Universidad Nacional Autónoma de México

Licenciatura en Matemáticas
Facultad de Ciencias

Principios de Correspondencia y Teoría Ergódica

TESIS

que para obtener el grado académico de:

Matemático

presenta:

Ricardo Edmundo Gómez Uribe

Director de tesis:
Dr. Gerónimo Uribe Bravo

México, D.F. agosto del 2014

Principios de Correspondencia y Teoría Ergódica

Ricardo Edmundo Gómez Uribe

Agradecemos al proyecto PAPIIT IA101014, *Procesos Infinitamente Divisibles*, por el apoyo económico para la impresión de esta tesis.

A Deni,

a mis padres y a mi carnal;

a la sonora banda gangrena;

al Comité;

a quienes compartimos trecho.

A los que no se quedan en papeles.

Índice general

Prólogo

El trabajo que tienes en tus manos surgió de una doble intención: por un lado, quisimos desarrollar a partir de algunos ejemplos de teoría ergódica, lo que quiere decir un *principio de correspondencia*; por otro lado, quisimos desarrollar toda la teoría necesaria para entender con claridad el artículo *Deducing the Density Hales Jewett from an infinitary removal lemma* y exponerlo de manera que fuese comprensible para un estudiante de licenciatura que haya llevado algún curso de teoría de la medida. Ambos propósitos convergen allí donde el artículo utiliza un principio de correspondencia para replantear el problema en los términos adecuados para su resolución. Sin embargo, pudiera parecer al lector que se pierde el hilo conductor, especialmente en el capítulo 3, ya que la resolución del Teorema de Densidad de Hales Jewett mediante las técnicas propuestas requiere de una explicación bastante más extensa de la que sugieren las 23 páginas del artículo original, y más allá del planteamiento del problema, y en la reducción al caso fuertemente estacionario (en particular, la equivalencia entre estacionariedad fuerte y estacionariedad finitaria), no parece utilizarse el principio de correspondencia como recurso de construcción teórica. Estos dos recursos son, sin embargo, críticos para la demostración.

La cantidad de espacio necesario para explicar y justificar el contenido del artículo nos sorprendió incluso a nosotros, pero consideramos que, como primer acercamiento, bien vale la pena sacarle jugo a un ejemplo particularmente interesante, antes que desarrollar de manera más precisa una caracterización de lo que quiere decir *principio de correspondencia.* Admitimos, pues, que el rumbo de nuestro trabajo se vió alterado un tanto, pero no demasiado.

Por otro lado, el primer capítulo tiene el objetivo de introducir algunas nociones y resultado básicos de teoría ergódica, algunos de los cuales se utilizan en el capítulo 2. Decidimos incluír una demostración de la Ley Fuerte de los Grandes Números a partir del Teorema Ergódico de Birkhoff, que además cumple el propósito de poner en acción las nociones y resultados discutidos al principio del capítulo. Si bien la segunda parte del capítulo esta desligada del hilo conductor del resto de la tesis, consideramos importante incluírla porque, esperamos, puede servir de referencia en cursos de probabilidad y teoría de la medida, ya que plantea una demostración poco conocida (al menos poco difundida) de un teorema ampliamente utilizado que, además, en varios textos como [**JP03**], es demostrado solamente para casos particulares. La demostración de la LFGN y del Teorema Ergódico de Birkhoff pueden tomarse como un complemento y leerse de forma separada.

En lo que respecta a la *interpretación* del artículo sobre el Teorema de Densidad de Hales-Jewett en el capítulo 3, hemos incluído algunos ejemplos en la sección de estacionariedad fuerte que ayudan a visualizar mejor la propiedad. A partir de estos ejemplos nos surgieron un par de preguntas que no hemos resuelto y que dejamos abiertas en esa sección. Cabe decir que la sección de estacionariedad fuerte fue la más dificil de comprender, y al final terminamos modificando ligeramente el argumento del artíclo original, o quizás simplemente llenando los huecos. En el proceso de comprender cómo Tim Austin utiliza la estacionariedad fuerte realizamos algunos desarrollos equivocados, que sin embargo nos generaron preguntas sobre la relación entre estacionariedad fuerte y una versión más débil de esa propiedad, que llamamos *cuasi-estacionariedad finitaria*. Pensamos que la relación entre estacionariedad fuerte y cuasi-estacionariedad finitaria puede ser un tema interesante de investigación y planteamos la pregunta en su respectiva sección.

Por último, queremos advertir al lector que utilizamos un cambio de notación en el capítulo 3: en algún punto pasamos de $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}, \mathbb{P}_\omega\right)$ a $(X^\omega, \Sigma^\omega, \mathbb{P})$, no sin advertirlo también en el cuerpo del texto. Aunque nos pareció un tanto incómodo este cambio, pensamos que facilita la lectura la segunda notación y que, por otro lado, era conveniente mantener la primera notación para la primera parte del capítulo, donde se discuten los subespacios de $[k]^\omega$ y los espacios de probabilidad indexados por éstos. También, hemos procurado homologar hasta cierto punto la notación a lo largo de nuestros capítulos y secciones, pero preservando razgos que permitan al lector consultar las fuentes primarias sin tener que hacer demasiada labor de traducción. Por eso, por ejemplo, usamos Ω para designar espacio de probabilidad en los primeros capítulos y X en el tercero. Esperamos que la conveniencia de faciliar al lector la consulta sobrepase la extrañeza que pudiera generar.

Esperamos que la exposición sea clara y el contenido motivador.

Introducción

El Teorema de Densidad de Hales Jewett ha sido demostrado de diferentes maneras. La demostración más sencilla es también una de las más recientes, y surgió a partir de un esfuerzo colaborativo en línea denominado *Polymath* en el año 2009. Antes de eso en 1991 Furstenberg y Katznelson aportaron una demostración con técnicas de teoría ergódica para este teorema combinatorio. La demostración forma parte de un cuerpo de resultados que vinculan resultados combinatorios, o incluso de teoría de gráficas, con resultados de teoría ergódica. En el año 2011 Tim Austin aportó una nueva demostración del teorema de Densidad de Hales Jewett que, haciendo un planteamiento similar al de Furstenberg y Katznelson, ponía el peso de la demostración en la estructura de una clase de distribuciones llamadas *saciadas* sobre un espacio estándar de Borel producto, indexado por objetos combinatorios. Dedicamos el capítulo tercero de este trabajo a estudiar esa demostración.

Un razgo común a este cuerpo de resultados es el uso de principios de correspondencia. Un principio de correspondencia es un recurso de investigación que, partiendo de un enunciados sobre algún sistema finitario (una estructra matemática con una colección finita de objetos), desarrolla un enunciado equivalente sobre un sistema infinitario (una estructura análoga con una colección infinita de objetos); la equivalencia se establece por medio de un argumento de extensión, en el que se observa que el enunciado el válido para sistemas finitarios crecientes que, en algún sentido preciso según la estructura matemática de que se esté hablando, convergen al sistema infinitario en cuestión. Dichos sistemas pueden ser conjuntos de los naturales, espacios combinatorios, espacios de probabilidad indexados sobre espacios combinatorios, gráficas, o algún otro. En esta tesis estudiaremos algunos ejemplos de principios de correspondencia; en particular, estudiamos aquellos que nos permiten plantear la demostración de Tim Austin del Teorema de Densidad de Hales Jewett.

Los teoremas de Van der Wärden, Szemerédi, Hales Jewett, y su generalización de densidad son parte de un esfuerzo por estudiar la existencia de sucesiones aritméticas arbitrariamente largas en subconjuntos infinitos de los naturales. En particular, esta agenda de estudio está ligada a la cuestión sobre la existencia de sucesiones aritméticas arbitrariamente largas de números primos, hasta ahora sin resolver. El teorema de Szemerédi, que es la generalización de densidad del de Van der Wärden, fue conjeturado por Erdos y Turan en 1936 antes de que Szemerédi lo demostrara en 1975. Fue un teorema que generó mucho interés, entre otras cosas, porque parecía dar un paso a la resolución de esta pregunta. El uso de técnicas de teoría ergódica y principios de correspondencia es un campo de desarrollo que pudiera dar más pistas en la resolución de problemas pendientes.

En vista de lo anterior, el artículo que decidimos estudiar de Tim Austin resulta interesante, no porque aporte nuevos resultados, sino porque sugiere nuevos *métodos* prometedores de desarrollo.

Capítulo 1

Antecedentes de Teoría Ergodica

1. La Ley fuerte de los grandes numeros

La ley fuerte de los grandes números es uno de los resultados fundamentales de la probabilidad que formaliza la noción frecuentista de la probabilidad. El calificativo *fuerte* viene de que el teorema involucra convergencia casi segura en un espacio de probabilidad. Existe también una llamada ley débil que involucra solamente convergencia en probabilidad. Como sabemos, la convergencia casi segura implica la convergencia en probabilidad, así que la ley fuerte implica la débil. En este trabajo nos concentraremos en la prueba de la ley fuerte mediante argumentos de teoría ergódica.

A grandes rasgos, la ley fuerte de los grande números nos dice que al realizar repetidamente una medición en un expermiento aleatorio, el promedio de las mediciones se acerca a la esperanza matemática. El enunciado formal es el siguiente:

TEOREMA 1 (Ley fuerte de los grandes números). *Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas integrables y con esperanza común μ. Entonces*

$$\frac{X_1 + \cdots + X_n}{n} \to \mu$$

casi seguramente conforme $n \to \infty$.

La prueba usual de la ley fuerte de los grandes números se puede encontrar en diversos textos de probabilidad (por ejemplo [**JP03**, Ch. 20]). Hay dos caminos clásicos: el que utiliza el lema de Borel-Cantelli y el que utiliza la teoría de martingalas. Nuestro interés es exponer una demostracion ergódica de la ley fuerte de los grandes números.

2. Definiciones básicas de teoría ergódica

A lo largo de esta tesis trabajaremos con transformaciones que preservan la medida. A continuación se enuncian las definiciones básicas importantes. Trabajaremos en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ fijo.

Definición. Decimos que $T : \Omega \to \Omega$ es una **transformación que preserva a la medida** $\mathbb{P}$ si

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

para toda $A \in \mathscr{F}$.

Definición. Decimos que un conjunto $I \subset \Omega$ es **invariante** bajo la transformación T si

$$T^{-1}(I) = I.$$

La **σ-álgebra invariante bajo** T es

$$\mathscr{I} = \{I \in \mathscr{F} : I \text{ es invariante bajo } T\ \}.$$

La clase de conjuntos $\mathscr{I}$ es una σ-álgebra. En efecto, claramente contiene a $\emptyset$ y si $A \in \mathscr{I}$ entonces

$$T^{-1}(A^c) = A^c$$

por lo que $A^c \in \mathscr{I}$. Por otra parte, si $A_1, A_2, \ldots \in \mathscr{I}$, la igualdad

$$T^{-1}\left(\bigcup_n A_n\right) = \bigcup_n T^{-1}(A_n) = \bigcup_n A_n$$

nos muestra que $\mathscr{I}$ es estable bajo uniones numerables, lo cual muestra que $\mathscr{I}$ es σ-álgebra. Además, notemos que $\mathscr{I} \subset \mathscr{F}$.

Definición. Una variable aleatoria $Z : \Omega \to \mathbb{R}$ es **T-invariante** o **invariante con respecto a T** cuando

$$Z \circ T = Z.$$

Proposición 1. *Sea $\mathscr{I}$ la sigma álgebra invariante bajo T. Una variable aleatoria $\mathscr{F}$-medible $Z : \Omega \to \mathbb{R}$ es $\mathscr{I}$- medible si y sólo si es T-invariante.*

Demostración. Si Z es T-invariante,

$$Z^{-1}(-\infty, x) = T^{-1}\big(Z^{-1}(-\infty, x)\big) \qquad \forall x \in \mathbb{R}$$

por lo que

$$I_x = Z^{-1}(-\infty, x) \in \mathscr{I} \qquad \forall x \in \mathbb{R}.$$

Luego, si

$$\mathscr{A} := \{I_x : \quad x \in \mathbb{R}\}$$

entonces
$$\sigma(\mathscr{A}) \subset \mathscr{I}.$$
Pero
$$\mathscr{C} := \{(-\infty, x) : \quad x \in \mathbb{R}\}$$
genera a $\mathscr{B}_{\mathbb{R}}$; es decir que
$$\sigma(\mathscr{C}) = \mathscr{B}_{\mathbb{R}}.$$
Entonces
$$Z^{-1}(B) \in \sigma(\mathscr{A}) \subset \mathscr{I} \qquad \forall B \in \mathscr{B}_{\mathbb{R}}.$$
Con esto concluímos que Z es $\mathscr{I}$-medible.

Para el regreso, suponemos que Z es $\mathscr{I}$-medible. Un resultado conocido de teoría de la medida nos garantiza que existe una sucesión de funciones $\mathscr{I}$-simples
$$Z_n : \Omega \to \mathbb{R} \quad n \geq 1$$
que convergen puntualmente a Z. Cada de esas simples es combinación lineal de funciones
$$k \cdot \mathbf{1}_A, \qquad k \in \mathbb{R} \quad A \in \mathscr{I},$$
y claramente
$$k \cdot \mathbf{1}_A \circ T = k \cdot \mathbf{1}_{T^{-1}(A)} = k \cdot \mathbf{1}_A.$$
Luego,
$$Z_n \circ T = Z_n \qquad \forall n \geq 1.$$
Entonces
$$Z \circ T = \lim_{n\to\infty} Z_n \circ T = \lim_{n\to\infty} Z_n = Z.$$
□

Definición. Decimos que $T : \Omega \to \Omega$ es una **transformación ergódica** si T preserva la medida $\mathbb{P}$ y su σ-álgebra invariante es trivial en el sentido de que todos sus conjuntos tienen probablidad cero o uno.

Hay una liga útil entre transformaciones que preservan medida y un concepto importante dentro de la teoría de la probabilidad: la estacionariedad. Para enunciarla, requerimos definir lo que significa una sucesión aleatoria y cuál es su distribución.

Sean $(\eta_n)_{n\geq 1}$ variables aleatorias definidas en Ω. Notemos que para cada $\omega \in \Omega$, la sucesión $(\eta_n(\omega)\,, n \geq 1)$ pertenece al espacio (comunmente denominado **espacio canónico de sucesiones reales**)
$$\tilde{\Omega} = \mathbb{R}^{\mathbb{Z}_+}.$$
En $\tilde{\Omega}$ definimos a las proyecciones
$$X_i(x_1, x_2, \ldots) = x_i$$

y mediante estas funciones, generamos a la σ-álgebra

$$\tilde{\mathscr{F}} = \sigma(X_1, X_2, \ldots) .$$

Definamos $\eta : \Omega \to \tilde{\Omega}$ por

$$\eta(\omega) = (\eta_i(\omega) \,, i \geq 1)$$

y notemos que es medible. Sea $\tilde{\mathbb{P}}$ la medida imagen de $\mathbb{P}$ bajo η. A $\tilde{\mathbb{P}}$ le llamamos la distribución de la sucesión aleatoria η. Afortunadamente hay un criterio sencillo para decidir cuándo dos sucesiones tienen la misma distribución.

Proposición 2. *Dos sucesiones aleatorias* $\eta^1 = \left(\eta^1_1, \eta^1_2, \ldots\right)$ *y* $\eta^2 = \left(\eta^2_1, \eta^2_2, \ldots\right)$ *tienen la misma distribución si y sólo si para cada* $n \geq 1$ *se tiene la igualdad en distribución siguiente entre vectores aleatorios:*

$$\left(\eta^1_1, \ldots, \eta^1_n\right) \overset{d}{=} \left(\eta^2_1, \ldots, \eta^2_n\right) .$$

Demostración. ($\Leftarrow$) Consideramos dos espacios de probabilidad, posiblemente idénticos, $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$ y $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$. Sean $\tilde{\mathbb{P}}_1$ y $\tilde{\mathbb{P}}_2$ las medidas imágen de $\mathbb{P}_1$ y $\mathbb{P}_2$ bajo η^1 y η^2, respectivamente. Notamos que $\tilde{\mathbb{P}}_1$ y $\tilde{\mathbb{P}}_2$ están definidas sobre $\left(\tilde{\Omega}, \tilde{\mathscr{F}}\right)$. Para establecer el resultado, basta exhibir una clase $\mathscr{C} \subset \tilde{\mathscr{F}}$ cerrada bajo intersecciones finitas tal que $\sigma(\mathscr{C}) = \tilde{\mathscr{F}}$ y

$$\tilde{\mathbb{P}}_1(A) = \tilde{\mathbb{P}}_2(A) \quad \forall A \in \mathscr{C} .$$

Demostrando lo anterior, uno de los corolarios más usados del teorema de las clases monótonas nos garantiza que $\tilde{\mathbb{P}}_1 = \tilde{\mathbb{P}}_2$.

Consideramos pues

$$\mathscr{C} = \{A_1 \times A_2 \times \ldots \times A_n \times \mathbb{R} \times \mathbb{R} \times \ldots : A_1, A_2, \ldots, A_n \in \mathscr{B}_{\mathbb{R}}, n \in \mathbb{N}\}.$$

donde $\mathscr{B}_{\mathbb{R}}$ es la σ-álgebra de Borel sobre los reales. Por nuestra hipótesis, $\tilde{\mathbb{P}}_1$ y $\tilde{\mathbb{P}}_2$ coinciden en $\mathscr{C}$. En efecto, dada $A \in \mathscr{C}$ podemos escribir

$$A = A_1 \times A_2 \times \ldots \times A_n \times \mathbb{R} \times \ldots \qquad A_j \in \mathscr{B}_{\mathbb{R}}$$

de donde

$$\begin{aligned}
\tilde{\mathbb{P}}_1(A) &= \mathbb{P}_1\Big(\left(\eta^1\right)^{-1}(A)\Big) \\
&= \mathbb{P}_1\big(\{\omega \in \Omega_1 : \eta^1_1(\omega) \in A_1, \ldots, \eta^1_n(\omega) \in A_n\}\big) \\
&= \mathbb{P}_2\big(\{\omega \in \Omega_2 : \eta^2_1(\omega) \in A_1, \ldots, \eta^2_n(\omega) \in A_n\}\big) \\
&= \mathbb{P}_2\Big(\left(\eta^2\right)^{-1}(A)\Big) \\
&= \tilde{\mathbb{P}}_2(A) .
\end{aligned}$$

Por otra parte, es claro que $\mathscr{C}$ es cerrada bajo intersecciones finitas. Nos resta demostrar que

$$\sigma(\mathscr{C}) = \tilde{\mathscr{F}}.$$

Una contención es directa:

$$\sigma(\mathscr{C}) \subset \tilde{\mathscr{F}},$$

ya que $\mathscr{C}$ está contenido en $\tilde{\mathscr{F}}$, que es σ-álgebra.

Por otro lado, vemos que

$$\tilde{\mathscr{F}} = \bigvee_{j=1}^{\infty} \sigma(X_1, \ldots, X_n)$$

donde

$$\bigvee_{n=1}^{\infty} \sigma(X_1, \ldots, X_n) = \sigma\left(\bigcup_{n=1}^{\infty} \sigma(X_1, \ldots, X_n)\right).$$

Así, para verificar la contención inversa, basta demostrar que

$$\sigma(X_1, \ldots, X_n) \subset \sigma(\mathscr{C})$$

para toda $n \geq 1$.

Sea

$$\mathscr{C}_n := \{B_1 \cap \ldots \cap B_n : \quad B_1 \in \sigma(X_1), \ldots, B_n \in \sigma(X_n)\}.$$

Notamos que

$$\sigma(\mathscr{C}_n) = \sigma(X_1, \ldots, X_n).$$

Pues bien, si $B \in \mathscr{C}_n$, entonces es de la forma

$$B = B_1 \times \ldots \times B_n \times \mathbb{R} \times \ldots \qquad B_j \in \mathscr{B}_{\mathbb{R}}.$$

Es decir, $B \in \mathscr{C}$. Luego, $\mathscr{C}_n \subset \mathscr{C}$, de donde concluimos

$$\sigma(X_1, \ldots, X_n) \subset \mathscr{C}$$

y

$$\tilde{\mathscr{F}} \subset \sigma(\mathscr{C}).$$

($\Rightarrow$) Para obtener la otra implicación suponemos que $\tilde{\mathbb{P}}_1 = \tilde{\mathbb{P}}_2$ y observamos que para n fija y $B = B_1 \times \ldots \times B_n \times \mathbb{R} \times \ldots$ medible en $\sigma(X_1, \ldots, X_n)$,

$$\mathbb{P}_1\big(\{\omega \in \Omega_1 : \eta_1^1(\omega) \in B_1, \ldots, \eta_n^1(\omega) \in B_n\}\big) = \tilde{\mathbb{P}}_1(B).$$

Simétricamente,

$$\mathbb{P}_2\big(\{\omega \in \Omega_2 : \eta_1^2(\omega) \in B_1, \ldots, \eta_n^2(\omega) \in B_n\}\big) = \tilde{\mathbb{P}}_2(B).$$

Considerando que

$$\sigma(X_1, \ldots, X_n) \subset \tilde{\mathscr{F}},$$

tenemos el resultado deseado. $\square$

DEFINICIÓN. Sean $\eta_1, \eta_2, \ldots$ variables aleatorias en Ω. Decimos que la sucesión $(\eta_n, n \geq 1)$ es una **sucesión estacionaria** si $(\eta_i, i \geq 1)$ tiene la misma distribución que $(\eta_{1+i}, i \geq 1)$.

TEOREMA 2. *Si η es una sucesión estacionaria entonces $\eta = (\eta_1, \eta_2, \ldots)$ es de la forma $(f \circ T^n(\eta); n \geq 1)$, donde $f : \tilde{\Omega} \to \mathbb{R}$ es función medible y $T : \tilde{\Omega} \to \tilde{\Omega}$ preserva a $\tilde{\mathbb{P}}$.*

Por otro lado, si $T : \Omega \to \Omega$ preserva a $\tilde{\mathbb{P}}$, la sucesión $(f \circ T^n(\omega); n \geq 1)$ es estacionaria.

DEMOSTRACIÓN. Observamos que

$$X_1(\eta_n, \eta_{n+1}, \ldots) = \eta_n$$

y definimos el operador *shift* $T : \tilde{\Omega} \to \tilde{\Omega}$:

$$T(\eta_1, \eta_2, \ldots) = (\eta_2, \eta_3, \ldots) .$$

Por la estacionariedad de η, $T\eta \stackrel{d}{=} \eta$ y por inducción $T^n\eta \stackrel{d}{=} \eta$. Así,

$$\eta = (X_1 \circ T^n(\eta);\ n \geq 1) .$$

Notamos que, por construcción de $\tilde{\Omega}$, X_1 es medible, y tenemos el primer resultado.

Para la segunda afirmación, suponemos que $T : \Omega \to \Omega$ preserva a $\mathbb{P}$. Vemos que

$$\begin{aligned} \big(f \circ T^i(\omega); i = 1, \ldots, n\big) &\stackrel{d}{=} \big(f \circ T^i T(\omega); i = 1, \ldots, n\big) \\ &= \big(f \circ T^{i+1}(\omega); i = 1, \ldots, n\big) \end{aligned}$$

para todo $n \geq 1$. Aplicando la Proposición 2 concluímos que $(f \circ T^n(\omega); n \geq 1)$ es estacionaria. □

El teorema anterior es útil en tanto que nos permite construir sucesiones estacionarias a partir de colecciones de variables aleatorias y utilizar la transformación *shift* para establecer ciertas propiedades. Una manera de construir sucesiones estacionarias a partir de variables aleatorias es hacerlas a todas independientes: justo eso haremos para demostrar la ley fuerte de los grandes números.

3. El Teorema Ergódico de Birkhoff

El teorema ergódico de Birkhoff nos resulta de interés, primeramente, porque nos permite proveer una demostración ergódica de la ley fuerte de los grandes números, además de ser un resultado fundamental de la teoría ergódica.

TEOREMA 3 (Teorema ergódico de Birkhoff). *Sea* $(\Omega, \mathscr{F}, \mathbb{P})$ *un espacio de probabilidad y* $T : \Omega \to \Omega$ *una transformación medible que preserva a* $\mathbb{P}$. *Entonces para cualquier variable aleatoria integrable* X *la sucesión*

$$\left(\frac{1}{n} \sum_{j=1}^{n} X \circ T^j \right)_{n \geq 1}$$

converge casi seguramente conforme $n \to \infty$ *hacia* $\mathbb{E}(X \mid \mathscr{I})$, *donde* $\mathscr{I}$ *es la* σ*-álgebra invariante bajo* T.

Antes de demostrar el teorema, veamos cómo implica a la ley fuerte de los grandes números.

3.1. Una prueba ergódica de la ley fuerte de los grandes numeros. Sean $(\xi_n)_{n \geq 1}$ variables aleatorias independientes e identicamente distribuidas (i.i.d.) e integrables definidas en el mismo espacio de probabilidad Ω. Sea $\tilde{\Omega} = \mathbb{R}^{\mathbb{Z}_+}$. En $\tilde{\Omega}$ definimos a las funciones $X_i(x_1, x_2, \ldots) = x_i$ y mediante ellas, generamos a la σ-álgebra

$$\tilde{\mathscr{F}} = \sigma(X_1, X_2, \ldots) .$$

Definamos $\nu : \Omega \to \tilde{\Omega}$ por $\nu(\omega) = (\xi_i(\omega), i \geq 1)$ y notemos que es medible con respecto a la σ-álgebra de Ω. Sea $\tilde{\mathbb{P}}$ la medida imagen de $\mathbb{P}$ bajo ν. . Así, bajo $\tilde{\mathbb{P}}$, las variables aleatorias $X_1, X_2, \ldots$ son independientes y con la misma distribución que ξ_1. Definimos a $T : \tilde{\Omega} \to \tilde{\Omega}$ con

$$T(x_1, x_2, \ldots) = (x_2, x_3, \ldots) .$$

Notemos que

$$X_i \circ T = X_{i+1}$$

y por lo tanto,

$$\frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{n} \sum_{i=0}^{n-1} X_1 \circ T^i .$$

Notemos además que T preserva a $\tilde{\mathbb{P}}$ sobre $\tilde{\Omega}$. Puesto que X_1 es integrable, el teorema ergódigo de Birkhoff implica que

$$\frac{1}{n} \sum_{i=1}^{n} X_i \to \mathbb{E}(X_1 \mid \mathscr{I})$$

casi seguramente, donde $\mathscr{I}$ es la sub-σ-álgebra de $\tilde{\mathscr{F}}$ invariante bajo T:

$$\mathscr{I} = \left\{ A \in \mathscr{F} : \quad T^{-1}(A) = A \right\} .$$

Notemos que la σ-álgebra invariante está contendida dentro de la σ-álgebra de eventos remotos

$$\mathscr{T} = \bigcap_n \sigma(X_n, X_{n+1}, \ldots)\,.$$

puesto que, para cada $n \in \mathbb{N}$, si $A \in \mathscr{I}$, tenemos que

$$A = T^{-n}(A) \in \sigma(X_n, X_{n+1}, \ldots)\,.$$

La célebre ley 0-1 de Kolmogorov afirma que $\mathscr{T}$ es trivial en el sentido de que todos sus eventos tienen probabilidad cero o uno. Por lo tanto

$$\mathbb{E}(\, X_1 \mid \mathscr{I}) = \mathbb{E}(X_1)\,,$$

La validez de la LFGN en $\tilde{\Omega}$ implica la validez en cualquier otro espacio de probabilidad. En efecto, para cada $A \in \tilde{\mathscr{F}}$ se tiene la igualdad

$$\mathbb{P}\big(\nu^{-1}(A)\big) = \tilde{\mathbb{P}}(A)$$

y la aplicamos con

$$A = \left\{ \lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mathbb{E}(X_1) \right\}.$$

3.2. Demostración del teorema ergodico de Birkhoff. .

Con la motivación anterior sirviéndonos de pretexto, exponemos una demostración del teorema de Birkhoff desarrollada en [**KW82**]. Consideraremos $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad y T una transformación sobre Ω medible y que preserva la medida $\mathbb{P}$. No suponemos que T sea invertible.

Demostraremos el teorema de Birkhoff para funciones integrables no negativas, ya que se extiende fácilmente al caso general, observando que si X es función integrable,

$$X = X^+ - X^-$$

con $X^+ = \max\{0, X\}$ y $X^- = \max\{0, -X\}$, siendo X^+ y X^- integrables no negativas.

Demostración. Sea X integrable no negativa sobre Ω. Supondremos, por ahora, que $|X| \le M$ para alguna M, y definimos

$$\overline{X}(\omega) = \varlimsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} X(T^j\omega) \qquad \underline{X}(\omega) = \varliminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} X(T^j\omega)$$

Queremos mostrar que

$$\int \overline{X}\, d\mathbb{P} \le \int X\, d\mathbb{P} \le \int \underline{X}\, d\mathbb{P} \tag{1}$$

ya que ello nos garantiza la igualdad $\overline{X} = \underline{X}$ casi seguramente, y por lo tanto tendremos que $X^*(\omega) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} X(T^j\omega)$ existe casi seguramente, y que $\int X^*\, d\mathbb{P} = \int X\, d\mathbb{P}$. Ésta última equivalencia será importante para establecer que X^* es la esperanza condicional que requerimos. Además, obervamos que tanto $\overline{X}$ como $\underline{X}$ son T-invariantes, ya que

$$\begin{aligned}
\frac{1}{n}\sum_{j=0}^{n-1} X(T^jT\omega) &= \frac{1}{n}\sum_{j=1}^{n} X(T^j\omega) \\
&= \frac{1}{n}\sum_{j=0}^{n} X(T^j\omega) - \frac{X(\omega)}{n} \\
&= \left(\frac{n+1}{n}\right)\left(\frac{1}{n+1}\sum_{j=0}^{n} X(T^j\omega)\right) - \frac{X(\omega)}{n}
\end{aligned}$$

de donde

$$\begin{aligned}
\overline{X}(T\omega) &= \overline{\lim_{n\to\infty}}\, \frac{1}{n}\sum_{j=0}^{n-1} X(T^{j+1}\omega) \\
&= \overline{\lim_{n\to\infty}} \left(\frac{n+1}{n}\right)\left(\frac{1}{n+1}\sum_{j=0}^{n} X(T^j\omega)\right) - \frac{X(\omega)}{n} \\
&= \overline{\lim_{n\to\infty}} \left(\frac{1}{n+1}\sum_{j=0}^{n} X(T^j\omega)\right) \\
&= \overline{X}(\omega).
\end{aligned}$$

De la misma manera tenemos que $\underline{X}(T\omega) = \underline{X}(\omega)$, e inductivamente $\underline{X}(T^j\omega) = \underline{X}(\omega)$, $\overline{X}(T^j\omega) = \overline{X}(\omega)$ para $j \in \mathbb{N}$. Luego, y ya que $\overline{X} = \underline{X}$ casi seguramente, tendremos también que X^* es T-invariante casi seguramente.

Nos abocamos pues a establecer (1), empezando por la primera desigualdad. Sean $\epsilon > 0$ y $\omega \in \Omega$. Definimos

$$n(\omega) = \min\{m \in \mathbb{N} : \frac{1}{m}\sum_{j=0}^{m-1} X(T^j\omega) > \overline{X} - \epsilon\}.$$

En probabilidad es usual utilizar tiempos de paro; nuestra función $n(\omega)$ no es un tiempo de paro pues depende de $\overline{X}$.

Como $\overline{X}$ es T-invariante, tenemos que

$$\sum_{j=0}^{n(\omega)-1} \overline{X}(T^j\omega) = n(\omega)\overline{X}(\omega) \leq \sum_{j=0}^{n(\omega)-1} X(T^j\omega) + \epsilon \cdot n(\omega). \tag{2}$$

Como $n(\omega)$ es finito para toda $\omega \in \Omega$, entonces $\exists N \in \mathbb{N}$ tal que el conjunto $A = \{\omega : n(\omega) > N\}$ tenga medida menor a $\frac{\epsilon}{M}$.

Definimos

$$\begin{aligned} \tilde{n}(\omega) &:= n(\omega)\mathbf{1}_{A^c}(\omega) + \mathbf{1}_A(\omega) \\ \tilde{X} &:= X\mathbf{1}_{A^c} + M\mathbf{1}_A \end{aligned}$$

a partir de las cuales establecemos una versión de (2) con la cantidad de sumandos acotada por N:

$$\sum_{j=0}^{\tilde{n}(\omega)-1} \overline{X}(T^j\omega) \leq \sum_{j=0}^{\tilde{n}(\omega)-1} \tilde{X}(T^j\omega) + \epsilon \cdot \tilde{n}(\omega). \tag{3}$$

Observamos que cuando $\omega \in A$ la desigualdad (3) se reduce a la afirmación $\overline{X}(\omega) \leq M + \epsilon$, que resulta trivial puesto que X está acotada por M. El caso complementario en que $\omega \in A^c$ se reduce a la desigualdad (2) al utilizar la cota superior $X \leq \tilde{X}$.

Notando entonces que

$$\begin{aligned} \int \tilde{X}\, d\mathbb{P} &= \int_{A^c} X\, d\mathbb{P} + \int_A M\, d\mathbb{P} \\ &\leq \int X\, d\mathbb{P} + \epsilon, \end{aligned}$$

establecemos la desigualdad

$$\int X\, d\mathbb{P} \leq \int \tilde{X}\, d\mathbb{P} \leq \int X\, d\mathbb{P} + \epsilon; \tag{4}$$

recordemos que la construcción de $\tilde{X}$ depende de la elección de ϵ.

Definimos ahora inductivamente

$$n_0(\omega) = 0 \quad \text{y} \quad n_{k+1}(\omega) = n_k(\omega) + \tilde{n}\Big(T^{n_k(\omega)}(\omega)\Big)$$

y observamos que $Y_k(\omega) = n_k(\omega) - n_{k-1}(\omega) = \tilde{n}(T^{n_{k-1}}\omega)$ es un tiempo aleatorio acotado por N, y que

$$\sum_{j=n_{k-1}}^{n_k} \overline{X}(T^j\omega) \le \sum_{j=n_{k-1}}^{n_k} \tilde{X}(T^j\omega) + Y_k(\omega)\cdot\epsilon \qquad \forall k \in \mathbb{N} \tag{5}$$

por la desigualdad (3).

Entonces bien, si $L > NM/\epsilon$, se puede descomponer

$$\sum_{j=0}^{L-1} \overline{X}(T^j\omega) \quad = \quad \sum_{k=1}^{K(\omega)} \sum_{n_{k-1}(\omega)}^{n_k(\omega)-1} \overline{X}(T^j\omega) \quad + \quad \sum_{n_K(\omega)}^{L-1} \overline{X}(T^j\omega) \tag{6}$$

donde $K(\omega) = \text{máx}\{k : n_k(\omega) \le L-1\}$.

Sustituyendo la desigualdad (5) en (6) obtenemos

$$\sum_{j=0}^{L-1} \overline{X}(T^j\omega) \quad \le \quad \sum_{k=1}^{K(\omega)} \left(\sum_{n_{k-1}(\omega)}^{n_k(\omega)-1} \tilde{X}(T^j\omega) + Y_k(\omega)\cdot\epsilon \right) \quad + \quad \sum_{n_K(\omega)}^{L-1} \overline{X}(T^j\omega). \tag{7}$$

Notamos que

$$\sum_{k=1}^{K(\omega)} Y_k(\omega) = n_{K(\omega)} \le L-1 < L$$

y que

$$\sum_{n_K(\omega)}^{L-1} \overline{X}(T^j\omega) \le N\cdot M,$$

pues

$$(L-1) - n_{K(\omega)}(\omega) \le n_{K(\omega)+1}(\omega) - n_{K(\omega)}(\omega) \le N \quad \text{y} \quad |\overline{X}| \le M.$$

Considerando lo anterior, tenemos

$$\sum_{j=0}^{L-1} \overline{X}(T^j\omega) \quad \le \quad \sum_{j=0}^{n_{K(\omega)}} \tilde{X}(T^j\omega) \quad + \quad L\cdot\epsilon \quad + \quad N\cdot M$$

y usando la no-negatividad de X, y por lo tanto la de $\tilde{X}$ podemos añadir sumandos al lado derecho de (7) para obtener:

$$\sum_{j=0}^{L-1} \overline{X}(T^j\omega) \quad \le \quad \sum_{j=0}^{L-1} \tilde{X}(T^j\omega) \quad + \quad L\cdot\epsilon \quad + \quad N\cdot M \tag{8}$$

Integrando ambos lados de la desigualdad sobre Ω obtenemos

$$\sum_{j=0}^{L-1} \int \overline{X} \circ T^j \, d\mathbb{P} \quad \leq \quad \sum_{j=0}^{L-1} \int \tilde{X} \circ T^j \, d\mathbb{P} \quad + \quad L \cdot \epsilon \quad + \quad N \cdot M \tag{9}$$

A continuación utilizamos la invarianza de medida de T sobre $(\Omega, \mathscr{F}, \mathbb{P})$ para obtener

$$\int \tilde{X} \circ T^j \, d\mathbb{P} = \int \tilde{X} \, d\mathbb{P} \qquad \forall j \in \mathbb{N}.$$

Luego, sustituyendo éstas equivalencias en la desigualdad (9), utilizando la T-invarianza de $\overline{X}$, y dividiendo ambos lados de la desigualdad entre L, obtenemos

$$\int \overline{X} \, d\mathbb{P} \quad \leq \quad \int \tilde{X} \, d\mathbb{P} \quad + \quad \epsilon \quad + \quad \frac{N \cdot M}{L}.$$

Utilizando la desigualdad (4) y la elección de L obtenemos finalmente

$$\int \overline{X} \, d\mathbb{P} \quad \leq \quad \int X \, d\mathbb{P} \quad + \quad 3\epsilon$$

donde ϵ es arbitraria. Se concluye

$$\int \overline{X} \, d\mathbb{P} \quad \leq \quad \int X \, d\mathbb{P}$$

que es la primera desigualdad que queríamos establecer en (1).

Para la segunda desigualdad en (1), seguimos un prodecimiento análogo. Elegimos nuevamente $\epsilon > 0$ y definimos $n'(\omega)$ y A':

$$\begin{aligned} n'(\omega) \quad &:= \quad \text{mín}\{m \geq 1 : \quad \frac{1}{m} \sum_{j=0}^{m-1} X(T^j \omega) \leq \underline{X}(\omega) + \epsilon\} \\ A' \quad &:= \quad \{\omega \in \Omega : \quad n'(\omega) > N\} \end{aligned}$$

con N suficientemente grande para que $\int_{A'} X(\omega) \, d\mathbb{P}(\omega) < \epsilon$; redefinimos N para ello si es necesario, tal que cumpla tanto con lo anterior como con la desigualdad requerida para la primera parte - esto es, $\mathbb{P}(A) < \frac{\epsilon}{M}$.

Ahora definimos

$$\begin{aligned} \hat{n} \quad &:= \quad n' \mathbf{1}_{A'^c} \quad + \quad \mathbf{1}_{A'} \\ \hat{X} \quad &:= \quad X \mathbf{1}_{A'^c} \end{aligned}$$

para establecer la desigualdad

$$\sum_{j=0}^{\hat{n}(\omega)-1} \hat{X}(T^j \omega) \quad \leq \quad \sum_{j=0}^{\hat{n}(\omega)-1} \underline{X}(T^j \omega) + \hat{n}(\omega) \cdot \epsilon \tag{10}$$

que es la versión con sumandos acotados de

$$\sum_{j=0}^{n'(\omega)-1} \hat{X}(T^j\omega) \quad \leq \quad \sum_{j=0}^{n'(\omega)-1} \underline{X}(T^j\omega) + n'(\omega)\cdot\epsilon. \tag{11}$$

Como anteriormente, establecemos (10) a partir de considerar tres casos. En el primero, si $\{\omega, T\omega, T^2\omega, \ldots, T^{\hat{n}(\omega)-1}\omega\} \subset A'^c$, la desigualdad (10) se reduca al caso (11). En el segundo caso, si $\omega \in A'$, la desigualdad se reduce a la afirmación $0 \leq \underline{X}(\omega)+\epsilon$, que es trivialmente cierta ya que X es no negativa. En el tercer caso, $\omega \in A'^c$ pero $\{\omega, T\omega, T^2\omega, \ldots, T^{\hat{n}(\omega)-1}\} \not\subseteq A'^c$, donde una combinación de las desigualdades en los dos casos anteriores establece (10).

Ahora bien, tenemos

$$\int X\,d\mathbb{P} = \int_{A'^c} X\,d\mathbb{P} + \int_{A'} X\,d\mathbb{P} \;<\; \int_{A'^c} \hat{X}\,d\mathbb{P} + \epsilon \tag{12}$$

por construcción de A'.

Definimos ahora inductivamente, como antes

$$n'_0(\omega) = 0 \quad \text{y} \quad n'_{k+1}(\omega) = n'_k(\omega) + \hat{n}\Big(T^{n'_k(\omega)}(\omega)\Big).$$

y le llamamos Z_k al tiempo aleatorio acotado por N dado por por

$$Z_k(\omega) = n'_k(\omega) - n'_{k-1}(\omega).$$

Luego, usando una descomposición de sumandos similar a la de (6) y aplicando la desigualdad (10), tenemos que

$$\begin{aligned}
\sum_{j=0}^{n'_{K'(\omega)-1}} \underline{X}(T^j\omega) \quad &= \quad \sum_{k=1}^{K(\omega)} \sum_{n'_{k-1}(\omega)}^{n'_k(\omega)-1} \underline{X}(T^j\omega) \\
&\geq \sum_{k=1}^{K(\omega)} \left(\sum_{n'_{k-1}(\omega)}^{n'_k(\omega)-1} \hat{X}(T^j\omega) \;-\; Z_k(\omega)\cdot\epsilon \right) \\
&= \sum_{j=0}^{n'_{K'(\omega)}-1} \hat{X}(T^j\omega) \quad - \quad n'_{K'(\omega)}(\omega)\cdot\epsilon
\end{aligned}$$

donde $K'(\omega) = \max\{m : n'_m(\omega) \leq L-1\}$, con L como antes.

Integrando ambos lados de la desigualdad y dividiendo entre $n'_{K'(\omega)}$ obtenemos:

$$\int \underline{X}\, d\mathbb{P} \quad \geq \quad \int \hat{X}\, d\mathbb{P} \; - \; \epsilon \tag{13}$$

Ahora, substituyendo (12) en (13) obtenemos

$$\int \underline{X}\, d\mathbb{P} \quad \geq \quad \int X\, d\mathbb{P} \; - \; 2\epsilon \tag{14}$$

donde ϵ es aribitraria. Haciendo que $\epsilon \to 0$ obtenemos la segunda desigualdad de (1).

Hemos establecido pues (1) para funciones integrables no negativas acotadas:

$$\int \overline{X}\, d\mathbb{P} \leq \int X\, d\mathbb{P} \leq \int \underline{X}\, d\mathbb{P}.$$

Para extender la demostración a funciones no acotadas, definimos $\overline{X}_M = \text{mín}\{M, \overline{X}\}$ y $\underline{X}_M = \text{mín}\{\underline{X}, M\}$ para obtener las desigualdes respectivas en términos de $\overline{X}_M$ y $\underline{X}_M$. Haciendo que $M \to \infty$ se obtiene el resultado general.

Resta ahora demostrar que el límite $X^* = \text{lím}_{n\to\infty} \sum_{j=0}^{n-1} X(T^j\omega)$, que ya demostramos que existe casi seguramente, es en efecto la esperanza condicional que requerimos.

Hemos demostrado ya que X^* es T-invariante casi seguramente, ya que $\overline{X}$ y $\underline{X}$ son T-invariantes en todas partes y $\overline{X} = \underline{X} \quad c.s.$ Sabemos, entonces, que X^* está en el subespacio de Hilbert $\mathscr{L}^1(\Omega, \mathscr{I}, \mathbb{P}) \subset \mathscr{L}^1(\Omega, \mathscr{F}, \mathbb{P})$, donde $\mathscr{I}$ es la sub-σ-álgebra de $\mathscr{F}$ invariante bajo T. Basta pues verificar que $\mathbb{E}\,(X^*Z) = \mathbb{E}\,(XZ)$ para $Z \in \mathscr{L}^1(\Omega, \mathscr{I}, \mathbb{P})$ acotada (que Z sea acotada nos garantiza que las esperanzas existan):

$$\begin{aligned}
\mathbb{E}(XZ) &= \frac{1}{n}\sum_{j=0}^{n-1}\mathbb{E}(XZ) \\
&= \frac{1}{n}\sum_{j=0}^{n-1}\mathbb{E}\left(\left(X\circ T^j\right)\left(Z\circ T^j\right)\right) \\
&= \mathbb{E}\left(\frac{1}{n}\sum_{j=0}^{n-1}\left(X\circ T^j\right)\left(Z\circ T^j\right)\right) \\
&= \mathbb{E}\left(\left(\frac{1}{n}\sum_{j=0}^{n-1}X\circ T^j\right)Z\right)
\end{aligned}$$

ya que, por la Proposición 1, Z es $\mathscr{I}$-medible si y sólo si es T-invariante.

Consideramos ahora la sucesión

$$\left(\mathbb{E}\left(\left(\frac{1}{n}\sum_{j=0}^{n-1}X\circ T^j\right)Z\right);\quad n\geq 1\right)$$

que es constante. Haciendo que $n\to\infty$ obtenemos

$$\mathbb{E}\left(\left(\frac{1}{n}\sum_{j=0}^{n-1}X\circ T^j\right)Z\right)\ \to\ \mathbb{E}\left(X^*Z\right)$$

con lo que concluímos que X^* es, en efecto, la esperanza condicional deseada.

□

4. El teorema de retorno de Poincaré

Terminamos el capítulo estableciendo un resultado de recurrencia para sistemas de transformaciones que preservan medida, que nos será de utilidad más adelante. En efecto, en el capiítulo 2 demostraremos el teorema de Szmeredi a partir de una versión multidimensional del presente.

4.1. Teorema de retorno de Poincaré en sistemas con transformaciones de medida invariante. [Fur81], p. 8

TEOREMA 4 (Teorema de Retorno de Poncaré). *Sea T una transformación de medida invariante sobre el espacio de medida $(\Omega,\mathscr{F},\mu)$, con $\mu(\Omega)<\infty$. Si $B\in\mathscr{F}$ es cualquier conjunto de medida positiva, entonces hay un punto $x\in B$ y un natural $n\geq 1$ tales que $T^n(x)\in B$*

Demostración. Consideramos las imágenes inversas de B bajo potencias de T: $T^{-1}B$, $T^{-2}B$, $T^{-3}B$, $T^{-4}B, \ldots$, tenemos una secuencia infinita de conjuntos con la misma medida. Como $\mu(\Omega) < \infty$, no pueden ser todos disjuntos, así que existen naturales $i < j$ tales que $T^{-i}B \cap T^{-j}B \neq \emptyset$. Fijamos entonces $x = T^{i} \in B$ y tenemos que $T^{j-i}x \in B$. $\square$

Capítulo 2

Principios de Correspondencia en Teoría Ergódica

El principio de correspondencia es un recurso que nos permite convertir enunciados sobre sistemas dinámicos finitarios (esto es, sistemas dinámicos definidos sobre espacios finitos) en enunciados acerca de sistemas dinámicos infinitarios (definidos sobre espacios infinitos). Es, por así decirlo, una estrategia para hacer demostraciones deduciendo comportamientos cuantitativos locales a partir de afirmaciones cualitativas generales en sistemas dinámicos infinitos; la mejor manera de explicarlo es por medio de los ejemplos que veremos en este capítulo.

1. Demostracion del Teorema de Van der Wärden mediante un principio de correspondencia

1.1. Formulaciones finitaria e infinitaria. El teorema de Van der Wärden tiene varias formulaciones equivalentes. En particular, nos interesan sus formulaciones finitaria e infinitaria, para establecer su equivalencia.

TEOREMA 5 (Teorema de Van der Waerden, formulación finitaria [**Fur81**]). *Para $q, l \in \mathbb{N}$ fijos, existe $N(q, l)$ tal que, si los números $\{1, 2, \ldots, N(q, l)\}$ se particionan en q conjuntos, alguno de éstos tendrá progresiones aritméticas de longitud l.*

La versión *infinitaria* del teorema versa así:

TEOREMA 6 (Teorema de Van der Waerden, formulación infinitaria [**Fur81**]). *Si $\mathbb{Z}$ se particiona en q conjuntos, para todo l existe un conjunto B_j en la partición que contenga una sucesión aritmética monocromática de longitud l.*

Demostramos la equivalencia entre ambas formulaciones por medio de un principio de correspondencia. Para este fin, necesitamos el siguiente resultado:

PROPOSICIÓN 3. *La versión infinitaria del teorema de Van der Wärden enunciada sobre $\mathbb{Z}$ es equivalente a la misma enunciada sobre $\mathbb{N}$.*

Demostración. Contener sucesiones aritméticas de longitud l es una propiedad *local* - depende sólo de un subconjunto finito de $\mathbb{Z}$ (resp. $\mathbb{N}$) y no de todo $\mathbb{Z}$ (resp. $\mathbb{N}$). Pensaremos en cada q-partición como una q-coloración - es decir, una función

$$\mathcal{C} : \mathbb{F} \to [q]$$

donde

$$\mathbb{F} = \mathbb{Z}, \mathbb{N}$$

según sea el caso, y

$$[q] = \{1, 2, \ldots, q\}.$$

Diremos que una sucesión aritmética es monocromática, si sus elementos están contenidos en el mismo conjunto de la partición o, equivalentemente, si

$$\mathcal{C}(j) = constante$$

para cada j en la sucesión.

($\Leftarrow$) Si una q-coloración de $\mathbb{N}$ presenta alguna sucesión aritmética monocromática de longitud l, cualquier extensión arbitaria de la coloración sobre $\mathbb{Z}$ también la presenta.

