

# SHIFTS DE BERNOULLI

## UNA INTRODUCCIÓN A LA TEORÍA ERGÓDICA

AUBIN ARROYO

Uno de los objetivos de la teoría de los sistemas dinámicos es descubrir y describir las características de la acción de una transformación sobre un espacio. Dado un espacio de puntos y una transformación del espacio en si mismo, a la dinámica le concierne averiguar cómo se mueven los puntos al aplicar sucesivamente la transformación y, en resumidas cuentas, cuál es el destino que ésta les depara. Convengamos desde un principio: un sistema dinámico es un espacio  $M$  y una transformación  $T : M \rightarrow M$ . Así, a cada un punto  $x \in M$  le podemos definir un sucesor y este será  $T(x)$ . El sucesor del sucesor de  $x$  será entonces  $T(T(x)) = T^2(x)$ , y así sucesivamente. El principal objeto de nuestro interés serán entonces las órbitas del sistema dinámico:

$$\mathcal{O}(x) = \{x, T(x), T^2(x), \dots, T^n(x), \dots\}$$

Los *Shifts de Bernoulli* son ejemplos de sistemas dinámicos que presentan algunas propiedades dinámicas interesantes y serán nuestro guía en un breve recorrido en una parte de la teoría de los sistemas dinámicos: la Teoría Ergódica. Por otra parte, los shifts de Bernoulli juegan un papel fundamental debido a que nos brindan valiosa información sobre el comportamiento de lo que algunos llaman "sistemas caóticos".

A lo largo de este minicurso construiremos los Shifts de Bernoulli y estudiaremos algunas de sus propiedades dinámicas básicas. Para esto introduciremos algunos conceptos de la Teoría Ergódica como las medidas invariantes y la ergodicidad. También presentaremos algunas nociones de la Dinámica Topológica. Durante este recorrido visitaremos dos teoremas importantes: el Teorema de Recurrencia de Poincaré y el Teorema Ergódico de Birkhoff.

Nos interesa estudiar las órbitas posibles de un determinado sistema dinámico. ¿Cuántos puntos tiene una órbita? ¿es posible que existan órbitas con un solo elemento? Esto es ¿existe cierto  $p \in M$  tal que  $T(p) = p$ ? A los puntos cuya órbita consta de un solo punto se les llama *puntos fijos* de  $T$ . Incluso, es posible también que:

$$p \neq T(p), T^2(p) \neq p, \dots, T^{k-1}(p) \neq p$$

sin embargo,  $T^k(p) = p$ . En este caso,  $p$  es un *punto periódico de período  $k$* . Otra posibilidad es que la órbita de  $p$  sea una sucesión infinita de puntos de  $M$ .

**Ejemplo:**  $M = \mathbb{R}$  y  $T: \mathbb{R} \rightarrow \mathbb{R}$  es  $T(x) = -x$ . El punto  $x = 0$  es un punto fijo y todos los demás puntos son puntos periódicos de periodo 2; esto es: si  $x \neq 0$   $T^2(x) = x$  y  $T(x) \neq x$ .

**0.1. Topología.** Es necesario, por otra parte, dotar al espacio  $M$  de alguna estructura que nos permita describir el comportamiento de las órbitas. Si contamos con una estructura topológica en el espacio  $M$ , entonces podremos preguntarnos si las órbitas infinitas convergen o no a determinado punto; y qué propiedades dinámicas tiene el dicho punto de convergencia. Para esto, es fundamental que la transformación preserve dicha estructura. Las funciones que preservan la topología son las funciones continuas, pero primero establezcamos qué es una topología en  $M$

Una topología no es más que una noción adecuada de vecindad entre los puntos de  $M$ , una manera de decidir, dado  $x \in M$ , cuales son los puntos "cercaños" a  $x$ . Esta noción es suficiente para poder hablar de convergencia de sucesiones, conjuntos abiertos y cerrados, y también de funciones continuas.

Si bien la noción técnica de topología es bastante general y un poco rebuscada, una manera sencilla de obtener una topología para  $M$  es a través de una función *distancia* en  $M$ , es decir, una función que nos informe la distancia entre cualesquiera dos puntos  $x, y \in M$ . Esto es, una función  $d: M \times M \rightarrow \mathbb{R}^+$  que verifique las siguientes propiedades:

- (1)  $d(x, y) = 0$  sí y sólo si  $x = y$
- (2)  $d(x, y) = d(y, x)$ , para todos  $x, y \in M$
- (3)  $d(x, z) \leq d(x, y) + d(y, z)$ , para todos  $x, y, z \in M$

Un espacio métrico es un espacio de puntos  $M$  dotado de una función distancia  $d$ . Denotaremos un espacio métrico por la pareja  $(M, d)$  o simplemente por  $M$ , cuando no sea necesario escribir explícitamente la función distancia.

Dado un punto  $x \in M$  y un número  $\varepsilon > 0$ , la bola *abierta* de radio  $\varepsilon > 0$  con centro en  $x$  es el conjunto:

$$B_\varepsilon(x) := \{y \in X \mid d(x, y) < \varepsilon\}$$

Las bolas abiertas nos permiten definir dos clases de subconjuntos de  $M$ : los conjuntos abiertos y los conjuntos cerrados.

**Definición 1.** Un subconjunto  $U \subset M$  es abierto si para cada punto  $x \in U$  existe  $\varepsilon > 0$  tal que  $B_\varepsilon(x) \subset U$ . Una vecindad de  $x$  es simplemente un abierto  $U$  que contiene a  $x$ .

**Propiedad 1.** La familia  $\tau$  de subconjuntos abiertos de  $M$  satisface las siguientes propiedades:

- (1)  $\emptyset \in \tau$  y  $M \in \tau$ .
- (2) Si  $U_j \in \tau$ ,  $j \in J$ , entonces  $\bigcup_{j \in J} U_j \in \tau$ ; para algún conjunto de índices  $J$ . ( $J$  puede ser infinito no numerable).
- (3) Si  $U_1, U_2, \dots, U_k \in \tau$  entonces  $\bigcap_{j=1}^k U_j \in \tau$ . (Finitos)

Los subconjuntos de  $M$  que son el complemento de un abierto reciben el nombre de subconjuntos *cerrados*.

En general, a una familia  $\tau$  de subconjuntos de  $M$  que satisface las características enunciadas en la Propiedad 1 se le llama formalmente una *topología* de  $M$ , y esta es suficiente para poder hablar de convergencia de sucesiones y funciones continuas sin tener que hacer uso de una función distancia. El libro de Munkres [2] es una excelente referencia para los conceptos de topología que aquí utilizaremos.

Supongamos que  $M$  es un espacio métrico entonces podemos definir cuándo una transformación  $T: M \rightarrow M$  es *continua*: si sucede que para cualquier abierto  $U$  de  $M$ ,  $f^{-1}(U)$

es un abierto de  $M$ . En este sentido, las funciones abiertas preservan los conjuntos abiertos de  $M$  al tomar su imagen inversa.

**Definición 2.** Una sucesión de puntos  $\{x_i\} \subset M$  converge a  $y \in M$  si para cualquier vecindad de  $y$  (i.e. cualquier abierto  $U$  tal que  $y \in U$ ) se cumple que  $\#\{(M-U) \cap \{x_i | i \in \mathbb{N}\}\}$  es finito; en este caso se escribe:

$$\lim_{n \rightarrow \infty} x_n = y$$

En otras palabras, la sucesión converge a  $y$  si para cualquier abierto que contenga a  $y$ , todos los puntos de la sucesión, salvo un número finito, están contenidos en él. Si bien el conjunto  $U$  puede ser muy grande; y así, el hecho de que sólo quede un número finito fuera no sea muy "difícil" de satisfacer, pero esto sucede para cualquier abierto; y por lo tanto no importa cuán pequeño sea  $\varepsilon > 0$ , fuera de la bola  $B_\varepsilon(y)$  sólo hay un número finito de puntos de la sucesión; y por ende, dentro están todos los demás.

**Teorema 1.** Sea  $M$  es un espacio métrico y  $T : M \rightarrow M$  una transformación continua. Si la órbita de  $x$  converge a  $y$ , entonces  $y$  es un punto fijo:  $T(y) = y$ .

*Demostración:* Considere la sucesión  $\{T^n(x)\}$ , con  $n \in \mathbb{N}$  y sea  $y \in M$  tal que:

$$\lim_{n \rightarrow \infty} T^n(x) = y$$

entonces,  $T(y) = T(\lim_{n \rightarrow \infty} T^n(x))$ . Como  $T$  es continua, "entra" al límite. Entonces:

$$T(y) = \lim_{n \rightarrow \infty} T(T^n(x)) = \lim_{n \rightarrow \infty} T^{n+1}(x) = y \quad \square$$

La *compacidad* es una característica del espacio que permite afirmar que cualquier sucesión de puntos contiene una subsucesión convergente, esto es, no está permitido que toda la sucesión se escape del espacio.

A pesar de que la definición estándar de subconjunto compacto de  $\mathbb{R}$  es más conocida:  $K \subset \mathbb{R}$  es un compacto si es un cerrado y está acotado, utilizaremos la siguiente definición, que es más general (no es necesario suponer que el espacio es métrico), pues nos será más útil para trabajar.

**Definición 3.** Decimos que  $M$  es un compacto si, dada cualquier familia de abiertos  $U_\alpha \subset M$ ,  $\alpha \in A$ , tales que  $M = \cup_{\alpha \in A} U_\alpha$ , siempre existe un número finito  $U_{\alpha_1}, U_{\alpha_2}, \dots, U_{\alpha_k}$  tales que

$$M = \cup_{j=1}^k U_{\alpha_j}$$

**Ejercicio:** Si  $K \subset \mathbb{R}$  es un subconjunto cerrado y acotado, entonces  $K$  es un compacto.

Un subconjunto  $A$  es *denso* en  $M$  si para cualquier punto  $x \in M$  y cualquier  $\varepsilon > 0$  se tiene que  $B_\varepsilon(x) \cap A \neq \emptyset$ ; esto es, los puntos de  $A$  están arbitrariamente cerca de cualquier punto de  $M$ .

**Lema 1.** Si  $M$  es un espacio métrico compacto, existe un subconjunto  $\{y_j | j \in \mathbb{N}\}$  denso en  $M$ .

