^u(x_0))$. Probemos que $y \in B$. Por definición de la holonomía estable h_s , se cumple $y \in B_\delta^*$ tal que $h_s(y) = z := W_\delta^s(y) \cap W_\delta^u(x_0) \in B$. Entonces, como $z \in B \cap W_\delta^s(y)$, obtenemos

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(z), f^n(y)) = 0, \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(z)} = \mu.$$

Aplicando el resultado probado en el Ejercicio 2.1.12, resulta

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(y)} = \mu,$$

es decir $y \in B$, como queríamos probar. \square

6.6. Lema de Distorsión Acotada

Al final de este capítulo probaremos el Teorema 6.3.4 que establece, en el contexto $C^{1+\alpha}$ -Anosov, la equivalencia entre las medidas SRB ergódicas y las medidas de Gibbs ergódicas, y la existencia de estas.

La demostración del Teorema 6.3.4 está fuertemente basada en el siguiente Lema 6.6.1 de Distorsión Acotada. La prueba de este lema se basa en la hipótesis $f \in \text{Diff}^{1+\alpha}(M)$.

Lema 6.6.1. Lema de Distorsión Acotada *Sea $f \in \text{Diff}^{1+\alpha}(M)$ de Anosov. Sea $TM = E^u \oplus E^s$ el splitting uniformemente hiperbólico del fibrado tangente, y sea $W_\delta^u(x)$ la variedad inestable local por el punto x para $\delta > 0$ constante, suficientemente pequeño. Denotamos con*

$$J^u(x) = |\det df_x|_{E_x^u}| > 0$$

al Jacobiano inestable en el punto x . Sean $x, y \in M$ dos puntos tales que $y \in W_\delta^u(x)$ y sean, para todo $n \geq 1$, las funciones

$$h_n(x, y) := \frac{\prod_{j=1}^n J^u(f^{-j}(x))}{\prod_{j=1}^n J^u(f^{-j}(y))} \in (0, +\infty)$$

$$h(x, y) := \frac{\prod_{j=1}^{+\infty} J^u(f^{-j}(x))}{\prod_{j=1}^{+\infty} J^u(f^{-j}(y))} \in [0, +\infty],$$

definidas solo para las parejas (x, y) de puntos en el conjunto:

$$H_\delta := \{(x, y) \in M \text{ tales que } y \in W_\delta^u(x)\}.$$

Entonces:

(i) *Existe una constante real $K > 0$ tal que*

$$\frac{1}{K} < h(x, y) < K \quad \forall (x, y) \in H_\delta$$

(ii) La función $h : H_\delta \rightarrow \mathbb{R}^+$ es continua.

(iii) Para todo $\epsilon > 0$ existe $N \geq 1$ (independiente de la pareja de puntos $(x, y) \in H_\delta$) tal que:

$$e^{-\epsilon} < \frac{h_n(x, y)}{h(x, y)} < e^\epsilon \quad \forall (x, y) \in H_\delta, \quad \forall n \geq N.$$

Demostración. (i) Definamos $\log : [0, +\infty] \mapsto [-\infty, +\infty]$ acordando que $\log 0 = -\infty$ y $\log(+\infty) = +\infty$. Para demostrar la afirmación (i) basta probar que existe una constante real $c > 0$ tal que $|\log h(x, y)| \leq c$. Por construcción de la función $h : H_\delta \mapsto [0, +\infty]$ tenemos:

$$\begin{aligned} |\log h(x, y)| &= \left| \sum_{j=1}^{+\infty} (\log J^u(f^{-j}(x)) - \log J^u(f^{-j}(y))) \right| \leq \\ &\sum_{j=1}^{+\infty} |\log J^u(f^{-j}(x)) - \log J^u(f^{-j}(y))|. \end{aligned} \quad (6.10)$$

$J^u : M \mapsto \mathbb{R}^+$ es continuo (porque df_x es continuo pues f es de clase C^1 y E_x^u depende continuamente de x). Entonces la función J^u está uniformemente acotada superiormente e inferiormente por constantes reales positivas. Luego, por el Teorema del valor medio del cálculo diferencial aplicado a la función real $\log t$ de variable real positiva t , existe una constante $c_1 > 0$ tal que

$$|\log J^u(z_1) - \log J^u(z_2)| \leq c_1 |J^u(z_1) - J^u(z_2)|$$

para toda pareja (z_1, z_2) de puntos en la variedad M . Sustituyendo en (6.10) resulta:

$$|\log h(x, y)| \leq c_1 \sum_{j=1}^{+\infty} |J^u(f^{-j}(x)) - J^u(f^{-j}(y))|. \quad (6.11)$$

Siendo f de clase $C^{1+\alpha}$, las variedades inestables son de clase $C^{1+\alpha}$ (cf. la última parte del enunciado del Teorema 3.6.1). Entonces el subespacio tangente $T_y W_\delta^u(x) = E_y^u$ depende α -Hölder continuamente del punto y cuando y varía a lo largo de la subvariedad $W_\delta^u(x)$. Es decir, para cada x existe una constante $c_2(x) > 0$ tal que

$$\text{dist}(E_y^u, E_z^u) \leq c_2(x) \text{dist}(y, z)^\alpha \quad \forall y, z \in W_\delta^u(x). \quad (6.12)$$

Afirmamos que para $\delta > 0$ fijo suficientemente pequeño, existe una constante c_2 uniforme tal que

$$c_2(x) \leq c_2 \quad \forall x \in M.$$

En efecto, fijando $x_1 \in M$ y una constante $C(x_1) > c_2(x_1)$, la desigualdad (6.12) se satisface para cualquier punto \hat{x} en un entorno suficientemente pequeño de x_1 , sustituyendo $c_2(x)$ por $C(x_1)$. Luego, cubriendo la variedad compacta M con una cantidad finita de tales entornos, centrados en puntos x_1, \dots, x_m respectivamente, y

tomando $c_2 = \max_{i=1}^m C(x_i)$, deducimos que se satisface la desigualdad (6.12) para todo $x \in M$, con la constante c_2 en lugar de $c_2(x)$.

Por otra parte, como f es de clase $C^{1+\alpha}$, su derivada df_x depende α -Hölder continuamente del punto x . Es decir, existe una constante $c_3 > 0$ tal que

$$\|df_y - df_z\| \leq c_3 \text{dist}(y, z)^\alpha \quad \forall y, z \in M.$$

Por otra parte existe una constante $c_4 > 0$ tal que

$$|\det A| \leq c_4 \|A\|$$

para toda aplicación lineal A de un espacio vectorial de dimensión igual a $\dim(E^u)$ en otro de la misma dimensión. Reuniendo las cuatro últimas desigualdades que involucran las constantes $c_2(x)$, c_2 , c_3 y c_4 , obtenemos

$$\begin{aligned} |J^u(y) - J^u(z)| &= \left| |\det(df_y|_{E_y^u})| - |\det(df_z|_{E_z^u})| \right| \leq \\ & \left| |\det(df_y|_{E_y^u})| - |\det(df_z|_{E_y^u})| \right| + \left| |\det(df_z|_{E_y^u})| - |\det(df_z|_{E_z^u})| \right| \leq \\ & \leq c_4 \|df_y - df_z\| + c_4 (\max_{z \in M} \|df_z\|) \text{dist}(E_y^u, E_z^u) \leq \\ & (c_4 c_3 + c_4 (\max_{z \in M} \|df_z\|) c_2) \cdot \text{dist}(y, z)^\alpha \quad \forall (y, z) \in W_\delta^u(x), \quad \forall x \in M. \end{aligned}$$

Sustituyendo en la desigualdad (6.11), obtenemos una constante $c_5 > 0$ tal que

$$|\log h(x, y)| \leq c_5 \sum_{j=1}^{+\infty} \text{dist}(f^{-j}(x), f^{-j}(y))^\alpha \quad \forall (x, y) \in H_\delta \quad (6.13)$$

Nota: Para obtener la desigualdad (6.13) observamos que no necesariamente se cumple $f^{-j}W_\delta^u(x) \subset W_\delta^u(f^j(x))$, y por lo tanto, aunque $(x, y) \in H_\delta$ no necesariamente son aplicables directamente todas las desigualdades anteriores para la pareja de puntos $(f^{-j}(x), f^{-j}(y))$ para cualquier $j \geq 1$. Sin embargo, fijado $\delta > 0$ suficientemente pequeño, aplicamos la parte B) del Teorema 3.6.1, y observamos que hay convergencia uniforme a cero de las distancias a lo largo de las variedades inestables locales al iterar hacia el pasado, pues los vectores tangentes $u \in E_x^u$ se contraen uniformemente hacia el pasado según la Definición 3.3.1 de difeomorfismo de Anosov. Luego, existe $N \geq 1$ uniforme tal que $f^{-n}W_\delta^u(x) \subset W_\delta^u(f^{-n}(x))$ para todo $n \geq N$ para todo $x \in M$. Entonces, basta elegir $0 < \delta' < \delta$ tal que $f^{-j}W_{\delta'}^u(x) \subset W_\delta^u(f^{-j}(x))$ para $j \in \{0, \dots, N-1\}$, para que esa misma inclusión valga también para todo $j \geq 0$. Finalmente renombramos δ' como δ para deducir la desigualdad (6.13).

Por la Definición 3.3.1 tenemos:

$$\text{dist}^u(f^{-j}(x), f^{-j}(y)) \leq C \sigma^{-j} \text{dist}^u(x, y), \quad (6.14)$$

donde $C > 0$ y $\sigma > 1$ son las constantes dadas en la Definición 3.3.1, y $\text{dist}^u(x, y)$ es la distancia entre los puntos x e y a lo largo de la variedad inestable local $W_\delta^u(x)$. Además la distancia dist en la variedad ambiente M entre dos puntos que están en la

misma variedad inestable local, es siempre el mínimo de las longitudes de curvas que unen esos puntos (con la métrica riemanniana dada en M). Por lo tanto es menor o igual que la distancia dist^u a lo largo de la variedad inestable local. Sustituyendo en la desigualdad (6.13) resulta:

$$|\log h(x, y)| \leq c_5 C^\alpha \sum_{j=1}^{+\infty} \sigma^{-\alpha j} (\text{dist}^u(x, y))^\alpha \leq c_6 (\text{dist}^u(x, y))^\alpha$$

$$\forall (x, y) \in H_\delta, \quad (6.15)$$

donde $c_6 = c_5 C^\alpha 1/(1 - \sigma^{-\alpha})$ (observar que $0 < \sigma^{-\alpha} < 1$ porque $\sigma > 1$ y $\alpha > 0$). Como $W_\delta^u(x)$ es una variedad C_1 -encajada en M que depende continuamente de x , existe

$$\text{Diam}(W_\delta^u(x)) := \sup_{y \in W_\delta^u(x)} \text{dist}^u(x, y) \in \mathbb{R}^+$$

y es una función continua real de $x \in M$. Luego, está acotada superiormente por una constante uniforme $c_7 > 0$. Sustituyendo en (6.15), concluimos:

$$|\log h(x, y)| \leq c_6 (\text{dist}^u(x, y))^\alpha \leq c_6 c_7^\alpha = c \quad \forall (x, y) \in H_\delta, \quad (6.16)$$

terminando la demostración de la parte (i) del Lema 6.6.1.

(iii) Fijemos $n \geq 1$. Por la definición de las funciones h y h_n tenemos

$$\frac{h_n(x, y)}{h(x, y)} =: \frac{\prod_{j=n+1}^{+\infty} J^u(f^{-j}(y))}{\prod_{j=n+1}^{+\infty} J^u(f^{-j}(x))}$$

Luego, aplicando la fórmula (6.15) a los puntos $f^{-n-1}(x)$ y $f^{-n-1}(y)$, en lugar de x e y respectivamente, obtenemos

$$\left| \log \left(\frac{h_n(x, y)}{h(x, y)} \right) \right| \leq c_6 (\text{dist}^u(f^{-n-1}(x), f^{-n-1}(y)))^\alpha.$$

Usando ahora la desigualdad (6.14), deducimos:

$$\left| \log \left(\frac{h_n(x, y)}{h(x, y)} \right) \right| \leq c_6 C^\alpha \sigma^{-(n+1)\alpha} (\text{dist}^u(x, y))^\alpha \leq c_6 C c_7^\alpha \sigma^{-(n+1)\alpha}.$$

Como $\sigma > 1$ y $\alpha > 0$, el término de la derecha tiende a cero cuando $n \rightarrow +\infty$. Luego, existe $N \geq 1$, independiente de (x, y) tal que

$$\left| \log \frac{h_n(x, y)}{h(x, y)} \right| < \epsilon \quad \forall n \geq N,$$

terminando de probar la afirmación (iii) del Lema 6.6.1.

(ii) Debido a la afirmación (ii), la sucesión de funciones $h_n(x, y)$ converge uniformemente a $h(x, y)$ para todo $(x, y) \in H_\delta$. Además $h_n(x, y)$ es continua en H_δ , porque el Jacobiano inestable $J^u(x) = |\det df_x^n|_{E_x^u}|$ es una función continua de x , pues df_x es continuo y E_x^u también. El límite uniforme de una sucesión de funciones continuas es continuo. Concluimos que h es una función continua, finalizando la prueba del Lema 6.6.1. \square

6.7. Demostración del Teorema 6.3.4

En la demostración del Teorema 6.3.4, usaremos, además del Lema de Distorsión Acotada, el siguiente resultado, válido para cualquier difeomorfismo de Anosov (no necesariamente de clase $C^{1+\alpha}$), y generalizable también (reformulando adecuadamente el enunciado) para los entornos locales de atractores topológicos uniformemente hiperbólicos.

Teorema 6.7.1. Estructura de producto local *Sea $f \in \text{Diff}^1(M)$ Anosov. Existen constantes $0 < \delta' < \delta$ suficientemente pequeñas, tales que f tiene estructura de producto local en entornos de radio δ' .*

La estructura de producto local significa, por definición, que para toda pareja de puntos (x, y) tales que $\text{dist}(x, y) < \delta'$, existen, son únicos y dependen continuamente de (x, y) los puntos $[x, y]$ y $[y, x]$ definidos por:

$$[x, y] = W_\delta^u(x) \cap W_\delta^s(y), \quad [y, x] = W_\delta^u(y) \cap W_\delta^s(x),$$

donde $W_\delta^{u,s}(x)$ denota la componente conexa de $W^{u,s}(x) \cap B_\delta(x)$ que contiene al punto x .

Una prueba del Teorema 6.7.1 puede encontrarse en [Bow 1975b, Theorem 3.12] o en [Kat-Has 1995, Proposition 6.4.21]. El Teorema 6.7.1 es consecuencia del llamado “shadowing lemma” (lema de sombreado).

Ahora estamos en condiciones de demostrar el Teorema 6.3.4 que establece la existencia, y mutua equivalencia, de las medidas ergódicas de Gibbs y de las medidas SRB, para los difeomorfismos de Anosov de clase $C^{1+\alpha}$.

6.7.2. Demostración del Teorema 6.3.4

Esta prueba consta de tres partes:

Afirmación I: *Existen medidas de probabilidad μ de Gibbs tales que:*

- *Las probabilidades condicionales inestables μ^u son equivalentes a las medidas de Lebesgue m^u a lo largo de las respectivas variedades inestables locales $W_\delta^u(x)$ para μ -c.t.p. $x \in M$. (Nota: La equivalencia entre μ^u y m^u se define como $\mu^u \ll m^u$ y $m^u \ll \mu^u$).*
- *La derivada de Radon-Nikodym $d\mu^u/dm^u$ (la densidad de las medidas condicionales inestables μ^u) es una función real continua y estrictamente positiva.*

Una vez probada la Afirmación I, como consecuencia, debido al Teorema 4.1.2 de Descomposición en componentes ergódicas, y por las partes (b) y (c) del Corolario 6.2.4 del Teorema de Rohlin, concluimos que existen medidas de probabilidad ergódicas de Gibbs μ tales que μ^u es equivalente a m^u para μ -c.t.p. Aplicando entonces el Teorema 6.3.1 ya demostrado, toda medida de probabilidad ergódica y de Gibbs, es SRB. Deducimos que existen medidas SRB ergódicas, es decir, concluimos la parte (a) del Teorema 6.3.4. Además la Afirmación I implica que existen medidas SRB ergódicas, que son de Gibbs, cuyas probabilidades condicionales inestables son equivalentes a la

medida de Lebesgue a lo largo de las variedades inestables respectivas para μ -c.t.p., y cuyas densidades son funciones reales continuas y estrictamente positivas.

Afirmación II: *Sea μ una medida de Gibbs ergódica μ tal que la derivada de Radon-Nikodym (densidad $d\mu^u/dm^u$ de las medidas condicionales inestables μ^u) es una función continua (no necesariamente estrictamente positiva). Entonces la cuenca de atracción estadística de μ cubre Lebesgue c.t.p. de un entorno abierto de su soporte.*

De esta propiedad, una vez demostrada, deducimos que dos medidas de Gibbs ergódicas diferentes cuyas densidades inestables sean continuas (en particular dos medidas de Gibbs ergódicas que satisfagan la Afirmación I), deben tener soportes disjuntos a distancia positiva. En efecto, por la Definición 5.5.1, las cuencas de atracción estadística de medidas diferentes deben ser disjuntas. Debido a la Afirmación II los soportes de esas dos medidas están mutuamente aislados. De aquí deducimos (usando que M es un espacio métrico compacto y por lo tanto existe base numerable de abiertos) que las medidas de Gibbs ergódicas que satisfacen la Afirmación I (y por lo tanto también la II) son a lo sumo una cantidad numerable.

Afirmación III: *Lebesgue casi todo punto de la variedad M pertenece a la unión de las cuencas de atracción estadística de las medidas de Gibbs ergódicas cuyas medidas condicionales inestables cumplen las dos propiedades de la Afirmación I.*

Supongamos probada la afirmación III. Tomando en particular una medida SRB o física cualquiera ν , con cuenca de atracción estadística A , deducimos que A está contenido en la unión de las cuencas de atracción estadística (que son disjuntas dos a dos) de una colección finita o infinita numerable de medidas de Gibbs ergódicas que satisfacen la Afirmación I. Como las cuencas de atracción estadística de medidas diferentes son disjuntas, deducimos que toda medida SRB o física es de Gibbs ergódica y satisface la Afirmación I, y por lo tanto también la II. Esto, junto con el Teorema 6.3.1, prueba la parte (b) del Teorema 6.3.4.

Además, la afirmación III implica inmediatamente la parte (c) del Teorema 6.3.4, y más aún, toda medida SRB no solo es de Gibbs ergódica, sino que además satisface las afirmaciones I y II.

Supongamos ahora además que f es topológicamente transitivo. Si hubiera dos medidas SRB (ergódicas) diferentes μ y ν , entonces sus cuencas de atracción estadística $B(\mu)$ y $B(\nu)$ serían disjuntas. Como μ y ν satisfacen la Afirmación II, existen abiertos disjuntos U y V tales que Lebesgue c.t.p. de U pertenece a $B(\mu)$ y Lebesgue c.t.p. de V pertenece a $B(\nu)$. Como $B(\mu)$ es f -invariante, entonces $m(f^n(U) \setminus B(\mu)) = 0$ para todo $n \geq 1$. Como f es topológicamente transitivo existe $n \geq 1$ tal que el abierto $f^n(U) \cap V \neq \emptyset$. Todo abierto (en particular $f^n(U) \cap V$) tiene medida de Lebesgue positiva, y Lebesgue c.t.p. de $f^n(U) \cap V$ pertenece a $B(\mu)$ (por pertenecer a $f^n(U)$), y pertenece también a $B(\nu)$ (por pertenecer a V). Luego las cuencas de atracción estadística $B(\mu)$ y $B(\nu)$ no son disjuntas, de donde $\mu = \nu$. Esto prueba la unicidad de la medida SRB, es decir, la parte (d) del Teorema 6.3.4.

En resumen, para demostrar completamente el Teorema 6.3.4, basta probar las Afir-

maciones I, II y III.

Primera parte de la demostración del Teorema 6.3.4

(Prueba de la Afirmación I de §6.7.2)

Demostración. .

Paso 1: Construcción de candidata a medida de Gibbs

Elijamos una variedad inestable local cualquiera W_0 y consideremos la medida m_0^u de Lebesgue a lo largo de W_0 . Tenemos $m_0^u(W_0) > 0$. Definimos la siguiente medida de probabilidad μ_0 en la variedad ambiente M :

$$\mu_0(A) := \frac{m^u(A \cap W_0)}{m^u(W_0)} \quad \forall A \in \mathcal{A},$$

donde \mathcal{A} es la sigma-álgebra de Borel. Para todo $n \geq 1$ definimos las siguientes medidas de probabilidad ν_n y μ_n :

$$\nu_n(A) := \mu_0(f^{-n}(A)), \quad \mu_n(A) := \frac{1}{n} \sum_{j=0}^{n-1} \nu_j(A) \quad \forall A \in \mathcal{A}. \quad (6.17)$$

Finalmente tomamos una subsucesión $\{\mu_{n_i}\}_{i \in \mathbb{N}}$ convergente a μ en el espacio \mathcal{M} de las medidas de probabilidad en (M, \mathcal{A}) con la topología débil*. Probaremos que la medida μ satisface la Afirmación I.

Paso 2: Descomposición de Rohlin de las medidas ν_n .

Consideremos un punto x_0 en el soporte de μ tal que existe una bola $B = B_\delta(x_0) \subset M$ de radio $\delta > 0$ suficientemente pequeño, tal que $\mu(B) > 0$. Entonces, por la definición de la topología débil estrella (tomando por ejemplo una función continua no negativa que vale 1 en x_0 y está soportada en B), existe $N \geq 1$ tal que $\mu_N(B) > 0$, de donde deducimos que, para todo $n \geq N$ existe $j \in \{0, \dots, n-1\}$ tal que $\nu_j(B) > 0$, y por lo tanto

$$\mu_n(B) > 0 \quad \forall n \geq N. \quad (6.18)$$

Esto implica, por la definición de la medida ν_j que $f^j(W_0) \cap B \neq \emptyset$. Fijemos $n \geq 1$ tal que

$$f^n(W_0) \cap B \neq \emptyset.$$

Por construcción, para todo boreliano $A \subset M$, se cumple:

$$\nu_n(A \cap B) = \frac{1}{m_0^u(W_0)} \int \chi_{f^{-n}(A)}(y) \chi_{W_0 \cap f^{-n}(B)}(y) dm_0^u(y).$$

Haciendo el cambio de variable $x = f^n(y)$ y denotando $J^u(y)$ al Jacobiano inestable en el punto y definido por

$$J^u(y) := |\det df_y|_{E^u(y)}|,$$

donde $E_y^u = T_y W^u(y)$, obtenemos:

$$\nu_n(A \cap B) = \frac{1}{m_0^u(W_0)} \int \chi_A(x) \chi_{f^n(W_0) \cap B}(x) |\det d(f^{-n})|_{E_x^u(x)} dm^u(x)$$

donde $m^u(x)$ denota la medida de Lebesgue a lo largo de la variedad inestable local $W_\delta^u(x)$ por el punto x . Tenemos

$$\begin{aligned} |\det d(f^{-n})|_{E_x^u(x)} &= \frac{1}{|\det d(f^n)|_{E_y^u(y)}} = \frac{1}{\prod_{i=0}^n J^u(f^i(y))} = \\ &= \frac{1}{\prod_{h=1}^n J^u(f^{-h}(x))}, \end{aligned}$$

de donde:

$$\nu_n(A \cap B) = \frac{1}{m_0^u(W_0)} \int \chi_A(x) \chi_{f^n(W_0) \cap B}(x) \frac{1}{\prod_{h=1}^n J^u(f^{-h}(x))} dm^u(x).$$

Denotamos k_n el número de componentes conexas de $f^n(W_0) \cap B$ (por convención, $k_n = 0$ si el conjunto es vacío y $\sum_{i=1}^0 \cdot = 0$). Para cada $n \geq 1$ fijo, si $k_n \geq 1$ denotamos $\{W_{i,n}\}_{1 \leq i \leq k_n}$ al conjunto de componentes conexas de $f^n(W_0) \cap B$. Observamos que para todo $x \in W_{i,n}$ la medida $m^u(x)$ es la de Lebesgue a lo largo de la subvariedad $W_{i,n}$. Entonces:

$$\nu_n(A \cap B) = \frac{1}{m_0^u(W_0)} \sum_{i=1}^{k_n} \int_{W_{i,n}} \chi_A(x) \frac{1}{\prod_{h=1}^n J^u(f^{-h}(x))} dm^u(x). \quad (6.19)$$

Fijamos $n \geq 1$ tal que $k_n \geq 1$, es decir $f^n(W_0) \cap B \neq \emptyset$. Luego, por construcción de la medida ν_n mediante la igualdad (6.17), tenemos $\nu_n(B) > 0$. Tomemos una subvariedad cualquiera V , encajada en B_δ , de dimensión complementaria a la inestable, que interseque en un solo punto cada una, a todas las variedades estables locales inestables $W^u \delta(z)$ (para todo $z \in B$). Por el Teorema 6.7.1 de estructura de producto local, tal variedad topológicamente transversal V existe, ya que puede tomarse igual a una variedad estable local.

Para cada $i \in \{1, \dots, k_n\}$ consideremos el punto $x_{i,n} = W_{i,n} \cap V$. Construimos, para cada $x \in \bigcup_{i=1}^{k_n} W_{i,n}$, el siguiente valor real positivo $h_n(x)$ de la que llamaremos función h_n :

$$h_n(x) := \frac{\prod_{h=1}^n J^u(f^{-h}(x_{i,n}))}{\prod_{h=1}^n J^u(f^{-h}(x))}, \quad (6.20)$$

donde $i \in \{1, \dots, k_n\}$ es el único índice tal que $x \in W_{i,n}$.

Consideramos la partición \mathcal{P}_n de la bola B cuyas piezas son $\{W_{i,n}\}_{1 \leq i \leq k_n}$ y además el complemento en B de $\bigcup_{i=1}^{k_n} W_{i,n}$. Denotamos con ρ_n la siguiente medida (finita) en el espacio medible cociente formado por las piezas de la partición \mathcal{P}_n :

$$\rho_n = \frac{1}{m_0^u(W_0)} \sum_{i=1}^{k_n} \frac{\int_{W_{i,n}} h_n dm^u}{\prod_{h=1}^n J^u \circ f^{-h}(x_{i,n})} \delta_{x_{i,n}}, \quad (6.21)$$

donde $\delta_{x_{i,n}}$ denota la Delta de Dirac soportada en la pieza $W_{i,n}$ (representada por el punto $x_{i,n}$). Sustituyendo (6.20) y (6.21) en la igualdad (6.19), obtenemos para todo boreliano A la siguiente descomposición de la medida de probabilidad ν_n restringida a la bola B :

$$\frac{\nu_n(A \cap B)}{\nu_n(B)} = \frac{1}{\nu_n(B)} \int d\rho_n \int_{W_{i,n}} \chi_A(x) \frac{h_n(x)}{\int_{W_{i,n}} h_n dm^u} dm^u(x). \quad (6.22)$$

La igualdad anterior da la descomposición de Rohlin con respecto de la partición \mathcal{P}_n , de la medida de probabilidad $\nu_n/\nu_n(B)$ en la bola B . En efecto, la medida de probabilidad ν_n^u condicionada a lo largo de las piezas de la partición (i.e. a lo largo de las variedades inestables locales $W_{i,n}$) está dada por la integral de la derecha en la igualdad (6.22). Esta medida condicionada ν_n^u cumple

$$d\nu_n^u = \frac{h_n}{\int_{W_{i,n}} h_n dm^u} dm^u,$$

y por lo tanto

$$\nu_n^u \ll m^u.$$

Decimos entonces que las medidas condicionales inestables de ν_n son absolutamente continuas (aún cuando ν_n no es necesariamente invariante con f). La integral de la izquierda en (6.22) da la medida de probabilidad $\hat{\nu}_n$ en el espacio cociente de la bola B con respecto a la partición \mathcal{P}_n . Esta medida cociente $\hat{\nu}_n$ es

$$\hat{\nu}_n = \frac{1}{\nu_n(B)} \rho_n.$$

El objetivo en los siguientes pasos de la demostración es usar la descomposición de Rohlin de las medidas ν_j dada por la igualdad (6.22) para todo $j \in \{0, \dots, n-1\}$, para encontrar la descomposición de Rohlin de la medida $\mu_n = \sum_{j=0}^{n-1} \nu_j$. Luego, pasando al límite para una subsucesión $\{n_i\}_{i \in \mathbb{N}}$, encontraremos la descomposición de Rohlin de la medida μ construida en el Paso 1, para demostrar que es medida de Gibbs y satisface la Afirmación I. El punto difícil en este argumento es justificar rigurosamente el pasaje al límite de los promedios de las medidas $\hat{\nu}_n$ y de las medidas condicionales ν_n^u . Para ello, necesitamos ϵ -aproximar la descomposición de Rohlin dada por la igualdad (6.22), para todo $\epsilon > 0$ arbitrariamente pequeño, y para n suficientemente grande.

Paso 3: ϵ - aproximación de la descomposición de Rohlin de ν_n

Fijemos, como en el paso 2, $n \geq 1$ tal que $f^n(W_0) \cap B \neq \emptyset$. Recordamos que $\{W_{i,n}\}_{1 \leq i \leq k_n}$ denota a las componentes conexas de $f^n(W_0) \cap B$, que son por lo tanto, variedades inestables locales contenidas en la bola B . Recordemos que para obtener la descomposición de Rohlin de ν_n dada por la igualdad (6.22), hemos elegido y dejado fijo un punto y uno solo $x_{i,n} \in W_{i,n}$. Definimos ahora, para todo $x \in \bigcup_{i=1}^{k_n} W_{i,n}$ el siguiente valor $h(x) \in [0, +\infty]$ de la que llamaremos función h restringida a $\bigcup_{i=1}^{k_n} W_{i,n}$:

$$h(x) := \frac{\prod_{h=1}^{+\infty} J^u(f^{-h}(x_{i,n}))}{\prod_{h=1}^{+\infty} J^u(f^{-h}(x))}, \quad (6.23)$$

donde i es el único índice en $\{1, \dots, k_n\}$ tal que $x \in W_{i,n}$. Ahora aplicamos el Lema 6.6.1 de Distorsión Acotada que establece que: Existe una constante real $K > 0$ tal que

$$\frac{1}{K} < h(x) < \frac{1}{K} \quad \forall x \in H := \bigcup_{n \geq 1} \bigcup_{i=1}^{k_n} W_{i,n} \quad (6.24)$$

$$h : H \rightarrow \mathbb{R}^+ \text{ es continua} \quad (6.25)$$

Para todo $\epsilon > 0$ existe $N \geq 1$ (independiente de $x \in H$) tal que:

$$e^{-\epsilon} < \frac{h_n(x)}{h(x)} < e^\epsilon \quad \forall x \in \bigcup_{i=1}^{k_n} W_{i,n} \quad \forall n \geq N. \quad (6.26)$$

Aplicando la afirmación (6.24) del Lema de Distorsión Acotada, construimos, para cada $n \geq 1$ fijo tal que $f^n(W_0) \cap B \neq \emptyset$, la siguiente medida finita I_n :

$$I_n(A) := \int d\rho_n \int_{W_{i,n}} \chi_A(x) \frac{h(x)}{\int_{W_{i,n}} h dm^u} dm^u(x), \quad \forall A \in \mathcal{A} \quad (6.27)$$

donde ρ_n es la medida definida en la igualdad (6.21). Comparando las igualdades (6.22) y (6.27), obtenemos, para todo boreliano A la siguiente igualdad:

$$\begin{aligned} \nu_n(A \cap B) &= \int d\rho_n \int_{W_{i,n}} \chi_A \frac{h_n}{\int_{W_{i,n}} h_n dm^u} dm^u = \\ &= \int d\rho_n \int_{W_{i,n}} \chi_A \frac{h_n}{h} \frac{h}{\int_{W_{i,n}} h(h_n/h) dm^u} dm^u \end{aligned}$$

Usando la propiedad (6.26) del Lema de Distorsión Acotada, deducimos, para todo $n \geq N$ tal que $f^n(W_0) \cap B \neq \emptyset$ y para todo Boreliano A :

$$e^{-2\epsilon} I_n(A) \leq \nu_n(A \cap B) \leq e^{2\epsilon} I_n(A)$$

En conclusión:

$$\begin{aligned} e^{-2\epsilon} \int d\rho_n \int_{W_{i,n}} \chi_A(x) \frac{h(x)}{\int_{W_{i,n}} h dm^u} dm^u(x) &\leq \nu_n(A \cap B) \leq \\ e^{2\epsilon} \int d\rho_n \int_{W_{i,n}} \chi_A(x) \frac{h(x)}{\int_{W_{i,n}} h dm^u} dm^u(x) &\quad \forall A \in \mathcal{A} \end{aligned} \quad (6.28)$$

Paso 4: Descomposición de Rohlin de la medida μ

Sea la medida de probabilidad μ construida en el Paso 1:

$$\mu = \lim_{i \rightarrow +\infty} \mu_{n_i}, \quad (6.29)$$

donde el límite es en la topología débil estrella, y μ_n está definida para todo $n \geq 1$ por la igualdad (6.17), como el promedio aritmético de las medidas ν_j para $j \in \{0, \dots, n-1\}$. Consideremos $n \geq N$ fijo y la partición medible \mathcal{Q} de la bola B formada por las variedades inestables locales. Denotemos B/\sim el espacio medible cociente de B con respecto a esa partición. Usando las desigualdades (6.28) obtenemos, para todo boreliano $A \subset M$ la siguiente ϵ -aproximación de $\mu_n(A)$

$$e^{-2\epsilon} \int_{B/\sim} d\left(\frac{1}{n} \sum_{j=0}^{n-1} \rho_j\right) \int_W \chi_A \frac{h}{\int_W h dm^u} dm^u(x) \leq \mu_n(A \cap B) \leq e^{2\epsilon} \int d\frac{1}{n} \left(\sum_{j=0}^{n-1} \rho_j\right) \int_W \chi_A \frac{h}{\int_W h dm^u} dm^u(x) \quad \forall A \in \mathcal{A}, \quad \forall n \geq N. \quad (6.30)$$

donde, por convención, ρ_j está definida por la igualdad (6.21) si $f^j(W_0) \cap B \neq \emptyset$, y $\rho_j := 0$ en caso contrario.

Denotamos

$$\rho_n^* := \frac{1}{n} \sum_{j=0}^{n-1} \rho_j. \quad (6.31)$$

Por construcción ρ_n^* es una medida finita en el espacio cociente B/\sim de la bola B con respecto a la partición \mathcal{Q} en subvariedades locales inestables. Por la igualdad (6.18) $\mu_n(B) > 0$ para todo n suficientemente grande (digamos $n \geq N$). Entonces, debido a las desigualdades (6.30), ρ_n^* es una medida no nula para todo $n \geq N$. Probemos que la sucesión de medidas $\{\rho_n^*\}_{n \geq N}$ está uniformemente acotada superiormente por e^ϵ . En efecto, por un lado la medida μ_n construida por la igualdad (6.17) es una medida de probabilidad en toda la variedad M , y por otro lado, de la desigualdad a la izquierda en (6.30), obtenemos

$$1 \geq \mu(B) \geq e^{-2\epsilon} \int_{B/\sim} d\rho_n^* = \rho^*(B/\sim).$$

El espacio de todas las medidas finitas en el espacio medible B/\sim que están uniformemente acotadas por $e^{2\epsilon}$, provisto de la topología débil estrella, de secuencialmente compacto. Luego, existe una subsucesión $\{n_{i_k}\}_{k \in \mathbb{N}}$ de $\{n_i\}_{i \in \mathbb{N}}$ (que por simplicidad seguimos denotando como n_i), tal que $\{\rho_{n_i}^*\}_i$ es convergente a una medida finita ρ^* . En resumen, hemos demostrado la existencia de una medida finita ρ^* en el espacio medible cociente B/\sim , tal que:

$$\rho^* = \lim_{i \rightarrow +\infty} \rho_{n_i}^* = \lim_{u \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \rho_j, \quad (6.32)$$

donde el límite de medidas es en la topología débil estrella.

Consideremos una función real continua no negativa $\psi : M \mapsto [0, 1]$. Las desigualdades (6.30) y la igualdad (6.31) implican

$$e^{-2\epsilon} \int_{B/\sim} d\rho_{n_i}^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u \leq \int_B \psi d\mu_{n_i} \leq$$

$$e^{2\epsilon} \int_{B/\sim} d\rho_{n_i}^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u \quad \forall n_i \geq N. \quad (6.33)$$

Aplicando la propiedad (6.25) del Lema de Distorsión Acotada, deducimos que el producto $\psi \cdot h$ es una función continua en B . Luego, como también $m^u(W_\delta^u(x))$ depende continuamente de x (porque el subespacio tangente E_x^u depende continuamente de x), deducimos que la integral a la derecha en (6.33) es un número real positivo que depende continuamente de la variedad inestable $W \in B/\sim$. Entonces, por la igualdad (6.32) y por la definición de la topología débil estrella, deducimos que:

$$\begin{aligned} \lim_{i \rightarrow +\infty} \int_{B/\sim} d\rho_{n_i}^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u &= \\ &= \int_{B/\sim} d\rho^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u. \end{aligned}$$

Tomando límite en las desigualdades (6.33) y recordando (6.29), obtenemos:

$$\begin{aligned} e^{-2\epsilon} \int_{B/\sim} d\rho^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u &\leq \int_B \psi d\mu \leq \\ e^{2\epsilon} \int_{B/\sim} d\rho^* \int_W \psi \frac{h}{\int_W h dm^u} dm^u & \end{aligned} \quad (6.34)$$

Por linealidad, la integral de la derecha en (6.34) define un funcional lineal acotado y positivo en el espacio $C^0(M, \mathbb{R})$ de todas las funciones continuas $\psi : M \mapsto \mathbb{R}$. Luego, por el Teorema de Representación de Riesz, define una medida finita $\tilde{\mu}$. Como las desigualdades (6.34) valen para toda $\psi \in C^0(M, [0, 1])$, y la función característica χ_A de cualquier boreliano A puede aproximarse en $L^1(\mu)$ y en $L^1(\tilde{\mu})$ por una función continua $\psi \in C^0(M, [0, 1])$, obtenemos, para todo boreliano A las siguientes desigualdades:

$$\begin{aligned} e^{-2\epsilon} \int_{B/\sim} d\rho^* \int_W \chi_A \frac{h}{\int_W h dm^u} dm^u &\leq \mu(A \cap B) \leq \\ \int_{B/\sim} d\rho^* \int_W \chi_A \frac{h}{\int_W h dm^u} dm^u & \end{aligned} \quad (6.35)$$

Por construcción, las medidas involucradas en las desigualdades (6.35) son independientes del número $\epsilon > 0$, pero valen para todo $\epsilon > 0$ suficientemente pequeño. Luego, haciendo $\epsilon \rightarrow 0^+$, deducimos:

$$\mu(A \cap B) = \int_{B/\sim} d\rho^* \int_W \chi_A \frac{h}{\int_W h dm^u} dm^u \quad \forall A \in \mathcal{A}. \quad (6.36)$$

Dividiendo la igualdad (6.36) entre $\mu(B) > 0$, hemos encontrado la descomposición de Rohlin de la medida de probabilidad μ restringida a la bola B , con respecto a la partición \mathcal{Q} de B en subvariedades inestables locales. En efecto, la medida $\hat{\mu}$ en el espacio cociente B/\sim con respecto a esa partición es $\rho^*/\mu(B)$ (es inmediato chequear

que esta es una medida de probabilidad en B/\sim ; y las medidas condicionales inestables μ^u de μ están definidas por las integrales de la derecha en la igualdad (6.36). Luego,

$$\mu^u \ll m^u$$

para $\hat{\mu}$ -casi toda variedad inestable $W \in B/\sim$. Además, la derivada de Radon-Nikodym de μ^u con respecto a m^u (la densidad) es:

$$\frac{d\mu^u}{dm^u}(x) = h.$$

Usando las propiedades (6.24) del Lema de Distorsión Acotada, $1/h$ es una función continua y acotada superiormente. Entonces

$$\frac{dm^u}{d\mu^u}(x) = \frac{1}{h},$$

lo cual implica que

$$m^u \ll \mu^u,$$

terminando de demostrar que μ es una medida de Gibbs, que sus medidas condicionales inestables μ^u son equivalentes a las medidas de Lebesgue m^u a lo largo de las variedades inestables, para μ -c.t.p., y que su función densidad inestable h es continua y estrictamente positiva. Esto completa la prueba de la Afirmación I del párrafo §6.7.2. \square

Nota: Además, usando la igualdad (6.23) hemos probado que la densidad de las medidas condicionales inestables de μ a lo largo de la variedad inestable local $W_\delta^u(x_i)$, es

$$h(x) = \frac{\prod_{h=0}^{+\infty} J^u(x_i)}{\prod_{h=0}^{+\infty} J^u(x)},$$

demostrando también lo afirmado en la Observación 6.3.6.

Segunda parte de la demostración del Teorema 6.3.4

(Prueba de la Afirmación II de §6.7.2)

Ejercicio 6.7.3. Probar la Afirmación II de §6.7.2.

Sugerencia: Sea dada μ , una medida de probabilidad de Gibbs ergódica con densidad inestable continua. Hay que probar que existe un conjunto medible con μ -medida 1 y un abierto que lo contiene, tal que Lebesgue c.t.p. de ese abierto está en la cuenca de atracción estadística B de μ . Para μ -c.t.p. $x_0 \in M$, toda bola $B_\delta(x_0)$ tiene μ -medida positiva. Probemos que Lebesgue casi todo punto de $B_\delta(x_0)$, para $\delta > 0$ suficientemente pequeño que depende de x_0 , pertenece a la cuenca B . Consideremos la densidad $h(x_0)$ de $\mu_{x_0}^u$ respecto la medida de Lebesgue inestable $m_{x_0}^u$. Cualquier función densidad es no negativa. Por definición de la función densidad inestable $h = d\mu^u/dm^u$, y por el teorema de descomposición de Rohlin, $h_{x_0} > 0$ para μ -c.t.p. $x_0 \in M$. Como por hipótesis h es continua, entonces $h_x > 0$ para todo $x \in W_\delta^u(x_0)$, si tomamos $\delta > 0$ suficientemente pequeño (dependiendo del punto x_0).

Ahora podemos aplicar los mismos argumentos de la demostración del Teorema 6.3.1, usando la condición (6.9) de continuidad absoluta de la holonomía h_s a lo largo de la foliación estable, establecida en el Teorema 6.5.2 de la Teoría de Pesin. En extenso, por la construcción sugerida antes, la medida condicionada inestable μ^u a lo largo de la variedad $W_\delta^u(x_0)$ tiene densidad estrictamente positiva. Entonces es equivalente a la medida de Lebesgue m^u en esa variedad inestable local. Usar esta propiedad, y la ergodicidad de μ para probar que m^u c.t.p. de $W_\delta^u(x_0)$ pertenece a B (i.e. μ -c.t.p. pertenece a la cuenca B , lo cual implica que μ^u c.t.p. de $W_\delta^u(x_0)$ pertenece a B debido al Teorema 6.1.7). Usar (6.9) en ambas direcciones, para deducir que la medida de Lebesgue m de $h_s^{-1}(W_\delta^u(x_0) \setminus B) \subset M$ es cero. Finalmente aplicar el Teorema 6.7.1 de estructura de producto local del difeomorfismo de Anosov f , para deducir que $h_s^{-1}(W_\delta^u(x_0)) = B_\delta(x_0)$, y el mismo argumento que en la prueba del Teorema 6.3.1 para demostrar que $h_s^{-1}(B \cap W_\delta^u(x_0)) \subset B$ y concluir que m -c.t.p. de $B_\delta(x_0)$ pertenece a B . Concluimos que para μ -c.t.p. $x_0 \in M$ existe $\delta > 0$ (que puede depender de x_0) tal que m -c.t.p. de $B_\delta(x_0)$ pertenece a la cuenca de atracción estadística B de la medida de Gibbs ergódica μ . Tomando la unión de esas bolas abiertas, se obtiene un abierto que contiene al soporte de μ y tal que m -c.t.p. de ese abierto está contenido en la cuenca B .

Tercera y última parte de la demostración del Teorema 6.3.4

(Prueba de la Afirmación III de §6.7.2)

Demostración. Debido a la Afirmación I ya probada, existen medidas de Gibbs μ cuyas medidas condicionales inestables μ^u satisfacen las dos propiedades en dicha afirmación. Aplicando el Teorema 6.3.1, la parte (b) del Corolario 6.2.4, y la Definición 5.5.6, deducimos que la cuenca de atracción estadística de las componentes ergódicas de μ tienen medida de Lebesgue positiva, y estas componentes ergódicas son medidas de Gibbs ergódicas cuyas condicionales inestables satisfacen las dos propiedades de la Afirmación I. Por lo tanto, el conjunto B formado por todos los puntos que pertenecen a las cuencas de atracción estadística de medidas de Gibbs ergódicas que satisfacen esas dos propiedades, cumple

$$m(B) > 0.$$

Nota: B es medible: En efecto, , las medidas de Gibbs ergódicas son a lo sumo, una cantidad numerable, porque por el Teorema 6.3.1 y la Definición 5.5.6, sus cuencas de atracción estadística son disjuntas dos a dos y tienen medida de Lebesgue positiva. La cuenca de atracción estadística de cualquier medida de probabilidad, por su construcción dada en la Definición 5.5.1, es un conjunto medible. Luego, la unión numerable B de conjuntos medibles, es medible.

Queremos probar que $m(B) = 1$, es decir $m(C) = 0$ donde

$$C = M \setminus B.$$

Como la cuenca de atracción estadística de cualquier medida de probabilidad es invariante con f , tenemos:

$$f(B) = B, \quad f(C) = C.$$

Tomemos $\delta > 0$ suficientemente pequeño. Como $m(B) > 0$, deducimos que:

$$\text{Existe una bola } B_\delta(x_0) \text{ tal que } m(B \cap B_\delta(x_0)) > 0.$$

Afirmamos que

$$\text{A probar: } \forall x_0 \in M, \quad m(B \cap B_\delta(x_0)) > 0 \Rightarrow m(C \cap B_\delta(x_0)) = 0. \quad (6.37)$$

Primero veamos que una vez probada la afirmación (6.37), se deduce la afirmación (III). En efecto, si la variedad M fuera conexa, entonces la afirmación (III) implica que las bolas abiertas de radio δ se clasifican en dos clases disjuntas: la de aquellas bolas tales que Lebesgue-c.t.p. de ella pertenece a B , y la de aquellas bolas tales que Lebesgue c.t.p. pertenece al complemento de B , o sea a C . Como la variedad es conexa, una de las dos clases es vacía. Como existe una bola $B_\delta(x_0)$ en la primera clase, entonces la segunda clase es vacía. Esto prueba que para toda bola de radio δ , $m(C \cap B_\delta) = 0$. Luego $0 = m(C) = m(M \setminus B)$, de donde deducimos que $m(B) = 0$, concluyendo la Afirmación III, cuando la variedad M es conexa. Y si la variedad M no es conexa, como es compacta por hipótesis, tiene una cantidad finita $k \geq 2$ de componentes conexas M_1, \dots, M_k . Siendo f un difeomorfismo de clase $C^{1+\alpha}$, en particular, es continuo. Entonces existe un $p \geq 1$ tal que $f^p : M_i \mapsto M_i$ para todo $1 \leq i \leq k$, y $f^p|_{M_i} \in \text{Diff}^{1+\alpha}(M_i)$. Como f es Anosov, $f^p|_{M_i}$ es Anosov, donde M_i es una variedad compacta y conexa. Por lo probado antes $m(C \cap M_i) = 0$ para todo $1 \leq i \leq k$. Luego $m(C) = 0$, terminando de probar la Afirmación III de §6.7.2, a partir de (6.37).

Ahora solo resta demostrar (6.37). Supongamos por absurdo que

$$m(B \cap B_\delta(x_0)) > 0, \quad m(C \cap B_\delta(x_0)) > 0.$$

Denotemos

$$W_0 = W_\delta^u(x_0)$$

a la variedad inestable local por x_0 en la bola $B_\delta(x_0)$ (es decir, $W_\delta^u(x_0)$ es la componente conexa que contiene al punto x_0 de la intersección $W^u(x_0) \cap B_\delta(x_0)$).

Debido a la propiedad (6.9), establecida en el Teorema 6.5.2 de continuidad absoluta de la holonomía estable, tenemos

$$m^u(W_0 \cap B) > 0, \quad m^u(W_0 \cap C) > 0$$

donde m^u es la medida de Lebesgue inestable a lo largo de la variedad inestable local W_0 . Construyamos, como en el paso 1 de la demostración de la Afirmación I en §6.7.2, las siguientes medidas de probabilidad en M . Para cualquier boreliano $A \subset M$, definimos:

$$\mu_0(A) := \frac{m^u(W_0 \cap A)}{m^u(W_0)} = \lambda \mu_{0,B}(A) + (1 - \lambda) \mu_{0,C}(A),$$

donde

$$\lambda := \frac{m^u(B \cap W_0)}{m^u(W_0)}, \quad 1 - \lambda := \frac{m^u(C \cap W_0)}{m^u(W_0)},$$

$$\mu_{0,B}(A) := \frac{m^u(W_0 \cap A \cap B)}{m^u(W_0 \cap B)}, \quad \frac{m^u(W_0 \cap A \cap C)}{m^u(W_0 \cap C)}.$$

$$\nu_n(A) := \mu_0(f^{-n}(A)) = \lambda \nu_{n,B}(A) + (1 - \lambda) \nu_{n,C}(A),$$

donde

$$\nu_{n,B}(A) := \mu_{0,B}(f^{-n}(A)), \quad \nu_{n,C}(A) := \mu_{0,C}(f^{-n}(A)),$$

$$\mu_n := \frac{1}{n} \sum_{j=0}^{n-1} \nu_j = \lambda \mu_{n,B} + (1 - \lambda) \mu_{n,C},$$

donde

$$\mu_{n,B} := \frac{1}{n} \sum_{j=0}^{n-1} \nu_{j,B}, \quad \mu_{n,C} := \frac{1}{n} \sum_{j=0}^{n-1} \nu_{j,C}.$$

Tomamos una subsucesión $\{n_i\}_{i \geq 1}$ tal que las sucesiones de medidas de probabilidad $\{\mu_{n_i}\}_{i \geq 1}$, $\{\mu_{n_i,B}\}_{i \geq 1}$ y $\{\mu_{n_i,C}\}_{i \geq 1}$ son convergentes en la topología débil estrella. Es decir, existen medidas de probabilidad μ , μ_B y μ_C tales que:

$$\mu = \lim_{i \rightarrow +\infty} \mu_{n_i}, \quad \mu_B = \lim_{i \rightarrow +\infty} \mu_{n_i,B}, \quad \mu_C = \lim_{i \rightarrow +\infty} \mu_{n_i,C}.$$

Entonces:

$$\begin{aligned} \mu &= \lim_{i \rightarrow +\infty} \mu_{n_i} = \lim_{i \rightarrow +\infty} \lambda \mu_{n_i,B} + (1 - \lambda) \mu_{n_i,C} = \\ &= \lambda \lim_{i \rightarrow +\infty} \mu_{n_i,B} + (1 - \lambda) \lim_{i \rightarrow +\infty} \mu_{n_i,C} = \lambda \mu_B + (1 - \lambda) \mu_C. \end{aligned}$$

Luego

$$\mu_C \ll \mu.$$

Por el teorema de Descomposición Ergódica, como $\mu_C \ll \mu$, deducimos que:

Las componentes ergódicas $\mu_{C,x}$ de μ_C coinciden con las componentes ergódicas μ_x de μ , para μ_C -c.t.p. $x \in M$.

Por lo demostrado en la Afirmación I de §6.7.2, la medida de probabilidad μ es de Gibbs y sus medidas condicionales inestables satisfacen las dos propiedades de la Afirmación I. Luego, usando las partes (b) y (d) del Corolario 6.2.4, obtenemos:

Las componentes ergódicas μ_x de μ son medidas ergódicas de Gibbs y sus medidas condicionales inestables μ_x^u satisfacen las dos propiedades de la Afirmación I, para μ -c.t.p. $x \in M$.

De las dos enunciados probados arriba, concluimos que:

Las componentes ergódicas de μ_C con medidas ergódicas de Gibbs cuyas condicionales inestables satisfacen las dos propiedades de la Afirmación I, para μ_C -c.t.p. $x \in M$.

Ahora aplicamos la Afirmación II ya probada, para deducir el siguiente resultado:

Enunciado (a) ya probado: *Lebesgue c.t.p. de un abierto V que contiene al soporte de μ_C , pertenece al conjunto $B = M \setminus C$.*

Pero, por construcción de la medida $\mu_C = \lim_i \mu_{n_i,C}$, si tomamos una función real continua no negativa $\psi : M \rightarrow [0, 1]$ soportada en V y tal que $\psi_V > 0$, se cumple:

$$0 < \int \chi_V \psi d\mu_C = \int \psi d\mu_C = \lim_{i \rightarrow +\infty} \int \psi d\mu_{n_i,C} =$$

$$= \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \int \psi d\nu_{j,C}.$$

Entonces, existe $j \geq 1$ tal que

$$0 < \int \psi d\nu_{j,C} = \int \psi \circ f^{-j} d\mu_{0,C} = \int (\chi_V \cdot \psi) \circ f^{-j} d\mu_{0,C}.$$

(La última igualdad proviene de que ψ está soportada en V , es decir $\chi_V \cdot \psi = \psi$). Por la construcción de $\mu_{0,C}$, y debido a la positividad establecida en la última desigualdad, deducimos que

$$m^u(f^{-j}(V) \cap C \cap W_0) > 0.$$

Como f es un difeomorfismo y $f^j(C) = C$, deducimos

$$m^u(V \cap C \cap f^j(W_0)) > 0. \quad (6.38)$$

Ahora tomamos una bola abierta $B_\delta(x_1) \subset V$, de radio $\delta > 0$ suficientemente pequeño, centrada en un punto $x_1 \in V \cap C \cap f^j(W_0)$ tal que

$$m^u(C \cap W_\delta^u(x_1)) > 0, \quad (6.39)$$

donde $W_\delta^u(x_1)$ es la componente conexa que contiene a x_1 de la intersección $B_\delta(x_1) \cap f^j(W^u(x_0))$, quien a su vez contiene a la componente conexa por el punto x_1 de $B_\delta(x_1) \cap f^j(W_0)$, cuya intersección con C , debido a (6.38), tiene medida de Lebesgue m^u positiva.

Aplicamos la propiedad (6.9) de continuidad absoluta de la holonomía estable $h_s : B_\delta(x_1) \mapsto W_\delta^u(x_1)$, establecida en el Teorema 6.5.2. De (6.39) deducimos

$$m(h_s^{-1}(C \cap W_\delta^u(x_1))) > 0.$$

Como $h_s^{-1}(C) \subset C$ y $h_s^{-1}(W_\delta^u(x_1)) \subset B_\delta(x_1) \subset V$, deducimos que $m(C \cap V) > 0$, lo cual contradice el enunciado (a) probado antes. Esta contradicción termina de demostrar (6.37) por absurdo, y por lo tanto, con ella termina la prueba de la Afirmación III de §6.7.2 y del Teorema 6.3.4. \square

Capítulo 7

Atractores estadísticos y medidas SRB-like

A lo largo de este capítulo $f : M \mapsto M$ es una transformación continua en una variedad compacta riemanniana M de dimensión finita. Denotamos con m a la medida de Lebesgue en M , re-escalada para que sea una medida de probabilidad: $m(M) = 1$ (i.e. si $0 < m(M) \neq 1$, sustituimos m por la probabilidad $m/m(M)$).

Definiremos primero atractor de Milnor. Este no es en general un atractor estadístico. Sin embargo, la demostración de su existencia es prácticamente la misma que la demostración de existencia de atractores estadísticos, que veremos en la sección siguiente.

7.1. Atractores de Milnor

Definición 7.1.1. Atractor de Milnor [Mi 1985] Se llama *atractor de Milnor* a un conjunto compacto no vacío $K \subset M$, invariante por f (i.e. $f^{-1}(K) = K$), tal que

$$m(E_K) = 1,$$

donde el conjunto $E_K \subset M$, llamado *cuenca de atracción (topológica)* de K , está definido por

$$E_K := \{x \in M : \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K) = 0\}. \quad (7.1)$$

Ejercicio 7.1.2. Probar que para cualquier conjunto compacto K no vacío, el conjunto E_K definido por la igualdad (7.1) es medible. Sugerencia: La función $d : M \mapsto \mathbb{R}$ definida por $d(x) = \text{dist}(x, K)$ es medible, y $f : M \mapsto M$ es medible porque es continua. El límite superior de una sucesión de funciones medibles es medible.

Ejercicio 7.1.3. Sean $K \subset K'$ dos atractores de Milnor. Probar que $E_K \subset E_{K'}$.

En la Definición 7.1.1 observamos que para un atractor de Milnor, el criterio de observabilidad de su cuenca es el criterio Lebesgue-medible (cf. Observación 5.4.7), mientras que el tipo de atracción es topológica (cf. Definiciones 5.4.8 y 5.4.9).

Es inmediato chequear que todo atractor topológico es de Milnor. Sin embargo, no todo atractor de Milnor es topológico, como veremos en los Ejemplos 7.1.7 y 7.1.9.

También es inmediato chequear que todo atractor ergódico es de Milnor. Pero no todo atractor de Milnor es ergódico, como muestra el Ejemplo 5.4.12, en que toda la esfera S^2 es un atractor topológico, y por lo tanto un atractor de Milnor, pero no existen atractores ergódicos.

