UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS

Sección de Posgrado y Segunda Especialización Profesional

Tesis para Optar el Grado Académico de
Maestro en Ciencias con mención en Matemática Aplicada

"EN TRANSFORMACIONES DE RANGO-UNO, 2-MEZCLANTE IMPLICA k-MEZCLANTE"

Presentada por:

Erick Nomberto Dávila Quesquén

Asesor:

Dr. Roger Metzger Alván

Lima-Perú
2010

Dedicado con mucho cariño a mis
padres *Popi y Carmen* por su apoyo
y paciencia en todo este tiempo.

AGRADECIMIENTO

- Mi sincero agradecimiento a los profesores *Dr. Cesar Silva* y *Dr. Roger Metzger*, por sus valiosos aportes con la orientación y corrección durante el desarrollo del presente trabajo.

- Agradezco a la Universidad Nacional de Ingenieria y al Instituto de Matemáticas y Ciencias Afines por la formación profesional que se me ha brindado.

Índice general

Capítulo 1

Introducción

La teoría ergódica es una rama de las matemáticas que usa instrumentos de teoría de la medida para estudiar el comportamiento a largo plazo de los sistemas dinámicos. Su desarrollo inicial fue motivado por problemas en mecánica celeste y mecánica estadística. En mecánica celeste, Henri Poincaré [16] mostró uno de los resultados más notables conocido como el Teorema de Recurrencia de Poincaré que esencialmente dice que para casi todas las órbitas, la órbita intersecta un abierto arbitrario no vacio en una sucesión de valores del tiempo tendiendo a infinito. En mecánica estadística la teoría ergódica se origina en problemas del estudio de gases en seudo-esferas. Un aspecto central de la teoría ergódica es el comportamiento de un sistema dinámico cuando se le permite correr durante mucho tiempo. Esto se expresa a través de teoremas ergódicos que afirman que, bajo ciertas condiciones, el promedio temporal y el promedio espacial de un sistema dinámico coinciden en casi todas partes. Dos de los ejemplos más importantes son los teoremas ergódicos de Birkhoff y von Neumann. Otra de las propiedades de los sistemas dinámicos que estudia la teoría ergódica es la entropía, que puede considerarse como el desorden de un sistema, es decir, cuán homogéneo está el sistema en cuestión. Coloquialmente, un sistema se dice que tiene mucha entropía si está altamente distribuido al azar. Así como la entropía, la propiedad de mezcla, ampliamente estudiada por la teoría ergódica, también cuantifica que tan desordenado se encuentra un sistema dinámico, como por ejemplo, el estudio de desorden de un sistema de gases.

Un papel destacado de la teoría ergódica está dada por sus aplicaciones a los procesos estocásticos usando las nociones de entropía de sistemas dinámicos. Estos métodos estocásticos son aplicados actualmente para predecir resultados estadísticos de los modelos estudiados En geometria, los métodos de la teoría ergódica han sido usados para estudiar los flujos geodésicos en variedades de Riemann. La teoría ergódica junto a las cadenas de Markov forman un contexto en común para aplicaciones en probabilidad. La teoría ergódica también tiene aplicaciones muy bien aprovechadas por el análisis armónico y teoría de números, por ejemplo, se ha probado con técnicas de teoría ergódica que los números primos contienen progresiones aritméticas de tamaño arbitrario. Estos últimos resultados han sido aplicados a la criptografía, que es la tecnología encargada de garantizar la confidencialidad y autenticidad de documentos electrónicos (seguridad informática).

Como se mencionó antes, la propiedad de mezcla de un sistema dinámico permite estudiar el desorden en que se encuentra dicho sistema. En 1949, Vladimir Abramovich Rohlin (1919-1984) en [18] generaliza la noción de transformación mezclante sobre un espacio de probabilidad. La definición de transformación mezclante T esencialmente dice que si A y B son dos eventos (representados por conjuntos medibles) cualquiera entonces A y $T^{-n}B$ se tornan cada vez más independientes cuando n crece. La generalización de transformación mezclante dada por Rohlin consiste en considerar transformaciones T tales que, al tomar tres eventos A, B y C entonces A, $T^{-n}B$ y $T^{-(n+m)}C$ se tornan cada vez más independientes a medida que n y m crecen; a una transformación que hace esto Rohlin las llama transformación 3-mezclante. En la definición de transformación 3-mezclante el número "3" representa la cantidad de eventos tomados; en este contexto, una transformación mezclante sería 2-mezclante. Siguiendo la idea de la definición de una transformación 3-mezclante se puede definir lo que es una transformación k-mezclante para cualquier entero k mayor o igual a 2. En la definición de transformación k-mezclante, al igual que en la definición de transformación 3-mezclante, el número k va a representar la cantidad de eventos tomados.

En 1949, Rohlin no solo generaliza la definición de transformación 2-mezclante a transformación 3-mezclante sino también plantea una interrogante, aún sin resolver, respecto a la relación que hay entre ambas definiciones, es decir pregunta ¿si ser una transformación 2-mezclante bastará para ser k-mezclante ?

En 1983, Steven Arthur Kalikow [9] prueba que 2-mezclante implica 3-mezclante para una familia de transformaciones invertibles llamadas transformaciones de rango-uno. Kalikow también menciona que el método utilizado en su trabajo para probar que 2-mezclante implica 3-mezclante en transformaciones de rango-uno también se puede utilizar para probar el caso general, es decir que también se puede usar para probar que k-mezclante implica $(k+1)$-mezclante en transformaciones de rango-uno. En el artículo de Kalikow solo se presenta la prueba de 2-mezclante implica 3-mezclante en transformaciones de rango-uno, pero no muestra que k-mezclante implica $(k+1)$-mezclante en transformaciones de rango-uno. En 1996, Sebastien Ferenczie en [6] también anuncia que el método usado por Kalikow en [9] se puede generalizar para probar que en transformaciones de rango-uno, k-mezclante implica $(k+1)$-mezclante, pero al igual que Kalikow no presenta detalles de dicha prueba. El presente trabajo está dedicado a dar la prueba minuciosa de la generalización del resultado de Kalikow en [9], es decir, se va a probar usando el método de Kalikow que en transformaciones de rango-uno 2-mezclante no solo implica 3-mezclante sino también k-mezclante, precisamente el objetivo es probar el siguiente teorema:

Teorema Principal. *Cada transformación* 2*-mezclante y de rango-uno también es* k*-mezclante, para todo* $k > 2$*.*

En 1993, Ryzhikov [17] prueba que para el caso de transformaciones de rango finito se tiene que 2-mezclante implica k-mezclante. Cabe resaltar que los metodos usados por Kalikow y Ryzhikov para sus respectivas demostraciones son totalmente distintos.

Otro resultado interesante respecto del problema planteado por Rohlin es el presentado en 1991 por Host en [8]. Host prueba que toda transformación 2-mezclante con espectro singular es k-mezclante. Una nueva interrogante, también sin resolver, que vale la pena mencionar es con respecto a la relación que pueda existir entre transformaciones de rango-uno y transformaciones con espectro singular, es decir, ¿si las transformaciones de rango-uno tienen espectro singular?

El presente trabajo se ha organizado en tres capítulos. El Capítulo 1, está dedicado a dar los conceptos básicos de Teoría de la Medida como por ejemplo espacio de medida e integral de Lebesgue; de Teoría de la Probabilidad como lo es esperanza y varianza y de Teoría Ergódica como lo es transformaciones ergódicas, el Teorema ergódico de Birkhoff, mezcla y mezcla de orden k. El Capítulo 2, aborda lo que es una Transformación de rango-uno para las cuales se va a probar el Teorema Principal. Finalmente, el Capítulo 3 no solo contiene la demostración del Teorema Principal, sino también una transformación que satisface tal teorema y otros avances respecto a la interrogante -aún sin resolver- planteada por Rohlin.

Capítulo 2

Teoría ergódica

En este primer capítulo se da a conocer las nociones básicas de Teoría Ergódica, necesarias para entender el presente trabajo, también recordaremos algunas notaciones de Teoría de la Medida y de Teoría de la Probabilidad.

2.1. Medida y probabilidad

Sea Ω un conjunto no vacío, una σ-**álgebra** de Ω es una colección $\mathcal{A}$ de subconjuntos de Ω tal que

(i) Ω está en $\mathcal{A}$.

(ii) Para todo $A \in \mathcal{A}$ se tiene $A^c \in \mathcal{A}$.

(iii) Si $A = \cup_{n\in\mathbb{N}} A_n$ y $A_n \in \mathcal{A}$ para todo $n \in \mathbb{N}$ entonces $A \in \mathcal{A}$.

A los elementos de la σ-algebra $\mathcal{A}$ los llamaremos **conjuntos medibles**.

Ejemplo 2.1.

1. Si Ω es cualquier conjunto $\mathcal{A} = \mathcal{P}(\Omega)$, la clase formada por todos los subconjuntos de Ω, siempre es una σ-álgebra.

2. También, $\mathcal{A} = \{\emptyset, \Omega\}$ es una σ-álgebra para cualquier conjunto Ω.

3. Sea $\mathcal{C} \subseteq \mathcal{P}(\Omega)$, diremos que $\mathcal{A}$ es la *σ-álgebra generada por $\mathcal{C}$* si es la menor σ-álgebra de Ω que contiene a $\mathcal{C}$.

4. Si $\mathcal{A}$ es una σ-álgebra de Ω y $\Omega' \subseteq \Omega$ entonces la colección $\mathcal{A}' = \{A \cap \Omega' : A \in \mathcal{A}\}$ es una σ-álgebra de Ω'.

5. Sea $\Omega = \mathbb{R}$, la σ-algebra generada por los conjuntos abiertos en $\mathbb{R}$ es llamada *σ-álgebra de Borel* y se denota por $\mathcal{B}$.

Si $\mathcal{A}$ es una σ-algebra de algún conjunto no vacío Ω, llamaremos **medida** sobre $\mathcal{A}$ a una función $\mu : \mathcal{A} \to [0, \infty]$ que cumpla las siguientes propiedades:

(i) La medida del conjunto vacío es cero, $\mu(\emptyset) = 0$.

(ii) La función μ es contablemente aditiva[1], es decir, si $\{A_i\}_{i\in\mathbb{N}}$ es una colección numerable de conjuntos medibles disjuntos dos a dos entonces

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i),$$

donde $\bigsqcup_{i=1}^{\infty} A_i$ denota la unión disjunta de los conjuntos A_i.

Ejemplo 2.2 (Medida de Lebesgue). Sea $A \subseteq \mathbb{R}$ se define la *medida exterior de Lebesgue* de A como

$$\lambda^*(A) = \text{ínf}\left\{\sum_{j=1}^{\infty} |I_j| : A \subseteq \bigcup_{j=1}^{\infty} I_j, \text{ donde } I_j \text{ son intervalos acotados}\right\}, \qquad (2.1)$$

donde $|I|$ denota la longitud del intervalo I.
Un conjunto $E \subseteq \mathbb{R}$ se dice *Lebesgue medible* si para todo $A \subseteq \mathbb{R}$ se cumple

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c) \qquad (2.2)$$

La ecuación (2.2) se conoce como la *condición de Caratheodory* y la clase $\mathcal{L}$ de los conjuntos Lebesgue medibles

$$\mathcal{L} = \{E \subseteq \mathbb{R} : E \text{ es Lebesgue medible}\}$$

es una σ-álgebra de $\mathbb{R}$ más aún la restricción $\lambda = \lambda^*|\mathcal{L}$ es una medida sobre $\mathcal{L}$.

La terna $(\Omega, \mathcal{A}, \mu)$ donde Ω es un conjunto no vacío, $\mathcal{A}$ es una σ-álgebra y μ es una medida sobre $\mathcal{A}$ se llama **espacio de medida**. Un espacio de medida se llama **espacio de medida finito** si $\mu(\Omega) < \infty$, en el caso particular en que $\mu(\Omega) = 1$ el espacio de medida es llamado **espacio de probabilidad**.

Ejemplo 2.3.

1. $(\mathbb{R}, \mathcal{L}, \lambda)$ es un espacio de medida.

2. $([0, 1], \mathcal{L}, \lambda)$ es un espacio de probabilidad.

[1]La condición (ii) también es conocida como σ-aditiva

Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida, un conjunto medible A con $\mu(A) > 0$ se llama **átomo** si para todo conjunto $B \subseteq A$ medible se tiene que $\mu(B) = 0$ ó $\mu(B) = \mu(A)$, un espacio de medida que no tiene átomos se llama **espacio de medida no-atómico**.

Ejemplo 2.4.

1. Los espacios de medida del Ejemplo 2.3 son no atómicos.

2. En $\Omega = \{a, b\}$ considere la σ-álgebra $\mathcal{A} = \{\emptyset, \{a\}, \{b\}, \Omega\}$ y definimos $\mu : \mathcal{A} \to \mathbb{R}$ como $\mu(\{a\}) = \mu(\{b\}) = 1/2$, entonces $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad donde los conjuntos $\{a\}$ y $\{b\}$ son átomos.

A continuación enumeramos algunas propiedades de los espacios de medida cuya prueba se encuentra en [12, pág. 25].

Proposición 2.1. *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida, se cumple lo siguiente:*

1. *Si A y B son conjuntos medibles tales que $A \subseteq B$ y $\mu(A) < \infty$ entonces se tiene que $\mu(B \setminus A) = \mu(B) - \mu(A)$.*

2. *Si $\{A_i\}_{i \in \mathbb{N}}$ es una sucesión de conjuntos medibles que crece hacia A, es decir, $A_1 \subseteq A_2 \subseteq \ldots \subseteq A$, entonces A es un conjunto medible y $\mu(A) = \lim_{i \to \infty} A_i$.*

3. *Si $\{A_i\}_{i \in \mathbb{N}}$ es una sucesión de conjuntos medibles que decrece hacia A, es decir, $A_1 \supseteq A_2 \supseteq \ldots \supseteq A$, y existe $i_0 \in \mathbb{N}$ tal que $\mu(A_{i_0}) < \infty$ entonces A es un conjunto medible y $\mu(A) = \lim_{i \to \infty} A_i$.*

4. *Si los A_i son conjuntos medibles para todo $i \in \mathbb{N}$ entonces $\mu(\cup_{i \in \mathbb{N}} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.*

A continuación se define semi-anillo para un espacio de medida, este semi-anillo cumple ciertas propiedades que serán de mucha utilidad a lo largo del presente trabajo.

Definición 2.1. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida, un *semi-anillo* para $(\Omega, \mathcal{A}, \mu)$ es una colección $\mathcal{C}$ de subconjuntos medibles de Ω tal que

(1) $\mathcal{C}$ es no vacío;

(2) Si $A, B \in \mathcal{C}$ entonces $A \cap B \in \mathcal{C}$;

(3) Si $A, B \in \mathcal{C}$ entonces

$$A \setminus B = \bigsqcup_{j=1}^{n} I_j \ ,$$

donde $I_j \in \mathcal{C}$ son dos a dos disjuntos.

(4) Para cada $A \in \mathcal{A}$

$$\mu(A) = \text{ínf}\left\{\sum_{j=1}^{\infty}\mu(I_j) : A \subset \bigcup_{j=1}^{\infty} I_j, \quad \text{y} \quad I_j \in \mathcal{C} \quad \text{para} \quad j \geq 1\right\}.$$

La siguiente proposición nos dice que si un conjunto es unión de elementos de un semi-anillo entonces también se puede expresar como unión disjunta de elemento de dicho semi-anillo.

Proposición 2.2. *Sea $\mathcal{C}$ un semi-anillo para $(\Omega, \mathcal{A}, \mu)$. Si $A = \bigcup_{n=1}^{\infty} A_n$ donde $A_n \in \mathcal{C}$ entonces A puede escribirse como*

$$A = \bigsqcup_{k=1}^{\infty} C_k \ ,$$

donde $C_k \in \mathcal{C}$.

Demostración. Defina la sucesión de conjuntos $\{B_n\}$ de la siguiente manera $B_1 = A_1$ y $B_n = A_n \setminus (A_1 \cup \ldots \cup A_{n-1})$ para $n > 1$. Luego, los conjuntos $\{B_n\}$ son disjuntos y

$$\bigcup_{n\geq 1} A_n = \bigsqcup_{n \geq 1} B_n.$$

Sin embargo los B_n, en general, no estáns en $\mathcal{C}$. Ahora probaremos que los B_n pueden escribirse como unión finita disjunta de elementos de $\mathcal{C}$. Sea $C_1 = B_1 \in \mathcal{C}$ el primer conjunto. Luego $B_2 = A_2 \setminus A_1$, donde $A_1, A_2 \in \mathcal{C}$, entonces por la propiedad (3) de anillo, B_2 puede escribirse como unión finita y disjunta de elementos de $\mathcal{C}$, que llamaremos $C_2, \ldots, C_{k_1}$.

Para $n \geq 3$, observe que

$$B_n = A_n \setminus (A_1 \cup \ldots \cup A_{n-1}) = (A_n \setminus A_1) \cap (A_n \setminus A_2) \cap \ldots \cap (A_n \setminus A_{n-1}).$$

Además, cada $A_n \setminus A_i$ puede escribirse como unión finita disjunta de elementos de $\mathcal{C}$

$$A_n \setminus A_i = \bigsqcup_{k=1}^{K_{n,i}} E_k^{n,i} \ ,$$

donde $E_k^{n,i} \in \mathcal{C}$.

Consideremos el caso $n = 3$; tenemos

$$B_3 = (A_3 \setminus A_1) \cap (A_3 \setminus A_2) \ ,$$

entonces

$$A_3 \setminus A_1 = \bigsqcup_{k=1}^{K_{3,1}} E_k^{3,1} \quad \text{y} \quad A_3 \setminus A_2 = \bigsqcup_{k=1}^{K_{3,2}} E_k^{3,2}$$

y sea

$$F_{k,l} = E_k^{3,1} \cap E_l^{3,2} \ ,$$

para $1 \leq k \leq K_{3,1}$ y $1 \leq l \leq K_{3,2}$, de donde $\{F_{k,l}\} \in \mathcal{C}$ son disjuntos y además

$$B_3 = \left(\bigsqcup_{k=1}^{K_{3,1}} E_k^{3,1} \right) \cap \left(\bigsqcup_{l=1}^{K_{3,2}} E_l^{3,2} \right) = \bigsqcup_{k,l} F_{k,l}$$

es unión finita disjunta de elementos de $\mathcal{C}$, que podemos escribir $C_{k_1+1}, \ldots, C_{k_2}$. Siguiendo el mismo procedimiento para $B_n, n \geq 3$ tenemos que

$$A = \bigsqcup_{k=1}^{\infty} C_k$$

con $C_k \in \mathcal{C}$ disjuntos. $\square$

Note que en la prueba de la proposición anterior no se ha usado la propiedad (4) de semi-anillo. El siguiente lema nos dice que el semi-anillo es una buena forma de aproximación de los conjuntos medibles.

Lema 2.1. *Sea $\mathcal{C}$ un semi-anillo para el espacio de medida $(\Omega, \mathcal{A}, \mu)$ y A un conjunto medible con $\mu(A) < \infty$ entonces existe un conjunto $H = H(A)$ de la forma*

$$H = \bigcap_{n=1}^{\infty} H_n,$$

tal que

(i) $A \subseteq H \subseteq \ldots H_n \subseteq \ldots \subseteq H_2 \subseteq H_1$.

(ii) $\mu(H_n) < \infty$.

(iii) Cada H_n es unión numerable disjunta de conjuntos en $\mathcal{C}$.

(iv) $\mu(H \setminus A) = 0$.

Demostración. De la definición de semi-anillo, propiedad (4), se sigue que para cualquier $\varepsilon > 0$ existe un conjunto $H(\varepsilon) = \cup_{i=1}^{\infty} I_i$ con $I_i \in \mathcal{C}$ tal que

$$A \subseteq H \text{ y } \mu(H(\varepsilon) \setminus A) < \varepsilon.$$

Tome

$$H_n = H(1) \cap H(1/2) \cap \ldots \cap H(1/n),$$

de la definición de semi-anillo, propiedad (2), y la Proposición 2.2 se sigue que cada H_n es unión disjunta de elementos de $\mathcal{C}$, además de la construcción se sigue que $H_{n+1} \subseteq H_n$, que $A \subseteq H_n$ y cada H_n es de medida finita para todo $n \geq 1$. Si

$$H = \bigcap_{n=1}^{\infty} H_n,$$

entonces H tiene la forma requerida y además

$$\mu(H \setminus A) \leq \mu(H_n \setminus A) < \frac{1}{n},$$

para todo $n \geq 1$, se sigue que $\mu(H \setminus A) = 0$. $\square$

Lema 2.2. *Sea $\mathcal{C}$ un semi-anillo para el espacio de medida $(\Omega, \mathcal{A}, \mu)$, A un conjunto medible con $\mu(A) < \infty$ y $\varepsilon > 0$ entonces existe un conjunto medible A' que es unión disjunta de elementos de $\mathcal{C}$ tal que*

$$\mu(A \triangle A') < \varepsilon.$$

Demostración. De la definición de semi-anillo, propiedad (4), se sigue que existe un conjunto

$$A \subseteq H[\varepsilon] = \bigcup_{i=1}^{\infty} I_i,$$

donde los I_i están en el semi-anillo $\mathcal{C}$ y $\mu(H[\varepsilon]) < \mu(A) + \varepsilon/2$, por otro lado, también se tiene que

$$\lim_{n \to \infty} \mu\left(\bigcup_{i=1}^{n} I_i\right) = \mu\left(\bigcup_{i=1}^{\infty} I_i\right)$$

entonces, existe un $N \in \mathbb{N}$ tal que

$$0 \leq \mu(H[\varepsilon]) - \mu\left(\bigcup_{i=1}^{N} I_i\right) < \frac{\varepsilon}{2}.$$

Tome $A' = \cup_{i=1}^{N} I_i$, entonces

$$\begin{aligned} \mu(A \triangle A') &= \mu(A' \setminus A) + \mu(A \setminus A') \\ &\leq \mu(H[\varepsilon] \setminus A) + \mu(H[\varepsilon] \setminus A) \\ &< \varepsilon. \end{aligned}$$

Finalmente de la Proposición 2.2 podemos suponer que A' es unión disjunta de elementos de $\mathcal{C}$. $\square$

Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $f : \Omega \to \mathbb{R}$ una función. Se dice que f **es medible** si $f^{-1}((a, \infty))$ es un conjunto medible para todo $a \in \mathbb{R}$.

La siguiente proposición nos da otras formas equivalentes de establecer cuando una función es medible y la prueba es inmediata a partir de la definicíon de σ-álgebra.

Proposición 2.3. *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $f : \Omega \to \mathbb{R}$ una función, entonces las siguientes afirmaciones son equivalentes*

(i) *f es medible.*

(ii) *El conjunto $f^{-1}([a, \infty))$ es medible para todo $a \in \mathbb{R}$.*

(iii) *El conjunto $f^{-1}((-\infty, a))$ es medible para todo $a \in \mathbb{R}$.*

(iv) *El conjunto $f^{-1}((-\infty, a])$ es medible para todo $a \in \mathbb{R}$.*

Ejemplo 2.5.

1. Sea $E \subseteq \Omega$ definimos la *función característica* de E como la función $\mathbb{I}_E : \Omega \to \mathbb{R}$ tal que
$$\mathbb{I}_E(\omega) = \begin{cases} 1 \text{ si } \omega \in E \\ 0 \text{ si } \omega \in \Omega \setminus E. \end{cases} \tag{2.3}$$
Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida entonces $\mathbb{I}_E$ es medible si, y sólo si, el conjunto E es un conjunto medible.

2. Decimos que $s : \Omega \to \mathbb{R}$ es una *función simple* si la imagen de s es un conjunto finito, es decir $s(\Omega) = \{a_1, a_2, \ldots, a_n\} \subseteq \mathbb{R}$, en tal caso s queda definida por
$$s(\omega) = \sum_{i=1}^{n} a_i \mathbb{I}_{E_i}(\omega) \tag{2.4}$$
donde $E_i = s^{-1}(a_i)$. Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida entonces s es medible si, y sólo si, los conjuntos E_i son conjuntos medibles.

El siguiente teorema nos dice que podemos aproximarnos a toda función medible positiva por funciones simples y la prueba puede encontrarla en [12, pág. 34].

Teorema 2.1 (Aproximación)**.** *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $f : \Omega \to [0, \infty)$ una función medible (positiva), entonces existe una sucesión de funciones simples medibles $\{s_n\}_{n \in \mathbb{N}}$ tal que*

(i) *La sucesión $\{s_n\}_{n \in \mathbb{N}}$ es monotona creciente, es decir $s_{n-1} \leq s_n \leq s_{n+1}$.*

(ii) *La sucesión $\{s_n\}_{n \in \mathbb{N}}$ convenge a la función f, es decir $\lim_{n \to \infty} s_n(\omega) = f(\omega)$ para todo $\omega \in \Omega$.*

2.1.1. Integral de Lebesgue

Ahora pasamos a definir la integral de funciones medibles, empezamos definiendo la integral de funciones simples no negativas. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $s : \Omega \to \mathbb{R}$ una función simple, medible tal que $s(\omega) \geq 0$ para todo $\omega \in \Omega$. Si $s(\Omega) = \{a_1, a_2, \ldots, a_n\}$ y $E_i = h^{-1}(a_i)$, definimos la **integral de** s, denotada por $\int_\Omega s d\mu$, como el número extendido

$$\int_\Omega s d\mu = \sum_{i=1}^{n} a_i \mu(E_i).$$

La siguiente proposición nos dice que la integral de funciones simples respeta desigualdades y la prueba se sigue de la definición.

Proposición 2.4. *Sea* $(\Omega, \mathcal{A}, \mu)$ *un espacio de medida y* $s_1, s_2 : \Omega \to \mathbb{R}$ *funciones simples, medibles y no negativas tales que* $s_1 \leq s_2$*, entonces* $\int_\Omega s_1 d\mu \leq \int_\Omega s_2 d\mu$*.*

Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $f : \Omega \to \mathbb{R}$ una función, medible no negativa. Definimos la **integral de Lebesgue** de f, denotada por $\int_\Omega f d\mu$, como

$$\int_\Omega f d\mu = \sup\left\{\int_\Omega s d\mu : s \in S\right\},$$

donde S es el conjunto de todas las funciones simples no negativas tal que $s \leq f$, es decir

$$S = \{s : \Omega \to \mathbb{R} : s \text{ es simple y } 0 \leq s \leq f\}.$$

Note que, con esta definición, habría dos formas de calcular la integral de una función simple, sin embargo estas dos definiciones coinciden en este caso.

La siguiente proposición es una generalización de la Proposición 2.4 y su demostración se sigue de la definición de integral de una función no negativa y de la Proposición 2.4.

Proposición 2.5. *Sea* $(\Omega, \mathcal{A}, \mu)$ *un espacio de medida y* $f_1, f_2 : \Omega \to \mathbb{R}$ *funciones medibles y no negativas tales que* $f_1 \leq f_2$*, entonces* $\int_\Omega f_1 d\mu \leq \int_\Omega f_2 d\mu$*.*

El siguiente es un teorema clásico de Teoría de la Medida y los detalles de su demostración puede encontrarlo en [12, pág. 38] ó [19, pág. 22].

Teorema 2.2 (Convergencia Monótona de Lebesgue)**.** *Sea* $(\Omega, \mathcal{A}, \mu)$ *un espacio de medida y* $\{f_n\}_{n\in\mathbb{N}}$ *una sucesión de funciones medibles no negativas. Suponga que*

(i) $f_n(\omega) \leq f_{n+1}(\omega)$ *para todo* $\omega \in \Omega$ *y todo* $n \in \mathbb{N}$*.*

(ii) Si definimos $f(\omega) = \lim_{n\to\infty} f_n(\omega)$ *para todo* $\omega \in \Omega$

entonces f es medible y

$$\int_\Omega f d\mu = \lim_{n\to\infty} f_n d\mu.$$

Si juntamos este teorema con el Teorema 2.1, tenemos que siempre es posible para todo f medible exhibir una sucesión de funciones simples s_n tal que $s_n \leq s_{n+1}$, $\lim_{n\to\infty} s_n = f$ y $\lim_{n\to\infty} \int_\Omega s_n d\mu = \int_\Omega f d\mu$.

Ahora definiremos la integral para una función medible cualquiera. Debido a que trataremos de funciones que pueden tomar valores positivos y negativos, necesitaremos las siguientes definiciones.