*Demostración:* Considera una sucesión de números  $\varepsilon_n > 0$  tales que  $\varepsilon_n \rightarrow 0$ , cuando  $n \rightarrow \infty$ . Con  $n$  fijo,  $\varepsilon_n$  nos determina una familia de abiertos:

$$B_{\varepsilon_n}(y) \text{ con } y \in M$$

Obviamente  $M = \cup_{y \in M} B_{\varepsilon_n}(y)$ . Como  $M$  es compacto son suficientes sólo un número finito de bolas para cubrir  $M$ , esto es, para cada  $n \in \mathbb{N}$  existen:  $y_1^n, y_2^n, \dots, y_{s_n}^n$  tales que

$$M = \cup_{j=1}^{s_n} B_{\varepsilon_n}(y_j^n)$$

Por lo tanto, el conjunto

$$\{x_1^1, \dots, x_{s_1}^1, x_1^2, \dots, x_{s_2}^2, \dots, x_1^n, \dots, x_{s_n}^n, \dots\}$$

es denso en  $M$ . Para verificarlo, toma cualquier  $x \in M$  y  $\varepsilon > 0$ . Ahora bien, como  $\varepsilon_n \rightarrow 0$  existe cierta  $\varepsilon_{n_0} < \varepsilon$ . Como  $y \in \cup_{j=1}^{s_{n_0}} B_{\varepsilon_{n_0}}(y_j^{n_0})$ , podemos afirmar que  $y \in B_{\varepsilon_{n_0}}(y_j^{n_0})$ , para alguna  $j$  y por lo tanto:

$$d(x, y_j^{n_0}) < \varepsilon_{n_0} < \varepsilon. \quad \square$$

En general, a la familia de las bolas de radio  $\varepsilon_n$  con centro en los puntos  $\{x_r^j\}$  que obtuvimos en el lema anterior se le llama: una *base numerable* de la topología de  $M$ . En el intervalo  $[0, 1]$ , por ejemplo, las bolas de radio racional sobre puntos racionales forman una base numerable de vecindades.

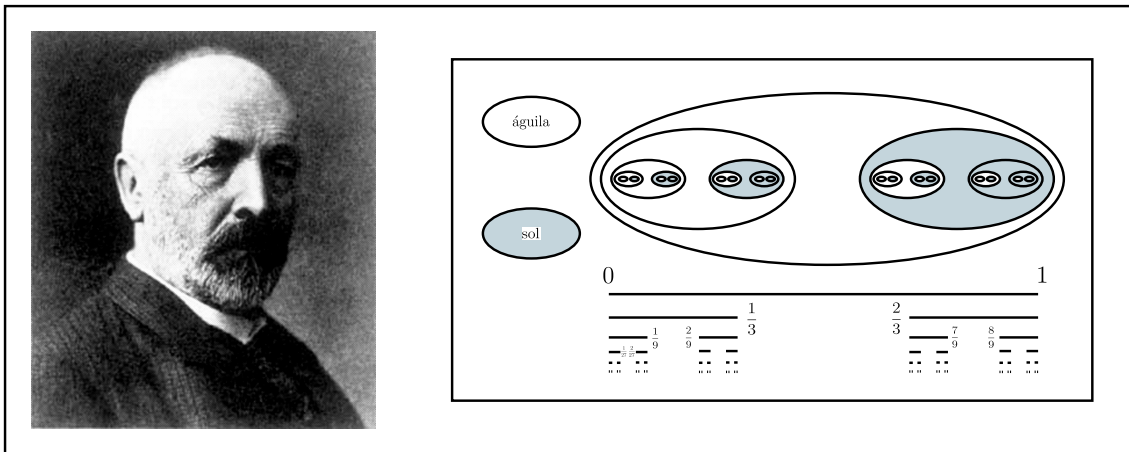


FIGURE 1. Georg Ferdinand Ludwig Philipp Cantor y el conjunto de Cantor

**El conjunto de Cantor ternario.** El famoso matemático Georg Cantor (1845-1918), al que le debemos la noción de función biyectiva, le dió su nombre a este conjunto que aparece en diversos contextos de las matemáticas. El conjunto de Cantor tiene diversas presentaciones, aunque bien todas son la misma.

En nuestro caso, podemos utilizar el conjunto de Cantor ternario como modelo. Este será el espacio ambiente de los sistemas dinámicos que nos interesan: *los Shifts de Bernoulli*. Una manera sucinta de describirlo es como el siguiente conjunto del intervalo  $[0, 1] \subset \mathbb{R}$ .

$$K = \{x \in [0, 1] \mid x = \sum_{i=1}^{\infty} \frac{x_i}{3^i}, \text{ donde } x_i \in \{0, 2\}\}$$

En resumidas cuentas, esta descripción del conjunto de Cantor se refiere a un subconjunto del intervalo  $[0, 1]$  el cual está determinado por aquellos puntos que no presentan el dígito

"1" en su expansión en base 3. La manera clásica de construirlo es hacerlo por etapas. Los puntos que en base 3 se escriben como

$$a = .a_1a_2a_3a_4a_5a_6a_7a_8 \dots \in [0, 1]$$

tales que  $a_1 \neq 1$  o bien  $a \in [0, 1/3]$ , cuando  $a_1 = 0$ , o  $a \in [2/3, 1]$ , cuando  $a_1 = 2$ . Si  $a_1 = 0$  y  $a_2 = 0$  entonces podemos ser más precisos sobre la ubicación de  $a$  en el intervalo:  $a \in [0, 1/9]$ , cuando  $a_2 = 0$  o bien,  $a \in [2/9, 1/3]$  cuando  $a_2 = 2$  y así sucesivamente. En general, el conjunto

$$K_n = \{.a_1a_2a_3a_4a_5a_6 \dots \in [0, 1] \mid a_j \neq 1 \text{ para } j \in \{1, \dots, n\}\}$$

está formado por  $2^n$  intervalos de longitud  $3^{-n}$ , y entonces  $K = \bigcap_{n=1}^{\infty} K_n$

Para definir una topología sobre el conjunto  $K$ , basta recordar que en el intervalo está bien definida una función distancia; o bien mediante los subconjuntos abiertos de  $[0, 1]$ , esto es,  $U$  es un abierto de  $K$  si y sólo si  $U = V \cap K$ , para algún abierto  $V$  del intervalo.

Otra manera de visualizar el conjunto de Cantor es mediante secuencias (infinitas) de volados. Si bien este espacio no tiene, en un principio mucha cara, podemos separarlo en dos partes: la parte que contiene a todos los volados que comienzan con "águila" y los que comienzan con "sol"; digamos  $C(0, A)$  y  $C(0, S)$ . Ahora bien, cada una de estas parte también la podemos dividir en dos; esto es,  $C(0, A) = (0, AA) \cup C(0, AS)$ . La primera parte corresponde a las secuencias de volados que comienzan con dos águilas seguidas y la otra primero cae águila y después sol. Análogamente  $C(0, S) = (0, SA) \cup C(0, SS)$ . De esta manera podemos caracterizar el conjunto de todas las secuencias de volados. De hecho, si reemplazamos los signos  $A$  por 0 y  $S$  por 2 e identificamos una sucesión de volados con el número real en base 3 definido de acuerdo a esta sustitución, por ejemplo:

$$AASSASSAS \dots \longleftrightarrow .002202202 \dots$$

tenemos una correspondencia entre los puntos del conjunto de Cantor con las secuencias infinitas de volados.

Tirar una moneda al aire inmediatamente nos lleva a pensar en la probabilidad. Si la moneda no está cargada podemos creer sin duda que en promedio, la mitad de las veces que nos echemos un volado, esta caerá en sol. Para poder hablar de probabilidad, de manera formal necesitamos revisar algunas nociones básicas de la teoría de la medida y es lo que haremos en la siguiente sección.

**0.2. Teoría de la Medida.** A la Teoría de la Medida le concierne, obviamente, la tarea de medir los subconjuntos de un cierto espacio  $M$ . En otras palabras, le interesa definir una función  $\mu$  que a los subconjuntos  $A \subset M$  les asigne un número real positivo de acuerdo a su "tamaño" y que se comporte como intuitivamente se comporta el tamaño: si  $A$  y  $B$  son subconjuntos disjuntos de  $M$  entonces  $\mu(A \cup B) = \mu(A) + \mu(B)$ , así como también  $\mu(A^c) = \mu(M) - \mu(A)$ . Sin embargo, proceder ingenuamente de esta manera nos lleva a paradojas muy curiosas como aquella de la manera de partir una naranja en pedazos y volvera a juntarlos de manera que el resultado es del tamaño del planeta Júpiter: El conjunto de Vitali. Para evitar esta incómoda consecuencia, debemos ser cuidadosos a la hora de escoger los conjuntos que podremos medir y resignarnos ante la idea de que no podremos medir todos los subconjuntos de un espacio  $M$ .

**Definición 4.** Una familia  $\mathcal{B}$  de subconjuntos de  $M$  es una familia de subconjuntos medibles de  $M$  si satisface las siguientes propiedades:

- (1)  $M \in \mathcal{B}$
- (2) Si  $A \in \mathcal{B}$  entonces  $A^c = (M \setminus A) \in \mathcal{B}$
- (3) Si  $A_i \in \mathcal{B}$ ,  $i \in \mathbb{N}$ , entonces  $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{B}$ .

A la pareja  $(M, \mathcal{B})$  la llamaremos un espacio medible.

El nombre técnico de una familia de subconjuntos que satisface las propiedades anteriores es el de  $\sigma$ -álgebra: la familia forma un *álgebra* con respecto a las operaciones de conjuntos: unión y complemento. Dado que en el inciso 3 consideramos uniones numerables, esta álgebra lleva el nombre de  $\sigma$ -álgebra. Si solamente garantizamos que la unión finita de elementos medibles es medible, a la familia  $\mathcal{B}$  se le llama simplemente: un álgebra de subconjuntos de  $M$ . Más adelante veremos que a partir de un algebra  $\mathcal{B}_0$  podemos construir la menor  $\sigma$ -álgebra  $\mathcal{B}$  que contiene a  $\mathcal{B}_0$ , considerando todas las uniones numerables de elementos de  $\mathcal{B}_0$ . En este caso,  $\mathcal{B}$  es la  $\sigma$ -álgebra generada por  $\mathcal{B}_0$  y será de bastante utilidad. Para garantizar que  $\mathcal{B}$  es la menor  $\sigma$ -álgebra que contiene a  $\mathcal{B}_0$  debemos definirla como la intersección de todas las  $\sigma$ -álgebras que contienen a  $\mathcal{B}_0$ .

Por otro lado, observe que si consideramos una cantidad numerable de elementos  $A_j$  de una familia de subconjuntos medibles  $\mathcal{B}$ , ( $j \in \mathbb{N}$ ), entonces también podemos garantizar que la intersección de todos ellos es medible:

$$\bigcap_{j=1}^{\infty} A_j = \left( \bigcup_{j=1}^{\infty} A_j^c \right)^c \in \mathcal{B}$$

**Funciones medibles.** Así como las funciones continuas respetan la topología (la pre-imagen de un abierto es un abierto), a las funciones que respetan a los conjuntos medibles las llamaremos *funciones medibles*.

**Definición 5.** *Dados dos espacios medibles  $(M, \mathcal{B})$  y  $(M', \mathcal{B}')$  Una función entre ellos,  $f: (M, \mathcal{B}) \rightarrow (M', \mathcal{B}')$  es medible si, para cualquier elemento  $B' \in \mathcal{B}'$  se tiene que  $f^{-1}(B') \in \mathcal{B}$ .*

Ahora bien, un *espacio de medida* es un espacio medible junto con una *medida*. Una medida es una función  $\mu: \mathcal{B} \rightarrow [0, +\infty]$  que cumple las siguientes propiedades:

- (1)  $\mu(\emptyset) = 0$
- (2) Si  $A_j \in \mathcal{B}$ ,  $j \in \mathbb{N}$ , disjuntos por pares, entonces  $\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$ .

