T- 25

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

TEORIA ERGODICA DE PROCESOS ESTOCASTICOS

T E S I S

Que para obtener el titulo de

A C T U A R I O S

P r e s e n t a n

JUAN RUIZ DE CHAVEZ SOMOZA

OSCAR VARGAS GARZA

1a ESTA

MEXICO, D. F. 1979

CAPITULO I

I.1 CONVERGENCIA CONMUTATIVA

I.1.1. DEFINICION: *Sea* **L** *un espacio lineal normado y suponer que* $\{X_\alpha\}$, $\alpha \in \Lambda$ *(Λ un conjunto de índices) es una colección de elementos en* **L**, *se dice que* $\{X_\alpha\}$ *es sumable (conmutativamente convergente) a* $X \in L$, *escrito* $\sum_{\alpha\in\Lambda} X_\alpha = X$ *o algunas veces solamente* $\sum_{\Lambda} X_\alpha = X$ *si* $\forall\varepsilon > 0$ $\exists$ *algún conjunto finito de índices* $H \subset \Lambda$ *tal que para cualquier conjunto finito de índices* $J \supset H$

$\| \sum_{\alpha\in J} X_\alpha - X \| < \varepsilon$

Antes de probar algunas consecuencias inmediatas de la definición; notemos la similaridad de esta noción con la convergencia incondicional de números reales: una convergencia suficientemente fuerte para garantizar que el reagrupamiento de términos no tiene efecto.

La primera observación que hacemos es que si $\sum_{\alpha\in\Lambda} X_\alpha = X$ *y* $\beta \in F$, *entonces*

$$\sum_{\alpha\in\Lambda} \beta X_\alpha = \beta X$$

Para probar esto, sea $\varepsilon > 0$ *dado, podemos elegir un conjunto finito de índices* H *tal que* $\forall$ *conjunto finito* $J \supset H$

$$\| \sum_{\alpha\in J} \beta X_\alpha - \beta X \| = |\beta| \; \| \sum_{\alpha\in J} X_\alpha - X \| < \varepsilon$$

Enseguida si tenemos que $\sum_{\alpha\in\Lambda} X_\alpha = X$ *y* $\sum_{\alpha\in\Lambda} Y_\alpha = Y$ *entonces* -

$$\sum_{\alpha\in\Lambda} (X_\alpha + Y_\alpha) = X + Y$$

Para probar esto elegimos H_1 *y* H_2 *tales que, para todos los conjuntos finitos de índices* J_1 *y* J_2 *donde* $J_1 \supset H_1$ *y* $J_2 \supset H_2$

$$\| \sum_{\alpha\in J_1} X_\alpha - X \| < \varepsilon/2 \text{ y } \| \sum_{\alpha\in J_2} Y_\alpha - Y \| < \varepsilon/2$$

y notamos que para cualquier conjunto finito $J \supset H_1 \cup H_2$ *implica* $J \supset H_1$ *y* $J \supset H_2$*, por tanto*

$$\| \sum_{\alpha\in J} X_\alpha + Y_\alpha - (X + Y) \| < \varepsilon$$

El siguiente resultado muestra que si una serie es sumable en el sentido definido antes, todos excepto un número numerable de términos en la serie deben ser cero.

I.1.2. TEOREMA: *Sea* L *un espacio lineal normado y suponer que* $\{X_\alpha\}$ *es sumable a* $X\in L$*, donde* $\alpha\in\Lambda$*. Entonces todos excepto un número numerable de las* X_α *deben ser cero.*

DEMOSTRACION: *Sea* $\varepsilon>0$ *dado y* H_ε *un conjunto finito de índices tal que* $\forall J_1 \supset H_\varepsilon$

$$\| \sum_{\alpha \in J_1} X_\alpha - X \| < \frac{\varepsilon}{2} .$$

Ahora consideramos otro conjunto finito de índices J *tal que* $J \cap H_\varepsilon = \phi$. *En este caso*

$$\| \sum_{\alpha \in J} X_\alpha \| = \| \sum_{\alpha \in J \cup H_\varepsilon} X_\alpha - \sum_{\alpha \in H\varepsilon} X_\alpha \|$$

$$\leq \| X - \sum_{\alpha \in J \cup H_\varepsilon} X_\alpha \| + \| X -$$

$$- \sum_{\alpha \in H_\varepsilon} X_\alpha \| < \varepsilon, \text{ ya que } J \cup H_\varepsilon \supset H_\varepsilon .$$

Ahora sea $\varepsilon = \frac{1}{n}$, $n = 1, 2, \ldots$. *Existen conjuntos finitos de índices* H_n *tal que, si* J *es un conjunto finito de índices y* $J \cap H_n = \phi$, *entonces*

$$\| \sum_{\alpha \in J} X_\alpha \| < \frac{1}{n}$$

Consideremos el conjunto numerable $\bigcup_{n=1}^{\infty} H_n$ *y algún índice* α *tal que*

$$\{\alpha\} \cap (\bigcup_{n=1}^{\infty} H_n) = \phi$$

esto implica

$$\{\alpha\} \cap H_n = \phi \text{ para toda } n,$$

lo que significa

$$\| X_\alpha \| < \frac{1}{n} \text{ para toda } n.$$

Consecuentemente $X_\alpha = 0$, *y así este debe ser el caso para cualquier* X_α *donde* $\alpha \notin \bigcup_{n=1}^{\infty} H_n$.

Como comentario final notemos que en ninguna parte de la discución anterior intentamos sumar más de un número finito de términos, a pesar de que en la notación lo podamos indicar; una examinación estricta revela que solo sumas finitas han sido distribuidas para caracterizar este tipo especial de suma infinita.

I.2. NORMAS Y PRODUCTOS INTERNOS SOBRE PRODUCTOS CARTESIANOS DE ESPACIOS CON PRODUCTO INTERNO NORMADOS.

Sean **L** *y* **M** *espacios lineales normados, deseamos introducir una norma sobre* **L X M**, *sin embargo, primero debemos introducir operaciones sobre* **L X M** *de tal forma que éste pueda verse como un espacio lineal. Para evitar confusión con productos internos, los elementos de* **L X M** *se denotarán por* $\langle X, Y \rangle$.

Usando esta notación, definimos suma de elementos de **L X M** *como*

$$< X_1, Y_1 > + < X_2, Y_2 > = < X_1 + X_2, Y_1 + Y_2 >$$

y la multiplicación por un escalar $\alpha \epsilon F$, *como*

$$\alpha < X, Y > = < \alpha X, \alpha Y >$$

Es fácil de verificar que **L X M** *con respecto a estas operaciones es un espacio lineal.*

Ahora afirmamos que los siguientes mapeos son normas para **L X M**

I.2.1. $\qquad \| < X, Y > \| = \| X \| + \| Y \|$

I.2.2. $\| < X, Y > \| = (\| X \|^{p} + \| Y \|^{p})^{1/p}, \; p > 1$

I.2.3. $\| < X, Y > \| = \max (\| X \|, \| Y \|)$.

La afirmación se puede verificar facilmente

Ahora definimos un producto interno para **L X M** *como*

I.2.4. $(< X_1, Y_1 >, < X_2, Y_2 >) = (X_1, X_2) + (Y_1, Y_2)$

en donde (X_1, X_2) *y* (Y_1, Y_2) *denotan él producto interno sobre* **L** *y* **M** *respectivamente. Se puede ver facilmente que* I.2.4. *es realmente un producto interno sobre* **L X M**.

Ademas notemos que si consideramos una correspondencia entre elementos **X** *de* **L** *y el elemento* $< X, 0 >$ *de* **L X M**, *se puede demostrar que esta correspondencia es un isomorfismo entre* **L** *y* - - $\{< X, 0 >: X \in L\}$, *y similarmente para* **M**. *Por tanto el producto interno definido aquí y las normas definidas antes extienden el producto interno original y la norma sobre cada uno de los espacios, es decir, las ecuaciones definidas para el producto cartesiano* **L X M** *se reducen a los mapeos originales cuando nos restringimos a uno de los espacios* **L** *o* **M**.

Ahora que tenemos una norma ó métrica sobre el producto cartesiano de espacios lineales, esto requiere un cambio para que hablemos de mapeos continuos de tales productos cartesianos en otros espacios.

Se dice que una función es continua en un punto **X** *si para cada*

ε- vecindad de $f(X)$ en el rango, se puede encontrar una δ-vecindad de X en el dominio, $S_\delta(X)$ tal que $f(S_\delta(X)) \subset S_\varepsilon(f(X))$. En la situación particular cuando el dominio es el producto cartesiano de 2 espacios, el punto X mencionado antes será una pareja ordenada de elementos en los espacios componentes y una δ-vecindad del punto será una vecindad en el espacio producto cartesiano.

Probaremos ahora que el producto interno es un mapeo continuo.

I.2.5. **TEOREMA:** *Si L es un espacio con producto interno, el producto interior (X, Y) es una función continua que mapea $L \times L$ en F.*

DEMOSTRACION: Consideremos un punto fijo en el rango (X_2, Y_2).

Ahora sean $X_3 = X_1 - X_2$ y $Y_3 = Y_1 - Y_2$ lo que implica

$$|(X_1, Y_1) - (X_2, Y_2)| = |(X_2 + X_3, Y_2 + Y_3) - (X_2, Y_2)|$$

desarrollando

$$|(X_2 + X_3, Y_2 + Y_3) - (X_2, Y_2)| =$$

$$= |(X_2, Y_2) + (X_2, Y_3) + (X_3, Y_2) + (X_3, Y_3) - (X_2, Y_2)|$$

usando la desigualdad de Cauchy-Schwarz, obtenemos

$$|(X_2, Y_3) + (X_3, Y_2) + (X_3, Y_3)| \leq$$

$$\leq \| X_2 \| \| Y_1 - Y_2 \| + \| Y_2 \| \| X_1 - X_2 \| +$$

$$+ \| X_1 - X_2 \| \| Y_1 - Y_2 \|$$

I.3. ESPACIOS DE HILBERT

Si L es un espacio con producto interno, podemos inmediatamente definir una norma en términos del producto interno. - Una vez que tenemos una norma, esto inmediatamente da surgimiento a una métrica sobre el espacio, y como usualmente, un hecho importante acerca del espacio es cuando éste es completo con - respecto a ésta métrica. Hacemos la siguiente definición.

I.3.1. DEFINICION: Si un espacio con producto interno H es completo con respecto a la métrica derivada del producto interno, entonces se dice que H es un espacio de Hilbert.

Para algunos ejemplos de espacios de Hilbert ver [7] p. 165-167.

Un ejemplo de un espacio con producto interno que no es un espacio de Hilbert. Consideremos el espacio lineal de las funciones continuas sobre un intervalo $[a, b]$ denotado por $C[a, b]$, con producto interno de f y $g \in C[a, b]$ definido como

$$(f, g) = \int_a^b f(X)\, \overline{g(X)}\, dX$$

Ahora si una sucesión de funciones $\{f_n\}$ en $C[a, b]$ converge con respecto a la norma definida por este producto interno, ellas sólo convergerán en media.

Y como convergencia media no necesariamente implica convergencia uniforme, por tanto este espacio no tiene que ser completo.

El siguiente ejemplo muestra que esto es así. ejemplo: Considerar $C[a, b]$ con la métrica

$$d(f, g) = \left(\int_a^b |f(X) - g(X)|^2 dX\right)^{1/2} \quad f, g \in C[a, b].$$

tomemos $a = -1$ y $b = 1$ y consideremos la siguiente sucesión de Cauchy

$$f_n(X) = \begin{cases} 0 & -1 \leq X \leq 0 \\ nX & 0 < X \leq \frac{1}{n} \\ 1 & \frac{1}{n} < X \leq 1 \end{cases}$$

Es claro que la función límite es una función discontinua y por tanto el espacio no es completo.

La ley del paralelógramo establece que si X y Y son 2 vectores en un espacio con producto interno L, entonces

I.3.2 $\quad \| X + Y \|^2 + \| X - Y \|^2 = 2\| X \|^2 + 2\| Y \|^2$

(ver [3], p.11)

Esta ley es de interés para nosotros porque da una condición necesaria y suficiente que cualquier norma determinada por un producto interno debe de satisfacer.

Si tenemos un espacio lineal normado y nos preguntamos cuando podemos introducir un producto interno, el cual determinará una norma que concuerde con la norma original, sabemos que nece sitamos checar primeramente si la norma original satisface la ley del paralelogramo; Cualquier norma derivada de un producto interno debe satisfacer I.3.2.

Como una aplicación inmediata de este hecho, consideremos el espacio lineal L^1 [0, 1], que consiste de clases de equivalencia de funciones integrales sobre [0, 1] con respecto a la medida de Lebesgue, con la norma de $f \in L^1$ [0, 1] definida como

$$\| f \| = \int_0^1 |f(X)| dX$$

Mostraremos que esta norma no satisface I.3.2. y así evitar la posibilidad de ver a éste espacio como un espacio de -- Hilbert (con respecto a esta norma).

Consideremos los conjuntos $A = [0, \frac{1}{2}]$ y $B = [\frac{1}{2}, 1]$ y las funciones características I_A, I_B.

Vemos ahora que

$$\| I_A + I_B \|^2 = 1 \text{ y } \| I_A - I_B \|^2 = 1$$

mientras que

$$2\| I_A \|^2 + 2\| I_B \|^2 = 2\tfrac{1}{4} + 2\tfrac{1}{4} = 1$$

por tanto no se cumple la ley del paralelógramo.

I.4 DESIGUALDAD DE BESSEL

I.4.1 TEOREMA. *Sea L un espacio con producto interno, A un conjunto ortonormal de vectores en L y Y un vector arbitra-*

rio en **L.**

Entonces

(1) (*Desigualdad de Bessel*) $\forall\ X_1, X_2, \ldots, X_n \in A$.

$$\sum_{i=1}^{n} |(Y, X_i)|^2 \leq \| Y \|^2;$$

(2) *El conjunto*

$$E = \{X \in A: \ (Y, X) \neq 0\} \text{ es numerable;}$$

(3) *Si* $Z \in L$, *entonces* $\sum_{X \in A} |(Y, X)(Z, X)| \leq \| Y \| \, \| Z \|$

DEMOSTRACION: *Sea* $\alpha_i = (Y, X_i)$. *Es claro que*

$$0 \leq (Y - \sum_{i=1}^{n} \alpha_i X_i, \ Y - \sum_{j=1}^{n} \alpha_j X_j) = \| Y \|^2 - \sum_{j=1}^{n} \overline{\alpha}_j (Y, X_j) -$$

$$- \sum_{i=1}^{n} \alpha_i (X_i, Y) + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \overline{\alpha}_j (X_i, X_j)$$

$$= \| Y \|^2 - \sum_{j=1}^{n} |\alpha_j|^2 - \sum_{i=1}^{n} |\alpha_i|^2 + \sum_{i=1}^{n} |\alpha_i|^2 = \| Y \|^2 - \sum_{j=1}^{n} |\alpha_j|^2$$

por tanto $\sum_{j=1}^{n} |(Y, X_j)|^2 \leq \| Y \|^2$, *lo cual prueba* (1)

Para probar (2), *consideremos el conjunto*

$$E_n = \{X \in A : \ |(Y, X)| \geq \tfrac{1}{n}\}$$

donde n *es algún entero positivo y suponer que* $X_1, X_2, \ldots, X_k \in E_n$.

En este caso por la desigualdad de Bessel y de la definición de E_n

$$K \cdot \frac{1}{n^2} < \sum_{i=1}^{n} |(Y, X_i)|^2 \leq \| Y \|^2 \text{ , implica } K < n^2 \| Y \|^2 < \infty$$

Así el conjunto $E = \{X \in A\colon (Y, X) \neq 0\} = \bigcup_{n=1}^{\infty} E_n$, es una unión numerable de conjuntos finitos y es por tanto numerable.

Por último, aplicando la parte (2) del teorema para asegurar que $(Y, X_i) = 0$ y $(Z, X_i) = 0$, para todos excepto un número numerable de $X_i \in A$, usando la desigualdad de Hölder y la desigualdad de Bessel, podemos decir que para una n cualquiera

$$\sum_{i=1}^{n} |(Y, X_i)(\overline{Z, X_i})| \leq \left(\sum_{i=1}^{n} |(Y, X_i)|^2\right)^{1/2} \left(\sum_{i=1}^{n} |(Z, X_i)|^2\right)^{1/2} \leq$$

$$\leq \| Y \| \, \| Z \| .$$

La serie $\sum_{X \in A} |(Y, X_i)\ (\overline{Z, X_i})|$ converge absolutamente y por tanto incondicionalmente, la cual, para series de números complejos significa convergencia conmutativa.

I.5. ALGUNOS RESULTADOS DE $L^2(0, 2\pi)$ Y EL TEOREMA DE RIESZ - FISCHER.

Consideremos el espacio de Hilbert real de las funciones reales de cuadrado integrable $L^2(0, 2\pi)$ (es un espacio completo ver [3], p. 115), y el conjunto ortonormal en este espacio

$$A: \frac{1}{\sqrt{2\pi}}, \frac{\text{sen } t}{\sqrt{\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\text{sen } 2t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \ldots = X_0, X_1, X_2, \ldots$$

(ver [9], 110)

es decir,

para $n>0$

$$X_n = \begin{cases} \frac{1}{\sqrt{\pi}} \cos(\frac{n}{2} t), & n \text{ par} \\ \frac{1}{\sqrt{\pi}} \text{sen } (\frac{n+1}{2} t), & n \text{ impar} \end{cases}$$

Ahora definimos, para $f \in L^2 (0, 2\pi)$, $\alpha_n = (f, X_n)$;

i.e.

$$\alpha_0 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t)dt$$

$$\alpha_1 = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(t) \text{ sen } t \, dt.$$

$$\vdots$$

Consideremos ahora los coeficientes de la serie de Fourier

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(t) \, dt$$

y para $n > 0$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos (\frac{n}{2} t) \, dt, \; n \text{ par}$$

y

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \text{ sen } (\frac{n+1}{2} t) \, dt, \; n \text{ impar}$$

entonces tenemos las relaciones

$$\alpha_0 = a_0\sqrt{\frac{\pi}{2}},\ \alpha_1 = b_1\sqrt{\pi},\ \alpha_2 = a_2\sqrt{\pi},\dots$$

Haciendo uso de la desigualdad de Bessel (teorema I.4.1. (1)) tenemos

$$\sum_{n=0}^{\infty} |\alpha_n|^2 \leq \| f \|^2$$

Substituyendo las anteriores relaciones, obtenemos

$$a_0^2\,\frac{\pi}{2} + \pi\Big(\sum_{n=1}^{\infty} a_n^2 + b_n^2\Big) \leq \int_0^{2\pi} f^2(t)\,dt$$

$$\therefore\ \lim_n a_n = 0 \ \text{y} \ \lim_n b_n = 0$$

lo cual es un resultado importante usualmente llamado el lema de Riemann-Lebesgue.

Regresando a la situación general una vez más, ahora deseamos probar el siguiente teorema.

I.5.1. TEOREMA. *Si* **L** *es un espacio con producto interno separable y* **A** *es un conjunto ortonormal en* **L**, *entonces* **A** *es numerable.*

DEMOSTRACION. *Sean* **X** *y* **Y** *dos elementos de* **A**. *Podemos afirmar que*

$$\|X - Y\|^2 = (X - Y,\ X - Y) = \| X \|^2 - (X,\ Y) - (Y,\ X) + \| Y \|^2 =$$

$$= \| X \|^2 + \| Y \|^2 = 2$$

$$\therefore\ \| X - Y \| = \sqrt{2}.$$

Ahora consideremos la colección de vecindades de radio $\sqrt{\frac{2}{2}}$ *alrededor de cada punto de* **A**. *Es claro que dos de éstas vecindades no pueden tener un punto en común de* **A**, *lo que implica -- que cada una de tales vecindades deben contener un punto distinto de un subconjunto denso numerable. Por tanto si* **A** *no fuera numerable tendríamos un subconjunto denso no numerable, lo que contradice la hipótesis.*

El siguiente resultado da un importante hecho acerca de convergencia (en el sentido de sumabilidad introducida en I.1) en espacios de Hilbert.