Definición 7.1.4. α -obs. minimalidad de un atractor de Milnor [Cat 2012]

Sea dado un número real $0 < \alpha \leq 1$. Un atractor de Milnor K se dice α -observable (escribimos “ K es α -obs”.) si su cuenca E_K de atracción (topológica) cumple

$$m(E_K) \geq \alpha.$$

Un atractor K de Milnor α -obs. se dice α -obs. *minimal* si no contiene subconjuntos compactos propios no vacíos que sean atractores de Milnor α -obs. para el mismo valor de α .

En particular, cuando $\alpha = 1$, tenemos definidos los atractores de Milnor 1-observables y 1-observables minimales.

Se observa que todo atractor de Milnor 1-observable es α -observable para cualquier $0 < \alpha \leq 1$. Pero un atractor de Milnor 1-obs. minimal no tiene por qué ser α -obs. minimal para todo $0 < \alpha < 1$ (ver Ejemplos 7.1.7 y 7.1.9).

Teorema 7.1.5. Existencia de atractores de Milnor

Sea $f : M \mapsto M$ continua en una variedad compacta y riemanniana M , de dimensión finita. Sea $0 < \alpha \leq 1$ dado. Entonces existen atractores de Milnor α -obs. minimales para f . Además, si $\alpha = 1$, el atractor de Milnor 1-obs. minimal es único.

La prueba del Teorema 7.1.5, en una versión más restringida que enuncia solo la existencia y unicidad del atractor de Milnor 1-observable minimal, fue dada primeramente en ([Mi 1985]). En el apéndice de [Cat 2012] se define α -obs. minimalidad de atractores de Milnor para $0 < \alpha \leq 1$, y se observa que la demostración de existencia de atractor de Milnor dada en ([Mi 1985]) puede aplicarse, con una inmediata adaptación, para probar la existencia de atractores de Milnor α -obs. minimales, para cualquier $0 < \alpha \leq 1$.

Demostración. del Teorema 7.1.5: Sea \aleph_α la familia de los atractores de Milnor α -obs. (no necesariamente minimales). Esta familia no es vacía pues, trivialmente, M es un atractor de Milnor 1-obs. En \aleph_α consideramos la relación de orden parcial $K_1 \subset K_2$. Entonces, las cuencas de atracción topológica E_{K_1} y E_{K_2} cumplen

$$E_{K_1} \subset E_{K_2}, \quad \alpha \leq m(E_{K_1}) \leq m(E_{K_2}).$$

Sea en \aleph_α una cadena $\{K_i\}_{i \in I}$ (no necesariamente numerable). Es decir, $\{K_i\}_{i \in I}$ es un subconjunto totalmente ordenado de \aleph_α , con la relación de orden \subset .

Probemos que:

Afirmación (i) (A probar) *Existe en \aleph_α un elemento K minimal de la cadena $\{K_i\}_{i \in I}$. Es decir, probemos que existe $K \in \aleph_\alpha$, $K \subset K_i$ para todo $i \in I$.*

En efecto, el conjunto $K = \bigcap_{i \in I} K_i$ es compacto no vacío, porque cualquier subcolección finita $K_1 \supset K_2 \supset \dots \supset K_l$ de la cadena dada $\{K_i\}_{i \in I}$, tiene intersección K_1 que es un compacto no vacío. Para probar que $K \in \aleph_\alpha$, debemos probar ahora que $m(E_K) \geq \alpha$. Sea $j \in \mathbb{N}^+$ y sea $V_j \supset K$ el abierto formado por todos los puntos de M que distan de K menos que $1/j$. Afirmamos que existe

$$K_{i_j} \subset V_j \text{ para algún } i_j \in I. \quad (7.2)$$

Por absurdo, supongamos que para cierto $j \in \mathbb{N}^+$ fijo, para todo $i \in I$, el compacto $K_i \setminus V_j$ es no vacío. Luego, como $\{K_i\}_{i \in \mathbb{N}}$ es totalmente ordenado con la relación de orden \subset , obtenemos que la familia de compactos $\{K_i \setminus V_j\}_{i \in I}$ es totalmente ordenada. Argumentando como hicimos más arriba, pero con esta nueva familia totalmente ordenada de compactos, en vez de con la familia $\{K_i\}_{i \in I}$, deducimos que el siguiente compacto es no vacío

$$\bigcap_{i \in I} (K_i \setminus V_j) = \left(\bigcap_{i \in I} K_i \right) \setminus V_j = K \setminus V_j,$$

contradiendo que $K \subset V_j$. Hemos probado la afirmación (7.2).

Como todo punto $x \in E_{K_{i_j}}$ cumple $\lim_n \text{dist}(f^n(x), K_{i_j}) = 0$, entonces $f^n(x) \in V_j$ para todo n suficientemente grande (que depende del punto x). Este argumento vale para cualquier $j \in \mathbb{N}^+$. Concluimos que todo punto en $\bigcap_{j \in \mathbb{N}^+} E_{K_{i_j}}$ pertenece a E_K . Recíprocamente, todo punto de E_K , por definición de la cuenca de atracción topológica, está contenido en E_{K_i} para todo $i \in I$ (porque $K \subset K_i$). En particular, esta afirmación se satisface para i_j , para todo $j \in \mathbb{N}^+$. Luego:

$$E_K := \bigcap_{j \in \mathbb{N}^+} E_{K_{i_j}}.$$

Como la colección numerable $E_{K_{i_j}}$ está totalmente ordenada, obtenemos

$$m(E_K) = m\left(\bigcap_{j \in \mathbb{N}^+} E_{K_{i_j}}\right) = \lim_{j \rightarrow +\infty} m(E_{K_{i_j}}) \geq \alpha,$$

terminando de demostrar la afirmación (i).

De la afirmación (i) se deduce que para toda cadena en \aleph_α existe algún elemento $K \in \aleph_\alpha$ minimal de la cadena. Aplicando el Lema de Zorn, existen en \aleph_α elementos minimales de \aleph_α . Es decir, existe $K \in \aleph_\alpha$ que no contiene subconjuntos propios que pertenezcan a \aleph_α . Esto es, existe K atractor de Milnor α -obs. minimal.

Ahora probemos la unicidad del atractor de Milnor 1-obs. minimal. Si existieran dos atractores de Milnor K_1 y K_2 que fueran 1-obs. minimales, entonces la intersección E de sus cuencas de atracción topológica $E := E_{K_1} \cap E_{K_2}$ cumple

$$m(E) = 1,$$

porque $m(E_{K_1}) = m(E_{K_2}) = 1$. Todo punto $x \in E$ verifica, por la definición de la cuenca E_{K_1} y E_{K_2} , la siguiente propiedad:

Para todo $\epsilon > 0$ existe $N \geq 1$ tal que

$$\text{dist}(f^n(x), K_1), \quad \text{dist}(f^n(x), K_2) < \epsilon \quad \forall n \geq N$$

Luego $\text{dist}(K_1, K_2) < 2\epsilon \quad \forall \epsilon > 0$, de donde $K := K_1 \cap K_2 \neq \emptyset$. Es estándar chequear que, siendo K_1 y K_2 compactos no vacíos con intersección no vacía, si un punto x cumple

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K_1) = \lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K_2) = 0,$$

entonces

$$\lim_{n \rightarrow +\infty} \text{dist}(f^n(x), K_1 \cap K_2) = 0.$$

(Chequear esta última afirmación en la parte (a) del Ejercicio 7.1.6.) Luego

$$E = E_{K_1} \cap E_{K_2} \subset E_{K_1 \cap K_2}$$

y como $m(E) = 1$, deducimos que $m(E_{K_1 \cap K_2}) = 1$. Entonces $K_1 \cap K_2$ es un atractor de Milnor 1-obs. Como K_1 y K_2 eran atractores de Milnor 1-obs. minimales, concluimos que $K_1 \cap K_2 = K_1 = K_2$, terminando de demostrar la unicidad del atractor de Milnor 1-obs. minimal, y el Teorema 7.1.5. \square

Ejercicio 7.1.6. (a) Demostrar que si K_1 y K_2 son compactos no vacíos, y si existe una sucesión de puntos $x_n \in M$ tal que

$$\lim_{n \rightarrow +\infty} \text{dist}(x_n, K_1) = \lim_{n \rightarrow +\infty} \text{dist}(x_n, K_2) = 0,$$

entonces $K_1 \cap K_2 \neq \emptyset$ y

$$\lim_{n \rightarrow +\infty} \text{dist}(x_n, K_1 \cap K_2) = 0.$$

(b) Demostrar que si K_1 y K_2 son dos atractores de Milnor tales $m(E_{K_1} \cap E_{K_2}) > 0$, entonces $K_1 \cap K_2$ es no vacío, y es un atractor de Milnor cuya cuenca de atracción topológica es

$$E_{K_1 \cap K_2} = E_{K_1} \cap E_{K_2}.$$

Ejemplo 7.1.7. Sea $X = [0, 4\pi] \times [0, 1]$ el rectángulo compacto de ancho 4π y altura 1, con un vértice en el origen, cuyo interior está contenido en el primer cuadrante y tiene lados paralelos a los ejes. Sea en el intervalo $[0, 4\pi]$ la transformación

$$f_1(x) = 1 + x - \cos x \quad \forall x \in [0, 4\pi]$$

y sea en X la transformación

$$f(x, y) = (f_1(x), y/2) \quad \forall (x, y) \in [0, 4\pi] \times [0, 1].$$

Sea m la medida de Lebesgue en X , re-escalada para que

$$m(X) = 1.$$

Es estándar chequear (ver Ejercicio 7.1.8), que

$$K_1 := \{(2\pi, 0)\}$$

es un atractor de Milnor con cuenca de atracción topológica

$$E_{K_1} = (0, 2\pi] \times [0, 1],$$

y por lo tanto K_1 es 1/2-obs. minimal como atractor de Milnor.

$$K_2 := \{(4\pi, 0)\}$$

es otro atractor de Milnor con cuenca de atracción topológica

$$E_{K_2} = (2\pi, 4\pi] \times [0, 1],$$

y por lo tanto K_2 es también 1/2-obs. minimal como atractor de Milnor.

K_1 no es atractor topológico, porque todo entorno de K_1 contiene puntos de la cuenca de atracción topológica de K_2 .

K_2 es atractor topológico.

$K_1 \cup K_2$ es el único atractor de Milnor 1-obs. minimal, y también es el único atractor α -obs, si $1/2 < \alpha \leq 1$.

K_1 y K_2 son los únicos atractores de Milnor α -obs. si $0 < \alpha \leq 1/2$ (y como contienen un solo punto cada uno, son α -obs. minimales).

$K_1 \cup K_2$ es atractor topológico, pero no es minimal como atractor topológico, pues contiene a K_2 que también es atractor topológico.

Ejercicio 7.1.8. Probar todas las afirmaciones del Ejemplo 7.1.7.

Ejemplo 7.1.9. Sea en el cuadrado $X = [0, 1]^2$ la transformación

$$f(x, y) = (x, (1/2)y) \quad \forall (x, y) \in [0, 1]^2.$$

Es inmediato chequear que todos los puntos del segmento $[0, 1] \times \{0\}$ son puntos fijos y que $\lim_{n \rightarrow +\infty} f^n(x, y) = (x, 0) \quad \forall (x, y) \in [0, 1]^2$. Entonces, para todo $0 < \alpha \leq 1$, cualquier segmento compacto $I \times \{0\}$ tal que la longitud del intervalo compacto $I \subset [0, 1]$ sea exactamente α , es un atractor de Milnor α -observable minimal. Sea $K = (I_1 \cup I_2 \cup \dots \cup I_k) \times \{0\}$ donde $I_k \subset [0, 1]$ es un intervalo compacto con interior no vacío, la colección de los intervalos I_k es disjunta dos a dos y tal que la suma de las longitudes de los I_k es exactamente α . Entonces, K también es un atractor de Milnor α -obs. minimal. Además, si $0 < \alpha < 1$ y si $I \subset [0, 1]$ es un conjunto de Cantor con medida de Lebesgue igual a α (tales conjuntos de Cantor siempre existen), entonces $I \times \{0\}$ es también un atractor de Milnor α -obs. minimal. Ninguno de los atractores α -obs. minimales construidos anteriormente, si $0 < \alpha < 1$, es atractor topológico, pues las cuencas de atracción topológica no contienen ningún entorno del atractor. El único atractor 1-observable minimal es el intervalo $[0, 1] \times \{0\}$, que sí es atractor topológico.

7.2. Atractores estadísticos o de Ilyashenko

En esta sección, al igual que en la anterior, $f : M \mapsto M$ es una transformación continua en una variedad compacta riemanniana M de dimensión finita. Denotamos con m a la medida de Lebesgue en M , re-escalada para que sea una medida de probabilidad: $m(M) = 1$, i.e. si $0 < m(M) \neq 1$, sustituimos m por la probabilidad $m/m(M)$.

Definición 7.2.1. .

Atractor estadístico o de Ilyashenko [Il 2005], [Go-II 1996] Se llama *atractor estadístico o de Ilyashenko* a un conjunto compacto no vacío $K \subset M$, invariante por f (i.e. $f^{-1}(K) = K$), tal que

$$m(A_K) = 1,$$

donde el conjunto $A_K \subset M$, llamado *cuenca de atracción estadística* de K , está definido por

$$A_K := \{x \in M : \lim_{n \rightarrow +\infty} \frac{1}{j} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) = 0\}. \quad (7.3)$$

Ejercicio 7.2.2. Probar que para cualquier conjunto compacto K no vacío, el conjunto A_K definido por la igualdad (7.1) es medible. Sugerencia: La misma que para el Ejercicio 7.1.2.

Ejercicio 7.2.3. Sean $K \subset K'$ dos atractores estadísticos. Probar que $A_K \subset A_{K'}$.

Es estándar chequear que todo atractor de Milnor (y en particular todo atractor topológico y todo atractor ergódico) es estadístico o de Ilyashenko (Ejercicio 7.2.4). Sin embargo, no todo atractor estadístico es de Milnor, como veremos en el Ejemplo 7.5.3. Además, no todo atractor estadístico es ergódico, como muestran los Ejemplos 5.4.12 y 7.1.9.

Ejercicio 7.2.4. (a) Sea $\{a_n\}_{n \in \mathbb{N}}$ una sucesión de reales que converge a $a \in \mathbb{R}$. Probar que la sucesión de promedios $\frac{1}{n} \sum_{j=0}^{n-1} a_j$ converge a a .

(b) Probar que todo atractor de Milnor K es atractor estadístico o de Ilyashenko. Sugerencia: probar, usando la parte (a), que la cuenca de atracción topológica de K está contenida en la cuenca de atracción estadística de K .

(c) Probar que en los Ejemplos 5.4.12 y 7.1.9, los atractores de Milnor K (que por la Parte (b) son también atractores estadísticos) no son atractores ergódicos. Sugerencia: el único atractor de Milnor K que cumple la condición (5.6) de la Definición 5.4.2 de atractor ergódico, no cumple la condición (5.7) de esa definición.

(d) Probar que en el Ejemplo 7.1.7, los atractores de Milnor K (que por la parte (b) también son atractores estadísticos) no son atractores ergódicos, aunque para dos de ellos se cumple la condición (5.7) de la Definición 5.4.2 de atractor ergódico.

7.3. Atracción estadística de un compacto

En la Definición 7.2.1 observamos que para un atractor estadístico, el criterio de observabilidad de su cuenca es el criterio Lebesgue-medible (cf. Observación 5.4.7), mientras que la atracción es estadística (cf. Definiciones 5.4.8 y 5.4.9). En efecto:

Proposición 7.3.1. Caracterización de la atracción estadística

Sea K un atractor estadístico o de Ilyashenko, y sea A_K su cuenca de atracción estadística. Entonces, para todo $x \in M$ se cumple $x \in A_K$ si y solo si toda medida de probabilidad μ que sea límite de alguna subsucesión convergente en la topología débil estrella de las probabilidades empíricas $\sigma_n := \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)}$ (cf. Definición 5.5.2), cumple

$$\mu(K) = 1.$$

Demostración. *Demostración del “solo si”:* Sea $x \in A_K$ y sea

$$\mu = \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \delta_{f^j(x)},$$

para una cierta subsucesión $\{n_i\}_{i \in \mathbb{N}}$ de los naturales tal que ese límite μ existe en la topología débil estrella del espacio de probabilidades. Debemos probar que $\mu(K) = 1$. Por la definición de la topología débil estrella, para toda función continua $\psi : M \rightarrow \mathbb{R}$ se cumple

$$\lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \psi \circ f^j(x) = \int \psi d\mu.$$

En particular, la igualdad anterior se cumple para la función continua

$$\psi(x) := \text{dist}(x, K).$$

Por la Definición 7.2.1, como $x \in A_K$, existe $n \geq 0$ tal que

$$\frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) < \epsilon$$

En particular, para $n = n_i$ para todo i suficientemente grande, tenemos:

$$\frac{1}{n_i} \sum_{j=0}^{n_i-1} \psi(f^j(x)) \geq \epsilon$$

De las igualdades anteriores, deducimos que $\int \psi d\mu = 0$, es decir

$$\int \text{dist}(x, K) d\mu = 0$$

y como $\text{dist}(\cdot, K)$ es una función no negativa, deducimos que $\text{dist}(x, K) = 0$ para μ -c.t.p. $x \in M$. Como K es compacto, deducimos que $x \in K$ para μ -c.t.p. $x \in M$, luego

$$\mu(K) = \int \chi_K(x) d\mu = \int 1 d\mu = 1,$$

como queríamos probar.

Demostración del “si”: Sea dado $x \in M$ tal que toda medida μ de probabilidad que sea el límite débil estrella de una subsucesión convergente de las probabilidades empíricas $\sigma_n(x)$, está soportada en K . Tenemos que probar que $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) = 0$. Sea una subsucesión n_i tal que existe el límite

$$\lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \text{dist}(f^j(x), K) = a.$$

No es restrictivo suponer que la subsucesión $\sigma_{n_i}(x)$ es convergente (en caso contrario, reemplazamos $\{n_i\}$ por una subsucesión adecuada de ella). Entonces, llamando μ al límite de σ_{n_i} , e integrando la función continua $\psi = \text{dist}(\cdot, K)$ obtenemos:

$$\begin{aligned} a &= \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \text{dist}(f^j(x), K) = \lim_{i \rightarrow +\infty} \int \psi(y) d(\sigma_{n_i}(x))(y) = \\ &= \int \psi d\mu = \int \text{dist}(y, K) d\mu(y). \end{aligned} \quad (7.4)$$

Pero la última integral en (7.4) es igual a cero, pues de la hipótesis $\mu(K) = 1$ deducimos que $y \in K$ para μ -c.t.p. Luego $\text{dist}(y, K) = 0$ para μ -c.t.p. $y \in M$. Entonces, la igualdad (7.4) implica que

$\lim_n \frac{1}{n} \text{dist} \sum_{j=0}^{n-1} (f^j(x), K) = 0$ (pues el límite de cualquier subsucesión convergente es $a = 0$). Entonces $x \in A_K$, terminando de demostrar la Proposición 7.3.1. \square

Definición 7.3.2. Sea μ una medida de probabilidad en los borelianos de M . Se llama *soporte compacto de μ* al mínimo compacto $K \subset M$ tal que $\mu(K) = 1$. El soporte compacto existe, debido al Lema de Zorn y a la propiedad de que cualquier familia de compactos tiene intersección no vacía si las intersecciones de todas las subfamilias finitas son no vacías. El soporte compacto de μ es único, pues si existieran dos diferentes K_1 y K_2 , entonces $\mu(K_1 \cap K_2) = 1$ y ni K_1 ni K_2 tendrían la propiedad de minimalidad.

Proposición 7.3.3. Medidas SRB y atractores estadísticos

Sea $f : M \mapsto M$ continua en la variedad compacta M . Si existe una medida μ que es SRB para f (cf. Definición 5.5.6), entonces el soporte compacto K de μ es un atractor estadístico o de Ilyashenko.

Observación 7.3.4. El recíproco de la Proposición 7.3.3 es falso. En efecto, las medidas SRB no siempre existen (cf. Ejemplos 5.4.12 y 7.1.9), pero los atractores de Ilyashenko siempre existen, como demostraremos en el Teorema 7.4.3.

A pesar de que el recíproco de la Proposición 7.3.3 es falso, se puede generalizar el enunciado de esta Proposición de forma que su recíproco es verdadero. Más precisamente, un compacto K no vacío e invariante es atractor estadístico o de Ilyashenko, si y solo sí, es el mínimo soporte compacto de una colección adecuada de medidas invariantes, que llamamos SRB-like, y que describen en forma óptima la estadística de las órbitas de un conjunto de medida de Lebesgue positivo. En efecto, en la próxima

sección definiremos las medidas de probabilidad “SRB-like” o “pseudo-físicas”, que incluyen a las medidas SRB o físicas (cuando éstas existen), pero también pueden incluir a otras medidas de probabilidad invariantes que no son necesariamente SRB. El nuevo enunciado generalizado de la Proposición 7.3.3 utilizando las medidas SRB-like, en lugar de las medidas SRB, y su recíproco, será establecido y demostrado en el Teorema 7.7.2.

Demostración. de la Proposición 7.3.3

Para probar que K es atractor estadístico, por la Definición 7.2.1, hay que probar que $m(A_K) > 0$, donde m denota la medida de Lebesgue y A_K denota la cuenca de atracción estadística de K definida en (7.3). Como μ es medida SRB, por la Definición 5.5.6, la cuenca de atracción estadística $B(\mu)$ de μ , dada en la Definición (5.5.1), cumple $m(B(\mu)) > 0$. Luego, basta probar que $B(\mu) \subset A_K$. Sea $x \in B(\mu)$. Entonces, por la Definición 5.5.1, para cualquier función continua $\psi : M \rightarrow \mathbb{R}$ se cumple:

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ f^j(x) = \int \psi d\mu.$$

En particular si tomamos $\psi(x) = \text{dist}(x, K)$ tenemos:

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) = \int \text{dist}(y, K) d\mu$$

Para concluir que $x \in A_K$ basta probar que la integral en la igualdad anterior es cero, y para ello basta chequear que $\text{dist}(y, K) = 0$ para μ -c.t.p. $y \in M$. En efecto, μ -c.t.p. $y \in M$ pertenece a K porque $\mu(K) = 1$. Luego, hemos probado que $x \in A_K$ para todo $x \in B(\mu)$, y por lo tanto $m(A_K) \geq m(B(\mu)) > 0$, y K es un atractor de Ilyashenko. \square

Definición 7.3.5. Tiempo medio de estadía Sea $V \subset M$ un conjunto medible no vacío. Sea $x \in M$ cualquiera. Llamamos *frecuencia de visitas a V* de la órbita por x hasta tiempo n , al siguiente número $\sigma_{n,x}(V)$ (cf. (5.9) en la Definición 5.5.2 de las probabilidades empíricas $\sigma_{n,x}$):

$$\begin{aligned} \sigma_{n,x}(V) &:= \int \chi_V d\sigma_{n,x} = \frac{1}{n} \sum_{j=0}^{n-1} \chi_V(f^j(x)) = \\ &= \frac{\#\{0 \leq j \leq n-1 : f^j(x) \in V\}}{n}, \end{aligned} \tag{7.5}$$

donde χ_V denota la función característica de V y $\#A$ denota cantidad de elementos de un conjunto finito A .

El último término de la igualdad (7.5) indica que el tiempo medio de estadía en V de la órbita por x hasta tiempo n es la cantidad relativa de iterados del punto x que caen dentro del conjunto V . Por el teorema ergódico de Birkhoff, si μ es una medida invariante, entonces para μ -c.t.p. $x \in M$ el tiempo medio de estadía en V de la órbita

por x hasta tiempo n , converge cuando $n \rightarrow +\infty$, a $\tilde{\chi}_V$, cuyo valor esperado $\int \tilde{\chi}_V d\mu$ es igual a $\mu(V)$. Además si la medida μ es ergódica, el tiempo medio de estadía en V tiende a $\mu(V)$ para μ -c.t.p. $x \in M$.

Sin embargo, el enunciado del Teorema de Birkhoff y las propiedades de las medidas invariantes ergódicas, no nos serán útiles a los propósitos en esta sección. Nos interesa considerar el tiempo medio de estadía en V para Lebesgue c.t.p., aunque la medida de Lebesgue no sea invariante con f , y aunque los puntos considerados x no pertenezcan al soporte de ninguna medida invariante μ .

Proposición 7.3.6. .

Caracterización de atracción estadística por el tiempo medio de estadía en un entorno

Sea K un compacto no vacío invariante con f . Para todo $\epsilon > 0$ denotamos con $V(\epsilon)$ al conjunto de puntos en la variedad M que distan de K menos que ϵ .

Entonces, la cuenca de atracción estadística A_K del conjunto K (definida en la igualdad (7.3) de la definición 7.2.1) está caracterizada por la siguiente igualdad:

$$A_K = \bigcap_{\epsilon > 0} \{x \in M : \lim_{n \rightarrow +\infty} \sigma_{n,x}(V(\epsilon)) = 1\}. \quad (7.6)$$

Notas: La igualdad (7.6) implica que $x \in A_K$ si y solo si, cuando n es suficientemente grande, la probabilidad (la frecuencia relativa) del suceso de encontrar al iterado $f^j(x)$, para algún $0 \leq j \leq n-1$, en un entorno $V(\epsilon)$ dado fijo arbitrariamente pequeño del atractor de Ilyashenko K , es 1. Sin embargo, la atracción de la órbita por x al conjunto K , no es necesariamente atracción topológica. Para ser atracción topológica, debe cumplirse con total certeza (no solo con probabilidad cercana a 1, ni tampoco solo con probabilidad 1) el suceso de encontrar al iterado $f^j(x)$ en el entorno $V(\epsilon)$, para todo j suficientemente grande.

La igualdad (7.6) da por lo tanto, la siguiente interpretación intuitiva del significado de la atracción estadística a un atractor de Ilyashenko K . Las órbitas en su cuenca de atracción estadística A_K se acercan al conjunto K tanto como se desee cuando el iterado n tiende a infinito, pero se admite que existan iterados excepcionales, con una frecuencia relativa despreciable para n arbitrariamente grande, en que la órbita “se toma una excursión relativamente muy breve (de vacaciones)” alejándose del atractor K durante la excursión (ver Ejemplo 7.5.3).

Demostración. de la Proposición 7.3.6: (i) Probemos que si $x \in A_K$ entonces $\lim_{n \rightarrow +\infty} \sigma_{n,x}(V(\epsilon)) = 1$ para todo $\epsilon > 0$. Fijemos $\epsilon > 0$. Por simplicidad, escribimos V en lugar de $V(\epsilon)$, ya que ϵ está fijo. Tenemos:

$$\epsilon \cdot \chi_{M \setminus V}(y) \leq \text{dist}(y, K) \quad \forall y \in M,$$

pues, cuando $y \in V$ el término de la izquierda es cero, y cuando $y \notin V$, el término de la izquierda es $\epsilon \leq \text{dist}(y, K)$. Entonces:

$$0 \leq \epsilon \cdot \frac{1}{n} \sum_{j=0}^{n-1} (1 - \chi_V(f^j(x))) \leq \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K).$$

Como $x \in A_K$ por hipótesis, el límite cuando $n \rightarrow +\infty$ del término a la derecha en la desigualdad anterior, es cero. Concluimos que

$$\begin{aligned} 0 &= \epsilon \cdot \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (1 - \chi_V(f^j(x))) = \epsilon \cdot \lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n} \sum_{j=0}^{n-1} \chi_V(f^j(x)) \right) = \\ &= \epsilon \cdot \lim_{n \rightarrow +\infty} \sigma_{n,x}(V). \end{aligned}$$

Como $\epsilon > 0$, concluimos que $\lim_{n \rightarrow +\infty} \sigma_{n,x}(V) = 0$, como queríamos demostrar.

(ii) Probemos que si $\lim_{n \rightarrow +\infty} \sigma_{n,x}(V(\epsilon))$ para todo $\epsilon > 0$ entonces $x \in A_K$. Fijemos $\epsilon > 0$ y por simplicidad denotemos $V = V(\epsilon)$. Sea

$$D := \text{diam}(M) = \max\{\text{dist}(x, y) : x, y \in M\} > 0.$$

Luego:

$$\begin{aligned} \text{dist}(y, K) &\leq D \cdot \chi_{M \setminus V}(y) \quad \forall y \notin V, \\ \text{dist}(y, K) &< \epsilon \cdot \chi_V(y) \quad \forall y \in V. \end{aligned}$$

Entonces

$$\text{dist}(y, K) \leq \epsilon \cdot \chi_V(y) + D \cdot \chi_{M \setminus V}(y) \quad \forall y \in M.$$

Consideramos x que satisface las hipótesis. De la última desigualdad obtenemos:

$$\begin{aligned} 0 &\leq \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) \leq \epsilon \cdot \frac{1}{n} \sum_{j=0}^{n-1} \chi_V(f^j(x)) + D \cdot \frac{1}{n} \sum_{j=0}^{n-1} \chi_{M \setminus V}(f^j(x)) = \\ &= \epsilon \cdot \sigma_{n,x}(V) + D \cdot (1 - \sigma_{n,x}(V)). \end{aligned}$$

Por hipótesis $\lim_{n \rightarrow +\infty} \sigma_{n,x}(V) = 1$. Como los números reales positivos ϵ y D están fijos (son independientes de n), deducimos que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) = 0,$$

La última igualdad significa, debido a la condición (7.3) de la Definición 7.2.1, que $x \in A_K$ como queríamos demostrar. \square

7.4. Existencia de atractores de Ilyashenko

Definición 7.4.1. .

α -obs. minimalidad de un atractor estadístico [Cat 2012]

Sea dado un número real $0 < \alpha \leq 1$. Un atractor estadístico o de Ilyashenko K se dice α -observable (escribimos “ K es α -obs”.) si su cuenca A_K de atracción estadística cumple

$$m(A_K) \geq \alpha.$$

Un atractor K estadístico o de Ilyashenko α -obs. se dice α -obs. *minimal* si no contiene subconjuntos compactos propios no vacíos que sean atractores estadísticos o de Ilyashenko α -obs. para el mismo valor de α .

En particular, cuando $\alpha = 1$, tenemos definidos los atractores estadísticos o de Ilyashenko 1-observables y 1-observables minimales.

Se observa que todo atractor estadístico o de Ilyashenko 1-observable es α -observable para cualquier $0 < \alpha \leq 1$. Pero un atractor estadístico o de Ilyashenko 1-obs. minimal no tiene por qué ser α -obs. minimal para todo $0 < \alpha < 1$.

Observación 7.4.2. Sobre conjuntos minimales

Recordemos la caracterización de los conjuntos minimales desde el punto de vista topológico, dada en la Definición 2.5.3: un conjunto K compacto, no vacío y f -invariante es minimal para f , desde el punto de vista topológico, si no contiene subconjuntos propios compactos no vacíos que sean invariantes por f hacia el futuro.

Observamos que no todo compacto K minimal para f desde el punto de vista topológico es un atractor de Ilyshenko α -obs. minimal para algún $0 \leq \alpha \leq 1$ (ver Ejemplo 7.5.3 o también [Kl 2006]).

Recíprocamente, no todo atractor de Ilyshenko α -obs. minimal es necesariamente un compacto minimal para f desde el punto de vista topológico (ver Ejemplo 7.5.1).

En [Go-II 1996] y [Kl 2006] se estudian relaciones entre minimalidad para f desde el punto de vista topológico y los atractores estadísticos o de Ilyashenko de f .

Teorema 7.4.3. Existencia de atractores de Ilyashenko

Sea $f : M \mapsto M$ continua en una variedad compacta y riemanniana M , de dimensión finita. Sea $0 < \alpha \leq 1$ dado. Entonces existen atractores estadísticos o de Ilyashenko α -obs. minimales para f . Además, si $\alpha = 1$, el atractor de Ilyashenko 1-obs. minimal es único.

La prueba del Teorema 7.4.3 sigue los mismos argumentos de la prueba del Teorema 7.1.5 de existencia de atractores de Milnor, con leves adaptaciones.

Demostración. Sea \aleph_α la familia de los atractores de Ilyashenko α -obs. (no necesariamente minimales). Esta familia no es vacía pues, trivialmente, M es un atractor de Ilyashenko 1-obs, y por lo tanto es α -obs. para cualquier $0 < \alpha \leq 1$. En \aleph_α consideramos la relación de orden parcial $K_1 \subset K_2$. De la condición (7.3), teniendo en cuenta que $\text{dist}(y, K_2) \leq \text{dist}(y, K_1)$ para todo $y \in M$, las cuencas de atracción estadística A_{K_1} y A_{K_2} cumplen

$$A_{K_1} \subset A_{K_2}, \quad \alpha \leq m(A_{K_1}) \leq m(A_{K_2}).$$

Sea en \aleph_α una cadena $\{K_i\}_{i \in I}$ (no necesariamente numerable). Es decir, $\{K_i\}_{i \in I}$ es un subconjunto totalmente ordenado de \aleph_α , con la relación de orden \subset .

Probemos que:

Afirmación (i) (A probar) *Existe en \aleph_α un elemento K minimal de la cadena $\{K_i\}_{i \in I}$.* Es decir, probemos que existe $K \in \aleph_\alpha$, $K \subset K_i$ para todo $i \in I$.

En efecto, el conjunto $K = \bigcap_{i \in I} K_i$ es compacto no vacío, porque cualquier subcolección finita $K_1 \supset K_2 \supset \dots \supset K_l$ de la cadena dada $\{K_i\}_{i \in I}$, tiene intersección K_1

que es un compacto no vacío. Para probar que $K \in \aleph_\alpha$, debemos probar ahora que $m(A_K) \geq \alpha$. Sea $j \in \mathbb{N}^+$ y sea $V_j \supset K$ el abierto formado por todos los puntos de M que distan de K menos que $1/j$. En la prueba del Teorema 7.1.5 demostramos la afirmación (7.2):

$$K_{i_j} \subset V_j \text{ para algún } i_j \in I.$$

Como todo punto $x \in A_{K_{i_j}}$ se cumple $\lim_n \sigma_{n,x}(V_j) = 0$ (cf. Proposición 7.3.6), y este argumento vale para todo $j \in \mathbb{N}^+$, entonces todo punto en $\bigcap_{j \in \mathbb{N}^+} A_{K_{i_j}}$ pertenece a A_K . Recíprocamente, todo punto de A_K está contenido en $A_{K_{i_j}}$ para todo $i \in I$ (porque $K \subset K_i$). En particular esta afirmación se satisface para i_j , para todo $j \in \mathbb{N}^+$. Luego:

$$A_K := \bigcap_{j \in \mathbb{N}^+} A_{K_{i_j}}.$$

Como la colección numerable $A_{K_{i_j}}$ está totalmente ordenada, obtenemos

$$m(A_K) = m\left(\bigcap_{j \in \mathbb{N}^+} A_{K_{i_j}}\right) = \lim_{j \rightarrow +\infty} m(A_{K_{i_j}}) \geq \alpha,$$

terminando de demostrar la afirmación (i).

De la afirmación (i) se deduce que para toda cadena en \aleph_α existe algún elemento $K \in \aleph_\alpha$ minimal de la cadena. Debido al Lema de Zorn, existen en \aleph_α elementos minimales de \aleph_α . Es decir, existe $K \in \aleph_\alpha$ que no contiene subconjuntos propios que pertenezcan a \aleph_α . Esto es, existe K atractor de Ilyashenko α -obs. minimal.

Ahora probemos la unicidad del atractor de Ilyashenko 1-obs. minimal. Si existieran dos atractores de Ilyashenko K_1 y K_2 que fueran 1-obs. minimales, entonces la intersección A de sus cuencas de atracción estadística $A := A_{K_1} \cap A_{K_2}$ cumple

$$m(A) = 1,$$

porque $m(A_{K_1}) = m(A_{K_2}) = 1$. Todo punto $x \in A$ verifica, por la Proposición 7.3.6 que caracteriza las cuencas A_{K_1} y A_{K_2} , la siguiente propiedad:

Para todo $0 < \epsilon < 1/2$ existe $N \geq 1$ tal que

$$\sigma_{n,x}(V_\epsilon(K_1)), \sigma_{n,x}(V_\epsilon(K_2)) > 1 - (\epsilon/2) \quad \forall n \geq N,$$

donde $V_\epsilon(K_i)$ denota el conjunto de puntos de la variedad M que distan del compacto K_i menos que ϵ , y $\sigma_{n,x}$ denota la probabilidad empírica definida en la igualdad (5.9), Definición 5.5.2.

Luego existen más de $N := n \cdot (1 - (\epsilon/2)) \geq (3/4)n$ iterados $f^j(x)$ con tiempos $j \in \{0, 1, \dots, n-1\}$ tales que $f^j(x) \in V_\epsilon(K_1)$. En efecto, si la cantidad de iterados $f^j(x)$ con tiempos $j \in \{0, 1, \dots, n-1\}$ fuera menor o igual que N , entonces tendríamos $\sigma_{n,x}(V_\epsilon(K_1)) \leq N/n = 1 - (\epsilon/2)$. Análogamente, existen más de N iterados $f^h(x)$ con tiempos $h \in \{0, 1, \dots, n-1\}$ tales que $f^h(x) \in V_\epsilon(K_2)$. Como la cantidad de iterados posibles con tiempos en $\{0, 1, \dots, n-1\}$ es a lo sumo n , y como $2N \geq (3/2)n > n$, deben existir algunos iterados comunes $f^j(x) = f^h(x) \in V_\epsilon(K_1) \cap V_\epsilon(K_2)$. Tomemos alguno de esos iterados comunes $f^j(x)$. Tenemos

$$\text{dist}(K_1, K_2) \leq \text{dist}(f^j(x), K_1) + \text{dist}(f^j(x), K_2) < 2\epsilon.$$

Entonces $\text{dist}(K_1, K_2) < 2\epsilon \quad \forall 0 < \epsilon < 1/2$, de donde

$$K := K_1 \cap K_2 \neq \emptyset.$$

Aunque no es inmediato, es estándar chequear que, siendo K_1 y K_2 compactos no vacíos con intersección no vacía, si un punto x cumple

$$\lim_{n \rightarrow +\infty} \sigma_{n,x} V_\epsilon(K_1) = \lim_{n \rightarrow +\infty} \sigma_{n,x} V_\epsilon(K_2) = 0 \quad \forall \epsilon > 0,$$

entonces

$$\lim_{n \rightarrow +\infty} \sigma_{n,x} V_\epsilon(K_1 \cap K_2) = 0 \quad \forall \epsilon > 0.$$

(Chequear esta última afirmación en la parte (a) del Ejercicio 7.4.4.) Luego

$$A := A_{K_1} \cap A_{K_2} \subset A_{K_1 \cap K_2}$$

y como $m(A) = 1$, deducimos que $m(A_{K_1 \cap K_2}) = 1$. Entonces $K_1 \cap K_2$ es un atractor de Ilyashenko 1-obs. Como K_1 y K_2 eran atractores de Ilyashenko 1-obs. minimales, concluimos que $K_1 \cap K_2 = K_1 = K_2$, terminando de demostrar la unicidad del atractor de Ilyashenko 1-obs. minimal, y el Teorema 7.4.3. \square

Ejercicio 7.4.4. Para un conjunto compacto no vacío $K \subset M$, y para $\epsilon > 0$, denotamos $V_\epsilon(K)$ al conjunto de puntos de M que distan de K menos que ϵ . Denotamos $\sigma_{n,x}$ las probabilidades empíricas de la órbita con estado inicial x hasta tiempo n , según la igualdad (5.9), dada por la Definición 5.5.2.

(a) Demostrar que si K_1 y K_2 son compactos no vacíos, y si existe un punto $x \in M$ tal que

$$\lim_{n \rightarrow +\infty} \sigma_{n,x}(V_\epsilon(K_1)) = \lim_{n \rightarrow +\infty} \sigma_{n,x}(V_\epsilon(K_2)) = 0 \quad \forall \epsilon > 0$$

entonces $K_1 \cap K_2 \neq \emptyset$ y

$$\lim_{n \rightarrow +\infty} \sigma_{n,x}(V_\epsilon(K_1 \cap K_2)) = 0 \quad \forall \epsilon > 0$$

(b) Demostrar que si K_1 y K_2 son dos atractores de Ilyashenko tales $m(A_{K_1} \cap A_{K_2}) > 0$, entonces $K_1 \cap K_2$ es no vacío, y es un atractor de Ilyashenko cuya cuenca de atracción estadística es

$$A_{K_1 \cap K_2} = A_{K_1} \cap A_{K_2}.$$

7.5. Ejemplos de atractores estadísticos

Ejemplo 7.5.1. Sea $f_0 \in \text{Diff}^\infty(\mathbb{T}^2)$ el mapa “Arnold’s cat” en el toro bidimensional dado en el ejemplo de la Sección 3.1, como el automorfismo lineal hiperbólico con matriz asociada

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \Big|_{\text{mód.}(1,1)}.$$

Recordando lo expuesto en la Sección 3.1: f_0 es un difeomorfismo de Anosov lineal, tiene a $(0,0)$ como punto fijo y preserva la medida de Lebesgue m (re-escalada para que $m(\mathbb{T}^2) = 1$).

Probemos que el único atractor estadístico (en particular el único atractor estadístico α -obs. minimal para cualquier $0 < \alpha \leq 1$), es $K = \mathbb{T}^2$. En efecto, en el Corolario 6.3.5 probamos que m es ergódica para f . Entonces, m -c.t.p. $x \in M$ está en la cuenca de atracción estadística $B(m)$ de m . (Recordar que una medida de probabilidad invariante μ es ergódica si y solo si $\mu(B(\mu)) = 1$).

Sea K un atractor estadístico. Por definición, su cuenca de atracción estadística A_K tiene medida de Lebesgue positiva. Luego

$$x \in B(m) \quad \text{para } m - \text{c.t.p. } x \in A_K.$$

Tomemos y dejemos fijo $x \in B(m) \cap A_K$. Como $x \in B(m)$ tenemos:

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} = m,$$

en la topología débil* del espacio \mathcal{M} de las probabilidades. Entonces para toda función continua ψ se cumple

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \psi \circ f^j(x) = \int \psi dm.$$

En particular si elegimos la función continua ψ definida por

$$\psi(x) = \text{dist}(x, K) \quad \forall x \in M,$$

se cumple

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K) = \int \psi dm$$

Como $x \in A_K$, por la Definición 7.2.1, el límite de la izquierda es cero. Entonces deducimos que

$$\int \psi dm = 0.$$

Pero $\psi \geq 0$. Entonces su integral da cero si y solo si $\psi = 0$ para m -c.t.p. Una función continua que es cero para Lebesgue casi todo punto, es idénticamente nula. Deducimos que

$$0 = \psi(x) = \text{dist}(x, K) \quad \forall x \in \mathbb{T}^2.$$

Como K es compacto, $\text{dist}(x, K) = 0$ si y solo si $x \in K$. Concluimos que $x \in K$ para todo $x \in \mathbb{T}^2$, es decir $K = \mathbb{T}^2$, probando que el único atractor de Ilyashenko en este ejemplo es todo el toro.

Entonces el conjunto $K_0 := \{(0,0)\}$ no es atractor de Ilyashenko. Pero K_0 es minimal para f desde el punto de vista topológico porque $(0,0)$ es un punto fijo por f (i.e.

K_0 es compacto no vacío, f -invariante, y no contiene subconjuntos propios con esas tres propiedades). Luego, en este ejemplo, hay un minimal (desde el punto de vista topológico) que no es atractor de Ilyashenko. Más aún, ningún minimal desde el punto de vista topológico puede ser atractor de Ilyashenko. En efecto, más arriba probamos que el único atractor estadístico es todo el toro. Pero todo el toro no es minimal desde el punto de vista topológico, pues contiene propiamente a $\{(0, 0)\}$ que es f -invariante.

Observación 7.5.2. La demostración que hicimos en el Ejemplo 7.5.1 de existencia de atractores de Ilyashenko α -obs. minimales que no son minimales desde el punto de vista topológico, se basó en el uso del Corolario 6.3.5, que a su vez se obtiene de la Teoría de Pesin y de la Teoría de medidas SRB para difeomorfismos de clase $C^{1+\alpha}$. Más precisamente, el argumento que utilizamos en el Ejemplo 7.5.1, pasó por deducir la convergencia de la sucesión de probabilidades empíricas

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

para m -c.t.p. x en la cuenca de atracción estadística A_K del atractor de Ilyashenko K . Sin embargo, en general (cuando no sean aplicables los argumentos de la Teoría de Pesin), *no es necesario que la sucesión $\{\sigma_{n,x}\}_{n \geq 1}$ sea convergente*, para un conjunto de medida de Lebesgue positiva contenido en la cuenca A_K de atracción estadística de K (ver por ejemplo [Go-Kl 2007]).

En la próxima sección, tendremos en cuenta la posible no convergencia de la sucesión de probabilidades empíricas para definir las medidas SRB-like (Definición 7.6.4) Finalmente, caracterizaremos a los atractores de Ilyashenko 1-obs. minimales, estudiando las medidas SRB-like. Otras relaciones entre la eventual convergencia o no convergencia de la sucesión de probabilidades empíricas, y los atractores estadísticos o de Ilyashenko se encuentran por ejemplo en [Ka-As 2011].

Ejemplo 7.5.3. Hu-Young [Hu-You 1995]

Sea $f_0 \in \text{Diff}^\infty(\mathbb{T}^2)$ el Anosov lineal en el toro bidimensional dado en el ejemplo de la Sección 3.1, con matriz asociada

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \Big|_{\text{mód.}(1,1)}.$$

En [Hu-You 1995] se construye una isotopía $\{f_t\}_{t \in [0,1]}$ de difeomorfismos $f_t \in \text{Diff}^2(\mathbb{T}^2)$ que transforma continuamente en el espacio $\text{Diff}^2(\mathbb{T}^2)$ el difeomorfismo de Anosov lineal f_0 en f_1 de modo que:

- f_t es un difeomorfismo de Anosov para todo $0 \leq t < 1$.
- $f_t(0, 0) = (0, 0)$ para todo $0 \leq t \leq 1$.
- f_t es conjugado a f_0 para todo $0 \leq t \leq 1$, es decir existe un homeomorfismo $h_t : \mathbb{T}^2 \mapsto \mathbb{T}^2$ tal que

$$f_t \circ h_t = h_t \circ f_0.$$

- Para $t = 1$ la derivada $df_1(0, 0)$ en el punto fijo $(0, 0)$ tiene matriz asociada diagonalizable, con un valor propio igual a 1, y el otro positivo, pero menor estrictamente que

1. Esto implica que el punto fijo $(0, 0)$ no es hiperbólico para f_1 (para el parámetro $t = 1$). Pierde hiperbolicidad en la dirección que era expansora para valores del parámetro $t < 1$, pero no la pierde en la dirección contractiva.

Además, en la construcción de [Hu-You 1995] se imponen otras condiciones a f_1 que permiten deducir la llamada quasi-hiperbolicidad en $\mathbb{T}^2 \setminus \{(0, 0)\}$.

En este ejemplo construido en [Hu-You 1995], los autores muestran que la medida $\delta_{(0,0)}$ concentrada en el origen (que es f_1 -invariante porque el $(0, 0)$ es punto fijo) es SRB o física para f_1 (de acuerdo a nuestra Definición 5.5.6, pero no de acuerdo con lo que los autores en [Hu-You 1995] llaman medida SRB). Además muestran que la cuenca de atracción estadística $B(\delta_{(0,0)})$ cubre Lebesgue c.t.p. del toro \mathbb{T}^2 . Entonces, $\delta_{(0,0)}$ es la única medida SRB o física, ya que ninguna otra medida puede tener cuenca de atracción estadística con medida de Lebesgue positiva (recordar que las cuencas de atracción estadística de medidas diferentes, por la Definición 5.5.1, son disjuntas). Aplicando la Proposición 7.3.3, deducimos que $\{(0, 0)\}$ es el único atractor de Ilyashenko de f_1 . Por lo tanto $\{(0, 0)\}$ es el único atractor de Ilyashenko α -obs. minimal para cualquier $0 < \alpha \leq 1$.

Como f_1 es conjugado a f_0 y f_0 es transitivo, entonces f_1 es transitivo. Aplicando lo probado en el Ejercicio 5.1.5, el único atractor topológico para f_1 es \mathbb{T}^2 . Luego $\{(0, 0)\}$ es un atractor estadístico o de Ilyashenko, que *no es atractor topológico*.

Por la Proposición 7.3.6, la frecuencia de visita a un entorno arbitrariamente pequeño del origen, para Lebesgue-casi toda órbita, tiende a 1, cuando el número n de iterados tiende a infinito. Sin embargo, el origen no es un atractor topológico. Es decir, no es un pozo. Por el contrario, la dinámica en un entorno del origen es localmente conjugado a la de un punto fijo hiperbólico tipo silla, pues el origen es una silla hiperbólica para $t = 0$ y f_1 es conjugado con f_0 . Entonces a pesar de que la frecuencia de estadía en un entorno del origen, para Lebesgue casi toda órbita, tiende a 1, todo punto que no esté en la variedad estable local del origen, termina saliendo de ese entorno, para hacer excursiones breves alejado de él.

Concluimos: Si el experimentador de la dinámica tiene como objetivo observar la estadística de Lebesgue casi toda órbita (es decir los promedios temporales de las distancias al atractor), entonces observará que el sistema dinámico por iterados de f_1 se comporta con un punto fijo atractor, que atrae Lebesgue casi toda órbita, como si fuera un pozo.

En cambio si el experimentador de la dinámica tiene como objetivo observar la topología dinámica de Lebesgue casi toda órbita (es decir los conjuntos ω -límite a donde estas órbitas tienden al iterar hacia el futuro), entonces observará que el sistema dinámico por iterados de f_1 se comporta como el automorfismo lineal hiperbólico en el toro, en que todo el toro es el único atractor transitivo de Lebesgue casi toda órbita.

En el primer caso, el experimentador estadístico, no calificará este sistema f_1 como caótico, pues es altamente previsible, desde el punto de vista estadístico (i.e. de los promedios temporales) para Lebesgue-casi toda órbita. En el segundo caso, el experimentador topológico, lo calificará como caótico, pues le resultará imposible predecir en qué abierto del espacio estará el iterado n -ésimo para Lebesgue-casi toda órbita.

7.6. Medidas SRB-like o pseudo-físicas

En esta sección, como en las dos anteriores, M es una variedad compacta y riemanniana, de dimensión finita, y $f : M \mapsto M$ es continua (no necesariamente invertible).

Definición 7.6.1. El conjunto $p\omega$ límite de probabilidades

Sea $x \in M$, sea \mathcal{M} el espacio de medidas de probabilidad de Borel en M con la topología débil estrella, y sea $\{\sigma_{n,x}\}_{n \geq 1} \subset \mathcal{M}$ la sucesión de probabilidades empíricas de la órbita futura de x hasta tiempo n , definida en (5.9) y Definición 5.5.2. Como \mathcal{M} es secuencialmente compacto, existen subsucesiones convergentes de $\{\sigma_{n,x}\}_{n \geq 1}$. Llamamos *p-omega-límite* de la órbita de x , u *omega-límite en el espacio de probabilidades* de la órbita de x al conjunto $p\omega(x)$ formado por los límites de todas las subsucesiones convergentes de $\{\sigma_{n,x}\}_{n \geq 1}$. Es decir

$$p\omega(x) := \{\mu \in \mathcal{M} : \exists n_i \rightarrow +\infty \text{ tal que } \lim_{i \rightarrow +\infty} \sigma_{n_i, x} = \mu\}, \quad (7.7)$$

donde el límite a la derecha es en la topología débil* del espacio \mathcal{M} de probabilidades. Es estándar chequear que, para todo $x \in M$, $p\omega(x) \subset \mathcal{M}$ es un conjunto no vacío y débil*-cerrado (y por lo tanto débil*-compacto, pues \mathcal{M} es débil*-compacto)

Ejercicio 7.6.2. Probar las dos últimas afirmaciones de la Definición 7.6.1.

Recordemos la Definición 5.5.1 de cuenca de atracción estadística $B(\mu)$ de una medida de probabilidad μ , y la Definición 5.5.6 de medida SRB o física. Ahora generalizaremos esas dos definiciones, agregando una ϵ - aproximación. Para ello elegimos y dejamos fija una métrica dist^* en el espacio \mathcal{M} de probabilidades, que induzca la topología débil* (cf. Teorema 1.2.6).

Definición 7.6.3. Cuenca de atracción estadística ϵ -débil

Dada una medida de probabilidad μ y dado $\epsilon > 0$, llamamos *cuenca de atracción estadística ϵ -débil* al conjunto $A_\epsilon(\mu)$ definido por:

$$A_\epsilon(\mu) := \{x \in M : \text{dist}(p\omega(x), \mu) < \epsilon\}. \quad (7.8)$$

Comparemos esta Definición 7.6.3 con la Definición 5.5.1 de la cuenca de atracción estadística (fuerte) B_μ de una medida μ . En efecto, combinando las igualdades (7.7) y (5.8), obtenemos:

$$B(\mu) := \{x \in M : p\omega(x) = \{\mu\}\}.$$

Luego

$$B(\mu) \subset A_\epsilon(\mu) \quad \forall \epsilon \geq 0. \quad (7.9)$$

En el segundo término de la Igualdad (7.8), observamos que la distancia entre el conjunto compacto $p\omega(x)$ y el punto $\mu \in \mathcal{M}$ es menor que ϵ . Pero esto no implica que todo el conjunto $p\omega(x)$ (cuando contiene más de un punto) deba estar contenido en la bola de centro μ y radio ϵ . Por lo tanto, aunque $\bigcap_{\epsilon > 0} A_\epsilon(\mu)$ pueda ser no vacío, este conjunto contiene a, *pero no necesariamente coincide con* $B(\mu)$, quien aún, puede ser vacío.

Definición 7.6.4. Medida SRB-like o pseudo-física

Llamaremos a una medida de probabilidad μ *SRB-like* o *física* si **para todo** $\epsilon > 0$ su cuenca de atracción estadística ϵ -débil $A_\epsilon(\mu)$ tiene medida de Lebesgue positiva. En breve:

$$m(A_\epsilon(\mu)) > 0 \quad \forall \epsilon > 0,$$

donde m denota la medida de Lebesgue en la variedad M .

Comparando la Definición 7.6.4 con la Definición 5.5.6 de medida SRB, observamos que debido a la inclusión (7.9):

Toda medida SRB es SRB-like.

Sin embargo el recíproco es falso. En efecto, toda f continua tiene medidas SRB-like (cf. Teorema 7.6.6 que demostraremos más adelante en esta sección). Sin embargo existen ejemplos para los que no existen medidas SRB (cf. Ejemplos 5.4.12 y 7.1.9, y en el caso hiperbólico, el mapa C^1 en un disco, atribuido a Bowen y estudiado en [Go-Kl 2007]).

Ejercicio 7.6.5.

- (a) Probar que toda medida SRB-like es invariante por f .
- (b) Probar que la definición de medida SRB-like no depende de la métrica elegida en el espacio \mathcal{M} de probabilidades con la topología débil*.

Notación: Denotamos con \mathcal{O}_f al conjunto de todas las medidas SRB-like para f . Esta notación proviene de [Cat-Enr 2011], que introduce la definición de medidas SRB-like, llamándolas también *medidas observables*.

Observaciones: \mathcal{O}_f está contenido en el conjunto \mathcal{M}_f de medidas de probabilidad f -invariantes, pero usualmente difiere mucho de \mathcal{M}_f , como veremos en el Ejemplo 7.8.1 C^1 genérico. Sin embargo, en el caso C^0 , y para los que llamamos *endomorfismos expansores de Misiurewicz en el círculo S^1* (para los que no existen medidas SRB), el conjunto de medidas SRB-like coincide con el conjunto \mathcal{M}_f de todas las medidas invariantes (ver [Mis 2005], o también el Corolario 7.6.11 y el Ejemplo 7.8.2 más adelante en esta sección).

Algunas propiedades que distinguen a las medidas SRB-like son:

- Existen medidas SRB-like, sin necesidad de agregar hipótesis adicionales a la continuidad de f (Teorema 7.6.6).
- El conjunto de medidas SRB-like describen en forma óptima (con un mínimo posible de medidas invariantes) la estadística en el futuro de Lebesgue casi toda órbita (Teorema 7.6.7).
- El mínimo soporte compacto común de medidas SRB-like (cf. Definiciones 7.3.2 y 7.7.1) caracteriza a los atractores estadísticos o de Ilyashenko α -obs. minimales. (Teorema 7.7.2).
- Bajo hipótesis adicionales de C^1 hiperbolicidad, las medidas SRB-like, satisfacen la Fórmula (6.7) de Pesin de la Entropía (cf. Ejemplos 7.8.1 y 7.8.3), aunque en el contexto general de regularidad C^1 no necesariamente tienen propiedades de continuidad absoluta respecto a Lebesgue.

Teorema 7.6.6. Existencia de medidas SRB-like [Cat-Enr 2011]

Sea $f : M \mapsto M$ continua. Entonces:

- (a) Existen medidas de probabilidad SRB-like para f .
 (b) El conjunto \mathcal{O}_f de las medidas SRB-like es débil*-compacto en el espacio de las probabilidades invariantes por f .

