TEORIA ERGODICA

Eleonora Catsigeras

Apuntes para el curso de Maestría en Matemática.
Abril de 2005.

CAPITULO 2-COMPLEMENTO
DEMOSTRACIONES DE LOS TEOREMAS ERGÓDICOS.

Las hipótesis generales para este capítulo complementario son las siguientes:

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

Observación 0.1 Como la medida μ es finita, entonces para todo p natural, $p \geq 1$, se cumple $L^p(\mu) \subset L^1(\mu)$.

Demostración: Basta demostrar que si $\int |f|^p \, d\mu < +\infty$ entonces $\int |f| \, d\mu < +\infty$. Sea $|f| = g+h$ donde $g(x) = |f|(x)$ si $|f|(x) \leq 1$ y $g(x) = 0$ en caso contrario. Basta probar que g y h están en $L^1(\mu)$

$$\int g \, d\mu \leq \int 1 \, d\mu = \mu(X) < +\infty$$

$$\int h \, d\mu = \int_{\{h>0\}} h \, d\mu = \int_{\{|f|>1\}} |f| \, d\mu \leq \int_{\{|f|>1\}} |f|^p \, d\mu \leq \int |f|^p \, d\mu < +\infty \qquad \square$$

1. Teorema ergódico maximal

El siguiente teorema es un lema para demostrar el teorema ergódico de Birkhoff-Khinchin. (Teorema II.1.2). A diferencia del teorema de Birkhoff-Khinchin, este no se refiere a los promedios temporales de una función f a lo largo de la órbita por un punto $x \in X$, sino al máximo o supremo de las sumas acumuladas de los valores de f a lo largo de la órbita por x.

Lema 1.1 Teorema ergódico maximal. Versión 1.

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

$$f_n = \max\{f, f + f \circ T, \ldots, f + f \circ T + \ldots + f \circ T^{n-1}\}$$

Sea

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

Entonces

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

Demostración: Sea $E_n = \{x \in X : f_n(x) > 0\}$. Se cumple $E_n \subset E_{n+1}$, y $E(f) = \cup_{n\geq 1} E_n$. Luego:

$$\int_{E(f)} f\, d\mu = \int_{E(f)} f^+\, d\mu - \int_{E(f)} f^-\, d\mu == \lim \int_{E_n} f^+\, d\mu - \lim \int_{E_n} f^-\, d\mu = \lim \int_{E_n} f\, d\mu$$

Luego basta demostrar que $\int_{E_n} f\, d\mu \geq 0$ para todo $n \geq 1$.

Sea

$$I = \int_{E_n} f\, d\mu = \int_{\{f_n > 0, f_n \circ T \leq 0\}} f\, d\mu + \int_{\{f_n > 0, f_n \circ T > 0\}} f\, d\mu \qquad [1]$$

Tenemos

$$f_n(x) = \max\{f(x), f(x) + f(Tx), \ldots, f(x) + f(Tx) + \ldots + f(T^{n-1}x)\}$$

$$f_n(Tx) = \max\{f(Tx), \ldots, f(Tx) + \ldots + f(T^{n-1}x), f(Tx) + \ldots + f(T^n x)\}$$

Por un lado, si $f_n(Tx) \leq 0$ las igualdades de arriba muestran que $f_n(x) = f(x)$

Por otro lado, si $f_n(Tx) > 0$ entonces

$$f(x) + f_n(Tx) = \max\{f(x) + f(Tx), \ldots, f(x) + f(Tx) + \ldots + f(T^{n-1}x),$$

$$, f(x) + f(Tx) + \ldots + f(T^n x)\} \geq$$

$$\geq \max\{f(x) + f(Tx), \ldots, f(x) + f(Tx) + \ldots + f(T^{n-1}x)\}$$

Además, si $f_n(Tx) > 0$ entonces

$$f(x) + f_n(Tx) > f(x)$$

Juntando ambas desigualdades cuando $f_n(Tx) > 0$ se obtiene que:

$$f(x) + f_n(Tx) \geq \max\{f(x), f(x) + f(Tx), \ldots, f(x) + f(Tx) + \ldots + f(T^{n-1}x)\} = f_n(x)$$

En resumen:

$$f_n(x) = f(x) \text{ si } f_n(Tx) \leq 0, \qquad f(x) \geq f_n(x) - f_n(Tx) \text{ si } f_n(Tx) > 0$$

