PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

“TÍTULO DE LA TESIS

ELEMENTOS DE LA DINÁMICA DE ITERACIÓN DE FUNCIONES ”

Tesis para optar el grado de Magíster

Autor:
Cesar Augusto Vergaray Albujar

Asesor:
Rudy José Rosas Bazan

Jurado:
Roland Rabanal Montoya
Percy Braulio Fernández Sánchez

LIMA-PERÚ

2015

“A mis padres,
a mis amigos y a mi asesor,
que a lo largo de este camino que llamamos vida
hicieron que todo sea más divertido.”

Índice general

1. Nociones de entropía topológica **1**

1.1. Cadenas de Markov 3

1.2. Teorema de Sharkovsky 9

2. Teorema de Sharkovsky **13**

3. Teorema de recurrencia de Poincaré y de Birkhoff **25**

3.1. Nociones preliminares 25

3.2. Transformaciones que preservan medida 27

3.2.1. Ejemplos 27

3.3. Teorema de recurrencia de Poincaré 28

3.4. Preliminares para el teorema de Birkhoff 29

3.5. Teorema de Birkhoff 35

3.5.1. Aplicaciones del teorema de Birkhoff 38

3.5.2. Aplicaciones del teorema de Birkhoff 41

Resumen

En este trabajo desarrollaremos dos aspectos de Dinámica: El primero que trata sobre la dinámica de funciones que van de un intervalo en si mismo, introduciremos las cadenas de Markov y algunos resultados previos para alcanzar al final el teorema de Sharkovsky demostrado con grafos, el cual lo haremos en la primera parte de este trabajo. La segunda parte de este trabajo tratará sobre la teoría ergódica, nos enfocaremos en dos de los teoremas fundamentales que son el teorema de recurrencia de Poincaré y el teorema de Birkhoff.

Introducción

En este trabajo se presenta, de forma autocontenida, una disertación escrita sobre uno de los teoremas más importantes y hermosos de la dinámica unidimensional. Se trata en realidad de una colección de resultados debidos al matemático ucraniano Oleksandr Sharkovsky (1936−) publicados inicialmente en ruso en el año 1964, y que en la actualidad son reunidos con el nombre de Teorema de Sharkovsky. El teorema de Sharkovsky permaneció sin conocerse fuera de la Europa Oriental hasta la segunda mitad de la década de 1970, cuando aparece publicado el artículo [LY75] de Tien Yien Li y James A. Yorke. En este artículo se demuestra parcialmente un caso particular del Teorema de Sharkovsky, no obstante, se introduce, sin nombre, la noción de conjuntos "scrambled", los cuales dan origen a lo que hoy se conoce con el nombre de Caos en el sentido de Li-Yorke.

Lo más llamativo de la teoría de sistemas dinámicos discretos es su novedad. Es en la segunda mitad del siglo XX (más específicamente en la década de los 60) cuando se despierta la curiosidad en la dinámica discreta, después del descubrimiento de el teorema de Sharkovsky y el caos en el sentido de Li-Yorke. Por otro lado, los ordenadores modernos jugaron también un importante rol, ayudando a descubrir impresionantes fenómenos matemáticos que habían estado ocultos hasta entonces. En la primera parte del trabajo veremos mas a fondo el teorema de Sharkovsky.

Para la segunda parte trataremos sobre el teorema de recurrencia de Poincaré para eso veremos una pequeña motivación: Supongamos que deseamos estimar la cantidad de peces de una laguna. Para tal efecto, echamos una red en un lugar determinado de la laguna, A, y sacamos por ejemplo 1000 peces, que pintaremos de un color, esta vez no tóxico, y volveremos a echar al agua. Esperaremos un tiempo razonable para que naden libremente, volvemos a echar la red en el lugar A y sacamos 1000 peces. La idea del algoritmo de estimación era contar cuántos peces de color se encontraban ahora en A, y de acuerdo a ese porcentaje, interpolar el porcentaje total de peces en la laguna, deduciendo el número de peces. Por ejemplo, si en A encontramos 10 peces, o sea el 1 de la muestra, eso nos estaría diciendo que 1000 eran aproximadamente el 1 de los peces de la laguna, que serian entonces aproximadamente 100.000. Buenísimo, ahora, qué nos asegura que algún pez vuelva a caer en la red?, y si ningún pez volviera a pasar nunca por A? En ese caso, nuestro algoritmo estaría mal, ya que nos estaría diciendo que los 1000 peces coloreados de la muestra es el 0 de los peces del lago Veamos si podemos contestar a las preguntas vuelve algún pez coloreado a caer en la red?, cuántos vuelven? cuántas veces vuelven?

El otro teorema que trataremos en esta parte, es el teorema de Birkhoff. Uno de los grandes objetivos de los sistemas dinámicos busca describir el comportamiento de las órbitas $\{T^n(x) : n \geq 0\}$ de una transformación $T : X \to X$, donde $T^0(x) = Id$ y $T^{n+1} = T \quad o \quad T^n$ para $n \geq 0$. Este estudio incide en medir las cantidades $f(T^n(x))$ para alguna función $f : X \to \mathbb{R}$ particularmente en términos de la media

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

Una cuestión básica de la teoría ergódica es la existencia de esta media cuando $n \to \infty$. Es claro que tal media existe siempre que x sea un punto periódico, esto es, cuando $T^k(x) = x$ para algun $k \geq 1$. En 1931 Birhkoff (Birkhoff, George David (1931), Proof of the ergodic theorem, Proc Natl Acad Sci USA 17(12): 656-660) probó un resultado que asegura que si T tiene una medida de probabilidad invariante μ, esto es, $\mu(T^{-1}(A)) = \mu(A)$ para todo A medible, y f es integrable con respecto a μ, entonces estas medias existen para todo $x \in X$(esto es exceptuando un conjunto de medida nula con respecto a la medida μ). Este resultado es conocido como el teorema ergódico de Birkhoff. Una condición necesaria y suficiente para que el valor del límite sea el mismo en casi todo punto de X es no exista ningún medible A con $0 < \mu(A) < 1$ y $T^{-1}(A) = A$. En estas condiciones

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

en casi todo punto de $x \in X$.
La transformación T se dice en este caso que es ergódica. De un punto de vista de aplicaciones prácticas, puede ser interesante saber si el recíproco de este resultado también vale: es decir si se da la igualdad, podremos decir que T es ergódica? Buczolich(Measures and functions with prescribed homogeneous multifractal spectrum Zoltán Buzcolich Volume 1, Issue 3, 2014, pp. $295 - 333$) muestra que la respuesta es en general, negativa. Pero si $f \leq 0$ el recíproco también es verdadero.

Capítulo 1

Nociones de entropía topológica

Deseamos conocer un poco sobre las cadenas de Markov, las cuales serán utilizadas mas adelante en nuestras demostraciones, y además dar una pequeña idea de lo que es la entropía topológica. La principal referencia de este capítulo son los libros [KH95, dMvS93] y además para análisis del libro [Lim81].

Definición 1.1. *Tomemos* $N \geq 2$ *el espacio*

$$\Omega_N = \{w = (\ldots, w_{-1}, w_0, w_1, \ldots)/w_i \in \{0, 1, \ldots, N-1\} \text{ para } i \in \mathbb{Z}\}.$$

Análogamente

$$\Omega_N^R = \{w = (w_0, w_1, w_2, \ldots)/w_i \in \{0, 1, \ldots, N-1\} \text{ para } i \in \{0, 1, \ldots\}\}.$$

Fijemos enteros $n_1 < n_2 < \ldots < n_k$ *y números*
$\alpha_1, \alpha_2, \ldots, \alpha_k \in \{0, 1, \ldots N-1\}$ *y llamemos al subconjunto*

$$C_{\alpha_1, \alpha_2, \ldots, \alpha_k}^{n_1, n_2, \ldots, n_k} = \{w \in \Omega_N / w_{n_i} = \alpha_i \text{ para } i = 1, 2, \ldots, k\}$$

un cilindro y el número k *de digitos fijos es el rango de ese cilindro. Los cilindros en el espacio* Ω_N^R *son definidos similarmente.*

Definición 1.2. *Decimos que un cilindro es simétrico si es de la forma* $C_{\alpha_{-n}, \ldots, \alpha_n}^{-n, \ldots, n}$

Una forma de definir la topología en el espacio Ω_N (y similarmente en Ω_N^R) es declarando que todos los cilindros son conjuntos abiertos y estos forman una base para la topología. Entonces cada cilindro también es cerrado debido a que

$$(C^{n_1,n_2,\ldots,n_k}_{\alpha_1,\alpha_2,\ldots,\alpha_k})^c = \{w \in \Omega_N / w_{n_i} \neq \alpha_i \text{ para } i = 1, 2, \ldots, k\}$$
$$= \bigcap_{i=1}^{k} \{w \in \Omega_N / w_{n_i} \neq \alpha_i\}$$

ahora bien veamos que $\{w \in \Omega_N / w_{n_i} \neq \alpha_i\}$ es abierto ya que

$$\{w \in \Omega_N / w_{n_i} \neq \alpha_i\} = \{w \in \Omega_N / w_{n_i} = 0\} \cup \{w \in \Omega_N / w_{n_i} = 1\} \ldots$$

$$\cup\{w \in \Omega_N / w_{n_i} = \alpha_i - 1\} \cup \{w \in \Omega_N / w_{n_i} = \alpha_i + 1\} \cup \ldots \cup \{w \in \Omega_N / w_{n_i} = N - 1\}$$

esto quiere decir que cada elemento de la intersección es unión numerable de cilindros, por lo tanto cada elemento de la intersección es un cilindro. Como el complemento de un cilindro es intersección finita de estos elementos que sabemos que son abiertos, entonces el complemento de un cilindro es abierto. Con lo cual concluimos que el complemento de un cilindro es abierto, entonces el cilindro es cerrado. Entonces un cilindro es abierto y cerrado a la vez.

Definición 1.3. *La transformación shift también conocida como el desplazamiento hacia la izquierda en Ω_N, se define de la manera siguiente*

$$\sigma_N : \Omega_N \to \Omega_N; \;\; \sigma_N(w) = w' = (\ldots, w'_0, w'_1, \ldots) \text{ donde } w'_n = w_{n+1}$$

Observe que σ_N es uno a uno (inyectiva) pues si

$$\sigma(w) = \sigma(t) \text{ esto significa que } w' = t' \text{ donde } w'_n = w_{n+1} \text{ y } t'_n = t_{n+1}$$

esto significa que $w_{n+1} = t_{n+1}; \;\; \forall n \in \mathbb{Z}$ de donde
$w_n = t_n; \;\; \forall n \in \mathbb{Z}$ de donde $w = t$. Y lleva cilindros en cilindros ya que

$$\sigma(C^{n_1,n_2,\ldots,n_k}_{\alpha_1,\alpha_2,\ldots,\alpha_k}) = \sigma(\{w \in \Omega_N / w_{n_i} = \alpha_i \text{ para } i = 1, 2, \ldots, k\})$$

$$= \{w \in \Omega_N / w_{n_i+1} = \alpha_i \text{ para } i = 1, 2, \ldots, k\} = C^{n_1+1,n_2+1,\ldots,n_k+1}_{\alpha_1,\alpha_2,\ldots,\alpha_k}$$

Como σ es continua entonces es un homeomorfismo ya que es continua, inyectiva, lleva abiertos en abiertos y cerrados en cerrados, además su inversa tambien es continua. A veces el σ_N es llamado "topological Bernoulli shift".

Definición 1.4. *Definimos el shift de un solo lado como*

$$\sigma_N^R : \Omega_N^R \to \Omega_N^R \qquad por \quad \Omega_N^R(w_0, w_1, \ldots) = (w_1, w_2, \ldots)$$

Lema 1.0.1. *Los puntos periódicos de los shifts σ_N y σ_N^R de período n son densos en Ω_N y Ω_N^R correspondientemente, además $P_N(\sigma_N) = P_N(\sigma_N^R) = N^n$.*

Demostración. Las órbitas periódicas para shift son secuencias periódicas, esto es,

$$(\sigma_N)^m w = w \text{ si y solo si } w_{n+m} = w_n \ \forall n \in \mathbb{Z} \text{ y similarmente para } \sigma_N^R$$

Con el fin de probar la densidad de los puntos periódicos es suficiente encontrar un punto periódico en cualquier cilindro. Desde que cualquier cilindro en Ω_N contiene un cilindro simétrico de rango $2m+1$ para algún m como $C^{-m,\ldots,m}_{\alpha_{-m},\ldots,\alpha_m} = C^m_\alpha$, donde $\alpha = \alpha_{-m}, \ldots, \alpha_m$, basta con considerar solo estos cilindros. Pero la sucesión se obtiene por simple repetición de la sucesión finita $\alpha_{-m}, \ldots, \alpha_m$, es decir, w donde $w_n = \alpha_{n'}$ para $\mid n' \mid \leq m,\ n' = n(mod\ 2m+1)$, obviamente se encuentra en nuestro cilindro y tiene periodo $2m+1$. Cualquier sucesión periódica w de período n es unicamente determinada por las coordenadas $w_0, \ldots, w_{n-1}$.
Y como sabemos existen N^n diferentes sucesiones finitas $(w_0, \ldots, w_{n-1})$. □

Definición 1.5. *La restricción de shift σ_N o σ_N^R a cualquier subconjunto cerrado, invariante de Ω_N o Ω_N^R, respectivamente es llamado sistema dinámico simbolico.*

1.1. Cadenas de Markov

Aquí vamos a considerar una clase en especial (aunque probablemente la más importante) de sistemas dinámicos simbólico. Sea $A = (a_{ij})_{i,j=0}^{N-1}$ una matriz $N \times N$ cuyas entradas a_{ij} son o bien ceros, o bien unos (que llamamos $0-1$ matriz).
Sea

$$\Omega_A = \{w \in \Omega_N / a_{w_n w_{n+1}} = 1 \text{ para } n \in \mathbb{Z}\}.$$

En otras palabras la matriz A determina todas las transiciones admisibles entre los símbolos $0, 1, \ldots, N-1$. El conjunto Ω_A es obviamente invariante por shift.