Extraemos la prueba de [Cat-Enr 2011]:

Demostración. (a) Supongamos por absurdo que \mathcal{O}_f es vacío. Entonces toda probabilidad μ es no SRB-like. Es decir, μ no satisface la Definición 7.6.4. Recordamos la Definición 7.6.1 del conjunto $p\omega(x)$ (p -omega-límite de la órbita por cada punto $x \in M$). Luego, toda $\mu \in \mathcal{M}$ está contenida en un entorno abierto $B(\mu) \subset \mathcal{M}$ (con la topología débil* del espacio de probabilidades \mathcal{M}) tal que

$$m\left(\{x \in M : p\omega(x) \cap B(\mu) \neq \emptyset\}\right) = 0$$

Como \mathcal{M} es débil* compacto, existe un subcubrimiento finito de M

$$\{B_1, B_2, \dots, B_k\}$$

tal que $B_i = B(\mu_i)$ para alguna medida de probabilidad μ_i . Luego:

$$\begin{aligned} m\left(\{x \in M : p\omega(x) \cap \left(\bigcup_{i=1}^k B_i\right) \neq \emptyset\}\right) &= \\ &= m\left(\bigcup_{i=1}^k \{x \in M : p\omega(x) \cap B_i \neq \emptyset\}\right) \leq \\ &\leq \sum_{i=1}^k m\left(\{x \in M : p\omega(x) \cap B_i \neq \emptyset\}\right) = 0. \end{aligned}$$

Como $\bigcup_{i=1}^k B_i = \mathcal{M}$, deducimos que

$$m\left(\{x \in M : p\omega(x) \cap \mathcal{M} \neq \emptyset\}\right) = 0.$$

Como $p\omega(x) \subset \mathcal{M}$ para todo $x \in M$, deducimos que para Lebesgue c.t.p. $x \in M$, $p\omega(x) = \emptyset$, lo cual es una contradicción porque toda sucesión de probabilidades tiene alguna subsucesión convergente.

(b) Para probar que \mathcal{O}_f es débil* compacto, basta probar que es débil* cerrado (pues $\mathcal{O}_f \subset \mathcal{M}$ y \mathcal{M} es un espacio metrizable débil* compacto). Sea $\mu_n \in \mathcal{O}_f$ convergente a μ . Debemos probar que $\mu \in \mathcal{O}_f$.

Sea $\epsilon > 0$ arbitrario. Sea $B_\epsilon(\mu)$ la bola de centro μ y radio ϵ . Sea n tal que $\mu_n \in B_\epsilon(\mu)$ y sea $\epsilon_n > 0$ tal que la bola $B_{\epsilon_n}(\mu_n)$ de centro μ_n y radio ϵ_n satisface

$$B_{\epsilon_n}(\mu_n) \subset B_\epsilon(\mu).$$

Como $\mu_n \in \mathcal{O}_f$, por la Definición 7.6.4 de medida SRB-like, se cumple

$$m\left(\{x \in M : p\omega(x) \cap B_{\epsilon_n}(\mu_n) \neq \emptyset\}\right) > 0.$$

Como

$$\{x \in M : p\omega(x) \cap B_{\epsilon_n}(\mu_n) \neq \emptyset\} \subset \{x \in M : p\omega(x) \cap B_\epsilon(\mu) \neq \emptyset\},$$

deducimos que

$$m\left(\{x \in M : p\omega(x) \cap B_\epsilon(\mu) \neq \emptyset\}\right) > 0.$$

Siendo $\epsilon > 0$, esto prueba que μ es SRB-like, como queríamos demostrar. \square

Para enunciar el siguiente teorema, recordamos la Definición 7.6.1 del conjunto $p\omega(x)$ (p -omega-límite de la órbita por el punto x). De la igualdad (7.7), observamos que $p\omega(x)$ es el mínimo conjunto de probabilidades que describen completamente la estadística (es decir los promedios temporales asintóticos) de la órbita por el punto x .

Teorema 7.6.7. . Optimalidad estadística del conjunto de medidas SRB-like [Cat-Enr 2011]

Para toda $f : M \mapsto M$ continua el conjunto \mathcal{O}_f de las medidas SRB-like para f es el mínimo conjunto débil*-compacto \mathcal{K} del espacio de probabilidades tal que

$$p\omega(x) \subset \mathcal{K} \text{ para Lebesgue-c.t.p. } x \in M.$$

Demostración. Por la parte (b) del Teorema 7.6.6, el conjunto \mathcal{O}_f de todas las medidas SRB-like para f es no vacío y débil*-compacto.

Primero probemos que $p\omega(x) \subset \mathcal{O}_f$ para Lebesgue c.t.p. $x \in M$. Para $\epsilon > 0$ arbitrario, denotamos con $B_\epsilon(\mathcal{O}_f)$ al conjunto abierto de todas las probabilidades ν que distan del compacto no vacío \mathcal{O}_f menos que ϵ . Consideremos el complemento

$$\mathcal{C} := \mathcal{M} \setminus B_\epsilon(\mathcal{O}_f).$$

El conjunto \mathcal{C} es compacto porque es el complemento del abierto $B_\epsilon(\mathcal{O}_f)$ en el espacio compacto \mathcal{M} . Por construcción $\mathcal{C} \cap \mathcal{O}_f = \emptyset$. Luego, toda $\nu \in \mathcal{C}$ es no SRB-like. Entonces existe un entorno abierto $B(\nu)$ de ν tal que:

$$m\left(\{x \in M : p\omega(x) \cap B(\nu) \neq \emptyset\}\right) = 0.$$

Al igual que al final de la demostración de la parte (a) del Teorema 7.6.6, pero escribiendo \mathcal{C} en el rol de \mathcal{M} , deducimos que existe un cubrimiento finito $\{B_1, B_2, \dots, B_k\}$ de \mathcal{C} tal que

$$m\left(\{x \in M : p\omega(x) \cap \left(\bigcup_{i=1}^k B_i\right) \neq \emptyset\}\right) = 0$$

Como $\bigcup_{i=1}^k B_i \supset \mathcal{C}$, deducimos

$$m\left(\{x \in M : p\omega(x) \cap \mathcal{C} \neq \emptyset\}\right) = 0.$$

Dicho de otra forma, para Lebesgue-casi todo punto $x \in M$, el conjunto $p\omega(x)$ está contenido en $\mathcal{M} \setminus \mathcal{C} = B_\epsilon(\mathcal{O}_f)$. Luego, tomando $\epsilon = 1/h$ para $h \in \mathbb{N}^+$, hemos probado que

$$\forall h \in \mathbb{N}^+ : p\omega(x) \subset B_{1/h}(\mathcal{O}_f) \quad m - \text{c.t.p. } x \in M.$$

Como la unión numerable de conjuntos con medida de Lebesgue nula, tiene medida de Lebesgue nula, la intersección numerable de conjuntos con medida de Lebesgue total, tiene medida de Lebesgue total. Concluimos que:

$$p\omega(x) \subset \bigcap_{h=1}^{+\infty} B_{1/h}(\mathcal{O}_f) = \mathcal{O}_f \quad m - \text{c.t.p. } x \in M,$$

como queríamos demostrar.

Segundo, probemos que \mathcal{O}_f es minimal en el conjunto de compactos no vacíos $\mathcal{K} \subset \mathcal{M}$ que tienen la propiedad $p\omega(x) \subset \mathcal{K}$ para Lebesgue-casi todo punto $x \in M$. Tomemos cualquier compacto no vacío $\mathcal{K} \subset \mathcal{O}_f$, tal que $\mathcal{K} \neq \mathcal{O}_f$. Probemos que tal \mathcal{K} no tiene la propiedad mencionada. Es decir, probemos que $m\left(p\omega(x) \cap (\mathcal{M} \setminus \mathcal{K})\right) > 0$.

En efecto, como $\mathcal{K} \subsetneq \mathcal{O}_f$, existe una medida $\mu \in \mathcal{O}_f \setminus \mathcal{K}$. Sea $\epsilon > 0$ tal que la bola $B_\epsilon(\mu)$ es disjunta con el compacto \mathcal{K} . Como μ es SRB-like, por las Definiciones 7.6.4 y 7.6.3, se cumple:

$$m\left(\{x \in M : p\omega(x) \cap B_\epsilon(\mu) \neq \emptyset\}\right) > 0. \quad (7.10)$$

Siendo $B_\epsilon(\mu) \cap \mathcal{K} = \emptyset$, deducimos que $B_\epsilon(\mu) \subset (\mathcal{M} \setminus \mathcal{K})$. Entonces, sustituyendo en (7.10) concluimos:

$$m\left(\{x \in M : p\omega(x) \cap (\mathcal{M} \setminus \mathcal{K}) \neq \emptyset\}\right) > 0,$$

como queríamos demostrar. \square

Corolario 7.6.8. *Sea $f : M \mapsto M$ continua.*

(a) *Si la medida SRB-like μ es única, entonces μ es SRB y su cuenca de atracción estadística (fuerte) $B(\mu)$ cubre M Lebesgue c.t.p.*

(b) *Recíprocamente, si existe una medida SRB μ cuya cuenca de atracción estadística (fuerte) $B(\mu)$ cubre M Lebesgue c.t.p., entonces μ es la única medida SRB-like.*

(c) *Si el conjunto de medidas SRB-like es finito, entonces todas las medidas SRB-like son SRB y la unión de las cuencas de atracción estadística de las medidas SRB cubre M Lebesgue c.t.p.*

(d) *Recíprocamente, si existe una cantidad finita de medidas SRB tales que la unión de sus cuencas de atracción estadística cubre M Lebesgue c.t.p., entonces estas son las únicas medidas SRB-like, y por lo tanto el conjunto de medidas SRB-like es finito.*

Ejercicio 7.6.9. Probar el Corolario 7.6.8. Sugerencia: Basta probar (c) y (d), pues estas implican (a) y (b). Probar primero que una medida SRB-like es aislada en el conjunto de las medidas SRB-like si y solo si es SRB. Combinar las Definiciones 5.5.6 y 7.6.4 de medidas SRB y SRB-like respectivamente, junto con las Definiciones 5.5.1 y 7.6.3 de las cuencas de atracción estadística fuerte y ϵ -débil, respectivamente.

Conjetura 7.6.10. Palis [Pa 1999] Para $r \geq 1$ suficientemente grande C^r -genéricamente en $\text{Diff}^r(M)$ existe una cantidad finita de medidas SRB tales que la unión de sus cuencas de atracción estadística cubre M Lebesgue c.t.p.

Del Corolario 7.6.8 deducimos el siguiente enunciado equivalente de la Conjetura de Palis:

C^r -genéricamente para $f \in \text{Diff}^r(M)$ el conjunto débil* no vacío \mathcal{O}_f de medidas de probabilidad SRB-like para f , carece de puntos de acumulación.

Corolario 7.6.11. (del Teorema 7.6.7)

Sea $f : M \mapsto M$ continua, no únicamente ergódica. Si para Lebesgue c.t.p. $x \in M$ el conjunto de probabilidades $p\omega(x)$ (cf. Definición 7.6.1) coincide con el conjunto de todas las medidas invariantes, entonces no existen medidas SRB y el conjunto de las medidas SRB-like coincide con el conjunto de todas las medidas invariantes.

Nota: Existen transformaciones f que cumplen las hipótesis del Corolario 7.6.11. En efecto, el mapa C^0 expansor en el círculo construido por Misiurevicz en [Mis 2005], y los mapas C^0 -expansores genéricos encontrados recientemente por Abdenur y Andersson en [Ab-An 2012], satisfacen las hipótesis de este Corolario. Ver también el Ejemplo 7.8.2 más adelante en esta sección.

Demostración. del Corolario 7.6.11:

Por hipótesis, el mínimo conjunto compacto \mathcal{K} de medidas de probabilidad tal que $p\omega(x) \subset \mathcal{K}$ es el conjunto \mathcal{M}_f de todas las medidas f -invariantes. Por el Teorema 7.6.7, \mathcal{K} es el conjunto \mathcal{O}_f de las medidas SRB-like. Luego $\mathcal{O}_f = \mathcal{M}_f$ como queríamos demostrar. \square

7.7. Relación entre atractor estadístico y medidas SRB-like

Definición 7.7.1. Sea \mathcal{K} un conjunto no vacío de medidas de probabilidad. Se llama *soporte compacto de \mathcal{K}* al mínimo compacto $K \subset M$ tal que

$$\mu(K) = 1 \quad \forall \mu \in \mathcal{K}.$$

El mínimo soporte compacto existe como resultado de aplicar el lema de Zorn a la familia de compactos no vacíos con μ -medida igual a 1 para toda $\mu \in \mathcal{K}$, y de usar también la propiedad de que es no vacía la intersección de una familia de compactos tal que toda subfamilia finita tiene intersección no vacío. El soporte compacto de \mathcal{K} es único, debido a su propiedad de minimalidad y a que la intersección de dos compactos con μ -medida igual a 1 es un compacto con μ -medida igual a 1.

Teorema 7.7.2.**Medidas SRB-like y atractores de Ilyashenko**

Para toda $f: M \mapsto M$ continua, el atractor 1-obs. minimal de Ilyashenko K (cf. Theorem 7.4.3) es el soporte compacto del conjunto \mathcal{O}_f de medidas SRB-like para f .

Nota: El Teorema 7.7.2 puede generalizarse, adaptando el enunciado adecuadamente para caracterizar todo atractor de Ilyashenko α -obs. minimal (para cualquier $0 < \alpha \leq 1$) como el soporte compacto de un subconjunto adecuado de medidas SRB-like para f (ver [Cat 2012]).

Demostración. del Teorema 7.7.2: Sea K_1 el atractor de Ilyashenko 1-obs. minimal. Por el Teorema 7.4.3, el compacto $K_1 \neq \emptyset$ existe y es único. Sea K_2 el soporte compacto del conjunto \mathcal{O}_f de medidas SRB-like para f , según la Definición 7.7.1. De acuerdo a lo observado al final de dicha definición, el compacto $K_2 \neq \emptyset$ existe y es único.

Demostración de $K_1 \subset K_2$: Como K_1 es 1-obs. minimal, por la Definición 7.4.1 basta probar que

$$\text{A probar: } \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K_2) = 0 \text{ para } m\text{-c.t.p. } x \in M. \quad (7.11)$$

Para todo $x \in M$ consideremos la sucesión $\{\sigma_{n,x}\}_{n \geq 1}$ de probabilidades empíricas de la órbita por x , según la Definición 5.5.2. Aplicando el Teorema 7.6.7, tenemos para m -c.t.p. $x \in M$ la siguiente propiedad:

$$\lim_{i \rightarrow +\infty} \sigma_{n_i, x} \in \mathcal{O}_f \quad (7.12)$$

para toda subsucesión $n_i \rightarrow +\infty$ tal que existe ese límite (donde dicho límite se toma en la topología débil* del espacio de probabilidades).

Consideremos la función continua ψ definida por

$$\psi(x) := \text{dist}(x, K_2) \quad \forall x \in M. \quad (7.13)$$

Integrando ψ en la inclusión (7.12), y teniendo en cuenta la definición de la topología débil*, deducimos que para m -c.t.p. $x \in M$, y para toda subsucesión convergente $\{\sigma_{n_i, x}\}_{i \in \mathbb{N}}$ de probabilidades empíricas, existe $\mu \in \mathcal{O}_f$ tal que

$$\int \psi d\mu = \lim_{i \rightarrow +\infty} \int \psi d(\sigma_{n_i, x}) = \lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \text{dist}(f^j(x), K_2). \quad (7.14)$$

Pero $\mu(K_2) = 1$, porque por hipótesis, K_2 es soporte compacto común de todas las medidas de probabilidad en \mathcal{O}_f . Luego $\psi(y) = \text{dist}(y, K_2) = 0$ para μ -c.t.p. $y \in M$, de donde $\int \psi d\mu = 0$. Sustituyendo en la igualdad (7.14) obtenemos

$$\lim_{i \rightarrow +\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} \text{dist}(f^j(x), K_2) = 0 \quad (7.15)$$

para m -c.t.p. $x \in M$, para toda subsucesión $n_i \rightarrow +\infty$ tal que $\{\sigma_{n_i, x}\}_{i \in \mathbb{N}}$ sea convergente.

Fijado un tal punto x , sea dada una sucesión cualquiera $n_i \rightarrow +\infty$ tal que la subsucesión $\{d_{n_i}\}_{i \in \mathbb{N}}$ de promedios de distancias a K_2 , dada por

$$d_{n_i} := \frac{1}{n_i} \sum_{j=0}^{n_i-1} \text{dist}(f^j(x), K_2),$$

es convergente. Por la compacidad del espacio de probabilidades, siempre existe una subsucesión de esta subsucesión $\{d_{n_i}\}_i$ (para índices $i = i_h$) tal que $\{\sigma_{n_{i_h}, x}\}_h$ es también convergente. Entonces, la igualdad (7.15) aplicada a n_{i_h} en lugar de n_i , implica (para la subsucesión de índices $\{n_i\}_i$), que el límite del promedio d_{n_i} de las distancias a K_2 es cero. Concluimos que toda subsucesión convergente de $\{d_n\}_{n \geq 1}$ converge a cero. Luego, la afirmación (7.11) está probada.

Demostración de $K_2 \subset K_1$: Por hipótesis K_2 es el mínimo compacto tal que $\mu(K_2) = 1$ para toda $\mu \in \mathcal{O}_f$. Luego, basta demostrar que

$$\text{A probar: } \mu(K_1) = 1 \quad \forall \mu \in \mathcal{O}_f. \quad (7.16)$$

Sean dados $\mu \in \mathcal{O}_f$ y $\epsilon > 0$ arbitrario. Sea φ la función continua no negativa definida por

$$\varphi(y) := \text{dist}(y, K_1). \quad (7.17)$$

Por la definición de la topología débil*, con la métrica dist^* que se haya elegido en el espacio \mathcal{M} de las probabilidades, existe $\delta > 0$ tal que

$$\text{dist}^*(\nu, \mu) < \delta \Rightarrow \left| \int \varphi d\mu - \int \varphi d\nu \right| < \epsilon. \quad (7.18)$$

Por la Definición 7.6.4 de medida SRB-like, la medida de Lebesgue m de la cuenca $A_\delta(\mu)$ de atracción estadística δ -débil de μ , es positiva. Como K_1 es un atractor de Ilyashenko 1-obs., su cuenca de atracción estadística $A(K_1)$ tiene medida de Lebesgue total. Luego deducimos que

$$m(A(K_1) \cap A_\delta(\mu)) > 0.$$

Tomemos un punto $x \in A(K_1) \cap A_\delta(\mu)$. Por la igualdad (7.3) que define $A(K_1)$ tenemos:

$$0 = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \text{dist}(f^j(x), K_1) = \lim_{n \rightarrow +\infty} \int \varphi d(\sigma_{n, x}), \quad (7.19)$$

donde $\sigma_{n, x}$ es la probabilidad empírica dada definida por (5.9). Por la Definición 7.6.3 de $A_\delta(\mu)$, existe una subsucesión $n_i \rightarrow +\infty$ tal que $\{\sigma_{n_i}\}_{i \in \mathbb{N}}$ converge a una medida $\nu \in p\omega(x)$ tal que $\text{dist}^*(\mu, \nu) < \delta$. Luego, combinando la afirmación (7.18) con la igualdad (7.19), deducimos:

$$0 = \lim_{i \rightarrow +\infty} \int \varphi d\sigma_{n_i, x} = \int \varphi d\nu, \quad \text{dist}^*(\nu, \mu) < \delta, \Rightarrow$$

$$\int \varphi d\nu = 0, \quad \left| \int \varphi d\mu - \int \varphi d\nu \right| < \epsilon \quad \Rightarrow$$

$$\left| \int \varphi d\mu \right| < \epsilon.$$

Como esta desigualdad vale para todo $\epsilon > 0$, deducimos que $0 = \int \varphi d\mu$. Siendo $\varphi \geq 0$, concluimos que $\varphi(y) = \text{dist}(y, K_1) = 0$ para μ -c.t.p. $y \in M$. Como K_1 es compacto, la distancia de un punto y a K_1 es cero si y solo si $y \in K_1$. Hemos probado que $y \in K_1$ para μ -c.t.p. $y \in M$, o dicho de otra forma, $\mu(K_1) = 1$ como queríamos demostrar. \square

7.8. Ejemplos de medidas SRB-like

Ejemplo 7.8.1. Endomorfismos C^1 -expansores en S^1

Sea $f : S^1 \mapsto S^1$ en el círculo S^1 un endomorfismo de clase C^1 expansor, esto es, existe una constante $\sigma > 1$ tal que $|f'(z)| \geq \sigma > 1 \quad \forall z \in S^1$. Por ejemplo en el círculo

$$S^1 := \{z \in \mathbb{C} : |z| = 1\}$$

la transformación $f(z) = z^2$ es expansora. Un resultado clásico de dinámica topológica (ver por ejemplo [Kat-Has 1995, Theorem 2.4.6]), establece que todo endomorfismo expansor en el círculo es sobreyectivo, el número de preimágenes de cualquier punto $x \in S^1$ es constante igual a $k \geq 2$ (llamado *grado* de f) y f es topológicamente conjugado a g_k definido por $g_k(z) = z^k$ para todo $z \in S^1 = \{|z| = 1\}$.

El endomorfismo expansor, puede mirarse como un endomorfismo (no invertible) uniformemente hiperbólico (cf. Definition 3.3.1), en que el subespacio inestable U_x es todo el espacio tangente $T_x S^1$, y el subespacio estable $E_x = \{0\}$. La variedad inestable $W^u(x_0)$ por un punto cualquiera $x_0 \in S^1$ se define por

$$W^u(x_0) := \left\{ y_0 \in S^1 : \exists x_{-n}, y_{-n} \in S^1 \text{ tales que} \right.$$

$$f(x_{-n}) = x_{-n+1}, \quad f(y_{-n}) = y_{-n+1} \quad \forall n \geq 0,$$

$$\left. \lim_{-n \rightarrow -\infty} \text{dist}(y_{-n}, x_{-n}) = 0 \right\}.$$

La variedad inestable de cualquier punto coincide con todo S^1 :

$$W^u(x_0) = S^1 \quad \forall x_0 \in S^1.$$

Por esta razón, para los endomorfismos expansores en el círculo, la medida condicional inestable de μ (cf. Definición 6.2.2) es la misma μ . Entonces, definimos:

Definición (medida de Gibbs): Si f es un endomorfismo C^1 expansor en el círculo S^1 , decimos que una medida f -invariante μ es de Gibbs si $\mu \ll m$, donde m es la medida de Lebesgue en S^1 .

Si f es de clase $C^{1+\alpha}$, se tiene el siguiente resultado clásico, que da una versión del Teorema 6.3.4 aplicable a los endomorfismos expansores del círculo (en vez de a los difeomorfismos de Anosov en variedades de dimensión mayor que 1).

Teorema (Ruelle) [Rue 1989]:

Sea $f : S^1 \mapsto S^1$ de clase $C^{1+\alpha}$ expansor en el círculo S^1 . Entonces:

- (a) Existe una única medida de Gibbs μ (i.e. μ absolutamente continua respecto a la medida de Lebesgue) y esta medida μ es ergódica.
- (b) Toda medida SRB es de Gibbs y recíprocamente. (Por lo tanto existe una única medida SRB y es ergódica).
- (c) La medida de Gibbs μ es equivalente a la medida de Lebesgue m (i.e. $\mu \ll m$ y $m \ll \mu$).
- (d) La cuenca de atracción estadística de μ cubre Lebesgue c.t.p. $z \in S^1$.

La prueba de este Teorema de Ruelle para mapas $C^{1+\alpha}$ expansores en S^1 , puede encontrarse por ejemplo en [Kat-Has 1995, Theorem 5.1.16].

Nota: Un resultado más general, que establece la existencia de medidas de Gibbs para mapas C^1 a trozos en el círculo, que sean expansores en cada trozo, y que cumplan hipótesis de variación acotada, fue demostrado recientemente por Liverani en [Liv 2013].

Como consecuencia de la parte (d) del Teorema de Ruelle, en el caso $C^{1+\alpha}$ expansor, no existen otras medidas SRB-like que no sean la medida SRB. En efecto, otra medida $\nu \neq \mu$ tendría una cuenca de atracción estadística ϵ -débil $A_\epsilon(\nu)$ que debe ser disjunta con la cuenca de atracción estadística (fuerte) $B(\mu)$ de μ (porque $\nu \neq \mu$). Pero como $m(M \setminus B(\mu)) = 0$, entonces $m(A_\epsilon(\nu)) = 0$ y ν no puede ser SRB-like.

Ahora veremos que la relación entre medidas de Gibbs y la Fórmula (6.7) de Pesin para la entropía (que vale para difeomorfismos de clase $C^{1+\alpha}$ según vimos en el Teorema 6.4.1 de la sección anterior), también se generaliza para endomorfismos expansores de clase $C^{1+\alpha}$ en el círculo.

Teorema

Fórmula de Pesin para la Entropía. Endomorfismos expansores Sea $f : S^1 \mapsto S^1$ expansor de clase C^2 . Entonces, la única medida SRB μ para f (que por el teorema de Ruelle es de Gibbs y equivalente a la medida de Lebesgue), satisface la Fórmula (6.7) de Pesin para la Entropía, y es la única medida que satisface tal fórmula.

Una prueba de este Teorema, fue dada en [Pe 1977] por Pesin. Una versión para endomorfismos de clase C^2 se encuentra en [Qi-Zh 2002] o en [Qi-Xi-Zh 2009]. Más aún, en [Qi-Zh 2002] se prueba la versión del Teorema de Ledrappier-Young [Led-You 1985] para endomorfismos, que establece la equivalencia, para todo endomorfismo de clase C^2 , entre las medidas de Gibbs y las medidas que satisfacen la Fórmula de Pesin para la Entropía (cf. Teorema 6.4.3).

En el caso que el endomorfismo f sea de clase C^1 pero no $C^{1+\alpha}$, las afirmaciones del

Teorema de Ruelle fallan. En efecto, en [Avi-Boc 2007] se prueba que los endomorfismos expansores C^1 -genéricos no poseen medidas invariantes absolutamente continuas respecto de m (no poseen medidas de Gibbs). Más en general, en [Sch-Gor 1989] se prueba (con un ejemplo explícito) que, aunque f sea discontinua, si f es C^1 expansora a trozos, entonces la existencia de medidas invariantes absolutamente continuas respecto de m no es necesaria.

Sin embargo, en [Cam-Qua 2001] Campbell y Quas probaron que C^1 -genéricamente, un endomorfismo expansor en el círculo (que es C^1 pero no $C^{1+\alpha}$), posee una única medida SRB μ y que esta medida es ergódica, satisface la Fórmula de Pesin para la Entropía (6.7) y tiene cuenca de atracción estadística que cubre Lebesgue c.t.p. Pero en vez de ser μ absolutamente continua respecto a la medida de Lebesgue m , se cumple todo lo contrario:

$$\mu \perp m,$$

es decir, μ es mutuamente singular respecto a la medida de Lebesgue m .

Del resultado de Campbell y Quas deducimos que C^1 -genéricamente para los endomorfismos expansores en el círculo, existe una única medida SRB-like que es la medida SRB μ encontrada por Campbell y Quas. En efecto, argumentamos igual que antes, como lo hicimos a partir del Teorema de Ruelle para los expansores de clase $C^{1+\alpha}$. Como la cuenca de atracción estadística $B(\mu)$ de la medida SRB μ cubre Lebesgue c.t.p., entonces ninguna otra medida $\nu \neq \mu$ puede ser SRB-like.

Sin necesidad de asumir hipótesis de C^1 -genericidad, todo endomorfismo expansor de clase C^1 en el círculo tiene medidas SRB-like, todas sus medidas SRB-like satisfacen la Fórmula de Pesin para la Entropía (6.7), y la unión de sus cuencas de atracción estadística ϵ -débil cubre Lebesgue c.t.p. para todo $\epsilon > 0$. Estos resultados fueron probados en [Cat-Enr 2012]. Sin embargo, la unicidad de la medida SRB-like (que es cierta para los expansores de clase $C^{1+\alpha}$ debido al Teorema de Ruelle), es falsa en general para los expansores de clase C^1 que no son $C^{1+\alpha}$ ni son C^1 -genéricos. En efecto, en [Qua 1996] Quas construyó un expansor de clase C^1 que exhibe más de una medida SRB-like.

Ejemplo 7.8.2. Endomorfismos C^0 -expansores en S^1

Un mapa continuo $f : S^1 \mapsto S^1$ en el círculo S^1 se llama *endomorfismo C^0 -expansor* (cf. [Kat-Has 1995, Definition 2.4.1]), si existen constantes $\delta > 0$ y $\sigma > 1$ tales que

$$\text{dist}(f(x), f(y)) \geq \sigma \text{dist}(x, y) \quad \forall x, y \in S^1 \text{ tales que } \text{dist}(x, y) < \delta.$$

En [Mis 2005], Misiurewicz construyó un endomorfismo C^0 -expansor f en S^1 , que satisface las hipótesis del Corolario 7.6.11. Por lo tanto, en ese ejemplo no existen medidas SRB y toda medida f -invariante es SRB-like. Existe entonces una cantidad infinita no numerable de medidas SRB-like. Llamaremos a los endomorfismos que tienen esta propiedad *endomorfismos de Misiurewicz*. En [Ab-An 2012] se prueba que los endomorfismos de Misiurewicz son C^0 -genéricos en el espacio de los endomorfismos C^0 -expansores del círculo S^1 .

Ejemplo 7.8.3. Anosov C^1

Sea f un difeomorfismo de Anosov transitivo en una variedad compacta y Riemanniana M (cf. Definition 3.3.1).

Primero repasemos el caso en que f es además de clase $C^{1+\alpha}$ o en particular si es de clase C^2 :

En el Teorema 6.3.4 de Sinai [Si 1972], vimos que si f es un difeomorfismo de Anosov de clase $C^{1+\alpha}$, entonces existe una única medida de probabilidad SRB μ , es de Gibbs ergódica, y su cuenca de atracción estadística cubre Lebesgue c.t.p. Deducimos que tal medida μ es la única medida SRB-like, argumentando de igual forma que en el ejemplo 7.8.1.

Además, debido al Teorema 6.3.1, la única medida SRB-like de un difeomorfismo de Anosov transitivo de clase $C^{1+\alpha}$, satisface la Fórmula (6.7) de Pesin para la Entropía.

Si el difeomorfismo f de Anosov transitivo es de clase C^2 , entonces su única medida SRB μ , es también la única medida de probabilidad que satisface la Fórmula de Pesin para la Entropía. En efecto, por el Teorema 6.4.3 de Ledrappier-Young [Led-You 1985], toda medida ν que satisfaga la Fórmula de Pesin para la Entropía es de Gibbs. Y por el Teorema 6.3.1 las componentes ergódicas ν_x de toda medida de Gibbs es SRB, y por lo tanto es SRB-like. Luego, como existe una única medida SRB-like μ , ν tiene una única componente ergódica que es μ , y por lo tanto ν es ergódica y coincide con μ .

Ahora pasemos al caso de difeomorfismo de Anosov transitivo f de clase C^1 pero no $C^{1+\alpha}$. En este caso las demostraciones conocidas de los teoremas mencionados en el repaso anterior, no son aplicables, porque utilizan la Teoría de Pesin. Por ejemplo, las pruebas utilizan el Teorema 6.5.2 que establece la continuidad absoluta de la holonomía de la foliación estable. Este resultado sería falso si f no fuera de clase $C^{1+\alpha}$ (ver [Rob-You 1980], o también [Bow 1975a]).

Sin embargo, en los últimos años se han obtenido algunos resultados parciales, aplicables a los difeomorfismos de Anosov de clase C^1 , que generalizan el Teorema 6.3.4: En [Qiu 2011] Qiu y Zhu demostraron que C^1 -genéricamente los difeomorfismos de Anosov transitivos tienen una única medida SRB, esta medida es ergódica, satisface la Fórmula de Pesin para la Entropía, es la única medida que satisface tal fórmula, y su cuenca de atracción estadística cubre Lebesgue c.t.p. de la variedad. Por lo tanto μ es la única medida SRB-like. La prueba de Qiu y Zhu no pasa (a diferencia de la prueba del Teorema 6.3.4) por la construcción de medidas de Gibbs. Más aún (tanto como la autora de este libro conoce) no se sabe si la única medida SRB μ de un difeomorfismo de Anosov transitivo y C^1 -genérico, es medida de Gibbs. Pero se sabe que tal μ no es absolutamente continua respecto a la medida de Lebesgue m en toda la variedad, pues Ávila y Bochi [Avi-Boc 2006] probaron que C^1 genéricamente no existen medidas invariantes absolutamente continuas respecto de m .

Cuando la medida SRB-like μ es única deducimos, como consecuencia del Teorema 7.6.7, que μ es SRB y que su cuenca de atracción estadística $B(\mu)$ cubre Lebesgue c.t.p. En este caso, por ejemplo para los difeomorfismos de Anosov transitivos C^1 -genéricos del Teorema de Qiu y Zhu [Qiu 2011], es válido un teoremas ergódico, probado

recientemente por Kleptsyn y Ryzhov en [Kl-Ry-Mi 2012], que estima la velocidad de convergencia de las probabilidades empíricas $\sigma_{n,x}$ (en la topología débil*) a la medida SRB- μ , para Lebesgue c.t.p. $x \in M$.

En el caso de difeomorfismo de Anosov C^1 no genérico y no $C^{1+\alpha}$, cuando f preserva una medida invariante μ absolutamente continua respecto a la medida de Lebesgue, Sun y Tian probaron en [Su-Ti 2012] que esta medida μ satisface la Fórmula de Pesin para la entropía. Esto implica que, bajo las hipótesis adicional de existencia de medida invariante μ equivalente a la medida de Lebesgue m , toda medida SRB-like satisface tal fórmula. En efecto, por el Teorema 4.1.2 de Descomposición Ergódica, μ -c.t.p. x pertenece a la cuenca de atracción estadística $B(\mu_x)$ de una componente ergódica μ_x de μ . Como μ es equivalente a la medida de Lebesgue, deducimos que Lebesgue c.t.p. x pertenece a $B(\mu_x)$. Luego, por el Teorema 7.6.7, estas componentes ergódicas μ_x son las medidas SRB-like. Como μ satisface la Fórmula de Pesin para la entropía, entonces toda componente ergódica μ_x de μ también satisface tal fórmula (cf. [Ke 1998]). Luego, del resultado de Sun y Tian concluimos que toda medida SRB-like para f , satisface la Fórmula de Pesin para la Entropía, si f es un difeomorfismo de Anosov de clase C^1 que preserve una medida equivalente a la medida de Lebesgue.

Más en general, sin necesidad de asumir hipótesis de C^1 -genericidad ni de existencia de medida invariante absolutamente continua respecto a la medida de Lebesgue, en [Cat-Cer-Enr 2012] se prueba que para todo f de Anosov de clase C^1 , toda medida SRB-like μ satisface la Fórmula de Pesin para la Entropía. Pero no sabemos si es necesario que además μ sea medida de Gibbs, ni que sea ergódica.

Capítulo 8

Espectro de Lebesgue y Medidas de Bernoulli

En este capítulo X es un espacio métrico compacto y $T : X \mapsto X$ es una transformación medible tal que existe alguna medida de probabilidad invariante por T .

8.1. El operador U_T en L^2

Definición 8.1.1. Operador U_T Sea $T : X \mapsto X$ medible que preserva la probabilidad μ . El operador $U_T : L^2(\mu) \mapsto L^2(\mu)$, inducido por T en $L^2(\mu)$ es:

$$U_T(f) = f \circ T \quad \forall f \in L^2(\mu)$$

Observación 8.1.2. Es inmediato verificar que:

U_T tiene autovalor 1 y el subespacio propio correspondiente es el de las funciones de L^2 que son invariantes por T . (Se recuerda que $f : X \mapsto C$ es invariante por T si $f \circ T = f$ μ -c.t.p.)

Observación 8.1.3. U_T es un operador unitario en $L^2(\mu)$, es decir: preserva el producto interno. En efecto, denotando con \bar{z} al conjugado de un número complejo z , tenemos:

$$\begin{aligned} \langle U_T f, U_T g \rangle &= \int (U_T f)(\overline{U_T g}) d\mu = \\ &= \int (f \circ T)(\overline{g \circ T}) d\mu = \int (f \bar{g}) \circ T d\mu = \\ &= \int f \bar{g} d\mu = \langle f, g \rangle \quad \square \end{aligned}$$

Como corolario se obtiene que U_T es una isometría (es decir $\|U_T f\|_{L^2(\mu)} = \|f\|_{L^2(\mu)}$ para todo $f \in L^2(\mu)$). Luego, todos los valores propios de U_T tienen módulo 1. En efecto, si $f \in L^2$ es un vector propio de autovalor λ entonces:

$$\|f\| = \|U_T f\| = \|\lambda f\| = |\lambda| \|f\| \Rightarrow |\lambda| = 1$$

Teorema 8.1.4. Caracterización espectral de la ergodicidad (ver por ejemplo [Si 1994, Lecture 2, Theorem 3])

Sea X espacio métrico compacto y sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ .

a) T es ergódica respecto de μ si y solo si el subespacio propio de autovalor 1 de U_T es el de las funciones constantes μ -c.t.p.

b) Si T es ergódica con respecto de μ entonces los subespacios propios de U_T son todos unidimensionales.

Demostración: a) Si las únicas funciones del subespacio propio de autovalor 1 son las constantes c.t.p., entonces usando 8.1.2, las únicas funciones invariantes con T son las constantes c.t.p., y recíprocamente. Esta última propiedad caracteriza la ergodicidad como se vio en el capítulo 2, teorema II-2.1.

b) Si λ es valor propio de U_T entonces $|\lambda| = 1$ porque U_T es una isometría (ver 8.1.3). Si f es vector propio de autovalor λ entonces $|U_T f| = |\lambda f| = |\lambda||f| = |f|$ de donde $|f|$ es propio con autovalor 1. Por la parte a) se obtiene $|f| = cte$ -c.t.p. Pero siendo f un vector propio de valor propio λ no puede ser idénticamente nula. Luego $|f| = cte \neq 0$, f no se anula μ -c.t.p.

Consideremos ahora g otra función propia con el mismo autovalor λ . El cociente g/f pertenece a $L^2(\mu)$ porque su módulo es constante μ -c.t.p. Aplicando $U_T(g/f) = (g \circ T)/(f \circ T) = (\lambda g)/(\lambda f) = g/f$. Luego $g/f \in L^2(\mu)$ es un vector propio de valor propio 1. Por la parte a) se tiene $g/f = cte$ μ -c.t.p. de donde el subespacio propio de autovalor λ está generado por f . \square

Ejercicio 8.1.5. Probar el recíproco de la parte b) del teorema anterior.

Teorema 8.1.6. Caracterización espectral del mixing

Sean X espacio métrico compacto y $T : X \mapsto X$ medible que preserva una probabilidad μ . T es mixing respecto de μ si y solo si

$$\lim_{n \rightarrow +\infty} \langle U_T^n f, g \rangle = \langle f, 1 \rangle \langle 1, g \rangle \quad \forall f, g \in L^2(\mu)$$

Esta prueba está extraída de [Mañ 1983a, págs. 182-183] Ver también [Mañ 1987] o [Si 1994, pages 21-22]

Demostración: Si

$$\lim_{n \rightarrow +\infty} \int (f \circ T^n) \bar{g} d\mu = \int f d\mu \int \bar{g} d\mu \quad \forall f, g \in L^2(\mu) \quad (8.1)$$

entonces en particular tomando $f = \chi_A, g = \chi_B$ se obtiene

$$\lim_{n \rightarrow +\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B) \quad \forall A, B \text{ medibles} \quad (8.2)$$

Recíprocamente, si vale (8.2), entonces, fijando el conjunto B medible, se obtiene (8.1) cuando $f = \chi_A, g = \chi_B$ para todo conjunto A medible. Luego, por la linealidad del límite y de la integral, también vale (8.1) para f función simple y $g = \chi_B$. Por la densidad de las funciones simples s en $L^2(\mu)$, si $\|f - s\|_2 < \epsilon$ y $g = \chi_B$, entonces

$$| \langle U_T^n f, g \rangle - \langle f, 1 \rangle \langle 1, g \rangle | \leq$$

$$\begin{aligned} &\leq | \langle U_T^n(f-s), g \rangle | + | \langle U_T^n s, g \rangle - \langle s, 1 \rangle \langle 1, g \rangle | \\ &\quad + | \langle s-f, 1 \rangle \langle 1, g \rangle | < 3\epsilon \end{aligned}$$

para todo n suficientemente grande. Luego vale (8.1) para $f \in L^2(\mu)$ y $g = \chi_B$. Ahora fijando $f \in L^2(\mu)$, acabamos de probar que ((8.1)) vale para $g = \chi_B$ donde B es cualquier conjunto medible. Por linealidad también vale (8.1) para g función simple, y usando el mismo método que antes se deduce que también vale (8.1) para toda $g \in L^2(\mu)$. \square

Ejercicio 8.1.7. Sea $T : X \mapsto X$ medible que preserve la probabilidad μ . Probar que:

- T es mixing respecto de μ si y solo si la igualdad del teorema 8.1.6 se cumple para un conjunto denso de funciones en $L^2(\mu)$
- T es mixing si $L^2(\mu)$ tiene base numerable (una familia numerable \mathcal{B} de funciones linealmente independientes tal que el subespacio de todas las combinaciones lineales finitas de funciones de \mathcal{B} es denso en L^2) y la igualdad del teorema 8.1.6 se cumple para toda pareja de funciones de esa base.

Ejercicio 8.1.8. Sea $T : X \mapsto X$ una transformación mixing respecto a una probabilidad invariante μ . Sea λ otra medida de probabilidad (no necesariamente invariante por T) en el mismo espacio medible, tal que $\lambda \ll \mu$. Probar que:

$$\lim_{n \rightarrow +\infty} \lambda(T^{-n}(A)) = \mu(A) \quad \forall A \text{ conjunto medible. Sugerencia:}$$

Sea $h \in L^1(\mu)$ real no negativa, la derivada de Radon- Nykodom $h = d\lambda/d\mu$. Como L^2 es denso en L^1 , dado $\epsilon > 0$ existe $g \in L^2(\mu)$ real no negativa (por ejemplo una función simple) tal que $\int |g-h| d\mu < \epsilon$. Siendo μ una medida mixing, para todo n suficientemente grande se tiene:

$$\left| \int (\chi_A \circ T^n) g d\mu - \mu(A) \int g d\mu \right| < \epsilon \quad (8.3)$$

Observando que

$$\lambda(T^{-n}(A)) = \int (\chi_A \circ T^n) h d\mu = \int (\chi_A \circ T^n) g d\mu - \int (\chi_A \circ T^n) (h-g) d\mu,$$

aplicando la propiedad triangular y (8.3), se deduce que, para todo n suficientemente grande, $|\lambda(T^{-n}(A)) - \mu(A)| < 3\epsilon$.

8.2. Equivalencia espectral

Definición 8.2.1. Equivalencia espectral Sean (X, \mathcal{A}, μ) y (Y, \mathcal{B}, ν) espacios de medida de probabilidad. Sean $T : X \mapsto X$ y $S : Y \mapsto Y$ transformaciones medibles que preservan μ y ν respectivamente. Se dice que T y S son *espectralmente equivalentes* si existe un operador lineal invertible

$$L : L^2(\mu) \mapsto L^2(\nu)$$

que preserva el producto interno y que cumple

$$U_S \circ L = L \circ U_T$$

Ejercicio 8.2.2. Probar que si dos transformaciones T y S son espectralmente equivalentes entonces:

- a) T es ergódica si y solo si S lo es.
 b) T es mixing si y solo si S lo es.

8.3. Bases de Fourier en $L^2(m)$

Sea $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ con la probabilidad de Lebesgue m , definida en los borelianos y que cumple:

$$\int_{z \in S^1} f(z) dm(z) = \int_{x \in [0,1)} f(e^{2\pi i x}) dx$$

Proposición 8.3.1. El conjunto de funciones $\{\varphi_n : n \in \mathbb{Z}\}$ definidas como

$$\varphi_n(z) = z^n = e^{2\pi i n x} \quad \forall z = e^{2\pi i x} \in S^1$$

es un conjunto ortonormal. (Es llamado Base de Fourier de $L^2(S^1, m)$.)

Demostración:

$$\langle \varphi_n, \varphi_m \rangle = \int \varphi_n \overline{\varphi_m} dm = \int_0^1 e^{2\pi i n x} e^{-2\pi i m x} dx = \int_0^1 e^{2\pi i (n-m)x} dx$$

La integral anterior es cero si $n \neq m \in \mathbb{Z}$ y es uno si $n = m \in \mathbb{Z}$. □

Usaremos el siguiente teorema, clásico del análisis espectral:

Teorema 8.3.2. Teorema de Parseval. Serie de Fourier. Sea el sistema ortonormal de Fourier $\{\varphi_n\}_{n \in \mathbb{Z}}$ definido en la proposición 8.3.1. Sea $f \in L^2(S^1, m)$ y sean $c_n(f) = \langle f, \varphi_n \rangle$ los llamados *coeficientes de Fourier de f* . Entonces la siguiente serie converge a f en $L^2(S^1, m)$:

$$\sum_{n \in \mathbb{Z}} c_n(f) \varphi_n = f \quad \text{en } L^2(S^1, m)$$

El Teorema 8.3.2 es parte de un teorema más general del Análisis Espectral, no solo aplicable al sistema ortonormal de Fourier de las funciones L^2 en el círculo. Su demostración se encuentra por ejemplo en [Rud 1979b, Teorema 7.9].

Corolario 8.3.3. El subespacio S formado por todas las combinaciones lineales finitas de funciones del sistema ortonormal de Fourier es denso en $L^2(\mu)$.

Sea $\mathbb{T}^2 = \mathbb{R}^2 / \mathbb{Z}^2$, con la medida de Borel de probabilidad de Lebesgue m que cumple:

$$\int_{z \in \mathbb{T}^2} f(z) dm(z) = \int_{(x,y) \in [0,1)^2} \widehat{f}(x,y) dx dy$$

siendo $\widehat{f}(x,y) := f(z) \quad \forall (x,y) \in \mathbb{R}^2$ tal que $z = (x,y)_{\text{mod } \mathbb{Z}^2} \in \mathbb{T}^2$.

Definición 8.3.4. Base de Fourier En $L^2(\mathbb{T}^2, m)$ se define el conjunto ortonormal $\{\varphi_{n,m} : n \in \mathbb{Z}, m \in \mathbb{Z}\}$ llamado *Base de Fourier* del siguiente modo:

$$\varphi_{n,m}(z) = e^{2\pi i(nx+my)} \text{ donde } (x, y) = \mathbb{R}^2, z = (x, y)_{\text{mod}\mathbb{Z}^2}$$

Corolario 8.3.5. *El subespacio S de las combinaciones lineales finitas de funciones en la base de Fourier de $L^2(\mathbb{T}^2, m)$ es denso.*

8.4. Rotación irracional revisitada

En el Teorema 2.6.1 ya probamos que la rotación irracional T del círculo es ergódica, pero no mixing, respecto a la medida de Lebesgue, y que esta es su única medida invariante. La teoría espectral permite una demostración de la ergodicidad de T , mucho más breve y elegante que la que hicimos a continuación del enunciado del Teorema 2.6.1. La generalización del siguiente teorema a toros n -dimensionales y más en general, a otras variedades con estructura de grupo topológico compacto, se encuentra por ejemplo en [Si 1994, Lecture 4, Theorem 2]

Teorema 8.4.1. *La rotación irracional del círculo es ergódica respecto a la medida de Lebesgue.*

Demostración: Llamemos $T : S^1 \mapsto S^1$ a la rotación irracional del círculo S^1 definida por $T(e^{2\pi ix}) = e^{2\pi i(x+\alpha)} \quad \forall x \in \mathbb{R}, \alpha \in (0, 1)$ irracional. Sea m la probabilidad de Lebesgue en el círculo. Tenemos

$$\int_{z \in S^1} f(z) dm(z) = \int_0^1 f(e^{2\pi ix}) dx$$

Sea $f : L^2(m)$ tal que $U_T(f) = f$. Basta probar que $f = \text{cte } \mu\text{-c.t.p.}$ Consideremos $\{\varphi_n\}_n \in \mathbb{Z} \subset L^2(m)$ la base de Fourier de $L^2(m)$ definida por

$$\varphi_n(e^{2\pi ix}) = e^{2\pi inx}.$$

Sea $c_n(f) = \langle f, \varphi_n \rangle$ el coeficiente de Fourier de f . Como

$$f =_{L^2(m)} \sum_{n \in \mathbb{Z}} c_n(f) \varphi_n$$

basta probar que si $U_T(f) = f$ entonces $c_n = \langle f, \varphi_n \rangle = 0 \quad \forall n \neq 0$.

En efecto

$$\begin{aligned} \langle f, \varphi_n \rangle &= \langle U_T f, \varphi_n \rangle = \int_{z \in S^1} f(Tz) \overline{\varphi_n(z)} dm(z) = \\ &= \int_0^1 f(e^{2\pi i(\alpha+x)}) e^{-2\pi inx} dx = \int_\alpha^{\alpha+1} f(e^{2\pi iy}) e^{-2\pi in(y-\alpha)} dy = \\ &= \int_0^1 f(e^{2\pi iy}) e^{-2\pi in(y-\alpha)} dy \end{aligned}$$

La penúltima igualdad se obtiene haciendo el cambio de variables $y = x + \alpha$ y la última igualdad se obtiene porque el integrando es una función periódica de período 1.

$$\begin{aligned} \int_0^1 f(e^{2\pi iy})e^{-2\pi ni(y-\alpha)} dy &= e^{2\pi in\alpha} \int_0^1 f(e^{2\pi iy})e^{-2\pi iny} dy = \\ &= e^{2\pi in\alpha} \langle f, \varphi_n \rangle \end{aligned}$$

Deducimos que

$$\langle f, \varphi_n \rangle = e^{2\pi in\alpha} \langle f, \varphi_n \rangle$$

de donde, si $n \neq 0$, se deduce que $\langle f, \varphi_n \rangle = 0$. □

La medida de Lebesgue en S^1 es positiva sobre abiertos y la topología en S^1 tiene base numerable de abiertos. En virtud a lo visto en el capítulo I, la ergodicidad de $T : S^1 \mapsto S^1$ respecto a la medida de Lebesgue, implica que T sea transitivo topológicamente y por lo tanto tenga una órbita densa. En la rotación del círculo una órbita cualquiera se obtiene de otra mediante una rotación. Luego, si una órbita es densa todas lo son. Esto demuestra lo siguiente:

Corolario 8.4.2. *La rotación irracional en el círculo es transitiva (topológicamente) y todas sus órbitas son densas.*

Ejercicio 8.4.3. En el toro k -dimensional $T^k = \mathbb{R}^k / \mathbb{Z}^k$ se define la traslación de vector $\alpha = (\alpha_1, \dots, \alpha_k) \in [0, 1)^k$ del modo siguiente:

$$T(z) = (x + \alpha)_{\text{mod}\mathbb{Z}^k}$$

donde $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ verifica $z = x_{\text{mod}\mathbb{Z}^k}$ (Es decir z es la clase de equivalencia en la relación $(x_1, \dots, x_k) \sim (y_1, \dots, y_k) \in \mathbb{R}^k$ cuando $y_j - x_j$ es entero para todo $1 \leq j \leq k$.)

Probar que si $\langle \alpha, p \rangle \notin \mathbb{Z} \quad \forall p \in \mathbb{Z}^k$ entonces T es ergódica respecto a la medida de Lebesgue m en T^k y recíprocamente. Nota: m es la medida de probabilidad de Borel en T^k , con la topología cociente inducida en T^k por la topología usual de \mathbb{R}^k , que cumple:

$$\int_{z \in T^k} f(z) dm(z) = \int_{x \in [0,1)^k} \widehat{f}(x) dx$$

donde dx indica integración respecto a la medida de Lebesgue en \mathbb{R}^k y $\widehat{f} : \mathbb{R}^k \mapsto \mathbb{C}$ es la función definida como $\widehat{f}(x) := f(z)$ si $z = x_{\text{mod}\mathbb{Z}^k}$.

Sugerencia: Imitar la prueba de la ergodicidad de la rotación irracional del círculo en el teorema anterior. Usar (no se pide demostrar esta afirmación) que el sistema de Fourier en $L^2(T^k, m)$: $\{\varphi_p\}_{p \in \mathbb{Z}^k}$ definido como

$$\varphi_p(z) = e^{2\pi i \langle p, z \rangle} \quad \forall z \in T^k \quad \forall p \in \mathbb{Z}^k$$

es una base ortonormal de $L^2(T^k, m)$. Es decir: es un sistema ortonormal y toda $f \in L^2(T^k, m)$ es la suma infinita numerable (tomando límite en la topología de $L^2(T^k, m)$) siguiente:

$$\sum_{p \in \mathbb{Z}^k} \langle f, \varphi_p \rangle \varphi_p$$

Teorema 8.4.4. Teorema de Katznelson [Katz 1968]

Sea $L : T^k \mapsto T^k$ una transformación lineal en el toro $T^k = \mathbb{R}^k / \mathbb{Z}^k$ $k \geq 1$ (i.e. L es la transformación inducida en $T^k = \mathbb{R}^k / \mathbb{Z}^k$ por la transformación lineal en \mathbb{R}^k dada por una matriz A $n \times n$, de coeficientes enteros y determinante igual a 1).

L preserva la medida de Lebesgue m en el toro. L es ergódica si y solo si ningún autovalor es raíz de la unidad.

Ejercicio 8.4.5. Probar el teorema de Katznelson. Sugerencia:

Usar la definición de la medida de Lebesgue dada en el ejercicio 8.4.3. Para probar que es invariante usar que $\det A = 1$, e imitar la prueba de la proposición I.4.1 del capítulo 1.

Para probar la caracterización de ergodicidad considerar la base de Fourier de $L^2(T^k, m)$ dada en la sugerencia del ejercicio 8.4.3. Probar que $\varphi_p \circ L^n = \varphi_{(A^*)^n p}$ $\forall p \in \mathbb{Z}^k$ $\forall n \in \mathbb{N}$ donde A^* denota la matriz simétrica de A .

Si $\det((A^*)^n - I) = 0$ para algún natural $n \geq 1$ probar que existe $p \in \mathbb{Z}^k, p \neq 0$ tal que $(A^*)^n p = p$. Sea $f = \sum_{j=0}^{n-1} \varphi_{(A^*)^j p}$. Probar que no es constante c.t.p. y es L invariante. Deducir que en ese caso L no es ergódica.

Si $\det((A^*)^n - I) \neq 0$ $\forall n \geq 1$ probar que entonces dado $p \in \mathbb{Z}^k, p \neq 0$ la sucesión $\{(A^*)^n p\}_{n \in \mathbb{N}}$ no repite sus valores en \mathbb{Z}^k . Probar que

$$\forall f \in L^2(T^k, m) \quad \forall n \in \mathbb{N} \quad \forall p \in \mathbb{Z}^k : \langle f \circ L^n, \varphi_p \rangle = \langle f, \varphi_{(A^*)^n p} \rangle$$

Sea $f \in L^2(T^k, m)$ tal que $f \circ L = f$ c.t.p. Considerar para $p \in \mathbb{Z}^k, p \neq 0$:

$$\begin{aligned} \sum_{n=1}^{\infty} |\langle f, \varphi_p \rangle|^2 &= \sum_{n=1}^{\infty} |\langle f \circ L^n, \varphi_p \rangle|^2 = \sum_{n=1}^{\infty} |\langle f, \varphi_{(A^*)^n p} \rangle|^2 \leq \\ &\leq \sum_{q \in \mathbb{Z}^k} |\langle f, \varphi_q \rangle|^2 = \|f\|_{L^2}^2 < +\infty \end{aligned}$$

De lo anterior se deduce que $\langle f, \varphi_p \rangle = 0$ para todo $p \neq 0$ en \mathbb{Z}^k . Concluir, como en la demostración del teorema 8.4.1, que f es constante m -c.t.p. y que L es ergódica en ese caso.

8.5. Espectro de Lebesgue

Sea $T : X \mapsto X$ medible, invertible con inversa medible, que preserva una probabilidad μ .

Definición 8.5.1. Espectro de Lebesgue Se dice que T tiene espectro de Lebesgue (con base numerable) respecto a la medida de probabilidad invariante μ , si existe $\{f_i\}_{i \in I} \subset L^2(\mu)$, donde I es un conjunto finito o infinito numerable, tal que:

$$\mathcal{B} = \{1, \{U_T^j f_i\}_{i \in I, j \in \mathbb{Z}}\}$$

es una base ortonormal de $L^2(\mu)$. (Es decir: \mathcal{B} es un conjunto ortonormal y el subespacio de todas las combinaciones lineales finitas de funciones de \mathcal{B} es denso en $L^2(\mu)$.)

Teorema 8.5.2.**Propiedad mixing del espectro de Lebesgue**

Sea T que preserva la medida de probabilidad μ . Si T tiene espectro de Lebesgue entonces T es mixing respecto de μ . Luego es ergódica.

Demostración: En virtud de lo enunciado en la parte b) del ejercicio 8.1.7 basta probar que la igualdad del teorema 8.1.6 vale para $f, g \in \{1, \{U_T^j f_i\}_{i \in I, j \in \mathbb{Z}}\} \subset L^2(\mu)$. Por un lado tenemos

$$\langle 1, 1 \rangle = 1, \quad \langle U_T^j f_i, 1 \rangle = 0 \quad \forall j \in \mathbb{Z} \quad \forall i \in I$$

Basta demostrar entonces que

$$\lim_{n \rightarrow +\infty} \langle U_T^n U_T^j f_i, 1 \rangle = 0, \text{ y que}$$

$$\lim_{n \rightarrow +\infty} \langle U_T^n U_T^j f_i, U_T^p f_k \rangle = 0 \quad \forall i, k \in I, \quad \forall p, j \in \mathbb{Z},$$

para todo $n \in \mathbb{N}$ suficientemente grande.