Sustituyendo en [1] resulta

$$I \geq \int_{\{f_n > 0, f_n \circ T \leq 0\}} f_n\, d\mu + \int_{\{f_n > 0, f_n \circ T > 0\}} f_n - f_n \circ T\, d\mu = \int_{\{f_n > 0\}} f_n\, d\mu - \int_{\{f_n > 0, f_n \circ T > 0\}} f_n \circ T\, d\mu \qquad [2]$$

Por otro lado como μ es T-invariante resulta

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

De donde

$$\int_{\{f_n > 0\}} f_n\, d\mu = \int (\chi_{\{f_n > 0\}} \circ T) \cdot (f_n \circ T)\, d\mu = \int (\chi_{\{f_n \circ T > 0\}}) \cdot (f_n \circ T)\, d\mu = \int_{\{f_n \circ T > 0\}} f_n \circ T\, d\mu$$

Sustituyendo en [2] resulta:

$$I \geq \int_{\{f_n \circ T > 0\}} f_n \circ T\, d\mu - \int_{\{f_n > 0, f_n \circ T > 0\}} f_n \circ T\, d\mu = \int_{\{f_n \leq 0, f_n \circ T > 0\}} f_n \circ T\, d\mu \geq 0 \qquad \square$$

Lema 1.2 Teorema ergódico maximal. Versión 2.

Sea $f \in L^1(\mu)$ real. Sea, como en el lema 1.1 el siguiente conjunto:

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

Sea $A \subset E(f)$ medible, tal que $T^{-1}(A) = A$. Entonces

$$\int_A f \, d\mu \geq 0$$

Demostración: Sea $g = \chi_A f$. Por el lema 1.1 se tiene

$$\int_{E(g)} \chi_A f \, d\mu = \int_{E(g)} g \, d\mu \geq 0$$

Basta demostrar que $E(g) = A$. En efecto, siendo $T^{-1}A = A$ se cumple

$$\chi_A \circ T^j = \chi_{T^{-j}(A)} = \chi_A \quad \forall j \geq 0$$

Luego:

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

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

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

Ahora veamos que el teorema ergódico maximal permite obtener resultados sobre los promedios temporales de la función f.

Lema 1.3 Corolario del teorema ergódico maximal.

Sea $f \in L^1(\mu)$ real. Sea α un número real cualquiera. Se define el siguiente conjunto:

$$\overline{G}_\alpha(f) = \{x \in X : \limsup_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) > \alpha\}$$

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

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

Lema 1.4 Corolario del teorema ergódico maximal.
Sea $f \in L^1(\mu)$ real. Sea β un número real cualquiera. Se define el siguiente conjunto:

$$\underline{G}_\beta(f) = \{x \in X : \liminf_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x) < \beta\}$$

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

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

Demostración de los lemas 1.3 y 1.4:

Demostración del lema 1.3: Basta probar que $\overline{G}_\alpha(f) \subset E(f-\alpha)$ ya que por el lema 1.2 $T^{-1}(A) = A \subset \overline{G}_\alpha(f) \subset E(f-\alpha)$ implica

$$\int_A (f-\alpha)\,d\mu \geq 0 \quad \Rightarrow \quad \int_A f\,d\mu \geq \alpha\mu(A)$$

En efecto $x \in \overline{G}_\alpha(f)$ implica $\limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1}(f-\alpha)\circ T^j(x) > 0$, de donde

$$\sup_{n\geq 1}\sum_{j=0}^{n-1}(f-\alpha)\circ T^j(x) > 0$$

Luego $x \in E(f-\alpha)$ como queríamos probar. □

Demostración del lema 1.4: Observemos que $\underline{G}_\beta(f) = \overline{G}_{-\beta}(-f)$. Por el lema 1.3 se tiene entonces

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

Lema 1.5 Los conjuntos $A_1 = \overline{G}_\alpha(f)$ y $A_2 = \underline{G}_\beta(f)$ definidos en los lemas 1.3 y 1.4 son invariantes con T, es decir $T^{-1}(A_i) = A_i$.