I.5.2. TEOREMA: *Sea* **L** *un espacio de Hilbert y sea -* **A**: $X_1, X_2, \ldots,$ *un conjunto ortonormal en* **L**. *Entonces, si* $\alpha_n \varepsilon F$, $n = 1, 2, \ldots,$ *tenemos*

(1) $\sum_{n=1}^{\infty} \alpha_n X_n$ *converge si y sólo si* $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$ *y*

(2) *Si* $\sum_{n=1}^{\infty} \alpha_n X_n$ *converge a* **X**, *entonces* $\alpha_n = (X, X_n)$.

DEMOSTRACION: (1) *Considerar la suma parcial* $S_n = \sum_{i=1}^{n} \alpha_i X_i$ *y suponer que* $n > m$; *entonces*

I.5.2.1 $$\| S_n - S_m \|^2 = \| \sum_{i=m+1}^{n} \alpha_i X_i \|^2 = \sum_{i=m+1}^{n} |\alpha_i|^2$$

Claramente, si la serie mencionada en (1) converge, entonces $\sum_{i=m+1}^{n} |\alpha_i|^2$ *debe tender a cero, por tanto la serie* $\sum_{n=1}^{\infty} |\alpha_n|^2$ *debe converger.*

Inversamente, suponer que $n, m \to \infty$ *y* $\sum_{i=m+1}^{n} |\alpha_i|^2 \to 0$.

Esto implica que la sucesión de sumas parciales $\sum_{i=1}^{n} \alpha_i X_i$ *es es una sucesión de Cauchy, y como* L *es un espacio de Hilbert* $\sum_{n=1}^{\infty} \alpha_n X_n$ *debe converger a algún* $X \in L$.

DEMOSTRACION DE (2). *Considerar que* $X = \sum_{n=1}^{\infty} \alpha_n X_n$, *y la suma parcial* $S_n = \sum_{i=1}^{n} \alpha_i X_i$.

Para $n > j$, *definimos* $\alpha_j = (S_n, X_j)$. *Como esta relación se cumple dada la ortonormalidad del conjunto* A, *para toda* $n > j$, *esto debe ser cierto en el límite, y podemos escribir* $\lim_n (S_n, X_j) = \alpha_j$.

Usando el teorema I.2.5. *podemos asegurar la continuidad del mapeo producto interno, lo cual nos permite intercambiar las operaciones de límite y producto interno en la ecuación anterior, lo que nos da* $(X, X_j) = \alpha_j$, *lo que es el resultado deseado.*

Así, conociendo que las series dadas convergen a algun **X**, *tenemos una relación fuerte entre los coeficientes en la serie y* **X**.

Regresando ahora al caso para $L^2(0, 2\pi)$ *mencionado el inicio de esta sección, tenemos el siguiente teorema.*

I.5.3. **TEOREMA: (RIESZ-FISCHER).** *Sea* $c_0, d_1, c_2, \ldots,$ *una colección de números reales tales que* $(c_0^2 + \sum_{n\ par} c_n^2 + \sum_{n\ impar} d_n^2) < \infty$ *entonces existe una función* $f \in L^2(0, 2\pi)$ *tal que los números* c_k *y* d_k *son los coeficientes de fourier de* **f**.

DEMOSTRACION: *Para probar esto necesitamos sólo hacer uso de la parte (1) del teorema* I.5.2. *donde el conjunto ortonormal es* $A: \frac{1}{\sqrt{2\pi}}, \frac{\text{sen } t}{\sqrt{\pi}}, \frac{\cos t}{\sqrt{\pi}}, \ldots = X_0, X_1, X_2, \ldots$

para afirmar que [*con respecto a la norma sobre* $L^2(0, 2\pi)$]

$$\frac{c_0}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}}\left(\sum_{n\ par} c_n \cos nt + \sum_{n\ impar} d_n \text{sen } nt\right)$$ *converge a alguna* $f \in L^2(0, 2\pi)$ *y* $f/\sqrt{\pi}$ *tiene a* c_n *y* $d_n (n = 0, 1, 2, \ldots)$ *como los coeficientes de fourier. Esto se sigue de* I.5.2.(2) *y de la definición de* (f, X_n) *para* $f \in L^2(0, 2\pi)$.

I.6. CONJUNTOS ORTONORMALES COMPLETOS.

I.6.1. DEFINICION. *Sea* **A** *un conjunto ortonormal en el es-*

pacio con producto interno **L**. Se dice que **A** es completo si no existe otro conjunto ortonormal que contiene a **A**, es decir, **A** debe ser un conjunto ortonormal maximal.

CRITERIO: Un conjunto ortonormal **A** es completo si y sólo si para toda **X** tal que $X \perp A$ (**X** es ortogonal a **A**), **X** debe ser cero.

DEMOSTRACION: Suponer que **A** es completo y que **X** es un vector diferente de cero tal que $X \perp A$. Claramente esto es una contradicción porque el conjunto ortonormal $A \cup \{\frac{X}{\| X \|}\}$ contiene a **A** propiamente y contradice la maximalidad de **A**.

Inversamente, suponer que $X \perp A$ implica $X = 0$. Si **A** no es completo debe de existir algún conjunto ortonormal **B** tal que $B \supset A$ propiamente.

Si esto es así sea $X \in B - A$. Como $\| X \| = 1$ y $X \perp A$, entonces la suposición de que existe un conjunto ortonormal **B** debe ser falsa y por tanto **A** debe ser completo.

El siguiente teorema muestra que existen conjuntos ortonormales completos en cualquier espacio con producto interno.

I.6.2. TEOREMA: Sea **L** un espacio con producto interno (**L** no es el espacio trivial que consiste sólo del vector cero). Entonces (1) Existen conjuntos ortonormales en **L**.

(2) Cualquier conjunto ortonormal puede ser extendido a un conjunto ortonormal completo.

DEMOSTRACION: *Es claro que si probamos (2) esto implicará (1) en virtud del hecho que en cualquier espacio deben existir conjuntos ortonormales; para un vector* **X** *diferente de cero, el conjunto* $\{\frac{X}{\|X\|}\}$ *es un conjunto ortonormal, por tanto probaremos (2).*

Sea **E** *un conjunto ortonormal dado y sea* **S** *la colección de todos los conjuntos ortonormales que contienen a* **E**. *Es claro que la colección* **S** *esta parcialmente ordenada por inclusión de conjuntos. Deseamos mostrar ahora que cada subconjunto totalmente ordenado de* **S** *tiene una cota superior en* **S**; **i. e.**, **S** *esta inductivamente ordenado. Podemos entonces aplicar el lema de Zorn, el cual establece que todo conjunto ordenado inductivamente debe tener un elemento máximo para obtener un conjunto ortonormal maximal.*

Sea **T** *un subconjunto totalmente ordenado de* **S**; *sea* $T = \{A_\alpha\}$ $(\alpha \in \Lambda)$.

Es claro que para cualquier α, $A_\alpha \subset \bigcup_\alpha A_\alpha$; *Ademas* $E \subset \bigcup_\alpha A_\alpha$.

Por tanto para mostrar que $\bigcup_\alpha A_\alpha \in S$ *sólo resta probar que éste es un conjunto ortonormal. Sean* **X** *y* **Y** *dos elementos de* $\bigcup_\alpha A_\alpha$, *implica que existen* A_α, A_β *tales que* $X \in A_\alpha$ *y* $Y \in A_\beta$. *Como* **T** *es totalmente ordenado, sin embargo sucede que* $A_\alpha \subset A_\beta$ *ó que* $A_\beta \subset A_\alpha$.

Suponiendo que la primera inclusión es cierta podemos decir

$X, Y \in A_\beta$, *implica* $X \perp Y$ ($\| X \| = \| Y \| = 1$) *y por tanto* $\bigcup_\alpha A_\alpha \in S$

Así el conjunto $\bigcup_\alpha A_\alpha$ *es una cota superior para* **T** *en* **S** *y hemos visto que* **S** *está inductivamente ordenado. Por el lema de Zorn debe existir un elemento maximal en* **S**. *Es claro que ningún otro conjunto ortonormal en el espacio contendría este elemento máximo, porque esto contradice su maximalidad.*

I.6.3. TEOREMA: *Sea* **H** *un espacio de Hilbert, suponer que* $A = \{X_\alpha\}$, $\alpha \in \Lambda$, *es un conjunto ortonormal en* **H** *y sea* **X** *un vector arbitrario en* **H**. *Entonces tenemos las siguientes proposiciones*

(1) $Y = \sum_{\alpha \in \Lambda} (X, X_\alpha) X_\alpha$ *existe*, i.e. *la serie es sumable;*

(2) *El vector* **Y** *pertenece a* $\overline{[A]}$;

(3) $X \in \overline{[A]}$ *si y sólo si* $X = Y$; i. e. **X** *se puede escribir como la serie anterior.*

(4) $X - Y \perp \overline{[A]}$.

DEMOSTRACION: (1) *Notemos que* $\sum_\alpha \| (X, X_\alpha) X_\alpha \|^2$ *converge, ya que* $\sum_\alpha \| (X, X_\alpha) X_\alpha \|^2 = \sum_\alpha |(X, X_\alpha)|^2 \leq \| X \|^2$, *por la desigualdad de Bessel (Teorema I.4.1. (1)).*

Como esto es así, usando el mismo argumento al probar el teorema I.1.2., para $\varepsilon > 0$ *debe existir un conjunto finito* $\kappa \subset \Lambda$ *tal que, para un subconjunto finito* $J \subset \Lambda$ *donde* $J \cap \kappa = \phi$,

$$\sum_{\alpha \in J} \| (X, X_\alpha) X_\alpha \|^2 < \varepsilon \quad \textit{pero}$$

$\| \sum_{\alpha \in J} (X, X_\alpha) X_\alpha \|^2 = \sum_{\alpha \in J} \| (X, X_\alpha) X_\alpha \|^2$ *entonces podemos decir*

$$\| \sum_{\alpha \in J} (X, X_\alpha) X_\alpha \|^2 < \varepsilon.$$

y $\sum_\alpha (X, X_\alpha) X_\alpha$ *existe por 1, p. 160 de* [3].

DEMOSTRACION (2). *Es claro que cualquier suma parcial de* $\sum_\alpha (X, X_\alpha) X_\alpha$ *debe pertenecer a* $[A]$, *esto implica que el límite* Y *debe pertenecer a* $\overline{[A]}$.

DEMOSTRACION (4). *Sea* $X_\beta \in A$ *y consideremos*

$$(X - Y, X_\beta) = (X, X_\beta) - (Y, X_\beta) = (X, X_\beta) - (\sum_\alpha (X, X_\alpha) X_\alpha, X_\beta)$$

$$= (X, X_\beta) - (X, X_\beta) = 0 \ (\textit{ver } [3] \text{ p. } 160)$$

así $X - Y \perp A$ *implica* $X - Y \perp [A]$

Ahora para cualquier $Z \in \overline{[A]}$ *debe existir una sucesión de elementos de* $[A]$, $\{U_n\}$ *tal que*

$$Z = \lim_n U_n$$

lo que siginifica $(X - Y, U_n) = 0$ *para toda* n. *Como esto se -- cumple para cualquier* n, *esto también es cierto en el límite; así*

$$\lim_{n} (X - Y, U_n) = 0$$

Por la continuidad del mapeo producto interno (teorema 1.2.5.) las operaciones de límite pueden intercambiarse para obtener

$$(X - Y, Z) = 0,$$

lo que quiere decir que

$X - Y \perp \overline{[A]}$ lo que prueba (4).

DEMOSTRACION (3). Es claro que si $X = Y$, entonces $X \varepsilon \overline{[A]}$ por (2). Inversamente, suponer que $X \varepsilon \overline{[A]}$. Usando (2) $Y \varepsilon \overline{[A]}$, entonces $X - Y \varepsilon \overline{[A]}$. Pero por (4) $X - Y \perp \overline{[A]}$, lo que implica $X - Y = 0$, y se termina la demostración.

Adelantamos que en el siguiente teorema no tenemos necesidad de la estructura de un espacio de Hilbert; un espacio con producto interno ordinario sera suficiente. En palabras, el siguiente teorema dice que, si el espacio generado por un conjunto ortonormal es denso en el espacio total, el conjunto ortonormal debe ser completo.

I.6.4. TEOREMA: Sea **L** un espacio con producto interno y suponer que el conjunto ortonormal **A** tiene la propiedad que

$\overline{[A]} = L$. Entonces **A** es completo.

DEMOSTRACION: La demostración sera por contradicción. Se probará que suponiendo que las hipótesis son ciertas la conclu--

sión es falsa, lo que lleva a resultados incompatibles. *Si* **A** *no es completo, debe existir un vector* **X** *diferente de cero tal que* $\mathbf{X} \perp \mathbf{A}$ *implica* $\mathbf{X} \perp [\mathbf{A}]$*, lo que en cambio implica* $\mathbf{X} \perp \overline{[\mathbf{A}]}$*, pero como* $\overline{[\mathbf{A}]} = \mathbf{L}$ *implica* $\mathbf{X} \perp \mathbf{X}$.

En consecuencia **X** *debe ser* 0 *y llegamos a la contradicción deseada.*

El siguiente teorema prueba que si estamos trabajando en un espacio de Hilbert el inverso del teorema anterior es cierto.

I.6.5. TEOREMA: *Suponer que* $\mathbf{A} = \{\mathbf{X}_\alpha\}$, $\alpha \in \Lambda$ *es un conjunto ortonormal completo en el espacio de Hilbert* **H**. *Entonces*

$$\overline{[\mathbf{A}]} = \mathbf{H}.$$

DEMOSTRACION: *Probaremos la contraposición de este resultado. Supondremos que* **A** *es un conjunto ortonormal en* **H** *y que* $\overline{[\mathbf{A}]} \neq \mathbf{H}$*, lo que significa que existe algún elemento* $\mathbf{X} \in \mathbf{H} - \overline{[\mathbf{A}]}$. *Ahora, como* **H** *es un espacio de Hilbert podemos aplicar la parte (1) del teorema I.6.3. para garantizar la existencia de*

$$\sum_\alpha (\mathbf{X}, \mathbf{X}_\alpha)\, \mathbf{X}_\alpha = \mathbf{Y}.$$

Haciendo uso de la parte (4) del mismo teorema, podemos decir que $\dfrac{\mathbf{X} - \mathbf{Y}}{\|\mathbf{X} - \mathbf{Y}\|} \perp \overline{[\mathbf{A}]}$ *implica* $\dfrac{\mathbf{X} - \mathbf{Y}}{\|\mathbf{X} - \mathbf{Y}\|} \perp \mathbf{A}$

por la parte (3) del teorema I.6.3, $\| X - Y \| \neq 0$.

Hemos encontrado un vector diferente de cero ortogonal a todo **A**. *En vista de esto podemos decir que* **A** *no es completo porque*

$$A \subset \left\{\frac{X - Y}{\| X - Y \|}\right\} \cup A.$$

I.7 CONJUNTOS ORTONORMALES COMPLETOS Y LA IDENTIDAD DE PARSEVAL.

El siguiente resultado da una descripción equivalente de un conjunto ortonormal completo en un espacio de Hilbert. Notemos, sin embargo, que para probar la parte (1) del siguiente teorema necesitamos sólo un espacio con producto interno.

I.7.1. TEOREMA: *Sea* **L** *un espacio con producto interno y sea* $A=\{X_\alpha\}, \alpha \in \Lambda$, *un conjunto ortonormal en* **L**. *Entonces (1) Si para cualquier* $X \in L$.

(I.7.1.1.) $\| X \|^2 = \sum_\alpha |(X, X_\alpha)|^2$, *entonces* **A** *es completo.*

(2) Si **L** *es un espacio de Hilbert,* **A** *es un conjunto ortonormal completo, entonces la ecuación (I.7.1.1) vale para cualquier* $X \in L$.

OBSERVACION: *La igualdad de (I.7.1.1) es llamada identidad de Parseval.*

DEMOSTRACION (1). *Si suponemos que* **A** *no es completo, debe* existir un vector diferente de cero X tal que $X \perp A$. Substituyendo esta X en la identidad de Parseval, obtenemos

$$\| X \|^2 = \sum_\alpha |(X, X_\alpha)|^2 = 0,$$

lo que es una contradicción, consecuentemente A debe ser completo.

DEMOSTRACION. (2). Por el teorema I.6.5, la cerradura del subespacio generado por un conjunto ortonormal completo en un espacio de Hilbert debe ser el mismo espacio, por tanto

$$\overline{[A]} = L.$$

Para cualquier vector $X \in L$ tenemos que $X \in \overline{[A]}$. Aplicando el teorema I.6.3.(3), podemos decir que

$$X = \sum_\alpha (X, X_\alpha) X_\alpha$$

Para ver que la identidad de Parseval es cierta, calculemos.

$$\| X \|^2 = (X, X) = (\sum_\alpha (X, X_\alpha) X_\alpha, \sum_\beta (X, X_\beta) X_\beta) =$$

$$= \sum_\alpha \sum_\beta (X, X_\alpha) \overline{(X, X_\beta)} (X_\alpha, X_\beta)$$

(ver [3], p.160)

$$= \sum_\alpha |(X \;\; X_\alpha)|^2$$

Combinando el primero y último término, vemos que la identidad de Parseval se satisface.

Deseamos resumir todos los resultados anteriores con el si *guiente teorema. Notar que no obstante que se enuncie un espacio de Hilbert, esto no es necesario en todos los casos.*

I.7.2. TEOREMA. *Sea* $\mathbf{A} = \{X_\alpha\}\ \alpha \in \Lambda$ *un conjunto ortonormal en el espacio de Hilbert* $\mathbf{H}$ *. Entonces:*

(1) $\mathbf{A}$ *es completo*

si y sólo si (2) $X \perp \mathbf{A}$ *implica* $X = 0$;

si y sólo si (3) $X \in \mathbf{H}$ *implica* $X = \sum_\alpha (X, X_\alpha) X_\alpha$;

si y sólo si (4) $\overline{[\mathbf{A}]} = \mathbf{H}$

si y sólo si (5) $\| X \|^2 = \sum_\alpha |(X, X_\alpha)|^2$

si y sólo si (6) *para* $X, Y \in \mathbf{H}$,

$$(X, Y) = \sum_\alpha (X\ \ X_\alpha)\ (X_\alpha, Y).$$

CAPITULO II

PROCESOS ESTOCASTICOS

II.1 DEFINICION: *Sean* (Ω, β, P) *un espacio de probabilidad y* (S, φ) *un espacio medible y* T *un conjunto arbitrario. Un proceso estocástico en* (Ω, β, P) *con espacio de estados* (S, φ) *y un conjunto de indices* T, *es una familia de funciones medibles*

$$X_t : (\Omega, \beta) \to (S, \varphi), \; t \in T.$$

Notamos que si T *es el conjunto de los enteros positivos y* $S = \mathbb{R}$ *y* $\varphi = \beta(\mathbb{R})$ *el proceso estocástico* $\{X_t, t \in T\}$ *es una sucesión de variables aleatorias reales.*

Similarmente podemos obtener un vector aleatorio n-dimensional $(T = \{1, 2 \ldots n\})$, $S = \mathbb{R}$, $\varphi = \beta(\mathbb{R})$, *e igualmente una sola variable aleatoria* $(n = 1)$.

Un sinónimo para procesos estocásticos es variable aleatoria. Ahora se tratara de explicar ésta terminologia.

Sea S^T *la colección de todas las funciones de* T *en* S *y sea* φ^T *el producto de σ-álgebras en* S^T *(Recordemos que* φ^T *es la mínima σ-álgebra que contiene a los rectángulos*

$$\{W \in S^T : (W(t_1) \in B_1, \ldots W(t_n) \in B_n)\}$$

$t_1, t_2, \ldots t_n \in T$, $B_1, B_2 \ldots B_n \in \varphi$, $n = 1, 2 \ldots$ *ver* [1] *pag.* 189).