En efecto

$$\langle U_T^n U_T^j f_i, 1 \rangle = \langle U_T^{n+j} f_i, 1 \rangle = 0 \quad \forall n \in \mathbb{N}$$

pues $U_T^{n+j} f_i$ y 1 son funciones diferentes de la base ortonormal de $L^2(\mu)$. Análogamente

$$\langle U_T^n U_T^j f_i, U_T^p f_k \rangle = \langle U_T^{n+j} f_i, U_T^p f_k \rangle = 0 \quad \forall n > p - j.$$

□

Ejercicio 8.5.3. Sea T y S transformaciones que preservan probabilidades μ y ν respectivamente. Probar que si T y S son espectralmente equivalentes entonces T tiene espectro de Lebesgue respecto de μ si y solo si S lo tiene respecto de ν .

8.6. Automorfismos lineales en el toro con espectro de Lebesgue.

Teorema 8.6.1. La transformación $T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} : \mathbb{T}^2 \mapsto \mathbb{T}^2$ tiene espectro de Lebesgue respecto a la medida de Lebesgue m . Luego, m es mixing y por lo tanto también ergódica para la transformación T .

Demostración: Sea la base ortonormal de Fourier $\mathcal{B} = \{\varphi_{n,m}\}_{(n,m) \in \mathbb{Z}^2}$ de $L^2(\mathbb{T}^2, m)$ definida en 8.3.4.

$$\varphi_{n,m}((x, y)_{\text{mod } \mathbb{Z}^2}) = e^{2\pi i(n x + m y)} \quad \forall (x, y) \in \mathbb{R}^2, \quad \forall (n, m) \in \mathbb{Z}^2$$

Tenemos $U_T \varphi_{n,m} = \varphi_{2n+m, n+m}$.

En \mathbb{Z}^2 elegimos un conjunto de índices formado por un representante de cada órbita de la transformación lineal $A^* = (n, m) \in \mathbb{Z}^2 \mapsto (2n + m, n + m) \in \mathbb{Z}^2$. Sea I el conjunto

de todos los representantes que no son $(0, 0)$. Como la matriz A^* es invertible y no tiene valores propios que sean raíces de la unidad, entonces $\det((A^*)^j - Id) \neq 0$ para todo $j \in \mathbb{N}$. Se deduce que para todo $j_1 \neq j_2 \in \mathbb{Z}$ se cumple

$$(A^*)^{j_1}(n, m) \neq (A^*)^{j_2}(n, m) \quad \forall (n, m) \neq (0, 0) \in \mathbb{Z}^2$$

Entonces la familia de funciones

$$\{\varphi_{(A^*)^j(n, m)}\}_{j \in \mathbb{Z}, (n, m) \in I}$$

es un reordenamiento de la base de Fourier \mathcal{B} sacando el 1. Luego

$$\mathcal{B} = \{1, \{U_T^j \varphi_{(n, m)}\}_{j \in \mathbb{Z}, (n, m) \in I}\}$$

es una base ortonormal de $L^2(\mu)$. \square

El teorema anterior da una demostración de la ergodicidad del $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ en el toro, que ya probamos de otras formas en la Sección 3.1 y en el Corolario 6.3.5. El teorema 8.6.1 sin embargo tiene una tesis más fuerte que la de la ergodicidad de T :

Corolario 8.6.2. *La transformación $T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} : \mathbb{T}^2 \mapsto \mathbb{T}^2$ respecto a la medida de Lebesgue es mixing. Luego es ergódica, topológicamente mixing, y topológicamente transitiva.*

Ejercicio 8.6.3. Teorema de Katznelson II.

Sea $L : T^k \mapsto T^k$ una transformación lineal en el toro $T^k = \mathbb{R}^k / \mathbb{Z}^k$ $k \geq 1$ que preserva la medida de Lebesgue m en T^k . (i.e. L es la transformación inducida en $T^k = \mathbb{R}^k / \mathbb{Z}^k$ por la transformación lineal en \mathbb{R}^k dada por una matriz A $n \times n$, de coeficientes enteros y determinante igual a 1).

Probar que *si los autovalores de A no son raíces de la unidad entonces T tiene espectro de Lebesgue respecto de m .* Sugerencia:

Imitar la prueba del teorema 8.6.1, usando las mismas sugerencias que en el ejercicio 8.4.5.

8.7. El espacio del shift y su topología.

Sea $K = \{0, 1, \dots, k-1\}$, donde $k \geq 2$, un conjunto finito llamado *alfabeto*. A los elementos del alfabeto se le llama *símbolos*.

Definición 8.7.1. Espacio del shift

Se denota con $k^{\mathbb{Z}}$ o $K^{\mathbb{Z}}$ al conjunto, llamado *espacio del shift*, formado por todas las sucesiones bi-infinitas $a : \mathbb{Z} \mapsto K$, es decir:

$$k^{\mathbb{Z}} = \{\{a_n\}_{n \in \mathbb{Z}} : a_n \in K \quad \forall n \in \mathbb{Z}\}$$

Se llama *shift* (para la izquierda) a la aplicación $\sigma : k^{\mathbb{Z}} \mapsto k^{\mathbb{Z}}$ dada por

$$\sigma(\{a_n\}_{n \in \mathbb{Z}}) = \{b_n\}_{n \in \mathbb{Z}} \quad \text{donde } b_n = a_{n+1}$$

El shift es un corrimiento de los términos de la sucesión un lugar para la izquierda.

Definición 8.7.2. Cilindros en el espacio del shift

Dados $l \geq 1$ natural, dados $a_1, \dots, a_l \in K$ y dados l enteros diferentes n_1, \dots, n_l se llama *cilindro de radio l , posiciones n_i y símbolos a_i* al subconjunto $C_{n_1, \dots, n_l}^{a_1, \dots, a_l} \subset k^{\mathbb{Z}}$ definido por:

$$C_{n_1, \dots, n_l}^{a_1, \dots, a_l} = \{ \{b_n\}_{n \in \mathbb{Z}} \in k^{\mathbb{Z}} : b_{n_j} = a_j \forall 1 \leq j \leq l \} = \bigcap_{j=1}^l C_{n_j}^{a_j}$$

Se observa que la familia \mathcal{C} de todos los cilindros es numerable.

Ejercicio 8.7.3. Probar que dos cilindros son disjuntos si tienen el mismo radio, el mismo vector de posiciones y los vectores de símbolos respectivos son diferentes. Probar que la intersección de cilindros o bien es vacía o bien es otro cilindro. Probar que el complemento de un cilindro C es la unión finita (disjunta) de todos los cilindros con el mismo vector de posiciones y diferente vector de símbolos que C . Probar que toda unión finita de cilindros puede descomponerse como unión finita disjunta de cilindros. Probar que la familia \mathcal{A}_0 formada por el conjunto vacío y por todas las uniones finitas de cilindros disjuntos dos a dos, es un álgebra (es cerrada en el complemento, en intersecciones finitas y en uniones finitas).

Definición 8.7.4. Métrica en el espacio del shift

En $k^{\mathbb{Z}}$ se define la estructura de *espacio métrico* mediante la distancia siguiente:

$$\text{dist}(\{a_n\}_{n \in \mathbb{Z}}, \{b_n\}_{n \in \mathbb{Z}}) = |b_0 - a_0| + \frac{1}{2} \sum_{n \in \mathbb{Z}} \frac{|b_n - a_n|}{k^{|n|}}$$

Denotamos con $B_\epsilon(a)$ a la bola abierta de centro $a \in k^{\mathbb{Z}}$ y radio $\epsilon > 0$.

Lema 8.7.5. Para todo $a = \{a_n\}_{n \in \mathbb{Z}} \in k^{\mathbb{Z}}$ y para todo $\epsilon > 0$ suficientemente pequeño se cumple

$$B_\epsilon(a) \subset \bigcap_{j=-N}^{j=N} C_j^{a_j}$$

donde $N < -1 + (-\log \epsilon)/(\log k)$.

Para todo $a = \{a_n\}_{n \in \mathbb{Z}} \in k^{\mathbb{Z}}$ y para todo $N \geq 1$ se cumple

$$\bigcap_{j=-N}^{j=N} C_j^{a_j} \subset B_\epsilon(a)$$

donde $\epsilon > 1/k^N$

Ejercicio 8.7.6. Demostrar el lema 8.7.5. Sugerencia: Para la parte a) $\epsilon < 1/k^{N+1}$. Cada sumando de la serie que define $\text{dist}(a, b)$ es o bien nulo (si $a_n = b_n$), o bien mayor o igual que $1/(2k^{|n|})$ (si $a_n \neq b_n$). Luego $\text{dist}(a, b) < \epsilon < 1/2^{N+1} \Rightarrow a_n = b_n \forall |n| \leq N$.

Para la parte b), si $a_n = b_n \forall |n| \leq N$ entonces

$$\text{dist}(a, b) \leq (1/2) \sum_{|n| \geq N+1} \frac{k-1}{k^{|n|}} = \frac{1}{2k^N} < \epsilon$$

Ejercicio 8.7.7. Probar que todo cilindro es abierto. Probar que todo abierto es unión de cilindros. Probar que los cilindros son una base numerable de la topología en $k^{\mathbb{Z}}$. (Sugerencia: usar el lema 8.7.5)

Ejercicio 8.7.8. Probar que la transformación shift es continua con inversa continua (es un homeomorfismo). Sugerencia: usar el resultado del ejercicio 8.7.7.

Teorema 8.7.9. Compacidad del espacio del shift

$k^{\mathbb{Z}}$ con la estructura de espacio métrico dada en 8.7.4, es compacto. Los cilindros son abiertos y cerrados. La familia \mathcal{C} de todos los cilindros es una base numerable de la topología, y genera la σ -álgebra de Borel.

Demostración: Que todo cilindro es abierto y que la familia \mathcal{C} es una base de la topología se prueba en el ejercicio 8.7.7. El complemento de un cilindro es unión (finita) de cilindros. Luego todo cilindro es cerrado.

Como todo abierto es unión de cilindros a lo sumo numerable porque \mathcal{C} es numerable, entonces la σ -álgebra \mathcal{A} generada por los cilindros contiene a todos los abiertos. Luego contiene a la sigma-álgebra de Borel. Por otra parte los cilindros son abiertos, luego la sigma-álgebra de Borel contiene a \mathcal{A} .

Sólo resta probar que $k^{\mathbb{Z}}$ con la estructura de espacio métrico dada en 8.7.4 es compacto. Siendo un espacio métrico con base numerable de abiertos, es suficiente demostrar que toda sucesión

$$\{a^{(m)}\}_{m \in \mathbb{N}} = \{\{a_n^{(m)}\}_{n \in \mathbb{Z}}\}_{m \in \mathbb{N}} \subset k^{\mathbb{Z}}$$

tiene una subsucesión convergente.

Sea $\{b^{(m,0)}\}_{m \in \mathbb{N}}$ una subsucesión de $\{a^{(m)}\}_{m \in \mathbb{N}}$ tal que los símbolos en las posiciones $0 : b_0^{(m,0)}$ sean una constante $c_0 \in K$ respecto de $m \in \mathbb{N}$. Esto se puede conseguir porque la sucesión de primeros términos $a_0^{(m)} \in K$ toma valores en el conjunto finito K para todo $m \in \mathbb{N}$.

Sea

$$\{b^{(m,1)}\}_{m \in \mathbb{N}}$$

una subsucesión de $\{b^{(m,0)}\}_{m \in \mathbb{N}}$ tal que sus símbolos en las posiciones 1 y -1 sean constantes respecto de m . Esto es posible porque la sucesión de parejas de símbolos en esas posiciones toma valores en el conjunto finito K^2 .

Sea para todo $j \geq 1$ la sucesión $\{b^{(m,j)}\}_{m \in \mathbb{N}}$, subsucesión de $\{b^{(m,j-1)}\}_{m \in \mathbb{N}}$ tal que sus símbolos en las posiciones $-j, \dots, -1, 0, 1, \dots, j$ sean constantes en relación a m , respectivamente iguales a símbolos

$$c_{-j}, \dots, c_{-1}, c_0, c_1, \dots, c_j$$

del alfabeto.

Sea $c = \{c_n\}_{n \in \mathbb{Z}} \in k^{\mathbb{Z}}$

Por el procedimiento diagonal tomemos la subsucesión

$\{q^{(m)}\}_{m \in \mathbb{N}} = \{\{q_n^{(m)}\}_{n \in \mathbb{Z}}\}_{m \in \mathbb{N}} \subset k^{\mathbb{Z}}$ definida como $q^{(m)} = b^{(m,m)} \forall j \in \mathbb{N}$. Por construcción, dado $N \geq 1$ se cumple $q_n^{(m)} = c_n \forall |n| \leq N, \forall m \geq N$. Dicho de otra forma, dado N la sucesión $\{q^{(m)}\}_{m \in \mathbb{N}}$ en $k^{\mathbb{Z}}$ para todo $m \geq N$ está en el cilindro $\bigcap_{|n| \leq N} C_n^{c_n}$. Por lo tanto converge a $c \in k^{\mathbb{Z}}$ como queríamos demostrar. \square

8.8. Espacio de medida del shift de Bernoulli

Consideremos en $k^{\mathbb{Z}}$ con la sigma-álgebra de Borel, el homeomorfismo shift $\sigma : k^{\mathbb{Z}} \mapsto k^{\mathbb{Z}}$ definido en la sección anterior.

Definición 8.8.1. Vector de probabilidad

Se llama *vector de probabilidad* a un vector

$p = (p_0, p_1, \dots, p_{k-1}) \in (0, 1)^k$, tal que $\sum_{j=0}^{k-1} p_j = 1$. El valor p_j se llama *probabilidad del símbolo* $j \in K = \{0, 1, \dots, k-1\}$. Observar que solo llamamos vector de probabilidad a p cuando todas sus componentes *son estrictamente positivas*. (Algunos autores llaman vector de probabilidad a p aunque alguna de sus componentes sean nulas.)

Dado un vector de probabilidad p se define en los borelianos de $k^{\mathbb{Z}}$ una medida de probabilidad μ_p del siguiente modo:

Definición 8.8.2. Sea \mathcal{C} la colección de los cilindros en $k^{\mathbb{Z}}$. Sea \mathcal{A}_0 el álgebra de las uniones finitas de cilindros disjuntos dos a dos. Sea \mathcal{B} la sigma álgebra de Borel en $k^{\mathbb{Z}}$.

Dado un vector de probabilidad $p = (p_0, \dots, p_{k-1})$:

1) Se define $\mu_p : \mathcal{C} \mapsto [0, 1]$ definida del siguiente modo:

$$\mu_p\left(\bigcap_{j=1}^l C_{n_j}^{a_j}\right) = \prod_{j=1}^l p_{a_j}$$

2) Se define la premedida $\mu_p : \mathcal{A}_0 \mapsto [0, 1]$, como

$$\mu_p(\emptyset) = 0, \mu_p(k^{\mathbb{Z}}) = 1, \mu_p(A) = \sum_{i=1}^N \mu_p(C_i) \quad \forall A \in \mathcal{A} \text{ tal que}$$

$$A = \bigcup_{i=1}^N C_i$$

donde $\{C_i\}_{1 \leq i \leq N}$ es una familia de cilindros disjuntos dos a dos.

Ejercicio 8.8.3. Probar que $\mu_p : \mathcal{A}_0 \mapsto [0, 1]$ definida anteriormente es una premedida en el álgebra \mathcal{A}_0 .

En virtud del enunciado del ejercicio anterior, existe una única extensión de la premedida $\mu_p : \mathcal{A}_0 \mapsto [0, 1]$ definida en 8.8.2, a la sigma álgebra de Borel \mathcal{B} en el espacio $k^{\mathbb{Z}}$, que es la sigma-álgebra generada por \mathcal{A}_0 .

Definición 8.8.4. Medida de Bernoulli

Se llama *medida de Bernoulli de vector* p a la única extensión μ_p a la sigma álgebra de Borel de $k^{\mathbb{Z}}$ de la premedida definida en 8.8.2. Es una medida de probabilidad ya que

$$\mu_p(k^{\mathbb{Z}}) = \mu_p\left(\bigcup_{j=0}^{k-1} C_0^j\right) = \sum_{j_0}^{k-1} p_j = 1$$

Observación 8.8.5. *Es fácil verificar que la medida de Bernoulli es invariante por el shift.*

Definición 8.8.6. Shift de Bernoulli

Se llama *shift de Bernoulli* en el espacio $k^{\mathbb{Z}}$, a la pareja (σ, μ_p) donde $\sigma : k^{\mathbb{Z}} \mapsto k^{\mathbb{Z}}$ es la transformación shift definida en la sección anterior, y μ_p es la medida de Bernoulli para cierto vector de probabilidad p dado.

8.9. Ergodicidad del shift de Bernoulli.**Teorema 8.9.1. Propiedad mixing del shift de Bernoulli**

El shift de Bernoulli (σ, μ_p) es mixing. Luego, es ergódico.

Demostración: El shift $\sigma : k^{\mathbb{Z}} \mapsto k^{\mathbb{Z}}$ preserva la medida de Bernoulli μ_p con vector de probabilidad $p = (p_0, \dots, p_k)$ en los conjuntos de la sigma álgebra de Borel \mathcal{B} . Denotemos con \mathcal{A}_0 al álgebra de las uniones finitas disjuntas de cilindros en $k^{\mathbb{Z}}$.

Hay que probar que

$$\lim_{n \rightarrow +\infty} \mu_p(\sigma^{-n}(A) \cap B) = \mu_p(A)\mu_p(B) \quad \forall A, B \in \mathcal{B}$$

En primer lugar, es fácil verificar que la igualdad anterior es cierta para los cilindros (ver ejercicio 8.9.2) y para los conjuntos del álgebra \mathcal{A}_0 formada por las uniones finitas de cilindros disjuntos dos a dos.

Por construcción la medida μ_p es la única extensión de la premedida en \mathcal{A}_0 , a la sigma álgebra por ella generada, que es la sigma álgebra de Borel \mathcal{B} . Por el teorema de extensión de premedidas:

$$\forall A \in \mathcal{B} \quad \mu_p(A) = \inf\left\{\sum_{i \in \mathbb{N}} \mu_p(A_i), A_i \in \mathcal{A}_0, \bigcup_{i \in \mathbb{N}} A_i \supset A\right\}$$

(Véase Teorema 1.14 e igualdad (1.12) en la página 30 del libro de G. Folland: *Real Analysis*, edit. J. Wiley, 1984.)

Luego, dados $A, B \in \mathcal{B}$ y dado $\epsilon > 0$, existen $A_0, B_0 \in \mathcal{A}_0$ tales que

$$\mu_p(A \Delta A_0) < \epsilon, \quad \mu_p(B \Delta B_0) < \epsilon$$

Usando la propiedad triangular:

$$|\mu_p(\sigma^{-n}(A) \cap B) - \mu(A)\mu(B)| \leq |\mu_p(\sigma^{-n}(A) \cap B) - \mu_p(\sigma^{-n}(A_0) \cap B_0)| + \quad (8.4)$$

$$+ |\mu_p(\sigma^{-n}(A_0) \cap B_0) - \mu(A_0)\mu(B_0)| + \quad (8.5)$$

$$+ |\mu(A_0)\mu(B_0) - \mu(A_0)\mu(B)| + |\mu(A_0)\mu(B) - \mu(A)\mu(B)| \quad (8.6)$$

El sumando (8.4) es menor que 2ϵ para todo $n \in \mathbb{N}$: para verificarlo usar que $|\mu(C) - \mu(D)| \leq \mu(C\Delta D)$ y que $(C_1 \cap C_2)\Delta(D_1 \cap D_2) \subset (C_1\Delta D_1) \cup (C_2\Delta D_2)$.

La expresión (8.5) tiende a cero cuando $n \rightarrow +\infty$ porque $A_0, B_0 \in \mathcal{A}_0$.

Finalmente cada sumando de (8.6) es menor que ϵ para todo n .

Se concluye entonces que el primer miembro de (8.4) tiende a cero cuando $n \rightarrow +\infty$ como queríamos demostrar. \square

Ejercicio 8.9.2. Sea (σ, μ_p) el shift de Bernoulli en $k^{\mathbb{Z}}$. Sean dados $m_1, n_1 \in \mathbb{Z}$, el natural $n > |n_1| + |m_1|$ y los símbolos $a_1, b_1 \in K$. Verificar que

$$\mu_p(\sigma^{-n}(C_{n_1}^{a_1}) \cap C_{m_1}^{a_2}) = p_{a_1}p_{a_2}$$

Deducir que

$$\lim_{n \rightarrow +\infty} \mu_p(\sigma^{-n}(C_{n_1}^{a_1}) \cap C_{m_1}^{a_2}) = \mu_p(C_{n_1}^{a_1})\mu_p(C_{m_1}^{a_2})$$

Enunciar y probar una igualdad similar para cualquier pareja de cilindros, aunque tengan radio mayor que 1.

Probar que también vale una igualdad como la anterior para cualquier pareja A_0, B_0 en el álgebra \mathcal{A}_0 formada por todas las uniones finitas de cilindros disjuntos dos a dos.

Ejercicio 8.9.3. Probar que el shift es topológicamente mixing y transitivo. Sugerencia: Verificar que la medida de Bernoulli es positiva sobre abiertos, y usar el teorema 8.9.1.

Observación 8.9.4. Espectro de Lebesgue del shift de Bernoulli

Aunque no lo demostraremos en este curso también vale el siguiente enunciado más fuerte que el teorema 8.9.1:

El shift de Bernoulli (σ, μ_p) tiene espectro de Lebesgue.

Definición 8.9.5. Shift unilateral

Sea $k^{\mathbb{N}}$ el conjunto formado por todas las sucesiones $a : \mathbb{N} \mapsto K = \{0, 1, \dots, k-1\}$. Se define el shift (unilateral)

$$\sigma : k^{\mathbb{N}} \mapsto k^{\mathbb{N}} : \sigma(\{a_n\}_{n \in \mathbb{N}}) = \{a_{n+1}\}_{n \in \mathbb{N}}$$

Observación 8.9.6. La teoría del shift unilateral es la misma que la del shift bilateral definido en las secciones anteriores, excepto que ahora σ no es invertible. De la misma forma que antes, se construye la medida de Bernoulli μ_p dado un vector de probabilidad para el espacio de símbolos, y se demuestra que el shift unilateral de Bernoulli (σ, μ_p) es mixing, y por lo tanto, ergódico.

Ejercicio 8.9.7. a) Sea $T : [0, 1] \mapsto [0, 1]$ el tent map definido en el ejercicio I.4 del capítulo 1. Sea m la medida de probabilidad de Lebesgue en $[0, 1]$. Sea $\Lambda \subset [0, 1]$ el conjunto obtenido al retirar del intervalo $[0, 1]$ el punto $1/2$ y todas las preimágenes de $1/2$ (es decir: retirando todos aquellos puntos cuyas órbitas futuras pasan por el punto $1/2$). Probar que $m(\Lambda) = 0$.

b) Llámese $I_0 = [0, 1/2]$, $I_1 = [1/2, 1]$. Se construye la aplicación medible $h : \Lambda \mapsto 2^{\mathbb{N}}$ (llamada *itinerario*) del siguiente modo:

$$h(x) = \{a_n\}_{n \in \mathbb{N}} \Leftrightarrow T^n(x) \in I_{a_n} \forall n \in \mathbb{N}$$

Demostrar que h es inyectiva. Demostrar que

$$\sigma \circ h = h \circ T$$

donde σ es el shift unilateral.

c) Considérese el vector de probabilidad $p = (1/2, 1/2)$, y la medida de Bernoulli μ_p en $2^{\mathbb{N}}$. Demostrar que $m(h^{-1}(A)) = \mu_p(A)$ para todo boreliano $A \subset 2^{\mathbb{N}}$. (Sugerencia: Basta probarlo para uniones finitas disjuntas de cilindros). Deducir que $\mu_p(h(\Lambda)) = 1$ (h es invertible c.t.p.)

d) Demostrar que el tent map es ergódico y mixing respecto de m . (Sugerencia: usar que el shift unilateral de Bernoulli es mixing).

e) Deducir que el tent map es transitivo y topológicamente mixing.

Ejercicio 8.9.8. Sea $T : S^1 \mapsto S^1$ la transformación definida por

$$T(z) = T(e^{2\pi ix}) = e^{4\pi ix} = z^2 \quad \forall x \in [0, 1)$$

a) Demostrar que T preserva la probabilidad de Lebesgue m en S^1 .

b) Considérese el conjunto Λ que se obtiene del círculo retirando todas las preimágenes de $e^{2\pi ix_0} = -1$ para $x_0 = 1/2 \in [0, 1)$. Llámese $I_0 = \{e^{2\pi ix} : x \in [0, 1/2)\}$, $I_1 = \{e^{2\pi ix} : x \in (1/2, 1)\}$. Defínase la transformación itinerario $h : \Lambda \mapsto 2^{\mathbb{N}}$ como en el ejercicio anterior, y probar que cumple las mismas propiedades enunciadas en ese ejercicio.

c) Demostrar que $T : S^1 \mapsto S^1$ es ergódica y mixing respecto de m , transitiva y topológicamente mixing.

Ejercicio 8.9.9. Investigar si es ergódica y mixing respecto a la probabilidad de Lebesgue, transitiva y topológicamente mixing, la transformación $T : S^1 \mapsto S^1 = \{z \in \mathbb{C} : |z| = 1\}$ definida por $T(z) = z^k$, donde $k \geq 2$ es un número natural.

Equivalencia de automorfismos de espacios de medida.

Definición 8.9.10. Sean $T : X \mapsto X$, $S : Y \mapsto Y$, medibles que preservan probabilidades μ en X y ν en Y respectivamente. Se dice que (T, μ) y (S, ν) son *equivalentes (como automorfismos de medida)* si existe $h : X \mapsto Y$ bimedible c.t.p. (es decir medible, y tal que existe $g : Y \mapsto X$ medible que cumple $g \circ h = id_X \mu - c.t.p.$, $h \circ g = id_Y \nu - c.t.p.$) y tal que:

a) $\mu(h^{-1}(B)) = \nu(B) \forall B \subset Y$ medible.

b) $S \circ h = h \circ T \mu - c.t.p.$

Ejercicio 8.9.11. Probar que la equivalencia de automorfismos de medida es una relación de equivalencia.

Ejercicio 8.9.12. Demostrar que si (T, μ) y (S, ν) , son equivalentes entonces:

- a) T es ergódica si y solo si S lo es.
- b) T es mixing si y solo si S lo es.
- c) T y S son espectralmente equivalentes.
- d) T tiene espectro de Lebesgue si y solo si S lo tiene.

Observación 8.9.13. Equivalencia espectral y equivalencia de automorfismos de espacios de medida

Se pueden encontrar dos shift de Bernoulli que son espectralmente equivalentes pero que no son equivalentes como automorfismos de medida. Aunque no lo demostraremos, eso prueba que la equivalencia espectral es diferente de la equivalencia de automorfismos de medida: las clases de equivalencia espectral son más amplias que las clases de equivalencia de automorfismos de medida.

Observación 8.9.14. Conjugación y equivalencia de automorfismos de espacios de medida

Recordemos la definición de *conjugación* dada en 3.6.3: Dos transformaciones $T : X \mapsto X$ y $S : Y \mapsto Y$ en espacios métricos (o más en general en espacios topológicos) se llaman conjugadas si existe un homeomorfismo $h : X \mapsto Y$ tal que

$$h \circ T = S \circ h.$$

(Se recuerda que un homeomorfismo es una transformación continua, invertible y con inversa continua).

La conjugación define una relación de equivalencia topológica entre los sistemas dinámicos por iterados de T y de S respectivamente. Pero además, es inmediato chequear que si existe una conjugación, y μ es una medida de probabilidad en X invariante por T , entonces la medida $h^*\mu$ es una probabilidad en Y invariante por S , y h es un isomorfismo de los espacios de medida (X, μ) e $(Y, h^*\mu)$, con las sigma-álgebras de Borel respectivas. (Recordamos que, por definición del operador h^* se cumple $(h^*\mu)(B) := \mu(h^{-1}(B))$ para todo boreliano $B \subset Y$).

En resumen, la conjugación h o equivalencia topológica entre sistemas dinámicos, los clasifica en clases para las cuales las propiedades topológicas y medibles, como automorfismos de espacios de medida, se preservan. Dos sistemas dinámicos conjugados (o topológicamente equivalentes) son también equivalentes como automorfismos de espacios de medida (X, μ) e $(Y, h^*\mu)$. Sin embargo el recíproco es falso: por ejemplo un difeomorfismo de Anosov f transitivo en el toro \mathbb{T}^2 , provisto de una medida invariante μ se demuestra ser equivalente como automorfismo de espacio de medida, a un shift σ en el espacio de las sucesiones con una cantidad finita de símbolos, provisto de cierta medida σ -invariante, llamada de Markov (ver por ejemplo [Kat-Has 1995, Theorem 18.7.4]). Pero f no puede ser conjugado a σ ya que los espacios métricos donde actúan (el toro \mathbb{T}^2 y el espacio del shift respectivamente) no son homeomorfos. En efecto, mientras el toro \mathbb{T}^2 es conexo, el espacio del shift es totalmente desconexo.

Por lo tanto, en general, la equivalencia topológica entre sistemas dinámicos (i.e. la existencia de conjugación) es estrictamente más fuerte que la equivalencia como automorfismos de espacios de medida. Pero en el caso particular en que los dos sistemas dinámicos sean shifts de Bernoulli, la equivalencia como automorfismos de espacios de medida implica también la equivalencia topológica (Teorema de Orstein que enunciamos más abajo, extraído de [Jo 2005, Theorem 31, page 137]). Más precisamente: En el próximo capítulo definiremos *entropía métrica* de una transformación T que preserva una medida de probabilidad μ : esta es un número no negativo $h_\mu(T)$ que depende solo de T y de μ , y que mide el grado de desorden espacial que producen los iterados de T en relación a la medida de probabilidad μ , y que es invariante por automorfismos de espacios de medida. Si dos sistemas dinámicos son equivalentes como automorfismos de espacios de medida, entonces sus entropías métricas son iguales. En general, sistemas dinámicos que tienen la misma entropía métrica, pueden ser muy diferentes desde el punto de vista de su dinámica topológica. Pero, en el caso particular de los shifts de Bernoulli, el siguiente resultado muestra que la entropía métrica $h_\mu(T)$ caracteriza la dinámica topológica del sistema:

Teorema de Orstein

Dos shifts de Bernoulli son conjugados si y solo si tienen la misma entropía métrica. La demostración de un enunciado más general, que incluye como caso particular la prueba del Teorema de Orstein, se encuentra por ejemplo en [Kea-Smo 1979].

8.10. Transformaciones de Bernoulli

Definición 8.10.1. Transformación de Bernoulli

Sea $T : X \mapsto X$ que preserva una medida de probabilidad μ . Se dice que (T, μ) es una *transformación de Bernoulli* si es equivalente como automorfismo de medida a un shift de Bernoulli (σ, μ_p) en el espacio $k^{\mathbb{Z}}$ o en el espacio $k^{\mathbb{N}}$ para cierto $k \geq 2$ y cierto vector de probabilidad μ .

Teorema 8.10.2. *Si (T, μ) es una transformación de Bernoulli entonces T es mixing respecto de μ . (Luego es ergódica).*

Demostración: Por el teorema 8.9.1 el shift de Bernoulli es mixing, y (T, μ) es equivalente a él. Luego, según el enunciado en el ejercicio 8.9.12, T es mixing respecto de μ . □

Observación 8.10.3. Enunciamos que los shift de Bernoulli tienen espectro de Lebesgue. Luego, las transformaciones de Bernoulli también tienen espectro de Lebesgue, ya que son equivalentes a los shift. La existencia de espectro de Lebesgue es más fuerte que la propiedad mixing.

Ejemplo 8.10.4. Ejemplos de transformaciones de Bernoulli son el *tent map* en el intervalo y la transformación $z \mapsto z^k$ con $z \in S^1 = \{z \in C : |z| = 1\}$ y $k \geq 2$ fijo natural (ver ejercicios 8.9.8 y 8.9.9). En dimensión 2 un ejemplo de transformación de Bernoulli es la herradura de Smale que veremos en la última sección de este capítulo.

Jerarquía ergódica

En resumen tenemos la siguiente jerarquía de propiedades ergódicas cada vez más restrictivas:

- a) $\{(T, \mu) \text{ ergódica}\} \supsetneq \{(T, \mu) \text{ mixing}\}$.
- b) $\{(T, \mu) \text{ mixing}\} \supsetneq \{(T, \mu) \text{ tiene espectro de Lebesgue}\}$.
- c) $\{(T, \mu) \text{ tiene espectro de Lebesgue}\} \supsetneq \{(T, \mu) \text{ es transformación de Bernoulli}\}$.

En la Sección 2.7 probamos a). El ejemplo clásico de transformación ergódica no mixing es la rotación irracional del círculo.

En el Teorema 8.5.2 probamos parte de b): toda transformación con espectro de Lebesgue es mixing. No dimos ningún ejemplo para refutar el recíproco. Los llamados *shifts gaussianos*, que no estudiaremos en este curso, son transformaciones mixing que no tienen espectro de Lebesgue (ver por ejemplo [Mañ 1983a, pág. 194–201]).

Ejemplo 8.10.5. Medidas de Bernoulli en la Herradura de Smale

Recordamos la definición 3.5.1 de herradura de Smale lineal (en dimensión 2) $T : Q \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$, donde $Q = [0, 1]^2$. Recordamos la definición 3.5.3 del conjunto Λ maximal invariante de T . En el ejercicio 3.5.4 se probó que Λ es un conjunto uniformemente hiperbólico.

Definición 8.10.6. Itinerario

En el conjunto invariante maximal Λ de la herradura de Smale, se define la *topología inducida* por la de \mathbb{R}^2 : un conjunto en Λ es abierto si y solo si es la intersección con Λ de un abierto de \mathbb{R}^2 .

Se llama *itinerario* $h : \Lambda \mapsto 2^{\mathbb{Z}}$ a la función definida como:

$$h(x) = \{a_n\}_{n \in \mathbb{Z}} : T^n(x) \in Q_{a_n} \quad \forall n \in \mathbb{Z}$$

Ejercicio 8.10.7. Sea $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ una herradura de Smale lineal. Probar que el itinerario h es un homeomorfismo del conjunto Λ , invariante maximal de T en $Q = [0, 1]^2$ a $2^{\mathbb{Z}}$. Probar que h es una conjugación entre $T|_{\Lambda}$ y el shift $\sigma : 2^{\mathbb{Z}} \mapsto 2^{\mathbb{Z}}$, es decir: $\sigma \circ h(x) = h \circ T(x) \quad \forall x \in \Lambda$.

Definición 8.10.8. Medida de Bernoulli para la herradura de Smale.

Dado un vector de probabilidad $p = (p_1, p_2)$ (es decir: $p_1 + p_2 = 1$, $0 < p_1, p_2 < 1$), se llama *medida de Bernoulli* m_p para la herradura de Smale lineal T , a la medida:

$$m_p(A) = \mu_p(h(A \cap \Lambda))$$

donde Λ es el conjunto invariante maximal de $Q = [0, 1]^2 \subset \mathbb{R}^2$, $h : \Lambda \mapsto 2^{\mathbb{Z}}$ es la función itinerario, y μ_p es la medida de Bernoulli para el shift en $2^{\mathbb{Z}}$ con vector de probabilidad p .

Definición 8.10.9. Herradura de Smale topológica.

Sea $f : X \mapsto X$ un homeomorfismo en un espacio topológico X . Se dice que f presenta una *herradura de Smale topológica* (f, Q) , si existe $Q \subset X$ homeomorfo al cuadrado $[0, 1]^2 \subset \mathbb{R}^2$ tal que f restringida a Q es conjugada con la herradura de Smale lineal T . Esto es: existe un homeomorfismo h tal que $h \circ f(x) = T \circ h(x)$ para todo $x \in Q$.

Nota: El dominio del homeomorfismo h es $f(Q) \cup Q$ y su imagen en \mathbb{R}^2 es $T([0, 1]^2) \cup [0, 1]^2$

Nota: El espacio topológico X no puede ser cualquiera, ya que debe tener subconjuntos homeomorfos al cuadrado de \mathbb{R}^2 . Puede ser por ejemplo una variedad diferenciable de dimensión 2.

Teorema 8.10.10. *Sea $f : X \mapsto X$ un homeomorfismo en un espacio topológico X que presenta una herradura de Smale topológica (f, Q) . Entonces f es una transformación de Bernoulli para ciertas medidas de probabilidad μ (llamadas medidas de Bernoulli para f).*

Demostración: Sea $\Lambda = \bigcap_{j \in \mathbb{Z}} f^{-j}(Q)$. Toda herradura de Smale topológica es, por definición, conjugada a la herradura de Smale lineal, y ésta es conjugada por la transformación itinerario con el shift $\sigma : 2^{\mathbb{Z}} \mapsto 2^{\mathbb{Z}}$. Entonces existe un homeomorfismo conjugación $h : \Lambda \mapsto 2^{\mathbb{Z}}$ tal que $\sigma \circ h = h \circ f|_{\Lambda}$. Sea m_p una medida de Bernoulli para el shift. Sea para todo boreliano $A \subset X$, la medida $\mu(A)$ definida como $\mu(A) = m_p(h(A \cap \Lambda))$. Solo falta ver que $h : \Lambda \mapsto 2^{\mathbb{Z}}$ puede extenderse a un $h : X \mapsto 2^{\mathbb{Z}}$ bimedible c.t.p. y que lleva la medida μ a m_p .

Siendo $h : \Lambda \mapsto 2^{\mathbb{Z}}$ un homeomorfismo, es invertible con inversa medible. Por construcción de μ , se tiene que $\mu(X \setminus \Lambda) = 0$; luego la extensión medible $h : X \mapsto 2^{\mathbb{Z}}$ que define $h(x) = cte$ para $x \in X \setminus \Lambda$ es bimedible casi todo punto, y por construcción de la medida μ lleva esta a m_p . \square

Recordamos que el soporte compacto de una medida de probabilidad de Borel en un espacio topológico compacto, es el mínimo compacto con medida 1.

Corolario 8.10.11. *La herradura de Smale topológica (f, μ) es ergódica y mixing respecto a la medida de Bernoulli μ . Además, restringida al soporte compacto de μ , la herradura de Smale es transitiva y topológicamente mixing.*

Demostración: La ergodicidad y propiedad mixing son consecuencia de que (T, μ) es una transformación de Bernoulli. La transitividad y propiedad mixing topológica se obtienen porque μ es positiva en todos los abiertos de Λ . \square

Capítulo 9

Entropía métrica y topológica

En este capítulo veremos los conceptos de entropía métrica y topológica. Comenzaremos con la entropía métrica. Su definición es debida a Kolmogorov [Ko 1958] y a Sinai [Si 1959]. El estudio de sus propiedades puede encontrarse en muchos textos de teoría ergódica, por ejemplo en [Wa 2000, §4.4], [Mañ 1983a, Capítulo 4], [Kat-Has 1995, pag. 168-170], [Si 2007], [Si 1994, pag. 55-76], [Jo 2005, §4.1-4.2], [Ke 1998, Chapter 3], o [Ki 2011]. En la exposición de propiedades, seguiremos esencialmente el libro [Mañ 1983a] (o también su versión en inglés [Mañ 1987]), y el libro [Kat-Has 1995].

Sea (X, \mathcal{A}, μ) un espacio de probabilidad y sea $T : X \mapsto X$ una transformación medible.

Cantidad de información

Veamos primero una introducción intuitiva de lo que se entiende por cantidad de información.

Sea $A \in \mathcal{A}$ tal que $\mu(A) > 0$. La cantidad de información i de la sentencia “ $x \in A$ ” es, dicho en lenguaje corriente, el espacio que merecería ocupar la noticia “ $x \in A$ ” en un informativo.

Pediremos lo siguiente:

1. La cantidad de información es nula si A ocurre casi siempre, es mayor cuanto menos probable es A y siempre es un número real ≥ 0 , que depende solo de la probabilidad $\mu(A)$ de A . Es decir,

$$i : (0, 1] \mapsto \mathbb{R}^+ \quad i(p) \text{ decreciente con } p \quad i(1) = 0$$

2. Si dos sucesos A y B son independientes, entonces la cantidad de información de la sentencia conjunta “ $x \in A$ y $x \in B$ ” es igual a la suma de las cantidades de información de “ $x \in A$ ” y de “ $x \in B$ ”. Es decir,

$$i(pq) = i(p) + i(q) \quad \text{para todos } p \text{ y } q \in (0, 1]$$

Derivando la condición 2 respecto a p , y multiplicando por p , se obtiene $pqi'(pq) = pi'(p)$, para todos p y q en $(0, 1]$. Luego $pi'(p) = -k$ constante para todo p en $(0, 1]$. Integrando esta ecuación, y considerando la condición (1) se obtiene

$$i(p) = -k \log p \quad \text{para } k \text{ constante } > 0$$

Elegir diferentes valores para la constante k , equivale a elegir diferentes unidades para medir la cantidad de información, y equivale también a elegir diferentes bases para el logaritmo. En Teoría Ergódica se utiliza $k = 1$. En Ingeniería se utiliza $k = 1/\log 2$, o lo que es lo mismo, se calculan los logaritmos en base 2. La unidad de medida así obtenida se llama “*bit*”.

Definición 9.0.12. Cantidad de información

La cantidad de información de un suceso con probabilidad $p > 0$ es

$$i(p) = -\log p$$

La cantidad de *bits* es $i(p)/\log 2 = -\log_2(p)$.

La palabra *bit* proviene del siguiente ejemplo:

Ejemplo: Sea el shift de Bernoulli de dos símbolos equiprobables $\Sigma(1/2, 1/2)$. Sea A el cilindro que se obtiene al fijar los primeros N símbolos:

$$A = C(a_1, a_2, \dots, a_N).$$

La probabilidad de A es $\mu(A) = \frac{1}{2^N}$. La cantidad de información de la sentencia “ $x \in A$ ” es

$$i\left(\frac{1}{2^N}\right) = -\log\left(\frac{1}{2^N}\right) = N \log 2 = N \text{ bits}$$

Sean A y B dos sucesos con probabilidad positiva. La probabilidad condicional de B dado A se define como

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

La cantidad de información del suceso B dado A , se define como

$$i(\mu(B|A)) = -\log \mu(B|A) = -\log \mu(A \cap B) + \log \mu(A)$$

Luego

$$i(\mu(A \cap B)) = i(\mu(A)) + i(\mu(B|A))$$

Interpretación: El espacio necesario para transmitir la información “ $x \in A$ y $x \in B$ ” es la suma del espacio necesario para transmitir “ $x \in A$ ”, más el necesario para transmitir “ $x \in B$ dado que ya se sabe $x \in A$ ”.

9.1. Entropía de particiones

Definición 9.1.1. Partición

Una *partición* \mathcal{P} de (X, \mathcal{A}, μ) es una colección finita o infinita numerable de conjuntos $P \in \mathcal{A}$, dos a dos disjuntos, y medibles tales que $\mu(\bigcup P) = 1$.

Los conjuntos P de \mathcal{P} se llaman *átomos* o *piezas* o *pedazos* de la partición. No es restrictivo asumir que todos tienen medida positiva.

Se llama *entropía* de la partición \mathcal{P} al promedio ponderado con la medida de cada átomo, de la cantidad de información de los átomos de \mathcal{P} . Es decir:

Definición 9.1.2. Entropía de una partición

Entropía de la partición \mathcal{P} con respecto a la probabilidad μ es

$$H(\mathcal{P}, \mu) = - \sum_{P \in \mathcal{P}} \mu(P) \log \mu(P)$$

(Usamos la convención $0 \cdot \log 0 = 0$). Obsérvese que $H(\mathcal{P}, \mu)$ es un número mayor o igual que cero, o infinito. Además $H(\mathcal{P}, \mu) = 0$ si y solo si \mathcal{P} tiene un único átomo con medida positiva; luego, con medida uno.

Por simplicidad, cuando queda claro del contexto cuál es la probabilidad μ , o cuando es la misma probabilidad que se considera para diferentes particiones \mathcal{P} , escribimos simplemente

$$H(\mathcal{P})$$

para denotar a la entropía de la partición \mathcal{P} con respecto a la medida μ .

La entropía de la partición es el valor esperado, (es decir la integral respecto a μ), de la función que a casi todo punto $x \in X$ le asigna la cantidad de información del átomo $P_x \in \mathcal{P}$ que contiene a x . Más precisamente:

$$f(x) = -\log \mu(P_x), \quad H(\mathcal{P}) = \int f d\mu$$

Proposición 9.1.3. Cota para la entropía de partición finita

Sea $\mathcal{P} = \{P_1, P_2, \dots, P_k\}$ una partición finita formada por $k \geq 1$ conjuntos medibles. Entonces, para cualquier medida de probabilidad μ se cumple

$$H(\mathcal{P}, \mu) \leq \log k,$$

y la igualdad se verifica si y solo si $\mu(P_i) = 1/k$ para todo $i \in \{1, \dots, k\}$.

Demostración. Denotemos $x_i = \mu(P_i)$ para todo $1 \leq i \leq k$. Se cumple $\sum_{i=1}^k x_i = 1$. Por definición de entropía

$$H(\mathcal{P}, \mu) = - \sum_{i=1}^k x_i \log x_i.$$

Luego, para demostrar la proposición, basta probar que la siguiente función F de k variables reales x_i en el intervalo $[0, 1] \ni x_i$:

$$F(x_1, x_2, \dots, x_k) := - \sum_{i=1}^k x_i \log x_i,$$

tiene un único máximo condicionado a

$$G(x_1, x_2, \dots, x_k) := \sum_{i=1}^k x_i - 1 = 0,$$

que ese máximo es igual a $\log k$ y que se alcanza solamente en el punto $(x_1, \dots, x_k) = (1/k, \dots, 1/k)$.

En efecto, F es no negativa y toma valores positivos en el interior del cubo compacto $[0, 1]^k$. Su máximo condicionado a $G = 0$ se encuentra por el método de Lagrange, con la siguiente función de Lagrange L con un coeficiente indeterminado λ :

$$L := F + \lambda G = - \sum_{i=1}^k x_i \log x_i + \lambda(-1 + \sum_{i=1}^k x_i).$$

Luego los candidatos a lugar del máximo condicionado y al valor de coeficiente λ , son las soluciones del siguiente sistema de $k + 1$ ecuaciones:

$$\frac{\partial L}{\partial x_i} = 0 \quad \forall 1 \leq i \leq k, \quad \frac{\partial L}{\partial \lambda} = 0.$$

Calculando las derivadas obtenemos:

$$-\log x_i - 1 + \lambda = 0 \quad \forall 1 \leq i \leq k, \quad \sum_{i=1}^k x_i = 1.$$

Entonces

$$\lambda = \log x_i + 1 \quad \forall i \in 1, \dots, k,$$

de donde $x_i = a$ constante para todo i . Luego $1 = \sum_{i=1}^k x_i = ka$, de donde $x_i = a = 1/k$ para todo i . Hemos encontrado un solo punto donde se alcanza el máximo, que corresponde a $\mu(P_i) = 1/k$ para todo i . El máximo es entonces

$$F(1/k, \dots, 1/k) = - \sum_{i=1}^k (1/k) \log(1/k) = \log k,$$

como queríamos demostrar. \square

Definición 9.1.4. Producto de particiones

Dadas dos particiones \mathcal{P} y \mathcal{Q} se define la *partición producto* $\mathcal{P} \vee \mathcal{Q}$ como la formada por todas las intersecciones $P \cap Q$, con medida no nula, de $P \in \mathcal{P}$ y $Q \in \mathcal{Q}$.

Notamos que el producto de particiones es una operación conmutativa, asociativa y que tiene como neutro la partición $\{X\}$ (o también cualquier partición con entropía nula).

Se llama entropía condicional de la partición \mathcal{Q} dada la partición \mathcal{P} al promedio, ponderado con la probabilidad $\mu(P \cap Q)$, de la cantidad de información de Q dado P , para $P \in \mathcal{P}$ y $Q \in \mathcal{Q}$, tales que $\mu(P \cap Q) \neq 0$. Es decir:

Definición 9.1.5. Entropía condicional

La entropía condicional de \mathcal{Q} dada \mathcal{P} es

$$H(\mathcal{Q}|\mathcal{P}) = - \sum_{P \in \mathcal{P}, Q \in \mathcal{Q}, \mu(P \cap Q) \neq 0} \mu(P \cap Q) \log \frac{\mu(P \cap Q)}{\mu(P)}$$

Nótese que es siempre mayor o igual que cero, o infinito.

Para simplificar la notación convenimos en que la letra imprenta indica un átomo de la partición denotada con la misma letra, pero caligráfica. También convenimos en que, si aparece indicado $0 \cdot \log 0$, entonces el término correspondiente en la suma es nulo. Con esta notación se tiene:

$$H(\mathcal{Q}|\mathcal{P}) = - \sum_P \sum_Q \mu(P \cap Q) \log \frac{\mu(P \cap Q)}{\mu(P)} \quad (9.1)$$

$$H(\mathcal{Q} \vee \mathcal{P}) = - \sum_P \sum_Q \mu(P \cap Q) \log \mu(P \cap Q) \quad (9.2)$$

Proposición 9.1.6. Sean \mathcal{P} , \mathcal{Q} y \mathcal{R} particiones. Se cumple:

1. $H(\mathcal{Q}|\mathcal{P}) \leq H(\mathcal{Q})$
2. $H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q}|\mathcal{P})$
3. $H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) \leq H(\mathcal{Q}|\mathcal{R})$

Para probar la proposición anterior precisaremos del siguiente lema:

Lema 9.1.7. Desigualdad de Jensen

Sea $\phi : [0, 1] \mapsto \mathbb{R}$ una función continua, derivable hasta segundo orden en $(0, 1)$ y con derivada segunda negativa.

Sea λ_i una sucesión de números reales ≥ 0 , tales que $\sum_{i=1}^{\infty} \lambda_i = 1$, y sea x_i una sucesión de puntos en $[0, 1]$.

Se cumple:

$$\phi \left(\sum_{i=1}^{\infty} \lambda_i x_i \right) \geq \sum_{i=1}^{\infty} \lambda_i \phi(x_i)$$

Demostración: Las series en el enunciado son absolutamente convergentes porque sus términos i -ésimos son menores o iguales, en valor absoluto, que una constante por λ_i , y la serie de los λ_i converge por hipótesis.

Probemos primero que, para todo $n \geq 1$, si $\sum_{i=1}^n \rho_i = 1$, con ρ_i dados ≥ 0 , entonces se cumple:

$$\phi \left(\sum_{i=1}^n \rho_i x_i \right) \geq \sum_{i=1}^n \rho_i \phi(x_i) \quad (9.3)$$

Para $n = 1$ es trivial, porque, en ese caso, $\rho_1 = 1$.

Para $n = 2$: sean x_1 y $x_2 \in [0,1]$. Como la concavidad de la gráfica de la función ϕ es negativa, porque la derivada segunda es negativa, entonces la gráfica entre x_1 y x_2 está por arriba de la cuerda, es decir, del segmento que tiene por extremos los puntos $(x_1, \phi(x_1))$ y $(x_2, \phi(x_2))$. Esto significa que, para todo $0 \leq \rho \leq 1$ se cumple:

$$\phi(\rho x_1 + (1 - \rho)x_2) \geq \rho \phi(x_1) + (1 - \rho)\phi(x_2) \quad (9.4)$$

Por lo tanto se sumple la desigualdad (9.3) cuando $n = 2$.

Supongamos cierta la desigualdad (9.3) para n . Demostrémosla para $n + 1$. Se tiene $\sum_{i=1}^{n+1} \rho_i = 1$. Sea $\rho = \sum_{i=1}^n \rho_i$. Entonces se tiene $\rho_{n+1} = 1 - \rho$. Sea $X = \sum_{i=1}^n \rho_i x_i / \rho$. Se cumple

$$\sum_{i=1}^{n+1} \rho_i x_i = \rho X + (1 - \rho)x_{n+1}$$

Aplicando la desigualdad (9.4), y luego la hipótesis de inducción, se tiene

$$\begin{aligned} \phi \left(\sum_{i=1}^{n+1} \rho_i x_i \right) &= \phi(\rho X + (1 - \rho)x_{n+1}) \geq \rho \phi(X) + (1 - \rho)\phi(x_{n+1}) = \\ &= \rho \phi \left(\sum_{i=1}^n \frac{\rho_i x_i}{\rho} \right) + \rho_{n+1} \phi(x_{n+1}) \geq \\ &\geq \rho \sum_{i=1}^n \frac{\rho_i \phi(x_i)}{\rho} + \rho_{n+1} \phi(x_{n+1}) = \sum_{i=1}^{n+1} \rho_i \phi(x_i). \end{aligned}$$

Por lo tanto vale (9.3) para todo n natural.

Ahora terminemos la demostración. Como las series son convergentes, son el límite de sus reducidas n -ésimas. Siendo ϕ continua:

$$\begin{aligned} \phi \left(\sum_{i=1}^{\infty} \lambda_i x_i \right) &= \lim_{n \rightarrow \infty} \phi \left(\frac{\sum_{i=1}^n \lambda_i x_i}{\sum_{i=1}^n \lambda_i} \right) \geq \\ &\geq \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \lambda_i \phi(x_i)}{\sum_{i=1}^n \lambda_i} = \sum_{i=1}^{\infty} \lambda_i \phi(x_i). \quad \square \end{aligned}$$

Ahora estamos en condiciones de demostrar la proposición 9.1.6.

Demostración de la proposición 9.1.6:

La afirmación 1 es consecuencia inmediata de la afirmación 3 tomando $\mathcal{R} = \{X\}$.
Veamos la afirmación 2:

$$\begin{aligned} H(\mathcal{Q} \vee \mathcal{P}) &= - \sum_P \sum_Q \mu(P \cap Q) \log \mu(P \cap Q) \\ &= - \sum_P \sum_Q \mu(P \cap Q) \log \frac{\mu(P \cap Q)}{\mu(P)} - \sum_P \sum_Q \mu(P \cap Q) \log \mu(P) = \\ &= H(\mathcal{Q}|\mathcal{P}) - \sum_P \mu(P) \log \mu(P) = H(\mathcal{Q}|\mathcal{P}) + H(\mathcal{P}) \end{aligned}$$

Ahora probemos la afirmación 3: Sea

$$\phi(x) = -x \log x \text{ si } x \in (0, 1] \text{ y } \phi(0) = 0$$

Es una función continua en $[0,1]$, y tiene derivada segunda negativa en $(0,1)$. Luego, ϕ cumple la desigualdad de Jensen, probada en el lema anterior.

Usando como convenio de notación que son nulos los términos denotados como $0 \log 0$, o $0 \log(0/0)$, se tiene:

$$\begin{aligned} H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) &= - \sum_R \sum_Q \sum_P \mu(P \cap Q \cap R) \log \frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)} \\ &= \sum_R \mu(R) \sum_Q \sum_P \frac{\mu(P \cap R)}{\mu(R)} \phi \left(\frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)} \right) \end{aligned}$$

Tomemos $Q \in \mathcal{Q}$ y $R \in \mathcal{R}$ fijos. Enumeremos con subíndice $i \geq 1$ todos los conjuntos $P_i \in \mathcal{P}$, tales que $\mu(P_i \cap R) \neq 0$. Definamos los números $\lambda_i = \mu(P_i \cap R)/\mu(R)$ y $x_i = \mu(P_i \cap Q \cap R)/\mu(P_i \cap R)$.

Se cumple que $x_i \in [0,1]$, que $\lambda_i \geq 0$ y que $\sum_{i=1}^{\infty} \lambda_i = 1$. Luego, aplicando la desigualdad de Jensen, se tiene:

$$H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) = \sum_R \mu(R) \sum_Q \sum_{i \geq 1} \lambda_i \phi(x_i) \leq \sum_R \mu(R) \sum_Q \phi \left(\sum_{i \geq 1} \lambda_i x_i \right)$$

Observando que $\sum_{i \geq 1} \lambda_i x_i = \mu(Q \cap R)/\mu(R)$ se concluye:

$$\begin{aligned} H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) &\leq \\ &- \sum_R \mu(R) \sum_Q \frac{\mu(Q \cap R)}{\mu(R)} \log \left(\frac{\mu(Q \cap R)}{\mu(R)} \right) = H(\mathcal{Q}|\mathcal{R}) \end{aligned}$$

□

Corolario 9.1.8. Entropía del producto de particiones

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

Demostración: Basta recordar que $H(\mathcal{Q}|\mathcal{P}) \geq 0$, y usar la igualdad 2) y la desigualdad 1) de la proposición 9.1.6. \square

Corolario 9.1.9. Entropía condicional del producto de particiones

Si $H(\mathcal{R}) < \infty$ entonces:

- i) $H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) = H(\mathcal{P}|\mathcal{R}) + H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R})$
- ii) $H(\mathcal{P}|\mathcal{R}) \leq H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) \leq H(\mathcal{P}|\mathcal{R}) + H(\mathcal{Q}|\mathcal{R})$

Demostración: i) Aplicando tres veces la igualdad (2) de 9.1.6:

$$H(\mathcal{R} \vee (\mathcal{P} \vee \mathcal{Q})) = H(\mathcal{R}) + H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R})$$

$$H((\mathcal{P} \vee \mathcal{R}) \vee \mathcal{Q}) = H(\mathcal{P} \vee \mathcal{R}) + H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) = H(\mathcal{R}) + H(\mathcal{P}|\mathcal{R}) + H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R})$$

Como $\mathcal{R} \vee \mathcal{P} \vee \mathcal{Q} = \mathcal{P} \vee \mathcal{R} \vee \mathcal{Q}$, igualando las expresiones de los últimos miembros, se tiene el resultado enunciado.

ii) De i), siendo siempre $H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}) \geq 0$, se tiene

$$H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) \geq H(\mathcal{P}|\mathcal{R}).$$

Aplicando la desigualdad (3) de la proposición 9.1.6, a (i), se deduce

$$H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) \leq H(\mathcal{P}|\mathcal{R}) + H(\mathcal{Q}|\mathcal{R}). \quad \square$$

Nota: Las partes i) y ii) del corolario 9.1.9 son también ciertas aunque $H(\mathcal{R}) = \infty$, pero para probarlas hay que remitirse a las definiciones.