Demostración:

$$A_1 = \{x \in X : \limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x) > \alpha\}$$

$$T^{-1}(A_1) = \{x \in X : T(x) \in A_1\} = \{x \in X : \limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(T(x)) > \alpha\}$$

Basta demostrar que

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x) = \limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(T(x))$$

o, lo que es lo mismo, probar que

$$\limsup_{n\to+\infty} \frac{1}{n+1}\sum_{j=0}^{n} f\circ T^j(x) = \limsup_{n\to+\infty} \frac{1}{n}\sum_{j=1}^{n} f\circ T^j(x) \qquad [1]$$

En efecto

$$\frac{1}{n+1}\sum_{j=0}^{n} f\circ T^j(x) = \frac{f(x)}{n+1} + \left(\frac{n}{n+1}\right)\frac{1}{n}\sum_{j=1}^{n} f\circ T^j(x)$$

Como $n/(n+1)\to 1$ y $f(x)/(n+1)\to 0$ se cumple [1] como queríamos probar.

Por otro lado para probar que A_2 es también invariante con T basta observar que también vale [1] escribiendo límite inferior en lugar de límite superior. $\square$

2. Demostración del teorema II.1.2 de Birkhoff-Khinchin.

Parte a) *Para toda $f\in L^1(\mu)$ existe μ -c.t.p.*

$$\widetilde{f}(x) = \lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x)$$

Demostración:

Basta demostrarlo para $f\in L^1(\mu)$ real.

Sea para todo $n\geq 1$:

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

Sea $F=\{x\in X : \liminf f_n(x) < \limsup f_n(x)\}$. Basta probar que $\mu(F)=0$.

Se cumple

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

donde los conjuntos $\overline{G}_\alpha(f)$ y $\underline{G}_\beta(f)$ son los definidos en los lemas 1.3 y 1.4 respectivamente.

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

El conjunto A es invariante con T, es decir $T^{-1}(A)=A$, ya que $\overline{G}_\alpha(f)$ y $\underline{G}_\beta(f)$ lo son (ver lema 1.5).

Por el lema 1.3 y por el lema 1.4 se tiene entonces:

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

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

Parte b) *$\widetilde{f}(x)$ existe si y solo si existe $\widetilde{f}(Tx)$ y en ese caso $\widetilde{f}(Tx)=\widetilde{f}(x)$. Luego $\widetilde{f}=\widetilde{f}\circ T$ μ-c.t.p.*

Demostración: Basta considerar f real. Por el lema 1.5 se tiene, para todo $x\in X$:

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

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

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

Parte c1) *Para toda $f \in L^1(\mu)$ la función $\widetilde{f}$ también pertenece a $L^1(\mu)$ y además la convergencia de f_n a $\widetilde{f}$ es también en $L^1(\mu)$.*

Demostración: Basta demostrar la siguiente parte c2):

Parte c2) *Si $f \in L^p(\mu)$ para algún $p \geq 1$ natural entonces la función $\widetilde{f}$ también pertenece a $L^p(\mu)$ y además la convergencia a $\widetilde{f}$ de $f_n = (1/n)\sum_{j=0}^{n-1} f \circ T^j$ es también en $L^p(\mu)$.*

Lema 2.1 *Si $f \in L^p(\mu)$ para algún $p \geq 1$ natural entonces las funciones $\widetilde{f}$ y $f_n = (1/n)\sum_{j=0}^{n-1} f \circ T^j$ también pertenecen a $L^p(\mu)$ y cumplen*

$$\|f_n\|_p \leq \|f\|_p, \qquad \|\widetilde{f}\|_p \leq \|f\|_p$$

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

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

Basta observar que $\|f \circ T^j\|_p = \|f\|_p$ para todo $n \geq 1$. Resulta $I_n \leq \|f\|_p$.

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

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

Demostración de la parte c2):

Por el lema 2.1 se tiene $\widetilde{f} \in L^p(\mu)$. Solo resta probar que $f_n \to \widetilde{f}$ en $L^p(\mu)$.

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

$$\lim \int |f_n - \widetilde{f}|^p \, d\mu = \int \lim |f_n - \widetilde{f}|^p \, d\mu = 0$$

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

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

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

$$\|f_n - \widetilde{f}\|_p \leq \|f_n - g_n\|_p + \|g_n - \widetilde{g}\|_p + \|\widetilde{g} - \widetilde{f}\|_p \qquad [1]$$

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

$$\|g_n - \widetilde{g}\|_p < \epsilon \quad \forall n \text{ suficientemente grande}$$

Por otra parte, el lema 2.1 implica:

$$\|f_n - g_n\|_p = \|(f-g)_n\|_p \leq \|f-g\|_p < \epsilon,$$

$$\|\widetilde{f} - \widetilde{g}\|_p = \|\operatorname{lím}(f_n - g_n)\|_p = \|\operatorname{lím}(f-g)_n\|_p = \|\widetilde{(f-g)}\|_p \leq \|f-g\|_p < \epsilon$$