Dada una función $f : \Omega \to \mathbb{R}$ definimos:

(i) La **parte positiva de** f, como la función $f_+ : \Omega \to \mathbb{R}$ tal que

$$f_+ = \max\{f, 0\}.$$

(ii) La **parte negativa de** f, como la función $f_- : \Omega \to \mathbb{R}$ tal que

$$f_- = \max\{-f, 0\} = -\min\{f, 0\}.$$

(iii) El **valor absoluto de** f, como la función $|f| : \Omega \to \mathbb{R}$ tal que

$$|f|(\omega) = |f(\omega)|.$$

Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $f : \Omega \to \mathbb{R}$ una función medible. Decimos que f es una función **Lebesgue integrable** si

$$\int_\Omega |f| d\mu < \infty,$$

en tal caso se define la integral de f como

$$\int_\Omega f d\mu = \int_\Omega f_+ d\mu - \int_\Omega f_- d\mu.$$

El siguiente teorema, uno de los resultados más importantes en Teoría de la Medida, nos da una condición para que el límite de una sucesión de funciones medibles integrables sea también una función integrable, la prueba del teorema puede encontrarla en [12, pág. 45] ó [19, pág. 27].

Teorema 2.3 (Convergencia Dominada)**.** *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $\{f_n\}_{n\in\mathbb{N}}$ una sucesión de funciones medibles tal que $\lim_{n\to\infty} f_n(\omega)$ existe para todo $\omega \in \Omega$. Si existe una función $g : \Omega \to \mathbb{R}$ integrable tal que para todo $n \in \mathbb{N}$ se tiene $|f_n(\omega)| \leq g(\omega)$ para todo $\omega \in \Omega$, entonces*

(i) La función $f = \lim_{n\to\infty} f_n$ es integrable,

(ii) $\lim_{n\to\infty} \int_\Omega |f_n - f| d\mu = 0$ y

(iii) $\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu$.

2.1.2. Esperanza y varianza

A continuación se da a conocer algunas definiciones y teoremas, mayormente vistos en un curso de Teoría de la Probabilidad, que serán de utilidad para enunciar y probar más adelante los Lemas 2.6 y 2.7 que aplicaremos en el Capítulo 3.

Si $f : \Omega \to \mathbb{R}$ es una función medible no negativa definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, la **Esperanza Matemática** de f se define como el número real

$$E(f) = \int_\Omega f d\mu.$$

Se sigue de la definición que si la función f es simple, es decir $f(\Omega) = \{a_1, a_2, \ldots, a_n\}$, entonces

$$E(f) = \sum_{i=1}^{n} a_i \mu(f^{-1}(\{a_i\})).$$

Teorema 2.4 (Desigualdad de Márkov)**.** *Si $f : \Omega \to \mathbb{R}$ es una función medible no negativa definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ y $\varepsilon > 0$ entonces*

$$\mu\{\omega \in \Omega : f(\omega) \geq \varepsilon\} \leq \frac{E(f)}{\varepsilon}. \tag{2.5}$$

Demostración. En efecto, se tiene que

$$\begin{aligned} \varepsilon\mu\{\omega \in \Omega : f(\omega) \geq \varepsilon\} &= \varepsilon E(\mathbb{I}_{\{\omega\in\Omega: f(\omega)\geq\varepsilon\}}) \\ &= E(\varepsilon\mathbb{I}_{\{\omega\in\Omega: f(\omega)\geq\varepsilon\}}) \\ &\leq E(f), \end{aligned}$$

donde la última desigualdad se sigue desde que $\varepsilon\mathbb{I}_{\{\omega\in\Omega: f(\omega)\geq\varepsilon\}}(\omega) \leq f(\omega)$ para todo $\omega \in \Omega$. □

La **Varianza** de f se define como el número real

$$Var(f) = (E(f^2) - E(f))^2.$$

Teorema 2.5 (Desigualdad de Chebychev)**.** *Si $f : \Omega \to \mathbb{R}$ una función medible no negativa definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ y $\varepsilon > 0$ entonces*

$$\mu\left(\{\omega \in \Omega : |f(\omega) - E(f)| \geq \varepsilon\}\right) \leq \frac{Var(f)}{\varepsilon^2}. \tag{2.6}$$

La prueba de la Desigualdad de Chebychev puede encontrarla en [7, pág. 194-195].

2.2. Transformaciones ergódicas

Una **transformación** es una función para la cual el dominio y el rango son los mismos, es decir la imagen de cada punto esta en el dominio de la función. Si $T : \Omega \to \Omega$ es una transformación en Ω, el n–**ésimo iterado** de $\omega \in \Omega$, escrito por $T^n(\omega)$, se define por

$$\begin{aligned} T^0(\omega) &= \omega \\ T^{n+1}(\omega) &= T \circ T^n(\omega) \text{ para } n \geq 0. \end{aligned}$$

Ademas, se dice que T es una **transformación invertible** si es una biyección, en tal caso T^{-1} también es una transformación.

Si $T : \Omega \to \Omega$ es una transformación y $\omega \in \Omega$, se llama **órbita positiva** de ω al conjunto $\{T^n(\omega)\}_{n\geq 0}$. Si T es una transformación invertible se llama **órbita completa**, o simplemente **órbita**, de ω al conjunto $\{T^n(\omega)\}_{n\in\mathbb{Z}}$. Un punto $\omega \in \Omega$ se dice **periódico** bajo $T : \Omega \to \Omega$, ó T-**periódico**, si el conjunto $\{T^n(\omega)\}_{n\geq 0}$ es finito, en tal caso el **periodo** de ω es $\#\{T^n(\omega)\}_{n\geq 0}$.

Ejemplo 2.6.

1. La función identidad, $Id(\omega) = \omega$ es una transformación.

2. Dado $a \in (0, 1)$ definimos la transformación *rotación por a* como

$$\begin{aligned} R_a : [0,1) &\to [0,1) \\ \omega &\mapsto R_a(\omega) = \omega + a - [\![\omega + a]\!] \end{aligned} \tag{2.7}$$

   Observe que $R_a(\omega) \equiv \omega \mod a$.

3. Sea $T : [0, 1) \to [0, 1)$ la transformación definida por

$$T(\omega) = \begin{cases} 2\omega & \text{si } 0 \leq \omega < \frac{1}{2} \\ 2\omega - 1 & \text{si } \frac{1}{2} \leq \omega < 1 \end{cases} \tag{2.8}$$

   es no invertible y se le conoce como *transformación doble.*

Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida, una transformación $T : \Omega \to \Omega$ se dice que es una **transformación medible** si el conjunto

$$T^{-1}(A) = \{\omega \in \Omega : T(\omega) \in A\}$$

está en $\mathcal{A}$ para todo $A \in \mathcal{A}$ y se dice que T **preserva la medida** si es medible y además

$$\mu(T^{-1}(A)) = \mu(A) \text{ para todo } A \in \mathcal{A}, \tag{2.9}$$

en este caso se dice que μ es una **medida invariante** por T.

Se sigue de la definición que si T es una transformación invertible medible entonces, T preserva la medida si, y sólo si, T^{-1} preserva la medida.

Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida, finito y $T : \Omega \to \Omega$ es una transformación que preserva la medida entonces decimos que $(\Omega, \mathcal{A}, \mu, T)$ es un **sistema dinámico**.

El siguiente teorema nos dice que si una transformación medible satisface la Ecuación 2.9 para un semi-anillo del espacio de probabilidad entonces esta transformación preserva la medida.

Teorema 2.6. *Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad completo*[2] *con semi-anillo $\mathcal{C}$, $T : \Omega \to \Omega$ es una tranformación y si para todo $I \in \mathcal{C}$*

(i) $T^{-1}(I)$ es medible, y

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

entonces T preserva la medida.

Demostración. Sea $H(A)$ del Lema 2.1, es decir

$$H(A) = \bigcap_{n_1}^{\infty} H_n,$$

con $H_{n+1} \subseteq H_n$ y $A \subseteq H_n$ para todo $n \geq 1$, cada H_n es de la forma

$$H_n = \bigsqcup_{j=0}^{\infty} C_{n,j},$$

donde $C_{n,j} \in \mathcal{C}$. Así, el conjunto

$$T^{-1}H_n = \bigsqcup_{j\geq 1} T^{-1}C_{n,j}$$

es medible, además

$$\begin{aligned} \mu(T^{-1}H_n) &= \sum_{j\geq 1} \mu(T^{-1}C_{n,j}) \\ &= \sum_{j\geq 1} \mu(C_{n,j}) \\ &= \mu(H_n). \end{aligned}$$

[2] $(\Omega, \mathcal{A}, \mu)$ es *completo* si: $A \in \mathcal{A}$ y $\mu(A) = 0$ entonces $B \in \mathcal{A}$ para todo $B \subseteq A$.

También se tiene que $T^{-1}H(A) = \cap_{n\geq 1}T^{-1}H_n$ es medible y

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

Por otro lado, si denotamos $N = H(A) \setminus A$, recuerde que $A \subseteq H(A)$, entonces $\mu(N) = 0$, usando nuevamente el Lema 2.1, esta vez para N, se tiene que existe $H(N)$ tal que $N \subseteq H(N)$ y $\mu(T^{-1}H(N)) = 0$, por lo tanto $\mu(T^{-1}N) = 0$. Luego como

$$T^{-1}A = T^{-1}H(A) \setminus T^{-1}N$$

entonces

$$\begin{aligned} \mu(T^{-1}A) &= \mu(T^{-1}H(A)) - \mu(T^{-1}N) \\ &= \mu(H(A)) = \mu(A). \end{aligned}$$

□

Ejemplo 2.7. La transformación doble es una transformación no invertible que preserva la medida en $([0,1), \mathcal{L}, \lambda)$, es suficiente verificar en el semi-anillo de los intervalos diádicos en $[0,1)$, es decir intervalos de la forma

$$\left[\frac{k}{2^i}, \frac{k+1}{2^i}\right).$$

donde $i \in \mathbb{N}$ y $0 \leq k \leq 2^i - 1$.

Una transformación que preserva la medida T definida en un espacio de medida $(\Omega, \mathcal{A}, \mu)$ se dice que es **recurrente** si para todo conjunto medible A de medida positiva existe un conjunto de medida nula $N \subseteq A$ tal que si $\omega \in A \setminus N$ entonces $T^n(\omega) \in A$ para infinitos valores $n \geq 0$.

El siguiente teorema fue probado por Poincaré en 1899, que es usado en el estudio de mecánica celeste, dice esencialmente que toda transformación que preserva medida sobre un espacio de probabilidad es siempre recurrente.

Teorema 2.7 (Recurrencia de Poincaré)**.** *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad. Si $T : \Omega \to \Omega$ es una transformación que preserva la medida entonces T es recurrente.*

Los detalles de la prueba del teorema de Recurrencia de Poincaré puede encontrarlo en [11, pag. 32].

Sea $T : \Omega \to \Omega$ una transformación. Decimos que $A \subseteq \Omega$ es **positivamente invariante** si $A \subseteq T^{-1}(A)$, es decir

$$\text{si } \omega \in A \text{ entonces } T(\omega) \in A,$$

en tal caso T restricto a A también define una transformación $T : A \to A$, y decimos que $A \subseteq \Omega$ es **estrictamente invariante** si $A = T^{-1}(A)$, es decir

$$\omega \in A \text{ si y sólo si } T(\omega) \in A.$$

Note que la definición de positivamente invariante e invariante dependen de la transformación T tomada.

Ejemplo 2.8.

1. Sea $\Omega = [0, \infty)$ y $T : \Omega \to \Omega$ definido por $T(\omega) = \omega + 1$, entonces $A = [1, \infty)$ es positivamente invariante pero no estrictamente invariante.

2. Sea $T : \Omega \to \Omega$ una transformación cualquiera, los conjuntos $\emptyset$ y Ω son positivamente invariantes y Ω es estrictamente invariante si y sólo si T es sobreyectiva.

Ergodicidad es uno de los conceptos más importantes y fue introducido por Birkhoff y Smith en 1928 [2], hay varias formas de definir ergodicidad, a continuación se dá una definición básica.

Definición 2.2. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y $T : \Omega \to \Omega$ una transformación que preserva la medida. Decimos que T es una *transformación ergódica*, ó simplemente que T es ergódico, si para todo conjunto medible A estrictamente invariante, se tiene que $\mu(A) = 0$ ó $\mu(A^c) = 0$.

La definición de transformación ergódica se puede entender como un sistema que no tiene subconjunto propio estrictamente invariante, es decir como un sistema básico que no se puede descomponer.

Ejemplo 2.9. Si $a \in \mathbb{R} \setminus \mathbb{Q}$, la rotación por a (ver Ejemplo 2.6) es una transformación ergódica.

Ahora introduciremos una técnica de aproximación que nos será útil para probar ergodicidad. Si $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y A e I (donde I es tomado mayormente en algún semi-anillo para $\mathcal{A}$) son conjuntos medibles, y $1 > \delta > 0$ diremos que I está $(1-\delta)$-**lleno** de A si

$$\mu(A \cap I) > (1 - \delta)\mu(I).$$

Lema 2.3. *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida no atómico con semi-anillo $\mathcal{C}$. Dado un conjunto A de medida positiva y $\varepsilon > 0$ entonces existe $I \in \mathcal{C}$ tal que I está $(1 - \varepsilon)$-lleno de A.*

Demostración. Considere $\varepsilon > 0$ suficientemente pequeño tal que

$$\varepsilon < \frac{\delta}{2-\delta}.$$

Se sigue del Lema 2.2 que existe $A' = \sqcup_{i=1}^{N} I_i$, con $I_i \in \mathcal{C}$, tal que

$$\mu(A \triangle A') < \varepsilon \mu(A),$$

entonces

$$\mu(A' \cap A) > (1-\varepsilon)\mu(A),$$

y también

$$\mu(A') < \mu(A) + \varepsilon\mu(A) = (1+\varepsilon)\mu(A).$$

Si suponemos que para todo $i \in \{1, \ldots, N\}$ se tiene que

$$\mu(A \cap I_i) \leq (1-\delta)\mu(I_i),$$

entonces $\mu(A \cap A') \leq (1-\delta)\mu(A')$, se sigue que

$$\begin{aligned} \mu(A \cap A') &\leq (1-\delta)(1+\varepsilon)\mu(A) \\ &< (1-\varepsilon)\mu(A) \\ &< \mu(A \cap A'), \end{aligned}$$

donde la segunda desigualdad se sigue de la elección del ε tomado, llegando así a una contradicción; lo que prueba el lema. □

2.3. El Teorema ergódico

El Teorema Ergódico, probado en 1931, primero por J. von Neumann [14] en el caso de convergencia en media y por G. D. Birkhoff [3] en el caso de convergencia puntual, resuelve una interrogante importante que surgió en mecánica estadística: *dar una condición bajo la cual el promedio temporal y el promedio espacial de un sistema dinámico coinciden.* La versión del Teorema Ergódico en convergencia puntual, conocido como el Teorema Ergódico de Birkhoff, en su versión más simple dice que una transformación T que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ es ergódico si y sólo si

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{I}_A(T^i(\omega)) = \mu(A)$$

para todo conjunto medible A en Ω y cada $\omega \in \Omega \setminus N$ donde $N = N(A)$ es un conjunto de medida nula que depende de A.

Ahora introducimos algunas notaciones que usaremos en esta sección

$$
\begin{aligned}
f_n(\omega) &= \sum_{i=0}^{n-1} f(T^i\omega),\ n \geq 1, & (2.10)\\
f_*(\omega) &= \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega), & (2.11)\\
f^*(\omega) &= \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega). & (2.12)
\end{aligned}
$$

Lema 2.4. *Las funciones $f_*(\omega)$ y $f^*(\omega)$ son T-invariantes, es decir*

$$
\begin{aligned}
f_*(\omega) &= f_*(T\omega),\\
f^*(\omega) &= f^*(T\omega).
\end{aligned}
$$

Demostración. De la definición se tiene

$$
\begin{aligned}
\frac{1}{n} f_n(T\omega) &= \frac{1}{n}\sum_{i=0}^{n-1} f(T^{i+1}\omega)\\
&= \frac{1}{n}\sum_{i=0}^{n} f(T^i\omega) - \frac{1}{n} f(\omega)\\
&= \frac{n+1}{n}\frac{1}{n+1} f_{n+1}(\omega) - \frac{1}{n} f(\omega),
\end{aligned}
$$

luego

$$
\begin{aligned}
f_*(T\omega) &= \liminf_{n\to\infty} \frac{1}{n} f_n(T\omega)\\
&= \liminf_{n\to\infty} \left[\frac{n+1}{n}\frac{1}{n+1} f_{n+1}(\omega) - \frac{1}{n} f(\omega)\right]\\
&= \liminf_{n\to\infty} f_{n+1}(\omega)\\
&= f_*(\omega).
\end{aligned}
$$

Analogamente se prueba que $f^*(T\omega) = f^*(\omega)$. □

El siguiente lema cuya prueba se encuentra en [20, pág. 179], será usado para la demostración del Teorema Ergódico de Birkhoff.

Lema 2.5. *Sea f una función medible y $p \in \mathbb{N}$*

1. Si definimos

$$E_p = \{\omega \in \Omega :\ f_n(\omega) \geq 0 \text{ para todo } n, 1 \leq n \leq p\},$$

entonces para todo $n \geq p$ y para c.t.p. $\omega \in \Omega$

$$f_n(\omega) \leq \sum_{i=0}^{n-1} f(T^i\omega)\mathbb{I}_{E_p}(T^i\omega) + \sum_{i=n-p}^{n-1} |f(T^i\omega)|. \tag{2.13}$$

2. Si para cada $r \in \mathbb{R}$ definimos

$$E_p^r = \{\omega \in \Omega : \ \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega) \geq r \ \textit{para todo} \ 1 \leq n \leq p\},$$

entonces

$$\int f d\mu \leq \int_{E_p^r} f d\mu + r\left(1 - \mu\left(E_p^r\right)\right). \tag{2.14}$$

A continuación se da el Teorema Ergódico de Birkhoff y una de sus demostraciones, que puede encontrarla en "Invitation to Ergodic Theory" de Cesar Silva (ver [20, pág. 177]).

Teorema 2.8 (Teorema Ergódico de Birkhoff)**.** *Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de probabilidad y $T : \Omega \to \Omega$ una transformación ergódica que preserva la medida. Si $f : \Omega \to \mathbb{R}$ es una función integrable entonces*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega) = \int f d\mu \ \textit{para c.t.p.} \tag{2.15}$$

Demostración. Primero probaremos que para toda función integrable f

$$\int f d\mu \leq f_*(\omega) \text{ c.t.p.} \tag{2.16}$$

Sea

$$A = \left\{\omega \in \Omega : \ f_*(\omega) < \int f d\mu\right\}.$$

Para probar (2.16) es suficiente probar que $\mu(A) = 0$. Dado $r \in \mathbb{Q}$, sea

$$C_r = \{\omega \in \Omega : f_*(\omega) < r < \int f d\mu\}, \tag{2.17}$$

entonces podemos escribir

$$A = \bigcup_{r\in\mathbb{Q}} C_r.$$

Si suponemos que $\mu(A) > 0$ entonces existe un $r \in \mathbb{Q}$ para el cual $\mu(C_r) > 0$.

Afirmación. Dado $r \in \mathbb{Q}$, el conjunto $C_r \triangle T^{-1}C_r$ es de medida nula.
En efecto, pues si $\omega \in C_r$ entonces del Lema 2.4 se sigue que $\omega \in T^{-1}(C_r)$, luego $C_r \subseteq T^{-1}(C_r)$, así

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

Lo que prueba la afirmación.

Como T es ergódico se sigue que $\mu(C_r) = 1$. Desde que

$$E_p^r = \{\omega \in \Omega : \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega) \geq r, \text{ para todo } 1 \leq n \leq p\},$$

entonces el hecho que $\mu(C_r) = 1$ implica que

$$\mu\left(\bigcap_{p=1}^{\infty} E_p^r\right) = 0.$$

Así

$$\lim_{p\to\infty} \mu(E_p^r) = 0,$$

y por (2.14)

$$\int f d\mu \leq r,$$

lo que contradice la elección del r. Por lo tanto $\mu(A) = 0$. Lo que prueba (2.16). Aplicando (2.16) a $-f$, se tiene

$$\begin{aligned} \int -f d\mu &\leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} -f(T^i\omega) \text{ c.t.p.} \quad \text{ó} \\ \int f d\mu &\geq \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega) \text{ c.t.p.} \end{aligned}$$

De esto y (2.16) se tiene

$$\int f d\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega) \text{ c.t.p.},$$

lo que completa la prueba del teorema. $\square$

Una aplicación interesante del Teorema Ergódico de Birkhoff es la siguiente caracterización para ergodicidad.

Teorema 2.9. *Si T es una transformación que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ entonces las siguientes afirmaciones son equivalentes*

1. *T es ergódico.*

2. *Si A es conjunto medible*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega) = \mu(A) \ \textit{c.t.p.}. \tag{2.18}$$

3. Dados A y B conjuntos medibles

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}(A)\cap B) = \mu(A)\mu(B). \tag{2.19}$$

Demostración. Suponga que T es ergódico y sean A, B conjuntos medibles entonces $\mathbb{I}_A$ es integrable y por el teorema de Birkhoff (Teorema 2.8) se tiene

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega) = \mu(A) \text{ c.t.p..}$$

Si se cumple (2.18) entonces

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega)\mathbb{I}_B(\omega) = \mu(A)\mathbb{I}_B(\omega) \text{ c.t.p.}$$

Observe que para todo $n > 0$

$$\left|\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega)\mathbb{I}_B(\omega)\right| \leq 1, \text{ c.t.p.}$$

Luego por el teorema de convergencia dominada (Teorema 2.3) se tiene

$$\begin{aligned}
\lim_{n\to\infty}\int \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega)\mathbb{I}_B(\omega)d\mu(\omega) &= \int \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{I}_A(T^i\omega)\mathbb{I}_B(\omega)d\mu(\omega) \\
&= \int \mu(A)\mathbb{I}_B(\omega)d\mu(\omega) \\
&= \mu(A)\mu(B).
\end{aligned}$$

Ahora suponga que se cumple (2.19), observe que si A es un conjunto medible T-invariante entonces

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

Por otro lado, si se hace $B = A$ en (2.19), se tiene

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}(A)\cap A) = \mu(A)^2.$$

Así, $\mu(A) = \mu(A)^2$ entonces $\mu(A) = 0$ ó $\mu(A) = 1$ y por lo tanto T es ergódico. □

2.4. Mezcla

En el Teorema 2.9 vemos que una transformación que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ es ergódico si, y sólo si, para todo par de conjuntos medibles

A, B se tiene

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

Ahora, cabe preguntar si el límite de la sucesión $\mu(T^{-i}(A) \cap B)$ existe y si puede ser $\mu(A)\mu(B)$; esto motiva la siguiente definición. Una transformación T que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ se dice que es **mezclante** si para cada par de conjuntos medibles A_0, A_1 se tiene

$$\lim_{N\to\infty} \mu(A_0 \cap T^{-N}(A_1)) = \mu(A_0)\mu(A_1). \tag{2.20}$$

De Teoría de la Probabilidad se tiene que dos *eventos*, representados por conjuntos medibles A_0 y A_1, son llamados *independientes* si $\mu(A_0 \cap A_1) = \mu(A_0)\mu(A_1)$. Si T es una transformación mezclante entonces $T^{-N}(A_0)$ es "*eventualmente independiente*" de A_1.

Ejemplo 2.10. La transformación doble (Ejemplo 2.6) es mezclante no invertible.

2.4.1. Mezcla de orden k

La definición de transformación mezclante dada anteriormente es también conocida como 2-mezclante, en la siguiente generalización; sea $k \in \mathbb{N}$ y T una transformación que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, se dice que T es k-**mezclante** ó **mezcla de orden** k si dado conjuntos medibles A_0, A_1, $\ldots A_{k-1}$ se tiene que

$$\lim_{\substack{N_j\to\infty \\ 1\le j\le k-1}} \mu(A_0 \cap T^{-N_1}A_1 \cap T^{-(N_1+N_2)}A_2 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}) = \prod_{j=0}^{k-1} \mu(A_j). \tag{2.21}$$

Note de la definición, que si T es una transformación k-mezclante entonces es $(k-1)$-mezclante, pues basta considerar en la ecuación (2.21) a $A_{k-1} = \Omega$.

El siguiente teorema nos presenta la relación entre transformaciones mezclantes y ergódicas.

Teorema 2.10. *Toda transformación mezclante es ergódica.*

Demostración. Sea A un conjunto estrictamente invariante respecto a una transformación mezclante T, entonces

$$\begin{aligned} \mu(A)\mu(A^c) &= \lim_{N\to\infty} \mu(T^{-N}A \cap A^c) \\ &= \mu(A \cap A^c) \\ &= 0, \end{aligned}$$

por lo tanto $\mu(A) = 0$ ó $\mu(A^c) = 0$, lo que muestra que T es ergódico. □

La siguiente proposición da una condición necesaria sobre la medida para que la transformación sea mezclante.

Proposición 2.6. *Si T es una transformación mezclante sobre el espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ no trivial entonces la medida es no atómica.*

Demostración. Suponga por el contrario que existe $A \in \mathcal{A}$ tal que $0 < \mu(A) < 1$ y que para todo $B \in \mathcal{A}$ con $B \subseteq A$ se tiene que $\mu(B) = 0$ ó $\mu(B) = \mu(A)$, es decir A es un átomo; luego por ser T una transformación mezclante se tiene

$$\lim_{N\to\infty} \mu(T^{-N}(A) \cap A) = \mu(A)^2 \tag{2.22}$$

pero $(T^{-N}(A) \cap A) \subseteq A$ entonces $\mu(T^{-N}(A) \cap A) \to 0$ ó 1, lo que contradice 2.22. □

El siguiente teorema nos dice que es suficiente verificar la ecuación (2.21) en un semi-anillo para garantizar que la transformación es k-mezclante.

Teorema 2.11. *Si T es una transformación que preserva medida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ con semi-anillo $\mathcal{C}$ y si para todo $I_0, I_1, \ldots, I_{k-1} \in \mathcal{C}$ se tiene*

$$\lim_{\substack{N_j\to\infty \\ 1\le j\le k-1}} \mu(I_0 \cap T^{-N_1} I_1 \cap T^{-(N_1+N_2)} I_2 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})} I_{k-1}) = \prod_{j=0}^{k-1} \mu(I_j)$$

entonces T es k-mezclante.

Demostración. Sea $\varepsilon > 0$ y $A_0, A_1, \ldots, A_{k-1}$ conjuntos medibles entonces, por el Lema 2.2, existen conjuntos $A'_j = \sqcup_{\ell=1}^{n(j)} I_{j,\ell}$ con $I_{j,\ell} \in \mathcal{C}$, tales que

$$\mu(A_j \triangle A'_j) < \varepsilon$$

para todo $0 \le j \le k-1$. Entonces como la propiedad de k-mezclante se cumple para conjuntos medibles en el semi-anillo $\mathcal{C}$, también se cumple para conjuntos que son unión disjunta de elementos de dicho semi-anillo, así se tiene que

$$\lim_{\substack{N_j\to\infty \\ 1\le j\le k-1}} \mu(A'_0 \cap T^{-N_1} A'_1 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})} A'_{k-1}) = \prod_{j=0}^{k-1} \mu(A'_j)$$

Luego, existe un $M \in \mathbb{N}$ tal que si $N_1 \ldots N_{k-1} > M$ entonces

$$\left| \mu\left(A'_0 \cap T^{-N_1} A'_1 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})} A'_{k-1}\right) - \prod_{j=0}^{k-1} \mu(A'_j) \right| < \varepsilon \tag{2.23}$$

Por otro lado, se tiene que

$$\begin{aligned}
&\left|\mu\left(A_0 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}\right)\right. \\
&\left.-\mu\left(A_0' \cap T^{-N_1}A_1' \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}'\right)\right| \\
&\qquad\leq \left|\mu\left(\left[A_0 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}\right]\right.\right. \\
&\qquad\qquad \left.\left.\triangle \left[A_0' \cap T^{-N_1}A_1' \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}'\right]\right)\right| \\
&\qquad\leq \sum_{j=0}^{k-1} \mu(A_j \triangle A_j') \\
&\qquad< k\varepsilon.
\end{aligned}$$

Luego

$$\begin{aligned}
&\left|\mu\left(A_0 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}\right) - \prod_{j=0}^{k-1}\mu(A_j)\right| \\
&\leq \left|\mu\left(A_0 \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}\right) - \mu\left(A_0' \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}'\right)\right| \\
&\quad + \left|\mu\left(A_0' \cap \ldots \cap T^{-(N_1+\ldots+N_{k-1})}A_{k-1}'\right) - \prod_{j=0}^{k-1}\mu(A_j')\right| + \left|\prod_{j=0}^{k-1}\mu(A_j') - \mu(A_0)\prod_{j=1}^{k-1}\mu(A_j')\right| \\
&\quad + \left|\mu(A_0)\prod_{j=1}^{k-1}\mu(A_j') - \mu(A_0)\mu(A_1)\prod_{j=2}^{k-1}\mu(A_j')\right| + \ldots + \left|\prod_{j=0}^{k-2}\mu(A_j)\mu(A_{k-1}') - \prod_{j=1}^{k-1}\mu(A_j)\right| \\
&< k\varepsilon + \varepsilon + k\varepsilon = (2k+1)\varepsilon.
\end{aligned}$$

□

A continuación se presentan los Lemas 2.6 y 2.7 que serán aplicados en el Capítulo 3.