Ahora supongamos que para cada $t \in T$ *tenemos una función* $X_t: \Omega \to S$; *y definimos a*

$$X: \Omega \to S^T \quad \textit{por}$$

$$X(W) = (X_t(W), \ t \in T)$$

Entonces X *es una función medible de* (Ω, β) *en* (S^T, φ^T) *si y solo si cada* X_t *es medible de* (Ω, β) *en* (S, φ) (*ver* [8] *pag.* 72). *Así un proceso estocásticoesunmapeo medible de* Ω *en el espacio de funciones* S^T.

Intuitivamente la realización de un experimento produce un punto muestral W *y si repetimos varias veces el experimento nos determinará una colección de elementos* $X(t, W) = X_t(W)$, $t \in T$. *En otras palabras el resultado del experimento determinará una función de* T *en* S, *llamada la función de muestra correspondiente al punto* W.

Si T *es un intervalo de la recta real nos ayudará para visualizar a* t *como un parámetro del tiempo y pensar a un proceso como una variable aleatoria del tiempo.*

II.2 PROCESOS ESTACIONARIOS

II.2.1. DEFINICION: *Un proceso* $X_1, X_2, \ldots,$ *es estacionario si* $\forall K > 0$, *el proceso* $X_{k+1}, X_{k+2} \ldots,$ *tiene la misma distri* *bución que* X_1, X_2,..., i.e. *para toda* $B \in \beta_\infty$

$$(1)\ldots P((X_1, X_2 \ldots) \in B) = P((X_{k+1}, X_{k+2}, \ldots) \in B)$$

Como la distribución está determinada por las funciones de distribución (1) es equivalente a: Para toda $X_1, X_2, \ldots, X_n$ *y* $K > 0$

$$(2)\ldots\ P(X_1 < x_1 \ldots X_n < x_n) = P(X_{k+1} < x_1 \ldots X_{k+n} < x_n)$$

ahora ésto es cierto ya que:

$$P(X_1 < x_1,\ X_2 < x_2 \ \ldots\ X_n < x_n) = P(X_1 < x_1 \ldots X_n < x_n,\ X_{n+1} < \infty,$$

$$X_{n+2} < \infty \ldots)$$

$$= P(X_1 \in (-\infty, x_1),\ X_2 \in (-\infty, x_2),\ \ldots X_n \in (-\infty, x_n),\ X_{n+1} \in (-\infty, \infty),$$

$$X_{n+2} \in (-\infty, \infty) \ldots)$$

$$= P((X_1, X_2, \ldots) \in \prod_{i=1}^{n} I_i \cdot \prod_{i=n+1}^{\infty} \mathbb{R}_i)$$ *donde* $I_i = (-\infty, x_i)$ *además*

$$\prod_{i=1}^{n} I_i \ \prod_{i=n+1}^{\infty} \mathbb{R}_i = B \in \beta_\infty$$

$$\therefore P(X_1 < x_1 \ \ldots\ X_n < x_n) = P((X_1\ X_2 \ldots) \in B)$$

Inversamente

Sea $B = I^\infty \in \beta_\infty$, $I^\infty = \prod_{i=1}^{\infty} I_i$, *donde los* I_i *son intervalos de la forma* $(-\infty, x_i)$ *y como sabemos que* $\beta_\infty = \sigma((-\infty, x_i),\ i \in \mathbb{N})$.

Sin perder generalidad podemos suponer que $I_1, I_2 \ldots I_n \subsetneq \mathbb{R}$ *y* $I_{n+1}, I_{n+2} \ldots = \mathbb{R}$

$\therefore P((X_1, X_2 \ldots) \in B) = P(X_1 \in I_1, X_2 \in I_2, \ldots X_n \in I_n, X_{n+1} \in I_{n+1} \ldots)$

$= P(X_1 \in (-\infty, x_1), X_2 \in (-\infty, x_2) \ldots X_n \in (-\infty, x_n), X_{n+1} \in (-\infty, \infty) \ldots) =$

$= P(X_1 < x_1, X_2 < x_1 \ldots, X_n < x_n).$

En particular si un proceso es estacionario, entonces todas las funciones de distribución unidimensionales son iguales i.e.

$$P(X_1 < x) = P(X_k < x) \;\forall\; K = 1, 2, \ldots$$

demostraremos esta afirmación.

Por hipótesis

$P(X_1 < x_1, X_2 < x_2, \ldots X_n < x_n) = P(X_{k+1} < x_1 \ldots X_{k+n} < x_n)$

ahora usando la propiedad de consistencia (ver II.2.5.) de funciones de distribución obtenemos que:

$$P(X_k < x_1) = \lim_{\substack{x_i \uparrow \infty \\ i = 2,3,\ldots}} P(X_{k+1} < x_1, \ldots X_{k+n} < x_n) =$$

$$= \lim_{\substack{x_i \uparrow \infty \\ i = 2,3,\ldots}} P(X_1 < x_1, X_2 < x_2 \ldots X_n < x_n) = P(X_1 < x_1)$$

II.2.2. PROPOSICION. *Un proceso* $X_1, X_2, \ldots,$ *es estaciona_rio, si el proceso* $X_2, X_3, \ldots,$ *tiene la misma distribución que* $X_1, X_2, \ldots$

DEMOSTRACION: *Sea* $X'_k = X_{k+1}$, $K = 1, 2, \ldots$ *y sea* $B \in \beta_\infty$, - *entonces por hipótesis tenemos que*

$$P[(X_1, X_2, \ldots) \in B] = P[(X'_1, X'_2 \ldots) \in B],$$

o sea, que $X_1, X_2, \ldots$ *tiene la misma distribución que* $X'_1, X'_2, \ldots$ *también* $P[(X'_2, X'_3, \ldots) \in B] = P[(X'_1, X'_2, \ldots) \in B]$; *por la definición de las* X'_k *i. e.* $X'_1, \ldots$ *tiene la misma distribución que* - $X'_2, X'_3, \ldots$ *y así sucesivamente.*

$$\therefore P[(X_1, X_2, \ldots) \in B] = P[(X_{k+1}, X_{k+2}, \ldots) \in B]$$

$$\therefore X_1, X_2, \ldots \text{ es estacionario.}$$

Algunas veces es mas conveniente ver procesos estacionarios que consisten de variables aleatorias bilaterales, $\ldots X_{-1}, X_0, X_1, \ldots$ *En el contexto que se tiene, esto se podría ver como una sucesión de lecturas empezando en el infinito pasado y continuado hasta el futuro infinito.*

II.2.3. DEFINICION: *Un proceso bilateral es estacionario, - si su distribución no depende de la elección del origen, i.e., en términos de distribuciones,*

$$P(X_1 < x_1 \ldots, X_n < x_n) = P(X_{k+1} < x_1 \ldots X_{k+n} < x_n)$$

$$\forall\ x_1, \ldots, x_n \ y \ \forall K \in \mathbb{Z}.$$

II.2.4. **PROPOSICION:** *Dado un proceso estacionario unilateral* $X_1, X_2, \ldots,$ *existe un proceso estacionario bilateral* ... $\hat{X}_{-1}, \hat{X}_0, \hat{X}_1 \ldots$ *tal que* $\hat{X}_1, \hat{X}_2 \ldots$ *y* $X_1, X_2, \ldots$ *tienen la misma distribución.*

Para la demostración de este resultado, enunciaremos las siguientes proposiciones, las cuales se encuentran en[4] pag, 27-28.

II.2.5. **PROPOSICION:** *Las funciones de distribucion* $F_n(x)$ *satisfacen las siguientes condiciones:*

i) *No-negatividad; para intervalos finitos*

$$P(X_1 \in I_1, \ldots, X_n \in I_n) \geq 0$$

ii) *Continuidad por abajo:* *Si* $X^{(k)} = (X_1^{(k)} \ldots X_n^{(k)})$ *y* $X_j^{(k)} \uparrow X_j$, $j = 1, 2, \ldots, n$ *entonces*

$$F_n(x^{(k)}) \uparrow F_n(x), \quad x = (x_1, \ldots, x_n)$$

iii) *Normalización:* *Todos los límites de* F_n *existen, cuando* $x_j \uparrow +\infty$ *ó* $x_j \downarrow -\infty$.

Si $x_j \downarrow -\infty$, *entonces* $F_n(x) \to 0$. *Si todas las* $x_j \uparrow +\infty$, $j = 1,\ldots, n$, *entonces* $F_n(x) \to 1$.

El conjunto de funciones de distribución estan conectadas por:

iv) Consistencia:

$$\lim_{x_n \uparrow \infty} F_n(x_1,\ldots, x_n) = F_{n-1}(x_1,\ldots, x_{n-1}), \quad n > 1.$$

II.2.6. TEOREMA: *Dado un conjunto de funciones* $\{F_n(x)\}$ *que satisfacen i), ii), iii) y iv) de* II.2.5., *existen un proceso* $\{X_n\}$ *sobre* (Ω, β, P) *tal que*

$$P(X_1 < x_1,\ldots, X_n < x_n) = F_n(x_1, x_2,\ldots, x_n).$$

Ahora pasaremos a demostrar la proposición II.2.4.

Para cada n *se tiene*

$$F_n(x_1,\ldots, x_n) = P(X_1 < x_1,\ldots, X_n < x_n)$$

$\therefore$ *tenemos un conjunto de funciones* $\{F_n(x)\}$ *que resultan ser consistentes.*

$\therefore \exists \ \{\hat{X}_{-m}\}$ *tal que*

$$P(\hat{X}_{-m} < x_1,\ldots, \hat{X}_{-(m+1)+n} < x_n) = F_n(x_1,\ldots, x_n)$$

$\therefore$ *Se tienen* $\{F_{\hat{X}_{-m}\ldots\hat{X}_{-(m+1)+n}}\}$, *aplicando otra vez el teorema -*

tenemos que $\exists$ $\{\hat{X}_n\}$ *tal que.*

$$P(\hat{X}_1 < x_1, \ldots, \hat{X}_n < x_n) = F_{\hat{X}_{-m} \cdots \hat{X}_{-(m+1)+n}}(x_1, x_2, \ldots, x_n)$$

$$\therefore P(\hat{X}_1 < x_1, \ldots, \hat{X}_n < x_n) = P(\hat{X}_{-m} < x_1, \ldots, \hat{X}_{-(m+1)+n} < x_n) =$$

$$= P(X_1 < x_1, \ldots, X_n < x_n)$$

$\therefore$ *el proceso* $\{\hat{X}_{-m}\}$ *es estacionario y además*

$$P(\hat{X}_1 < x_1, \ldots, \hat{X}_n < x_n) = P(X_1 < x_1, \ldots, X_n < x_n)$$

i. e. $X_1, X_2, \ldots$ *y* $\hat{X}_1, \hat{X}_2, \ldots$ *, tienen la misma distribución.*

II.2.7. **PROPOSICION:** *Sea* $X_1, X_2, \ldots$ *un proceso estacionario* $\varphi(x)$ *una función* β_∞ *medible, entonces el proceso* $Y_1, Y_2, \ldots$ *definido por*

$$Y_k = \varphi(X_k, X_{k+1}, \ldots) \text{ es estacionario.}$$

DEMOSTRACION: *Sobre* $\mathbb{R}^\infty$ definir $\varphi_k(x)$ *como*

$$\varphi_k(x) = \varphi(x_k, x_{k+1}, \ldots) = Y_k$$

$$\text{i.e. } \varphi_1(x) = \varphi(x_1, x_2, \ldots) = Y_1$$

$$\varphi_2(x) = \varphi(x_2, x_3, \ldots) = Y_2$$

$$\vdots$$

Sea $\mathbf{B} \in \beta_\infty$, *el conjunto* $\mathbf{A} = \{x: (\varphi_1(x), \ldots, \varphi_n(x), \ldots) \in \mathbf{B}\}$ *i. e.* $\mathbf{A} \subset \beta_\infty$, *porque cada* $\varphi_k(x)$ *es una variable aleatoria sobre* $(\mathbb{R}^\infty, \beta_\infty)$.

Los conjuntos $\mathbf{R} = \{W: (Y_1, Y_2, \ldots) \in \mathbf{B}\}$ *y* $\mathbf{S} = \{W: (X_1, X_2, \ldots) \in \mathbf{A}\}$ *son iguales, ya que:*

Sea $\mathbf{W} \in \mathbf{S} \Rightarrow (X_1(W), X_2(W), \ldots) \in \mathbf{A}$, *Sea* $X = (X_1(W), X_2(W), \ldots)$ $\therefore X \in \mathbf{A} \Rightarrow (\varphi_1(x), \varphi_2(x), \ldots) \in \mathbf{B} \Rightarrow (Y_1, Y_2, \ldots) \in \mathbf{B} \Rightarrow \mathbf{W} \in \mathbf{R} \therefore$ $\mathbf{S} \subset \mathbf{R}$.

Ahora sea $W \in R \Rightarrow (Y_1(W), Y_2(W), \ldots) \in B \Rightarrow (\varphi_1(X(W)), \varphi_2(X(W)), \ldots) \in B$, *ahora* $X = (X_1, X_2, \ldots) \Rightarrow (\varphi(X_1(W), X_2(W), \ldots), \varphi(X_2(W), X_3(W), \ldots), \ldots) \in \mathbf{B} \Rightarrow (X_1(W), X_2(W), \ldots) \in \mathbf{A}$, *i.e.* $X(W) \in \mathbf{A} \Rightarrow W \in S$

$$\therefore \quad R \subset S$$

$$\therefore \quad R = S.$$

tambien se tiene $\{W: (Y_2, Y_3, \ldots) \in B\} = \{W: (X_2, X_3, \ldots) \in A\}$, *con una demostración similar a la hecha para* $R = S$.

Y así sucesivamente obtenemos que $\{W: (Y_k, Y_{k+1}, \ldots) \in B\} = \{W: (X_k, X_{k+1}, \ldots) \in A\}$ *además sabemos que el proceso* $X_1, X_2, \ldots$ *es estacionario, por tanto tenemos que*

$$P((X_1, X_2, \ldots) \in A) = P((X_{k+1}, \ldots, X_{k+n}, \ldots) \in A\}$$

$$= P((Y_1, Y_2, \ldots) \in B) = P((Y_{k+1}, Y_{k+2}, \ldots) \in B\}$$

de donde tenemos que la sucesión $Y_1, Y_2, \ldots$ *es estacionaria.*

II.2.8. COROLARIO: *Sean* $X_1, X_2, \ldots,$ *variables aleatorias independientes identicamente distribuidas,* $\varphi(x)$ *una función* B_∞ *medible, entonces*

$$Y_k = \varphi(X_k, X_{k+1}, \ldots) \text{ es estacionaria.}$$

DEMOSTRACION: *La sucesión* $X_1, X_2, \ldots$ *es estacionaria.*

CAPÍTULO III: TEORÍA ERGÓDICA

III.1. MOTIVACION Y DEFINICIONES

El punto de partida de la teoría Ergódica es la noción de preservación de una medida, bajo una transformación definida como sigue:

III.1.1. DEFINICION. Sea (Ω,β,μ) un espacio con medida y T una transformación medible del espacio (Ω,β) en sí mismo, ésto es $T: (\Omega,\beta) \to (\Omega,\beta)$. Se dice que la transformación T preserva la medida μ si y solo si $\mu(T^{-1} A) = \mu(A)$ para todo $A \in \beta$. Esto implica que $\mu(T^{-K} A) = \mu(A), K=1,2,\ldots$ donde $T^{-K} A = \{W: T^k W \in A\}$ y T^K es la composición de T con sí misma k veces.

El concepto físico de un flujo puede ser usado para motivar el estudio de las transformaciones que preservan medidas. Un flujo puede ser considerado como un proceso en el cual un sistema de partículas de un fluído (cada punto del fluído corresponde a una partícula) se mueve por la acción de una fuerza externa. La fuerza se supone que es independiente del tiempo y tal que, a lo más en tiempos discretos $t=0,1,2,\ldots$ el flujo puede ser descrito por una función simple T (medible). Si x es un punto del fluido $T\,x$ será la posición de la partícula después de que ha transcurrido un segundo; así $T^2\,x = T(T\,x)$ es la posición después de dos segundos y así sucesivamente. Si A es un subconjunto (Boreliano) del fluido, entonces $T^{-1}\,A$ corresponde al conjunto de partículas que estarán en A después de un segundo.

Si μ es la medida de Lebesgue (volumen en este caso) y el fluido es no compresible es razonable esperar que $\mu(T^{-1}A) = \mu(A)$.

III.1.2. EJEMPLOS

a). Sea Ω un conjunto finito $\{x_1, x_2, \ldots x_n\}$, $n > 2$ donde β consiste de todos los subconjuntos de Ω. Sea T una permutación cíclica de Ω, decimos $T\,x_i = x_{i+1}$, con índices reducidos módulo n, esto es $T\,x_i = x_{i+1}$ para $i = 1, \ldots n-1$ y $T\,x_n = x_1$. Como $T^{-1}\{x_i\} = \{x_{i-1}\}$, T preserva μ si y solo si $\mu\{x_i\}$ es constante $\forall i$. Así si μ es una medida de probabilidad P, entonces $P\{x_i\}$ debe ser $1/n$, $\forall i$. Pasaremos a demostrar tal afirmación:

Dem.- Demostraremos la necesidad de la afirmación ya que la suficiencia es inmediata.

$$\mu(T^{-1}\{x_i\}) = \mu(T^{-K}\{x_i\}) = \mu(x_{i-K}) \;\forall\, k = 1,2,\ldots$$

Por tanto μ *es constante.*

b). *Sea* $\Omega = \mathbb{R}$, $\beta = \beta(\mathbb{R})$, *y sea* μ *la medida de Lebesgue.* - *Si* $T\,x = x + c$, *donde* c *es una constante, entonces* T *preserva* μ *ya que* μ *es invariante bajo traslaciones. Pasaremos a demostrar tal afirmación:*

$$\mu(T^{-1}[a,b]) = \mu([a-c,\ b-c]) = (b-c) - (a-c) =$$

$$= b - a = \mu([a,b])$$

c). *Sea* Ω *el círculo unitario en el plano* $\mathbb{R}^2$ *(*Ω *puede - identificarse con el intervalo* $[0,2\pi)$ *bajo la correspondencia* $e^{i\theta} \to \theta$*). Tomar* β *como los conjuntos Borel y* $\mu = P = \frac{\lambda}{2\pi}$, *donde* λ *es la longitud de arco del círculo (o la medida de Lebesgue sobre* $[0,2\pi)$*). Así, si* A *es un subconjunto de Borel de* $[0,2\pi)$, $\mu(A) =$

$$= \int_A (2\pi)^{-1}\, d\theta.$$

Sea α *fijo en* $[0,2\pi)$ *y sea* T *la rotación por* α*; es decir,* T *está definida en* Ω *por* $T(e^{i\theta}) = e^{i(\theta+\alpha)}$ *o equivalente sobre* $[0,2\pi)$ *por* $T(\theta) = \theta + \alpha$ *(módulo* 2π*). Como en* b) T *preserva a* μ *por la invarianza de la medida de Lebesgue bajo traslaciones.*

III.1.3. PROPOSICION. *Sea* T *una transformación que preser-*

va medida sobre (Ω, β, P), i.e. $T:(\Omega,\beta) \to (\Omega,\beta)$, X *una variable aleatoria sobre* (Ω,β), *entonces la sucesión* $X_n(\omega) = X(T^{n-1}\omega)$; $n=1,2,\ldots$ *es una sucesión estacionaria de variables aleatorias.*

DEMOSTRACION. *Antes que todo* $X_n(\omega)$ *es variable aleatoria por que* $\{X_n(\omega)\varepsilon B\} = \{X(T^{n-1}\omega)\varepsilon B\}$ *-La composición de funciones medibles es una función medible-* Con $B \in \beta$.

Sea $A = \{X \varepsilon B\}$ *entonces los conjuntos*

$R = \{X(T^{n-1}\omega)\varepsilon B\}$ *y* $S = \{\omega | T^{n-1}\omega \varepsilon A\} = \{\omega | X(\omega)\varepsilon B\}$

son iguales.

DEMOSTRACION: $R = S$.