Si requerimos además que  $\mu(M) = 1$ , a la medida la llamaremos una *probabilidad* y en este caso  $(M, \mathcal{B}, \mu)$  es un espacio de probabilidad. Supongamos que  $(M, \mu)$  es un espacio de probabilidad y tomemos un subconjunto medible  $A$ . Podemos pensar que  $\mu(A)$  es la probabilidad de que un punto aleatorio de  $M$  pertenezca a  $A$ . Por lo tanto, los conjuntos de medida cero no importan, pues algo que tiene probabilidad 0 de ocurrir, desde esta perspectiva no ocurre. Incluso, no importa que una transformación medible no esté definida en un conjunto de medida 0.

Así, si dos conjuntos difieren en un conjunto de medida cero, esto es si  $A$  y  $B \in \mathcal{B}$  son tales que la diferencia simétrica<sup>1</sup> entre ellos es tal que  $\mu(A \Delta B) = 0$ ; entonces decimos que ellos son iguales *módulo 0* y lo denotamos

$$A = B \text{ mod } 1, \text{ ó simplemente } A = B.$$

<sup>1</sup>La diferencia simétrica de dos conjuntos  $A$  y  $B$  se define como  $A \Delta B := (A - B) \cup (B - A)$ .

Una propiedad es válida para  $\mu$ -casi todo punto si el conjunto de puntos donde dicha propiedad no es válida tiene medida cero.

Una  $\sigma$ -álgebra  $\mathcal{B}$  de subconjuntos de  $M$  es una familia relativamente grande, sobre todo, a la hora de definir una función en cada uno de sus elementos. Sin embargo, si  $\mathcal{B}_0$  genera a  $\mathcal{B}$  podemos simplificar el trabajo. Los dos teoremas siguientes nos permiten considerar únicamente a los elementos del álgebra  $\mathcal{B}_0$ . El primero nos afirma que, desde el punto de vista de una medida, podemos aproximar arbitrariamente cualquier conjunto medible por algún elemento del álgebra que la genera.

**Teorema de aproximación.** *Sea  $(M, \mathcal{B}, \mu)$  un espacio de probabilidad y sea  $\mathcal{B}_0$  un álgebra que genera la  $\sigma$ -álgebra  $\mathcal{B}$ . Entonces, para todo  $\varepsilon > 0$  y para todo  $B \in \mathcal{B}$  existe  $B_0 \in \mathcal{B}_0$  tal que  $\mu(B \Delta B_0) < \varepsilon$ .*

Por otro lado, construir medidas a partir de familias tan grandes como lo es la familia  $\mathcal{B}$  puede resultar complicado. Sin embargo, si  $\mathcal{B}$  está generada por un álgebra  $\mathcal{B}_0$ , podemos recurrir al siguiente teorema que garantiza que basta definir la medida  $\mu$  en  $\mathcal{B}_0$  para obtener una medida.

**Teorema de extensión.** *Si  $\mathcal{B}_0$  es un álgebra de subconjuntos de  $M$  y  $\mu_0: \mathcal{B}_0 \rightarrow [0, +\infty]$  es una función que cumple las propiedades de medida para elementos de  $\mathcal{B}_0$ , entonces existe una única función  $\mu: \mathcal{B} \rightarrow [0, +\infty]$  que es una extensión de  $\mu_0$  a la  $\sigma$ -álgebra generada por  $\mathcal{B}_0$ . (Esto es,  $\mu_0$  coincide con  $\mu$  en los elementos de  $\mathcal{B}_0$  y está bien definida en todo  $\mathcal{B}$ ). Además, si  $\mu_0(M) = 1$  entonces  $\mu$  es una probabilidad.*

En un espacio métrico  $M$  los conjuntos borelianos de  $M$  son los elementos de una familia de subconjuntos medibles de  $M$  que contiene a los abiertos de  $M$ ; esto es, la  $\sigma$ -álgebra generada por los abiertos de  $M$ . Con esta noción podremos relacionar las propiedades topológicas con las de la Teoría de la Medida y nos será harto útil en nuestro recorrido. Trabajar con los conjuntos borelianos nos garantiza precisamente que los conjuntos abiertos son conjuntos medibles. Una medida  $\mu$ , definida sobre la familia de los borelianos es *positiva sobre abiertos* si  $\mu(A) > 0$ , para cualquier abierto  $A \neq \emptyset$ .

**La Medida de Lebesgue en el intervalo:** Supongamos que  $M$  es el intervalo  $[0, 1]$ . Considere entonces  $\mathcal{B}_0$  el álgebra de todos los subconjuntos de la forma  $B = I_1 \cup \dots \cup I_k$ , donde  $I_1, \dots, I_k$  son intervalos disjuntos por pares. Podemos definir fácilmente una medida para este tipo de subconjuntos:

$$\text{Leb}(B) = |I_1| + |I_2| + \dots + |I_k|$$

Donde  $|I|$  denota la longitud del intervalo. A esta medida se le conoce como la *medida de Lebesgue* del intervalo.

**Otras medidas en el intervalo:** Fije  $x \in [0, 1]$ . Así, definimos:

$$\delta_x(A) = 1 \text{ si } x \in A \text{ y } \delta_x(A) = 0 \text{ si no.}$$

A cada una de estas medidas se les conoce como la *delta de Dirac* soportada en  $x$ . De hecho, estas medidas están definidas en cualquier espacio medible.

**0.3. Espacios de sucesiones.** Consideremos el espacio de sucesiones infinitas de ceros y unos, esto es, puntos de la forma:

$$w = (\dots, w_{-n}, \dots, w_{-1}, w_0, w_1, \dots, w_n, \dots)$$

Este conjunto, formalmente, se describe como el producto cartesiano de infinitas (numerables) copias del conjunto  $\{0, 1\}$ , indicando cada una de sus coordenadas por un número entero:  $\Sigma_2 = \{0, 1\}^{\mathbb{Z}}$ . Cabe resaltar que si bien este conjunto es bastante parecido al conjunto de las secuencias de volados que definimos anteriormente, cada punto  $w \in \Sigma_2$  corresponde a tener dos secuencias de volados: una hacia el futuro  $w_i$  con  $i \geq 0$  y su historia  $w_i$  con  $i < 0$ .

Este conjunto es un espacio métrico junto con la siguiente función distancia:

$$d(x, y) = \frac{1}{2} \sum_{i=-\infty}^{\infty} \frac{|x_i - y_i|}{2^{|i|}}$$

Observe que la bola de radio  $\frac{1}{2^n}$  alrededor del punto  $w$  corresponde a todos los puntos  $x \in \Sigma_2$  cuyas coordenadas  $x_i = w_i$  para  $-n \geq i \geq n$ .

**Ejercicio:** Demuestre que  $(\Sigma_2, d)$  es un espacio métrico.

Ahora bien, al conjunto  $\Sigma_2$  le podemos dar una estructura de espacio medible. Para esto debemos escoger una familia de conjuntos medibles que contenga a las bolas abiertas y luego una medida.

Consideremos los siguientes subconjuntos de  $\Sigma_2$ : Dado un entero  $j \in \mathbb{Z}$  y un valor  $x \in \{0, 1\}$  el conjunto:

$$C(j, x) = \{w \in \Sigma_2 | w_j = x\}$$

es decir, todos los puntos de  $\Sigma_2$  que tienen en la coordenada  $j$  el valor  $x$ . De hecho  $\Sigma_2 = C(j, 0) \cup C(j, 1)$  para cualquier  $j \in \mathbb{Z}$ . Los conjuntos  $C(j, x)$  son los cilindros más sencillos. Un *cilindro* es un subconjunto de  $\Sigma_2$  de la siguiente forma: Dado un número entero  $i \in \mathbb{Z}$  y una sucesión  $\{x_0, x_2, \dots, x_k | x_i \in \{0, 1\}\}$  el cilindro:

$$C(i, \{x_0, x_2, \dots, x_k\}) = \{w \in \Sigma_2 | w_i + j = x_j, \text{ con } j \in \{0, \dots, k\}\}$$

La familia que contiene todas las intersecciones finitas y uniones de conjuntos  $C(i, x)$  es un álgebra  $\mathcal{B}_0$ : *álgebra de los cilindros*. Ésta, por su parte, genera una familia de subconjuntos medibles de  $\Sigma_2$  que denotaremos por  $\mathcal{B}$ . Vale notar que todas las bolas de radio  $2^{-n}$ , alrededor de cualquier punto  $w \in \Sigma_2$  están contenidas en  $\mathcal{B}$ , y por lo tanto, se trata de la  $\sigma$ -álgebra de Borel (los borelianos).

**La medida de Bernoulli en  $\Sigma_2$ :** Dados dos números reales positivos:  $p_0$  y  $p_1$ , tales que  $p_0 + p_1 = 1$ , podemos definir la siguiente función:

$$\mu_{p_0, p_1}: \mathcal{B}_0 \longrightarrow [0, 1]$$

de la siguiente manera

- (1)  $\mu(C(i, 0)) = p_0$  y  $\mu(C(i, 1)) = p_1$ .
- (2) Si  $i \neq j$ , entonces  $\mu(C(i, x) \cap C(j, y)) = \mu(C(i, x))\mu(C(j, y)) = p_x p_y$ ; donde  $x, y \in \{0, 1\}$ .
- (3) De hecho,  $\mu(C(i, \{x_0, x_2, \dots, x_k\})) = \prod_{j=0}^k p_j$

**Propiedad 2.** *El espacio  $(\Sigma_2, \mu_{p_0, p_1})$  es un espacio de probabilidad.*

*Demostración:* Por el teorema de extensión, basta definir la medida en  $\mathcal{B}_0$ . Ahora bien, dado  $i \in \mathbb{Z}$ , como  $C(i, 0) \cup C(i, 1) = \Sigma_2$  y como  $C(i, 0) \cap C(i, 1) = \emptyset$  entonces:

$$\mu(C(i, 0) \cup C(i, 1)) = \mu(C(i, 0)) + \mu(C(i, 1)) = p_0 + p_1 = 1$$

Por lo tanto  $(\Sigma_2, \mu_{p_0, p_1})$  es un espacio de probabilidad.  $\square$