($\Rightarrow$) Ahora suponemos que cualquier q-partición de $\mathbb{Z}$ contiene una sucesión aritmética monocromática de longitud l. Supongamos, para obtener una contradicción, que la restricción de esa partición sobre $\mathbb{N}$ no contiene ninguna sucesión aritmética monocromática de longitud l. Por consiguiente, tampoco contendrá ninguna sucesión aritmética monocromática de longitud $l+m$ para cualquier $m \geq 1$. Así, $-\mathbb{N} = \{\ldots, -1, 0\}$ tendrá sucesiones aritméticas monocromáticas de longitud arbitraria, y $\mathbb{N}$ tendrá sólo sucesiones aritméricas monocromáticas de longitud $s = 1, 2, \ldots, M$ para algún $M \leq l$. Pero entonces, si copiamos la q-partición de $\mathbb{N}$ sobre $-\mathbb{N}$ obtendremos una q-partición de $\mathbb{Z}$ que a lo más tendrá sucesiones aritmáticas monocromáticas de longitud $s = 1, 2, \ldots, 2M$, contradiciendo nuestra hipótesis. □

Ahora estamos en condiciones de demostrar la equivalencia:

Proposición 4. *Las versiones finitaria e infinitaria del teorema de Van der Wärden son equivalentes.*

Demostración. ($\Rightarrow$) La versión finitaria implica a la infinitaria de manera casi directa. En efecto, si particionamos a $\mathbb{Z}$ en q conjuntos

$$\mathbb{Z} = B_1 \cup \ldots \cup B_q$$

y elegimos un número l fijo, la versión finitaria del teorema nos garantiza que existe un natural N suficientemente grande tal que $B_j \cap \{1, 2, \ldots, N\}$ contenga una sucesión aritmética de longitud l, para algún $j = 1, \ldots, q$. Pero

l es arbitrario, y dejando que $l \to \infty$ obtenemos la versión infinitaria del teorema.

($\Leftarrow$) Por contrapositiva.

Suponemos la negación de la versión finitaria: dado q fijo, existe una sucesión creciente de naturales $(N_n; n \geq 1)$, $N_n \geq q$, con sus respectivas particiones en q conuntos:

$$(\mathcal{P}_n; n \geq 1) \qquad \mathcal{P}_n = \left\{B_1^n, \ldots, B_q^n\right\}$$

donde

$$\{1, \ldots, N_n\} = B_1^n \cup \ldots \cup B_q^n,$$

tales que ningún B_j^n contenga una sucesión aritmética de longitud l.

Por construcción cada partición tiene exactamente q conjuntos y podemos pensar en cada una de ellas como una q-coloración:

$$\mathcal{C}^n : [N_n] \to [q] \qquad \mathcal{C}_j^n := \mathcal{C}^n(j)\,,$$

donde

$$[N_n] = \{1, \ldots, N_n\}\,.$$

Así, la sucesión de particiones

$$(\mathcal{P}_n; n \geq 1)$$

corresponde a una sucesión de coloraciones

$$(\mathcal{C}^n; n \geq 1)\,.$$

Nos fijamos en el color del número 1 según la coloraciónes $\mathcal{C}^n$; es decir, nos fijamos en los colores $\mathcal{C}_1^n$. Dado que hay una cantidad finita de colores, existe un k_1 tal que

$$\mathcal{C}^n(1) = \mathcal{C}_1^{k_1}$$

para una infinidad de n. Entonces bien, podemos extraer una subsucesión

$$(\mathcal{C}^n; n \in \mathcal{A}_1) \subset (\mathcal{C}^n; n \geq 1)$$

con $\mathcal{A}_1 \subset \mathbb{N}$ tal que

$$\mathcal{C}^n(1) = \mathcal{C}_1^{k_1} \qquad \forall n \in \mathcal{A}_1.$$

De la misma manera existe $k_2 > k_1$ tal que

$$k_2 \in \mathcal{A}_1$$

y

$$\mathcal{C}^n(2) = \mathcal{C}_2^{k_2}$$

para una infinidad de n en $\mathcal{A}_1$. Siendo así, existe una segunda refinación del conjunto de índices

$$\mathcal{A}_2 \subset \mathcal{A}_1$$

que satisface

$$\mathcal{C}^n(2) = \mathcal{C}_2^{k_2} \qquad \forall n \in \mathcal{A}_2.$$

Es decir, el 2 tiene el color constante en la segunda subsucesión de coloraciones.

Procediendo recursivamente, construimos una cadena de refinamientos de conjuntos de índices

$$\ldots \mathcal{A}_m \subset \ldots \subset \mathcal{A}_2 \subset \mathcal{A}_1$$

con sus respectivos refinamientos de sucesiones de particiones

$$\ldots \subset (\mathcal{C}^n; n \in \mathcal{A}_m) \subset \ldots \subset (\mathcal{C}^n; n \in \mathcal{A}_1) \subset (\mathcal{C}^n; n \in \mathcal{A})$$

de tal manera que, para cada $m \geq 1$ se tiene que

$$\mathcal{C}^n(j) = \mathcal{C}_j^{k_m} \qquad \forall n \in \mathcal{A}_m \quad j = 1, \cdots, m.$$

Las constancias en las coloraciones de los números $1, \cdots, m$ se preservan tras la m-ésima refinación y todas las refinaciones subsecuentes.

Ahora, consideramos la subsucesión *diagonal*

$$\mathcal{D} = \left(\mathcal{C}^{k_m}; m \geq 1\right)$$

y planteamos la coloración límite sobre los naturales

$$\mathcal{C}^*$$

donde a cada natural j le corresponde el color $\mathcal{C}_j^{k_j}$. Le llamamos partición *límite*, puesto que si extendemos las coloraciones de manera arbitraira sobre los naturales, a modo que

$$\mathcal{C}^{k_m *} : \mathbb{N} \to [q] \qquad \mathcal{C}^{k_m *}|_{N_n} = \mathcal{C}^{k_m} \quad m \geq 1,$$

entonces para cada natural n la sucesión

$$\left(\mathcal{C}^{k_m}(n)\right)_m$$

converge a $\mathcal{C}_n^{k_n}$. Es decir,

$$\mathcal{C}^{k_m *} \to \mathcal{C}^*$$

puntualmente.

Alternativamente, podemos definir una métrica sobre el espacio de q-coloraciones de $\mathbb{N}$, verificando que efectivamente

$$\mathcal{C}^* = \lim_{m \to \infty} \mathcal{C}^{k_m *}$$

bajo esa métrica. Reservamos ese desarrollo para la demostración del Teorema 8, más adelante. Los detalles se pueden consultar en [**Fur81**], pgs. 6, 11.

Para terminar la demostración falta sólo evidenciar que $\mathcal{C}^*$ es una q-coloración de los naturales que no contiene sucesiones aritméticas monocromáticas de longitud l. Por construcción, ninguna coloración $\mathcal{C}^{k_m}$ en $\mathcal{D}$ las contiene. Supongamos que $\mathcal{C}^*$ presentara alguna sucesión aritmética monocromática de longitud l; ella estaría dentro de $[N_{k_m}]$ para m suficientemente grande. Pero $\mathcal{C}^*$ restringido a $[N_{k_m}]$ es simplemente $\mathcal{C}^{k_m}$, que no presenta sucesiones aritméticas monocromáticas de longitud l, de manera que la suposición es insostenible.

Retomando la Proposición 3 nuestro resultado es extensible a coloraciones sobre $\mathbb{Z}$. □

1.2. El teorema de Van der Wärden como recurrencia múltiple, versión topológica. Hay dos maneras de demostrar el teorema de Van der Wärden a partir de un teorema de recurrecia múltiple. La primera es construir un espacio métrico sobre el espacio de q-particiones de $\mathbb{Z}$ y así demostrar una formulación topológica del teorema de Van der Wärden a partir del teorema de recurrencia múltiple de Birkhoff, que trata con espacios métricos compactos. Este enfoque se desarrolla en [**Fur81**], pgs. 9-10.

Para enunciar el teorema de recurrencia múltiple que nos interesa necesitamos primero el siguiente concepto:

DEFINICIÓN. Si para un espacio métrico tenemos l transformaciones

$$T_1, \ldots, T_l : X \to X$$

diremos que un punto

$$x_0 \in X$$

es **conjuntamente recurrente** si existe una sucesiíon

$$(n_k)_k \subset \mathbb{N} \qquad n_k \to \infty$$

tal que

$$T_1^{n_k}(x_0) \to x_0, \ldots, T_l^{n_k}(x_0) \to x_0$$

conforme $k \to \infty$.

TEOREMA 7 (Teorema de recurrencia múltiple de Birkhoff [**Fur81**], p. 9). *Si X es un espacio métrico compacto y las transformaciones*

$$T_1, \ldots, T_l : X \to X$$

conmutan, entonces X tiene un punto conjuntamente recurrente.

No demostraremos aquí el teorema de recurrencia múltiple de Birkhoff, pero la demostración está disponible en el capítulo 2 de [**Fur81**]. Pero tampoco queremos pasar por alto la relación entre tan importante teorema topológico y el teorema de Van der Wärden. Empezamos exponiendo una versión más del teorema de Van der Wärden:

TEOREMA 8 (Van der Wärden, versión tolopógica). *Si X es un espacio métrico compacto, T un homeomorfismo de X, $x_0 \in X$, entonces para todo entero $l \geq 1$ y $\epsilon > 0$, hay un punto en la órbita de x_0, $y = T^m(x_0)$, tal que para algún $n \geq 1$, los puntos*

$$y, T^n(y)\,, T^{2n}(y)\,, \ldots, T^{ln}(y)$$

están a una distancia menor a ϵ entre sí.

Veamos cómo esta versión es equivalente a las que ya conocemos:

DEMOSTRACIÓN. ($\Rightarrow$) Usamos la versión infinitaria para establecer el resultado anterior.

Ya que X es compacto, damos una cubierta de bolas abiertas de radio $\frac{\epsilon}{2}$ y elegimos una subcubierta finita. Numeramos a las bolas abiertas de la subcubierta finita

$$X = B_1 \cup \ldots \cup B_q,$$

y establecemos una coloración de los naturales según la función

$$\mathcal{C} : \mathbb{N} \to [q]$$

definida por

$$\mathcal{C}(n) = \text{máx}\left\{j : T^n(x_0) \in B_j\right\}.$$

Por la versión infinitaria del teorema, esta coloración tiene una sucesión aritmética monocromática de longitud $l+1$, lo que nos dice que, para algunos m, n,

$$T^n(x_0)\,, T^{n+m}(x_0)\,, \ldots, T^{n+lm}(x_0) \in B_j \qquad j \in [q],$$

y por lo tanto distan menos de ϵ entre sí.

($\Leftarrow$) Para ver la implicación inversa consideramos un sistema dinámico sobre el espacio de sucesiones dobles compuestas de q símbolos

$$\Lambda^{\mathbb{Z}} \qquad \Lambda = [q].$$

Es decir, cada punto del espacio es una sucesión doble $(x(n)\,; n \in \mathbb{Z})$ donde

$$x : \mathbb{Z} \to \Lambda$$

es suprayectiva. Observamos que a cada punto del espacio corresponde una q-partición

$$\mathbb{Z} = \bigcup_{i=1}^{q} D_i$$

con

$$D_i = \{n :\, x(n) = i\}\,.$$

Definimos la siguiente métrica sobre el espacio:

$$d(x,y) = \text{ínf}\left\{\frac{1}{k+1} : x(i) = y(i) \quad \forall\, |i| < k\right\}.$$

Verifiquemos que d es, en efecto, una métrica:

a) La no negatividad es evidente al observar que d toma sus valores (una cantidad numerable de ellos) en el intervalo $[0,1]$

b) Computando directamente se observa que

$$d(x,x) = 0;$$

por otro lado, si

$$x \neq y$$

entonces existe

$$k = \text{máx}\,\{|n| :\ n \in \mathbb{Z}, \quad x(i) = y(i) \quad \forall |i| < |n|\}$$

finito, por lo que

$$d(x,y) > 0.$$

c) Para establecer la desigualdad del triángulo definimos

$$k_{x,y} = \text{máx}\,\{k \geq 1 : \quad x(i) = y(i), \quad |i| < k\}\,.$$

Para puntos x, y y z arbitrarios y distintos consideramos dos casos: si $k_{x,y} \leq k_{x,z}$ entonces

$$d(x,z) \leq d(x,y) \leq d(x,y) + d(y,z)\,.$$

Por otro lado si $k_{x,y} > k_{x,z}$ entonces $k_{y,z} \leq k_{x,z}$ y

$$d(x,z) \leq d(y,z) \leq d(x,y) + d(y,z)\,.$$

Con esto hemos demostrado que d es, en efecto, métrica. Notamos, además, que el espacio es compacto bajo la topología inducida por esa métrica. En efecto, sabemos que un espacio métrico es compacto si y sólo si toda sucesión acotada de puntos admite una subsucesión convergente bajo su métrica. Pues bien, sea

$$(x_j; j \geq 1)$$

una sucesión acotada de puntos en

$$\Lambda^{\mathbb{Z}}.$$

Es decir, existe una bola abierta

$$B_i = \{x : d(x,z) < \epsilon\}$$

alrededor de algún punto z, tal que

$$(x_j; j \geq 1) \subset B_i.$$

Si $\epsilon \geq 1$, entonces B_i es todo el espacio, así que debemos demostrar que toda sucesión arbitraria en el espacio admite una subsucesión convergente.

Para ello basta con reproducir el argumento que desarrollamos para demostrar que una sucesión arbitraria de coloraciones sobre $\mathbb{N}$ admite una subsucesión convergente, pero esta vez, a cada paso del refinamiento inductivo, nos aseguramos que la subsucesión de coloraciones obtenida sea constante sobre

$$\{-n, \ldots, 0 \ldots, n\}$$

en lugar de sólo sobre $[n]$.

Ya que estamos concientes de tener un espacio métrico compacto, definimos sobre este espacio el grupo de transformaciones compuesto por iteraciones del operador *shift*:

$$T(x(n)) = x(n+1)$$

con sus respectivos inversos. Para aplicar la versión topológica del teorema de Van der Wärden, resta convencernos de que las transformaciones del grupo son homeomorfismos. Es claro que cada una de ellas tiene inverso (por algo constituyen un grupo - y por algo construimos un espacio de sucesiones *dobles* y no sencillas, es decir, sobre $\mathbb{Z}$ y no sobre $\mathbb{N}$).

Para ver que cada transformación tiene inversa continua o, lo que es lo mismo, que cada transformación del grupo es continua, elegimos

$$B_\epsilon(z)$$

una bola abierta arbitraria de radio ϵ alrededor de algún punto z, y

$$T^u \qquad u \in \mathbb{Z}$$

una transformación del grupo. Ahora, por construción de la métrica,

$$B_\epsilon(z) = B_{\frac{1}{K+1}}(z)$$

para algún $K \geq 0$. Es decir, para cada punto x de la bola, los símbolos

$$\{x(-K+1)\,, \ldots, x(0)\,, \ldots, x(K-1)\}$$

son constantes. Observamos entonces que

$$T^{-u}(B_\epsilon(z))$$

consiste en todos los puntos y que conservan estos símbolos, pero trasladados $-u$ posiciones. Así pues,

$$T^u\big(B_{|K-1|+|u|}(y)\big) \subset B_\epsilon(z)\,,$$

con lo que afirmamos la continuidad de T^u.

Ya sin el menor atisbo de duda, podemos aplicar la versión topológica del teorema sobre nuestro sistema dinámico, para concluir que para cualquiera $x_0 \in \Lambda^{\mathbb{Z}}$, $l \geq 1$, $\epsilon > 0$, existe $y = T^m(x_0)$ tal que

$$y,\, T^n(y)\,,\, T^{2n}(y)\,,\, \ldots,\, T^{ln}(y)$$

estén a ϵ de distancia entre sí, para algún $n \geq 1$. En particular, si $\epsilon < 1$, esto quiere decir que

$$y(0) = y(n) = \ldots = y(ln)\,.$$

En otras palabras, hemos exhibido una sucesión aritmética monocromática de longitud arbitraria, para una una coloración arbitraria de $\mathbb{Z}$. $\square$

Esta equivalencia nos permite demostrar el siguiente resultado.

PROPOSICIÓN 5. *El teorema de recurrencia múltiple de Birkhoff implica el teorema de Van der Wärden.*

DEMOSTRACIÓN. [**Fur81**], p. 10. Demostramos la versión topológica del teorema de Van der Wärden.

Sean X un espacio métrico compacto, $x_0 \in X$, y Y la cerradura en X de la órbita de x_0:

$$Y = \overline{\{T^m(x_0) :\, m \in \mathbb{Z}\}};$$

notemos que Y es invariante bajo T, y es compacto por ser cerrado dentro de un compacto. Definimos

$$T_1 = T,\, T_2 = T^2,\, \ldots,\, T_l = T^l$$

y observamos que conmutan entre sí.

Por el teorema de Birkhoff de recurrencia múltiple, existen $y' \in Y$ y

$$(n_k)_k \subset \mathbb{N} \qquad n_k \to \infty$$

tales que

$$T_1^{n_k}(y') \to y', \ldots, T_l^{n_k}(y') \to y'$$

conforme $k \to \infty$. Entonces, para $N \in (n_k)_k$ suficientemente grande, los puntos

$$y',\, T_1^N(y')\,,\, T_2^N(y')\,,\, \ldots,\, T_l^N(y')$$

están a ϵ de distancia entre sí. Para y suficientemente cercano a y', el mismo resultado será válido para

$$y,\, T_1^N(y)\,,\, T_2^N(y)\,,\, \ldots,\, T_l^N(y)\,,$$

por continuidad de las T_i, y por tratarse de un número finito de puntos. Pero

$$y' \in \overline{\{T^m(x_0)\}_{m\in\mathbb{Z}}},$$

así que obtenemos el resultado deseado para algún y en la órbita de x_0. $\square$

1.3. Versión ergódica. El teorema de recurrencia múltiple de Birkhoff es para espacios métricos lo que es el teorema de recurrencia múltiple de Poincaré para espacios de medida:

TEOREMA 9 (Teorema de recurrencia múltiple de Poincaré [**Fur81**], pgs. 14, 77). *Si* $(X, \mathscr{F}, \mu)$ *es un espacio de medida y* $T_1, \ldots, T_l$ *son transformaciones sobre* X *que preservan a* μ *y que conmutan entre sí, entonces para cualquier conjunto* $A \in \mathscr{F}$ *tal que* $\mu(A) > 0$ *existe un entero* $n \geq 1$ *que satisface*

$$\mu\big(A \cap T_1^{-n}(A) \cap \ldots \cap T_l^{-n}(A)\big) > 0.$$

No demostraremos aquí tampoco el teorema de recurrencia múltiple de Poincaré, que nos desviaría de nuestros propósitos, pero el lector decepcionado por esta omisión puede referirse a [**Fur81**], capítulo 7.

Lo que nos interesa desarrollar es lo siguiente: la segunda manera de demostrar el teorema de Van der Wärden a partir de un teorema de recurrencia múltiple es observando que el teorema de recurrencia múltiple de Poincaré implica el teorema de Szemerédi, que a su vez implica el teorema de Van der Wärden. Estamos pues ante una demostración ergódica del teorema de Van der Wärden.

Aún más interesante para nosotros, veremos rumbo al final del presente capítulo cómo el teorema de Szemerédi se puede replantear a partir de un principio de correspondencia.

2. El Teorema de Szemerédi

El teorema de Szemerédi, en una de sus formulaciones, es la extensión de densidad del teorema de Van der Wärden. Quedará más claro qué queremos decir con esto después de enunciarlo:

TEOREMA 10 (Teorema de Szemerédi, versión finitaria [**Pol12**], p. 1). *Para cada entero* k *y para todo* $\delta > 0$ *existe* $N \in \mathbb{N}$ *tal que todo conjunto*

$$A \subset [N]$$

que satisface

$$\frac{|A|}{N} \geq \delta$$

contiene una sucesión aritmética de longitud k.

Decimos que $|A|/N$ es la densidad de A en $[N]$. No es difícil ver que el teorema de Van der Wärden es un corolario del teorema de Szemerédi. Para k y r fijos, elegimos

$$\delta \in \left(0, \frac{1}{r}\right)$$

y obtenemos N tal que todo subconjunto A de $[N]$ de densidad al menos δ contiene una sucesión aritmética de longitud k. El resultado se sigue de observar que toda r-coloración de $[N]$ tiene al menos un conjunto monocromático de densidad al menos N/r en $[N]$.

Como en el caso del teorema de Van der Wärden, el teorema de Szemerédi se puede enunciar sobre $\mathbb{N}$ o sobre $\mathbb{Z}$, y la equivalencia se establece por un procedimiento análogo.

2.1. Versiones finitaria e infinitaria. Aquí también tenemos una formulación finitaria y una infinitaria, equivalentes. Veremos, de hecho, que la demostración de su equivalencia es muy parecida al caso del teorema de Van der Wärden.

TEOREMA 11 (Teorema de Szemerédi, versión infinitaria [**Fur81**], p. 13). *Si $B \subset \mathbb{N}$ es tal que, para alguna sucesión de intervalos $([0, a_n]; n \geq 1)$ tales que $a_n \to \infty$ se cumple*

$$\frac{|B \cap [a_n]|}{a_n} \to d > 0, \qquad n \to \infty,$$

entonces B contiene progresiones aritméticas arbitrariamente largas.

En este caso B tiene densidad superior positiva en $\mathbb{N}$, en el sentido de la siguiente definición:

DEFINICIÓN. La **densidad superior** de un subconjunto de los naturales se define como

$$D^*(B) = \limsup_{N \to \infty} \frac{|B \cap [N]|}{N}.$$

PROPOSICIÓN 6. *Las versiones finitaria e infinitaria del teorema de Szemerédi son equivalentes.*

DEMOSTRACIÓN. ($\Rightarrow$) Sean k fijo y $A \subset \mathbb{N}$ con

$$D^*(A) = \delta.$$

Sea $N\left(\frac{\delta}{2}, k\right)$ el natural que por la versión finitaria del teorema garantiza que todo subconjunto de densidad al menos $\frac{\delta}{2}$ de $[N\left(\frac{\delta}{2}, k\right)]$ contiene una sucesión aritmética de longitud l. Para algún $n \geq N\left(\frac{\delta}{2}, k\right)$ tenemos

$$\frac{1}{n}|A \cap [n]| > \frac{\delta}{2}.$$

Luego, la versión finitaria nos garantiza la existencia de una sucesión aritmética de longitud l en

$$A \cap [n].$$

($\Leftarrow$) Procedemos por contrapositiva. Suponemos la negación de la versión finitaria y elegimos una sucesión $(n_j; j \geq 1) \subset \mathbb{N}$ tal que, para cada j, $[n_j]$ tiene un subconjunto A_j de densidad al menos δ y que no contiene ninguna sucesión aritmética de longitud l. Ahora, cada partición

$$[n_j] = A_j \cup ([n_j] \setminus A_j)$$

establece una 2-coloración, y la sucesión

$$(A_j; j \geq 1)$$

admite una subsucesión $\big(A_{k_j}; j \geq 1\big)$ para la cual la sucesión de funciones $\big(C^{k_j}; j \geq 1\big)$, definida por

$$C^{k_j} = \mathbf{1}_{A_{k_j}},$$

converge puntualmente a

$$C^*(l) := \lim_{j\to\infty} C^{k_j}(l)\,.$$

En efecto, basta utilizar el procedimiento diagonal que usamos para demostrar el Teorema 4. La subsucesión preserva la propiedad de densidad,

$$\frac{|A_{k_j}|}{n_{k_j}} \geq \delta \qquad \forall j,$$

y la propiedad de que ningún A_{k_j} contenga una sucesión aritmética de longitud l.

El subconjunto de $\mathbb{N}$ definido por C^*,

$$\{n \in \mathbb{N} : \quad C^*(n) = 1\} = \bigcup A_{k_j},$$

no contiene sucesiones aritméticas de longitud l, pues en ese caso lo haría también el conjunto definido por C^{k_j}, con j suficientemente grande. Para concluir, basta observar que

$$\frac{|A_{k_j}|}{k_j} \geq \delta \qquad \forall j$$

implica

$$D^*\Big(\bigcup A_{k_j}\Big) \geq \delta.$$

Hemos exhibido un conjunto en los naturales de densidad superior positiva sin sucesiones aritméticas de longitud l, contradiciendo la versión infinitaria del teorema. $\square$

2.2. Prueba del teorema de Szemerédi mediante un principio de correspondencia y el teorema de recurrencia multiple. Si bien a la fecha se han desarrollado ya varias demostraciones del teorema de Szemerédi, echando mano de diversas ramas de las matemáticas - análisis de Fourier, combinatoria, hipergráficas, entre otras -, nos ocuparemos en este trabajo sólo en la demostración ergódica. Para una breve pero ilustrativa discusión de las distintas rutas seguidas para demostrar el teorema de Szemerédi, referimos al lector a [**Tao06**], donde, además, se desarrolla una demostración a partir de la teoría ergódica distinta a la que trataremos a continuación.

Antes de entrar en materia, requeriremos de un resultado cuya prueba es desarrollada en [**Fur81**], capítulo 3, y que no desarrollaremos aquí.

Definición. A un conjunto finito de símbolos

$$\Lambda = \{a_1, \ldots, a_n\}$$

le llamamos un **alfabeto**. Lo denotamos usualmente con la letra Λ.

Definición. Si S es un subconjunto de $\mathbb{Z}$ (resp. $\mathbb{N}$), la **densidad superior de Banach** del conjunto S es

$$BD^*(S) = \limsup_{|I|\to\infty} \frac{|S \cap I|}{|I|}$$

donde I varía sobre intervalos de $\mathbb{Z}$ (resp. $\mathbb{N}$). Es decir, hay alguna secuencia de intervalos (I_k) con

$$|I_k| \to \infty$$

para los cuales

$$\frac{|S \cap I_k|}{|I_k|} \to BD^*(S)\,;$$

y para cualquier otra secuencia (I_n) con

$$|I_n| \to \infty$$

se verifica la desigualdad

$$\limsup \frac{|S \cap I_n|}{I_n} \leq BD^*(S)\,.$$

Teorema 12. *Consideramos el espacio $\Omega = \Lambda^{\mathbb{Z}}$, construído a partir de un alfabeto Λ, y una sucesión de símbolos $\xi \in \Omega$; sea*

$$X = \{T^n(\xi) : n \in \mathbb{Z}\}$$

la órbita de ξ con respecto al operador shift*, definido como anteriormente; es decir,*

$$T(\omega(n)) = \omega(n+1) \quad \omega \in \Omega.$$

Sean $a \in \Lambda$ *y*

$$A(a) = \{\omega : \omega(0) = a\}.$$

El símbolo a *ocurre en* ξ *con densidad superior de Banach positiva sí y sólo si existe una medida invariante* μ *sobre* X *tal que* $\mu(A(a)) > 0$*.*

Mencionamos, por interés general, que la demostración del teorema depende de un resultado general sobre espacios compactos y convexos, conocido como el teorema de Krein-Milman, aplicado en este caso a un espacio de medidas T-invariantes sobre X.

Para demostrar el teorema de Szemerédi, se necesita poco más que ligar el resultado anterior con el teorema de recurrencia múltiple de Poincaré:

Demostración ergódica del teorema de Szemerédi. Supongamos que $S \subset \mathbb{Z}$ tiene densidad superior de Banach positiva, y definimos

$$\Lambda = \{0, 1\} \qquad \Omega = \{0, 1\}^{\mathbb{Z}}.$$

El punto $1_S \in \Omega$ está definido por la función característica de S sobre $\mathbb{Z}$. Sea $X \subset \Omega$ la cerradura de la órbita de 1_S bajo el *shift*, y sea

$$A = A(1) = \{\omega : \omega(0) = 1\}.$$

Por el Teorema 12, hay una medida μ sobre X invariante bajo el *shift* y tal que

$$\mu(A(1)) > 0.$$

Entonces bien, se verifican las hipótesis del teorema de recurrencia múltiple de Poincaré (Teorema 9) con

$$T_1 = T, \quad T_2 = T^2, \quad \ldots \quad T_l = T^l$$

donde T es el operador *shift*. Así, tenemos que para algún $n \geq 1$

$$\mu\big(A \cap T^{-n}(A) \cap T^{-2n}(A) \cap \ldots \cap T^{-ln}(A)\big) > 0,$$

lo que implica que hay algún punto $\omega \in A \cap X$ tal que

$$T^{in}(\omega) \in A \cap X \qquad i = 1, 2, \ldots, l.$$

Entonces, por definición de $A = A(1)$,

$$\omega(0) = \omega(n) = \omega(2n) = \ldots = \omega(ln) = 1.$$

Además, como $\omega \in X$ es un límite de traslaciones de ξ, está arbitrariamente cerca de una traslación de ξ o, lo que es lo mismo, equivale a una traslación de ξ en un intervalo de $\mathbb{Z}$ arbitrariamente grande alrededor del 0. Así, para algún h,

$$1_S(h) = 1_S(h+n) = 1_S(h+2n) = \ldots = 1_S(h+ln) = 1.$$

Por definición de 1_S, esto signnifica que

$$h,\, h+n,\, h+2n,\, \ldots,\, h+ln\ \in S.$$

$\square$

3. El teorema de densidad de Hales-Jewett

3.1. Una demostracion alternativa del teorema de Szemerédi. El lector que sienta que nos hemos saltado demasiados pasos para llegar a la demostración ergódica del teorema de Szemerédi -no sin razón - será compensado en el capítulo siguiente, con otra demostración ergódica, que sí estudiaremos de manera completa, y que pasa por el teorema de densidad de Hales-Jewett, y cuyo último paso exponemos aquí: es decir, la demostración de que éste implica el teorema de Szemerédi.

El teorema de densidad de Hales-Jewett generaliza el teorema de Hales-Jewett de la misma manera que el teorema de Szemerédi generaliza el teorema de Van der Wärden. En lo que sigue, denotaremos por

$$[k]^n$$

al conjunto de secuencias de símbolos de longitud n que toman sus valores en

$$[k] := \{1, 2, \ldots, k\}\,.$$

A cada elemento de $[k]^n$ le llamaremos **palabra**.

Definición. Una **línea combinatoria** de $[k]^n$ es un conjunto

$$A \subset [k]^n$$

de la forma

$$l([k]) = \{l(i) : i \in [k]\}$$

donde

$$l : [k] \to [k]^n$$

se define sustituyendo en

$$w_l \in ([k] \cup *)^n \smallsetminus [k]^n$$

la letra respectiva por cada ocurrencia de $*$. Requerimos, pues, que w_l contenga de hecho al menos un $*$.

Definición. Para una línea combinatoria en $[k]^n$ que corresponde a $w_l \in ([k] \cup *)^n \smallsetminus [k]^n$, el conjunto de posiciones donde aparece $*$ se llama el **conjunto comodín**. Es decir, es el conjunto

$$\{m \in [n] : w_l(m) = *\}\,.$$

Ahora estamos en condiciones de enunciar el primer teorema.

TEOREMA 13 (Teorema de Hales-Jewett). *Para todo par de enteros $k \geq 2$ y $r \geq 2$ existe $N \geq 0$ tal que para todo $n \geq N$ se cumple que toda r-coloración de $[k]^n$ contiene una línea combinatoria monocromática.*

Su generalización de densidad es la siguiente:

TEOREMA 14. *[Teorema de densidad de Hales-Jewett, versión finitaria] Para cada entero positivo k y cada real $\delta > 0$, existe $N \geq 0$ tal que para todo $n \geq N$, todo subconjunto de $[k]^n$ con densidad al menos δ contiene una línea combinatoria.*

La demostración de que la versión de densidad de Hales-Jewett implica el primer teorema es exactamente análoga a la que ya hicimos para establecer que el teorema de Szemerédi implica el de Van der Wärden.

Por otro lado nos resulta necasario establecer el siguiente resultado:

PROPOSICIÓN 7. *El teorema de densidad de Hales-Jewett implica el teorema de Szemerédi.*

DEMOSTRACIÓN. Fijamos k, entero positivo, y $\delta > 0$. Por el teorema de densidad de Hales-Jewett, existe un entero $N > 0$ tal que para todo $n \geq N$, todo subconjunto

$$A \subset [k]^n$$

de densidad al menos δ contiene una línea combinatoria. En particular, consideramos

$$A \subset [k]^N$$

tal que

$$\frac{|A|}{k^N} \geq \delta$$

y observamos que contiene una línea combinatoria: le llamaremos

$$L = \{l(i) : i \in [k]\} .$$

Ahora bien, definimos la función

$$\phi : [k]^N \to [k^N]$$

que manda a cada palabra en $[k]^N$ al número que representa escrita en base k, interpretando a $[k]$ como

$$[k] := \{0, 1, \ldots, k-1\}$$

y a $[k^N]$ como

$$[k^N] := \left\{0, 1, \ldots, k^N - 1\right\} .$$

Es decir,

$$\phi((a_1, a_2, \dots, a_N)) = \sum_{j=0}^{N-1} a_j \cdot k^j \qquad a_j \in [k].$$

Claramente, ϕ es biyectiva.

Notamos que $\phi(L)$ es una sucesión algebráica en $[k^N]$ contenida en $\phi(A)$. En efecto, si S es el conjunto comodín de L, entonces

$$\phi(l(i)) - \phi(l(i-1)) = \sum_{j \in S} k^j \qquad i = 2, \dots, k.$$

Nótese que en general hay muchas más sucesiones aritméticas que las que se pueden expresar de esta manera, por lo que es más difícil encontrar líneas combinatorias de lo que es encontrar sucesiones aritméticas.

Siendo biyectiva, la función ϕ preserva densidades. Así, cualquier conjunto A' en $[k^N]$ de densidad al menos δ tiene una sucesión aritmética de longitud k, que corresponde a una línea combinatoria en

$$\phi^{-1}(A')\,.$$

Por último, hacemos la observación de que, para cualquier $n > N$ se puede repetir el procedimiento anterior, eligiendo un conjunto de densidad al menos δ en $[k^n]$ para exhibir una sucesión aritmética de longitud k contenida en él. $\square$

Con el mismo procedimiento se puede verificar:

PROPOSICIÓN 8. *El teorema de Hales-Jewett implica el teorema de Van der Wärden.*

Omitimos la demostración pues resultaría repetitiva.

3.2. Versiones finitaria e infinitara. También para el teorema de densidad de Hales-Jewett hay una versión finitaria y una infinitaria. Ya hemos enunciado la formulación finitaria.

Para armonizar nuestra notación con la de [**Aus11**], a quien habremos de referirnos más adelante, denotamos por ω a un conjunto de la cardinalidad de los naturales.

DEFINICIÓN. El **espacio combinatorio infinito-dimensional** sobre $[k]$ es

$$[k]^\omega := \bigcup_{n \geq 1} [k]^n.$$

TEOREMA 15 (Teorema de densidad de Hales-Jewett, formulación infinitaria). *Si*

$$A \subset [k]^\omega,$$

y

$$\limsup_{n\to\infty} \frac{|A \cap [k]^n|}{k^n} > 0,$$

entonces A contiene una línea combinatoria.

Como bien sospechará el lector a estas alturas, las versiones finitaria e infinitaria del teorema son equivalentes, y la demostración sigue el mismo principio que hemos usado para los teoremas de Szemerédi y Van der Wärden:

DEMOSTRACIÓN. ($\Rightarrow$) Suponemos la versión finitaria del teorema, y elegimos un conjunto

$$A \subset [k]^\omega$$

tal que

$$\limsup_{n\to\infty} \frac{|A \cap [k]^n|}{k^n} = \epsilon \qquad \epsilon > \delta > 0.$$

Por la versión finitaria, existe un entero positivo N tal que para toda $n \geq N$, cualquier conjunto

$$B \subset [k]^n \quad tal \quad que \quad \frac{|B|}{k^n} > \delta$$

contiene una línea combinatoria.

Pero por definición de límite superior podemos encontrar M arbitrariamente grande tal que

$$\frac{|A \cap [k]^M|}{k^M} > \delta,$$

por lo que, eligiendo M mayor a N,

$$A \cap [k]^M \subset A$$

contiene una línea combinatoria.

($\Leftarrow$) Procedemos por contrapositiva. Suponemos la negación de la versión finitaria del teorema y elegimos una sucesión $(n_j : j \geq 1)$ junto con una sucesión de conjuntos

$$(A_j : j \geq 1) \qquad A_j \subset [k]^j$$

tal que

$$\frac{|A_j|}{k^j} \geq \delta$$

y ningún A_j contenga líneas combinatorias.

Consideramos el conjunto

$$A^* = \bigcup_{j\geq 1} A_j \quad \subset [k]^\omega.$$

Por construcción, A^* no contiene líneas combinatorias, y además

$$\limsup_{n\to\infty} \frac{|A^* \cap [k]^n|}{k^n} \geq \delta > 0,$$

contradiciendo la versión infinitaria del teorema. □

Habrá notado el lector un toque de familiaridad entre esta demostración y las que desarrollamos para establecer las equivalencias entre las distintas formulaciones de los teoremas de Szemerédi y Van der Wärden, con la singularidad de que esta vez no requerimos de ningún truco de diagonalización. Esto se debe a que, a diferencia de los casos anteriores, no requerimos que exista algún tipo de proyección de las A_j sobre $[k]^{j-1}$ que sea equivalente a A_{j-1}.

3.3. Otra formulación del teorema de densidad de Hales-Jewett. Para el teorema de densidad de Hales-Jewett, existe otra formulación finitaria, con su respectiva equivalencia infinitaria, que es más cercana a la que demostraremos en el siguiente capítulo.

Teorema 16. *[Densidad de Hales-Jewett, versión ergódica finitaria] Para todo $\delta > 0$ existe $N \geq 1$ tal que, si $n \geq N$ y*

$$f : [k]^n \to \mathscr{B}$$

es una función hacia la σ-álgebra de algún espacio de probabilidad $(X, \mathscr{B}, \mathbb{P})$, con

$$f(w) = B_w \quad \in \mathscr{B},$$

y tal que

$$\mathbb{P}(B_w) \geq \delta \qquad \forall w \in [k]^n;$$

entonces hay una línea combinatoria

$$l([k]) \subset [k]^n$$

tal que

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l(j)}\right) > 0.$$

Proposición 9. *Las formulaciones 14 y 16 del teorema de densidad de Hales-Jewett son equivalentes.*

Demostración. ($14 \Rightarrow 16$): Suponemos el teorema de densidad de Hales-Jewett enunciado como en 14. Sea $\delta > 0$ y consideramos una colección de funciones

$$f_n : [k]^n \to \mathscr{B}$$

que envían palabras de longitud n en conjuntos medibles del espacio de probabilidad, cada uno con probabilidad mayor o igual a δ. Escribimos

$$f_n(w) = B_w^n$$

para cualquier n.

Queremos evidenciar que, para toda $n \geq 1$, existe un conjunto de palabras de longitud n

$$A_n \subset [k]^n$$

tal que

$$\frac{|A_n|}{k^n} \geq \delta \quad \text{y} \quad \mathbb{P}\left(\bigcap_{w \in A_n} B_w^n \right) > 0.$$

Estableciendo esto, ya habremos terminado, puesto que el Teorema 14 nos garantiza que existe $N > 0$ tal que para todo $n \geq N$, A_n contiene una línea combinatoria

$$l_n([k]) \subset A_n.$$

En ese caso

$$\bigcap_{w \in A_n} B_w^n \quad \subset \quad \bigcap_{i=1}^{k} B_{l_n(i)},$$

de donde

$$\mathbb{P}\left(\bigcap_{i=1}^{k} B_{l_n(i)} \right) > 0.$$

Llamaremos conjuntos de *tipo delta* a aquellos medibles B_α en X que satisfacen

$$\mathbb{P}(B_\alpha) > 0.$$

Si particionáramos el espacio X en conjuntos de tipo delta, cabrían a lo más

$$M = \left\lfloor \frac{1}{\delta} \right\rfloor$$

de ellos. Para una colección de $s \cdot M$ medibles tipo delta, existe al menos una intersección positiva

$$\bigcap_{i=1}^{s} B_{\alpha_i}$$

donde coinciden s medibles de la colección: en particular cuando respetamos una partición inicial de X en M medibles tipo delta, y con mayor razón si no lo hacemos.

Para n fija, tenemos una colección de k^n medibles tipo delta; a saber,

$$\{B_w^n : w \in [k]^n\} =: \mathscr{A}.$$

Luego, existe una intersección de probabilidad positiva

$$S = \bigcap_{B_\alpha \in \mathscr{E}} B_\alpha \qquad \mathscr{E} \subset \mathscr{A}$$

donde coinciden

$$\frac{k^n}{M} = \frac{k^n}{\left\lfloor \frac{1}{\delta} \right\rfloor} \geq \delta \cdot k^n$$

medibles de $\mathscr{A}$. Llamémosle A_n a la colección maximal de medibles en $\mathscr{A}$ tales que

$$\bigcap_{w \in A_n} B_w^n = S.$$

Claramente

$$\frac{|A_n|}{k^n} \geq \delta.$$

(14 $\Leftarrow$ 16) Observamos primero que 16 es cierto para $\delta > 1 - \frac{1}{k}$. En efecto, podemos elegir $N = 1$ y $n \geq N$; si tomamos

$$A \subset [k]^n$$

de densidad al menos δ, podemos particionar A en los subconjuntos

$$A_j = \{w \in A : w(1) = j\}.$$

Consideramos las proyecciones de las A_j sobre las últimas $n-1$ letras:

$$A_j' \subset [k]^{n-1}$$

y observamos que

$$\left| \bigcup_{j=1}^{k} A_j' \right| = |A| > \delta \cdot k^n > k^{n-1}(k-1).$$

Dado que cada A_j' tiene a lo más k^{n-1} elementos, que hay k conjuntos A_j', y que la desigualdad anterior es estricta, concluímos que

$$\bigcap_{j=1}^{k} A_j' \neq \emptyset.$$

Dicha intersección corresponde a una línea combinatoria en A.

Sea ϵ_0 el ínfimo de las δ para las cuales 14 es válido. Basta demostrar $\epsilon_0 = 0$: suponemos, para obtener una contradicción, que $\epsilon_0 > 0$. Tomamos $m > N_{\epsilon_0}$, donde N_{ϵ_0} es el natural que garantiza se cumpla 16 para conjuntos de densidad mayor a ϵ_0.

Sean

$$\epsilon_1 = \epsilon_0 \left(1 - \frac{1}{k^{m+2}}\right)$$

$$\epsilon_2 = \epsilon_1 + \frac{\epsilon_0}{2}\frac{1}{k^m};$$

notamos que

$$\epsilon_2 > \epsilon_0 > \epsilon_1.$$

Sea, por otra parte, N_{ϵ_2} el natural que garantiza la conclusión de 14 para conjuntos de palabras de densidad al menos ϵ_2, y sea $M \geq N_{\epsilon_2}$. Demostramos a continuación que la conclusión de 14 es válida para conjuntos de densidad al menos ϵ_1 y naturales mayores que $M + m$, lo cual contradice la definicón de ϵ_0. Concluímos así que $\epsilon_0 = 0$.

Elegimos $n > M + m$ y consideramos la medida de probabilidad *uniforme* sobre $[k]^n$ dada por

$$\mathbb{P}(B) = \frac{|B|}{k^n} \qquad B \subset [k]^n.$$

Consideramos las distribuciones uniformes sobre $[k]^m$ y $[k]^{n-m}$ de la misma manera.

Sea $A \subset [k]^n$ de medida mayor a ϵ_1 y definimos la funcion

$$g : [k]^m \to \mathscr{P}\left([k]^{n-m}\right),$$

donde $\mathscr{P}([k]^{n-m})$ es el conjunto potencia de $[k]^{n-m}$:

$$g(w) = A_w$$

$$A_w = \left\{u \in [k]^{n-m} : wu \in A\right\}.$$

Si se cumple que

$$\mathbb{P}(A_w) > \frac{\epsilon_0}{2} \qquad \forall w \in [k]^m,$$

entonces el Teorema 16 nos garantiza que existe una línea combinatoria $l_m([k]) \in [k]^m$ tal que

$$\mathbb{P}\left(\bigcap_{j=1}^{k} A_{l(j)}\right) > 0.$$

Si u es elemento de dicha intersección, entonces

$$l_m([k])u = \{l_m(1)u, \ldots, l_m(k)u\}$$

es línea combinatoria en A.

Por otro lado, si algún A_{w_j} tiene medida menor a $\frac{\epsilon_0}{2}$, demostramos que algún otro A_{w_i} debe tener medida mayor a ϵ_2; entonces, dado que $\epsilon_2 > \epsilon_0$, el conjunto

$$w_i A_{w_i} = \{w_i z : z \in A_{w_i}\}$$

tiene una línea combinatoria que, por definición de A_{w_i}, está en A.

En efecto, la medida de A es igual al promedio de las medidas de los A_w:

$$\frac{|A|}{k^n} = \frac{\sum_{w\in[k]^m} |A_w|}{k^n} = \frac{1}{k^m} \sum_{w\in[k]^m} \frac{|A_w|}{k^{n-m}}.$$

Luego, dado que

$$\frac{|A|}{k^n} > \epsilon_1,$$

tenemos

$$\sum_{w\in[k]^m} \frac{|A_w|}{k^{n-m}} > \epsilon_1 \cdot k^m.$$

Ahora bien, si para algún $w_j \in [k]^m$ se cumple

$$\frac{|A_{w_j}|}{k^{n-m}} < \frac{\epsilon_0}{2},$$

entonces

$$\sum_{[k]^m \setminus \{w_j\}} \frac{|A_w|}{k^{n-m}} > \epsilon_1 \cdot k^m - \frac{\epsilon_0}{2} = \epsilon_0 \left(k^m \left(1 - \frac{1}{k^{m+2}} \right) - \frac{1}{2} \right).$$

Tomando el promedio de las medidas sobre $g([k]^m \setminus \{w_j\})$, obtenemos

$$\begin{aligned} \frac{1}{k^m - 1} \sum_{[k]^m \setminus \{w_j\}} \frac{|A_w|}{k^{n-m}} &> \epsilon_0 \cdot \frac{1}{k^m - 1} \left(k^m \left(1 - \frac{1}{k^{m+2}} \right) - \frac{1}{2} \right) \\ &= \epsilon_0 \left(\frac{2k^{m+2} - 2 - k^2}{2k^2(k^m - 1)} \right) \\ &= \epsilon_0 \left(1 + \frac{k^2 - 2}{2k^2(k^m - 1)} \right) \\ &> \epsilon_2, \end{aligned}$$

de donde concluímos que al menos para un $w_i \in [k]^m$, $g(w_i)$ tiene medida mayor a ϵ_2 en $[k]^{n-m}$.

$\square$

3.4. Versiones finitaria e infinitaria, formulación ergódica de Hales-Jewett. Como antes, tenemos una versón infinitara, equivaente a la finitaria, de la formulación ergódica del teorema de densidad de Hales-Jewett.

TEOREMA 17. *Si a cada $w \in [k]^\omega$ le asignamos un conjunto medible en un espacio dado de probabilidad, $(X, \mathscr{B}, \mathbb{P})$ bajo una función*

$$f : [k]^\omega \to \mathscr{B},$$

de tal manera que

$$f(w) = B_w, \qquad \mathbb{P}(B_w) \geq \delta > 0 \quad \forall w \in [k]^\omega;$$

entonces hay una línea combinatoria

$$l([k]) \subset [k]^\omega$$

tal que

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l(j)}\right) > 0.$$

La demostración de la equivalencia sigue el mismo principio que hemos venido usando hasta ahora para equiparar las formulciones finitarias e infinitarias de los teoremas.

PROPOSICIÓN 10. *Los Teoremas 16 y 17 son equivalentes.*

DEMOSTRACIÓN. (16 ⇒ 17) Sea $(X, \mathscr{B}, \mathbb{P})$ un espacio de probabilidad fijo. Suponemos la formulación finitaria y elegimos $\delta > 0$. Por hipótesis, hay N_δ tal que para todo $n \geq N_\delta$ y toda función

$$g : [k]^n \to \mathscr{B}$$

$$g(w) = B_w$$

que satisface

$$\mathbb{P}(B_w) \geq \delta$$

hay una línea combinatoria

$$l_g([k]) \in [k]^n$$

tal que

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l_g(j)}\right) > 0.$$

Pero $l_g([k]) \in [k]^\omega$.

(16 $\Leftarrow$ 17) Procedemos por contrapositiva. Suponemos la negación del Teorema 16; elegimos $\delta > 0$ y una sucesión de naturales $(n_i; i \geq 1)$ con su respectiva sucesión de funciones

$$g_i : [k]^{n_i} \to \mathscr{B}$$

tal que, para todo n_i en la sucesión, toda línea combinatoria $l_{n_i}([k])$ en $[k]^{n_i}$ satisface

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l_{n_i}(j)}\right) = 0.$$

El teorema se sigue de la siguiente observación: Si n_i es parte de la sucesión anterior y $m \leq n_i$ entonces hay una función

$$g_m : [k]^m \to \mathscr{B}$$

tal que para toda línea combinatoria

$$l_m([k]) \subset [k]^m$$

se verifica

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l_m(j)}\right) = 0.$$

En efecto, considerando lo anterior, podemos construir

$$g_\omega : [k]^\omega \to \mathscr{B}$$

a partir de las

$$g_n \qquad n \in \mathbb{N};$$

basta establecer

$$w \in [k]^n \cap [k]^\omega \Rightarrow g_\omega(w) = g_n(w)\,.$$

Por construcción g_ω cumple que, para toda línea combinatoria $l([k]) \subset [k]^\omega$, se tiene

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l(j)}\right) = 0,$$

otorgándonos la contradicción del Teorema 17.

