



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE CIENCIAS

LA TEORÍA DE LÍMITES DE GRÁFICAS Y EL LEMA  
DE REGULARIDAD DE SZEMERÉDI

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A:

CÉSAR ZARCO ROMERO

DIRECTOR DE TESIS:

DR. GERÓNIMO URIBE BRAVO



CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO,  
2021



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

### Hoja de datos del jurado

1. Datos del alumno	1. Datos del alumno
Apellido paterno	Zarco
Apellido materno	Romero
Nombre	César
Teléfono	5541686208
Universidad	Universidad Nacional Autónoma de México
Facultad	Facultad de Ciencias
Carrera	Matemáticas
Número de cuenta	313296163
2. Datos del tutor	2. Datos del tutor
Grado	Doctor
Nombre	Gerónimo
Apellido paterno	Uribe
Apellido materno	Bravo
3. Datos del sinodal 1	3. Datos del sinodal 1
Grado	Doctora
Nombre	María
Apellido paterno	Clara
Apellido materno	Fittipaldi
4. Datos del sinodal 2	4. Datos del sinodal 2
Grado	Doctor
Nombre	Ricardo
Apellido paterno	Gómez
Apellido materno	Aiza
5. Datos del sinodal 3	5. Datos del sinodal 3
Grado	Doctor
Nombres	Sergio Iván
Apellido paterno	López
Apellido materno	Ortega
6. Datos del sinodal 4	6. Datos del sinodal 4
Grado	Doctora
Nombre	Laura Clementina
Apellido paterno	Eslava
Apellido materno	Fernández
7. Datos del trabajo escrito	7. Datos del trabajo escrito
Título	La teoría de límites de gráficas y el lema de regularidad de Szemerédi
Número de páginas	110 p.
Año	2021

Este —es *mi* camino, —¿dónde está el  
vuestro?, así respondía yo a quienes me  
preguntaban, ‘por el camino’. ¡*El* camino,  
en efecto, —no existe!

---

F. Nietzsche, *Así habló Zaratustra*



## Agradecimientos

Me encuentro dando el último paso en una etapa que me definió como ser humano. Es increíblemente inmenso el apoyo que he recibido siempre, en todo momento, por seres extraordinarios, y que sin ello muy probablemente no habría trabajado que presentar. Me será difícil mencionar a cada persona que en algo aportó a la creación de este trabajo. Aquí es donde quiero disculparme con quienes su nombre no está escrito en este apartado, pero sepan que les agradezco demasiado.

A mis padres, Graciela y César, por su infinito amor y paciencia; por darme el mejor regalo que un hijo puede pedir: la libertad de ser uno mismo. A mis hermanos, Mariana y José, entre muchas cosas, por su comprensión durante la redacción de este trabajo. A mis abuelos, Josefina y Agustín. A mis padrinos, Susana y Carmelo. A toda mi familia, por ser mi soporte principal, por todo su apoyo; porque, como alguna vez alguien dijo, un individuo, solo, no puede trascender.

A Gerónimo Uribe Bravo. Es imposible agradecer en palabras todo el apoyo brindado. Gracias por su confianza al aceptar dirigir mi tesis de licenciatura. Infinitas gracias por sus enseñanzas durante la escritura de esta tesis, por escucharme cada semana, por brindarme las herramientas necesarias para escribir este trabajo y darme una llave de acceso a una puerta del conocimiento.

Al Instituto de Matemáticas, por abrirme sus puertas como becario. Dentro de sus instalaciones pasé muchas horas escribiendo, borrando, editando y escribiendo una y otra vez los párrafos de mi tesis.

A mis sinodales, María Clara, Ricardo Gómez, Sergio López y Laura Eslava, por su tiempo dedicado a la lectura de este trabajo, y por sus correcciones que mejoraron en buena parte la presentación de este escrito.

A Daniel Labardini Fragoso, por las grandes enseñanzas que me forjaron como matemático, por su interés en una buena redacción y en siempre hacer las cosas bien.