Lema 2.6. *Si T es una transformación mezclante sobre el espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, A es un conjunto medible y $\varepsilon > 0$ entonces existe un $K \in \mathbb{N}$ tal que para todo $k \geq K$ y para toda sucesión $n_1, n_2, \ldots, n_k$ con k términos se tiene*

$$\mu\left(\left\{\omega \in \Omega : \left|\frac{1}{k}\#(A \cap \{T^{n_1}(\omega), T^{n_2}(\omega), \ldots, T^{n_k}(\omega)\}) - \mu(A)\right| \geq \varepsilon\right\}\right) < \varepsilon. \qquad (2.24)$$

Demostración. Para cada m, sea $X_m : \Omega \to \mathbb{R}$ definido por

$$X_m(\omega) = \begin{cases} 1, \text{ si } T^m(\omega) \in A \\ 0, \text{ en otro caso.} \end{cases}$$

Entonces

$$X_nX_m(\omega) = \begin{cases} 1, \text{ si } \omega \in T^{-n}(A) \cap T^{-m}(A) \\ 0, \text{ en otro caso.} \end{cases}$$

Dado $\delta > 0$, considere k_0 tal que si $\mid m - n \mid > k_0$ entonces

$$\mid E(X_n X_m) - (\mu(A))^2 \mid \;=\; \mid \mu(T^{-n}(A) \cap T^{-m}(A)) - (\mu(A))^2 \mid < \delta.$$

Sea k y $n_1 < n_2 < \ldots < n_k$.
Puesto que $n_{i+1} - n_i \geq 1$, si $\mid i - j \mid > k_0$ entonces $\mid n_i - n_j \mid > k_0$.
Sea

$$\begin{aligned} \Gamma(\omega) &= \frac{1}{k}\#(A \cap \{T^{n_1}(\omega), T^{n_2}(\omega), \ldots, T^{n_k}(\omega)\}) \\ &= \frac{1}{k}\sum_{i=1}^{k} X_{n_i}(\omega). \end{aligned}$$

Luego

$$\begin{aligned} E(\Gamma) &= \int_\Omega \Gamma d\mu \\ &= \frac{1}{k}\sum_{i=1}^{k}\int_\Omega X_{n_i} d\mu \\ &= \frac{1}{k}\sum_{i=1}^{k}\mu(T^{-n_i}(A)) \\ &= \mu(A), \end{aligned}$$

ahora bien

$$\begin{aligned} \sqrt{Var(\Gamma)} &= E(\Gamma^2) - (E(\Gamma))^2 \\ &= E\left(\frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k} X_{n_i}X_{n_j}\right) - (\mu(A))^2 \\ &= \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k} E(X_{n_i}X_{n_j}) - (\mu(A))^2 \\ &= \frac{1}{k^2}\left[\sum_{i=1}^{k}\sum_{j=1}^{k}\big(E(X_{n_i}X_{n_j}) - (\mu(A))^2\big)\right] \\ &= \frac{1}{k^2}\left[\sum_{\substack{i,j=1 \\ |i-j|>k_0}}^{k}\big(E(X_{n_i}X_{n_j}) - (\mu(A))^2\big) + \sum_{\substack{i,j=1 \\ |i-j|\leq k_0}}^{k}\big(E(X_{n_i}X_{n_j}) - (\mu(A))^2\big)\right] \\ &< \frac{1}{k^2}\left[k^2\delta + (2k_0+1)k\right] \\ &= \delta + \frac{2k_0+1}{k} \\ &\leq \delta + \frac{2k_0+1}{K} \end{aligned}$$

Luego por la desigualdad de Chebychev (Teorema 2.5) se tiene

$$\mu\left(\{\omega \in \Omega :\mid \Gamma(\omega) - E(\Gamma) \mid \geq \varepsilon\}\right) \leq \frac{Var(\Gamma)}{\varepsilon^2}$$

entonces

$$\begin{aligned} \mu\left(\{\omega \in \Omega : |\Gamma(\omega) - \mu(A)| \geq \varepsilon\}\right) &\leq \frac{Var(\Gamma)}{\varepsilon^2} \\ &< \frac{\left(\delta + \frac{2k_0+1}{K}\right)^2}{\varepsilon^2} \\ &< \varepsilon. \end{aligned}$$

Donde la última desigualdad se sigue de considerar δ suficientemente pequeño y K suficientemente grande. $\square$

Lema 2.7. *Si T es una transformación invertible k-mezclante sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$, dados $B_0, B_1 \ldots, B_{k-1}$ conjuntos medibles y $M_1, \ldots, M_{k-1}$ enteros positivos, se define $\Gamma_n[M_1, \ldots, M_{k-1}] : \Omega \to \mathbb{R}$, haciendo $\Gamma_n[M_1, \ldots, M_{k-1}](\omega)$ igual a*

$$\frac{1}{n}\#(B_0 \cap T^{-M_1}B_1 \cap \ldots \cap T^{-(M_1+\ldots+M_{k-1})}B_{k-1} \cap \{\omega, T\omega, \ldots, T^{n-1}\omega\}).$$

Si $\varepsilon > 0$ entonces existe un $M_0 \in \mathbb{N}$ tal que si $M_1 \ldots, M_{k-1} \in \mathbb{Z}$ son tales que $M_j > M_0$ y para $n > M_0$ se tiene

$$E\left(\left|\Gamma_n[M_1, \ldots, M_{k-1}] - \prod_{j=0}^{k-1} \mu(B_j)\right|\right) < \varepsilon. \tag{2.25}$$

Demostración. Como T es mezclante, $\underbrace{T \times T \times \ldots \times T}_{k\text{–veces}}$ es ergódico. Por el Teorema ergódico de Birkhoff (Teorema 2.8), para $(\omega_0, \ldots, \omega_{k-1}) \in \Omega^k$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \prod_{j=0}^{k-1} \mathbb{I}_{B_j}(T^i\omega_j) = \prod_{j=0}^{k-1} \mu(B_j)$$

entonces se escoge $N_1 \in \mathbb{N}$ tal que

$$E\left(\left|\tilde{\Gamma}_{N_1} - \prod_{j=1}^{k} \mu(B_j)\right|\right) < \frac{\varepsilon}{2},$$

donde

$$\tilde{\Gamma}_{N_1}(\omega_0, \omega_1, \ldots, \omega_{k-1}) = \frac{1}{N_1}\#\{0 \leq i \leq N_1 - 1 : T^i(\omega_j) \in B_j \text{ para todo } 0 \leq j \leq k-1\}.$$

Ya que T es k-mezclante se tiene que para $N_1, \ldots, N_{k-1}$ suficientemente grandes, la distribución conjunta de los eventos

$$\{T^i\omega_0 \in B_0, T^i\omega_1 \in B_1 \ldots, T^i\omega_{k-1} \in B_{k-1}\}_{0 \leq i \leq N_1-1}$$

está muy cerca de la distribución conjunta de los eventos

$$\{T^i\omega \in B_0, T^{i+N_1}\omega \in B_1, \ldots, T^{i+N_1+\ldots+N_{k-1}}\omega \in B_{k-1}\}_{0\leq i\leq N_1-1},$$

de modo que se puede elegir N_2 tal que para $M_1 \ldots M_{k-1} > N_2$

$$E\left(\left|\Gamma_{N_1}[M_1, \ldots, M_{k-1}] - \prod_{j=0}^{k-1}\mu(B_j)\right|\right) < \frac{\varepsilon}{2}.$$

Sea N_3 un entero positivo tal que $N_3 > \frac{2}{\varepsilon}$ y sea $M_0 = \text{máx}\{N_1, N_2, N_3\}$. Supongase $M_1, \ldots, M_{k-1} > M_0$ y $n > M_0$, lo suficiente, tal que $n = N_1L + R$ donde L y R son enteros tales que $0 \leq R \leq N_1$ y $L \geq N_3 \geq \frac{2}{\varepsilon}$, entonces

$$\begin{aligned}
\left|\Gamma_n(\omega) - \prod_{j=0}^{k-1}\mu(B_j)\right| &= \left|\frac{1}{n}\left(\sum_{i=1}^{L} N_1\Gamma_{N_1}(T^{(i-1)N_1}\omega) + R\Gamma(T^{N_1L}\omega)\right) - \prod_{j=0}^{k-1}\mu(B_j)\right| \\
&\leq \left|\frac{1}{LN_1}\sum_{i=1}^{L} N_1\Gamma_{N_1}(T^{(i-1)N_1}\omega) - \prod_{j=0}^{k-1}\mu(B_j)\right| + \frac{R}{n} \\
&\leq \frac{1}{L}\sum_{i=1}^{L}\left|\Gamma_{N_1}(T^{(i-1)N_1}\omega) - \prod_{j=0}^{k-1}\mu(B_j)\right| + \frac{1}{L}.
\end{aligned}$$

Así

$$\begin{aligned}
E\left(\left|\Gamma_n - \prod_{j=0}^{k-1}\mu(B_j)\right|\right) &\leq \frac{1}{L}\sum_{i=1}^{L} E\left(\left|\Gamma_{N_1} - \prod_{j=0}^{k-1}\mu(B_j)\right|\right) + \frac{1}{N_3} \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

$\square$

Capítulo 3

Transformaciones de rango-uno

Este capítulo está dedicado a las Transformaciones de Rango-uno que son el objeto de estudio de la presente tesis, ya que se desea probar que 2-mezclante implica k-mezclante para este tipo de transformaciones.

3.1. Definición de Transformación de rango-uno

Sea $\{r_i\}_{i\in\mathbb{N}} \subseteq \mathbb{Z}^+$, y sea $\{s_{i,j}\}_{i\in\mathbb{N},j\in\{1,2,\ldots,r_i\}} \subseteq \mathbb{Z}_0^+$. Un n-**bloque**, con $n = 0, 1, 2, \ldots$ se define de la siguiente manera; el 0-bloque es la palabra 'a' y para $n > 0$, un n-bloque es la palabra resultante del siguiente proceso: escriba el $(n-1)$-bloque, luego la letra 'b' $s_{n,1}$ veces, seguidamente anote el $(n-1)$-bloque y la letra 'b' $s_{n,2}$ veces, $\ldots$, concluyendo el listado con la letra 'b' s_{n,r_n} veces, es decir

$$B_n = B_{n-1} \underbrace{b\ldots b}_{s_{n,1}\text{veces}} B_{n-1} \underbrace{b\ldots b}_{s_{n,2}\text{veces}} B_{n-1} \ldots B_{n-1} \underbrace{b\ldots b}_{s_{n,r_n}\text{veces}} .$$

Sea $\Omega = \{a, b\}^{\mathbb{Z}}$ el espacio de sucesiones bi-infinitas de a's y b's, es decir

$$\Omega = \{\omega = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) : \omega_i \in \{a, b\} \text{ para todo } i \in \mathbb{Z}\}.$$

Un *cilindro finito* en Ω es un conjunto de la siguiente forma: Para cada número $i \in \mathbb{Z}$ y sucesión finita $\{x_0, x_1, \ldots, x_k : x_j \in \{a, b\}\}$ dada, definimos el cilindro

$$C(i, \{x_0, x_1, \ldots, x_k\}) = \{\omega \in \Omega : \omega_{i+j} = x_i \text{ para todo } 0 \leq j \leq k\}.$$

Sea $\mathcal{A}$ la σ-álgebra generada por el álgebra $\mathcal{C}$ de todos los cilindros finitos, es decir, $\mathcal{A}$ es la menor σ-álgebra que contiene a $\mathcal{C}$. Ahora, se va a definir una medida sobre la σ-álgebra $\mathcal{A}$; sea $B_n = \beta_1 \ldots \beta_{h_n}$ el n-bloque, dado $0 \leq k \leq h_n - 1$ definimos

$$\begin{array}{rcl} \mu_n : \{a, b\}^{k+1} & \to & [0, 1] \\ F & \mapsto & \mu_n(F) = \dfrac{1}{h_n - k} \displaystyle\sum_{j=1}^{h_n-k} \mathbb{I}_F(\beta_j \beta_{j+1} \ldots \beta_{j+k}). \end{array}$$

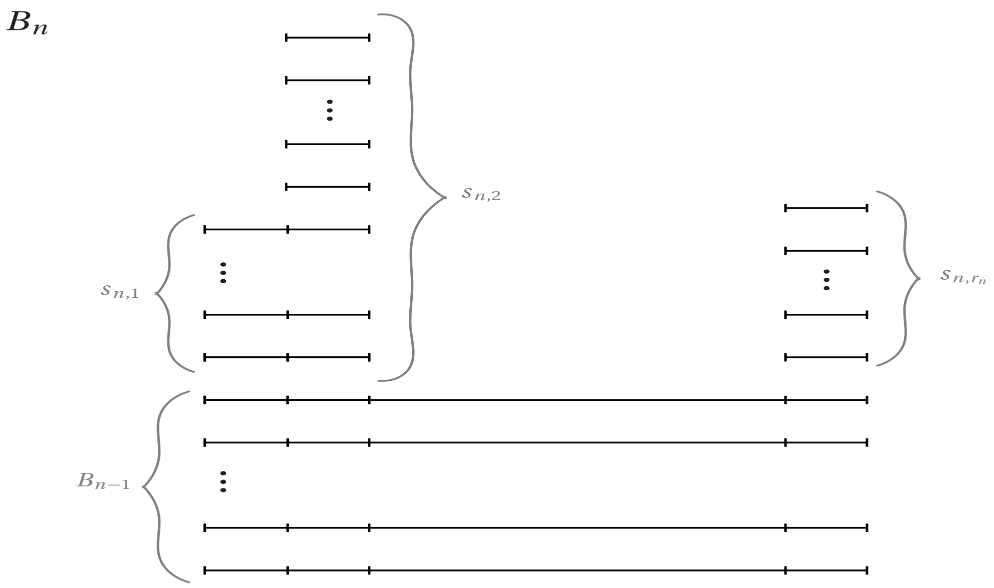

Figura 3.1: n-bloque

El número $\mu_n(F)$ denota la frecuencia con que aparece la palabra F en el n-bloque B_n, por ejemplo:

> *La frecuencia de la palabra 'aa' en 'aaab' es* $2/3$*, ya que 2 de 3 "subpalabras" de dos letras son aa.*

Ahora bien, definimos

$$\begin{array}{rcl} \mu : \mathcal{C} & \to & [0,1] \\ C(i, \{x_0, x_1, \ldots, x_k\}) & \mapsto & \mu(C(i, \{x_0, x_1, \ldots, x_k\})) = \lim\limits_{n\to\infty} \mu_n(x_0 x_1 \ldots x_k). \end{array}$$

Note que la medida del cilindro $C(i, \{x_0, x_1, \ldots, x_k\})$ sólo depende de la sucesión $\{x_0, x_1, \ldots, x_k\}$, más no depende del $i \in \mathbb{Z}$ tomado. Luego, μ se extiende de forma única a una medida de probabilidad sobre $\mathcal{A}$ [5, pág 19], que seguiremos denotando por μ, por lo tanto tenemos que $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad.

Definición 3.1. Una *Transformación de rango-uno* está dada por $(\Omega, \mathcal{A}, \mu, T)$, donde $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad, como antes, y $T : \Omega \longrightarrow \Omega$ es la transformación definida por:

$$(T(\omega))_i = \omega_{i+1} \text{ para todo } \omega \in \Omega.$$

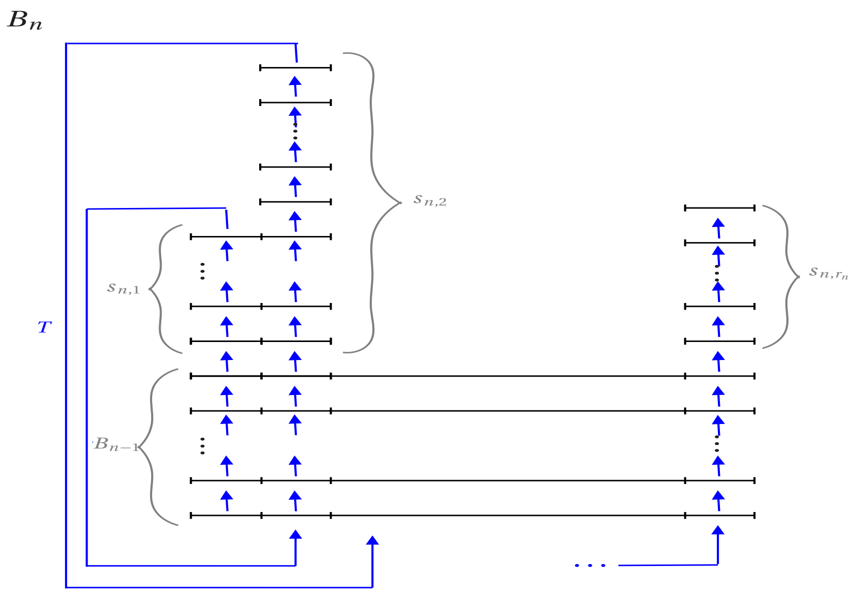

Figura 3.2: Transformación de rango-uno

Note que la transformación de rango-uno T está determinada por los valores $\{r_i\}$ y $\{s_{i,j}\}$ dados, además, puesto que $T(C(i,\{x_0,x_1,\ldots,x_k\})) = C(i+1,\{x_0,x_1,\ldots,x_k\})$, se sigue que T preserva medida.

Sea $B_\infty = x_0x_1x_2\ldots x_j\ldots$ la palabra infinita límite de la construcción inductiva de los n-bloques, es decir

$$B_\infty = \lim_{n\to\infty} B_n,$$

donde B_n es el n-bloque. Dado $k \in \mathbb{Z}$, denotemos

$$C(k,B_\infty) = \{\omega \in \Omega : \omega_{k+j} = x_j \text{ para todo } j \geq 0\}.$$

Dado $w \in \Omega$, diremos que su cero-entrada ω_0 es la *i-ésima letra del n-bloque* si $\omega \in C(-k,B_\infty)$ para algún $k \geq 0$ y x_k es la i-ésima entrada de alguno de los n-bloques B_n cuando expresamos

$$B_\infty = B_nb\ldots bB_nb\ldots\ldots bB_nb\ldots, \tag{3.1}$$

si por el contrario x_k es uno de los b's que no está en ningún B_n de (3.1) entonces decimos que ω_0 *es uno de los b's que no está en ningún n-bloque.*

Sea h_n el largo (o altura) del n-bloque, como palabra finita, y considere la regla de correspondencia $f_n : \Omega \longrightarrow \{1,2,\ldots,h_n,\text{'}error\text{'},@\}$ dada por:

$$f_n(\omega) = \begin{cases} i & \text{si } \omega_0 \text{ es la } i\text{-ésima letra del } n\text{-bloque,} \\ \text{'}error\text{'} & \text{si } \omega_0 \text{ es uno de los } b\text{'s que no está en ningún } n\text{-bloque.} \\ @ & \text{en otro caso.} \end{cases}$$

Lema 3.1. *Si ω es una palabra no periódica y existe $i \in \{1, \ldots, h_n\}$ tal que $f_n(\omega) = i$ entonces este i es único.*

Demostración. Primero note que se puede suponer sin pérdida de generalidad que cada n-bloque comienza y termina con una 'a'. Caso contrario, simplemente se remueve los comienzos y finales 'b' en cada n-bloque y se hácen intermedios 'b' del $(n+1)$-bloque, obteniendo así la misma palabra resultante.
Se probara el resultado por contradicción. Suponga que es posible ver a ω_0 en la i-ésima y j-ésima entrada de algún n-bloque, con $i < j$.
Si B denota al n-bloque, entonces deben haber dos extensiones doblemente infinitas de B's y b's, que serán llamadas formas 1 y 2 respectivamente, tales que cuando se escriba cada B en sus componentes a's y b's ambas formas se reducen a la misma palabra ω, cuya cero-entrada, o sea ω_0, es el i-ésimo término de B en su forma 1 y el j-ésimo término de B en su forma 2.
Sea $B_{1,0}$ el B en la forma 1 que contiene a ω_0, $B_{1,1}$,$B_{1,2}$,..., los B's siguientes a $B_{1,0}$ en la forma 1, en ese orden y $B_{1,-1}$,$B_{1,-2}$,..., los B's anteriores a $B_{1,0}$ pero en orden inverso. Análogamente la forma 2; ..., $B_{2,-2}$, $B_{2,-1}$, $B_{2,0}$, $B_{2,1}$, $B_{2,2}$,..., donde $B_{2,0}$ contiene a ω_0. Sea r el número de b's entre $B_{1,0}$ y $B_{1,1}$, y s el número de b's entre $B_{2,0}$ y $B_{2,1}$. Dado que ω_0 es la i-ésima entrada de $B_{1,0}$, entonces ω_{1-i} es la primera entrada de $B_{1,0}$ y por lo tanto una a; pero también ω_{1-i} es la $(j+1-i)$-ésima entrada de $B_{2,0}$, puesto que ω_{1-i} es una a, entonces la $(j+1-i)$-ésima entrada de B es también una a.
Note que ω_{1-i+h_n+s} es la $(j+1-i)$-ésima entrada de $B_{2,1}$ y por lo tanto una a. Sin embargo, como ω_{h_n-i} es la última entrada de $B_{1,0}$, para que ω_{h_n-i+1}, ω_{h_n-i+2}, ..., ω_{h_n-i+r} sean todos b's, se debe tener

$$1 - i + h_n + s > h_n - i + r,$$

de donde se tiene que

$$s \geq r. \tag{3.2}$$

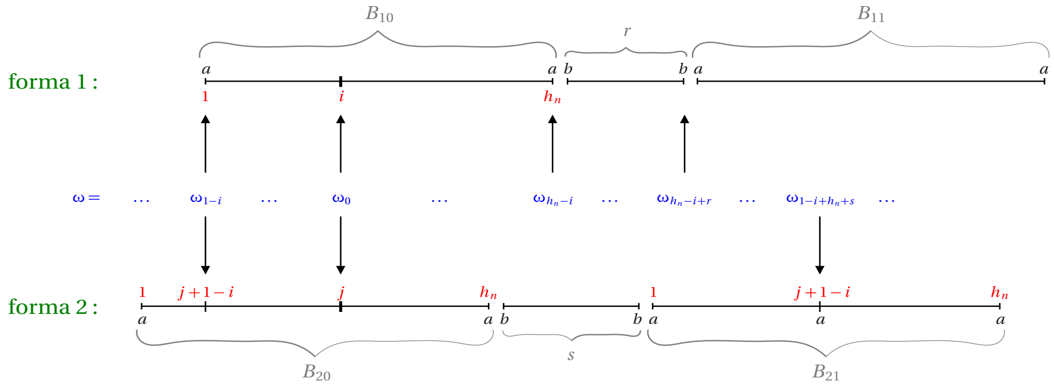

Figura 3.3:

Por otro lado note que ω_{2h_n+s-j} es la última entrada de $B_{2,1}$ y por lo tanto una a.

Afirmación. El término ω_{2h_n+s-j} también está en $B_{1,1}$.

En efecto, puesto que ω_{2h_n+s-j} es una a, debe estar en $B_{1,k}$, para algún $k \in \mathbb{Z}$. No puede estar en $B_{1,k}$ para $k \leq 0$ puesto que

$$2h_n + s - j > h_n - i,$$

y ω_{h_n-i} es la última letra de $B_{1,0}$.
Ahora suponga que ω_{2h_n+s-j} está en $B_{1,k}$ para algún $k \geq 2$, entonces ω_{h_n+s-j} a lo menos está en $B_{1,1}$ y como $\omega_{h_n-i+r+1}$ es la primera entrada de $B_{1,1}$ se deduce que

$$h_n + s - j \geq h_n - i + r + 1, \tag{3.3}$$

también se tiene

$$h_n - i + r + 1 \geq h_n - i + 1 > h_n - j + 1. \tag{3.4}$$

Como ω_{h_n-j} es la última entrada de $B_{2,0}$ entonces $\omega_{h_n-j+1}, \omega_{h_n-j+2}, \ldots, \omega_{h_n-j+s}$ son todos b's, en particular de (3.3) y (3.4) se sigue que $\omega_{h_n-i+r+1}$ es una b, lo que es imposible, ya que esta es la primera entrada de $B_{1,1}$ y por tanto una a. Lo que prueba la afirmación.

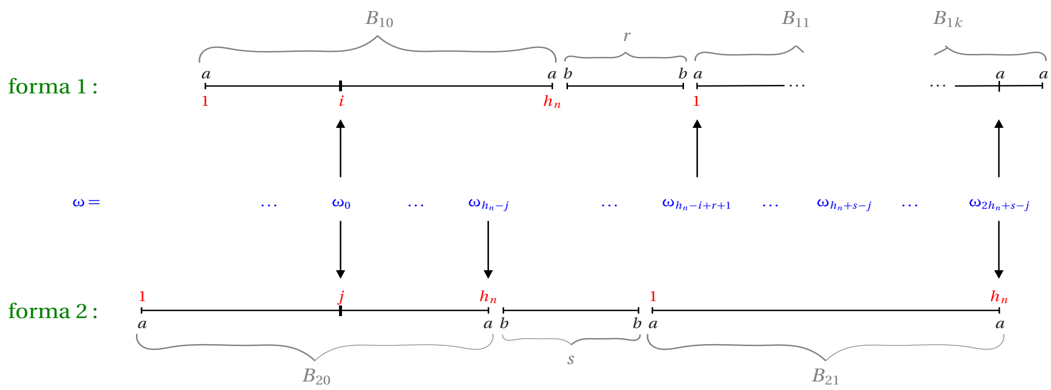

Figura 3.4:

Luego se tiene que ω_{2h_n+s-j} es la $(h_n+s-r+i-j)$-ésima entrada de $B_{1,1}$. Por lo tanto, desde que ω_{2h_n+s-j} es una a y $\omega_{h_n+s-r-j}$ es la $(h_n+s-r+i-j)$-ésima entrada de $B_{1,0}$, se deduce que $\omega_{h_n+s-r-j}$ es una a. Como ω_{h_n-j} es la última entrada de $B_{2,0}$, por definición de s, $\omega_{h_n-j+1}, \omega_{h_n-j+2}, \ldots, \omega_{h_n-j+s}$ son todos b's, de donde se tiene que

$$h_n+s-r-j<h_n-j+1,$$

y por lo tanto

$$s \leq r. \tag{3.5}$$

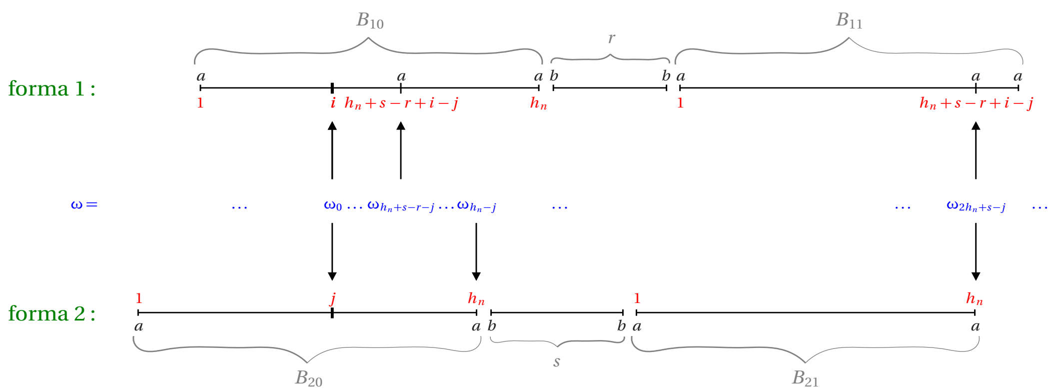

Figura 3.5:

De (3.2) y (3.5) se obtiene que $r=s$.