Ejemplo:
Si tenemos la siguiente matriz

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Esto dice que, el punto x_1 esta relacionado con x_2, x_3 y x_4; el punto x_2 solo esta relacionado con el punto x_3; el punto x_3 solo esta relacionado consigo mismo; y el punto x_4 esta relacionado con x_1 y x_3

Definición 1.6. *Sea X un espacio métrico compacto con la función distancia d y $f : X \to X$. La sucesión creciente de métricas $d_n^f, n = 1, 2, \ldots,$ iniciando desde $d_1^f = d$ por $d_n^f = \max\limits_{0 \le i \le n-1} d(f^i(x), f^i(y))$.*

En otras palabras, d_n^f mide la distancia entre los segmentos de la órbita $I_x^n = \{x, f(x), \ldots, f^{n-1}(x)\}$ y I_y^n.
Denotemos la bola abierta $\{y \in X / d_n^f(x, y) < \epsilon\}$ por $B_f(x, \epsilon, n)$. El conjunto $E \subset X$ es llamado (n, ϵ) cubridor si $X \subset \bigcup\limits_{x \in E} B_f(x, \epsilon, n)$.
Sea $S_d(f, \epsilon, n)$ la cardinalidad mínima de un (n, ϵ) cubridor,

$$h_d(f, \epsilon) = \overline{\lim_{n \to \infty}} \frac{1}{n} \log S_d(f, \epsilon, n)$$

Definición 1.7. *Se define la entropía topológica como:*

$$h_d(f) = \lim_{\epsilon \to 0} h_d(f, \epsilon)$$

Lema 1.1.1. *Sea d' otra métrica en X. Entonces $h_{d'}(f) = h_d(f)$*

Demostración. Sea $D_\epsilon = \{(x_1, x_2) \in X \times X / d(x_1, x_2) \ge \epsilon\}$ (como D_ϵ es cerrado y esta incluido en $X \times X$ entonces se concluye que D_ϵ es compacto). Como d' es continua en $X \times X$ entonces posee un minimo en D_ϵ

$$\Rightarrow \exists\ x_1^*, x_2^* / d'(x_1^*, x_2^*) = \delta(\epsilon) \le d'(x_1, x_2)\ \forall (x_1, x_2) \in D_\epsilon$$

Afirmo que $\delta(\epsilon) > 0$, ya que si

$$\delta(\epsilon) = 0 \Rightarrow d'(x_1^*, x_2^*) = 0 \Rightarrow x_1^* = x_2^*.$$

Como X es compacto, entonces alcanza su mínimo en D_ϵ. Entonces

$$d(x_1^*, x_2^*) \geq \epsilon \Rightarrow 0 \geq \epsilon$$

lo cual es una contradicción.
Afirmo que $d'(x_1, x_2) < \delta(\epsilon) \Rightarrow d(x_1, x_2) < \epsilon$, ya que caso contrario si

$$d(x_1, x_2) \geq \epsilon \Rightarrow (x_1, x_2) \in D_\epsilon \Rightarrow \delta(\epsilon) > d'(x_1, x_2) \geq \delta(\epsilon)$$

lo cual es una contradicción.

Esto significa que

$$B_{d'}(x_1, \delta(\epsilon)) \subset B_d(x_1, \epsilon) \Rightarrow B_{d'^f_n}(x_1, \delta(\epsilon)) \subset B_{d^f_n}(x_1, \epsilon)$$

$$\Rightarrow S_{d'}(f, \delta(\epsilon), n) \geq S_d(f, \epsilon, n) \Rightarrow h_{d'}(f, \delta(\epsilon)) \geq h_d(f, \epsilon)$$

Ahora si vemos a δ como una función esta sería monotóna, decreciente acotada inferiormente y continua. Entonces existe $\lim\limits_{\epsilon \to 0} = l$ y $\delta(\epsilon) \geq l$. Además nos damos cuenta que si $\epsilon \leq \epsilon'$ entonces $h_d(f, \epsilon) \geq h_d(f, \epsilon')$. En particular

$$\lim_{\epsilon \to 0} h_d(f, \epsilon) \geq h_d(f, \epsilon') \Rightarrow h_{d'}(f, l) \geq h_{d'}(f, \delta(\epsilon))$$

$$\lim_{\epsilon \to 0} h_{d'}(f, \epsilon) \geq h_{d'}(f, \delta(\epsilon))$$

En resumen:
a) $\exists \lim\limits_{\epsilon \to 0} \delta(\epsilon) = l,\ \delta(\epsilon) \geq l$
b) Si $l = 0$ se cumple que $h_{d'}(f) \geq h_d(f)$, basta con tomar limite
c) Si $l > 0$ como $\delta(\epsilon) \geq l \Rightarrow h_{d'}(f, \delta(\epsilon)) \geq h_d(f, \epsilon)$ además por las observaciones hechas tenemos que

$$h_{d'}(f) = \lim_{\epsilon \to 0} h_{d'}(f, \epsilon) \geq h_d(f, l) \geq h_d(f, \epsilon)$$

Con lo que demostramos una desigualdad, para demostrar la otra, basta invertir los papeles de d y d', asi podemos decir que la entropía topológica está bien definida.

□

Corolario 1.1.1. *La entropía topológica es invariante por conjugación topológica*

Demostración. Sea $f : X \to X,\ g : Y \to Y$ topologicamente conjugado vía el homeomorfismo $h : X \to Y$. Fije la métrica d en X y defina d' en Y como

$$d'(y_1, y_2) = d(h^{-1}(y_1), h^{-1}(y_2)).$$

Entonces h se convierte en una isometría, ya que, $\exists x_1, x_2 / h(x_1) = y_1; h(x_2) = y_2$ tendriamos que $d'(h(x_1), h(x_2)) = d(y_1, y_2)$ así $h_d(f) = h_{d'}(g)$ $\square$

Definición 1.8. *El mapeo $g : N \to N$ es un factor (o factor topológico) de $f : M \to M$ si existe un mapeo continuo sobreyectivo $h : M \to N$ tal que $h\ o\ f = g\ o\ h$.*

Lema 1.1.2. *Si g es un factor de f, entonces $h_{top}(g) \leq h_{top}(f)$.*

Demostración. Sea $f : X \to X,\ g : Y \to Y,\ h : X \to Y,\ \ h \text{ o } f = g \text{ o } h$, $h(X) = Y$ y d_X, d_Y sus metricas correspondientes, h es uniformemente continua (ya que es h continua en el espacio compacto X) esto es

$$\forall \epsilon > 0\ \exists \delta(\epsilon > 0) / d_X(x_1, x_2) < \delta(\epsilon) \Rightarrow d_Y(h(x_1), h(x_2))$$

$$\Rightarrow B_X(x_1, \delta(\epsilon)) \subset B_Y(h(x_1), \epsilon) \Rightarrow S_{d_X}(f, \delta(\epsilon), n) \geq S_{d_Y}(g, \epsilon, n)$$

$$\therefore h_{top}(X) \geq h_{top}(Y).$$

$\square$

Definición 1.9. *Definimos $D_d(f, \epsilon, n)$ como el minimo número de conjuntos cuyo diametro en la metrica d_n^f es menor que ϵ y cuya union cubre a X.*

Vemos que

$$D_d(f, 2\epsilon, n) \leq S_d(f, \epsilon, n)$$

y

$$S_d(f, \epsilon, n) \leq D_d(f, \epsilon, n)$$

lo primero es obvio ya que las bolas de radio ϵ son conjuntos de radio 2ϵ. Para lo segundo si suponemos que $S_d(f, \epsilon, n) > D_d(f, \epsilon, n)$ y sabiendo que $B(x, \frac{\epsilon}{2})$ tienen diámetro ϵ tendriamos que

$$S_d(f, \epsilon, n) > D_d(f, \epsilon, n) \geq S_d(f, \epsilon, n)$$

lo cual es una contradicción.

Lema 1.1.3. *Para todo $\epsilon > 0$ el límite* $\lim_{n\to\infty} \frac{1}{n}\log D_d(f,\epsilon,n)$ *existe.*

Demostración. Vamos a demostrar la desigualdad

$$D_d(f,\epsilon,m+n) \leq D_d(f,\epsilon,n)\cdot D_d(f,\epsilon,m) \ \ \forall m,n\in\mathbb{N}.$$

La sucesión $a_n = \log D_d(f,\epsilon,n)$ es subaditiva, es decir, $a_{m+n} \leq a_m + a_n$ lo cual nos diría que $\lim_{n\to\infty}\frac{a_n}{n}$ existe.
Sea A un conjunto con d_n^f-diametro menor que ϵ y B un conjunto con d_m^f-diametro menor que ϵ entonces mostraré que $A\cap f^{-n}(B)$ es un conjunto de d_{m+n}^f-diametro menor que ϵ. Sea $x,y\in A\cap f^{-n}(B)$ como $x,y\in A \Rightarrow d_n^f(x,y)<\epsilon$
(es decir $d(f^i(x),f^i(y))<\epsilon$ con $0\leq i\leq n-1$). Como $x,y\in f^{-n}(B)$ lo que es lo mismo decir que $f^n(x),f^n(y)\in B$ entonces

$$d_m^f(f^n(x),f^n(y))<\epsilon \Rightarrow d(f^{i+n}(x),f^{i+n}(y))<\epsilon \text{ con } 0\leq i\leq m-1$$

$$\Rightarrow d(f^k(x),f^k(y))<\epsilon \text{ con } 0\leq k\leq m+n-1$$

$$\Rightarrow d_{m+n}^f(x,y)<\epsilon.$$

Entonces $A\cap f^{-n}(B)$ es un conjunto d_{m+n}^f-diametro menor que ϵ.
Sea U una cobertura de X por $D_d(f,\epsilon,n)$ conjuntos de d_n^f-diametro menor que ϵ y D una cobertura de X por $D_d(f,\epsilon,m)$ conjuntos de d_m^f-diametro menor que ϵ, la cobertura de todos los conjuntos $A\cap f^{-n}(B)$, donde
$A\in U, B\in D$ contiene no más que $D_d(f,\epsilon,n)\cdot D_d(f,\epsilon,m)$ conjuntos, y estos cubren X por conjuntos con d_{m+n}^f-diametro menor que ϵ, ya que

$$\bigcup_{A\in U,B\in D} A\cap f^{-n}(B) = \bigcup_{A\in U} A\cap f^{-n}(\bigcup_{B\in D} B) = X\cap X = X.$$

Es decir

$$D_d(f,\epsilon,n)\cdot D_d(f,\epsilon,m) \geq D_d(f,\epsilon,m+n).$$

De aquí se concluye lo que se quiere. □

Definición 1.10. *Definimos* $\lim_{n\to\infty}\frac{1}{n}\log D_d(f,\epsilon,n) = \widetilde{h}_d(f,\epsilon)$.

De las observaciones hechas tenemos que

$$\widetilde{h}_d(f,\epsilon) \geq h_d(f,\epsilon) \geq \widetilde{h}_d(f,2\epsilon).$$

Lema 1.1.4.

(1) Sea $\Lambda \subset X$ cerrado. Entonces $h_{top}(f) \geq h_{top}(f \mid \Lambda)$.
(2) Sea Λ_i con $i = 1 : m$ cerrados, tal que $\bigcup_{i=1}^{m} \Lambda_i = X$. Entonces $\max\limits_{1 \leq i \leq m} h_{top}(f \mid \Lambda_i) = h_{top}(f)$.

Demostración. (1) Como $\Lambda \subset X$, Λ cerrado, y X compacto entonces Λ compacto, sabemos que

$$S_d(f, \epsilon, n) \geq S_d(f \mid \Lambda, \epsilon, n).$$

Dividimos entre $\frac{1}{n}$, tomamos límite superior, y hacemos $\epsilon \to 0$ y obtenemos

$$h_{top}(f \mid \Lambda) \leq h_{top}(f)$$

(2) Sea $D_d(f \mid \Lambda_i, \epsilon, n)$. Es claro que

$$D_d(f, \epsilon, n) \leq \sum_{i=1}^{n} D_d(f \mid \Lambda_i, \epsilon, n)$$

ya que en la segunda expresión tenemos un número de conjuntos que cubre a $\bigcup\limits_{i=1}^{n} \Lambda_i = X$(ya que cada sumando cubre a Λ_i) y en la primera expresión tenemos el mínimo número de conjuntos que cubren a X. Es decir

$$\exists i / D_d(f \mid \Lambda_i, \epsilon, n) \geq \frac{1}{m} D_d(f, \epsilon, n)$$

$$\widetilde{h}(f \mid \Lambda_i, \epsilon) = \overline{\lim_{n\to\infty}} \frac{\log D_d(f \mid \Lambda_i, \epsilon, n)}{n} \geq \overline{\lim_{n\to\infty}} \frac{D_d(f, \epsilon, n) - \log m}{n} = \widetilde{h}_d(f, \epsilon)$$

$$\Rightarrow \max_{1 \leq i \leq m} h_{top}(f \mid \Lambda_i) \geq h_{top}(f).$$

La otra de desigualdad se obtiene aplicando la parte (1) del lema, con lo cual tendriamos que

$$\max_{1 \leq i \leq m} h_{top}(f \mid \Lambda_i) = h_{top}(f).$$

□

1.2. Teorema de Sharkovsky

Definición 1.11. *Considere el mapeo continuo $f : I \to I$, decimos que $J \subset I$ cubre a $K \subset I$ (bajo f) si $K \subset f(J)$ y lo denotamos por $J \to K$. Si J cubre exactamente a K lo denotamos por $J \rightarrowtail K$. Un intervalo cuyos extremos están en un ciclo O de f es llamado O-intervalo. Si contiene solo dos puntos de O entonces diremos que es un O-intervalo básico, y que estos puntos son adyacentes.*

Lema 1.2.1. *Si J,K son intervalos, K es cerrado y $J \to K$ entonces existe un intervalo cerrado $L \subset J$ tal que $f(L) = K$.*

Demostración. Sea $K = [c, d]$ sabemos que $[c, d] \subset f(J)$ entonces definimos el siguiente conjunto

$$C = \{w \in J : f(w) \in \{c, d\}\} = (f^{-1}(\{c\}) \cup f^{-1}(\{d\})) \cap J.$$