Corolario 9.1.10. Diferencia de entropías

Si $H(\mathcal{P}) < \infty$, entonces, para cualquier otra partición \mathcal{Q} se cumple

$$H(\mathcal{Q}) - H(\mathcal{P}) \leq H(\mathcal{Q}|\mathcal{P})$$

Demostración: Por el corolario 9.1.8 y la igualdad (2) de la proposición 9.1.6, se tiene:

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

Restando $H(\mathcal{P})$ se obtiene la tesis. \square

9.2. Entropía de partición como función de la medida

Ahora vamos a dejar fija la partición \mathcal{P} y estudiar la función $\mu \mapsto H(\mathcal{P}, \mu)$.

Probaremos primero una propiedad de convexidad de la entropía de una partición con respecto a la medida:

Teorema 9.2.1. Convexidad de la entropía

Sea (X, \mathcal{A}) un espacio medible. Sea $k \geq 1$ un número natural. Sea $\{\mu_i\}_{1 \leq i \leq k}$ una k -upla de medidas de probabilidad y sea $\{\lambda_i\}_{1 \leq i \leq k}$ una k -upla de números reales no negativos y no mayores que 1, tales que $\sum_{i=1}^k \lambda_i = 1$. Entonces, para toda partición finita \mathcal{P} se cumple:

$$H(\mathcal{P}, \sum_{i=1}^k \lambda_i \mu_i) \geq \sum_{i=1}^k \lambda_i H(\mathcal{P}, \mu_i).$$

Dicho de otra forma:

La entropía de una partición con respecto a una combinación lineal convexa de probabilidades es mayor o igual que la combinación lineal convexa de las entropías con respecto a cada una de esas probabilidades.

Demostración. Sea la función $\phi : [0, 1] \mapsto \mathbb{R}$ definida por

$$\phi(x) := -x \log x \quad \forall x \in [0, 1] \text{ con la convención } 0 \log 0 := 0.$$

Es inmediato chequear que ϕ es continua en $[0, 1]$, derivable en $(0, 1]$ y que $\phi'(x) = -1 - \log x < 0 \quad \forall x \in (0, 1]$. Luego, por el Lema 9.1.7 de Jensen se cumple

$$\phi\left(\sum_{i=1}^k \lambda_i \mu_i(P)\right) \geq \sum_{i=1}^k \lambda_i \phi(\mu_i(P)) \quad \forall P \in \mathcal{P}.$$

Además, por definición de entropía de la partición \mathcal{P} con respecto a la medida de probabilidad $\sum_{i=1}^k \lambda_i \mu_i$, se cumple

$$H(\mathcal{P}, \sum_{i=1}^k \lambda_i \mu_i) = \sum_{P \in \mathcal{P}} \phi\left(\sum_{i=1}^k \lambda_i \mu_i(P)\right) \geq \sum_{P \in \mathcal{P}} \sum_{i=1}^k \lambda_i \phi(\mu_i(P)).$$

Como la partición \mathcal{P} es finita, tenemos:

$$H(\mathcal{P}, \sum_{i=1}^k \lambda_i \mu_i) \geq \sum_{P \in \mathcal{P}} \sum_{i=1}^k \lambda_i \phi(\mu_i(P)) = \sum_{i=1}^k \lambda_i \sum_{P \in \mathcal{P}} \phi(\mu_i(P)) = \sum_{i=1}^k \lambda_i H(\mathcal{P}, \mu_i),$$

terminando de demostrar el Teorema 9.2.1. □

De ahora en adelante, en esta sección, vamos a considerar que el espacio medible X es un espacio métrico compacto y que la sigma-álgebra \mathcal{A} es la de Borel. Denotamos con \mathcal{M} al conjunto de todas las medidas de probabilidad en (X, \mathcal{A}) dotado de la topología débil*.

Definición 9.2.2. Borde de una partición finita

Sea \mathcal{P} una partición finita. Denotamos con $\partial\mathcal{P}$ al conjunto que llamamos borde de la partición, definido por

$$\partial\mathcal{P} := \bigcup_P \partial P,$$

donde ∂P denota el borde topológico de la pieza P .

Definición 9.2.3. Diámetro de una partición finita Sea \mathcal{P} una partición finita. Denotamos con $\text{diam}(\mathcal{P})$ al número real que llamamos diámetro de la partición, definido por

$$\text{diam}(\mathcal{P}) := \max_P \text{diam}(P) = \max_P \sup_{x, y \in P} \text{dist}(x, y).$$

donde $\text{diam}(P)$ denota el diámetro de la pieza P .

Proposición 9.2.4. Existencia de partición con borde de medida nula

Dado $\epsilon > 0$ y dada $\mu \in \mathcal{M}$, existe una partición finita \mathcal{P} tal que $\mu(\partial\mathcal{P}) = 0$ y $\text{diam}(\mathcal{P}) < \epsilon$.

Demostración. Fijado un punto $x \in X$, probemos que existe $0 < \delta(x) < \epsilon/2$ tal que

$$\mu(\partial(B_{\delta(x)}(x))) = 0,$$

donde $B_{\delta(x)}(x)$ denota la bola de centro x y radio $\delta(x)$. En efecto, si para todo $0 < \delta < \epsilon/2$ el borde $\partial(B_{\delta}(x))$ tuviera μ -medida positiva, la siguiente familia no numerable de compactos (con x fijo)

$$\{\partial(B_{\delta}(x))\}_{0 < \delta < \epsilon/2}$$

estaría formada por compactos disjuntos dos a dos, y todos de medida positiva, lo que es una contradicción.

Ahora cubrimos X con una cantidad finita de tales bolas

$$B_1, \dots, B_k,$$

donde $B_i = B_{\delta(x_i)}(x_i)$ para todo $i \in \{1, \dots, k\}$. Construimos la partición $\mathcal{P} := \{P_1, \dots, P_k\}$ definiendo

$$P_1 := B_1, \quad P_2 := B_2 \setminus B_1, \quad P_3 := B_3 \setminus (B_1 \cup B_2), \quad \dots$$

$$P_i := B_i \setminus (B_1 \cup \dots \cup B_{i-1}), \quad \dots, P_k := B_k \setminus (B_1 \cup \dots \cup B_{k-1}).$$

Por construcción se cumple $\partial P_i \subset \bigcup_{j=1}^k \partial B_j$, $P_i \subset B_i$. Luego

$$0 \leq \mu(\partial P_i) \leq \bigcup_{j=1}^k \mu(\partial B_{\delta(x_j)}(x_j)) = 0$$

y $\text{diam}(P_i) \leq \text{diam} B_i \leq 2\delta(x_i) < \epsilon$. Entonces $\mu(\mathcal{P}) = 0$ y $\text{diam}(\mathcal{P}) < \epsilon$, como queríamos probar. \square

Fijada una partición \mathcal{P} , en general la entropía $H(\mathcal{P}, \mu)$ como función de la medida de probabilidad μ no es continua. Sin embargo, cuando la partición \mathcal{P} es tal que $\mu(\partial\mathcal{P}) = 0$, entonces esta medida μ es punto de continuidad de la función $H(\mathcal{P}, \cdot)$, como demostraremos a continuación:

Teorema 9.2.5. Medidas de continuidad de la entropía de partición

Sea X un espacio métrico compacto con la sigma-álgebra de Borel. Sea μ una medida de probabilidad. Sea \mathcal{P} una partición finita tal que

$$\mu(\partial\mathcal{P}) = 0.$$

Entonces, para toda sucesión μ_n de medidas de probabilidad tal que

$$\lim_{n \rightarrow +\infty} \mu_n = \mu$$

en la topología débil*, se cumple:

$$\lim_{n \rightarrow +\infty} H(\mathcal{P}, \mu_n) = H(\mathcal{P}, \mu).$$

Para probar el Teorema 9.2.5 necesitamos la siguiente proposición, sobre la topología débil* del espacio \mathcal{M} de probabilidades:

Proposición 9.2.6. Medida límite de conjuntos

Sea X un espacio métrico compacto con la sigma-álgebra de Borel \mathcal{A} y sea \mathcal{M} el espacio de medidas de probabilidad en (X, \mathcal{A}) , dotado de la topología débil*. Sea en \mathcal{M} una sucesión de medidas tales que

$$\lim_{n \rightarrow +\infty} \mu_n = \mu.$$

- (a) Si $K \subset X$ es compacto entonces $\limsup_{n \rightarrow +\infty} \mu_n(K) \leq \mu(K)$.
 (b) Si $V \subset X$ es abierto entonces $\liminf_{n \rightarrow +\infty} \mu_n(V) \geq \mu(V)$.
 (c) Si $A \in \mathcal{A}$ cumple $\mu(\partial A) = 0$ entonces: $\lim_{n \rightarrow +\infty} \mu_n(A) = \mu(A)$.

Demostración. (a) Sea $\epsilon > 0$ y $V \supset K$ tal que $\mu(V \setminus K) < \epsilon$. Sea $\psi : X \mapsto [0, 1]$ continua tal que $\psi|_K = 1$ y $\psi|_{X \setminus V} = 0$. Entonces

$$\mu_n(K) \leq \int \psi d\mu_n \rightarrow \int \psi, d\mu \leq \mu(V) = \mu(K) + \mu(V \setminus K) < \mu(K) + \epsilon.$$

Concluimos que

$$\limsup_{n \rightarrow +\infty} \mu_n(K) \leq \mu(K) + \epsilon.$$

Como esta desigualdad vale para todo $\epsilon > 0$ implica (a).

(b) Dado V abierto, consideramos $K = X \setminus V$. Tenemos

$$\mu_n(V) = 1 - \mu_n(K), \quad \liminf_{n \rightarrow +\infty} \mu_n(V) = 1 - \limsup_{n \rightarrow +\infty} \mu_n(K).$$

Como K es compacto, podemos aplicar lo probado en la parte (a), y deducimos:

$$\liminf_{n \rightarrow +\infty} \mu_n(V) = 1 - \limsup_{n \rightarrow +\infty} \mu_n(K) \geq 1 - \mu(K) = \mu(V),$$

terminando de probar la parte (b).

(c) Tenemos:

$$\text{int}(A) = A \setminus \partial A \subset A \subset A \cup \partial A = \bar{A}.$$

Por hipótesis $\mu(\partial A) = 0$; entonces:

$$\mu(\text{int}(A)) = \mu(A \setminus \partial A) = \mu(A) - \mu(A \cup \partial A) = \mu(A)$$

$$\mu(\bar{A}) = \mu(A \cup \partial A) = \mu(A) + \mu(\partial A \setminus A) = \mu(A).$$

Además

$$\mu_n(\text{int}(A)) \leq \mu_n(A) \leq \mu_n(\bar{A}).$$

Como \bar{A} es compacto e $\text{int}(A)$ es abierto, podemos aplicar a ellos las propiedades (a) y (b). Deducimos:

$$\mu(A) = \mu(\text{int}(A)) \leq \liminf_{n \rightarrow +\infty} \mu_n(\text{int}(A)) \leq \liminf_{n \rightarrow +\infty} \mu_n(A) \leq$$

$$\leq \limsup_{n \rightarrow +\infty} \mu_n(A) \leq \limsup_{n \rightarrow +\infty} \mu_n(\bar{A}) \leq \mu(\bar{A}) = \mu(A),$$

de donde se concluye que las desigualdades son todas igualdades, terminando de demostrar (c). \square

Ahora estamos en condiciones de probar el Teorema 9.2.5:

Demostración. del Teorema 9.2.5: Sea $\phi : [0, 1] \rightarrow \mathbb{R}$ la función no negativa definida por

$$\phi(x) := -x \log x \quad \forall x \in [0, 1], \quad \text{donde por convención } 0 \log 0 := 0$$

Es inmediato chequear que ϕ es continua. Por definición de entropía de la partición finita $\mathcal{P} = \{P_1, \dots, P_k\}$ tenemos

$$H(\mathcal{P}, \mu_n) = \sum_{i=1}^k \phi(\mu_n(P_i)).$$

Por hipótesis $\mu(\partial P_i) = 0$ para todo $1 \leq i \leq k$. Entonces, aplicando la Proposición 9.2.6 (c), se cumple: $\lim_{n \rightarrow +\infty} \mu_n(P_i) = \mu(P_i)$, y como ϕ es continua, tenemos

$$\lim_{n \rightarrow +\infty} \phi(\mu_n(P_i)) = \phi(\mu(P_i)).$$

Finalmente, sumando en i , concluimos

$$\begin{aligned} \lim_{n \rightarrow +\infty} H(\mathcal{P}, \mu_n) &= \lim_{n \rightarrow +\infty} \sum_{i=1}^k \phi(\mu_n(P_i)) = \\ &= \sum_{i=1}^k \lim_{n \rightarrow +\infty} \phi(\mu_n(P_i)) = \sum_{i=1}^k \phi(\mu(P_i)) = H(\mathcal{P}, \mu), \end{aligned}$$

terminando de probar el Teorema 9.2.5. \square

9.3. Refinamiento de particiones

Definición 9.3.1. Refinamiento de particiones

Sean \mathcal{P} y \mathcal{Q} dos particiones. Decimos que \mathcal{Q} es más fina que \mathcal{P} , o \mathcal{P} más gruesa que \mathcal{Q} , y escribimos

$$\mathcal{Q} \prec \mathcal{P}$$

si todo átomo de \mathcal{Q} está contenido μ -c.t.p. en un átomo de \mathcal{P} . Esto es equivalente a decir que todo átomo de \mathcal{P} coincide μ -c.t.p. con una unión de átomos de \mathcal{Q} .

El producto de dos particiones es, obviamente, más fino que cada una de las particiones.

Nótese que \mathcal{Q} es más fina que \mathcal{P} si y solo si $\mathcal{P} \vee \mathcal{Q}$ coincide μ -c.t.p. con \mathcal{Q} . Entonces en todo cálculo de entropías, que solo depende de la medida de los átomos, donde aparezca la partición \mathcal{Q} , podemos poner $\mathcal{P} \vee \mathcal{Q}$. Tenemos $H(\mathcal{Q} \vee \mathcal{P}) = H(\mathcal{Q})$. Luego, aplicando la propiedad 2) de la proposición 9.1.6,

$$H(\mathcal{Q}) = H(\mathcal{Q} \vee \mathcal{P}) = H(\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q})$$

se concluye que si \mathcal{Q} tiene entropía finita y es más fina que \mathcal{P} , entonces $H(\mathcal{P}|\mathcal{Q}) = 0$. (Aplicando la definición directamente, se puede demostrar que el resultado también vale aunque $H(\mathcal{Q}) = \infty$).

Veamos ahora que la entropía crece cuanto más fina es la partición. Si se condiciona, la entropía crece cuanto más fina es la partición condicionada, y decrece cuanto más fina es la partición condicionante:

Proposición 9.3.2. Sean $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ tres particiones. Si la partición \mathcal{Q} es más fina que \mathcal{P} , entonces:

1. $H(\mathcal{Q}) \geq H(\mathcal{P})$
2. $H(\mathcal{Q}|\mathcal{R}) \geq H(\mathcal{P}|\mathcal{R})$
3. $H(\mathcal{R}|\mathcal{Q}) \leq H(\mathcal{R}|\mathcal{P})$

Demostración: Para demostrar 1) apliquemos el corolario 9.1.8:

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

Para demostrar 2) usemos la desigualdad ii) del corolario 9.1.9:

$$H(\mathcal{Q}|\mathcal{R}) = H(\mathcal{Q} \vee \mathcal{P}|\mathcal{R}) \geq H(\mathcal{P}|\mathcal{R})$$

Para demostrar 3) usemos la parte 3) de la proposición 9.1.6:

$$H(\mathcal{R}|\mathcal{Q}) = H(\mathcal{R}|\mathcal{P} \vee \mathcal{Q}) \leq H(\mathcal{R}|\mathcal{P})$$

□

Definición 9.3.3. Sea $\{\mathcal{P}_n\}_{n \geq 1}$ una sucesión de particiones. Indicamos con

$$\bigvee_{n \geq 1} \mathcal{P}_n$$

a la mínima σ -álgebra que contiene a todos los átomos de todas las particiones de la sucesión. Es decir contiene a $\bigcup_{n \geq 1} \mathcal{P}_n$.

Por lo tanto, una partición \mathcal{P} está contenida en $\bigvee_{n \geq 1} \mathcal{P}_n$, si todos sus átomos pertenecen a esa σ -álgebra.

Decimos que una partición *está contenida* μ -c.t.p. en $\bigvee_{n \geq 1} \mathcal{P}_n$, si todos sus átomos son iguales μ -c.t.p. a algún subconjunto miembro de esa σ -álgebra.

Ahora probemos un lema para demostrar el teorema de Kolmogorov-Sinai que veremos en las próximas secciones:

Lema 9.3.4. *Sea $\{\mathcal{P}_n\}_{n \geq 1}$ una sucesión de particiones tales que, para todo $n \geq 1$ la partición \mathcal{P}_{n+1} es mas fina que \mathcal{P}_n .*

Sea \mathcal{P} una partición finita contenida μ -c.t.p. en $\bigvee_{n \geq 1} \mathcal{P}_n$.

Entonces

$$\lim_{n \rightarrow \infty} H(\mathcal{P} | \mathcal{P}_n) = 0$$

Demostración: Todas las particiones consideradas lo son del espacio de probabilidad (X, \mathcal{A}, μ) .

Definamos la familia $\mathcal{F} = \{A \in \mathcal{A} : \text{Para todo } \epsilon > 0 \text{ existen } N \geq 1 \text{ y alguna unión } B \text{ de átomos de } \mathcal{P}_N \text{ que cumple } \mu(A \Delta B) < \epsilon\}$.

Nótese que, si $A \in \mathcal{F}$, entonces, para todo $n \geq N$, existe alguna unión $B^{(n)}$ de átomos de \mathcal{P}_n que cumple $\mu(A \Delta B^{(n)}) < \epsilon$. Esto se debe a que la partición \mathcal{P}_n es mas fina que \mathcal{P}_N . Por lo tanto, si N es adecuado, cualquier natural mayor que N también lo es.

Es inmediato ver que la familia \mathcal{F} contiene a \mathcal{P}_N , para todo $N \geq 1$.

Afirmamos que \mathcal{F} es una σ -álgebra. En efecto, para probar que el complemento de A está en \mathcal{F} , basta usar los mismos N y B que para A , y comparar A^c con la unión de los átomos de la partición \mathcal{P}_N que no están en B .

Ahora, probemos que la unión numerable de conjuntos $A_i \in \mathcal{F}$ también pertenece a \mathcal{F} :

Sea $\epsilon > 0$ dado. Para cada $i \geq 1$ sean $N_i \geq 1$ y B_i , unión de átomos de \mathcal{P}_{N_i} , que cumple $\mu(A_i \Delta B_i) < \epsilon/2^{i+1}$. Sea $B = \bigcup_{i \geq 1} B_i$. Se tiene:

$$\mu(A \Delta B) \leq \sum_{i=1}^{\infty} \mu(A_i \Delta B_i) < \sum_{i=1}^{\infty} \epsilon/2^{i+1} = \epsilon/2 \quad (9.5)$$

Para todo $k \geq 1$ defínase $C_k = \bigcup_{i=1}^k B_i$. Se cumple $C_1 \subset C_2 \subset \dots \subset C_k \subset \dots \bigcup_{i \geq 1} C_i = B$. Luego $\mu(B) = \lim \mu(C_k)$, y existe $K \geq 1$ tal que $\mu(B) - \mu(C_K) < \epsilon/2$. Como $C_K \subset B$, se tiene:

$$\mu(B \Delta C_K) < \epsilon/2 \quad (9.6)$$

Sea $N = \max_{1 \leq i \leq K} N_i$. Como B_i es una unión de átomos de \mathcal{P}_{N_i} , coincide μ -c.t.p. con una unión de átomos de \mathcal{P}_N . Luego $C_K = \bigcup_{i=1}^K B_i$ también coincide μ -c.t.p. con una unión H de átomos de \mathcal{P}_N .

Por (9.5) y (9.6):

$$\mu(A \Delta H) = \mu(A \Delta C_K) \leq \mu(A \Delta B) + \mu(B \Delta C_K) < \epsilon$$

Luego $A \in \mathcal{F}$ y \mathcal{F} es una σ -álgebra.

Por lo tanto la familia \mathcal{F} contiene a la σ -álgebra $\bigvee_{n \geq 1} \mathcal{P}_n$. Esta, por hipótesis, contiene μ -c.t.p. a la partición \mathcal{P} . Si $A \in \mathcal{F}$, entonces cualquier otro conjunto que coincida μ -c.t.p. con A también pertenece a \mathcal{F} . Luego, todos los átomos de \mathcal{P} pertenecen a la familia \mathcal{F} .

Siendo \mathcal{P} una partición finita, se tiene $\mathcal{P} = \{P_1, P_2, \dots, P_r\}$, con $P_i \in \mathcal{F}$, para todo $i = 1, 2, \dots, r$. Dado ϵ existe $N = N(\epsilon) \geq 1$ (el mismo para todo $i = 1, 2, \dots, r$) y existe, para cada $i = 1, 2, \dots, r$, alguna unión Q_i de átomos de \mathcal{P}_N , tal que

$$\mu(P_i \Delta Q_i) < \epsilon$$

Los conjuntos Q_i no forman necesariamente una partición, porque pueden intersectarse, aunque los P_i no lo hacen. Pero aproximan a los átomos de la partición \mathcal{P} y son uniones de átomos de \mathcal{P}_N .

Construyamos, usando los Q_i , una partición $\mathcal{S} = \mathcal{S}(\epsilon)$ cuyos átomos continúen aproximando a los P_i y sigan siendo μ -c.t.p. uniones de átomos de \mathcal{P}_N . Entonces \mathcal{S} será mas gruesa que \mathcal{P}_N , y por la proposición 9.3.2 se cumple:

$$0 \leq H(\mathcal{P}|\mathcal{P}_n) \leq H(\mathcal{P}|\mathcal{P}_N) \leq H(\mathcal{P}|\mathcal{S}(\epsilon)) \quad \forall n \geq N = N(\epsilon)$$

Para demostrar la tesis

$$H(\mathcal{P}|\mathcal{P}_N) \rightarrow 0,$$

bastará entonces probar que $H(\mathcal{P}|\mathcal{S}(\epsilon))_{\epsilon \rightarrow \infty} \rightarrow 0$

Pasemos a la construcción de $\mathcal{S}(\epsilon)$, definiendo

$$S_1 = Q_1, \quad S_i = Q_i - \bigcup_{j=1}^{i-1} Q_j, \quad \text{para } 2 \leq i \leq r-1, \quad S_r = \left(\bigcup_{j=1}^{r-1} Q_j \right)^c$$

$$\mathcal{S}(\epsilon) = \{S_i : 1 \leq i \leq r, \quad \mu(S_i) \neq 0\}$$

Por construcción $\mathcal{S}(\epsilon)$ es una partición, y sus átomos son μ -c.t.p. uniones de átomos de \mathcal{P}_N , porque los Q_i lo eran.

Afirmamos que para todo $i = 1, 2, \dots, r$ se cumple

$$\mu(P_i \Delta S_i) < r\epsilon$$

En efecto, para $i = 1$: $S_1 = Q_1$ y $\mu(P_1 \Delta Q_1) < \epsilon$.

Para $2 \leq i \leq r-1$, si $1 \leq j \leq i-1$ se tiene:

$$S_i \subset Q_i, \quad S_i^c = Q_i^c \bigcup_{j=1}^{i-1} Q_j, \quad P_i \subset P_j^c$$

Luego

$$\begin{aligned} \mu(P_i \Delta S_i) &= \mu(P_i^c \cap S_i) + \mu(P_i \cap S_i^c) \leq \\ &\leq \mu(P_i^c \cap Q_i) + \mu(P_i \cap Q_i^c) + \sum_{j=1}^{i-1} \mu(P_j^c \cap Q_j) \leq \\ &\leq \sum_{j=1}^i \mu(P_j \Delta Q_j) < r\epsilon. \end{aligned}$$

Finalmente, para $i = r$, si $1 \leq j \leq r - 1$, se tiene

$$S_r \subset Q_j^c, \quad S_r^c = \bigcup_{j=1}^{r-1} Q_j, \quad P_i \subset P_j^c, \quad P_r^c = \bigcup_{j=1}^{r-1} P_j \text{ } \mu\text{-c.t.p.}$$

Luego

$$\begin{aligned} \mu(P_r \Delta S_r) &= \mu(P_r^c \cap S_r) + \mu(P_r \cap S_r^c) \leq \\ &\leq \sum_{j=1}^{r-1} \mu(P_j \cap Q_j^c) + \sum_{j=1}^{r-1} \mu(P_j^c \cap Q_j) \leq \\ &\leq \sum_{j=1}^{r-1} \mu(P_j \Delta Q_j) < r\epsilon. \end{aligned}$$

Con esto acabamos de demostrar que los átomos de la partición $\mathcal{S}(\epsilon)$ aproximan a los átomos de la partición dada \mathcal{P} .

Para terminar, falta probar que $H(\mathcal{P}|\mathcal{S}(\epsilon)) \rightarrow 0$ cuando $\epsilon \rightarrow 0$.

$$\begin{aligned} H(\mathcal{P}|\mathcal{S}(\epsilon)) &= - \sum_i \sum_j \mu(P_j \cap S_i) \log \left(\frac{\mu(P_j \cap S_i)}{\mu(S_i)} \right) = \\ &= \sum_i \mu(S_i) \sum_j \phi \left(\frac{\mu(P_j \cap S_i)}{\mu(S_i)} \right), \end{aligned}$$

donde $\phi(x) = -x \log x$, la primera suma es para aquellos $i \in \{1, 2, \dots, r\}$ tales que $\mu(S_i) \neq 0$, y la segunda suma para aquellos $j \in \{1, 2, \dots, r\}$ tales que $\mu(P_j \cap S_i) \neq 0$. Como hay una cantidad finita acotada de sumandos, basta probar que cada uno converge a cero cuando $\epsilon \rightarrow 0$.

Primero veamos que $\mu(S_i) \rightarrow \mu(P_i) > 0$. En efecto,

$$|\mu(S_i) - \mu(P_i)| \leq \mu(S_i \Delta P_i) < r\epsilon \rightarrow 0.$$

Por un lado, si $j = i$: $\mu(P_i \cap S_i) = \mu(S_i) - \mu(P_i^c \cap S_i)$,

$$0 \leq \mu(P_i^c \cap S_i) \leq \mu(P_i \Delta S_i) < r\epsilon \rightarrow 0$$

Entonces

$$\lim_{\epsilon \rightarrow 0} \phi \left(\frac{\mu(P_i \cap S_i)}{\mu(S_i)} \right) = \phi(1) = 0$$

Por otro lado, si $j \neq i$, entonces $P_j \subset P_i^c$, y

$$0 \leq \mu(P_j \cap S_i) \leq \mu(P_i^c \cap S_i) \leq \mu(P_i \Delta S_i) < r\epsilon \rightarrow 0$$

Luego,

$$\lim_{\epsilon \rightarrow 0} \phi \left(\frac{\mu(P_j \cap S_i)}{\mu(S_i)} \right) = \lim_{x \rightarrow 0} \phi(x) = 0$$

□

9.4. Entropía de una transformación respecto a una partición

Sea $T : X \mapsto X$ medible que preserva una medida de probabilidad μ . Sea \mathcal{P} una partición de X con entropía $H(\mathcal{P})$ finita. Es fácil ver que la familia formada por las preimágenes por T de los átomos de \mathcal{P} forman una nueva partición, que denotamos como $T^{-1}\mathcal{P}$, que tiene la misma entropía que \mathcal{P} .

$$H(T^{-1}\mathcal{P}) = H(\mathcal{P})$$

Si se tienen dos particiones \mathcal{P} y \mathcal{Q} , se cumple:

$$H(T^{-1}\mathcal{P}|T^{-1}\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q})$$

Considérese para cada $n \geq 1$, la partición producto:

$$\mathcal{P}_n = \mathcal{P} \vee T^{-1}\mathcal{P} \vee \dots \vee T^{-(n-1)}\mathcal{P}$$

Al aumentar n se tienen particiones cada vez mas finas, luego $H(\mathcal{P}_n)$ es creciente con n .

Además, por el corolario 9.1.8

$$H(\mathcal{P}_n) \leq H(\mathcal{P}) + H(T^{-1}\mathcal{P}) + \dots + H(T^{-(n-1)}\mathcal{P}) = nH(\mathcal{P})$$

Luego:

$$\frac{H(\mathcal{P}_n)}{n} \leq H(\mathcal{P}) < \infty \quad \forall n \geq 1$$

Proposición 9.4.1. La sucesión $\{H(\mathcal{P}_n)\}_{n \geq 1}$ es subaditiva, esto es:

$$H(\mathcal{P}_{n+m}) \leq H(\mathcal{P}_n) + H(\mathcal{P}_m) \quad \forall n, m \geq 1$$

Demostración:

$$\begin{aligned} H(\mathcal{P}_{n+m}) &= H((\vee_{j=0}^{n-1} T^{-j}\mathcal{P}) \vee (\vee_{j=n}^{n+m-1} T^{-j}\mathcal{P})) \\ &\leq H(\mathcal{P}_n) + H(T^{-n}\mathcal{P}_m) = H(\mathcal{P}_n) + H(\mathcal{P}_m) \quad \forall n, m \geq 1 \end{aligned}$$

□

Teorema 9.4.2.

$$\lim_{n \rightarrow \infty} \frac{H(\mathcal{P}_n)}{n} = \inf_{n \geq 1} \frac{H(\mathcal{P}_n)}{n} \leq H(\mathcal{P})$$

Demostración: Basta demostrar que toda sucesión $\{a_n\}_{n \geq 1}$ de reales, subaditiva (es decir $a_{n+m} \leq a_n + a_m$ para todos n y $m \geq 1$), con a_n/n acotada inferiormente, cumple

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} = \inf_{n \geq 1} \frac{a_n}{n}$$

Sea $c = \inf_{n \geq 1} a_n/n$. Sea, dado $\epsilon > 0$, un natural $n_0 \geq 1$ tal que $a_{n_0}/n_0 < c + \epsilon$. Si $n \geq n_0$, haciendo la división entera $n = qn_0 + r$, con $0 \leq r < n_0$, y aplicando la subaditividad, se obtiene:

$$\frac{a_n}{n} = \frac{a_{qn_0+r}}{n} \leq \frac{qa_{n_0}}{qn_0+r} + \frac{a_r}{n} \leq \frac{q}{qn_0} a_{n_0} + \frac{\max_{0 \leq i < n_0} a_i}{n} < c + 2\epsilon \quad \forall n \geq N$$

si N es suficientemente grande.

Luego

$$c \leq \frac{a_n}{n} < c + 2\epsilon \quad \forall n \geq N$$

Es decir

$$\frac{a_n}{n} \xrightarrow{n \rightarrow \infty} c$$

□

Definición 9.4.3. Entropía de una transformación respecto a una partición

Si T es medible y \mathcal{P} es una partición con entropía finita, se llama *entropía de la transformación T respecto a la partición \mathcal{P}* al número real

$$h(T, \mathcal{P}) = \lim_{n \rightarrow \infty} \frac{H(\mathcal{P}_n)}{n} = \inf_{n \geq 1} \frac{H(\mathcal{P}_n)}{n}$$

Interpretación:

$T : X \mapsto X$ define un *sistema dinámico determinístico*: Da una ley conocida, que en función de cada *estado inicial* $x \in X$, determina la sucesión de estados futuros: $T(x), T^2(x), \dots, T^j(x), \dots$ (la órbita por x). Supongamos que la transformación T se aplica a cada segundo: $T^j(x)$ es el estado al comienzo del segundo j -ésimo.

El físico o el ingeniero quiere informar sobre el estado del sistema en cada segundo futuro, a partir de su estado inicial x , por lo menos para casi todo punto x . Modestamente, no requiere determinar exactamente el punto $T^j(x)$ del espacio, sino que dispone de una cierta partición \mathcal{P} del espacio, predeterminada, y solo requiere saber en cuál átomo A_j de la partición está el punto $T^j(x)$, para cada instante futuro $j \geq 0$. (Por ejemplo si X es un subconjunto de \mathbb{R}^2 , quiere tener el dato de qué cuadrante contiene a $T^j(x)$ para cada $j \geq 0$.) La sucesión de átomos A_j , se llama *itinerario* de x . También modestamente, puede tolerar atrasos muy grandes en transmitir la información, con tal de que, cada tanto, puedan darse los datos acumulados.

El valor esperado de la cantidad de información inicial es $H(\mathcal{P})$, según la interpretación que dimos al definir entropía de una partición, en 9.1.2.

Al cabo de $n - 1$ segundos tendrá la información acumulada:

$$x \in A_0, T(x) \in A_1, T^2(x) \in A_2, \dots, T^{n-1}(x) \in A_{n-1}$$

Esto es lo mismo que el dato de qué átomo de la partición \mathcal{P}_n contiene al punto x . Luego, el valor esperado de la cantidad de información acumulada al cabo de $n - 1$ segundos, será $H(\mathcal{P}_n)$.

Tuvo n segundos para transmitir esa información acumulada. Entonces debió hacerlo con una *velocidad* promedio $H(\mathcal{P}_n)/n$. Como puede tolerar atrasos tan grandes como

necesite, puede acercarse tanto al límite de $H(\mathcal{P}_n)/n$ como desee. Este límite es un ínfimo, y es, por definición la entropía de T respecto a la partición \mathcal{P} .

Luego, la entropía de T respecto a la partición \mathcal{P} , es el límite inferior teórico de la velocidad de transmisión de datos requerida para informar sobre el itinerario del sistema.

$$\text{Cantidad de bits/segundo} = h(T, \mathcal{P}) / \log 2$$

Si la entropía es positiva, quiere decir que se requiere de una velocidad promedio positiva, para ir transmitiendo los datos, a medida que se acumulan. Usando mayor velocidad pueden adelantarse algunos datos. Pero no se pueden transmitir en tiempo acotado los datos para *todo* el futuro. Si la entropía es positiva, entonces no se pueden hacer previsiones en el estado asintótico futuro, transmitiendo solo una cantidad finita de datos en tiempo acotado N (porque, si así fuera, la velocidad promedio tendería a cero cuando n tiende a infinito, $n \geq N$).

En este sentido, cuando tiene entropía positiva el sistema es *imprevisible*, aunque sea determinístico.

También podemos dar la siguiente interpretación: si la entropía $h = h(T, \mathcal{P})$ es positiva el sistema es *caótico*, esto es, sensible a las condiciones iniciales, aunque en un sentido débil: hay itinerarios asintóticamente diferentes. En efecto, si el itinerario a partir de cierto instante N no dependiera más del estado inicial x , entonces las particiones \mathcal{P}_n , para $n \geq N$ coincidirían con \mathcal{P}_N . Al dividir su entropía $H(\mathcal{P}_n)$ entre n , y hacer n tender a infinito, la entropía h de la transformación sería nula.

También podemos dar la siguiente interpretación a la entropía $h = h(T, \mathcal{P})$ de la transformación: Es el grado de *desorden* que provoca la transformación T en el espacio. En efecto, nótese que $H(\mathcal{P}_n)$ es creciente con n , y h es la velocidad con la que crece en relación a n . Si h es positivo, entonces la partición inicial \mathcal{P} debe sufrir más y más subatomizaciones al aplicarle T sucesivas veces (pues de lo contrario, las particiones \mathcal{P}_n serían todas iguales a \mathcal{P}_N para $n \geq N$ y luego h sería cero).

Proposición 9.4.4. *Cualesquiera sean las particiones \mathcal{P} y \mathcal{Q} con entropías finitas, se cumple:*

1. $h(T, \mathcal{Q}) - h(T, \mathcal{P}) \leq H(\mathcal{Q}|\mathcal{P})$
2. Si \mathcal{Q} es mas fina que \mathcal{P} entonces $0 \leq h(T, \mathcal{Q}) - h(T, \mathcal{P})$.

Demostración: 1) Según el corolario 9.1.10 que acota la diferencia de entropías de particiones, se tiene

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

Aplicando la parte 2) del corolario 9.1.9, sobre la entropía del producto condicionada:

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

Como al condicionar con una partición mas fina disminuye la entropía, según la afirmación 3) de la proposición 9.3.2:

$$\begin{aligned} \sum_{j=0}^{n-1} H(T^{-j}\mathcal{Q} | \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}) &\leq \sum_{j=0}^{n-1} H(T^{-j}\mathcal{Q} | T^{-j}\mathcal{P}) = \\ &= \sum_{j=0}^{n-1} H(\mathcal{Q} | \mathcal{P}) = nH(\mathcal{Q} | \mathcal{P}). \end{aligned}$$

Luego:

$$\frac{1}{n}H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) - \frac{1}{n}H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}) \leq H(\mathcal{Q} | \mathcal{P}),$$

y haciendo n tender a infinito, se tiene

$$h(T, \mathcal{Q}) - h(T, \mathcal{P}) \leq H(\mathcal{Q} | \mathcal{P}).$$

2) Si \mathcal{Q} es mas fina que \mathcal{P} , entonces $\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}$ es mas fina que $\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}$, y entonces por la proposición 9.3.2

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

Dividiendo entre n y haciendo n tender a infinito se concluye

$$0 \leq h(T, \mathcal{Q}) - h(T, \mathcal{P}).$$

□

Proposición 9.4.5. .

- 1) $h(T, \mathcal{P}) = h(T, T^{-1}\mathcal{P}) = h(T, T^{-n}\mathcal{P})$ para todo $n \geq 0$
- 2) $h(T, \mathcal{P}) = h(T, \bigvee_{j=0}^m T^{-j}\mathcal{P})$ para todo $m \geq 0$.
- 3) Si T es invertible con inversa medible: $h(T, \mathcal{P}) = h(T, T\mathcal{P}) = h(T, T^n\mathcal{P})$ para todo $n \in \mathbb{Z}$.

Demostración:

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

Dividiendo entre n y haciendo n tender a infinito se tiene $h(T, \mathcal{P}) = h(T, T^{-1}\mathcal{P})$. La otra igualdad se obtiene aplicando n veces la primera.

$$2) \bigvee_{i=0}^{n-1} T^{-i}(\bigvee_{j=0}^m T^{-j}\mathcal{P}) = \bigvee_{i=0}^{n+m-1} T^{-i}\mathcal{P}.$$

Calculando su entropía H , dividiendo entre n y haciendo n tender a infinito, con m fijo, se tiene la igualdad 2).

3) Si T es invertible y preserva μ , entonces T^{-1} también preserva μ . Luego, para toda partición \mathcal{P} se cumple $H(T\mathcal{P}) = H(\mathcal{P})$. La prueba sigue análogamente a la de 1), tomando la imagen por T de \mathcal{P} en lugar de su preimagen.

□

9.5. Entropía métrica de una transformación

Definición 9.5.1. Entropía métrica

Sea $T : X \mapsto X$ medible que preserva la medida de probabilidad μ . Se llama *entropía métrica de T con respecto de la medida μ* a

$$h_\mu(T) = \sup\{h(T, \mathcal{P}) : \mathcal{P} \text{ partición tal que } H(\mathcal{P}) < \infty\} \in [0, \infty]$$

Teorema 9.5.2. Entropía métrica

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

Demostración: Como toda partición finita tiene entropía finita, se cumple

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

Para probar la desigualdad contraria, basta demostrar que dada una partición \mathcal{P} con $H(\mathcal{P}) < \infty$, existe una sucesión \mathcal{P}_n de particiones finitas tales que $h(T, \mathcal{P}_n) \rightarrow h(T, \mathcal{P})$.

Dada $\mathcal{P} = \{P_i\}_{i \geq 1}$ sea $\mathcal{P}_n = \{P_1, \dots, P_{n-1}, \bigcup_{i \geq n} P_i\}$. \mathcal{P} es mas fina que \mathcal{P}_n . Luego vale la proposición 9.4.4:

$$0 \leq h(T, \mathcal{P}) - h(T, \mathcal{P}_n) \leq H(\mathcal{P}|\mathcal{P}_n)$$

y basta probar que $H(\mathcal{P}|\mathcal{P}_n) \rightarrow 0$.

Aplicando la definición de entropía condicionada, y observando que para $i = 1, 2, \dots, n-1$, los átomos P_i de las particiones \mathcal{P} y \mathcal{P}_n coinciden, se tiene:

$$H(\mathcal{P}|\mathcal{P}_n) = - \sum_{i \geq n} \mu(P_i) \log \frac{\mu(P_i)}{\mu(\bigcup_{j \geq 1} P_j)} = \sum_{i \geq n} \phi(\mu(P_i)) - \phi\left(\sum_{j \geq n} \mu(P_j)\right)$$

donde $\phi(x) = -x \log x$.

Como $H(\mathcal{P}) = \sum_{i \geq 1} \phi(\mu(P_i)) < \infty$, la cola de la serie $\sum_{i \geq n} \phi(\mu(P_i))$ converge a cero cuando n tiende a infinito.

Por otro lado, como $\sum_{j \geq 1} \mu(P_j) = 1$, la cola de la serie $\sum_{j \geq n} \mu(P_j)$ converge a cero cuando n tiende a infinito. Siendo $\lim_{x \rightarrow 0} \phi(x) = 0$, se tiene que $\phi\left(\sum_{j \geq n} \mu(P_j)\right)$ converge a cero cuando n tiende a infinito. \square

Teorema 9.5.3. Invariancia de la entropía métrica por isomorfismos de espacios de medida

Sean $T_1 : X_1 \mapsto X_1$ que preserva la medida μ_1 , y $T_2 : X_2 \mapsto X_2$ que preserva la medida μ_2 .

Si T_1 y T_2 son equivalentes por isomorfismo de espacios de medida, entonces las entropías métricas son iguales:

$$h_{\mu_1}(T_1) = h_{\mu_2}(T_2)$$

Demostración: Sea $f : X_1 \mapsto X_2$ el isomorfismo de espacios de medida tal que $f \circ T_1 = T_2 \circ f$. Para toda partición $\mathcal{P}_2 = \{P_i\}_{i \geq 1}$ de X_2 , la partición $f^{-1}\mathcal{P}_2 = \{f^{-1}(P_i)\}_{i \geq 1}$ de X_1 , cumple $H(f^{-1}\mathcal{P}_2) = H(\mathcal{P}_2)$, porque f preserva medidas. Además, como $f \circ T_1^j = T_2^j \circ f$ para todo $j \geq 0$, se cumple

$$f^{-1} \bigvee_{j=0}^{n-1} T_2^{-j} \mathcal{P}_2 = \bigvee_{j=0}^{n-1} T_1^{-j} f^{-1} \mathcal{P}_2$$

de donde se obtiene que

$$H(\bigvee_{j=0}^{n-1} T_2^{-j} \mathcal{P}_2) = H(\bigvee_{j=0}^{n-1} T_1^{-j} f^{-1} \mathcal{P}_2)$$

Dividiendo entre n y haciendo n tender a infinito se obtiene $h(T_2, \mathcal{P}_2) = h(T_1, f^{-1}\mathcal{P}_2)$. Como vale para cualquier partición \mathcal{P}_2 de X_2 , se tiene:

$$\begin{aligned} h_{\mu_2}(T_2) &= \sup\{h(T_2, \mathcal{P}_2) : H(\mathcal{P}_2) < \infty\} \leq \\ &\leq \sup\{h(T_1, \mathcal{P}_1) : H(\mathcal{P}_1) < \infty\} = h_{\mu_1}(T_1) \end{aligned}$$

Como la relación de ser equivalentes por isomorfismo de espacios de medida es simétrica, usando f^{-1} en lugar de f , se obtiene la desigualdad contraria. \square

La entropía métrica de una transformación es un supremo de las entropías de la transformación respecto a particiones. Para calcularla puede ser muy útil saber cuando es un máximo, y para que partición se alcanza ese máximo. Condiciones que aseguran eso estarán dadas en el teorema de Kolmogorov-Sinai, para el cual necesitamos la siguiente definición:

Definición 9.5.4. Partición generadora

Sea T medible en un espacio de medida (X, \mathcal{A}, μ) . Una partición \mathcal{P} se llama *generadora* si se cumple alguna de las siguientes condiciones:

- i) o bien $\bigvee_{j=0}^{\infty} T^{-j} \mathcal{P} = \mathcal{A}$ μ -c.t.p.
- ii) o bien $\bigvee_{j=0}^{\infty} T^{-j} \mathcal{P} \neq \mathcal{A}$ μ -c.t.p., pero T es invertible con inversa medible y $\bigvee_{j=-\infty}^{\infty} T^{-j} \mathcal{P} = \mathcal{A}$ μ -c.t.p.

Se aclara que cuando T es invertible, tanto son particiones generadoras las que cumplen i) como las que cumplen ii).

Teorema 9.5.5. Kolmogorov-Sinai

Si \mathcal{P} es una partición generadora con $H(\mathcal{P}) < \infty$, entonces

$$h_{\mu}(T) = h(T, \mathcal{P})$$

Demostración: Consideraremos los dos casos i) y ii) de la definición de partición generadora

Caso i): Sea $\mathcal{P}_n = \bigvee_{j=0}^{n-1} T^{-j} \mathcal{P}$. Tenemos una sucesión de particiones cada vez mas finas, y que generan la σ -álgebra, porque \mathcal{P} es partición generadora.

Aplicando la proposición 9.4.4 y el lema 9.3.4, para toda partición finita \mathcal{Q} se cumple:

$$h(T, \mathcal{Q}) - h(T, \mathcal{P}_n) \leq H(\mathcal{Q}|\mathcal{P}_n) \rightarrow 0$$

Por otro lado, usando la proposición 9.4.5 se tiene

$$h(T, \mathcal{P}_n) = h(T, \mathcal{P}) \quad \text{para todo } n \geq 1.$$

Luego, para toda partición finita \mathcal{Q} se cumple

$$h(T, \mathcal{Q}) \leq h(T, \mathcal{P}).$$

Finalmente

$$\begin{aligned} h_\mu(T) &= \sup\{h(T, \mathcal{Q}) : \mathcal{Q} \text{ partición finita}\} \leq h(T, \mathcal{P}) \leq \\ &\sup\{h(T, \mathcal{R}) : H(\mathcal{R}) < \infty\} = h_\mu(T). \end{aligned}$$

Caso ii): Sea $\mathcal{P}_n = \bigvee_{j=-n}^n T^{-j}\mathcal{P} = T^n \left(\bigvee_{j=0}^{2n} T^{-j}\mathcal{P} \right)$. Usando la proposición 9.4.5 se tiene $h(T, \mathcal{P}_n) = h(T, \mathcal{P})$ para todo $n \geq 1$.

La prueba continúa como en el caso anterior, sin mas cambios. \square

Ejemplos: Entropía métrica del shift de Bernoulli.

Sea $\sigma : B(m) \mapsto B(m)$ el shift unilateral de m símbolos.

Tomemos la partición formada por los m cilindros de radio 1:

$$\mathcal{P} = \{C(1), C(2), \dots, C(m)\}$$

Obsérvese que para todo $n \geq 1$ la partición $\mathcal{P}_n = \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}$ es la de los cilindros de radio n . Como la σ -álgebra de Borel es la generada por los cilindros, la partición \mathcal{P} es generadora.

Luego, para cualquier medida μ que sea invariante por el shift, la entropía métrica $h_\mu(T)$ es igual a $h(T, \mathcal{P})$.

Si la medida μ es de Bernoulli, con vector de probabilidad

$$p = (p(1), p(2), \dots, p(m)),$$

entonces

$$\begin{aligned} H(\mathcal{P}_n) &= - \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \mu(C(i_1, i_2, \dots, i_n)) \log \mu(C(i_1, i_2, \dots, i_n)) \\ &= - \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m p(i_1)p(i_2) \dots p(i_n) \log p(i_1)p(i_2) \dots p(i_n) = \\ &= - \sum_{j=1}^n \sum_{i_1=1}^m \dots \sum_{i_j=1}^m \dots \sum_{i_n=1}^m p(i_1) \dots p(i_j) \dots p(i_n) \log p(i_j) \end{aligned}$$

Como $\sum_{i=1}^m p(i) = 1$, se tiene:

$$H(\mathcal{P}_n) = - \sum_{j=1}^n \sum_{i_j=1}^m p(i_j) \log p(i_j) = -n \sum_{i=1}^m p(i) \log p(i)$$

Dividiendo entre n y haciendo n tender a infinito, se tiene

$$h_\mu(T) = h(T, \mathcal{P}) = - \sum_{i=1}^m p(i) \log p(i)$$

Veamos que, entre las medidas que son de Bernoulli, la que da mayor entropía métrica es la que tiene vector de probabilidad

$$p = (1/m, 1/m, \dots, 1/m),$$

y que el shift con esa medida de Bernoulli equidistribuida tiene entropía métrica igual a $\log m$.

En efecto, para el vector de probabilidad

$$(1/m, 1/m, \dots, 1/m),$$

la entropía métrica es

$$-m \frac{1}{m} \log \frac{1}{m} = \log m$$

Para cualquier otro vector de probabilidad, usando la función $\phi(x) = -x \log x$, y la desigualdad de Jensen, se tiene

$$h_\mu(T) = \sum_{i=1}^m \phi(p_i) = m \sum_{i=1}^m \phi(p_i)/m \leq$$

$$m \phi \left(\sum_{i=1}^m p_i/m \right) = m \phi(1/m) = \log m.$$

Ejercicio 9.5.6. Probar que la entropía métrica del shift de Bernoulli bilateral es

$$- \sum_{i=1}^m p_i \log p_i.$$

Ejemplo 9.5.7. La entropía métrica de la transformación $T(z) = z^k$, con k natural mayor o igual que 2, en el círculo $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, para la medida de Lebesgue, es $\log k$. En efecto, vemos que T es equivalente por isomorfismo de espacios de medida al shift de Bernoulli con k símbolos, y con vector de probabilidad $(1/k, \dots, 1/k)$.

Ejemplo 9.5.8. Entropía métrica de la herradura de Smale

La herradura de Smale de dos patas definida en la Sección 3.5 tiene entropía métrica, respecto a la medida de Bernoulli de vector de probabilidad $(1/2, 1/2)$, igual a $\log 2$. En efecto, es conjugada al shift bilateral de 2 símbolos.

Teorema 9.5.9. Entropía métrica de los iterados de T

Sea $T : X \mapsto X$ medible. Denotamos con \mathcal{M}_T el conjunto de las probabilidades invariantes con T . Asumimos por hipótesis que $\mathcal{M}_T \neq \emptyset$. Entonces para todo $n \in \mathbb{N}^+$ y para toda $\mu \in \mathcal{M}_T$ se cumple

$$h_\mu(T^n) = nh_\mu(T).$$

Demostración. Dejemos fijo $n \in \mathbb{N}^+$. Como $\mu \in \mathcal{M}_T$ y $\mathcal{M}_T \subset \mathcal{M}_{T^n}$ entonces $\mu \in \mathcal{M}_{T^n}$. Sea \mathcal{P} una partición finita. Denotamos $h(\mathcal{P}, \mu, T)$ y $h(\mathcal{P}, \mu, T^n)$ a las entropías de la partición \mathcal{P} con respecto a la medida μ , para las transformaciones T y T^n respectivamente. Por definición se cumple:

$$h(\mathcal{P}, \mu, T) = \lim_{m \rightarrow +\infty} \frac{H(\bigvee_{j=0}^{m-1} T^{-j} \mathcal{P}, \mu)}{m} = \lim_{q \rightarrow +\infty} \frac{H(\bigvee_{j=0}^{qn-1} T^{-j} \mathcal{P}, \mu)}{qn}. \quad (9.7)$$

Para todo $q \geq 2$ se cumple:

$$\bigvee_{j=0}^{qn-1} T^{-j} \mathcal{P} = \bigvee_{i=0}^{q-1} \left(\bigvee_{j=in}^{(i+1)n-1} T^{-j} \mathcal{P} \right) = \bigvee_{i=0}^{q-1} T^{-in} \left(\bigvee_{h=0}^{n-1} T^{-h} \mathcal{P} \right).$$

Luego, llamando

$$\mathcal{P}_n := \bigvee_{h=0}^{n-1} T^{-h} \mathcal{P},$$

tenemos:

$$H\left(\bigvee_{j=0}^{qn-1} T^{-j} \mathcal{P}, \mu\right) = H\left(\bigvee_{i=0}^{q-1} (T^n)^{-i} \mathcal{P}_n, \mu\right)$$

Sustituyendo en la igualdad (9.7), obtenemos:

$$nh(\mathcal{P}, \mu, T) = \lim_{q \rightarrow +\infty} \frac{H\left(\bigvee_{i=0}^{q-1} (T^n)^{-i} \mathcal{P}_n, \mu\right)}{q} = h(\mathcal{P}_n, \mu, T^n). \quad (9.8)$$

Tomando el supremo en la colección \aleph de todas las particiones \mathcal{P} finitas en la igualdad (9.8), siempre con n fijo, obtenemos:

$$nh_\mu(T) = n \sup_{P \in \aleph} h(\mathcal{P}, \mu, T) = \sup_{P \in \aleph} h(\mathcal{P}_n, \mu, T^n) \leq \sup_{Q \in \aleph} h(Q, \mu, T^n) = h_\mu(T^n)$$

Por otro lado, como \mathcal{P}_n es más fina que \mathcal{P} , tenemos:

$$h(\mathcal{P}, \mu, T^n) \leq h(\mathcal{P}_n, \mu, T),$$

de donde, usando nuevamente la igualdad (9.8), obtenemos:

$$h_\mu(T^n) = \sup_{P \in \aleph} h(\mathcal{P}, \mu, T^n) \leq \sup_{P \in \aleph} h(\mathcal{P}_n, \mu, T^n) = \sup_{P \in \aleph} nh(\mathcal{P}, \mu, T) = nh_\mu(T),$$

terminando la demostración del Teorema 9.5.9. \square

9.6. Entropía topológica.

En lo que sigue, $T : X \mapsto X$ será una transformación continua en un espacio métrico compacto X .

Definición 9.6.1. Generadores y separadores

Sean n natural ≥ 1 y $\epsilon > 0$ real, dados.

Dos puntos x e y de X se (ϵ, n) -acompañan si $d(T^j(x), T^j(y)) < \epsilon$ para todo $j = 0, 1, \dots, n-1$.

Dos puntos x e y de X se (ϵ, n) -separan si no se (ϵ, n) -acompañan. Es decir, existe algún $j = 0, 1, \dots, n-1$ tal que $d(T^j(x), T^j(y)) \geq \epsilon$.

Un subconjunto E de X es un (ϵ, n) -generador si dado $x \in X$ existe algún $y \in E$ tal que x e y se (ϵ, n) -acompañan.

Un subconjunto no vacío S de X es un (ϵ, n) -separador si dos puntos x e y de S , diferentes, se (ϵ, n) -separan

Usaremos la siguiente notación: $\sharp A =$ cantidad de elementos del conjunto $A \in [0, \infty]$.

Proposición 9.6.2. Dado $\epsilon > 0$ existe $M = M(\epsilon) \geq 1$ tal que para todo $n \geq 1$ existe un (ϵ, n) -generador E con $\sharp E \leq M^n$.

Demostración: Sea \mathcal{V} un cubrimiento finito de X con bolas abiertas de radio $\epsilon/2$, formado por, digamos, M bolas.

Dada una n -upla $(V_0, V_1, \dots, V_{n-1})$ de bolas de \mathcal{V} , elijamos, cuando no es vacío, un y solo un punto en el conjunto $\bigcap_{j=0}^{n-1} T^{-j}V_j$. Sea E el conjunto de los puntos así elegidos. Se tiene $\sharp E \leq M^n$.

Veamos que E es un (ϵ, n) -generador. En efecto, dado $x \in X$, existe para cada $j = 0, 1, \dots, n-1$, algún V_j del cubrimiento \mathcal{V} que contiene al punto $T^j(x)$. Luego $x \in \bigcap_{j=0}^{n-1} T^{-j}V_j \neq \emptyset$, y existe $y \in E$ en el mismo conjunto intersección. Entonces, $d(T^j(x), T^j(y)) < \text{diam}V_j = \epsilon$ para todo $j = 0, 1, \dots, n-1$. \square

Definición 9.6.3. Como consecuencia de la proposición anterior, existe, para todo $\epsilon > 0$ y todo $n \geq 1$, el número natural

$$r(\epsilon, n) = \text{mín}\{\sharp E : E \text{ es un } (\epsilon, n)\text{-generador}\} \leq M(\epsilon)^n$$

Además

$$r(\epsilon, n) \geq 1$$

Consecuencias:

Existe, para todo $\epsilon > 0$, el número real

$$\limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} \geq 0$$

Esta expresión decrece cuando ϵ crece. En efecto, si $\epsilon' < \epsilon$, todo (ϵ', n) -generador es un (ϵ, n) -generador. Entonces $r(\epsilon', n) \geq r(\epsilon, n)$. Calculando el logaritmo, dividiendo entre n , y tomando el límite superior cuando n tiende a infinito, se mantiene la desigualdad.