Basta, pues, demostrar la observación anterior. Notamos que si $m \leq n_i$ y $l_m([k])$ es una línea combinatoria en $[k]^m$, entonces $l_m([k])$ es la proyección de una línea combinatoria l_{n_i} en $[k]^{n_i}$ sobre las primeras m letras. Elegimos un elemento arbitrario u en $[k]^{n_i - m}$ y definimos

$$g_m : [k]^m \to \mathscr{B}$$

por

$$g_m(w) = g_{n_i}(wu)\,.$$

Así,

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{l_m(j)}\right) = \mathbb{P}\left(\bigcap_{j=1}^{k} B_{l_m(j)u}\right) = 0,$$

ya que

$$\{l_m(j)u\}_{j=1}^{k}$$

es línea combinatoria en $[k]^{n_i}$.

$\square$

Capítulo 3

Una demostración ergódica del teorema de densidad de Hales-Jewett

Ya que hemos formulado la *versión ergódica* del teorema de densidad de Hales-Jewett (de ahora en adelante DHJ), dedicamos este capítulo a su demostración. Seguiremos a grandes razgos el argumento de [**Aus11**] con algunas modificaciones.

Antes de entrar en materia, adelantamos un poco de la ruta a seguir. Demostraremos una formulación más específica, que a primera vista pudiera parecer más débil, de la versión ergódica de DHJ. No es difícil, si bien un poco técnico, ver que dicha formulación es equivalente a todas las demás que hemos hecho de DHJ; tiene, además, la ventaja de permitirnos trabajar en *espacios estándar de Borel*, y usar el potente *teorema de consistencia de Kolmogorov*. Así, en lugar de asociarle a cada palabra un medible, sin más, desarrollaremos nuestro problema en términos de

$$\left(X^{[k]^n}, \Sigma^{[k]^n}\right),$$

el espacio de medida producto de un espacio estándar de Borel (X, Σ) indexado sobre $[k]^n$. La notación se extiende de manera natural a espacios producto indexados sobre $[k]^\omega$: recordemos que por ω nos referimos a la cardinalidad de los naturales y con $[k]^\omega$ denotamos al espacio de todas las palabras compuestas por k letras.

La idea general que está detrás de todo lo que desarrollaremos de aquí en adelante es la siguiente: refinar suficiente y adecuadamente a $[k]^\omega$ para obtener subespacios combinatorios infinito-dimensionales, tales que los espacios producto con medida de probabilidad indexados por ellos verifiquen las condiciones de DHJ; verificar que, entonces, se tiene que haber satisfecho DHJ en el espacio original indexado por $[k]^\omega$. Después, tomar una sucesión de extensiones de (X, Σ) hasta llegar a un espacio *suficientemente y adecuadamente extendido* para el cual se cumpla DHJ, y verificar que entonces se debe cumplir DHJ en el espacio original.

Los *refinamientos adecuados y suficientes* serán aquellos tendientes a la *estacionariedad fuerte* y las extensiones *adecuadas y suficientes* aquellas tendientes a la *saciedad.* Ambas propiedades sentarán la bases para deducir DHJ a partir de la estructura de las distribuciones saciadas. Recordemos que ha sido la equivalencia entre las formulaciones finitaria e infinitaria de DHJ, a partir de un principio de correspondencia, que nos ha permitido aproximarnos al problema desde esta perspectiva.

Necesitaremos entonces definir con precisión qué es un subespacio de $[k]^\omega$, cuáles son las medidas de probabilidad definidas sobre los subespacios de $\left(X^{[k]^n}, \Sigma^{[k]^n}\right)$ que nos interesan, y qué relación tienen con la medida sobre el espacio original. Tal será el objeto de las siguientes secciones.

1. Subespacios

1.1. Subespacios de $[k]^\omega$.

Definición. Sean $l_1, l_2, \cdots, l_n$ líneas combinatorias y sea $N(n)$ la suma de las extensiones de las l_i, $i = 1, \cdots, n$. El **n-subespacio**, o **espacio combinatorio n-dimensional** definido por estas líneas es la función $\phi : [k]^n \to [k]^{N(n)}$ dada por

$$\phi(w) = l_1(x_1)\, l_2(x_2) \cdots l_n(x_n) \quad \text{si} \quad w = x_1 \cdots x_n.$$

Definición. Sea

$$(l_n)_{n \geq 1}$$

una sucesión de líneas combinatorias. El **subespacio combinatorio** (infinito-dimensional) definido por estas líneas es la función $\psi : [k]^\omega \to [k]^\omega$ dada por

$$\psi(w) = l_1(x_1) \cdots l_n(x_n) \quad \text{si} \quad w = x_1 \cdots x_n, \quad n \in \mathbb{N}.$$

Cuando el contexto no se preste a confusión, le llamaremos también n-subespacio a la imagen de un n-subespacio, y subespacio a la imagen de un subespacio.

1.2. Espacios polacos y espacios estándar de Borel.

Definición. La **σ-álgebra de Borel** sobre un espacio topológico es la σ-álgebra generada por todos los abiertos de la topología - o, equivalentemente, por todos los cerrados.

Definición. Un **espacio de Borel** es un espacio topológico junto con su σ-álgebra de Borel.

Definición. Decimos que un espacio métrico es **separable** cuando contiene un subconjunto denso y numerable.

Definición. Un **espacio polaco** es un espacio métrico separable y completo.

Definición. Si (X, Σ_X) y (Y, Σ_Y) son espacios medibles y existe

$$\lambda : \Sigma_X \to \Sigma_Y$$

biyección de conjuntos que preserva operaciones de σ-álgebra; es decir si

$$\lambda\left(\bigcap_{n\geq 1} A_n\right) = \bigcap_{n\geq 1} \lambda(A_n) \quad (A_n)_n \subset \Sigma_X,$$

$$\lambda\left(\bigcup_{n\geq 1} A_n\right) = \bigcup_{n\geq 1} \lambda(A_n) \quad (A_n)_n \subset \Sigma_X$$

y

$$\lambda(A^c) = \lambda(A)^c \quad A \in \Sigma_X,$$

entonces decimos que Σ_X y Σ_Y son **σ-álgebras isomorfas**. Decimos en este caso que X y Y son **σ-isomorfos** o, en símbolos,

$$X \overset{\sigma}{\backsim} Y.$$

Definición. Decimos que un espacio de Borel (X, Σ_X) es **numerablemente generado** si existe una familia numerable $\mathscr{D}$ de subconjuntos de X tal que

$$\sigma(\mathscr{D}) = \Sigma_X.$$

Definición. Si (X, Σ_X) es espacio de Borel numerablemente generado y X es σ-isomorfo a un espacio polaco Y con su respectiva σ-álgebra de Borel Σ_Y, decimos que (X, Σ_X) es **espacio estándar de Borel**.

A un espacio estándar de Borel bien podemos asociarle una medida de probabilidad. En tal caso, y en términos intuitivos, tiene una estructura análoga a la de un espacio de probabilidad sobre un Boreliano en $\mathbb{R}^n$. No pretendemos sintetizar aquí un estudio sobre espacios estándar de Borel; tomamos de los capítulos 1 y 5 de [**Par05**] algunos resultados y conceptos que exponemos a continuación.

1.3. Espacios producto y subespacios indexados.

Definición. Si X_i son espacios polacos para $i = 1, 2, \ldots$, el **espacio producto**

$$\prod_{i\in\mathbb{N}} X_i$$

es el espacio de sucesiones

$$(x_1, x_2, \ldots) \quad x_i \in X_i.$$

Escribimos también $X^{\mathbb{N}}$ para denotar tal espacio. También son espacios producto los de sucesiones

$$(x_j)_{j \in J} \qquad J \subset \mathbb{N},$$

con J finito o infinito. Los denotamos por X^J.

Definición. Las **proyecciones** de un espacio producto sobre sus coordenadas son las funciones

$$\pi_j : X^{\mathbb{N}} \to X_j$$

definidas por

$$\pi_j(x_1, x_2, \cdots, x_j, \ldots) = x_j.$$

También hay proyecciones sobre conjuntos de coordenadas $J \subset \mathbb{N}$; son las funciones

$$\pi_J : X^{\mathbb{N}} \to X^J$$

definidas por

$$\pi_J(x_1, x_2, \ldots) = (x_j)_{j \in J}.$$

Para $I \subset J$ subconjuntos de índices, la proyección

$$\pi_{JI} : X^J \to X^I$$

se define por

$$\pi_{JI}\Big((x_j)_{j \in J}\Big) = (x_i)_{i \in I}.$$

Definición. Si X_i es un espacio estándar de Borel y Σ_i su σ-álgebra para $i = 1, 2, \ldots$, entonces el **espacio medible producto**

$$\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right) = \prod_{i \in \mathbb{N}} (X_i, \Sigma_i)$$

es el espacio producto

$$X^{\mathbb{N}} = \prod_{i \in \mathbb{N}} X_i$$

junto con la σ-álgebra generada por los cilindros del tipo

$$\pi_j^{-1}(A) \quad \text{con} \quad A \in \Sigma_j.$$

Esto es,

$$\Sigma^{\mathbb{N}} = \sigma\big(\big\{\pi_j^{-1}(A) : \quad j \in \mathbb{N},\, A \in \Sigma_j\big\}\big).$$

Si además le asignamos una medida de probabilidad a $\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$, entonces tenemos un **espacio de probabilidad producto**.

De manera análoga definimos para cada $J \subset \mathbb{N}$ el espacio medible producto

$$\left(X^J, \Sigma^J\right) = \prod_{i \in J} (X_i, \Sigma_i)\,.$$

En particular, $\left(X^{\{m\}}, \Sigma^{\{m\}}\right) = (X_m, \Sigma_m)$.

Proposición 11. *Si X es polaco, (X, Σ_X) es espacio medible, y $\mathscr{T}_X$ la topología sobre X que induce su métrica, entonces existe una familia numerable $\mathscr{D}_X$ de conjuntos en Σ_X que generan a $\mathscr{T}_X$ bajo operaciones de σ-álgebra.*

Demostración. X es numerablemente generado, por lo que existe una familia numerable $\mathscr{D} \subset \Sigma_X$ tal que

$$\sigma(\mathscr{D}) = \Sigma_X.$$

Pero

$$\sigma(\mathscr{T}_X) = \Sigma_X;$$

entonces $\mathscr{T}_X$ se obtiene a partir de $\mathscr{D}$ bajo operaciones de σ-álgebra. $\mathscr{D}_X = \mathscr{D}$ satisface los requerimientos del teorema.

En particular, el conjunto de bolas de radio racional centradas en elementos de un conjunto numerable denso en X satisface los requerimientos. Además está contenido en $\mathscr{T}_X$. □

Definición. Si

$$(X_n, \mathscr{T}_n)_{n \geq 1}$$

son espacios polacos, la **topología producto** $\mathscr{T}^{\mathbb{N}}$ sobre $X^{\mathbb{N}}$ es aquella generada por los cilindros

$$\pi_n^{-1}(A) \quad A \in \mathscr{T}_n \quad n \geq 1;$$

es la topología más pequeña que hace continuas a las proyecciones

$$\pi_n : X^{\mathbb{N}} \to X_n.$$

La σ-álgebra de Borel sobre $X^{\mathbb{N}}$ es entonces

$$\sigma\left(\mathscr{T}^{\mathbb{N}}\right).$$

Teorema 18. *$\Sigma^{\mathbb{N}}$ es la σ-álgebra de Borel sobre $X^{\mathbb{N}}$*

Demostración. Denotamos por

$$\mathscr{B} = \mathscr{B}_{X^{\mathbb{N}}},$$

a la σ-álgebra de Borel sobre $X^{\mathbb{N}}$. Como cada (X_j, Σ_j) es estándar de Borel, podemos asumir sin pérdida de generalidad que X_j es polaco.

Para cada $j \geq 1$, denotamos por $\mathscr{T}_j$ a la topología que le corresponde a X_j. Por definición,

$$\sigma\big(\mathscr{T}^{\mathbb{N}}\big) = \mathscr{B},$$

así que queremos demostrar que

$$\sigma\big(\mathscr{T}^{\mathbb{N}}\big) = \Sigma^{\mathbb{N}}.$$

Empezamos con la contención más directa,

$$\Sigma^{\mathbb{N}} \subset \mathscr{B}.$$

$\Sigma^{\mathbb{N}}$ es generada por cilindros del tipo

$$\pi_j^{-1}(A_j) \quad A_j \in \Sigma_j = \sigma(\mathscr{T}_j)\,.$$

Como los X_j son polacos, para cada j hay una familia numerable $\mathscr{D}_j$ en $\mathscr{T}_j$ tal que

$$\sigma(\mathscr{D}_j) = \Sigma_j.$$

Luego,

$$\Sigma^{\mathbb{N}} = \sigma\left(\bigcup_{j\in\mathbb{N}} \pi_j^{-1}(\mathscr{D}_j)\right).$$

Notamos que

$$\pi_j^{-1}(\mathscr{D}_j) \subset \mathscr{T}^{\mathbb{N}} \quad \forall j \in \mathbb{N}$$

y tenemos la contención deseada.

Para establecer la inclusión inversa,

$$\mathscr{B} \subset \Sigma^{\mathbb{N}},$$

basta con demostrar que

$$\mathscr{T}^{\mathbb{N}} \subset \Sigma^{\mathbb{N}}.$$

Por la Proposición 11, la familia

$$\bigcup_{j\in\mathbb{N}} \pi_j^{-1}(\mathscr{D}_j)$$

genera a $\mathscr{T}^{\mathbb{N}}$ bajo operaciones de σ-álgebra. Como

$$\pi_j^{-1}(\mathscr{D}_j) \subset \Sigma^{\mathbb{N}} \quad \forall j \in \mathbb{N},$$

concluimos la segunda contención. $\square$

Definición. Si $\mathbb{P}_{\mathbb{N}}$ es medida de probabilidad sobre $X^{\mathbb{N}}$ y $J \subset \mathbb{N}$, decimos que la medida de probabilidad

$$\mathbb{P}_J = \mathbb{P}_{\mathbb{N}} \circ \pi_J^{-1}$$

sobre X^J es la **probabilidad proyectada** de $\mathbb{P}_{\mathbb{N}}$ sobre X^J. Si $I \subset J$, la probabilidad proyectada de $\mathbb{P}_J$ sobre X^I es

$$\mathbb{P}_I = \mathbb{P}_J \circ \pi_{JI}^{-1}.$$

Naturalmente,

$$\mathbb{P}_I = \mathbb{P}_{\mathbb{N}} \circ \pi_I^{-1}.$$

Definición. Sea $\mathbb{P}_J$ una medida de probabilidad sobre $\left(X^J, \Sigma^J\right)$ para toda $J \subset \mathbb{N}$. Decimos que la familia $(\mathbb{P}_J)$ es **consistente** si $\mathbb{P}_I = \mathbb{P}_J \circ \pi_{JI}^{-1}$ siempre que $I \subset J$.

Consideramos el siguiente problema: tenemos una sucesión de conjuntos de índices en $\mathbb{N}$

$$J_1 \subset J_2 \subset J_3 \subset \dots$$

tal que

$$\mathbb{N} = \bigcup_{i=1}^{\infty} J_i;$$

suponemos que la sucesión de medidas

$$\mathbb{P}_{J_1}, \mathbb{P}_{J_2}, \mathbb{P}_{J_3}, \dots$$

es consistente sobre sus respectivos espacios

$$\left(X^{J_1}, \Sigma^{J_1}\right), \left(X^{J_2}, \Sigma^{J_2}\right), \left(X^{J_3}, \Sigma^{J_3}\right), \dots.$$

Queremos saber si existe una distribución $\mathbb{P}_{\mathbb{N}}$ sobre $\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$ tal que

$$\mathbb{P}_{J_n} = \mathbb{P}_{\mathbb{N}} \circ \pi_{J_n}^{-1} \quad \forall n;$$

es decir, que $\mathbb{P}_{\mathbb{N}}$ sea consistente con todas las medidas de la sucesión.

Motivamos la resolución de este problema: Sean

$$\mathscr{B}_n = \left\{\pi_{J_n}^{-1}(A) : \quad A \in \Sigma^{J_n}\right\}.$$

Notamos que

$$\mathscr{B}_1 \subset \mathscr{B}_2 \subset \mathscr{B}_3 \subset \dots,$$

y que cada $\mathscr{B}_n \subset \Sigma^{\mathbb{N}}$ es σ-álgebra. Notamos además que

$$\Sigma^{\mathbb{N}} = \sigma\left(\bigcup_{n=1}^{\infty} \mathscr{B}_n\right).$$

Definimos, para cada n, una medida sobre $\mathscr{B}_n$:

$$\nu_n\big(\pi_{J_n}^{-1}(A)\big) = \mathbb{P}_{J_n}(A) \quad \forall A \in \Sigma^{J_n};$$

dada la consistencia de las $\mathbb{P}_{J_n}$, tenemos que

$$\nu_n(A) = \nu_m(A) \quad \forall A \in \Sigma^{J_n}$$

siempre que $m \geq n$. De esta manera,

$$\nu(A) = \nu_n(A)\,, \quad A \in \Sigma^{J_n}$$

es una función bien definida sobre

$$\bigcup_{n=1}^{\infty} \mathscr{B}_n.$$

Por otra parte, $\nu\big(X^{\mathbb{N}}\big) = \nu_1\big(X^{\mathbb{N}}\big) = 1$ y de igual manera $\nu(\emptyset) = 0$. Además, ν es finitamente aditiva puesto que cada $\mathbb{P}_{J_n}$ lo es. Si logramos establecer que ν es σ-aditiva, entonces será medida sobre $\Sigma^{\mathbb{N}}$, y por el Teorema de Extensión de Carathéodory será única, puesto que está definida sobre una familia que genera a la σ-álgebra. La dificultad principal está entonces en demostrar la σ-aditividad de ν. Ya que en [**Par05**] se le dedica buena parte del quinto capítulo a ello, nos limitamos a referir al lector.

Mencionamos que dicho estudio culmina con un teorema importante, que generaliza el resultado anterior:

TEOREMA 19 (Teorema de consistencia de Kolmogorov). *Sea I un conjunto de índices cualquiera, posiblemente no numerable. Sean (X, Σ) un espacio estándar de Borel y*

$$\big(X^I, \Sigma^I\big) = \prod_{\alpha \in I} (X_\alpha, \Sigma_\alpha)$$

el espacio producto sobre I con la σ-álgebra generada por los cilindros

$$\pi_\alpha^{-1}(A) \quad A \in \Sigma_\alpha.$$

Si

$$\{\mu_F : \quad F \subset I,\ F \text{ finito}\}$$

es una familia consistente de medidas sobre los espacios

$$\prod_{\alpha \in F} (X_\alpha, \Sigma_\alpha)\,, \quad F \text{ finito},$$

entonces existe una única medida μ_I sobre Σ^I tal que

$$\mu_F(A) = \mu_I\big(\pi_{IF}^{-1}(A)\big) \quad \forall A \in \Sigma^F$$

para todo $F \subset I$ finito.

No requerimos de tanta potencia para nuestros objetivos, ya que sólo trabajaremos con espacios-producto numerables.

En virtud del teorema de consistencia de Kolmogorov, una familia consistente de medidas

$$\mathscr{A} := \{\mathbb{P}_J : \quad J \subset \mathbb{N} \quad \text{finito}\}$$

induce su extensión

$$\mathscr{A}' := \{\mathbb{P}_J : \quad J \subset \mathbb{N}\}.$$

En efecto, $\mathbb{P}_{\mathbb{N}}$ existe y es único, y para conjuntos infinitos de índices $J \subset \mathbb{N}$ las medidas

$$\mathbb{P}_J := \mathbb{P}_{\mathbb{N}} \circ \pi_J^{-1}$$

están bien definidas. Además, para $J \subset I \subset \mathbb{N}$ no necesariamente finitos tenemos

$$\pi_J = \pi_{IJ} \circ \pi_I,$$

de modo que

$$\mathbb{P}_J = \mathbb{P}_{\mathbb{N}} \circ (\pi_{IJ} \circ \pi_I)^{-1} = \mathbb{P}_{\mathbb{N}} \circ \pi_I^{-1} \circ \pi_{IJ}^{-1} = \mathbb{P}_I \circ \pi_{IJ}^{-1}$$

y la familia $\mathscr{A}'$ también es consistente.

1.4. Reformulación del problema. Si ahora consideramos el espacio producto de espacios estándar de Borel indexados sobre $[k]^\omega$, y a cada espacio producto

$$\prod_{w \in F} (X_w, \Sigma_w)$$

con F finito le asignamos una distribución de manera que sean consistentes, habrá una única distribución $\mathbb{P}_\omega$ sobre

$$\prod_{w \in [k]^\omega} (X_w, \Sigma_w)$$

tal que

$$\mathbb{P}_F = \mathbb{P}_\omega \circ \pi_F^{-1} \quad \text{para todo} \quad F \subset [k]^\omega.$$

A continuación hacemos algunas observaciones acerca de la estructura de probabilidad sobre $X^{[k]^\omega}$ que requeriremos más adelante.

Definición. Sean (X, Σ_X) y (Y, Σ_Y) espacios estándar de Borel con medidas de probabilidad $\mathbb{P}_X$ y $\mathbb{P}_Y$, respectivamente. Supongamos que existe biyección bimedible

$$\lambda : X \to Y$$

tal que

$$\mathbb{P}_Y = \mathbb{P}_X \circ \lambda^{-1}.$$

Entonces también

$$\mathbb{P}_X = \mathbb{P}_Y \circ \lambda$$

y decimos que $(X, \Sigma_X, \mathbb{P}_X)$ y $(Y, \Sigma_Y, \mathbb{P}_Y)$ son **$\mathbb{P}$-isomorfos**.

Es inmediato que si dos espacios de probabilidad son $\mathbb{P}$-isomorfos entonces son σ-isomorfos.

De este punto en adelante haremos un cambio de notación para distinguir entre dos espacios $\mathbb{P}$-isomorfos pero distintos.

Definición. Dados $\mathbb{P}_\omega$ e $I \subset [k]^\omega$, escribimos $\mathbb{P}_{\#I}$, para denotar a la **probabilidad proyectada** sobre $\left(X^I, \Sigma^I\right)$. Es decir,

$$\mathbb{P}_{\#I} = \mathbb{P}_\omega \circ \pi_I^{-1}.$$

Por otro lado, cuando I sea de cardinalidad infinita y

$$\psi : [k]^\omega \to I$$

sea biyección, designaremos con $\mathbb{P}_I$ a la medida de probabilidad sobre $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}\right)$ determinada por sus distribuciones finito-dimensionales

$$\mathbb{P}_I(X_{w_1} \in A_1, \ldots, X_{w_n} \in A_n) = \mathbb{P}_\omega\left(X_{\psi(w_1)} \in A_1, \ldots, X_{\psi(w_n)} \in A_n\right).$$

Notamos que

$$\left(X^I, \Sigma^I, \mathbb{P}_{\#I}\right) \overset{\mathbb{P}}{\backsim} \left(X^{[k]^\omega}, \Sigma^{[k]^\omega}, \mathbb{P}_I\right).$$

Nos interesa en especial el caso en el que $I = \psi([k]^\omega)$ es subespacio de $[k]^\omega$.

Establecemos una distinción similar para subespacios finito-dimensionales.

Definición. Si $\phi \subset [k]^\omega$ es n-subespacio, escribimos $\mathbb{P}_{\#\phi}$ para denotar a la probabilidad proyectada sobre X^ϕ y $\mathbb{P}_\phi$ para denotar a la probabilidad sobre $X^{[k]^n}$ determinada por

$$\mathbb{P}_\phi(X_{w_1} \in A_1, \cdots, X_{w_n} \in A_n) = \mathbb{P}_\omega\left(X_{\phi(w_1)} \in A_1, \cdots, X_{\phi(w_n)} \in A_n\right)$$

para $w_1, \cdots, w_n \in [k]^n$.

Para $\phi = [k]^n$ abreviamos

$$\mathbb{P}_n := \mathbb{P}_{[k]^n}.$$

Definición. El **proceso estocástico correspondiente a I** es el generado por las proyecciones a coordenadas, es decir

$$\left(\pi_i(x)\right)_{i \in I} \quad x \in X^I.$$

1.5. Límites inversos. Una noción íntimamente ligada al teorema de consistencia de Kolmogorov es la de límite inverso o límite proyectivo. Antes de entrar en materia necesitamos algunas observaciones adicionales sobre nuestros espacios estándar de Borel.

Definición. Un **átomo** de un espacio medible $(X, \mathscr{B})$ es un conjunto $A \in \mathscr{B}$ tal que si $B \subset A$ y $B \in \mathscr{B}$, entonces $B = A$ ó $B = \emptyset$.

Definición. Decimos que un espacio estándar de Borel $(X, \mathscr{B}_X)$ es **separable** si

$$\{x\} \in \mathscr{B}_X \quad \forall x \in X.$$

Es decir, todos los singuletes son medibles.

No se confunda el lector con la noción de *separabilidad* sobre espacios polacos, que es un concepto netamente topológico. Si bien la separabilidad de un espacio polaco y la separabilidad de un espacio estándar de Borel definido sobre el polaco están relacionadas, no son lo mismo.

Proposición 12. *Todo espacio estándar de Borel es σ-isomorfo a un espacio estándar de Borel separable.*

Antes de demostrar la Proposición exhibimos un espacio estándar de Borel separable que es particularmente útil e invocamos un resultado al respecto.

Le llamamos a M al espacio de funciones

$$x : \mathbb{N} \to \{0, 1\}$$

con la topología generada por los abiertos básicos

$$M_j = \{x : \, x(j) = 1\}.$$

M es un espacio métrico compacto, por ejemplo bajo la métrica

$$d(x, y) = \text{ínf} \left\{ \frac{1}{k+1} : x(i) = y(i) \quad \forall \, i < k \right\}.$$

Para verificar compacidad, basta observar que para cualquier sucesión $(x_n)_n$ de puntos en M existe una subsucesión convergente. Para un natural N arbitrariamente grande, siempre habrá una subsucesión $(x_{n_m})_m$ tal que todos sus elementos coincidan en los los primeros N lugares, ya que 2^N es finito. A esa subsucesión se le pueda refinar de manera que todos sus elementos coincidan en los primeros $2N$ lugares, y así sucesivamente. Ahora construímos una sucesión con el truco de *diagonalización* que ya empleamos en el Capítulo 2: elegimos el primer elemento de la primer refinación, el segundo elemento de la segunda refinación, y así sucesivamente. Ésta subsucesión *diagonal* es convergente.

Escribimos $(M, \mathscr{B}_M)$ para designar al espacio estándar de Borel generado por M con su topología. Para ver que es separable, observamos que todo singulete $\{x\} \in M$ es una intersección numerable de M_j's y complementos de M_j's, según si $x(j) = 0$ ó 1, respectivamente.

Demostración de la Proposición 12. Sea $(X, \mathscr{B}_X)$ espacio estándar de Borel, y sea

$$(A_n)_{n\geq 1}$$

una familia numerable de medibles que genera a $\mathscr{B}_X$. Definimos a

$$\tau : X \to M$$

por

$$\tau(x) = (\mathbf{1}_{A_1}(x), \mathbf{1}_{A_2}(x), \ldots).$$

Notamos que τ es medible ya que

$$\tau^{-1}(M_j) = A_j \quad \forall j \in \mathbb{N}.$$

Por otro lado,

$$\left\{\tau^{-1}(A) : A \in \mathscr{B}_M\right\} = \mathscr{B}_X.$$

En efecto, el lado izquierdo de la ecuación es una σ-álgebra por ser imagen inversa bajo una función medible de una σ-álgebra; además todo A_j pertenece a esa σ-álgebra. Como la familia $(A_j)_j$ genera a $\mathscr{B}_X$ se tiene la igualdad.

Consideramos ahora los conjuntos

$$e_x = \tau^{-1}(\tau(x)) \quad x \in X.$$

Afirmamos que cada e_x es un átomo. Notamos primero que $e_x \in \mathscr{B}_X$ puesto que $\mathscr{B}_M$ es separable y $\tau(x)$ es un punto en M. Ahora, supongamos que

$$B \subset e_x \quad B \in \mathscr{B}_X.$$

Entonces

$$B = \tau^{-1}(A) \text{ con } A \in \mathscr{B},\ A \subset \tau(x).$$

Luego, $A = \tau(x)$ ó $\emptyset$. Esto es, $B = e_x$ ó $\emptyset$. Tenemos pues que e_x es átomo.

Consideramos el espacio

$$X^0 = \{e_x : x \in X\}$$

y definimos la función

$$\theta : X \to X^0$$

con el mapeo natural

$$x \to e_x.$$

A X^0 le asignamos la σ-álgebra

$$\mathscr{B}^0 = \{\theta(A) : A \in \mathscr{B}_X\}.$$

Claramente θ es bimedible y

$$X \overset{\sigma}{\backsim} X^0.$$

Además X^0 es separable ya que para todo $e \in X^0$ se tiene que $\theta^{-1}(e)$ es un átomo y, en particular, está en $\mathscr{B}_X$. A $\left(X^0, \mathscr{B}^0\right)$ le llamamos el espacio *sepárable canónico* de $(X, \mathscr{B}_X)$. □

Ahora que sabemos que todo espacio estándar de Borel tiene un separable canónico que preseva su estructura de medida intacta, podemos enunciar con precisión qué es un límite proyectivo.

Definición. Sea

$$(X_n, \Sigma_n)_{n\geq 1}$$

una sucesión de espacios estándar de Borel separables y para cada n,

$$\psi_{n+1} : X_{n+1} \to X_n$$

una función medible suprayectiva. El **límite inverso** o **límite proyectivo** correspondiente a esas funciones es el espacio medible $\left(\tilde{X}, \tilde{\Sigma}\right)$ que satisface:

i) $\tilde{X}$ es el espacio de todas las sucesiones $(x_1, x_2, \ldots)$ con $x_i \in X_i$, tales que $\psi_n(x_{n+1}) = x_n$ para toda $n \geq 1$.

ii) $\tilde{\Sigma}$ es la σ-álgebra más pequeña de subespacios de $\tilde{X}$ para la cual las proyecciones $\pi_n(x_1, x_2, \ldots) \to x_n$ son medibles para todo $n \geq 1$.

En términos intuitivos, el límite proyectivo es el espacio medible que contiene a todos los (X_n, Σ_n). Un ejemplo que indirectamente ya discutimos en la subsección sobre espacios producto es en el que

$$(X_n, \Sigma_n) = \prod_{j=1}^{n} (X_j, \Sigma_j)$$

y

$$\psi_n(x_1, \cdots, x_n, x_{n+1}) = (x_1, \cdots, x_n) \quad \forall n \geq 1.$$

En este caso $\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$ es el límite proyectivo - o, mejor dicho, el límite proyectivo $\left(\tilde{X}, \tilde{\Sigma}\right)$ satisface

$$\tilde{X} \overset{\sigma}{\backsim} X^{\mathbb{N}}.$$

Invitamos al lector a que confirme esta afirmación usando la biyección natural

$$\theta(x_1, x_2, \ldots) = \left((x_1), (x_1, x_2), \ldots\right).$$

Hay otro límite proyectivo que nos resultará de importancia en nuestra exposición: se trata de $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}\right)$ como límite proyectivo de

$$\left(X^{[k]^n}, \Sigma^{[k]^n}\right)_{n\geq 1}$$

bajo ciertas proyecciones

$$\psi_n : X^{[k]^{n+1}} \to X^{[k]^\phi}$$

donde ϕ es n-subespacio en $[k]^{n+1}$. Abordamos esta discusión en la sección de estacionariedad fuerte.

Introducimos un término de uso común que nos facilitará la exposición.

DEFINICIÓN. Sean (Y, Σ_Y) y (X, Σ_X) espacios estándar de Borel con

$$\psi : Y \to X$$

función medible suprayectiva. Entonces

$$\psi^{-1}(\Sigma_X)$$

es subσ-álgebra de Σ_Y y decimos que Y es **extensión** de X o, equivalentemente, que X es **factor** de Y.

Cuando X y Y son espacios de probabilidad, es decir que $\mathbb{P}_Y$ es probabilidad sobre (Y, Σ_Y) y $\mathbb{P}_X$ sobre (X, Σ_X), decimos que Y es **extensión** de X (o que X es **factor** de Y) solamente si

$$\mathbb{P}_X = \mathbb{P}_Y \circ \psi^{-1}.$$

En ese caso decimos también que $\mathbb{P}_Y$ es extensión de $\mathbb{P}_X$.

A una sucesión de extensiones

$$(X_0, \Sigma_0) \leftarrow (X_1, \Sigma_1) \leftarrow (X_2, \Sigma_2) \leftarrow \dots$$

le llamamos también **torre de extensiones**.

El límite proyectivo de una sucesión de extensiones es el espacio medible más pequeño que las contiene a todas ellas como factores. Por ahora, y hasta la sección de saciedad, consideramos solamente extensiones que son productos cartesianos; no todo límite proyectivo es de este tipo. Gracias al teorema de consistencia de Kolmogorov sabemos que existe una única distribución sobre un límite proyectivo que sea consistente con las medidas de sus factores.

Las características de un espacio estándar de Borel se preservan bajo límites proyectivos:

TEOREMA 20 ([**Par05**], p. 136). *Sea*

$$(X_n, \Sigma_n)_{n \geq 1}$$

una torre de extensiones donde cada X_n *espacio estándar de Borel separable. Entonces su límite inverso* $\left(\tilde{X}, \tilde{\Sigma}\right)$ *es espacio estándar de Borel separable.*

No demostramos aquí este resultado.

TEOREMA 21 (de Tikhonov). *Si*

$$(X_\alpha, \mathscr{T}_\alpha)_{\alpha\in\mathscr{A}}$$

es una sucesión de espacios topológicos compactos, entonces el espacio producto

$$\prod_{\alpha\in\mathscr{A}} X_\alpha$$

con la topología producto también es compacto.

Como corolario inmediato, si

$$(X_n, \Sigma_n)_{n\geq 1}$$

es sucesión de espacios estándar de Borel compactos, entonces el espacio medible producto

$$\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$$

es estándar de Borel compacto.

2. Otra versión ergódica de DHJ

Estamos ahora en condiciones de reformular el teorema de densidad de Hales-Jewett:

TEOREMA 22. *Sean $\delta > 0$, (X, Σ) un espacio estándar de Borel compacto, $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}, \mathbb{P}_\omega\right)$ un espacio de probabilidad indexado por $[k]^\omega$, y $A \in \Sigma$ tal que*

$\mathbb{P}_\omega \circ \pi_w^{-1}(A) \geq \delta$ para todo $w \in [k]^\omega$; entonces para algún $m \geq 1$ hay una línea combinatoria $l([k]) \subset [k]^m$ tal que

$$\mathbb{P}_\omega\left(\bigcap_{i=1}^{k} \pi_{l(i)}^{-1}(A)\right) > 0.$$

Como antes, el teorema tiene una contraparte finitaria:

TEOREMA 23. *Para todo $\delta > 0$ existe $N \geq 0$ suficientemente grande tal que si $n \geq N$,*

$$\left(X^{[k]^n}, \Sigma^{[k]^n}, \mathbb{P}_n\right)$$

es espacio producto de un estándar de Borel compacto indexado por $[k]^n$ y $A \in \Sigma$ es tal que

$$\mathbb{P}_n \circ \pi_w^{-1}(A) \geq \delta \quad \forall w \in [k]^n;$$

entonces hay una línea combinatoria $l([k]) \subset [k]^n$ tal que

$$\mathbb{P}_n\left(\bigcap_{i=1}^{k} \pi_{l(i)}^{-1}(A)\right) > 0.$$

La equivalencia entre estas dos formulaciones se demuestra mediante un principio de correspondencia, análogo al que utilizamos para demostrar que los Teoremas 17 y 16 son equivalentes. La similitud es tal que resultaría repetitivo escribir la demostración.

Ya que de aquí en adelante nos enfocaremos en demostrar el Teorema 22, cabe explicitar la equivalencia entre éste y su formulación anterior, el Teorema 17.

PROPOSICIÓN 13. *Los teoremas 22 y 17 son equivalentes.*

DEMOSTRACIÓN. ($22 \Leftarrow 17$) $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}, \mathbb{P}_\omega\right)$ es un espacio de medida donde se cumplen las hipótesis del Teorema 17 con

$$f(w) = \pi_w^{-1}(A)\,.$$

($22 \Rightarrow 17$) Sabemos que los Teoremas 22 y 23 son equivalentes, así como los Teoremas 16 y 17; por otro lado ya demostramos en el Capítulo 2 que los Teoremas 16 y 14 son equivalentes. Basta entonces con demostrar $23 \Rightarrow 14$. Pero ésta demostración es idéntica a la que expusimos para establecer $16 \Rightarrow 14$. En efecto, consideramos

$$\left(X^{[k]^n}, \Sigma^{[k]^n}, \mathbb{P}_n\right)$$

donde

$$X = \{0, 1\}$$

y Σ es el conjunto potencia de X, que es la σ-álgebra de Borel sobre la topología discreta. Definimos a $\mathbb{P}_n$ de la siguiente manera: si $B \in \Sigma^{[k]^n}$,

$$\mathbb{P}_n(B) = \frac{\#B}{k^n},$$

donde $\#B$ es la cantidad de índices $w \in [k]^n$ tales que

$$1 \in \pi_w(B)\,.$$

El espacio X es trivialmente separable y numerablemente generado: es estándar de Borel. Además es compacto. Claramente

$$\left(X^{[k]^n}, \Sigma^{[k]^n}, \mathbb{P}_n\right) \overset{\mathbb{P}}{\backsim} \left([k]^n, \mathscr{P}([k]^n), \mathbb{P}_{unif}\right),$$

con $\mathbb{P}_{unif}$ la distribución uniforme definida sobre el conjunto potencia de $[k]^n$, al cual denotamos con $\mathscr{P}([k]^n)$. Luego, si $A \subset [k]^n$ es de densidad al menos $\delta > 0$ entonces

$$\mathbb{P}_n\left(\bigcap_{w \in A} \pi_w^{-1}(1)\right) \geq \delta > 0$$

y por el Teorema 23 existe N suficientemente grande tal que si $n \geq N$, existe línea combinatoria $l([k]) \subset [k]^n$ con

$$\mathbb{P}_n\left(\bigcap_{i=1}^{k} \pi_{l(i)}^{-1}(1)\right) > 0.$$

Con estas observaciones la demostración de 16 $\Rightarrow$ 14 explicada en el Capítulo 2 aplica tal cual. $\square$

Habiendo establecido nuestro problema, necesitaremos un poco más de maquinaria teórica para enfocarlo a una resolución con teoría ergódica y probabilidad de DHJ.

3. La métrica de Prohorov

El propósito de esta sección es establecer la completud de el espacio de medidas de probabilidad sobre $\left(X^{[k]^n}, \Sigma^{[k]^n}\right)$ y sobre $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}\right)$ bajo cierta métrica, que llamamos de Prohorov. Utilizaremos esta propiedad más adelante para reducir DHJ al caso de estacionariedad fuerte.

Si X es un espacio abstracto de Borel, denotamos por $M(X)$ al conjunto de medidas de probabilidad definidas sobre X. Una topología sobre $M(X)$ es la generada por los siguientes básicos: si $\mathbf{P}$ es medida de probabilidad sobre X, una vecindad de $\mathbf{P}$ es el conjunto de medidas de probabilidad $\mathbf{Q}$ que satisfacen

$$\left\|\int f_i \, d\mathbf{P} - \int f_i \, d\mathbf{Q}\right\| < \epsilon$$

para todo conjunto finito de funciones

$$f_i \quad i = 1, \cdots, s \quad s \in \mathbb{N}$$

continuas y acotadas. A esta topología se le llama la *topología débil* sobre $M(X)$.

Cuando X es polaco y consideramos el espacio de medidas de probabilidad sobre $(X, \mathscr{B}_X)$ donde $\mathscr{B}_X$ es la σ-álgebra de Borel, la topología débil es metrizable.

Para $A \subset X$, escribimos A^ϵ para denotar

$$\{x \in X : d(x, A) < \epsilon\}$$

donde d es métrica en X y

$$d(x, A) = \text{ínf}\left\{d(x, y) : y \in A\right\}.$$

Definición. La **métrica de Prohorov**

$$d_P : M(X) \times M(X) \to \mathbb{R}$$

manda a cada par de medidas $\{\mathbf{P}, \mathbf{Q}\}$ al ínfimo de las $\epsilon > 0$ que satisfacen

$$\mathbf{P}(A) \leq \mathbf{Q}(A^\epsilon) + \epsilon \quad \mathbf{Q}(A) \leq \mathbf{P}(A^\epsilon) + \epsilon$$

para todo A medible.

Teorema 24 ([**Bil99**], p. 72-73.). *La métrica de Prohorov está bien definida sobre cualquier espacio de Borel. Si X es polaco, la topología débil coincide con la topología inducida por la métrica de Prohorov sobre $M(X)$. En ese caso, $M(X)$ es un espacio polaco bajo la métrica de Prohorov.*

No reproducimos aquí la demostración. Necesitamos otros resultados para poder utilizar la métrica de Prohorov.

Definición. Sea X espacio de Borel. Una familia de medidas

$$\mathscr{F} := \{\mathbf{P}_\alpha : \quad \alpha \in \mathscr{A}\}$$

en $M(X)$ es **relativamente compacta** si toda sucesión

$$(\mathbf{P}_n)_{n \geq 1} \subset \mathscr{F}$$

contiene una subsucesión

$$(\mathbf{P}_{n_i})_{i \geq 1}$$

que converge bajo la topología débil a algún $\mathbf{P} \in M(X)$, no necesariamente en $\mathscr{F}$.

Si X es polaco o estándar de Borel, la convergencia también se da bajo la métrica de Prohorov.

Definición. Sea X un espacio de Borel y $\mathscr{T}$ su topología. Decimos que una famila de medidas $\mathscr{F}$ en $M(X)$ es **tensa** si para todo $\epsilon > 0$ existe un compacto $K_\epsilon \in \mathscr{T}$ tal que

$$\mathbf{P}(K_\epsilon) > 1 - \epsilon \qquad \forall \mathbf{P} \in \mathscr{F}.$$

Decimos también que $\mathscr{F}$ tiene tensión.

Teorema 25 ([**Bil99**], p. 58-63.). *Si una familia de medidas $\mathscr{F}$ en $M(X)$ es tensa, entonces es relativamente compacta.*

No reproducimos tampoco la demostración de este resultado.

Proposición 14. *Si X es polaco, una familia $\mathscr{F}$ en $M(X)$ es relativamente compacta si y sólo si su cerradura bajo d_P es d_P-compacta.*

Demostración. Es consecuencia del Teorema 24. □

Proposición 15. *Si X es espacio estándar de Borel compacto, entonces el espacio $M\left(X^{\mathbb{N}}\right)$ de distribuciones sobre $\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$ es d_P-compacto.*

Demostración. Por el teorema de Tikhonov, $X^{\mathbb{N}}$ es compacto bajo la topología producto $\mathscr{T}^{\mathbb{N}}$ y, por el Teorema 18,

$$\Sigma^{\mathbb{N}} = \sigma\left(\mathscr{T}^{\mathbb{N}}\right).$$

Luego, $M\left(X^{\mathbb{N}}\right)$ es trivialmente tenso puesto que

$$\mathbf{P}\left(X^{\mathbb{N}}\right) > 1 - \epsilon \qquad \forall \epsilon > 0 \quad \forall \mathbf{P} \in M\left(X^{\mathbb{N}}\right)$$

y $X^{\mathbb{N}}$ es compacto. Por el Teorema 25, $M\left(X^{\mathbb{N}}\right)$ es relativamente compacto y, por la Proposición 14, resta sólo ver que es cerrado. Por el Teorema 24 sabemos que $M\left(X^{\mathbb{N}}\right)$ es completo; por otra parte, no puede haber sucesiones que se escapen del espacio. En efecto, la compacidad relativa ya nos asegura para toda sucesión la existencia de una subsucesión que converge a algún elemento de $M\left(X^{\mathbb{N}}\right)$. □

4. Estacionariedad fuerte

Ahora tenemos las herramientas teóricas para acotar el problema, al considerar una clase de distribuciones que llamamos fuertemente estacionarias. Veremos que es suficiente con demostrar el teorema de Densidad de Hales-Jewett para esta clase.

Por comodidad de ahora en adelante abreviamos

$$\begin{aligned}
(X^{\omega}, \Sigma^{\omega}) &:= \left(X^{[k]^{\omega}}, \Sigma^{[k]^{\omega}}\right), \\
\left(X^{\psi}, \Sigma^{\psi}\right) &:= \left(X^{\psi([k]^{\omega})}, \Sigma^{\psi([k]^{\omega})}\right), \quad \psi([k]^{\omega}) \quad \text{subespacio}
\end{aligned}$$

y, cuando el contexto no se preste a confusión, escribimos ψ en lugar de $\psi([k]^{\omega})$.

Definición. Si ψ y γ son subespacios, decimos que γ es una **refinación** de ψ siempre que

$$\gamma \subset \psi.$$

Por ejemplo, para un subespacio generado por la sucesión de líneas combinatorias

$$(l_n)_{n\geq 1}$$

una refinación posible es

$$\{l_1(k) l_2(1) \dots l_m(x_m) : \quad m \in \mathbb{N}\}.$$

Es decir, *fijamos*

$$x_1 = k$$

$$x_2 = 1.$$

En general toda refinación se obtiene fijando un número finito o numerable de variables en $[k]$ para las líneas combinatorias que definen al subespacio, de manera que se reducen las líneas combinatorias *libres*.

DEFINICIÓN. Si

$$\phi = \{l_1(x_1)l_2(x_2)\cdots l_n(x_n) : x_i \in [k]\}$$

es n-subespacio, le llamamos **subespacio acumulado** a

$$\phi^* = \bigcup_{j=1}^{n} \{l_1(x_1)\cdots l_j(x_j) : x_i \in [k]\}.$$

En particular, el **espacio acumulado canónico** de $[k]^n$ es

$$K_n := \bigcup_{j=1}^{n} [k]^j.$$

Dado $w \in K_n$ con

$$w = x_1x_2\cdots x_m \quad m \leq n,$$

definimos

$$\phi^*(w) := l_1(x_1)\cdots l_j(x_m).$$

En este sentido, si

$$N_j := \sum_{i=1}^{j} |l_i| \qquad 1 \leq j \leq n,$$

entonces

$$\phi^* : K_n \to \bigcup_{j=1}^{n} [k]^{N_j}$$

es función.

Como anteriormente, usamos la notación $\mathbb{P}_{\phi^*}$ para denotar a la distribución sobre $\left(X^{K_n}, \Sigma^{K_n}\right)$ determinada por sus distribuciones finito-dimensionales

$$\mathbb{P}_{\phi^*}\left(X_{w_1} \in A_1, \cdots, X_{w_j} \in A_j\right) = \mathbb{P}\left(X_{\phi^*(w_1)} \in A_1, \cdots, X_{\phi^*(w_j)} \in A_j\right)$$

donde

$$w_i \in K_n \quad i = 1, \cdots, j$$

y

$$(A_i)_{1\leq i\leq j}$$

son medibles arbitrarios. Por otro lado, denotamos por $\mathbb{P}_{\#\phi^*}$ a la probabilidad proyectada de $\mathbb{P}$ sobre $\left(X^{\phi^*}, \Sigma^{\phi^*}\right)$. Naturalmente,

$$\left(X^{K_n}, \Sigma^{K_n}, \mathbb{P}_{\phi^*}\right) \overset{\mathbb{P}}{\backsim} \left(X^{\phi^*}, \Sigma^{\phi^*}, \mathbb{P}_{\#\phi^*}\right).$$

Definición. Decimos que $\mathbb{P}$ es **fuertemente estacionaria** si, para todo subespacio, la medida $\mathbb{P}_\psi$ en $(X^\omega, \Sigma^\omega)$ determinada por sus distribuciones finito-dimensionales

$$\mathbb{P}_\psi(X_{w_1} \in A_1, \ldots, X_{w_n} \in A_n) = \mathbb{P}\big(X_{\psi(w_1)} \in A_1, \ldots, X_{\psi(w_n)} \in A_n\big)$$

coincide con $\mathbb{P}$.

Recordemos que

$$\left(X^\psi, \Sigma^\psi, \mathbb{P}_{\#\psi}\right) \overset{\mathbb{P}}{\backsim} (X^\omega, \Sigma^\omega, \mathbb{P}_\psi).$$

La condición es equivalente a la siguiente formulación finitaria:

Definición. Decimos que $\mathbb{P}$ es **finitamente estacionaria** si para todo $n \geq 1$ y para cualquier par de subespacios acumulados n-dimensionales ψ y γ, se tiene

$$\mathbb{P}_{\psi^*} = \mathbb{P}_{\gamma^*}.$$

Es decir,

$$\mathbb{P}\big(X_{\psi(w_1)} \in A_1, \ldots, X_{\psi(w_m)} \in A_m\big) = \mathbb{P}\big(X_{\gamma(w_1)} \in A_1, \ldots, X_{\gamma(w_m)} \in A_m\big)$$

para

$$w_1, \ldots, w_m \in K_n$$

y

$$(A_i)_{1 \leq i \leq m}$$

medibles arbitrarios en Σ.

Verificamos la equivalencia de ambas definiciones:

Teorema 26. *Sea $\mathbb{P}$ una distribución sobre $(X^\omega, \Sigma^\omega)$. $\mathbb{P}$ es fuertemente estacionaria si y sólo si es finitamente estacionaria.*

Demostración. (Fuertemente estacionaria $\Rightarrow$ Finitamente estacionaria)

Sea $\mathbb{P}$ medida de probabilidad fuertemente estacionaria sobre

$$(X^\omega, \Sigma^\omega)$$

y sean ϕ_1 y ϕ_2 dos subespacios n-dimensionales. Ambos ϕ_1 y ϕ_2 pueden extenderse a dos subespacios infinito-dimensionales, digamos ν_1 y ν_2.