Sabemos que $C \neq \phi$, tomamos $\alpha = \text{ínf}\, C$, como $\alpha \in \overline{C}$ y C es un conjunto cerrado (ya que f es continua) entonces $\alpha \in C$ de donde $f(\alpha) \in \{c, d\}$ ahora definimos

$$D = \{w \in J/f(w) = d, w \geq \alpha\}.$$

Observamos que $D \neq \phi$ ya que si $f(\alpha) = d$ no tenemos nada que probar. Si $f(\alpha) = c$; como sabemos que existe un $y \in J$ tal que $f(y) = d$, de donde $y \in C$ entonces $y \geq \text{ínf}\, C = \alpha$ de donde $y \in D$. Supongamos sin pérdida de generalidad que $f(\alpha) = c$. Tomo $\beta = \text{ínf}\, D$, como D es un conjunto cerrado entonces $\beta \in D$, de donde $f(\beta) = d$. Definimos $L = [\alpha, \beta]$ y afirmo que $f(L) = K = [c, d]$.
Sea $y \in [c, d]$, como $f(\alpha) = c$ y $f(\beta) = d$ por el teorema de valor intermedio existe $x \in [\alpha, \beta] = L$ tal que $f(x) = y$ entonces $K \subset f(L)$.
Sea $y = f(x)$ con $x \in <\alpha, \beta>$, si $y > d$, $f(\alpha) = c$, $f(x) > d$ entonces por el teorema de valor intermedio existe $x' \in [\alpha, x]$ tal que $f(x') = d$ pero $x' \leq x < \beta$ de donde $x' \in D(\rightarrow\leftarrow)$ ya que β es el mínimo elemento de D.
Si $y < c$ hacemos la prueba análoga y llegaremos a una contradicción de donde $f(<\alpha, \beta>) \subset [c, d]$ y sabiendo que $f(\alpha) = c$, $f(\beta) = d$ entonces $f([\alpha, \beta]) \subset [c, d]$. □

Lema 1.2.2. *Si $f : J \to J$ entonces f tiene un punto fijo $x \in J$.*

Demostración. Sea $J = [a, b]$ como $J \to J$ entonces $J \subset f(J)$ de donde como $a \in J$ entonces $a \in f(J)$, de aquí se tiene que existe $c \in J$ tal que $f(c) = a$, análogamente existe $d \in J$ tal que $f(d) = b$.
Definimos $g(x) = f(x) - x$, de aquí $g(c) = f(c) - c = a - c \leq 0$, analogamente $g(d) = f(d) - d = b - d \geq 0$ entonces tenemos que

$$g(c) \leq 0 \leq g(d),$$

de donde por el teorema de valor intermedio existe $e \in J$ tal que $g(e) = 0$, lo cual es equivalente a decir que $f(e) = e$, de donde se concluye lo pedido. □

Lema 1.2.3. *Si* $I_0 \to I_1 \to I_2 \ldots \to I_n$ *entonces* $\bigcap_{i=0}^{n} f^{-i}(I_i)$ *contiene un intervalo* Δ_n *tal que* $f^n(\Delta_n) = I_n$.

Demostración. Tenemos que $I_0 \to I_1$ lo cual es equivalente a decir que $I_1 \subset f(I_0)$ entonces existe $\Delta_1 \subset I_0$ tal que $f(\Delta_1) = I_1$. Como $I_1 \to I_2$ entonces $I_2 \subset f(I_1)$ entonces existe $\delta_2 \subset I_1$ tal que $f(\delta_2) = I_2$. Como $\delta_2 \subset I_1 = f(\Delta_1)$ entonces por definicion $\Delta_1 \to \delta_2$ entonces existe $\Delta_2 \subset \Delta_1$ tal que $f(\Delta_2) = \delta_2$ entonces $f^2(\Delta_2) = I_2$. Nuevamente repetimos el mismo procedimiento, como $I_2 \to I_3$ lo que es lo mismo decir que $I_3 \subset f(I_2)$ entonces existe $\delta_3 \subset I_2 = f^2(\Delta_2)$ tal que $f(\delta_3) = I_3$. Además sabemos que $\Delta_2 \to^{f^2} \delta_3$ entonces existe $\Delta_3 \subset \Delta_2$ tal que $f^2(\Delta_3) = \delta_3$ de lo cual se ve que $f^3(\Delta_3) = I_3$.
Así siguiendo inductivamente tendremos un Δ_n tal que $f^n(\Delta_n) = I_n$ y además dado un $0 \leq i \leq n$ tenemos que

$$\Delta_n \subset \Delta_i \Rightarrow f^i(\Delta_n) \subset f^i(\Delta_i) = I_i \Rightarrow f^i(\Delta_n) \subset I_i.$$

De donde tenemos que

$$\Delta_n \subset f^{-i}(I_i); \quad \forall i \in \{0, 1, \ldots, n\}.$$

De aquí concluimos que $\Delta_n \subset \bigcap_{i=0}^{n} f^{-i}(I_i)$ de donde lo cual demuestra el lema.

□

El siguiente resultado es llamado lema del itinerario, el cual nos dice que si tenemos un lazo de longitud n que comienza y termina en el mismo intervalo entonces existe un punto en ese intervalo tal que se cumple dos cosas, primero la órbita de ese punto recorre el lazo en orden y segundo ese punto tiene período n:

Lema 1.2.4. *Si $I_0 \to I_1 \to I_2 \to \ldots \to I_{n-1} \to I_0$ (esto es llamado un lazo o un n-lazo de intervalos) entonces existe x tal que $f^n(x) = x$ y además $f^i(x) \in I_i$ para $0 \leq i \leq n-1$*

Demostración. Aplicando el lema anterior tenemos que $\Delta_n \subset I_0$, $f^n(\Delta_n) = I_0$ y además que $\Delta_n \subset \bigcap_{i=0}^{n} f^{-i}(I_i)$ de donde $f^i(\Delta_n) \subset I_i \;\; \forall i$. Además como $\Delta_n \subset I_0 = f^n(\Delta_n)$ entonces $\Delta_n \to \Delta_n$ entonces existe $x \in \Delta_n$ tal que $f^n(x) = x$ y por la propia construcción tenemos que $f^i(x) \in I_i$.

□

Proposition 1.2.1. *Suponga que $f : [0,1] \to [0,1]$ tiene un punto periódico de período tres. Entonces f posee puntos de todos los períodos.*

Demostración. La prueba está en [KH95]. Supongamos que $\{x_1 < x_2 < x_3\}$ sea la órbita de ese punto periódico, tomamos $I_1 = [x_1, x_2]$ y $I_2 = [x_2, x_3]$. Supongamos sin pérdida de generalidad que $f(x_2) = x_3$ entonces $f^2(x_2) = x_1, f(x_3) = x_1$ entonces $I_2 = [x_2, x_3] \to I_1 = [x_1, x_2]$ ya que $f(x_2) = x_3, f(x_3) = x_1$ de donde se obtiene que $I_1 \subset f(I_2)$. Además $[x_2, x_3] = I_2 \to I_2 = [x_2, x_3]$ ya que $f(x_2) = x_3$ y $f(x_3) = x_1$ de donde se obtiene que $I_2 \subset f(I_2)$, $I_1 = [x_1, x_2] \to I_2 = [x_2, x_3]$ ya que $f(x_1) = x_2$ y $f(x_2) = x_3$. De aquí obtenemos que $I_2 \subset f(I_1)$ (si $f(x_2) = x_1$ llamamos a I_1 como I_2 y viceversa y obtendremos la misma conclusión), haciendo su diagrama de Markov obtenemos que

$$I_1 \leftrightarrows\circlearrowright I_2 \text{ para todo } n \in \mathbb{N}$$

entonces se tiene

$$I_1 \to I_2 \to I_2 \to \ldots \to I_2 \to I_1$$

(para n-1 ocurrencias de I_2),supongamos que $n > 3$, luego por el corolario existe

$$x \in I_1 \text{ y } f(x), f^2(x), \ldots, f^{n-1}(x) \in I_2$$

Caso 1: Supongamos que $x \notin \{x_1, x_2, x_3\}$, sabemos que para $i = 1, 2, \ldots, n-1$ tenemos que $f^i(x) \in I_2$. Si para alguno de estos "$i's$" tuvieramos que $f^i(x) \in$

$\{x_2, x_3\}$ entonces $f^j(x) \in \{x_1, x_2, x_3\}$ $\forall j \geq i$ entonces $f^n(x) \in \{x_1, x_2, x_3\}$ entonces $f^n(x) \neq x(\rightarrow\leftarrow)$

$$\therefore \forall i = 1, 2, \ldots, n-1 \text{ tenemos } f^i(x) \in I_2 \diagdown \{x_2, x_3\}$$

$$\Rightarrow f^i(x) \notin I_1 \Rightarrow f^i(x) \neq x; \quad \forall i = 1, 2, \ldots, n-1$$

pues $f^i(x) \notin I_1$ y $x \in I_1$ de donde se deduce que x tiene periodo n.
Caso 2: $x \in \{x_1, x_2, x_3\}$ entonces $x = x_1$ o $x = x_2$ ya que $x \in I_1$ y x tiene periodo tres. Como $f^n(x) = x$ deducimos que n es multiplo de 3, es decir $n = 3k, k \in \mathbb{N}$ (en particular $n \geq 6$ ya que $n \neq 3$)
Caso 2.1: $x = x_1 \Rightarrow f(x) = x_2, f^2(x) = x_3, f^3(x) = x_1 \notin I_2(\rightarrow\leftarrow)$.
Caso 2.2: $x = x_2 \Rightarrow f(x) = x_3, f^2(x) = x_1 \notin I_2(\rightarrow\leftarrow)$

$$\therefore \text{ Existen puntos de periodo } n, \ \forall n > 3.$$

Por lo tanto el caso 2 no se puede dar, con lo que nos quedamos con el caso 1.
Existe punto de periodo $n = 3$ por hipótesis, además existen puntos periódicos de periodo $n = 1$ ya que $I_2 \rightarrow I_2$ y aplicando el lema anterior encontraremos que f tiene punto fijo y de período $n = 2$ ya que $I_1 \rightarrow I_2 \rightarrow I_1$ mediante f. De donde $I_1 \rightarrow I_1$ mediante f^2, de aquí aplicando el corolario anterior tenemos que

$$\exists x \in I_1 \text{ tal que } f^2(x) = x; \quad f(x) \in I_2.$$

Si $f(x) = x$ y como $f(x) \in I_2$ y $x \in I_1$ tenemos que $x \in I_1 \cap I_2 = \{x_2\}$, de donde se tiene que $x = x_2$ pero esto no puede ser ya que x_2 tiene período tres. Por lo tanto

$$\exists x \in I_1; f^2(x) = x \text{ y } f(x) \neq x.$$

Esto implica que x tiene período 2, por lo tanto de todo ,lo hecho anteriormente podemos deducir que existe punto de cualquier período. $\square$

Capítulo 2

Teorema de Sharkovsky

En esta capítulo, usaré de referencia el libro [Rud87] y [BC92] Los resultados de la dinámica unidimensional que conforman lo que en la actualidad se conoce con el nombre de teorema de Sharkosvky involucran una función continua del intervalo en sí mismo, y un orden en los enteros positivos llamado el orden de Sharskovsky el cual esta dado por:

$$3 \rhd 5 \rhd 7 \rhd \ldots \rhd 2 \times 3 \rhd 2 \times 5 \rhd \ldots \rhd 2^2 \times 3 \rhd 2^2 \times 5 \rhd \ldots \rhd 2^3 \rhd 2^2 \rhd 2^1 \rhd 2^0 = 1.$$

Dado que todo número entero positivo puede ser escrito de manera única en la forma $2^m(2n+1)$ para algunos enteros $m, n \geq 0$, el orden definido anteriormente es un orden total. Además se escribe $m \rhd l$ o $l \lhd m$, si m está a la izquierda de l y este orden también puede ser definido formalmente por:

$$2^a(2b+1) \rhd 2^\alpha(2\beta+1) \text{ si y solo si}$$

$$a < \alpha \text{ y } 0 < b, \beta \text{ o}$$

$$a = \alpha \text{ y } 0 < b < \beta \text{ o}$$

$$a > \alpha \text{ y } 0 = b = \beta.$$

La lista comienza con los números impares mayores que 1 ordenados de forma creciente. Luego se repite la secuencia con cada impar multiplicado por 2, después la secuencia inicial es multiplicada por 2^2, después por 2^3 y así sucesivamente. Al final se colocan las potencias de 2 en orden decreciente.