Luego,

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} \in [0, \infty]$$

Proposición 9.6.4. Si S es un (ϵ, n) -separador y E es un $(\epsilon/2, n)$ -generador, entonces $\#S \leq \#E$.

Demostración: Veamos que puede establecerse una correspondencia inyectiva que va de S a E .

Dado $x \in S$, como E es un $(\epsilon/2, n)$ -generador, existe $y \in E$ tal que x e y se $(\epsilon/2, n)$ -acompañan.

La correspondencia así establecida de S a E es inyectiva, pues si no lo fuera, existirían $x \neq z$ en S que son $(\epsilon/2, n)$ -acompañados por el mismo $x \in E$. Entonces, por la propiedad triangular de la distancia, los puntos x y z se (ϵ, n) acompañan entre sí. Esto contradice la hipótesis de que S es un (ϵ, n) -separador. \square

Definición 9.6.5. Como consecuencia de la proposición anterior, existe para todo $\epsilon > 0$ y todo $n \geq 1$ el número natural

$$s(\epsilon, n) = \max\{\#S : S \text{ es un } (\epsilon, n)\text{-separador}\} \leq r(\epsilon/2, n) \leq M(\epsilon/2)^n$$

Además

$$s(\epsilon, n) \geq 1$$

Consecuencias:

Existe, para todo $\epsilon > 0$, el número real

$$\limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} \geq 0$$

Esta expresión decrece cuando ϵ crece. En efecto, si $\epsilon > \epsilon'$, todo (ϵ, n) -separador es un (ϵ', n) -separador. Entonces $s(\epsilon, n) \leq s(\epsilon', n)$. Calculando el logaritmo, dividiendo entre n , y tomando el límite superior cuando n tiende a infinito, se mantiene la desigualdad. Luego,

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} \in [0, \infty]$$

Proposición 9.6.6. Si S es un (ϵ, n) -separador con la cantidad máxima de elementos, entonces también es un (ϵ, n) -generador.

Demostración: Por absurdo, supongamos que existe $x \in X$ tal que para todo $y \in S$, los puntos x e y no se (ϵ, n) -acompañan. Entonces agregando el punto x al conjunto S se tiene otro (ϵ, n) -separador, con más elementos que S . \square

Consecuencias:

$$r(\epsilon, n) \leq s(\epsilon, n) \leq r(\epsilon/2, n)$$

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} = \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} \in [0, \infty]$$

Definición 9.6.7. Entropía topológica

Se llama *entropía topológica de la transformación T* a

$$\begin{aligned} h_{top}(T) &:= \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{\log s(\epsilon, n)}{n} \\ &= \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n)}{n} \in [0, \infty]. \end{aligned}$$

Notación: Cuando quede claro del contexto que nos estamos refiriendo a la entropía topológica, escribiremos:

$$h(T) := h_{top}(T).$$

El siguiente teorema vincula la entropía métrica con la entropía topológica:

Teorema 9.6.8. Principio variacional de la entropía

La entropía topológica de una transformación $T : X \mapsto X$ continua del espacio métrico compacto X , es el supremo de las entropías métricas de T con respecto a las medidas de probabilidad T invariantes:

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

Demostraremos el Teorema 9.6.8 más adelante en este capítulo, en la Sección 9.8. La prueba del Principio Variacional de la entropía puede encontrarse también, por ejemplo, en [Wa 2000, Theorem 8.6] o en [Kat-Has 1995, Theorem 4.5.3].

Interpretación de la entropía topológica

Por un lado, $h_{top}(T) > 0$ si y solo si para alguna medida μ que sea T -invariante, la entropía métrica $h_{\mu}(T) > 0$. Ya dimos un significado heurístico a la entropía métrica, como el grado de desorden que provoca T en el espacio.

Demos ahora otra interpretación: Decimos que estudiamos la dinámica con error $\epsilon > 0$ si no distinguimos puntos del espacio que distan menos que ϵ . Así, si x e y se (ϵ, n) acompañan, no distinguiremos la órbita de x de la de y , por lo menos hasta tiempo n .

Luego $s(\epsilon, n)$ es el número máximo de órbitas distinguibles hasta tiempo n con error ϵ . Es el número de datos necesarios para describir la dinámica del sistema hasta tiempo n con error ϵ .

La entropía topológica positiva, quiere decir que el número de datos necesarios para describir la dinámica con error pequeño, crece exponencialmente con el tiempo n .

Veamos que la entropía topológica es invariante por conjugaciones. Es entonces un invariante topológico.

Definición 9.6.9. Semiconjugación

Sean $T_1 : X_1 \mapsto X_1$ y $T_2 : X_2 \mapsto X_2$ transformaciones continuas en espacios métricos compactos.

Se dice que T_1 es *semiconjugada con* T_2 si existe $f : X_1 \mapsto X_2$ continua y sobreyectiva que cumple $f \circ T_1 = T_2 \circ f$.

Se dice que T_1 es *conjugada con* T_2 si existe $f : X_1 \mapsto X_2$ homeomorfismo que cumple $f \circ T_1 = T_2 \circ f$.

Se observa que dos transformaciones son conjugadas si y solo si cada una de ellas es semiconjugada con la otra.

Nótese que $f \circ T_1 = T_2 \circ f$ implica $f \circ T_1^j = T_2^j \circ f$, para todo $j \geq 0$.

Teorema 9.6.10. Invariancia de la entropía topológica

Si T_1 es semiconjugada con T_2 entonces

$$h(T_1) \geq h(T_2).$$

Luego, si T_1 es conjugada con T_2 entonces $h(T_1) = h(T_2)$.

Demostración: La primera afirmación implica la segunda.

Sea $f : X_1 \mapsto X_2$ continua y sobreyectiva, según la definición de semiconjugación.

Como f es sobreyectiva, existe $g : X_2 \mapsto X_1$, inversa a la derecha de f . Luego g es inyectiva.

Como f es continua, dado $\epsilon > 0$, existe $\delta > 0$ tal que $d(x, y) < \delta$ en X_1 implica $d(f(x), f(y)) < \epsilon$ en X_2 .

Como $f \circ T_1^j = T_2^j \circ f$, para todo $j \geq 0$, todo (ϵ, n) -separador en X_2 para T_2 , es transformado por g en un (δ, n) -separador en X_1 para T_1 . Luego:

$$\begin{aligned} s_2(\epsilon, n) &= \max\{\#S : S \text{ es un } (\epsilon, n)\text{-separador en } X_2 \text{ para } T_2\} \\ &\leq \max\{\#S : S \text{ es un } (\delta, n)\text{-separador en } X_1 \text{ para } T_1\} = s_1(\delta, n) \end{aligned}$$

Entonces, para todo $\epsilon > 0$:

$$\limsup_{n \rightarrow \infty} \frac{\log s_2(\epsilon, n)}{n} \leq \limsup_{n \rightarrow \infty} \frac{\log s_1(\delta, n)}{n} \leq \sup_{\rho > 0} \limsup_{n \rightarrow \infty} \frac{\log s_1(\rho, n)}{n} = h(T_1)$$

Haciendo ϵ tender a cero:

$$h(T_2) \leq h(T_1)$$

□

Proposición 9.6.11. Si $A \subset X$ es compacto invariante para el futuro, o para adelante (esto es $T(A) \subset A$), entonces

$$h(T) \geq h(T|_A)$$

Demostración: Todo (ϵ, n) separador para $T|_A$ es un (ϵ, n) separador para T . Entonces

$$\begin{aligned} s_A(\epsilon, n) &= \max\{\#S : S \text{ es un } (\epsilon, n)\text{-separador para } T|_A\} \\ &\leq \max\{\#S : S \text{ es un } (\epsilon, n)\text{-separador para } T\} = s(\epsilon, n) \end{aligned}$$

Por definición de entropía topológica, lo anterior implica la tesis. □

Veamos una condición muy utilizada para demostrar que una transformación tiene entropía topológica positiva:

Definición 9.6.12. Un transformación $T : X \mapsto X$ contiene una herradura si existe un compacto A invariante hacia adelante tal que $T|_A$ es semiconjugada con la herradura de Smale.

Corolario 9.6.13. Si una transformación contiene una herradura, entonces tiene entropía topológica mayor o igual que $\log 2$.

Demostración: Para la medida de Bernoulli con vector de probabilidad $(1/2, 1/2)$, la herradura de Smale σ tiene entropía métrica igual a $\log 2$. Por el principio variacional de la entropía, la entropía topológica de la herradura de Smale es mayor o igual que $\log 2$. (Mas adelante veremos que es igual a $\log 2$).

Luego

$$h(T) \geq h(T|_A) \geq h(\sigma) \geq \log 2$$

□

El siguiente teorema muestra como la entropía topológica crece al iterar T :

Teorema 9.6.14. Entropía topológica de los iterados de T

Sea T continua en un espacio métrico compacto. Se cumple:

1. $h(T^m) = mh(T)$ para todo $m \geq 1$.
2. Si T es un homeomorfismo, entonces $h(T^{-1}) = h(T)$ y $h(T^{-m}) = mh(T)$ para todo $m \geq 1$.

Demostración:

1) Todo (ϵ, n) -separador de T^m es un (ϵ, mn) -separador de T . Entonces:

$$\begin{aligned} s_m(\epsilon, n) &= \max\{\#\{S : S \text{ es un } (\epsilon, n)\text{-separador para } T^m\}\} \\ &\leq \max\{\#\{S : S \text{ es un } (\epsilon, mn)\text{-separador para } T\}\} = s(\epsilon, mn) \end{aligned}$$

Luego:

$$\limsup_{n \rightarrow \infty} \frac{\log s_m(\epsilon, n)}{n} \leq \limsup_{n \rightarrow \infty} \frac{m \log s(\epsilon, mn)}{mn} \leq m \limsup_{k \rightarrow \infty} \frac{\log s(\epsilon, k)}{k}$$

Tomando $\epsilon \rightarrow 0$, se obtiene:

$$h(T^m) \leq mh(T)$$

Para probar la desigualdad contraria, usemos la definición de entropía topológica basada en los generadores.

Dado $\epsilon > 0$, por la continuidad de T , sea $\delta > 0$ tal que, si $d(x, y) < \delta$, entonces, para $j = 0, 1, \dots, m-1$, se cumple $d(T^j(x), T^j(y)) < \epsilon$.

Todo (δ, n) -generador para T^m será un $(\epsilon, mn + j)$ -generador para T . Luego

$$\begin{aligned} r_m(\delta, n) &= \min\{\#\{E : E \text{ es un } (\delta, n)\text{-generador para } T^m\}\} \\ &\geq \min\{\#\{E : E \text{ es un } (\epsilon, mn + j)\text{-generador para } T\}\} = r(\epsilon, mn + j) \end{aligned}$$

$$r_m(\delta, n) \geq \max_{0 \leq j \leq m-1} r(\epsilon, mn + j)$$

Luego, para todo $\epsilon > 0$, se cumple:

$$\begin{aligned} h(T^m) &= \sup_{\rho > 0} \limsup_{n \rightarrow \infty} \frac{\log r_m(\rho, n)}{n} \geq \limsup_{n \rightarrow \infty} \frac{\log r_m(\delta, n)}{n} \\ &\geq \limsup_{n \rightarrow \infty} \frac{\log (\max\{r(\epsilon, mn + j) : 0 \leq j \leq m-1\})}{n} \\ &\geq \limsup_{N \rightarrow \infty} \frac{\log r(\epsilon, N)}{\text{ent}(N/m)} = m \limsup_{N \rightarrow \infty} \frac{\log r(\epsilon, N)}{N} \end{aligned}$$

Haciendo $\epsilon \rightarrow 0$, se tiene

$$h(T^m) \geq mh(T)$$

2) Todo (ϵ, n) -generador de T es transformado por T^{n-1} en un (ϵ, n) -generador de T^{-1} , con la misma cantidad de elementos. Y recíprocamente, todo (ϵ, n) -generador de T^{-1} , es transformado por $T^{-(n-1)}$ en un (ϵ, n) -generador de T .

Luego la cantidad mínima de elementos de los (ϵ, n) -generadores de T y de T^{-1} son iguales. Por definición de entropía topológica, $h(T^{-1}) = h(T)$.

La última igualdad es una consecuencia inmediata de las anteriores. \square

9.7. Ejemplos de cálculo de la entropía topológica

Ejemplo 9.7.1. La entropía topológica del shift de m símbolos es $\log m$.

Demostremoslo para el shift unilateral σ . Se deja como ejercicio, adaptar la prueba, para verificar que lo mismo vale para el shift bilateral.

1a. parte) $h(\sigma) \geq \log m$. Usando la entropía métrica del shift de Bernoulli y el principio variacional de la entropía, se deduce esta desigualdad. Pero, ya que no demostramos este principio, probemos la desigualdad directamente a partir de la definición de entropía topológica.

Elijamos para cada $n \geq 1$ un conjunto S formado por un y solo un punto en cada cilindro $C(a_0, a_1, \dots, a_{n-1})$. Se tiene $\#S = m^n$, y S es un (ϵ, n) -separador, para todo $\epsilon > 0, \epsilon < 1$. En efecto: $x \neq y \in S$ implica que existe $j = 0, 1, \dots, n-1$ tal que $x_j \neq y_j$. Luego:

$$d(\sigma^j x, \sigma^j y) = \sum_{i=1}^{\infty} \frac{|x_{i+j} - y_{i+j}|}{\lambda^i} \geq |x_j - y_j| \geq 1 > \epsilon$$

Entonces $s(\epsilon, n) \geq m^n$ para todo $\epsilon < 1$, de donde $h(\sigma) \geq \log m$.

2a. parte) $h(\sigma) \leq \log m$.

Dado $\epsilon > 0$ sea N tal que todos los cilindros de radio N tienen diámetro menor que ϵ .

Elijamos para cada $n > N$ un conjunto E formado por un y solo un punto en cada cilindro $C(a_0, a_1, \dots, a_{n-1})$. Se tiene $\#E = m^n$, y E es un $(\epsilon, n - N)$ -generador, para todo $\epsilon > 0$. En efecto: dado $x \in B(m)$, existe $y \in E$ en el cilindro $C(x_0, x_1, \dots, x_{n-1})$. Luego, para todo $j = 0, 1, \dots, n - N$, los puntos $\sigma^j x$ y $\sigma^j y$ están en el cilindro $C(x_j, \dots, x_{n-1})$ de radio $n - j \geq N$. Entonces están en el mismo cilindro de radio N , y distan menos que ϵ . Los puntos x e y se $(\epsilon, n - N)$ -acompañan. Concluimos que $r(\epsilon, n - N) \leq m^n$, de donde

$$\limsup_{n \rightarrow \infty} \frac{\log r(\epsilon, n - N)}{n - N} \leq \lim_{n \rightarrow \infty} \frac{n}{n - N} \log m = \log m$$

Luego $h(\sigma) \leq \log m$.

Ejemplo 9.7.2. La entropía topológica de la herradura de Smale (con 2 patas) es $\log 2$.

La herradura de Smale (con dos patas) es conjugada al shift bilateral de dos símbolos, que tiene entropía topológica $\log 2$. La entropía topológica es invariante por conjugaciones.

Ejemplo 9.7.3. La entropía topológica de la transformación $T(z) = z^k$ en el círculo S^1 , es $\log k$.

La entropía métrica para la medida de Lebesgue es $\log k$, porque T es equivalente, por isomorfismo de espacios de medida, al shift de Bernoulli con vector de probabilidad $(1/k, \dots, 1/k)$. Luego $h(T) \geq \log k$, por el principio variacional de la entropía.

Pero veámoslo directamente, usando separadores: Dado $n \geq 1$, sea S el conjunto de puntos extremos de $2k^n$ intervalos iguales en S^1 , de longitud $1/2k^n$ cada uno. Se tiene $\#S = 2k^n$, y S es un (ϵ, n) separador para todo $\epsilon < 1/2k$. En efecto, dos puntos $x \neq y \in S$, distan por lo menos $1/2k^n$. Aplicando T , la distancia, mientras sea menor que $1/2k$, se multiplica por k . Luego, para algún $j = 0, 1, \dots, n - 2$ se cumple $d(T^j x, T^j y) \geq 1/2k$, o de lo contrario

$$d(T^{n-1} x, T^{n-1} y) = k^{n-1} d(x, y) \geq \frac{k^{n-1}}{2k^n} = \frac{1}{2k} > \epsilon$$

Luego $s(\epsilon, n) \geq 2k^n$, y aplicando la definición de entropía topológica, se tiene $h(T) \geq \log k$.

Ahora demostremos que $h(T) \leq \log k$. Dado $\epsilon > 0$, dividamos S^1 en $M = M(\epsilon)$ intervalos iguales de longitud $1/M < \epsilon$. Dado $n \geq 1$, dividamos cada intervalo en k^n intervalitos iguales de longitud $< \epsilon/k^n$. Sea E el conjunto de puntos extremos de esos intervalitos. Se tiene $\#E = Mk^n$. Afirmamos que E es un (ϵ, n) -generador. En efecto, dado $x \in S^1$, sea $y \in E$ el punto mas próximo a x . Para todo $j = 0, 1, \dots, n - 1$ se cumple:

$$d(T^j x, T^j y) \leq k^j d(x, y) \leq \frac{k^j \epsilon}{k^n} < \epsilon$$

Luego, $r(\epsilon, n) \leq Mk^n$. Aplicando la definición de entropía topológica, se deduce $h(T) \leq \log k$.

Ejemplo 9.7.4. La entropía topológica de cualquier traslación (o rotación rígida) en el toro es nula.

La traslación T en el toro conserva la distancia. Dado ϵ sea \mathcal{V} un cubrimiento del toro con $M = M(\epsilon)$ bolas de radio ϵ . Sea E el conjunto de centros de las bolas de \mathcal{V} . Se tiene $\#E = M$. Además E es un (ϵ, n) -generador, para todo $n \geq 1$. En efecto, dado x en el toro, existe alguna bola de \mathcal{V} que contiene a x . Luego, existe algún $y \in E$, tal que $d(x, y) < \epsilon$. Pero como al aplicar T no se modifica la distancia, se tiene $d(T^j x, T^j y) < \epsilon$ para todo $j \geq 0$.

De lo anterior se deduce que $1 \leq r(\epsilon, n) \leq M$ independiente de n . Luego, tomando logaritmo, y dividiendo entre n , su límite es cero, para todo $\epsilon > 0$. Luego $h(T) = 0$. Nótese que lo anterior demuestra que la entropía topológica de cualquier isometría en un espacio métrico compacto es nula.

Ejemplo 9.7.5. La entropía topológica de cualquier homeomorfismo T del círculo S^1 es nula.

Consideraremos el círculo S^1 como el cociente \mathbb{R}/\sim , con $x \sim y$ en \mathbb{R} si $y - x$ es entero. Orientaremos el círculo, según valores crecientes de x en \mathbb{R} .

Basta demostrar la tesis asumiendo que T preserva la orientación del círculo. En efecto, como $h(\mathbb{T}^2) = 2h(T)$, si T no preserva la orientación de S^1 , podemos sustituirlo por \mathbb{T}^2 que preserva la orientación.

Sea dado ϵ positivo, $\epsilon < 1/2$.

Sea F el conjunto de puntos extremos de $M = M(\epsilon)$ intervalos iguales en S^1 , de longitud $1/M < \epsilon$ cada uno. Tenemos que F es un $(\epsilon, 1)$ generador tal que, dado $x \in S^1$, existe $y \geq x, y \in F$ tal que $y - x < \epsilon$. Demostremos que para todo $n \geq 1$ existe un (ϵ, n) -generador con a lo sumo nM elementos, y tal que dado $x \in S^1$, existe $y \geq x, y \in E$ tal que para todo $j = 0, 1, \dots, n - 1$ se cumple $0 \leq T^j y - T^j x < \epsilon$. La demostración se hará por inducción en n :

Sea E un (ϵ, n) -generador que cumple lo anterior.

Afirmamos que $E \cup T^n F$ es un $(\epsilon, n + 1)$ -generador que también cumple lo anterior. En efecto, dado $x \in S^1$, sea $y \geq x, y \in E$ tal que, para $j = 0, 1, \dots, n - 1$ se cumple $T^j y - T^j x < \epsilon$. Ahora, hay dos posibilidades:

- 1) o bien $T^n y - T^n x < \epsilon$, en cuyo caso, x e y se $(\epsilon, n + 1)$ -acompañan,
- 2) o bien $T^n y - T^n x \geq \epsilon$. En este caso, existe $z \in F$ tal que $T^n x \leq z < T^n y$ y $z - T^n x < \epsilon$. Como T preserva la orientación, para todo $i = 0, 1, \dots, n$ se cumple $T^{n-i} x \leq T^{-i} z < T^{n-i} y$; luego $T^{-i} z - T^{n-i} x < T^{n-1} y - T^{n-i} x < \epsilon$. Entonces, el punto $T^n z \in T^n F$ $(\epsilon, n + 1)$ -acompaña a x .

Como $\#(E \cup T^n F) \leq \#E + M \leq (n + 1)M$, se tiene demostrada la afirmación.

Entonces $r(\epsilon, n) \leq nM$. Tomando logaritmo y dividiendo entre n , su límite, cuando $n \rightarrow \infty$ es cero. Luego $h(T) = 0$.

Ejemplo 9.7.6. La entropía topológica del Anosov lineal $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ en el toro

\mathbb{T}^2 es $\log\left(\frac{3 + \sqrt{5}}{2}\right)$.

El toro \mathbb{T}^2 es el cociente \mathbb{R}^2/\sim con la relación de equivalencia $x \sim y$ si $x - y \in \mathbb{Z}^2$. Denotamos como $\pi(x)$ o como \bar{x} al punto del toro que es la clase de equivalencia de

$x \in \mathbb{R}^2$.

La distancia entre dos puntos \bar{x} y \bar{y} en el toro es igual a la mínima distancia en \mathbb{R}^2 de sus representantes x e y . Luego $d(\bar{x}, \bar{y}) \leq d(x, y)$, para todos x e y en \mathbb{R}^2 . Pero además, si $d(x, y) < 1/2$ entonces $d(\bar{x}, \bar{y}) = d(x, y)$.

Llamando A a la matriz $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, se tiene $T(\bar{x}) = \pi(Ax)$.

Sean $\lambda = (3 + \sqrt{5})/2 > 1$ y $\mu = (3 - \sqrt{5})/2 < 1$ los dos valores propios de A . Sean u y v dos vectores propios de A de valores propios respectivos λ y μ . En \mathbb{R}^2 , la dirección de u , que se dilata con factor $\lambda > 1$, se llama *inestable*, y la dirección de v , que se contrae con factor $\mu < 1$, se llama *estable*.

Para comprender la prueba que sigue, por favor vaya dibujando en \mathbb{R}^2 los conjuntos que se definen:

1a. parte: Prueba de $h(T) \geq \log \lambda$.

Dado $\epsilon > 0$, $\epsilon < 1/2\lambda$, sea I un segmento en \mathbb{R}^2 , según la dirección inestable, con longitud igual a ϵ . Como su longitud es menor que $1/2$, se proyecta inyectivamente en el toro.

Dado $n \geq 1$ se partirá I en s partes iguales, donde $s = \text{ent}(\lambda^n)$. La longitud de cada subintervalo obtenido es $\epsilon/s \geq \epsilon/\lambda^n$.

Sea S el subconjunto formado por los $s + 1$ puntos extremos de los subintervalos construidos en I . Se tiene $\#S = \text{ent}(\lambda^n) + 1 \geq \lambda^n$.

Afirmamos que $\pi(S)$ es un $(\epsilon, n + 1)$ -separador en el toro.

En efecto, aplicando A a dos puntos x e y diferentes de S , como son extremos de un segmento con la dirección inestable, se alejarán multiplicando su distancia por $\lambda > 1$. Luego $d(A^n x, A^n y) = \lambda^n d(x, y) \geq \epsilon$.

Pero esa distancia es en \mathbb{R}^2 y no en el toro. Para ver que se distancian por lo menos ϵ al aplicar T en el toro, tomemos el primer $J \geq 0$ tal que en el plano $d(A^J x, A^J y) \geq \epsilon$. Se tiene $0 \leq J \leq n$.

Si $J = 0$, entonces $d(x, y) = \epsilon < 1/2$, porque la longitud de I es ϵ . Entonces, en el toro $d(\bar{x}, \bar{y}) = d(x, y) = \epsilon$.

Si $J \geq 1$, entonces $d(A^{J-1} x, A^{J-1} y) < \epsilon < 1/2\lambda$. Al aplicar A una vez más, la distancia se multiplica por λ , luego $d(A^J x, A^J y) < 1/2$, y se tiene:

$$d(T^J \bar{x}, T^J \bar{y}) = d(A^J x, A^J y) \geq \epsilon$$

Esto prueba que $\pi(S)$ es un $(\epsilon, n + 1)$ separador en el toro.

Entonces $s(\epsilon, n + 1) \geq \#\pi(S) = \#S \geq \lambda^n$. Tomando logaritmo, dividiendo entre n y haciendo tender n a infinito, se tiene $h(T) \geq \log \lambda$.

2a. parte: Prueba de $h(T) \leq \log \lambda$.

Sea dado $\epsilon > 0$, $\epsilon < 1/2$. Sea en \mathbb{R}^2 un *cuadrículado* formado por infinitas rectas paralelas a la dirección inestable, e infinitas rectas paralelas a la dirección estable, tales que los lados de los rombitos que forman sean exactamente $\epsilon/2$.

Sea Q el conjunto (finito) de los vértices del cuadrículado que están en el rectángulo $[-2, 3]^2$. Sea $q = \#Q$.

Dado $n \geq 1$ dividamos cada lado inestable de los rombitos en r segmentos iguales con $r = \text{ent}(\lambda^n) + 1$. Los lados estables no los dividimos. Cada subintervalo inestable así obtenido tiene longitud $\epsilon/2r < \epsilon/2\lambda^n$.

Sea E el conjunto finito de los puntos extremos de los intervalos así construidos, que están en el rectángulo $[-1, 2]^2$. Se tiene $\#E \leq rq \leq (\lambda^n + 1)q$.

El conjunto $\pi(E)$ en el toro tiene a lo sumo $(\lambda^n + 1)q$ puntos, con q independiente de n .

Afirmamos que $\pi(E)$ es un (ϵ, n) -generador. En efecto, dado $\bar{x} \in \mathbb{T}^2$, sea $x \in \mathbb{R}^2$ su único representante en $[0, 1]^2$. Sea $y \in E$ el punto más próximo de x . Tracemos la recta estable por x y la inestable por y , que se cortan en z . Por construcción, el segmento $[x, z]$ es estable y su longitud es menor que $\epsilon/2$. Análogamente el segmento $[z, y]$ es inestable y su longitud es menor que $\epsilon/2\lambda^n$. Aplicando la matriz A , los segmentos inestables se multiplican por $\lambda > 1$ y los estables por $\mu < 1$, de donde, en \mathbb{R}^2 se obtiene:

$$\begin{aligned} d(A^j x, A^j y) &\leq d(A^j x, A^j z) + d(A^j z, A^j y) = \\ &= \mu^j d(x, z) + \lambda^j d(z, y) \leq \\ &\leq d(x, z) + \frac{\lambda^j}{\lambda^n} \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

si $j = 0, 1, \dots, n-1$. Como $d(T^j \bar{x}, T^j \bar{y}) \leq d(A^j x, A^j y)$, el conjunto $\pi(E)$ es un (ϵ, n) -generador.

Entonces $r(\epsilon, n) \leq (\lambda^n + 1)q$. Tomando logaritmo, dividiendo entre n y haciendo $n \rightarrow \infty$ se obtiene $h(T) \leq \log \lambda$.

9.8. Prueba del Principio Variacional de la Entropía

En esta sección asumimos que X es un espacio métrico compacto y $T : X \mapsto X$ es continua. Denotamos con \mathcal{M} el conjunto de medidas de probabilidad en el espacio medible (X, \mathcal{A}) donde \mathcal{A} es la sigma-álgebra de Borel. Dotamos \mathcal{M} de la topología débil *, y denotamos con $\mathcal{M}_T \subset \mathcal{M}$ el compacto no vacío formado por las medidas de probabilidad invariantes con T . El propósito de esta sección es demostrar el Teorema 9.6.8 que establece el Principio Variacional de la entropía:

$$\sup_{\mu \in \mathcal{M}_T} h_\mu(T) = h_{top}(T). \quad (9.9)$$

Dividiremos la prueba en dos partes:

Lema 9.8.1.

$$\sup_{\mu \in \mathcal{M}_T} h_\mu(T) \leq h_{top}(T)$$

Lema 9.8.2.

$$\sup_{\mu \in \mathcal{M}_T} h_\mu(T) \geq h_{top}(T)$$

Demostración del Lema 9.8.1:

Demostración. Por la definición de entropía topológica, existe $\delta_0 > 0$ tal que

$$\limsup_{n \rightarrow +\infty} \frac{\log s(n, \delta)}{n} \leq h_{top}(T) + 1 \quad \forall 0 < \delta \leq \delta_0, \quad (9.10)$$

donde $s(n, \delta) = \max\{\#S(n, \delta_0)\}$, es la máxima cardinalidad de los (n, δ_0) -separadores $S(n, \delta_0)$ por la transformación T .

Sea $\mu \in \mathcal{M}_T$. Si la única partición finita \mathcal{P} con todas sus piezas de medida positiva, tiene 1 solo elemento, entonces $h_\mu(\mathcal{P}) = 0 \leq h_{top}(T)$ trivialmente. Entonces consideremos el caso en que existe una partición finita $\mathcal{P} = \{P_1, P_2, \dots, P_k\}$ de X con $k \geq 2$ piezas medibles P_i todas con medida positiva. Como P_i es boreliano en un espacio métrico compacto, cualquier medida de probabilidad es regular. Entonces existe $Q_i \subset P_i$ compacto tal que $\mu(Q_i) > 0$ y

$$\mu(P_i \setminus Q_i) < \frac{1}{k \log k}.$$

Sea $0 < \delta < \delta_0$ tal que

$$\text{dist}(Q_i, Q_j) > \delta \text{ para todos } i \neq j \text{ en } \{1, \dots, k\}. \quad (9.11)$$

Consideramos la partición finita

$$\mathcal{Q} := \{Q_0, Q_1, \dots, Q_k\},$$

donde

$$Q_0 = X \setminus (\cup_i Q_i) = \cup_i (P_i \setminus Q_i).$$

Entonces calculamos la entropía condicional de las partición \mathcal{P} dada \mathcal{Q} , siempre con respecto a la misma medida μ :

$$H(\mathcal{P}/\mathcal{Q}) = H(\mathcal{Q}) + \sum_{j=0}^k \sum_{i=1}^k \mu(P_i \cap Q_j) \log \left(\frac{\mu(P_i \cap Q_j)}{\mu(Q_j)} \right). \quad (9.12)$$

Como $Q_i \subset P_i$ si $i \in \{1, \dots, k\}$, tenemos $\log(\mu(P_i \cap Q_i)/\mu(Q_i)) = 1$. Además como las piezas P_i son disjuntas dos a dos, tenemos $Q_j \cap P_i = \emptyset$ si $j \neq i$, y $j \in \{1, \dots, k\}$. Entonces del último término en la Igualdad (9.12) son no nulos solo los sumandos para $j = 0$. Resulta:

$$H(\mathcal{P}/\mathcal{Q}) = \sum_{i=1}^k \mu(P_i \cap Q_0) \log \left(\frac{\mu(P_i \cap Q_0)}{\mu(Q_0)} \right) = \mu(Q_0) \sum_{i=1}^k x_i \log x_i$$

donde $x_i = \mu(P_i \cap Q_0)/\mu(Q_0)$. Entonces $\sum_{i=1}^k x_i = k$ y por la Proposición (9.1.3) deducimos

$$H(\mathcal{P}/\mathcal{Q}) = \mu(Q_0) \sum_{i=1}^k x_i \log x_i \leq \mu(Q_0) \log k.$$

Como $Q_0 = X \setminus (\bigcup_i Q_i)$, tenemos $P_i \cap Q_0 = P_i \setminus Q_i$. Luego por construcción de los Q_i , obtenemos:

$$\mu(P_i \cap Q_0) = \mu(P_i \setminus Q_i) < \frac{1}{k \log k},$$

de donde

$$\mu(Q_0) = \sum_{i=1}^k \mu(P_i \cap Q_0) < \frac{1}{\log k}.$$

Concluimos que

$$H(\mathcal{P}/\mathcal{Q}) \leq \mu(Q_0) \log k < 1. \quad (9.13)$$

Ahora nos proponemos acotar por arriba $h_\mu(\mathcal{P})$, para lo cual debemos acotar por arriba $(1/n)H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P})$. Como $(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}) \vee (\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q})$ es más fina que $\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}$, tenemos:

$$\begin{aligned} H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}) &\leq H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}) \vee (\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) = \\ &= H((\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) + H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}/(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q})) \leq \\ &\leq H((\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) + \sum_{j=0}^{n-1} H(T^{-j}\mathcal{P}/\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) \leq \\ &\leq H((\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) + \sum_{j=0}^{n-1} H(T^{-j}\mathcal{P}/T^{-j}\mathcal{Q}) = \\ &= H((\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) + \sum_{j=0}^{n-1} H(\mathcal{P}/\mathcal{Q}) \leq \\ &\leq H((\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q}) + n. \end{aligned}$$

En la cadena de desigualdades de arriba usamos que la entropía del producto de particiones es menor o igual que la suma de las entropías. Después usamos que la entropía condicional dada una partición más fina, es menor o igual que la entropía condicional dada una partición más gruesa. Luego, aplicamos que μ es T -invariante, por lo que la entropía condicional de $T^{-j}\mathcal{P}$ dada $T^{-j}\mathcal{Q}$ es igual a la entropía condicional de \mathcal{P} dada \mathcal{Q} . Finalmente, en la última desigualdad usamos (9.13). Entonces

$$\begin{aligned} h_\mu(\mathcal{P}) &= \lim_{n \rightarrow +\infty} \frac{H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P})}{n} \leq \\ &\leq \lim_{n \rightarrow +\infty} \frac{n + H(\bigvee_{j=0}^{n-1} T^{-j}\mathcal{Q})}{n} = 1 + h_\mu(\mathcal{Q}). \end{aligned} \quad (9.14)$$

Ahora afirmamos que

$$\mathbf{AProbar} : h_\mu(\mathcal{Q}) \leq 1 + \log 2 + h_{top}(T). \quad (9.15)$$

Primero demostremos que una vez probada la afirmación (9.15), se deduce la desigualdad del Lema 9.8.1. En efecto, reuniendo la afirmación (9.15) con la desigualdad (9.14), obtenemos:

$$h_\mu(\mathcal{P}) \leq 2 + \log 2 + h_{top}(T).$$

para toda partición \mathcal{P} finita. Tomando el supremo en las particiones finitas, deducimos que

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

para toda transformación T continua, y para toda medida T -invariante μ . Entonces, dejando fija la medida T -invariante μ (que también es T^n -invariante cualquiera sea el natural $n \geq 1$), y tomando la transformación T^n en el rol de T , deducimos que

$$h_\mu(T^n) \leq 2 + \log 2 + h_{top}(T^n).$$

Ahora usando los Teoremas 9.5.9 y 9.6.14, obtenemos:

$$nh_\mu(T) \leq 2 + \log 2 + nh_{top}(T) \quad \forall n \geq 1.$$

Dividiendo entre n y haciendo $n \rightarrow +\infty$, deducimos

$$h_\mu(T) \leq h_{top}(T) \quad \forall \mu \in \mathcal{M}_T.$$

Luego

$$\sup_{\mu \in \mathcal{M}_T} h_\mu(T) \leq h_{top}(T)$$

como afirma el Lema 9.8.1. Entonces basta demostrar la afirmación (9.15) para terminar de probar el Lema 9.8.1 como queremos.

Segundo, probemos la afirmación (9.15). Por definición de entropía métrica de una partición, y por la Proposición 9.1.3, tenemos

$$h_\mu(\mathcal{Q}) = \lim_{n \rightarrow +\infty} \frac{H(\vee_{j=0}^{n-1} T^{-j} \mathcal{Q})}{n} \leq \limsup_{n \rightarrow +\infty} \frac{\log \#(\vee_{j=0}^{n-1} T^{-j} \mathcal{Q})}{n}. \quad (9.16)$$

Elijamos un punto x_R en cada pieza no vacía $R \in \vee_{j=0}^{n-1} T^{-j} \mathcal{Q}$. Por definición del producto \vee de particiones, cada pieza R puede escribirse como

$$R = Q_{i_0} \cap T^{-1}Q_{i_1} \cap T^{-2}Q_{i_2} \cap \dots \cap T^{-n}Q_{i_n}, \quad (9.17)$$

donde $i_j \in \{0, 1, 2, \dots, k\}$. Clasifiquemos las diferentes piezas R de la partición $\vee_{j=0}^{n-1} T^{-j} \mathcal{Q}$ (y los respectivos puntos x_R) en las siguientes clases de piezas, dos a dos disjuntas:

- Clase C_0 : Formada por las piezas R tales que ninguno de los Q_{i_j} en la igualdad (9.17) es Q_0 .
- Clase C_1 : Uno y solo uno de los Q_{i_j} es Q_0
- Clase C_2 : Exactamente dos de los Q_{i_j} son Q_0
- ...
- Clase C_h : Exactamente h , con $0 \leq h \leq n$ fijo dado, de las piezas Q_{i_j} , son Q_0 .

...

• Clase C_n : Todos los n conjuntos Q_{i_j} son Q_0 .

A su vez, cada clase C_h , con h fijo, se clasifica que $n!/h!(n-h)!$ subclases dos a dos disjuntas. En efecto, cada combinación $\gamma \in C_h^n$ de h elementos tomado del conjunto de n índices posibles $\{0, 1, 2, \dots, n-1\}$ determina la subclase $C_{h,\gamma}$ formada por las piezas R tales que $Q_{i_j} = Q_0$ si y solo si $i_j \in \gamma$.

Fijemos una de estas subclases $C_{h,\gamma}$. O bien no hay dos piezas diferentes $R \neq R'$ de la partición $\bigvee_{j=0}^{n-1} T^{-j} \mathcal{Q}$ en esta subclase, o bien sí las hay. En el primer caso, la cantidad de piezas en esta subclase es menor o igual que 1, que a su vez es menor o igual que $s_{n,\delta}$ (recordar que $s_{n,\delta}$ denota la cardinalidad máxima de los (n, δ) separadores para T). En el segundo caso, consideremos dos piezas $R \neq R'$ en la subclase $C_{h,\gamma}$. Denotamos con $x \in R$ y $x' \in R'$ los puntos seleccionados en estas dos piezas. Escribimos la igualdad (9.17) para R y R' denotando sin primas los conjuntos Q_{i_j} que corresponden a R y con primas los conjuntos Q'_{i_j} que corresponden a R' . Entonces $R \neq R'$ se diferencian entre sí porque existe algún $i_j \notin \gamma$ tal que $Q'_{i_j} \neq Q_{i_j}$, y ambos conjuntos son diferentes de Q_0 porque $i_j \notin \gamma$. Luego,

$$x \in R \subset T^{-j}(Q_{i_j}), \quad x' \in R' \subset T^{-j}(Q'_{i_j}),$$

de donde, usando la desigualdad (9.11), obtenemos

$$\text{dist}(T^j(x), T^j(x')) \geq \text{dist}(Q_{i_j}, Q'_{i_j}) > \delta.$$

Deducimos que para h, γ fijos, el conjunto de puntos seleccionados en las piezas $R \in C_{h,\gamma}$ es un (n, δ) -separador. Entonces

$$\#C_{h,\gamma} \leq s(n, \delta) \quad \forall 0 \leq h \leq n, \quad \forall \gamma \in C_h^n$$

Luego:

$$\#C_h = \sum_{\gamma \in C_h^n} \#C_{h,\gamma} \leq \sum_{\gamma \in C_h^n} s_{n,\delta} = \frac{n!}{h!(n-h)!} s(n, \delta).$$

Ahora sumamos en h , para contar todas las piezas de la partición $\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{Q})$:

$$\begin{aligned} \#(\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{Q})) &= \sum_{h=0}^n \#C_h \leq \sum_{h=0}^{n-1} \frac{n!}{h!(n-h)!} s(n, \delta) = \\ &= (1+1)^n s(n, \delta) = 2^n s(n, \delta). \end{aligned}$$

Ahora, sustituyendo en la igualdad (9.16), resulta:

$$\begin{aligned} h_\mu(\mathcal{Q}) &\leq \limsup_{n \rightarrow +\infty} \frac{\log \#(\bigvee_{j=0}^{n-1} T^{-j} \mathcal{Q})}{n} \leq \limsup_{n \rightarrow +\infty} \frac{2^n s(n, \delta)}{n} = \\ &= \log 2 + \limsup_{n \rightarrow +\infty} \frac{s(n, \delta)}{n}. \end{aligned}$$

Finalmente aplicando la desigualdad (9.10) establecida al principio, concluimos que

$$h_\mu(\mathcal{Q}) \leq \log 2 + \limsup_{n \rightarrow +\infty} \frac{s(n, \delta)}{n} \leq 1 + \log 2 + h_{top}(T),$$

terminando de demostrar la afirmación (9.15) y el Lema 9.8.1. \square

Demostración del Lema 9.8.2:

Demostración. Si $h_{top}(T) = 0$ la desigualdad de la tesis es trivial. Si $h_{top}(T) > 0$, por definición de entropía topológica, para todo $\delta > 0$ suficientemente pequeño y para todo $n \geq 1$ natural suficientemente grande, existe un (n, δ) generador $S(n, \delta)$ con cardinalidad máxima $s(n, \delta)$ tal que $\log s(n, \delta)/n > 0$.

Para cada $\delta > 0$ y n fijos, elijamos tal generador maximal

$$S(n, \delta) = \{x_1, \dots, x_s\},$$

donde $s = s(n, \delta) \geq 2$. Con $\delta > 0$ fijo, elijamos una subsucesión $n_i \rightarrow +\infty$ tal que

$$\lim_{i \rightarrow +\infty} \frac{\log s(n_i, \delta)}{n_i} = \limsup_{n \rightarrow +\infty} \frac{\log s(n, \delta)}{n}. \quad (9.18)$$

Construyamos las siguientes medidas de probabilidad, no necesariamente T -invariantes:

$$\nu_n := \frac{1}{s} \sum_{h=1}^s \delta_{x_h}, \quad \mu_n := \frac{1}{n} \sum_{j=0}^{n-1} (T^j)^* \nu_n, \quad \mu := \lim_{i \rightarrow +\infty} \mu_{n_i}, \quad (9.19)$$

donde n_i es una subsucesión tal que satisface (9.18) y además tal que $\{\mu_{n_i}\}_i$ es convergente en el espacio \mathcal{M} de las probabilidades con la topología débil*. La medida de probabilidad μ es T -invariante porque $|\mu_n - T^* \mu_n| \leq (\nu_n - (T^*)^n \nu_n)/n$ para todo $n \geq 1$.

Aplicando la Proposición 9.2.4, existe una partición finita \mathcal{P} con piezas medibles tal que $\text{diam}(\mathcal{P}) < \delta$ y el borde $\partial(\mathcal{P})$ tiene μ -medida nula. Como μ es T -invariante, entonces

$$\mu(\partial(T^{-j}(\mathcal{P}))) = 0 \quad \forall j \geq 0, \quad \mu(\partial(\bigvee_{j=0}^{n-1} T^{-j}(\mathcal{P}))) = 0 \quad \forall n \geq 1.$$

Afirmamos que

$$\mathbf{Aprobar} : \limsup_{n \rightarrow +\infty} \frac{\log s(n, \delta)}{n} \leq \frac{H\left(\bigvee_{j=0}^{q-1} T^{-j} \mathcal{P}, \mu\right)}{q} \quad \forall q \geq 1. \quad (9.20)$$

Primero probemos que la afirmación (9.20) implica la desigualdad del Lema 9.8.2 que queremos demostrar. En efecto, tomando límite cuando $q \rightarrow +\infty$ en la igualdad (9.20) y aplicadno la definición de entropía de T para la partición \mathcal{P} y la definición de entropía métrica de T respecto a la medida μ , obtenemos:

$$\limsup_{n \rightarrow +\infty} \frac{\log s(n, \delta)}{n} \leq \lim_{q \rightarrow +\infty} \frac{H\left(\bigvee_{j=0}^{q-1} T^{-j} \mathcal{P}, \mu\right)}{q} = h(\mathcal{P}, \mu) \leq$$

$$\leq \sup_{\mathcal{Q}} h(\mathcal{Q}, \mu) = h_{\mu}(T) \leq \sup_{\mu \in \mathcal{M}_T} h_{\mu}(T).$$

Luego, tomando límite cuando $\delta \rightarrow 0^+$ y aplicando la definición de entropía topológica, concluimos:

$$h_{top}(T) = \lim_{\delta \rightarrow 0^+} \limsup_{n \rightarrow +\infty} \frac{\log s(n, \delta)}{n} \leq \sup_{\mu \in \mathcal{M}_T} h_{\mu}(T),$$

terminando de demostrar el Lema 9.8.2 una vez probada la afirmación (9.20).

Segundo, probemos la afirmación (9.20). Comenzamos fijando los naturales

$$q \geq 1, \quad 1 \leq r \leq q, \quad m \geq 2, \quad 1 \leq (m-1)q \leq n < mq$$

para calcular

$$H(\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}, \nu_n).$$

Como la partición \mathcal{P} tiene diámetro menor que δ , todos los puntos de cada pieza de la partición $\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}$ se δ -acompañan en sus iterados de 0 a $mq+r-1$; en particular hasta el iterado $n-1$. Luego, cada pieza R de la partición $\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}$ contiene a lo sumo, un solo punto del (n, δ) separador maximal $S(n, \delta)$. Entonces $\nu_n(R)$ es o bien igual a cero, o bien igual a $1/s$, donde $s = s(n, \delta) = \#(S(n, \delta))$. Como $\nu_n(X) = 1$ y $\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}$ es una partición, deducimos que la cantidad de sus piezas R tales que $\nu_n(R) = 1/s$, es exactamente s , y las restantes de sus piezas tienen ν_n -medida nula. Entonces:

$$H(\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}, \nu_n) = - \sum_R \nu_n(R) \log \nu(R) = -s \left((1/s) \log(1/s) \right) = \log s. \quad (9.21)$$

Si denotamos

$$\mathcal{P}_q := \bigvee_{i=0}^{q-1} T^{-i} \mathcal{P}$$

tenemos:

$$\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P} = \left(\bigvee_{j=0}^{r-1} T^{-j} \mathcal{P} \right) \vee \left(\bigvee_{j=0}^{m-1} T^{-jq} T^{-r} \mathcal{P}_q \right).$$

Luego, de la igualdad (9.21) obtenemos:

$$\begin{aligned} \log s(n, \delta) &= H(\bigvee_{j=0}^{mq+r-1} T^{-j} \mathcal{P}, \nu_n) = \\ &= H\left(\left(\bigvee_{j=0}^{r-1} T^{-j} \mathcal{P}\right) \vee \left(\bigvee_{j=0}^{m-1} T^{-jq} T^{-r} \mathcal{P}_q\right), \nu_n\right) \leq \\ &\leq H(\bigvee_{j=0}^{r-1} T^{-j} \mathcal{P}, \nu_n) + H(\bigvee_{j=0}^{m-1} T^{-jq} T^{-r} \mathcal{P}_q, \nu_n) \leq \\ &\leq \sum_{j=0}^{r-1} H(T^{-j} \mathcal{P}, \nu_n) + \sum_{j=0}^{m-1} H(T^{-jq} T^{-r} \mathcal{P}_q, \nu_n) \leq \\ &\leq r \cdot \log(\#\mathcal{P}) + \sum_{j=0}^{m-1} H(\mathcal{P}_q, (T^{jq+r})^* \nu_n). \end{aligned}$$

Sumando la desigualdad anterior para $1 \leq r \leq q$, obtenemos:

$$\begin{aligned}
q \cdot \log s(n, \delta) &\leq q^2 \cdot \log(\#\mathcal{P}) + \sum_{r=1}^q \sum_{j=0}^{m-1} H(\mathcal{P}_q, (T^{jq+r})^* \nu_n) = \\
& q^2 \cdot \log(\#\mathcal{P}) + \sum_{j=1}^{qm} \sum_{j=0}^{m-1} H(\mathcal{P}_q, (T^j)^* \nu_n) = \\
&= q^2 \cdot \log(\#\mathcal{P}) + \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n) + \sum_{j=n+1}^{qm} H(\mathcal{P}_q, (T^j)^* \nu_n) \leq \\
&\leq q^2 \cdot \log(\#\mathcal{P}) + \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n) + \sum_{j=n+1}^{qm} \log(\#\mathcal{P}_q) \leq \\
&\leq q^2 \cdot \log(\#\mathcal{P}) + \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n) + (qm - n)q \log(\#\mathcal{P}) \leq \\
&\leq 2 \cdot q^2 \cdot \log(\#\mathcal{P}) + \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n).
\end{aligned}$$

Para la última desigualdad recordamos que $\#\mathcal{P}_q \leq (\#\mathcal{P})^q$ y que $(m-1)q \leq n < mq$, luego $qm - n \leq q$. Dividiendo la desigualdad anterior entre n obtenemos

$$q \cdot \frac{\log s(n, \delta)}{n} \leq \frac{2 \cdot q^2 \cdot \log(\#\mathcal{P})}{n} + \frac{1}{n} \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n).$$

Usando el Teorema 9.2.1 que establece que la entropía de una combinación convexa de medidas es mayor o igual que la combinación convexa de las entropías, deducimos que

$$\begin{aligned}
q \cdot \frac{\log s(n, \delta)}{n} &\leq \frac{2 \cdot q^2 \cdot \log(\#\mathcal{P})}{n} + \frac{1}{n} \sum_{j=1}^n H(\mathcal{P}_q, (T^j)^* \nu_n) \leq \\
&\frac{2 \cdot q^2 \cdot \log(\#\mathcal{P})}{n} + H\left(\mathcal{P}_q, \frac{1}{n} \sum_{j=1}^n (T^j)^* \nu_n\right) = \\
&\leq \frac{2 \cdot q^2 \cdot \log(\#\mathcal{P})}{n} + H(\mathcal{P}_q, \mu_n). \tag{9.22}
\end{aligned}$$

Ahora, consideramos la subsucesión $n_i \rightarrow +\infty$ que cumple las condiciones (9.18) y (9.19). Con $q \geq 1$ fijo, definimos

$$m_i := 1 + \text{parte entera}(n_i/q).$$

Entonces es aplicable a n_i la desigualdad (9.22). De esta igualdad (9.22) tomando $n = n_i \rightarrow +\infty$, con q fijo, deducimos que

$$q \cdot \limsup_{n \rightarrow +\infty} \frac{s(n, \delta)}{n} = q \cdot \lim_{i \rightarrow +\infty} \frac{s(n_i, \delta)}{n_i} \leq \limsup_{i \rightarrow +\infty} H(\mathcal{P}_q, \mu_{n_i}) \quad (9.23)$$

Como $\mu(\partial(\mathcal{P}_q)) = 0$, aplicando el Teorema 9.2.5, μ es un punto de continuidad de la aplicación $\nu \in \mathcal{M} \mapsto H(\mathcal{P}_q, \nu)$. Entonces, como por construcción de μ se cumple $\lim_i \mu_{n_i} = \mu$, deducimos que existe el siguiente límite

$$\lim_{i \rightarrow +\infty} H(\mathcal{P}_q, \mu_{n_i}) = H(\mathcal{P}, \mu).$$

Entonces, sustituyendo en (9.23) concluimos

$$q \cdot \limsup_{n \rightarrow +\infty} \frac{s(n, \delta)}{n} \leq H(\mathcal{P}_q, \mu). \quad (9.24)$$

Esto termina de probar la Afirmación (9.20) como queríamos, y por lo tanto termina la demostración del Lema 9.8.2. \square

Capítulo 10

Apéndice Existencia del primer subespacio de Oseledets

El objetivo de este apéndice es demostrar el Teorema 4.5.1 que afirma la existencia de un primer fibrado no nulo de Oseledets, de los exponentes de Lyapunov en todas las direcciones de este fibrado, y la existencia de su fibrado invariante complementario. Este Teorema 4.5.1, constituye la parte esencial no trivial de la prueba, expuesta en la Sección 4.6, del Teorema 3.9.8 de Oseledets o Teorema Ergódico Multiplicativo 4.6.2.

10.1. Demostración del Teorema 4.5.1: Primer Subespacio de Oseledets

Sea $f : M \mapsto M$ un difeomorfismo en una variedad compacta y Riemanniana de dimensión finita. Sea μ una medida de probabilidad f -invariante. Sea E un subfibrado medible y df -invariante del fibrado tangente tal que $E_x \neq \{0\}$ para un conjunto de puntos $x \in M$ con medida μ positiva.

Observación 10.1.1. Para probar el Teorema 4.5.1, primero consideramos el conjunto $A_0 = \{x \in M : E_x \neq \{0\}\}$. Como E es medible y df -invariante, entonces $A_0 \subset M$ es un conjunto medible e invariante por f . Por hipótesis $\mu(A_0) > 0$. Construimos la medida de probabilidad μ_1 definida por

$$\mu_1(B) = \frac{\mu(A_0 \cap B)}{\mu(A_0)} \quad \forall B \subset M \text{ medible .}$$

Entonces μ_1 es invariante por f (porque el conjunto fijo A_0 lo es) y cumple $\mu_1(A_0) = 1$. En lo que sigue sustituimos la medida μ por μ_1 , así no es restrictivo asumir que

$$E_x \neq \{0\} \quad \mu - \text{c.t.p. } x \in X.$$

Ahora enunciamos los siguientes lemmas:

Lema 10.1.2. *En las hipótesis del Teorema 4.4.8, para todo $\epsilon > 0$, definimos el “cono” $F(2\epsilon)$ por*

$$F_x(2\epsilon) := \{0\} \cup \left\{ 0 \neq v \in E_x : \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}(v)\|/\|v\|)}{n} \leq -\chi_E^+(x) + 2\epsilon \right\}. \quad (10.1)$$

Entonces:

- (A) $F_x(2\epsilon)$ es un subespacio vectorial de E_x para todo x .
- (B) $F_x(2\epsilon) \neq \{0\}$ para μ -c.t.p. x .
- (C) Para todo $\epsilon' > 0$ existe $N = N(\epsilon', x) \geq 1$ (uniforme en $0 \neq v \in F_x(2\epsilon)$) tal que

$$\frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} \leq -\chi_E^+(x) + 2\epsilon + \epsilon' \quad \forall n \geq N \quad \forall 0 \neq v \in F_x(2\epsilon).$$

Recordemos la Definición 4.4.11 de transformación lineal invertible medible $L : E \mapsto E$ en un subfibrado medible E (no necesariamente invariante por df). El Lema 10.1.2 es un caso particular del siguiente Lema, tomando $L = df|_E$:

Lema 10.1.3. *En las hipótesis del Teorema 4.4.12, para todo $\epsilon > 0$ definimos el “cono” $F(2\epsilon)$ por*

$$F_x(2\epsilon) := \{0\} \cup \left\{ 0 \neq v \in E_x : \limsup_{n \rightarrow +\infty} \frac{\log(\|L_x^{-n}(v)\|/\|v\|)}{n} \leq -\chi_{E,L}^+(x) + 2\epsilon \right\}. \quad (10.2)$$

Entonces

- (A) $F_x(2\epsilon)$ es un subespacio vectorial de E_x para todo x .
- (B) $F_x(2\epsilon) \neq \{0\}$ para μ -c.t.p. x .
- (C) Para todo $\epsilon' > 0$ existe $N = N(\epsilon', x) \geq 1$ (uniforme en $0 \neq v \in F_x(2\epsilon)$) tal que

$$\frac{\log(\|L^{-n}v\|/\|v\|)}{n} \leq -\chi_{E,L}^+(x) + 2\epsilon + \epsilon' \quad \forall n \geq N \quad \forall 0 \neq v \in F_x(2\epsilon).$$

El Lema 10.1.2 es el punto crucial en la demostración del Teorema 4.5.1. Lo probaremos a a partir del párrafo 10.2.1, como consecuencia del Lema de Pliss 10.2.2.

En esta sección admitimos los Lemas 10.1.2 y 10.1.3, y solo veremos cómo ellos implican el Teorema 4.5.1.

10.1.4. Demostración del Teorema 4.5.1 (admitiendo los Lema 10.1.2 y 10.1.3):

Paso 1: Prueba de partes (a), (b) y (c):

Lema 10.1.5. *Existe un primer subespacio de Oseledets E_x^1 para μ -c.t.p. $x \in M$, que satisface las condiciones (a), (b) y (c) del Teorema 4.5.1.*

Demostración. Del Lemma 10.1.2, parte (A), teniendo en cuenta que la dimensión de $T_x M$ es finita, y observando que $F(2\epsilon') \subset F(2\epsilon)$ para todos los reales $0 < \epsilon' < \epsilon$, se deduce que existe $\epsilon_0 > 0$ tal que

$$F_x(2\epsilon) = F_x(2\epsilon_0) \quad \forall 0 < \epsilon \leq \epsilon_0. \quad (10.3)$$

En efecto, no puede haber infinitos subespacios diferentes, uno contenido en el otro, en un espacio vectorial de dimensión finita. Entonces $F_x(2\epsilon)$ es un subespacio vectorial *constante* (independiente de ϵ) y además de dimensión mayor o igual que 1 para μ -c.t.p., porque coincide con $F_x(2\epsilon_0)$ que no es nulo (cf. Lema 10.1.2). Luego, el subfibrado E^1 definido por

$$E_x^1 := \bigcap_{0 < \epsilon < \epsilon_0} F_x(2\epsilon) = F_x(2\epsilon_0) \quad \mu - \text{c.t.p.} \quad (10.4)$$

tiene dimensión mayor o igual que 1 porque existe $0 \neq v \in F_x(2\epsilon_0)$ debido al Lema 10.1.2. Además, como E_x^1 es la intersección de los subespacios $F_{2\epsilon}$, para todo $\epsilon > 0$ arbitrariamente pequeño, entonces está caracterizado por la siguiente igualdad:

$$E_x^1 = \{0\} \cup \left\{ 0 \neq v \in E_x : \limsup_{n \rightarrow +\infty} \frac{\log(\|df^{-n}(v)\|/\|v\|)}{n} \leq -\chi_E^+ \right\}. \quad (10.5)$$

Entonces E^1 es df -invariante, porque χ_E^+ es una función f -invariante, el fibrado dado E es df -invariante, y el límite superior en (10.5) para $0 \neq v \in E_x$ es el mismo que para $0 \neq u = df_x v \in df_x E_x = E_{f(x)}$.