En efecto, sean

$$(l_i)_{1 \leq i \leq n}$$

las líneas combinatorias que generan a ϕ_1 y

$$(s_i)_{1\leq i\leq n}$$

las que generan a ϕ_2. Definimos dos sucesiones de líneas combinatorias

$$(u_i)_{i\geq 1}$$

y

$$(v_i)_{i\geq 1}$$

con

$$u_i = l_i \quad 1 \leq i \leq n$$
$$u_i = \mathrm{Id} \quad i > n$$

y

$$v_i = s_i \quad 1 \leq i \leq n$$
$$v_i = \mathrm{Id} \quad i > n.$$

Les llamamos ν_1 y ν_2 a los subespacios generados por $(u_i)_i$ y $(v_i)_i$, respectivamente.

Por hipótesis,

$$\mathbb{P} = \mathbb{P}_{\nu_1} = \mathbb{P}_{\nu_2}.$$

Tomando proyecciones sobre X^{K_n} concluímos que

$$\mathbb{P}_{\#K_n} = \mathbb{P}_{\phi_1^*} = \mathbb{P}_{\phi_2^*}.$$

(Fuertemente estacionaria $\Leftarrow$ Finitamente estacionaria)

Sea $\mathbb{P}$ distribución sobre $(X^\omega, \Sigma^\omega)$ finitamente estacionaria. Consideramos una cadena de extensiones

$$\left(X^{\phi_n^*}, \Sigma^{\phi_n^*}\right)_{n\geq 1}$$

donde cada ϕ_n es n-subespacio y

$$\ldots \subset \phi_n^* \subset \phi_{n+1}^* \subset \ldots.$$

La sucesión

$$(\phi_n^*)_{n\geq 1}$$

determina al subespacio

$$\bigcup_n \phi_n^*,$$

y cada subespacio determina una sucesión bajo proyecciones sobre las K_n. A cada $X^{\phi_n^*}$ le asignamos la medida de probabilidad $\mathbb{P}_{\#\phi_n^*}$, de manera que nuestra cadena de extensiones a cada paso es $\mathbb{P}$-isomorfa a

$$\left(X^{K_n}, \Sigma^{K_n}, \mathbb{P}_{\phi_n^*}\right).$$

Ya que $\mathbb{P}$ es finitamente estacionaria, la sucesión de medidas $\left(\mathbb{P}_{\phi_n^*}\right)_n$ no depende de la elección de n-subespacios.

Las funciones

$$\pi_{mn} : X^{K_m} \to X^{K_n} \quad m \geq n$$

que determinan a las extensiones son las proyecciones naturales.

El límite proyectivo correspondiente a las π_{mn} es el de las sucesiones

$$(\bar{x}_{a_1}, \bar{x}_{a_2}, \ldots) \quad \bar{x}_{a_j} \in [k]^j,$$

donde

$$\pi_{(n+1)n}\left(\bar{x}_{a_{n+1}}\right) = \bar{x}_{a_n} \quad \forall n \geq 1.$$

Es decir, que la proyección de $\bar{x}_{a_{n+1}}$ sobre K_n es justamente $\bar{x}_{a_n}$. Entonces, el límite proyectivo ¡no es más que X^ω! O, para ser precisos, el límite proyectivo $\tilde{X}$ satisface

$$\tilde{X} \overset{\mathbb{P}}{\backsim} X^\omega.$$

Por el teorema de consistencia de Kolmogorov existe una única distribución $\mathbb{P}_\omega$ tal que

$$\mathbb{P}_{\# K_n} = \mathbb{P}_\omega \circ \pi_n^{-1}$$

donde π_n es la proyección de X^ω sobre X^{K_n}.

Acabamos de demostrar que el subespacio

$$\psi = \bigcup_n \phi_n^*$$

satisface

$$X^\psi \overset{\mathbb{P}}{\backsim} X^\omega.$$

Por la unicidad de $\mathbb{P}_\omega$ en el párrafo anterior, el resultado es extensible a cualquier subespacio infinito-dimensional de $[k]^\omega$ y la cadena de n-subespacios acumulados que lo generan. □

En la demostración anterior, para la segunda implicación nos apoyamos de la construcción de X^ω como límite proyectivo de manera un tanto innecesaria. La siguiente demostración es más directa. Conservamos las dos versiones pues nos parecen ilustrativas.

Finitamente estacionaria $\Leftarrow$ Fuertemente estacionaria. Sea $\mathbb{P}$ distribución sobre $(X^\omega, \Sigma^\omega)$ finitamente estacionaria. Sea ψ subespacio de $[k]^\omega$. Para cada colección

$$w_1, \cdots, w_m \in [k]^\omega$$

existe N suficientemente grande tal que

$$w_1, \cdots, w_m \in K_N.$$

Por hipótesis,

$$\mathbb{P}_{\#K_N}(X_{w_1} \in A_1, \cdots, X_{w_m} \in A_m) = \mathbb{P}_{\psi\#K_N}(X_{w_1} \in A_1, \cdots, X_{w_m} \in A_m)$$

para $A_1, \cdots, A_m$ medibles arbitrarios. Pero el lado izquierdo de la ecuación es simplemente

$$\mathbb{P}(X_{w_1} \in A_1, \cdots, X_{w_m} \in A_m)$$

y el derecho

$$\mathbb{P}_\psi(X_{w_1} \in A_1, \cdots, X_{w_m} \in A_m)\,.$$

Recordamos que la distribución $\mathbb{P}_\psi$ está determinada por las medidas de sus cilindros finito-dimensionales

$$\mathbb{P}_\psi(X_{w_1} \in A_1, \cdots, X_{w_m} \in A_m) = \mathbb{P}\big(X_{\psi(w_1)} \in A_1, \cdots, X_{\psi(w_m)} \in A_m\big)$$

con

$$(A_i)_{1\leq i\leq m}$$

medibles arbitrarios. $\square$

Planteamos una pregunta abierta: si en la definición de estacionariedad finitaria, sustituyéramos subespacios acumulados n-dimensionales simplemente por n-subespacios, tendríamos una propiedad más débil. Si a esta propiedad la bautizamos como **cuasi-estacionariedad finitaria**, cabe la pregunta: ¿Qué relación guardan las propiedades de estacionariedad fuerte y cuasi-estacionariedad finitaria? No intentamos aquí responder esta pregunta, pero nos parece que no son equivalentes.

4.1. Ejemplos de procesos fuertemente estacionarios. Para entender mejor qué es un proceso fuertemente estacionario, conviene brindar algunos ejemplos. Comenzamos con dos ejemplos triviales.

Ejemplo 1 Consideramos el proceso X^ω donde las X_w son independientes e idénticamente distribuídas. En este caso, cualquier subconjunto infinito J de $[k]^\omega$ induce un proceso X^J con la misma distribución que $X^{[k]^\omega}$.

Ejemplo 2 Otro proceso fuertemente estacionario es aquél donde $X_w = X$ para toda $w \in [k]^\omega$. Es decir, todas son la misma variable aleatoria.

Los anteriores ejemplos no nos dicen gran cosa acerca de la estructura de un proceso fuertemente estacionario; las hipótesis son demasiado exigentes como para dejarnos ver la estructra que queremos comprender. Antes de dar un ejemplo más sustancioso, definimos un objeto que nos ha de ser útil.

DEFINICIÓN. Entendemos por **árbol** una gráfica conexa sin ciclos simples, y por **árbol dirigido** una gráfica dirigida que sería un árbol si ignorásemos las direcciones de sus aristas, a la que además le pedimos que haya un vértice distinguido llamado **ancestro**, o **vértice** 0, tal que todas las aristas se dirigen en la dirección opuesta al ancestro.

DEFINICIÓN. Dado un vértice λ en un árbol, denotamos por $\pi(0,\lambda)$ al trayecto único entre el ancestro y λ. Al número de aristas en $\pi(0,\lambda)$ le llamamos el **nivel** de λ y le abreviamos con $|\lambda|$.

DEFINICIÓN. El **padre** de λ es el único vértice de nivel $|\lambda|-1$ que comparte arista con λ. Los **hijos** de λ son aquellos vértices de nivel $|\lambda|+1$ que comparten arista con λ. Los hijos del mismo padre son **hermanos**. Escribimos

$$\overleftarrow{\lambda}$$

para denotar al padre de λ.

Los hijos de los hermanos del padre de un vértice son sus **primos**. Dos primos cuya última palabra sea la misma son **primos afines**. Al conjunto de vértices con la misma cantidad de letras le llamamos **generación**.

DEFINICIÓN. Un k**-árbol** es un árbol donde todos los vértices tienen exactamente k hijos.

Llamaremos **rama** a una subgráfica conexa de un k-árbol. Llamaremos **ruta** a una sucesión de vértices

$$(\lambda_n)_{n\geq 0}$$

donde λ_{n+1} es hijo de λ_n para toda n.

A un k-árbol lo podemos asociar con $[k]^\omega$ de la siguiente manera. Al ancestro lo asociamos con la palabra "vacía", que en estricto sentido no está en $[k]^\omega$. Cada hijo del ancestro corresponde a alguna de las k palabras consistentes en una sóla letra. La segunda generación consiste en las k^2 palabras formadas por dos letras, de manera que cada una de ellas sea "nieta"del ancestro e hija de de la palabra de una letra que coincide con ella en la primera posición. Recursivamente, el n-ésimo nivel del árbol consiste en las k^n palabras compuestas de n letras, y cada una de ellas es hija de la palabra correspondiente de $n-1$ letras que coincide con ella en las primeras $n-1$ posiciones.

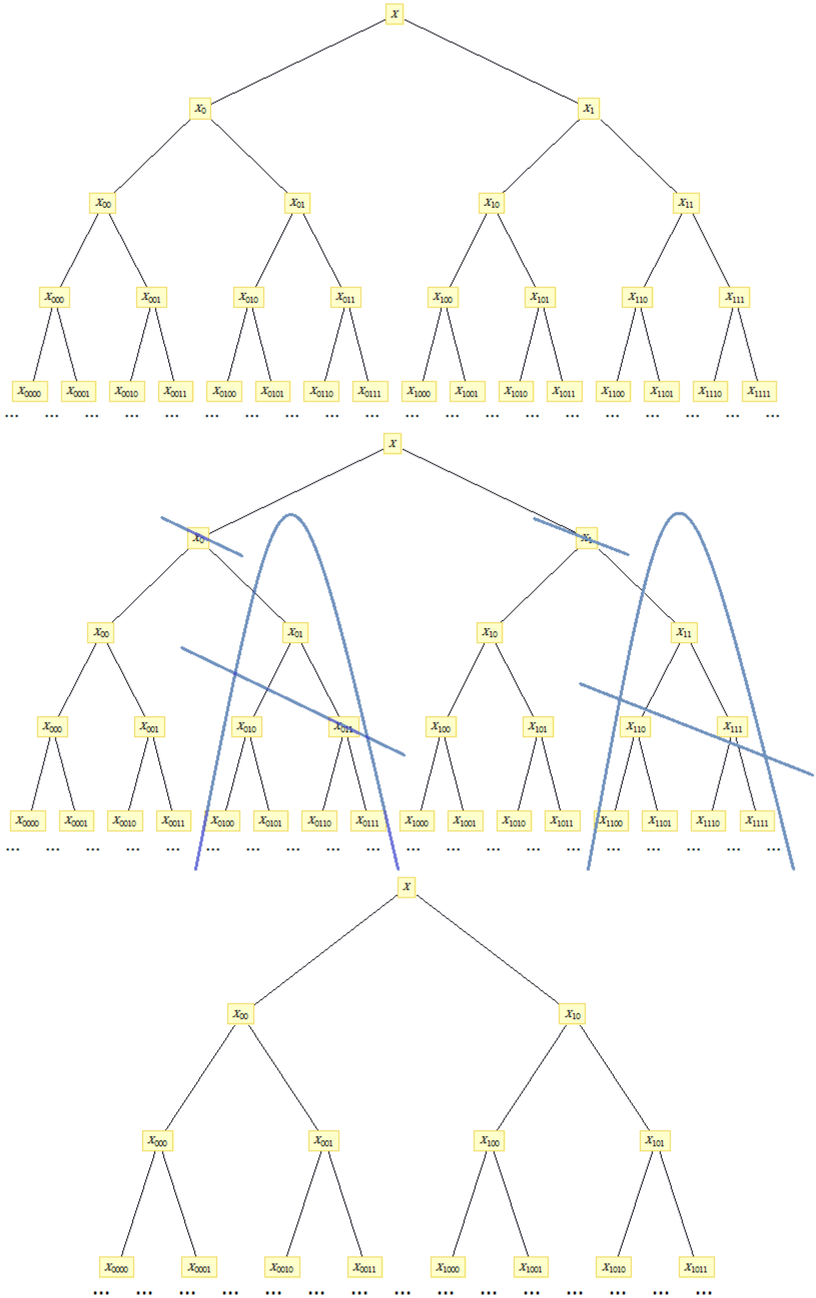
Ilustramos el caso $k=2$ en la figura 1. Ejemplificamos, además, la generación de un k-árbol que representa a un subespacio: en este caso, el subespacio en $[2]^\omega$ generado por

$$(l_n)_{n\geq 1}$$

con

$$l_1(j) = j0 \quad \forall j \in [k]$$
$$l_m = \mathrm{Id}, \quad \forall m \geq 2.$$

Figura 1. 2-árbol bien podado



Notamos que el resultado no es, en estricto sentido, un subarbol del k-árbol inicial, ya que tuvimos que agregar nuevas arista para conectar al vértice X con los vértice X_{00} y X_{10}, respectivamente. Esto ocurrirá toda vez que nuestro subespacio sea propio, y para cada ocasión en que una de las líneas combinatorias que generan el subespacio impongan requisitos o cortes. Nótese también que los k-árboles construidos de esta manera tienen cortes simétricos: siempre que cortamos una rama a partir de un vértice, hay que cortar las ramas correspondientes a sus primos afines.

Definición. Si a un k-árbol le realizamos cortes respetando las observaciones anteriores y añadiendo las aristas necesarias, obtenemos un k-árbol cuyos vértices son los elementos de un subespacio de $[k]^\omega$. Decimos que éste nuevo k-árbol está **bien podado**.

No todo k-árbol obtenido con cortes de un k-árbol original está bien podado. La figura 2 ilustra un ejemplo de un k-árbol que no representa a un subespacio de $[k]^\omega$. Nótese que no se respeta que todas las palabras en el mismo nivel tengan la misma extensión, cosa que sí debe cumplir todo k-árbol bien podado.

Si consideramos un k-árbol donde los vértices son variables aleatorias, y por la distribución de un k-árbol entendemos la distribución conjunta de sus variables aleatorias, podemos enunciar una formulación gráfica de estacionariedad fuerte:

Definición. Un proceso indexado en $[k]^\omega$ es **fuertemente estacionario** si para el k-árbol que lo representa, todo k-árbol bien podado tiene la misma distribución.

La definición no nos dice nada nuevo sobre estacionariedad fuerte, pero nos ayuda a visualizar mejor los ejemplos.

Definición. Diremos que una distribución $\mathbb{P}$ sobre $\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}\right)$ es **esparcible** siempre que

$$\mathbb{P}_J = \mathbb{P}$$

para cualquier subconjunto infinito

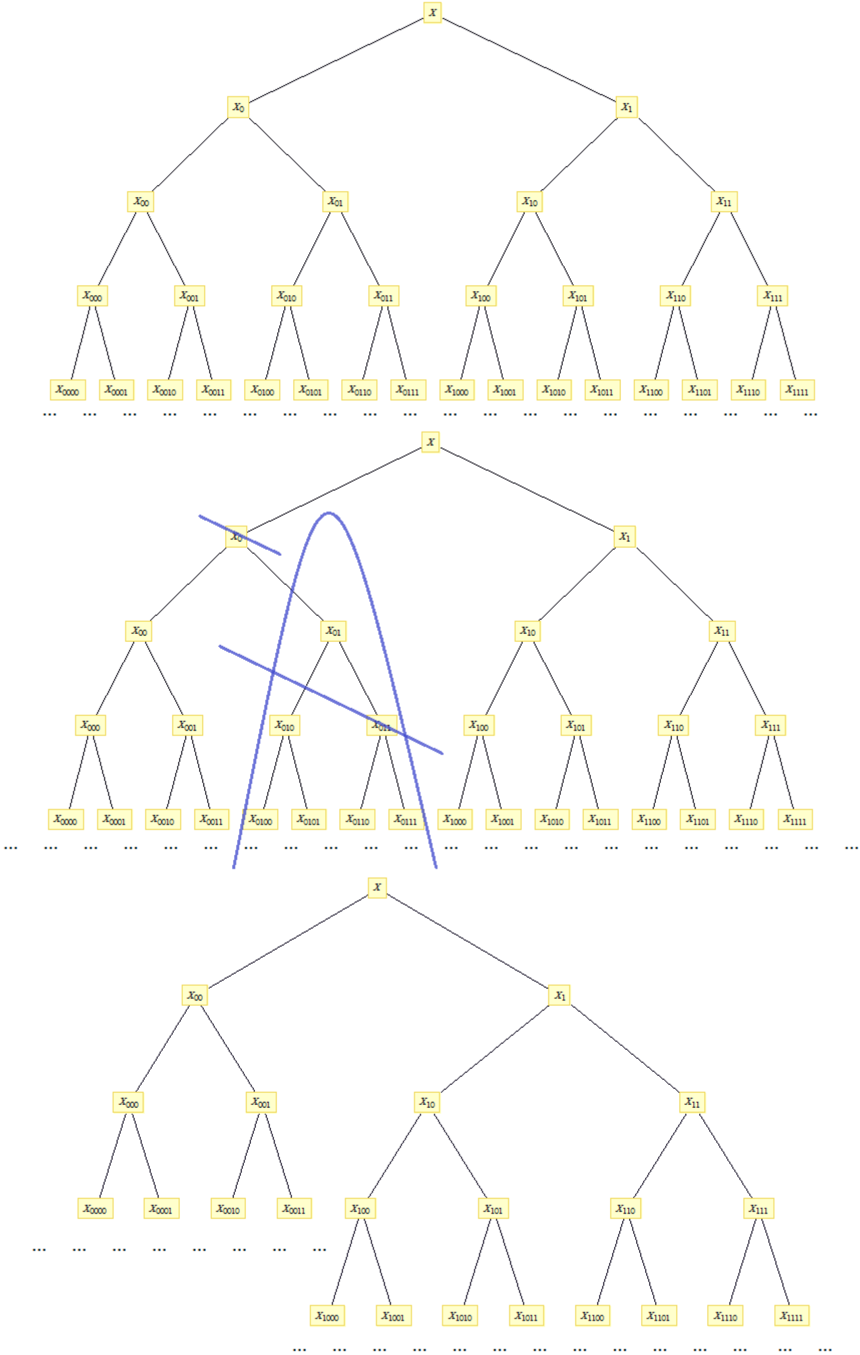
$$J \subset \mathbb{N}.$$

Equivalentemente, diremos que el proceso estocástico

$$(\pi_n(x))_{n\geq 1} \qquad x \in X^{\mathbb{N}}$$

es esparcible.

FIGURA 2. 2-árbol mal podado



PROPOSICIÓN 16. *Sea $\mathscr{A}$ un k-árbol fuertemente estacionario, y sean*

$$X_\lambda \quad \lambda \in ver(\mathscr{A})$$

los espacios de probabilidad asociados a cada vértice. Sea η una ruta de $\mathscr{A}$. La probabilidad $\mathbb{P}_\eta$ sobre el espacio

$$(X^\eta, \Sigma^\eta)$$

es esparcible.

DEMOSTRACIÓN. La demostración se comprende mejor gráficamente e invitamos al lector a que desarrolle en papel los siguientes pasos:

1) Elegir un subconjunto $J \subset \eta$ tal que $\eta \setminus J$ sea infinito.

2) Construir un k-árbol bien podado cortando cada uno de los vértices en J, más los respectivos vértices necesarios a cada nivel para conservar la simetría del k-árbol.

3) Verificar que $\eta \setminus J$ es una rama del k-árbol bien podado, y que el proceso definido por los espacios de probabilidad correspondientes a sus vértices tiene la misma distribución que el proceso correspondiete a η en $\mathscr{A}$. $\square$

DEFINICIÓN. Diremos que el proceso

$$(\pi_n(x))_{n \geq 1} \qquad x \in X^{\mathbb{N}}$$

es **intercambiable** siempre que

$$\mathbb{P}(X_1 \in A_1, \ldots X_n \in A_n) = \mathbb{P}\big(X_{\rho(1)} \in A_1, \ldots, X_{\rho(n)} \in A_n\big)$$

para cualquier permutación finita

$$\rho : [n] \to [n]$$

y

$$(A_i)_{1 \leq i \leq n}$$

medibles arbitrarios.

DEFINICIÓN. Además, diremos que los espacios de probabilidad

$$(X_n)_{n \geq 1}$$

son **condicionalmente independientes e idénticamente distribuídos** si para alguna subσ-álgebra

$$\mathscr{G} \subset \Sigma^{\mathbb{N}}$$

se cumple

$$\mathbb{P}(X_{\rho_1} \in A, \ldots, X_{\rho_n} \in A | \mathscr{G}) = \mathbb{P}(X_{\rho_1} \in A | \mathscr{G})^n$$

para cualesquiera

$$\rho_1, \ldots, \rho_n \in \mathbb{N}$$

y

$$A \in \Sigma.$$

El siguiente teorema nos permite dar una condición necesaria para procesos fuertemente estacionarios.

TEOREMA 27 (de Finetti, [**Kal02**], p. 212). *Para un proceso estocástico*

$$\left(X^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \mathbb{P}\right)$$

definido sobre un espacio de Borel las siguientes condiciones son equivalentes:

i) $\left(\pi_n(x)\right)_n \quad x \in X^{\mathbb{N}}$ *es esparcible*

ii) $\left(\pi_n(x)\right)_n \quad x \in X^{\mathbb{N}}$ *es intercambiable*

iii) Los espacios de probabilidad $\left(X_n\right)_n$ *son condicionalmente independientes e idénticamente distribuídas.*

La Proposición 16 y el Teorema de de Finetti implican que cada ruta de un k-árbol que represente a un proceso fuertemente estacionario induce un proceso intercambiable, esparcible y condicionalmente i. i. d.

Una pregunta interesante cuya respuesta no abarcaremos en este trabajo, es si el inverso es cierto. Es decir, un k-árbol tal que todas sus rutas inducen procesos esparcibles ¿será siempre fuertemente estacionario? Nos parece posible.

Pasemos a otros ejemplos. Por facilidad de la exposición en lo que resta de los ejemplos adoptamos la notación

$$\left(X_w\right)_{w \in [k]^\omega}$$

para designar al proceso estocástico que hasta ahora hemos designado con

$$\left(\pi_w(x)\right)_{w \in [k]^\omega} \qquad x \in X^{[k]^\omega},$$

y ya no al espacio de probabilidad.

Ejemplo 3.

Fijamos $k = 2$. Sean

$$U_0,$$

$$U_1,$$

$$U_2^0, U_2^1,$$

$$U_3^{00}, U_3^{01}, U_3^{10}, U_3^{11},$$

$$U_4^{000}, U_4^{001}, U_4^{010}, U_4^{011}, U_4^{100}, U_4^{101}, U_4^{110}, U_4^{111},$$

$$\ldots$$

variables aleatorias independientes e idénticamente distribuídas, uniformes en el intervalo $[0,1]$. Es decir que para toda $n \geq 1$ consideramos la lista

$$(U_n^w)_{w\in[2]^{n-1}}\,.$$

Definimos

$$X_\emptyset = U_0$$
$$X_0 = \mathbf{1}_{U_0\in[0,\frac{1}{2})}2U_0 + \mathbf{1}_{U_0\in[\frac{1}{2},1]}U_1$$

y de manera análoga

$$X_1 = \mathbf{1}_{U_0\in[0,\frac{1}{2})}U_1 + \mathbf{1}_{U_0\in[\frac{1}{2},1]}\left(2U_0 - 1\right).$$

Pensando en el k-árbol que representa al proceso $X^{[k]^\omega}$, U_0 se hereda con su respectiva expansión a alguno de los hijos de $X_\emptyset$ en la primera generación, X_0 o X_1, según en la mitad del intervalo en donde haya caído. El hermano que no haya heredado a U_0 inicia una nueva uniforme independiente.

El proceso se sigue recursivamente; así, por ejemplo,

$$X_{00} = \mathbf{1}_{X_0\in[0,\frac{1}{2})}2X_0 + \mathbf{1}_{X_0\in[\frac{1}{2},1]}U_2^0$$

y

$$X_{01} = \mathbf{1}_{X_0\in[0,\frac{1}{2})}U_2^0 + \mathbf{1}_{X_0\in[\frac{1}{2},1]}\left(2X_0 - 1\right).$$

Aquí, U_2^0 designa a la uniforme independiente del segundo nivel del k-árbol que determinará a alguno de los dos hijos del vértice indexado por 0. En general, U_n^w designa a la uniforme independiente del n-ésimo nivel del k-árbol que determinará a alguno de los dos hijos del vértice indexado por w, con $w \in [2]^{n-1}$.

De manera recursiva, para

$$w \in [2]^{\omega,}, \quad |w| = n-1$$

definimos

$$X_{w0} = \mathbf{1}_{X_w\in[0,\frac{1}{2})}2X_w + \mathbf{1}_{X_w\in[\frac{1}{2},1]}U_n^w$$

y

$$X_{w1} = \mathbf{1}_{X_w\in[0,\frac{1}{2})}U_n^w + \mathbf{1}_{X_w\in[\frac{1}{2},1]}\left(2X_w - 1\right).$$

Dejamos al lector verificar que cualquier 2-árbol bien podado obtenido a partir del 2-árbol que corresponde al proceso $X^{[2]^\omega}$ tiene la misma distribución conjunta que el 2-árbol original. En efecto, esto es consecuencia de que las U_w sean independientes e idénticamente distribuidas.

Hacemos una observación: sean

$$\eta := (\eta_n)_{n\geq 1}$$

y

$$\theta := (\theta_n)_{n\geq 1}$$

dos rutas de un k-árbol y sea

$$m = \text{máx}\{n : \eta_n = \theta_n\}.$$

Escribimos en este caso

$$\eta \wedge \theta$$

para designar al vértice de máximo nivel que tienen en común η y θ, a saber

$$\eta_m = \theta_m.$$

Como consecuencia de la independencia de las U_w^n, los procesos

$$\left(X_{\eta_{m+n}}\right)_{n\geq 1}$$

y

$$\left(X_{\theta_{m+n}}\right)_{n\geq 1}$$

son independientes e idénticamente distribuidos.

En efecto, uno esperaría que la obsevación anterior fuese cierta, ya que de otra manera no podrían ser esparcibles las rutas del 2-árbol.

Ejemplo 4.

Repetimos la construcción anterior pero fraccionando el intervalo $[0,1]$ en k partes iguales y definiendo las uniformes i. i .d. necesarias para los vértices de un k-árbol, de manera que la definición recursiva de nuestras variables aleatorias para

$$w \in [k]^{\omega}, \quad |w| = n-1$$

se vuelve

$$X_{wj} = k\mathbf{1}_{X_w \in [\frac{j}{k}, \frac{j+1}{k})}\left(X_w - \frac{j}{k}\right) + \mathbf{1}_{X_w \notin [\frac{j}{k}, \frac{j+1}{k})} U_n^w, \quad j = 1, \cdots, k.$$

De manera que $k-1$ de los k hijos de cada vértice inician nuevas uniformes i.i.d.

Ejemplo 5 [**MSW05**].

Consideramos el proceso de secuencias de Pólya definidas sobre las rutas de un k-árbol de la siguiente manera:

Si $w \in [k]^n$ y $m \leq n$ denotamos por $w\#m$ a la restricción canónica de w a $[k]^m$; es decir

$$w\#m = w_1 w_2 \cdots w_m$$

si

$$w = w_1 w_2 \cdots w_m \cdots w_n.$$

Sea $X_{\#0}$ una Bernoulli de parámetro

$$\frac{a}{a+b}.$$

Al ancestro le asociamos el espacio $\{0,1\}$ con la distribución de $X_{\#0}$. A cada hijo del ancestro,

$$X_w \quad w = 1, \cdots, k$$

le asociamos el espacio $\{0,1\}$ con la distribución Bernoulli de parámetro

$$\frac{a + X_{\#0}}{a+b+1}.$$

Le llamamos X_w la variable aleatoria asociada a w. Recursivamente, a los hijos de $w \in [k]^n$ les asociamos a cada uno la Bernoulli

$$X_{wj} \quad j = 1, \cdots, k$$

de parámetro

$$\frac{a + \sum_{i=0}^{n} X_{w\#i}}{a+b+n+1}$$

donde adoptamos la convención

$$X_{\#0} = X_{w\#0} \quad \forall w.$$

De esta manera, cada camada de hermanos tienen Bernoullis independientes e idénticamente distribuídas, condicionadas a la Bernoulli del padre.

Se comprende mejor el proceso con la siguiente imagen mental: cada vértice del árbol tiene colgada una urna de a bolas rojas y b bolas verdes. Como regla general, para cada urna, cada una de sus bolas tiene la misma probabilidad de ser extraída. Comenzamos extrayendo una bola de la urna que cuelga del ancestro y añadimos, a cada una de las urnas que cuelgan de sus hijos, una bola de ese color. Para cada hijo, extraemos una bola de su urna aleatoriamente, y añadimos, a las urnas de los hijos de éste, las bolas necesarias (dos) para que la composición inicial de sus urnas sea igual a la de su padre. Es decir, a cada paso retiramos al azar una bola de la urna del padre, y las urnas de los hijos serán iguales a la del padre antes de la extracción, más una bola del color de la que fue retirada.

De esta manera,

$$P_w = \frac{a + \sum_{i=1}^{n} X_{w\#i}}{a+b+n+1}$$

representa la proporción de bolas rojas en las urnas de los hijos de $w \in [k]^n$, antes de la extracción. Notamos que, para una palabra infinita w^* el proceso

$$(P_{w^*\#n})_{n\geq 1}$$

es una martingala acotada. Lo que es decir que, para una ruta en el árbol $\eta = (\eta_n)_n$, el proceso

$$(P_{\eta_n})_{n\geq 1}$$

es martingala acotada. Entonces converge casi seguramente y en L_∞ a una variable aleatoria P_{η_∞} que se puede interpretar como la proporción límite de bolas rojas en la urna.

El proceso

$$(X_{\eta_n})_{n\geq 0}$$

es una secuencia de Pólya. Un resultado conocido de urnas de Pólya nos dice que las

$$X_{\eta_0}, X_{\eta_1}, X_{\eta_2}, \ldots$$

son independientes e idénticamente distribuídas, condicionadas a P_{η_∞}. Por de Finetti, el proceso que definen es esparcible. Además, P_{η_∞} tiene distribución Beta con parámetros (a, b).

Notamos por último que si dos rutas coinciden en los primeros N vértices; es decir

$$|\eta \wedge \theta| = N,$$

entonces los procesos

$$\left(X_{\eta_{N+1}}, X_{\eta_{N+2}}, \ldots\right)$$

y

$$\left(X_{\theta_{N+1}}, X_{\theta_{N+2}}, \ldots\right)$$

son independientes e idénticamente distribuídos, condicionados a

$$(X_{\eta_1}, \cdots, X_{\eta_N}) = (X_{\theta_1}, \cdots, X_{\theta_N}).$$

Considerando lo anterior podemos ver que si $\mathscr{A}$ es el árbol original y $\mathscr{R}$ un árbol bien podado del mismo,

$$\{X_\lambda : \lambda \in \mathscr{A}\} \overset{\mathbb{P}}{\backsim} \{X_\lambda : \lambda \in \mathscr{R}\}.$$

4.2. Basta demostrar DHJ para procesos fuertemente estacionarios. El siguiente teorema nos permite hacer la reducción deseada del problema:

TEOREMA 28. *Para toda distribución $\mathbb{P}_\omega$ sobre $X^{[k]^\omega}$, existe una sucesión de subespacios $(\psi_m)_m$ tal que las probabilidadas proyectadas sobre ellos, $(\mathbb{P}_{\psi_m})_m$, convergen en la topología débil de medidas, a una distribución fuertemente estacionaria.*

Es importante observar que la distribución límite no necesariamente corresponde a un subespacio límite de la sucesión de refinaciones. De hecho, podemos construir una infinidad de sucesiones $(\psi_m([k]^\omega))_{m\geq 1}$ donde

$$\lim_{m\to\infty} \psi_m([k]^\omega)$$

es un subespacio *degenerado* que consiste en una sóla palabra. Por ejemplo, consideramos la sucesión

$$\begin{aligned}
\psi_1([2]^\omega) &= \{1x_2x_3\dots \quad : x_i \in [2]\}\\
\psi_2([2]^\omega) &= \{11x_3x_4\dots \quad : x_i \in [2]\}\\
&\dots\\
\psi_n([2]^\omega) &= \{11\dots 1x_{n+1}x_{n+2}\dots \quad : x_i \in [2]\}\\
&\dots
\end{aligned}$$

donde, claramente,

$$\lim_{m\to\infty} \psi_m([k]^\omega) = \bigcap_{m\geq 1} \psi_m([k]^\omega) = \{111111\dots\},$$

¡una sola palabra!

Antes de proceder a demostrar el teorema, exponemos dos resultados preliminares de los que habremos de hechar mano.

Teorema 29. *Sea $(X,\Sigma,\mathbb{P})$ un espacio de probabilidad y $(\mathscr{F}_n)_n$ una filtración en Σ. Definimos*

$$\mathscr{F}_\infty = \sigma(\mathscr{F}_1, \mathscr{F}_2, \dots).$$

Sea f una función Σ-medible en L_1. Entonces, conforme $k\to\infty$,

$$\mathbb{E}(f|\mathscr{F}_k) \to \mathbb{E}(f|\mathscr{F}_\infty)$$

$\mathbb{P}$-casi seguramente y en L_1.

Teorema 30 ([**McC99**], p. 61.). *Sea $n\in\mathbb{N}$ y ψ un subespacio. Para toda coloración finita de los n-subespacios en ψ, existe un subespacio $\gamma\subset\psi$, tal que todos los n-subespacios en γ tienen el mismo color.*

A cada n-subespacio ϕ le corresponde un y sólo un subespacio acumulado ϕ^*, ya que está completamente determinado por la sucesión de n líneas combinatorias que lo generan. De esta observación se desprende que cada coloración de n-subespacios en un subespacio infinito-dimensional ψ induce una coloración análoga de n-subespacios acumulados en ψ. De ahí el siguiente resultado:

TEOREMA 31 (Corolario). *Sea $n \in \mathbb{N}$ y ψ un subespacio. Para toda coloración finita de los n-subespacios acumulados en ψ, existe un subespacio $\gamma \subset \psi$, tal que todos los n-subespacios acumulados en γ tienen el mismo color.*

DEMOSTRACIÓN DEL TEOREMA 28. Comenzamos con el espacio de probabilidad

$$(X^\omega, \Sigma^\omega, \mathbb{P}) .$$

Elegimos una sucesión de natrurales donde cada natural se repite una infinidad de veces: por ejemplo,

$$(n_m)_{m\geq 0} = (1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots) .$$

El argumento de la demstración es el siguiente: a cada paso m de la sucesión, construímos un subespacio refinado ψ_m tal que todo n_m-subespacio acumulado perteneciente a la refinación tenga una distribución cada vez más parecida. La sucesión de subespacios obtenida de esta manera induce una sucesión de distribuciones que converge bajo la métrica de Prohorov a una distribución fuertemente estacionaria.

Construimos la sucesión de manera inductiva, comenzando con $\psi_0 = \mathrm{Id}([k]^\omega)$. En el paso $m > 0$ particionamos a $M\big(X^{K_{n_m}}\big)$, el espacio métrico de medidas de probabilidad sobre

$$\big(X^{K_{n_m}}, \Sigma^{K_{n_m}}\big) ,$$

en bolas de radio 2^{-m}. Como $M\big(X^{K_{n_m}}\big)$ es compacto, elegimos una cubierta abierta finita de bolas de radio 2^{-m}, y le llamamos

$$(B_i)_{i=1}^{r_m}$$

a esa colección de bolas. Establecemos una coloración del espacio asignando un color a cada uno de los conjuntos disjuntos

$$B_1, \quad B_j \smallsetminus \bigcup_{i=1}^{j-1} B_i, \quad j = 1, \cdots, r_m.$$

Ahora bien, notamos que el conjunto

$$\mathscr{P}_m = \Big\{\mathbb{P}_{\phi^*_{n_m}} : \phi^*_{n_m} \subset \psi_{m-1} \text{ es } n_m\text{-subespacio acumulado}\Big\}$$

está contenido en $M\big(X^{K_{n_m}}\big)$, de manera que tenemos una coloración finita de $\mathscr{P}_m$.

Más aún, la misma corresponde a una coloración análoga de

$$\mathscr{C}_m = \big\{\phi^*_{n_m} : \phi^*_{n_m} \subset \psi_{m-1} \text{ es } n_m\text{-subespacio acumulado}\big\} .$$

En efecto, a cada n_m-subespacio acumulado ϕ^* contenido en ψ_{m-1} le asociamos el color de la probabilidad $\mathbb{P}_{\phi^*}$. Que dos subespacios acumulados sean del mismo color indica que sus distribuciones distan a lo más 2^{-m} bajo la métrica de Prohorov.

Por el Teorema 31, existe un subespacio infinito-dimensional ψ_m contenido en ψ_{m-1} tal que todos los n_m-subespacios acumulados contenidos en él sean del mismo color.

Ya que cualquier natural n ocurre una infinidad de veces en la sucesión $(n_m)_m$, las medidas de probabilidad correspondientes a n-subespacios acumulados eventualmente caen dentro de una bola de radio 2^{-m}, para toda $m \in \mathbb{N}$. En particular,

$$(\mathbb{P}_{\psi_m \# K_n})_{m \geq 1}$$

es de Cauchy bajo la métrica de Prohorov sobre $M\big(X^{K_n}\big)$, para toda $n \geq 1$. Ya que por el Teorema 24 $M\big(X^{K_n}\big)$ es polaco, la sucesión converge para n arbitrario.

Por otro lado X^ω es límite proyectivo de espacios polacos medibles, y por el Teorema 20 también es polaco medible. Luego, aplicando una vez más el Teorema 24, $M(X^\omega)$ es espacio polaco. Queremos demostrar que

$$(\mathbb{P}_{\psi_m})_{m \geq 1}$$

converge bajo la métrica de Prohorov sobre $M(X^\omega)$; basta, entonces, demostrar que es de Cauchy.

Dado que X^ω es polaco, la topología que induce la métrica de Prohorov sobre $M(X^\omega)$ coincide con la topología débil. Queremos demostrar entonces que para cualquier $\epsilon > 0$ existe $N \geq 1$ suficientemente grande tal que para $n, q \geq N$ se satisface

$$\left\| \int f_i \, d\mathbb{P}_{\psi_n} - \int f_i \, d\mathbb{P}_{\psi_q} \right\| < \epsilon \tag{15}$$

para todo conjunto finito

$$\{f_i : \quad i \in \mathscr{A}\}$$

de funciones continuas y acotadas. Abreviamos la ecuación anterior con

$$\left\| \mathbb{E}_{\mathbb{P}_{\psi_n}}(f_i) - \mathbb{E}_{\mathbb{P}_{\psi_q}}(f_i) \right\| < \epsilon.$$

Consideramos la sucesión de subσ-álgebras

$$\mathscr{B}_n = \pi_n^{-1}\big(\Sigma^{K_n}\big) \subset \Sigma^\omega,$$

donde π_n es la proyección

$$\pi_n : \Sigma^\omega \to \Sigma^{K_n}$$

y notamos que

$$\Sigma^\omega = \mathscr{B}_1 \vee \mathscr{B}_2 \vee \ldots.$$

La sucesión

$$(\mathscr{F}_n)_{n\geq 1} \qquad \mathscr{F}_n = \mathscr{B}_1 \vee \cdots \vee \mathscr{B}_n$$

es filtración y

$$\mathscr{F}_\infty = \Sigma^\omega.$$

Por el Teorema 29, para f continua y acotada

$$\mathbb{E}_{\mathbb{P}_{\psi_k}}(f|\mathscr{F}_n) \to \mathbb{E}_{\mathbb{P}_{\psi_k}}(f), \quad n \to \infty$$

con $k \geq 0$ arbitrario. Recordamos que $\mathbb{P}_{\psi_0} = \mathbb{P}$.

Ahora bien, dado $N \geq 1$ tenemos

$$\begin{aligned}\left\|\mathbb{E}_{\mathbb{P}_{\psi_n}}(f_i) - \mathbb{E}_{\mathbb{P}_{\psi_q}}(f_i)\right\| \quad &\leq \quad \left\|\mathbb{E}_{\mathbb{P}_{\psi_n}}(f_i) - \mathbb{E}_{\mathbb{P}_{\psi_n}}(f_i|\mathscr{B}_N)\right\| \\ &+ \quad \left\|\mathbb{E}_{\mathbb{P}_{\psi_n}}(f_i|\mathscr{B}_N) - \mathbb{E}_{\mathbb{P}_{\psi_q}}(f_i|\mathscr{B}_N)\right\| \\ &+ \quad \left\|\mathbb{E}_{\mathbb{P}_{\psi_q}}(f_i|\mathscr{B}_N) - \mathbb{E}_{\mathbb{P}_{\psi_q}}(f_i)\right\|\end{aligned}$$

El primer y último término del lado derecho de la desigualdad anterior se pueden hacer arbitrariamente pequeños si elegimos N sufiicientemente grande, en virtud del Teorema 29. El segundo término es equivalente a

$$\left\|\mathbb{E}_{\mathbb{P}_{\psi_n \# K_N}}(f_i') - \mathbb{E}_{\mathbb{P}_{\psi_q \# K_N}}(f_i')\right\|$$

para

$$f_i' : X^{K_N} \to \mathbb{R}$$

continua y acotada que satisface

$$f_i = f_i' \circ \pi_n.$$

Pero la sucesión

$$(\mathbb{P}_{\psi_n \# K_N})_{n\geq 1}$$

converge, y dado que $n \geq N$, el segundo término del lado derecho de la desigualdad se puede hacer arbitrariamente pequeño eligiendo N suficientemente grande. Con esto concluímos que $(\mathbb{P}_{\psi_m})_m$ es de Cauchy y converge a

$$\mathbb{P}^* = \lim_{m\to\infty} \mathbb{P}_{\psi_m}.$$

Por construcción $\mathbb{P}^*$ es finitamente estacionaria y, por el Teorema 26, es fuertemente estacionaria. □

La importancia del teorema anterior radica en la siguiente implicación:

PROPOSICIÓN 17. *Si el Teorema 22 es válido para distribuciones fuertemente estacionarias, entonces es válido en general.*

Para establecer este resultado clave, usamos una de las equivalencias del bien conocido Teorema de Portmanteau:

Teorema 32 (Teorema de Portmanteau, [**Bil99**], p. 16). *Sea (X, Σ) un espacio polaco, donde Σ es la σ-álgebra generada por la topología $\mathcal{T}_d$, inducida por la métrica d sobre X. Sea $(\mathbb{P}_n)_n$ una sucesión de medidas de probabilidad sobre (X, Σ). Entonces $\mathbb{P}_n$ converge débilmente a $\mathbb{P}$ sí y sólo si*

$$\liminf_{n\geq 1} \mathbb{P}_n(A) \geq \mathbb{P}(A)$$

para todo $A \in \mathcal{T}_d$.

No revisaremos aquí la demostración del Teorema de Portmanteau.

Demostración de la Proposición 17. Sean $\delta > 0$ y $A \in \Sigma$ tal que $\mathbb{P}_\omega \circ \pi_w^{-1}(A) > \delta$ para todo $w \in [k]^\omega$. Es decir, A satisface las hipótesis del Teorema 22. Ahora bien, el espacio X con la σ-álgebra trivial inducida por A,

$$D := \{\emptyset, A, A^c, X\},$$

es isomorfo al espacio

$$\{0, 1\}$$

con la σ-álgebra discreta. Consideramos pues al espacio de probabilidad

$$Y^{[k]^\omega} := \{0, 1\}^{[k]^\omega}$$

con la medida de probabilidad inducida por $\mathbb{P}_\omega$; es decir,

$$\mathbb{P}_A(Y_{w_1} = 1, \cdots, Y_{w_m} = 1) := \mathbb{P}_\omega(X_{w_1} \in A, \cdots, X_{w_m} \in A)$$

$$w_1, \cdots, w_m \in [k]^\omega.$$

Tenemos así al espacio de probabilidad

$$\left(Y^{[k]^\omega}, D^{[k]^\omega}, \mathbb{P}_A\right),$$

donde $D^{[k]^\omega}$ es la σ-álgebra producto generada por los cilindros

$$\pi_w^{-1}(1).$$

Notamos que $Y^{[k]^\omega}$ es un espacio estándar de Borel sobre la topología que generan los mismos cilindros.

Por el Teorema 28, existe una sucesión de subespacios $(\psi_n)_{n\geq 1}$ con su respectiva sucesión de medidas de probabilidad inducidas $(\mathbb{P}_{\psi_n})$ sobre $\left(X^{[k]^\omega}, \Sigma^{[k]^\omega}, \mathbb{P}_\omega\right)$, tal que

$$\mathbb{P}_{\psi_n} \to \mathbb{P}^*$$

débilmente, con $\mathbb{P}^*$ fuertemente estacionaria.

A cada subespacio ψ_n corresponde también una medida de probabilidad $\mathbb{P}_{A,\psi_n}$ sobre el proceso Y^{ψ_n}, donde

$$\mathbb{P}_{A,\psi_n}(Y_{w_1} = 1, \cdots, Y_{w_m} = 1) = \mathbb{P}_{\psi_n}(X_{w_1} \in A, \cdots, X_{w_m} \in A)$$

$$w_1, \cdots, w_m \in [k]^\omega.$$

De la misma manera,

$$\mathbb{P}_{A,\psi_n} \to \mathbb{P}^*_A$$

débilmente, donde $\mathbb{P}^*_A$ es fuertemente estacionaria. Por unicidad de la convergencia, sabemos que

$$\mathbb{P}^*_A = \mathbb{P}^* \circ \mathbf{1}_A^{-1}.$$

Por hipótesis, el Teorema 22 es válido para procesos fuertemente estacionarios. Entonces, para algún $m \geq 1$ existe línea combinatoria

$$l([k]) \subset [k]^m$$

tal que

$$\mathbb{P}^*\left(\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(A)\right) > 0.$$

Luego también,

$$\mathbb{P}^*_A\left(\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(1)\right) > 0$$

para la misma línea combinatoria.

Más aún, el conjunto

$$\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(1)$$

pertenece a la topología sobre $Y^{[k]^\omega}$ generada por los cilindros

$$\left\{\pi^{-1}_w(1) : \quad w \in [k]^\omega\right\}.$$

Aplicando el Teorema de Portmanteau, tenemos

$$\liminf_{n \geq 1} \mathbb{P}_{A,\psi_n}\left(\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(1)\right) \geq \mathbb{P}^*_A\left(\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(1)\right),$$

así que para una infinidad de n's se cumple

$$\mathbb{P}_{A,\psi_n}\left(\bigcap_{i=1}^{k} \pi^{-1}_{l(i)}(1)\right) > 0.$$

Elegimos alguna n arbitraria que satisfaga la desigualdad anterior, y notamos que
$$\psi_n^{-1}(l([k]))$$
es línea combinatoria en $[k]^\omega$. Para concluír el resultado, resta sólo observar que
$$\mathbb{P}_{A,\mathrm{Id}}\left(\bigcap_{i=1}^{k}\pi^{-1}_{\psi_n^{-1}(l(i))}(1)\right)=\mathbb{P}_{A,\psi_n}\left(\bigcap_{i=1}^{k}\pi^{-1}_{l(i)}(1)\right),$$
y que
$$\mathbb{P}_\omega\left(\bigcap_{i=1}^{k}\pi^{-1}_{\psi_n^{-1}(l(i))}(A)\right)=\mathbb{P}_{A,\mathrm{Id}}\left(\bigcap_{i=1}^{k}\pi^{-1}_{\psi_n^{-1}(l(i))}(1)\right)>0.$$
□

5. Saciedad

La propiedad de saciedad nos permitirá demostrar DHJ con técnicas de teoría ergódica. Así como toda distribución $\mathbb{P}$ sobre (X^ω,Σ^ω) admite una sucesión de subespacios ψ_m cuyas probabilidades inducidas $\mathbb{P}_{\psi_m}$ tienden hacia la estacionariedad fuerte, X^ω admite una torre de extensiones que tienden hacia la saciedad. Antes de definir el concepto comenzamos con algunos preliminares.

5.1. σ-álgerbas invariantes.

DEFINICIÓN. Dado un subconjunto no vacío del abecedario $e\subset[k]$ y $\mathbb{P}$ una distribución fuertemente estacionaria sobre (X^ω,Σ^ω), la **σ-álgebra e-insensible** es la subσ-álgebra en Σ definida por
$$\Phi_e=\left\{A\in\Sigma:\ \mathbf{1}_A\circ\pi_{l(i)}=\mathbf{1}_A\circ\pi_{l(j)}\quad \mathbb{P}\text{-casi seguramente}\quad \forall i,j\in e\right\}$$
para cualquier línea combinatoria $l([k])$.

Notamos que por ser $\mathbb{P}$ fuertemente estacionaria, la distribución $\mathbb{P}_l$ es indiferente a la elección de línea combinatoria, y por ello la σ-álgebra está bien definida. Veamos:

Adoptamos la notación
$$x_w:=\pi_w(x)\quad x\in X^\omega.$$

Φ_e ES, EN EFECTO, σ-ÁLGEBRA. Si $A\in\Phi_e$, $l([k])$ es línea combinatoria y $i,j\in e$ entonces
$$\mathbf{1}_{A^c}\big(x_{l(i)}\big)=1-\mathbf{1}_A\big(x_{l(i)}\big)=1-\mathbf{1}_A\big(x_{l(j)}\big)=\mathbf{1}_{A^c}\big(x_{l(j)}\big)\quad \mathbb{P}\text{-c. s.},$$
de donde $A^c\in\Phi_e$

Si

$$\{A_n\}_{n\geq 1} \subset \Phi_e$$

y

$$A^* = \bigcup_{n\geq 1} A_n, \qquad A^k = \bigcup_{n=1}^{k} A_n$$

tenemos

$$\mathbf{1}_{A^*}\left(x_{l(i)}\right) = \sup_{k\geq 1}\left\{\mathbf{1}_{A^k}\left(x_{l(i)}\right)\right\} = \sup_{k\geq 1}\left\{\mathbf{1}_{A^k}\left(x_{l(j)}\right)\right\}$$

de donde

$$\mathbf{1}_{A^*}\left(x_{l(i)}\right) = \mathbf{1}_{A^*}\left(x_{l(j)}\right) \quad \mathbb{P}\text{-c.s.} \quad \forall i,j \in e$$

y por lo tanto $A^* \in \Phi_e$.