Definición 2.1. *Una cola del orden de Sharkovsky es un conjunto $\Gamma \subset \mathbb{N}$ tal que $s \rhd t$ para todo $s \notin \Gamma$.*

Hay tres tipos de colas: $\{m\} \cup \{l \in \mathbb{N} : l \lhd m\}$ para algún $m \in \mathbb{N}$, el conjunto $\{..., 2^4, 2^3, 2^2, 2, 1\}$ de todas las potencias de 2 y ϕ.
Ya habíamos definido anteriormente $I \rightarrow J$ mediante una función f, si $J \subset f(I)$, ahora complementaremos la anterior definición

Definición 2.2. *Decimos que I cubre exactamaente a J y lo denotaremos por $I \rightarrowtail J$ si $J = f(I)$.*

Definición 2.3. *Un intervalo cuyos extremos están en un ciclo O es llamado $O-$intervalo. Si contiene solo dos puntos de O diremos que es un $O-$intervalo básico, y que estos dos puntos son adyacentes.*

Nuestro objetivo en esta parte es demostrar que si tenemos una función continua de un intervalo sobre si mismo tenemos que

Teorema 2.0.1. *Si m es un período para f y $m \rhd l$, entonces l también es un período para f.*

Este es el resultado principal de esta parte, y para hacerla necesitamos varias herramientas que iremos demostrando a lo largo de esta tesis, comencemos por:

Lema 2.0.1. *Sean $g : J \rightarrow \mathbb{R}$ una función continua e $I \subset J$ un intervalo compacto. Si $I \rightarrow I$, entonces g tiene un punto fijo en I.*

Demostración. Sea $I = [\beta_0, \beta_1]$. Dado que $I \subset g(I)$, existen α_0, $\alpha_1 \in I$ tales que $g(\alpha_i) = \beta_i$ para $i = 0, 1$. Así $\alpha_0 - g(\alpha_0) = \alpha_0 - \beta_0 \geq 0$ y $\alpha_1 - g(\alpha_1) \leq 0$. Dado que g es continua por el Teorema de valor intermedio, se tiene que $g(x) - x = 0$ para algún x entre α_0 y α_1. Por lo tanto, $g(x) = x$ para algún $x \in I$ □

Lema 2.0.2. *Sean $f : J \rightarrow J$ una función continua y $I_1, I_2, ...$ una sucesión de intervalos compactos con $I_n \subset J$ e $I_{n+1} \subset f(I_n)$ para todo $n \geq 0$. Entonces existe una sucesión de intervalos compactos $\{Q_n\}$ tal que $Q_{n+1} \subset Q_n \subset I_0$ y $f^n(Q_n) = I_n$ para todo $n \geq 0$. Además para cualquier $x \in Q = \cap Q_n$ se tiene $f^n(x) \in I_n$ para todo $n \geq 0$*

Demostración. La demostración la realizaremos por inducción.
Definamos $Q_0 = I_0$. Entonces $f^0(Q_0) = I_0$.
Primero veamos que se cumple para $k = 1$, sabemos por hipótesis que $I_0 \subset J$, e $I_1 \subset f(I_0)$, entonces por un lema anterior existe $Q_1 \subset I_0 = Q_0$ tal que $f(Q_1) = I_1$.
Supongamos ahora que el lema se cumpla para $k \leq n-1$, esto es, que existe Q_{n-1} tal que $f^{n-1}(Q_{n-1}) = I_{n-1}$ ($Q_{n-1} \subset Q_{n-2} \subset \ldots \subset I_0$). Veamos que se cumple para $k = n$. Como $f^{n-1}(Q_{n-1}) = I_{n-1}$ entonces por hipótesis sabemos que $I_n \subset f(I_{n-1}) = f^n(Q_{n-1})$. Aplicando un lema anterior a $g = f^n$ en Q_{n-1}, se tiene que existe $Q_n \subset Q_{n-1}$ tal que $f^n(Q_n) = I_n$ y

$$Q_n \subset Q_{n-1} \subset \ldots \subset Q_1 \subset Q_0 = I_0.$$

Ahora sea $x \in Q = \cap_{n=0}^{\infty} Q_n$, entonces $x \in Q_n$ para todo $n \geq 0$, y dado que $f^n(Q_n) = I_n$ para todo $n \geq 0$. Entonces $f^n(x) \in I_n$ para todo $n \geq 0$. □

Definición 2.4. *Diremos que un punto p sigue al lazo $J_0 \to \ldots \to J_{n-1} \to J_0$ si satisface el lema del itinerario. Note que si p sigue a un lazo de longitud n entonces su periodo debe ser divisor de n.*

Deseamos asegurar que el periodo del punto p encontrado en el lema del itinerario es n y no un divisor propio de n, tal como ocurre en el siguiente caso. Sea el 2-lazo $[-1, 0] \rightleftarrows [0, 1]$ de $f(x) = -2x$ el cual es seguido solo por el punto fijo 0, entonces para evitar esto añadimos un nuevo concepto.

Definición 2.5. *Se dice que un lazo de intervalos $J_0 \to \ldots \to J_{n-1} \to J_0$ es elemental si cada punto p que lo sigue tiene periodo n.*

Con la definición anterior, empleando el lema del itinerario obtenemos el siguiente resultado:
Si tenemos un lazo elemental $J_0 \to \cdots \to J_{n-1} \to J_0$ implica la existencia de un punto periodico p con periodo n que sigue al lazo.

Lema 2.0.3. *Un lazo $J_0 \to \ldots \to J_{n-1} \to J_0$ de $O-$intervalos es elemental si no es seguido por un punto de O y el interior de J_0 ($int(J_0)$) es disjunto de cada J_i con $1 \leq i \leq n-1$ ($Int(J_0) \cap J_i = \phi$; $1 \leq i \leq n-1$).*

Demostración. Supongamos que el lazo $J_0 \to ... \to J_{n-1} \to J_0$ no es seguido por puntos de O. Así, si $f^n(p) = p$ y p sigue al lazo entonces $p \in J_0$ y $p \notin O$, además como es un lazo de $O-$intervalos, significa que los extremos de J_0 son puntos de O. Esto significa que p no es ninguno de los extremos de J_0. Por tanto, $p \in int(J_0)$. Si $1 \leq i \leq n-1$ entonces $f^i(p) \notin int(J_0)$ (ya que $f^i(p) \in J_i$ para $i = 1, 2, ..., n-1$). Así, $f^i(p) \neq p$ para $1 \leq i \leq n$. Luego, p tiene periodo n. $\square$

Sea O un ciclo de f con longitud m. Demostraré que hay órbitas de periodo l para todo $l \lhd m$, para ello encontraré $l-$lazos elementales de $O-$intervalos para todo $l \lhd m$ y luego deduciremos la existencia de ciclos de esas longitudes. Si O es un ciclo no trivial; es decir, si $m \geq 2$. Sean $p = \max\{x \in O : f(x) > x\}$ y q el punto de O que está inmediatamente a la derecha de p. Entonces, sabemos que $f(p) > p$ y como q es el punto que está inmediatamente a la derecha de p entonces $f(p) \geq q \geq p$. Además $q \geq f(q)$ ya que caso contrario $q > f(q)$ lo cual contradeciría la maximalidad de p. Sabiendo que $f(q) \in O$, recordando además que p es el punto de la órbita que está inmediatamente a la derecha de q entonces $f(p) \geq q \geq p \geq f(q)$. De aquí tenemos que $f(p) \geq q$ y $p \geq f(q)$. Consideremos $I = [p, q]$, de donde se ve claramente que $I \to I$. Tomemos un punto fijo y arbitrario $c \in Int(I)$(por ejemplo podemos tomar el punto medio de I). Para $x \in O$ denotamos por O_x el conjunto de puntos de O en el intervalo cerrado acotado por x y c.

Definición 2.6. *Diremos que $x \in O$ cambia de lado si x y $f(x)$ están a distintos lados de c.*

Los extremos del intervalo I, p y q cambian de lado, ya que

$$f(p) \geq q > c = \frac{p+q}{2} > p \geq f(q).$$

Observación: Si todos los puntos de O cambian de lado, entonces f es una biyección entre O_L y O_R donde $L := \min O$ y $L := \max O$. Para demostrar estos veamos, primero recordemos que $c \notin O$ pues c está entre p y q y estos eran puntos de la órbita consecutivos. Ahora veamos
$O_L = \{z \in O : z \in [L, c]\}$ y $O_R = \{z \in O : z \in [c, R]\}$ como todos los puntos cambian de lado, y $c \notin O$ entonces $f(O_L) \subset O_R$. Ya que si sea z un punto de O_L y sabiendo que este cambia de lado entonces $f(z) \notin$

O_L de donde se concluye que $f(z) \in O_R$. Además sean $x, y \in O$ tal que $f(x) = f(y)$, recordando que O es una órbita periódica de período m entonces $f^{m-1}(f(x)) = f^{m-1}(f(y))$ de donde
$f^m(x) = f^m(y)$ y como $x, y \in O$, se tiene $f^m(x) = x = y = f^m(y)$, por lo tanto f es inyectiva. Sea $z \in O_R$ entonces en particular como $z \in O$ y los elementos de O tienen período m entonces

$$z = f^m(z) = f(f^{m-1}(z)) = f(w) \text{ donde } w = f^{m-1}(z).$$

Ahora si $w \in O_R$ habría una contradicción, ya que $w, f(w) = z \in O_R$ y esto no se puede dar, pues sabemos que todos los puntos cambian de lado. Por lo tanto $w \in O_L$, en resumen hemos tenemos que

$$\forall z \in O_R \;\; \exists w \in O_L \text{ tal que } z = f(w),$$

lo cual con lo anterior quiere decir que es sobreyectiva. Por lo tanto
$f : O_L \to O_R$ es biyectiva. Ahora bien sabemos que la cantidad de puntos de la órbita es la cantidad de puntos de la órbita que están en O_L unido con la cantidad de los puntos de la órbita que están en O_R, recordar que $c \notin O$, pero como f es una biyección entre O_L y O_R. En particular tienen la misma cantidad de elementos entonces la cantidad de elementos de O, que hemos llamado m, es el doble de la cantidad de puntos de la órbita que están en O_L en particular m es par.

Proposición 2.1. *Si un $m-$ciclo O con $m \geq 2$ contiene un punto que no cambia de lado, entonces hay un $l-$lazo elemental de $O-$intervalos para cada $l \lhd m$*

Demostración. Comencemos contruyendo puntos $x_0, x_1, ..., x_k$ de puntos de O que "salen en espiral"tan rápido como sea posible.
Seleccionemos x_0 y x_1 tales que sean los extremos de I (osea p y q) tal que $f(x_1) \neq x_0$. Esto es posible ya que supongamos sin pérdida de generalidad que $x_0 = p$ y $x_1 = q$, si $f(q) = p$ entonces $O = \{p, q\}$. Entonces todos los puntos cambiarían de lado, pues $f(p) \geq q > c = \frac{p+q}{2} > p \geq f(q)$ lo cual claramente es una contradicción.
Extendemos por inducción, de la siguiente manera:
Si $i \geq 1$ y todos los puntos de O_{x_i} cambian de lado, entonces x_{i+1} es el punto de $f(O_{x_i})$ que está mas lejos de c; en otro caso, x_{i+1} no está definida.

Nótese que este proceso acaba en algún momento, ya que hay al menos un punto de O que no cambia de lado. Además, se ve que los términos consecutivos de esta secuencia esta en lados opuestos con respecto a c, ya que supongamos que $x_i < c$ entonces $O_{x_i} = \{z \in O : z \in [x_i, c]\}$ y sabemos que todos los puntos de O_{x_i} cambian de lado eso significa que $\forall z \in O_{x_i} \ \ f(z) > c$. En particular $x_{i+1} > c$, ya que $x_{i+1} \in f(O_{x_i})$. Es decir $x_0, x_2, ..., x_{2i}, ...$ están a un lado de c, y $x_1, x_3, ..., x_{2i+1}, ...$ están al otro lado de c.
Para continuar, comencemos haciendo un lema.

Lema 2.0.4. *El punto x_{i+2} está más lejos de c que x_i si ambos están definidos.*

Demostración. Para $i = 0$ se ve ya que recordando que $x_0 = p, x_1 = q$, como sabemos

$$O_{x_1} = O_q = \{z \in O : z \in [c, q]\} = \{q\}.$$

Y como q cambia de lado entonces $x_2 = f(q)$. Además sabemos que entre c y $p = x_0$ no existe ningún punto de O. Como $x_2 = f(q)$ está en el mismo lado que $x_0 = p$ entonces $x_2 = f(q)$ tiene que ir detrás que p. Si $i \geq 1$ con x_{i+1} y x_{i+2} están definidos, todos los puntos de O_{x_i} y $O_{x_{i+1}}$ cambian de lado y tenemos la inclusión $f(O_{x_i}) \subset O_{x_{i+1}}$ y $f(O_{x_{i+1}}) \subset O_{x_{i+2}}$ se cumple. Ya que supongamos que $x_i < c$ esto significa que todos los elementos de O_{x_i} se ubican a la izquierda de c y sea $w \in f(O_{x_i})$. Como todos los elementos de O_{x_i} cambian de lado entonces todos los elementos de $f(O_{x_i})$ se ubican a la derecha de c. En particular w, además $w \leq x_{i+1}$ ya que $x_{i+1} = \max f(O_{x_i})$ esto significa que $w \in [c, x_{i+1}]$. De aquí $f(O_{x_i}) \subset x_{i+1}$, y además $f(O_{x_{i+1}}) \subset x_{i+2}$ de donde $f^2(O_{x_i}) \subset O_{x_{i+2}}$. Como f^2 es inyectiva en O, pues si
$f^2(x) = f^2(y),$ con $x, y \in O$ y recordando que O tiene periodo m entonces

$$x = f^m(x) = f^{m-2}(f^2(x)) = f^{m-2}(f^2(y)) = f^m(y) = y.$$

Se tiene que $O_{x_{i+2}}$ tiene tantos puntos como O_{x_i}, por lo tanto x_{i+2} está al menos tan lejos de c como x_i.
Por otro lado no podemos tener que $x_i = x_{i+2}$ porque sino entonces tendriamos que

$$f(O_{x_i} \cup O_{x_{i+1}}) = f(O_{x_i}) \cup f(O_{x_{i+1}}) \subset O_{x_{i+1}} \cup O_{x_{i+2}} = O_{x_i} \cup O_{x_{i+1}}.$$

Pero $O_{x_i} \cup O_{x_{i+1}} \neq O$ pues sabemos que todos los puntos de O_{x_i} y $O_{x_{i+1}}$ cambian de lado, pues x_{i+1} y x_{i+2} están definidos, pero O posee un punto que no cambia de lado. Por lo tanto no pueden ser iguales, $\square$

Corolario 2.0.1. *Los puntos $x_0, x_1, ...$ son distintos y este proceso termina para algun x_k con $k < m$.*

Ahora construimos una secuencia de $O-$intervalos para la cual tenemos relaciones de cubrimiento que producirán lazos de longitud específica. Serán lazos de longitud 1, longitud l para todo número para $l \leq k$ y longitud l para todo $l \geq k$ excepto posiblemente para m.
Emplearemos el lema para verificar que estos lazos son elementales.
Como $k < m$ $(m \geq k+1)$, este conjunto de longitudes incluye 1, todo número par $l < m$, y todo $l > m$, esto es, $l \lhd m$.
Una elección simple de intervalo que produce las relaciones de cubrimientos deseadas (pero no permite aplicar el lema) es la siguiente:
Para $1 \leq i < k$, sea J_i el $O-$intervalo más corto que contiene a O_{x_i} y a ambos extremos del intervalo I.
Del lema anterior, se sigue que como $O_{x_{k-1}} \supset O_{x_{k-3}} \supset O_{x_{k-5}} \cdots$ y $O_{x_{k-2}} \supset O_{x_{k-4}} \supset \ldots$ entonces $J_{k-1} \supset J_{k-3} \supset J_{k-5} \supset ...$ y $J_{k-2} \supset J_{k-4} \supset ...$
Sea J_k el $O-$intervalo mas corto que contiene a O_{x_k}. Con esta elección tenemos:

1. $J_1 \to J_1$,

2. $J_i \to J_{i+1}$ para $1 \leq i < k$,

3. $J_k \to J_{k-1}$.