**El espacio de sucesiones unilaterales:** Considere  $\Sigma_2^+ = \{0, 1\}^{\mathbb{N}}$ ; ahora sí, el conjunto de volados que definimos en la sección 0.1:

$$\Sigma_2^+ \ni w = (w_i)_{i \in \mathbb{N}} = (w_0, w_1, \dots, w_n, \dots)$$

Observe que el conjunto de Cantor ternario  $K$ , con la topología heredada del intervalo es homeomorfo a  $\Sigma_2^+$  con la topología definida por la función distancia

$$d(x, y) = \sum_{i=0}^{\infty} \frac{|x_i - y_i|}{2^i}$$

A través de un procedimiento idéntico al que utilizamos para  $\Sigma_2$ , podemos construir una familia de conjuntos medibles y una medida de Bernoulli para  $\Sigma_2^+$ . Con esta medida podemos decir, ahora formalmente, que la probabilidad de que en el primer tiro la moneda caiga águila es la medida del cilindro  $C(0, 0)$ :  $\mu(C(0, 0)) = p_0$ , y la probabilidad de que en el primer tiro caiga sol es la medida del cilindro  $C(0, 1)$ :  $\mu(C(0, 1)) = p_1$ . Ahora bien  $\mu(C(0, 0)) = \mu(C(i, 0)) = p_0$ . Esto quiere decir que la probabilidad que caiga águila en el turno  $i \in \mathbb{N}$  es también  $1/2$ , por lo tanto, no importa en qué momento, la probabilidad que caiga águila es siempre la misma. De alguna manera esto habla de fenómenos independientes. Por otro lado ¿cuál es la probabilidad que caiga primero águila y luego sol?

$$\mu(C(0, 0) \cap C(1, 1)) = \mu(C(0, 0))\mu(C(1, 1)) = p_0 p_1$$

Incluso podemos considerar espacios más grandes. En lugar de trabajar con sucesiones de ceros y unos, podemos reemplazar el conjunto de símbolos  $\{0, 1\}$  por uno más grande:  $\{0, 1, \dots, n-1\}$  y obtener así los espacios de sucesiones, bilaterales y unilaterales, en  $n$  símbolos. A estos espacios los denotaremos por:  $\Sigma_n$  y  $\Sigma_n^+$ , respectivamente. Para construir las medidas de Bernoulli en  $\Sigma_n$  y  $\Sigma_n^+$ , podemos repetir el mismo procedimiento, sólo que ahora será necesario un vector de probabilidad:  $(p_0, p_1, \dots, p_{n-1})$  donde  $p_0 + p_1 + \dots + p_{n-1} = 1$ .

## 1. MEDIDAS INVARIANTES

Consideremos un espacio de probabilidad  $(M, \mathcal{B}, \mu)$  y una transformación medible:

$$T: M \longrightarrow M$$

La medida  $\mu$  es *una medida invariante* por  $T$ , o bien  $T$  preserva la medida  $\mu$  si para cualquier subconjunto medible  $A \in \mathcal{B}$  se cumple lo siguiente:

$$\mu(T^{-1}(A)) = \mu(A)$$

esto es, la imagen inversa de cualquier conjunto medible, mide, según  $\mu$  exactamente lo mismo. Es equivalente decir que  $\mu$  es una medida invariante por  $T$ , o bien es  $T$ -invariante.

**Ejemplos:** Sea  $T$  una transformación medible  $T: M \rightarrow M$ , definida en cualquier espacio de medida  $M$ .

- La medida de Dirac, soportada en un punto fijo es invariante; esto es, si  $T(x) = x$ : la medida  $\delta_x$  es invariante por  $T$ .
- Si  $x$  es un punto periódico de período  $k$ , esto es:  $T^k(x) = x$  y  $T^j(x) \neq x$  si  $1 \leq j < k$ . La medida:

$$\sum_{j=0}^{k-1} \delta_{T^j(x)}$$

es invariante.

Cada transformación continua de un espacio topológico tiene, en efecto, al menos una medida invariante y estas dependen de la transformación y del espacio. Probar que siempre existe una medida invariante requiere conocimientos del análisis funcional. Por el momento nos conformaremos con algunos ejemplos.

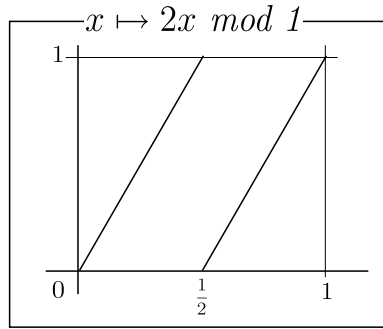


FIGURE 2. La transformación  $D(x)$

Considere al intervalo junto con la medida de Lebesgue, y la siguiente transformación

$$D: [0, 1] \rightarrow [0, 1]$$

$$D(x) = \begin{cases} 2x & \text{si } x \in [0, \frac{1}{2}] \\ 2x - 1 & \text{si } x \in [\frac{1}{2}, 1] \end{cases}$$

$$x \mapsto 2x \bmod 1$$

**Teorema 2.** *La medida de Lebesgue del intervalo es una medida invariante por  $D$ .*

*Demostración:* Observe que  $D$  tiene dos ramos de inversa:  $\varphi_0: I \rightarrow [0, \frac{1}{2}]$  y  $\varphi_1: I \rightarrow [\frac{1}{2}, 1]$ ; esto es  $D \circ \varphi_0 = D \circ \varphi_1 = \text{Id}$ . La derivada  $\varphi_0'(x) = \varphi_1'(x) = \frac{1}{2}$ , para todo  $x \in I$ . Utilizando el teorema del cambio de variable de integración podemos concluir que la medida de Lebesgue es invariante: Tome  $J = [a, b]$  cualquier intervalo. Observe que:

$$D^{-1}(J) = \varphi_0(J) \cup \varphi_1(J),$$

entonces:

$$\begin{aligned}
 |D^{-1}(J)| &= \int_{D^{-1}(J)} dx = \int_{\varphi_0(J)} dx + \int_{\varphi_1(J)} dx \\
 &= \int_a^b \varphi'_0(x) dx + \int_a^b \varphi'_0(x) dx \\
 &= \frac{1}{2} \left( \int_a^b dx + \int_a^b dx \right) \\
 &= \int_a^b dx = b - a = |J|
 \end{aligned}$$

y como cualquier boreliano de  $I$  lo podemos aproximar por una unión finita de intervalos, terminamos.  $\square$

**1.1. El Teorema de la Recurrencia de Poincaré.** Si bien ya nos preguntamos si la órbita de un punto converge o no a un punto, o si esta es periódica o no, también podemos considerar un punto en un subconjunto medible  $E \subset M$  y preguntarnos si su órbita regresa a  $E$  o no. Desde la perspectiva de la teoría de la medida, hablar de un punto es insignificante, pues en algunos casos, los puntos tienen medida cero. Una manera más acertada de plantear esta cuestión es determinar cuántos puntos de un conjunto  $E \subset M$  eventualmente regresan a  $E$ .

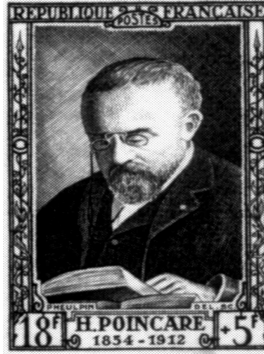


FIGURE 3. Jules Henri Poincaré

Esta pregunta apunta a nuestro primer resultado sorprendente de los sistemas dinámicos: El teorema de Recurrencia de Poincaré, en el cual intervienen conceptos de teoría de la medida como de topología. El primer paso es el siguiente teorema que nos indica cuántos puntos en efecto retornan infinitas veces a  $E$ , en términos de una medida invariante.

**Teorema 3.** *Sea  $T: M \rightarrow M$  una transformación medible y  $\mu$  una medida de probabilidad, invariante. Sea  $E \subset M$  cualquier conjunto medible ( $E \in \mathcal{B}$ ) con  $\mu(E) > 0$ , para  $\mu$ -casi todo punto  $x \in E$  existen infinitos valores de  $n \geq 1$  tales que  $T^n(x) \in E$ .*

En otras palabras, el teorema anterior afirma que la medida del conjunto de puntos  $E_0 \subset E$  que retornan infinitas veces a  $E$  es tal que  $\mu(E_0) = \mu(E)$ ; para cualquier subconjunto medible de  $M$ .

Para demostrar el Teorema 3, podemos reducirlo al siguiente Teorema, el cual resuelve el punto fundamental: Dado  $E \in \mathcal{B}$ ,  $\mu$ -casi todo punto regresa al menos una vez.

**Teorema 4.** *Bajo las mismas hipótesis del teorema 3, dado cualquier conjunto  $E \in \mathcal{B}$ , para  $\mu$ -casi todo punto  $x \in E$  existe  $n \geq 1$  tal que  $T^n(x) \in E$ ,  $n \geq 1$ .*

Supongamos por un momento que el teorema anterior es verdadero y probemos pues el Teorema 3, a partir de él.

*Demostración:* Para cada número entero  $k \geq 1$  podemos definir el siguiente conjunto:

$$E_k = \{x \in E \mid \#\{m \geq 1 \mid f^m(x) \in E\} = k\}$$

El conjunto  $E_k$  se refiere a los puntos en  $E$  que regresan exactamente  $k$  veces a  $E$  (¿es un conjunto medible?). Luego, el conjunto de puntos que regresan sólo una cantidad finita de veces lo podemos describir como:

$$\bigcup_{k=1}^{\infty} E_k$$

Para mostrar el corolario, tenemos que asegurarnos que  $\mu(E_k) = 0$ , para toda  $k \geq 1$ . Argumentaremos por contradicción, es decir, supongamos que  $\mu(E_k) > 0$  para algún  $k \geq 1$ .

Como  $\mu(E_k) > 0$ , podemos aplicar el Teorema 4 sobre el conjunto  $E_k$ , esto nos indica que  $\mu$ -casi todo punto  $x \in E_k$  regresa a  $E_k$ .

Tomemos un punto  $x \in E_k$  tal que existe  $n \geq 1$  tal que  $f^n(x) \in E_k$ . Llamemos  $y = f^n(x)$ . Por definición  $y$  tiene  $k$  iterados en  $E$ , esto es

$$\#\{m \geq 1 \mid f^m(y) \in E\} = k$$

Por lo tanto,  $x$  tiene  $k + 1$  iterados en  $E$ , pues  $f^n(x) = y$ , lo cual es una contradicción. Entonces,  $\mu(E_k) = 0$ , para cualquier  $k \geq 1$ ; y así no podemos aplicar el teorema 4.  $\square$

**Demostración del teorema 4:** Considere el conjunto de puntos en  $E$  que no regresan nunca a  $E$ , es decir:

$$E^0 = \{x \in E \mid f^n(x) \notin E, \forall n \geq 1\}$$

Queremos demostrar que  $\mu(E^0) = 0$ . Para esto, primero observaremos lo siguiente:

- Los subconjuntos  $\{E^0, f^{-1}(E^0), f^{-2}(E^0), \dots\}$  son disjuntos por pares.

De hecho, si existen  $m > n$  números naturales tales que:

$$f^{-n}(E^0) \cap f^{-m}(E^0) \neq \emptyset$$

entonces existe  $x \in f^{-n}(E^0) \cap f^{-m}(E^0)$ . Por lo tanto:

- $y = f^n(x) \in E^0$ , y también
- $f^{m-n}(y) = f^{m-n}(f^n(x)) = f^m(x) \in E^0$ .

Por lo tanto  $y \in E^0$  regresa en  $m - n > 0$  iterados a  $E$ . Una contradicción.