Se desprende de lo anterior que Φ_e es cerrado bajo intersecciones arbitrarias.

Por último notamos que Φ_e es siempre no vacía pues contiene al menos a la subσ-álgebra trivial $\{X, \emptyset\}$. $\square$

Observamos que si $e \subset e'$ entonces $\Phi_{e'} \subset \Phi_e$.

Definición. Decimos que una **distribución** $\mathbb{P}$ fuertemente estacionaria es e**-insensible** si

$$\Phi_e = \Sigma.$$

Es decir,

$$\mathbf{1}_A \circ \pi_{l(i)} = \mathbf{1}_A \circ \pi_{l(j)} \quad i,j \in e \quad \mathbb{P}\text{-c s.}, \quad A \in \Sigma.$$

Como vimos en la sección de espacios estándar de Borel, todo espacio X estándar de Borel tiene un canónico separable que le es σ-isomorfo, y al cual se le puede asignar la medida de probabilidad que mantenga el isomorfismo. Es decir, si $X \overset{\sigma}{\backsim} Y$ con

$$\lambda : \Sigma_X \to \Sigma_Y$$

biyección que preserva operaciones de σ-álgebra, entonces la distribución

$$\mathbb{P}_Y(A) := \mathbb{P}_X \circ \lambda^{-1}(A) \quad A \in \Sigma_Y$$

reproduce la estructura de probabilidad de X sobre Y.

Luego, suponemos sin pérdida de generalidad que X es separable y vemos que la condición de e-insensibilidad es equivalente a

$$x_{l(i)} = x_{l(j)} \quad \mathbb{P}\text{-c.s.}$$

ya que

$$\{x_w\} \in \Sigma \quad \forall x_w \in X.$$

Definición. Una **familia ascendente** $\mathscr{I}$ es es una colección de subconjuntos de $[k]$ tal que todos sus elementos constan de al menos dos letras en $[k]$ y

$$u \in \mathscr{I}, \quad u \subset v \Rightarrow v \in \mathscr{I}.$$

Para una familia ascendente $\mathscr{I}$, la **σ-álgebra invariante asociada a $\mathscr{I}$** es

$$\bigvee_{e \in \mathscr{I}} \Phi_e.$$

La **familia ascendente asociada a** $e \subset [k]$ es

$$\langle e \rangle := \{u \subset [k] : |u| \geq 2,\, e \subset u\}.$$

Para $i \in [k]$, abreviamos $\langle \{i\} \rangle$ con $\langle i \rangle$.

Nótese que $e \in \langle e \rangle$ si y sólo si $|e| \geq 2$.

5.2. Independencia condicional o relativa. La propiedad de saciedad se define con base en la independencia condicional de ciertas subσ-álgebras de un espacio producto.

Definición. Si $(X, \Sigma, \mathbb{P})$ es un espacio de probabilidad y $\mathscr{H}_1$, $\mathscr{H}_2$, $\mathscr{G}$ subσ-álgebras de Σ, decimos que $\mathscr{H}_1$ y $\mathscr{H}_2$ son **condicionalmente independientes con repecto a $\mathscr{G}$** si para cualesquiera variables aleatorias acotadas H_1 y H_2 tales que H_i es $\mathscr{H}_i$-medible se tiene

$$\mathbb{E}(H_1 H_2 | \mathscr{G}) = \mathbb{E}(H_1 | \mathscr{G}) \cdot \mathbb{E}(H_2 | \mathscr{G}).$$

En símbolos,

$$\mathscr{H}_1 \perp_{\mathscr{G}} \mathscr{H}_2.$$

De ahora en adelante usamos la notación común

$$\mathbb{E}(f \,|\, \mathscr{G}, \mathscr{H}) := \mathbb{E}(f \,|\, \sigma(\mathscr{G}, \mathscr{H})).$$

Podemos demostrar más directamente el resultado principal de esta sección utilizando el Lema de Clases Monótonoas de Dynkin.

Definición. Una familia de conjuntos $\mathscr{C} \subset \mathscr{P}(X)$ se llama **π-sistema** si es cerrada bajo intersecciones finitas.

Definición. Una familia de conjuntos $\mathcal{L} \subset \mathscr{P}(X)$ se llama **λ-sistema** si:

i) $X \in \mathcal{L}$
ii) $\mathcal{L}$ es cerrada bajo diferencias:

$$E, F \in \mathcal{L} \quad \text{y} \quad F \subset E \quad \Rightarrow \quad F \setminus E \in \mathcal{L}$$

iii) $\mathcal{L}$ es cerrada bajo uniones numerables crecientes:

$$(E_n)_{n\geq 1} \subset \mathcal{L}, \quad E_n \subset E_{n+1} \ \forall n \geq 1 \quad \Rightarrow \quad \bigcup_{n\geq 1} E_n \in \mathcal{L}.$$

TEOREMA 33 (Lema de Clases Monótonas de Dynkin). *Sean $\mathscr{C} \subset \mathscr{P}(X)$ un π-sistema y $\mathcal{L}_0$ un λ-sistema tales que $\mathscr{C} \subset \mathcal{L}_0$; entonces $\sigma(\mathscr{C}) \subset \mathcal{L}_0$.*

La demostración se imparte en cursos de teoría de la medida y la omitimos aquí. Nos interesa en particular una consecuencia de éste teorema:

PROPOSICIÓN 18 (Corolario). *Si $\mathscr{C} \subset \mathscr{P}(X)$ es π-sistema, entonces*

$$\sigma(\mathscr{C}) = \bigcap \{\mathcal{L} : \quad \mathcal{L} \quad \textit{es} \quad \lambda\textit{-sistema, y} \quad \mathscr{C} \subset \mathcal{L}\} =: \lambda(\mathscr{C})\,;$$

es decir que $\sigma(\mathscr{C})$ es el λ-sistema más pequeño que contiene a $\mathscr{C}$.

La siguiente proposición nos permite trabajar mejor con independencia condicional.

PROPOSICIÓN 19. *Las subσ-álgebras $\mathscr{H}_1$ y $\mathscr{H}_2$ son condicionalmente independientes relativas a $\mathscr{G}$ si y sólo si para toda variable aleatoria H_1 que sea $\mathscr{H}_1$-medible y acotada se tiene*

$$\mathbb{E}(\, H_1 \mid \mathscr{H}_2, \mathscr{G}) = \mathbb{E}(\, H_1 \mid \mathscr{G})\,.$$

DEMOSTRACIÓN. ($\Rightarrow$) Supongamos que

$$\mathscr{H}_1 \perp_{\mathscr{G}} \mathscr{H}_2.$$

Sea

$$\mathscr{L} := \left\{ A \in \Sigma : \quad \int_A H_1 \, d\mathbb{P} = \int_A \mathbb{E}(\, H_1 \mid \mathscr{G}) \; d\mathbb{P} \right\}.$$

Demostraremos que $\mathscr{L}$ es λ-sistema con

$$\mathscr{G}, \mathscr{H}_2 \subset \mathscr{L},$$

y que

$$\mathscr{A} := \{A \cap B : \quad A \in \mathscr{G}, \quad B \in \mathscr{H}_2\,\}$$

está contenido en $\mathscr{L}$ y es π-sistema. Entonces, por el lema de clases monótonas de Dynkin,

$$\sigma(\mathscr{A}) = \lambda(\mathscr{A}) \subset \lambda(\mathscr{L}) = \mathscr{L},$$

de donde

$$\sigma(\mathscr{G}, \mathscr{H}_2) \subset \mathscr{L}.$$

Concluímos de ello que

$$\mathbb{E}(\, H_1 \mid \mathscr{H}_2, \mathscr{G}) = \mathbb{E}(\, H_1 \mid \mathscr{G}) \quad \mathbb{P}\text{-c.s.}$$

Demostremos entonces lo necesario:

i) $\mathscr{L}$ es λ-sistema.

Claramente $X \in \mathscr{L}$. Si E, $F \in \mathscr{L}$ y $F \subset E$ tenemos

$$\begin{aligned}
\int_{E \smallsetminus F} H_1 \, d\mathbb{P} &= \int_E H_1 \, d\mathbb{P} - \int_F H_1 \, d\mathbb{P} \\
&= \int_E \mathbb{E}(H_1 \mid \mathscr{G}) \, d\mathbb{P} - \int_F \mathbb{E}(H_1 \mid \mathscr{G}) \, d\mathbb{P} \\
&= \int_{E \smallsetminus F} \mathbb{E}(H_1 \mid \mathscr{G}) \, d\mathbb{P}
\end{aligned}$$

de modo que $E \smallsetminus F \in \mathscr{L}$.

Sea

$$(E_n)_{n \geq 1} \subset \mathscr{L}$$

sucesión creciente. Definimos

$$\begin{aligned}
E'_1 &:= E_1 \\
E'_n &:= E_n \smallsetminus \bigcup_{j=1}^{n-1} E_j \quad n \geq 2
\end{aligned}$$

y notamos que

$$H_1^+ := \text{máx}\,\{H_1, 0\}$$

también es $\mathscr{H}_1$-medible y acotada. Tenemos

$$\begin{aligned}
T_n := \int_{\bigcup_{i=1}^n E_i} H_1^+ \, d\mathbb{P} &= \sum_{i=1}^n \int_{E'_i} H_1^+ \, d\mathbb{P} \\
&= \sum_{i=1}^n \int_{E'_i} \mathbb{E}\big(H_1^+ \mid \mathscr{G}\big) \, d\mathbb{P} \\
&= \int_{\bigcup_{i=1}^n E_i} \mathbb{E}\big(H_1^+ \mid \mathscr{G}\big) \, d\mathbb{P} =: S_n
\end{aligned}$$

Las sucesiones $(T_n)_n$ y $(S_n)_n$ son monótonas crecientes y acotadas con $T_n = S_n$ para toda n; por lo tanto convergen ambas a

$$T := \lim_{n \to \infty} T_n = \int_{\bigcup_n E_n} H_1^+ \, d\mathbb{P}.$$

Lo mismo se demuestra para

$$H_1^- := \text{máx}\,\{-H_1, 0\}$$

y a partir de la descomposición

$$H_1 = H_1^+ - H_1^-$$

obtenemos

$$\bigcup_{n\geq 1} E_n \in \mathscr{L}.$$

ii) $\mathscr{G}, \mathscr{H}_2 \subset \mathscr{L}$.

Claramente $\mathscr{G} \subset \mathscr{L}$. Sea $D \in \mathscr{H}_2$; luego

$$\begin{aligned}
\int_D H_1\, d\mathbb{P} = \int_X H_1 \cdot \mathbf{1}_D\, d\mathbb{P} &= \int_X \mathbb{E}(\, H_1 \mathbf{1}_D \mid \mathscr{G})\; d\mathbb{P} \\
&= \int_X \mathbb{E}(\, H_1 \mid \mathscr{G}) \cdot \mathbb{E}(\, \mathbf{1}_D \mid \mathscr{G})\; d\mathbb{P} \\
&= \int_X \mathbf{1}_D \mathbb{E}(\, H_1 \mid \mathscr{G})\; d\mathbb{P} = \int_D \mathbb{E}(\, H_1 \mid \mathscr{G})\; d\mathbb{P}
\end{aligned}$$

iii) $\mathscr{A}$ es π-sistema contenido en $\mathscr{L}$.

Claramente $\mathscr{A}$ es π-sistema. Sean $A \in \mathscr{G}$ y $B \in \mathscr{H}_2$; luego

$$\begin{aligned}
\int_{A\cap B} H_1\, d\mathbb{P} = \int_A H_1 \cdot \mathbf{1}_B\, d\mathbb{P} &= \int_A \mathbb{E}(\, H_1 \mathbf{1}_B \mid \mathscr{G})\; d\mathbb{P} \\
&= \int_A \mathbb{E}(\, H_1 \mid \mathscr{G}) \cdot \mathbb{E}(\, \mathbf{1}_B \mid \mathscr{G})\; d\mathbb{P} \\
&= \int_A \mathbb{E}(\, H_1 \mid \mathscr{G}) \cdot \mathbf{1}_B\, d\mathbb{P} \\
&= \int_{A\cap B} \mathbb{E}(\, H_1 \mid \mathscr{G})\; d\mathbb{P}
\end{aligned}$$

donde la igualdad de la segunda línea se cumple por ser $\mathbf{1}_B$ $\mathscr{H}_2$-medible y por la hipótesis de independencia condicional. Por lo tanto $A \cap B \in \mathscr{L}$ y

$$\mathscr{A} \subset \mathscr{L}.$$

($\Leftarrow$) Supongamos que

$$\mathbb{E}(\, H_1 \mid \mathscr{G}, \mathscr{H}_2) = \mathbb{E}(\, H_1 \mid \mathscr{G})$$

para H_1 acotada y $\mathscr{H}_1$-medible. Elegimos H_2 acotada y $\mathscr{H}_2$-medible. Notamos que

$$\begin{aligned}
\mathbb{E}(\, H_1 \cdot H_2 \mid \mathscr{G}) &= \mathbb{E}(\; \mathbb{E}(\, H_1 \cdot H_2 \mid \mathscr{G},\, \mathscr{H}_2) \;\mid \mathscr{G}) \\
&= \mathbb{E}(\; H_2 \cdot \mathbb{E}(\, H_1 \mid \mathscr{G},\, \mathscr{H}_2) \;\mid \mathscr{G}) \\
&= \mathbb{E}(\; H_2 \cdot \mathbb{E}(\, H_1 \mid \mathscr{G}) \mid \mathscr{G}) \\
&= \mathbb{E}(\, H_1 \mid \mathscr{G}) \cdot \mathbb{E}(\, H_2 \mid \mathscr{G})\, .
\end{aligned}$$

La primera igualdad es una aplicación directa de la propiedad de la torre de la esperanza condicional observando que

$$\mathscr{G} \subset \sigma(\mathscr{G}, \mathscr{H}_2)\,,$$

la segunda se debe a que H_2 es $\mathscr{H}_2$-medible y a la $\mathscr{H}_2$-homogeneidad de la esperanza condicional, la tercera se debe a la hipótesis y la cuarta es una aplicación de $\mathscr{G}$-homogeneidad, dado que $\mathbb{E}(H_1 \,|\, \mathscr{G})$ es $\mathscr{G}$-medible. $\square$

5.3. Los espacios $L_\infty(\mathbb{P}_w)$ y $L_2(\mathbb{P}_w)$. Cuando tengamos una distribución $\mathbb{P}$ sobre $(X^\omega, \Sigma^\omega)$, escribimos $\mathbb{P}_w$ para referirnos a la medida de probabilidad

$$\mathbb{P}_w = \mathbb{P} \circ \pi_w^{-1}$$

sobre X_w y distinguirlo de la medida sobre X^ω. Si el contexto no se presta a confusión también escribimos $\mathbb{P}$ para una medida de probabilidad sobre un espacio X no indexado.

Sea $w \in [k]^\omega$. Denotamos por $L_\infty(\mathbb{P}_w)$ al espacio de funciones

$$f : X_w \to \mathbb{R}$$

Σ_w-medibles y tales que para algún $K > 0$ se tiene

$$\mathbb{P}_w(\{x \in X_w : |f(x)\,| \geq K\}) = 0.$$

A este espacio le asignamos la norma

$$\|f\|_\infty = \text{ínf}\,\{k : \mathbb{P}_w(|f| \geq k) = 0\}\,.$$

$L_\infty(\mathbb{P}_w)$ es completo bajo esta norma. En estricto sentido, cada elemento de $L_\infty(\mathbb{P}_w)$ no es una función sino una clase de equivalencia de funciones que son iguales $\mathbb{P}_w$-casi seguramente. Así, al escribir $f \in L_\infty(\mathbb{P}_w)$ en realidad queremos decir

$$\{g : g = f \quad \mathbb{P}_w\text{-c.s.}\} \in L_\infty(\mathbb{P}_w)\,.$$

El abuso de notación que esto representa es práctica común y no causará mayor confusión.

El espacio $L_2(\mathbb{P}_w)$ consiste en las clases de equivalencia formadas por funciones Σ_w-medibles equivalentes $\mathbb{P}_w$-casi seguramente que satisfacen

$$\mathbb{E}_{\mathbb{P}_w}\left(|f|^2\right) < \infty.$$

$L_2(\mathbb{P}_w)$ es espacio normado completo bajo la norma

$$\|f\|_2 = \mathbb{E}_{\mathbb{P}_w}\left(|f|^2\right)^{\frac{1}{2}}\,.$$

Aquí también, no debe generar confusión que escribamos $f \in L_2(\mathbb{P}_w)$ en lugar de $[f] \in L_2(\mathbb{P}_w)$; es decir, denotamos con un representante de la clase de equivalencia a la clase en sí.

Los espacios $L_\infty(\mathbb{P}_w)$ y $L_2(\mathbb{P}_w)$ son normados y completos - ver [**Bar95**], capítulo 6.

Notamos que todo conjunto acotado en $L_\infty(\mathbb{P}_w)$ está acotado en $L_2(\mathbb{P}_w)$:

Demostración. Sea $A \subset L_\infty(\mathbb{P}_w)$ acotado. Luego, A está contenido en una bola de radio suficientemente grande, digamos $R > 0$, alrededor de la función constante cero. Denotamos por $B_R(0)$ a dicha bola. Para $f \in A$ tenemos

$$\mathbb{E}_{\mathbb{P}_w}\left(|f|^2\right) \leq R^2,$$

de donde $f \in L_2(\mathbb{P}_w)$. □

El siguiente resultado es necesario para reducir DHJ al caso de distribuciones saciadas.

Proposición 20. *Sea $(X, \Sigma, \mathbb{P})$ es estándar de Borel separable. La bola unitaria en $L_\infty(\mathbb{P})$,*

$$B_1(0) = \{f : \|f\|_\infty \leq 1\} \subset L_2(\mathbb{P})$$

contiene un denso numerable $(f_r)_{r\geq 1}$ bajo la métrica $\|\cdot\|_2$.

Demostración. Consideramos la clase $\mathcal{S}$ de todas las funciones simples dentro de la bola $B_1(0)$ con coeficientes racionales; es decir funciones de la forma

$$\sum_{j=1}^{n} r_j \cdot \mathbf{1}_{A_j}, \quad A_j \in \Sigma_w, \quad r_j \in \mathbb{Q} \cap [0,1).$$

Sabemos que $\mathcal{S}$ es denso en $B_1(0)$, puesto que cualquier función medible es aproximable por simples y cualquier simple es aproximable por funciones en $\mathcal{S}$. Basta entonces con demostrar que $\mathcal{S}$ contiene un denso numerable.

Sea $\mathscr{D}$ familia numerable en Σ_w tal que

$$\sigma(\mathscr{D}) = \Sigma_w$$

y $\mathscr{D}'$ el álgebra generada por $\mathscr{D}$. Afirmamos que la familia

$$\mathscr{C} = \left\{ \sum_{j=0}^{n} r_j \mathbf{1}_{A_j} : \quad r_j \in \mathbb{Q} \cap [0,1), \quad n \in \mathbb{N}, \quad (A_j)_{j=1}^{n} \subset \mathscr{D}' \right\}$$

es densa en $\mathcal{S}$.

En efecto, sea $C \in \Sigma_w$ de la forma

$$C = \bigcup_{j=1}^{\infty} A_j, \quad (A_j)_{j\geq 1} \subset \mathscr{D}'.$$

Definimos

$$A'_1 := A_1$$
$$A'_n := A_n \setminus A_{n-1} \quad n \geq 2,$$
$$E_n := \bigcup_{j=1}^{n} A'_j \quad \in \mathscr{D}',$$

de manera que C es unión disjunta de las A'_j y

$$\mathbb{P}(C \setminus E_n) = \mathbb{P}\left(\bigcup_{j=n+1}^{\infty} A'_j \right) \leq \sum_{j=n+1}^{\infty} \mathbb{P}\big(A'_j\big) \to 0$$

conforme $n \to \infty$, ya que

$$\sum_{j=1}^{\infty} \mathbb{P}\big(A'_j\big) = \mathbb{P}(C) \leq 1.$$

Tenemos

$$\begin{aligned} \|\mathbf{1}_C - \mathbf{1}_{E_n}\|_2 &= \mathbb{E}_{\mathbb{P}}\big(|\mathbf{1}_C - \mathbf{1}_{E_n}|^2\big)^{\frac{1}{2}} \\ &= \mathbb{P}(C \Delta E_n)^{\frac{1}{2}} \\ &= \mathbb{P}(C \setminus E_n)^{\frac{1}{2}} \\ &\to 0. \end{aligned}$$

Ya que $\mathbf{1}_{E_n} \in \mathscr{C}$ para $n \geq 1$, hemos demostrado que cada indicadora en $\mathcal{S}$ es aproximable por funciones en $\mathscr{C}$ bajo la norma $\| \cdot \|_2$. Pero toda función en $\mathcal{S}$ es suma finita de indicadoras y entonces es aproximable por sumas finitas de funciones en $\mathscr{C}$. Para concluír basta notar que $\mathscr{C}$ es cerrado bajo sumas finitas.

□

5.4. DHJ en saciedad. En esta subsección demostramos que DHJ es reducible al caso saciado. Recordemos que cuando un espacio de probabilidad $(Y, \Sigma_Y, \mathbb{P}_Y)$ es extensión de $(X, \Sigma_X, \mathbb{P}_X)$ decimos simplemente que $\mathbb{P}_Y$ es extensión de $\mathbb{P}_X$.

Consideramos solamente extensiones de la siguiente forma: sea $\mathbb{P}$ distribución sobre $(X^\omega, \Sigma^\omega)$ y $\tilde{\mathbb{P}}$ una extensión con

$$\mathbb{P} = \tilde{\mathbb{P}} \circ \pi^{-1}$$
$$\pi\Big(\tilde{X}^\omega\Big) = X^\omega,$$

entonces $\tilde{\mathbb{P}}$ induce la extensión índice por índice

$$\mathbb{P}_w = \tilde{\mathbb{P}}_w \circ \pi^{-1}$$
$$\pi\Big(\tilde{X}_w\Big) = X_w$$

indiferente a la elección de $w \in [k]^\omega$. En efecto, la siguiente proposición nos garantiza que toda distribución fuertemente estacionaria es de esta forma.

PROPOSICIÓN 21. *Si $\mathbb{P}$ es fuertemente estacionaria sobre $(X^\omega, \Sigma^\omega)$, la distribución $\mathbb{P}_w$ sobre (X, Σ) es indiferente a la elección de $\mathbb{P}_w$.*

DEMOSTRACIÓN. Si dos palabras w y v tienen alguna letra en común j, existen dos líneas combinatorias l_w y l_v tales que

$$l_w(j) = w \quad l_v(j) = v;$$

de donde

$$\mathbb{P}_w = \mathbb{P}_{l_w} \circ \pi_j^{-1} = \mathbb{P}_{l_v} \circ \pi_j^{-1} = \mathbb{P}_v.$$

En caso contrario, de todos modos existe una palabra que tenga letras en común con ambas w y v, y repetimos el argumento anterior dos veces. $\square$

Dadas una extensión $\tilde{\mathbb{P}}$ de $\mathbb{P}$ y una familia ascendente $\mathscr{I}$, denotamos con $\tilde{\Sigma}$ a la σ-álgebra asociada a $\tilde{X}$ y con $\tilde{\Phi}_{\mathscr{I}}$ a la subσ-álgebra invariante de $\tilde{\Sigma}$ asociada a $\mathscr{I}$.

DEFINICIÓN. Dadas una familia ascendente $\mathscr{I}$ y una distribución fuertemente estacionaria $\mathbb{P}$ sobre $(X^\omega, \Sigma^\omega)$ con σ-álgebras parcialmente insensibles

$$(\Phi_e)_{e\in[k]}\,,$$

decimos que la distribución $\mathbb{P}$ es **$\mathscr{I}$-saciada** si para toda extensión $\tilde{\mathbb{P}}$ de $\mathbb{P}$ **que conserve estacionariedad fuerte**, las subσ-álgebras $\pi^{-1}(\Sigma)$ y $\tilde{\Phi}_{\mathscr{I}}$ son condicionalmente independientes dada $\pi^{-1}(\Phi_{\mathscr{I}})$.

Intuitivamente, el resultado anterior nos dice que ninguna extensión de $\mathbb{P}$ nos puede inducir ninguna mayor información sobre la σ-álgebra $\mathscr{I}$-invariante. Esto se ve más claro con la siguiente proposición.

PROPOSICIÓN 22. *$\mathbb{P}$ es $\mathscr{I}$-saciada si y sólo si para toda extensión fuertemente estacionaria $\tilde{\mathbb{P}}$ y toda función*

$$f : \tilde{X} \to \mathbb{R}$$

acotada y $\pi^{-1}(\Sigma)$-medible se tiene

$$\mathbb{E}\Big(f \;\Big|\; \tilde{\Phi}_{\mathscr{I}},\, \pi^{-1}(\Phi_{\mathscr{I}})\Big) = \mathbb{E}\big(f \;\big|\; \pi^{-1}(\Phi_{\mathscr{I}})\big) \tag{16}$$

Demostración. El resultado es corolario inmediato de la Proposición 19. $\square$

En vista de la Proposición 20, es suficiente que se cumpla la igualdad (16) para toda $f \in L_2(\mathbb{P}_w)$ que sea $\pi^{-1}(\Sigma)$-medible.

Notamos, por otra parte, que el lado izquierdo de la ecuación (16) es redundante puesto que

$$\pi^{-1}(\Phi_{\mathscr{I}}) \subset \tilde{\Phi}_{\mathscr{I}}.$$

Así, podemos reescribir la ecuación (16) como

$$\mathbb{E}\Big(f \,\Big|\, \tilde{\Phi}_{\mathscr{I}} \Big) = \mathbb{E}\big(f \,|\pi^{-1}(\Phi_{\mathscr{I}})\big)\,. \tag{17}$$

Demostración. En efecto, si $e \in \mathscr{I}$, $\{i, j\} \subset e$ y $A \in \Phi_e$ tenemos

$$\tilde{\mathbb{P}}\Big(\Big\{(\tilde{x}_w)_{w\in[k]^\omega} : \mathbf{1}_{\pi^{-1}(A)}\big(\tilde{x}_{l(i)}\big) = \mathbf{1}_{\pi^{-1}(A)}\big(\tilde{x}_{l(j)}\big)\Big\}\Big) = \mathbb{P}\Big(\Big\{(x_w)_{w\in[k]^\omega} : \mathbf{1}_A\big(x_{l(i)}\big) = \mathbf{1}_A\big(x_{l(j)}\big)\Big\}\Big) = 1,$$

de donde

$$\pi^{-1}(A) \in \tilde{\Phi}_e.$$

Recordemos que

$$\tilde{\Phi}_{\mathscr{I}} = \bigvee_{e\in\mathscr{I}} \tilde{\Phi}_e;$$

ahora es inmediato el resultado. $\square$

Establecemos aún otra equivalencia que nos resultará necesaria:

Proposición 23. *Si $\tilde{\mathbb{P}}$ es extensión de $\mathbb{P}$ y*

$$f' : X \to \mathbb{R}$$

es Σ-medible, entonces

$$\mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big) = \mathbb{E}(f' \,|\, \Phi_{\mathscr{I}}) \circ \pi.$$

Demostración. Para todo $A \in \Phi_{\mathscr{I}}$

$$\int_A f'\, d\mathbb{P} = \int_{\pi^{-1}(A)} f' \circ \pi \, d\tilde{\mathbb{P}}.$$

Por otro lado,

$$\begin{aligned} \int_A f'\, d\mathbb{P} &= \int_A \mathbb{E}(f' \,|\, \Phi_{\mathscr{I}})\; d\mathbb{P} \\ &= \int_{\pi^{-1}(A)} \mathbb{E}(f' \,|\, \Phi_{\mathscr{I}}) \circ \pi \, d\tilde{\mathbb{P}} \end{aligned}$$

y

$$\int_{\pi^{-1}(A)} f' \circ \pi \, d\tilde{\mathbb{P}} = \int_{\pi^{-1}(A)} \mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big) \; d\tilde{\mathbb{P}}.$$

Juntando las equivalencias anteriores obtenemos el resultado. □

PROPOSICIÓN 24. *$\mathbb{P}$ es $\mathscr{I}$-saciada si y sólo si para toda extensión fuertemente estacionaria $\tilde{\mathbb{P}}$ y para toda función*

$$f' : X \to \mathbb{R}$$

acotada y Σ-medible se tiene

$$\mathbb{E}\Big(f' \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\Big) = \mathbb{E}(f' \,|\, \Phi_{\mathscr{I}}) \circ \pi. \tag{18}$$

DEMOSTRACIÓN. Toda función

$$f : \tilde{X} \to \mathbb{R}$$

acotada y $\pi^{-1}(\Sigma)$-medible puede escribirse como

$$f = f' \circ \pi \quad \text{con} \quad f' : X \to \mathbb{R}$$

acotada y Σ-medible. Por la Proposición 22, $\mathbb{P}$ es $\mathscr{I}$-saciada si y sólo si

$$\mathbb{E}\Big(f' \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\Big) = \mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big) \tag{19}$$

para toda f'. Pero, por la Proposición 23,

$$\mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big) = \mathbb{E}(f' \,|\, \Phi_{\mathscr{I}}) \circ \pi$$ □

A continuación demostramos dos resultados generales sobre independencia condicional que nos permitirán entender cómo es posible llegar a una distribución saciada a través de una cadena de extensiones.

PROPOSICIÓN 25. *Si $(X, \Sigma, \mathbb{P})$ es espacio de probabilidad, $\mathscr{G}, \mathscr{D} \subset \Sigma$ son subσ-álgebras tales que*

$$\mathscr{G} \subset \mathscr{D}$$

y $f : X \to \mathbb{R}$ es Σ-medible e integrable, entonces

$$\|\mathbb{E}(f \,|\, \mathscr{G})\|_2^2 \leq \|\mathbb{E}(f \,|\, \mathscr{D})\|_2^2 . \tag{20}$$

DEMOSTRACIÓN. Por la propiedad de la torre de la esperanza condicional tenemos

$$\mathbb{E}(f \,|\, \mathscr{G}) = \mathbb{E}(\; \mathbb{E}(f \,|\, \mathscr{D}) \;\; | \, \mathscr{G} \;) \, .$$

Aplicando la bien conocida *desigualdad de Jensen*,

$$\begin{aligned}\mathbb{E}(f\,|\,\mathscr{G})^2 = \mathbb{E}(\ \mathbb{E}(f\,|\,\mathscr{D})\ \ |\ \mathscr{G}\ )^2 \quad &\leq \quad \mathbb{E}\Big(\ \mathbb{E}(f\,|\,\mathscr{D})^2\ \ |\ \mathscr{G}\ \Big),\\ \text{e integrando:}\quad \mathbb{E}\Big(\ \mathbb{E}(f\,|\,\mathscr{G})^2\ \Big) \quad &\leq \quad \mathbb{E}\Big(\ \mathbb{E}\Big(\ \mathbb{E}(f\,|\,\mathscr{D})^2\ \ |\ \mathscr{G}\ \Big)\ \Big)\\ &= \quad \mathbb{E}\Big(\ \mathbb{E}(f\,|\,\mathscr{D})^2\ \Big) \quad \square\end{aligned}$$

TEOREMA 34. *Sea* $(X, \Sigma, \mathbb{P})$ *espacio de probabilidad. Dados* f *en* $L_2(\mathbb{P})$ *y*

$$\mathscr{G} \subset \mathscr{D} \qquad \textit{sub}\sigma\textit{-álgebras de } \Sigma,$$

las identidades

$$\|\mathbb{E}(\,f\ |\mathscr{G})\|_2^2 = \|\mathbb{E}(\,f\ |\mathscr{D})\|_2^2$$

y

$$\mathbb{E}(\,f\ |\mathscr{G}) = \mathbb{E}(\,f\ |\mathscr{D}) \quad \mathbb{P}\textit{-c.s.}$$

son equivalentes.

DEMOSTRACIÓN. Definimos

$$Y := \mathbb{E}(\,f\ |\mathscr{D})$$

y escribimos

$$Y = \mathbb{E}(\,Y\ |\mathscr{G}) + (Y - \mathbb{E}(\,Y\ |\mathscr{G}))\,,$$

que es una descomposición ortogonal de Y. No hace falta entrar en detalles de lo que esto significa, basta notar que, dado $A \in \mathscr{G}$

$$\begin{aligned}\mathbb{E}(\mathbf{1}_A \cdot (Y - \mathbb{E}(\,Y\ |\mathscr{G}))) \quad &= \quad \mathbb{E}(\mathbf{1}_A \cdot Y) - \mathbb{E}(\mathbf{1}_A \cdot \mathbb{E}(\,Y\ |\mathscr{G}))\\ &= \quad 0.\end{aligned}$$

Por linealidad de la integral, para toda función $\mathscr{G}$-simple S también

$$\mathbb{E}(S \cdot (Y - \mathbb{E}(\,Y\ |\mathscr{G}))) = 0.$$

Dada una función Z en $L_2(\mathbb{P})$ que sea $\mathscr{G}$-medible tenemos también

$$\mathbb{E}(Z \cdot (Y - \mathbb{E}(Y\,|\,\mathscr{G}))) = 0.$$

En efecto, existe una sucesión de funciones $\mathscr{G}$-simples

$$(Z_n)_{n\geq 1} \quad \text{tales que} \quad Z_n \to Z \quad \text{c.s.} \quad \text{y} \quad |Z_n| \leq |Z|.$$

Para cada $n \geq 1$ tenemos

$$\mathbb{E}(Z_n Y) = \mathbb{E}(Z_n \mathbb{E}(\,Y\ |\mathscr{G}))\,.$$

Notemos que

$$|Z_n Y| \leq |ZY|\,, \qquad |Z_n \mathbb{E}(\,Y\ |\mathscr{G})| \leq |Z\mathbb{E}(\,Y\ |\mathscr{G})|\,;$$

de donde

$$\mathbb{E}(|Z_n|\,|Y|) \le \mathbb{E}(|Z|\,|Y|) < \infty$$

y análogamente

$$\mathbb{E}(|Z_n|\,|\mathbb{E}(Y\mid\mathscr{G})|) \le \mathbb{E}(|Z|\,|\mathbb{E}(Y\mid\mathscr{G})|) < \infty,$$

de modo que

$$Z_nY, \quad Z_n\mathbb{E}(Y\mid\mathscr{G}) \in L_1(\mathbb{P})\,.$$

Ya que

$$Z_nY \to ZY \quad \text{c.s.} \quad \text{y} \qquad Z_n\mathbb{E}(Y\mid\mathscr{G}) \to Z\mathbb{E}(Y\mid\mathscr{G}) \quad \text{c.s.}$$

por el *teorema de convergencia dominada de Lebesgue* tenemos

$$\lim_{n\to\infty}\mathbb{E}(Z_nY) = \mathbb{E}(ZY)$$

$$\lim_{n\to\infty}\mathbb{E}(Z_n\mathbb{E}(Y\mid\mathscr{G})) = \mathbb{E}(Z\,\mathbb{E}(Y\mid\mathscr{G}))\,.$$

Entonces

$$\mathbb{E}(ZY) = \mathbb{E}(Z\,\mathbb{E}(Y\mid\mathscr{G}))$$

y con esto queda demostrada la afirmación.

Regresando a nuestra descomposición de Y tenemos

$$Y^2 = \mathbb{E}(Y\mid\mathscr{G})^2 + (Y-\mathbb{E}(Y\mid\mathscr{G}))^2 + 2\cdot\mathbb{E}(Y\mid\mathscr{G})\cdot(Y-\mathbb{E}(Y\mid\mathscr{G}))\,,$$

e integrando ambos lados de la igualdad y distribuyendo sumas,

$$\begin{aligned}\mathbb{E}\big(Y^2\big) &= \mathbb{E}\Big(\,\mathbb{E}(Y\mid\mathscr{G})^2\,\Big) + \mathbb{E}\Big(\,(Y-\mathbb{E}(Y\mid\mathscr{G}))^2\,\Big)\\ &+ \;2\cdot\mathbb{E}(\,\mathbb{E}(Y\mid\mathscr{G})\cdot(Y-\mathbb{E}(Y\mid\mathscr{G}))\,)\,.\end{aligned}$$

Pero $\mathbb{E}(Y\mid\mathscr{G})$ por definición es $\mathscr{G}$-medible y además está en $L_2(\mathbb{P})$; entonces

$$\mathbb{E}(\,\mathbb{E}(Y\mid\mathscr{G})\cdot(Y-\mathbb{E}(Y\mid\mathscr{G}))\,) = 0;$$

de ahí que

$$\mathbb{E}\big(Y^2\big) = \mathbb{E}\Big(\,\mathbb{E}(Y\mid\mathscr{G})^2\,\Big) + \mathbb{E}\Big(\,(Y-\mathbb{E}(Y\mid\mathscr{G}))^2\,\Big)\,. \tag{21}$$

Recordando quién es Y, la igualdad en la ecuación (20) es exactamente

$$\mathbb{E}\big(Y^2\big) = \mathbb{E}\Big(\,\mathbb{E}(Y\mid\mathscr{G})^2\,\Big)\,,$$

y se cumple si y sólo si

$$\mathbb{E}\Big(\,(Y-\mathbb{E}(Y\mid\mathscr{G}))^2\,\Big) = 0,$$

lo cual es cierto si y sólo si

$$Y = \mathbb{E}(Y\mid\mathscr{G}) \quad \mathbb{P}\text{-c.s.}$$

Es decir,

$$\mathbb{E}(f \mid \mathscr{D}) = \mathbb{E}(\mathbb{E}(f \mid \mathscr{D}) \mid \mathscr{G}) = \mathbb{E}(f \mid \mathscr{G})\,. \qquad \square$$

En particular,

$$\mathscr{D} := \sigma(\mathscr{G}, \mathscr{H}) \qquad \text{con} \quad \mathscr{H} \quad \text{sub}\sigma\text{-álgebra}$$

satisface las hipótesis de la Proposición 25 y el Teorema 34.

Ambos resultados tienen traducción en términos de saciedad. Veamos:

Definición (Notación). Trabajamos con dos espacios L_2 distintos. Uno comprende las funciones

$$f' : X \to \mathbb{R} \quad \text{que satisfacen} \quad \int |f'|^2 \, d\mathbb{P} < \infty$$

y el otro las funciones

$$f : \tilde{X} \to \mathbb{R} \quad \text{que satisfacen} \quad \int |f|^2 \, d\tilde{\mathbb{P}} < \infty;$$

los denotamos con $L_2(\mathbb{P})$ y $L_2(\tilde{\mathbb{P}})$, respectivamente. A sus normas las denotamos con

$$\|\cdot\|_{2,\mathbb{P}} \quad \text{y} \quad \|\cdot\|_{2,\tilde{\mathbb{P}}}\,.$$

Proposición 26. *Sean $\tilde{\mathbb{P}}$ extensión de $\mathbb{P}$ y*

$$f' : X \to \mathbb{R} \qquad en \quad L_2(\mathbb{P})\,;$$

entonces

$$\left\|\mathbb{E}\Big(f' \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\Big)\right\|^2_{2,\tilde{\mathbb{P}}} \geq \|\mathbb{E}(f' \,|\, \Phi_{\mathscr{I}})\|^2_{2,\mathbb{P}}\,, \tag{22}$$

Demostración. Primero notamos que

$$\|\mathbb{E}(f' \,|\, \Phi_{\mathscr{I}}) \circ \pi\|^2_{2,\tilde{\mathbb{P}}} = \|\mathbb{E}(f' \,|\, \Phi_{\mathscr{I}})\|^2_{2,\mathbb{P}}\,.$$

Luego, por la Proposición 23,

$$\left\|\mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big)\right\|^2_{2,\tilde{\mathbb{P}}} = \|\mathbb{E}(f' \,|\, \Phi_{\mathscr{I}})\|^2_{2,\mathbb{P}}\,.$$

Entonces queremos demostrar

$$\left\|\mathbb{E}\Big(f' \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\Big)\right\|^2_{2,\tilde{\mathbb{P}}} \geq \left\|\mathbb{E}\big(f' \circ \pi \,|\, \pi^{-1}(\Phi_{\mathscr{I}})\big)\right\|^2_{2,\tilde{\mathbb{P}}}\,, \tag{23}$$

pero esto consecuencia directa de la Proposición 25, dado que

$$\pi^{-1}(\Phi_{\mathscr{I}}) \subset \tilde{\Phi}_{\mathscr{I}}\,. \qquad \square$$

PROPOSICIÓN 27 (Corolario). *Si g y h son dos funciones Σ-medibles de X en $\mathbb{R}$, toda sucesión de extensiones*

$$\left(X_{(1)}, \Sigma_{(1)}, \mathbb{P}_{(1)}\right) \leftarrow \left(X_{(2)}, \Sigma_{(2)}, \mathbb{P}_{(2)}\right) \leftarrow \cdots \leftarrow \left(X_{(n)}, \Sigma_{(n)}, \mathbb{P}_{(n)}\right) \leftarrow \ldots,$$

que abreviamos con

$$\mathbb{P}_{(1)} \leftarrow \mathbb{P}_{(2)} \leftarrow \cdots \leftarrow \mathbb{P}_{(n)} \ldots,$$

induce sucesiones crecientes

$$\left(\left\|\mathbb{E}\big(g \mid \Phi_{\mathscr{I},(n)}\big)\right\|^2_{2,\mathbb{P}_{(n)}}\right)_{n\geq 0}, \quad \left(\left\|\mathbb{E}\big(h \mid \Phi_{\mathscr{I},(n)}\big)\right\|^2_{2,\mathbb{P}_{(n)}}\right)_{n\geq 0},$$

donde

$$\Phi_{\mathscr{I},(n)}$$

es la subσ-álgebra invariante de $\Sigma_{(n)}$ asociada a $\mathscr{I}$.

Si g y h son acotadas en $L_\infty(\mathbb{P})$, las sucesiones convergen.

DEMOSTRACIÓN. La primera parte es una aplicación recursiva de la Proposición 26. La segunda viene de observar que, si $\|g\|_\infty \leq K$, entonces

$$\left\|\mathbb{E}\big(g \mid \Phi_{\mathscr{I},(n)}\big)\right\|^2_{2,\mathbb{P}_{(n)}} \leq K \quad \forall n \geq 1.$$

Luego, por ser monótona creciente y acotada, la sucesión converge. Lo mismo se puede decir de h. □

PROPOSICIÓN 28. *$\mathbb{P}$ es $\mathscr{I}$-saciada si y sólo si para toda extensón fuertemente estacionaria $\tilde{\mathbb{P}}$ y para toda función*

$$f' : X \to \mathbb{R}$$

acotada y Σ-medible se tiene

$$\left\|\mathbb{E}\left(f' \circ \pi \,\middle|\, \tilde{\Phi}_{\mathscr{I}}\right)\right\|^2_{2,\tilde{\mathbb{P}}} = \left\|\mathbb{E}(f' \mid \Phi_{\mathscr{I}})\right\|^2_{2,\mathbb{P}}. \tag{24}$$

DEMOSTRACIÓN. Por el Teorema 34, dado $f_n \in L_2(\mathbb{P})$ la igualdad en (23), que es equivalente a (24), se cumple si y sólo si

$$\mathbb{E}\left(f_n \circ \pi \,\middle|\, \tilde{\Phi}_{\mathscr{I}}\right) = \mathbb{E}\big(f_n \circ \pi \mid \pi^{-1}(\Phi_{\mathscr{I}})\big) \quad \tilde{\mathbb{P}}\text{-c.s.}$$

Por la Proposición 20, dada f' acotada en $L_\infty(\mathbb{P})$ y Σ-medible, existe una sucesión $(f_n)_n$ de funciones en $L_2(\mathbb{P})$ que convergen a f' bajo la norma $\|\cdot\|_{2,\mathbb{P}}$;

de manera que si $\mathbb{P}$ es saciada

$$\begin{aligned}\mathbb{E}\Big(f'\circ\pi \,\Big|\, \tilde{\Phi}_{\mathscr{I}}\Big) &= \lim_{n\to\infty}\mathbb{E}\Big(f_n\circ\pi \,\Big|\, \tilde{\Phi}_{\mathscr{I}}\Big) \quad \text{en} \quad L_2\Big(\tilde{\mathbb{P}}\Big)\\ &= \lim_{n\to\infty}\mathbb{E}\big(f_n\circ\pi \,\big|\, \pi^{-1}(\Phi_{\mathscr{I}})\big) \quad \text{en} \quad L_2\Big(\tilde{\mathbb{P}}\Big)\\ &= \mathbb{E}\big(f'\circ\pi \,\big|\, \pi^{-1}(\Phi_{\mathscr{I}})\big)\end{aligned}$$

En vista de lo anterior el resultado es consecuencia de la Proposición 34. El regreso sigue los mismos pasos. □

El siguiente teorema es el resultado decisivo que nos permite reducir DHJ al caso saciado; si bien es más complicada, la demostración sigue el mismo espíritu de la del Teorema 28, que nos permitió reducir DHJ al caso fuertemente estacionario.

Teorema 35. *Sea $\mathscr{I}$ familia ascendente. Toda distribución fuertemente estacionaria sobre (X^ω,Σ^ω) tiene una extensión $\mathscr{I}$-saciada.*

Demostración. Sean $\mathbb{P}$ distribución fuertemente estacionaria sobre (X^ω,Σ^ω) y

$$(f_r)_{r\geq 1}$$

una sucesión de funciones densa en la bola unitaria

$$B_{1,\infty}(0)=\{f\in L_\infty(\mathbb{P}) : \|f\|_\infty\leq 1\},$$

bajo la norma $\|\cdot\|_{2,\mathbb{P}}$. Tal sucesión existe por la Proposición 20. Sean $(r_n)_n$ y $(m_n)_n$ sucesiones de naturales donde cada natural aparece una infinidad de veces y donde, además, todo par ordenado de naturales $(r,m)\in\mathbb{N}\times\mathbb{N}$ ocurre una infinidad de veces en

$$(r_n,m_n)_{n\geq 1}\,.$$

Por ejemplo, dada $(r_n)_n$ consideramos las subsucesiones

$$\left(r_{n_i^s}\right)_{i\geq 1} \qquad r_{n_i^s}=s \quad \forall i\geq 1, \quad s\geq 1$$

y fijamos

$$\left(m_{n_i^s}\right)_{i\geq 1}=(1,1,2,1,2,3,1,2,3,4,1,2,3,4,5,\ldots) \quad \forall s\geq 1.$$

Así

$$\mathbb{N}=\dot{\bigcup_{s\geq 1}}\,(n_i^s)_{i\geq 1}$$

y cada par ordenado $(s,m)\in\{s\}\times\mathbb{N}$ ocurre una infinidad de veces en $\left(r_{n_i^s},m_{n_i^s}\right)_{i\geq 1}$, para cada $s\in\mathbb{N}$.