A los grandes profesores con quienes cursé en la Facultad de Ciencias, por orden de aparición, a Isabel Hubard, Carmen Arrillaga, Daniel Labardini, Gerónimo Uribe, Mónica Clapp, Pierre M. Bayard, Lizbeth Naranjo, María Clara, Sergio López, en fin, entre muchos muchos más, que se esfuerzan por dar en cada clase lo mejor de sí y mantienen un excelente nivel competitivo, pero sobre todo, por difundir el amor a las matemáticas.

A Carlos Nathanael Chávez y a Imanol Nuñez, por dar un excelente curso de Teoría de la medida y por su introducción formal a la teoría de la probabilidad. Sin su curso, es probable que este trabajo no hubiera alcanzado el suficiente rigor matemático. Algunas de las proposiciones expuestas en su curso fueron tomadas y enunciadas en los apéndices A y B de este trabajo. Varias de sus proposiciones son los pilares que sostienen muchos teoremas en esta tesis.

A Sandra Palau Calderón, por las pequeñas pláticas que tuvimos. Gracias por tenerme en cuenta.

A Alejandro Santoyo Cano, por escucharme cuando todavía divagaba sobre el lema de Szemerédi.

---

A mis amigos desde primer semestre, Carlos, Cristian y Linda, con quienes compartí largas horas de estudio, otras de música y cine, y de quienes aprendí constantemente.

A la UNAM. Todo lo que soy se lo debo a mi universidad.

Agradezco al Consejo Nacional de Ciencia y Tecnología por la beca otorgada, en la extensión del proyecto 2016-01-1946 titulado “*Intercambiabilidad, divisibilidad infinita, cambios de tiempo y estructuras combinatorias de gran tamaño*”. Sin duda, este apoyo económico permitió dedicarme a fondo a este trabajo.

# Índice general

Agradecimientos	V
Introducción: El lema de Szemerédi	IX
Capítulo 1. El lema de regularidad de Szemerédi	1
1. Semianillos: Construcciones y propiedades	1
2. Particiones y refinamientos	11
3. Funciones de crecimiento	13
4. Prueba del lema fuerte de regularidad vía martingalas	14
5. Prueba del lema de regularidad en su versión de la teoría de gráficas	26
Capítulo 2. Elementos de la teoría de límites de gráficas	33
1. Conceptos básicos	33
2. Densidad de homomorfismos	36
3. Algunas relaciones entre grafones y gráficas	37
4. Métrica de corte y otras distancias en el espacio de gráficas	38
5. Norma de corte y distancia de corte para núcleos	42
6. El lema de regularidad como un teorema de compacidad en el espacio de grafones	49
Capítulo 3. Aplicaciones del lema fuerte de regularidad	53
1. Eventos uniformes	53
2. Regularidad en Probabilidad	55
3. Lema fuerte de regularidad para grafones	58
Capítulo 4. Gráficas con sucesión de grados dada	59
1. Introducción	59
2. Límites de escalamiento	60
3. Simulación mediante el Modelo de configuración	64
4. Estadísticas con sucesiones de grados.	65
5. Convergencia al estimador máximo verosímil del Modelo $\beta$	67
Apéndice A. Nociones de la teoría de la medida	75
1. Espacios de probabilidad discretos	78
2. Funciones medibles	78
3. Propiedades de la integral	80
4. Espacios $L^p$ y desigualdades	82
Apéndice B. Martingalas	83
1. Esperanza condicional	83
2. Martingalas	85
3. Sucesión de diferencias de martingala	85
Apéndice C. Códigos	87
Símbolos	89
Índice de figuras	91
Bibliografía	93



## Introducción: El lema de Szemerédi

En 1936 Erdős y Turán [ET36] conjeturaron que cualquier subconjunto  $A$  de enteros positivos que tenga densidad positiva, esto es,

$$\limsup_{N \rightarrow \infty} \frac{|A \cap \{1, \dots, N\}|}{N} > 0,$$

debe contener una *progresión aritmética*<sup>1</sup> de tamaño  $k$  para cualquier número natural  $k$  mayor o igual a tres.<sup>2</sup>

Los primeros pasos hacia la aseveración de esta conjetura se dan en 1953. En este año Roth [Rot53] muestra mediante análisis de Fourier que el enunciado es cierto para una progresión aritmética de tamaño tres; no obstante, resulta complicado extender esta técnica para asegurar la existencia de progresiones aritméticas de cualquier tamaño finito. Es hasta 1975 que se publica el artículo donde Endre Szemerédi demostraría la conjetura de manera general en [Sze75b].