Ahora bien, puesto que los conjuntos  $\{f^{-n}(E^0)\}$  son disjuntos, podemos calcular su medida de la siguiente manera:

$$\begin{aligned} \mu\left(\bigcup_{n=0}^{\infty} f^{-n}(E^0)\right) &= \sum_{n=0}^{\infty} \mu(f^{-n}(E^0)) = \\ &= \sum_{n=0}^{\infty} \mu(E^0) < 1 \end{aligned}$$

Observe que la primera igualdad se dá porque los subconjuntos son disjuntos dos a dos. La segunda porque la medida es invariante bajo  $f$  y la tercera porque  $\bigcup_{n=0}^{\infty} f^{-n}(E^0) \subset M$ .

De modo que, para que una suma de infinitos términos iguales esté acotada, es necesario que  $\mu(E^0) = 0$ . Que es lo que queríamos demostrar.  $\square$

### Versión topológica del Teorema de Recurrencia de Poincaré.

Cuando la órbita de un punto  $x \in M$  regresa arbitrariamente cerca de sí mismo decimos que el punto  $x$  es *recurrente*. El hecho de un punto ser recurrente o no le concierne únicamente a la topología, pues, en resumidas cuentas, un punto  $x$  es recurrente si su órbita contiene una subsucesión que converge a  $x$ . Considere un espacio métrico compacto  $M$  y una transformación  $f: M \rightarrow M$  continua.

**Definición 6.** Decimos que un punto  $x$  es recurrente de  $f$  para sí para cualquier vecindad  $V$  de  $x$  existe algún iterado de  $x$  que pertenece a  $V$ :  $f^n(x) \in V$ ,  $n > 0$ .

Los puntos periódicos, por ejemplo, son recurrentes. De hecho, si  $p$  es tal que  $f^k(p) = p$ , para algún  $k \in \mathbb{N}$ , no importa cual vecindad  $U$  de  $p$  tomemos, es verdad que  $f^{nk}(p) \in U$ , para todo  $n \in \mathbb{N}$ . Sin embargo, es posible que una órbita no sea periódica y sí recurrente.

Los puntos recurrentes nos interesan porque son ellos los que cargan con toda la dinámica. Si  $x$  es un punto no recurrente, su dinámica es relativamente sencilla, de hecho, su órbita se aleja y quizás converge a un punto fijo o a una órbita periódica, o quizás se acumula en alguna región de  $M$  alejada de  $x$ . El conjunto de puntos donde la órbita de  $x$  se acumula en el futuro se llama  $\omega$ -límite de  $x$  ("omega límite") y se define de la siguiente manera:

$$\omega(x) = \{y \in M \mid \exists \{n_j\} \subset \mathbb{N} \text{ tal que } f^{n_j}(x) \rightarrow y \text{ cuando } j \rightarrow +\infty\}$$

Esto es, corresponde al conjunto de puntos a los cuales una subsucesión de la órbita de  $x$  converge. Como el espacio  $M$  es compacto, y entonces toda sucesión tiene una subsucesión convergente, podemos afirmar que  $\omega(x) \neq \emptyset$  para cualquier  $x \in M$ . ¿Es verdad que si  $y \in \omega(x)$  para algún  $x \in M$ , entonces  $y$  es recurrente? ¿si no, quién es  $\omega(y)$ ?

El siguiente resultado, enunciado por Poincaré a finales del siglo XIX, mezcla la noción de recurrencia y la noción de medida invariante y nos brinda un resultado algo sorprendente. En pocas palabras el Teorema de la Recurrencia de Poincaré afirma que *casi todo punto es recurrente*, con respecto a cualquier medida de probabilidad, invariante del sistema dinámico.

**Teorema 5.** Sea  $M$  un espacio métrico compacto y  $f: M \rightarrow M$  una transformación medible y  $\mu$  una medida de probabilidad invariante y positiva sobre abiertos; entonces,  $\mu$ -casi todo punto  $x \in M$  es recurrente para  $f$ .

*Demostración:* Sea  $U_k$  una base numerable de vecindades de  $M$ . Para cada  $k \in \mathbb{N}$  representaremos por  $U_k^0$  el conjunto de puntos  $x \in U_k$  que nunca regresan a  $U_k$ . De acuerdo con el teorema 3, cada  $U_k^0$  tiene medida cero. Consecuentemente, la unión numerable de todos ellos

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} U_k^0$$

tiene medida cero. Para demostrar el teorema, basta observar que todo punto que no pertenece a  $\tilde{U}$  es recurrente.

Sea  $x \in M - \tilde{U}$ . Sea también  $V$  una vecindad arbitraria de  $x$ . Como  $\{U_k \mid k \in \mathbb{N}\}$  es una base de vecindades, sabemos que existe algún  $k \in \mathbb{N}$  talque  $x \in U_k \subset V$ . Como  $x \notin \tilde{U}$ ,

en particular no pertenece a  $U_k^0$ ; es decir, existe algún  $n \in \mathbb{N}$  tal que  $f^n(x) \in U_k$ . Por lo tanto  $f^n(x) \in V$ . Puesto que  $V$  es una vecindad arbitraria, acabamos de demostrar que  $x$  es un punto recurrente.  $\square$

Si bien el teorema 3 es muy fuerte en lo que afirma, cabe recalcar que la afirmación es sobre *cualquier medida invariante* del sistema. Ésta puede decir mucho o no, de acuerdo a su naturaleza. Hemos visto que las medidas soportadas en una órbita periódica, inclusive en un punto fijo, son invariantes. Ejemplo  $x \mapsto \frac{x}{2}$  y  $\delta_0$  como medida invariante.  $\delta_0$ -casi todo punto significa sólo el 0. El resto de los puntos no son recurrentes.

Por otro lado, la aplicación  $D(x) = 2x \bmod 1$  preserva la medida de Lebesgue del intervalo. Por lo tanto, podemos concluir que: Lebesgue casi todo punto es recurrente.

En realidad, no es necesario garantizar que el espacio sea de probabilidad; esto es  $\mu(M) = 1$ . De hecho, basta con que la medida total del espacio sea finita para poder concluir el Teorema de Recurrencia de Poincaré, aunque eso no es una gran generalización, pues si definimos una nueva medida  $\rho(\cdot) := \frac{\mu(\cdot)}{\mu(M)}$  obtenemos rápidamente una medida de probabilidad. Por otro lado, si la medida del espacio es infinita, por ejemplo  $\mathbb{R}$  este teorema no es verdad. Un contraejemplo es el siguiente: Considere  $f: \mathbb{R} \rightarrow \mathbb{R}$  tal que  $f(x) = x + 1$ . Obviamente preserva la medida de Lebesgue y ningún punto es recurrente: Todos se escapan a  $+\infty$ .

**1.2. El shift o el dislocamiento.** Ahora si, introduciremos la transformación que nos atañe en el curso: *el shift o dislocamiento*. Consideremos el espacio  $\Sigma_2$  que definimos anteriormente. En él, podemos definir la siguiente transformación:

$$\sigma: \Sigma_2 \rightarrow \Sigma_2$$

$$(w_i)_{i \in \mathbb{Z}} \mapsto (w_{i-1})_{i \in \mathbb{Z}}$$

Lo podemos escribir mejor, si resaltamos con una  $*$  la casilla 0

$$\sigma(\dots, w_{-2}, w_{-1}, w_{-0}^*, w_1, w_2, \dots) = (\dots, w_{-1}, w_{-0}, w_1^*, w_2, w_3, \dots)$$

### Propiedades:

- a)  $\sigma: \Sigma_2 \rightarrow \Sigma_2$  es biyectivo
- b)  $\sigma$  preserva los cilindros, así que es una transformación medible.
- c) Es continua (ejercicio).
- d) Preserva la medida de Bernoulli,  $\mu = \mu_{p_0, p_1}$ , para cualesquiera números positivos  $p_0, p_1$  tales que  $p_0 + p_1 = 1$ .

Para demostrar la última propiedad necesitamos el siguiente lema que nos permite considerar únicamente el álgebra de los cilindros, pues esta genera a la familia de los subconjuntos medibles de  $\Sigma_2$ .

**Lema 2.** *Sea  $T: M \rightarrow M$  una transformación medible y  $\mu$  una medida en  $M$ . Supongamos que existe una sub-familia  $\mathcal{I}$  que genera la  $\sigma$ -álgebra de  $M$ , en la cual, para cualquier  $E \in \mathcal{I}$  se tiene que*

$$\mu(E) = \mu(T^{-1}(E))$$

*entonces, se vale lo mismo para todo subconjunto medible de  $M$ . Esto es, la medida  $\mu$  es invariante por  $T$ .*

Usando este lema, basta verificar que el shift preserva la medida en los cilindros y, como éstos generan toda la  $\sigma$ -álgebra de  $M$ , para concluir que el shift preserva cualquier medida  $\mu_{p_0, p_1}$ .

## 2. EL TEOREMA DE BIRKHOFF

El Teorema de Recurrencia de Poincaré sostiene que de acuerdo a cualquier medida invariante, casi todo punto regresa infinitas a un conjunto de medida positiva que lo contenga. La siguiente pregunta que podríamos formularnos es ¿Qué porcentaje del tiempo pasa en determinado conjunto? ¿podemos calcular ese número?

Tomemos cualquier conjunto medible  $E \subset M$  y un punto  $x \in E$ . El número:

$$\tau_n(E, x) = \frac{\#\{j \in \{0, 1, \dots, n-1\} \mid f^j(x) \in E\}}{n}$$

nos indica el promedio de las veces que  $f^j(x)$  cae en  $E$  en los primeros  $n$  iterados. Otra manera de calcular este número es mediante una fórmula:

$$(1) \quad \tau_n(x, E) = \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x))$$

donde  $\chi_E$  denota la función característica del conjunto  $E$ . La *función característica* de  $E \in \mathcal{B}$  es una función medible que se define:  $\chi_E(x) = 1$  si  $x \in E$  y  $\chi_E(x) = 0$  si  $x \notin E$ .

El límite de la expresión (1), cuando  $n$  tiende a infinito, lo llamaremos *el tiempo medio de estadía* de la órbita de  $x$  en el conjunto  $E$ :

$$\tau(x, E) = \lim_{n \rightarrow \infty} \tau_n(x, E)$$

En principio, este límite puede existir o no, dependiendo del punto  $x$  y del conjunto  $E$ . El Teorema Ergódico de Birkhoff sostiene, precisamente, que dicho límite existe para  $\mu$ -casi todo punto; como siempre, de acuerdo a cualquier medida  $\mu$  invariante.

**Teorema 6.** (*Birkhoff, version 1*) Sea  $f: M \rightarrow M$  una transformación medible y  $\mu$  una probabilidad invariante bajo  $f$ . Dado cualquier conjunto medible  $E \subset M$ , el tiempo de permanencia  $\tau(x, E)$  existe para  $\mu$ -casi todo punto  $x \in M$ ; además:

$$\int \tau(x, E) d\mu(x) = \mu(E)$$

La demostración del teorema de Birkhoff, en esta nuestra primera lectura de los sistemas dinámicos carece de importancia. Una excelente referencia para esto es el libro de Ricardo Mañé [1] o las notas de Marcelo Viana [3]. Por el momento nos contentaremos con entender el enunciado y alcanzar algunas de sus consecuencias.