A continuación construímos una cadena de extensiones fuertemente estacionarias

$$\left(\left(X_{(n)}^{\omega}, \Sigma_{(n)}^{\omega}, \mathbb{P}_n\right)\right)_{n\geq 0}$$

fijando

$$\left(X_{(0)}^{\omega}, \Sigma_{(0)}^{\omega}, \mathbb{P}_0\right) := (X^{\omega}, \Sigma^{\omega}, \mathbb{P})$$

y de manera que el límite proyectivo tenga una distribución $\mathscr{I}$-saciada. La construcción es recursiva de la siguiente manera:

Dados

$$\left(X_{(n-1)}^{\omega}, \Sigma_{(n-1)}^{\omega}, \mathbb{P}_{n-1}\right)$$

y la función Σ-medible y suprayectiva que determina la extensión,

$$\psi_{n-1} : X_{(n-1)} \to X$$
$$\mathbb{P}_{n-1} \circ \psi_{(n-1)}^{-1} = \mathbb{P},$$

consideramos al natural m_n y la función f_{r_n} que le corresponde al r_n-ésimo lugar en la sucesión $(f_r)_r$. Para determinar

$$\psi_n : X_{(n)} \to X$$

consideramos dos casos:

En el primer caso, existe una extensión fuertemente estacionaria

$$\pi : \tilde{X} \to X_{(n-1)}$$
$$\tilde{\mathbb{P}} \circ \pi^{-1} = \mathbb{P}_{n-1}$$

tal que

$$\left\|\mathbb{E}\left(f_{r_n} \circ \psi_{n-1} \circ \pi \mid \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 > \left\|\mathbb{E}\left(f_{r_n} \circ \psi_{n-1} \mid \Phi_{\mathscr{I},(n-1)}\right)\right\|_{2,\mathbb{P}_{n-1}}^2 + 2^{-m_n}.$$

Si $\mathscr{E}_{n-1}$ es el conjunto de las extensiones fuertemente estacionarias $\tilde{\mathbb{P}}$ que satisfacen la desigualdad anterior, y

$$D_n\left(\tilde{\mathbb{P}}\right) := \left|\left\|\mathbb{E}\left(f_{r_n} \circ \psi_{n-1} \circ \pi \mid \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 - \left\|\mathbb{E}\left(f_{r_n} \circ \psi_{n-1} \mid \Phi_{\mathscr{I},(n-1)}\right)\right\|_{2,\mathbb{P}_{n-1}}^2\right|,$$

elegimos $\tilde{\mathbb{P}} \in \mathscr{E}_{n-1}$ que satisfaga

$$D_n\left(\tilde{\mathbb{P}}\right) > \frac{1}{2} \sup\left\{D_n(\mathbb{P}') : \ \mathbb{P}' \in \mathscr{E}_{n-1}\right\}$$

y fijamos

$$\tilde{\mathbb{P}} \circ \pi^{-1} = \mathbb{P}_{n-1}$$
$$\psi_n := \psi_{n-1} \circ \pi$$
$$\left(X_{(n)}, \Sigma_{(n)}, \mathbb{P}_n\right) := \left(\tilde{X}, \tilde{\Sigma}, \tilde{\mathbb{P}}\right).$$

En el segunda caso, el conjunto $\mathscr{E}_{n-1}$ es vacío. Es decir, para toda extensión fuertemente estacionaria $\tilde{\mathbb{P}}$ de $\mathbb{P}_{n-1}$ se tiene

$$\left\|\mathbb{E}\left(f_{r_n} \circ \psi_{n-1} \circ \pi \mid \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 \leq \left\|\mathbb{E}\big(f_{r_n} \circ \psi_{n-1} \mid \Phi_{\mathscr{I},(n-1)}\big)\right\|_{2,\mathbb{P}_{n-1}}^2 + 2^{-m_n}.$$

En este caso fijamos

$$\begin{aligned} \mathbb{P}_n &:= \mathbb{P}_{n-1} \circ \mathrm{Id}_{X_{(n-1)}}^{-1} = \mathbb{P}_{n-1} \\ \psi_{(n)} &:= \psi_{(n-1)} \\ \big(X_{(n)}, \Sigma_{(n)}, \mathbb{P}_n\big) &:= \big(X_{(n-1)}, \Sigma_{(n-1)}, \mathbb{P}_{n-1}\big)\,. \end{aligned}$$

Es necesario que hagamos algunas observaciones para explicitar por qué nuestra construcción genera un límite proyectivo $\mathscr{I}$-saciado:

Denotamos con $\mathscr{K}$ al conjunto de todas las extensiones fuertemente estacionarias de $\mathbb{P}$ y con $\mathscr{K}_n$ al conjunto de todas las extensiones fuertemente estacionarias de $\mathbb{P}_n$. Observamos, primeramente, que

$$\begin{aligned} \sup_{\tilde{\mathbb{P}} \in \mathscr{K}} \left\|\mathbb{E}\left(f_n \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 &\leq \|f_n\|_\infty \cdot \tilde{\mathbb{P}}\left(\tilde{X}\right) \\ &\leq 1 \end{aligned}$$

ya que $f_n \in B_{1,\infty}(0)$. Definimos

$$M_n := \sup_{\tilde{\mathbb{P}} \in \mathscr{K}} \left\|\mathbb{E}\left(f_n \circ \pi \,|\, \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2, \quad 0 \leq M_n \leq 1.$$

Consideramos una función $f_s \in (f_{r_n})_{n\geq 1}$ y al subconjunto

$$(n_i^s)_{i\geq 1} \subset \mathbb{N}, \quad r_{n_i^s} = s \;\forall\, i \geq 1.$$

Queremos demostrar que

$$\left\|\mathbb{E}\left(f_s \circ \psi_{(n_i^s)} \,|\, \Phi_{\mathscr{I},(n_i^s)}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 \to M_s, \quad i \to \infty.$$

Por construcción, la sucesión

$$\left(\left\|\mathbb{E}\left(f_s \circ \psi_{(n_i^s)} \,|\, \Phi_{\mathscr{I},(n_i^s)}\right)\right\|_{2,\tilde{\mathbb{P}}}^2\right)_{i\geq 1}$$

es acotada y monótona creciente, por lo tanto converge. Consideramos dos casos: En el primero, $\mathscr{E}_{n_i^s} \neq \emptyset$ para una cantidad finita de índices $i \geq 1$. Esto significa que a partir de cierta extensión $\mathbb{P}_t$ se tiene

$$\left\|\mathbb{E}\left(f_s \circ \pi \mid \tilde{\Phi}_{\mathscr{I}}\right)\right\|_{2,\tilde{\mathbb{P}}}^2 = \left\|\mathbb{E}\big(f_s \circ \psi_{(t)} \mid \Phi_{\mathscr{I},(t)}\big)\right\|_{2,\mathbb{P}_t}^2 \quad \forall \tilde{\mathbb{P}} \in \mathscr{K}_t$$

y no hay nada que hacer.

En el segundo caso, $\mathscr{E}_{n_i^s} \neq \emptyset$ para una cantidad infinita de índices $i \geq 1$. Sea

$$d_s := M_s - \|\mathbb{E}(f_s \mid \Phi_{\mathscr{I}})\|_{2,\mathbb{P}}^2 .$$

Por construcción,

$$\begin{aligned} \left\|\mathbb{E}\big(f_s \circ \psi_{(n_u^s)} \mid \Phi_{\mathscr{I},(n_u^s)}\big)\right\|_{2,\mathbb{P}_{n_u^s}}^2 &= \|\mathbb{E}(f_s \mid \Phi_{\mathscr{I}})\|_{2,\mathbb{P}}^2 + \sum_{i=1}^{u} D_{n_i^s}\big(\mathbb{P}_{n_i^s}\big) \\ &\geq \|\mathbb{E}(f_s \mid \Phi_{\mathscr{I}})\|_{2,\mathbb{P}}^2 + \sum_{i=1}^{u} \frac{d_s}{2^i}, \end{aligned}$$

de modo que

$$\begin{aligned} M_s - \left\|\mathbb{E}\big(f_s \circ \psi_{(n_u^s)} \mid \Phi_{\mathscr{I},(n_u^s)}\big)\right\|_{2,\mathbb{P}_{n_u^s}}^2 &= d_s - \sum_{i=1}^{u} D_{n_i^s}\big(\mathbb{P}_{n_i^s}\big) \\ &\leq d_s - \sum_{i=1}^{u} \frac{d_s}{2^i} \\ &= d_s \cdot \sum_{i=u+1}^{\infty} \frac{1}{2^i} \\ &\to 0 \end{aligned}$$

conforme $u \to \infty$.

Pero $s \geq 1$ es arbitrario, así que para cualquier función en $(f_r)_r$ la esperanza condicional de f_r dada la subσ-álgebra invariante asciada a $\mathscr{I}$ converge bajo la torre de extensiones al supremo M_r.

Ahora, si $f \in B_{1,\infty}(0) \setminus (f_r)_r \subset L_2(\mathbb{P})$, $\epsilon > 0$ y $\tilde{\mathbb{P}}$ es extensión de $\mathbb{P}$ en la torre de extensiones que construímos, tenemos que para algún $f_r \in (f_r)_r$

$$\|f \circ \pi - f_r \circ \pi\|_{2,\tilde{\mathbb{P}}} < \epsilon.$$

Luego,

$$\left| \left\|\mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}} - \left\|\mathbb{E}\Big(f_r \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}} \right|$$

$$
\begin{aligned}
&\leq \left\| \mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) - \mathbb{E}\Big(f_r \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|_{2,\tilde{\mathbb{P}}} \\
&= \left(\int \left| \mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) - \mathbb{E}\Big(f_r \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right|^2 d\tilde{\mathbb{P}} \right)^{\frac{1}{2}} \\
&= \left(\int \left| \mathbb{E}\Big(f \circ \pi - f_r \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right|^2 d\tilde{\mathbb{P}} \right)^{\frac{1}{2}} \\
&= \left(\int \left| f \circ \pi - f_r \circ \pi \right|^2 d\tilde{\mathbb{P}} \right)^{\frac{1}{2}} \\
&< \epsilon
\end{aligned}
$$

de modo que hay una subsucesión

$$(f_{r_i})_{i \geq 1} \subset (f_r)_{r \geq 1}$$

tal que

$$\left\| \mathbb{E}\Big(f_{r_i} \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|_{2,\tilde{\mathbb{P}}} \quad \to \quad \left\| \mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|_{2,\tilde{\mathbb{P}}}, \quad i \to \infty$$

Por continuidad de

$$y \to y^2 \quad y \in \mathbb{R},$$

tenemos

$$\left\| \mathbb{E}\Big(f_{r_i} \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|^2_{2,\tilde{\mathbb{P}}} \quad \to \quad \left\| \mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|^2_{2,\tilde{\mathbb{P}}}, \quad i \to \infty.$$

Siendo así, definimos

$$M_f := \sup_{\tilde{\mathbb{P}} \in \mathscr{K}} \left\| \mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}} \Big) \right\|^2_{2,\tilde{\mathbb{P}}}$$

con miras a establecer que

$$\left\| \mathbb{E}\big(f \circ \psi_{(n)} \;\big|\; \Phi_{\mathscr{I},(n)} \big) \right\|^2_{2,\mathbb{P}_n} \to M_f, \quad n \to \infty.$$

Sea $(f_{s_i})_i \subset (f_r)_r$ una subsucesión que converge a f en $L_2(\mathbb{P})$. Por la desigualdad del triángulo,

(25)

$$
\begin{aligned}
M_f - &\left\| \mathbb{E}\big(f \circ \psi_{(n)} \;\big|\; \Phi_{\mathscr{I},(n)} \big) \right\|^2_{2,\mathbb{P}_n} \\
&\leq |M_f - M_s| \\
&+ \left| M_s - \left\| \mathbb{E}\big(f_s \circ \phi_{(n)} \;\big|\; \phi_{\mathscr{I},(n)} \big) \right\|^2_{2,\mathbb{P}_n} \right| \\
&+ \left| \left\| \mathbb{E}\big(f_s \circ \phi_{(n)} \;\big|\; \phi_{\mathscr{I},(n)} \big) \right\|^2_{2,\mathbb{P}_n} - \left\| \mathbb{E}\big(f \circ \psi_{(n)} \;\big|\; \Phi_{\mathscr{I},(n)} \big) \right\|^2_{2,\mathbb{P}_n} \right|
\end{aligned}
$$

y cada uno de estos términos tiende a cero conforme

$$n \to \infty \quad \text{y} \quad f_{s_i} \to f \quad \text{en} \quad L_2(\mathbb{P})\,, \quad i \to \infty.$$

En efecto, ya demostramos que los dos últimos términos son menores, digamos, a $\epsilon/3$ para n suficientemente grande y f_s suficientemente cercano a f. Resta ver qué pasa con el primer término. Veamos:

Por definición de M_f y por la Proposicion 27, existe una extensión fuertemente estacionaria $\tilde{\mathbb{P}}$ de $\mathbb{P}$ tal que

$$M_f - \left\|\mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2 < \frac{\epsilon}{9}, \quad M_s - \left\|\mathbb{E}\Big(f_s \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2 < \frac{\epsilon}{9}.$$

También podemos elegir f_s de la sucesión $(f_{r_i})_i$ de manera que

$$\left|\left\|\mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2 - \left\|\mathbb{E}\Big(f_s \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2\right| < \frac{\epsilon}{9};$$

luego

$$\begin{aligned}
|M_s - M_f| \quad &\leq \quad \left|M_f - \left\|\mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2\right| \\
&+ \quad \left|\left\|\mathbb{E}\Big(f \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2 - \left\|\mathbb{E}\Big(f_s \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2\right| \\
&+ \quad \left|\left\|\mathbb{E}\Big(f_s \circ \pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|_{2,\tilde{\mathbb{P}}}^2 - M_s\right| \\
&\leq \quad 3 \cdot \frac{\epsilon}{9} = \epsilon \cdot \frac{1}{3}.
\end{aligned}$$

Regresando a la ecuación 25,

$$M_f - \left\|\mathbb{E}\big(f \circ \psi_{(n)} \;\big|\; \Phi_{\mathscr{I},(n)}\big)\right\|_{2,\mathbb{P}_n}^2 \leq \epsilon$$

para ϵ arbitraria.

Hemos demostrado que toda función en $B_{1,\infty}(0)$ induce segundos momentos de la esperanza condicionada a la subσ-álgebra invariante asociada a $\mathscr{I}$ que convergen bajo la torre de extensiones a un supremo. Por linealidad de la esperanza condicional, el resultado es extensible a toda bola centrada

$$B_{R,\infty}(0) = R \cdot B_{1,\infty}(0) \subset L_\infty(\mathbb{P})$$

y, por lo tanto, a toda función acotada y Σ-medible de X en $\mathbb{R}$.

Para concluír falta sólo echar un vistazo a la estructura del límite proyectivo. Tenemos una torre de extensiones

$$X \leftarrow X_{(1)} \leftarrow X_{(2)} \leftarrow \cdots \leftarrow X_{(n)} \leftarrow \ldots$$

con

$$\psi_n\big(X_{(n)}\big) = X;$$

denotamos por

$$\psi_n^m : X_{(n)} \to X_{(m)} \quad m < n$$

a las funciones suprayectivas que definen las extensiones entre sus respectivos espacios. El límite proyectivo es el espacio de probabilidad

$$\left(X_{(\infty)}, \Sigma_{(\infty)}, \mathbb{P}_\infty\right),$$

donde

$$X_{(\infty)} = \left\{ (x_1, x_2, \ldots) \in \prod_{j=0}^{\infty} X_{(j)} \ : \quad x_n = \psi_{n+1}^n(x_{n+1}) \quad \forall n \geq 0 \right\}$$

con

$$X_{(0)} := X.$$

Si denotamos por

$$\begin{aligned} \rho_n : X_{(\infty)} &\to X_{(n)} \qquad n \geq 1 \\ \rho_0 : X_{(\infty)} &\to X \\ \psi_{(\infty)} &:= \rho_0 \end{aligned}$$

a las proyecciones naturales

$$\rho_n(x_1, x_2, \cdots, x_n, \ldots) = x_n$$

entonces

$$\Sigma_{(\infty)} = \sigma\big(\rho_n^{-1}\big(\Sigma_{(n)}\big)\,; n \geq 0\big)\,;$$

de modo que $\Sigma_{(\infty)}$ contiene una copia de cada $\Sigma_{(n)}$, que además conserva la masa de probabilidad en el sentido de que

$$\mathbb{P}_\infty\big(\rho_n^{-1}(A)\big) = \mathbb{P}_n(A) \quad \forall A \in \Sigma_{(n)}.$$

Es decir,

$$\mathbb{P}_\infty \circ \rho_n^{-1} = \mathbb{P}_n \quad \forall n \geq 0.$$

Por el teorema de consistencia de Kolmogorov, $\mathbb{P}_\infty$ es única y está bien definida.

Pues bien, retomando la Proposición 25, si f es Σ-medible y acotada tenemos

$$\left\|\mathbb{E}\big(f \circ \psi_{(\infty)} \;\big|\; \Phi_{\mathscr{I},(\infty)}\big)\right\|_{2,\mathbb{P}_\infty}^2 \geq \left\|\mathbb{E}\big(f \circ \psi_{(n)} \;\big|\; \Phi_{\mathscr{I},(n)}\big)\right\|_{2,\mathbb{P}_n}^2 \quad \forall n \geq 0,$$

lo cual, como ya demostramos, implica

$$\left\|\mathbb{E}\big(f \circ \psi_{(\infty)} \;\big|\; \Phi_{\mathscr{I},(\infty)}\big)\right\|_{2,\mathbb{P}_\infty}^2 = M_f.$$

Sea $\tilde{\mathbb{P}}$ extensión de $\mathbb{P}_\infty$. Por la Proposición 25 otra vez, y por definición de M_f,

$$\left\|\mathbb{E}\big(f\circ\psi_{(\infty)} \;\big|\; \Phi_{\mathscr{I},(\infty)}\big)\right\|^2_{2,\mathbb{P}_\infty} \leq \left\|\mathbb{E}\Big(f\circ\pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|^2_{2,\tilde{\mathbb{P}}} \leq M_f;$$

es decir,

$$\left\|\mathbb{E}\big(f\circ\psi_{(\infty)} \;\big|\; \Phi_{\mathscr{I},(\infty)}\big)\right\|^2_{2,\mathbb{P}_\infty} = \left\|\mathbb{E}\Big(f\circ\pi \;\Big|\; \tilde{\Phi}_{\mathscr{I}}\Big)\right\|^2_{2,\tilde{\mathbb{P}}}.$$

Recordando la Proposición 28, hemos demostrado que $\mathbb{P}_\infty$ es $\mathscr{I}$-saciada.

□

Definición. Si una distribución $\mathbb{P}$ es $\mathscr{I}$-saciada para toda familia ascendente $\mathscr{I}$, decimos que $\mathbb{P}$ es **netamente saciada**. Si el contexto no se presta a confusión, diremos indistintamente que dicha distribución es **saciada**.

Teorema 36. *Toda distribución $\mathbb{P}$ fuertemente estacionaria sobre $(X^\omega, \Sigma^\omega)$ admite una extensión netamente saciada.*

Hay dos argumentos paralelos para establecer este teorema a partir del anterior. En ambos indexamos a las familias ascendentes de conjuntos en $[k]$:

$$\mathscr{I}_1, \cdots, \mathscr{I}_T.$$

Primera demostración. Consideramos una sucesión

$$(r_n, m_n, t_n)_n$$

de elementos en $\mathbb{N}\times\mathbb{N}\times[T]$ donde cada trío posible ocurre una infinidad de veces. Dejamos al lector verificar que esto es posible. Repetimos la construcción de la torre de extensiones que hicimos para demostrar el Teorema 35, salvo que a cada paso consideramos las diferencias específicas a $\mathscr{I}_{t_n}$

$$D_n\Big(\tilde{\mathbb{P}}\Big) := \left|\left\|\mathbb{E}\Big(f_{r_n}\circ\psi_{n-1}\circ\pi \mid \tilde{\Phi}_{\mathscr{I}_{t_n}}\Big)\right\|^2_{2,\tilde{\mathbb{P}}} - \left\|\mathbb{E}\big(f_{r_n}\circ\psi_{n-1} \mid \Phi_{\mathscr{I}_{t_n},(n-1)}\big)\right\|^2_{2,\mathbb{P}_{n-1}}\right|$$

y la extensión correspondiente.

De esta forma,

$$\left\|\mathbb{E}\big(f \;\big|\; \Phi_{\mathscr{I}_t,(n)}\big)\right\|^2_{2,\mathbb{P}_n} \to M_{f,t} =: \sup_{\tilde{\mathbb{P}}\in\mathscr{K}} \left\|\mathbb{E}\Big(f \;\Big|\; \tilde{\Phi}_{\mathscr{I}_t}\Big)\right\|^2_{2,\tilde{\mathbb{P}}}, \quad n\to\infty$$

para toda $t\in[T]$ y para toda f acotada y Σ-medible, y el límite proyectivo es netamente saciado. □

SEGUNDA DEMOSTRACIÓN. A partir del Teorema 35 obtenemos un límite proyectivo $\mathscr{I}_1$-saciado de $\mathbb{P}$, al que denotamos por

$$\big(X_{(\infty)}, \Sigma_{(\infty)}, \mathbb{P}_\infty\big) =: (Y, \mathscr{F}, \mathbb{P}_Y)$$

y el cual es, a su vez, una extensión de $\mathbb{P}$ que admite una torre de extensiones

$$\mathbb{P}_Y \leftarrow \mathbb{P}_{Y,(1)} \leftarrow \mathbb{P}_{Y,(2)} \leftarrow \dots$$

tales que el límite proyectivo sea $\mathscr{I}_2$-saciado. Repetimos el argumento T veces para obtener un límite proyectivo $\mathscr{I}_t$-saciado para toda $t = 1, \cdots, T$. $\square$

Gracias a este resultado podemos reducir DHJ al caso netamente saciado:

PROPOSICIÓN 29. *Si DHJ (Teorema 22) es cierto para toda distribución netamente saciada entonces es cierto en general.*

DEMOSTRACIÓN. Por la Proposición 17 será suficiente considerar distribuciones fuertemente estacionarias. Sea pues

$$(X^\omega, \Sigma^\omega, \mathbb{P})$$

fuertemente estacionario y $A \in \Sigma$ tal que

$$\mathbb{P}_w(A) \geq \delta \qquad \forall w \in [k]^\omega \quad \text{con} \quad \delta > 0.$$

Sea $\tilde{\mathbb{P}}$ extensión netamente saciada de $\mathbb{P}$ con

$$\psi : \tilde{X} \to X,$$

$$\mathbb{P} = \tilde{\mathbb{P}} \circ \psi^{-1}.$$

Tenemos así que

$$\tilde{\mathbb{P}}\big(\psi^{-1}(A)\big) \geq \delta \qquad \forall w \in [k]^\omega$$

y por hipótesis

$$\tilde{\mathbb{P}}\left(\bigcap_{i=1}^{k} \pi_{l(i)}^{-1}\big(\psi^{-1}(A)\big)\right) > 0$$

o, lo que es lo mismo,

$$\tilde{\mathbb{P}}_l\big(\psi^{-1}(A) \times \psi^{-1}(A) \times \cdots \times \psi^{-1}(A)\big) > 0$$

para cualquier línea combinatoria l.

Pero

$$\tilde{\mathbb{P}}_l\big(\psi^{-1}(A) \times \psi^{-1}(A) \times \cdots \times \psi^{-1}(A)\big) = \mathbb{P}_l(A \times A \times \cdots \times A)\,.$$

$\square$

5.5. La estructura de una distribución saciada. Los siguientes resultados nos describen algunas simplificaciones no triviales que induce la saciedad de una distribución.

DEFINICIÓN. Sea $e \subset [k]$ y sea i letra en $[k]$, posiblemente contenida en e. Denotamos por

$$r_{e,i} : [k]^\omega \to (([k] \setminus e) \cup \{i\})^\omega$$

a la función que impone la letra i a cada lugar ocupado por letras en $e \setminus \{i\}$. Cuando $e = \{j\}$ escribimos simplemente

$$r_{j,i} := r_{\{j\},i}.$$

DEFINICIÓN. Si $(X, \Sigma, \mathbb{P})$ es espacio de probabilidad y $\mathscr{G} \subset \Sigma$, decimos que una función f Σ-medible es **ortogonal** a $\mathscr{G}$ siempre que

$$\mathbb{E}(f \mid \mathscr{G}) = 0 \quad \mathbb{P}\text{-c.s.}.$$

Para una función g Σ-medible cualquiera, a la descomposición

$$g = \mathbb{E}(g \mid \mathscr{G}) + (g - \mathbb{E}(g \mid \mathscr{G}))$$

le llamamos **descomposición ortogonal**, notando que $(g - \mathbb{E}(g \mid \mathscr{G}))$ es ortogonal a $\mathscr{G}$. Si $\mathscr{H} \subset \Sigma$ es σ-álgebra y toda función $\mathscr{H}$-medible es ortogonal a $\mathscr{G}$ decimos que las σ-álgebras $\mathscr{H}$ y $\mathscr{G}$ son ortogonales.

DEFINICIÓN. Decimos que una función

$$f : X \to \mathbb{R}$$

es **e-insensible** si es Φ_e-medible.

La siguiente caracterización será útil.

PROPOSICIÓN 30. *Sea $\mathbb{P}$ una distribución fuertemente estacionaria sobre $(X^\omega, \Sigma^\omega)$ y sea $\mathbb{P}_l$ la distribución inducida por cualquier línea combinatoria. Una función*

$$f : X \to \mathbb{R}$$

es e-insensible si y sólo si las funciones

$$f \circ \pi_w : X^k \to \mathbb{R}$$

y

$$f \circ \pi_{r_{e,i}(w)} : X^k \to \mathbb{R}$$

son equivalentes $\mathbb{P}_l$-casi seguramente, para cada $i \in e$.

Demostración. ($\Rightarrow$) Si f es e-insensible, entonces es límite de alguna sucesión $(S_n)_n$ de funciones Φ_e-simples, y para cada S_n tenemos

$$S_n \circ \pi_w = S_n \circ \pi_{r_{e,i}(w)} \quad \mathbb{P}_l\text{-c.s.}, \quad i \in e.$$

Tomando límites de ambos lados de la equivalencia obtenemos la primera implicación.

($\Leftarrow$) Suponemos que para cada i en e

$$f \circ \pi_w = f \circ \pi_{r_{e,i}(w)} \quad \mathbb{P}_l\text{-c.s..}$$

Sea A boreliano en $\mathbb{R}$. Los conjuntos

$$(f \circ \pi_w)^{-1}(A) = \pi_w^{-1}\big(f^{-1}(A)\big)$$

y

$$\big(f \circ \pi_{r_{e,i}(w)}\big)^{-1}(A) = \pi_{r_{e,i}(w)}^{-1}\big(f^{-1}(A)\big) \quad i \in e$$

coinciden salvo por conjuntos $\mathbb{P}_l$-nulos. Entonces

$$f^{-1}(A) \in \Phi_e. \qquad \square$$

El siguiente resultado nos da la primera simplificación no trivial que permite una distribución saciada.

Teorema 37. *Si $\mathbb{P}$ es netamente saciada, $f \in L_\infty(\mathbb{P}_w)$ es e-insensible y $i \in e \subset [k]$, entonces*

$$\mathbb{E}\left(f \,\middle|\, \bigvee_{j \in [k] \setminus e} \Phi_{\{i,j\}}\right) = \mathbb{E}\left(f \,\middle|\, \bigvee_{j \in [k] \setminus e} \Phi_{e \cup \{j\}}\right)$$

Antes de demostrar el teorema desarrollamos con [**Aus11**] un ejemplo que ilustra el espíritu de la demostración.

Ejemplo

Consideramos el caso en el que $k = 3$, $i = 2$ y $e = \{1, 2\}$. Si $\mathbb{P}$ es saciada sobre $X^{[3]^\omega}$ y $f \in L_\infty(\mathbb{P}_w)$ es $\{1, 2\}$-insensible, queremos demostrar que

$$\mathbb{E}\big(f \,|\, \Phi_{\{2,3\}}\big) = \mathbb{E}\big(f \,|\, \Phi_{\{1,2,3\}}\big)\,. \tag{26}$$

Como

$$\mathbb{E}\big(f \,|\, \Phi_{\{1,2,3\}}\big)$$

es siempre $\Phi_{\{2,3\}}$-medible, basta con demostrar que siempre que f sea ortogonal a $\Phi_{\{1,2,3\}}$, también será ortogonal a $\Phi_{\{2,3\}}$. En efecto, consideramos por un momento las proyecciones ortogonales

$$h := f - \mathbb{E}\big(f \,\big|\, \Phi_{\{1,2,3\}}\big)$$

y

$$g := f - \mathbb{E}\big(f \bigm| \Phi_{\{2,3\}}\big)\,;$$

notamos que

$$f = \mathbb{E}\big(f \bigm| \Phi_{\{1,2,3\}}\big) + h$$

y

$$f = \mathbb{E}\big(f \bigm| \Phi_{\{2,3\}}\big) + g.$$

Luego, si las proyecciones ortogonales coinciden, es decir si $g = h$, entonces necesariamente

$$\mathbb{E}\big(f \bigm| \Phi_{\{1,2,3\}}\big) = \mathbb{E}\big(f \bigm| \Phi_{\{2,3\}}\big)\,.$$

Así pues, suponemos que el lado derecho de la ecuación (26) es cero para demostrar que el lado izquierdo también lo es.

Suponemos entonces, para obtener una contradicción, que

$$h := \mathbb{E}\big(f \,|\, \Phi_{\{2,3\}}\big) \neq 0.$$

Por definición de esperanza condicional,

$$\int_X Z \cdot h\, d\mathbb{P}_w = \int_X Z \cdot f\, d\mathbb{P}_w$$

para toda $Z \in L_\infty(\mathbb{P}_w)$ que sea $\Phi_{\{2,3\}}$-medible. En particular,

$$\int_X h^2\, d\mathbb{P}_w = \int_w f \cdot h\, d\mathbb{P}_w,$$

de modo que

$$\int_w f \cdot h\, d\mathbb{P}_w = \kappa \neq 0.$$

Por estacionariedad fuerte, la distribución $\mathbb{P}_w$ sobre X es indiferente a la elección de $w \in [k]^\omega$. Tenemos así que

$$\begin{aligned}
0 \neq \int_X f \cdot h\, d\mathbb{P}_w &= \int_{X^\omega} (f \circ \pi_w) \cdot (h \circ \pi_w)\, d\mathbb{P} \\
&= \int_{X^\omega} \big(f \circ \pi_{r_{1,2}(w)}\big) \cdot \big(h \circ \pi_{r_{1,2}(w)}\big)\; d\mathbb{P} \\
&= \int_{X^\omega} (f \circ \pi_w) \cdot \big(h \circ \pi_{r_{1,2}(w)}\big)\; d\mathbb{P} \\
&= \int_{X^\omega} (f \circ \pi_w) \cdot \big(h \circ \pi_{r_{3,2}(r_{1,2}(w))}\big)\; d\mathbb{P} \\
&= \int_{X^\omega} (f \circ \pi_w) \cdot (h \circ \pi_{22\cdots 2})\; d\mathbb{P}
\end{aligned}$$

donde la tercera igualdad se da por ser f $\{1,2\}$-insensible, la cuarta por ser h $\{2,3\}$-insensible, y la quinta se debe simplemente a que para toda palabra w,

$$r_{3,2} \circ r_{1,2}(w) = 22 \cdots 2, \quad |r_{3,2} \circ r_{1,2}(w)\,| = |w|.$$

La última igualdad nos sugiere la siguiente extensión: definimos

$$\tau_2(w) = 22 \cdots 2 \quad \text{donde} \quad |22 \cdots 2| = |w|$$

y consideramos el espacio

$$\left(\tilde{X}^\omega, \tilde{\Sigma}^\omega\right) := \left((X \times X)^{[k]^\omega}, (\Sigma \otimes \Sigma)^{[k]^\omega}\right)$$

con la distribución $\tilde{\mathbb{P}}$ tal que su proyección sobre cada índice w está dada por los cilindros

$$\tilde{\mathbb{P}}_w(A, B) = \mathbb{P}\big(\ X_w \in A,\ X_{\tau_2(w)} \in B\ \big) \quad A, B \in \Sigma,$$

es decir que

$$\tilde{\mathbb{P}}_w = \mathbb{P}_{\{w, \tau_2(w)\}} \quad \forall w \in [k]^\omega;$$

y tal que si

$$\rho : \tilde{X} \to X$$

es la proyección sobre el primer término índice por índice,

$$\rho((x_w, y_w)_w) = (x_w)_w$$

entonces

$$\tilde{\mathbb{P}} \circ \rho^{-1} = \mathbb{P}.$$

Para determinar completamente a $\tilde{\mathbb{P}}$ definimos su masa sobre los cilindros

$$\tilde{\mathbb{P}}(\ (x_{w_1}, y_{w_1}) \in A_1 \times B_1, \cdots, (x_{w_n}, y_{w_n}) \in A_n \times B_n\ ) =$$
$$\mathbb{P}(\ w_1 \in A_1,\ \tau_1(w_1) \in B_1, \cdots, w_n \in A_n,\ \tau_2(w_n) \in B_n\ )$$

para cualesquiera

$$w_1, \cdots, w_n \in [k]^\omega \quad \text{y} \quad (A_i)_{1 \le i \le n}, (B_i)_{1 \le i \le n} \subset \Sigma.$$

Pues bien, si $\rho_1 : \tilde{X} \to X$ y $\rho_2 : \tilde{X} \to X$ son las proyecciones unidimensionales a la primera y segunda coordenadas,

$$\rho_1(x, y) = x \quad \rho_2(x, y) = y,$$

tenemos que

$$f \circ \rho_1 : \tilde{X} \to \mathbb{R} \quad \text{y} \quad h \circ \rho_2 : \tilde{X} \to \mathbb{R}$$

son acotadas y $\tilde{\Sigma}$-medibles. Más aún,

$$f \circ \rho_1 \quad \text{es} \quad \rho_1^{-1}\big(\Phi_{\{1,2\}}\big)\text{-medible}$$

y en particular es $\rho_1^{-1}(\Sigma)$-medible;

$$h \circ \rho_2 \quad \text{es} \quad \rho_2^{-1}\big(\Phi_{\{2,3\}}\big)\text{-medible}$$

y en particular es $\rho_2^{-1}(\Sigma)$-medible. De modo que

$$\int_{\tilde{X}_w} (f \circ \rho_1) \cdot (h \circ \rho_2)\, d\tilde{\mathbb{P}}_w = \int_{X^\omega} (f \circ \pi_w) \cdot \big(h \circ \pi_{\tau_2(w)}\big)\, d\mathbb{P}$$

con A y B medibles arbitrarios en Σ , así que

$$\mathbb{E}_{\tilde{\mathbb{P}}_w}\big(\ (f \circ \rho_1) \cdot (h \circ \rho_2) \mid \mathscr{G}\ \big) = \mathbb{E}\Big(\ (f \circ \pi_w) \cdot \big(h \circ \pi_{\tau_2(w)}\big) \mid \mathscr{G}\ \Big)$$

para toda $\mathscr{G} \subset \Sigma \otimes \Sigma$ subσ-álgebra.

Notamos, por otro lado, que $h \circ \rho_2(x)_w$ es invariante bajo cambios de letras, pues

$$h \circ \rho_2\big((x_w, x_{\tau_2(w)})\big) = h \circ \rho_2\big((x_v, x_{\tau_2(v)})\big) \quad \forall w, v \in [k]^n \quad \forall n \geq 1.$$

Retomando la Proposición 30,

$$h \circ \rho_2 \quad \text{es} \quad \tilde{\Phi}_{\{1,2,3\}}\text{-medible.}$$

Por ser $\mathbb{P}$ saciada, de ser $\tilde{\mathbb{P}}$ extensión fuertemente estacionaria tendríamos que

$$\rho_1^{-1}(\Sigma) \perp_{\rho_1^{-1}\left(\Phi_{\{1,2,3\}}\right)} \tilde{\Phi}_{\{1,2,3\}}.$$

Luego entonces,

$$\mathbb{E}_{\tilde{\mathbb{P}}_w}\Big(\ (f \circ \rho_1) \cdot (h \circ \rho_2) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\ \Big) = \tag{27}$$

$$\mathbb{E}_{\tilde{\mathbb{P}}_w}\Big((f \circ \rho_1) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\Big) \cdot \mathbb{E}_{\tilde{\mathbb{P}}}\Big((h \circ \rho_2) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\Big) =$$

$$\mathbb{E}\big(f \mid \Phi_{\{1,2,3\}}\big) \cdot \mathbb{E}\big(h \mid \Phi_{\{1,2,3\}}\big) = 0$$

ya que por hipótesis

$$\mathbb{E}\big(f \mid \Phi_{\{1,2,3\}}\big) = 0.$$

Pero entonces la expresión (27) es cero, mientras que

$$\int_{X^\omega} \mathbb{E}\Big((f \circ \pi_w) \cdot \big(h \circ \pi_{\tau_2(w)}\big) \ \mid \Phi_{\{1,2,3\}}\Big)\, d\mathbb{P} =$$

$$\int_{X^\omega} (f \circ \pi_w) \cdot \big(h \circ \pi_{\tau_2(w)}\big)\, d\mathbb{P} =$$

$$\int_X f \cdot h\, d\mathbb{P}_w = \kappa \neq 0,$$

claramente una contradicción.

Concluiríamos así que f debe ser ortogonal a $\Phi_{\{2,3\}}$.

Entonces sólo debemos justificar por qué y en qué sentido podemos suponer que $\tilde{\mathbb{P}}$ sea fuertemente estacionaria siendo que, en general, $\tilde{\mathbb{P}}$ no es fuertemente estacionaria ya que las distribuciones

$$\tilde{\mathbb{P}}_w = \mathbb{P}_{\{w,\tau_2(w)\}}, \quad \tilde{\mathbb{P}}_v = \mathbb{P}_{\{v,\tau_2(w)\}} \quad w, v \in [k]^n$$

no siempre coinciden. Sin embargo, por el Teorema 28 sabemos que existe una sucesión de subespacios refinados de $[k]^\omega$,

$$(\psi_m)_{m\geq 1}\,,$$

tal que la sucesión

$$\left(\tilde{\mathbb{P}}_{\psi_m}\right)_{m\geq 1}$$

converge bajo la métrica de Prohorov a una distribución fuertemente estacionaria, a la que llamaremos $\mathbb{P}^*$, sobre

$$\left(\tilde{X}, \tilde{\Sigma}\right).$$

Si sustituímos $\tilde{\mathbb{P}}$ por $\mathbb{P}^*$ en la ecuación (27) tenemos que

$$\mathbb{E}_{\mathbb{P}^*_w}\big(\ (f\circ\rho_1)\cdot(h\circ\rho_2) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\ \big) = 0.$$

Veremos que aunque f no sea continua, por convergencia débil podemos elegir $N > 0$ tal que si $m \geq N$ entonces

$$\int_{\tilde{X}_w} (f\circ\rho_1)\cdot(h\circ\rho_2)\ d\tilde{\mathbb{P}}_{\psi_m(w)} \in$$

$$\left\{\int_{\tilde{X}^\omega} (f\circ\pi_w)\cdot\big(h\circ\pi_{\tau_2(w)}\big)\ d\tilde{\mathbb{P}}_* + r\ :\ -\epsilon < r < \epsilon\right\},$$

de donde, para $\mathscr{G} \subset \Sigma$ subσ-álgebra,

$$\left\|\tilde{\mathbb{E}}_{\tilde{\mathbb{P}}_{\psi_m(w)}}\big(\ (f\circ\rho_1)\cdot(h\circ\rho_2) \mid \mathscr{G}\ \big) - \mathbb{E}_{\tilde{\mathbb{P}}_*}\big(\ (f\circ\rho_1)\cdot(h\circ\rho_2) \mid \mathscr{G}\ \big)\right\|_\infty \leq \epsilon.$$

Fijando

$$\mathscr{G} = \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)$$

tenemos que

$$\mathbb{E}_{\tilde{\mathbb{P}}_{\psi_m(w)}}\big(\ (f\circ\rho_1)\cdot(h\circ\rho_2) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\ \big) \to 0, \quad m\to\infty \quad \tilde{\mathbb{P}}_*\text{-c.s.}$$

Por otro lado, para cualquier m y cualquier palabra w,

$$\int_{\tilde{X}_w} \mathbb{E}_{\tilde{\mathbb{P}}_{\psi_m(w)}}\big(\ (f\circ\rho_1)\cdot(h\circ\rho_2) \mid \rho_1^{-1}\big(\Phi_{\{1,2,3\}}\big)\ \big)\ d\tilde{\mathbb{P}}_{\psi_m(w)} =$$

$$\int_{X^\omega} \mathbb{E}\big((f \circ \pi_w) \cdot \big(h \circ \pi_{\tau_2(w)}\big) \mid \Phi_{\{1,2,3\}}\big)\, d\mathbb{P} =$$
$$\int_X f \cdot h\, d\mathbb{P} = \kappa \neq 0$$

Hemos arribado a la misma contradicción, pero en términos de límites.

La demostración del Teorema 37 es una generalización de este procedimiento. Para formalizar echamos mano de dos resultados preliminares.

Proposición 31.
$$\bigvee_{j \in [k] \setminus e} \Phi_{e \cup \{j\}} = \Phi_{\mathscr{I}}$$
para alguna familia ascendente $\mathscr{I}$.

Demostración. Recordemos que
$$\Phi_{\langle e \cup \{j\} \rangle} = \Phi_{e \cup \{j\}} \quad j \in [k] \setminus e.$$
Luego,
$$\bigvee_j \Phi_{\langle e \cup \{j\} \rangle} = \bigvee_j \Phi_{e \cup \{j\}} \quad j \in [k] \setminus e.$$
□

Definición. Le llamamos **soporte** de una función
$$f : X \to \mathbb{R}$$
a la cerradura del conjunto
$$\{x \in X : \quad f(x) \neq 0\}.$$

Teorema 38 ([**Rud87**], p. 69). *Sea* (X, Σ) *espacio estándar de Borel compacto y* $\mathbb{P}$ *medida de probabilidad sobre él. Sean* $1 \leq p < \infty$ *y* $C_X(\mathbb{R})$ *la familia de funciones continuas de* X *en* $\mathbb{R}$ *con soporte compacto. La familia* $C_X(\mathbb{R})$ *es densa en* $L_p(\mathbb{P})$.

No reproducimos la demostración; es un resultado muy cercano al Teorema de Lusin. Hemos añadido hipótesis para adaptar el teorema a nuestra situación. A saber, usamos que un espacio polaco es Hausdorff localmente compacto.

Proposición 32. *Si* $(X, \Sigma, \mathbb{P})$ *es estándar de Borel compacto y*
$$f : X \to \mathbb{R}$$
es Σ*-medible y acotada, entonces existe una sucesión de funciones continuas*
$$(g_n)_{n \geq 1}$$

de X en $\mathbb{R}$ tal que

$$\int g_n \, d\mathbb{P} \to \int f \, d\mathbb{P}, \quad n \to \infty.$$

DEMOSTRACIÓN. Consideramos el Teorema 38 para $p = 1$; toda función Σ-medible acotada está en $L_1(\mathbb{P})$. $\square$

PROPOSICIÓN 33. *Si (X, Σ) es espacio estándar de Borel y $(\mathbb{P}_n)_n$ una sucesión de medidas que convergen bajo la métrica de Prohorov a $\mathbb{P}$, y si*

$$f : X \to \mathbb{R}$$

es Σ-medible y acotada, entonces

$$\int f \, d\mathbb{P}_n \to \int f \, d\mathbb{P}, \quad n \to \infty.$$

DEMOSTRACIÓN. Sea $(g_n)_n$ sucesión de funciones continuas que convergen a f en $L_1(\mathbb{P})$. Tenemos

$$\begin{aligned} \left| \int f \, d\mathbb{P} - \int f \, d\mathbb{P}_n \right| &\leq \left| \int f \, d\mathbb{P} - \int g_n \, d\mathbb{P} \right| + \left| \int g_n \, d\mathbb{P} - \int g_n \, d\mathbb{P}_n \right| \\ &+ \left| \int g_n \, d\mathbb{P}_n - \int f \, d\mathbb{P}_n \right|. \end{aligned}$$

El primer término del lado derecho de la desigualdad tiende a cero por la Proposición 32 y por elección de las g_n; el segundo término tiende a cero por convergencia débil de las $\mathbb{P}_n$. Acotamos al tercer término usndo la desigualdad del triángulo:

$$\begin{aligned} \left| \int g_n \, d\mathbb{P}_n - \int f \, d\mathbb{P}_n \right| &\leq \left| \int g_n \, d\mathbb{P}_n - \int g_n \, d\mathbb{P} \right| + \left| \int g_n \, d\mathbb{P} - \int f \, d\mathbb{P} \right| \\ &+ \left| \int f \, d\mathbb{P} - \int f \, d\mathbb{P}_n \right|. \end{aligned}$$

El tercer término tiende a cero por convergencia débil de las $\mathbb{P}_n$; a los otros dos términos ya los conocemos. $\square$

PROPOSICIÓN 34. *Sea*

$$\mathscr{G} := \bigvee_{i=1}^{t} \mathscr{H}_i$$

donde cada $\mathscr{H}_i$ es σ-álgebra. Entonces para toda función $\mathscr{G}$-medible y acotada f existe una sucesión

$$(G_n)_{n \geq 1}$$

de funciones de la forma

$$G_n = a_1 \cdot g_{n,1} + \cdots + a_m \cdot g_{n,r} \qquad r \geq 1 \quad a_j \in \mathbb{R}$$

con

$$g_{n,j} = \prod_{i=1}^{t} h_i^j \qquad h_i^j \quad \mathscr{H}_i\textit{-medibles}$$

tal que

$$G_n \to f \quad en \quad L_2(\mathscr{G})\,.$$

Demostración. Dada f $\mathscr{G}$-medible, sabemos que existe una sucesión

$$(S_n)_{n\geq 1}$$

de $\mathscr{G}$-simples tal que

$$S_n \to f$$

en $L_2(\mathscr{G})$. Afirmamos que para cada S_m existe una sucesión

$$(G_n^m)_{n\geq 1}$$

de funciones de la forma descrita en el teorema - es decir, combinaciones lineales de productos de funciones $\mathscr{H}_i$-medibles con $i = 1, \cdots, t$, tal que

$$G_n^m \to S_m$$

en $L_2(\mathscr{G})$. Una vez demostrando esto podemos ver, con un argumento de diagonalización similar al que usamos en varias ocasiones en el Capítulo 2, que

$$(G_n^n)_{n\geq 1}$$

converge a f en $L_2(\mathscr{G})$.

Sea

$$S := \sum_{1\leq s\leq m} k_s \mathbf{1}_{A_s} \qquad A_s \in \mathscr{G} \quad k_s \in \mathbb{R}$$

función $\mathscr{G}$-simple. Denominamos con $\mathscr{C}$ a la familia de todos los conjuntos A en $\mathscr{G}$ tales que

$$k\mathbf{1}_A \quad k \in \mathbb{R}$$

puede expresarse como combinación lineal de productos de funciones $\mathscr{H}_i$-medibles para $i = 1, \cdots, t$. Suponemos que las $\mathscr{H}_i$-medibles son acotadas, sin perder generalidad puesto que su producto lo es. Veamos que $\mathscr{C}$ es π-sistema: sean

$$\mathbf{1}_A = \prod_{i=1}^{t} h_i^1 + \cdots + \prod_{i=1}^{t} h_i^u \qquad h_i^j \quad \mathscr{H}_i\text{-medibles}$$

y

$$\mathbf{1}_B = \prod_{i=1}^t r_i^1 + \cdots + \prod_{i=1}^t r_i^v \qquad r_i^j \quad \mathscr{H}_i\text{-medibles;}$$

entonces

$$k\mathbf{1}_{A\cap B} = k\cdot\left(\prod_{i=1}^t h_i^1 + \cdots + \prod_{i=1}^t h_i^u\right)\left(\prod_{i=1}^t r_i^1 + \cdots + \prod_{i=1}^t r_i^v\right),$$

que también es combinación lineal.

Por el Lema de Clases Monótonas de Dynkin, si demostramos que $\mathscr{C}$ es λ-sistema, sabremos que

$$\mathscr{G} := \sigma(\mathscr{H}_1, \cdots, \mathscr{H}_n) \subset \sigma(\mathscr{C}) = \lambda(\mathscr{C}) = \mathscr{C}.$$

Claramente $X \in \mathscr{C}$. Sean A y B conjuntos en $\mathscr{C}$ como antes, con el requerimiento adicional de que

$$A \setminus B \neq \emptyset;$$

tenemos

$$\mathbf{1}_{B^c} = 1 - \left(\prod_{i=1}^t r_i^1 + \cdots + \prod_{i=1}^t r_i^v\right)$$

y

$$k\mathbf{1}_{A\setminus B} = k\cdot\left(\prod_{i=1}^t h_i^1 + \cdots + \prod_{i=1}^t h_i^u\right)\left(1 - \left(\prod_{i=1}^t r_i^1 + \cdots + \prod_{i=1}^t r_i^v\right)\right),$$

de donde

$$A \setminus B \in \mathscr{C}.$$

Si

$$(A_n)_{n\geq 1} \qquad A_n \subset A_{n+1}$$

es sucesión creciente de elementos en $\mathscr{C}$ con

$$\mathbf{1}_{A_n} = \prod_{i=1}^t h_i^{1,(n)} + \cdots + \prod_{i=1}^t h_i^{u,(n)}$$

tenemos

$$\begin{aligned}
\mathbf{1}_{\cup A_n} &= \lim_{n\to\infty} \mathbf{1}_{A_n}\\
&= \lim_{n\to\infty}\left(\prod_{i=1}^{t} h_i^{1,(n)} + \cdots + \prod_{i=1}^{t} h_i^{u,(n)}\right)\\
&= \lim_{n\to\infty}\prod_{i=1}^{t} h_i^{1,(n)} + \cdots + \lim_{n\to\infty}\prod_{i=1}^{t} h_i^{u,(n)}\\
&= \prod_{i=1}^{t}\lim_{n\to\infty} h_i^{1,(n)} + \cdots + \prod_{i=1}^{t}\lim_{n\to\infty} h_i^{u,(n)}
\end{aligned}$$

y cada $\lim_{n\to\infty} h_i^{j(n)}$ es $\mathscr{H}_i$-medible, ya que la sucesión

$$\left(h_i^{j,(n)}\right)_{n\geq 1}$$

es acotada y monótona creciente. Concluímos que

$$\bigcup_{n=1}^{\infty} A_n \in \mathscr{C}.$$

Hemos demostrado que para cada $s = 1, \cdots, m$ la función

$$k_s \mathbf{1}_{A_s}$$

es aproximable por combinaciones lineales de productos de $\mathscr{H}_i$-medibles. No es difícil ver que S lo es también; omitimos los detalles por motivos de espacio. $\square$

Demostración del Teorema 37. Elegimos $i \in e$. Claramente

$$\Phi_{e\cup\{j\}} \subset \Phi_{\{i,j\}} \quad \forall j \in [k] \setminus e$$

de donde

$$\bigvee_{j\in[k]\setminus e} \Phi_{e\cup\{j\}} \subset \bigvee_{j\in[k]\setminus e} \Phi_{\{i,j\}}.$$

Entonces

$$\mathbb{E}\left(f \;\middle|\; \bigvee_{j\in[k]\setminus e} \Phi_{e\cup\{j\}}\right) \quad \text{es} \quad \bigvee_{j\in[k]\setminus e} \Phi_{\{i,j\}}\text{-medible.}$$

Queremos demostrar que si f es acotada, Σ-medible y e-insensible, entonces

$$\int_G f\, d\mathbb{P} = \int_G \mathbb{E}\left(f \,\middle|\, \bigvee_{j\in[k]\smallsetminus e} \Phi_{e\cup\{j\}} \right) d\mathbb{P} \quad \forall G \in \bigvee_{j\in[k]\smallsetminus e} \Phi_{\{i,j\}},$$

que es equivalente a demostrar que

$$\int f\cdot h\, d\mathbb{P} = \int \mathbb{E}\left(f \,\middle|\, \bigvee_{j\in[k]\smallsetminus e} \Phi_{e\cup\{j\}} \right)\cdot h\, d\mathbb{P}$$

para todo

$$h \quad \bigvee_{j\in[k]\smallsetminus e} \Phi_{\{i,j\}}\text{-medible} \quad \text{en} \quad L_2(\mathbb{P})\,;$$

es decir que

$$f - \mathbb{E}\left(f \,\middle|\, \bigvee_{j\in[k]\smallsetminus e} \Phi_{e\cup\{j\}} \right)$$

es ortogonal a toda tal h.