Actualmente la conjetura de Erdős y Turán es comúnmente conocida como el *teorema de regularidad de Szemerédi*. La demostración de este teorema es un argumento combinatorio sofisticado,<sup>3</sup> y un componente importante de la prueba es un resultado en teoría de gráficas nombrado por primera vez, en palabras de Szemerédi, como “un lema sobre gráficas bipartitas”. Posteriormente el lema sería conocido como *lema de regularidad de Szemerédi*.

Otra demostración del teorema de regularidad de Szemerédi se publica en 1977 por Furstenberg [Fur77], ahora desde la teoría ergódica. Komlós y Simonóvits comentan que las subsecuentes generalizaciones de la prueba de Furstenberg, por el mismo autor y Katznelson [FK89], así como por Bergelson y Leibman [BL96], muestran que la prueba de Furstenberg es distinta a la de Szemerédi, y con ello se confirma que la teoría ergódica es una “herramienta natural para atacar preguntas combinatorias” [KS96, p. 7].

En abril de 1975 Szemerédi dedica un artículo completo [Sze75a] a su lema de regularidad, mejorando el enunciado y ajustando la prueba para exhibir un resultado más limpio y más fuerte que atañe a cualquier gráfica, no solo bipartita, con la condición de tener una cantidad de vértices muy grande. Para entender el lema de regularidad de Szemerédi en su versión cercana a la original requerimos del lenguaje de la teoría de gráficas.

En primer lugar necesitamos hablar del concepto de densidad de aristas en una gráfica, que puede pensarse como la proporción de aristas entre dos subconjuntos, ajenos, del conjunto de vértices. Más precisamente, supongamos que tenemos una gráfica  $G = (V, E)$  y dos subconjuntos ajenos  $S$  y  $T$  de  $V$ . Definimos  $e_G(S, T)$  como la cantidad de aristas que tienen un vértice en  $S$  y otro vértice en  $T$ . La densidad de aristas entre  $S$  y  $T$ , cuando ambos conjuntos son no vacíos, se define como

$$d_G(S, T) = \frac{e_G(S, T)}{|S||T|}.$$

---

<sup>1</sup>Una progresión aritmética es una sucesión de enteros con diferencia constante. Por ejemplo, una sucesión de la forma 3, 7, 11, 15, 19 es una progresión aritmética de tamaño 5 y diferencia 4.

<sup>2</sup>La condición límite en el tamaño de  $A$  significa que nuestro subconjunto contiene una fracción positiva de los números naturales.

<sup>3</sup>Como lo comentan Terence Tao y Van Vu en [TV06, p. 369]

Como segundo punto se encuentra la noción de regularidad. En términos simples, una gráfica bipartita  $G$  es regular si sucede que para cualquier subconjunto  $A_1$  de  $S$  y cualquier subconjunto  $A_2$  de  $T$ , bajo ciertas condiciones, la densidad de aristas entre  $A_1$  y  $A_2$  es cercana a la densidad de aristas entre  $S$  y  $T$ . Formalmente se enuncia esta idea en la siguiente definición.