Se ha probado que la distancia entre $B_{2,0}$ y $B_{2,1}$, y la distancia entre $B_{1,0}$ y $B_{1,1}$, son la misma, a saber s.

Trasladando ω, se puede utilizar el mismo argumento para demostrar que la distancia entre $B_{1,k}$ y $B_{1,k+1}$ es la misma que la distancia entre $B_{2,k}$ y $B_{2,k+1}$, para todo $k \in \mathbb{Z}$, luego se tiene que

$$B_{10}b\dots bB_{11}b\dots B_{2-1}\underbrace{b\dots b}_{\hat{s}},$$

donde $\hat{s}$ es el número de b's entre B_{2-1} y B_{20}, es un periodo para ω, por lo tanto ω es periódica, lo que contradice la hipótesis. $\square$

Denotemos $I_{n,i} = \{\omega \in \Omega; f_n(\omega) = i\}$, para todo $i \in \{1, 2, \dots, h_n\}$, la sucesión finita $(I_{n,1}, \dots, I_{n,h_n})$ es llamada **n-torre** e $I_{n,i}$ el **i-ésimo nivel** de la n-torre y el conjunto $E_n = \{\omega \in \Omega; f_n(\omega) = error\}$ es llamado **conjunto error para la n-torre** o simplemente **conjunto n-error**.

De la definición se sigue que la (n-torre) $\subseteq$ (($n+1$)-torre) para todo $n \geq 0$, pues si ω_0 es una letra del n-bloque entonces es una letra del $(n+1)$-bloque que lo contiene, más aún cada nivel de la n-torre es unión de niveles de la $(n+1)$-torre.

La siguiente proposición nos dice que las n-torres son casi todo el espacio Ω.

Proposición 3.1. *La unión de las n-torres es un subconjunto de Ω de medida* 1.

Demostración. Sea

$$A_k = \{\vec{\alpha} = (\alpha_0, \dots, \alpha_k) \in \{a, b\}^{k+1} : \alpha_0 \dots \alpha_k \text{ no es subpalabra de ningún } n\text{-bloque}\}.$$

Primero se probara que

$$\left(\bigcup_{n\geq 0}(n\text{-torre})\right)^c \subseteq \bigcup_{k\in\mathbb{N}}\bigcup_{\vec{\alpha}\in A_k}\bigcup_{i=-k}^{0} C(i, \vec{\alpha}). \tag{3.6}$$

En efecto, si ω no está en ninguna n-torre entonces w_0 no es ninguna entrada del n-bloque para cualquier $n \in \mathbb{N}$, por lo tanto existe un cilindro $C(i, \alpha_0, \dots, \alpha_k)$ lo suficientemente pequeño que contiene a ω tal que $\alpha_0, \dots, \alpha_k$ no es subpalabra de ningún n-bloque, por lo tanto ω se encuentra en el conjunto del lado derecho de la ecuación (3.6).
Ahora bien, de la definición de A_k se sigue que $\mu(C(i, \vec{\alpha})) = 0$ para todo $\vec{\alpha} \in A_k$ y todo $i \in \mathbb{Z}$, lo que concluye la prueba.

$\square$

De ahora en adelante cuando se diga $\omega \in \Omega$ solo se estará considerando ω en la unión de las n-torres, es decir

$$\omega \in \bigcup_{n\geq 0}(n\text{-torre}),$$

esto sera suficiente, ya que por la proposición anterior se tiene que $\mu(\cup_{n\geq 0}(n\text{-torre})) = 1$.

Teorema 3.1. *Si la transformación de rango-uno T es mezclante entonces las palabras ω's en Ω son no-periódicas.*

Demostración. Suponga que existe $\omega \in \Omega$ que es periódica

$$\omega = \ldots J_0 J_0 J_0 \ldots$$

donde $J_0 = x_0 x_1 \ldots x_k$, entonces existe un $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$, el n-bloque, que se denota B_n, es de la forma

$$B_n = J_0 J_0 J_0 \ldots J_0.$$

Si $J_i = x_i \ldots x_k x_0 \ldots x_{i-1}$ para $1 \leq i \leq k$, evidentemente $TC(0, J_{i-1}) \cap C(0, J_i)$ es no vacío, así también $TC(0, J_k) \cap C(0, J_0)$ es no vacío.

Afirmación. Para cada $i \in \{1, \ldots, k\}$ se tiene que $TC(0, J_{i-1}) \subseteq C(0, J_i)$ y $TC(0, J_k) \subseteq C(0, J_0)$.

En efecto, suponga que existe $\hat{\omega} \in C(0, J_{i-1})$ tal que $T(\hat{\omega}) \notin C(0, J_i)$ entonces $\hat{\omega} \notin C(0, J_{i-1} J_{i-1})$ y por lo tanto $\hat{\omega} \notin B_n$ para $n \geq n_0$, lo que es una contradicción.

De la afirmación se sigue que

$$0 < \mu(C(0, J_0)) = \mu(C(0, J_1)) = \ldots = \mu(C(0, J_k)) < 1.$$

Pero T es mezclante entonces

$$\mu(C(0, J_0))^2 = \lim_{N \to \infty} \mu(T^{-N(k+1)} C(0, J_0) \cap C(0, J_0)) = \lim_{N \to \infty} \mu(C(0, J_0)) = \mu(C(0, J_0))$$

Lo que es una contradicción. $\square$

Como la hipótesis general del trabajo es que T sea mezclante, de ahora en adelante, se va a considerar Ω, osea la únion de las n-torres, solo de palabras no periódicas, en tal caso, note que, para $n \in \mathbb{N}$ fijo los niveles de la n-torre son disjuntos y el espacio Ω es unión disjunta de los niveles de la n-torre y el conjunto n-error E_n, de esto y de la Proposición 3.1 se sigue que

$$\lim_{n \to \infty} \mu(E_n) = 0.$$

Observación 3.1. Puesto que la σ-álgebra $\mathcal{A}$, es generada por los n-bloques y cada n-bloque es unión disjunta de niveles de la $(n+1)$-torre entonces se tiene que el conjunto $\mathcal{C}$ formado por todos los niveles de todas las torres, es decir

$$\mathcal{C} = \{I_{n,i} : n \geq 0 \text{ y } 1 \leq i \leq h_n\},$$

es un semi-anillo para la σ-álgebra $\mathcal{A}$.

También se tiene que

$$T(I_{n,i}) = I_{n,i+1} \text{ para todo } 1 \leq i < h_n,$$

además, si $\omega \in \Omega$ es tal que $f_n(\omega) \neq error$ y $1 - f_n(\omega) \leq m \leq h_n - f_n(\omega)$ entonces

$$f_n(T^m\omega) = f_n(\omega) + m. \tag{3.7}$$

La siguiente proposición muestra que éstas transformaciones de rango-uno son ergódicas.

Proposición 3.2. *Toda transformación de rango-uno es ergódica.*

Demostración. Sea A un conjunto medible tal que $T(A) = A$ y $\mu(A) > 0$, el objetivo es probar que $\mu(A) = 1$. En efecto, dado $\varepsilon > 0$, existe $I_{n,k}$ en el semi-anillo $\mathcal{C}$ tal que

$$\mu(A \cap I_{n,k}) > (1 - \varepsilon)\mu(I_{n,k}), \tag{3.8}$$

donde $n \in \mathbb{N}$ y $1 \leq k \leq h_n$ y se puede suponer que $\mu(E_n) < \varepsilon$, pues de no ser así recuerde que $I_{n,k}$ es unión disjunta de niveles de cualquier m-torre con $m \geq n$ y para uno de estos niveles también se cumple (3.8), entonces existen $p, q \in \mathbb{Z}$ tales que $T^i(I_{n,k})$ son disjuntos dos a dos para $p \leq i \leq q$, $q - p + 1 = h_n$ y $\cup_{i=p}^{q} T^i(I_{n,k})$ es la n-torre, que se denotará por B. Así

$$\begin{aligned}
\mu(A) &= \mu(\bigcup_{i=p}^{q} A \cap T^i(I_{n,k})) + \mu(A \cap (E_n)) \\
&\geq \sum_{i=p}^{q} \mu(A \cap T^i(I_{n,k})) \\
&= \sum_{i=p}^{q} \mu(T^i(A \cap I_{n,k})) \quad \text{esto de suponer que } A \text{ es invariante} \\
&= \sum_{i=p}^{q} \mu(A \cap I_{n,k}) \\
&> (q - p + 1)(1 - \varepsilon)\mu(I_{n,k}) \\
&= h_n(1 - \varepsilon)\mu(I_{n,k}) \\
&= \mu(B)(1 - \varepsilon) \\
&> (1 - \varepsilon)^2.
\end{aligned}$$

Puesto que $\varepsilon > 0$ se escogió arbitrario, se sigue que $\mu(A) = 1$. Por lo tanto T es ergódica.

□

Sea $\omega \in \Omega$ tal que $f_n(\omega) \neq error$, definimos $s_n(\omega)$ como el número de $b's$ entre el n-bloque que contiene a ω_0 y el siguiente n-bloque (ver figura 3.6), es decir, si $f_n(\omega) = i$ entonces

$$s_n(\omega) = \text{mín}\{j \in \mathbb{N} : f_n(T^{j+1+h_n-i}\omega) = 1\}.$$

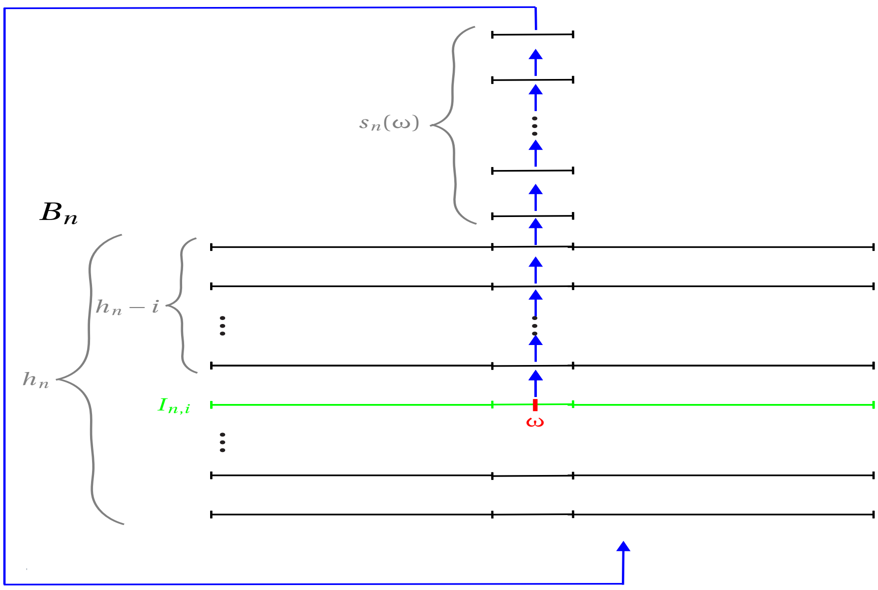

Figura 3.6: $s_n(\omega)$

Observación 3.2. Así como se ha definido la función $s_n : \{\omega \in \Omega : f_n(\omega) \neq error\} \to \mathbb{Z}_0^+$ como el número de b's entre el n-bloque que contiene a ω_0 y el siguiente n-bloque, también se puede definir la función $r_n : \{\omega \in \Omega : f_n(\omega) \neq error\} \to \mathbb{Z}_0^+$ como el número de b's entre el n-bloque que contiene a ω_0 y el n-bloque anterior, es decir, si $f_n(\omega) = i$ entonces

$$r_n(\omega) = \text{mín}\{j \in \mathbb{N} : f_n(T^{-(i+j)}(\omega)) = h_n\}.$$

Sea $\omega \in \Omega$ y $L \in \mathbb{Z}$ con $L > 0$, considere el conjunto

$$\Theta = \{i : 0 \leq i \leq L-1 \text{ y } f_n(T^i\omega) \neq error\}.$$

Note que Θ depende de ω y L, si $f_n(\omega) = f_n(T^L\omega) = 1$ entonces

$$\sum_{i\in\Theta} \left(\frac{s_n(T^i\omega)}{h_n}\right) = L - \#\Theta.$$

Así, el valor promedio de $s_n(T^i\omega)$ para todo $i \in \Theta$ es

$$\frac{1}{\#\Theta}\sum_{i\in\Theta} s_n(T^i\omega) = h_n\left(\frac{L-\#\Theta}{\#\Theta}\right). \tag{3.9}$$

Si L es escogido arbitrariamente grande de modo que $f_n(T^L\omega) = 1$, puesto que T es una transformación de rango-uno, de la Proposción 3.2 se sigue que es ergódica entonces por el Teorema ergódico de Birkhoff (Teorema 2.8) se tiene

$$\lim_{L\to\infty} \frac{1}{L}\sum_{i=0}^{L-1} \mathbb{I}_{E_n}(T^i\omega) = \mu(E_n),$$

luego

$$\lim_{L\to\infty} \frac{L - \#\Theta}{L} = \mu(E_n) \tag{3.10}$$

Así mismo

$$\lim_{L\to\infty} \frac{1}{L}\sum_{i=0}^{L-1} \mathbb{I}_{E_n^c}(T^i\omega) = 1 - \mu(E_n),$$

entonces

$$\lim_{L\to\infty} \frac{\#\Theta}{L} = 1 - \mu(E_n). \tag{3.11}$$

De (3.10) y (3.11) se sigue que

$$\lim_{L\to\infty} \frac{L - \#\Theta}{\#\Theta} = \frac{\mu(E_n)}{1 - \mu(E_n)}. \tag{3.12}$$

Luego de (3.9) y (3.12) se tiene

$$\begin{aligned} E(s_n|\{\omega \in \Omega : f_n(\omega) \neq error\}) &= \lim_{L\to\infty} \frac{1}{\#\Theta}\sum_{i\in\Theta} s_n(T^i\omega) \\ &= h_n\left(\frac{\mu(E_n)}{1 - \mu(E_n)}\right). \end{aligned} \tag{3.13}$$

entonces

$$\lim_{n\to\infty} \frac{1}{h_n} E(s_n|\{\omega \in \Omega : f_n(\omega) \neq error\}) = 0, \tag{3.14}$$

esto último se sigue desde que $\mu(E_n) \to 0$ cuando $n \to \infty$.

3.2. Las n-generaciones, definición y propiedades

Ahora se va a definir ciertos intervalos de números enteros, que serán llamados n-generaciones; para n fijo, las n-generaciones serán conjuntos disjuntos y cubrirán casi todos los enteros, además cada n-generación será subconjunto de una $(n+1)$-generación. El resultado a resaltar en esta sección es que para n fijo, cada dos n-generaciones sucesivas almenos una de ellas cumple cierta propiedad llamada “buena”.

Dados $n, N_1, \ldots, N_k \in \mathbb{N}$, denotaremos $l_j[n, N_1, \ldots, N_k] : \Omega \to \{1, \ldots h_n, error\}$ para $0 \leq j \leq k$ a las funciones definidas como sigue

$$\begin{aligned}
l_0[n, N_1, \ldots, N_k](\omega) &= f_n(\omega), \\
l_1[n, N_1, \ldots, N_k](\omega) &= f_n \circ T^{N_1}(\omega), \\
\vdots &= \vdots \\
l_k[n, N_1, \ldots, N_k](\omega) &= f_n \circ T^{N_1+\ldots+N_k}(\omega).
\end{aligned}$$

Escribiremos simplemente $l_0, l_1, \ldots, l_k$, si $n, N_1, \ldots, N_k$ y $\omega \in \Omega$ estén fijados.

Teorema 3.2. *Sean $\omega \in \Omega$, $A_0, A_1, \ldots, A_k$ unión de niveles de una m_0-torre y $n \geq m_0$ entonces son equivalentes*

(i) $I_{n,l_0} \subseteq A_0, I_{n,l_1} \subseteq A_1, \ldots, I_{n,l_{k-1}} \subseteq A_{k-1}$ *y* $I_{n,l_k} \subseteq A_k$.

(ii) $\omega \in A_0 \cap T^{-N_1} A_1 \cap \ldots \cap T^{-(N_1+\ldots+N_k)} A_k$.

donde $l_j = l_j[n, N_1, \ldots, N_k](\omega)$ *para todo* $j \in \{0, 1, \ldots, k\}$.

Demostración. Por definición se tiene

$$f_n(\omega) = l_0,\ f_n\left(T^{N_1}\omega\right) = l_1, \ldots, f_n\left(T^{\sum_{j=1}^{k-1} N_j}\omega\right) = l_{k-1} \text{ y } f_n\left(T^{\sum_{j=1}^{k} N_j}\omega\right) = l_k,$$

entonces

$$\omega \in I_{n,l_0} \subseteq A_0,\ T^{N_1}\omega \in I_{n,l_1} \subseteq A_1, \ldots \text{ y } T^{\sum_{j=1}^{k} N_j}\omega \in I_{n,l_k} \subseteq A_k.$$

Luego

$$\omega \in A_0 \cap T^{-N_1} A_1 \cap \ldots \cap T^{-\sum_{j=1}^{k} N_j} A_k.$$

Recíprocamente, se tiene

$$\omega \in A_0 \cap I_{n,l_0},\ T^{N_1}\omega \in A_1 \cap I_{n,l_1}, \ldots \text{ y } T^{\sum_{j=1}^{k} N_j}\omega \in A_k \cap I_{n,l_k}$$

entonces $A_j \cap I_{n,l_j} \neq \emptyset$ y como $n \geq m_0$ se sigue que $I_{n,l_j} \subseteq A_j$ para todo $j \in \{0, 1, \ldots, k\}$.

□

Observación 3.3. Se sigue de la Observación 3.1 que es suficiente considerar desde ahora a los conjuntos medibles $A_0, A_1, \ldots, A_k$ como en el Teorema 3.2, es decir, unión de niveles de una m_0-torre fija y n siempre será mayor o igual a este m_0.

Diremos que el número entero i es *inicio de una generación* si el conjunto $\{l_0(T^i\omega), l_1(T^i\omega), \ldots l_k(T^i\omega)\}$ contiene 1 pero no contiene *error* y diremos que es *fin de una generación* si el conjunto $\{l_0(T^i\omega), l_1(T^i\omega), \ldots l_k(T^i\omega)\}$ contiene h_n pero no contiene *error*.

Definición 3.2. Dados c y d números enteros, el conjunto $[c, d] \cap \mathbb{Z}$ es llamado *n-generación* o simplemente *generación* si c es el inicio de una generación y d es el menor entero tal que d es el fin de una generación y $d \geq c$.

Note que la definición de n-generación depende de los enteros positivos n, $N_1, \ldots, N_k$ y de $\omega \in \Omega$.

Proposición 3.3. *Para n fijo, las n-generaciones son todas disjuntas.*

Demostración. Sean $c_1 < c_2$ inicios de generaciones, se debe probar que existe un $d \in \mathbb{Z}$, fin de generación, tal que $c_1 \leq d < c_2$.
Como c_1 es inicio de generación entonces $1 \in \{l_0(T^{c_1}\omega), l_1(T^{c_1}\omega), \ldots, l_k(T^{c_1}\omega)\}$, luego tenemos dos casos

i) Si $h_n \in \{l_0(T^{c_1}\omega), l_1(T^{c_1}\omega), \ldots, l_k(T^{c_1}\omega)\}$ entonces $d = c_1$.

ii) Si $h_n \notin \{l_0(T^{c_1}\omega), l_1(T^{c_1}\omega), \ldots, l_k(T^{c_1}\omega)\}$, entonces se procede como sigue

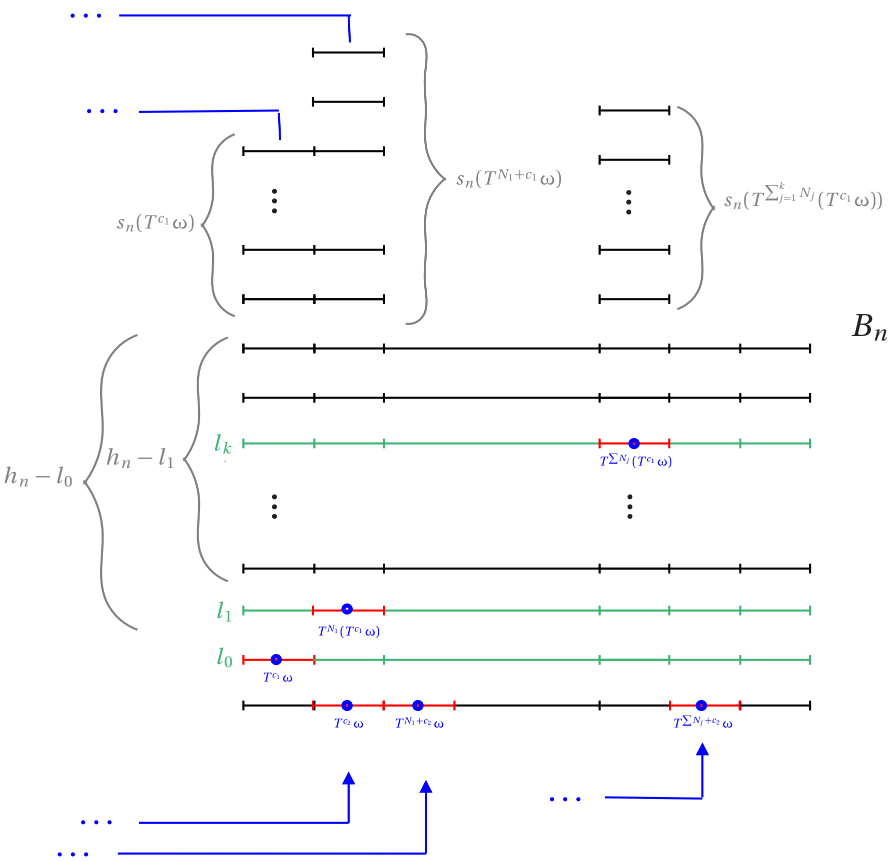

Figura 3.7: caso (ii) para $\sigma = Id$.

Sea σ una permutación de $\{0,1,\ldots,k\}$ tal que

$$l_{\sigma(0)}(T^{c_1}\omega) \leq l_{\sigma(1)}(T^{c_1}\omega) \leq \ldots \leq l_{\sigma(k)}(T^{c_1}\omega),$$

si $d = c_1 + h_n - l_{\sigma(k)}$ entonces este d es fin de generación, ya que

$$h_n = l_{\sigma(k)}\left(T^{c_1+h_n-l_{\sigma(k)}}\omega\right) \in \{l_0\left(T^{c_1+h_n-l_{\sigma(k)}}\omega\right), l_1\left(T^{c_1+h_n-l_{\sigma(k)}}\omega\right), \ldots, \\ l_k\left(T^{c_1+h_n-l_{\sigma(k)}}\omega\right)\},$$

y

$$1 < l_{\sigma(0)}(T^{c_1+h_n-l_{\sigma(k)}}\omega) \leq l_{\sigma(1)}(T^{c_1+h_n-l_{\sigma(k)}}\omega) \leq \ldots \leq l_{\sigma(k)}(T^{c_1+h_n-l_{\sigma(k)}}\omega) = h_n,$$

es decir, no contiene '*error*'. Por otro lado, si c_2 es inicio de generación entonces

$$c_2 \in \{c_1 + h_n - l_0 + s_n(T^{c_1}\omega) + 1, c_1 + h_n - l_1 + s_n\left(T^{N_1}(T^{c_1}\omega)\right) + 1, \\ \ldots, c_1 + h_n - l_k + s_n\left(T^{\sum_{j=1}^{k} N_j}(T^{c_1}\omega)\right) + 1\}.$$

En cualquier caso, $d < c_2$.

$\square$

Una n-generación es un intervalo de enteros, es decir un conjunto de la forma $[c,d]\cap\mathbb{Z}$ que de ahora en adelante, por comodidad, denotaremos simplemente $[c,d]$.

Observación 3.4. Un número entero i está en la generación $[c,d]$ donde $c = i + 1 - \text{mín}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}$ y $d = i + h_n - \text{máx}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}$.

En efecto, de la Ecuación 3.7 se sigue que para todo $0 \leq j \leq k$ se tiene que

$$\begin{aligned} l_j(T^c\omega) &= l_j(T^{i+1-\text{mín}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}}\omega) \\ &= l_j(T^i\omega) + 1 - \text{mín}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}, \end{aligned}$$

entonces $\{l_0(T^c\omega),\ldots,l_k(T^c\omega)\}$ no contiene *error* y además

$$\text{mín}\{l_0(T^c\omega),\ldots,l_k(T^c\omega)\} = 1.$$

Por lo tanto $c = i + 1 - \text{mín}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}$ es inicio de generación. De la misma manera se sigue que $d = i + h_n - \text{máx}\{l_0(T^i\omega),\ldots,l_k(T^i\omega)\}$ es fin de generación. Además de la Ecuación 3.7 también se sigue que si i, i' están contenidos en la misma generación entonces

$$l_j(T^{i'}\omega) = l_j(T^i\omega) + i' - i \text{ para } j \in \{0,1,\ldots,k\}.$$

De esto y el Teorema 3.2 se sigue que el conocimiento de $l_0(T^i\omega), l_1(T^i\omega), \ldots, l_k(T^i\omega)$, no solo determina a la generación que contiene a i sino que también determina qué valores i' en dicha generación tienen la propiedad

$$T^{i'}(\omega) \in A_0 \cap T^{-N_1}(A_1) \cap, \ldots, \cap T^{-\sum_{j=1}^{k} N_j}(A_k).$$

Llamaremos **generación actual** a aquella que contiene al 0, a la generación anterior a la actual, será llamada **generación predecesora** y a la generación siguiente a la actual, será llamada **generación sucesora**.

Reemplazando i por 0 en la observación anterior, se sigue que el conocimiento de $l_0 = l_0(\omega)$, $l_1 = l_1(\omega), \ldots, l_k = l_k(\omega)$ provee suficiente información para completar la generación actual. ¿Podremos saber algo de la generación sucesora?.

Sean $n, N_1, \ldots, N_k \in \mathbb{N}$, denotaremos $s_{n,j}[N_1, \ldots, N_k] : \Omega \to \{1, \ldots h_n, error\}$ con $0 \leq j \leq k$ a las funciones definidas como sigue

$$\begin{aligned}
s_{n,0}[N_1, \ldots, N_k](\omega) &= s_n(\omega), \\
s_{n,1}[N_1, \ldots, N_k](\omega) &= s_n(T^{N_1}\omega), \\
&\vdots \\
s_{n,k}[N_1, \ldots, N_k](\omega) &= s_n(T^{\sum_{j=1}^{k} N_j}\omega).
\end{aligned}$$

Escribiremos simplemente $s_0, s_1, \ldots, s_k$ si $n, N_1, \ldots, N_k \in \mathbb{N}$ y $\omega \in \Omega$ estén fijados.

Sea σ una permutación de $\{0, 1, \ldots, k\}$ que ordena de forma creciente al conjunto $\{l_0, l_1, \ldots, l_k\}$, es decir $l_{\sigma(0)} \leq l_{\sigma(1)} \leq \ldots \leq l_{\sigma(k-1)} \leq l_{\sigma(k)}$. Si definimos $\widehat{s} = s_{\sigma(k)}$, entonces $\widehat{s}$ estará bien definido siempre que el máximo de $\{l_0, l_1, \ldots, l_k\}$ sea único. Diremos que $(l_0, l_1, \ldots, l_k, \widehat{s})$ es **aceptable** si $\{l_0, l_1, \ldots, l_k\}$ tiene un único máximo y $\widehat{s} < l_{\sigma(k)} - l_{\sigma(k-1)}$, donde σ es la permutación anterior (ver Figura 3.8).