Obtenemos 1, pues $J_1 = I \to I$, como vimos anteriormente,
pues $O_{x_1} = O_q = \{z \in O : z \in [c, q]\} = \{q\}$ como J_1 es el menor intervalo que contiene a $O_{x_1} = \{q\}$. Entonces $J_1 = I = [p, q]$, ya que como sabemos p, q son puntos de O adyacentes.
Obtenemos 2 pues primero, supongamos sin perdida de generalidad que
$x_i < c$, J_i contiene los puntos extremos de I y el punto $y_i \in O_{x_i}$ para el cual $f(y_i) = x_{i+1}$. Como f es continua y J_i un intervalo, entonces $f(J_i)$ es un intervalo. Además recordando que todos los puntos de O_{x_i} cambian de lado obtenemos que $[c, x_{i+1}] \subset f(J_i)$ de donde $O_{x_{i+1}} \subset f(J_i)$. Además como

$$p, q \in J_i \Rightarrow f(p), f(q) \in f(J_i) \Rightarrow [p, q] \subset [f(q), f(p)] \subset f(J_i),$$

de donde $p, q \in f(J_i)$, de aquí $f(J_i)$ contiene a $O_{x_{i+1}}$ y a p, q. Así $f(J_i) \supset J_{i+1}$. Obtenemos 3 pues, supongamos sin pérdida de generalidad que $x_k < c$, sabemos que $O_{x_k} \supset O_{x_{k-2}}$. Además J_k contiene un punto y_{k-2} que cambia de lado, y tal que $f(y_{k-2}) = x_{k-1}$. Pero J tiene un punto $z \in O$ que no cambia de lado, se sigue que $f(J_k)$ contiene a x_{k-1} y a $f(z)$ el cual está al otro lado de c desde x_{k-1}. Ya que como $z < c$ no cambia de lado entonces $f(z) < c$ y como $y_{k-2} \in O_{x_{k-2}} \subset O_{x_k} \subset J_k$, tenemos de aquí que $y_{k-2}, z \in J_k$. De aquí

$$O_{x_{k-1}} = [c, x_{k-1}] \subset [f(z), f(y_{k-2})] \subset f(J_k)$$

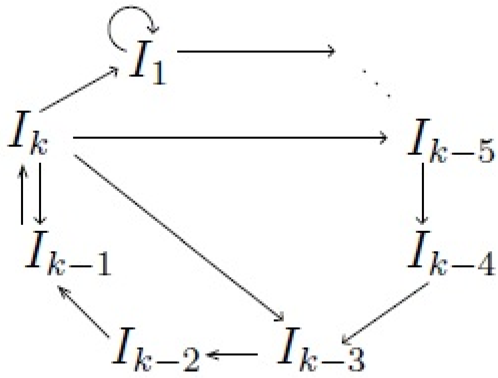
ya que f es continua y J_k es un intervalo. Además recordando que
$f(z) \in O \Rightarrow f(z) \leq p$ entonces $p, q \in f(J_k)$. Por lo tanto $J_k \to J_{k-1}$. $\square$

Ahora restringimos los intervalos J_i a intervalos I_i que tienen todas las propiedades deseadas. Para $1 \leq i < k$ sea I_i el $O-$intervalo más corto que contiene a y_i y a ambos extremos del intervalo I. Sea I_k el $O-$intervalo acotado por y_{k-2} y el punto z señalado anteriormente. De la definición se tiene que $I_i \subset J_i$ para $1 \leq i < k$ y $I_1 = J_1 = I$.
Con la misma idea que hicimos para los J_i podemos demostrar algo similar para los I_i, tendríamos que
4. $I_1 \to I_1$
5. $I_1 \to I_2 \to \cdots \to I_k \to I_1$
6. $I_k \to I_{k-1}, I_{k-3}, \ldots$
Ahora demostraremos un lema

Lema 2.0.5. $I_i \cap int(I_k) = \phi$ *para* $1 \leq i < k$.

Demostración. El punto z esta mas lejos de c que x_{k-2} pues el punto z no cambia de lado y sabemos que todos los puntos de O que están entre c y x_{k-2} si cambian de lado. Por lo tanto $Int(I_k)$ está al lado opuesto de x_{k-2} respecto a c. Por otro lado $J_{k-2} \cup J_{k-1}$ estan en el mismo lado que x_{k-2} con respecto a c e $I_i \subset J_{k-2} \cup J_{k-1}$ para $1 \leq i < k (J_{k-1} \supset J_{k-3} \supset \ldots, J_{k-2} \supset J_{k-4} \supset J_{k-6} \ldots)$. $\square$

Las relaciones $4 - 6$ serán graficadas a continuación:



Del cual obtenemos los siguientes lazos:

7. $I_1 \to I_1$
8. $I_k \to I_{k-(l-1)} \to I_{k-(l-2)} \to \cdots \to I_{k-2} \to I_{k-1} \to I_k$ para $l \leq k$ par
9. $I_k \to I_1 \to I_1 \cdots \to I_1 \to I_2 \to \cdots \to I_{k-1} \to I_k$ con j apariciones de I_1

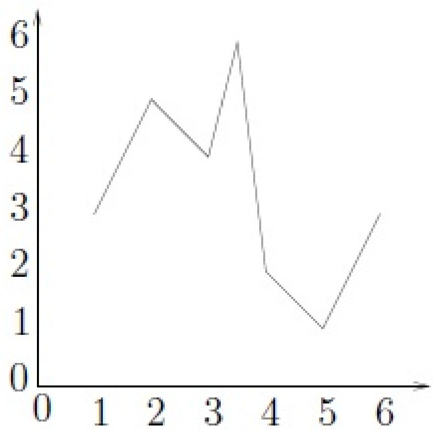
El lazo de 7 es elemental pues tiene longitud 1.
Los lemas anteriores nos dicen que los lazos en 8 y 9 son elementales una vez hayamos probado que no pueden ser seguidos por un punto de O.
Este es el caso para los lazos en 8 porque tienen longitud $l \leq k < m$.
Los lazos en 9 son seguidos por un punto de O en los siguientes casos:

- $j = 1$, ya que este ciclo tiene longitud $k < m$,
- $j = 2$ y $k < m - 1$, ya que este lazo tiene longitud $k + 1 < m$,
- $j > 2$, ya que estos lazos tienen al menos tres repeticiones de I_1.

El caso excepcional es $j = 2$ y $k = m - 1$ en el que el lazo de 9 tiene longitud m y no necesariamente produce un punto periódico. El lazo en 7 tiene longitud 1. Los lazos en 8 tienen todas las longitudes pares hasta k(menores o iguales a k). Los lazos elementales en 9 tienen todas longitudes $l \geq k$, excepto posiblemente m.
Note que si $l \lhd m$ entonces ó $l = 1, l > m$ ó $l < m$ y l es par. Luego, hay lazos elementales de longitud l para todo $l \lhd m$.

Definición 2.7. *Una relación de cubrimiento de $I \to J$ de $O-$intervalos se dice que es $O-$forzada si J esta contenida en el intervalo cuyos puntos extremos son los puntos mas a la izquierda y mas a la derecha de $f(I \cap O)$; es decir, $J \subset [\min f(I \cap O), \max f(I \cap O)]$. Un lazo de $O-$intervalos se dice que es $O-$forzado si cada relación de cubrimiento es $O-$forzada.*

Ejemplo: Consideremos la función $f : [1,6] \to [1,6]$ cuyo gráfico es:



Entonces f es continua, y viendo el grafico tenemos que $f(1) = 3; f(2) = 5; f(3) = 4; f(4) = 2; f(5) = 1; f(6) = 3$. De aquí vemos que $f^5(1) = f^4(3) = f^3(4) = f^2(2) = f(5) = 1$. De donde decimos que 1 tiene periodo 5 y $O = \{1,3,4,2,5\}$. Además:

- $[1,2] \to [3,4]$, es una relación de curbimiento $O-$forzada pues
$$[3,4] \subset [3,5] = [\min f([1,2] \cap O), \max f([1,2] \cap O)], \text{ y}$$
- $[3,4] \not\to [2,6]$ pero esta relación de cubrimiento no es $O-$forzada pues
$$[2,6] \not\subseteq [2,4] = [\min f([3,4] \cap O), \max f([3,4] \cap O)].$$

La siguiente proposición lo haremos mediante un argumento inductivo, en el cual es importante notar que todos los lazos obtenidos en la anterior proposición son $O-$forzados. Es más cualquier relación de cubrimiento derivada sobre la dinámica de O será $O-$forzada.

Ahora para terminar nuestra prueba comenzaremos haciendo una proposición:

Proposición 2.2. *Un $m-$ciclo O tiene un $l-$lazo elemental $O-$forzado de $O-$intervalos para cada $l \lhd m$.*

Demostración. Usaremos inducción sobre m. El enunciado es cierto para $m = 1$ pues no existe l tal que $l \lhd 1$.
Supongamos que la proposición se cumple para todos los ciclos de longitud menor que m.
Sea O un $m-$ciclo. Si hay un punto que no cambia de lado, entonces la conclusión de la proposición se sigue de la proposición anterior y de nuestra observación que la proposición produce lazos $O-$forzados.
Por otro lado, si todos los puntos cambian de lado, por la observación echa anteriormente, se tiene que m es par y f es una biyección de O_L y O_R donde $L := \min O$ y $R = \max O$.
Para la segunda iterada, f^2, O_L y O_R son ciclos de longitud $\frac{m}{2}$, y por la hipótesis inductiva podemos aplicar la proposición a cualquiera de estos, en particular a O_R. Por lo tanto f^2 tiene un $k-$lazo elemental O_R-forzado de O_R-intervalos para cada $k \lhd \frac{m}{2}$.
Resta deducir a partir de esto que f tiene un $2k-$lazo elemental $O-$forzado de $O-$intervalos para cada $k \lhd \frac{m}{2}$ además de un $1-$lazo elemental.
En consecuencia la siguiente proposición concluye la inducción:

Proposición 2.3. *Sea O un ciclo de f tal que todos sus puntos cambian de lado, y supongamos que el ciclo O_R de f^2 da origen a un $k-$lazo elemental O_R-intervalos para f^2. Entonces hay un $2k-$lazo elemental $O-$forzado de $O-$intervalos para f. Además, hay un $1-$lazo elemental $O-$forzado para f.*

Demostración. El $1-$lazo se obtiene del $O-$intervalo que está en el medio, que está comprendido entre el punto de O_L que esta mas a la derecha ($\max O_L = p$) y el punto de O_R mas a la izquierda ($\min O_R = q$).
Como $f(\max O_L) \geq \min O_R$ y $f(\min O_R) \leq \max O_L$, se tiene dado que $I = [\max O_L, \min O_R] = [p, q]$ se cumple que $I \leftrightarrows I$.
Para un $k-$lazo elemental

$$J_0 \to J_1 \to J_2 \cdots \to J_{k-1} \to J_k \to J_0$$

de O_R-intervalos para f^2.
Sea $J_i^{'}$ el intervalo mas corto que contiene a $f(J_i \cap O) \subset O_L$. Los intervalos $J_1^{'}, \ldots, J_{k-1}^{'}$ estan a la izquierda de c, donde c es el punto medio del intervalo $I = [p, q]$ con $p = \max\{x \in O : f(x) > x\}$ y q el punto de O que esta inmediatamente a la derecha de p.
Como la relación de cubrimiento $J_i \to J_{i+1}$ es O_R-forzada para f^2 se tiene que $J_{i+1} \subset [\min f^2(J_i \cap O), \max f^2(J_i \cap O)]$.
Dado que $J_i^{'} = [\min f(J_i \cap O), \max f(J_i \cap O)]$, entonces

$$\min f^2(J_i \cap O), \max f^2(J_i \cap O) \in f(J_i^{'}).$$

Por tanto, si $y \in [\min f^2(J_i \cap O), \max f^2(J_i \cap O)]$, existe $x \in J_i^{'}$ tal que $f(x) = y$. Luego $J_{i+1} \subset [\min f^2(J_i \cap O), \max f^2(J_i \cap O)] \subset f(J_i^{'})$. En consecuencia $J_i^{'} \to J_{i+1}$ para f, y esta relación de cubrimiento es $O-$forzada. Así, obtenemos un $2k-$lazo $O-$forzado para f:

$$J_0 \to J_0^{'} \to J_1 \to J_1^{'} \to \cdots \to J_{k-1} \to J_{k-1}^{'} \to J_0 \qquad (*)$$

Un punto periódico p para f que sigue el lazo anterior es un punto periódico para f^2 que sigue el lazo elemental

$$J_0 \to J_1 \to J_2 \cdots \to J_{k-1} \to J_k \to J_0$$

y por tanto tiene periodo k con respecto a f^2.
Como los intervalos en $(*)$ están alternativamente a la derecha y a la izquierda del centro, también lo están los iterados de p bajo f. Por lo tanto, p tiene período $2k$ con respecto a f, y el lazo en $(*)$ es elemental. $\square$

Por tanto hemos dividido la prueba en dos grandes casos, primero si existe al menos un punto que cambia de lado, y luego cuando ningún punto cambia de lado. En ambos casos concluimos que existe un lazo elemental $O-$forzado, lo cual termina la prueba. $\square$

Capítulo 3

Teorema de recurrencia de Poincaré y de Birkhoff

3.1. Nociones preliminares

Para este capítulo, usaré de referencia inicial el libro [Ash72] Primero vamos a recordar algunas nociones fundamentales de la teoría de la medida. Sea X un conjunto. Un $\sigma-$algebra de subconjuntos de X es la colección Σ de subconjuntos de X que satisface:

1. $X \in \Sigma$
2. $B \in \Sigma \Rightarrow B^c \in \Sigma$
3. $B_n \in \Sigma$; n $\geq 1 \Rightarrow \bigcup_{n=1}^{+\infty} B_n \in \Sigma$.