Sustituyendo en [1], se deduce que

$$\|f_n - \widetilde{f}\|_p < 3\epsilon \quad \forall n \text{ suficientemente grande}$$

de donde $\|f_n - \widetilde{f}\|_p \to 0$ como queríamos probar. □

Parte d) *Para toda $f \in L^1(\mu)$ se cumple*

$$\int f\, d\mu = \int \widetilde{f}\, d\mu$$

Demostración: Por la parte c1) se tiene que $\|f_n - \widetilde{f}\|_1 \to 0$. Luego:

$$\left|\int f_n\, d\mu - \int \widetilde{f}\, d\mu\right| \leq \int |f_n - \widetilde{f}|\, d\mu = \|f_n - \widetilde{f}\|_1 \to 0$$

Concluimos por un lado que

$$\int f_n\, d\mu \to \int \widetilde{f}\, d\mu$$

Por otro lado:

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

Reuniendo ambos resultados se obtiene $\int f\, d\mu = \int \widetilde{f}\, d\mu$, como queríamos. □

3. Caracterización de la ergodicidad.

En el capítulo I, Definición I.2.10 definimos:
Definición I.2.10
La medida de probabilidad μ invariante con T, es *ergódica* para T si todo conjunto A medible tal que $T^{-1}(A) = A$ tiene μ medida igual a uno o a cero.

Recordemos que en el capítulo I, teorema I.2.11 demostramos lo siguiente:

Teorema I.2.11

La medida de probabilidad μ invariante con T es ergódica para T si y solo si para todos $A, B \subset X$ medibles con μ medida positiva, existe $n \geq 1$ tal que $T^{-n}(A) \cap B \neq \emptyset$.

Demostración: Ver capítulo I, Teorema I.2.11.

Ahora demostraremos el teorema II.2.1, que caracteriza la ergodicidad de varias otras formas:

Teorema II.2.1.

Son equivalentes las siguientes afirmaciones:

(a) μ es ergódica respecto de T.

(b) Para toda $f \in L^1(\mu)$ el promedio temporal $\widetilde{f}(x)$ es igual al promedio espacial $\int f\, d\mu$ para $\mu-$c.t.p. $x \in X$.

(c) Si $f \in L^1(\mu)$ cumple $f \circ T = f$ $\mu-$c.t.p. entonces $f =$cte $\mu-$c.t.p.

(d) Para algún natural $p \geq 1$ las únicas funciones $f \in L^p(\mu)$ que son invariantes con T (es decir que cumplen $f \circ T = f$ $\mu-$c.t.p) son las funciones constantes μ-c.t.p.

(e) Para todos $A, B \subset X$ medibles se cumple

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

Primero probaremos la siguiente:

Afirmación: *Si μ es ergódica y $f : X \mapsto C$ es una función medible tal que $f(T(x)) = f(x)$ para todo $x \in X$, entonces $f = cte$ $\mu-$c.t.p.*

Demostración: Basta probarlo para f real. Sea α un número real cualquiera y sea $A_\alpha = \{x \in X : f(x) \leq \alpha\} = \{x \in X : f(Tx) \leq \alpha\} = T^{-1}(A_\alpha)$. Como μ es ergódica, por la definición I.2.10, $\mu(A_\alpha)$ es cero o uno.

Sea

$$\alpha_0 = \inf\{\alpha \in I\!R : \mu(A_\alpha) = 1\}$$

(A priori α_0 podría ser $-\infty$ o $+\infty$.)

Por construcción $\mu(A_\alpha)$ es creciente con α. Entonces se cumple $\mu(A_\alpha) = 1$ para todo α racional mayor que α_0. En otras palabras, μ -c.t.p. $x \in X$ verifica $f(x) \leq \alpha$ para todo α racional mayor que α_0. De lo anterior se deduce:

$$f(x) \leq \alpha_0 \quad \mu - c.t.p.$$

Por otro lado de la construcción de α_0 se cumple $\mu(A_\alpha) = 0$ para todo α racional menor que α_0. En otras palabras, μ-c.t.p. $x \in X$ verifica $f(x) > \alpha$ para todo α racional menor que α_0. De lo anterior se deduce:

$$f(x) \geq \alpha_0 \quad \mu - c.t.p. \quad \square$$

Demostración de que a) $\Rightarrow$ b) en el teorema II.2.1: Por el teorema de Birkhoff-Khinchine $\int \widetilde{f}\, d\mu = \int f\, d\mu$. Entonces basta demostrar que $\widetilde{f}$ es constante $\mu-$c.t.p.