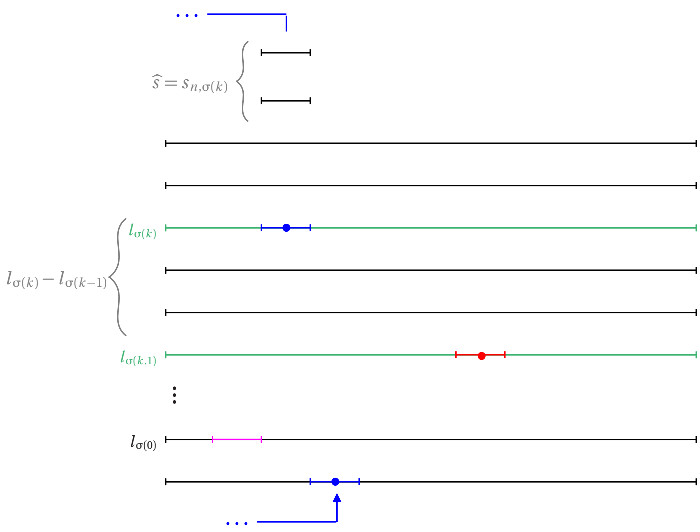

Figura 3.8: aceptable.

Observación 3.5. Si $(l_0, l_1, \ldots, l_k, \widehat{s})$ es aceptable y σ es la permutación de $\{0, 1, \ldots, k\}$ tal que $l_{\sigma(0)} \leq l_{\sigma(1)} \leq \ldots \leq l_{\sigma(k-1)} < l_{\sigma(k)}$ entonces $[c, d]$ es la generación sucesora si $c = h_n + 1 + \widehat{s} - l_{\sigma(k)}$ y $d = h_n - l_{\sigma(k-1)}$ (ver Figura 3.9).

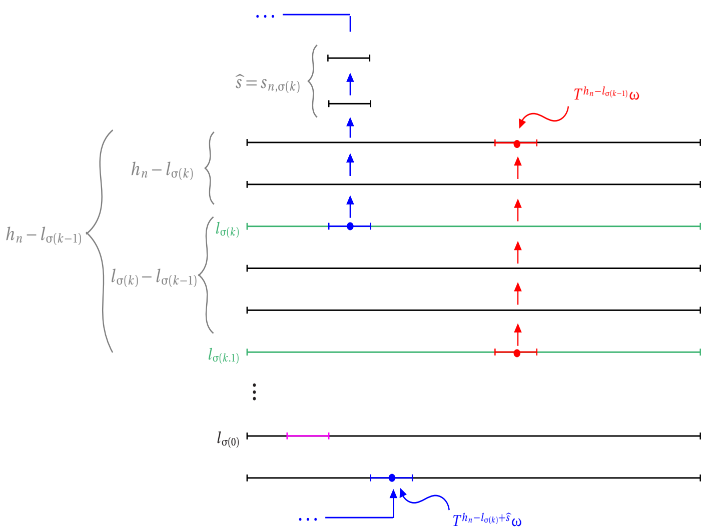

Figura 3.9:

Además para cualquier $i \in [c,d]$,

$$\begin{aligned} l_j(T^i\omega) &= l_j(\omega) + i \text{ para todo } j \in \{\sigma(0), \sigma(1), \ldots, \sigma(k-1)\} \\ \text{y } l_{\sigma(k)}(T^i\omega) &= l_{\sigma(k)}(\omega) + i - h_n - \widehat{s} \end{aligned}$$

De esto y el Teorema 3.2 podemos determinar que valores i en la generación sucesora $[c,d]$ satisfacen

$$T^i(\omega) \in A_0 \cap T^{-N_1}(A_1) \cap \ldots \cap T^{-\sum_{j=1}^k N_j}(A_k).$$

A continuación definimos cuando una generación tiene la propiedad de ser buena. Dados $N_1, \ldots, N_k \in \mathbb{N}$, denotaremos $I[N_1, \ldots, N_k] : \Omega \to \{0,1\}$ a la función definida por

$$I[N_1, \ldots, N_k](\omega) = \mathbb{I}_{A_0 \cap T^{-N_1} A_1 \cap \ldots \cap T^{-\sum_{j=1}^k N_j} A_k}(\omega).$$

Escribiremos simplemente I cuando $N_1, \ldots, N_k$ estén fijados.

Definición 3.3. Dado $\varepsilon > 0$, una generación $[c,d]$ es llamada ε-**buena** si

$$\left| \frac{1}{d-c+1} \sum_{i=c}^{d} I(T^i\omega) - \prod_{j=0}^{k} \mu(A) \right| < \varepsilon.$$

caso contrario será llamada ε-**mala**.

De las Observaciones 3.4 y 3.5 se sigue que, si $(l_0, l_1, \ldots, l_k, \widehat{s})$ es aceptable entonces se puede determinar si la generación actual y la generación sucesora son ε-buena o ε-mala.

Ahora el trabajo va a consistir en reducir el estudio de $l_0, l_1, \ldots, l_k$, que determina si T es $(k+1)$-mezclante, a simplemente estudiar $l_{\sigma(0)}, \ldots, l_{\sigma(k-1)}, \widehat{s}$ que determina si T es k-mezclante. Primero denotemos los vectores $(l_0, l_1, \ldots, l_k)$ por $\overrightarrow{l}$ y $(l_0, \ldots, l_k, s_0, \ldots, s_k)$ por $\left(\overrightarrow{l}, \overrightarrow{s}\right)$; tenga en cuenta que estos valores dependen de $n, N_1, \ldots, N_k$ y ω.

Dado $\delta > 0$, diremos que $\overrightarrow{l}$ es δ-**razonable** si cumple las siguientes condiciones

(i) El término *error* $\notin \{l_0, l_1, \ldots, l_k\}$,

(ii) $(1 - \sqrt{\delta})h_n \geq \text{máx}\{|l_j - l_{j'}| : j, j' \in \{0, 1, \ldots, k\} \text{ y } j \neq j'\}$ y

(iii) $\sqrt{\delta} h_n \leq \text{mín}\{|l_j - l_{j'}| : j, j' \in \{0, 1, \ldots, k\} \text{ y } j \neq j'\}$.

Analogamente, $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ es δ-**razonable** cuando

(i) El vector $\overrightarrow{l}$ es δ-**razonable** y

(ii) $\delta h_n > \text{máx}\{s_0, , s_1 \ldots, s_k\}$.

Si $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ (respectivamente $\overrightarrow{l}$) no es δ-razonable, diremos que es δ-**irrrazonable**.

Note que si $\overrightarrow{l}$ es δ-razonable entonces los l_j's son todos distintos, además si $0 < \delta \leq 1$ y $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ es δ-razonable, se tiene

$$\begin{aligned} \widehat{s} &\leq \max\{s_0, s_1, \ldots, s_k\} \\ &< \delta h_n \\ &\leq \sqrt{\delta} h_n \\ &\leq \min\{|l_j - l_{j'}| : j, j' \in \{0, 1, \ldots, k\} \text{ y } j \neq j'\} \\ &\leq l_{\sigma(k)} - l_{\sigma(k-1)}, \end{aligned}$$

entonces $(l_0, l_1, \ldots, l_k, \widehat{s})$ es aceptable.

El siguiente lema nos dice que el conjunto de los ω's para los cuales $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ es δ-irrazonable es pequeño en medida.

Lema 3.2. *Si la transformación de rango-uno es mezclante y $0 < \delta < 1$, entonces para cada n suficientemente pequeño de modo que*

$$\mu(E_n) < \frac{\delta^2}{1 + \delta^2},$$

existe $\hat{N}_n$ tal que

$$\mu\left(\left\{\omega \in \Omega : \left(\overrightarrow{l}, \overrightarrow{s}\right) \text{ es } \delta\text{-irrazonable}\right\}\right) < (k+1)\delta^2 + (k+1)\delta + 3k(k+1)\sqrt{\delta}$$

siempre que $N_1, \ldots, N_k > \hat{N}_n$.

Demostración. Primero, note que para todo $0 \leq j \leq k$ se tiene

$$\begin{aligned} \mu(\{\omega : l_j(\omega) = error\}) &= \mu(\{\omega : f_n(T^{N_1 + \ldots + N_j}\omega) = error\}) \\ &= \mu(\{\omega : f_n(\omega) = error\}) \\ &= \mu(E_n), \end{aligned}$$

entonces

$$\mu(\{\omega : error \in \{l_0, l_1, \ldots, l_k\}\}) \leq (k+1)\mu(E_n) < (k+1)\delta^2. \tag{3.15}$$

Por otro lado, condicionado sobre el evento $\{\omega : error \notin \{l_0, l_1, \ldots, l_k\}\}$, cada l_j toma valores de 1 a h_n con distribución uniforme.

Afirmación. Si los l_j's fuesen independientes dos a dos, entonces

$$\mu\left(\left\{\omega : |l_j - l_{j'}| < \sqrt{\delta} h_n\right\}\right) < 2\sqrt{\delta} \quad \text{y} \quad \mu\left(\left\{\omega : |l_j - l_{j'}| > \left(1 - \sqrt{\delta}\right) h_n\right\}\right) < 2\sqrt{\delta}$$

para $j, j' \in \{0, 1, \ldots, k\}, j \neq j'$.
En efecto, si $\sqrt{\delta} h_n \in \mathbb{Z}$,

$$\begin{aligned} \mu\left(\left\{\omega : |l_j - l_{j'}| < \sqrt{\delta} h_n\right\}\right) &\leq \frac{\left(2h_n\sqrt{\delta} - 2\right) h_n}{h_n^2} \\ &< 2\sqrt{\delta}, \end{aligned}$$

y si $\sqrt{\delta} h_n \notin \mathbb{Z}$,

$$\begin{aligned} \mu\left(\left\{\omega : |l_j - l_{j'}| < \sqrt{\delta} h_n\right\}\right) &\leq \frac{2\left[\!\left[\sqrt{\delta} h_n\right]\!\right] h_n}{h_n^2} \\ &< 2\sqrt{\delta}. \end{aligned}$$

Análogamente, si $(1 - \sqrt{\delta}) h_n \in \mathbb{Z}$

$$\begin{aligned} \mu\left(\left\{\omega : |l_j - l_{j'}| > \left(1 - \sqrt{\delta}\right) h_n\right\}\right) &= \frac{\left(h_n - 1 - \left(1 - \sqrt{\delta}\right) h_n\right) h_n}{h_n^2} \\ &= \sqrt{\delta} - \frac{1}{h_n}, \end{aligned}$$

y si $(1 - \sqrt{\delta}) h_n \notin \mathbb{Z}$,

$$\begin{aligned} \mu\left(\left\{\omega : |l_j - l_{j'}| > \left(1 - \sqrt{\delta}\right) h_n\right\}\right) &= \frac{\left(h_n - 1 - \left[\!\left[\left(1 - \sqrt{\delta}\right) h_n\right]\!\right]\right) h_n}{h_n^2} \\ &< 1 - \frac{1}{h_n} - \frac{\left(1 - \sqrt{\delta}\right) h_n - 1}{h_n} \\ &= \sqrt{\delta}. \end{aligned}$$

Cuando $N_1, \ldots, N_k$ son suficientemente grandes, los l_j's son casi independientes dos a dos, por lo tanto existe $\hat{N}_n \in \mathbb{N}$ tal que si $N_1, \ldots, N_k > \hat{N}_n$ entonces

$$\mu\left(\left\{\omega : |l_j - l_{j'}| < \sqrt{\delta} h_n\right\}\right) < 3\sqrt{\delta} \text{ y } \mu\left(\left\{\omega : |l_j - l_{j'}| > \left(1 - \sqrt{\delta}\right) h_n\right\}\right) < 3\sqrt{\delta}.$$

Así, si $N_1, \ldots, N_k > \hat{N}_n$ entonces

$$\begin{aligned} &\mu(\{\omega : (1 - \sqrt{\delta}) h_n < \max\{|l_j - l_{j'}|\} \text{ ó } \sqrt{\delta} h_n > \min\{|l_j - l_{j'}|\} : j \neq j'\} \\ &\qquad |\{\omega : \mathit{error} \notin \{l_0, \ldots, l_k\}\}) < 3k(k+1)\sqrt{\delta}. \quad (3.16) \end{aligned}$$

Desde que $\mu(E_n) < \frac{\delta^2}{1+\delta^2}$ y de (3.13), para todo $j \in \{0, 1, \ldots, k\}$ se tiene que

$$\begin{aligned} E(s_j|\{\omega : l_j \neq \mathit{error}\}) &= h_n \left(\frac{\mu(E_n)}{1 - \mu(E_n)}\right) \\ &< h_n \delta^2. \end{aligned}$$

Así, de la desigualdad de Márkov (Teorema 2.4), se sigue que

$$\mu(\{\omega : s_j \geq h_n\delta\}|\{\omega : l_j \neq error\}) < \delta \text{ para todo } j \in \{0, 1, \ldots, k\}. \tag{3.17}$$

De (3.15),(3.16) y (3.17) se sigue que existe $\hat{N}_n \in \mathbb{N}$ tal que si $N_1, \ldots, N_k > \hat{N}_n$ entonces

$$\mu\left(\left\{\omega \in \Omega : \left(\overrightarrow{l}, \overrightarrow{s}\right) \text{ es } \delta\text{-irrazonable}\right\}\right) < (k+1)\delta^2 + (k+1)\delta + 3k(k+1)\sqrt{\delta}.$$

□

El siguiente lema esencialmente dice que si $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ es δ-razonable, para δ pequeño, entonces de $T^i, T^{i+N_1}, \ldots, T^{i+\sum_{j=1}^{k} N_j}$, k de estos se mezclan mientras i recorre la generación actual y otros k de ellos se mezclan cuando i recorre la generación sucesora. Sea σ la permutación de $\{0, 1, \ldots, k\}$ tal que $l_{\sigma(0)} \leq l_{\sigma(1)} \leq \ldots \leq l_{\sigma(k)}$, considere los siguientes conjuntos:

$$X_{\sigma(j)} = \{i : 1 \leq i \leq h_n - l_{\sigma(k)} + l_{\sigma(0)} \text{ y } I_{n, i-l_{\sigma(0)}+l_{\sigma(j')}} \subseteq A_{\sigma(j')} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\}\}$$
$$Y = \{i : 1 \leq i \leq l_{\sigma(k)} - l_{\sigma(k-1)} \text{ y } I_{n, i-h_n+l_{\sigma(j)}+l_{\sigma(k)}} \subseteq A_{\sigma(j)} \text{ para todo } j \in \{0, \ldots, k-1\}\}$$

Observe que si $i \in X_{\sigma(j)}$ entonces $i - l_{\sigma(0)}$ está en la generación actual y

$$T^{i-l_{\sigma(0)}}\omega \in \bigcap_{\substack{j'=0 \\ j' \neq j}}^{k} T^{-\sum_{\ell=0}^{\sigma(j')} N_\ell} A_{\sigma(j')}.$$

Lema 3.3. *Sea T una transformación de rango-uno k-mezclante y sea $\delta > 0$, si n es suficientemente grande, $\overrightarrow{l}$ δ-razonable entonces para la permutación σ de $\{0, 1, \ldots, k\}$ que satisface $l_{\sigma(0)} < l_{\sigma(1)} < \ldots < l_{\sigma(k)}$, se tiene*

$$\left| \frac{\# X_{\sigma(j)}}{h_n - l_{\sigma(k)} + l_{\sigma(0)}} - \prod_{\substack{j'=0 \\ j' \neq j}}^{k} \mu(A_{\sigma(j')}) \right| < \delta \tag{3.18}$$

y

$$\left| \frac{\# Y}{l_{\sigma(k)} - l_{\sigma(k-1)}} - \prod_{j=0}^{k-1} \mu(A_{\sigma(j)}) \right| < \delta. \tag{3.19}$$

Demostración. Sin pérdida de generalidad se puede asumir que $\sigma = id$.
Sea $\widetilde{\delta} > 0$ tal que $\widetilde{\delta} < \frac{\delta^2}{5}$, por el Lema 2.7, considere n_1 de modo que para $n_2 \geq n_1$ y $m_0, \ldots, \widehat{m}_j, \ldots, m_k \geq n_1$, se tiene

$$E\left(\left| \Gamma_{n_2}[m_0, \ldots, \widehat{m}_j, \ldots, m_k] - \prod_{\substack{j'=0 \\ j' \neq j}}^{k} \mu(A_{j'}) \right|\right) < \widetilde{\delta}, \tag{3.20}$$

donde $\Gamma_n[m_0,\ldots,\widehat{m}_j,\ldots,m_k]:\Omega\to\mathbb{R}$ es la función definida por

$$\Gamma_n[m_0,\ldots,\widehat{m}_j,\ldots,m_k](\omega)=\frac{1}{n}\#\left\{\bigcap_{\substack{j'=0\\ j'\neq j}}^{k} T^{-\sum_{\ell=0}^{j'} m_\ell} A_{j'}\cap\{\omega,\ldots,T^{n_2-1}\omega\}\right\},$$

ahora escoja n suficientemente grande tal que

$$\mu(E_n)<\widetilde{\delta} \qquad \text{y} \qquad h_n>\frac{n_1}{\widetilde{\delta}}. \tag{3.21}$$

Dado $i\in\mathbb{Z}$ tal que $1\leq i\leq h_n-l_k+l_0-n_1+1$, sea

$$\begin{aligned} Q_i &= \{p:\ i\leq p\leq i+n_1-1,\ \text{y } I_{n,p-l_0+l_{j'}}\subseteq A_{j'} \text{ para todo } j'\in\{0,\ldots,\widehat{j},\ldots,k\}\}\\ &= \{p:\ i\leq p\leq i+n_1-1,\ \ p\in X_j\}. \end{aligned}$$

Ahora se considera la suma

$$\Sigma=\frac{1}{n_1}\sum_{i=1}^{h_n-l_k+l_0-n_1+1}\#Q_i,$$

Note que cada elemento de X_j se cuenta exactamente una vez en Σ excepto los $p\in X_j\cap([1,n_1]\cup[h_n-l_k+l_0-n_1+1,h_n-l_k+l_0])$, por lo tanto

$$|\#X_j-\Sigma|<2n_1. \tag{3.22}$$

Por otro lado, sea $N:\Omega\to\mathbb{N}$ la función definida como

$$N(\widetilde{\omega})=\#\{q:\ 0\leq q\leq n_1-1 \text{ y } T^q\widetilde{\omega}\in\bigcap_{\substack{j'=0\\ j'\neq j}}^{k} T^{l_j-l_{j'}}A_j\},$$

desde que $\overrightarrow{l}$ es δ-razonable y de (3.21) se tiene

$$\begin{aligned} |l_j-l_{j'}| &> \sqrt{\delta}h_n\\ &> \widetilde{\delta}h_n>n_1, \end{aligned}$$

entonces aplicando (3.20) para $n_2=n_1$ y $\sum_{\ell=0}^{j'} m_{j'}=l_j-l_{j'}$, se tiene

$$E\left(\left|\frac{1}{n_1}N-\prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu\left(A_{j'}\right)\right|\right)<\widetilde{\delta}. \tag{3.23}$$

Afirmación. Si $\widetilde{\omega} \in I_{n,i-l_0+l_j}$ y $1 \leq i \leq h_n - l_k + l_0 - n_1 + 1$ entonces

$$N(\widetilde{\omega}) = \#Q_i.$$

En efecto, si q se cuenta en $N(\widetilde{\omega})$ entonces $T^q(\widetilde{\omega}) \in T^{l_j - l_{j'}} A_{j'}$ para todo $j' \in \{0, \ldots, \widehat{j}, \ldots, k\}$, así

$$\widetilde{\omega} \in T^{l_j - l_{j'} - q} A_{j'} \cap I_{n,i-l_0+l_j} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\},$$

se sigue que

$$I_{n,i-l_0+l_j} \subseteq T^{l_j - l_{j'} - q} A_{j'} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\},$$

entonces

$$I_{n,i+q-l_0+l_{j'}} \subseteq A_{j'} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\}.$$

Por lo tanto $i + q \in Q_i$.
Análogamente, sea $p \in Q_i$ entonces $p = i + q$ con $0 \leq q \leq n_1 - 1$, luego $T^{q + l_{j'} - l_j} \widetilde{\omega} \in I_{n,i+q-l_0+l_{j'}} = I_{n,p-l_0+l_{j'}}$ entonces

$$T^{q + l_{j'} - l_j} \widetilde{\omega} \in A_{j'} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\},$$

se sigue que

$$T^q \widetilde{\omega} \in T^{l_j - l_{j'}} A_{j'} \text{ para todo } j' \in \{0, \ldots, \widehat{j}, \ldots, k\}.$$

Por lo tanto q se cuenta en $N(\widetilde{\omega})$, lo que prueba la afirmación.

De (3.21) se sigue que

$$\mu(I_{n,i}) > \frac{1 - \widetilde{\delta}}{h_n} \text{ para } 1 \leq i \leq h_n.$$

De lo último, la afirmación y aplicando (3.23) se tiene

$$\begin{aligned}
\widetilde{\delta} &> \left(\frac{1 - \widetilde{\delta}}{h_n}\right) \sum_{i=1}^{h_n - l_k + l_0 - n_1 + 1} \left| \frac{1}{n_1} \#Q_i - \prod_{\substack{j'=0 \\ j' \neq j}}^{k} \mu(A_{j'}) \right| \\
&\geq \left(\frac{1 - \widetilde{\delta}}{h_n}\right) \left| \frac{1}{n_1} \sum_{i=1}^{h_n - l_k + l_0 - n_1 + 1} \#Q_i - (h_n - l_k + l_0 - n_1 + 1) \prod_{\substack{j'=0 \\ j' \neq j}}^{k} \mu(A_{j'}) \right|
\end{aligned}$$

$$
\begin{aligned}
&\geq \left(\frac{1-\widetilde{\delta}}{h_n}\right)\left(\left|(h_n - l_k + l_0)\prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'}) - \#X_j\right| - |n_1 - 1|\prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'})\right)\\
&\quad -\left(\frac{1-\widetilde{\delta}}{h_n}\right)\left|\frac{1}{n_1}\sum_{i=1}^{h_n-l_k+l_0-n_1+1}\#Q_i - \#X_j\right|\\
&\geq \left(\frac{1-\widetilde{\delta}}{h_n}\right)\left(\left|\#X_j - (h_n - l_k + l_0)\prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'})\right| - 3n_1\right),
\end{aligned}
$$

entonces

$$
\left|\#X_j - (h_n - l_k + l_0)\prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'})\right| \leq \widetilde{\delta}\left(\frac{h_n}{1-\widetilde{\delta}}\right) + 3n_1
$$

por lo tanto

$$
\left|\frac{\#X_j}{h_n - l_k + l_0} - \prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'})\right| \leq \frac{\widetilde{\delta}h_n}{\left(1-\widetilde{\delta}\right)(h_n - l_k + l_0)} + \frac{3n_1}{h_n - l_k + l_0} \tag{3.24}
$$

Por otro lado, puesto que $\overrightarrow{l}$ es δ-razonable, se tiene que $l_k - l_0 \leq \left(1-\sqrt{\delta}\right)h_n$ entonces

$$
\sqrt{\delta}h_n \leq h_n - l_k + l_0.
$$

Así,

$$
\frac{\widehat{\delta}h_n}{\left(1-\widetilde{\delta}\right)(h_n - l_k + l_0)} < \frac{\widetilde{\delta}}{\left(1-\widetilde{\delta}\right)\sqrt{\delta}} < 2\frac{\widetilde{\delta}}{\sqrt{\delta}}, \tag{3.25}
$$

y, otra vez aplicando (3.21) se tiene

$$
\frac{3n_1}{h_n - l_k + l_0} < \frac{3n_1}{\sqrt{\delta}h_n} < 3\frac{\widetilde{\delta}}{\sqrt{\delta}}. \tag{3.26}
$$

Así, de (3.25) y (3.26) en (3.24) se sigue que

$$
\left|\frac{\#X_j}{h_n - l_k + l_0} - \prod_{\substack{j'=0\\ j'\neq j}}^{k}\mu(A_{j'})\right| < 5\frac{\widetilde{\delta}}{\sqrt{\delta}} < \frac{\delta^2}{\sqrt{\delta}} < \delta.
$$

Lo que demuestra (3.18). De manera análoga se prueba (3.19). □

A continuación se presentan los Lemas 3.4 y 3.5 en cuyas pruebas se usa el Lema2.6 para mostrar un número Q que depende de ε y dice que si conocemos k valores de los

$k+1$ valores $l_0, l_1, \ldots, l_k$ entonces a lo más hay $2Q$ valores para el l_j faltante que hace a la generación actual ε-mala y si conocemos los $k+1$ valores entonces a lo más hay $2Q$ valores para $\widehat{s}$ que hace a la generación sucesora ε-mala; en conclusión si conocemos k valores de $l_0, l_1, \ldots, l_k$, entonces a lo más hay $4Q^2$ valores para $\widehat{s}$ que hace a ambas generaciones ε-malas.

Lema 3.4. *Sea T una transformación de rango-uno k-mezclante, $\varepsilon > 0$ y $j \in \{1, \ldots k-1\}$, si $\delta > 0$ es lo suficientemente pequeño tal que*

$$\delta + \frac{\sqrt{\delta}}{1-\delta} < \varepsilon$$

entonces existe $Q \in \mathbb{N}$ de modo que si $n \in \mathbb{N}$ es lo suficientemente grande tal que $\mu(E_n) < \delta$ y que cumpla las condiciones del Lema 3.3, si σ es definido por $l_{\sigma(0)} < \ldots < l_{\sigma(k)}$, si $\sigma, l_{\sigma(0)}, \ldots, l_{\sigma(j-1)}, l_{\sigma(j+1)} \ldots, l_{\sigma(k)}$ son conocidos y se sabe que $\overrightarrow{l}$ es δ-razonable, entonces existen menos de $2Q$ valores posibles para $l_{\sigma(j)}$ (entre $l_{\sigma(j-1)}$ y $l_{\sigma(j+1)}$) tal que la generación actual es forzada a ser ε-mala.