Una manera de obtener información de un determinado fenómeno es como lo hacen los físicos: a través de mediciones. En cada punto del espacio podemos medir valores como la temperatura, la presión, etcétera. Estas mediciones no son otra cosa sino números, o mejor dicho funciones  $\varphi: M \rightarrow \mathbb{R}$  que a cada punto del espacio le asignan un valor real. A estas funciones se les llama *observables* y la información que podamos obtener sobre estas funciones en general, es información sobre nuestro sistema dinámico.

**Definición 7.** Una función medible  $\varphi: M \rightarrow \mathbb{R}$  es invariante si  $\varphi(x) = \varphi(f(x))$ , para toda  $x \in M$ .

Las funciones invariantes asumen el mismo valor para cada elemento de la órbita.

**Propiedad 3.** La función  $\tau(x, E)$  es una función invariante, esto es:  $\tau(x, E) = \tau(f(x), E)$

*Demostración:*

$$\begin{aligned}\tau(f(x), E) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \chi_E(f^j(x)) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_E(f^j(x)) - \frac{1}{n} [\chi_E(x) - \chi_E(f^n(x))] \\ &= \tau(x, E) - \lim_{n \rightarrow \infty} \frac{1}{n} [\chi_E(x) - \chi_E(f^n(x))]\end{aligned}$$

Como  $\chi(x) \leq 1$ , para toda  $x \in M$ , el último límite es igual a 0. □

A partir de este resultado, podemos concluir un resultado más fuerte:

**Teorema 7.** (Birkhoff, version 2) Sea  $f: M \rightarrow M$  una transformación medible y  $\mu$  una probabilidad invariante por  $f$ . Dada cualquier función integrable  $\varphi: M \rightarrow \mathbb{R}$ , el límite

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

existe, para  $\mu$ -casi todo punto  $x \in M$ . Además,

$$\int \tilde{\varphi}(x) d\mu(x) = \int \varphi(x) d\mu(x)$$

### Observaciones:

- Observe que si  $\varphi$  es la función característica de un conjunto medible  $E \subset M$ , el enunciado es el mismo que el del Teorema 6.
- Como cualquier función medible  $\varphi: M \rightarrow \mathbb{R}$  la podemos aproximar por una suma finita de funciones características, en realidad, el toda la fuerza del teorema de Birkhoff está concentrada en el Teorema 6.

**Ejercicio:** Muestre que para cualquier función integrable, la media temporal  $\tilde{\varphi}$  satisface que  $\tilde{\varphi} \circ f = \tilde{\varphi}$  en  $\mu$ -casi todo punto; esto es, la media temporal es invariante por  $f$ .

**2.1. Ergodicidad.** La palabra *ergodico* proviene de la yuxtaposición de dos vocablos latinos: *ergos* y *odos*. *Ergos* significa trabajo; *odos* significa camino. Fue el físico Ludwig Boltzman, en el siglo XIX, quién introdujo por primera vez este término. En esa época él, así como también Maxwell y Gibbs, estudiaban la teoría cinética de los gases, a partir de un sistema dinámico (un flujo). Boltzman en particular creía en la hipótesis ergódica, la cual afirma que las órbitas típicas de dicho sistema dinámico llenaban por completo cada superficies de energía constante; y entonces el promedio de los observables (la integral) coincide con las medias temporales (las medias de Birkhoff):

Desde el punto de vista matemático, podemos reformular esta hipótesis en un nuevo concepto: la ergodicidad.

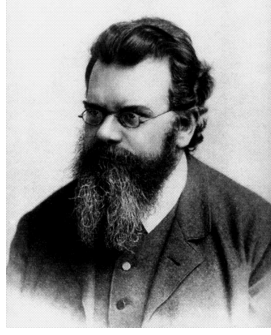


FIGURE 4. Ludwig Boltzmann

**Definición 8.** Una transformación  $f: M \rightarrow M$  medible es ergódica para una probabilidad invariante  $\mu$  ( $\mu$  es una medida ergódica de  $f$ , o bien, el sistema  $(f, \mu)$  es ergódico) si las medias temporales coinciden en  $\mu$  casi todo punto con las medias espaciales, esto es:

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

para cualquier función integrable  $\varphi: M \rightarrow \mathbb{R}$ , en  $\mu$  casi todo punto  $x \in M$ .

Esto por el lado de los observables, pero la ergodicidad también atañe a los conjuntos formados por órbitas completas del sistema: los conjuntos invariantes.

**Definición 9.** Un conjunto medible  $A \subset M$  es invariante si  $f^{-1}(A) = A$ .

La siguiente proposición relaciona las medidas ergódicas con los conjuntos invariantes.

**Propiedad 4.** Sea  $f: M \rightarrow M$  una aplicación medible y  $\mu$  una probabilidad invariante por  $f$ , entonces las siguientes afirmaciones son equivalentes:

- (1) El sistema  $(f, \mu)$  es ergódico.
- (2) Cualquier subconjunto invariante  $A$  cumple que:  $\mu(A) = 1$  o  $\mu(A) = 0$ .
- (3) Toda función  $\psi: M \rightarrow \mathbb{R}$  invariante es constante en un conjunto de medida total.

El hecho de que si  $\mu$  es ergódica y  $A \subset M$  un subconjunto invariante, desde el punto de vista de la medida  $A$  es todo el espacio o es un conjunto insignificante. Dicho de otra manera, no podemos descomponer el espacio en dos regiones donde la dinámica sea independiente y que ambas partes tengan medida positiva.

*Demostración:*

[1  $\Rightarrow$  2] Sea  $A$  un conjunto invariante y considere  $\varphi(x) = \chi_A(x)$ . Por un lado, la hipótesis implica que

$$\tilde{\varphi}(x) = \int \varphi d\mu = \mu(A)$$

para  $\mu$ -casi todo punto  $x \in M$ . Ahora bien,

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

Observe que  $x \in A$  sí y sólo si  $f(x) \in A$ , esto implica que  $\varphi(f^j(x)) = \varphi(x)$ , para todo  $j \in \mathbb{N}$ . Por lo tanto

$$\tilde{\varphi}(x) = \varphi(x) = \chi_A(x)$$

para todo  $x \in M$ . Sin embargo, como la función característica sólo toma valores 0 o 1, las únicas posibilidades para  $\mu(A)$  son 0 ó 1

[2  $\Rightarrow$  3] Sea  $\psi: M \rightarrow \mathbb{R}$  una función invariante. Tome  $I \subset \mathbb{R}$  cualquier intervalo. Observe que  $\psi^{-1}(I) = \{x \in M | \psi(x) \in I\}$  es un conjunto invariante. Por hipótesis,  $\psi^{-1}(I)$  tiene medida 0 ó 1. Como el intervalo  $I$  es arbitrario, esto demuestra que  $\psi$  es constante en un conjunto de medida total.

[3  $\Rightarrow$  1] Sea  $\varphi$  una función integrable cualquiera. Observe que  $\tilde{\varphi}$  es una función invariante por  $f$  ( $\tilde{\varphi} \circ f = \tilde{\varphi}$ ). Por hipótesis,  $\tilde{\varphi}$  es constante en  $\mu$ -casi todo punto. Usando el Teorema Ergódico de Birkhoff obtenemos que:

$$\tilde{\varphi}(x) = \int \tilde{\varphi} d\mu = \int \varphi d\mu$$

para  $\mu$  casi todo punto  $x \in M$ . □

Ahora demostraremos el siguiente teorema.

**Teorema 8.** *El shift bilateral es ergódico.*

Para probar esta afirmación necesitamos el siguiente lema:

**Lema 3.** *Si  $A$  y  $B$  son uniones finitas de cilindros disjuntos. Entonces existe  $m \in \mathbb{N}$  tal que:*

$$\mu(A \cap \sigma^{-m}(B)) = \mu(A)\mu(\sigma^{-m}(B)) = \mu(A)\mu(B)$$

*Demostración del lema:* Veamos qué sucede cuando se trata de cilindros. Un cilindro es de la forma:

$$A = C(i, \{x_0, x_2, \dots, x_k\}) = \{w \in \Sigma_2 | w_i = x_0, w_{i+1} = x_1, \dots, w_{i+k} = x_k\}$$

Observe además que para cada  $n \in \mathbb{N}$  se tiene que

$$\sigma^{-n}(C(i, \{x_0, x_2, \dots, x_k\})) = C(i+n, \{x_0, x_2, \dots, x_k\})$$

Si  $B = C(j, \{y_0, \dots, y_l\})$  tomando  $m$  tal que  $i+m > j+l$  los cilindros  $A$  y  $\sigma^{-m}(B)$  son independientes, y por lo tanto

$$\mu(A \cap \sigma^{-m}(B)) = \mu(A)\mu(\sigma^{-m}(B)) = \mu(A)\mu(B). \quad \square$$

**Demostración de la Proposición:** Dado un conjunto invariante  $A$ , ( $\sigma^{-1}(A) = A$ ), queremos probar que  $\mu(A) = 0$  ó  $\mu(A) = 1$ . Para clarificar el argumento, supongamos primero que  $A$  está en el álgebra de los cilindros (como en el lema anterior). Entonces, existe  $m \in \mathbb{N}$  tal que:

$$\mu(A \cap \sigma^{-m}(A)) = \mu(A)\mu(\sigma^{-m}(A)) = \mu(A)^2$$

Puesto que  $A$  es invariante. Esto sólo es posible cuando  $\mu(A) = 0$  ó 1.

Ahora tomemos cualquier conjunto invariante  $A \in \mathcal{B}$ . Lo que haremos será aproximar  $A$  por elementos del álgebra de los cilindros  $\mathcal{B}_0$ ; recuerde el Teorema de aproximación (Sección 1) que afirma que, dado  $\varepsilon > 0$ , existe  $A_0 \in \mathcal{B}_0$ , tal que  $\mu(A \Delta A_0) < \varepsilon$ .