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

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

Demostración de que b) ⇒ c) en el teorema II.2.1: Si $f \circ T = f \quad \mu - c.t.p.$ entonces como μ es invariante con T se cumple para todo $j \geq 0$ que $f \circ T^{j+1} = f \circ T^j \quad \mu c.t.p.$. Luego $f \circ T^j = f \quad \mu c.t.p.$, de donde $f_n = (1/n) \sum_{j=0}^{n-1} f \circ T^j = f \quad \mu - c.t.p. \quad \forall n \geq 1$. Entonces $\widetilde{f} = \lim f_n$ es igual a $f \quad \mu - c.t.p.$. Pero debido a la afirmación b) la función $\widetilde{f}$ es constante $\mu - c.t.p.$, de donde f también lo es. □

Demostración de que c) ⇒ d) en el teorema II.2.1: Como para todo natural $p \geq 1$ se cumple $L^p(\mu) \subset L^1(\mu)$ entonces c) implica la siguiente propiedad aparentemente más fuerte que d): *Para todo natural $p \geq 1$ las únicas funciones $f \in L^p(\mu)$ que cumplen $f = f \circ T \quad \mu - c.t.p$ son las funciones constantes $\mu - c.t.p.$* □

Demostración de que d) ⇒ e) en el teorema II.2.1:
Obsérvese que $\mu(T^{-j}A \cap B) = \int \chi_{T^{-j}(A)} \chi_B \, d\mu = \int (\chi_A \circ T^j) \chi_B \, d\mu$, de donde:

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

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

$$\lim I_n = \int \lim_{n \to +\infty} (\frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ T^j) \chi_B \, d\mu = \int (\widetilde{\chi}_A) \chi_B \, d\mu \quad [1]$$

Por la afirmación d), y el teorema de Bikhoff-Khinchine, observando que $\chi_A \in L^p(\mu)$ y $\widetilde{\chi}_A$ es invariante con T, luego contante $\mu - c.t.p$, se tiene que

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

Sustituyendo en [1] resulta

$$\lim I_n = \int (\mu(A)) \chi_B \, d\mu = \mu(A)\mu(B) \quad \square$$

Demostración de que e) ⇒ a) en el teorema II.2.1: Por el teorema I.2.11 basta demostrar que si $A, B \subset X$ son medibles con μ medida positiva, entonces existe $j \geq 1$ tal que $\mu(T^{-j}(A) \cap B) > 0$. Por la hipótesis e):

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

Entonces

$$\sum_{j=1}^{n-1} \mu(T^{-j}(A) \cap B) > 0 \quad \forall n \text{ suficientemente grande}$$

de donde $\mu(T^{-j}(A) \cap B) > 0$ para algún $j \geq 1$ como queríamos. □

4. Otras caracterizaciones de la ergodicidad.

Recordemos la definición de tiempo de estadía $\tau_A(x)$ de la órbita por $x \in X$ en un conjunto medible A (definición II.1.6):

$$\tau_A(x) = \lim_{n\to+\infty} \frac{\#\{j \in I\!N : 0 \leq j \leq n-1, T^j(x) \in A\}}{n} = \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \chi_A \circ T^j(x) = \widetilde{\chi}_A(x)$$

Recordamos ahora el teorema II.2.3 que enuncia la siguiente caracterización de la ergodicidad:

Teorema II.2.3
μ es ergódica respecto de T si y solo si para todo conjunto medible A el tiempo de estadía $\tau_A(x) = \widetilde{chi}_A(x)$ es constante μ-c.t.p. $x \in X$. En ese caso $\tau_A(x) = \mu(A)$ $\mu - c.t.p.$

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

$$\widetilde{\chi}_A(x) = \int \widetilde{\chi}_A \, d\mu = \int \chi_A \, d\mu = \mu(A) \quad \mu - c.t.p. \qquad [1]$$

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

$$\widetilde{\chi}_A = \lim \frac{1}{n}\sum_{j=0}^{n-1} \chi_A \circ T^j$$

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

$$\widetilde{\chi}_A(x) = \chi_A(x) \in \{0,1\} \quad \mu - c.t.p.x \in X \qquad [2]$$