Sea $\omega \varepsilon S$, *entonces* $T^{n-1}\omega \varepsilon A \Rightarrow X(T^{n-1}\omega)\varepsilon B \subset R$ *por -- tanto* $S \subset R$. *Sea* $X(T^{n-1}\omega)\varepsilon B$ *y* $\omega^1 = T^{n-1}\omega$, *por tanto* $X(\omega^1)\varepsilon B \Rightarrow$ *por la definición de* $A, \omega^1 \varepsilon A$, i.e. $(T^{n-1}\omega)\varepsilon A \Rightarrow R \subset S$, $\therefore R = S$.
También tenemos que si T *es medible, entonces* T^{n-1} *es medible, es decir;* $T^{(-n+1)} A \varepsilon \beta$.

Sean $A = \{\omega | (X_1, X_2, \ldots)\varepsilon B\}$, $B \varepsilon \beta_\infty$, *o sea*

$A = \{\omega | (X(\omega), X(T\omega), \ldots)\varepsilon B\}$ *y*

$A_1 = \{\omega | (X_2, X_3, \ldots)\varepsilon B\} = \{\omega | (X(T\omega), X(T^2\omega), \ldots)\varepsilon B\}$.

Así se tiene que $\omega \varepsilon A_1 \Leftrightarrow T\omega \varepsilon A$.

$\omega \in A_1 \Longleftrightarrow (X(T\omega), X(T^2\omega), \ldots) \in B \Longleftrightarrow T\omega \in A$, *es decir;* $A_1 = T^{-1} A = \{\omega | T\omega \in A\}$, *pero por hipótesis* $P(T^{-1}A) = P(A)$ *es decir;* $P(A_1) = P((X_1, X_2, \ldots) \in B)$ *-por la proposición II.2.3-*

Por lo tanto la sucesión es estacionaria.

III.1.4. DEFINICION: *Sea* $(\mathbb{R}^\infty, \beta_\infty, P)$ *un espacio con probabilidad y* T *la transformación definida sobre* $(\mathbb{R}^\infty, \beta_\infty)$, -- $(T : \mathbb{R}^\infty \to \mathbb{R}^\infty)$ *por* $T(X_0, X_1 \ldots) = (X_1, X_2, \ldots)$. *A esta transformación se le llama el* <u>*corrimiento unilateral*</u>.

III.1.5. PROPOSICION. *La transformación corrimiento es medible y preserva a* P *si y sólo si* P *es estacionaria. Es decir, si y sólo si* $P(X | (X_0, X_1, \ldots X_{n-1}) \in B_n) = P(X | (X_k, X_{k+1}, \ldots X_{k+n-1}) \in B_n)$; $\forall n$, $k = 1, 2 \ldots$ *y todo* $B_n \in \beta_n$.

DEMOSTRACION. *Sea* $X_k\ (x) = x_k$, $k = 0, 1, 2, \ldots$ *La estacionalidad de* P *es equivalente a la estacionalidad del proceso* $X_0, X_1, \ldots$

Ahora notemos que

$T^{-k}\{X | (X_0, X_1, \ldots, X_{n-1}) \in B_n\} = \{X_k, X_{k+1}, \ldots X_{k+n-1}) \in B_n\}$ *si* T *preserva a* P *entonces* $P(A) = P(T^{-1} A) = \ldots = P(T^{-k} A)$, $A \in \beta_\infty$. *Por tanto* P *es estacionaria, inversamente si* P *es estacionaria y* -- $A = \{X | (X_0, X_1 \ldots X_{n-1}) \in B_n\}$, $A \in \beta_\infty$. *Entonces*

$T^{-1}A = \{X \mid (X_1, \ldots, X_n) \varepsilon B_n\}$ *por lo tanto*

$P(A) = P(T^{-1} A)$.

y como la clase M *de conjuntos* $A\varepsilon\ \beta_\infty$ *tal que* $P(A) = P(T^{-1} A)$ *es una clase monótona que contiene a las cilindros medibles y como* β_∞ *es la mínima σ-álgebra que contiene a los cilindros, entonces* $M = \beta_\infty$. *(Ver [8] página 14).*

Sea $D = T^{-1}\{X \mid X_0 \varepsilon I_0, \ldots, X_n \varepsilon I_n\} = T^{-1}\ C \varepsilon \beta_\infty$.

Por tanto T *es medible.*

Ahora consideraremos un ejemplo físico para motivar el concepto de una transformación ergódica.

Supongamos que los datos de la cantidad de lluvia son coleccionados de un número grande de puntos de observación a_0, $a_1, \ldots$ *en los tiempos* $t=0,1,\ldots$. *Supongamos también que el carácter estadístico de las observaciones en* a_i *es el mismo para toda* i *y están representados para cada* a_i *por una sucesión de variables aleatorias estacionarias* $X_0, X_1 \ldots$ *donde* X_n *denota la cantidad de lluvia en el tiempo* n *para un punto* a_i. *Supongamos que las* a_i *son "independientes" en otras palabras que a las* a_i *les corresponden una sucesión de ejecuciones independientes de un experimento no determinístico, donde una ejecución significa una observación de una sucesión de variables aleatorias* $\{X_0, X_1 \ldots\}$.

Supongamos que el problema es medir el promedio de la cantidad de lluvia. Un científico que denotaremos por A puede tomar la siguiente aproximación: tomar medidas en cada punto de observación en un tiempo $t = 0$ y tomar el promedio de los resultados. Otro científico denotado por B puede razonar como sigue: Como todos los puntos de observación tienen las mismas características estadísticas, podemos simplemente ir a un punto de observación en los tiempos $t = 0, 1 \ldots n$ y promediar los resultados. El científico A está usando el hecho de que puede aplicar un esquema de medibilidad vertical, y el científico B un esquema horizontal como se ilustra en el siguiente esquema

Puntos de observación	Medidas t = 0	1	2...
a_0	X_{00}	X_{01}	$X_{02} \ldots$
a_1	X_{10}	X_{11}	$X_{12} \ldots$
a_2	X_{20}	X_{21}	$X_{22} \ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Las observaciones de A corresponden a la primera columna y las de B al primer renglón.

Ahora A y B no obtendrán necesariamente el mismo resultado. Por ejemplo supongamos que la "naturaleza" lanza una moneda honesta en cada punto de observación. Si la moneda cae - sol no hay lluvia nunca. Si la moneda cae águila la cantidad de lluvia es de un centímetro en todos los tiempos de observación. La mitad de las veces "aproximadamente" de las observaciones que haga A obtendrá un centímetro y la otra mitad cero. Pero B obtendrá un promedio de un centímetro o de cero, de donde no obtendrán el mismo resultado.

Matemáticamente estamos en la situación descrita en III.1.4 donde B está calculando un promedio en el tiempo o sea

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k\ (W) \text{ para una particular } W$$

(Notemos también que si T es el corrimiento unilateral, - $X_k(W) = X_0(T^k W)$). El científico A observa y toma un promedio en un tiempo particular, llamemosle a ése promedio:

$$n^{-1}(Y_0 + Y_1 \ldots + Y_{n-1})$$

Donde las Y_i son variables aleatorias independientes todas teniendo la misma distribución que X_0. Así el resultado de A se aproximará a:

$$EX_0 = \int_\Omega X_0 \, dP$$

Debido a la ley débil de los grandes números.

Para que A *y* B *obtuviesen el mismo resultado deberíamos tener*

$$\frac{1}{n} \sum_{k=0}^{n-1} X_0 \, (T^k W) \longrightarrow \int_\Omega X_0 \, dP.$$

En general nos podemos preguntar bajo que condiciones para una función f *integrable en* $L^1(\Omega,\beta,P)$ *se tiene*

(1)...
$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k W) \xrightarrow{c.d} \int_\Omega f \, dP.$$

En particular, si f *es el indicador* I_A *de un conjunto medible* A, *la propiedad* (1) *nos dice simplemente que el número promedio de visitas de* W *a* A *en los primeros* n *pasos converje a* P(A) *esto es*

(2)...
$$\frac{1}{n} \sum_{k=0}^{n-1} I_A(T^k W) \longrightarrow \int_\Omega I_A \, dP = P(A)$$

Ahora supongamos que A *es un conjunto casi invariante, en otras palabras* A *y* T^{-1} A *difieren sólo por un conjunto de medida cero. (Cuando la naturaleza lanza una moneda para determi--*

nar la cantidad de lluvia, podemos tomar $A = \{W : X_0(W) = 1\}$ *y como tenemos las siguientes relaciones*

$$T^{-1}A = \{W : TW \in A\} = \{W : TW \in \{W : X_0(W) = 1\}\} = \{W : X_0(TW) = 1\} =$$

$$= \{W : X_1(W) = 1\}$$

$$\therefore T^{-1}A = \{W : X_1(W) = 1\}$$

$$\therefore W \in A \iff TW \in A$$

$$\therefore \frac{1}{n}\sum_{k=0}^{n-1} I_A(T^k W) = I_A(W) \;\forall n.$$

Así la frecuencia relativa de visitas a A *no puede converger a* P(A) *excepto cuando* P(A) = 0 *o* P(A) = 1, *inversamente el teorema Ergódico puntual III.3. implica que los anicos conjuntos invariantes son aquellos que tienen probabilidad* 0 *o* 1 *cuando la convergencia (1) se da. En la siguiente sección nos prepararemos para la demostración de éste resultado básico dentro de la Teoría Ergódica.*

III.2. CONJUNTOS INVARIANTES Y ERGODICIDAD

III.2.1. DEFINICION: *Sea* T *una transformación que preserva medida sobre* (Ω, β, P). *Un conjunto es* ***invariante*** *(bajo* T*) si y solo si* $T^{-1}A = A$, *ésto es* $W \in A$ *si y sólo si* $TW \in A$. *O* -

sea si A *es invariante entonces* $T: \Omega \to \Omega$ *lleva (manda)* A *en* - A.

De igual forma un conjunto A *es casi invariante si y so_lo si* A *y* $T^{-1}A$ *difieren por un conjunto de medida cero, en - otras palabras* $P(A \Delta T^{-1}A) = 0$.

III.2.2. PROPOSICION. *La clase de conjuntos invariantes es una* σ-*álgebra* I.

DEMOSTRACION. *Sea* $I(T)$ *la colección de conjuntos invariantes bajo* T

a) $\Phi \in I(T)$ *ya que* $T^{-1}(\Phi) = \{W : TW \in \Phi\} = \Phi$

b) *Sea* $A \in I(T) \therefore T^{-1}A = A \therefore T^{-1}(A^c) = (T^{-1}A)^c = A^c$

$\therefore A^c \in I(T)$.

c) *Sea* $\{A_n\} \in I(T)$, *entonces* $T^{-1}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} T^{-1} A_n =$

$= \bigcup_{n=1}^{\infty} A_n \therefore \bigcup_{n=1}^{\infty} A_n \in I(T) \therefore I(T)$ *es* σ-*álgebra.*

III.2.3. DEFINICION. *Sea* T *una transformación que preser_va medida sobre* (Ω, β, μ), *se dice que* T *es* *ergódica* *si para cada* $A \in I(T)$ $\mu(A) = 0$ *o* $\mu(\Omega - A) = 0$ *es decir si* $I(T) = \{\Phi, \Omega\}$.

Una pregunta que pudiera surgir con respecto a conjuntos *invariantes y conjuntos casi invariantes es ¿Es la clase de - conjuntos invariantes considerablemente diferente de la clase de conjuntos casi invariantes?.*

La respuesta es no como se muestra en el siguiente lema.

III.2.4. LEMA.

a) *Si* A *es un conjunto casi invariante, entonces, existe un conjunto invariante* B *tal que*

$$P(A \Delta B) = 0$$

b) *Una transformación* T *es ergódica si y sólo si para cada conjunto* A *casi invariante se tiene que*

$$\mu(A) = 0 \text{ o } \mu(\Omega - A) = 0$$

DEMOSTRACION

a) *Sea* $B = \lim_n \operatorname{Sup}(T^{-n} A) = \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} T^{-k} A$ *y sea* $A_n =$ $= \bigcup_{k=n}^{\infty} T^{-k} A$, $n = 0,1,\dots$ $\therefore B = \bigcap_{n=0}^{\infty} A_n$ *y* $T^{-1} B = T^{-1}(\bigcap_{n=0}^{\infty} A_n) =$ $= \bigcap_{n=0}^{\infty} T^{-1} A_n = \bigcap_{n=0}^{\infty} T^{-1} (\bigcup_{k=n}^{\infty} T^{-k} A) = \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} T^{-1}(T^{-k} A) =$ $= \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} T^{-(k+1)} A = \bigcap_{n=0}^{\infty} A_{n+1}$ *pero tenemos que* $\{A_n\}$ *es decreciente* $\therefore \bigcap_{n=0}^{\infty} A_{n+1} = \bigcap_{n=0}^{\infty} A_n$

$\therefore T^{-1} B = \bigcap_{n=0}^{\infty} A_{n+1} = \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} T^{-k} A = B$

$\therefore$ B *es invariante*

$\therefore B = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} T^{-k} A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{W : T^k W \in A\}$

ahora $(A \Delta B) = (A - B) \cup (B - A) \subset \bigcup_{k=0}^{\infty} (T^{-k} A \Delta T^{-(k+1)} A) =$

$= \bigcup_{k=0}^{\infty} \left[(T^{-k}A - T^{-(k+1)}) \cup (T^{-(k+1)} A - T^{-k} A) \right]$

ya que si $W \in A - B$ *entonces* $W \in T^{-n} A$ *para un número finito de* n - s *incluyendo al cero.* (*ya que si* $W \in T^{-n} A$ - *para un número infinito de* n - s *entonces* $W \in B$); *Así* -- $W \in T^{-k} A - T^{-(k+1)} A$ *para alguna* $k = 0, 1 \ldots$ *ahora si* - $W \in B - A \Rightarrow T^n W \in A$ *para un número infinito de* n - s - *pero* $W \notin A$. *Si* $k + 1$ *es el entero más pequeño tal que* $T^{k+1} W \in A$, *entonces* $W \in T^{-(k+1)} A - T^{-k} A$ (*éste entero* - *existe ya que* $W \notin A$).

Como $P(T^{-k} A \Delta T^{-(k+1)} A) = P(A \Delta T^{-1} A) = 0$ *ya que* -- $(P(A \Delta T^{-1} A) = P(T^{-1} A \Delta T^{-2} A)$ *y así sucesivamente*).

Se sigue que $P(A \Delta B) = 0$ *debido a la subaditividad de la medida* P

$P(A \Delta B) \leq P(\bigcup_{k=0}^{\infty} (T^{-k} A \Delta T^{-(k+1)} A)) \leq$

$\sum_{k=0}^{\infty} P(T^{-k} A \Delta T^{-(k+1)} A) = 0$

$$\therefore P(A \Delta B) = 0.$$

DEMOSTRACION.

b) *Sea* **T** *ergódica y sea* **A** *un conjunto casi invariante, entonces existe por* (**a**) *un conjunto* **B** *invariante tal que* -

$$P(A \Delta B) = 0$$

$$\therefore P(A) = P(B)$$

$$\therefore P(B) = 0 \text{ o } P(\Omega - B) = 0 \quad \textit{pero}$$

$$P(\Omega - A) = P(\Omega - B) \text{ y } P(A) = P(B)$$

$$\therefore P(A) = 0 \text{ o } P(\Omega - A) = 0.$$

El inverso es claro ya que todo conjunto invariante es - *casi invariante.*

III.2.5. DEFINICION. *Sean* $X(W)$ *una variable aleatoria sobre* (Ω, β, P), T *una transformación que preserva medida; entonces* $X(W)$ *es una variable aleatoria casi invariante si y solo si* $X(W) = X(TW)$, *para casi todo* $W \in \Omega$.

III.2.6. X *es invariante si y solo si* X *es* $I(T)$ *medible.*

DEMOSTRACION. *Si* X *es invariante, entonces para cada* -- $x, \{X < x\} \in I$ *ya que si* $A = \{W : X(W) < x\}$ *entonces*

$$T^{-1} A = \{W : TW \in A\} = \{W : TW \in \{W : X(W) < x\}\} =$$

$$= \{W : X(TW) < x\} = \{W : X(W) < x\} = A.$$

$$\therefore X \text{ es } I \text{ medible}$$

Inversamente, si $X(W) = I_A(W)$, $A \in I$ *entonces*

$$X(TW) = I_A(TW) = I_{T^{-1}A}(W) = I_A(W).$$

ahora consideramos la clase $\mathcal{L}$ *de todas las funciones medibles sobre* (Ω, I) *que son invariantes.*

$\mathcal{L}$ *es cerrada bajo combinaciones lineales y* $X_n \in \mathcal{L}$, -- $X_n(W) \uparrow X(W)$ $\forall W$ *implica* $X(TW) = \lim_n X_n(TW) = \lim_n X_n(W) = X(W)$

$\therefore X \in \mathcal{L}$.

$\therefore$ $\mathcal{L}$ *contiene a todas las funciones medibles no negativas sobre* (Ω, I). (*ver* [4] *pag.* 32).

$\therefore$ *Cada función medible sobre* (Ω, I) *es invariante.*

III.2.7. PROPOSICION. *Un conjunto* $A \in \beta$ *es invariante si y solo si* I_A *es una función invariante.*

DEMOSTRACION.

$$A \in I(T) \text{ si y solo si } T^{-1}A = A \text{ si y solo si}$$

$$I_A(TW) = I_{T^{-1}A}(W) = I_A(W) \text{ si y solo si}$$

$$I_A \text{ es invariante.}$$

III.2.8. DEFINICION. *La distancia entre dos conjuntos medibles* $A,B \in \beta$ *es:*

$$d(A,B) = P(A \Delta B)$$

Demostraremos que en realidad ésta relación define una métrica.

i) $d(A,B) \geq 0$ *inmediata de la definición*

ii) $d(A,B) = 0 \iff A = B$ *modulo* P.