Nosotros entonces llamaremos (X, Σ) un espacio medible. Un espacio de medida es la tripleta (X, Σ, m), donde X es un conjunto, Σ es un $\sigma-$álgebra de subconjuntos de X, y m es una función $m : \Sigma \to \mathbb{R}_0^+$ que satisface

$$m(\bigcup_{n=1}^{+\infty} B_n) = \sum_{n=1}^{+\infty} m(B_n).$$

donde B_n es una sucesión de elementos disjuntos dos a dos de Σ. Nosotros llamaremos (X, Σ, m) un espacio de probabilidad o espacio de medida nor-

malizado, si $m(X) = 1$. Usualmente trabajaremos en estos espacios.
La colección Υ de subconjuntos de X es un álgebra si:

1. $X \in \Upsilon$
2. $A \in \Upsilon \Rightarrow A^c \in \Upsilon$
3. $A_1, ..., A_n \in \Upsilon \Rightarrow \bigcup_{i=1}^{n} A_i \in \Upsilon$.

Una clase monótona de subconjuntos de X es la colección C de subconjuntos de X tal que si

$$E_1 \subset E_2 \subset \dots \text{ donde } E_i \in C, \quad \forall i \geq 1 \text{ entonces } \bigcup_{i=1}^{+\infty} E_i \in C$$

y si

$$F_1 \supset F_2 \dots \text{ donde } F_j \in C, \quad \forall j \geq 1 \text{ entonces } \bigcap_{j=1}^{+\infty} F_j \in C.$$

Un subconjunto Borel de $\mathbb{R}$ es un elemento del $\sigma-$álgebra generada por los conjuntos abiertos.
Sea (X, Σ, m) un espacio de medida, $f : X \to \mathbb{R}$ es medible si

$$\forall c \in \mathbb{R},\ f^{-1}(< c, +\infty >) \in \Sigma.$$

Una función simple es una función de la forma $\sum_{i=1}^{n} a_i \chi_{A_i}$, donde $a_i \in \mathbb{R}$, además A_i son elementos disjuntos de Σ, y χ_{A_i} denota la función característica de A_i. Las funciones simples son medibles. Nosotros definiremos la integral sobre funciones simples como:

$$\int (\sum_{i=1}^{n} a_i \chi_{A_i}) dm = \sum_{i=1}^{n} a_i m(A_i).$$

Suponga que f es medible y que $f \geq 0$; entonces existe una secuencia de funciones simples $f_n \uparrow f$, por ejemplo nosotros tomaremos

$$f_n(x) = \begin{cases} \frac{i-1}{2^n}, & \text{si } \frac{i-1}{2^n} \leq x < \frac{i}{2^n} \ i = 1, 2, ..., n2^n \\ \\ n, & \text{si } f(x) \geq n. \end{cases}$$

Se define $\int f dm = \lim\limits_{n\to+\infty} \int f_n dm$, notemos que la definición es independiente de la secuencia f_n que se escoja.
Suponga que $f : X \to \mathbb{R}$ es medible; entonces $f = f^+ - f^-$ donde

$$f^+(x) = \text{máx}\{f(x), 0\} \geq 0$$

$$f^-(x) = \text{máx}\{-f(x), 0\} \geq 0.$$

Se dice que f es integrable si $\int f^+ dm,\ \int f^- dm\ < +\infty$, y se define

$$\int f dm = \int f^+ dm - \int f^- dm.$$

Teorema 3.1.1. *Teorema de la convergencia monótona de Lebesgue*
Suponga que f_n una sucesión de funciones medibles, $f_n \to f$ c.t.p. y existe una función integrable g tal que $| f_n(x) | \leq g(x)$ en c.t.p para todo n, entonces f y cada f_n son integrables y

$$\int f_n\, dm \to \int f dm.$$

3.2. Transformaciones que preservan medida

3.2.1. Ejemplos

Suponga que (X, Σ_1, m_1) y (X, Σ_2, m_2) son espacios de probabilidad.

Definición 3.1.
1. *$T : X_1 \to X_2$ es medible si $T^{-1}(\Sigma_2) \subset \Sigma_1$ (es decir, $B_2 \in \Sigma_2 \Rightarrow T^{-1}(B_2) \in \Sigma_1$)*

2. *$T : X_1 \to X_2$ preserva medida si T es medible y*

$$m_1(T^{-1}(B_2)) = m_2(B_2)\ \forall B_2 \in \Sigma_2.$$

3. *Diremos que $T : X_1 \to X_2$ es una transformación que preserva medida e invertible, si T preserva medida, es biyectiva y T^{-1} también preserva medida.*

Teorema 3.2.1. *Suponga que (X_1, Σ_1, m_1), (X_2, Σ_2, m_2) dos espacios de probabilidad y $T : X_1 \to X_2$ un mapeo. Sea A_2 un álgebra generado por Σ_2. Si $a_2 \in A_2 \Rightarrow T^{-1}(a_2) \in B_1$ entonces T preserva medida.*

3.3. Teorema de recurrencia de Poincaré

Teorema 3.3.1. *Sea T una transformacion que preserva medida del espacio de probabilidad (X, Σ, m). Sea $E \in \Sigma, m(E) > 0$. entonces casi todos los puntos de E retornan infinitamente a E para una iteración positiva de T.*

Demostración. Para $N \geq 0$ sea $E_N = \bigcup_{n=N}^{+\infty} T^{-n}(E)$. Se tiene que

$$T^{-1}(E_N) = T^{-1}(\bigcup_{n=N}^{+\infty} T^{-n}(E)) = \bigcup_{n=N}^{+\infty} T^{-n-1}(E) = \bigcup_{n=N+1}^{+\infty} T^{-n}(E) = E_{N+1}.$$

Y además para todo $n \geq 1$ se cumple que

$$E_{N-1} = \bigcup_{n=N-1}^{+\infty} T^{-n}(E) = T^{-(N-1)} \cup \bigcup_{n=N}^{+\infty} T^{-n}(E) = T^{-(N-1)}(E) \cup E_N.$$

Entonces $E_N \subset E_{N-1}$, en general aplicando repetidamente este resultado obtenemos que:

$$E_N \subset E_{N-1} \subset E_{N-2} \subset \cdots \subset E_0.$$

También que tenemos que $E \subset E_0$ pues $T^0(E) = E$, y

$$m(E_{N+1}) = m(T^{-1}(E_N)) = m(E_N), \,\forall N \geq 0,$$

la última se da gracias a que T preserva medida. De aquí para $N = 1$ se deduce que $m(E_2) = m(E_1)$; y para $N = 0$ tenemos que $m(E_1) = m(E_0)$ de donde $m(E_2) = m(E_0)$. Así aplicando repetidamente este proceso tenemos que $m(E_N) = m(E_0)$ $\forall N \geq 0$, además sea $F_N = E_N^c$. Como es un espacio de probabilidad y $1 = m(X) = m(E_N \cup E_N^c) = m(E_N \cup F_N) = m(E_N) + m(F_N)$ de donde $m(F_N) = 1 - m(E_N)$ y como $E_{N+1} \subset E_N$ entonces $F_N \subset F_{N+1}$. Además

$$m(F_N) = 1 - m(E_N) = 1 - m(E_0) = m(F_0)$$

entonces por propiedad de medida

$$m(\bigcup_{N=0}^{+\infty} F_N) = \lim_{N \to +\infty} m(F_N) = \lim_{N \to +\infty} m(F_0) = m(F_0).$$

De donde

$$m(E_0) = 1 - m(F_0) = 1 - m(\bigcup_{N=0}^{+\infty} F_N)$$

$$= m((\bigcup_{N=0}^{+\infty} F_N)^c) = m(\bigcap_{N=0}^{+\infty} F_N^c) = m(\bigcap_{N=0}^{+\infty} E_N).$$

De esto tenemos que

$$m(\bigcap_{N=0}^{+\infty} E_N) = m(E_0).$$

De aqui

$$\bigcap_{N=0}^{+\infty} E_N = \bigcap_{N=0}^{+\infty} \bigcup_{n=N}^{+\infty} T^{-n}(E)$$

que es el conjunto de todos los puntos que entran infinitamente a menudo bajo una iteración positiva de T.

Ahora llamo $F = E \cap (\bigcap_{N=0}^{+\infty} E_N)$ que es el conjunto de puntos de E que entran a E mediante una iteración positiva de T. Como $\bigcap_{N=0}^{+\infty} E_N \subset E_0$ y ambos conjuntos tienen la misma medida,

$$m(F) = m(E \cap \bigcap_{N=0}^{+\infty} E_N) = m(E \cap E_0) = m(E).$$

Falta demostrar que un punto de F vuelve infinitamente a F mediante T.
Sea $x \in F$ entonces $\exists\ 0 < n_1 < n_2 ... \ni T^{n_i}(x) \in E$, $\forall i$.
Consideremos $T^{n_1}(x) \in E$ y sigue en E infinitamente mediante iteraciones positivas de T, llamemosla $n_2 - n_1, n_3 - n_1, ...$ donde
$T^{n_i - n_1}(T^{n_1}(x)) = T^{n_i}(x) \in E$ esto es $T^{n_1} \in F$. Similarmente se muestra que $T(x)^{n_i} \in F$, $\forall i$. $\square$

3.4. Preliminares para el teorema de Birkhoff

Definición 3.2. $T : (X, \Sigma, m) \to (X, \Sigma, m)$ *es ergódica respecto a* m *si para todo* $B \in \Sigma$, *y que* $T^{-1}(B) = B \Rightarrow m(B) = 0$ *o* $m(B) = 1$.

Teorema 3.4.1. *Las siguientes proposiciones son equivalentes para* $T : X \to X$ *que preserva medida:*
(a) T *es ergodica,*

(b) $m(T^{-1}(B)\triangle B)$, $B \in \Sigma \Rightarrow m(B) = 0$ *o* $m(B) = 1$,
(c) *Para todo* $A, B \in \Sigma, m(A), m(B) > 0 \;\; \exists n > 0; \;\; m(T^{-n}(A) \cap B) > 0$.

Demostración. $(a) \Rightarrow (b)$ Suponga que $m(T^{-1}(B)\triangle B) = 0$. Sea
$B_\infty = \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}(B) \in \Sigma$. Entonces

$$B_\infty = [(B \cup T^{-1}(B) \cup T^{-2}(B) \ldots) \cap (T^{-1}(B) \cup T^{-2}(B) \cup \ldots)] \cap \bigcap_{n=2}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}(B)$$

Nos damos cuenta que uno de los dos elementos encerrados en corchete está incluido en el otro. Por lo tanto

$$B_\infty = (T^{-1}(B) \cup T^{-2}(B) \cup \ldots) \cap (T^{-2}(B) \cup T^{-3}(B) \cup \ldots) \cap \ldots$$

$$= \bigcap_{n=0}^{\infty} \bigcup_{i=n+1}^{\infty} T^{-i}(B) = \bigcap_{n=0}^{\infty} \bigcup_{j=n}^{\infty} T^{-j-1}(B) = T^{-1}(\bigcap_{n=0}^{\infty} \bigcup_{j=n}^{\infty} T^{-j}(B)) = T^{-1}(B_\infty).$$

Llamemos ahora $A_n = \bigcup_{i=n}^{\infty} T^{-i}(B)$ nosotros probamos en el teorema anterior que

$$m(B_\infty) = m(\bigcap_{n=0}^{\infty} A_n) = m(A_0) = m(B \cup T^{-1}(B) \cup T^{-2}(B) \cup \ldots).$$

Pero sea $Z_n = \bigcup_{j=0}^{n} T^{-j}(B)$, como $Z_n \subset Z_{n+1}$.
Primero como $m(T^{-1}(B)\triangle B) = 0 \Rightarrow m((T^{-1}(B) \setminus B) = m(B \setminus T^{-1}(B)) = 0$. Además demostraré por inducción que $m(T^{-n}(B)\triangle B) = 0$, primero para $n = 1$ cumple por dato, ahora supondré que cumple para n y demostraré que cumple para $n + 1$

$$m(T^{-n-1}(B)\triangle B) \leq m(T^{-n-1}(B)\triangle T^{-n}(B) + m(T^{-n}(B)\triangle B)$$

$$= m(T^{-1}(B)\triangle B) + m(T^{-n}(B)\triangle B) = 0 + 0 = 0.$$

Donde la primera igualdad se da ya que T preserva medida y la segunda por hipótesis inductiva. Ahora bien

$$m(Z_n) = m(\bigcup_{j=0}^{n} T^{-j}(B)) \leq m(B) + m(T^{-1}(B)\triangle B) + \cdots + m(T^{-n}(B)\triangle B)$$

$$m(Z_n) \leq m(B) + 0 + 0 + \cdots + 0 = 0 \Rightarrow \lim_{n\to+\infty} m(Z_n) \leq m(B)$$

y como $B \subset Z_n$ $\forall n$ entonces $m(B) \leq \lim_{n\to+\infty} m(Z_n)$ de donde $m(B) = \lim_{n\to+\infty} m(Z_n)$. De aquí que

$$m(B) = \lim_{n\to+\infty} m(Z_n) = m(B_\infty)$$

entonces $m(B_\infty) = 0$ o $m(B_\infty) = 1$, por lo tanto $m(B) = 0$ o $m(B) = 1$
(2) $\Rightarrow$ (3) Sea $m(A), m(B) > 0$ y supongamos que (3) es falso, es decir

$$\forall\ n > 0 \quad m(T^{-n}(A) \cap B) = 0.$$

Entonces $m((\bigcup_{n=1}^{\infty} T^{-n}(A)) \cap B) = 0$. Llamemos $A' = \bigcup_{n=1}^{\infty} T^{-n}(A) \in \Sigma$ ya que T es medible. Entonces

$$T^{-1}(A') = T^{-1}(\bigcup_{n=1}^{\infty} T^{-n}(A)) = \bigcup_{n=1}^{\infty} T^{-n-1}(A) = \bigcup_{n=2}^{\infty} T^{-n}(A).$$

De donde se obtiene que $T^{-1}(A') \subset A'$ y $m(A') = m(T^{-1}(A'))$ ya que T preserva medida, entonces

$$m(T^{-1}(A')\triangle A') = m(A' \setminus T^{-1}(A')) = m(A') - m(T^{-1}(A')) = 0.$$

Luego por (2) se tiene que $m(A') = 0$ o $m(A') = 1$. Pero $T^{-1}(A) \subset A'$ entonces $0 < m(A) = m(T^{-1}(A)) \leq m(A')$ entonces $m(A') = 1$. De aquí $m(A' \cup B) = 1$ ya que estamos en un espacio de probabilidad. Entonces

$$m(A' \cap B) = m(A') + m(B) - m(A \cup B) = 1 + m(B) - 1 = m(B) > 0$$

lo cual es una contradicción ya que antes habíamos obtenido que $m(A' \cap B) = m(\bigcup_{n=1}^{\infty} T^{-n}(A) \cap B) = 0$.
(3) $\Rightarrow$ (1) Supongamos que (1) es falso, es decir, $\exists\ B \in \Sigma, T^{-1}(B) = B$ y $0 < m(B) < 1$. Entonces por inducción probaré que $T^{-n}(B) = B$, $\forall n$. Para $n = 1$ se cumple la condición. Supongamos que se cumpla para n, ahora demostraré que se cumple para $n+1$;

$$T^{-(n+1)}(B) = T^{-n-1}(B) = T^{-1}(T^{-n}(B)) = T^{-1}(B) = B.$$

Entonces lo hemos demostrado.
Ahora $T^{-n}(B) \cap (X \setminus B) = B \cap (X \setminus B) = \phi$, $\forall n$ entonces $m(T^{-n}(B) \cap (X \setminus B)) = 0$, $\forall n$ lo cual contradice (3). $\square$

Sea $T : (X, \Sigma, m) \to (X, \Sigma, m)$ una transformacion que preserva medida. Definamos el operador U_T de funciones de valores complejos en X por

$$U_T(f(x)) = f(T(x)).$$

Se tiene $U_T(L^p(m)) \subset L^p(m)$.