Equivalentemente, suponemos que

$$f \quad \text{es ortogonal a} \quad \bigvee_{j\in[k]\smallsetminus e} \Phi_{e\cup\{j\}}$$

para demostrar que f debe también ser ortogonal a

$$\bigvee_{j\in[k]\smallsetminus e} \Phi_{\{i,j\}}.$$

Para obtener una contradicción, suponemos que pudieramos encontrar funciones acotadas h_j , $j\in[k]\smallsetminus e$ con

$$h_j \quad \Phi_{\{i,j\}}\text{-medible}$$

tales que

$$\int_X f\cdot \prod_{j\in[k]\smallsetminus e} h_j\, d\mathbb{P}_w = \kappa \neq 0.$$

Por la Proposición 34 basta considerar este caso.

Transfiriendo la integral al espacio producto X^ω,

$$\int_{X^\omega} (f\circ\pi_w)\cdot \prod_{j\in[k]\smallsetminus e} (h_j\circ\pi_w)\, d\mathbb{P} = \kappa\neq 0,$$

y haciendo valer la estacionariedad fuerte de $\mathbb{P}$ y la e-insensibilidad de f tenemos

$$\begin{aligned}\kappa &= \int_{X^\omega} \left(f\circ\pi_{r_{e,i}(w)}\right)\cdot \prod_{j\in[k]\smallsetminus e}\left(h_j\circ\pi_{r_{e,i}(w)}\right)\, d\mathbb{P}\\ &= \int_{X^\omega} \left(f\circ\pi_{w}\right)\cdot \prod_{j\in[k]\smallsetminus e}\left(h_j\circ\pi_{r_{e,i}(w)}\right)\, d\mathbb{P}\\ &= \int_{X^\omega} \left(f\circ\pi_{w}\right)\cdot \prod_{j\in[k]\smallsetminus e}\left(h_j\circ\pi_{r_{e,j}(w)}\right)\, d\mathbb{P}\end{aligned}$$

echando mano de la $\{i,j\}$-insensibilidad de cada h_j para la última igualdad. Notamos que las palabras

$$\{r_{e,j}(w):\ j\in[k]\smallsetminus e\quad w\in[k]^\omega\}$$

se escriben con el alfabeto reducido $[k]\smallsetminus e$.

Buscamos ahora construir una extensión de X que a partir de esta información nos exhiba una contradicción. Consideramos el espacio

$$\left(\tilde{X},\tilde{\Sigma}\right) := \left(X\times X^{[k]\smallsetminus e},\Sigma\otimes\Sigma^{[k]\smallsetminus e}\right) \overset{\sigma}{\backsim} \left(X^{k-|e|+1},\Sigma^{k-|e|+1}\right)$$

y le otorgamos la distribución $\tilde{\mathbb{P}}$ tal que si

$$A_m\in\tilde{\Sigma}\quad m=1,\cdots,n$$

con

$$A_m = A_{(m,0)}\times\prod_{j\in[k]\smallsetminus e} A_{(m,j)}$$

entonces

$$\begin{aligned}&\tilde{\mathbb{P}}(\tilde{x}_{w_1}\in A_1,\cdots,\tilde{x}_{w_n}\in A_n) =\\ &\qquad \mathbb{P}\big(x_{w_m}\in A_{(m,0)},\quad x_{r_{e,j}(w_m)}\in A_{(m,j)}\quad \forall j\in[k]\smallsetminus e\quad m=1,\cdots,n\big)\,.\end{aligned}$$

En particular, si para $(\tilde{x}_w)_w\in\tilde{X}^\omega$ escribimos

$$(\tilde{x}_w)_w = (x_w,\bar{x}_w)_w$$

con

$$\bar{x}_w = (r_{e,j}(w))_{j\in[k]\smallsetminus e}\,,$$

y

$$\rho:\tilde{X}^\omega\to X^\omega$$

es la proyección sobre la primera coordenada término a término,

$$\rho((x_w,\bar{x}_w)_w) = (x)_w\,,$$

tenemos
$$\mathbb{P} = \tilde{\mathbb{P}} \circ \rho^{-1},$$
de modo que $\tilde{\mathbb{P}}$ es extensión de $\mathbb{P}$.

Por otro lado, si definimos las proyecciones término a término sobre las coordenadas $j \in [k] \setminus e$,
$$\rho_j((x_w, \bar{x}_w)_w) = \big(x_{r_{e,j}(w)}\big)_w$$
podemos ver que ρ_j es indiferente a cambios de letras en $e \cup \{j\}$. Si escribimos
$$\rho_{(j,w)} : \tilde{X}_w \to X_{r_{e,j}(w)}$$
para denotar a la proyección unidimensional del índice w sobre la coordenada j, vemos que
$$\rho^{-1}_{(j,w)}(A) \in \tilde{\Phi}_{e\cup\{j\}} \quad \text{para} \quad A \in \Sigma_{r_{e,j}(w)};$$
es decir que $\rho_{(j,w)}$ es $\tilde{\Phi}_{e\cup\{j\}}$-medible para toda $j \in [k]\setminus e$ y $w \in [k]^\omega$. Entonces, si escribimos
$$\tilde{\pi}_w : \tilde{X}^\omega \to \tilde{X}_w$$
para denotar las proyecciones por índice en $\tilde{X}^\omega$ y $\rho_0 : \tilde{X}_w \to X_w$ para denotar a la proyección unidimensional
$$\rho_0(x_w, \bar{x}_w) = x_w$$
tenemos
$$\int_{\tilde{X}} (f \circ \rho_0) \cdot \prod_j \big(h_j \circ \rho_{(j,w)}\big)\, d\tilde{\mathbb{P}}_w =$$
$$\int_{\tilde{X}^\omega} (f \circ \rho_0 \circ \tilde{\pi}_w) \cdot \prod_j \big(h_j \circ \rho_{(j,w)} \circ \tilde{\pi}_w\big)\, d\tilde{\mathbb{P}} =$$
$$\int_{X^\omega} (f \circ \pi_w) \cdot \prod_j \big(h_j \circ \pi_{r_{e,j}(w)}\big)\, d\mathbb{P} = \kappa$$

Claramente $(f \circ \rho_0)$ es $\rho_0^{-1}(\Sigma)$-medible; veamos que cada $\big(h_j \circ \rho_{(j,w)}\big)$ es $\tilde{\Phi}_{e\cup\{j\}}$-medible: si $B \in \mathscr{B}_{\mathbb{R}}$ tenemos
$$\big(h_j \circ \rho_{(j,w)}\big)^{-1}(B) = \rho^{-1}_{(j,w)}\big(h_j^{-1}(B)\big) = \rho^{-1}_{(j,w)}(A) \in \tilde{\Phi}_{e\cup\{j\}} \quad A \in \Phi_{\{i,j\}}$$
por ser $\rho_{(j,w)}$ $\tilde{\Phi}_{e\cup\{j\}}$-medible y h_j $\Phi_{\{i,j\}}$-medible.

Aunque $\tilde{\mathbb{P}}_w$ es extensión de $\mathbb{P}_w$ para cada $w \in [k]^\omega$, no es indiferente a la elección de w, ya que en general
$$\mathbb{P}_{\{w,(r_{e,j}(w))_j\}} \neq \mathbb{P}_{\{v,(r_{e,j}(v))_j\}} \quad \text{para} \quad w, v \in [k]^n \quad n \geq 1.$$

Por lo mismo $\tilde{\mathbb{P}}$ no puede ser fuertemente estacionario. Sin embargo, por el Teorema 28 existe una sucesión de subespacios refinados de $[k]^\omega$,

$$(\psi_m)_{m\geq 1}\,,$$

tal que la sucesión

$$\left(\tilde{\mathbb{P}}_{\psi_m}\right)_{m\geq 1}$$

converge bajo la métrica de Prohorov a una distribución fuertemente estacionaria, a la que llamaremos $\tilde{\mathbb{P}}^*$. Elegimos $w\in[k]^\omega$ arbitraria y denotamos por $\tilde{\mathbb{E}}(g\,|\,\mathscr{G})$ a la esperanza condicional de una función

$$g:\left(\tilde{X},\tilde{\Sigma},\tilde{\mathbb{P}}^*_w\right)\to\mathbb{R}$$

$\tilde{\Sigma}$-medible, con respecto a una subσ-álgebra $\mathscr{G}\subset\tilde{\Sigma}$, y escribimos

$$\tilde{\mathbb{E}}(g):=\int_{\tilde{X}} g\, d\tilde{\mathbb{P}}^*_w.$$

Como consecuencia de la saciedad de $\mathbb{P}$ y de la estacionariedad fuerte de $\tilde{\mathbb{P}}^*$ tenemos

$$\tilde{\mathbb{E}}\left(\,(f\circ\rho_0)\cdot\prod_j\big(h_j\circ\rho_{(j,w)}\big)\;\middle|\;\rho_0^{-1}\big(\Phi_{e\cup\{j\}}\big)\right)=$$

$$\tilde{\mathbb{E}}\big(\,(f\circ\rho_0)\;\big|\;\rho_0^{-1}\big(\Phi_{e\cup\{j\}}\big)\big)\cdot\tilde{\mathbb{E}}\left(\,\prod_j\big(h_j\circ\rho_{(j,w)}\big)\;\middle|\;\rho_0^{-1}\big(\Phi_{e\cup\{j\}}\big)\right);$$

pero

$$\tilde{\mathbb{E}}\big(\,(f\circ\rho_0)\;\big|\;\rho_0^{-1}\big(\Phi_{e\cup\{j\}}\big)\big)=\mathbb{E}\big(f\;\big|\;\Phi_{e\cup\{j\}}\big)$$

que es igual a cero por hipótesis.

Entonces si denotamos por

$$\mathbb{E}_m(g\;|\mathscr{G})\quad g:\tilde{X}^\omega\to\mathbb{R},\quad \mathscr{G}\subset\tilde{\Sigma}^\omega\quad \text{sub}\sigma\text{-álgebra}$$

a la esperanza condicional de g correspondiente al espacio

$$\left(\tilde{X}^\omega,\tilde{\Sigma}^\omega,\tilde{\mathbb{P}}_{\psi_m}\right)$$

la Proposición 33 nos dice que

$$\mathbb{E}\left(\mathbb{E}_m\left(\,(f\circ\rho_0\circ\tilde{\pi}_w)\cdot\left(\prod_j h_j\circ\rho_{(j,w)}\circ\tilde{\pi}_w\right)\;\middle|\;(\rho_0\circ\tilde{\pi}_w)^{-1}\big(\Phi_{e\cup\{j\}}\big)\right)\right)\to 0$$

en $L_1\left(\tilde{\mathbb{P}}^*\right)$ conforme $m \to \infty$. Notemos que esto es equivalente a:

$$\mathbb{E}\left(\mathbb{E}\left(\left(f \circ \pi_{\psi_m(w)}\right) \cdot \left(\prod_j h_j \circ \pi_{r_{e,j} \circ \psi_m(w)}\right) \,\middle|\, \pi^{-1}_{\psi_m(w)}\left(\Phi_{e \cup \{j\}}\right)\right)\right) \to 0$$

en $L_1\left(\tilde{\mathbb{P}}^* \circ \rho^{-1}\right)$.

Pero por otro lado, tenemos que

$$\int_{\tilde{X}^\omega} \mathbb{E}_m\left(\left(f \circ \rho_0 \circ \tilde{\pi}_w\right) \cdot \left(\prod_j h_j \circ \rho_{(j,w)} \circ \tilde{\pi}_w\right) \,\middle|\, \left(\rho_0 \circ \tilde{\pi}_w\right)^{-1}\left(\Phi_{e \cup \{j\}}\right)\right) d\tilde{\mathbb{P}}_{\psi_m}$$

$$= \int_{\tilde{X}^\omega} \left(f \circ \rho_0 \circ \tilde{\pi}_w\right) \cdot \left(\prod_j h_j \circ \rho_{(j,w)} \circ \tilde{\pi}_w\right) d\tilde{\mathbb{P}}_{\psi_m}$$

$$= \int_{X^\omega} \left(f \circ \pi_{\psi_m(w)}\right) \cdot \left(\prod_j h_j \circ \pi_{r_{e,j} \circ \psi_m(w)}\right) d\mathbb{P} = \kappa$$

para todo $m \geq 1$, usando para la última igualdad la estacionariedad fuerte de $\mathbb{P}$. Siendo así, por convergencia de las $\mathbb{P}_{\psi_m}$ en la topología débil y la Proposición 33 tenemos

$$\int_{\tilde{X}} \tilde{\mathbb{E}}\left(\left(f \circ \rho_0\right) \cdot \prod_j \left(h_j \circ \rho_{(j,w)}\right) \,\middle|\, \rho_0^{-1}\left(\Phi_{e \cup \{j\}}\right)\right) d\tilde{\mathbb{P}}^* = \kappa,$$

claramente una contradicción, puesto que en el párrafo anterior concluímos que el integrando es cero $\tilde{\mathbb{P}}^*$-casi seguramente.

Retomando la Proposición 34, hemos demostrado que

$$f \quad \text{es ortogonal a} \quad \bigvee_{j \in [k] \smallsetminus e} \Phi_{\{i,j\}}. \qquad \square$$

Necesitamos aún más resultados sobre la estructura de una distribución saciada.

Teorema 39. *Si $e \subset [k]$ es no vacío, $\mathbb{P}$ es netamente saciada sobre $(X^\omega, \Sigma^\omega)$, $l := l([k])$ es línea combinatoria y $f_i \in L_\infty(\mathbb{P}_w)$ para $i \in e$, entonces*

$$\int_{X^k} \prod_{i \in e} f_i \circ \pi_i \, d\mathbb{P}_l = \int_{X^k} \prod_{i \in e} \mathbb{E}\left(f_i \,\middle|\, \bigvee_{t \in e \smallsetminus \{i\}} \Phi_{\{i,t\}}\right) \circ \pi_i \, d\mathbb{P}_l.$$

Demostración. Suponemos la negación de lo que afirma el teorema para obtener una contradicción. Supongamos, pues, que para alguna colección $(f_i)_{i\in e}$ falla la desigualdad, y a partir de eso construímos una extensión de X^ω que evidencíe la no saciedad de $\mathbb{P}$. Abreviamos

$$\Xi_i := \bigvee_{t\in e\smallsetminus\{i\}} \Phi_{\{i,t\}} = \bigvee_{t\in e\smallsetminus\{i\}} \Phi_{\langle\{i,t\}\rangle},$$

notando que

$$\Xi_i = \Phi_{\mathscr{I}_i},$$

donde $\mathscr{I}_i$ es la familia ascendente generada por

$$(\{i,t\})_{t\in e\smallsetminus\{i\}}\,.$$

Escribimos la diferencia

$$\int_{X^k}\prod_{i\in e} f_i\circ\pi_i\,d\mathbb{P}_l - \int_{X^k}\prod_{i\in e}\mathbb{E}(f_i\,|\,\Xi_i)\circ\pi_i\,d\mathbb{P}_l \tag{28}$$

como serie telescópica

$$\sum_{j\in e}\int_{X^k}\left(\prod_{i\in e,i<j} f_i\circ\pi_i\right)\cdot(f_j\circ\pi_j-\mathbb{E}(f_j\,|\,\Xi_j)\circ\pi_j)\cdot\left(\prod_{i\in e,i>j}\mathbb{E}(f_i\,|\,\Xi_i)\circ\pi_i\right)\,d\mathbb{P}_l$$

y notamos que si la desigualdad es distinta de cero entonces para algún $j\in e$ la integral

$$\int_{X^k}\left(\prod_{i\in e,i<j} f_i\circ\pi_i\right)\cdot(f_j\circ\pi_j-\mathbb{E}(f_j\,|\,\Xi_j)\circ\pi_j)\cdot\left(\prod_{i\in e,i>j}\mathbb{E}(f_i\,|\,\Xi_i)\circ\pi_i\right)\,d\mathbb{P}_l$$

es distinta a cero. Ahora, elegimos cualquier palabra w que contenga alguna j y consideramos la línea combinatoria

$$l_j := (r_{j,s}(w))_{s=1}^k\,.$$

Por estacioneriedad fuerte de $\mathbb{P}$,

$$\mathbb{P}_{l_j} = \mathbb{P}_l$$

y

$$\int_{X^\omega}\left(\prod_{i\in e,i<j} f_i\circ\pi_{r_{j,i}(w)}\right)\cdot(f_j\circ\pi_w-\mathbb{E}(f_j\,|\,\Xi_j)\circ\pi_w)$$
$$\cdot\left(\prod_{i\in e,i>j}\mathbb{E}(f_i\,|\,\Xi_i)\circ\pi_{r_{j,i}(w)}\right)\,d\mathbb{P}_l$$

es también distinta a cero.

Empleamos una vez más la técnica que usamos para demostrar el Teorema 37 y fijamos

$$\left(\tilde{X}, \tilde{\Sigma}\right) := \left(X \times X^{e \setminus \{j\}}, \Sigma \otimes \Sigma^{e \setminus \{j\}}\right) \overset{\sigma}{\backsim} \left(X^k, \Sigma^k\right)$$

con la probabilidad $\tilde{\mathbb{P}}$ sobre $\left(\tilde{X}^\omega, \tilde{\Sigma}^\omega\right)$ tal que si

$$A_m \in \tilde{\Sigma} \quad m = 1, \cdots, n$$

con

$$A_m = A_{(m,0)} \times \prod_{i \in [k] \setminus \{j\}} A_{(m,i)}$$

entonces

$$\tilde{\mathbb{P}}(\tilde{x}_{w_1} \in A_1, \cdots, \tilde{x}_{w_n} \in A_n) =$$

$$\mathbb{P}\big(x_{w_m} \in A_{(m,0)}, \quad x_{r_{j,i}(w_m)} \in A_{(m,i)} \quad \forall i \in [k] \setminus \{j\} \quad m = 1, \cdots, n\big)\,.$$

Si $\rho : \tilde{X}^\omega \to X^\omega$ es, como antes, la proyección índice por índice sobre la primera coordenada, tenemos

$$\mathbb{P} = \tilde{\mathbb{P}} \circ \rho^{-1}$$

y $\tilde{\mathbb{P}}$ es extensión de $\mathbb{P}$. Transfiriendo la integral anterior al espacio $\tilde{X}_w$ con w arbitraria que contenga al menos una letra j, tenemos

$$\int_{\tilde{X}} \left(\prod_{i \in e, i<j} f_i \circ \rho_{(i,w)}\right) \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \cdot \left(\prod_{i \in e, i>j} \mathbb{E}(f_i \,|\, \Xi_i) \circ \rho_{(i,w)}\right) d\tilde{\mathbb{P}}_w$$

distinto a cero, donde, como ya es costumbre,

$$\rho_{(i,w)} : \tilde{X}_w \to X_{r_{i,j}(w)} \quad i \in [k] \setminus \{j\}$$

son las proyecciones sobre la i-ésima coordenada de cada $\tilde{X}_w$ y

$$\rho_0 : \tilde{X}_w \to X_w$$

es la proyección unidimensional sobre la primera coordenada. Notamos que las proyecciones

$$\rho_{(i,w)} \quad \text{son} \quad \{i, j\}\text{-insensibles} \quad \forall i \in [k] \setminus \{j\}\,,$$

y por lo tanto son $\tilde{\Xi}_j$-medibles; las funciones compuestas

$$f_i \circ \rho_{(i,w)} \quad \text{son también} \quad \tilde{\Xi}_j\text{-medibles} \quad \forall i \in [k] \setminus \{j\}\,.$$

Abreviando
$$h_w := \left(\prod_{i\in e, i<j} f_i \circ \rho_{(i,w)}\right) \cdot \left(\prod_{i\in e, i>j} \mathbb{E}(f_i \,|\, \Xi_i) \circ \rho_{(i,w)}\right)$$
tenemos
$$\int_{\tilde{X}} h_w \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0)\ \tilde{\mathbb{P}}_w = \kappa \neq 0$$
con
$$h_w \quad \tilde{\Xi}_j\text{-medible.}$$

Ahora bien, si w contiene letras j y tomamos una sucesión de subespacios en refinación
$$(\psi_m)_{m\geq 1}$$
tendremos que $\psi_m(w)$ contiene letras j, para toda $m \geq 1$. Por el Teorema 28, podemos elegir los subespacios en refinación de tal manera que la sucesión
$$\left(\tilde{\mathbb{P}}_{\psi_m}\right)_{m\geq 1}$$
converge bajo la métrica de Prohorov a una distribución fuertemente estacionaria, a la que llamaremos $\tilde{\mathbb{P}}^*$. Tenemos así que
$$\begin{aligned} \int_{\tilde{X}} h_w \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0)\ \tilde{\mathbb{P}}_{\psi_m(w)} &= \\ \int_{\tilde{X}} h_{\psi_m(w)} \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0)\ \tilde{\mathbb{P}}_w &= \kappa \quad \forall m \geq 1; \end{aligned}$$
por convergencia en la topología débil y por la Proposición 33,
$$\int_{\tilde{X}} h_w \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0)\ \tilde{\mathbb{P}}^*_w = \kappa.$$

Por otro lado, al ser $\mathbb{P}$ saciada y $\tilde{\mathbb{P}}^*$ fuertemente estacionaria tenemos
$$\rho_0^{-1}(\Sigma) \quad \perp_{\rho_0^{-1}(\Xi_j)} \tilde{\Xi}_j$$
y, ya que
$$(f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \quad \text{es} \quad \rho_0^{-1}(\Sigma)\text{-medible,}$$
esto implica que
$$\begin{aligned} \tilde{\mathbb{E}}\big(h_w \cdot (f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \ \mid \rho_0^{-1}(\Xi_j)\big) &= \\ \tilde{\mathbb{E}}\big(h_w \mid \rho_0^{-1}(\Xi_j)\big) \cdot \tilde{\mathbb{E}}\big((f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \ \mid \rho_0^{-1}(\Xi_j)\big) \end{aligned} \tag{29}$$
pero
$$\begin{aligned} \tilde{\mathbb{E}}\big((f_j \circ \rho_0 - \mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \ \mid \rho_0^{-1}(\Xi_j)\big) &= \\ \tilde{\mathbb{E}}\big((f_j \circ \rho_0 \mid \rho_0^{-1}(\Xi_j))\big) - \tilde{\mathbb{E}}\big((\mathbb{E}(f_j \,|\, \Xi_j) \circ \rho_0) \ \mid \rho_0^{-1}(\Xi_j)\big), \end{aligned} \tag{30}$$

y

$$\begin{aligned}\tilde{\mathbb{E}}\big((\mathbb{E}(f_j\,|\,\Xi_j)\circ\rho_0)\ \mid \rho_0^{-1}(\Xi_j)\big) &= \mathbb{E}((\mathbb{E}(f_j\,|\,\Xi_j))\ \mid \Xi_j) \\ &= \mathbb{E}(f_j\,|\,\Xi_j) \\ &= \tilde{\mathbb{E}}\big((f_j\circ\rho_0 \mid \rho_0^{-1}(\Xi_j))\big)\end{aligned}$$

de donde las igualdades (30) y (29) son cero.

Pero entonces

$$0 = \int_{\tilde{X}} \tilde{\mathbb{E}}\big(h_w\cdot(f_j\circ\rho_0 - \mathbb{E}(f_j\,|\,\Xi_j)\circ\rho_0)\ \mid \rho_0^{-1}(\Xi_j)\big)\ d\tilde{\mathbb{P}}_w^*$$

$$= \int_{\tilde{X}} h_w\cdot(f_j\circ\rho_0 - \mathbb{E}(f_j\,|\,\Xi_j)\circ\rho_0)\ \tilde{\mathbb{P}}_w^* = \kappa,$$

expresamente una contradicción. Concluímos que la expresión (28) debe ser cero.

$\square$

Definición. Sean $\mathbb{P}$ fuertemente estacionaria sobre X^ω, l cualquier línea combinatoria y $i\in e\subset[k]$. Considerando las proyecciones unidimensionales

$$\pi_j : X^k \to X_j \qquad j\in[k]$$

decimos que la completación de $\pi_i^{-1}(\Phi_e)$ bajo la medida $\mathbb{P}_l$ es la **copia oblicua** de Φ_e y la denotamos por Φ_e^*.

De manera más general, dada una familia ascendente $\mathscr{I}$ decimos que la copia oblicua de su σ-álgebra invariante asociada es

$$\Phi_{\mathscr{I}}^* := \bigvee_{e\in\mathscr{I}} \Phi_e^*.$$

Veamos que este objeto está bien definido: dado $e\subset[k]$ y

$$\pi_i^{-1}(\Phi_e)\,,\quad \pi_j^{-1}(\Phi_e) \qquad i,j\in e$$

queremos ver que ambas σ-álgebras difieren solamente por conjuntos $\mathbb{P}_l$-nulos. Es decir, dado $A\in\pi_i^{-1}(\Phi_e)$ queremos exhibir $B\in\pi_j^{-1}(\Phi_e)$ tal que

$$A\Delta B\subset C \quad \text{para algún} \quad C\in\Sigma^k \quad \text{tal que} \quad \mathbb{P}_l(C)=0.$$

Digamos que

$$A := \pi_i^{-1}(A') \quad A'\in\Phi_e$$

y recordamos:

$$\mathbb{P}_l\Big(\Big\{(x_s)_{s=1}^k : \quad \mathbf{1}_{A'}(x_i) = \mathbf{1}_{A'}(x_j)\Big\}\Big) = 1.$$

Pero
$$\pi_i^{-1}(A')\,\Delta\pi_j^{-1}(A') = \left\{(x_s)_{s=1}^k : \quad \mathbf{1}_{A'}(x_i) = \mathbf{1}_{A'}(x_j)\right\}^c$$
y
$$\mathbb{P}_l\Big(\left\{(x_s)_{s=1}^k : \quad \mathbf{1}_{A'}(x_i) = \mathbf{1}_{A'}(x_j)\right\}^c\Big) = 0.$$
Siendo así, la copia oblicua de Φ_e no cambia según el índice en e que se eliga, y está bien definida.

En términos intuitivos, la copia oblicua de Φ_e reduce Σ^k a la información necesaria para considerar a los elementos de e como bloque homogéneo.

Definición. Denotamos por $\mathscr{P}_{\geq 2}([k])$ a la familia de todos los subconjuntos de $[k]$ que contienen al menos dos elementos.

Si $\mathscr{I} \subset \mathscr{P}_{\geq 2}([k])$ es familia ascendente, decimos que $e \in \mathscr{P}_{\geq 2}([k]) \setminus \mathscr{I}$ es **maximal ajeno** a $\mathscr{I}$ si
$$e \cup \{j\} \in \mathscr{I} \quad \forall\, j \in [k] \setminus e.$$

Dado $a \in \mathscr{I}$ decimos que a es **elemento minimal** de $\mathscr{I}$ si
$$a \setminus \{j\} \notin \mathscr{I} \quad \forall\, j \in a.$$

El **sustento de elementos minimales** o **sustento minimal** de $\mathscr{I}$ es una colección de elementos minimales
$$\{a_1, a_2, \cdots, a_q\}$$
tales que la familia ascendente más pequeña que los contiene es precisamente $\mathscr{I}$. En otras palabras,
$$\mathscr{I} = \bigcup_{i=1}^q \{\{a_i\}\} \cup \{\{a_i \cup b\} : \quad b \subset [k] \setminus a_i\}\,.$$

Existencia y unicidad. Veamos que el sustento minimal existe y es único: Sea $\mathscr{I}$ familia ascendente y $a \in \mathscr{I}$. Consideramos dos casos: en el primero, para todo simplejo
$$\{j\} \subset a$$
tenemos
$$a \setminus \{j\} \notin \mathscr{I},$$
y entonces a es elemento minimal de $\mathscr{I}$. En el segundo caso,
$$a \setminus \{j\} \in \mathscr{I} \qquad \text{para algún} \;\; j \in a.$$
Si el conjunto $a \setminus \{j\}$ cae en el primer caso, es minimal; de otra manera existe $i \in a$ tal que
$$a \setminus \{j, i\} \in \mathscr{I}.$$

Procediendo de manera recursiva - considerando todas las letras en a- existe un conjunto

$$a' \subset a \quad \text{tal que} \quad a' \quad \text{es minimal}$$

y

$$\langle a' \rangle \subset \mathscr{I}.$$

A continuación, suponemos que existe un elemento

$$b \in \mathscr{I} \smallsetminus \langle a' \rangle;$$

de no existir tal elemento tendríamos que

$$\mathscr{I} = \langle a' \rangle$$

y que

$$\{a'\}$$

es sustento minimal de $\mathscr{I}$. Entonces bien, por el argumento anterior existe

$$b' \subset b$$

tal que b' es elemento minimal de $\mathscr{I}$ y

$$\langle b' \rangle \in \mathscr{I}.$$

La cantidad de subconjuntos de $[k]$ es finita, por lo que procediendo de manera recursiva encontramos una colección exhaustiva de elementos minimales de $\mathscr{I}$; dicha colección es sustento minimal. Nótese que todo elemento de $\mathscr{I}$ contiene a algún minimal de la colección.

Para demostrar la unicidad, supongamos que

$$\{a_1, \cdots, a_n\} = \mathscr{A}$$

es sustento minimal de $\mathscr{I}$ y que $b \notin \mathscr{A}$ es elemento minimal de $\mathscr{I}$. Entonces

$$\langle b \rangle \in \mathscr{I},$$

de donde

$$\langle b \rangle \subset \langle a_j \rangle \qquad \text{para algún} \quad j \in \{1, \cdots, n\}.$$

Pero esto es imposible, ya que

$$b \neq a_j$$

y ambos son minimales. □

Proposición 35. *Si $\mathscr{G}$, $\mathscr{H}_1$ y $\mathscr{H}_2$ son subσ-álgebras de Σ, una condición suficiente para*

$$\mathscr{H}_1 \perp_{\mathscr{G}} \mathscr{H}_2$$

es que

$$\mathbb{E}(H_1 \cdot H_2) = \mathbb{E}(H_1 \cdot \mathbb{E}(\, H_2 \mid \mathscr{G}))$$

para cualesquiera H_i acotadas y $\mathscr{H}_i$-medibles.

DEMOSTRACIÓN. Por definición de esperanza condicional

$$\mathbb{E}(\mathbf{1}_A H_1 \cdot H_2) = \mathbb{E}(\mathbf{1}_A \mathbb{E}(H_1 \cdot H_2 \mid \mathscr{G})) \quad \forall A \in \mathscr{G}.$$

Por otro lado, dado $A \in \mathscr{G} \subset \mathscr{H}_1$ tenemos que $\mathbf{1}_A H_1$ es $\mathscr{H}_1$-medible y por hipótesis

$$\begin{aligned} \mathbb{E}(\mathbf{1}_A \cdot H_1 \cdot H_2) &= \mathbb{E}(\mathbf{1}_A \cdot H_1 \cdot \mathbb{E}(H_2 \mid \mathscr{G})) \\ &= \mathbb{E}(\mathbf{1}_A \cdot \mathbb{E}(H_1 \mid \mathscr{G}) \cdot \mathbb{E}(H_2 \mid \mathscr{G})) \end{aligned}$$

ya que $\mathbf{1}_A \cdot \mathbb{E}(H_2 \mid \mathscr{G})$ es $\mathscr{G}$-medible. Entonces

$$\mathbb{E}(H_1 \cdot H_2 \mid \mathscr{G}) = \mathbb{E}(H_1 \mid \mathscr{G}) \cdot \mathbb{E}(H_2 \mid \mathscr{G}) \quad \text{c.s.}$$ □

PROPOSICIÓN 36. *Sean $\mathbb{P}$ distribución saciada sobre*

$$(X^\omega, \Sigma^\omega)$$

y $i \in e \subset [k]$. La distribución

$$\mathbb{P}_{e,i} := \mathbb{P} \circ r_{e,i}^{-1}$$

sobre

$$\left(X^{([k]\smallsetminus e)\cup\{i\}^\omega}, \Sigma^{([k]\smallsetminus e)\cup\{i\}^\omega}\right)$$

es saciada.

DEMOSTRACIÓN. Afirmamos que la estacionariedad fuerte de $\mathbb{P}_{e,i}$ se hereda de $\mathbb{P}$, en el sentido de que

$$r_{e,i}([k]^\omega) \quad \text{y} \quad r_{e,i}(\psi([k]^\omega)) \quad \text{con} \quad \psi \quad \text{subespacio}$$

tienen la misma distribución, y cada subespacio de

$$\{[k] \smallsetminus (e \smallsetminus \{i\})\}^\omega$$

es equivalente a

$$r_{e,i}(\psi([k]^\omega))$$

para algún subespacio ψ de $[k]^\omega$.

Para comprobar esto conviene pensar el k-árbol correspondiente a X^ω. Generamos un árbol bien podado en dos etapas. Primero, cortamos todas las ramas necesarias para reducir el alfabeto de $[k]$ a $[k] \smallsetminus (e \smallsetminus \{i\})$. Después, cortamos todas las ramas que comiencen en vértices que no estén en ψ, junto con sus primos afines. Añadiendo las aristas necesarias, el resultado es un árbol bien podado correspondiente al espacio

$$r_{e,i}(\psi([k]^\omega)) .$$

Por otro lado, si invertimos el procedimiento, y primero cortamos las ramas correspondientes a los vértices no contenidos en ψ, con sus primos afines, y despúes las ramas de los vértices que tengan letras en $e \setminus \{i\}$, el resultado es el mismo. Como todo subespacio de

$$\{[k] \setminus (e \setminus \{i\})\}^\omega$$

está contenido en un subespacio de $[k]^\omega$, se sigue el resultado.

Demostremos pues que $\mathbb{P}_{e,i}$ es saciada: Toda extensión fuertemente estacionara $\tilde{\mathbb{P}}_{e,i}$ es una extensión coordenada a coordenada

$$\pi : \tilde{X} \to X$$

y, por lo tanto, permite una extensión fuertemente estacionaria $\tilde{\mathbb{P}}$ sobre

$$(X^\omega, \Sigma^\omega, \mathbb{P})$$

tal que

$$\tilde{\mathbb{P}} \circ r_{e,i}^{-1} = \tilde{\mathbb{P}}_{e,i}.$$

Entonces, si $\mathbb{P}_{e,i}$ no fuese saciada, existirían una familia ascendente $\mathscr{I}$ y una extensión $\tilde{\pi}$ tales que

$$\tilde{\Phi}_{r_{e,i}(\mathscr{I})} \quad \text{y} \quad \tilde{\pi}^{-1}\big(\Phi_{r_{e,i}(\mathscr{I})}\big)$$

no fuesen condicionalmente independientes dada

$$\tilde{\pi}^{-1}(\Sigma)\,.$$

Pero en ese caso

$$\tilde{\Phi}_{\mathscr{I}} \quad \text{y} \quad \tilde{\pi}^{-1}(\Phi_{\mathscr{I}})$$

no serían condicionalmente independientes dada

$$\tilde{\pi}^{-1}(\Sigma)\,,$$

lo cual es imposible ya que $\mathbb{P}$ es saciada. □

TEOREMA 40. *Si $\mathbb{P}$ es saciada, l es cualquier línea combinatoria, $\mathscr{I}$ es familia ascendente y $e \subset [k]$ es maximal ajeno a $\mathscr{I}$, entonces las σ-álgebras oblicuas Φ_e^* y $\Phi_{\mathscr{I}}^*$ son condicionalmente independientes dada $\Phi_{\langle e\rangle\cap\mathscr{I}}^*$ en el espacio*

$$\big(X^k, \Sigma^k, \mathbb{P}_l\big)\,.$$

DEMOSTRACIÓN. Sean F_1 y F_2 funciones en $L_\infty(\mathbb{P}_l)$ tales que

$$F_1 \ \text{ es } \ \Phi_e^*\text{-medible y } F_2 \ \text{ es } \ \Phi_{\mathscr{I}}^*\text{-medible.}$$

Por la Proposición 41 será suficiente demostrar que

$$\mathbb{E}_l(F_1F_2) = \mathbb{E}_l\Big(\mathbb{E}_l\Big(F_1 \ \Big|\ \Phi^*_{\mathscr{I}\cap\langle e\rangle}\Big) \cdot F_2\Big)\,.$$

Consideramos el sustento minimal de $\mathscr{I}$:

$$\{a_1, a_2, \cdots, a_q\}$$

y notamos que
$$a_s \setminus e \neq \emptyset, \qquad 1 \leq s \leq q.$$
Elegimos
$$i_s \in a_s \setminus e$$
de manera arbitraria y consideramos que la σ-álgebra oblicua asociada a a_s es la completación bajo $\mathbb{P}_l$ de
$$\pi_{i_s}^{-1}(\Phi_{a_s})\,.$$

Si recordamos que
$$\Phi^*_{\mathscr{I}} = \bigvee_{1\leq s\leq q} \Phi^*_{a_s}$$
y retomamos la Proposición 34, podemos suponer que F_2 es producto de funciones a_s-invariantes con $s = 1, \cdots, q$; es decir que
$$F_2 := \prod_{1\leq s\leq q} \eta_s \qquad \text{donde cada } \eta_s \quad \text{es } \Phi^*_{a_s}\text{-medible.}$$
En efecto, para cada η_s podemos elegir
$$f_s : X \to \mathbb{R} \quad \Phi_{a_s}\text{-medible}$$
tal que
$$\eta_s = f_s \circ \pi_{i_s}.$$

Elegimos $i \in e$ y $f_1 \in L_\infty(\mathbb{P}_w)$ e-invariante tales que
$$F_1 = f_1 \circ \pi_i \quad \mathbb{P}_l\text{-c.s.}$$
Entonces
$$\mathbb{E}_l(F_1F_2) = \int_{X^k} f_1 \circ \pi_i \cdot \prod_{1\leq s\leq q} f_s \circ \pi_{i_s}\, d\mathbb{P}_l. \tag{31}$$
Por otro lado, por la Proposición 36 la distribución $\mathbb{P}_{e,i}$ es fuertemente estacionaria y saciada sobre
$$\left(X^{r_{e,i}([k]^\omega)}, \Sigma^{r_{e,i}([k]^\omega)}\right).$$
Notamos que
$$r_{e,i}(l([k]))$$
es línea combinatoria en
$$X^{r_{e,i}([k])^\omega} = X^{([k]\setminus e)\cup\{i\}^\omega}$$
y consideramos el subespacio-línea de éste
$$\left(X^{k-|e|+1}, \Sigma^{k-|e|+1}, \mathbb{P}_{r_{e,i}(l([k]))}\right),$$

donde

$$\mathbb{P}_{r_{e,i}(l([k]))} := \mathbb{P}_{e,i} \circ r_{e,i}(l([k]))^{-1}.$$

Expresamos la integral (31) en este espacio

$$\mathbb{E}_l(F_1F_2) = \int_{X^{k-|e|+1}} f_1 \circ \pi_i \cdot \prod_{1\leq s\leq q} f_s \circ \pi_{i_s} \, d\mathbb{P}_{r_{e,i}(l)}$$

y por el Teorema 39

$$\mathbb{E}_l(F_1F_2) = \int_{X^{k-|e|+1}} \mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t\in[k]\setminus e} \Phi_{\{i,t\}} \right) \circ \pi_i \cdot \prod_{1\leq s\leq q} \mathbb{E}\left(f_s \,\middle|\, \bigvee_{t\in[k]\setminus\{i_s\}} \Phi_{\{i_s,t\}} \right) \circ \pi_{i_s} \, d\mathbb{P}_{r_{e,i}(l)}.$$

Pero cada f_s es Φ_{a_s}-medible para $s = 1, \cdots, q$ y

$$\Phi_{a_s} \subset \bigcap_{t\in a_s\setminus\{i_s\}} \Phi_{\{i_s,t\}} \subset \bigvee_{t\in[k]\setminus\{i_s\}} \Phi_{\{i_s,t\}}$$

de donde

$$\mathbb{E}\left(f_s \,\middle|\, \bigvee_{t\in[k]\setminus\{i_s\}} \Phi_{\{i_s,t\}} \right) = f_s \quad \mathbb{P}_l\text{-c.s.}$$

Regresando al espacio original tenemos que

$$\mathbb{E}_l(F_1F_2) = \int_{X^k} \mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t\in[k]\setminus e} \Phi_{\{i,t\}} \right) \circ \pi_i \cdot F_2 \, d\mathbb{P}_l.$$

Por el Teorema 37 y dado que f_1 es e-invariante,

$$\mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t\in[k]\setminus e} \Phi_{\{i,t\}} \right) = \mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t\in[k]\setminus e} \Phi_{e\cup\{t\}} \right).$$

Entonces, para concluír la demostración basta establecer que

$$\mathbb{E}_l\left(F_1 \,\middle|\, \Phi^*_{\mathscr{I}\cap\langle e\rangle} \right) = \mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t\in[k]\setminus e} \Phi_{e\cup\{t\}} \right) \circ \pi_i.$$

Notamos que

$$e\cup\{t\} \in \mathscr{I} \qquad \forall\, t \in [k]\setminus e$$

ya que e es maximal ajeno a $\mathscr{I}$. La familia

$$\{e \cup a_1, e \cup a_2, \cdots, e \cup a_q\}$$

es sustento minimal de $\mathscr{I} \cap \langle e \rangle$, pero

$$e \cup \{i_s\} \subset e \cup a_s \quad \text{y} \quad e \cup \{i_s\} \in \mathscr{I} \cap \langle e \rangle \qquad 1 \leq s \leq q;$$

entonces necesariamente

$$e \cup \{i_s\} = e \cup a_s \qquad 1 \leq s \leq q.$$

Por lo tanto,

$$\Phi^*_{\mathscr{I} \cap \langle e \rangle} = \bigvee_{1 \leq s \leq q} \Phi^*_{e \cup \{i_s\}}$$

de donde

$$\begin{aligned}
\mathbb{E}_l\Big(F_1 \,\Big|\, \Phi^*_{\mathscr{I} \cap \langle e \rangle}\Big) &= \mathbb{E}_l\left(F_1 \,\middle|\, \bigvee_{1 \leq s \leq q} \Phi^*_{e \cup \{i_s\}}\right) \\
&= \mathbb{E}_l\left(f_1 \circ \pi_i \,\middle|\, \bigvee_{1 \leq s \leq q} \Phi^*_{e \cup \{i_s\}}\right) \\
&= \mathbb{E}_l\left(f_1 \circ \pi_i \,\middle|\, \bigvee_{t \in [k] \smallsetminus e} \Phi^*_{e \cup \{t\}}\right) \\
&= \mathbb{E}_l\left(f_1 \circ \pi_i \,\middle|\, \bigvee_{t \in [k] \smallsetminus e} \pi_i^{-1}\big(\Phi_{e \cup \{t\}}\big)\right) \\
&= \mathbb{E}_l\left(f_1 \circ \pi_i \,\middle|\, \pi_i^{-1}\left(\bigvee_{t \in [k] \smallsetminus e} \Phi_{e \cup \{t\}}\right)\right) \\
&= \mathbb{E}\left(f_1 \,\middle|\, \bigvee_{t \in [k] \smallsetminus e} \Phi_{e \cup \{t\}}\right) \circ \pi_i.
\end{aligned}$$

$\square$

A partir del Teorema 40 llegamos a través de un argumento inductivo al siguiente resultado:

TEOREMA 41. *Si $\mathbb{P}$ es netamente saciada y $\mathscr{I}$, $\mathscr{I}'$ son familias ascendentes entonces las σ-álgebras oblicuas $\Phi^*_{\mathscr{I}}$ y $\Phi^*_{\mathscr{I}'}$ son condicionalmente independientes dada $\Phi^*_{\mathscr{I}\cap\mathscr{I}'}$ en el espacio*

$$\left(X^k, \Sigma^k, \mathbb{P}_l\right).$$

DEMOSTRACIÓN. Fijamos $\mathscr{I}$ y procedemos por un argumento de inducción fuerte sobre $\mathscr{I}'$. Digamos que

$$|\mathscr{I}'| = C$$

y que el resultado es cierto para toda familia ascendente $\mathscr{I}''$ tal que

$$|\mathscr{I}''| < C.$$

Como base inductiva consideramos

$$\mathscr{I}'' = \{[k]\}$$

y notamos que

$$\mathscr{I} \cap \{[k]\} = \{[k]\}.$$

Claramente

$$\Phi^*_{\mathscr{I}} \perp_{\Phi^*_{[k]}} \Phi^*_{[k]}.$$

Para el paso inductivo queremos demostrar entonces que

$$\Phi^*_{\mathscr{I}} \quad \perp_{\Phi^*_{\mathscr{I}\cap\mathscr{I}'}} \quad \Phi^*_{\mathscr{I}'}.$$

Sean F $\Phi^*_{\mathscr{I}'}$-medible y H $\Phi^*_{\mathscr{I}}$-medible. Si $\mathscr{I}' \subset \mathscr{I}$ entonces

$$\mathbb{E}_l(\,FH \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'}, \Phi^*_{\mathscr{I}'}) = \mathbb{E}_l(\,FH \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'})$$

y el resultado es inmediato. Suponemos pues

$$\mathscr{I}' \smallsetminus \mathscr{I} \neq \emptyset.$$

Por la Proposición 41 basta con demostrar que

$$\mathbb{E}_l(FH) = \mathbb{E}_l(H\ \mathbb{E}_l(\,F \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'}))\,;$$

por otro lado

$$\mathbb{E}_l(FH) = \mathbb{E}_l(H\ \mathbb{E}_l(\,F \mid \Phi^*_{\mathscr{I}}))\,.$$

Entonces será suficiente demostrar que

$$\mathbb{E}_l(\,F \mid \Phi^*_{\mathscr{I}}) = \mathbb{E}_l(\,F \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'}) \quad \mathbb{P}_l\text{-c.s.}$$

Consideramos un elemento minimal

$$e \in \mathscr{I}' \smallsetminus \mathscr{I}$$

de tamaño máximo de entre todos los elementos minimales contenidos en esa diferencia; definimos

$$\mathscr{I}'' := \mathscr{I}' \smallsetminus \{e\}$$

y notamos que

$$\Phi^*_{\mathscr{I}'} = \Phi^*_{\mathscr{I}''} \vee \Phi^*_{\langle e\rangle};$$

entonces por la Proposición 34 podemos suponer sin pérdida de generalidad que

$$F = F_1 \cdot F_2$$

con

$$F_1 \quad \Phi^*_{\langle e\rangle}\text{-medible} \quad \text{y} \quad F_2 \quad \Phi^*_{\mathscr{I}''}\text{-medible.}$$

De este modo tenemos

$$\begin{aligned} \mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}}) &= \mathbb{E}_l(\mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}\cup\mathscr{I}''}) \mid \Phi^*_{\mathscr{I}}) && (32)\\ &= \mathbb{E}_l(\mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}\cup\mathscr{I}''}) \cdot F_2 \mid \Phi^*_{\mathscr{I}})\,. && (33) \end{aligned}$$

Por el Teorema 40 y dado que e es maximal ajeno a $\mathscr{I}\cup\mathscr{I}''$,

$$\mathbb{E}_l\Big(\, F_1 \;\Big|\; \Phi^*_{\mathscr{I}\cup\mathscr{I}''},\, \Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle}\Big) = \mathbb{E}_l\Big(\, F_1 \;\Big|\; \Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle}\Big)\,,$$

pero

$$\Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle} = \bigvee_{a\in(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle} \Phi^*_a \quad\subset \bigvee_{a\in\mathscr{I}\cup\mathscr{I}''} \Phi^*_a \quad=\quad \Phi^*_{\mathscr{I}\cup\mathscr{I}''}$$

y entonces

$$\mathbb{E}_l\Big(\, F_1 \;\Big|\; \Phi^*_{\mathscr{I}\cup\mathscr{I}''},\, \Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle}\Big) = \mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}\cup\mathscr{I}''})\,,$$

de modo que

$$\mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}\cup\mathscr{I}''}) = \mathbb{E}_l\Big(\, F_1 \;\Big|\; \Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle}\Big)\,. \tag{34}$$

Por otro lado

$$(\mathscr{I}\,\cup\mathscr{I}'')\cap\langle e\rangle \subset \mathscr{I}''$$

por definición de $\mathscr{I}''$ y ya que $\mathscr{I}'$ es familia ascendente. Repitiendo el argumento anterior tenemos

$$\mathbb{E}_l\Big(\, F_1 \;\Big|\; \Phi^*_{(\mathscr{I}\cup\mathscr{I}'')\cap\langle e\rangle}\Big) = \mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}''})\,. \tag{35}$$

Juntando las ecuaciones (34) y (35) tenemos

$$\mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}\cup\mathscr{I}''}) = \mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}''})$$

y sustituyendo en la ecuación (32)

$$\begin{aligned} \mathbb{E}_l(\mathbb{E}_l(\, F_1 \mid \Phi^*_{\mathscr{I}''}) \cdot F_2 \mid \Phi^*_{\mathscr{I}}) &= \mathbb{E}_l(\mathbb{E}_l(\, F_1 \cdot F_2 \mid \Phi^*_{\mathscr{I}''}) \mid \Phi^*_{\mathscr{I}}) && (36)\\ &= \mathbb{E}_l(\mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}''}) \mid \Phi^*_{\mathscr{I}})\,. && (37) \end{aligned}$$

Ahora bien, la hipótesis inductiva nos asegura que

$$\Phi^*_{\mathscr{I}} \quad \perp_{\Phi^*_{\mathscr{I}\cap\mathscr{I}''}} \quad \Phi^*_{\mathscr{I}''}$$

y la función

$$H := \mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}''}) \quad \text{es} \quad \Phi^*_{\mathscr{I}''}\text{-medible,}$$

por lo que

$$\mathbb{E}_l(\, H \mid \Phi^*_{\mathscr{I}}, \Phi^*_{\mathscr{I}\cap\mathscr{I}''}) = \mathbb{E}_l(\, H \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}''}) \,.$$

Pero

$$\Phi^*_{\mathscr{I}\cap\mathscr{I}''} \subset \Phi^*_{\mathscr{I}};$$

luego

$$\mathbb{E}_l(\, H \mid \Phi^*_{\mathscr{I}}) = \mathbb{E}_l(\, H \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}''}) \,.$$

Regresando a la ecuación (36)

$$\begin{aligned} \mathbb{E}_l(\mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}''}) \mid \Phi^*_{\mathscr{I}}) &= \mathbb{E}_l(\mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}''}) \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}''}) \\ &= \mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}''}) \\ &= \mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'}) \end{aligned}$$

donde usamos la propiedad de la torre de la esperanza condicional para la segunda igualdad y la identidad

$$\mathscr{I} \cap \mathscr{I}'' = \mathscr{I} \cap \mathscr{I}'$$

para la tercera.