Por [1] y [2] se tiene $\mu(A) \in \{0,1\}$ y μ es ergódica como queríamos probar.
Recíprocamente, si μ es ergódica, entonces por la parte b) del teorema II.2.1. $\tau_A = \widetilde{\chi}_A = cte$ $\mu - c.t.p.$ $\square$

Ahora recordemos la definición II.2.4:

Definición II.2.4:

$$\mathcal{M}_T = \{\mu \text{ medidas de probabilidad en } (X, \mathcal{A}) \text{ invariantes con } T\}$$

μ es *extremal* en $\mathcal{M}_T$ si la única combinación convexa de dos medidas en $\mathcal{M}_T$ que da como resultado μ es $1\mu + 0$. Más precisamente, μ es extremal si $\mu = t\nu_1 + (1-t)\nu_2$, donde $\nu_1 \neq \nu_2$ son dos medidas en $\mathcal{M}_T$ y $t \in [0,1]$, implica que $t \in \{0,1\}$ y entonces o bien $\mu = \nu_1$ o bien $\mu = \nu_2$.

Ahora vamos a demostrar el teorema II.2.5 que da otra caracterización para la ergodicidad:

Teorema II.2.5 *μ es ergódica para T si y solo si μ es extremal en $\mathcal{M}_T$.*

Demostración:

Prueba de que μ extremal implica μ ergódica:

Por absurdo si μ no fuera ergódica existiría un conjunto medible A tal que $T^{-1}(A) = A$ y $0 < \mu(A) < 1$. Sea $t = \mu(A) \in (0,1)$. Sean $\nu_1(B)$ y $\nu_2(B)$ definidas para todo medible B como $\nu_1(B) = \mu(B \cap A)/\mu(A)$, $\nu_2(B) = \mu(B \cap A^c)/\mu(A^c)$. Como $T^{-1}(A) = A$ entonces ν_1 y ν_2 son invariantes con T. Además $\nu_1 \neq \nu_2$ porque en A una toma el valor 1 y la otra el valor 0. Por construcción se verifica para todo medible B:

$$\mu(B) = \mu(B \cap A) + \mu(B \cap A^c) = t\nu_1(B) + (1-t)\nu_2(B)$$

Luego μ no es extremal contradiciendo la hipótesis $\square$.

Antes de probar que μ ergódica implica μ extremal, veremos el siguiente lema:

Lema 4.1 *Si μ es una medida de probabilidad ergódica respecto de T, y si ν es una medida de probabilidad invariante con T tal que $\nu << \mu$ entonces $\nu = \mu$.*

Demostración: Sea $h \in L^1(\mu)$ la derivada de Radon-Nikodym de ν respecto de μ. Para todo A medible se tiene

$$\nu(A) = \int_A h\,d\mu, \qquad \nu(X) = 1 = \int h\,d\mu \qquad [2]$$

Siendo ν invariante con T se cumple:

$$\nu(A) = \frac{1}{n}\sum_{j=0}^{n-1} \nu(T^{-j}(A) = \frac{1}{n}\sum_{j=0}^{n-1}\int (\chi_A \circ T^j)h\,d\mu = \int (\frac{1}{n}\sum_{j=0}^{n-1}\chi_A \circ T^j)h\,d\mu$$

Por convergencia dominada la última integral tiende a $\int \widetilde{\chi}_A h\,d\mu$. Siendo μ ergódica aplicando II.2.3 y la igualdad [2], concluimos

$$\nu(A) = \int \widetilde{\chi}_A h\,d\mu = \int \mu(A) h\,d\mu = \mu(A)\int h\,d\mu = \mu(A) \quad \forall A \qquad \square$$

Prueba de que μ ergódica implica μ extremal:

Supongamos que existen medidas de probabilidad ν_1 y ν_2 invariantes con T tales que

$$\mu = t\nu_1 + (1-t)\nu_2 \quad \text{para } t \in (0,1) \qquad [2]$$

Para demostrar que μ es extremal en $\mathcal{M}_T$ basta probar que $\nu_1 = \nu_2 = \mu$.

Por [2] se tiene $\nu_1 << \mu, \nu_2 << \mu$. Siendo μ ergódica, por el lema 4.1, se tiene $\nu_1 = \nu_2 = \mu$ como queríamos. $\square$