Además, de la igualdad (10.5) se desprende la medibilidad del subfibrado E^1 . En efecto, primero obsérvese que E^1 es seccionalmente cerrado porque E_x es un subespacio vectorial de $T_x M$ y todo subespacio de $T_x M$ es cerrado en $T_x M$. Segundo, el límite superior en la igualdad (10.5) es una función medible de $TM \mapsto \mathbb{R}$ porque es el límite de funciones medibles; $\chi_E^+ : M \mapsto \mathbb{R}$ es una función medible por el Teorema 4.4.8. La preimagen de $(-\infty, a]$ por la suma de dos funciones reales medibles, es medible. Entonces, el subconjunto A en $TM \setminus \{0\}$ donde el límite superior de (10.5) más χ_E^+ es no positivo, es medible. Finalmente, como por hipótesis el fibrado dado E es medible, $E^1 = \{0\} \cup (A \cap E)$ es medible.

Aplicando las desigualdades (4.70) y (4.71), la propiedad de que el ínfimo de los límites inferiores es menor o igual que el supremo de los límites superiores, y la caracterización del subfibrado E^1 dada por la igualdad (10.5), obtenemos:

$$\begin{aligned} -\chi_{E^1}^+(x) &= \lim_{n \rightarrow +\infty} \frac{\log \inf \|df_x^{-n}|_{E_x^1}\|}{n} \leq \\ &\leq \inf_{0 \neq v \in E_x^1} \liminf_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}(v)\|/\|v\|)}{n} \leq \\ &\leq \sup_{0 \neq v \in E_x^1} \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}(v)\|/\|v\|)}{n} \leq -\chi_E^+(x) \end{aligned} \quad (10.6)$$

Además, usando nuevamente la igualdad (a) del Teorema 4.4.8, para E en lugar de E^1 , aplicando la definición de ínfimo y teniendo en cuenta que el ínfimo en un subconjunto $E^1 \subset E$ es mayor o igual que el ínfimo en E , se tiene:

$$\begin{aligned} -\chi_E^+ &= \lim_{n \rightarrow +\infty} \frac{\log \inf(\|df_x^{-n}|_{E_x}\|)}{n} \leq \\ &\leq \lim_{n \rightarrow +\infty} \frac{\log \inf(\|df_x^{-n}|_{E_x^1}\|)}{n} = -\chi_{E^1}^- \end{aligned}$$

Luego, deducimos que

$$\chi_{E^1}^+ = \chi_E^+ \mu - \text{c.t.p.},$$

y entonces, las desigualdades en (10.6) son igualdades. Deducimos que existe el siguiente límite (exponente de Lyapunov hacia el pasado) para todo $0 \neq v \in E_x^1$, es independiente de la dirección $v \in E_x^1$, e igual a $\chi_{E^1}^+ = \chi_E^+$:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}v\|/\|v\|)}{-n} = \chi_{E^1}^+(x) = \chi_E^+(x) \quad \forall 0 \neq v \in E_x^1. \quad (10.7)$$

De la parte (C) del Lema 10.1.2, deducimos que la convergencia por abajo al límite de la igualdad (10.7) es uniforme en $v \in E^1 \setminus \{0\}$. En efecto, dado $\epsilon > 0$, como $E^1 \subset F(2\epsilon)$, tomamos $\epsilon' = \epsilon$ en la parte (C) del Lema 10.1.2.

Aplicando el Corolario 4.4.9, deducimos que para todo $0 \neq v \in E_x^1$ existen los siguientes límites (exponentes de Lyapunov en el futuro y en el pasado), son iguales entre sí, e iguales a $\chi_{E^1}^-(x) = \chi_{E^1}^+(x) = \chi_E^+(x)$:

$$\chi_1(x) = \lim_{n \rightarrow +\infty} \frac{\log(\|df^n(v)\|/\|v\|)}{n} = \lim_{m \rightarrow -\infty} \frac{\log(\|df^m(v)\|/\|v\|)}{m}.$$

Además, por el Teorema 4.4.8, $\chi_E^+(x)$ es una función medible de x . Luego el exponente de Lyapunov $\chi_1(x)$ en el subespacio E_x^1 , es una función medible de x .

Esto termina de probar la existencia del primer subespacio de Oseledecs E_x^1 , demostrando así las partes (a), (b) y (c) del Teorema 4.5.1 (una vez que el Lema 10.1.2 esté probado). \square

Ahora probaremos las partes (d) y (e) del Teorema 4.5.1, referente al cono G complementario de E^1 en E . Recordamos que para μ -c.t.p. G es medible, pues está definido por

$$G = \{0\} \cup (E \setminus E^1).$$

Paso 2: Prueba de la parte (d) del Teorema 4.5.1

Lema 10.1.6. *G es invariante por df y para todo subfibrado vectorial medible $F^1 \subset G$ invariante con df y para μ -c.t.p. x tal que $F_x^1 \neq \{0\}$ se verifica la siguiente desigualdad:*

$$\sup_{0 \neq v \in F_x^1} \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^n v\|/\|v\|)}{n} \leq \chi_{F^1}^+(x) < \chi_{E^1}^+(x) = \chi_E^+ \quad (10.8)$$

Demostración. G es medible e invariante por df porque es, para cada punto $x \in M$, el complemento en E de E^1 , y estos dos últimos son fibrados medibles e invariantes por df . Ahora probemos que todo subfibrado medible F^1 invariante con df contenido en G satisface (10.8):

La desigualdad a la izquierda en (10.8) se cumple debido a (4.70). La igualdad a la derecha en (10.8) se cumple porque, por lo probado antes, el subfibrado E^1 satisface la parte (c) del Teorema 4.5.1. Asumamos, por absurdo, que la desigualdad del medio en (10.8) no se cumple para un conjunto de puntos x con μ medida positiva. Entonces,

$$\chi_{F^1}^+(x) \geq \chi_E^+(x). \tag{10.9}$$

Como $F_x^1 \subset G_x \subset E_x$, entonces el supremo de cualquier conjunto de números para los vectores en F_x^1 , es menor o igual que el supremo para los vectores en E_x . Luego, aplicando nuevamente la parte (a) del Teorema 4.4.8:

$$\chi_{F^1}^+(x) = \lim_{n \rightarrow +\infty} \frac{\log \sup \|df_x^n|_{F_x^1}\|}{n} \leq \lim_{n \rightarrow +\infty} \frac{\log \sup \|df_x^n|_{E_x}\|}{n} = \chi_E^+(x).$$

Juntando ésta con la desigualdad (10.9), resulta

$$\chi_{F^1}^+(x) = \chi_E^+(x).$$

Entonces, aplicando a F^1 en lugar de E , las partes (b) y (c) ya demostradas del Teorema 4.5.1, para μ -c.t.p. x tal que $F_x^1 \neq \{0\}$ donde se cumple la igualdad $\chi_{F^1}^+(x) = \chi_E^+(x)$, existiría un subespacio

$$E_x^2 \subset F_x^1 \subset G_x = \{0\} \cup (E_x \setminus E_x^1) \tag{10.10}$$

con $\dim(E_x^2) \geq 1$ y tal que:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} = -\chi_{F^1}^+(x) = -\chi_E^+(x) \quad \forall 0 \neq v \in E_x^2.$$

Pero, por la caracterización del subespacio $E^1(x)$ dada en la ecuación (10.5), resultaría $E_x^2 \subset E_x^1$, lo cual contradice (10.10).

Esto termina de demostrar la parte (d) del Teorema 4.5.1. □

Ahora probaremos la última parte (e) del Teorema 4.5.1. Para poder demostrarla necesitamos generalizar las partes (a), (b) y (c) ya probadas del Teorema 4.5.1, al caso en que la transformación lineal $L = E \mapsto E$ no es necesariamente el diferencial $df|_E$.

Paso 3. Generalización de partes (a), (b), (c):

Lema 10.1.7. *En las hipótesis del Teorema 4.4.12, existe un subfibrado $F^{(1)} \subset F$, invariante por L , es decir*

$$L_x F^{(1)} = F_{f(x)}^{(1)}$$

tal que:

(a') No nulidad de F^1 :

$$\dim(F_x^1) \geq 1 \quad \text{para } \mu\text{-c.t.p. tal que } F_x \neq \{0\}.$$

(b') Existencia del exponente de Lyapunov $\chi_{1,L}$ en F^1 :

Para μ -c.t.p. tal que $F_x \neq \{0\}$, para todo $0 \neq v \in F_x^1$, existen los siguientes límites (exponentes de Lyapunov $\chi_{1,L}^+(x)$ y $\chi_{1,L}^-(x)$, para el futuro y para el pasado respectivamente, en todas las direcciones de F_x^1), son independientes de la dirección v , e iguales entre sí:

$$\chi_{1,L}^+(x) = \lim_{n \rightarrow +\infty} \frac{\log(\|L^n(v)\|/\|v\|)}{n} = \chi_{1,L}^-(x) = \lim_{n \rightarrow +\infty} \frac{\log(\|L^{-n}(v)\|/\|v\|)}{-n}$$

(c') Maximalidad de $\chi_{1,L}$:

El exponente de Lyapunov $\chi_{1,L}^+(x) = \chi_{1,L}^-(x)$ en las direcciones de F_x^1 (que denotamos como $\chi_{1,L}(x)$), satisface:

$$\chi_{1,L}(x) := \chi_{F,L}^+(x) = \chi_{F^1,L}^-(x) = \chi_{F^1,L}^+(x),$$

(donde $\chi_{F,L}^+$, $\chi_{F^1,L}^+$ y $\chi_{F^1,L}^-$ están definidos en el Teorema 4.4.12) para los subfibrados F y F^1 respectivamente.

Demostración. Basta repetir el paso 1) de la demostración de las partes (a), (b), (c) del Teorema 4.5.1, escribiendo F en lugar de E , la transformación lineal $L_x : F_x \mapsto F_{f(x)}$ en lugar de $df_x : E_x \mapsto E_{f(x)}$, usando el Teorema 4.4.12 en vez del Teorema 4.4.8, y el Lema 10.1.3 en vez del Lema 10.1.2. \square

Finalmente, probemos ahora la última parte (e) del Teorema 4.5.1:

Paso 4. Construcción del subfibrado complementario F^1

Lema 10.1.8. *Existe un subfibrado F^1 medible e invariante con df tal que*

$$E_x^1 \oplus F_x^1 = E_x \quad \mu\text{-c.t.p. } x \in M.$$

Demostración. Si $E_x^1 = E_x$ μ -c.t.p., la parte (e) del Teorema 4.5.1 es trivial tomando $F_x = \{0\}$ para todo $x \in M$.

Consideremos el caso en que el conjunto $A_1 := \{x \in M : E_x^1 \subsetneq E_x\}$ cumple $\mu(A_1) > 0$. Como A_1 es un conjunto f invariante, porque el subfibrado E_x y el subfibrado E_x^1 lo son, entonces, no es restrictivo asumir que la medida invariante μ cumple $\mu(A_1) = 1$. (En caso contrario sustituimos μ por la medida $\mu_1(B) := \mu(B \cap A_1)/\mu(A_1)$ para todo boreliano $B \subset M$.)

Sea $E^{1,\perp}$ el subfibrado tal que para todo $x \in M$ el subespacio $E_x^{1,\perp}$ es el complemento ortogonal de E_x^1 en E_x . Entonces para μ -c.t.p. se cumple $E_x^{1,\perp} \neq \{0\}$. (En los puntos donde $E_x^1 = E_x$ definimos $E_x^{1,\perp} = \{0\}$.)

El fibrado $E^{1,\perp}$ es medible porque E y E^1 lo son. En general $E^{1,\perp}$ no es df -invariante, pero como E^1 lo es, entonces

$$\dim(E_x^{1,\perp}) = \dim(E_{f(x)}^{1,\perp}) \quad \forall x \in M.$$

La esencia de esta demostración radica en probar que existe el siguiente límite de subfibrados:

$$\text{A probar: } \exists F^1 = \lim_{n \rightarrow +\infty} df^{-n} E^{1,\perp} \quad \text{y} \quad F_x^1 \cap E_x^1 = \{0\} \quad \forall x \in M. \quad (10.11)$$

La primera afirmación de existencia del límite, en extenso, significa que para todo $x \in M$ existe en $T_x M$ el siguiente límite de subespacios

$$F_x^1 = \lim_{n \rightarrow +\infty} df_{f^n(x)}^{-n} E_{f^n(x)}^{1,\perp},$$

donde la topología para definir el límite de una sucesión de subfibrados es la del Grassmaniano $\mathcal{G}(T_x M)$ dada en la Definición 4.4.1.

De la afirmación (10.11) se deduce inmediatamente que $F^1 \oplus E^1 = E_x$ porque, para cada $x \in M$ fijo, si existe el límite de una sucesión de subespacios (todos de la misma dimensión complementaria a E_x^1 en E_x), el subespacio límite tiene esa misma dimensión. Además de la afirmación (10.11) se deduce que F^1 es medible, porque es el límite de una sucesión de subfibrados medibles. Finalmente, se deduce también que el subfibrado F^1 es df -invariante porque

$$df_x F_x^1 = df_x \lim_{n \rightarrow +\infty} df_{f^n(x)}^{-n} E_{f^n(x)}^{1,\perp} = \lim_{n \rightarrow +\infty} df_{f^n(x)}^{-n+1} E_{f^n(x)}^{1,\perp} = F_{f(x)}^1.$$

En la última igualdad hemos usado que, para $x \in M$ fijo, la transformación $df_x : T_x M \mapsto T_{f(x)} M$ es continua, por ser una transformación lineal.

Por lo tanto, para terminar de demostrar el Lema 10.1.8 y el Teorema 4.5.1, basta demostrar la afirmación (10.11).

Observamos que el grafo $G(A_x)$ de toda aplicación lineal $A_x : E_x^{1,\perp} \mapsto E_x^1$ es un subespacio de E_x con dimensión igual a $E_x^{1,\perp}$ y que interseca a E_x^1 solo en $\{0\}$. Es decir $G(A_x) \oplus E_x^1 = E_x$ para todo $x \in M$. Recíprocamente, todo subfibrado G tal que $G_x \oplus E_x^1 = E_x$ para todo $x \in M$, es tal que cada fibra G_x es el grafo de una aplicación lineal $A_x : E_x^{1,\perp} \mapsto E_x^1$. En efecto, si $G_x \oplus E_x^1 = E_x$, entonces la proyección ortogonal $\pi_x|_{G_x} : G_x \mapsto E_x^{1,\perp}$ sobre $E_x^{1,\perp}$ es una transformación lineal invertible. Definimos $A_x = \pi'_x|_{G_x} \circ (\pi_x|_{G_x})^{-1} : E_x^{1,\perp} \mapsto E_x^1$, donde π'_x es la proyección ortogonal sobre E_x^1 . Por construcción tenemos que la fibra G_x es el grafo de A_x .

Entonces, toda sucesión $A^{(n)}$ tal que $A_x^{(n)}$ es una transformación lineal $E_x^{1,\perp} \mapsto E_x^1$ para todo $x \in M$, define una sucesión de fibrados $G(A^{(n)})$ complementarios a E^1 (no necesariamente medibles). Supongamos que una tal sucesión $A_x^{(n)}$ cumple

$$\lim_{n \rightarrow +\infty} A_x^{(n)} = A_x \quad \forall x \in M$$

para cierta transformación lineal $A_x : E_x^{1,\perp} \mapsto E_x^1$, donde el límite de una sucesión de transformaciones lineales está definido por la topología inducida por la norma usual de transformaciones lineales:

$$\|A_x\| := \max_{0 \neq u \in E_x^{1,\perp}} \|A_x u\| \quad \forall A_x \in \mathcal{L}(E_x^{1,\perp}, E_x^1) \quad \text{si } E_x^{1,\perp} \neq \{0\}.$$

(Si $E_x^{1,\perp} = \{0\}$ entonces la transformación lineal es solo 0, y es siempre convergente, porque es constante.) Es estándar chequear que entonces el grafo $G(A_x)$ es el límite de los subespacios grafos $G(A_x^{(n)})$, con la topología del Grassmaniano $\mathcal{G}(E_x)$ dado en la Definición 4.4.1. Es decir:

$$\exists A_x = \lim_{n \rightarrow +\infty} A_x^{(n)} \Rightarrow G(A_x) = \lim_{n \rightarrow +\infty} G(A_x^{(n)}).$$

Deducimos que para probar la afirmación (10.11) (y por lo tanto para terminar de probar el Lema 10.1.8) es suficiente demostrar lo siguiente:

A probar: $\forall n \geq 0, \forall x \in M :$

$$\exists df_{f^n(x)}^{-n} E_{f^n(x)}^{1,\perp} = G(A_x^{(n)}), \quad \exists A_x = \lim_{n \rightarrow +\infty} A_x^{(n)}. \quad (10.12)$$

Es inmediato chequear que $G(0) = E^{1,\perp}$. Es decir, definimos

$$A_x^{(0)} = 0 \in \mathcal{L}(E_x^{1,\perp}, E_x^1) \quad \forall x \in M. \quad (10.13)$$

Veamos cómo construir $A_x^{(n)}$ para $n \geq 1$.

Consideremos la transformación lineal

$$L : E^{1,\perp} \mapsto E^{1,\perp} \quad L_x(u) := \pi_{f(x)} df_x(u) \quad \forall u \in E^{1,\perp}$$

donde $\pi_z : E_z \mapsto E_z^{1,\perp}$ es la proyección ortogonal sobre el subespacio $E_z^{1,\perp}$, para todo $z \in M$. (Si $E^{1,\perp} = \{0\}$, π_z es la transformación lineal nula). Siendo df continua y π medible (porque E^1 es medible), la transformación L es medible, de acuerdo a la Definición 4.4.11. Como el subfibrado E^1 , ortogonal a $E^{1,\perp}$, es df -invariante, deducimos

$$L^n = \pi df^n \quad \forall n \in \mathbb{Z}. \quad (10.14)$$

Sea $P : E^{1,\perp} \mapsto E^1$ tal que $P_x \in \mathcal{L}(E_x^{1,\perp} \mapsto E_{f^{-1}(x)}^1)$ definida por

$$P_x u := df_x^{-1} u - \pi_{f^{-1}(x)} df_x^{-1} u = df_x^{-1} u - L_x^{-1} u \quad \forall u \in E_x^{1,\perp}. \quad (10.15)$$

Entonces

$$\begin{aligned} df_x^{-1}(G(A_x)) &= \{df_x^{-1}(u + A_x u) : u \in E_x^{1,\perp}\} = \\ &= \{L_x^{-1} u + P_x u + df_x^{-1} A_x u : u \in E_x^{1,\perp}\} = \\ &= \{v + P_x L_{f^{-1}(x)} v + df_x^{-1} A_x L_{f^{-1}(x)} v : v \in E_{f^{-1}(x)}^{1,\perp}\} = \\ &= G(P_x L_{f^{-1}(x)} + df_x^{-1} A_x L_{f^{-1}(x)}) \end{aligned}$$

Hemos probado que para todo subespacio $G(A_x)$ complementario a E_x^1 en E^x , el subespacio $df_x^{-1}(G(A_x))$ (complementario a $E_{f^{-1}(x)}^1$ en $E_{f^{-1}(x)}$), es el grafo de la aplicación lineal $P_x L_{f^{-1}(x)} + df_x^{-1} A_x L_{f^{-1}(x)}$.

Denotamos como $F_x(A_x)$ a la siguiente aplicación lineal:

$$\begin{aligned} F_x(A_x) &:= P_x L_{f^{-1}(x)} + df_x^{-1} A_x L_{f^{-1}(x)} \in \mathcal{L}(E_{f^{-1}(x)}^{1,\perp}, E_{f^{-1}(x)}^1) \\ &\forall A_x \in \mathcal{L}(E_x^{1,\perp}, E_x^1). \end{aligned} \quad (10.16)$$

Tenemos

$$df_x^{-1}(G(A_x)) = G(F_x^* A_x) \quad \forall x \in M, \forall A_x \in \mathcal{L}(E_x^{1,\perp}, E_x^1).$$

Luego

$$df_{f^n(x)}^{-n} G(A_{f^n(x)}) = G(F_{f^n(x)}^n A_{f^n(x)}) \quad \forall n \geq 0,$$

$$\forall x \in M, \forall A_{f^n(x)} \in \mathcal{L}(E_{f^n(x)}^{1,\perp}, E_{f^n(x)}^1),$$

donde $F_{f^n(x)}^n := F_{f(x)} \circ F_{f^2(x)} \circ \dots \circ F_{f^n(x)}$. En particular tomando $A_z^{(0)} = 0$ para todo $z \in M$ deducimos

$$df_{f^n(x)}^{-n} E_{f^n(x)}^{1,\perp} = G(A_x^{(n)}) \quad \text{donde} \quad A_x^{(n)} = F_{f^n(x)}^n(A_{f^n(x)}^{(0)}) = F_{f^n(x)}^n(0). \quad (10.17)$$

Para terminar de demostrar la afirmación (10.12) (y por lo tanto para terminar de demostrar el Lema 10.1.8), basta ahora demostrar que la sucesión $A_x^{(n)} \in \mathcal{L}(E_x^{1,\perp}, E_x^1)$ es convergente con la norma de las transformaciones lineales en el espacio $\mathcal{L}(E_x^{1,\perp}, E_x^1)$ para cada punto $x \in M$.

Sean $A^{(0)}, B^{(0)}$ tales que $A_x^{(0)}, B_x^{(0)} \in \mathcal{L}(E_x^{1,\perp}, E_x^1)$ para todo $x \in M$ y tales que para todo $x \in M$ existe una constante $C(x) > 0$ tal que $\|A_{f^n(x)}^{(0)}\|, \|B_{f^n(x)}^{(0)}\| \leq C(x) \forall n \in \mathbb{Z}$. Sea $n \geq 1$. Usando la fórmula (10.17) aplicada también a $B_x^{(n)}$, y la fórmula (10.16), obtenemos:

$$\begin{aligned} A_x^{(1)} - B_x^{(1)} &= F_{f(x)}(A_{f(x)}^{(0)}) - F_{f(x)}(B_{f(x)}^{(0)}) = df_{f(x)}^{-1}(A_{f(x)}^{(0)} - B_{f(x)}^{(0)})L_x. \\ A_x^{(n)} - B_x^{(n)} &= F_{f(x)}(A_{f(x)}^{(n-1)}) - F_{f(x)}(B_{f(x)}^{(n-1)}) = \\ &= df_{f(x)}^{-1}(A_{f(x)}^{(n-1)} - B_{f(x)}^{(n-1)})L_x. \end{aligned}$$

Luego, por inducción en $n \geq 1$ tenemos:

$$\begin{aligned} A_x^{(n)} - B_x^{(n)} &= (\Pi_{j=1}^n df_{f^j(x)}^{-1})(A_{f^n(x)}^{(0)} - B_{f^n(x)}^{(0)})(\Pi_{j=0}^{n-1} L_{f^j(x)}) = \\ &= df_{f^n(x)}^{-n}(A_{f^n(x)}^{(0)} - B_{f^n(x)}^{(0)})L_x^n, \end{aligned}$$

donde $L_x^n = L_{f^{n-1}(x)} \dots L_{f(x)} L_x$. Luego:

$$\|A_x^{(n)} - B_x^{(n)}\| \leq \frac{\|A_{f^n(x)}^{(0)} - B_{f^n(x)}^{(0)}\| \cdot \sup_{0 \neq u \in E_x^{1,\perp}} \{\|L_x^n(u)\|/\|u\|\}}{\inf_{0 \neq v \in E_x^1} \{df_x^n(v)/\|v\|\}}. \quad (10.18)$$

Dado $\epsilon > 0$ arbitrariamente pequeño, aplicando la parte A) del Teorema 4.4.12 al subfibrado $E^{1,\perp}$, existe $N_1 = N_1(x, \epsilon)$ tal que

$$\sup_{0 \neq u \in E_x^{1,\perp}} \{\|L_x^n(u)\|/\|u\|\} \leq \exp n(\chi_{E^{1,\perp}, L}^+ + \epsilon) \quad \forall n \geq N_1. \quad (10.19)$$

Análogamente, aplicando la parte A) del Teorema 4.4.8 al subfibrado invariante E^1 , existe $N_2 = N_2(x, \epsilon)$ tal que

$$\inf_{0 \neq v \in E_x^1} \{\|df_x^n(v)\|/\|v\|\} \geq \exp n(\chi_{E^1, df}^- - \epsilon) \quad \forall n \geq N_1. \quad (10.20)$$

Tomando $N = N(\epsilon, x) = \max\{N_1(x, \epsilon), N_2(x, \epsilon)\}$ y sustituyendo (10.19) y (10.20) en la desigualdad (10.18), resulta:

$$\|A_x^{(n)} - B_x^{(n)}\| \leq \exp n(2\epsilon + \chi_{E^{1\perp}, L}^+ - \chi_{E^1, df}^-) \cdot \|A_x^{(0)} - B_x^{(0)}\| \quad \forall n \geq N. \quad (10.21)$$

Usando la Parte (c) ya demostrada del Teorema 4.5.1 tenemos

$$\chi_{E^1, df}^- = \chi_{E^1, df}^+ = \chi_E^+. \quad (10.22)$$

Afirmamos que

$$\chi_{E^{1\perp}, L}^+ < \chi_E^+ \quad (\text{a probar}). \quad (10.23)$$

Entonces, para terminar la demostración del Lema 10.1.8 basta probar los dos siguientes enunciados (i) y (ii):

ii) La desigualdad (10.23) es verdadera.

i) La desigualdad (10.23) implica que existe $A_x = \lim_{n \rightarrow +\infty} A_x^{(n)} \in \mathcal{L}(E_x^{1\perp}, E_x^1)$.

Por lo tanto, una vez demostrados los enunciados (i) y (ii), la afirmación (10.12) quedará probada, terminando la demostración del Lema 10.1.8.

i) (10.23) \Rightarrow $\exists A_x = \lim_{n \rightarrow +\infty} A_x^{(n)}$.

Para probar la afirmación (i), construimos

$$\epsilon := \frac{\chi_E^+ - \chi_{E^{1\perp}}^-}{3} > 0. \quad (10.24)$$

Recordamos que $\chi_E^+ - \chi_{E^{1\perp}}^-$ dependen de x , pero es invariante en la órbita por f . Por lo tanto, fijado $x \in M$, el número ϵ es constante a lo largo de toda la órbita de x . Sustituyendo (10.22) y (10.23) en la desigualdad (10.21), obtenemos:

$$\|A_x^{(n)} - B_x^{(n)}\| \leq e^{-n\epsilon} \|A_x^{(0)} - B_x^{(0)}\| \quad \forall n \geq N$$

para cierto $N \geq 1$. Entonces existe $N_0 \geq 1$ (que depende de x pero que para cada x fijo es constante a lo largo de toda la órbita de x) tal que

$$\|A_x^{(N_0)} - B_x^{(N_0)}\| \leq (1/2) \|A_{f^{N_0}(x)}^{(0)} - B_{f^{N_0}(x)}^{(0)}\| \quad (10.25)$$

La afirmación anterior vale para $f^n(x)$ en lugar de x para todo $n \in \mathbb{Z}$, debido a que N_0 es constante a lo largo de la órbita de x . Es decir:

$$\|A_{f^j(x)}^{(N_0)} - B_{f^j(x)}^{(N_0)}\| \leq (1/2) \|A_{f^{N_0+j}(x)}^{(0)} - B_{f^{N_0+j}(x)}^{(0)}\| \quad \forall j \in \mathbb{Z}. \quad (10.26)$$

La fórmula (10.26) vale para todas las aplicaciones $A^{(0)}, B^{(0)} : E^{1,\perp} \mapsto E^1$ tales que $A_z^{(0)}, B_z^{(0)} \in \mathcal{L}(E_z^{1,\perp}, E_x^1) \forall z \in M$. En la fórmula (10.13) tomamos, en particular, $A^{(0)} = 0$, es decir $A_z^{(0)} = 0$ para todo $z \in M$. Fijemos el punto $x \in M$ y definamos la constante

$$C(x) = \sup\{\|A_z^{(j)}\| : 0 \leq j \leq N_0 - 1 = N_0(x) - 1\}.$$

Tal supremo existe en \mathbb{R} porque para $j = 0$ la transformación lineal $A_z^{(0)}$ es 0, y la aplicación F_z que lleva $A_{f(z)}^{(j)}$ en $A_{f(z)}^{(j+1)}$ está acotada uniformemente para todo $z \in M$. Tomemos cualquier aplicación $B^{(0)} : E^{1\perp} \mapsto E^1$ tal que

$$B_z^{(0)} \in \mathcal{L}(E_z^{1\perp}, E_z^1) \quad \forall z \in M \quad \text{y} \quad \|B_{f^j(x)}^{(0)}\| \leq C(x) \quad \forall j \in \mathbb{Z}.$$

El espacio $\Sigma_{C(x)}$ de tales aplicaciones es una bola compacta con la norma $\|B\| = \sup_{j \in \mathbb{Z}} \|B_{f^j(x)}\|$. Aplicando la fórmula (10.26) y el teorema del punto fijo de Banach deducimos que existe $A_x \in \mathcal{L}(E_x^{1\perp}, E_x^1)$ tal que

$$\lim_{n \rightarrow +\infty} B_x^{(nN_0)} = A_x \quad \forall B^{(0)} \in \Sigma_{C(x)}.$$

En particular para cada $r \in \{0, 1, \dots, N_0 - 1\}$ si elegimos $B_z^{(0)} = A_z^{(r)}$ para todo $z \in M$, obtenemos

$$\lim_{n \rightarrow +\infty} A_x^{nN_0+r} = \lim_{n \rightarrow +\infty} B_x^{nN_0} = A_x.$$

Como todo $m \in \mathbb{N}$ es de la forma $nN_0 + r$ para algún $0 \leq r \leq N_0 - 1$, concluimos que

$$\lim_{m \rightarrow +\infty} A_x^m = A_x,$$

terminando la prueba de (i).

ii) Prueba de la desigualdad (10.23): Por la parte (A) del Teorema 4.4.12, usando la igualdad (10.14), y observando que la proyección ortogonal π_x sobre $E_x^{1\perp}$ tiene norma menor o igual que 1, obtenemos:

$$\begin{aligned} \chi_{E^{1\perp}, L}^+(x) &= \lim_{n \rightarrow +\infty} \frac{\sup \|L_x^n\|}{n} = \lim_{n \rightarrow +\infty} \frac{\sup \|\pi_{f^n(x)} df_x^n\|}{n} \leq \\ & \lim_{n \rightarrow +\infty} \frac{\sup \|df_x^n\|}{n} = \chi_E^+(x). \end{aligned}$$

Entonces, para probar (10.23), argumentamos por absurdo, suponiendo que $\chi_{E^{1\perp}, L} = \chi_E^+$. Usando el lema 10.1.7 (con $E^{1,\perp}$ en lugar de F), si $E_x^{1\perp} \neq \{0\}$, existe un vector u (que dejamos fijo) tal que

$$0 \neq u \in E_x^{1\perp}, \tag{10.27}$$

y

$$\lim_{n \rightarrow +\infty} \frac{\log(\|L_x^{-n}u\|/\|u\|)}{n} = -\chi_{E^{1\perp}, L}^+(x).$$

Por hipótesis de absurdo obtenemos:

$$\lim_{n \rightarrow +\infty} \frac{\log(\|L_x^{-n}u\|/\|u\|)}{n} = -\chi_E^+(x). \tag{10.28}$$

Aplicando recursivamente la igualdad (10.15) obtenemos:

$$df_x^{-1}u = P_x u + L_x^{-1}u, \tag{10.29}$$

$$\begin{aligned}
df_x^{-2}u &= df_{f^{-1}(x)}^{-1}P_xu + P_{f^{-1}(x)}L_x^{-1}u + L_x^{-2}u \\
df_x^{-n}u &= df_{f^{-1}(x)}^{-(n-1)}P_xu + df_{f^{-2}(x)}^{-(n-2)}P_{f^{-1}(x)}L_x^{-1}u + \dots + \\
&+ df_{f^{-(n-1)}(x)}^{-1}P_{f^{-(n-2)}(x)}L_x^{-(n-2)}u + P_{f^{-(n-1)}(x)}L_x^{-(n-1)}u + L_x^{-n}u \quad \forall n \geq 0. \quad (10.30)
\end{aligned}$$

Observando que $P_xu \in E_{f^{-1}(x)}^1$ y $L_x^{-1}u \in E_{f^{-1}(x)}^{1\perp}$ obtenemos, de la igualdad (10.29) y por Pitágoras $\|P_xu\| \leq \|df_x^{-1}u\|$. Luego

$$\|P_xu\| \leq k \|u\|,$$

donde $k = \max\{\|df_z^{-1}\|: z \in M\}$. Luego, tomando norma en la igualdad (10.30), deducimos:

$$\begin{aligned}
\frac{\|df_x^{-n}u\|}{\|u\|} &\leq \frac{\|L_x^{-n}u\|}{\|u\|} + \\
\sum_{j=1}^n \sup_{0 \neq v \in E_{f^{-j}(x)}^1} &\{ \|df_{f^{-j}(x)}^{-(n-j)}v\|/\|v\| \} k (\|L_x^{-j+1}u\|/\|u\|). \quad (10.31)
\end{aligned}$$

De la igualdad (10.28) deducimos que existe una constante $c(x) > 0$ tal que

$$\|L_x^{-n}u\|/\|u\| \leq c(x) \exp n(\chi_E^+ + \epsilon) \quad \forall n \geq 1. \quad (10.32)$$

Aplicando a f^{-1} en vez de f la Definición 4.4.2 de la función ψ_n y la Definición 4.4.5 de la función supremo $H(x)$ y usando las igualdades (4.68) y (4.69), obtenemos:

$$\sup_{0 \neq v \in E_{f^{-j}(x)}^1} \{ \|df_{f^{-j}(x)}^{-(n-j)}v\|/\|v\| \} \leq H(f^{-j}(x)) \exp((n-j)(\chi_{E^1}^- + \epsilon)). \quad (10.33)$$

De (10.22) obtenemos:

$$\sup_{0 \neq v \in E_{f^{-j}(x)}^1} \{ \|df_{f^{-j}(x)}^{-(n-j)}v\|/\|v\| \} \leq H(f^{-j}(x)) \exp((n-j)(\chi_E^+ + \epsilon)). \quad (10.34)$$

Sustituyendo (10.32) y (10.34) en la desigualdad (10.31):

$$\begin{aligned}
\frac{\|df_x^{-n}u\|}{\|u\|} &\leq c(x) \exp(n(\chi_E^+ + \epsilon)) + \\
k c(x) \sum_{j=1}^n \exp((j-1)(\chi_E^+ + \epsilon)) &H(f^{-j}(x)) \exp((n-j)(\chi_E^+ + \epsilon)) = \\
c(x) \exp(n(\chi_E^+ + \epsilon)) + k c(x) \exp((n-1)(\chi_E^+ + \epsilon)) &\sum_{j=1}^n H(f^{-j}(x)) \\
\forall n \geq 1. & \quad (10.35)
\end{aligned}$$

La función $H(x)$ tiene crecimiento subexponencial (c.f. Definición 4.3.1 y Proposición 4.4.6). Entonces, puede elegirse $N \geq 2$ suficientemente grande tal que

$$H(f^{-j}(x)) \leq \exp(j\epsilon) \quad \forall j \geq N.$$

Luego

$$\sum_{j=1}^n H(f^{-j}(x)) \leq \sum_{j=1}^{N-1} H(f^{-j}(x)) + \sum_{j=N}^n e^{j\epsilon} \leq C(x) + (n-1)e^{n\epsilon}, \quad \forall n \geq N$$

donde $C(x) = \sum_{j=1}^{N-1} H(f^{-j}(x))$. Para n suficientemente grande tenemos, además, $C(x) < e^{n\epsilon}$. Luego, existe $N \geq 1$ tal que:

$$\sum_{j=1}^n H(f^{-j}(x)) \leq ne^{n\epsilon}, \quad \forall n \geq N.$$

Sustituyendo en (10.35) obtenemos:

$$\frac{\|df_x^{-n}u\|}{\|u\|} \leq c(x) \exp(n(\chi_E^+ + \epsilon)) +$$

$$c(x) k n e^{-\chi_E^+} \exp(n(\chi_E^+ + 2\epsilon)) \quad \forall n \geq N.$$

Entonces, existe $c^*(x) > 0$ tal que

$$\frac{\|df_x^{-n}u\|}{\|u\|} \leq c^*(x) \exp(n(\chi_E^+ + 2\epsilon)) \quad \forall n \geq N.$$

Tomando logaritmo, dividiendo entre n y haciendo $n \rightarrow +\infty$, deducimos que

$$\limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}u\|/\|u\|)}{n} \leq \chi_E^+ + 2\epsilon.$$

Como $\epsilon > 0$ es arbitrario, deducimos que

$$\limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}u\|/\|u\|)}{n} \leq \chi_E^+ \quad (10.36)$$

Usando la caracterización del subespacio E^1 dada por la igualdad (10.5), de (10.36) deducimos que $u \in E_x^1$, lo cual contradice (10.27), terminando de probar, por absurdo, la desigualdad (10.23) y finalizando la prueba del Lema 10.1.8. \square

Hemos terminado de probar el Teorema 4.5.1 asumiendo verdaderos los Lemas 10.1.2 y 10.1.3. Entonces, para que quede completa la demostración de dicho Teorema solo resta probar estos dos Lemas en forma independiente, lo que haremos en la próxima sección.

10.2. Lema de Pliss y consecuencias

En esta sección demostraremos, sin usar los resultados obtenidos en la sección anterior, los Lemas 10.1.2 y 10.1.3. Estos dos Lemas son consecuencia del llamado Lema de Pliss. Primero reduciremos la prueba de los dos Lemas 10.1.2 y 10.1.3 al Lema de Pliss, y al final de esta sección demostraremos el Lema de Pliss.

Solo expondremos explícitamente la prueba del Lema 10.1.2. Su generalización (el Lema 10.1.3) se demuestra repitiendo la prueba del Lema 10.1.2, usando la transformación $L : E \mapsto E$ en vez de $df|_E$.

10.2.1. Demostración del Lema 10.1.2

Sea $F(2\epsilon)$ el cono definido por la igualdad (10.3):

$$F_x(2\epsilon) := \{0\} \cup \left\{ \limsup_{n \rightarrow +\infty} \frac{\log(\|df_x^{-n}(v)\|/\|v\|)}{n} \leq -\chi_E^+(x) + 2\epsilon \right\} \quad (10.37)$$

Demostración de la parte (A) del Lema 10.1.2:

$F(2\epsilon)$ es fibrado vectorial.

Demostración. Para demostrar que $F_x(2\epsilon)$ es un subespacio vectorial de E_x , basta repetir la demostración del paso 4 del Teorema 4.5.1, para el subfibrado $F(2\epsilon)$ en el rol de F^1 , usando la igualdad (10.3), usando $-\chi_E^+ + 2\epsilon$ (con $\epsilon > 0$ fijo), en lugar de χ_E^+ , y usando $\epsilon' > 0$ cualquiera, en lugar de $\epsilon = (1/2) \min\{\Delta_1, \Delta_2\}$. \square

Dejaremos la prueba de la parte (B) del Lema 10.1.2 para el final. Primero probaremos su parte (C):

Demostración de la parte (C) del Lema 10.1.2

Tenemos que demostrar la aproximación uniforme (en los vectores $0 \neq v \in F_x(2\epsilon)$ al límite superior acotado por $-\chi_E^+ + 2\epsilon$. Más precisamente:

Para μ -c.t.p. $x \in M$, dado $\epsilon' > 0$, existe $N = N(x)$ (independiente de $0 \neq v \in F_x(2\epsilon)$), tal que:

$$\frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} \leq -\chi_E^+(x) + 2\epsilon + \epsilon' \quad \forall n \geq N \quad \forall 0 \neq v \in F_x \quad (10.38)$$

Demostración. Por simplicidad, a lo largo de esta prueba, denotaremos F_x al subespacio $F_x(2\epsilon)$. Por la parte (A) F_x es un subespacio vectorial de E_x . En los puntos x tales que la dimensión de F_x es nula, no existen vectores $0 \neq v \in F_x$ y no hay nada que probar. Para cada uno de los demás puntos x , existe una base ortonormal v_1, \dots, v_k del subespacio F_x , donde

$$k = \dim(F_x)$$

depende de x . Elegimos la base ortonormal; es decir

$$\|v_i\| = 1, \quad \langle v_i, v_j \rangle = 0 \quad \forall 1 \leq i \neq j \leq k.$$

Dejamos fijo el punto x y tal base v_1, \dots, v_k . Por definición del subespacio F_x , para cada v_i se cumple:

$$\limsup_{n \rightarrow +\infty} \frac{\log \|df_x^{-n} v_i\|}{n} \leq -\chi_E^+(x) + 2\epsilon.$$

Luego, dado $\epsilon' > 0$ existe $N_i \geq 1$ tal que

$$\frac{\log \|df_x^{-n} v_i\|}{n} < -\chi_E^+(x) + 2\epsilon + \frac{\epsilon'}{2} \quad \forall n \geq N_i. \quad (10.39)$$

Sea

$$N = \max \left\{ \frac{2 \log \dim(F_x)}{\epsilon'}, \max\{N_i : 1 \leq i \leq k\} \right\} \geq 1. \quad (10.40)$$

Para probar la aproximación uniforme, demostraremos que para este natural N , se satisface la desigualdad (10.38) para todo $n \geq N$ y para todo $0 \neq v \in F_x$. No es restrictivo asumir que

$$\|v\| = 1,$$

ya que el límite superior a la izquierda de la desigualdad (10.38) no cambia si se multiplica el vector no nulo v por un escalar no nulo. Como $\{v_1, \dots, v_k\}$ generan F_x , existen reales $\{a_i\}$ tales que:

$$v := \sum_{i=1}^k a_i v_i$$

Como $\|v\| = 1$ y la base $\{v_i\}_{1 \leq i \leq k}$ es ortonormal, tenemos:

$$\langle v, v_j \rangle = a_j, \quad |a_i| \leq \|v\| \cdot \|v_i\| = 1, \quad \sum_{i=1}^k |a_i| \leq \sum_{i=1}^k 1 = k = \dim(F_x). \quad (10.41)$$

Por la linealidad de df_x^{-n} y la propiedad triangular de la norma tenemos

$$\|df_x^{-n} v\| = \left\| \sum_{i=1}^k a_i df_x^{-n} v_i \right\| \leq \sum_{i=1}^k |a_i| \|df_x^{-n} v_i\| \quad \forall n \geq 0 \quad (10.42)$$

Multiplicando por n y aplicando la función exponencial a ambos miembros de la desigualdad (10.39), se obtiene:

$$\|df_x^{-n} v_i\| < e^{n(-\chi_E^+(x) + 2\epsilon + (\epsilon'/2))} \quad \forall n \geq N \geq N_i, \quad \forall 1 \leq i \leq k. \quad (10.43)$$

Reuniendo las desigualdades (10.42) y (10.43), y luego aplicando la desigualdad (10.41) resulta, para todo $n \geq N$:

$$\begin{aligned} \|df_x^{-n} v\| &\leq \sum_{i=1}^k |a_i| \|df_x^{-n} v_i\| \leq e^{n(-\chi_E^+(x) + 2\epsilon + (\epsilon'/2))} \left(\sum_{i=1}^k |a_i| \right) \leq \\ &\leq e^{n(-\chi_E^+(x) + 2\epsilon + (\epsilon'/2))} \dim(E_x) = \end{aligned}$$

$$\begin{aligned}
&= e^{n(-\chi_E^+(x) + 2\epsilon + (\epsilon'/2))} e^{\log \dim(E_x)} \\
&\quad \forall n \geq N \quad \forall v \in F_x, \quad \|v\| = 1.
\end{aligned} \tag{10.44}$$

Tomando logaritmo en el primero y último miembro de la cadena de desigualdades (10.44), se deduce:

$$\begin{aligned}
\log \|df_x^{-n}v\| &\leq n(-\chi_E^+(x) + 2\epsilon + (\epsilon'/2)) + \log \dim(F_x) \\
&\quad \forall n \geq N \quad \forall v \in F_x, \quad \|v\| = 1.
\end{aligned}$$

Dividiendo entre n y usando la desigualdad (10.40), $(1/n) \leq (1/N) \leq \epsilon'/(2 \log \dim(F_x))$, resulta:

$$\begin{aligned}
\frac{\log \|df_x^{-n}v\|}{n} &\leq -\chi_E^+(x) + 2\epsilon + \frac{\epsilon'}{2} + \frac{\log \dim(F_x)}{n} \leq \\
&\leq -\chi_E^+(x) + 2\epsilon + \frac{\epsilon'}{2} + \frac{\epsilon'}{2} = -\chi_E^+(x) + 2\epsilon + \epsilon' \\
&\quad \forall n \geq N \quad \forall v \in F_x, \quad \|v\| = 1.
\end{aligned}$$

Entonces, para cualquier vector $0 \neq w \in F_x$, observando que $\|df^{-n}w/\|w\|\| = \|df^{-n}w/\|w\|\|$ y que $w/\|w\| = v$ cumple $\|v\| = 1$, de la desigualdad anterior se deduce:

$$\begin{aligned}
\frac{\log(\|df_x^{-n}w/\|w\|\|)}{n} &\leq -\chi_E^+(x) + 2\epsilon + \epsilon' \\
&\quad \forall n \geq N \quad \forall 0 \neq w \in F_x,
\end{aligned}$$

probando la desigualdad (10.38) como queríamos, y terminando la demostración de la parte (C) del Lema 10.1.2. \square

Demostración de la parte (B) del Lema 10.1.2:

No nulidad de $F(2\epsilon)$ para μ -c.t.p. donde $E \neq \{0\}$

En la Observación 10.1.1, vimos que no es restrictivo asumir que

$$E_x \neq \{0\} \quad \mu - \text{c.t.p.}$$

Entonces, en esta parte (A) probaremos que $F_x(2\epsilon) \neq \{0\}$ μ -c.t.p.

Paso 1: Conjuntos $A_n(\epsilon)$ de ϵ -aproximación.

Por la parte (a) del Teorema 4.4.8, tenemos:

$$\lim_{n \rightarrow +\infty} \frac{\log \inf \|df_x^{-n}|_{E_x}\|}{n} = -\chi_E^+(x) \quad \mu - \text{c.t.p.} \tag{10.45}$$

Fijamos $\epsilon > 0$. Usando la igualdad (10.45), para μ -c.t.p. $x \in M$ existen números naturales j , arbitrariamente grandes tales que

$$\frac{\log \inf \|df_x^{-j}|_{E_x}\|}{j} \leq -\chi_E^+(x) + \epsilon.$$

Como E_x es un subespacio vectorial, existe un vector $0 \neq v \in E_x$ tal que

$$\inf \|df_x^{-j}|_{E_x}\| = \frac{\|df_x^{-j}v\|}{\|v\|}.$$

Deducimos que el siguiente subfibrado $V_x(j, \epsilon) \subset E_x$ no es nulo para μ -c.t.p. $x \in M$ si $j \geq N_0(x)$:

$$V_x(j, \epsilon) := \{0\} \cup \{0 \neq v \in E_x : \frac{\log(\|df_x^{-j}v\|/\|v\|)}{j} \leq -\chi_E^+(x) + \epsilon\} \neq \{0\} \quad (10.46)$$

Para todo $n \geq 1$ definimos el siguiente conjunto:

$$A_n(\epsilon) := \{x \in M : V_x(j, \epsilon) \neq \{0\} \ \forall j \geq n\}. \quad (10.47)$$

$\{A_n\}_{n \geq 1}$ es una sucesión creciente de conjuntos. Por lo probado al principio, para μ -c.t.p. $x \in M$ existe $n = N_0(x) \geq 1$ tal que $V_x(j, \epsilon) \neq \{0\}$ para todo $j \geq n$. Deducimos que μ -c.t.p. $x \in A_n$ para algún $n \geq 1$, es decir:

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n(\epsilon)\right) = 1 \ \forall \ \mu \in \mathcal{M}_f \quad (10.48)$$

Paso 2: Conjuntos X_m de 2ϵ -aproximación por tramos.

Dados fijos $\epsilon > 0$ y un natural $m \geq 1$, construimos el siguiente conjunto $X_m(2\epsilon) \subset M$:

$$X_m(2\epsilon) := \left\{y \in M : \bigcap_{j=1}^m V_y(j, 2\epsilon) \neq \{0\}\right\}, \quad (10.49)$$

donde $V_y(j, 2\epsilon)$ es el subfibrado definido en (10.46). Dicho de otra forma, $y \in X_m(2\epsilon)$ si y solo si existe un vector no nulo $u \in E_y$ tal que:

$$\frac{\log(\|df_y^{-j}u\|/\|u\|)}{j} \leq -\chi_E^+(y) + 2\epsilon \ \forall j \in \{1, \dots, m\}. \quad (10.50)$$

Es decir, un *mismo* vector no nulo $v \in E_y$ está en todos los subfibrados $V_y(j, 2\epsilon)$ de 2ϵ -aproximación, para todos los valores de j en el tramo que va desde 1 hasta m . Entonces, $X_m(2\epsilon)$ es el conjunto de puntos donde existen vectores no nulos para los cuales la 2ϵ -aproximación al límite superior $-\chi_E^+$ se produce en todo el tramo (para cada uno de los *mismos* vectores) de valores naturales de j entre 1 y m . Se observa que, aunque en la cola para valores de $j > m$ no se afirma nada sobre la aproximación, la longitud finita del tramo donde se satisface la 2ϵ -aproximación es tanto mayor cuanto mayor es m .

Pero que no se afirme nada sobre la aproximación para valores de $j > m$, es un problema resoluble del siguiente modo:

Si existiera algún punto y :

$$y \in \bigcap_{m \geq N_0} X_m(2\epsilon), \quad (10.51)$$

entonces existiría, para cada $m \geq N_0$, un vector no nulo $u_m \in E_y$ que cumple (10.50) para cada m fijo que se elija.

Con y y m fijos, el conjunto $U_{m,y}$ de vectores $u \in E_y$ tales que $\|u\| = 1$ y que cumplen (10.50) para todo $j \in \{1, \dots, m\}$, es compacto en E_y (porque E_y es un subespacio y df_y^{-j} es una transformación lineal invertible para cada j). Además la sucesión $\{U_{m,y}\}_{m \geq N_0}$ (con y fijo) de conjuntos de vectores en E_y , es decreciente al crecer $m \geq N_0$. Entonces deducimos (usando la propiedad de intersecciones finitas de compactos) que $\bigcap_{m \geq N_0} U_{m,y} \neq \emptyset$. Luego, *si existe* algún punto y en el conjunto intersección definido en (10.51), entonces, para este punto y , existe un vector de $u \in \bigcap_{m \geq N_0} U_{m,y}$, con norma 1. Es decir, existe $0 \neq u \in E_y$ tal que

$$\frac{\log(\|df_y^{-j}u\|/\|u\|)}{j} \leq -\chi_E^+(y) + 2\epsilon \quad \forall j \in \{1, \dots, m\} \quad \forall m \geq N_0.$$

Esto significa que:

$$\frac{\log(\|df_y^{-j}u\|/\|u\|)}{j} \leq -\chi_E^+(y) + 2\epsilon \quad \forall j \in \mathbb{N}^+. \quad (10.52)$$

Luego

$$\limsup_{j \rightarrow +\infty} \frac{\log(\|df_y^{-j}u\|/\|u\|)}{j} \leq -\chi_E^+(y) + 2\epsilon,$$

probando la no nulidad del subespacio $F_y(2\epsilon)$, c.f. igualdad (10.37), como queríamos demostrar.

En resumen, hemos probado que:

$$\bigcap_{m \geq 1} X_m(2\epsilon) \subset B := \{y \in M : F_y(2\epsilon) \neq \{0\}\}. \quad (10.53)$$

El argumento desarrollado para probar la no nulidad del subfibrado $F_y(2\epsilon)$ vale solamente para los puntos y (si existe alguno) que están en el conjunto intersección $\bigcap_{m \geq 1} X_m(2\epsilon)$. Pero nosotros queremos demostrar que el subespacio $F_y(2\epsilon) \neq \{0\}$ para μ -c.t.p. $y \in M$.

Entonces, para demostrar la parte (B) del Lema 10.1.2, bastaría probar que el conjunto intersección $\bigcap_{m \geq 1} X_m(2\epsilon)$, tiene medida μ igual a 1. No probaremos tal afirmación tan fuerte, sino una aparentemente más débil, que es la siguiente:

$$\text{A demostrar: } \exists \delta > 0 \text{ tal que } \mu\left(\bigcap_{m \geq N_0} X_m(2\epsilon)\right) \geq \delta \quad \forall \mu \in \mathcal{M}_f. \quad (10.54)$$

donde $X_m(2\epsilon) \subset M$ es el conjunto de aproximación definido en la igualdad (10.49), y $\delta > 0$ es independiente de la medida de probabilidad μ invariante por f .

Afirmamos que basta probar (10.54) para poder deducir que $F(2\epsilon) \neq \{0\}$ μ -c.t.p. para toda $\mu \in \mathcal{M}_f$, lo cual terminaría de demostrar la parte (B) del Lema 10.1.2.

En efecto, si es cierta la afirmación (10.54), entonces en particular para toda medida de probabilidad ν ergódica para f , tenemos

$$\nu\left(\bigcap_{m \geq N_0} X_m(2\epsilon)\right) \geq \delta > 0.$$

Por la inclusión (10.53), deducimos que

$$\nu(B) = \nu\left(\{x \in M : F_x(2\epsilon) \neq \{0\}\}\right) \geq \delta > 0.$$

Pero el subespacio $F_x(2\epsilon)$ es df -invariante, porque el límite superior en la fórmula (10.1) es invariante con f y la función χ_E^+ también lo es (c.f. parte d) del Teorema 4.4.8).

Entonces, el conjunto B es invariante con f . Como tiene medida positiva $\nu(B) \geq \delta > 0$, para cualquier medida de probabilidad ergódica ν , entonces B tiene medida 1 para cualquier probabilidad ergódica ν . Luego, por el Teorema 4.1.2 de Desintegración Ergódica, el conjunto B donde el subfibrado $F(2\epsilon)$ no es nulo tiene probabilidad total (medida 1 para cualquier $\mu \in \mathcal{M}_f$), terminando de demostrar el Lema 10.1.2 (módulo la prueba de la afirmación (10.54)).

Paso 3: Reducción del problema al Lema de Pliss

Ahora el objetivo es probar (10.54). Para ello reduciremos la demostración a la prueba de un lema (Lema de Pliss) que enunciaremos y demostraremos al final.

Demostración. Denotamos

$$X(2\epsilon) := \bigcap_{m=1}^{+\infty} X_m(2\epsilon)$$

$$A(\epsilon) := \bigcup_{n=1}^{+\infty} A_n(\epsilon),$$

donde $X_m(2\epsilon)$ es el conjunto de 2ϵ -aproximación en el tramo $[1, m]$, definido por la igualdad (10.49), y $A_n(\epsilon)$ es el conjunto de ϵ -aproximación en la cola $[n, +\infty)$, definido por la igualdad (10.47). En el paso 1 probamos la afirmación (10.48):

$$\mu(A(\epsilon)) = 1 \quad \forall \mu \in \mathcal{M}_f.$$

Queremos probar (10.54), o sea estimar por abajo la medida $\mu(X(2\epsilon))$ para toda $\mu \in \mathcal{M}_f$. Como por ahora $\epsilon > 0$ está fijo, por simplicidad omitiremos provisoriamente la escritura de 2ϵ y ϵ en los conjuntos $X(2\epsilon)$ y $A(\epsilon)$ respectivamente.

De la Fórmula (10.49) que define los conjuntos X_m , observamos que la sucesión $\{X_m\}_{m \geq 1}$ es decreciente con m ; es decir, $X_{m+1} \subset X_m \quad \forall m \geq 1$. Denotemos ξ_{X_m} a la función característica del conjunto X_m . Tenemos

$$\xi_{X_{m+1}} \leq \xi_{X_m} \quad \forall m \geq 1.$$

Entonces, para cualquier medida de probabilidad μ se cumple

$$\mu(X) = \lim_{m \rightarrow +\infty} \mu(X_m).$$

Denotemos ξ_{X_m} a la función característica del conjunto X_m . Por el teorema de Birkhoff, y por convergencia dominada, para toda medida de probabilidad μ invariante con f , tenemos:

$$\begin{aligned} \mu(X) &= \lim_{m \rightarrow +\infty} \mu(X_m) = \lim_{m \rightarrow +\infty} \int \xi_{X_m} d\mu = \\ &= \lim_{m \rightarrow +\infty} \int \tilde{\xi}_{X_m} d\mu \end{aligned}$$

donde $\tilde{\xi}_{X_m}$ denota el límite de los promedios de Birkhoff de la función ξ_{X_m} con $m \geq 1$ fijo, para el difeomorfismo f^{-1} (si μ es f -invariante, entonces es también f^{-1} -invariante: luego podemos aplicar el teorema de Birkhoff para f^{-1} en lugar de f).