Teorema 3.4.2. *Sea $U : L^1_R(m) \to L^1_R$ un operador lineal positivo con norma menor o igual a 1. Sea $N > 0$ un entero, defino*

$$f_0 = 0, f_n = f + U(f) + U^2(f) + ... + U^{n-1}(f),$$

y $F_N = \max\limits_{0\leq n\leq N} f_n \geq 0$
Entonces

$$\int_{\{x:F_N(x)>0\}} f \; dm \geq 0.$$

Demostración. Claramente F_N es una función integrable (ya que es el máximo de funciones integrables) y es mayor o igual a 0. Entonces $F_N \in L^1_R(m)$. Para $0 \leq n \leq N$ se cumple que $f_n \leq F_N$ entonces $F_N - f_n \geq 0$ como U es un operador lineal positivo entonces $U(F_N - f_n) \geq 0$ y como U es lineal $U(F_N) \geq U(f_n)$. Entonces

$$U(F_N) \geq U(f) + U^2(f) + ... + U^n(f) = f_{n+1} - f,$$

de donde $U(F_N) + f \geq f_{n+1}$. Por lo tanto

$$U(F_N(x)) + f(x) \geq \max_{0\leq n\leq N} f_{n+1}(x) = \max_{1\leq n\leq N+1} f_n(x) \geq \max_{1\leq n\leq N} f_n(x)$$

cuando $F_N(x) > 0$ esto implica que existe $n_0 \in \{0, 1, ..., N\}$ tal que $f_{n_0}(x) > 0 = f_0(x)$ entonces $\max\limits_{1\leq n\leq N} f_n(x) = \max\limits_{0\leq n\leq N} f_n(x)$. Entonces

$$U(F_N(x)) + f(x) \geq \max_{0\leq n\leq N} f_n(x) = F_N(x) \text{ cuando } F_N(x) > 0.$$

Esto es $f \geq F_N - U(F_N)$ en $A = \{x : F_N(x) > 0\}$, esto es

$$\int_A f dm \geq \int_A F_N dm - \int_A U(F_N) dm.$$

Sabemos que como $F_N \geq 0$ y $F_N > 0$ en A entonces $F_N = 0$ en $X \setminus A$ de aqui que

$$\int_A f dm \geq \int_A F_N dm + 0 - \int_A U(F_N) dm$$

$$= \int_A F_N dm + \int_{X\setminus A} F_N dm - \int_A U(F_N) = \int_X F_N dm - \int_A U(F_N) dm.$$

Y como $F_N \geq 0$ y U es operador lineal positivo entonces $U(F_N) \geq 0$ de aquí

$$\int_X U(F_N) dm \geq \int_A U(F_N).$$

Colocando esto en la desigualdad anterior tenemos que

$$\int_A f dm \geq \int_X F_N dm - \int_X U(F_N) dm \geq \int_X \| F_N \| dm - \int_X \| U(F_N) \| dm$$

$$\geq \int_X \| F_N \| dm - \int_X \| U \| \| F_N \| dm \geq \int_X \| F_N \| (1 - \| U \|) dm.$$

Como $\| U \| \leq 1 \Rightarrow 1 - \| U \| \geq 0$ por lo tanto tenemos que:

$$\int_A f dm \geq \int_X \| F_N \| (1 - \| U \|) dm \geq 0.$$

□

Corolario 3.4.1. *Sea $T : X \to X$ una transformación que preserva medida. Si $g \in L^1_R(m)$ y*

$$B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} g(T^m(x)) > \alpha\}.$$

entonces

$$\int_{B_\alpha \cap A} g dm \geq \alpha m(B_\alpha \cap A)$$

si $T^{-1}(A) = A$ y $m(A) < \infty$.

Demostración. Primero probaremos el resultado bajo las siguientes hipótesis $m(X) < \infty$ y $A = X$. Sea $f = g - \alpha$ entonces

$$B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} g(T^m(x)) > \alpha\}$$

$$B_\alpha = \{x \in X : (\sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} g(T^m(x))) - \alpha > 0\}$$

$$B_\alpha = \{x \in X : (\sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} g(T^m(x))) - \alpha \times \frac{n}{n} > 0\}$$

$$B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} (g(T^m(x)) - \alpha) > 0\}.$$

Entonces por definición de f tenemos que $g(T^m) - \alpha = f(T^m)$

$$B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n} \sum_{m=0}^{n-1} f(T^m(x)) > 0\}.$$

Ahora recordando que $U(f) = f(T)$ entonces

$f_n = f + U(f) + U^2(f) + ... + U^{n-1}(f) = f + f(T) + f(T^2) + ... + f(T^{n-1})$.

Reescribiendo nuestra anterior desigualdad tendremos

$$B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n} f_n(x) > 0\}.$$

Sea $x \in B_\alpha$ entonces $\sup\limits_{n\geq 1} \frac{1}{n} f_n(x) > 0$ por lo tanto

$$\exists\ n_0 \geq 1 \text{ tal que } \frac{1}{n_0} f_{n_0}(x) > 0.$$

De donde $F_{n_0}(x) \geq f_{n_0}(x) > 0$ de donde obtenemos que

$$x \in \bigcup_{N=0}^{\infty} \{x : F_N(x) > 0\}.$$

Ahora sea $x \in \bigcup\limits_{N=0}^{\infty} \{x : F_N(x) > 0\}$ entonces

$$\exists\ n_0 \geq 0 \text{ tal que } f_{n_0}(x) > 0.$$

Primero $n_0 \neq 0$ pues si $n_0 = 0$ entonces $f_0(x) = 0 > 0$, lo cual es una contradicción, por lo tanto $n_0 \geq 1$, ahora bien entonces

$$\sup_{n\geq 1} \frac{1}{n} f_n(x) \geq f_{n_0}(x) > 0.$$

De donde concluimos que $x \in B_\alpha$, entonces

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

Entonces por el teorema anterior

$$\int_{B_\alpha} f dm \geq 0 \Rightarrow \int_{B_\alpha} (g-\alpha) dm \geq 0$$

$$\int_{B_\alpha} g dm \geq \int_{B_\alpha} \alpha dm = \alpha m(B_\alpha).$$

Para hacer el caso general solo basta, usamos simplemente $T \mid_A$ en lugar de T, el resto de la prueba se sigue de modo similar. Con lo cual demostraremos que

$$\int_{A\cap B_\alpha} g dm \geq \alpha m(A \cap B_\alpha).$$

□

3.5. Teorema de Birkhoff

Teorema 3.5.1. *Suponga $T : (X, \Sigma, m) \to (X, \Sigma, m)$ es un operador que preserva medida y $f \in L^1(m)$. Entonces $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ converge a casi todo punto a $g \in L^1(m)$. Más aún $foT = T$ en casi todo punto y si $m(X) < \infty$, $\int g \ dm = \int f \ dm$.*

Demostración. Nuestra referencia para este teorema es el libro [Ash72]. Es suficiente probar el teorema para $f \in L^1_R(m)$.

Sea $f^*(x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$ y $f_*(x) = \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$. Se tiene que $f^* oT = f^*$ y $f_* oT = f_*$ ya que sea $a_n(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$. Entonces $a_n(T(x)) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(T(x)) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^{i+1}(x)) = \frac{1}{n}\sum_{i=1}^{n} f(T^i(x))$. Y además $\frac{n+1}{n} a_{n+1} = \frac{n+1}{n} \times \frac{1}{n+1} \sum_{i=0}^{n+1-1} f(T^i(x)) = \frac{1}{n}\sum_{i=0}^{n} f(T^i(T(x))$. De donde

$\frac{n+1}{n}a_{n+1} - a_n(T(x)) = \frac{f(x)}{n}$. Tomando límite superior a esta última igualdad (se hace de forma análoga con el límite inferior) $\limsup_{n\to\infty} a_{n+1}(x) - a_n(T(x)) = 0$. De donde $\limsup_{n\to\infty} a_{n+1}(x) = \limsup_{n\to\infty} a_n T(x)$ pero como $\limsup_{n\to\infty} a_n(x) = \limsup_{n\to\infty} a_{n+1}(x)$. Entonces tenemos que

$$\limsup_{n\to\infty} a_n(x) = \limsup_{n\to\infty} a_n(T(x)) \text{ de donde } f^* = f^* oT.$$

Para los números reales $\beta < \alpha$. Sea

$$E_{\alpha,\beta} = \{x \in X : f_*(x) < \beta, \alpha < f^*(x)\}.$$

Entonces $T^{-1}(E_{\alpha,\beta}) = E_{\alpha,\beta}$ y

$$E_{\alpha,\beta} \cap \{x \in X : \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) > \alpha\} = E_{\alpha,\beta}.$$

Primero probaré que $m(E_{\alpha,\beta}) < \infty$ para aplicar el anterior corolario.
Suponga que $\alpha > 0$. Sea $C \subset E_{\alpha,\beta}$ con $m(C) < \infty$. Entonces
$h = f - \alpha\chi_C$ es integrable y por el teorema maximal ergódico $\int_{\bigcup_{N=0}^{\infty}\{x : H_N(x) > 0\}} (f-\alpha\chi_C)dm \geq 0$ (H_N está definido analogamente como F_N en el teorema maximal ergódico). Pero $C \subset \bigcup_{N=0}^{\infty}\{x : H_N(x) > 0\}$ esto es $\int_X \mid f \mid\geq \alpha m(C)$. Por lo tanto $m(C) \leq \frac{1}{\alpha}\int_X \leq\mid f \mid dm$ para todo subconjunto de $E_{\alpha,\beta}$ con medida finita y por lo tanto $m(E_{\alpha,\beta}) < \infty$. Si $\alpha < 0$ entonces $\beta < 0$ así que podemos aplicar lo anterior con $-f$ y $-\beta$ reemplazando f y α para obtener $m(E_{\alpha,\beta}) < \infty$.

Sea $B_\alpha = \{x \in X : \sup_{n\geq 1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) > \alpha\}$. Entonces aplicando el corolario anterior

$$\int_{E_{\alpha,\beta}} fdm = \int_{E_{\alpha,\beta}\cap B_\alpha} fdm \geq \alpha m(E_{\alpha,\beta} \cap B_\alpha) = \alpha m(E_{\alpha,\beta}).$$

Es decir,

$$\int_{E_{\alpha,\beta}} fdm \geq \alpha m(E_{\alpha,\beta})....(1)$$

Si reemplazamos f, α, β por $-f, -\alpha, -\beta$ respectivamente obtendremos que $(-f)^* = -f_*, (-f)_* = -f^*$ y

$$\int_{E_{\alpha,\beta}} f dm \leq \beta m(E_{\alpha,\beta})....(2)$$

Así, si $\alpha > \beta$ entonces $m(E_{\alpha,\beta}) = 0$, y desde

$$\{x : f_*(x) < f^*(x)\} \subset \bigcup_{\beta<\alpha} E_{\alpha,\beta} \text{ con } \alpha, \beta \text{ racionales}$$

tendremos $m\{x : f_*(x) < f^*(x)\} = 0$. Es decir, $f^*(x) = f_*(x)$ en casi todo punto. Por lo tanto $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ en casi todo punto.

Para mostrar que $f^* \in L^1(m)$ usamos la parte del lema de Fatou para funciones integrables no negativas g_n, $\liminf_{n\to\infty} \int g_n dm < \infty$ implica que $\liminf_{n\to\infty} g_n$ es integrable. Sea

$$g_n(x) = \mid \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \mid .$$

Entonces

$$\int g_n dm = \int \mid \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \mid dm \leq \int \mid f \mid dm,$$

esto es $\liminf_{n\to\infty} g_n dm < \infty$. Y por el lema de Fatou $\liminf_{n\to\infty} = \mid f_* \mid$ es integrable. Por lo tanto f_* es integrable.

Queda por demostrar que $\int f dm = \int f^* dm$ si $m(X) < \infty$.