DEMOSTRACION.

$$P(A \Delta B) = 0 = P((A - B) \cup (B - A))$$

$$\therefore P(B - A) = 0 = P(A - B)$$

$$\therefore A = B \text{ modulo } P$$

iii) $d(A,B) = d(B,A)$ *inmediata*

iv) $d(A_1,A_3) \leq d(A_1,A_2) + d(A_2,A_3)$

$$i.e.\ P(A_1 \Delta A_3) \leq P(A_1 \Delta A_2) + P(A_2 \Delta A_3)$$

DEMOSTRACION

$$A_1 \Delta A_2 = (A_1 - A_2) \cup (A_2 - A_1)$$

$$A_1 \Delta A_3 = (A_1 - A_3) \cup (A_3 - A_1)$$

$$A_2 \Delta A_3 = (A_2 - A_3) \cup (A_3 - A_2)$$

(1)... $A_1 - A_3 \subset (A_1 - A_2) \cup (A_2 - A_3)$ *ya que sea*

$X \in A_1 - A_3 \Rightarrow X \in A_1$ *y* $X \notin A_3 \Rightarrow (X \in A_1$ *y* $X \notin A_2$ *y*

$X \notin A_3)$ *ó* $(X \in (A_1$ *y* $A_2)$ *y* $X \notin A_3) \Rightarrow X \in A_1 - A_2 \cup A_3$

ó $X \in A_2 - A_3 \Rightarrow X \in (A_1 - A_2 \cup A_3) \cup (A_2 - A_3)$

$\therefore (A_1 - A_3) \subset (A_1 - (A_2 \cup A_3)) \cup (A_2 - A_3) \subset (A_1 - A_2) \cup$

$\cup (A_2 - A_3)$

de igual manera obtenemos que

(2)... $(A_3 - A_1) \subset (A_3 - (A_2 \cup A_1)) \cup (A_2 - A_1) \subset (A_3 - A_2) \cup$

$(A_2 - A_1)$ $\therefore$ *de* (1) *y* (2) *obtenemos*

$P((A_1 - A_3) \cup (A_3 - A_1)) \leq P([(A_1 - A_2) \cup (A_2 - A_3)] \cup$

$$[(A_3 - A_2) \cup (A_2 - A_1)]) = P([(A_1 - A_2) \cup (A_2 - A_1)] \cup$$

$$[(A_2 - A_3) \cup (A_3 - A_2)]) = P(A_1 \Delta A_2 \cup A_2 \Delta A_3) <$$

$$< P(A_1 \Delta A_2) + P(A_2 \Delta A_3)$$

i.e. $P(A_1 \Delta A_3) < P(A_1 \Delta A_2) + P(A_2 \Delta A_3)$

III.2.9. PROPOSICION. *Sea* $J(T)$ *el conjunto de todas las funciones invariantes bajo* T. *El espacio lineal* $J(T)$ *es un subespacio lineal de* $L^1(\Omega, \beta, p)$. *El espacio métrico* $X_I \equiv \equiv \{I_A : A \in I\}$ *es isométrico a* I *bajo la transformación*

$$T(A) = I_A : I \to I_I$$

DEMOSTRACION. *Sean* $f, g \in J(T) \Rightarrow af + bg \in J(T)$

Sea $h = af + bg$ *por demostrar que* $h(TW) = h(W)$

$h(TW) = af(TW) + bg(TW) = af(W) + bg(W) = h(W)$

por demostrar que es una isometría

$$\text{i.e. } d(TA, TB) = d(A, B)$$

$$d(A,B) = P(A \Delta B) = \int I_{A \Delta B}\, dp = \int |I_A - I_B|\, dp =$$

$$= \| I_A - I_B \| = \| TA - TB \| = d(TA, TB).$$

III.2.10. **LEMA.** *Sea* T *que preserva medida en* $(\Omega, \mathcal{B}, P)$ *entonces las siguientes condiciones son equivalentes.*

i) T *ergódica*

ii) *Toda función casi invariante es casi una constante.*

iii) *Toda función invariante es casi una constante*

DEMOSTRACION. a => b. *Sea* g *una función casi invariante entonces para* $\lambda \in \mathbb{R}$. *Sea* $A_\lambda = \{W|g(W) < \lambda\}$ *que es un conjunto casi invariante ya que* $g(TW) = g(W)$ *casi donde quiera. Ya que* T *es ergódica* =>

$$\mu(A_\lambda) = 0 \text{ ó } \mu(A_\lambda^c) = 0$$

Sea $c = \sup \{\lambda|\mu(A_\lambda) = 0\}$, c *es finito ya que* $A_\lambda \uparrow \Omega$ -- *cuando* $\lambda \uparrow \infty$ *y se tendria* $\mu(A_\lambda) = \mu(\Omega) = 1$ *contradicción ó* $A_\lambda^c \uparrow \Omega$ *cuando* $\lambda \downarrow -\infty$ *(mismo razonamiento que antes) entonces*

$$\mu\{W|g(W) < c\} = \mu[\bigcup_{n=1}^{\infty} \{W|g(W) < c - \frac{1}{n}\}] = 0$$

Similarmente sea $\lambda \in \mathbb{R}$, $A_\lambda^c = \{W|g(W) > \lambda\}$, *que es un conjunto casi invariante ya que* $g(TW) = g(W)$ *casi dondequiera y ya que* T *es ergódica* $\mu(A_\lambda^c) = 0$ *ó* $\mu(A_\lambda) = 0$. *Sea* $c' = \inf$ -- $\{\lambda|\mu(A_\lambda^c) = 0\}$, c' *es finito ya que* $A_\lambda^c \uparrow \Omega$ *cuando* $\lambda \downarrow -\infty$ *ó* $A_\lambda \uparrow \Omega$ *cuando* $\lambda \uparrow \infty$. *Entonces*

$$\mu\{W|g(W) > c'\} = \mu[\bigcup_{n=1}^{\infty}\{W|g(W) > c' + \frac{1}{n}\}] = 0$$

Ahora $c' = c$ *ya que por un lado se pide que* $c = \sup \{\lambda|\mu(A_\lambda) = 0\}$ *y por otro se pide que* $c' = \inf \{\lambda|\mu(A_\lambda) = 0\}$ - *donde* A_λ *y* A_λ^c *son como se especificó antes.*

Por tanto

$$\mu\{W|g(W) > c\} = \mu\{W|g(W) < c\} = 0$$

Entonces

$$g(W) = c$$

Casi dondequiera. Es decir g *es una constante casi dondequiera.*.

b) => c) *Toda función invariante es casi invariante*

c) => a) *Sea* $A \in I(T)$ *y* I_A *el indicador de* A, *que es una función invariante.*

Por tanto

I_A *es casi dondequiera una constante.*

Si $I_A = 0$ *casi dondequiera entonces*

$$\mu(A) = \int_\Omega I_A \, d\mu = 0$$

y si $I_A = 1$ *casi dondequiera. Entonces*

$$\mu(A^c) = \int_\Omega I_A{}^c \, d\mu = \int_\Omega 1 - I_A \, d\mu = 0$$

Por tanto

T *es ergódica.*

III.2.11. DEFINICION. *Sea* T *una transformación que preserva medida en el espacio de probabilidad* (Ω, β, P). T *es una mezcla si y solo si* $\forall A, B \in \beta$

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

La restricción a una medida de probabilidad es escencial aqui. Si por ejemplo T *fuese una mezcla con respecto a* μ, *sea* $A = B = \Omega$ *entonces obtenemos* $\mu(\Omega) = [\mu(\Omega)]^2$ *por tanto si* μ *es finita debe ser 1; si* $\mu(\Omega) = \infty$ *y* A *es un conjunto con medida finita y positiva, tomemos* $B = \Omega$, *entonces*

$\mu(A \cap T^{-n} B) = \mu(A) < \infty$ *pero*

$\mu(A)\ \mu(B) = \infty$ *contradicción*

La propiedad de ser una mezcla tiene la siguiente interpretación intuitiva. Este ejemplo es debido a Halmos.

Si vemos a T *como la descripción de un flujo como se hi-*

zo en la discusión posterior a III.1.1.

Supongamos por ejemplo que un vaso es llenado con un liquido que es 90% ginebra y 10% vermouth, las particulas de vermouth ocupan el conjunto A, las particulas de ginebra el conjunto $\Omega - A$.

Después aplicamos una fuerza externa con un agitador; las condiciones del contenido son observadas en los tiempos $t = 0,1,2,\ldots$

Si B es un conjunto de Borel del contenido, Sea P(B) el volumen de B dividido entre el volumen del contenido del total.

(Asi $P(A) = .1$).

Ahora es razonable esperar que si el proceso de mezclar se continua por un periodo de tiempo grande, el porcentaje de vermouth en B sera aproximadamente el mismo que el porcentaje en el contenido total, a saber un 10%.

Para traducir esto en terminos Matemáticos, notemos que si W es un punto del contenido y una particula esta inicialmente en W, entonces $T^n W$ es la posicion de la particula n segundos después, asi el conjunto de particulas de vermouth que estan en B en el tiempo $t = n$ es

$$\{W \in A : T^n W \in B\} = A \cap T^{-n} B.$$

La proporción de vermouth en **B** *en el tiempo* $t = n$ *es* $\frac{P(A \cap T^{-n} B)}{P(B)}$ *y la propiedad de ser mezcla la transformación* **T** *se expresa diciendo que*

$$\frac{P(A \cap T^{-n} B)}{P(B)} \longrightarrow P(A) = .1$$

(La propiedad de ser mezcla es mas fuerte que ergodicidad, como se probará mas adelante).

III.2.12. **TEOREMA.** *Sea* **T** *una transformación mezcla en* (Ω, β, P). *Entonces* **T** *es ergódica.*

DEMOSTRACION. *Sea* $B \in I(T)$. *Si* $A \in \beta$, *entonces como* $B = T^{-n} B$, *tenemos*

$$P(A \cap T^{-n} B) = P(A \cap B) \text{ para toda } n.$$

Si hacemos que $n \to \infty$ *y usamos la propiedad de que* **T** *es una mezcla, obtenemos*

$$P(A \cap B) = P(A) P(B).$$

Pero como A *es un conjunto arbitrario en* β *podemos tomar* $A = B$ *y por tanto*

$$P(B) = [P(B)]^2$$

Por tanto

$$P(B) = 0 \text{ ó } 1$$

III.2.13. TEOREMA. *Sea* **T** *una transformación que preserva medida sobre* (Ω, β, P). *Sea* β_0 *un álgebra de subconjuntos de* Ω - *tal que la* σ*-álgebra generada por* β_0 *es* β. *Si la condición de mezcla se cumple para todo* $A, B \in \beta_0$, *esta se cumple para todo* $A, B \in \beta$.

DEMOSTRACION. *Sean* $A, B \in \beta$ *y* $\{A_k\} \in \beta_0$ *y* $\{B_k\} \in \beta_0$, *sucesiones de conjuntos tales que*

$$A_k \to A \quad y \quad B_k \to B \quad implica$$

$$P(A \Delta A_k) \quad y \quad P(B \Delta B_k) \to 0 \text{ cuando } k \to \infty.$$

Ahora afirmamos que

$$C = (A \cap T^{-n}B) \Delta (A_k \cap T^{-n}B_k) \subset (A \Delta A_k) \cup (T^{-n}(B \Delta B_k)).$$

Demostración de la afirmación:

Sea $x \in C$ *implica*

$$x \in [(A \cap T^{-n}B) - (A_k \cap T^{-n}B_k)] \cup [(A_k \cap T^{-n}B_k) - (A \cap T^{-n}B)]$$

implica

$$x \in [(A \cap T^{-n}B) - (A_k \cap T^{-n}B_k)] \text{ ó}$$

$$x \in [(A_k \cap T^{-n}B_k) - (A \cap T^{-n}B)].$$

Primer caso

$$\text{si } x \in (A \cap T^{-n}B) - (A_k \cap T^{-n}B_k) \text{ implica}$$

$$x \in A \cap T^{-n}B \text{ y } x \notin A_k \cap T^{-n}B_k \text{ implica}$$

$$x \in A \text{ y } x \in T^{-n}B \text{ y } [x \notin A_k \text{ y } x \notin T^{-n}B_k] \text{ implica}$$

$$x \in (A - A_k) \text{ y } x \in (T^{-n}B - T^{-n}B_k) \text{ implica}$$

$$x \in (A - A_k) \cap (T^{-n}B - T^{-n}B_k) \subset (A - A_k) \cup (T^{-n}B - T^{-n}B_k)$$

Por tanto

(1) $$C_1 = (A \cap T^{-n}B) - (A_k \cap T^{-n}B_k) \subset (A - A_k) \cup (T^{-n}B - T^{-n}B_k) = D_1$$

Segundo caso

$$\text{si } x \in (A_k \cap T^{-n}B_k) - (A \cap T^{-n}B) \Rightarrow$$

$$x \in A_k \cap T^{-n}B_k \text{ y } x \notin A \cap T^{-n}B$$

$$\Rightarrow x \in A_k \text{ y } x \in T^{-n}B_k \text{ y } [x \notin A \text{ y } x \notin T^{-n}B]$$

$$\Rightarrow x \in (A_k - A) \text{ y } x \in (T^{-n}B_k - T^{-n}B)$$

$$\Rightarrow x \in (A_k - A) \cap (T^{-n}B_k - T^{-n}B) \subset (A_k - A) \cup (T^{-n}B_k - T^{-n}B)$$

Por tanto

(2) $$C_2 = (A_k \cap T^{-n}B_k) - (A \cap T^{-n}B) \subset (A_k - A) \cup (T^{-n}B_k - T^{-n}B) = D_2$$

de (1) y (2) obtenemos que

(3) $$C_1 \cup C_2 \subset D_1 \cup D_2$$

Desarrollando la contención (3) se demuestra la afirmación.

Continua la demostración del teorema.

Así la probabilidad del conjunto de la izquierda es menor ó igual que

$$P(A \Delta A_k) + P(B \Delta B_k)$$

la cual tiende a cero cuando $k \to \infty$, uniformemente en n.

Por tanto

$$P(A_k \cap T^{-n}B_k) \to P(A \cap T^{-n}B) \text{ cuando } k \to \infty,$$

Uniformemente en n. Por hipótesis

$$P(A_k \cap T^{-n}B_k) \to P(A_k)P(B_k) \text{ cuando } n \to \infty$$

Por tanto (usando el teorema del límite doble)

$$\lim_{n\to\infty} P(A\cap T^{-n}B) = \lim_{n\to\infty}\lim_{k\to\infty} P(A_k\cap T^{-n}B_k)$$

$$= \lim_{k\to\infty}\lim_{n\to\infty} P(A_k\cap T^{-n}B_k) =$$

$$= \lim_{k\to\infty} P(A_k)\ P(B_k) = P(A)\ P(B).$$

y el teorema queda demostrado.

El siguiente resultado básico en teoría de la medida se usara frecuentemente, para su demostración ver [1] p.50.

III.2.14.TEOREMA. Sea $T : (\Omega,\beta) \to (\Omega_0,\beta_0)$ una transformación medible, μ una medida sobre β. Se define la medida μ_T como

$$\mu_T = \mu T^{-1} \text{ en } \beta_0 \text{ por}$$

$$\mu_T(A) = \mu(T^{-1}A) \quad A\in\beta_0.$$

(Si $\Omega = \Omega_0$, $\beta = \beta_0$ y T preserva a μ, entonces $\mu_T = \mu$)

Ahora si

$$f : (\Omega_0,\beta_0) \to (\overline{\mathbb{R}},\beta(\overline{\mathbb{R}}))\ y\ A\in\beta_0$$

entonces

$$\int_{T^{-1}A} f(TW)\ d\mu(W) = \int_A f(W)\,d\mu_T(W)$$

En el sentido que si una de las integrales existe así también la otra y las integrales son iguales.

III.2.15. DEFINICION. *La isometría inducida por una transformación* $T: \Omega \to \Omega$ *es la función*

$$\hat{T}: L^1(\Omega, \beta, \mu) \to L^1(\Omega, \beta, \mu) \quad \text{tal que}$$

$$\hat{T} f = f o T$$

III.2.16. *La isometría inducida por una transformación* T *que preserva medida es una isometría lineal de* L^1 *en* L^1*. Si* f *es una función invariante entonces*

$$\int \hat{T}^n f\, dP = \int f\, dP$$

DEMOSTRACION

$$\|\hat{T} f\|_1 = \int |\hat{T} f| d P = \int |f(T)| d P = \int f\, d P_T = \int |f|\, d P = \|f\|_1$$

ahora

$$\hat{T}(af + bg) = a\,\hat{T} f + b\,\hat{T} g$$

$$\hat{T}(af + bg) = (af + bg)\, T = af(T) + bg(T)$$

$$= a\,\hat{T} f + b\,\hat{T} g$$

tambien

$$\int \hat{T}^2 f \; dP = \int f(T^2) dP = \int f \; dP_{T^2} = \int f \; dP$$

ya que

$$P_T = P = P_{T^2} = \ldots$$

$$P_{T^2}(A) = P(T^{-2}A) = P(T^{-1}(T^{-1}A)) = P_T(T^{-1}A) = P(A)$$

y así sucesivamente.

III.2.17.EJEMPLOS.

a) *Sea* T *una permutación de* $\Omega = \{x_1, \ldots, x_n\}$, $n \geq 2$, μ *una medida sobre todos los subconjuntos de* Ω *que asigna igual peso a cada punto dentro de un ciclo dado de* T. *(Cuando hablamos acerca de la propiedad mezcla suponemos que* μ *es una medida de probabilidad; suponer tambien que*

$$\mu\{x_i\} > 0 \quad \forall i).$$

Tenemos que T *es ergódica si y sólo si* T *tiene sólo un ciclo. Esto se sigue porque los únicos conjuntos invariantes son uniones de ciclos de* T *(y el conjunto vacio).*

Pero T *no puede ser una mezcla. Suponer que* $\{x_1, \ldots, x_k\}$

es un ciclo dado de T y que $Tx_i = x_{i+1}$, con índices reducidos módulo K.

Suponer que $K \geq 2$, si $K = 1$, $\{x_1\}$ es un conjunto invariante no trivial.

Sea $A = B = \{x_i\}$, entonces

$A \cap T^{-n}B = A$ si n es un multiplo de K y es el conjunto vacio en otro caso.

Así

$$\lim \mu(A \cap T^{-n}B) \text{ no existe}$$

Por tanto T no puede ser una mezcla.

b) Sea $Tx = x + c$ sobre $\mathbb{R}$, con conjuntos de borel y la medida de lebesgue. Entonces T no es ergódica ya que si tomamos

$$A = \bigcup_{n=-\infty}^{\infty} (nc, nc + \frac{c}{2}),$$

es un conjunto invariante que no es ni el vacio ni el total. Demostración de que A es invariante

Sea

$$x \in A \Rightarrow x \in \bigcup_{n=-\infty}^{\infty} (nc, nc + \frac{c}{2}) \Rightarrow$$

$$x \in (nc, nc + \frac{c}{2}) \text{ para alguna } n \Rightarrow$$

$$Tx = x + c \in ((n+1)c, (n+1)c + \frac{c}{2}) \subset A$$

c) *Sea* T *la transformación corrimiento unilateral (ó bilateral) definida sobre el espacio de todas las sucesiones infinitas con un sólo final (ó con doble final) de numeros reales.*

El problema es encontrar condiciones sobre P *tal que* T *sea ergódica ó mezcla.*

Por lo cual consideraremos el caso en el cual las variables aleatorias X_k *son independientes*

es decir

$$P\{W : X_i(W) \in A_i, i = 1, 2, \ldots, n\} = \prod_{i=1}^{n} P\{W : X_i(W) \in A_i\}$$

$$\forall A_1, A_2, \ldots, A_n, \text{con } A_i \in \beta \text{ y } \forall n = 1, 2 \ldots$$

En este caso T *resulta ser una mezcla (por tanto ergódica).*

DEMOSTRACION: *Sea* $A = \{W : (X_0(W), \ldots, X_{k-1}(W)) \in B_k\}$ *y* $B = \{W : (X_0(W), \ldots, X_{r-1}(W)) \in B'_r\}$ *cilindros medibles. Para* n *suficientemente grande tenemos que* $n > k - 1$, *por tanto los índi* ces involucrados en las definiciones de los conjuntos **A** y T^{-n} **B** son distintos.

Por independencia

$$P(A\cap T^{-n}B) = P\{W : (W_0,\ldots,W_{k-1})\varepsilon\ B_k, (W_n,\ldots,W_{n+r-1})\varepsilon B'_r\}$$

$$= P(A)\,P(T^{-n}B) = P(A)\ P(B)$$

por tanto la condición de mezcla se cumple para todos los ci-- lindros medibles y por III.2.13 la condición de mezcla se cumple $\forall A,B\ \varepsilon\ \beta$.

Por tanto **T** es mezcla.