**DEFINICIÓN 0.1** (Condición de regularidad). Sean  $G$  una gráfica bipartita con bipartición  $\{V_1, V_2\}$  y  $\epsilon > 0$ . Decimos que  $G$  es  $\epsilon$ -regular si para cualesquiera  $A_1 \subseteq V_1$  y  $A_2 \subseteq V_2$  con la propiedad  $|A_1| > \epsilon|V_1|$  y  $|A_2| > \epsilon|V_2|$  se satisface

$$|d_G(A_1, A_2) - d_G(V_1, V_2)| \leq \epsilon.$$

Ahora bien, cuando tenemos una gráfica  $G = (V, E)$  y cualesquiera dos subconjuntos ajenos del conjunto de vértices,  $S$  y  $T$ , podemos inducir una gráfica bipartita,  $G[S, T]$ , definida como la gráfica con conjunto de vértices  $S \cup T$  y cuyo conjunto de aristas se define como las aristas de  $G$  que tienen un vértice en  $S$  y otro vértice en  $T$ .

Una vez establecidas estas definiciones podemos enunciar el lema de regularidad de Szemerédi.

**TEOREMA 0.2** (Lema de regularidad de Szemerédi, versión casi original). Para cada  $\epsilon > 0$  existe un número natural  $N(\epsilon)$  con la propiedad de que si  $G = (V, E)$  es una gráfica de orden  $n \in \mathbb{Z}_{\geq N(\epsilon)}$ , entonces existe una partición  $\mathcal{P} = \{V_0, V_1, \dots, V_p\}$  de  $V$ , con  $1/\epsilon \leq p \leq N(\epsilon)$ , que satisface simultáneamente las siguientes propiedades:

- (i) (Conjunto excepcional)  $|V_0| < \epsilon n$ .
- (ii) (Partición uniforme)  $|V_1| = \dots = |V_p|$ .
- (iii) (Regularidad) Cualesquiera  $i, j \in [p] = \{1, \dots, p\}$  con  $i \neq j$ , salvo una cantidad  $\epsilon p^2$  de pares  $(i, j)$ , satisfacen que la gráfica bipartita  $G[V_i, V_j]$  es  $\epsilon$ -regular.

De manera que el lema de regularidad describe la estructura subyacente de una gráfica de un orden muy grande. Uno de los objetivos de la tesis será probar este teorema. Varias estructuras y fenómenos interesantes en el mundo real pueden ser modelados con gráficas de gran tamaño, por ejemplo el internet. Sin embargo, en los problemas de la teoría de gráficas tradicional una gráfica exacta nos es dada y esto es una desventaja, pues en ocasiones gráficas de gran tamaño nunca son conocidas completamente. Una manera de resolver este problema es describir una gráfica de orden grande con un objeto matemático más sencillo, y el lema de regularidad Szemerédi es una herramienta adecuada.

Una de las cosas sorprendentes en el lema de Szemerédi, descubierta en 1991 por Simonóvits y Sós, es que la condición de regularidad de una gráfica coincide con la noción de un comportamiento aleatorio.<sup>4</sup> Consideremos, por ejemplo, un número real  $p$  en el intervalo  $(0, 1)$  y un entero positivo  $n$ . En el modelo de gráficas aleatorias de Erdős-Rényi [ER60],  $\mathcal{G}(n, p)$  es el conjunto de gráficas con conjunto de vértices igual a  $[n]$  y tales que sus aristas son escogidas independientemente al azar y con probabilidad  $p$ . Denotemos por  $G_{n,p}$  a un elemento de  $\mathcal{G}(n, p)$ . En 1989 Chung, Graham y Wilson [CGW89] enuncian una lista de distintas propiedades, todas ellas equivalentes, que cumple una cierta sucesión de gráficas  $(G_n)$ . A tal sucesión  $(G_n)$  que cumple una propiedad de esta lista (y por lo tanto todas) la nombran cuasi-aleatoria, y puede pensarse como una sucesión de gráficas deterministas que tienen un comportamiento similar a las gráficas aleatorias. Esta última afirmación se debe a que las sucesiones de gráficas aleatorias —entre ellas, las sucesiones de la forma  $(G_{n,p})$  para un  $p \in (0, 1)$  fijo— comparten estas propiedades con las sucesiones de gráficas cuasi-aleatorias. Dos años más tarde, Simonóvits y Sós [SS91] descubren la conexión entre una sucesión cuasi-aleatoria y el lema de regularidad de Szemerédi.