Tomando  $m \in \mathbb{N}$  del lema anterior, aplicado en  $A_0$ , tenemos que:

$$(2) \quad \mu(A_0 \cap \sigma^{-m}(A_0)) = \mu(A_0)\mu(\sigma^{-m}(A_0)) = \mu(A_0)^2$$

Observe que

$$\begin{aligned} (A \cap \sigma^{-m}(A)) \Delta (A_0 \cap \sigma^{-m}(A_0)) &\subset (A \Delta A_0) \cup (\sigma^{-m}(A) \Delta \sigma^{-m}(A_0)) \\ &\subset (A \Delta A_0) \cup f^{-m}(A \Delta A_0) \end{aligned}$$

Esto implica, junto con el hecho de que  $\mu$  es invariante bajo  $\sigma$ , que:

$$(3) \quad |\mu(A \cap \sigma^{-m}(A)) - \mu(A_0 \cap \sigma^{-m}(A_0))| \leq 2\mu(A \Delta A_0) < 2\varepsilon$$

Por otro lado

$$(4) \quad |\mu(A)^2 - \mu(A_0)^2| \leq 2|\mu(A) - \mu(A_0)| < 2\varepsilon$$

Juntando las ecuaciones en (2), (3), (4) y recordando que  $A$  es un conjunto invariante, obtenemos que

$$|\mu(A) - \mu(A)^2| < 4\varepsilon$$

Como esto vale para  $\varepsilon > 0$  arbitrario concluimos que  $\mu(A) = \mu(A)^2$ , es decir  $\mu(A) = 0$  ó 1.  $\square$

### 3. DINÁMICA TOPOLOGICA

**3.1. El shift unilateral.** Habíamos definido el espacio  $\Sigma_2^+$  como el espacio de todas las sucesiones unilaterales de  $\{0, 1\}$ . Vimos además que se trata de un espacio de medida y también le construimos ciertas medidas  $\mu_{p_0 p_1}$ . En este espacio también podemos definir una transformación semejante al shift.

$$\sigma: \Sigma_2^+ \longrightarrow \Sigma_2^+$$

$$\sigma((x_1, x_2, x_3, \dots)) = (x_2, x_3, x_4, \dots)$$

Observe que esta transformación es continua, medible y preserva las medidas  $\mu_{p_0, p_1}$ . Observe también que no es biyectiva, sino 2 a 1.

**Definición:** Un punto  $w \in \Sigma_2^+$  es un punto periódico de período  $k$  si  $\sigma^k(w) = w$ .

**Propiedad 5.** Dado  $k \in \mathbb{N}$ , existe al menos una órbita periódica de período  $k$  para  $(\Sigma_2^+, \sigma)$ .

*Demostración:* Una palabra finita de longitud  $m \in \mathbb{N}$  con símbolos en  $\{0, 1\}$  es un elemento de  $\{0, 1\}^m$ . Obviamente, una palabra  $w$  de longitud  $k$  define un punto en  $\Sigma_2^+$  periódico:

$$x = (w_0, \dots, w_{k-1}, w_0, \dots, w_{k-1}, \dots)$$

$\square$

Otra información fácil de obtener es la siguiente proposición.

**Propiedad 6.** El conjunto de órbitas periódicas de  $(\Sigma_2^+, \sigma)$  es denso en  $\Sigma_2^+$ .

*Demostración:* Sabemos que dos puntos están una distancia menor que  $2^{-k}$  si coinciden en sus primeros  $k$  términos:  $x_0, \dots, x_{k-1}$ . Estos  $k$  términos definen una palabra de longitud  $k$ , que por su parte, repitiendo infinitas veces la palabra, esta define un punto periódico de período  $k$ .  $\square$

**Propiedad 7.** Existe un punto  $x \in \Sigma_2^+$  tal que su órbita es densa en  $\Sigma_2^+$ ; esto es

$$\overline{\{x, \sigma(x), \sigma^2(x), \dots\}} = \Sigma_2^+$$

*Demostración:* Para esto, escribámos todas las palabras de longitud 1 como  $w_0^1, w_1^1$ , e inductivamente las de longitud  $m \geq 1$  como:

$$w_0^m, w_1^m, \dots, w_{s^m}^m.$$

No es difícil verificar que la órbita del punto definido por la concatenación de todas las palabras finitas:

$$x = (w_0^1 w_1^1 \dots w_0^m w_1^m \dots w_{s^m}^m \dots)$$

es densa en  $\Sigma_2^+$ .  $\square$

Esta última propiedad nos advierte que la dinámica del shift unilateral no se puede descomponer en dos subconjuntos ajenos, cerrados e invariantes.

**Propiedad 8.** *El shift unilateral es ergódico.*

*Demostración:* El mismo argumento que prueba que el shift bilateral es ergódico.  $\square$

**3.2. ¿Qué relación hay con la aplicación  $D : 2x \bmod 1$ ?** Para comparar dos sistemas dinámicos necesitamos establecer cuándo estos son iguales o no. Como lo que nos interesa son las órbitas, una noción adecuada de igualdad debería garantizar que ambos sistemas tienen las mismas órbitas, módulo un cambio de coordenadas. La noción de *conjugación* es la adecuada para comparar sistemas dinámicos. Seamos más precisos.

**Definición 10.** Sean  $S : M \rightarrow M$  y  $T : N \rightarrow N$  dos transformaciones definidas en dos espacios topológicos, decimos que ambas son conjugadas si existe un homeomorfismo  $h : M \rightarrow N$  tal que transforma órbitas de  $S$  en órbitas de  $T$ . Esto es:

$$h \circ S = T \circ h$$

A dicho homeomorfismo se le llama una *conjugación*. Observe que una conjugación respeta órbitas periódicas.

**Teorema 9.** *Existe una conjugación en Lebesgue-casi todo punto entre el shift unilateral y la aplicación  $D : x \mapsto 2x \bmod 1$ .*

Para definir una conjugación entre ambos sistemas dinámicos, comencemos por dar direcciones a los puntos del intervalo: Denote por  $\varphi_0$  y  $\varphi_1$  las dos ramas de  $D^{-1}$ , de manera que  $\text{Img}(\varphi_0) = I_0 = [0, \frac{1}{2}]$  y  $\text{Img}(\varphi_1) = I_1 = [\frac{1}{2}, 1]$ .

De esta manera, dada una palabra finita de longitud  $m$ :  $w^m = (x_0, x_1, \dots, x_m)$ ,  $x_i \in \{0, 1\}$ , podemos definir un conjunto:

$$I(x_0, x_1, \dots, x_m) = \varphi_{x_0} \circ \varphi_{x_1} \circ \dots \circ \varphi_{x_m}(I)$$

Observe que el conjunto  $I(w^m) = I(x_0, x_1, \dots, x_m)$  es un intervalo, para cualquier palabra.

Cualquier punto  $x \in I(x_0, x_1, \dots, x_m)$  tiene la siguiente propiedad:

$$D^n(x) \in I_{x_n} \text{ para cada } n \in \{0, \dots, m\}$$

esto es, la palabra  $w$  es el itinerario de todos los puntos en  $I(w)$ . Observe también que  $|I(w^m)| \rightarrow 0$ , cuando  $m \rightarrow \infty$ . Como son intervalos cerrados y anidados, podemos concluir que:

$$\lim_{m \rightarrow \infty} I(w^m) = \bigcap_{m \rightarrow \infty} I(w^m) = \{x\}$$

Observe que algunos puntos racionales tienen doble etiqueta, por ejemplo: considere las siguientes dos sucesiones en  $\Sigma_2^+$ :

$$w = (0, 1, \bar{0}) \text{ y } v = (0, 0, \bar{1}).$$

Si bien  $w$  y  $v$  son dos itinerarios distintos, si denotamos por  $w^m$  las primeras  $m$  letras de  $w$  y por  $v^m$  las primeras las de  $v$  tenemos que:

$$I(w^m) \rightarrow \frac{1}{4} \text{ y } I(v^m) \rightarrow \frac{1}{4}.$$

Esto sucede solamente con los puntos de la forma  $\frac{r}{2^s}$  y carece de importancia, debido a que la medida de Lebesgue de un conjunto numerable es 0.

Una vez hecho esto, podemos definir una función  $h: \Sigma_2^+ \rightarrow I$  de manera que a cada  $w \in \Sigma_2^+$  le asocia  $\lim_{m \rightarrow \infty} I(w^m)$ .

**Ejercicio:** Fuera de los puntos  $\{\frac{r}{2^s} | r \leq 2^s, r, s \in \mathbb{N}\}$  la función  $h$  es biyectiva, continua y su inversa también es continua;  $h$  es un homeomorfismo.

No es difícil convencerse de que esta función conjuga las dos dinámicas. Y por lo tanto, podemos concluir que el sistema dinámico definido por

$$D: [0, 1] \rightarrow [0, 1]$$

$$x \mapsto 2x \text{ mod } 1$$

tiene órbitas periódicas de todos los períodos, así como también tiene una órbita densa.

**3.3. Las órbitas densas son típicas.** Volvamos al shift bilateral. Ya vimos que hay órbitas periódicas de todos los períodos. ¿Cuáles son las órbitas típicas? Para responder esta pregunta, consideremos un caso más general. Supongamos que  $M$  es un espacio métrico compacto y  $f: M \rightarrow M$  un homeomorfismo.

Dado un subconjunto  $A \subset M$  denote por  $\bar{A}$  su cerradura, esto es, el conjunto de todos los puntos de acumulación de  $A$ . Si  $A$  es denso en  $M$ , entonces  $\bar{A} = M$ .

**Teorema 10.** *Sea  $M$  un espacio métrico compacto y  $\mathcal{B}$  la  $\sigma$ -álgebra de Borel. Sea también  $f: M \rightarrow M$  un homeomorfismo y  $\mu$  una medida de Borel positiva sobre abiertos. Si  $f$  es ergódica con respecto a  $\mu$ , entonces*

$$\mu(\{x \in M | \overline{\text{órbita}(x)} = M\}) = 1$$

*En particular, existe una órbita densa.*

*Demostración:* Sea  $\{U_i | i \in \mathbb{N}\}$  una base numerable de la topología de  $M$ . Observe que

$$\{x \in M | \overline{\text{órbita}(x)} = M\} = \bigcap_{k=1}^{\infty} \bigcup_{n=-\infty}^{\infty} T^n(U_k)$$

De hecho, la órbita de  $x$  es densa, sí y sólo si para cada  $U_k$  existe  $n \in \mathbb{Z}$  tal que  $f^n(x) \in U_k$ . Ahora bien, observe que para  $k$  fijo, el conjunto

$$\bigcup_{n=-\infty}^{\infty} T^n(U_k)$$

es invariante. La ergodicidad implica que:

$$\mu \left( \bigcup_{n=-\infty}^{\infty} T^n(U_k) \right) = 0 \text{ ó } 1$$

Sin embargo,  $U_k \subset \bigcup_n T^n(U_k)$ . Por lo tanto,  $\mu(\bigcup_n T^n(U_k)) = 1$ .  $\square$

Con este teorema podemos concluir que en el shift bilateral, casi todas las órbitas son densas.

**Ejercicio:** Demuestre lo mismo para la aplicación  $\sigma : \Sigma_2^+ \rightarrow \Sigma_2^+$ .

**3.4. Los subshifts de tipo finito.** Dentro de  $\Sigma_k$  podemos encontrar subconjuntos invariantes. Consideremos una matriz

$$A = (a_{ij}) \in \mathcal{M}_{k \times k}$$

cuyos coeficientes  $a_{ij} \in \{0, 1\}$ , para toda  $i, j$ . Definamos entonces

$$\Sigma_2(A) = \{(x_n) \in \Sigma_k \mid a_{x_n x_{n+1}} = 1\}$$

- $\sigma(\Sigma_k(A)) = \Sigma_k(A)$ , es decir, se trata de un subconjunto invariante; además,  $\sigma|_{\Sigma_k(A)}$  es biyectiva y es un homeomorfismo.
- $\Sigma_k(A)$  es un conjunto cerrado, pues su complemento es abierto en  $\Sigma_2$ .
- Si  $a_{ij} = 1$ , para toda  $i, j \in \{1, \dots, k\}$ , entonces  $\Sigma_k(A) = \Sigma_k$ .
- Si  $A$  es la matriz identidad,  $\Sigma_k(A)$  consta únicamente de  $k$  puntos. ¿cuáles?
- También  $\Sigma_k(A)$  puede ser vacío e incluso dos matrices diferentes pueden dar el mismo subconjunto. Ejemplo en  $\Sigma_2$ , las matrices:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Sea  $k \in \mathbb{N}$ . Otra manera de estudiar los subshifts de tipo finito  $\Sigma_k(A)$  es utilizando una gráfica dirigida. Esta consiste de  $k$  vértices llamados  $v_1, v_2, \dots, v_k$ , y tiene la arista  $(v_i, v_j)$  si y sólo si el elemento el coeficiente de la matriz  $a_{ij} = 1$ .