Demostración. Sin pérdida de generalidad se puede asumir que $\sigma = id$.
Del Lema 2.6, escoja Q tal que para cualquier sucesión de Q enteros $n_1, n_2, \ldots, n_Q$ se tiene

$$E\left(|\Gamma - \mu(A_j)|\right) < \delta, \tag{3.27}$$

donde $\Gamma : \Omega \to \mathbb{R}$ está definido por

$$\Gamma(\omega) = \frac{1}{Q}\#(A_j \cap \{T^{n_1}\omega, T^{n_2}\omega, \ldots, T^{n_Q}\omega\}).$$

Se va a suponer lo contrario, es decir que que se pueden tener $2Q$ opciones $n_1, n_2, \ldots, n_{2Q}$ de valores para l_j entre l_{j-1} y l_{j+1}, tal que $(l_0, \ldots, l_{j-1}, l_j, l_{j+1}, \ldots, l_k)$ fuerza a la generación actual a ser ε-mala. Para cualquier elección de l_j, considere el conjunto

$$\theta_{l_j} = \left\{i : 1 \leq i \leq h_n - l_k + l_0; I_{n,i} \subseteq A_0, \ldots, I_{n,i+l_j-l_0} \subseteq A_j, \ldots, I_{n,i+l_k-l_0} \subseteq A_k\right\}.$$

Note que un entero i' en la generación actual $[1-l_0, h_n - l_k]$ cumple que

$$T^{i'}\omega \in A_0 \cap \ldots \cap T^{-\sum_{\ell=0}^{j} N_\ell} A_j \ldots \cap T^{-\sum_{\ell=0}^{k} N_\ell} A_k$$

si y sólo si

$$I_{n,l_0+i'} \subseteq A_0, \ldots, I_{n,l_j+i'} \subseteq A_j, \ldots \text{ y } I_{n,l_k+i'} \subseteq A_k,$$

es decir, si y sólo si $l_0 + i' \in \theta_{l_j}$, entonces

$$\sum_{i'=1-l_0}^{h_n+l_k} I(T^{i'}\omega) = \#\theta_{l_j},$$

así, la generación actual $[1-l_0, h_n-l_k]$ es ε-buena si y sólo si

$$\left|\frac{1}{h_n-l_k+l_0}\#\theta_{l_j}-\prod_{j'=0}^{k}\mu(A_{j'})\right|\leq\varepsilon,$$

entonces, para $1\leq p\leq 2Q$ se tiene

$$\left|\frac{1}{h_n-l_k+l_0}\#\theta_{n_p}-\prod_{j'=0}^{k}\mu(A_{j'})\right|>\varepsilon,$$

por lo tanto, ya sea que, hay Q valores para p tales que

$$\frac{1}{h_n-l_k+l_0}\#\theta_{n_p}-\prod_{j'=0}^{k}\mu(A_{j'})>\varepsilon, \tag{3.28}$$

ó, hay Q valores para p tales que

$$\frac{1}{h_n-l_k+l_0}\#\theta_{n_p}-\prod_{j'=0}^{k}\mu(A_{j'})<-\varepsilon.$$

En cualquier caso, la prueba por contradicción es esencialmente la misma. Suponga sin pérdida de generalidad que (3.28) se tiene para $1\leq p\leq Q$. Considere el conjunto

$$X=\{i:1\leq i\leq h_n-l_k+l_0 \text{ y } I_{n,i+l_{j'}-l_0}\subseteq A_{j'} \text{ para todo } j'\in\{0,\ldots,\widehat{j}\ldots,k\}\}.$$

Observe que para cualquier valor de l_j escogido

$$\theta_{l_j}\subseteq X.$$

Luego, considere la función $c:X\to\mathbb{N}$ definida como

$$c(i)=\sum_{p=1}^{Q}\mathbb{I}_{\theta_{n_p}}(i),$$

es decir, $c(i)$ cuenta los $p\in\{1,\ldots,Q\}$ para los cuales $i\in\theta_{n_p}$, entonces de (3.28) se tiene

$$\begin{aligned}\sum_{i\in X}c(i) &= \sum_{p=1}^{Q}\#\theta_{n_p}\\ &> Q(h_n-l_k+l_0)\left(\prod_{j'=0}^{k}\mu(A_{j'})+\varepsilon\right),\end{aligned} \tag{3.29}$$

también se tiene que si $i\in X$, entonces $i\in\theta_{n_p}$ si y sólo si $I_{n,i+n_p-l_0}\subseteq A_j$ y esto sucede si y sólo si

$$T^{n_p-l_0}I_{n,i}\subseteq A_j$$

Así, para $i \in X$ y $\widetilde{\omega} \in I_{n,i}$

$$c(i) = \sum_{p=1}^{Q} \mathbb{I}_{A_j}(T^{n_p - l_0}\widetilde{\omega}) = \#\{p : 1 \leq p \leq Q \text{ y } T^{n_p - l_0}\widetilde{\omega} \in A_j\}. \tag{3.30}$$

Desde que $\mu(E_n) < \delta$, se sigue que para cada $i \in \{1, 2, \ldots, h_n\}$

$$\mu(I_{n,i}) \geq \frac{1-\delta}{h_n}.$$

Puesto que $T^{-l_0}\widetilde{\omega}$ tiene la misma distribución que $\widetilde{\omega}$, de (3.27) y (3.30) se sigue que

$$\begin{aligned}
\delta &> E\left(|\Gamma - \mu(A_j)|\right) \\
&> \left(\frac{1-\delta}{h_n}\right) \sum_{i \in X} \left| \frac{1}{Q} c(i) - \mu(A_j) \right| \\
&\geq \left(\frac{1-\delta}{h_n}\right) \left(\left| \frac{1}{Q} \sum_{i \in X} c(i) - \#(X)\mu(A_j) \right| \right) \\
&\geq \left(\frac{1-\delta}{h_n}\right) \left| \left[\frac{1}{Q} \sum_{i \in X} c(i) \right] - \prod_{j'=0}^{k} \mu(A_{j'})(h_n - l_k + l_0) \right| \\
&\quad - \left(\frac{1-\delta}{h_n}\right) \left| \prod_{j'=0}^{k} \mu(A_{j'})(h_n - l_k + l_0) - \#(X)\mu(A_j) \right|
\end{aligned} \tag{3.31}$$

De (3.29) tenemos

$$\left| \frac{1}{Q} \sum_{i \in X} c(i) - \prod_{j'=0}^{k} \mu(A_{j'})(h_n - l_k + l_0) \right| > \varepsilon(h_n - l_k + l_0) \tag{3.32}$$

y de (3.18) se sigue que

$$\begin{aligned}
\left| \prod_{j'=0}^{k} \mu(A_{j'})(h_n - l_k + l_0) - \#(X)\mu(A_j) \right| &= \mu(A_j)(h_n - l_k + l_0) \left| \prod_{\substack{j'=0 \\ j' \neq j}}^{k} \mu(A_{j'}) - \frac{\#(X)}{h_n - l_k + l_0} \right| \\
&\leq \delta(h_n - l_k + l_0).
\end{aligned} \tag{3.33}$$

Por hipótesis tenemos que $\overrightarrow{l}$ es δ-razonable, entonces $\left(1 - \sqrt{\delta}\right) h_n \geq l_k - l_0$, así

$$\frac{h_n - l_k + l_0}{h_n} \geq \sqrt{\delta},$$

de esto, de (3.32) y de (3.33) en (3.31) tenemos

$$
\begin{aligned}
\delta &> \left(\frac{1-\delta}{h_n}\right)\left|\left[\frac{1}{Q}\sum_{i\in X} c(i)\right] - \prod_{j'=0}^{k}\mu(A_{j'})(h_n - l_k + l_0)\right| \\
&\quad -\left(\frac{1-\delta}{h_n}\right)\left|\prod_{j'=0}^{k}\mu(A_{j'})(h_n - l_k + l_0)\#(X)\mu(A_j)\right| \\
&> \left(\frac{1-\delta}{h_n}\right)(\varepsilon-\delta)(h_n - l_k + l_0) \\
&\geq \sqrt{\delta}(1-\delta)(\varepsilon-\delta).
\end{aligned}
$$

De aquí se sigue que

$$\varepsilon < \frac{\sqrt{\delta}}{(1-\delta)} + \delta,$$

lo que se contradice con la elección del δ tomado. $\square$

Lema 3.5. *Sea T una transformación de rango-uno k-mezclante, $\varepsilon > 0$, si $\delta > 0$ es lo suficientemente pequeño tal que*

$$\frac{\delta}{\sqrt{\delta}-\delta}\left(\frac{1}{1-\delta} + 4\right) < \varepsilon$$

entonces existe $Q \in \mathbb{N}$ tal que si $n \in \mathbb{N}$ es lo suficientemente grande tal que $\mu(E_n) < \delta$ y que cumpla las condiciones del Lema 3.3, entonces si $l_0, \ldots, l_k$ son conocidos y se sabe que $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ es δ-razonable, entonces existen menos de $2Q$ valores posibles para $\widehat{s}$ (desde que suponemos $\left(\overrightarrow{l}, \overrightarrow{s}\right)$ δ-razonable solo se considera los valores $\widehat{s}$ con $\widehat{s} < \delta h_n$) tal que $(l_0, \ldots, l_k, \widehat{s})$ fuerza a la generación sucesora a ser ε-mala.

Demostración. Sin pérdida de generalidad, se puede asumir que $l_0 < l_1 < \ldots < l_k$ entonces $\widehat{s} = s_k$.

Del Lema 2.6 escoja Q tal que para cualquier sucesión de Q enteros $n_1, n_2, \ldots, n_Q$ se tiene

$$E\left(|\Gamma - \mu(A_k)|\right) < \delta, \tag{3.34}$$

donde $\Gamma : \Omega \to \mathbb{R}$ está definido por

$$\Gamma(\omega) = \frac{1}{Q}\#(A_k \cap \{T^{n_1}\omega, T^{n_2}\omega, \ldots, T^{n_Q}\omega\}).$$

Se va a suponer lo contrario, es decir que se puede tener $2Q$ opciones $n_1, n_2, \ldots, n_{2Q}$ de valores para s_k, todos menores que δh_n, cada uno de los cuales fuerza a la generación sucesora a ser ε-mala.

Para cualquier elección de s_k, considere el conjunto

$$\theta_{s_k} = \{i : s_k + 1 \leq i \leq l_k - l_{k-1}; I_{n,i+h_n+l_j-l_k} \subseteq A_j \text{ para todo } 0 \leq j < k \text{ y } I_{n,i-s_k} \subseteq A_k\}.$$

Luego la generación sucesora $[h_n + 1 + s_k - l_k, h_n - l_{k-1}]$ es ε-mala si y sólo si

$$\left| \left(\frac{1}{l_k - l_{k-1} - s_k} \right) \# \theta_{s_k} - \prod_{j=0}^{k} \mu(A_j) \right| > \varepsilon.$$

De manera análoga al Lema 3.4 asuma que para $1 \leq i \leq Q$ se tiene que

$$\frac{1}{l_k - l_{k-1} - n_i} \# \theta_{n_i} - \prod_{j=0}^{k} \mu(A_j) > \varepsilon. \tag{3.35}$$

Considere el conjunto

$$Y = \{ i : 1 \leq i \leq l_k - l_{k-1}; I_{n, i + h_n + l_j - l_k} \subseteq A_j \text{ para todo } 0 \leq j < k \},$$

entonces para cualquier i se tiene que

$$i \in \theta_{s_k} \text{ si y sólo si } i \in Y,\ i \geq s_k + 1 \text{ y } I_{n, i - s_k} \subseteq A_k. \tag{3.36}$$

Ahora considere los subconjunto

$$\begin{array}{rrcl} & \widetilde{\theta}_{s_k} & = & \theta_{s_k} \setminus \{ i : 1 \leq i \leq \delta h_n \} \\ \text{y} & \widetilde{Y} & = & Y \setminus \{ i : 1 \leq i \leq \delta h_n \}, \end{array}$$

entonces si $s_k < \delta h_n$ de (3.36) se tiene que

$$i \in \widetilde{\theta}_{s_k} \text{ si y sólo si } i \in \widetilde{Y},\ \text{ y } I_{n, i - s_k} \subseteq A_k. \tag{3.37}$$

Luego, considere la función $c : \widetilde{Y} \to \mathbb{N}$ definida como

$$c(i) = \sum_{p=1}^{Q} \mathbb{I}_{\widetilde{\theta}_{n_p}}(i),$$

entonces de (3.35) y desde que $n_p < \delta h_n$ para cada p con $1 \leq p \leq Q$, se sigue que

$$\begin{aligned} \sum_{i \in \widetilde{Y}} c(i) &= \sum_{p=1}^{Q} \# \widetilde{\theta}_{n_p} \\ &\geq \sum_{p=1}^{Q} (\# \theta_{n_p} - \delta h_n) \\ &> \sum_{p=1}^{Q} \left((l_k - l_{k-1} - n_p) \left(\prod_{j=0}^{k} \mu(A_j) + \varepsilon \right) - \delta h_n \right) \\ &> \left((l_k - l_{k-1} - \delta h_n) \left(\prod_{j=0}^{k} \mu(A_j) + \varepsilon \right) - \delta h_n \right) Q. \end{aligned} \tag{3.38}$$

Suponga que $i \in \widetilde{Y}$, entonces de (3.37) se tiene que $i \in \widetilde{\theta}_{n_p}$ si y sólo si

$$T^{-n_p} I_{n,i} \subseteq A_k$$

Así, para $i \in \widetilde{Y}$ y $\widetilde{\omega} \in I_{n,i}$

$$c(i) = \sum_{p=1}^{Q} \mathbb{I}_{A_k}(T^{-n_p}\widetilde{\omega}) = \#\{p : 1 \leq p \leq Q \text{ y } T^{-n_p}\widetilde{\omega} \in A_k\}. \tag{3.39}$$

Desde que $\mu(E_n) < \delta$, se sigue que para cada $i \in \{1, 2, \ldots, h_n\}$

$$\mu(I_{n,i}) \geq \frac{1-\delta}{h_n}.$$

De (3.34) y (3.39) se sigue que

$$\begin{aligned}
\delta &> E\left(|\Gamma - \mu(A_k)|\right) \\
&> \left(\frac{1-\delta}{h_n}\right) \sum_{i \in \widetilde{Y}} \left|\frac{1}{Q} c(i) - \mu(A_k)\right| \\
&\geq \left(\frac{1-\delta}{h_n}\right) \left|\frac{1}{Q} \sum_{i \in \widetilde{Y}} c(i) - \#(\widetilde{Y})\mu(A_k)\right| \\
&\geq \left(\frac{1-\delta}{h_n}\right) \left[\left|\frac{1}{Q} \sum_{i \in \widetilde{Y}} c(i) - (l_k - l_{k-1} - \delta h_n) \prod_{j=0}^{k} \mu(A_j)\right| - \delta h_n \prod_{j=0}^{k} \mu(A_j)\right] \\
&\quad - \left(\frac{1-\delta}{h_n}\right) |(l_k - l_{k-1}) \prod_{j=0}^{k} \mu(A_j) - \#(Y)\mu(A_k)| - \left(\frac{1-\delta}{h_n}\right) \left|\mu(A_k)\left(\#(Y) - \#(\widetilde{Y})\right)\right|.
\end{aligned} \tag{3.40}$$

De (3.38) se tiene que

$$\frac{1}{Q} \sum_{i \in \widetilde{Y}} c(i) - (l_k - l_{k-1} - \delta h_n) \prod_{j=0}^{k} \mu(A_j) > \varepsilon(l_k - l_{k-1} - \delta h_n) - \delta h_n.$$

Así

$$\left|\frac{1}{Q} \sum_{i \in \widetilde{Y}} c_i - (l_k - l_{k-1} - \delta h_n) \prod_{j=0}^{k} \mu(A_j)\right| > \varepsilon(l_k - l_{k-1} - \delta h_n) - \delta h_n. \tag{3.41}$$

Por otro lado, aplicando (3.19)

$$\begin{aligned}
\left|(l_k - l_{k-1}) \prod_{j=0}^{k} \mu(A_j) - \#(Y)\mu(A_k)\right| &= \mu(A_k)(l_k - l_{k-1}) \left|\prod_{j=0}^{k} \mu(A_j) - \frac{\#(Y_2)}{l_k - l_{k-1}}\right| \\
&< \mu(A_k)(l_k - l_{k-1})\delta.
\end{aligned} \tag{3.42}$$

De la definición de $\widetilde{Y}$ se sigue que

$$|\#(Y) - \#(\widetilde{Y})| \leq \delta h_n. \tag{3.43}$$

Juntando (3.41),(3.42) y (3.43) en (3.40) se sigue que

$$\begin{aligned}
\delta &> \left(\frac{1-\delta}{h_n}\right)\left(\varepsilon(l_k - l_{k-1} - \delta h_n) - \delta h_n - \delta h_n \prod_{j=0}^{k} \mu(A_j) - \delta(l_k - l_{k-1})\mu(A_k) - \mu(A_k)\delta h_n\right) \\
&\geq \left(\frac{1-\delta}{h_n}\right)\left(\varepsilon(\sqrt{\delta} h_n - \delta h_n) - \delta h_n - \delta h_n \prod_{j=0}^{k} \mu(A_j) - \delta(1-\sqrt{\delta})h_n\mu(A_k) - \mu(A_k)\delta h_n\right) \\
&\geq (1-\delta)(\varepsilon(\sqrt{\delta} - \delta) - \delta - \delta - \delta + \delta\sqrt{\delta}\mu(A_k) - \delta) \\
&> (1-\delta)(\varepsilon(\sqrt{\delta} - \delta) - 4\delta).
\end{aligned}$$

Así

$$\frac{\delta}{\sqrt{\delta} - \delta}\left(\frac{1}{1-\delta} + 4\right) > \varepsilon,$$

lo que se contradice con la elección del δ tomado. □

El siguiente lema nos dice que la probabilidad de que $l_{\sigma(1)}, \ldots, l_{\sigma(k-1)}$ tomen determinados valores no varía mucho a la probabilidad de que $l_{\sigma(1)}, \ldots, l_{\sigma(k-1)}$ tomen esos mismos valores pero ya fijando los valores de $l_{\sigma(k)}$ y $s_{\sigma(k)}$.

Lema 3.6. *Sea T una transformación de rango-uno k-mezclante, se cumple*

1. *Bajo la condición $l_j \neq$error, l_j es independiente de s_j para todo $0 \leq j \leq k$.*

2. *Para cada permutación σ de $\{0, 1, \ldots, k\}$. Si $\delta > 0$, existe $M_0 \in \mathbb{N}$ tal que si $N_1, \ldots, N_k > M_0$ entonces*

$$\begin{aligned}
|\mu(\{\omega : l_{\sigma(1)} = \alpha_1, \ldots, l_{\sigma(k-1)} = \alpha_{k-1}\} | \{\omega : l_{\sigma(k)}, s_{\sigma(k)}\}) \\
- \mu(\{\omega : l_{\sigma(1)} = \alpha_1, \ldots, l_{\sigma(k-1)} = \alpha_{k-1}\})| < \delta,
\end{aligned}$$

   para todo α_j.

Demostración.

1. Por definición s_j es el número de b's entre el n-bloque que contiene a $T^{\sum_{\ell=0}^{j} N_\ell}\omega$ y el siguiente n-bloque, es decir que depende de $N_1 \ldots, N_j$ y n; mas no depende del nivel en que se encuentre, osea no depende de l_j.

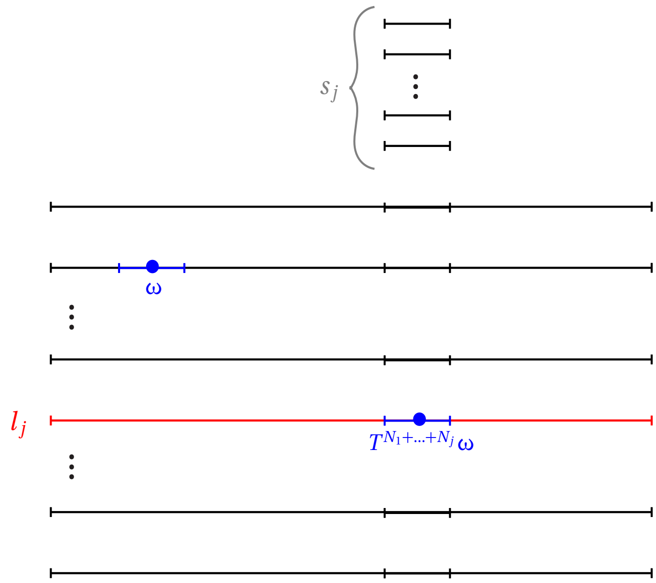

Figura 3.10:

2. Se puede suponer, sin pérdida de generalidad que $\sigma = id$, en tal caso queremos probar que

   $$|\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\}|\{\omega : l_k,\ s_k\}) - \mu(\{l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})| < \delta.$$

   Sabemos que $0 < \mu(E_n) \leq \mu(E_1) < 1$, entonces

   $$\frac{1}{1-\mu(E_n)} \leq \frac{1}{1-\mu(E_1)}.$$

   Como T es k-mezclante, existe $M_0 \in \mathbb{N}$ tal que si $N_1, \ldots, N_k > M_0$ entonces

   $$\left|\mu\left(\bigcap_{j=1}^{k-1} T^{-\sum_{\ell=1}^{j} N_\ell} I_{n,\alpha_j} \cap T^{-\sum_{\ell=1}^{k} N_\ell} I_{n,l_k}\right) - \mu(I_{n,l_k})\prod_{j=1}^{k-1}\mu(I_{n,\alpha_j})\right| < \frac{1-\mu(E_1)}{h_n}\delta. \tag{3.44}$$

   Luego

   $$\begin{aligned}
   &|\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\}|\{\omega : l_k, s_k\}) - \mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})| \\
   &= \left|\frac{\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}, l_k, s_k\})}{\mu(\{\omega : l_k, s_k\})} - \mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})\right| \\
   &= \left|\frac{\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}, l_k\})\mu(\{\omega : s_k\})}{\mu(\{\omega : l_k\})\mu(\{\omega : s_k\})}\right. \\
   &\qquad\qquad \left. - \mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})\right|
   \end{aligned}$$

$$
\begin{aligned}
&= \left| \frac{\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}, l_k\})}{\mu(\{\omega : l_k\})} \right. \\
&\qquad\qquad \left. - \frac{\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})\mu(\{\omega : l_k\})}{\mu(\{\omega : l_k\})} \right| \\
&= \left| \frac{\mu\left(\{\omega : \omega \in \bigcap_{j=1}^{k-1} T^{-\sum_{\ell=1}^{j} N_\ell} I_{n,\alpha_j} \cap T^{-\sum_{\ell=1}^{k} N_\ell} I_{n,l_k}\}\right) - \mu(I_{n,l_k}) \prod_{j=1}^{k-1} \mu(I_{n,\alpha_j})}{\frac{1}{h_n}(1 - \mu(E_n))} \right| \\
&= \left| \frac{\mu\left(\bigcap_{j=1}^{k-1} T^{-\sum_{\ell=1}^{j} N_\ell} I_{n,\alpha_j} \cap T^{-\sum_{\ell=1}^{k} N_\ell} I_{n,l_k}\right) - \mu(I_{n,l_k}) \prod_{j=1}^{k-1} \mu(I_{n,\alpha_j})}{\frac{1}{h_n}(1 - \mu(E_n))} \right| \\
&< \delta
\end{aligned}
$$

□

Observación 3.6. De la demostración del lema anterior, también se obtiene

$$|\mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\}|\{\omega : l_k\}) - \mu(\{\omega : l_1 = \alpha_1, \ldots, l_{k-1} = \alpha_{k-1}\})| < \delta,$$

siempre que $N_1, \ldots, N_k > M_0$.

Lema 3.7. *Sea T una transformación de rango-uno mezclante y $\delta > 0$ entonces para $n \in \mathbb{N}$ suficientemente grande,*

$$\mu(\{\omega \in \Omega : s_j = i\}) < \delta$$

para todo i y para todo $j \in \{0, 1, \ldots, k\}$.

Demostración. Sea $m \in \mathbb{N}$ tal que

$$\frac{1}{h_m} < \frac{\delta}{2}$$

y sea $I = I_{m,1}$ el primer nivel de la m-torre, entonces

$$\mu(I) = \frac{1}{h_m}\mu(m\text{-torre}) < \frac{\delta}{2}.$$

Como T es una transformación mezclante, podemos escoger L_0 tal que

$$\mu(T^L(I) \cap I) < \frac{\delta}{2} + \mu(I)^2 < \delta \text{ para todo } L \geq L_0.$$

Así

$$\mu(\{\omega \in I : T^L\omega \in I\}) < \delta \text{ para todo } L \geq L_0. \tag{3.45}$$

Ahora, se considera $n > m$ tal que $h_n > L_0$ entonces para cualquiera de estos n tenemos

$$\mu(\{\omega \in \Omega : s_j = i\}) < \delta \text{ para todo } i \text{ y todo } j \in \{0, 1, \ldots, k\}.$$

En efecto, como s_j no depende en que nivel de la n-torre se encuentra $T^{N_1+\ldots+N_j}\omega$, se puede asumir que está en I puesto que, por ser $n > m$, I es unión de niveles de la n-torre, así

$$\mu(\{\omega : s_j = i\}|\{\omega : s_j \text{ está definido }\}) = \mu(\{\omega : s_j = i\}|\{\omega : T^{N_1+\ldots+N_j}\omega \in I\}).$$

Si $s_j = i$, entonces

$$T^{h_n+i}(T^{N_1+\ldots+N_j}\omega) \in I_{n,r} \text{ si y sólo si } T^{N_1+\ldots+N_j}\omega \in I_{n,r}.$$

Así, si $T^{N_1+\ldots+N_j}\omega \in I$ entonces $T^{h_n+j}(T^{N_1+\ldots+N_j}\omega) \in I$, por lo tanto

$$\begin{aligned}
\mu(\{\omega \in \Omega : s_j = i\}) &\leq \mu(\{\omega : s_j = i\}|\{\omega : s_j \text{ esta definido }\}) \\
&= \mu(\{\omega : s_j = i\}|\{\omega : T^{N_1+\ldots+N_j}\omega \in I\}) \\
&\leq \mu(\{\omega : T^{h_n+j}(T^{N_1+\ldots+N_j}\omega) \in I\}|\{\omega : T^{N_1+\ldots+N_j}\omega \in I\}) \\
&< \delta
\end{aligned}$$

donde la última desigualdad se sigue de (3.45). □

Teorema 3.3. *Sea T una transformación de rango-uno k-mezclante y $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para cada $n > n_0$ la probabilidad de que ω se encuentre entre n-generaciones o que ambas, la n-generación actual y la n-generación sucesora sean ε-malas es menor que ε, cuando $N_1, \ldots, N_k > M_n$ para algún M_n.*

Demostración. Considere $\delta > 0$ lo suficientemente pequeño tal que obedezca las condiciones de los Lemas 3.4 y 3.5 y además que

$$(k+1)\delta^2 + (k+1)(1 + 24(k-1)k!)\delta + 3k(k+1)\sqrt{\delta} < \varepsilon \tag{3.46}$$

Sea σ una permutación cualquiera de $\{0, 1, \ldots, k\}$ y $j \in \{1, \ldots, k-1\}$ diremos que $(l_{\sigma(0)}, \ldots, l_{\sigma(j-1)}, l_{\sigma(j+1)} \ldots, l_{\sigma(k)}, s_{\sigma(k)})$ es **doblemente malo** si $l_{\sigma(0)} < \ldots < l_{\sigma(j-1)} < l_{\sigma(j+1)} < \ldots < l_{\sigma(k)}$ y si existe un valor para $l_{\sigma(j)}$ tal que $l_{\sigma(j-1)} < l_{\sigma(j)} < l_{\sigma(j+1)}$ y $(l_{\sigma(0)}, \ldots, l_{\sigma(j-1)}, l_{\sigma(j)}, l_{\sigma(j+1)} \ldots, l_{\sigma(k)}, s_{\sigma(k)})$ es δ-razonable (es decir que existe $s_{\sigma(j)}$ tal que $(\overrightarrow{l}, \overrightarrow{s})$ es δ-razonable), pero ambas, la generación actual y la generación sucesora son ε-malas.

Considere los siguientes conjuntos

$$\begin{aligned}
Z &= \{\omega : (\overrightarrow{l}, \overrightarrow{s}) \text{ es } \delta\text{-irrazonable}\} \\
Z_{\sigma,j} &= \{\omega : (l_{\sigma(0)}, \ldots, l_{\sigma(j-1)}, l_{\sigma(j+1)} \ldots, l_{\sigma(k)}, s_{\sigma(k)}) \text{ es doblemente malo}\}.
\end{aligned}$$

Donde σ varía en el conjunto de permutaciones de $\{0,1,\ldots,k\}$ y $j \in \{1,\ldots,k-1\}$, es decir, los $Z_{\sigma,j}$ son un total de $[(k-1)(k+1)!]$ conjuntos.

Observe que si ω evita estos $[(k-1)(k+1)!+1]$ conjuntos entonces ω esta en una generación y esta generación o la generación sucesora es ε-buena. Entonces la prueba va a consistir en encontrar n_0 tal que para $n > n_0$ existe M_n tal que si $N_1,\ldots,N_k > M_n$ entonces $N_1,\ldots,N_k,n$ fuerzan a estos conjuntos a ser pequeños.
Escoja Q y n_0 tales que si $n > n_0$, la conclusión de los Lemas 3.4 y 3.5 se cumplan para este Q, también que

$$\mu(E_n) < \frac{\delta^2}{1+\delta^2}$$

y que se cumpla el Lema 3.7 pero con δ reemplazado por $\frac{\delta}{Q^2}$.

Luego escoja M_n tal que si $N_1,\ldots,N_k > M_n$ entonces se cumplen los Lemas 3.2 y 3.6 en este último se reeplaza δ por $\frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}$.
Si fijamos $n > n_0$ y $N_1,\ldots,N_k > M_n$, por el Lema 3.2 se tiene

$$\mu(Z) < (k+1)\delta^2 + (k+1)\delta + 3k(k+1)\sqrt{\delta} \tag{3.47}$$

Ahora se va a calcular $\mu(Z_{id,1})$. Por el Lema 3.4 hay menos de $2Q$ valores para l_1 tal que $l_0 < l_1 < l_2$, $(l_0,l_1,\ldots,l_k)$ sea δ-razonable y que la generación actual sea ε-mala. Una vez escogido l_1, por el Lema 3.5 hay menos de $2Q$ valores a escoger para $\widehat{s}$ (en este caso s_k) tal que $(l_0,l_1,\ldots,l_k,s_k)$ es razonable y para el cual la generción sucesora es ε-mala. Así, si $l_0 < l_2 < \ldots < l_k$ hay a lo más $(2Q)(2Q) = 4Q^2$ valores de s_k tal que $(l_0,l_2,\ldots,l_k,s_k)$ es doblemente malo.