**TEOREMA 0.3.** Una sucesión de gráficas  $(G_n)$  es cuasi-aleatoria si y sólo si para todo entero positivo  $\ell$  y todo número real  $\epsilon > 0$  existen enteros positivos  $k(\epsilon, \ell)$  y  $n_0(\epsilon, \ell)$  tales que, para todo  $n > n_0(\epsilon, \ell)$ ,  $V(G_n)$  tiene una partición en  $k$  clases  $U_0, \dots, U_k$  ( $\ell < k < k(\epsilon, \ell)$ ) donde, para

<sup>4</sup>En términos exactos, la literatura se refiere a este comportamiento de gráficas como *cuasi-aleatorio*. En 1991 Szimonóvits y Sós descubrieron la conexión entre el lema de Szemerédi y sucesiones de gráficas cuasi-aleatorias en [SS91]. Véase también [Bol01, p. 380].

cada  $i, j \in [k]$  con  $i \neq j$ , salvo a lo más  $\epsilon k^2$ ,  $G[U_i, U_j]$  es  $\epsilon$ -regular y además con densidades que satisfacen

$$|d(U_i, U_j) - p| < \epsilon,$$

para algún  $p \in (0, 1)$ .

Como consecuencia, para un  $p \in (0, 1)$  fijo, una sucesión  $(G_{n,p})$  es regular en el sentido del Teorema 0.3. Chung [Gra91, p. 242] resume el descubrimiento de Simonóvits y Sós [SS91] en una frase:

Szemerédi's regularity lemma asserts that if the vertex set of a graph is partitioned in the right way, then the edges joining the sets behave in some random-like manner.

Terence Tao, por su parte, muestra en 2006 una versión puramente probabilista del lema de regularidad [Tao06]. En esta nueva perspectiva uno ve el lema de Szemerédi no como un teorema de estructura para gráficas grandes, sino como un teorema de estructura para eventos o variables aleatorias en un espacio de probabilidad. Con esta nueva visión, evidentemente, es posible utilizar herramientas de la teoría de la probabilidad como  $\sigma$ -álgebras y esperanza condicional. Mediante la probabilidad Tao consigue una versión más fuerte y flexible del lema de regularidad que encuentra fácilmente aplicaciones a hipergráficas. La prueba que analizaremos contiene esta generalización.

A pesar de que Tao menciona que su versión probabilista del lema de regularidad se extiende sin complicaciones a cualquier cantidad finita de variables aleatorias, deja de lado esta generalización y trabaja con una sola variable aleatoria para enfocarse en los argumentos combinatorios.

Dos años después, en 2008, Bollobás y Nikiforov abstraen la versión probabilista de Tao [BN08]. En su artículo trabajan constantemente con un “semianillo de conjuntos”, el cual típicamente es descrito, informalmente, como una familia de conjuntos estructurados. Introducen además el concepto de sistema de regularidad de Szemerédi (sistema-RS) definido como una cuarteta  $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{S})$ , donde  $\Omega$  es un conjunto no vacío;  $\mathcal{A}$  es un álgebra de subconjuntos de  $\Omega$ ;  $\mathbb{P}$  es una medida en  $\mathcal{A}$  que es completa, no negativa, finitamente aditiva y cumple  $\mathbb{P}(\Omega) = 1$ ; y, para algún número natural  $k$ ,  $\mathcal{S}$  es un  $k$ -semianillo<sup>5</sup> en  $\Omega$  que satisface  $\mathcal{S} \subseteq \mathcal{A}$ .

La abstracción de Bollobás y Nikiforov exhibe la esencia del lema de regularidad de Szemerédi en su versión original. Para convencernos de ello, veamos brevemente esta generalización. En adelante suponemos que  $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{S})$  es un sistema-RS. El análogo en probabilidad al concepto de densidad de aristas es el siguiente. Para cada  $A, V \in \mathcal{A}$  definimos

$$d(A, V) = \frac{\mathbb{P}(A \cap V)}{\mathbb{P}(V)}$$

si  $\mathbb{P}(V) > 0$ , y  $d(A, V) = 0$  si  $\mathbb{P}(V) = 0$ .