Sustituyendo una vez más en la ecuación (32) hemos demostrado que

$$\mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}}) = \mathbb{E}_l(\, F \mid \Phi^*_{\mathscr{I}\cap\mathscr{I}'}) \,,$$

terminando con esto el argumento inductivo. □

6. Un argumento recursivo sobre familias ascendentes.

Finalmente estamos en condiciones de dar una demostración ergódica-probabilística del Teorema de Densidad de Hales Jewett; el eslabón faltante es el siguiente resultado.

Teorema 42. *Sean $\mathbb{P}$ una distribución netamente saciada, $\mathbb{P}_l$ su proyección sobre cualquier línea combinatoria y $\mathbb{P}_w$ su proyección sobre cualquier palabra. Sea*

$$(\mathscr{I}_{i,j})_{i,j} \qquad i = 1, 2, \cdots, k \quad j = 1, 2, \cdots, c_i$$

colección de familias ascendentes tal que

$$\mathscr{I}_{i,j} \subset \langle i \rangle \qquad \forall i, j.$$

Si una familia de conjuntos

$$(A_{i,j})_{(i,j)} \qquad A_{i,j} \in \Phi_{\mathscr{I}_{i,j}}$$

satisface

$$\mathbb{P}_l\left(\bigcap_{j=1}^{c_1} A_{1,j} \times \bigcap_{j=1}^{c_2} A_{2,j} \times \cdots \times \bigcap_{j=1}^{c_k} A_{k,j}\right) = 0$$

entonces

$$\mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} A_{i,j}\right) = 0.$$

Antes de proceder a demostrar el Teorema 42 veamos cómo, a partir de él, podemos demostrar DHJ (Teorema 22). Retomando el teorema 29 basta considerar distribuciones saciadas.

Demostración de DHJ a partir del Teorema 42. Sea $\mathbb{P}$ distribución saciada sobre

$$(X^\omega, \Sigma^\omega)$$

y $A \in \Sigma$ tal que

$$\mathbb{P}_l(A \times A \times \cdots \times A) = 0.$$

Queremos demostrar que entonces

$$\mathbb{P}_w(A) = 0$$

para establecer el resultado por contrapositiva.

Entonces bien, por el Teorema 39

$$\begin{aligned}
\int_{X^k} \prod_{i=1}^{k} \mathbb{E}\big(\mathbf{1}_A \bigm| \Phi_{\langle i\rangle}\big) \circ \pi_i \, d\mathbb{P}_l &= \int_{X^k} \prod_{i=1}^{k} \mathbf{1}_A \circ \pi_i \, d\mathbb{P}_l \\
&= \mathbb{P}_l\big(A^k\big) \\
&= 0
\end{aligned}$$

ya que

$$\langle i\rangle = \bigcup_{j\in[k]\smallsetminus\{i\}} \langle i,j\rangle$$

y entonces

$$\Phi_{\langle i\rangle} = \bigvee_{j\in[k]\smallsetminus\{i\}} \Phi_{\{i,j\}} \qquad \forall\, i \in [k].$$

El conjunto

$$B_i := \big\{\mathbb{E}\big(\mathbf{1}_A \bigm| \Phi_{\langle i\rangle}\big) > 0\big\}$$

pertenece a $\Phi_{\langle i\rangle}$ y por la igualdad anterior tenemos

$$\mathbb{P}_l(B_1 \times B_2 \times \cdots \times B_n) = 0.$$

Entonces el Teorema 42 nos dice, considerando

$$\mathscr{I}_{i,1} := \langle i \rangle \qquad c_i := 1 \quad \forall\, i \in [k],$$

que

$$\mathbb{P}_w(B_1 \cap B_2 \cap \cdots \cap B_n) = 0.$$

Por otra parte afirmamos que

$$\mathbb{P}_w(A \setminus B_i) = 0 \qquad i \in [k]$$

pues, en efecto,

$$\mathbb{P}_w(B_i \cap A) = \int_{B_i} \mathbf{1}_A \, d\mathbb{P}_w = \int_{B_i} \mathbb{E}\big(\mathbf{1}_A \;\big|\; \Phi_{\langle i \rangle}\big)\; d\mathbb{P}_w =$$
$$\int_X \mathbb{E}\big(\mathbf{1}_A \;\big|\; \Phi_{\langle i \rangle}\big)\; d\mathbb{P}_w = \int_X \mathbf{1}_A \, d\mathbb{P}_w = \mathbb{P}_w(A)\,.$$

Siendo así,

$$\mathbb{P}_w(A) \leq \mathbb{P}_w(B_1 \cap B_2 \cap \cdots \cap B_k) + \sum_{i=1}^{k} \mathbb{P}_w(A \setminus B_i) = 0.$$

$\square$

El peso de la demostración descansa ahora sobre el Teorema 42, el cual procedemos a demostrar a partir de un argumento inductivo.

6.1. Transfiriendo nulidad en el espacio-línea a nulidad en el espacio unidimensional para familias de σ-álgebras invariantes. Haremos un argumento inductivo de dos pasos sobre la *profunidad* de colecciones de familias ascendentes en $[k]$.

DEFINICIÓN. La **profundidad** de una familia ascendente $\mathscr{I}$ es la cardinalidad mínima de los conjuntos en $\mathscr{I}$. En símbolos

$$p(\mathscr{I}) := \min \{|e| : \quad e \in \mathscr{I}\}\,.$$

DEFINICIÓN. Decimos que una colección de familias ascendentes

$$(\mathscr{I}_{i,j})_{i,j}$$

tiene la propiedad $\mathcal{P}$ si satisface la afirmación del Teorema 42; es decir si

$$\mathbb{P}_l\left(\bigcap_{j=1}^{c_1} A_{1,j} \times \cdots \times \bigcap_{j=1}^{c_k} A_{k,j}\right) = 0$$

implica

$$\mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} A_{i,j}\right) = 0$$

para cualquier distribución saciada $\mathbb{P}$ y siempre que

$$A_{i,j} \in \mathscr{I}_{i,j} \qquad \forall\,(i,j)\,.$$

De aquí en adelante siempre indexamos nuestras colecciones de tal manera que

$$\mathscr{I}_{i,j} \subset \langle i\rangle.$$

Recordemos que una familia ascendente $\mathscr{I}$ es **principal** simpre que

$$\mathscr{I} = \langle e\rangle$$

para algún $e \subset [k]$.

Las siguientes dos proposiciones nos permitirán hacer el paso inductivo de la demostración. La base inductiva la exponemos después.

PROPOSICIÓN 37. *Supongamos que todas las familias ascendentes en la colección*

$$\mathscr{C} := \left(\mathscr{I}_{i,j}\right)_{i,j}$$

tienen profundidad mayor o igual a M, que hay L familias con profundidad M, y que todas esas L familias son principales. Si toda colección semejante que tenga solamente $L-1$ familias ascendentes de profundidad M, todas ellas principales, satisface la propiedad $\mathcal{P}$, entonces la colección $\mathscr{C}$ también la satisface.

DEMOSTRACIÓN. Sean

$$\mathscr{I}_{i_1,j_1} = \langle e_1\rangle, \cdots, \mathscr{I}_{i_L,j_L} = \langle e_L\rangle$$

las familias ascendentes de profundidad M en $\mathscr{C}$.

Consideramos dos casos: en el primero,

$$e_s = e_r \quad \text{para algunos} \quad r, s \in \{1, \cdots, L\}\,.$$

Sean

$$A_{i,j} \in \Phi_{\mathscr{I}_{i,j}} \quad i = 1, \cdots, k \quad j = 1, \cdots, c_i$$

tales que

$$\mathbb{P}_l\left(\bigcap_{j=1}^{c_1} A_{1,j} \times \cdots \times \bigcap_{j=1}^{c_k} A_{k,j}\right) = 0. \tag{38}$$

Reemplazamos a
$$A_{i_s,j_s} \quad \in \Phi_{\mathscr{I}_{i_s,j_s}}$$
por
$$A'_{i_s,j_s} := A_{i_s,j_s} \cap A_{i_r,j_r} \quad \in \Phi_{\mathscr{I}_{i_s,j_s}}$$
y a
$$A_{i_r,j_r} \quad \in \Phi_{\mathscr{I}_{i_r,j_r}}$$
por
$$A'_{i_r,j_r} := X.$$
Escribimos
$$A'_{i,j} := A_{i,j} \qquad \forall\,(i,j) \neq (i_s, j_s)\,.$$
Consideramos ahora dos subcasos: en el primero,
$$i_s = i_r$$
y podemos omitir $\mathscr{I}_{i_r,j_r}$ ya que
$$\bigcap_{j\in[c_{i_s}]} A_{i_s,j} \quad = \bigcap_{j\in[c_{i_s}]\smallsetminus\{j_r\}} A'_{i_s,j};$$
de modo que se preserva la nulidad en (38) cambiando las $A_{i,j}$ por $A'_{i,j}$ y omitiendo A'_{i_r,j_r}. En la colección
$$\mathscr{C}' := \mathscr{C} \smallsetminus \{\mathscr{I}_{i_r,j_r}\}$$
tenemos $L-1$ familias ascendentes principales de profundidad minimal M, y caemos dentro de los supuestos de la hipótesis inductiva. Para el segundo subcaso,
$$i_s \neq i_r,$$
debemos justificar cómo
$$\mathbb{P}_l\left(\bigcap_{j\in[c_1]} A'_{1,j} \times \cdots \times \bigcap_{j\in[c_{i_r}]\smallsetminus\{j_r\}} A'_{i_r,j} \times \cdots \times \bigcap_{j\in[c_k]} A'_{k,j}\right) = 0. \tag{39}$$
Recordemos que
$$\mathscr{I}_{i_s,j_s} \subset \langle i_s\rangle \qquad \mathscr{I}_{i_r,j_r} \subset \langle i_r\rangle$$
y ya que estamos en el caso donde
$$\mathscr{I}_{i_s,j_s} = \mathscr{I}_{i_r,j_r},$$
entonces
$$\mathscr{I}_{i_s,j_s} \subset \langle i_s\rangle \cap \langle i_r\rangle = \langle\{i_s, i_r\}\rangle.$$
Se desprende que
$$\Phi_{\mathscr{I}_{i_s,j_s}} \subset \Phi_{\langle\{i_s,i_r\}\rangle}$$

y de ahí que

$$\mathbb{P}_l\big(\pi_{i_s}^{-1}(A_{i_s,j_s})\,\Delta\pi_{i_r}^{-1}(A_{i_s,j_s})\big) = 0,$$

$$\mathbb{P}_l\big(\pi_{i_s}^{-1}(A_{i_r,j_r})\,\Delta\pi_{i_r}^{-1}(A_{i_r,j_r})\big) = 0.$$

Siendo así,

$$\mathbb{P}_l\Big(\pi_{\{i_s,i_r\}}^{-1}(A_{i_s,j_s}\times A_{i_r,j_r}) \quad \Delta \quad \pi_{i_s}^{-1}(A_{i_s,j_s}\cap A_{i_r,j_r})\Big) = 0,$$

y concluímos que las expresiones en (39) y (38) son equivalentes. Una vez más caemos dentro de los supuestos de la hipótesis inductiva.

Consideremos ahora el segundo caso, es decir cuando no hay repeticiones:

$$e_s \neq e_r \qquad \forall\,(s,r)\in[L]\times[L] \quad \text{con} \quad s\neq r.$$

Suponemos, sin pérdida de generalidad y solamente para facilitar la notación, que $(i_L, j_L) = (1,1)$. Para hacer uso de la hipótesis inductiva definimos la colección

$$\mathscr{C}' := \big(\mathscr{I}'_{i,j}\big)_{i,j}$$

con

$$\mathscr{I}'_{i,j} := \begin{cases} \langle e_L\rangle \smallsetminus \{e_L\}\,, & \text{si} \quad (i,j) = (1,1) \\ \mathscr{I}_{i,j}, & \text{en otro caso} \end{cases}.$$

Hemos reducido en uno el número de familias ascendentes de profundidad M, con el costo de cambiar una familia principal ($\mathscr{I}_{1,1}$) por una no principal ($\mathscr{I}'_{1,1}$); esta observación será importante para deteminar la necesidad de la siguiente proposición que exponemos. La colección $\mathscr{C}'$ cae dentro de los supuestos de la hipótesis inductiva y queremos demostrar a partir de esto que la colección $\mathscr{C}$ satisface la propiedad $\mathcal{P}$.

Ahora bien,

$$0 = \mathbb{P}_l\left(\prod_{i=1}^{k}\bigcap_{j=1}^{c_i} A_{i,j}\right)$$

$$
\begin{aligned}
&= \int_{X^k} \mathbf{1}_{\cap_j A_{1,j} \times \cdots \times \cap_j A_{k,j}} \, d\mathbb{P}_l \\
&= \int_{X^k} \mathbf{1}_{\pi_1^{-1}(\cap_j A_{1,j})} \cdot \cdots \cdot \mathbf{1}_{\pi_k^{-1}(\cap_j A_{k,j})} \, d\mathbb{P}_l \\
&= \int_{X^k} \prod_{i=1}^{k} \prod_{j=1}^{c_i} \mathbf{1}_{\pi_i^{-1}(A_{i,j})} \, d\mathbb{P}_l \\
&= \int_{X^k} \mathbb{E}_l \left(\prod_{i=1}^{k} \prod_{j=1}^{c_i} \mathbf{1}_{\pi_i^{-1}(A_{i,j})} \,\middle|\, \Phi^*_{\langle e_L \rangle \smallsetminus \{e_L\}} \right) d\mathbb{P}_l \\
&= \int_{X^k} \mathbb{E}_l \left(\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \cdot \left(\prod_{j=2}^{c_1} \mathbf{1}_{\pi_1^{-1}(A_{1,j})} \cdot \prod_{i=2}^{k} \prod_{j=1}^{c_i} \mathbf{1}_{\pi_i^{-1}(A_{i,j})} \right) \,\middle|\, \Phi^*_{\langle e_L \rangle \smallsetminus \{e_L\}} \right) d\mathbb{P}_l,
\end{aligned}
$$

una expresión sugerente para retomar el Teorema 41:

$$\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \qquad \text{es} \qquad \Phi^*_{\langle e_L \rangle}\text{-medible}$$

y

$$H := \prod_{j=2}^{c_1} \mathbf{1}_{\pi_1^{-1}(A_{1,j})} \cdot \prod_{i=2}^{k} \prod_{j=1}^{c_i} \mathbf{1}_{\pi_i^{-1}(A_{i,j})} \quad \text{es} \quad \bigvee_{(i,j) \neq (1,1)} \Phi^*_{\mathscr{I}_{i,j}}\text{-medible.}$$

Pero

$$\bigvee_{(i,j) \neq (1,1)} \Phi^*_{\mathscr{I}_{i,j}} \subset \left(\Phi^*_{\langle e_L \rangle \smallsetminus \{e_L\}} \right) \vee \left(\bigvee_{j=1}^{L-1} \Phi^*_{\langle e_j \rangle} \right) =: \mathscr{M}$$

de donde H es también $\mathscr{M}$-medible.

Definimos

$$\mathscr{I}^* := (\langle e_L \rangle \smallsetminus \{e_L\}) \cup \left(\bigcup_{j=1}^{L-1} \langle e_j \rangle \right)$$

para escribir

$$\Phi^*_{\mathscr{I}^*} = \mathscr{M},$$

y por el Teorema 41

$$\Phi^*_{\mathscr{I}^*} \quad \perp_{\Phi^*_{\mathscr{I}^* \cap \langle e_L \rangle}} \Phi^*_{\langle e_L \rangle};$$

pero

$$\mathscr{I}^* \cap \langle e_L \rangle = \langle e_L \rangle \smallsetminus \{e_L\}$$

y entonces

$$\mathbb{E}_l\Big(\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \ \cdot \ H \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big)$$

$$= \mathbb{E}_l\Big(\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big) \cdot \mathbb{E}_l\Big(H \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big) .$$

Por otro lado, recordando la definición de esperanza condicional

$$\int_{X^k} \mathbb{E}_l\Big(\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big) \cdot \mathbb{E}_l\Big(H \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big) \, d\mathbb{P}_l$$

$$= \int_{X^k} \mathbb{E}_l\Big(\mathbf{1}_{\pi_1^{-1}(A_{1,1})} \ \Big| \ \Phi^*_{\langle e_L\rangle \smallsetminus \{e_L\}} \Big) \cdot H \, d\mathbb{P}_l$$

y por la Proposición 23 esta última integral es igual a

$$\int_{X^k} \mathbb{E}_l\big(\mathbf{1}_{A_{1,1}} \ \big| \ \Phi_{\langle e_L\rangle \smallsetminus \{e_L\}} \big) \circ \pi_1 \ \cdot \ H \, d\mathbb{P}_l$$

que, junto con todas las anteriores, se anula en cero.

Fijamos

$$A'_{1,1} := \big\{ \mathbb{E}_l\big(\mathbf{1}_{A_{1,1}} \ \big| \ \Phi_{\langle e_L\rangle \smallsetminus \{e_L\}} \big) > 0 \big\}$$

y

$$A'_{i,j} := A_{i,j} \qquad \forall \, (i,j) \neq (1,1) .$$

Dado que

$$\mathbb{E}_l\big(\mathbf{1}_{A_{1,1}} \ \big| \ \Phi_{\langle e_L\rangle \smallsetminus \{e_L\}} \big)$$

es no negativa tenemos

$$\begin{aligned} 0 = \int_{X^k} \mathbb{E}_l\big(\mathbf{1}_{A_{1,1}} \ \big| \ \Phi_{\langle e_L\rangle \smallsetminus \{e_L\}} \big) \circ \pi_1 \ \cdot \ H \, d\mathbb{P}_l \quad &= \quad \int_{X^k} \mathbf{1}_{A'_{1,1}} \circ \pi_1 \ \cdot \ H \, d\mathbb{P}_l \\ &= \quad \mathbb{P}_l \left(\prod_{i=1}^{k} \bigcap_{j=1}^{c_i} A'_{i,j} \right) \end{aligned}$$

con

$$A'_{i,j} \in \mathscr{I}'_{i,j} \qquad \forall \, (i,j)$$

y donde la colección de las $\mathscr{I}'_{i,j}$ satisface las condiciones de la hipótesis inductiva; entonces bien, tenemos por hipótesis que

$$\mathbb{P}_w \left(\bigcap_{i=1}^{k} \bigcap_{j=1}^{c_i} A'_{i,j} \right) = 0.$$

Ahora,

$$\begin{aligned}\mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} A_{i,j}\right) &\leq \mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} A'_{i,j}\right)\\ &+ \mathbb{P}_w\left(A_{1,1}\smallsetminus A'_{1,1}\cap\left(\bigcap_{j=2} A'_{1,j}\right)\cap\left(\bigcap_{i=2}^{k}\bigcap_{j=1}^{c_i} A'_{i,j}\right)\right)\\ &\leq 0+\mathbb{P}_w\big(A_{1,1}\smallsetminus A'_{1,1}\big)\,;\end{aligned}$$

pero

$$\begin{aligned}\mathbb{P}_w\big(A_{1,1}\cap A'_{1,1}\big) = \int_{A'_{1,1}} \mathbf{1}_{A_{1,1}}\, d\mathbb{P}_l &= \int_{A'_{1,1}} \mathbb{E}_l\big(\mathbf{1}_{A_{1,1}}\ \big|\ \Phi_{\langle e_L\rangle\smallsetminus\{e_L\}}\big)\ d\mathbb{P}_l\\ = \int_X \mathbb{E}_l\big(\mathbf{1}_{A_{1,1}}\ \big|\ \Phi_{\langle e_L\rangle\smallsetminus\{e_L\}}\big)\ d\mathbb{P}_l &= \int_X \mathbf{1}_{A'_{1,1}}\, d\mathbb{P}_l = \mathbb{P}_w(A_{1,1})\,,\end{aligned}$$

de donde

$$\mathbb{P}_w\big(A_{1,1}\smallsetminus A'_{1,1}\big)=0$$

y con esto hemos concluído el resultado. □

Para la demostración de DHJ bastaría con utilizar una versión reducida del Teorema 42 que considerase solamente familias ascendentes principales. Sin embargo, demostrar esta versión reducida no parece ser más sencillo; en la demostración anterior vemos cómo para extender el resultado sobre colecciones de familias ascendentes donde sólo haya familias principales de profundidad mínima, nos vemos obligados a considerar familias no principales de profundidad mayor - a saber, a la hora de considerar $\langle e_L\rangle \smallsetminus \{e_L\}$ -, para caer en la hipótesis inductiva. Siendo así, no podríamos completar el argumento inductivo que requerimos considerando solamente familias principales, y la siguiente Proposición es imprescindible para nuestro argumento:

Proposición 38. *Supongamos que todas las familias en la colección $\mathscr{C} := (\mathscr{I}_{i,j})_{i,j}$ tienen profundidad al menos M y que de estas hay $L \geq 0$ familias no principales de profundidad exactamente M. Si la propiedad $\mathcal{P}$ se cumple para toda colección semejante que tenga sólo $L-1$ familias no principales de profundidad exactamente M, entonces también se cumple para $\mathscr{C}$.*

Que pueda haber familias principales de profundidad exactamente M por el momento no nos interesa.

Antes de demostrar la proposición enunciamos una formulación del Teorema de Convergencia de Martingalas que nos permitirá acceder al argumento de [**Aus10**]:

Definición. Dado un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ decimos que una sucesión de σ-álgebras

$$(\mathscr{F}_n)_{n \geq 1}$$

tales que

$$\mathscr{F}_n \subset \mathscr{F}_{n+1} \subset \mathscr{F} \quad \forall n \geq 1$$

es una **filtración**.

Definición. A un proceso estocástico

$$(Y_n)_{n \geq 1}$$

definido en $(\Omega, \mathscr{F}, \mathbb{P})$ le llamamos **martingala** cuando satisface las siguientes condiciones:

(i) $\mathbb{E}(|Y_n|) < \infty$ para toda n
(ii) Y_n es $\mathscr{F}_n$-medible para toda n
(iii) $\mathbb{E}(Y_n \mid \mathscr{F}_m) = Y_m$ para cada $m \leq n$

Teorema 43 (de Convergencia de Martingalas). *Si $(Y_n)_n$ es una martingala tal que*

$$|Y_n| \leq Y \quad c.s. \quad \forall n \geq 1$$

para alguna variable aleatoria Y tal que

$$\mathbb{E}(Y) < \infty,$$

entonces

$$Y_\infty := \lim_{n \to \infty} Y_n$$

existe casi seguramente, pertenece a $\mathscr{L}_1(\Omega)$, y la convergencia también se da bajo la norma de $\mathscr{L}_1(\Omega)$. Además

$$Y_n = \mathbb{E}(Y_\infty \mid \mathscr{F}_n) \qquad \forall n \geq 1.$$

Omitimos incluír una demostración del Teorema, pero recomendamos consultar [**JP03**], página 231.

Demostración de la Proposición 38. Sean

$$\mathscr{I}_{i_1,j_1}, \mathscr{I}_{i_2,j_2}, \cdots, \mathscr{I}_{i_L,j_L}$$

las familias ascendentes no principales de profundidad M en $\mathscr{C}$ y, por comodidad notacional y sin pérdida de generalidad suponemos

$$\mathscr{I}_{i_L,j_L} = \mathscr{I}_{1,1}.$$

Sean, además,

$$e_1, e_2, \cdots, e_r$$

todos los conjuntos en $\mathscr{I}_{1,1}$ de tamaño M. Demostramos primero el teorema para un caso base y después argumentamos la generalización.

Caso base.

Consideramos una colección de conjuntos

$$(A_{i,j})_{i,j} \qquad A_{i,j} \in \Phi_{\mathscr{I}_{i,j}}$$

tal que podemos encontrar subálgebras *finitas*

$$\mathscr{B}_s \subset \Phi_{\langle e_s \rangle} \qquad s = 1, 2, \cdots, r$$

de manera que

$$A_{1,1} \in \mathscr{B}_1 \vee \mathscr{B}_2 \vee \cdots \vee \mathscr{B}_r \vee \Phi_{\mathscr{I}_{1,1} \cap \binom{[k]}{\geq M+1}}$$

donde

$$\binom{[k]}{\geq M+1}$$

es la familia compuesta de conjuntos de tamaño mayor o igual a $M+1$ y con

$$\Phi_{\mathscr{I}_{1,1} \cap \binom{[k]}{\geq M+1}}$$

denotamos a la σ-álgebra invariante bajo

$$\mathscr{I}_{1,1} \setminus \{e_1, e_2, \cdots, e_r\}.$$

Notamos que podemos generar a $A_{1,1}$ bajo operaciones conjuntistas de unión e intersección considerando sólo una cantidad finita de conjuntos en las $\Phi_{\langle e_s \rangle}$.

Es decir, si elegimos $K \geq \sup_{1 \leq s \leq r} |\mathscr{B}_s|$ podemos escribir

$$A_{1,1} = \bigcup_{m=1}^{K^r} (B_{m,1} \cap B_{m,2} \cap \cdots \cap B_{m,r} \cap C_m)$$

con

$$B_{m,s} \in \mathscr{B}_s \quad s = 1, 2, \cdots, r \qquad C_m \in \Phi_{\mathscr{I}_{1,1} \cap \binom{[k]}{\geq M+1}}.$$

En efecto, no existen más de K^r colecciones de conjuntos de la forma

$$\{B_1, B_2, \cdots, B_r\} \qquad B_s \in \mathscr{B}_s.$$

Ahora, sustituyendo esta descomposición de $A_{1,1}$ en la ecuación

$$\mathbb{P}_l \left(A_{1,1} \cap \left(\bigcap_{j=2}^{c_1} A_{1,j} \right) \times \bigcap_{j=1}^{c_2} A_{2,j} \times \cdots \times \bigcap_{j=1}^{c_k} A_{k,j} \right) = 0$$

podemos ver que cada conjunto

$$A^{(m)} := (B_{m,1} \cap B_{m,2} \cap \cdots \cap B_{m,r} \cap C_m) \cap \left(\bigcap_{j=2}^{c_1} A_{1,j} \right) \times \bigcap_{j=1}^{c_2} A_{2,j} \times \cdots \times \bigcap_{j=1}^{c_k} A_{k,j}$$

con

$$m = 1, 2, \cdots, r$$

es $\mathbb{P}_l$-nulo.

Consideramos la colección $\mathscr{C}'$ que obtenemos a partir de $\mathscr{C}$ al omitir $\mathscr{I}_{1,1}$ y en su lugar añadir

$$\{\langle e_1 \rangle, \langle e_2 \rangle, \cdots, \langle e_r \rangle\}.$$

La colección $\mathscr{C}'$ satisface los supuestos de la hipótesis inductiva ya que contiene $L-1$ familias no principales de profundidad M. Pero los conjuntos $\mathbb{P}_l$-nulos del párrafo anterior están asociados a $\mathscr{C}'$, puesto que

$$B_{m,s} \in \Phi_{\langle e_s \rangle} \qquad 1 \leq m \leq K^r \quad 1 \leq s \leq r$$

y

$$C_m \in \Phi_{\mathscr{I}_{1,1} \cap \left(\binom{[k]}{M+1}\right)} \in \mathscr{C}'.$$

Entonces, por hipótesis, $A^{(m)}$ es $\mathbb{P}_w$-nulo para cada m. Luego,

$$\mathbb{P}_w \left(\bigcap_{i=1}^{k} \bigcap_{j=1}^{c_i} A_{i,j} \right) \leq \sum_{m=1}^{r} \mathbb{P}_w \left(A^{(m)} \right) = 0.$$

Con esto queda establecida la propiedad $\mathcal{P}$ en el caso base.

Caso general.

No todas las elecciones de conjuntos

$$\left(A_{i,j}\right)_{i,j} \qquad A_{i,j} \in \mathscr{I}_{i,j}$$

satisfacen el requerimiento del caso base. Tal elección pudiera incluso no exisitr para $\mathscr{C}$. Afortunadamente, podemos generalizar el procedimiento anterior.

Recordemos que (X, Σ) es espacio estándar de Borel y, en particular, $\Phi_{\langle e \rangle}$ es numerablemente generado para cualquier $e \subset [k]$. Consideramos sólo aquellos subconjuntos e del alfabeto de cardinalidad M y, para cada uno de ellos, podemos encontrar una sucesión de subálgebras finitas de $\Phi_{\langle e \rangle}$,

$$\mathscr{B}_{e,1} \subset \mathscr{B}_{e,2} \subset \ldots,$$

de tal manera que

$$\bigvee_{n \geq 1} \mathscr{B}_{e,n} = \Phi_{\langle e \rangle}.$$

Con base en estas subálgebras que se aproximan a $\Phi_{\langle e\rangle}$ para cada $e \subset [k]$ de cardinalidad M, definimos ahora subσ-álgebras que se aproximan a $\Phi_{\mathscr{I}_{i,j}}$ para cada $\mathscr{I}_{i,j}$ en $\mathscr{C}$:

$$\Xi_{i,j}^{(n)} := \Phi_{\mathscr{I}_{i,j}\cap\binom{[k]}{\geq M+1}} \vee \bigvee_{e\in\mathscr{I}_{i,j}\cap\binom{[k]}{M}} \mathscr{B}_{e,n}.$$

Claramente

$$\Phi_{\mathscr{I}_{i,j}} = \bigvee_{n\geq 1} \Xi_{i,j}^{(n)}.$$

Observamos que para $\mathscr{I}_{i,j} \in \mathscr{C}$ de profundidad mayor a M la construcción anterior se reduce a

$$\Xi_{i,j}^{(n)} = \mathscr{I}_{i,j} \qquad \forall n \geq 1.$$

Ahora bien, consideramos los conjuntos

$$B_{i,j}^{(n)} := \left\{\mathbb{E}\left(\mathbf{1}_{A_{i,j}} \,\middle|\, \Xi_{i,j}^{(n)}\right) > 1-\delta\right\}$$

para

$$\delta \in (0,1) \qquad \text{fijo}$$

y definimos

$$F := \bigcap_{j=1}^{c_1} B_{1,j}^{(n)} \times \bigcap_{j=1}^{c_2} B_{2,j}^{(n)} \times \cdots \times \bigcap_{j=1}^{c_k} B_{k,j}^{(n)}.$$

Elegimos cualquier $\mathscr{I}_{i^*,j^*}$ en $\mathscr{C}$ y calculamos

$$\mathbb{P}_l\big(F \setminus \pi_{i^*}^{-1}(A_{i^*,j^*})\big) = \mathbb{P}_l(F) - \mathbb{P}_l\big(F \cap \pi_{i^*}^{-1}(A_{i^*,j^*})\big)$$

$$\begin{aligned}
&= \int_{X^k} \left(\prod_{(i,j)} \mathbf{1}_{B_{i,j}^{(n)}} \circ \pi_i\right) \\
&- \left(\mathbf{1}_{A_{i^*,j^*}\cap B_{i^*,j^*}^{(n)}} \circ \pi_{i^*}\right) \cdot \left(\prod_{(i,j)\neq(i^*,j^*)} \mathbf{1}_{B_{i,j}^{(n)}} \circ \pi_i\right) d\mathbb{P}_l \\
&= \int_{X^k} \left(\mathbf{1}_{B_{i^*,j^*}^{(n)} \setminus A_{i^*,j^*}} \circ \pi_{i^*}\right) \cdot \left(\prod_{(i,j)\neq(i^*,j^*)} \mathbf{1}_{B_{i,j}^{(n)}} \circ \pi_i\right) d\mathbb{P}_l \\
&= \int_{X^k} \mathbb{E}\left(\mathbf{1}_{B_{i^*,j^*}^{(n)} \setminus A_{i^*,j^*}} \,\middle|\, \Xi_{i^*,j^*}^{(n)}\right) \circ \pi_{i^*} \cdot \left(\prod_{(i,j)\neq(i^*,j^*)} \mathbf{1}_{B_{i,j}^{(n)}} \circ \pi_i\right) d\mathbb{P}_l.
\end{aligned}$$

La última igualdades válida porque

$$B^{(n)}_{i^*,j^*} \quad \text{es} \quad \Xi^{(n)}_{i^*,j^*}\text{-medible.}$$

Por otro lado tenemos

$$\begin{aligned}
\mathbb{E}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}\smallsetminus A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right) &= \mathbb{E}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}} - \mathbf{1}_{B^{(n)}_{i^*,j^*}\cap A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right)\\
&= \mathbf{1}_{B^{(n)}_{i^*,j^*}} - \mathbb{E}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}\cap A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right)\\
&= \mathbf{1}_{B^{(n)}_{i^*,j^*}} - \mathbb{E}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}}\cdot \mathbf{1}_{A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right)\\
&= \mathbf{1}_{B^{(n)}_{i^*,j^*}} - \mathbf{1}_{B^{(n)}_{i^*,j^*}}\cdot \mathbb{E}\left(\mathbf{1}_{A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right)\\
&= \mathbf{1}_{B^{(n)}_{i^*,j^*}}\cdot\left(1 - \mathbb{E}\left(\mathbf{1}_{A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right)\right)
\end{aligned}$$

y, recordando la definición de $B^{(n)}_{i^*,j^*}$,

$$\mathbb{E}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}\smallsetminus A_{i^*,j^*}} \,\middle|\, \Xi^{(n)}_{i^*,j^*}\right) \leq \mathbf{1}_{B^{(n)}_{i^*,j^*}}\cdot(1-(1-\delta)) = \delta\cdot\mathbf{1}_{B^{(n)}_{i^*,j^*}}. \tag{40}$$

Juntando lo anterior tenemos

$$\begin{aligned}
\mathbb{P}_l\big(F\setminus \pi_i^{-1}(A_{i,j})\big) &\leq \delta\int_{X^k}\left(\mathbf{1}_{B^{(n)}_{i^*,j^*}}\circ\pi_{i^*}\right)\cdot\left(\prod_{(i,j)\neq(i^*,j^*)}\mathbf{1}_{B^{(n)}_{i,j}}\circ\pi_i\right)d\mathbb{P}_l\\
&= \delta\cdot\mathbb{P}_l(F)\,.
\end{aligned}$$

Siendo así,

$$\begin{aligned}
\mathbb{P}_l(F) &\leq \mathbb{P}_l\left(\prod_{i=1}^{k}\bigcap_{j=1}^{c_i}A_{i,j}\right) + \sum_{(i,j)}\mathbb{P}_l\big(F\setminus\pi_i^{-1}(A_{i,j})\big)\\
&\leq 0 + \left(\sum_{i=1}^{k}c_i\right)\delta\cdot\mathbb{P}_l(F)\,,
\end{aligned}$$

de manera que si elegimos

$$\delta < \sum_{i=1}^{k}c_i$$

nos vemos obligados a concluír que

$$\mathbb{P}_l(F) = 0.$$

La importancia de la construcción anterior radica en la siguiente observación. Dijimos ya que

$$B_{i,j}^{(n)} \quad \text{es} \quad \Xi_{i,j}^{(n)}\text{-medible}$$

para cada par (i,j); en particular

$$B_{1,1}^{(n)} \quad \text{es} \quad \Xi_{1,1}^{(n)}\text{-medible}$$

y la colección de conjuntos

$$\mathscr{D}^{(n)} := \left(B_{i,j}^{(n)}\right)_{i,j}$$

satisface los supuestos del caso base. Aplicando el resultado anterior tenemos

$$\mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} B_{i,j}^{(n)}\right) = 0.$$

Luego,

$$\begin{aligned}
\mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} A_{i,j}\right) &\leq \mathbb{P}_w\left(\bigcap_{i=1}^{k}\bigcap_{j=1}^{c_i} B_{i,j}^{(n)}\right) + \sum_{i,j}\mathbb{P}_w\Big(A_{i,j} \setminus B_{i,j}^{(n)}\Big) \\
&= \sum_{i,j}\mathbb{P}_w\Big(A_{i,j} \setminus B_{i,j}^{(n)}\Big)
\end{aligned}$$

y sólo resta demostrar que

$$\sum_{i,j}\mathbb{P}_w\Big(B_{i,j}^{(n)} \setminus A_{i,j}\Big) \to 0, \quad n \to \infty.$$

Con este propósito, consideramos un par abitrario (i,j) con el proceso estocástico

$$(Y_n)_{n\geq 1} \qquad Y_n := \mathbb{E}\Big(\mathbf{1}_{A_{i,j}} \,\Big|\, \Xi_{i,j}^{(n)}\Big),$$

que expresamente es medible con respecto a la filtración

$$\Big(\Xi_{i,j}^{(n)}\Big)_{n\geq 1}$$

y además satisface, dados $m \leq n$,

$$\begin{aligned}
\mathbb{E}\Big(Y_n \,\Big|\, \Xi_{i,j}^{(m)}\Big) &= \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{A_{i,j}} \,\Big|\, \Xi_{i,j}^{(n)}\Big) \,\Big|\, \Xi_{i,j}^{(m)}\Big) \\
&= \mathbb{E}\Big(\mathbf{1}_{A_{i,j}} \,\Big|\, \Xi_{i,j}^{(m)}\Big) \\
&= Y_m
\end{aligned}$$

y

$$|Y_n| \leq 1 \quad \forall n \geq 1.$$

De manera que $(Y_n)_n$ es una martingala acotada, y por el *Teorema de Convergencia de Martingalas*

$$Y_\infty := \lim_{n\to\infty} Y_n \quad \text{c.s.}$$

existe, pertenece a $\mathscr{L}_1(X)$, Y_n converge a Y_∞ bajo la norma de $\mathscr{L}_1(X)$, y

$$Y_n = \mathbb{E}\Big(Y_\infty \ \Big|\ \Xi_{i,j}^{(n)} \Big) \qquad \forall n \geq 1$$

Pero Y_∞ es único salvo por conjuntos $\Xi_{i,j}^{(\infty)}$-nulos, con

$$\Xi_{i,j}^{(\infty)} := \bigvee_{n\geq 1} \Xi_{i,j}^{(n)} = \Phi_{\mathscr{I}_{i,j}}.$$

Ya que $A_{i,j} \in \Phi_{\mathscr{I}_{i,j}}$, concluímos que

$$Y_\infty = \mathbf{1}_{A_{i,j}} \qquad \mathbb{P}_w|_{\Phi_{\mathscr{I}_{i,j}}}\text{-c.s.}$$

Entonces

$$Y_n \to \mathbf{1}_{A_{i,j}} \quad \text{en} \quad L_1(\mathbb{P}_w)\,.$$

Por otro lado afirmamos que

$$\int \mathbf{1}_{A_{i,j}}\, d\mathbb{P}_w - \int Y_n\, d\mathbb{P}_w \geq \mathbb{P}_w\Big(A_{i,j} \smallsetminus B_{i,j}^{(n)} \Big)\,.$$

Efectivamente,

$$\begin{aligned}
\int Y_n\, d\mathbb{P}_w &= \int \mathbb{E}\Big(\mathbf{1}_{A_{i,j}} \ \Big|\ \Xi_{i,j}^{(n)} \Big)\, d\mathbb{P}_w \\
&\geq \int \mathbf{1}_{B_{i,j}^{(n)}} \cdot \mathbb{E}\Big(\mathbf{1}_{A_{i,j}} \ \Big|\ \Xi_{i,j}^{(n)} \Big)\, d\mathbb{P}_w \\
&= \int \mathbf{1}_{B_{i,j}^{(n)}} \cdot \mathbf{1}_{A_{i,j}}\, d\mathbb{P}_w \\
&= \mathbb{P}_w\Big(A_{i,j} \cap B_{i,j}^{(n)} \Big)\,,
\end{aligned}$$

de donde

$$\begin{aligned}
\int \mathbf{1}_{A_{i,j}}\, d\mathbb{P}_w - \int Y_n\, d\mathbb{P}_w &\geq \mathbb{P}_w(A_{i,j}) - \mathbb{P}_w\Big(A_{i,j} \cap B_{i,j}^{(n)} \Big) \\
&= \mathbb{P}_w\Big(A_{i,j} \smallsetminus B_{i,j}^{(n)} \Big)\,.
\end{aligned}$$

Siendo así,

$$\mathbb{P}_w\Big(B_{i,j}^{(n)} \smallsetminus A_{i,j} \Big) \to 0, \quad n \to \infty.$$

Pero el par (i, j) es arbitrario y entonces

$$\sum_{i,j} \mathbb{P}_w\Big(A_{i,j} \smallsetminus B_{i,j}^{(n)}\Big) \to 0, \quad n \to \infty. \qquad \square$$

Ponendo las piezas en su lugar podemos finalmente completar la demostración.

Demostración del Teorema 42. Hacemos un argumento recursivo ulitizando las proposiciones 37 y 38. Para el caso base, consideramos

$$c_i := 1 \quad \mathscr{I}_{i,1} := \{[k]\}, \qquad i = 1, 2, \cdots, k,$$

donde todas las familias ascendentes son la trivial generada por la totalidad del alfabeto. En este caso sabemos que las imágenes inversas

$$\pi_i^{-1}(A) \quad A \in \Phi_{\{[k]\}}$$

difieren solamente por conjuntos $\mathbb{P}_l$-nulos; de manera que

$$\begin{aligned} \mathbb{P}_l(A_{1,1} \times A_{2,1} \times \cdots \times A_{k,1}) &= \mathbb{P}_l\Big((A_{1,1} \cap A_{2,1} \cap \cdots \cap A_{k,1})^k\Big) \\ &= \mathbb{P}_w(A_{1,1} \cap A_{2,1} \cap \cdots \cap A_{k,1}) \end{aligned}$$

para

$$A_{i,1} \in \Phi_{\{[k]\}} \qquad i = 1, 2, \cdots, k.$$

Así que, dados

$$A_1, A_2, \cdots, A_k \in \Phi_{\{[k]\}}$$

tales que

$$\mathbb{P}_l(A_1 \times A_2 \times \cdots \times A_k) = 0,$$

tenemos

$$\mathbb{P}_w(A_1 \cap A_2 \cap \cdots \cap A_k) = 0.$$

Aplicando la Proposición 38, extendemos la propiedad $\mathcal{P}$ a toda colección que tenga una familia ascendente no principal de profundidad $k-1$ y donde las demás familias tengan profunidad k; recursivamente, a toda colección que tenga familias no principales de profundidad $k-1$. De manera análoga, aplicando la Proposición 37 extendemos ahora el resultado a toda colección que tenga sólo una familia ascendente de profundidad $k-1$, que ésta sea principal, y donde todas las demas tengan profundidad mayor; recursivamente, extendemos el resultado a toda colección donde haya sólo familias principales de profundidad $k-1$ y las demás tengan profunidad k. Así consideramos todas las colecciones posibles de profunidad mínima $k-1$.

Repitiendo el argumento para las profundidades

$$k-2, \cdots, 1$$

abarcamos todas las colecciones posibles de familias ascendentes. $\square$

Epílogo

Hemos estudiado varios casos del principio de correspondencia sin mencionarlo siempre explícitamente. En el capítulo 2 establecimos un princio de correspondencia entre las versiones finitaria e infinitaria de Van der Wärden, para traducir esta última en una formulación topológica y demostrarla mediante el Teorema de Recurrencia Múltiple de Birkhoff; a partir de un principio de correspondencia entre las versiones finitaria e infinitaria del Teorema de Szémeredi planteamos una demostración con base en el Teorema de Recurrencia Múltimple de Poincaré; mediante un principio de correspondencia planteamos la equivalencia entre las versiones finitaria e infinitaria del Teorema de Densidad de Hales-Jewett, y esto nos permitió afinar su traducción en lenguaje probabilístico y nos abrió la posibilidad de realizar una demostración ergódica de DHJ, misma que desarrollamos a lo largo del Capítulo 3. En el capítulo 3 utilizamos sin explicitarlo un principio de correspondencia para establecer la equivalencia entre estacionariedad fuerte y estacionariedad finitaria, misma que utilizamos para reducir la demostración al caso de distribuciones fuertemente estacionarias sobre $(X^\omega, \Sigma^\omega, \mathbb{P})$.

Hay aún otro principio de correspondencia del que echamos mano, sin siquiera evidenciarlo hasta ahora. En el capítulo 3 hicimos uso crítico del Teorema 30 para demostrar el Teorema 28 y reducir la formulación ergódica de DHJ al caso de distribuciones fuertemente estacionarias. No lo mencionamos, pero el Teorema 30 es una generalización del Teorema de Carslon:

Teorema 44 (Teorema de Carlson, versión finitaria; [**McC99**], p. 58). *Dados*

$$k, r, n, t \in \mathbb{N} \quad t > n$$

existe $M(k, r, n, t)$ suficientemente grande tal que para toda r-coloración de los n-subespacios en $[k]^M$, existe un t-subespacio tal que todos los n-subespacios contenidos en él tienen el mismo color.

La demostración tiene técnicas desarrolladas en [**FK89**] cuya explicación excede los objetivos de este trabajo. Lo que nos interesa subrayar es que, mediante un principio de correspondencia, se puede demostrar que el teorema anterior es equivalente al siguiente:

TEOREMA 45 (Teorema de Carlson, versión infinitaria; [**McC99**], p. 59). *Sea*

$$\psi \subset [k]^\omega$$

subespacio y $r \in \mathbb{N}$. Dada una r-coloración de las líneas combinatorias en ψ, existe un subespacio

$$\gamma \subset \psi$$

tal que todas sus líneas combinatorias tienen el mismo color.

Este último es el caso particular del Teorema 30 que considera solamente n-subespacios con $n = 1$. El argumento de generalización puede consultarse en la página 61 de [**McC99**].

¿Qué tienen en común todas estas equivalencias entre formulaciones finitarias e infinitarias? En todas ellas hay alguna propiedad que se preserva para alguna secuencia (o subsecuencia) de espacios o sistemas finitos de cardinalidad creciente que convergen en algún sentido preciso, según el objeto, a un espacio o sistema infinitario que conserva la propiedad. En esa convergencias usualmente la propiedad adquiere un sabor cualitativo, mientras que en los sistemas finitarios tenía que expresarse cuantitativamente - con enes, epsilons y deltas. Éstos son los ingredientes clave para un principio de correspondencia, y esperamos haber demostrado que esta perspectiva general abre posibilidades interesantes de desarrollo.

Como nota final, existen formulaciones precisas de lo que es es un principio de correspondencia utilizando análisis no estándar y ultrafiltros. No abordamos este estudio, pero para consultar una discusión al respecto recomendamos como lectura inicial el capítulo 4 del libro *Compactness and Contradiction*, de Terrence Tao.

Bibliografía

[Aus10] Tim Austin, *Deducing the multidimensional Szemerédi theorem from an infinitary removal lemma*, J. Anal. Math. **111** (2010), 131–150. MR 2747063 (2012d:37010)

[Aus11] ———, *Deducing the density Hales-Jewett theorem from an infinitary removal lemma*, J. Theoret. Probab. **24** (2011), no. 3, 615–633. MR 2822475 (2012h:60109)

[Bar95] Robert G. Bartle, *The elements of integration and Lebesgue measure*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1995, Containing a corrected reprint of the 1966 original [ıt The elements of integration, Wiley, New York; MR0200398 (34 #293)], A Wiley-Interscience Publication. MR 1312157 (95k:28001)

[Bil99] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749 (2000e:60008)

[FK89] H. Furstenberg and Y. Katznelson, *Idempotents in compact semigroups and Ramsey theory*, Israel J. Math. **68** (1989), no. 3, 257–270. MR 1039473 (92d:05170)

[Fur81] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures. MR 603625 (82j:28010)

[JP03] Jean Jacod and Philip Protter, *Probability essentials*, second ed., Universitext, Springer-Verlag, Berlin, 2003. MR 1956867 (2003m:60002)

[Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169 (2002m:60002)

[KW82] Yitzhak Katznelson and Benjamin Weiss, *A simple proof of some ergodic theorems*, Israel J. Math. **42** (1982), no. 4, 291–296. MR 682312 (84i:28020)

[McC99] Randall McCutcheon, *Elemental methods in ergodic Ramsey theory*, Lecture Notes in Mathematics, vol. 1722, Springer-Verlag, Berlin, 1999. MR 1738544 (2001c:05141)

[MSW05] Pietro Muliere, Piercesare Secchi, and Stephen Walker, *Partially exchangeable processes indexed by the vertices of a k-tree constructed via reinforcement*, Stochastic Process. Appl. **115** (2005), no. 4, 661–677. MR 2128635 (2006e:60053)

[Par05] K. R. Parthasarathy, *Probability measures on metric spaces*, AMS Chelsea Publishing, Providence, RI, 2005, Reprint of the 1967 original. MR 2169627 (2006d:60004)

[Pol12] D. H. J. Polymath, *A new proof of the density Hales-Jewett theorem*, Ann. of Math. (2) **175** (2012), no. 3, 1283–1327. MR 2912706

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

[Tao06] Terence Tao, *A quantitative ergodic theory proof of Szemerédi's theorem*, Electron. J. Combin. **13** (2006), no. 1, Research Paper 99, 49. MR 2274314 (2007i:37016)