Sea $D_k^n = \{x \in X : \frac{k}{n} \leq f^*(x) < \frac{k+1}{n}\}$ donde $k \in \mathbb{Z}, n \geq 1$. Para cada $\epsilon > 0$ se tiene que $D_k^n \cap B_{\frac{k}{n}-\epsilon} = D_k^n$ y por el corolario anterior

$$\int_{D_k^n} f dm \geq (\frac{k}{n} - \epsilon) m(D_k^n)$$

esto es

$$\int_{D_k^n} f dm \geq \frac{k}{n} m(D_k^n)....(3)$$

Entonces

$$\int_{D_k^n} f^* dm \leq \frac{k+1}{n} m(D_k^n) \leq \frac{1}{n} m(D_k^n) + \int_{D_k^n} f dm \text{ por (3).}$$

Sumando sobre $k \in (Z)$ obtenemos que

$$\int_X f^* dm \leq \frac{m(X)}{n} + \int_X f dm \ \forall n \geq 1.$$

Por tanto $\int_X f^* dm \leq \int_X f dm$ ya que $m(X) < \infty$. Aplicando esto para $-f$ en lugar de f obtenemos que
$\int_X (-f)^* dm \leq \int_X -f dm$, es decir, $-\int_X f_* dm \leq -\int_X f dm$. Como $f_* = f^*$ en casi todo punto, tenemos que $\int_X f^* dm \geq \int_X f dm$.
Por lo tanto, $\int f^* dm = \int f dm$.

□

3.5.1. Aplicaciones del teorema de Birkhoff

Teorema 3.5.2. *Sea (X, Σ, m) un espacio de probabilidad y $T : X \to X$ una funcion que preserva medida, entonces T es ergodica si y solo si para todo $A, B \in \Sigma$ se cumple que*

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

Demostración. $(\Rightarrow)$ Suponga que T es ergódica. Poniendo $f = \chi_A$ en el teorema anterior, se tiene $\frac{1}{n}\sum_{i=0}^{n-1} \chi_A(T^i(x)) \to m(A)$ en casi todo punto. Multiplicando por χ_B tenemos que

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_A(T^i(x))\chi_B \to m(A)\chi_B \text{ en casi todo punto.}$$

Por el teorema de la convergencia dominada si nosotros integramos obtenemos que

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

$(\Leftarrow)$ Sea $T^{-1}(E) = E, E \in \Sigma$. Sea $A = B = E$. Entonces

$$\frac{1}{n}\sum_{i=0}^{n-1} m(E) \to (m(E))^2.$$

Pero también sabemos que

$$\frac{1}{n}\sum_{i=0}^{n-1} m(E) \to m(E)$$

de donde $m(E) = (m(E))^2$ entonces $m(E) = 0$ o $m(E) = 1$. $\square$

Teorema 3.5.3. *Sea $1 \leq p < \infty$. Sea T una funcion que preserva medida en un espacio de probabilidad (X, Σ, m). Si $f \in L^p(m), \exists g \in L^p(m); goT = T$ en casi todo punto y*

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) - g(x) \right\|_p \to 0.$$

Demostración. Si g es acotada y medible entonces $g \in L^p$ $\forall p$ y por el teorema ergódico tenemos que

$$\frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) \to g^* \text{ en casi todo punto.}$$

Claramente $g \in L^\infty(m)$ y por lo tanto $g^* \in L^p(m)$.

También, $\mid \frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - g^*(x) \mid^p \to 0$ en casi todo punto. Y por el teorema de convergencia dominada(acotada), $\left\| \frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - g^*(x) \right\|_p \to 0$ es decir, $\forall\ \epsilon > 0 \exists N(\epsilon, g)$, si $n > N(\epsilon, g)$ y $k > 0$

$$\left\| \frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) - \frac{1}{n+k}\sum_{i=0}^{n+k-1} g(T^i(x)) \right\|_p < \epsilon.$$

Sea $f \in L^p(m)$, y $M_n(f)(x) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$. Nosotros vamos a mostrar que $\{M_n(f)\}$ es una secuencia de Cauchy en $L^p(m)$. Note que $\| M_n(f) \|_p \leq \| f \|_p$. Elegimos $g \in L^\infty(m), \| f - g \|_p < \frac{\epsilon}{4}$; entonces

$$\| M_n(f) - M_{n+k}(f) \|_p \leq \| M_n(f) - M_n(g) \|_p + \| M_n(g) - M_{n+k}(g) \|_p$$

$$+ \| M_{n+k}(g) - M_{n+k}(f) \|_p \leq \frac{\epsilon}{4} + \frac{\epsilon}{2} + \frac{\epsilon}{4} = \epsilon.$$

Si $n > N(\frac{\epsilon}{2}, g)$ y $k > 0$, tenemos que $f^* oT = T$ en casi todo punto ya que

$$(\frac{n+1}{n})(M_{n+1}(f)) - (M_n(f))(T(x)) = \frac{f(x)}{n}$$

con lo cual concluye la prueba. $\square$

3.5.2. Aplicaciones del teorema de Birkhoff

Se dice que un número irracional x es normal en base 10 si la frecuencia relativa en cada dígito en su expresión decimal es $\frac{1}{10}$. Dicho de otro modo, si $x = \sum_{i=1}^{\infty} \frac{a_i(x)}{10^i}$ con $a_i(x) \in \{0, 1, \ldots, 9\}$ decimos que x es normal si para todo $l \in \{0, 1, \ldots, 9\}$ se tiene que

$$\lim_{n\to\infty} \frac{\text{Card}(\{i : 1 \leq i \leq n, a_i(x) = l\})}{n} = \frac{1}{10},$$

donde Card(A) denota el número de elementos de A.

Definición 3.3. *Sea* $n_A(x) = \min\{n \in \mathbb{N} : T^n(x) \in A\}$.

Aplicación 1:
Casi todos los números del intervalo $[0, 1]$ son normales

Demostración. Tomamos X el conjunto de números irracionales con la σ-álgebra de Lebesgue y la medida de Lebesgue y consideramos la transformación

$$\tau : X \to X, \ \tau x = 10x(\text{mod } 1)$$

como se sabe τ es ergódica y conserva medida.
Por otra parte, si $x = \sum_{i=1}^{\infty} \frac{a_i(x)}{10^i}$ entonces $10x = \sum_{i=1}^{\infty} \frac{a_i(x)}{10^{i-1}}$, y por lo tanto $\tau x = \sum_{i=2}^{\infty} \frac{a_i(x)}{10^{i-1}}$ y generalizando este hecho tenemos que $\tau^j x = \sum_{i=j+1}^{\infty} \frac{a_i(x)}{10^{i-j}}$.
Por lo tanto

$$\text{Card}\{i : 1 \leq i \leq n = a_i(x) = l\} = \text{Card}\{i : 1 \leq i \leq n, \ \tau^{i-1}x \in \left[\frac{l}{10}, \frac{l+1}{10}\right)\}.$$

De aquí decimos que

$$\frac{\text{Card}\{i : 1 \leq i \leq n, a_i(x) = l\}}{n} = \frac{1}{n}\sum_{i=1}^{n} \chi_{[\frac{l}{10}, \frac{l+1}{10})}(\tau^{i-1}x) = \frac{1}{n}\sum_{i=0}^{n-1} \chi_{[\frac{l}{10}, \frac{l+1}{10})}.$$

Tomando límite cuando $n \to \infty$ y aplicando el teorema de Birkhoff (para τ ergódica) tenemos que

$$\lim_{n\to\infty} \frac{\text{Card}\{i : 1 \leq i \leq n, a_i(x) = l\}}{n} = \int_X \chi_{[\frac{l}{10}, \frac{l+1}{10})}(x)dx = \frac{1}{10} \text{ para casi todo } x.$$

□

Aplicación 2:

Sea T una aplicación que preserva medida en el espacio de probabilidad (X, Σ, μ). Suponga que $B \in \Sigma$. Entonces para casi todo punto $x \in X$, la frecuencia con que la órbita de X se encuentra en B es $\mu(B)$, es decir

$$\lim_{n\to\infty} \frac{1}{n}\text{Card}\{j \in \{0, 1, ..., n-1\} \text{ tal que } T^j(x) \in B\} = \mu(B) \text{ en casi todo punto}$$

Demostración. Sólo basta aplicar el teorema de Birkhoff a $f = \chi_B$. □

Aplicación 3:

Sea $T : X \to X$ una transformación ergódica medible en el espacio de probabilidad (X, Σ, μ) y A un conjunto medible, $\mu(A) > 0$. Entonces $n_A \in L^1$ y, para el primer retorno $n_A(x)$ para cualquier $x \in A$,

$$\int_A n_A(x)d\mu_A(x) = \frac{1}{\mu(A)} \quad \text{resp.} \quad \int_A n_A(x)d\mu(x) = 1$$

y

$$\lim_{N\to\infty} \frac{1}{n} \sum_{0\leq n<N} n_A(T^n(x)) = \frac{1}{\mu(A)}.$$

Demostración. Para todo $x \in A$ consideremos la órbita de x bajo T_A, esto es

$$x, T_A(x), \ldots, T_A^n(x), \ldots, T_A^N(x) \ldots.$$

La cantidad $\tau := \sum_{0\leq n<N} n_A(T_A^n(x))$ mide el tiempo de los primeros N retornos de x bajo T al conjunto A, es decir:

$$\sum_{0\leq n<\tau} \chi_A(T^n(x)) = N.$$

Aplicando el teorema de Birkhoff para T y T_A (cuando $N \to \infty$ y $t \to \infty$), tenemos que

$$\int_A n_A(x)d\mu_A(x) = \lim_{n\to\infty} \frac{1}{N} \sum_{0\leq n<N} n_A(T_A^n(x))$$

$$= \lim_{\tau\to\infty} \frac{\tau}{\sum_{0\leq n<\tau} \chi_A(T^n(x))} = \left(\int_X \chi_A d\mu\right)^{-1} = \frac{1}{\mu(A)}$$

□

Aplicación 4:

Sea $T : [0,1> \to [0,1>$ la transformación de Gauss definida como $T(x) = 0$ si $x = 0$, y $T(x) = \frac{1}{x} - [\frac{1}{x}]$ si $x \neq 0$. Si $x \neq 0$ tenemos que

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

Sea

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

con $p_n(x), q_n(x)$ primos relativos. Para casi todo x con respecto a la medida de Lebesgue se tiene que:

1. $\lim\limits_{n\to\infty} \sqrt[n]{a_1 \dots a_n} = \prod\limits_{k=1}^{\infty} (1 + \dfrac{1}{k^2 + 2k})^{\frac{\log k}{\log 2}}$

2. $\lim\limits_{n\to\infty} \dfrac{\log q_n}{n} = \dfrac{\pi^2}{12 \log 2}$.

Para demostrar estas ecuaciones usaremos el hecho de que T es $\mu-$ergódica donde μ es definida en los borelianos por

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

Para hallar 1 vemos que $a_i(x) = k \Leftrightarrow T^{i-1}(x) \in \left[\frac{1}{k+1}, \frac{1}{k}\right)$. Sea $f :< 0, 1 > \to \mathbb{R}$ definida por $f(x) = \log k$ para $x \in \left[\frac{1}{k+1}, \frac{1}{k}\right)$.

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

converge a

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

Y para hallar 2 haremos lo siguiente

$$\frac{p_n(x)}{q_n(x)} = [a_1, \dots, a_n] = \frac{1}{a_1 + [a_2, \dots, a_n]}$$

$$= \frac{1}{1 + \frac{p_{n-1}(T(x))}{q_{n-1}(T(x))}} = \frac{q_{n-1}(T(x))}{a_1 q_{n-1}(T(x)) + p_{n-1}(T(x))}.$$

Las fracciones de la izquierda y la derecha no se pueden simplificar por lo tanto $p_n(x) = q_{n-1}(T(x))$. Entonces

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

de donde

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

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

3- $p_n(x) = a_n(x) p_{n-1}(x) + p_{n-2}(x)$

4- $q_n(x) = a_n(x) q_{n-1}(x) + q_{n-2}(x)$

lo que implica que las sucesiones $p_n(x), q_n(x)$ son crecientes. Hagamos el paso inductivo

$$p_{n+1}(x) = q_n(T(x)) = a_n(T(x)) q_{n-1}(T(x)) + q_{n-2}(T(x))$$
$$= a_{n+1}(x) p_n(x) + p_{n-1}(T(x))$$

$$q_{n+1}(x) = a_1(x) q_n(T(x)) + p_n(T(x)) = a_1(x)(a_{n+1}(x) q_{n-1}(T(x)) + q_{n-2}(T(x)))$$
$$+ a_{n+1}(x) p_{n-1}(T(x)) + p_{n-2}(T(x)) = a_{n-1}(x) q_n(x) + q_{n-1}(x).$$

Utilizando 3 y 4 se demuestra fácilmente por inducción que $p_k \geq 2^{\frac{k-2}{2}}$ y $q_k \geq 2^{\frac{k-1}{2}}$ para $k \geq 1$. Como

5- $\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{q_n p_{n-1} - p_n q_{n-1}}{q_{n+1} q_n} = \frac{q_{n-1}}{q_{n+1}} (\frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n}),$

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

Utilizando 5 se sigue fácilmente por inducción que

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

Similarmente se prueba por inducción que

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

y luego

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

De donde

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

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

Por lo tanto

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

Bibliografía

[Ash72] Robert B. Ash, *Real analysis and probability*, Academic Press, New York-London, 1972, Probability and Mathematical Statistics, No. 11. MR 0435320 (55 #8280)

[BC92] L. S. Block and W. A. Coppel, *Dynamics in one dimension*, Lecture Notes in Mathematics, vol. 1513, Springer-Verlag, Berlin, 1992. MR 1176513 (93g:58091)

[dMvS93] Welington de Melo and Sebastian van Strien, *One-dimensional dynamics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 25, Springer-Verlag, Berlin, 1993. MR 1239171 (95a:58035)

[KH95] Anatole Katok and Boris Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge University Press, Cambridge, 1995, With a supplementary chapter by Katok and Leonardo Mendoza. MR 1326374 (96c:58055)

[Lim81] Elon Lages Lima, *Curso de análise. Vol. 2*, Projeto Euclides [Euclid Project], vol. 13, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1981. MR 654862 (83h:26002b)

[LY75] Tien Yien Li and James A. Yorke, *Period three implies chaos*, Amer. Math. Monthly **82** (1975), no. 10, 985–992. MR 0385028 (52 #5898)

[Rud87] Walter Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987. MR 924157 (88k:00002)