Como $\mu(A) = 1$ para toda $\mu \in \mathcal{M}_f$, entonces:

$$\mu(X) = \lim_{m \rightarrow +\infty} \int_{x \in A} \tilde{\xi}_{X_m}(x) d\mu.$$

Ahora tenemos en cuenta que el promedio de Birkhoff de la función característica $\xi_{X_m}(x)$ hasta tiempo n por f^{-1} , en cada punto $x \in M$, es la frecuencia con la que los puntos del tramo de órbita $\{f^{-k}(x)\}_{0 \leq k \leq n-1}$ cayeron en X_m . Entonces:

$$\begin{aligned} \tilde{\xi}_{X_m}(x) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{m-1} \chi_{X_m}(f^{-k}(x)) = \\ &= \lim_{n \rightarrow +\infty} \frac{\#\{1 \leq k \leq n: f^{-k}(x) \in X_m\}}{n}. \end{aligned}$$

Juntando esto con lo anterior, obtenemos:

$$\begin{aligned} \mu(X) &= \\ &= \lim_{m \rightarrow +\infty} \int_{x \in A} \lim_{n \rightarrow +\infty} \frac{\#\{1 \leq k \leq n: f^{-k}(x) \in X_m\}}{n} d\mu \\ &\quad \forall \mu \in \mathcal{M}_f \end{aligned} \tag{10.55}$$

Supongamos demostrada la siguiente propiedad:

Afirmación X:

Existe constante $0 < \delta \leq 1$, independiente del punto x , de la medida de probabilidad μ , de m y de n , tal que:

Para todo $m \geq 1$, para todo $x \in A$, existe $n_1 = n_1(x, m)$ tal que, para todo $n \geq n_1$ se cumple:

$$\#\{1 \leq k \leq n: f^{-k}(x) \in X_m\} \geq \delta n \quad (10.56)$$

Si demostramos la afirmación X, entonces, sustituyendo en la igualdad (10.55), obtendríamos inmediatamente que

$$\mu(X) \geq \delta \quad \forall \mu \in \mathcal{M}_f,$$

demostrando (10.54) como queríamos, y terminando de probar el Lema 10.1.2.

Entonces ahora el objetivo es demostrar la Afirmación X. La reformularemos como sigue, incorporando ahora de vuelta el parámetro $\epsilon > 0$ en la notación y los enunciados:

Afirmación Y:

Para todo $\epsilon > 0$ existen constantes $N_0 = N_0(\epsilon) \geq 1$ y $0 < \delta = \delta(\epsilon) \leq 1$ (independientes del punto x , de la medida de probabilidad μ , de n y de m) tales que:

Para todo $x \in A(\epsilon)$, existe $n_0 = n_0(x) \geq N_0$ tal que para todo $n \geq n_0$ y para todo $m \leq \delta n$ se cumple:

$$\#\{1 \leq k \leq n: f^{-k}(x) \in X_m(2\epsilon)\} \geq \delta n$$

Si demostramos la Afirmación Y, entonces deducimos inmediatamente la Afirmación X, eligiendo para cada $x \in A(\epsilon)$ y cada $m \geq 1$ dados:

$$n_1 = n_1(x, m) := \max\{n_0(x), m/\delta\}.$$

Entonces, ahora el objetivo final es probar la Afirmación Y, que es justamente el Lema de Pliss, que demostraremos en el próximo párrafo.

Hemos terminado de probar la afirmación (10.54) como queríamos (módulo el Lema de Pliss), y con ella termina la demostración del Lema 10.1.2 \square

El siguiente resultado reformula, en un contexto adaptado a nuestra notación, el Lema de Pliss [Pl 1972]. Este lema es un resultado aritmético que se refiere a promedios de sucesiones cualesquiera de números reales.

Sea (fijo) $\epsilon > 0$. Sea, para cada $n \geq 1$, el conjunto $A_n(\epsilon)$ definido por la igualdad (10.47).

Denotemos

$$A(\epsilon) := \bigcup_{n=1}^{+\infty} A_n(\epsilon) \quad (10.57)$$

Sea, para cada $m \geq 1$, el conjunto $X_m(2\epsilon)$ definido por la igualdad (10.49).

Lema 10.2.2. Una reformulación del Lema de Pliss**Hipótesis:** Dado $\epsilon > 0$ sea

$$x \in A(\epsilon).$$

En extenso, la hipótesis asume que existe $n(x)$ tal que para cada $n \geq n(x)$, existe un vector no nulo $0 \neq v \in E_x$ que cumple:

$$\frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} \leq -\chi_E^+(x) + \epsilon \quad (10.58)$$

Tesis: Para todo $\epsilon > 0$, existen $N = N(\epsilon) \geq 1$ y $0 < \delta = \delta(\epsilon) \leq 1$ tales que:

Para todo punto $x \in A(\epsilon)$, y para todo $n \geq \max\{n(x), N(\epsilon)\}$, se cumple lo siguiente:
Existen

$$r \geq \delta n$$

puntos diferentes en el tramo de órbita

$$\left\{ f^{-k}(x) \right\}_{0 \leq k \leq n-1}$$

que pertenecen al conjunto $X_m(2\epsilon)$, definido en la igualdad (10.49), para cualquier natural

$$1 \leq m \leq \delta n.$$

En otras palabras, existen r índices $\{k_1, k_2, \dots, k_r\} \subset \{k \in \mathbb{N} : 0 \leq k \leq n-1\}$, con $r \geq \delta n$ y $k_i \neq k_j$ si $i \neq j$, tales que para todo $i = 1, \dots, r$ el punto $f^{-k_i}(x)$ posee un vector no nulo $u_i \in E_{f^{-k_i}(x)}$ que satisface la siguiente condición de 2ϵ -aproximación en tramos largos:

$$\frac{\log(\|df_{f^{-k_i}(x)}^{-j}u_i\|/\|u_i\|)}{j} \leq -\chi_E^+(x) + 2\epsilon \quad \forall j \in \{1, \dots, m\}. \quad (10.59)$$

Se resalta que $\delta = \delta(\epsilon)$ y $N = N(\epsilon)$ no dependen del punto $x \in A(\epsilon)$, ni de ninguna medida de probabilidad μ , ni de los naturales n y m , y que la tesis vale para cualquier punto $x \in A(\epsilon)$ (no solo para un subconjunto con probabilidad total de $A(\epsilon)$).

Interpretación: Se asume por hipótesis, que en un instante al final de por lo menos n iterados (para cualquier n suficientemente grande), la aproximación al límite superior χ_E^+ de los posibles exponentes de Lyapunov, es por lo menos ϵ . La tesis afirma que entonces, en el camino recorrido entre el iterado 1 y el n , existieron por lo menos $r \geq \delta n$ instantes (una proporción no menor que $\delta > 0$), en que la aproximación fue de por lo menos 2ϵ y no solo en vectores diferentes para cada instante ni en instantes aislados, sino para un mismo vector, y en tramos de longitud m , que crece y tiende infinito con n .

Ahora demostraremos el Lema 10.2.2. La formulación original de Pliss de este lema y su demostración se encuentra en [Pl 1972]. Otras formulaciones del Lema de Pliss y sus demostraciones pueden encontrarse, por ejemplo, en [Mañ 1983a, Lema on page 356], [Mañ 1987], [Bon-Dia-Via 2005, Lemma 11.5], [Via 2012a] y [Po 2007].

Demostración. del Lema 10.2.2: Recordamos del Teorema 4.4.8, que

$$|\chi_E^+| \leq \log k \quad (10.60)$$

donde $k = \max_{z \in M} \{\|df_z\|, \|df_z^{-1}\|\} \geq 1$. Además, por la fórmula de derivada de función compuesta y de función inversa: 4.4.3:

$$\left| \frac{\log(\|df_x^{-n}v\|/\|v\|)}{n} \right| \leq \log k \quad \forall x \in X, \forall v \in T_x M. \quad (10.61)$$

Sean

$$N = N(\epsilon) \geq \frac{2 \log k + 2\epsilon}{\epsilon} \quad (10.62)$$

$$0 < \delta = \delta(\epsilon) := \frac{\epsilon}{4 \log k + 4\epsilon} \quad (10.63)$$

Dado $x \in A(\epsilon)$, debido a las definiciones (10.47 y (10.57), existe

$$n_0 = n_0(x) \geq N$$

tal que para todo $n \geq n_0$ existe algún vector $0 \neq v_0 \in E_x$ que satisface la desigualdad (10.58). Fijemos cualquier $n \geq n_0$ y el vector $0 \neq v_0 \in E_x$ que satisface la desigualdad (10.58).

Aplicando la regla de la cadena tenemos:

$$df_x^n v_0 = df_{f^{-(n-1)}(x)}^{-1} \cdots df_{f^{-j}(x)}^{-1} \cdots df_{f^{-2}(x)}^{-1} df_{f^{-1}(x)}^{-1} df_x^{-1} v_0.$$

Denotamos:

$$v_j = df_x^{-j} v_0 \quad \forall j \in \{0, 1, \dots, n-1\}.$$

Tomando norma en la igualdad obtenida de la regla de la cadena, y dividiendo entre $\|v_0\|$, resulta:

$$\frac{\|df_x^{-n} v_0\|}{\|v_0\|} = \prod_{j=0}^{n-1} \frac{\|df_{f^{-j}(x)}^{-1} v_j\|}{\|v_j\|}.$$

Tomando logaritmo:

$$\log(\|df_x^{-n} v_0\|/\|v_0\|) = \sum_{j=0}^{n-1} \log(\|df_{f^{-j}(x)}^{-1} v_j\|/\|v_j\|).$$

Por la hipótesis (10.58):

$$\sum_{j=0}^{n-1} \log(\|df_{f^{-j}(x)}^{-1} v_j\|/\|v_j\|) \leq n(-\chi_E^+ + \epsilon)$$

Por simplicidad en la notación, escribimos:

$$a_j := \log(\|df_{f^{-j}(x)}^{-1} v_j\|/\|v_j\|) \quad (10.64)$$

Entonces tenemos:

$$\sum_{j=0}^{n-1} a_j \leq n(-\chi_E^+ + \epsilon). \quad (10.65)$$

Definimos para todo

$$k \in \{1, \dots, n\}$$

la siguiente suma auxiliar:

$$S_k := \left(\sum_{j=0}^{k-1} a_j \right) + k\chi_E^+ - 2k\epsilon. \quad (10.66)$$

Entonces, si $1 \leq k \leq n-2$, tenemos:

$$\begin{aligned} S_{k+1} - S_k &= \sum_{j=0}^k a_j + (k+1)\chi_E^+ - 2(k+1)\epsilon - \sum_{j=0}^{k-1} a_j + k\chi_E^+ - 2k\epsilon = \\ &= a_k + \chi_E^+ - \epsilon \geq -|a_k| - |\chi_E^+| - 2\epsilon. \end{aligned}$$

Usando las cotas (10.60) y (10.61), resulta:

$$S_{k+1} - S_k \geq -2 \log k - 2\epsilon \quad \forall k \in \{1, \dots, n-2\} \quad (10.67)$$

Además, reuniendo (10.65) y (10.66):

$$S_n = \sum_{j=0}^{n-1} a_j + n\chi_E^+ - 2n\epsilon \leq -n\epsilon \quad (10.68)$$

Usando que $n \geq N$ por hipótesis, y la fórmula (10.62), deducimos:

$$S_n \leq -n\epsilon \leq -N\epsilon \leq -2(\log k + 2\epsilon) < -\log k \quad (10.69)$$

De la definición de la suma auxiliar S_k en (10.66), aplicada al caso $k=0$, tenemos:

$$S_1 = \sum_{j=0}^0 a_j = a_0 \geq -|a_0| \geq -\log k > S_N \quad (10.70)$$

Luego,

$$S_1 > S_N.$$

Ahora recorremos los naturales $k \in \{1, 2, \dots, n\}$ en forma regresiva, comenzando por

$$k_1 := n$$

del siguiente modo: Tenemos $k_1 = n$, luego $S_{k_1} = S_n$. Si $n-1$ cumple que $S_{n-1} \geq S_n$, entonces nombramos $k_2 := n-1$, $S_{k_2} = S_{n-1}$ y pasamos al natural siguiente en orden descendente. En caso contrario, no nombramos todavía a k_2 ni a S_{k_2} y

pasamos al natural siguiente en orden descendente $n - 2$. Digamos que en este paso tenemos nombrados hasta el índice k_i . Si $S_{n-2} \geq S_{n-1}, S_n$, entonces nombramos $k_{i+1} = n - 2, S_{k_{i+1}} = S_{n-2}$ y pasamos a $n - 3$. En caso contrario (o sea si $S_{n-2} < S_{n-1}$ o si $S_{n-2} < S_n$), no nombramos todavía a k_{i+1} y pasamos al natural siguiente en orden descendente. Dicho con precisión:

Construimos todos los naturales

$$1 \leq k_s < k_{s-1} < \dots < k_i < \dots < k_2 < k_1 = n \quad (10.71)$$

que cumplen

$$S_{k_i} \geq S_k \quad \forall k \in \{k_i, \dots, n\}. \quad (10.72)$$

Se observa que k_s no tiene porqué ser 1, ya que si bien $S_1 > S_n$, puede haber otro valor del índice $k_i > 1$, tal que $S_{k_i} \geq S_1$. El último k_s en ser definido (es decir, el menor índice k_s) siempre cumple:

$$S_{k_s} \geq S_1. \quad (10.73)$$

También observamos que la cantidad s de índices así construidos, es por lo menos 2. En efecto, por construcción, el primero (el mayor índice) es $k_1 = n$, por convención. Por otro lado, ya probamos que $S_1 > S_n$. Luego, existe un primer natural (primero en orden descendente) después del mayor, un $k_2 < n$ que cumple $S_{k_2} \geq S_k \quad \forall k \in \{k_2, \dots, n\}$. En caso contrario serían todas las sumas auxiliares S_k menores estrictamente que S_n , para $1 \leq k < n$, contradiciendo que $S_1 > S_n$.

Como hay por lo menos 2 índices diferentes en la colección (10.71), para todo $i \in \{1, \dots, s - 1\}$ existe $S_{k_{i+1}}$. Por el procedimiento de construcción con la condición (10.72), deducimos que:

$$S_{k_i} \geq S_{k_{i+1}+1}, \quad (10.74)$$

pues si fuera $S_{k_{i+1}+1} > S_{k_i}$, entonces habríamos asignado el siguiente índice k_{i+1} antes (es decir mayor o igual) de llegar índice $k_{i+1} + 1$ (y k_{i+1} no es mayor o igual que $k_{i+1} + 1$).

Usando la desigualdad (10.67):

$$S_{k_{i+1}+1} - S_{k_{i+1}} \geq -2 \log k - 2\epsilon \quad \forall i \in \{1, \dots, s - 1\}$$

Luego, reuniendo esta última desigualdad con (10.74), obtenemos:

$$S_{k_i} - S_{k_{i+1}} \geq S_{k_{i+1}+1} - S_{k_{i+1}} \geq -2 \log k - 2\epsilon \quad \forall i \in \{1, \dots, s - 1\}.$$

Entonces:

$$\sum_{i=1}^{s-1} (S_{k_i} - S_{k_{i+1}}) \geq (s - 1)(-2 \log k - 2\epsilon),$$

$$S_1 - S_s \geq (s - 1)(-2 \log k - 2\epsilon).$$

Reuniendo esta última desigualdad con (10.68) y (10.73):

$$-n\epsilon - S_0 \geq S_1 - S_s \geq (s - 1)(-2 \log k - 2\epsilon).$$

Finalmente empleando la cota (10.70), resulta:

$$-n\epsilon + \log k \geq -n\epsilon - S_0 \geq (s-1)(-2\log k - 2\epsilon),$$

de donde

$$s-1 \geq \frac{n\epsilon - \log k}{2\log k + 2\epsilon} \geq \frac{n\epsilon}{2\log k + 2\epsilon} - 1.$$

Entonces,

$$\frac{s}{n} \geq \frac{\epsilon}{2\log + 2\epsilon} = 2\delta. \quad (10.75)$$

Con esta construcción hemos encontrado una proporción s/n no menor que 2δ del conjunto inicial de n índices. Vamos a interpretar ahora el resultado en términos de las aproximaciones de $(1/h)\log(\|df_{f^{k_i}}^{-h}v_{k_i}\|/\|v_{k_i}\|)$ al límite superior χ_E^+ :

Para todo k_i con

$$i \in \{2, \dots, s\}$$

y para todo natural

$$h \in \{k_i + 1, \dots, n\}$$

se cumple:

$$\begin{aligned} \sum_{j=k_i}^{h-1} a_j &= \sum_{j=0}^{h-1} a_j - \sum_{j=0}^{k_i-1} a_j = S_h + h(-\chi_E^+ + 2\epsilon) - S_{k_i} - k_i(-\chi_E^+ + 2\epsilon) = \\ &= S_h - S_{k_i} + (h - k_i)(-\chi_E^+ + 2\epsilon). \end{aligned}$$

Como $S_h \leq S_{k_i}$:

$$\sum_{j=k_i}^{h-1} a_j \leq (h - k_i)(-\chi_E^+ + 2\epsilon).$$

Reemplazando a_j por la igualdad (10.64):

$$\sum_{j=k_i}^{h-1} \log(\|df_{f^{-j}(x)}^{-1}v_j\|/\|v_j\|) \leq (h - k_i)(-\chi_E^+ + 2\epsilon).$$

$$\log \prod_{j=k_i}^{h-1} (\|df_{f^{-j}(x)}^{-1}v_j\|/\|v_j\|) \leq (h - k_i)(-\chi_E^+ + 2\epsilon),$$

y por la regla de la cadena:

$$\log(\|df_{f^{-k_i}(x)}^{-(h-k_i)}v_{k_i}\|/\|v_{k_i}\|) \leq (h - k_i)(-\chi_E^+ + 2\epsilon).$$

$$\frac{\log(\|df_{f^{-k_i}(x)}^{-(h-k_i)}v_{k_i}\|/\|v_{k_i}\|)}{h - k_i} \leq -\chi_E^+ + 2\epsilon \quad \forall h \in \{k_i + 1, \dots, n\}. \quad (10.76)$$

(Hemos descartado a $k_1 = n$ para que el cálculo anterior sea válido pudiendo computar $k_i + 1$). Luego, hemos encontrado $s - 1 \geq (2\delta n) - 1$ puntos

$$f^{-k_i}(x) : \quad 2 \leq i \leq s$$

a lo largo del tramo finito de órbita $\{f^{-j}(x)\}_{1 \leq j \leq n-1}$ del punto x , tales que tienen vectores no nulos $v_{k_i} \in E_{f^{k_i}(x)}$ que 2ϵ aproximan, por la desigualdad (10.76), al límite χ_E^+ , durante todo un tramo (de los $df^{-(h-k_i)}$ con k_i fijo):

$$h - k_i \in \{1, \dots, n - k_i\}.$$

Para terminar de probar el Lema 10.2.2, contemos cuántos, de estos $s - 1$ puntos $\{f^{-k_i}(x)\}$, están en el conjunto $X_m(2\epsilon)$ dado por la igualdad (10.49); es decir, cuántos 2ϵ -aproximan durante un tramo que incluya por lo menos a todos los $df^{-(h-k_i)}$ en el siguiente intervalo de naturales

$$h - k_i \in \{1, \dots, m\},$$

donde m es un natural que puede tomar cualquier valor, siempre que no supere a $n\delta$. Para asegurarse que esto suceda, basta imponer la condición

$$n - k_i \geq m.$$

Esto es, $k_i \leq n - m \leq n(1 - \delta)$. Entonces contemos (o mejor dicho, estimemos en cota inferior) cuántos k_i diferentes, de los $s \geq \delta n$ índices diferentes que construimos al principio, cumplen $k_i \leq n(1 - \delta)$. Contemos cuántos a lo sumo no lo cumplen, y restémoslos de s :

Si $k_i > n(1 - \delta)$, como por construcción (después de descartar el último) $1 \leq k_i \leq n - 1$, entonces $k_i \in (n - n\delta, n) \cap \mathbb{N}$. En el intervalo de reales $(n - n\delta, n - 1)$ hay a lo sumo $(\delta n) - 1$ naturales. Entonces los índices k_i que debemos descartar son a lo sumo δn (incluyendo $k_1 = n$, que ya lo descartamos una vez). Como tenemos $s \geq 2\delta n$ índices diferentes en el intervalo $[1, n]$, restando δn quedan:

$$r = s - \delta n \geq 2\delta n - \delta n = \delta n$$

índices diferentes $k_i \in [0, n - 1]$ tales que los puntos $f^{-k_i}(x)$ satisfacen la desigualdad (10.76), para todo $1 \leq h - k_i \leq m$. Luego $f^{-k_i}(x) \in X_m(2\epsilon)$, como queríamos demostrar.

Finalmente, notamos que, en el enunciado del Lema 10.2.2, los índices k_i están renumerados con $i = 1, \dots, r$. Para ser precisos, los índices del enunciado son en realidad la subcolección $\{k_{i_l}\}_{1 \leq l \leq r}$, obtenida de la colección $\{k_i\}_{1 \leq i \leq s}$ construida a lo largo de esa prueba, después de descartar los del conjunto $\{1 \leq i \leq s : k_i \in (n - n\delta, n)\}$.

Hemos terminado de probar el Lema 10.2.2 de Pliss. \square

Bibliografía

- [Ab-An 2012] Abdenur F., Andersson M.: *Ergodic theory of generic continuous maps* Preprint ArXiv:1201.0632[math.DS], 2012
- [Al-Ar 2003] Alves J.F., Araújo V., Saussol B.: *On the uniform hyperbolicity of some nonuniformly hyperbolic systems* Proc. Amer. Math. Soc. **131** (2003) pp. 1303–1309
- [An 1967] Anosov D.V.: *Geodesic flows on closed Riemannian manifolds of negative curvature* Proc. Steklov Inst. **90** (1967) pp. 1–235
- [Arn-Ave 1967] Arnold V.I., Avez A.: *Problèmes ergodiques de la mécanique classique* Gauthier-Villars, Paris, 1967
- [Avi-Boc 2006] Ávila A., Bochi J.: *A generic C^1 map has no absolutely continuous invariant probability measure* Nonlinearity **19** (2006) pp. 2717–2725
- [Avi-Boc 2007] —, —: *Generic Expanding maps without absolutely continuous invariant σ -finite measure* Math. Research Letters **14** (2007), pp. 721–730
- [Bac 1999] Bachurin P.S.: *The connection between time averages and minimal attractors* Russ. Math. Surv. **54** (1999) pp. 1233–1235
- [Ba-Pe 2001] Barreira L., Pesin Y.: *Lyapunov Exponents and Smooth Ergodic Theory* University Lectures Series Vol. 23, ISBN 978-08218-2921-1, Amer. Math. Soc., Providence 2001
- [Bi 1931] Birkhoff G.D.: *Proof of the Ergodic Theorem* Proc. Nat. Acad. of Sciences **17** (12) (1931) pp 656–660
- [Boc 2012] —, —: *On the Subadditive Ergodic Theorem* Teaching Notes of PUC-Rio www.mat.puc-rio.br/~jairo/kingbirk.pdf (Last retrieved May 11th., 2012)
- [Bog-Kry 1937] Bogliubov N., Krylov N.: *La théorie générale de la mesure dans son application a l'étude de systemes dynamiques de la mécanique non-linéaire* Annals Math. **38** (1937) pp. 65–113

- [Bon 2011] Bonatti C.: *Towards a global view of dynamical systems, for the C^1 -topology* Ergod. Th. & Dynam. Sys. **31** (4) (2011) pp.959–993
- [Bon-Dia-Via 2005] —, Diaz L., Viana M.: *Dynamics beyond uniform hyperbolicity. A global geometric and probabilistic perspective* Encyclopaedia of Mathematical Sciences, **102** Editors: J. Fröhlich, S.P. Novikov, D. Ruelle. Springer, Berlin- Heidelberg-New York, 2005
- [Bow 1971] Bowen R.: *Periodic points and measures for Axiom A diffeomorphisms* Trans. Amer. Math. Soc. **154** (1971) pp. 377–397
- [Bow 1975a] —: *A horseshoe with positive measure* Invent. Math. **29** (3) (1975) pp. 203–204
- [Bow 1975b] —: *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms* Lect. Notes in Math. **470** ISBN: 3-450-07187-3, Springer-Verlag, New York, 1975
- [Bow-Rue 1975] —, Ruelle D.: *The ergodic theory of Axiom A flows* Invent. Math. **29** (1975) pp. 181–202
- [Bu 2011] Buzzi J.: *Chaos and Ergodic Theory* In the book: Mathematics of Complexity and Dynamical Systems, R. A. Meyers (Ed.), ISBN: 978-1-4614-1806-1, pp. 63–87, Springer, New York, 2011
- [Cam-Qua 2001] Campbell J., Quas A.: *A generic C^1 expanding map has a singular S - R - B measure* Commun. Math. Phys. **221** (2001) pp. 335–349
- [Car 1991] Carvalho M.: *Medidas de Bowen-Ruelle-Sinai para derivados de Anosov n -dimensionales* PhD Thesis, IMPA, Informes de Matemática Série F, 044/91, Rio de Janeiro, 1991.
- [Car 1992] —: *Sinai-Ruelle-Bowen measures for N -dimensional derived from Anosov diffeomorphisms* Ergod. Th. & Dynam. Sys. **13** (1992), pp. 21–44
- [Cat 2012] Catsigeras E.: *On Ilyashenko’s Statistical Attractors* (Submitted) Preprint arXiv:1106.4072v3 [math.DS] (2012)
- [Cat-Cer-Enr 2012] —, Cerminara M., Enrich H.: *Pesin’s Entropy Formula for C^1 Diffeomorphisms with Dominated Splitting* (Submitted) Preprint arXiv:1209.5784v1 [math.DS] (2012)
- [Cat-Enr 2001] —, Enrich H.: *SRB measures of certain almost hyperbolic diffeomorphisms with a tangency* Discr. & Cont. Dynam. Sys. **7** (2001) pp. 177–202
- [Cat-Enr 2011] —, —: *SRB-like measures for C^0 dynamics* Bull. Polish Acad. Sci. - Math. **59** (2011) pp. 151-164

- [Cat-Enr 2012] —, —: *Equilibrium States and SRB-like measures of C^1 expanding maps of the circle* Portug. Math. **69** (2012), pp. 193–212
- [Enr 1998] Enrich, H.: *A heteroclinic bifurcation of Anosov diffeomorphisms* PhD Thesis IMPA, Informes de Matemática Série F-079/95, Rio de Janeiro, 1995 and Ergod. Th. & Dynam. Sys. **18** (1998) pp. 567–608
- [Fo 1984] Folland G. B.: *Real Analysis. Modern Techniques and Their Applications*. ISBN 0-471-80958-6. Wiley-Interscience, U.S.A, 1984.
- [Fr 1969] Franks J.: *Anosov diffeomorphisms on tori* Trans. of the Amer. Math. Soc. **145** (1969) pp. 117–124.
- [Fra-McC 2011] Frantzikinakis N., McCutcheon R. *Ergodic Theory: Recurrence* In the book: Mathematics of Complexity and Dynamical Systems, R. A. Meyers (Ed.), ISBN: 978-1-4614-1806-1, pp. 357–368 Springer, New York, 2011
- [Fu 1961] Furstenberg H.: *Strict ergodicity and transformations of the torus* Amer. Journ. of Math. **83** (1961) pp.573-601
- [Go-Kl 2007] Golenishcheva-Kutuzova T., Kleptsyn V.: *Convergence of the Krylov-Bogolyubov procedure in Bowen's example* Mat. Zametki **82** (5) (2007) pp. 678–689; English Translation in Math. Notes **82** (5-6) (2007) pp. 608–618
- [Hu 2000] Hu H.: *Conditions for the existence of SRB measures for almost Anosov diffeomorphisms* Trans. Amer. Math. Soc **352** (2000) pp. 2331–2367
- [Go-Il 1996] Gorodetski A., Ilyashenko Yu.S.: *Minimal and strange attractors* Int. Journ. of Bif. and Chaos **6** (1996) pp. 1177–1183
- [Ha 2013] Hammerlindl A.: *Partial hyperbolyity on 3-dimensional nilmanifolds* Discr. & Cont. Dynam. Sys. **33** (8) (2013) pp. 3641–3669
- [Hi-Pu-Sh 1977] Hirsch M, Pugh C., Shub M.: *Invariant manifolds* Lect. Notes in Math **587**, Springer-Verlag, Berlin, 1977
- [Hu-You 1995] Hu H., Young L.S.: *Nonexistence of SRB measures for some diffeomorphisms that are almost Anosov* Ergod. Th. & Dynam. Sys. **15** (1995) pp. 67–76
- [Il 2005] Ilyashenko, Yu.S.: *Minimal attractors* In Proceedings of EQUAD-IFF 2003, International Conference on Differential Equations in Husselt, Belgium, pp. 421–428, ISBN: 978-981-256-169-5, World Scientific Publishing, Singapore, 2005

- [Jo 2005] Jost J.: *Dynamical Systems* ISBN 978-3-540-22908-7, Springer-Verlag, Berlin-Heidelberg, 2005.
- [Ka-As 2011] Karabacak O., Ashwin P.: *On statistical attractors and the convergence of time averages* Math. Proc. Camb. Phil. Soc. **150** (2011) pp. 353–365.
- [Kar-Mar 1999] Karlsson A., Margulis G.A.: *A multiplicative ergodic theorem and nonpositively curved spaces* Comm. Math. Phys. **208**(1)(1999) pp. 107–123
- [Kat-Has 1995] Katok A., Hasselblatt B.: *Introduction to the modern theory of dynamical systems* ISBN 0-521-57557-5, Cambridge Univ. Press, Cambridge, 1995.
- [Katz 1968] Katznelson Y. *An Introduction to Harmonic Analysis* 3rd. Edition ISBN:9780521543590 (1st. edition 1968), Cambridge Math. Library, Cambridge, 2004
- [Kea-Smo 1979] Keane M., Smorodinsky M.: *Bernoulli schemes of the same entropy are finitarily isomorphic* Annals of Math. **109** (1979) pp.397–406
- [Ke 1998] Keller, G.: *Equilibrium States in Ergodic Theory* London Math. Soc. Student Texts, Vol. 42, ISBN 978-052159534-6, Cambridge University Press, Cambridge, 1998
- [Ki 2011] King J.L.F.: *Entropy in Ergodic Theory* In the book: Mathematics of Complexity and Dynamical Systems, R. A. Meyers (Ed.), ISBN: 978-1-4614-1806-1, pp. 205–224 Springer, New York, 2011
- [Kin 1973] Kingmann J.F.C. *Subadditive ergodic theory* Ann. Prob. **1** (1973) pp. 889–909
- [Kl 2006] Kleptsyn V.: *An example of non-coincidence of minimal and statistical attractors* Ergod. Th. & Dynam. Sys. **26** (2006) 759–768.
- [Kl-Ry-Mi 2012] —, Ryzhov D., Minkov S.: *Special ergodic theorems and dynamical large deviations* Nonlinearity **25** (11) (2012) pp. 3189–3196
- [Ko 1958] Kolmogorov, A.N.: *New metric invariants of transitive dynamical systems and autormorphisms of Lebesgue spaces* Dokl. Akad. Nauk. SSSR **119**(5) (1958) pp. 861–864
- [Pi 2007] Kuznetsov S.P., Pikovsky A.: *Autonomous coupled oscillators with hyperbolic strange attractors* Physica D, **232** (2007) (2) pp.87–102
- [Led 1984] Ledrappier F.: *Propriétés Ergodiques des Mesures de Sinai* Publ. Math. I.H.E.S **59** (1984) pp.163–188

- [Led-You 1985] —, Young L-S: *The metric entropy of diffeomorphisms. Part I: Characterization of measures satisfying Pesin's formula. Part II: Relations between entropy, exponents and dimensions* Ann. Math. **122** (1985) pp 509–539, pp. 540–574
- [Lew 1980] Lewowicz J: *Lyapunov functions and topological stability* Journ. of Diff. Eq. **38** (1980) pp. 192–209
- [Lew 1989] —: *Expansive homeomorphisms of surfaces* Bol. Soc. Bras. Mat. **20** (1989), pp. 113–133
- [Liv 2013] Liverani C.: *Multidimensional expanding maps with singularities: a pedestrian approach* Ergod. Th. & Dynam. Sys. **33** (1) (2013) pp. 168–182
- [Ma 1974] Manning A.: *There are no new Anosov diffeomorphisms on tori* Amer. J. Math. **96** (1974) pp. 422–429
- [Mañ 1983a] Mañé R.: *Introdução a Teoria Ergódica*, Projeto Euclides, I.M.P.A., Rio de Janeiro, 1983.
- [Mañ 1983b] —: *A proof of Pesin's formula* Ergod. Th. & Dynam. Sys. **1** (1985) pp. 95–102 and *Errata to "A proof of Pesin's formula"* Ergod. Th. & Dynam. Sys. **3** (1983) pp. 159–160
- [Mañ 1987] —: *Ergodic theory and differentiable dynamics* Ergebnisse der Mathematik und ihrer Grenzgebiete Vol. 3, ISBN 978- 0387152783, Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1987
- [Mar 1966] Margulis G.A.: Unpublished proof, 1966, cited in Hasselblatt B, Pesin Ya.B: *Pesin entropy formula* Scholarpedia **3** (3) (2008): 3733
- [Mi 1985] Milnor J.: *On the concept of attractor* Comm. Math. Phys. **99** (1985) pp. 177–195, *On the concept of attractor: Correction and remarks* Comm. Math. Phys. **103** (1985) pp. 517-519
- [Mis 2005] Misiurewicz, M.: *Ergodic natural measures* in Contemporary Mathematics **385** Algebraic and topological dynamics. Edts: Kolyada S., Manin Y., Ward T., 978-0-8218-3751-1, pp. 1–6, Amer. Math. Soc., Providence, R.I., 2005.
- [Ni-Pe 2011] Nicol M., Petersen K.: *Ergodic Theory: Basic Examples and Constructions* In the book: Mathematics of Complexity and Dynamical Systems, R. A. Meyers (Ed.), ISBN: 978-1-4614-1806-1, pp. 264–287 Springer, New York, 2011
- [Os 1966] Oseledets V.I.: *On spectra of ergodic automorphisms* Dokl. Acad. Sci. USSR **168** (5) (1966) pp. 1009–1011

- [Os 1968] —: *A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems* Trans. Moscow Math. Soc. **19** (1968) pp. 197–231
- [Pa 1999] Palis J.: *A global view of Dynamics and a conjecture on the denseness of finitude of attractors*. Astérisque **261** (1999) pp. 339–351
- [Pe 1976] Pesin Ya.B.: *Families of invariant manifolds corresponding to nonzero characteristic exponents* Math. USSR- Izv. Vol. **40** (1976), No. 6, pp. 1261–1305
- [Pe 1977] —: *Characteristic Lyapunov exponents and smooth ergodic theory* Russian Math. Surveys **32** (1977) pp. 55–114
- [Pe 1992] —: *Dynamical Systems with generalized hyperbolic attractors: hyperbolic, ergodic and topological properties* Ergod. Th. & Dynam. Sys. **12** (part I) (1992) pp. 123–152
- [Pe-Si 1982] —, Sinai, Ya.G.: *Gibbs measures for partially hyperbolic attractors* Ergod. Th. & Dynam. Sys. **2** (1982), pp. 417–438
- [Pl 1972] Pliss V.A.: *On a conjecture due to Smale* Differencial'nye Uravnenija 8 (1972) (Russian), pp. 268–282
- [Po 2007] Potrie R.: *Lema de Pliss* Notas de estudio, Univ. de la República, Montevideo, 2007 <http://www.cmat.edu.uy/cmat/docentes/rpotrie/documentos/Pliss.pdf/view> (Last retrieved: April 30th., 2013)
- [Pug 1984] Pugh C.: *The $C^{1+\alpha}$ hypothesis in Pesin theory* Inst. Hautes Études Sci. Publ. Math. Vol **59** (1984), pp. 143–161
- [Pug-Shu 1989] —, Shub M.: *Ergodic attractors* Trans. Amer. Math. Soc. **312** (1989) pp. 1–54
- [Puj-Sam 2009] Pujals E., Sambarino M.: *On the dynamics of dominated splitting* Annals of Math. **169** (2009) pp. 675–740
- [Qi-Xi-Zh 2009] Qian M., Xie J-S., Zhu S. *Smooth Ergodic Theory for Endomorphisms* Lect. Notes in Math. **1978** ISBN: 978-3-642-01953-1, Springer-Verlag, Berlin, 2009
- [Qi-Zh 2002] Qian M., Zhu S.: *SRB measures and Pesin's entropy formula for endomorphisms* Trans. Amer. Math. Soc. **354** (4) (2002) pp. 1453–1471
- [Qiu 2011] Qiu H.: *Existence and uniqueness of SRB measure on C^1 generic hyperbolic attractors* Commun. Math. Phys. **302** (2011) pp. 345–357

- [Qua 1996] Quas A.N.: *Non-ergodicity for C^1 expanding maps and g -measures* Ergod. Th. & Dynam. Sys. **16** (1996) pp. 531–543
- [Qua 2011] —: *Ergodicity and Mixing Properties* In the book: Mathematics of Complexity and Dynamical Systems, R. A. Meyers (Ed.), ISBN: 978-1-4614-1806-1, pp. 225–240 Springer, New York, 2011
- [Ra 1979] Raghunathan M.S.: *A proof of Oseledec's multiplicative ergodic theorem* Isr. J. Math. **32** (1979) pp. 356–362
- [Rob-You 1980] Robinson C., Young L.S.: *Nonabsolutely Continuous Foliations for an Anosov Diffeomorphism* Inventiones Math. **61** (1980) pp. 159–176
- [Roh 1962] Rohlin V.: *On the fundamental ideas of measure theory* Amer. Math. Soc. Transl. **1** (1962) pp. 1–52
- [Roh 1966] —: *Selected topics in metric theory of dynamical systems* Uspekhi Mat. Nauk. **4** (1949) pp.57–128 English translation in Amer. Math. Soc. Transl. (Ser. 2) **49** (1966), pp. 171–240
- [Rud 1979a] Rudin W.: *Análisis Real y Complejo* Spanish translation of *Real and Complex Analysis* ISBN 84-205-0651-6, Editorial Alhambra, Madrid, 1979
- [Rud 1979b] —: *Análisis Funcional* Spanish translation of *Functional Analysis* ISBN 82-291-5115-X, Editorial Reverté, Barcelona, 1979
- [Rue 1976] Ruelle D.: *A measure associated with axiom A attractors*. Amer. Journ. of Math. **98** (1976) pp. 619–654
- [Rue 1978] —: *An inequality for the entropy of differentiable maps*. Bol. Soc. Bras. Mat. **9** (1978) pp. 83–87
- [Rue 1989] —: *The thermodynamic formalism for expanding maps* Commun. Math. Phys. **125** (2) (1989) pp. 239–262.
- [Sa 2011] Sarig O.: *Bernoulli Equilibrium States for Surface Diffeomorphisms* J. Modern Dynam. **5** (3) (2011) pp 593–608
- [Si 1959] Sinai Ya.G.: *On the Notion of Entropy of a Dynamical System* Dokl. Acad. Sci. USSR **124** (4) (1959) pp. 768–771
- [Si 1972] —: *Gibbs measure in ergodic theory* Russ. Math. Surveys **27**(4) (1972) pp. 21–69
- [Si 1994] —: *Topics in Ergodic Theory* ISBN 0-691-03277-7, Princeton University Press, 1994

- [Si 2007] —: *Metric Entropy of Dynamical System* Preprint Princeton, 2007 <http://web.math.princeton.edu/facultypapers/Sinai/MetricEntropy2.pdf> (last retrieved April 16th., 2013)
- [Sm 1967] Smale S.: *Differentiable dynamical systems* Bull. Amer. Math. Soc. **73** (1967) pp. 747–817
- [Sch-Gor 1989] Schmitt B., Gora P.: *Un exemple de transformation dilatante et C^1 par morceaux de l'intervalle, sans probabilité absolument continue invariante* Ergod. Th. & Dynam. Sys. **9** (1) (1989) pp. 101–113
- [Ste 2005] Stein E.M, Shakarchi R.: *Real Analysis - Measure Theory, Integration and Hilbert Spaces* ISBN 978-0-691-11386-9, Princeton University Press, Princeton-Oxford, 2005
- [Su-Ti 2012] Sun W., Tian X.: *Dominated splitting and Pesin's entropy formula* Discr. & Cont. Dyn. Sys. **32** (4) (2012) pp. 1421–1434
- [Ta 2002] Tahzibi A.: *C^1 -generic Pesin's entropy formula* Compt. Rend. Acad. Sci. Paris (Ser. I) **335** (2002) pp. 1057–1062
- [Via 1998] Viana M: *Dynamics: a probabilistic and geometric perspective*. In Proceedings of the International Congress of Mathematicians at Berlin. Documenta Mathematica. Extra Vol. I (1998) pp. 395–416
- [Via 2012a] —: *A proof of Oseledets' theorem* Teaching notes of IMPA www.impa.br/viana/out/oseledets.pdf (Last retrieved August 22th., 2012)
- [Via 2012b] —: *Disintegration into conditional measures: Rokhlin's theorem* Teaching notes of IMPA www.impa.br/viana/out/rokhlin.pdf (Last retrieved August 22th., 2012)
- [Via-You 2013] —, Yang J.: *Physical measures and absolute continuity for one-dimensional center direction* Ann. de l'I.H.P.-Analyse non linéaire (in press) (2013) doi: 10.1016/j.anihpc.2012.11.002
- [Wa 2000] Walters P.: *An Introduction to Ergodic Theory*, ISBN 0387951529, Springer, New York-Heidelberg-Berlin, 2000
- [You 2002] L.S. Young: *What are SRB measures, and which dynamical systems have them?* Journ. Stat. Physics **108** (2002) pp. 733–754

Índice alfabético

- A_K cuenca de atracción
 - estadística del compacto K , 204
- $A_\epsilon(\mu)$ cuenca de atracción
 - estadística ϵ -débil
 - de la medida μ , 216
- $B(\mu)$ cuenca de atracción
 - estadística de la medida μ , 160
- $C(K)$ o E_K cuenca de atracción
 - topológica del compacto K , 146, 199
- $C^0(X, \mathbb{R})$ espacio de
 - funciones reales continuas, 3
- T^* pull back, 7
- $\omega(x)$ omega-límite, 13
- \prec relación “más fina que”
 - entre particiones, 164, 260
- σ transformación shift, 237
- $\sigma_{n,x}$ probabilidad empírica, 160
- \vee operación producto
 - entre particiones, 164, 252
- $k^{\mathbb{Z}}$ o $k^{\mathbb{N}}$ espacio del shift, 237
- $p\omega(x)$ p-omega límite en \mathcal{M} , 216
- \mathcal{M} (espacio de medidas de probabilidad),
 - 4
- \mathcal{O}_f conjunto de medidas SRB-like
 - (o pseudo-físicas u observables) para f , 217
- alfa-límite, 13
- átomos, 251
- atracción
 - estadística, 158
 - topológica, 158
- atractor
 - α -observable, 200, 209
 - caótico, 148, 157
 - de Ilyashenko, 204, 210, 217, 222
 - α -obs minimal, 209
 - de Milnor, 199
 - α -obs minimal, 200
 - de Smale-Williams, 149
 - ergódico, 155
 - ergódico caótico, 157
 - estadístico, 204–206, 208, 210, 217, 222
 - α -obs minimal, 209, 210
 - hiperbólico, 148, 150, 152, 153, 157, 178
 - periódico, 10
 - solenoides, 149
 - topológico, 139, 140, 142, 147, 150, 156, 178
 - no orbitalmente estable, 141
- automorfismo
 - de esp. de medida, 1
 - equivalencia de, 243, 244
 - lineal del toro, 57, 63, 67, 75, 140
 - lineal en el toro, 159, 165, 212, 227, 236, 237, 281
- autovalor del operador U_T , 229
- base de Fourier, 232, 233
- bit, 250
- Bogliubov-Krylov
 - procedimiento de, 8
- cantidad de información, 249, 250
- caos, 74, 148, 157, 176, 177
- caracterización espectral, 230, 231, 236
- cilindro, 238
- combinación convexa, 23
- compacidad
 - de \mathcal{M} , 6

- de \mathcal{O}_f , 218
- del espacio del shift, 239
- componentes ergódicas, 90, 169, 171
- conjetura
 - Palis, 221
 - Viana, 175
- conjugación, 73, 244, 245, 277
- conjunto
 - de probabilidad total, 31, 90
 - estable, 31
 - hiperbólico
 - no uniforme, 82, 85
 - uniforme, 70, 72
 - inestable, 31
 - Lyapunov estable, 139
 - maximal invariante, 72, 142
 - minimal, 46, 139, 210
 - minimal α -obs, 200, 209
 - no errante, 14
 - no orbitalmente estable, 141
 - orbitalmente estable, 139, 142
- continuidad
 - de la función densidad, 186
 - del operador pull back, 7
- continuidad absoluta, 224
 - de foliación, 179
 - de holonomía, 179
 - de medidas, 21
 - de medidas condicionales, 167–169, 178
- crecimiento subexponencial, 117, 123
- criterio de crecimiento subexponencial, 117
- cuenca de atracción
 - abierta, 146
 - de Ilyashenko, 204
 - de Milnor, 199
 - estadística, 160–162, 186, 204, 205, 208, 216, 220, 225, 227
 - estadística ϵ -débil, 216
 - global, 146
 - invariante, 146
 - observabilidad de, 157, 158
 - topológica, 139, 146, 157, 199
- densidad, 175, 185, 186
- derivada de Radon-Nykodim, 175, 185, 186, 231
- descomposición
 - de Fubini, 66
 - de Radon-Nikodym, 22
 - de Rohlin, 187, 189, 190
 - ergódica, 89, 90
- desigualdad de Jensen, 253
- desigualdad Margulis-Ruelle, 178
- difeomorfismos
 - casi-Anosov, 173
 - conjugados, 73
 - de Anosov, 57, 63, 67, 69, 70, 73, 75, 227, 281
 - derivados de Anosov, 173
 - Morse-Smale, 13, 17
 - parcialmente hiperbólicos, 172
- diferencia de entropías, 256
- dinámica topológica, 13
- distorsión acotada, 181
- endomorfismo expansor, 224–226
- endormorfismo expansor, 225
- entropía
 - condicional, 253, 256
 - convexidad de, 256
 - cota de, 251
 - de la herradura de Smale, 272
 - de los iterados, 273, 278
 - de partición finita, 251
 - de particiones, 256
 - de una partición, 251, 258
 - de una transformación, 266, 269
 - del producto de particiones, 255
 - del shift, 271
 - desigualdad de, 178
 - diferencia de, 256
 - fórmula de Pesin, 177, 178, 225
 - invariancia de, 269, 277
 - métrica, 176–178, 245, 269, 271–273
 - principio variacional de, 276, 283
 - topológica, 276–278
 - de Anosov lineal, 281
 - de homeomorfismo del círculo, 281
 - de la herradura de Smale, 280

- de rotación, 281
- del shift, 279
- envolvente convexa, 23
- equivalencia
 - de espacios de medida, 269
- equivalencia de definiciones
 - de ergodicidad, 17, 18, 23, 43, 44, 161, 230
- equivalencia espectral, 231, 236
- ergodicidad, 16, 18, 22, 23, 43, 44, 63, 69, 230, 233, 241, 246
 - de la rotación irracional, 48
 - descomposición, 89, 90
- ergodicidad única, 44, 47
- espacio
 - cociente por partición, 165
 - de medidas de probabilidad, 4, 6
 - del shift, 237
- espectro
 - de Fourier, 232, 233
 - de Lebesgue, 235, 236
- espectro de Lebesgue, 242, 246
- estabilidad
 - Lyapunov, 139
 - orbital, 139, 142
- estados de equilibrio, 177
- estructura de producto local, 64, 185
- expansividad, 74, 75
- exponentes de Lyapunov
 - generadores de, 120
- exponentes de Lyapunov, 62, 68, 69, 78–81, 85, 132, 133, 298
 - cotas de, 127
 - de punto periódico, 12
 - de puntos regulares, 78–80
 - no nulos, 62, 69, 81, 85
 - positivos, 62, 69, 177, 178
- fórmula de Pesin, 177, 178, 225
- fibrado, 119
 - complementario, 132, 298
 - de Oseledec, 132
 - inestable, 67, 70, 73, 82
 - medible, 119
 - puntos regulares de, 134
- transformación lineal de, 130
- flujo polo norte-polo sur, 12, 158
- foliación
 - absolutamente continua, 179
 - de clase C^0 , 76
 - dinámicamente definida, 76
 - Hölder continua, 76, 77
 - holonomía de, 179
 - invariante, 179
 - estable, 61, 75, 77
 - inestable, 61, 75, 77
 - Lipschitz, 77
 - regular, 76, 77
 - trivialización de, 76
- Fourier
 - base de, 233
 - coeficientes, 232
 - espectro de, 233
- fuelle, 11, 70
- función generadora de exp. de Lyapunov, 120
- funciones supremo, 123
- generador, 274
- generadores de exp. de Lyapunov, 120
- grupo topológico, 55
- herradura de Smale, 70, 71, 150, 152, 153, 156, 246, 247, 272, 280
- hipótesis de Boltzmann, 18
- hiperbolicidad
 - en región de Pesin, 85
 - no uniforme, 82, 85
 - parcial, 172
 - uniforme, 57, 67, 70, 72, 178
- holonomía, 179
- homeomorfismos conjugados, 73
- información, 249
- itinerario, 246
- jacobiano, 58
- jacobiano inestable, 175, 181
- jerarquía ergódica, 246
- lema de distorsión acotada, 181

- lema de Pliss, 314
- métrica
 - en \mathcal{M} , 6
 - en el espacio del shift, 238
- mapa expansor, 221, 224–226
- maximal invariante, 72, 142
- medida
 - absolutamente continua, 21, 167–169
 - atracción estadística de, 160–162
 - condicional, 166–168
 - condicional inestable, 167–169, 178, 185, 186
 - de Bernoulli, 241, 246
 - de conjuntos
 - sucesión de, 259
 - de Gibbs, 168, 169, 171, 172, 174, 175, 177, 178, 185, 224, 225
 - de Gibbs ergódica, 186, 227
 - de Haar, 54
 - de Lebesgue, 58, 63, 233, 237
 - de probabilidad empírica, 160, 214
 - densidad de, 175, 185, 186
 - entropía respecto de, 269
 - equivalencia de, 185, 225
 - ergódica, 16, 18, 22, 23, 43, 44, 63, 90, 161, 230, 233, 241
 - extremal, 23
 - física, 156, 162, 171
 - hiperbólica, 85, 177
 - hiperbólica ergódica, 87
 - invariante, 2, 27
 - ejemplo de no existencia, 3
 - teorema de existencia, 2
 - mixing, 51, 230, 236, 237, 241
 - mutuamente singulares, 21
 - positiva sobre abiertos, 17
 - pseudo-física, 217, 218
 - soportada en atractor, 147, 156, 157
 - soporte compacto de, 221
 - SRB, 156, 157, 162, 171, 172, 174, 175, 178, 206, 220, 225
 - SRB ergódica, 172, 174
 - SRB no ergódica, 162
 - SRB-like, 217–220, 222, 227
- no uniformemente hiperbólico, 82
- o`rbita, 1
- observabilidad, 155, 157, 158
- omega-límite, 13
- operador U_T , 229–231
- operador unitario, 229
- partición, 251
 - borde de, 257
 - con borde de medida nula, 258
 - diámetro de, 257
 - generadora, 270
 - más fina que, 164, 260
 - medible, 163, 165
 - operación \vee , 164
- Pesin
 - fórmula de, 177, 178, 225
 - región de, 85
 - teoría de, 163, 171, 179
- pozo, 11, 70, 140
- principio variacional de la entropía, 276, 283
- probabilidad
 - empírica, 160, 214, 216
 - total, 31, 90
- producto de particiones, 252, 255, 256
- producto local, 64, 185
- promedio
 - de Birkhoff, 27, 160
 - de medida de transitividad, 29
 - de sucesión de funciones, 30
 - de tiempo de estadía, 31, 43
 - temporal, 18, 27, 28, 154, 160
- propiedad mixing, 246, 247
- pull back, 7
- punto
 - débilmente regular, 78, 79
 - fijo, 2
 - Lyapunov regular, 78, 80
 - periódico, 2
 - atractor, 10
 - hiperbólico, 9, 11, 70
 - recurrente, 13
 - regular, 78–80, 133, 134

- recurrencia, 15
 - topológica, 13
- refinamiento de particiones, 260
- región de Pesin, 85
- regularidad
 - débil, 78, 79
 - de puntos, 78–80
 - Lyapunov, 78, 80
- repulsor periódico, 10
- rotación, 281
 - irracional, 9, 48, 75, 158, 233
 - racional, 9, 75
- semiconjugación, 276
- separador, 274
- serie de Fourier, 232
- shif
 - de Bernoulli, 241
- shift, 279
 - de Bernoulli, 241, 242, 245, 250, 271
 - espacio del, 237
 - métrica del, 238
 - unilateral, 242
- silla, 11, 70
- singularidad mutua, 21, 22
- sistema dinámico, 1
- sistemas dinámicos conjugados, 73
- solenoides, 149
- soporte compacto, 221, 222
- splitting
 - de Oseledets, 80
 - dominado, 173
 - hiperbólico, 67, 68, 70, 82
- subaditividad, 33, 102, 103
- subespacio
 - complementario, 132, 298
 - de Oseledets, 132
 - de Oseledets, 80, 293, 294
 - inestable, 67, 70, 73, 82
- subexponencial, 117, 123
 - criterio, 117
- tent map, 10, 75, 158, 245
- teoría de Pesin, 163, 171, 179
- teorema
 - Birkhoff-Khinchin, 28, 38
 - de caracterización de
 - atractor estadístico, 222
 - de compacidad de $2^{\mathbb{Z}}$, 239
 - de compacidad de \mathcal{M} , 6
 - de compacidad de \mathcal{O}_f , 218
 - de continuidad de
 - entropía de una partición, 258
 - de cotas de exp. de Lyapunov, 127
 - de descomposición ergódica, 90
 - de descomposición medible, 166
 - de entropía de los iterados, 278
 - de existencia de
 - atractor de Ilyashenko, 210
 - atractor de Milnor, 200
 - atractor estadístico, 210
 - foliaciones invariantes, 77
 - medidas condicionales, 166
 - medidas de Gibbs, 174
 - medidas ergódicas, 21, 23, 90
 - medidas invariantes, 2
 - medidas SRB, 174
 - medidas SRB-like, 218
 - variedades invariantes, 72
 - de hiperbolicidad no uniforme, 85
 - de invariancia de la entropía, 269, 277
 - de medida de Gibbs, 171
 - de medida SRB ergódica, 172, 174
 - de medidas ergódicas
 - existencia, 21, 90
 - extremalidad, 23
 - singularidad mutua, 22
 - de metrizable de \mathcal{M} , 6
 - de producto local, 185
 - de propiedad mixing
 - del shift de Bernoulli, 241
 - de representación de Riesz, 5
 - de unicidad de medida SRB, 174, 225
 - de unicidad de medida SRB-like, 220
 - ergódico, 28, 32, 34, 36–38, 103, 135
 - ergódico maximal, 32, 34, 36, 37
 - ergódico multiplicativo, 135
 - ergódico subaditivo, 103
 - Franks, 73
 - Hopf, 16

- Jensen - desigualdad de, 253
- Katznelson, 235, 237
- Kingmann, 103
- Kolmogorov-Sinai, 270
- Lebesgue-Radon-Nikodym, 22
- Ledrappier-Young, 178
- lema de Poincaré, 15
- Lewowicz, 75
- Margulis-Ruelle, 178
- optimalidad estadística, 219
- Orstein, 245
- Oseledets, 69, 78, 81, 135, 293
- Parseval, 232
- Pesin-Sinai, 172, 174
- Pliss - Lema de, 314
- Riesz, 5
- Rohlin, 166, 169
- Ruelle, 225
- Tichonov, 5
- variacional de la entropía, 276, 283
- tiempo medio de estadía, 31, 43, 207, 208
- topología débil estrella, 4, 258, 259
- transformación
 - únicamente ergódica, 44, 47
 - caótica, 74
 - conjugada, 277
 - de Bernoulli, 245
 - entropía de, 266, 269
 - ergódica, 16, 18, 22, 23, 43, 44, 63, 69, 230, 233, 241, 246
 - expansiva, 74, 75
 - hiperbólica
 - no uniforme, 82
 - uniforme, 57, 63, 67, 70, 72
 - mixing, 51, 230, 236, 237, 241, 246
 - semiconjugada, 276
 - transitiva, 227, 234, 237, 242
- transitividad, 14, 17, 69, 227, 234, 247
- traslación en el toro, 281
- variedad invariante
 - estable, 59, 72, 73
 - inestable, 59, 72, 73, 87, 152, 153
 - local, 60, 87
- vector de probabilidad, 240