III.2.13. **DEFINICIONES.** Sea T una transformación medible (no necesariamente preserva a μ) en (Ω,β,μ).

a) T es recurrente si y sólo si $\forall A\varepsilon\beta$ y casi toda $W\ \varepsilon\ A$ - $T^nW\varepsilon A$, para alguna $n \geq 1$.

b) T es infinitamente recurrente si y sólo si $\forall A\varepsilon\beta$ y casi toda $W\varepsilon A$, $T^nW\varepsilon A$ para un número infinito de n-s.

c) Un conjunto $B\varepsilon\beta$ es viajero si y sólo si $B, T^{-1}B, T^{-2}B,\ldots,$ son disjuntos.

d) T es conservativa si y sólo si todos sus conjuntos viajeros tienen medida cero.

e) T *es incompresible si y sólo si* $A \subset T^{-1}A \Rightarrow$

$$\mu(T^{-1}A - A) = 0.$$

III.2.19. TEOREMA. *Sea* $T: \Omega \to \Omega$ *una transformación medible.* T *es incompresible si y sólo si* $T^{-1}A \subset A \Rightarrow \mu(A - T^{-1}A) = 0.$

DEMOSTRACION. $\Rightarrow$) *Suponer que* T *es incompresible y que* $T^{-1}A \subset A$ *implica* $A^c \subset T^{-1}A^c = (T^{-1}A)^c \Rightarrow \mu(T^{-1}A^c - A^c) = 0 = \mu(A - T^{-1}A)$ *por tanto* $\mu(A - T^{-1}A) = 0.$

$\Leftarrow$) *Suponer* $T^{-1}A \subset A \Rightarrow \mu(A - T^{-1}A) = 0$. *Ahora sea* B *un conjunto tal que*

$$B \subset T^{-1}B, \text{ como } T^{-1}B^c \subset B^c$$

$\Rightarrow \mu(B^c - T^{-1}B^c) = 0$ *pero* $\mu(B^c - T^{-1}B^c) = \mu(T^{-1}B - B),$

por tanto

$$\mu(T^{-1}B - B) = 0.$$

III.2.20. TEOREMA. *Las siguientes relaciones son equivalentes.*

1) T *es incompresible*

2) T *es conservativa*

3) T *es recurrente*

4) T *es infinitamente recurrente*

El esquema de la demostración es el siguiente:

1) => 2) => 3) => 1), 4) => 3), 1) => 4)

DEMOSTRACION:

1) => 2). *Sea* A *un conjunto viajero y sea*

$$B = \bigcup_{n=0}^{\infty} T^{-n} A.$$

Entonces

$$T^{-1}B = \bigcup_{n=1}^{\infty} T^{-n} A \subset B$$

Y además

$$\mu(B - T^{-1}B) = 0 \ldots \textit{por (1)}$$

Pero

$$B - T^{-1}B = A$$

Ya que los conjuntos $T^{-n} A$ *son disjuntos*

$$\therefore \mu(A) = 0$$

2) => 3) *Sea* $A^{(r)} = A \cap \bigcup_{n=1}^{\infty} T^{-n} A = \{W \in A \mid T^n W \in A,$ *para alguna* $n > 1\}$.

Sea $C = A - A^{(r)} = \{W \in A \mid T^n W \notin A, \forall n > 1\}$

P.D.

$$\mu(C) = 0$$

C es un conjunto viajero ya que

$$T^{-1}C = \{W \mid T^{n}W \in A \text{ pero } T^{k}W \notin A \quad \forall k > 1\}$$

En general

$$T^{-n}C = \{W \mid T^{n}W \in A \text{ pero } T^{k}W \notin A \quad k \neq n\}$$

ó sea C es un conjunto viajero y por hipotesis (2) la medida - de todo conjunto viajero es 0.

Por tanto

$$\mu(C) = 0.$$

i.e. para casi toda $W \in A$ se tiene $T^{n}W \in A$ para alguna - $n > 1$ ó sea T es recurrente.

3) => 1) sea $T^{-1}A \subset A$

P. D. que $\mu(A - T^{-1}A) = 0$

Como

$T^{-1}A \subset A$ => (por inducción) $T^{-1}(T^{-1}A) \subset T^{-1}A$ y $T^{-n}A \subset T^{-1}A$, $n > 1$ por tanto $T^{-1}A = \bigcup_{n=1}^{\infty} T^{-n}A$.

Ahora

$$A - T^{-1}A = A - \bigcup_{n=1}^{\infty} T^{-n}A = A - (A \cap \bigcup_{n=1}^{\infty} T^{-n}A) = A - A^{(r)} = C$$

El cual tiene medida 0 por ...(3)

Por tanto

$$\mu(A - T^{-1}A) = 0,$$

con lo cual queda demostrado (1).

4) => 3) *es claro*

1) => 4)

Sea $A^1 = A \cap \limsup_{n\geq 1} T^{-n}A = \{W \in A \mid T^n W \in A$ *para un númer*o *infinito de* n-s$\}$.

Si $A \in \beta$, *sea*

$$B = \bigcup_{n=0}^{\infty} T^{-n}A.$$

Entonces

$$T^{-1}B \subset B$$

Ya que

$$T^{-1}B = \bigcup_{n=1}^{\infty} T^{-n}A \subset \bigcup_{n=0}^{\infty} T^{-n}A$$

Por lo tanto

Por (1)

$$\mu(B - T^{-1}B) = 0.$$

Similarmente,

$$T^{-(k+1)}B \subset T^{-k}B \Rightarrow \mu(T^{-k}B - T^{-(k+1)}B) = 0.$$

Pero

$$T^{-k}B - T^{-(k+1)}B = \bigcup_{n=k}^{\infty} T^{-n}A - \bigcup_{n=k+1}^{\infty} T^{-n}A =$$

$$= \{W \mid T^nW \in A \text{ por última vez en } n = k\}$$

Como:

$$A - A^i = A - \{W \in A \mid T^nW \in A \text{ para un número infinito de } n\text{-s}\}$$

$$= \{W \in A \mid T^nW \in A \text{ para un número finito de } n\text{-s}\}$$

$$= A \cap \bigcup_{k=0}^{\infty} \{W \mid T^nW \in A \text{ por última vez en } n = k\}$$

Por lo tanto.

$$A - A^i \subset \bigcup_{k=0}^{\infty} (T^{-k}B - T^{-(k+1)}B) \Rightarrow \mu(A - A^i) <$$

$$< \mu\left(\bigcup_{k=0}^{\infty} (T^{-k}B - T^{-(k+1)}B)\right) < \sum_{k=0}^{\infty} \mu(T^{-k}B - T^{-(k+1)}B) = 0$$

Por lo tanto

T *es infinitamente recurrente*

III.2.21. TEOREMA. *Sea* T *una transformación que preserva medida sobre* (Ω, β, μ) *donde* μ *es una medida finita, entonces* T *es infinitamente recurrente.*

DEMOSTRACION. *Suponer que* T *no es infinitamente recurrente, entonces el conjunto*

$$F = \{W \in E \mid T^n W \notin E, \forall n\}, \forall E \in \beta,$$

es un conjunto de medida positiva.

El conjunto F *es medible y*

$$F = E \cap T^{-1}(\Omega - E) \cap T^{-2}(\Omega - E) \cap \ldots$$

$$= E \cap \{W \mid TW \in (\Omega - E)\} \cap \{W \mid T^2 W \in (\Omega - E)\} \cap \ldots$$

si $W \in F$*, entonces ninguno de los puntos*

$$TW,\ TW^2,\ TW^3, \ldots \text{pertenece a } F.$$

(Ya que si esto fuera así, se tendría, que W *es recurrente) en otras palabras* F *es disjunto de*

$$T^{-n}F \quad \forall n \text{ positivo}$$

Como

$$T^{-n}(F) \cap T^{-(k+n)}F = T^{-n}(F \cap T^{-k}F)$$

Por lo tanto

$$\mu\left(\bigcup_{n=0}^{\infty} T^{-n}F\right) = \sum_{n=0}^{\infty} \mu(T^{-n}F) = \sum_{n=0}^{\infty} \mu(F) = \infty$$

Contradicción ya que el espacio tiene medida finita.

III.3. TEOREMA ERGODICO PUNTUAL.

En seguida probaremos el teorema ergódico puntual, que establece que si T *es una transformación que preserva medida sobre* (Ω,β,μ) *y* $f \in L^1(\Omega,\beta,\mu)$ *entonces:*

$$n^{-1}[f(W) + f(TW) + \ldots + f(T^{n-1})]$$

converge a una función integrable $\hat{f}(W)$ *casi dondequiera.*

III.3.1. TEOREMA. *Sea* f *una función borel medible sobre* (Ω,β,μ) *y* T *preserva a* μ, *sea* $\hat{T}f = f \circ T$. *Entonces para* $p \in [0,\infty]$ *se tiene* $\|\hat{T} f\|_p = \|f\|_p$. *Por tanto* $\hat{T}$ *es un operador lineal sobre el espacio de Banach* $L^p(\Omega,\beta,\mu)$ *(ver [8], pag. 58). Entonces* $\hat{T}$ *es una isometría (ver teorema III.2.16). Además,* $\hat{T}$ *es positiva, i.e., si* $f > 0$ c.d. *entonces* $\hat{T}f > 0$ c.d.

DEMOSTRACION:

Por III.2.14 $\int_\Omega |\hat{T}(f(W))|^p \, d\mu(W) = \int_\Omega |f(TW)|^p \, d\mu(W) =$

$$= \int_\Omega |f(W)|^p \, d\mu(W)$$

$$\therefore \ \|\hat{T} f\|_p = \|f\|_p, \text{ para } 0 < p < \infty .$$

Si $p = \infty$, *tenemos*

$$\|\hat{T}f\|_\infty = \inf\{C : \mu\{|\hat{T}f| > C\} = 0\} = \inf\{C : \mu\{W : |f(TW)| > C\} = 0\} =$$

$$= \inf\{C : \mu T^{-1}\{W : |f(W)| > C\} = 0\} = \inf\{C : \mu\{|f| > C\} = 0\}$$

ya que T *preserva a* μ

$$\therefore \|\hat{T}f\|_\infty = \|f\|_\infty$$

Ahora sea $N \in \beta$ *tal que* $\mu(N) = 0$ *y* $f \geq 0$ *sobre* N^c, *entonces* $\mu(T^{-1}N) = \mu(N) = 0$.

Para $W \in N^c$, $\hat{T}f(W) = f(TW) \geq 0$ *sobre* $T^{-1}(N^c) = (T^{-1}N)^c$ *lo que prueba la positividad de* $\hat{T}$.

La sucesión de promedios $f^{(n)}(W) = n^{-1}[f(W) + f(TW) + \ldots + f(T^{n-1}W)]$ *puede entonces expresarse en términos de* $\hat{T}$ *como sigue:* $f^{(n)} = n^{-1}[f + \hat{T}f + \ldots + \hat{T}^{n-1}f]$, *donde* $\hat{T}^0 f = f$ *y* $\hat{T}^k$ *es la composición de* $\hat{T}$ *con sí misma* k *veces*.

En los resultados que siguen, T *es una transformación que preserva medida sobre* (Ω, β, μ) *y* $f : (\Omega, \beta) \to (\overline{\mathbb{R}}, \beta(\overline{\mathbb{R}}))$. *Observando las demostraciones se verá, sin embargo, que los resultados siguen siendo ciertos si reemplazamos a* $\hat{T}$ *por una contracción arbitraria positiva sobre* L^1, *no necesariamente tratándose de una transformación que preserva medida.*

III.3.2. LEMA. *Si* $f \in L^1(\Omega,\beta,\mu)$, *sea* $f_0 = f$ *y* $f_n = \max(f, f + \hat{T}f, \ldots, f + \hat{T}f + \ldots + \hat{T}^n f)$, $n \geq 1$ *entonces* $f_{n+1} \leq f + \hat{T}f_n^+$ $n = 0,1,2,\ldots$

DEMOSTRACION.

Si $0 \leq m \leq n$ *entonces* $\sum_{k=0}^{m+1} \hat{T}^k f = f + \hat{T}\left(\sum_{k=0}^{m} \hat{T}^k f\right)$,

Además:

$$\sum_{k=0}^{m} \hat{T}^k f \leq \max(f, f + \hat{T}f, f + \hat{T}f + \hat{T}^2 f, \ldots, f + \hat{T}f + \ldots + \hat{T}^n f) = f_n$$

Como $\hat{T}$ *es positiva, tenemos:*

$$\hat{T}\left(\sum_{k=0}^{m} \hat{T}^k f\right) \leq \hat{T}f_n \leq \hat{T}f_n^+$$

$$\therefore\ f + \hat{T}\left(\sum_{k=0}^{m} \hat{T}^k f\right) = \sum_{k=0}^{m+1} \hat{T}^k f \leq f + \hat{T}f_n^+,\ 0 \leq m \leq n$$

También $\quad f \leq f + \hat{T}f_n^+ \implies$

$$f_{n+1} = \max\{f, f + \hat{T}f, \ldots, f + \hat{T}f + \ldots + \hat{T}^{n+1} f\} \leq f + \hat{T}f_n^+$$

$$= f + \hat{T}[\max(f^+, f^+ + \hat{T}f^+, \ldots, f^+ + \ldots + \hat{T}^n f^+)]$$

$$\therefore\ f_{n+1} \leq f + \hat{T}f_n^+$$

III.3.3. LEMA. *Sea* f_n *como se definieron antes, y suponer*

$$f \in L^1(\Omega,\beta,\mu).$$

Si $A_n = \{f_n > 0\}$, *entonces* $\int_{A_n} f\, d\mu > 0$

DEMOSTRACION: *Por el lema anterior tenemos* $f_{n+1} < f + \hat{T} f_n^+$

$$\therefore \int_{A_n} f\, d\mu > \int_{A_n} f_{n+1}\, d\mu - \int_{A_n} \hat{T} f_n^+\, d\mu$$

$$> \int_{A_n} f_n\, d\mu - \int_{A_n} \hat{T} f_n^+\, d\mu = \int_{\Omega} f_n^+\, d\mu - \int_{A_n} \hat{T} f_n^+\, d\mu$$

$$> \int_{\Omega} f_n^+\, d\mu - \int_{\Omega} \hat{T} f_n^+\, d\mu = \| f_n^+ \|_1 - \| \hat{T} f_n^+ \|_1$$

$$> \| f_n^+ \|_1 - \| \hat{T} \|_1 \| f_n^+ \|_1 > 0$$

$$\therefore \int_{A_n} f\, d\mu > 0$$

En particular esto sigue siendo válido si $\hat{T}$ *es una con-- tracción positiva, i.e.* $\| \hat{T} \| < 1$.

III.3.4. TEOREMA. (*Ergódico Maximal*). *Si* $f \in L^1(\Omega, \beta, \mu)$ *y*

$A = \{W : \sup_{n \geq 1} \sum_{k=0}^{n-1} (\hat{T}^k f)(W) > 0\} = [\{W : \sup_{n \geq 1} f^{(n)}(W) > 0\}$ *don- de* $f^{(n)}(W) = n^{-1} \sum_{k=0}^{n-1} \hat{T}^k f] = \{W : \sup_{n \geq 1} n^{-1} \sum_{k=0}^{n-1} \hat{T}^k f(W) > 0\}$

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

DEMOSTRACION: *Los conjuntos* A_n *del lema anterior converge crecientemente a* A.

En seguida definimos algunos de los conjuntos que utilizaremos en la demostración del teorema ergódico puntual.

Sea $f \in L^1(\Omega,\beta,\mu)$ *y* $f^{(n)}(W) = n^{-1}\sum_{k=0}^{n-1} f(T^k W)$, $n=1,2,\ldots$ *y* T *preserva a* μ. *Si* $a<b$, *definamos:*

$$C_{ab} = C_{ab}(f) = \{W : \lim_{n\to\infty} \inf f^{(n)}(W) < a < b < \lim_{n\to\infty} \sup f^{(n)}(W)\} \text{ y}$$

$$N_b = \{W : \sup f^{(n)}(W) > b\}$$

veremos que C_{ab} *es un subconjunto de* N_b.

Sea $W \in C_{ab} \Rightarrow b < \lim_{n\to\infty} \sup f^{(n)}(W) = \inf_n \sup_{k>n} f^{(n)}(W) \leq \sup_n f^{(n)}(W)$

$$\therefore b < \sup_n f^{(n)}(W) \Rightarrow W \in N_b \therefore C_{ab} \subset N_b$$

Por tanto estableceremos convergencia casi donde quiera de la sucesión $f^{(n)}$, *si probamos que* $\mu(C_{ab})$ *es cero. Para ver esto podemos suponer sin pérdida de generalidad que* $b > 0$.

Ya que $C_{ab}(f) = \{W : \lim \inf -f^{(n)}(W) < -b < -a < \lim\sup -f^{(n)}(W)\}$

$$= C_{-b,-a}(-f)$$

Si $b \leq 0$. *entonces* $-a > 0$, *y el argumento siguiente mos--*

trará que:

$C_{-b,-a}(-f)$ *tiene medida cero y* $\therefore \mu(C_{ab}(f)) = 0$

III.3.5. LEMA.

a) *El conjunto* C_{ab} *es casi invariante.*

b) $\mu(C_{ab}) < \infty$

c) *En realidad* $\mu(C_{ab}) = 0$

DEMOSTRACION: *a) Sea*

$$f^{(n)}(TW) = \frac{1}{n}\sum_{k=0}^{n-1} f(T^{k+1}W) = \frac{n+1}{n}\left[\frac{1}{n+1}\sum_{k=0}^{n} f(T^k W)\right] - \frac{f(W)}{n}$$

$$= \frac{n+1}{n} f^{(n+1)}(W) - \frac{f(W)}{n}$$

ahora $f \in L^1$, *entonces* $f(W)$ *es finita para casi toda* W. *(ver - [1], pag. 47).*

$$\therefore \liminf_{n\to\infty} f^{(n)}(TW) = \liminf_{n\to\infty}\left[\frac{n+1}{n} f^{(n+1)}W - \frac{f(W)}{n}\right]$$

$$= \liminf_{n\to\infty} f^{(n+1)}(W) = \liminf_{n\to\infty} f^{(n)}(W) \quad \ldots(1)$$

Similarmente:

$$\limsup_{n\to\infty} f^{(n)}(TW) = \limsup_{n\to\infty} \left[\frac{n+1}{n} f^{(n+1)}(W) - \frac{f(W)}{n}\right]$$

$$= \limsup_{n\to\infty} f^{(n+1)}(W) = \limsup f^{(n)}(W) \ldots (2)$$

∴ *tenemos las igualdades* (1) *y* (2) *excepto, posiblemente para un conjunto de medida cero.*

∴ *fuera de un conjunto de medida cero, tenemos que*

$W \in C_{ab}$ *si y solo si* $TW \in C_{ab}$

i.e. C_{ab} *es casi invariante.*

DEMOSTRACION: *b) Sea* $C \in \beta$ *tal que* $C \subset C_{ab}$ *y* $\mu(C) < \infty$, - *definamos* $F_b = \{W : \sup_{n\geq 1} (f-bI_c)^{(n)}(W) > 0\} = \{W : \sup_{n\geq 1} n^{-1} \sum_{k=0}^{n-1} (f-bI_c)(T^kW) > 0\}$ *tenemos que si* $W \in C_{ab}$ *entonces*

$$W \in N_b = \{W : \sup_{n\geq 1} f^{(n)}(W) > 0\}$$

$$\therefore \frac{1}{n} \sum_{k=0}^{n-1} f(T^kW) > b (\textit{para alguna } n)$$

$$> \frac{1}{n} \sum_{k=0}^{n-1} bI_c (T^kW)$$

$$\therefore \frac{1}{n} \sum_{k=0}^{n-1} f(T^kW) - \frac{1}{n} \sum_{k=0}^{n-1} bI_c (T^kW) > 0$$

$$\therefore \frac{1}{n} \sum_{k=0}^{n-1} (f - bI_c)(T^kW) > 0, \text{ i.e. } (f-bI_c)^{(n)} > 0$$

$\therefore$ $W \in F_b$ *y* $C_{ab} \subset F_b$, *en particular* C *es un subconjunto de* F_b.