Suponga que $l_0,l_2,\ldots,l_k$ son conocidos, digamos $l_j = a_j$ para todo $j \in \{0,2,\ldots,k\}$. Sean $d_1,d_2,\ldots,d_{4Q^2}$ los $4Q^2$ valores incluyendo los valores de s_k donde $(l_0,l_2,\ldots,l_k,s_k)$ es doblemente malo. Recuerde que escogimos n_0 tal que si $n > n_0$, se cumple el Lema 3.7 con δ reemplazado por $\frac{\delta}{Q^2}$. Para tal n

$$\mu(\{\omega : s_k = d_i\}) < \frac{\delta}{Q^2} \quad \text{para todo } i,\ 1 \le i \le 4Q^2 \tag{3.48}$$

Por otro lado se tiene

$$\mu(\{\omega : s_k{=}d_i\}|\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_k{=}a_k\}) = \frac{\mu(\{\omega : s_k{=}d_i, l_0{=}a_0, l_2{=}a_2, \ldots, l_k{=}a_k\})}{\mu(\{\omega : l_0{=}a_0, l_2 = a_2, \ldots, l_k{=}a_k\})}$$

$$= \frac{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_{k-1}{=}a_{k-1}\}|\{\omega : l_k{=}a_k, s_k{=}d_i\})\mu(\{\omega : l_k{=}a_k, s_k{=}d_i\})}{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_k{=}a_k\})} \quad (3.49)$$

Puesto que los eventos $l_k = a_k$, $s_k = d_i$, son independientes bajo la condición $l_k \neq error$, se sigue que

$$\mu(\{\omega : l_k{=}a_k, s_k{=}d_i\}|\{\omega : l_k{\neq}error\}) = \mu(\{\omega : l_k{=}a_k\}|\{\omega : l_k{\neq}error\})$$
$$\mu(\{\omega : s_k{=}d_i\}|\{\omega : l_k{\neq}error\}) \quad (3.50)$$

Pero $\{\omega : l_k = a_k\} \subseteq \{\omega : l_k \neq error\}$ y $\{\omega : s_k = d_i\} \subseteq \{\omega : l_k \neq error\}$ entonces en (3.50) se tiene

$$\frac{\mu(\{\omega : l_k = a_k, s_k = d_i, l_k \neq error\})}{\mu(\{\omega : l_k \neq error\})} = \frac{\mu(\{\omega : l_k = a_k, l_k \neq error\})}{\mu(\{\omega : l_k \neq error\})}$$
$$\frac{\mu(\{\omega : s_s = d_i, l_k \neq error\})}{\mu(\{\omega : l_k \neq error\})}$$

Así

$$\begin{aligned}\mu(\{\omega : l_k = a_k, s_k = d_i\}) &= \frac{\mu(\{\omega : l_k = a_k\})\mu(\{\omega : s_k = d_i\})}{\mu(\{\omega : l_k \neq error\})} \\ &< 2\mu(\{\omega : l_k = a_k\})\mu(\{\omega : s_k = d_i\}) \quad (3.51)\end{aligned}$$

donde la última desigualdad se sigue desde que $\mu(\{\omega : l_k \neq error\}) > \frac{1}{2}$. Usando (3.51) en (3.49) tenemos

$$\begin{aligned}&\mu(\{\omega : s_k = d_i\}|\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_k = a_k\}) \\ &< 2\frac{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_{k-1}{=}a_{k-1}\}|\{\omega : l_k{=}a_k, s_k{=}d_i\})\mu(\{\omega : l_k{=}a_k\})\mu(\{\omega : s_k{=}d_i\})}{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_k{=}a_k\})} \\ &= 2\frac{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_{k-1}{=}a_{k-1}\}|\{\omega : l_k{=}a_k, s_k{=}d_i\})\mu(\{\omega : s_k{=}d_i\})}{\mu(\{\omega : l_0{=}a_0, l_2{=}a_2, \ldots, l_{k-1}{=}a_{k-1}\}|\{\omega : l_k{=}a_k\})} \quad (3.52)\end{aligned}$$

Recuerde que M_n se escogió de manera que si $N_1, \ldots, N_k > M_n$ entonces se cumple el Lema 3.6 con δ reemplazado por $\frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}$, entonces

$$\begin{aligned}
&\mu(\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_{k-1}\}|\{\omega : l_k = a_k, s_k = d_i\}) \\
&\quad < \mu(\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_{k-1}\}) + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1} \\
&\quad = \mu(\{\omega : l_2 = a_2, \ldots, l_{k-1}\}|\{\omega : l_0 = a_0\})\mu(\{\omega : l_0 = a_0\}) + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1} \\
&\quad < \mu(\{\omega : l_0 = a_0\})\left(\mu(\{\omega : l_2 = a_2, \ldots, l_{k-1}\}) + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}\right) + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1} \\
&\quad = \mu(\{\omega : l_0 = a_0\})\mu(\{\omega : l_2 = a_2, \ldots, l_{k-1}\}) + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}(1 + \mu(\{\omega : l_0 = a_0\}))
\end{aligned}$$

Siguiendo así por inducción y sabiendo que

$$\mu(\{\omega : l_0 = a_0\}) = \ldots = \mu(\{\omega : l_k = a_k\}) = \frac{1-\mu(E_n)}{h_n}$$

se tiene que

$$\begin{aligned}
&\mu(\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_{k-1}\}|\{\omega : l_k = a_k, s_k = d_i\}) \\
&\qquad < (\mu(\{\omega : l_0 = a_0\}))^{k-1} + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}\sum_{j=0}^{k-2}\mu(\{\omega : l_0 = a_0\})^j \\
&\qquad < (\mu(\{\omega : l_0 = a_0\}))^{k-1} + \frac{1}{2k}\left(\frac{1-\delta}{h_n}\right)^{k-1}(k-1) \\
&\qquad < (\mu(\{\omega : l_0 = a_0\}))^{k-1} + \frac{1}{2}\left(\frac{1-\delta}{h_n}\right)^{k-1} \\
&\qquad < \frac{3}{2}(\mu(\{\omega : l_0 = a_0\}))^{k-1}
\end{aligned}$$

Análogamente, se tiene

$$\begin{aligned}
\mu(\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_{k-1}\}|\{\omega : l_k = a_k\}) &> (\mu(\{\omega : l_0 = a_0\}))^{k-1} - \frac{1}{2}\left(\frac{1-\delta}{h_n}\right)^{k-1} \\
&> \frac{1}{2}(\mu(\{\omega : l_0 = a_0\}))^{k-1}
\end{aligned}$$

De esto y (3.48) en (3.52) tenemos

$$\begin{aligned}
\mu(\{\omega : s_k = d_i\}|\{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_k = a_k\}) &< 2 \cdot \frac{\frac{3}{2}(\mu(\{\omega : l_1 = a\}))^{k-1}}{\frac{1}{2}(\mu(\{\omega : l_1 = a\}))^{k-1}} \cdot \frac{\delta}{Q^2} \\
&= \frac{6\delta}{Q^2} \qquad (3.53)
\end{aligned}$$

Para que $(l_0, l_2, \ldots, l_k, s_k)$ sea doblemente malo , donde $l_j = a_j$ para cada $j \in \{0, 2, \ldots, k\}$, es necesario que $s_k = d_i$ para cualquier $1 \leq i \leq 4Q^2$. Así por (3.53)

$$\mu(\{\omega : (l_0, l_2, \ldots, l_k, s_k) \text{ doblemente malo}\} | \{\omega : l_0 = a_0, l_2 = a_2, \ldots, l_k = a_k\}) < \frac{6\delta}{Q^2} \cdot 4Q^2 = 24\delta$$

donde los a_j son arbitrarios, por lo tanto $\mu(Z_{id,1}) < 24\delta$.

Análogamente se sigue $\mu(Z_{\sigma,j}) < 24\delta$ para toda permutación σ y todo $1 \leq j \leq k-1$. De esto, (3.46) y (3.47) se sigue que

$$\begin{aligned}
\mu\left(Z \cup \bigcup_{\sigma,j} Z_{\sigma,j}\right) &\leq \mu(Z) + \sum_{\sigma,j} \mu(Z_{\sigma,j}) \\
&< (k+1)\delta^2 + (k+1)\delta + 3k(k+1)\sqrt{\delta} + [(k-1)(k+1)!]24\delta \\
&= (k+1)\delta^2 + (1 + 24(k-1)k!)\delta + 3k(k+1)\sqrt{\delta} \\
&< \varepsilon
\end{aligned}$$

□

Teorema 3.4. *Sea T una transformación de rango-uno k-mezclante y $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para cada $n > n_0$ la probabilidad de que ω se encuentre entre n-generaciones o que ambas, la n-generación predecesora y la n-generación actual sean ε-malas es menor que ε, cuando $N_1, \ldots, N_k > M_n$ para algún M_n.*

Demostración. La prueba es ánaloga a la del Teorema 3.3 teniendo en cuenta la Observación 3.2 la cual dice que también podemos definir el número de b's entre el n-bloque que contiene a ω_0 el n-bloque anterior. □

Capítulo 4

En transformaciones de rango-uno, 2-mezclante implica k-mezclante

El resultado principal de la tesis se presenta en este capítulo, Sección 3.2, la prueba será hecha por inducción, de modo que primero probaremos que k-mezclante implica $(k+1)$-mezclante.

4.1. En transformaciones de rango-uno, k-mezclante implica $(k+1)$-mezclante

En esta sección vamos a suponer que la transformación de rango-uno ya es k-mezclante para algún $k \geq 2$ y probaremos que también es $(k+1)$-mezclante, así el objetivo es encontrar $M_0 \in \mathbb{N}$ tal que si $N_1, \ldots, N_k > M_0$ entonces

$$\left| \mu \left(A_0 \cap T^{N_1} A_1 \cap \ldots T^{N_1 + \ldots + N_k} A_k \right) - \prod_{j=0}^{k} \mu(A_j) \right|$$

es pequeño.

Se empezara por encontrar una sucesión de términos, el último de los cuales es M_0. Primero fije $\varepsilon > 0$, y sea $g \in \mathbb{N}$ un número par de modo que

$$\left(\frac{1}{1+\varepsilon} \right)^{\frac{g}{2}} < \sqrt{\varepsilon}, \tag{4.1}$$

considere

$$\delta = \left(\frac{\varepsilon}{g} \right)^2. \tag{4.2}$$

Ahora se define una sucesión de $(g+1)$ números enteros $n_0, n_1, \ldots, n_g$, de la siguiente manera; para empezar escoja n_0 mayor que el encontrado en los Teoremas 3.3 y 3.4 , con

ε reemplazado por el δ de (4.2). Para $i < g$ escoja $n_{i+1} > n_i$, lo suficiente, tal que

$$\delta^2 h_{n_{i+1}} > h_{n_i} \tag{4.3}$$

donde h_n es el tamaño de la n-torre. Observe que $n_i \geq n_0$ para todo $0 \leq i \leq g$ y por lo tanto mayores que el de los Teoremas 3.3 y 3.4, con ε reemplazado por δ.

Ahora se va a escoger el M_0. Sean $j \neq j'$ ambos en $\{0, 1, \ldots, k\}$ y algún n fijo, puesto que la transformación T es 2-mezclante entonces $l_j(n, N_1, \ldots, N_k)$ y $l_{j'}(n, N_1, \ldots, N_k)$ son casi independientes a medida que $N_1, \ldots, N_k$ crecen, en particular, si l_j y $l_{j'}$ fuesen valores completamente independientes entre 1 y h_n entonces

$$\mu(\{\omega : |l_j - l_{j'}| < 2\varepsilon h_n\}) < 4\varepsilon.$$

Así, existe M_0 suficientemente grande tal que si $N_1, \ldots, N_k > M_0$

$$\mu(\{\omega : |l_{j'}(n_i, N_1, \ldots, N_k) - l_j(n_i, N_1, \ldots, N_k)| < 2\varepsilon h_{n_i}\}) < 4\varepsilon \tag{4.4}$$

para todo $j \neq j'$ ambos en $\{0, 1, \ldots, k\}$ y todo $i \in \{0, 1, \ldots, g\}$.
Dado $i \in \{0, 1, \ldots, g\}$, considere los siguientes eventos

$$\begin{aligned}
G_{i,1} &= \{\omega : \omega \text{ esta entre } (n_i, N_1, \ldots, N_k)\text{-generaciones}\}.\\
G_{i,2} &= \{\omega : \text{la } (n_i, N_1, \ldots, N_k)\text{-generación actual y la } (n_i, N_1, \ldots, N_k)\text{-generación}\\
&\quad\ \text{sucesora son } \delta\text{-malos}\}.\\
G_{i,3} &= \{\omega : \text{la } (n_i, N_1, \ldots, N_k)\text{-generación actual y la } (n_i, N_1, \ldots, N_k)\text{-generación}\\
&\quad\ \text{predecesora son } \delta\text{-malos}\}.
\end{aligned}$$

Sea G_i^0 a la unión de estos eventos, es decir

$$G_i^0 = G_{i,1} \cup G_{i,2} \cup G_{i,3}.$$

Se sigue de los Teoremas 3.3 y 3.4 que si M_0 es lo suficientemente grande entonces para $N_1, \ldots, N_k > M_0$

$$\mu(G_i^0) < 2\delta \text{ para cada } i \in \{0, 1, \ldots, g\} \tag{4.5}$$

Se toma M_0 lo suficientemente grande tal que se cumpla (4.4) y (4.5), de ahora en adelante consideraremos $N_1, \ldots, N_k$ fijos mayores que M_0.

Definición 4.1. Sea $S \subseteq \mathbb{Z}$ finito. Diremos que S es ω–*bueno* si

$$\left| \frac{1}{\#(S)} \sum_{i \in S} I(T^i \omega) - \prod_{j=0}^{k} \mu(A_j) \right| < \delta,$$

donde $I:\Omega\to\{0,1\}$ es la función definida por

$$I(\omega)=\mathbb{I}_{A_0\cap T^{-N_1}A_1\cap\ldots\cap T^{-\sum_{j=1}^{k}N_j}A_k}(\omega).$$

Vamos a decir que S es *bueno* si ω está fijado.

Se sigue de la definición que una n-generación δ-buena para ω es un conjunto, en este caso un intervalo, ω-bueno, además, la unión disjunta de conjuntos ω-buenos es ω-bueno, en efecto, pues si S, R son buenos entonces

$$\begin{aligned}
&\left|\frac{1}{\#(S\sqcup R)}\sum_{i\in S\sqcup R}I(T^i\omega)-\prod_{j=0}^{k}\mu(A_j)\right| \\
&\quad\leq \frac{1}{\#(S\sqcup R)}\left(\left|\sum_{i\in S}I(T^i\omega)-\#(S)\prod_{j=0}^{k}(A_j)\right|+\left|\sum_{i\in R}I(T^i\omega)-\#(R)\prod_{j=0}^{k}(A_j)\right|\right)\\
&\quad< \frac{1}{\#(S\sqcup R)}(\#(S)\delta+\#(R)\delta)\\
&\quad= \delta.
\end{aligned}$$

Sea $L=h_{n_g}$ se va a probar que $\mu(A_0\cap T^{N_1}A_1\cap\ldots\cap T^{N_1+\ldots+N_k}A_k)$ esta cerca de $\prod_{j=0}^{k}\mu(A_j)$ encontrando, para la mayoria de ω's, un subconjunto '**grande**' de $\{0,1,\ldots,L-1\}$ el cual es bueno (por grande se refiere a que su cardinal está muy serca de ser igual a L). Ahora, para cada ω, se va a construir una sucesión de particiones del conjunto $\{0,1,\ldots,L-1\}$ en tres subconjuntos, cada partición usada para construir la siguiente, tal que para la mayoria de ω's uno de los conjuntos en la última partición es nuestro conjunto grande y bueno buscado.

Sea

$$(J_g^1(\omega),J_g^2(\omega),J_g^3(\omega)),(J_{g-1}^1(\omega),J_{g-1}^2(\omega),J_{g-1}^3(\omega)),\ldots,(J_0^1(\omega),J_0^2(\omega),J_0^3(\omega))$$

la sucesión de particiones en tres conjuntos de $\{0,1,\ldots,L-1\}$ construidos en este orden por inducción. Se Denotara J_i^1,J_i^2,J_i^3 en vez de $J_i^1(\omega),J_i^2(\omega),J_i^3(\omega)$, cuando ω está fijado. Primero se considera

$$J_g^3=J_g^2=\emptyset \text{ y } J_g^1=\{0,1,\ldots,L-1\}.$$

Sea $\mathfrak{J}_i=\{R\subseteq J_{i+1}^1:\ R \text{ es } n_i\text{-generación } \delta\text{-malo para } \omega \text{ y } \#R<\delta h_{n_i}\}$ para $0\leq i\leq g-1$;

entonces se define

$$
\begin{aligned}
J_{g-1}^3 &= \{R \subseteq J_g^1 : \ R \text{ es } n_{g-1}\text{-generación } \delta\text{-bueno para } \omega\}, \\
J_{g-1}^2 &= \{q \in J_g^1 : \ q \text{ no está en ninguna } n_{g-1}\text{-generación para } \omega\} \cup \mathfrak{I}_{g-1} \\
&\quad \cup \{q \in J_g^1 : \ q \text{ está en una } n_{g-1}\text{-generación para } \omega \text{ que intersecta a } J_g^1 \\
&\quad \ \text{pero no está contenida en } J_g^1\} \ \text{ y} \\
J_{g-1}^1 &= \{0, 1, \ldots, L-1\} \setminus (J_{g-1}^2 \cup J_{g-1}^3).
\end{aligned}
$$

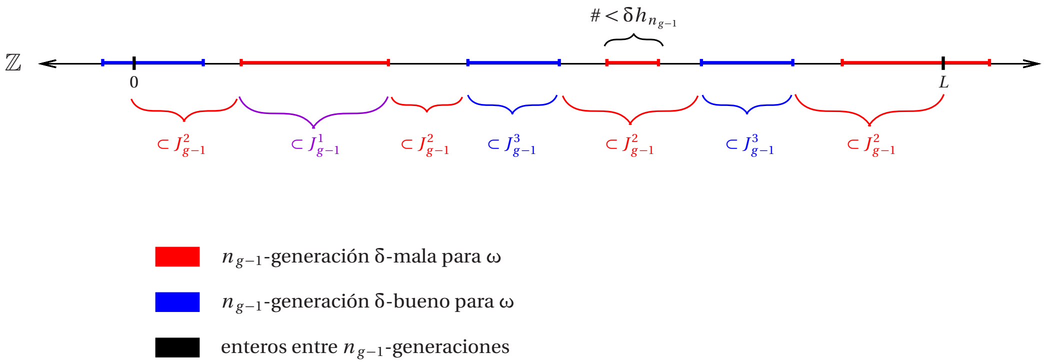

Figura 4.1:

Ahora suponga que J_{i+1}^1, J_{i+1}^2, J_{i+1}^3 ya están construidos, entonces se define

$$
\begin{aligned}
J_i^3 &= J_{i+1}^3 \cup \{R \subseteq J_{i+1}^1 : \ R \text{ es } n_i\text{-generación } \delta\text{-bueno para } \omega\}, \\
J_i^2 &= J_{i+1}^2 \cup \{q \in J_{i+1}^1 : \ q \text{ no esta en ninguna } n_i\text{-generación para } \omega\} \cup \mathfrak{I}_i, \\
J_i^1 &= \{0, 1, \ldots, L-1\} \setminus (J_i^2 \cup J_i^3).
\end{aligned}
$$

Observación 4.1.

1. Para el caso particular de $i = g-1$ la definición de J_{g-1}^2 varía ligeramente en comparación con J_i^2 para $0 \leq i \leq g-2$.

2. El conjunto $J_{g-1}^2 \cup J_{g-1}^3$, es unión de n_{g-1}-generaciones, los enteros entre $(g-1)$-generaciones y las n_{g-1}-generaciones incompletas, que a lo más son dos, que están en los extremos de $\{0, 1, \ldots, L-1\}$. El complemento de esta unión, J_{g-1}^1, es unión de n_{g-1}-generaciones, cada una de estas es unión de n_{g-2}-generaciones con enteros que están entre n_{g-2}-generaciones; continuando por inducción se tiene que cada J_i^1 es unión de n_i-generaciones.

3. Se ver, por inducción, que J_i^3 es unión disjunta de n_i-generaciones δ-buenos entonces se sigue que J_i^3 es un conjunto bueno. En particular J_0^3 es un conjunto bueno.

El objetivo es establecer que para la mayoria de ω's, $J_0^3(\omega)$ es el conjunto grande y bueno buscado que se mencionó anteriormente. Por lo tanto solo resta establecer que para la mayoría de ω's el $J_0^3(\omega)$ es grande.
Dado $i \in \{0, 1, \ldots, g\}$, considere el siguiente evento

$$\begin{aligned} C_i^0 &= \{\omega : \text{la } n_i\text{-generación predecesora, la } n_i\text{-generación actual o la } n_i\text{-generación} \\ &\quad \text{sucesora para } \omega \text{ es de tamaño (cardinal) menor que } \varepsilon h_{n_i}\}. \end{aligned}$$

Observe que de la definición de C_i^0 se tiene

$$C_i^0 \subseteq \{\omega : |l_j - l_{j'}| < 2\varepsilon h_{n_i} \text{ para algún par } j \neq j' \text{ó } l_j < \varepsilon h_{n_i} \text{ para algún } j\}. \tag{4.6}$$

Dado j fijo, tenemos que

$$\mu(\{\omega : l_j < \varepsilon h_{n_i}\}) \approx \varepsilon,$$

así

$$\mu(\{\omega : l_j < \varepsilon h_{n_i} \text{ para algún } j\}) \approx (k+1)\varepsilon; \tag{4.7}$$

por otro lado, de (4.4) se tienes que

$$\mu(\{\omega : |l_j - l_{j'}| < \varepsilon h_{n_i} \text{ para algún } j, j'\}) < 2k(k+1)\varepsilon, \tag{4.8}$$

entonces, de (4.12), (4.7) y (4.8) se sigue que

$$\mu(C_i^0) < (2k+1)(k+1)\varepsilon. \tag{4.9}$$

Puesto que T es una transformación que preserva medida, se sigue que

$$\mu(C_i^q) < (2k+1)(k+1)\varepsilon, \tag{4.10}$$

donde $C_i^q = \{\omega : T^q\omega \in C_i^0\}$ para cualquier $q \in \mathbb{Z}$.
Considere la función $\mathfrak{C}_i : \Omega \to \mathbb{N}$ definida como

$$\mathfrak{C}_i(\omega) = \sum_{q=0}^{L-1} \mathbb{I}_{C_i^q}\omega = \#\{q :\ 0 \leq q \leq L-1 \text{ y } \omega \in C_i^q\}.$$

De (4.10), se tiene

$$\begin{aligned} E\left(\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{C}_i\right) &= \frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}E(\mathfrak{C}_i) \\ &\leq \frac{1}{g}\sum_{i=0}^{g-1}\mu(C_i^0) \\ &< (2k+1)(k+1)\varepsilon, \end{aligned}$$

así, de la desigualdad de Márkov (Teorema 2.4) se sigue que

$$\mu\left(\left\{\omega : \left[\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{C}_i(\omega)\right] > \sqrt{\varepsilon}\right\}\right) < (2k+1)(k+1)\sqrt{\varepsilon}. \tag{4.11}$$

Ahora, considere el evento

$$D_i^0 \;=\; \{\omega : \omega \text{ está en una } n_i\text{-generacion de tamaño menor que } \delta h_{n_i}\}.$$

Observe que de la definición de D_i^0 se tiene

$$D_i^0 \subseteq \{\omega : l_j < \delta h_{n_i} \text{ para algún } j\}. \tag{4.12}$$

Como

$$\mu(\{\omega : l_j < \delta h_{n_i} \text{ para algún } j\}) \approx (k+1)\delta,$$

entonces, definitivamente

$$\mu(\{\omega : l_j < \delta h_{n_i} \text{ para algún } j\}) < (k+2)\delta. \tag{4.13}$$

Considere la función $\mathfrak{D}_i : \Omega \to \mathbb{N}$ definida como

$$\mathfrak{D}_i(\omega) = \sum_{q=0}^{L-1}\mathbb{I}_{D_i^q}\omega = \#\{q :\ 0 \leq q \leq L-1 \text{ y } \omega \in D_i^q\},$$

donde $D_i^q = \{\omega : T^q\omega \in D_i^0\}$ para cualquier $q \in \mathbb{Z}$, entonces de (4.13)

$$E\left(\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{D}_i\right) < (k+2)\delta,$$

y así

$$\mu\left(\left\{\omega : \left[\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{D}_i(\omega)\right] > \sqrt{\delta}\right\}\right) < (k+2)\sqrt{\delta}. \tag{4.14}$$

Análogamente de (4.5), si definimos $\mathfrak{G}_i : \Omega \to \mathbb{N}$ como

$$\mathfrak{G}_i(\omega) = \sum_{q=0}^{L-1}\mathbb{I}_{G_i^q}\omega = \#\{q :\ 0 \leq q \leq L-1 \text{ y } \omega \in G_i^q\},$$

entonces

$$\mu\left(\left\{\omega : \left[\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{G}_i(\omega)\right] > \sqrt{\delta}\right\}\right) < 2\sqrt{\delta}. \tag{4.15}$$

Definición 4.2. Diremos que ω es **especial** si

$$\begin{aligned}\frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{C}_i(\omega) &\leq \sqrt{\varepsilon},\\ \frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{D}_i(\omega) &\leq \sqrt{\delta} \text{ y}\\ \frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}\mathfrak{G}_i(\omega) &\leq \sqrt{\delta}.\end{aligned}$$

De (4.11), (4.14) y (4.15) Se sigue que

$$\mu(\{\omega : \omega \text{ no es especial}\}) < (2k+1)(k+1)\sqrt{\varepsilon} + (k+4)\sqrt{\delta} < (2k^2+4k+5)\sqrt{\varepsilon}. \quad (4.16)$$

El siguiente lema establece que si ω es especial entonces su respectivo $J_0^3(\omega)$, que ya sabemos que es ω-bueno, es grande.

Lema 4.1. *Si ω es especial entonces $J_0^3(\omega)$ es grande.*

Demostración. Sea ω un punto fijo especial. Probaremos que J_0^3 es grande, probando que J_0^2 y J_0^1 son pequeños.

1. Probaremos que J_0^2 es pequeño.
   Sea $q \in \{0, 1, \ldots, L-1\}$, diremos que q es **añadido** a J_i^2 si $q \in J_i^2 \setminus J_{i+1}^2$, luego se tiene

   $$\begin{aligned}\bigcup_{i=0}^{g-1}(J_i^2 \setminus J_{i+1}^2) &= \bigcup_{i=0}^{g-1} J_i^2 \setminus \bigcap_{i=0}^{g-1} J_{i+1}^2\\ &= J_0^2 \setminus J_g^2\\ &= J_0^2,\end{aligned}$$

   entonces J_0^2 es exactamente la unión de los enteros añadidos a J_i^2 para algún $i \in \{0, 1, \ldots, g-1\}$.
   Ahora, si fijamos q en $\{0, 1, \ldots, L-1\}$ añadido a J_i^2, entonces

   a) Si q está en una n_i-generación de tamaño menor que δh_{n_i}, entonces se tiene que ω está en D_i^q. Así, el número de enteros añadidos a J_i^2 que están en una n_i-generación de tamaño menor que δh_{n_i} para algún $i \in \{0, \ldots, g-1\}$ es a lo más

   $$\sum_{i=0}^{g-1}\mathfrak{D}_i(\omega),$$

de (4.2) y desde que ω es especial, se tiene

$$\sum_{i=0}^{g-1} \mathfrak{D}_i(\omega) \leq \sqrt{\delta} g L = \varepsilon L. \tag{4.17}$$

b) Si q está entre n_i-generaciones entonces se tiene que ω está en G_i^q. Así, el número de enteros añadidos a J_i^2 que están entre n_i-generaciones para algún $i \in \{0, \ldots, g-1\}$ es a lo más

$$\sum_{i=0}^{g-1} \mathfrak{G}_i(\omega) \leq \varepsilon L. \tag{4.18}$$

c) El número de enteros añadidos a J_{g-1}^2 que están en las n_{g-1}-generaciones incompletas que intersectan a $\{0, 1, \ldots, L-1\}$, pero no están contenidas en $\{0, 1, \ldots, L-1\}$, es menor que $2h_{n_{g-1}}$, puesto que a lo más son dos. De (4.3) se tiene

$$2h_{n_{g-1}} < 2\delta^2 h_{n_g} = 2\delta^2 L.$$

Así, de los items anteriores tenemos

$$\# J_0^2 \leq (2\delta^2 + 2\varepsilon) L < 4\varepsilon L. \tag{4.19}$$

2. Probaremos que J_0^1 es pequeño.

Como ya mencionamos, para cualquier $i \in \{0, 1, \ldots, g-1\}$, J_{i+1}^1 consiste de n_{i+1}-generaciones, todas estas de tamaño mayor o igual que $\delta h_{n_{i+1}}$, ya que por definición las de tamaño menor que $\delta h_{n_{i+1}}$ están en J_{i+1}^2 ó J_{i+1}^3.