Para ver un ejemplo de esto consideremos la matriz:

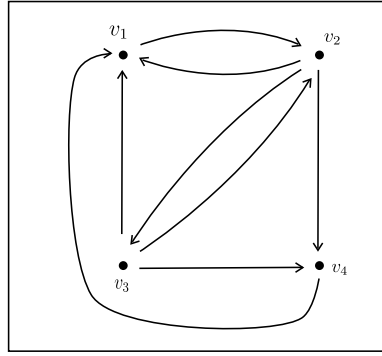
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

El conjunto  $\Sigma_k(A)$  está representado por los caminos posibles dentro de esta gráfica, a partir de un punto dado. Observe que son caminos con "futuro" (en el sentido de las flechas) e "historia" (en sentido contrario). Cada camino de estos representa un punto en  $\Sigma_k(A)$ .

- Observe que si la matriz  $A$  tiene todos sus coeficientes = 1, entonces la gráfica correspondiente es la que contiene todas las aristas dirigidas posibles.

Una propiedad interesante que nos permitirá obtener algunas propiedades dinámicas de este tipo de sistemas está enunciada en el siguiente lema:

**Lema 4.** *Dados cualesquiera  $i, j \in \{0, \dots, k-1\}$  si denotamos por  $N_{ij}^m$  el número de caminos admisibles que comienzan en  $v_i$ , terminan en  $v_j$  y de longitud  $m$ . Si denotamos*


 FIGURE 5. Grafica dirigida asociada a la matriz  $A$ .

también  $a_{ij}^m$  la entrada  $i, j$  de la matriz  $A^m$ , entonces:

$$N_{ij}^m = a_{ij}^m$$

*Demostración:* Primero mostraremos la siguiente fórmula:

$$N_{ij}^{m+1} = \sum_{s=0}^{k-1} N_{is}^m a_{sj}$$

Para esto, observe que si tomamos  $s \in \{0, \dots, k-1\}$ , cualquier camino de longitud  $m+1$  que une  $v_i$  a  $v_s$  produce un camino de longitud  $m+2$  que une  $v_i$  a  $v_j$  siempre que  $a_{sj} = 1$ , simplemente agregando al final  $v_j$ . Esto prueba la fórmula.

Ahora bien, demostraremos el lema por inducción: Para  $m = 1$ , es decir,  $N_{ij}^1 = a_{ij}^1$ , es verdad por definición. Así que podemos suponer que  $N_{ij}^m = a_{ij}^m$ . Pero la fórmula que probamos, no es más que la fórmula que nos indica la entrada  $i, j$  de la matriz  $(N_{ij}^m) \circ (a_{ij})$ .  $\square$

A partir de un vértice dado cada camino cerrado de longitud  $m+1$  nos indica una órbita periódica de periodo  $m$ . Esto nos muestra el siguiente corolario:

### Corolario

$$\#\{\text{órbitas periódicas de periodo } m \text{ de } \sigma_A\} = \text{traza}(A^m)$$

**3.5. La transformación de Gauss.** Consideremos la siguiente transformación del intervalo:

$$G: (0, 1] \longrightarrow (0, 1]$$

$$G(x) = \frac{1}{x} - \left[ \frac{1}{x} \right]$$

Donde  $[y]$  denota la parte entera de  $y$ . Observe que  $G$  no está definida en  $x = 0$ .

**Ejercicio:** Dibuje la gráfica de  $G$ , considerando que si  $x \in (\frac{1}{k+1}, \frac{1}{k})$ , entonces,  $G(x) = x^{-1} - k$ .

Vamos a restringir el dominio de  $G$  a  $[0, 1] \cap (\mathbb{R} \setminus \mathbb{Q})$ , es decir, a los puntos irracionales en el intervalo. Esto con el fin de evitar problemas de definición, pues, por ejemplo  $G^2(1/6)$  tampoco está definido. Una vez más, esto no representa un problema, pues los número racionales del intervalo tienen medida de Lebesgue 0.

Una de las características interesantes de esta transformación es preserva una medida equivalente a la medida de Lebesgue.

**Definición:** Dos medidas  $\mu$  y  $\nu$  son equivalentes si tienen los mismos conjuntos de medida cero; esto es, si  $A$  es un conjunto medible, entonces

$$\mu(A) = 0 \Leftrightarrow \nu(A) = 0$$

De hecho, la medida invariante por la transformación de Gauss se define de la siguiente manera

$$\mu(E) = \int_E \frac{c}{1+x} dx$$

para cualquier conjunto medible  $E$  y para alguna constante  $c > 0$ . Note que la integral está bien definida, pues el integrando es una función continua en el intervalo. De hecho, se cumple la siguiente desigualdad, para cualquier conjunto medible:

$$\frac{c}{2} \text{leb}(E) \leq \mu(E) \leq c \text{leb}(E).$$

Por lo tanto,  $\mu$  es equivalente a la medida de Lebesgue.

**Propiedad 9.**  $\mu$  es una medida  $G$ -invariante.

Para demostrar esta proposición apelaremos al siguiente teorema:

**Teorema** Denote por  $I = [0, 1]$ . Sea  $f: I \rightarrow I$  una función derivable, cuya derivada es continua y tal que  $f(x) \neq 0$  para toda  $x \in I$ . Sea también  $\rho: I \rightarrow \mathbb{R}$  una función continua, entonces la medida  $\mu = \rho \text{leb}$  es  $f$ -invariante si y sólo si

$$(5) \quad \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|f'(x)|} = \rho(y); \forall y \in I$$

Su demostración se obtiene de utilizar el Teorema de cambio de variable de integración.

*Demostración de la Proposición:* En vistas del lema anterior, queremos demostrar que la igualdad en (5) es válida para  $\rho = \frac{c}{1+x}$ . Para esto, primero observe que, dado un número  $y \in I$ , este tiene una única preimagen en cada intervalo de la forma  $(\frac{1}{k+1}, \frac{1}{k})$ . Sea  $x_k$  esta preimagen. De hecho

$$G(x_k) = \frac{1}{x_k} - k = y.$$

O lo que es lo mismo:  $x_k = \frac{1}{y+k}$ . Por otro lado  $G'(x) = -\frac{1}{x^2}$ . Por lo tanto, podemos reescribir la ecuación en 5 de la siguiente manera:

$$(5) \Leftrightarrow \sum_{k=1}^{\infty} \frac{cx_k^2}{1+x_k} = \frac{x}{1+y}$$

$$\Leftrightarrow \sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \frac{x}{1+y}$$

No es difícil verificar que la última igualdad es cierta debido a que

$$\frac{1}{(y+k)(y+k+1)} = \frac{1}{y+k} - \frac{1}{y+k+1}$$

y por lo tanto se trata de una suma telescópica.

Finalmente, integrando la función primitiva de  $\rho$  podemos calcular:

$$\mu([0, 1]) = \int_0^1 \frac{c}{1+x} dx = c \log(2).$$

Así, tomando  $c = (\log(2))^{-1}$ , obtenemos que  $\mu$  es una probabilidad.

3.5.1. *Fraciones continuas.* Considere  $x_1 \in (0, 1)$  y denote por  $a_1 = [1/x_1]$ . Tome entonces  $x_1 = G(x_0) = \frac{1}{x_0} - a_1$ . De esta manera  $a_1 \in \mathbb{N}$  y  $x_1 \in I$ . Además:

$$x_0 = \frac{1}{a_1 + x_1}.$$

Si  $x_1 \neq 0$ , podemos repetir el proceso:  $a_2 = [1/x_1]$  y  $x_2 = G(x_1) = \frac{1}{x_1} - a_2$ . De esta manera obtenemos que:

$$x_1 = \frac{1}{a_2 + x_2} \text{ y también } x_0 = \frac{1}{a_1 + \frac{1}{a_2 + x_2}}$$

Para cada  $n \geq 1$  tal que  $x_{n-1} \neq 0$  definimos  $a_n = [1/x_{n-1}]$  y  $x_n = G(x_{n-1})$ . De esta manera obtenemos que:

$$x_0 = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n + x_n}}}}$$

**Ejercicio:** La sucesión definida por

$$z_n = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}}$$

converge a  $x_0$  cuando  $n \rightarrow \infty$ .

**Observación:** Los números  $z_n \in \mathbb{Q}$  son los números racionales que mejor aproximan a  $x_0$ , en el sentido de que  $z_n$  es el racional más cercano a  $x_0$  con el denominador menor o igual a  $q_n \in \mathbb{N}$ , donde  $z_n = \frac{p_n}{q_n}$  y  $(p_n, q_n) = 1$ . Es decir, si  $|\frac{l}{k} - \alpha| < |\frac{p_n}{q_n} - \alpha|$ , entonces  $q_n < k$ .

Ahora bien, si  $x_0 \notin \mathbb{Q}$ , entonces  $x_n \neq 0$  para toda  $n \in \mathbb{N}$ . De esta manera, podemos asociar a cada  $x_0$  la sucesión infinita de números naturales que corresponde a la expresión de  $x_0$  en fracciones continuas:

$$x_0 \mapsto [a_1, a_2, \dots, a_n, \dots] \in \mathbb{N}^{\mathbb{N}}$$

Observe que:

$$G(x_0) = G([a_1, a_2, a_3, \dots]) = [a_2, a_3, \dots]$$

Esto es, se trata de un shift en el espacio  $\Sigma_{\infty}^+ = \mathbb{N}^{\mathbb{N}}$ .

**Nota final:** Estas notas están en construcción. Cualquier aclaración, sugerencia o comentario es bienvenida en la dirección electrónica: aubin@matcuer.unam.mx.

## REFERENCIAS

- [1] Mañé, Ricardo. *Ergodic Theory and differentiable dynamics*, Springer Verlag, 1987.
- [2] Munkres, James R. *Topology*, Prentice Hall, 1975.
- [3] Viana, Marcelo. *Introdução à Teoria Ergódica*, Notas de la Escuela de Verano de Recife. Enero 2003.  
(<http://w3.impa.br/~viana/#notes>)
- [4] Walters, Peter. *An Introduction to Ergodic Theory*, Springer Verlag, 1982.



INSTITUTO DE MATEMÁTICAS - UNAM, UNIDAD CUERNAVACA