Por el teorema ergódico maximal, tenemos: $\int_{F_b} (f - I_c b) d\mu \geq 0$

$$\therefore \int_\Omega |f| \, d\mu \geq \int_{F_b} f \, d\mu \geq b \int_{F_b} I_c \, d\mu = b\mu(C \cap F_b) = b\mu(C)$$

o sea

$$\mu(C) \leq b^{-1} \int_\Omega |f| \, d\mu < \infty$$

veamos ahora que C_{ab} *es un subconjunto de* $\bigcup_{n=0}^{\infty} \{W : |f(T^n W)| > 0\} = B$

Suponer que $W \in C_{ab}$ *tal que* $W \notin B \Rightarrow |f(T^n W)| = 0 \ \forall n \Rightarrow$ $f(T^n W) = 0 \ \forall n \ \therefore \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W) = 0$ *una contradicción ya que la serie convergería.*

Ahora por III.2.14. $\int_\Omega |f(T^n W)| \, d\mu(W) = \int |f(W)| \, d\mu(W) < \infty$

Tenemos

$$C_{ab} \subset B, \text{ además}$$

$\{W : |f(T^n W)| > 0\} = \bigcup_{k=1}^{\infty} A_k$ *donde* $A_k = \{W : |f(T^n W)| \geq 1/k\}$ *y*

$\frac{1}{k} \int_{A_k} I_{A_k} \, d\mu = \frac{1}{k} \mu(A_k) \leq \int_{A_k} |f(T^n W)| \, d\mu < \infty \therefore \mu(A_k) < \infty$ *es decir* $\{W : |f(T^n W)| > 0\}$ *es la union numerable de conjuntos de medida finita.*

$\therefore C_{ab} \subset \bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty} A_{k,n}$ *y podemos poner a* C_{ab} *en la forma*

$$C_{ab} = C_{ab} \cap \left[\bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{k,n}\right] = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} (A_{k,n} \cap C_{ab})$$

$$\therefore C_{ab} = \bigcup_{n=1}^{\infty} B_n \text{ donde } B_n = \bigcup_{k=1}^{\infty} (A_{k,n} \cap C_{ab})$$

ó sea C_{ab} *es la union numerable de conjuntos de medida finita además* $\mu(B_n) < \infty$. *Ahora hagamos* $D_j = \bigcup_{n=1}^{j} B_n$, $\mu(D_j) < \infty$ *entonces*

$$\mu(D_j) \uparrow \mu(C_{ab})$$

$$\therefore \mu(C_{ab}) = \operatorname{Sup}\ \{\mu(D_j) : D_j \in \beta,\ D_j \subset C_{ab},\ \mu(D_j) < \infty\}$$

$$\leq b^{-1} \int_{\Omega} |f|\ d\mu < \infty$$

$$\therefore \mu(C_{ab}) < \infty$$

DEMOSTRACION: c) *Sabemos que* C_{ab} *es casi invariante,* T *una transformación que preserva medida que esta bién definida sobre,* $(C_{ab}, \beta_{ab}, \mu_{ab})$ *donde:*

$\beta_{ab} = \{A \in \beta : A \subset C_{ab}\} = \{B \cap C_{ab} : B \in \beta\}$ *y* $\mu_{ab} = \mu$ *restringida a* C_{ab}.

[*Estrictamente hablando , si* $D = C_{ab} \Delta T^{-1} C_{ab}$, *entonces* T *está bien definida sobre* $C_{ab} - D$. *Como* $\mu(D) = 0$, *esto no causa* *dificultad ya que podemos redefinir a* **T** *como la identidad so-bre* D, *y entonces estará bien definida y preserva medida sobre* C_{ab}] .

Podemos aplicar el argumento de (b) *a* **T** *sobre* C_{ab}. *En particular la ecuación* $\int_{F_b} (f-bI_c)\, d\mu > 0$ *se trasforma en* $\int_{C_{ab}\cap F_b} (f - bI_c)\, d\mu > 0$.

Como C_{ab} *tiene medida finita podemos hacer* $C = C_{ab}$ *y ya que* $C_{ab} \subset F_b$, *se obtiene:*

$$\int_{C_{ab}} (f-b)\, d\mu > 0 \qquad \ldots(1)$$

Ahora sea:

$F'_{ab} = \{W \in C_{ab} : \sup_{n\geq 1} (a-f)^{(n)}(W)>0\} = \{W \in C_{ab} : \sup_{n\geq 1} n^{-1} \sum_{k=0}^{n-1} (a-f)(T^kW) > 0\}$. *Si* $W \in C_{ab}$, *entonces* $f^{(n)}(W) = n^{-1} \sum_{k=0}^{n-1} f(T^kW) < a$, *para al menos una* n, *entonces* $W \in F'_{ab}$, $\therefore C_{ab} = F'_{ab}$. *Como* C_{ab} *tiene medida finita, las funciones constantes son integrables sobre* C_{ab}, *Por tanto podemos aplicar el teorema ergódico maximal para obtener*

$$\int_{C_{ab}} (a-f)\, d\mu > 0 \left[\text{i.e.} \int_{C_{ab}} a\, d\mu - \int_{C_{ab}} f\, d\mu > 0 \Rightarrow a\,\mu\,(C_{ab}) =\right.$$

$$= \int_{C_{ab}} a\, d\mu > \int_{C_{ab}} f\, d\mu \Big]$$

Con este resultado y usando (1) para $a < b$, se tiene - $b\mu(C_{ab}) < \int_{C_{ab}} f\, d\mu < a\mu(C_{ab})$ una contradicción, a menos -- que

$$\mu(C_{ab}) = 0 \quad \therefore \mu(C_{ab}) = 0.$$

Ahora probaremos el principal resultado.

III.3.6. TEOREMA ERGODICO PUNTUAL. *Sea* T *una transformación que preserva medida sobre* (Ω, β, μ) *y* $f \in L^1(\Omega, \beta, \mu)$. *Entonces existe una función* $\hat{f} \in L^1$ *tal que:*

$$n^{-1} \sum_{k=0}^{n-1} f(T^k W) \to \hat{f}(W) \text{ c.d.}$$

DEMOSTRACION. *Sea* $D = \{W : f^{(n)}(W)$ *no converge a un límite finito o infinito*$\}$. *Entonces* $D = \cup\{C_{ab}(f) : a < b,\ a,b$ *racionales*$\}$. *Por el lema anterior* $\mu(D) = 0$ *y por tanto* $f^{(n)}(W)$ *converge para casi toda* W; *llamaremos al límite* $\hat{f}(W)$, *y definamos* $\hat{f}(W) = 0$ *sobre el conjunto* D.

Solo resta demostrar que $\hat{f} \in L^1$

Usando el lema de Fatou.

$$\int_{\Omega}|\hat{f}|d\mu = \int_{\Omega}\lim_{n\to\infty}|f^{(n)}(W)|d\mu(W) \leq \lim_{n\to\infty}\inf\int_{\Omega}|f^{(n)}(W)|\ d\mu\ (W)$$

Pero

$$\int_{\Omega}|f^{(n)}|d\mu - \int_{\Omega}|n^{-1}\sum_{k=0}^{n-1} f(T^kW)|d\mu \leq \frac{1}{n}\sum_{k=0}^{n-1}\int_{\Omega}|f(T^kW)|d\mu =$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\int_{\Omega}|f|\ d\mu \quad \textit{por III.2.14.}$$

$$= \int_{\Omega}|f|\ d\mu < \infty$$

$$\therefore \int_{\Omega}|\hat{f}|\ d\mu \leq \int_{\Omega}|f|\ d\mu < \infty$$

el teorema queda demostrado.

Ahora veremos mas estrictamente la convergencia de la sucesión $\{f^{(n)}\}$.

III.3.7. TEOREMA. *Si* $\mu(\Omega) < \infty$ *y* $f \in L^P (1 \leq P < \infty)$, *entonces* $\hat{f} \in L^P$ *y* $f^{(n)} \to \hat{f}$ *en* L^P.

DEMOSTRACION: *Como las funciones simples son densas en* L^P, *para cada* $\varepsilon > 0$, *existe una función medible acotada* g *tal que* $\| f - g\|_P < \varepsilon$ *(Ver [1] pag. 88).*

Sea $f_k(W) = f(T^kW)$, $g_k(W) = g(T^kW)$ *y* $|g| \leq M$ *entonces*

$$\| f^{(n)} - \hat{f}\|_p = \| f^{(n)} - g^{(n)} + g^{(n)} - \hat{g} + \hat{g} - \hat{f}\| = \| \frac{1}{n} \sum_{k=0}^{n-1} f_k(W) -$$

$$- n^{-1} \sum_{k=0}^{n-1} g_k(W) + n^{-1} \sum_{k=0}^{n-1} g_k(W) - \hat{g} + \hat{g} - \hat{f} \| \leq$$

$$\leq \| n^{-1} \sum_{k=0}^{n-1} (f_k - g_k) \|_p + \|(\frac{1}{n} \sum_{k=0}^{n-1} g_k) - \hat{g}\|_p + \|\hat{g} - \hat{f}\|_p \quad \ldots(1)$$

Como $|n^{-1} \sum_{k=0}^{n-1} g_k| \leq M$ $(\therefore |\hat{g}| \leq M$ c.d.$)$, *el segundo término del lado derecho de (1) tiende a cero cuando* $n \to \infty$, *por el teorema de convergencia dominada.*

Por III.2.14 $\|f_k - g_k\|_p = \|f - g\|_p < \varepsilon$, *por tanto el primer término es menor que* ε.

Ahora

$$\int_\Omega |\hat{f} - \hat{g}|^p \, d\mu = \int_\Omega \lim_{n\to\infty} |\frac{1}{n} \sum_{k=0}^{n-1} (f_k - g_k)|^p \, d\mu \leq$$

$$\leq \liminf_{n\to\infty} \int_\Omega |\frac{1}{n} \sum_{k=0}^{n-1} (f_k - g_k)|^p \, d\mu =$$

$$= \liminf_{n\to\infty} \|\frac{1}{n} \sum_{k=0}^{n-1} (f_k - g_k)\|_p^p < \varepsilon^p$$

así $\|f^{(n)} - \hat{f}\|_p < 2\varepsilon$, *para* n *suficientemente grande y el teorema queda demostrado.*

Si $p = 1$ *en III.3.7, la hipótesis de que* $\mu(\Omega) < \infty$ *no puede ser eliminada.*

CONTRAEJEMPLO:

p.e.(*1*). *Sea* $T(W) = W + 1$ *sobre* $\mathbb{R}$ (*con los conjuntos de borel y la medida de lebesgue*). *Sea* f *la función característica del intervalo* $[0,1]$, *entonces* $f^{(n)}$ *no converge en* L^1 *a* $\hat{f}$.

DEMOSTRACION: $\forall W \in \mathbb{R},\ f(T^n W) = 0$

para n *grande* $\therefore\ \hat{f} \equiv 0$.

Pero si $-(n-1) < W < 1$, *entonces* $T^k W \in [0,1]$ *para exactamente un entero* $k \in [0, n-1]$ *entonces* $\frac{1}{n}\sum_{k=0}^{n-1} f(T^k W) = \frac{1}{n}$ $\forall W \in (-(n-1),1]$.

Para otros valores de W *tenemos*

$$n^{-1}\sum_{k=0}^{n-1} f(T^k W) = 0$$

Por tanto:

$$\int_\Omega \left|\frac{1}{n}\sum_{k=0}^{n-1} f(T^k W) - \hat{f}(W)\right|^p d\mu(W) = \int_{-(n-1)}^{1} \left(\frac{1}{n}\right)^p d\mu = \left(\frac{1}{n}\right)^p \left[1-(-n+1)\right] =$$

$$= \left(\frac{1}{n}\right)^p n = n^{1-p}$$

$\therefore \| f^{(n)} - \hat{f} \|_1 = 1 \ \forall n$, *por consiguiente no hay convergencia en* L^1.

Ahora veremos un ejemplo en el cual la hipótesis de que $p < \infty$ *no puede eliminarse aunque se tenga que* $\mu(\Omega) < \infty$.

Sea $\Omega = \mathbb{R}^\infty$ *y* $\beta = [\beta(\mathbb{R}^\infty)]$ $X_n(W_0, W_1, \ldots) = W_n$, *y* P *la medida de probabilidad que hace a las* X_{n-s} *independientes con* $P(X_n = 0) = P(X_n = 1) = \frac{1}{2}$ *(en otras palabras, consideramos una sucesión infinita de ensayos bernoulli con probabilidad de éxito* $= \frac{1}{2}$ *en cada ensayo).*

Si T *es el corriente unilateral y* $f(W) = W_0$, *entonces* $n^{-1} \sum_{k=0}^{n-1} f(T^k W)$ *no converge en* L^∞ *a* $\hat{f}(W) = \frac{1}{2}$.

DEMOSTRACION. *Si* $W_0 = W_1 = \ldots = W_{n-1} = 1$, *entonces* $n^{-1} \sum_{k=0}^{n-1} f(T^k W) = 1$, *por tanto* $|n^{-1} \sum_{k=0}^{n-1} f(T^k W) - \frac{1}{2}| = \frac{1}{2}$, *sobre un conjunto de probabilidad* 2^{-n}.

$$\therefore \left\| \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W) - \frac{1}{2} \right\|_\infty =$$

$$= \inf \{ C : P\{ W : |n^{-1} \sum_{k=0}^{n-1} f(T^k W) - \frac{1}{2}| > C \} = 0 \} = \frac{1}{2} \not\to 0$$

III.3.8. LEMA. *Si* $f \in L^1$, *entonces* $\hat{f}$ *es casi invariante. Así si* G *es la* σ*-álgebra de los conjuntos casi invariantes,* -

$f : (\Omega, G) \to (\overline{\mathbb{R}}, \beta(\overline{\mathbb{R}}))$.

DEMOSTRACION: *Si* $f^{(n)}(W) \to \hat{f}(W)$ *para* $W \notin N$, *donde* $\mu(N)=0$, *entonces* $f^{(n)}(TW) \to \hat{f}(TW)$ *para* $W \notin T^{-1}N$, *donde* $\mu(T^{-1}N) = 0$. -

Pero $f^{(n)}(TW) = \frac{n+1}{n} f^{(n+1)}(W) - \frac{f(W)}{n}$ *(ver teorema III.3.5) y* $\frac{f(W)}{n} \to 0$ *casi dondequiera, como* $f \in L^1$, *se tiene* -

$$\lim_{n\to\infty} f^{(n)}(TW) = \lim_{n\to\infty} f^{(n+1)}(W) = \lim_{n} f^{(n)}(W) = \hat{f}(W).$$

$$\therefore\ f^{(n)}(TW) \to \hat{f}(W) \text{ casi dondequiera,}$$

$$\therefore\ \hat{f}(W) = \hat{f}(TW) \quad \text{casi dondequiera}$$

lo que prueba que $\hat{f}$ *es casi invariante.*

Ahora si $B \in \beta(\overline{\mathbb{R}})$ *y* $C = \{W : \hat{f}(W) \in B\} = \hat{f}^{-1}(B)$, *entonces si* $\hat{f}(W) = \hat{f}(TW)$ *tenemos que* $W \in C$, *si y sólo si* $TW \in C$ *por tanto* C *es casi invariante o sea* $C \in G$, *con esto se completa la demostración.*

III.3.9. TEOREMA. *Si* $f \in L^1$ *y* $A \in G$ *entonces*

$$\int_A f \, d\mu = \int_A \hat{f} \, d\mu .$$

Así en un espacio de probabilidad

$$\hat{f} = E(f|G)$$

DEMOSTRACION. *Como en el Lema III.3.5., restringimos a* T *a* (A, β_A, μ_A) *donde* $\beta_A = \{B \in G : B \subset A\} = \{B \cap A : B \in G\}$ *y* $\mu_A = \mu$ *restringida a* A *y redefinimos* $T = I$ *sobre* $G - A$. *Como* $\mu(A) < \infty$, *tenemos convergencia de* $f^{(n)}(W)$ *en* L^1 *(sobre* A*) por III.3.7, - por tanto*

$$\int_A f^{(n)}\, d\mu \to \int_A \hat{f}\, d\mu, \text{ pero } \int_A f^{(n)} d\mu = \int_A f\, d\mu$$

por III.2.14.

$$\therefore \int_A f = \int_A \hat{f}\, d\mu$$

Si estamos en un espacio de probabilidad (Ω, β, P) *y* G *es - la* σ*-álgebra de los eventos casi invariantes en* β, *se tiene*

$$\int_A f\, dP = \int_A \hat{f}\, dP_G \quad \forall A \in G$$

$$= \int_A E(f|G)\, d\, P_G$$

$\therefore \hat{f} = E\,(f|G)$ *por definición (ver [8], pág. 121).*

III.3.10. **TEOREMA.** *Si* T *es ergodica y* $f \in L^1$, *entonces* $\hat{f}$ *es una constante casi dondequiera.*

Si $\mu(\Omega) = 0$, *la constante es* $C = 0$, *si* $\mu(\Omega) < \infty$, *entonces*

$$C = \frac{1}{\mu(\Omega)} \int_\Omega f\, d\mu \ .$$

Así sobre un espacio de probabilidad

$$\hat{f} = E\,(f) \quad \textit{casi dondequiera}$$

DEMOSTRACION: *Por III.3.8,* $\hat{f}$ *es casi invariante y por III.2.3.(b),* $\hat{f} = C$*, casi dondequiera.*

Ahora si $\mu(\Omega) = \infty$*, entonces* C *debe ser* 0 *ya que* $\hat{f} \in L^1$ *(y la única constante integrable en* L^1 *es la constante* $C = 0$*); por III.3.6. Si* $\mu(\Omega) < \infty$*, si hacemos* $A = \Omega$ *en III.3.9. obtenemos*

$$\int_\Omega f\, d\mu = \int_\Omega \hat{f}\, d\mu = C\int_\Omega d\mu = C\mu\,(\Omega) \ .$$

$$\therefore\ C = \frac{1}{\mu(\Omega)} \int_\Omega f\, d\mu \ .$$

Por tanto si estamos en un espacio de probabilidad $\hat{f} = C = E(f)$ *casi dondequiera.*

Si T *es ergódica y* μ *es finita, consideremos el caso en que* $f = I_A$, $A \in \beta$.

Entonces por III.3.10 $\hat{f} = \frac{\mu(A)}{\mu(\Omega)}$ *casi dondequiera, así que* $n^{-1} \sum_{k=0}^{n-1} I_A(T^k W)$.

La frecuencia relativa de visitas a **A**, *converge casi dondequiera a la masa relativa de* **A** (*La probabilidad de* **A** *si* $\mu(\Omega) = 1$).

Aparentemente tenemos una versión de la ley fuerte de los grandes números, y en realidad el teorema ergódico puntual puede ser visto como una generalización de este resultado.

Sea **T** *la transformación corrimiento unilateral, (ver III.1.4) con variables aleatorias coordenadas* X_k.

Si **Z** *es una variable aleatoria integrable sobre* (Ω, β, P), *entonces*

$$Z^{(n)} = n^{-1}\left[Z(T^0(W_0, W_1, \dots)) + Z(T(W_0, W_1, \dots)) + \dots + Z(T^{n-1}(W_0, \dots))\right]$$

$$= n^{-1}\left[Z(W_0, W_1, \dots) + Z(W_1, W_2, \dots) + \dots + Z(W_{n-1}, W_n, \dots)\right]$$

en otras palabras

$$Z^{(n)}(W) = n^{-1}\left[Z(X_0, X_1, \dots) + Z(X_1, X_2, \dots) + \dots + Z(X_{n-1}, X_n, \dots)\right]$$

Por III.3.9. $Z^{(n)} \xrightarrow{c.d.} E(Z|G)$, *donde* G *es la* σ*-álgebra de los conjuntos casi invariantes; Si* T *es ergódica el límite es* $E(Z)$ *por III.3.10.*

En particular sean $X_0, X_1 \dots$, *variables aleatorias identi* *camente distribuidas con esperanza finita.*

Si $Z(W) = W_0$, *o sea* $Z = X_0$, *obtenemos* $n^{-1}(X_0 + \ldots + X_{n-1}) \to \to E X_0$ *casi dondequiera. Si además las variables aleatorias son independientes, obtenemos el caso de la ley fuerte de los grandes números. (ver [1], pag. 275).*

En la sección siguiente, será necesario considerar medidas de probabilidad P_1 *y* P_2 *que son preservadas por una transformación sobre* (Ω, β, μ)

III.3.11. DEFINICION. P_i *es ergódica (relativa a* T*) si y solo si* T *es ergódica sobre* $(\Omega, \beta, P_i), i = 1, 2.$

III.3.12. TEOREMA. *Sea* $T : \Omega \to \Omega$ *una transformación que preserva medida y ergódica con respecto a* (Ω, β, P_1) *y* (Ω, β, P_2).

Entonces $P_1 = P_2$ *ó* P_1 *y* P_2 *son ortogonales en el sentido de que* $\exists$ *un conjunto* $A \in I(T)$ *tal que* $P_1(A) = 1$ *y* $P_2(A^c) = 1$

DEMOSTRACION. *Si* $P_1 \neq P_2$, *tomar* $B \in \beta$ *tal que* $P_1(B) \neq P_2(B)$ *y sea* $f(W) = I_A(W)$.