Sea R una n_{i+1}-generación fija, arbitraria, en J_{i+1}^1. Si $H_{0,R} = \# R$ entonces

$$H_{0,R} \geq \delta h_{n_{i+1}}. \tag{4.20}$$

Sea

$$H_{1,R} = \#\{r \in R : \ r \text{ no está en ninguna } n_i\text{-generación}\}.$$

Si ordenamos las n_i-generaciones, contenidas en R en pares

$$R_{1,1}, R_{1,2}, R_{2,1}, R_{2,2}, R_{3,1}, R_{3,2}, \ldots ;$$

y si en caso el número de n_i-generaciones contenidas en J es impar llamamos a la última R_0.

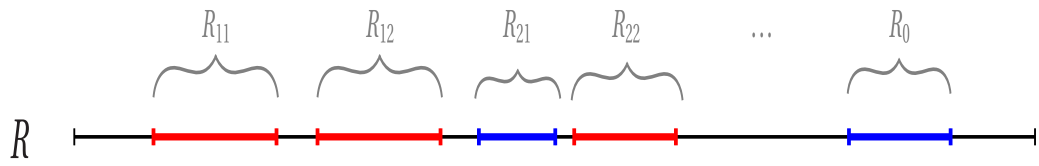

Figura 4.2:

Sea

$$H_{2,R} = \#R_0,$$

si el número de n_i-generaciones contenidas en R es par entonces $H_{2,R} = 0$. Evidentemente

$$H_{2,R} \leq h_{n_i}, \tag{4.21}$$

entonces

$$H_{0,R} - H_{1,R} - H_{2,R} = \# \bigcup_t (R_{t,1} \cup R_{t,2}).$$

Sean

$$\begin{aligned} H_{3,R} &= \#\{r \in R : \ r \in R_{t,1} \cup R_{t,2} \text{ tales que } R_{t,1} \text{ y } R_{t,2} \text{ son } \delta\text{-malos}\}, \\ H_{4,R} &= \#\{r \in R : \ r \in R_{t,1} \cup R_{t,2} \text{ tales que } R_{t,1} \text{ ó } R_{t,2} \text{ es de tamaño} \\ &\qquad \text{menor que } \varepsilon h_{n_i}\}. \end{aligned}$$

Observación 4.2. Note que $H_{3,R}$ y $H_{4,R}$ pueden contar a los mismos pares de generaciones, esto ocurrirá en el caso que para algún t, $R_{t,1}$ y $R_{t,2}$ sean δ-malos y alguno de ellos de tamaño menor que εh_{n_i}.

Si en R retiramos los enteros que están entre n_i-generaciones, retiramos R_0 si es que existe y retiramos los pares de generaciones en que ambos son δ-malos o en el que al menos uno de ellos es de tamaño menor que εh_{n_i} y llamamos a lo que resta E. Entonces

$$\#E \geq H_{0,R} - H_{1,R} - H_{2,R} - H_{3,R} - H_{4,R}, \tag{4.22}$$

donde no siempre se tiene la igualdad debido a la Observación 4.2.

Fijemos $R_{t,1}, R_{t,2} \subseteq E$, por lo tanto ambos de tamaño mayor o igual que εh_{n_i}, y sea η la permutación de $\{1,2\}$ tal que $R_{t,\eta(1)}$ es siempre una n_i-generación δ-bueno. Puesto que $R_{k,\eta(2)}$ es de tamaño menor que h_{n_i}, se sigue que

$$\frac{\#R_{t,\eta(1)}}{\#R_{t,1} + \#R_{t,2}} \geq \frac{\varepsilon}{1+\varepsilon}.$$

Puesto que, por definición de J_i^3, $R_{t,\eta(1)} \subseteq J_i^3$ y como esto se cumple para cada par de n_i-generaciones en E entonces

$$\#(R \cap J_i^3) \geq \frac{\varepsilon}{1+\varepsilon}(H_{0,R} - H_{1,R} - H_{2,R} - H_{3,R} - H_{4,R}).$$

Si ahora hacemos variar R entre las n_{i+1}-generaciones en J_{i+1}^1 tenemos

$$\#(J_{i+1}^1 \cap J_i^3) \geq \frac{\varepsilon}{1+\varepsilon}\sum_R (H_{0,R} - H_{1,R} - H_{2,R} - H_{3,R} - H_{4,R}). \tag{4.23}$$

Además

$$\#J_{i+1}^1 = \sum_R H_{0,R}. \tag{4.24}$$

Sea $r \in J_{i+1}^1$

a) Si r es contado por $H_{1,R}$ entonces se tiene que $T^r\omega$ está en $G_{i,1}$.

b) Si r es contado por $H_{3,R}$ entonces se tiene que $T^r\omega$ está en $G_{i,2}$ ó en $G_{i,3}$.

Así, si r es contado por $H_{1,R} + H_{3,R}$ entonces se tiene que ω está en G_i^r, se sigue que

$$\sum_R H_{1,R} + H_{3,R} \leq \mathfrak{G}_i(\omega). \tag{4.25}$$

De (4.3), (4.20) y (4.21), tenemos

$$H_{2,R} \leq h_{n_i} < \delta^2 h_{n_{i+1}} < \delta H_{0,R},$$

de esto y (4.24), se sigue que

$$\sum_R H_{2,R} < \delta \# J_{i+1}^1 \leq \delta L. \tag{4.26}$$

Si r es contado por $H_{4,R}$ entonces se tiene que ω está en C_i^r, se sigue que

$$\sum_R H_{4,R} \leq \mathfrak{C}_i(\omega). \tag{4.27}$$

De (4.23), (4.24), (4.25), (4.26) y (4.27) se sigue que

$$\#(J_{i+1}^1 \cap J_i^3) \geq \frac{\varepsilon}{1+\varepsilon}(\#J_{i+1}^1 - \mathfrak{G}_i(\omega) - \delta L - \mathfrak{C}_i(\omega)), \tag{4.28}$$

entonces

$$\#J_i^1 \leq \#J_{i+1}^1 - \frac{\varepsilon}{1+\varepsilon}(\#J_{i+1}^1 - (\mathfrak{G}_i(\omega) + \delta L + \mathfrak{C}_i(\omega))). \tag{4.29}$$

Dado que ω es especial

$$\begin{aligned} \frac{1}{g}\sum_{i=0}^{g-1}\frac{1}{L}(\mathfrak{G}_i(\omega) + \delta L + \mathfrak{C}_i(\omega)) &\leq \sqrt{\delta} + \delta + \sqrt{\varepsilon} \\ &< 3\sqrt{\varepsilon}, \end{aligned}$$

así, hay al menos $\frac{g}{2}$ (recuerde que g es par) valores de i, digamos $i_1 > i_2 >, \ldots, > i_{\frac{g}{2}}$, tales que para todo $t \in \{1, \ldots, \frac{g}{2}\}$, $0 \leq i_t \leq g-1$ y

$$\frac{1}{L}(\mathfrak{G}_{i_t}(\omega) + \delta L + \mathfrak{C}_{i_t}(\omega)) < 6\sqrt{\varepsilon}. \tag{4.30}$$

De (4.29) y (4.30)

$$\begin{aligned} \#J^1_{i_1} &\leq \#J^1_{i_1+1} - \frac{\varepsilon}{1+\varepsilon}(\#J^1_{i_1+1} - (\mathfrak{G}_{i_1}(\omega) + \delta L + \mathfrak{C}_{i_1}(\omega))) \\ &\leq \#J^1_{i_1+1} - \frac{\varepsilon}{1+\varepsilon}(\#J^1_{i_1+1} - 6\sqrt{\varepsilon}L) \\ &= \frac{1}{1+\varepsilon}\#J^1_{i_1+1} + \frac{\varepsilon}{1+\varepsilon}6\sqrt{\varepsilon}L \\ &\leq L\left(\frac{1}{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon}6\sqrt{\varepsilon}\right). \end{aligned} \tag{4.31}$$

Puesto que $\#J^1_{i_2+1} \leq \#J^1_{i_1}$, de (4.29), (4.30) y (4.31) se tiene

$$\begin{aligned} \#J^1_{i_2} &\leq \#J^1_{i_2+1} - \frac{\varepsilon}{1+\varepsilon}(\#J^1_{i_2+1} - 6\sqrt{\varepsilon}L) \\ &= \frac{1}{1+\varepsilon}\#J^1_{i_2+1} + \frac{\varepsilon}{1+\varepsilon}6\sqrt{\varepsilon}L \\ &\leq \frac{L}{1+\varepsilon}\left(\frac{1}{1+\varepsilon} + \frac{\varepsilon}{1+\varepsilon}6\sqrt{\varepsilon}\right) + \frac{\varepsilon}{1+\varepsilon}6\sqrt{\varepsilon}L \\ &= L\left(\left(\frac{1}{1+\varepsilon}\right)^2 + \frac{\varepsilon}{1+\varepsilon}\left(\frac{1}{1+\varepsilon} + 1\right)6\sqrt{\varepsilon}\right), \end{aligned}$$

si continuamos así por inducción y de (4.1) tenemos

$$\begin{aligned} \#J^1_{i_{\frac{g}{2}}} &\leq L\left(\left(\frac{1}{1+\varepsilon}\right)^{\frac{g}{2}} + \frac{\varepsilon}{1+\varepsilon}\left(\left(\frac{1}{1+\varepsilon}\right)^{\frac{g}{2}-1} + \left(\frac{1}{1+\varepsilon}\right)^{\frac{g}{2}-2} + \ldots + 1\right)6\sqrt{\varepsilon}\right) \\ &\leq L\left(\left(\frac{1}{1+\varepsilon}\right)^{\frac{g}{2}} + 6\sqrt{\varepsilon}\frac{\varepsilon}{1+\varepsilon}\sum_{t=0}^{\infty}\left(\frac{1}{1+\varepsilon}\right)^t\right) \\ &= L\left(\left(\frac{1}{1+\varepsilon}\right)^{\frac{g}{2}} + 6\sqrt{\varepsilon}\right) \\ &< L(\sqrt{\varepsilon} + 6\sqrt{\varepsilon}) \\ &< 7\sqrt{\varepsilon}L. \end{aligned}$$

Así

$$\#J^1_0 \leq J^1_{i_{\frac{g}{2}}} < 7\sqrt{\varepsilon}L. \tag{4.32}$$

Por lo tanto, de (4.19) y (4.32) se sigue que

$$\begin{aligned} \#J^3_0 &= L - \#J^1_0 - \#J^2_0 \\ &\geq L - 7\sqrt{\varepsilon}L - 4\varepsilon L \\ &\geq L(1 - 11\sqrt{\varepsilon}). \end{aligned} \tag{4.33}$$

Esto prueba que J_0^3 es un subconjunto de $\{0,1,\ldots,L-1\}$ grande y bueno para la mayoría de ω's. $\square$

Teorema 4.1. *Si T es una transformación de rango-uno k-mezclante entonces también es $(k+1)$-mezclante para todo $k \geq 2$.*

Demostración. Nuestro objetivo es probar que T es $(k+1)$-mezclante, suponiendo que ya es k-mezclante, esto será mostrando que la existencia de un subconjunto grande bueno en $\{0,1,\ldots,L-1\}$ para la mayoría de ω's implica que

$$\left|\mu\left(A_0 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1+\ldots+N_k)}A_k\right) - \prod_{j=0}^{k}\mu(A_j)\right|$$

es pequeño.

Desde que $J_0^3(\omega)$ es bueno, por definición se tiene

$$\left|\frac{1}{\#J_0^3}\sum_{i\in J_0^3} I(T^i\omega) - \prod_{j=0}^{k}\mu(A_j)\right| < \delta. \tag{4.34}$$

Además

$$\left|\frac{1}{L-\#J_0^3}\sum_{\substack{i=0\\ i\notin J_0^3}}^{L-1} I(T^i\omega) - \prod_{j=0}^{k}\mu(A_j)\right| < 1, \tag{4.35}$$

Esto último, por ser el valor absoluto de la diferencia de dos valores entre 0 y 1.

De (4.33), (4.34) y (4.35) se sigue

$$\begin{aligned}
\left|\frac{1}{L}\sum_{i=0}^{L-1} I(T^i\omega) - \prod_{j=0}^{k}\mu(A_j)\right| &\leq \frac{\#J_0^3}{L}\left|\frac{1}{\#J_0^3}\sum_{i\in J_0^3} I(T^i\omega) - \prod_{j=0}^{k}\mu(A_j)\right| \\
&\quad + \frac{L-\#J_0^3}{L}\left|\frac{1}{L-\#J_0^3}\sum_{\substack{i=0\\ i\notin J_0^3}}^{L-1} I(T^i\omega) - \prod_{j=0}^{k}\mu(A_j)\right| \\
&\leq \frac{\#J_0^3}{L}\delta + \frac{L-\#J_0^3}{L} \\
&< \delta + 11\sqrt{\varepsilon} < 12\sqrt{\varepsilon}.
\end{aligned} \tag{4.36}$$

Note que (4.36) solo es valido para ω especial.

Por otro lado, dado i arbitrario

$$\begin{aligned}
\mu\left(A_1 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1\ldots+N_k)}A_k\right) &= \mu\left(T^{-i}\left(A_1 \cap T^{-N_1}A_1 \cap \ldots \cap T^{-(N_1\ldots+N_k)}A_k\right)\right) \\
&= E\left(I \circ T^i\right) \\
&= E\left(\frac{1}{L}\sum_{i=0}^{L-1} I \circ T^i\right).
\end{aligned}$$

Así, de (4.16), (4.36) tenemos

$$
\begin{aligned}
&\left|\mu\left(A_1\cap T^{-N_1}A_1\cap\ldots\cap T^{-(N_1\ldots+N_k)}A_k\right)-\prod_{j=0}^{k}\mu(A_j)\right| \\
&\qquad = \left|E\left(\frac{1}{L}\sum_{i=0}^{L-1} I\circ T^i\right)-\prod_{j=0}^{k}\mu(A_j)\right| \\
&\qquad = \left|E\left(\frac{1}{L}\sum_{i=0}^{L-1} I\circ T^i-\prod_{j=0}^{k}\mu(A_j)\right)\right| \\
&\qquad \leq E\left(\left|\frac{1}{L}\sum_{i=0}^{L-1} I\circ T^i-\prod_{j=0}^{k}\mu(A_j)\right|\right) \\
&\qquad = E\left(\left|\frac{1}{L}\sum_{i=0}^{L-1} I\circ T^i-\prod_{j=0}^{k}\mu(A_j)\right|\cdot \mathbb{I}_{\{\omega:\omega\text{ especial}\}}\right) \\
&\qquad\quad +E\left(\left|\frac{1}{L}\sum_{i=0}^{L-1} I\circ T^i-\prod_{j=0}^{k}\mu(A_j)\right|\cdot \mathbb{I}_{\{\omega:\omega\text{ no especial}\}}\right) \\
&\qquad \leq 12\sqrt{\varepsilon}+\mu(\{\omega:\omega\text{ no es especial}\}) \\
&\qquad \leq 12\sqrt{\varepsilon}+(2k^2+4k+5)\sqrt{\varepsilon} \\
&\qquad = (2k^2+4k+17)\sqrt{\varepsilon}.
\end{aligned}
$$

□

4.2. Teorema principal

El resultado principal de esta tesis se presenta a continuación, la demostración queda como un corolario del Teorema 4.1

Teorema 4.2. *Si T es una transformación de rango-uno 2-mezclante entonces también es k-mezclante para todo $k > 2$.*

Demostración. Si la transformación de rango-uno es 2-mezclante entonces cumple las hipotesis del Teorema 4.1, para $k = 2$, luego se sigue que T es 3-mezclante. Ahora suponga que el enunciado del teorema es válido para un k cualquiera, probaremos que es valido para $(k + 1)$, en efecto esto se sigue de suponer que ya es válido para k y nuevamente aplicando el Teorema 4.1. □

4.3. Ejemplo de transformación mezclante de rango-uno

En 1972, Ornstein [15] presenta por primera vez un ejemplo de transformación mezclante de rango-uno. En esta sección se da a conocer una condición suficiente para que una transformación de rango-uno sea mezclante.
Recordemos de la definición de transformación de rango-uno que cada T es determinada por las sucesiones $\{r_n\}_{n\in\mathbb{N}} \subseteq \mathbb{N}$ y $\{s_{i,j}\}_{i\in\mathbb{N},j\in\{1,2,\ldots,r_i\}} \subseteq \mathbb{Z}_0^+$. Una transformación de rango-uno T es llamada **construcción escalera** si $s_{i,j} = j-1$, para todo $i \in \mathbb{N}$ y todo $1 \leq j \leq r_i$ (ver figura 4.3).

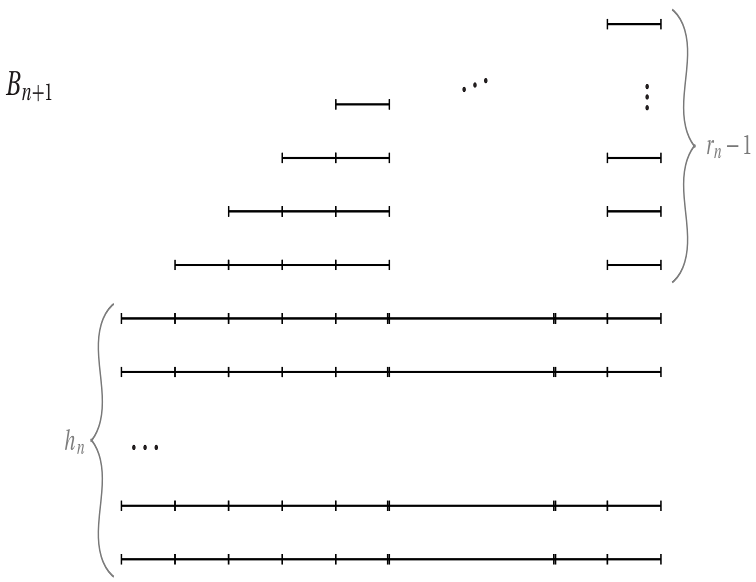

Figura 4.3: construcción Escalera

Observación 4.3. Sea T la construcción escalera dada por la sucesión $\{r_n\}$.

1. Si $\{r_n\}_{n\in\mathbb{N}}$ es una sucesión acotada, diremos que T es una *construcción escalera acotada.*

2. Si $r_n \to \infty$ diremos que T es una *construcción escalera* infinita

3. La *construcción escalera clásica* es dada por $r_n = n$ para todo $n \in \mathbb{N}$.

El teorema de Terrence Adams [1] presenta una condición suficiente para que una construcción escalera sea mezclante.

Teorema 4.3 (T. Adams). *Sea T la construcción escalera dada por la sucesión divergente $\{r_n\}_{n\in\mathbb{N}}$. Si*

$$\lim_{n\to\infty} \frac{r_n^2}{h_n} = 0,$$

donde h_n es la altura del n-bloque, entonces T es mezclante.

Ejemplo 4.1. La construcción escalera clásica es k–mezclante para todo $k \geq 2$. Esto se sigue desde que $r_n = n$ y $h_n = nh_{n-1} + \frac{(n-1)n}{2} \geq n!$, entonces

$$\lim_{n\to\infty} \frac{r_n^2}{h_n} \leq \lim_{n\to\infty} \frac{n^2}{n!} = 0.$$

4.4. Contra-ejemplo de Ledrappier

En 1978, F. Ledrappier en su artículo "*Un champ markovien peut être nulle et mélangeant*" prueba que 2-mezclante no implica 3-mezclante para acciones de $\mathbb{Z}^r$ con $r > 1$, las transformaciones son acciones de $\mathbb{Z}$, en esta sección introducimos algunos alcances del mencionado contra-ejemplo.
Sean $T_1, \ldots, T_r$ una familia de transformaciones conmutativa invertibles que preservan medida sobre un espacio de medida $(\Omega, \mathcal{A}, \mu)$, una $\mathbb{Z}^r$-**acción**, asociada a $T_1, \ldots, T_r$, es una aplicación $T : \mathbb{Z}^r \times \Omega \to \Omega$ definida por

$$T(n_1, \ldots, n_r, \omega) = T_1^{n_1} \circ \ldots \circ T_r^{n_r}\omega.$$

Denotemos $T^{\mathbf{n}} = T_1^{n_1} \circ \ldots \circ T_r^{n_r}$, donde $\mathbf{n} = (n_1, \ldots, n_r) \in \mathbb{Z}^r$. Así la $\mathbb{Z}^r$-acción queda denotada por

$$T^{\mathbf{n}}\omega = T(n_1, \ldots, n_r, \omega).$$

Sea $(T^{\mathbf{n}})$ una $\mathbb{Z}^r$-acción sobre un espacio de probabilidad $(\Omega, \mathcal{A}, \mu)$ y $k \geq 2$, decimos que $(T^{\mathbf{n}})$ es k-**mezclante** si para $A_1, \ldots, A_k$ conjuntos medibles se tiene

$$\mu\left(\bigcap_{i=1}^{k} T^{\mathbf{n}_i} A_i\right) \longrightarrow \prod_{i=1}^{k} \mu(A_i)$$

cuando $\mathbf{n}_i - \mathbf{n}_j \to \infty$ para todo $i \neq j$.

Denotemos por $\mathbb{F}_2 = \mathbb{Z}_2$ el cuerpo de dos elementos. Considere el conjunto $\mathbb{F}_2^{\mathbb{Z}^2}$ de todas las sucesiones dobles sobre $\mathbb{F}_2$. Equipado con la *topología producto* y la suma coordenada a coordenada, $\mathbb{F}_2^{\mathbb{Z}^2}$ forma un *grupo abeliano compacto*. La *medida de Haar* sobre $\mathbb{F}_2^{\mathbb{Z}^2}$ es el producto de medidas obtenido al tomar el *normalizador* de la medida de conteo de $\mathbb{F}_2$ (ver [11, pág. 57]). En $\mathbb{F}_2^{\mathbb{Z}^2}$ tenemos un shift hacia la izquierda σ y un shift hacia abajo τ, definidos de la siguiente manera

$$\begin{aligned}\sigma((\nu_{mn})_{m,n=-\infty}^{\infty}) &= (\nu_{(m+1)n})_{m,n=-\infty}^{\infty}\\ \tau((\nu_{mn})_{m,n=-\infty}^{\infty}) &= (\nu_{m(n+1)})_{m,n=-\infty}^{\infty}\end{aligned}$$

Note que, el conjunto

$$G = \{(\nu_{mn})_{m,n=-\infty}^{\infty} :\ \nu_{mn} + \nu_{(m+1)n} + \nu_{m(n+1)} = 0,\ (m,n) \in \mathbb{Z}^2\}$$

es un subrupo compacto de $\mathbb{F}_2^{\mathbb{Z}^2}$, σ y τ invariante.

François Ledrappier muestra en [10] que el par de automorfismos conmutativos σ y τ son mezclantes sobre el espacio de probabilidad $(G, \mathcal{B}, \mu)$ donde $\mathcal{B}$ es el cuerpo de Borel de G y μ es la medida de Haar normalizada en G y presenta un conjunto de medida positiva A para el cual

$$\mu(A \cap \sigma^{2^n} A \cap \tau^{2^n} A) \nrightarrow \mu(A)^3,$$

cuando $n \to \infty$. Lo que muestra que para acciones de $\mathbb{Z}^2$, 2-mezclante no implica 3-mezclante.

Capítulo 5

Conclusiones y recomendaciones

5.1. Conclusiones

En transformaciones de rango-uno se encuentran ejemplos de transformaciones que son ergódicas pero no son mezclantes, es por eso que en algún momento uno podría pensar que en esta clase de transformaciones se podría encontrar un ejemplo de una transformación k-mezclante que no sea $(k+1)$-mezclante, este trabajo descarta esa posibilidad pues se ha probado que la interrogante planteado por Rohlin, ¿Si 2-mezclante implica k-mezclante?, tiene respuesta afirmativa para el caso particular de transformaciones de rango-uno.

Para la prueba del teorema principal ha sido de gran ayuda el haber conseguido expresar a las transformaciones de rango-uno como un shift, el cual cumple propiedades que facilitan el desarrollo de este trabajo algo que no siempre es posible conseguir con todas las transformaciones.

5.2. Recomendaciones

En la Sección 3 del Capítulo 1 se ha dado una prueba del teorema ergódico de Birkhoff para el caso particular de que la transformación es ergódica, mencionado teorema se enuncia también para transformaciones que no necesariamente son ergódicas, el estudio de mencionada generalización será importante para trabajos posteriores , pues su aplicación se encuentra en diversos problemas del área.

Es conveniente recordar que en el presente trabajo se ha generalizado el método usado por Kalikow, entonces cabe preguntarse si este método también se puede usar para probar el mismo resultado para el caso de transformaciones de rango-q con $q > 1$, que es una generalización de la definición de transformación de rango-uno que se ha visto en este trabajo.

Otros aportes importantes respecto al tema son los hechos por Ryzhikov y Host, un trabajo posterior sera el revisar los métodos y resultados de mencionados trabajos.

Bibliografía

[1] T.M. Adams, *Smorodinsky's conjeture on rank-one mixing*, Proc. Amer. Math. Soc.**126** (1998), 739-744.

[2] G.D. Birkhoff & P.A. Smith, *Structure analysis of surface transformations*, J. Math. pures appl. **7** (1928), 345-379.

[3] G.D. Birkhoff *Proof of the Ergodic Theorem*, Proc. Nat.Acad. Sci., U.S.A. **17** (1931), 656-660.

[4] L. Arenas-Carmona, D. Berend and V. Bergelson, *Ledrappier's system is almost mixing of all orders*, Ergod. Th & Dynam. Sys.**28** (2008), 339-365.

[5] J.L. Doob, *Measure Theory*, Springer (1994).

[6] S. Ferenczi, *Systems of finite rank*, Colloq. Math. **73** (1997), no. 1, 35-65.

[7] B. V. Gnedenko, *Theory of Probability*, Sexta edición, Gordon and Breach Science Publishers.

[8] B. Host, *Mixing of all orders and pairwise independent joinings of systems with singular spectrum.* Israel J. Math. **76** (1991), no. 3, 289-298.

[9] S.A. Kalikow, *Twofold mixing implies threefold mixing for rank one transformations*, Ergod. Th & Dynam. Sys.**4** (1984), 237-259.

[10] F. Ledrappier, *Un champ markovien peut être nulle et mélangeant*, C. R. Acad. Sci. Paris **287** (1978), 561-563.

[11] R. Mañe, *Teoria Ergódica*, Projeto Euclides (1983).

[12] R. Metzger, *Curso Básico de Teoría de la Medida*, SMP (2008).

[13] M.G. Nadkarni, *Basic Ergodic Theory*, Segunda edición, Birkhäuser(1998)

[14] von Neumann J. *Proof of the Quasi-ergodic Hypothesis* Proc. Nat. Acad. Sci., U.S.A. **18** (1932), 70-82.

[15] D. Ornstein. *On the root problem in ergodic theory*, Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA, (1972) pp. 347-356.

[16] H. PoincarÃ©, *Les MÃ©thodes nouvelles de la MÃ©canique CÃ©leste*. Gauthier-Villars, Tome I. Paris. (1892). 8vo

[17] V.V. Ryzhikov, *Joinings and multiple mixing of the actions of finite rank*, Funct. Anal. Appl **27** (1993), no 2 63â78; translation in Funct. Anal. Appl. 27 (1993), no. 2, 128â140.

[18] V.A. Rohlin, *On endomirphisms of compact commutative groups*, Izv. Acad. Nauk. SSSR.**13** (1949), 329-340.

[19] W. Rudin, *Real and Complex Analysis*, Mc. Graw-Hill. Singapore (1987)

[20] C.E. Silva, *Invitation to Ergodic Theory*, Oxford (2007).