Sea $A_1 = \{W : \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W) \to P_1(B)\}$. *Por III.3.6. y III.3.10.*

$$P_1(A) = 1.$$

Pero

$$A^c = \{W : \frac{1}{n}\sum_{k=0}^{n-1} f(T^k W) \to P_2(B)\}$$

$$\therefore P_2(A^c) = 1$$

$$\therefore P_1 \perp P_2.$$

III.3.13. TEOREMA. *La medida de probabilidad* P *es ergódica relativa a* T *si y solo si no existe una medida de probabilidad* $P_1 \not\equiv P$ *preservada por* T *tal que sea absolutamente continua con respecto a* P.

DEMOSTRACION: *Suponer que* P *es ergódica y que* $\exists P_1$ *tal que* $P_1 \ll P$, $P_1 \not\equiv P$, *entonces* P_1 *tambien es ergódica, ya que si* $A \in I(T)$, *entonces* $P(A) = 0$ *ó* $P(A^c) = 0$ *por ergodicidad, por tanto* $P_1(A) = 0$ *ó* $P_1(A^c) = 0$ *por la continuidad absoluta usando III.3.12* $P_1 \perp P$.

Pero entonces P_1 *es absolutamente continua con respecto a* P *y tambien es singular con respecto a* P_1 *entonces* P_1 *debe ser la medida* 0, *una contradicción ya que se tiene que* P_1 *es ergódica.*

$\therefore$ *No puede existir* P_1

Inversamente, si P *no es ergódica, sea* $A \in I(T)$ *tal que* $0 < P(A) < 1$.

Definamos $P_1(B) = P(B|A) = \dfrac{P(B \cap A)}{P(A)}$, $B \in \beta$, *entonces* $P_1 \ll P$ *y como* $P_1(A) = \dfrac{P(A \cap A)}{P(A)} = 1 \neq P(A)$, *tenemos que* $P_1 \not\equiv P$, *ahora*

$$P_1(T^{-1}E) = \frac{P(T^{-1}E \cap A)}{P(A)} = \frac{P(T^{-1}E \cap T^{-1}A)}{P(A)}$$

ya que A *es invariante*

$$= \frac{P(T^{-1}(E \cap A))}{P(A)} = \frac{P(E \cap A)}{P(A)} = P(E|A) =$$

$$= P_1(E)$$

entonces T *preserva a* P_1.

Sería natural preguntar si las condiciones del teorema ergódico son necesarias y suficientes.

Si f *es una función medible no negativa sobre* (Ω, β, μ) *de medida finita y* $\int_\Omega f\, d\mu = \infty$ *es fácil probar que para* T *que preserva medida y ergódica.*

$$\frac{1}{n}\sum_{k=0}^{n-1} f(T^k W) \to \infty \quad \textit{casi dondequiera}$$

La demostración se hará definiendo para $\alpha > 0$

$$f_\alpha(W) = \begin{cases} f(W) & \textit{si } f(W) < \alpha \\ 0 & \textit{en otro caso} \end{cases}$$

de hecho

$$\int_\Omega |f_\alpha| \, d\mu < \int_\Omega \alpha \, d\mu = \alpha\mu(\Omega) < \infty$$

Entonces podemos usar el teorema ergódico para obtener

$$\lim_{n\to\infty} \inf \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W) > \lim \frac{1}{n} \sum_{k=0}^{n-1} f_\alpha(T^k W) = E(f_\alpha) \text{ casi}$$

dondequiera.

Si hacemos que $\alpha \uparrow \infty$, obtenemos la conclusión.

Para transformaciones ergódicas sobre espacios de medida finita y funciones medibles no negativas el inverso del teorema ergódico es cierto en el siguiente sentido.

III.3.14. TEOREMA: Si $\frac{1}{n} \sum_{k=0}^{n-1} f(T^k W)$ converge casi dondequiera. Entonces f es integrable.

DEMOSTRACION: Comenzamos la demostración observando que el límite es finito e igual a una constante C por ergodicidad.

Si f_j es la función obtenida a partir de f truncandola en j o sea:

$$f_j(W) = f(W) \quad si \quad f(W) < j$$

$$f_j(W) = j \quad si \quad f(W) > j$$

Entonces f_j es acotada y por tanto integrable ya que $\mu(\Omega) < \infty$. El teorema de la convergencia monótona implica que

$$\int f_j \, d\mu \to \int f \, d\mu ;$$

por tanto demostraremos la integrabilidad de f si mostramos que la sucesión $\{\int f_j \, d\mu\}$ es acotada.

Como $f_j \leq f$ se sigue que

$$\frac{1}{n} \sum_{k=0}^{n-1} f_j(T^k W) \leq \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W)$$

entonces

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_j(T^k W) = \hat{f}_j \leq C = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k W)$$

$$\therefore \hat{f}_j \leq C \quad \text{casi dondequiera}$$

En consecuencia $\int f_j \, d\mu \leq C\mu(\Omega) \; \forall_j$ y por tanto las integrales $\int f_j \, d\mu$ tienen la misma cota de donde concluimos que f es integrable.

III.4. TEOREMA ERGODICO MEDIO EN UN ESPACIO DE HILBERT.

Para enunciar el teorema y demostrarlo necesitamos algunos resultados importantes, los cuales enunciamos y demostramos en seguida.

III.4.1. DEFINICION: *Sea* U *un operador lineal acotado sobre un espacio de Hilbert* H *(o sea* $U : H \to H$*) se define el operador adjunto de* U *por el requerimiento de que*

$$(U f, g) = (f, U^* g) \quad \forall f, g \in H.$$

(La existencia de este operador adjunto nos lo asegura - el teorema 2 pág. 39 del libro introduction to Hilbert Spaces de Halmos)

III.4.2. PROPOSICION: *Las siguientes condiciones son -- equivalentes (*U *un operador unitario)*

1) $UU^* = U^* U = I$ *. El operador identidad sobre* H.

2) U *es uno a uno, sobre, y* $(f, g) = (Uf, Ug)$ $\forall f, g \in H$.

3) U *es uno a uno, sobre, y* $\| Uf \| = \| f \| \ \forall f \in H$

DEMOSTRACION: *1) ==> 2) Por hipotesis* U *tiene una inversa izquierda y una derecha por tanto* U *es uno a uno y sobre. - También* $(f, g) = (f, U^* U g) = (Uf, Ug)$

2) ==> 3) Si tomamos $f = g$ *entonces*

$$(f , f) = (Uf , Uf)$$

$$\therefore \| f \| = \| Uf \|$$

3) => 2) *De la identidad de polarización tenemos:*

$$4 (Uf, Ug) = \| Uf + Ug \|^2 - \| Uf - Ug \|^2 + i \| Uf + i Ug \|^2 - i \| Uf - i Ug \|^2 =$$

$$= \| U(f+g) \|^2 - \| U(f-g) \|^2 + i \| U(f+ig) \|^2 - i \| U(f-ig) \|^2 =$$

$$= \| f+g \|^2 - \| f-g \|^2 + i \| f+ig \|^2 - i \| f-ig \|^2$$

(por hipótesis)

$$= 4 \ (f , g)$$

$$\therefore (Uf, Ug) = (f , g)$$

2) => 1) *Sea*

$$(U^* Uf, g) = (Uf, U^{**} g) = (Uf, Ug) =$$

$$= (f, g) \ \forall f, g \in H$$

$$\therefore (U^* Uf, g) = (f, g) \Rightarrow U^* Uf = f \Rightarrow U^* U = I$$

pero tambien

$U(U^* U) = (U U^*) U = U \therefore$ *si* U *es sobre entonces* $UU^* = I$

III.4.3. PROPOSICION: *Las siguientes condiciones son equivalentes (U una isometría)*

1) $U^* U = I$

2) $(f, g) = (Uf, Ug) \quad \forall f, g \in H$

3) $\| Uf \| = \| f \| \quad \forall f \in H.$

DEMOSTRACION: *Similar a la demostración de III.4.2.*

En lo siguiente U *representará una isometría sobre* H.

III.4.4. PROPOSICION: *Si* $f \in H$, *entonces* $Uf = f$ *si y solo si* $U^* f = f$

DEMOSTRACION: *Si* $Uf = f$ *entonces*

$$U^* Uf = U^* f, \text{ por ser U una isometría}$$

se tiene $U^* U = I \therefore f = U^* f$. *Inversamente, si* $U^* f = f$, *entonces*

$$\begin{aligned} \| Uf - f \|^2 &= (Uf - f, Uf - f) \\ &= (Uf, Uf) - (f, Uf) - (Uf, f) + (f, f) \\ &= \| Uf \|^2 - (f, Uf) - (Uf, f) + \| f \|^2 \\ &= 2 \| f \|^2 - (U^* f, f) - (f, U^* f) \end{aligned}$$

$= 2\| f \|^2 - 2\ (f, f)$ *por hipotesis*

$= 0$

$$\therefore\ Uf = f$$

III.4.5. PROPOSICION: *Sea* $A_n = n^{-1}(I + U + \ldots + U^{n-1})$ *notemos que* $\| A_n \| = \| \frac{1}{n}(I + U + \ldots + U^{n-1}) \| \leq \frac{1}{n} \sum_{k=0}^{n-1} \| U \|^k = 1.$

Si $E = \{f \in H : \lim_{n\to\infty} A_n f$ *existe (en H)*$\}$, *entonces* E *es un subespacio cerrado de* H.

DEMOSTRACION: *Sea* f_k *una sucesión en* E *tal que* $f_k \to f$, *entonces*

$$\begin{aligned}
\| A_m f - A_n f \| &= \| A_m f - A_m f_k + A_m f_k - A_n f_k + A_n f_k - A_n f \| \\
&\leq \| A_m f - A_m f_k \| + \| A_m f_k - A_n f_k \| + \| A_n f_k - A_n f \| \\
&= \| A_m (f - f_k) \| + \| A_m f_k - A_n f_k \| + \| A_n (f_k - f) \| \\
&\leq \| A_m \| \ \| f - f_k \| + \| A_m f_k - A_n f_k \| + \| A_n \| \ \| f_k - f \| \\
&\leq \| f - f_k \| + \| A_m f_k - A_n f_k \| + \| f_k - f \| \leq \\
&\leq 2 \| f - f_k \| + \| A_m f_k - A_n f_k \|
\end{aligned}$$

$\therefore$ *tenemos que* $\{A_n f\}$ *es una sucesión de Cauchy, entonces* $f \in E$, *lo que prueba que* E *es cerrado.*

Es inmediato que E *es un subespacio.*

III.4.6. PROPOSICION: *Sea* M *el conjunto de elementos de* H *que son invariantes bajo* U, *o sea* $M = \{f \in H : Uf = f\}$.

Notar que M *es un subespacio cerrado de* H, *por la continuidad de* U.

Sea $N_0 = \{g - Ug : g \in H\}$, *si definimos* $\hat{f} = \lim_{n\to\infty} A_n f$ *(donde el límite exista), entonces*

$$f \in M \Rightarrow f \in E \text{ y } \hat{f} = f$$

$$f \in N_0 \Rightarrow f \in E \text{ y } \hat{f} = 0$$

(donde E *es el conjunto de III.4.5).*

DEMOSTRACION: *Si* $f \in M$, *entonces*

$$A_n f = \frac{1}{n}\sum_{k=0}^{n-1} U^k f = \frac{f + Uf + \ldots U^k f}{n} = \frac{n f}{n} = f$$

$$\therefore \hat{f} = f$$

si $f = g - Ug \in N_0$, *entonces*

$$A_n f = \frac{f + Uf + \ldots + U^{n-1} f}{n} = \frac{(g - Ug) + (Ug - U^2 g) + (U^2 g - U^3 g) + \ldots + (U^{n-1} g - U^n g)}{n} =$$

$$= \frac{g - U^n g}{n}$$

$$\therefore \| A_n f \| = \| \frac{g - U^n g}{n} \| < \frac{\| g \| + \| U^n g \|}{n} <$$

$$< \frac{\| g \| + \| U^n \| \, \| g \|}{n} < \frac{\| g \| + \| U \|^n \| g \|}{n}$$

$$< \frac{2 \| g \|}{n} \to 0$$

ya que U *es una isometría y*

$$\| U \| < 1$$

$$\therefore \; \hat{f} = 0$$

III.4.7. PROPOSICION: *Si* $N = \overline{N}_0$, *la cerradura de* N_0, *entonces* H *es la suma directa ortogonal de* M *y* N *(ver [1] pag. 121). entonces por III.4.5 y III.4.6.,* E = H.

DEMOSTRACION: $M \subset H$ *y* N *es un subespacio cerrado de* H. - *Por demostrar que* $M = N^{\perp}$.

Sea $h \in H$ *entonces* $h \perp N$ *si y solo si* $h \perp N_0$

$$\iff (h, g - Ug) = 0 \;\forall g \in H$$

$$\iff (h - U^* h, g) = 0 \;\forall g \in H$$

$$\iff U^* h = h$$

$$\iff Uh = h \text{ por III.4.4.}$$

$$\iff h \in M$$

por tanto

$$M = N^{\perp}$$

entonces

$$H = M \oplus N$$

ahora tenemos que $E \subset H.$

Sea $f \in H$, *entonces* $f = f_1 + f_2$, *donde* $f_1 \in M$ *y* $f_2 \in N$ *por tanto* $f \in E$ *ya que* E *es un subespacio.*

$$\therefore H = E$$

III.4.8. **TEOREMA ERGODICO MEDIO.** *Sea* U *una isometría sobre un espacio de Hilbert* H *y sea* P *la proyección de* H *sobre el espacio* M *de todos los elementos invariantes bajo* U. *Entonces* $\forall f \in H$

$\frac{1}{n}\sum_{k=0}^{n-1} U^k f \to Pf$. *Si* T *es una transformación que preserva medida y* $U = \hat{T}$, *obtenemos la convergencia en* L^2 *de* $f^{(n)}$ *a* $\hat{f}$.

DEMOSTRACION: *Como* $E = H$, $A_n f$ *converge a un límite* $\hat{f}$. *Sea* $f = f_1 + f_2$ *donde* $f_1 \in M$ *y* $f_2 \in N$. *Ahora* $A_n f_1 \to f_1$ *por* III.4.6 *y también* $A_n f_2 \to 0$, *ya que sea* $g \in N_0$ *tal que* $\|f_2 - g\| < \varepsilon$ *entonces*

$$\|A_n f_2\| = \|A_n f_2 - A_n g + A_n g\| \leq \|A_n f_2 - A_n g\| + \|A_n g\|.$$

$$= \|A_n(f_2 - g)\| + \|A_n g\| \leq \|A_n\| \, \|f_2 - g\| + \|A_n - g\|$$

$$< \varepsilon + \|A_n g\|, \quad \textit{y por III.4.6} \quad A_n g \to 0$$

$$\therefore A_n f_2 \to 0$$

entonces

$$A_n f \to f_1 = Pf.$$

BIBLIOGRAFIA

[1] R. ASH. REAL ANALYSIS AND PROBABILITY.
ACADEMIC PRESS 1972

[2] R. B. ASH. AND MELVIN F. GARDNER. TOPICS IN STOCHASTIC PROCESSES

[3] BACHMAN AND. NARICI. FUNCTIONAL ANALYSIS

[4] L. BREIMAN.PROBABILITY. ADDISON WESLEY, 1968

[5] L. BREIMAN. PROBABILITY AND STOCHASTIC PROCESSES;
HOUGHTON MIFFLIN 1969

[6] P. R. HALMOS. ERGODIC THEORY. CHELSEA 1960

[7] A. N. KOLMOGOROV S. V. FOMIN. ELEMENTOS DE LA TEORIA DE FUNCIONES Y DEL ANALISIS FUNCIONAL MIR. 1975

[8] J. NEVEU. MATHEMATICAL FOUNDATIONS OF THE CALCULUS OF PROBABILITY

[9] A. E. TAYLOR. FUNCTIONAL ANALYSIS.

CONTENIDO

INTRODUCCION

El presente trabajo tiene el objeto de presentar una parte de la teoría ergódica de procesos estocásticos, el cual incluye tres capítulos.

En realidad, el desarrollo de éste trabajo esta basado en los cursos a nivel maestría: "Procesos estocásticos y Teoría - Ergódica", impartidos por los profesores Samuel Escarela Cornejo y Joaquín Curiel Cañedo respectivamente. El primero basandose en el libro de J. Neveu, "Mathematical Foundations of the - - Calculus of Probability", y el segundo en los libros: (1) Topics in stochastic processes, de R. B. Ash and Melvin F. Gardner, (2) Probability, de L. Breiman, y (3) Ergodic Theory, de P. R. Hal-- mos..

El primer capítulo lo dedicamos a algunos resultados sobre espacios de Hilbert. Hacemos notar, sin embargo, que este capítulo sólo tiene el fin de familiarizarse con las propiedades de estos espacios, ya que en la sección III.4 se enuncia el " Teorema Ergódico Medio en un espacio de Hilbert". Los Teoremas principales de este capítulo son: "La desigualdad de Bessel "(I.4.1), " El teorema de Riesz-Fischer "(I.5.3) y "La identidad de Parseval "(I.7.1).

En el segundo capítulo se tratarán unicamente los procesos estacionarios y algunas de sus propiedades, que serán útiles en el desarrollo del capítulo III, teniendo como principales resultados la estacionalidad de un proceso y la existencia de procesos estacionarios bilaterales.

El tercer capítulo trata sobre teoría ergódica, el cual es el tema principal de este trabajo. Dentro de este capítulo tienen un papel relevante las transformaciones que preservan medida, funciones invariantes, transformaciones ergódicas, transformaciones mezcla, transformaciones conservativas o recurrentes.

Los teoremas principales de este capítulo son: "Teoremas de recurrencia "(III.2.20, III.2.21, el primero demostrado por Jacobs (1962) y el segundo por Poincaré), " El Teorema Ergódico Maximal" (III.3.4.), " El Teorema Ergódico Puntual" que es el - Teorema principal de esta sección y el "Teorema Ergódico Medio en un espacio de Hilbert".

TERMINOLOGIA: *Un espacio medible* (Ω, β) *es un conjunto no vacío* Ω *con una* σ*-álgebra* β *de subconjuntos de* Ω. *Un espacio de probabilidad* (Ω, β, P) *es un espacio medible* (Ω, β) *con una medida de probabilidad* $\mathbf{P}$ *sobre* β. *Si* $\mathbf{X}$ *es una función de* Ω *en* $\mathbf{S}$, *se dice que* $\mathbf{X}$ *es medible (notacion* $\mathbf{X}: (\Omega, \beta) \to (\mathbf{S}, \varphi)$*) si y sólo si* $\mathbf{X}^{-1}(\mathbf{A}) \in \beta$ *para todo* $\mathbf{A} \in \varphi$. *Si* $\mathbf{S}$ *es el conjunto* $\mathbb{R}$ *de los reales y* $\varphi = \beta(\mathbb{R})$ *la* σ*-álgebra de los conjuntos de Borel en* $\mathbb{R}$, $\mathbf{X}$ *es llamada una variable aleatoria real; Si* $\mathbf{S} = \overline{\mathbb{R}}$ *y* $\varphi = \beta(\overline{\mathbb{R}})$, $\mathbf{X}$ *es llamada una variable aleatoria real extendida. Más general si* $\mathbf{S}$ *es el conjunto* $\mathbb{R}^n$ *de n-adas de números reales,* $\varphi = \beta(\mathbb{R}^n) = \beta_n$, $\mathbf{X}$ *es llamada un vector aleatorio n-dimensional;* $\mathbf{X}$ *puede considerarse como una n-ada* $(X_1, X_2 \ldots X_n)$ *de variables aleatorias (ver [1] pag. 212).*

Si $\mathbf{S} = \mathbb{R}^\infty$, *el conjunto de todas las sucesiones de números reales, y* $\varphi = \beta(\mathbb{R}^\infty) = \beta_\infty$, $\mathbf{X}$ *es llamada una sucesión aleatoria;* $\mathbf{X}$ *puede verse como una sucesión* $(X_1, X_2, \ldots)$ *de variables aleatorias reales.*