Escrito de esta manera es inmediato reconocer que esta definición de densidad entre  $A$  y  $V$  coincide con la probabilidad condicional del evento  $A$  dado  $V$ . Con esto en mente, se vislumbra la relación con esperanza condicional.<sup>6</sup> Ahora veamos la generalización de regularidad.

**DEFINICIÓN 0.4** (Condición general de regularidad). *Sea  $0 < \epsilon < 1$  y  $V \in \mathcal{S}$  tal que  $\mathbb{P}(V) > 0$ . Un conjunto  $A \in \mathcal{A}$  es llamado  $\epsilon$ -regular en  $V$  si*

$$|d(A, U) - d(A, V)| < \epsilon$$

para cada  $U \in \mathcal{S}$  tal que  $U \subseteq V$  y  $\mathbb{P}(U) > \epsilon \mathbb{P}(V)$ .

<sup>5</sup>Un  $k$ -semianillo,  $\mathcal{S}$ , en  $\Omega$  es una familia de subconjuntos de  $\Omega$  que satisface tres propiedades: (i)  $\Omega, \emptyset \in \mathcal{S}$ ; (ii) para cada  $S, T \in \mathcal{S}$  se cumple  $S \cap T \in \mathcal{S}$  y también (iii) para cada  $S, T \in \mathcal{S}$  existe  $\ell \in [k] = \{1, \dots, k\}$  tal que  $S \setminus T$  es igual a la unión ajena de  $\ell$  elementos no vacíos que pertenecen a  $\mathcal{S}$ .

<sup>6</sup>Tao [Tao06, p. 19] llama *energía* de una  $\sigma$ -álgebra al concepto que relaciona esperanza condicional y densidad de aristas, como se detalla en el Capítulo 1 de esta tesis.

Finalmente, como paso previo a enunciar la abstracción probabilista, requerimos de la siguiente definición de regularidad en una partición.

**DEFINICIÓN 0.5.** Sean  $0 < \epsilon < 1$  y  $\mathcal{P}$  una partición de  $\Omega$  con la propiedad  $\mathcal{P} \subseteq \mathcal{S}$ . Fijemos  $A \in \mathcal{A}$  y denotemos  $N_A = \{V \in \mathcal{P} \mid A \text{ no es } \epsilon\text{-regular en } V\}$ . Decimos que  $A$  es  $\epsilon$ -regular en  $\mathcal{P}$  si

$$\sum_{V \in N_A} \mathbb{P}(V) < \epsilon.$$

Así, estamos listos para enunciar la versión abstracta del lema de regularidad de Szemerédi en Probabilidad [BN08].

**TEOREMA 0.6** (Lema de regularidad, versión abstracta). Sea  $0 < \epsilon < 1$ . Sea  $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{S})$  un sistema-RS. Para cada  $\mathcal{C}$  subconjunto finito de  $\mathcal{A}$  y cada partición finita  $\mathcal{P}$  de  $\Omega$  tal que  $\mathcal{P} \subseteq \mathcal{S}$ , existe una partición finita  $\mathcal{Q}$  de  $\Omega$  tal que  $\mathcal{Q} \subseteq \mathcal{S}$  y satisface

- (i)  $\mathcal{Q}$  refina a  $\mathcal{P}$ ; y
- (ii) cada  $A \in \mathcal{C}$  es  $\epsilon$ -regular en  $\mathcal{Q}$ .

La estructura de los semianillos es utilizada para arribar a las aplicaciones a partir del teorema anterior. Como consecuencia tenemos el lema de regularidad de Szemerédi para una familia finita de gráficas.

**TEOREMA 0.7.** Sean  $0 < \epsilon < 1$  y  $\ell$  un entero positivo. Existen enteros positivos  $N(\epsilon)$  y  $q(\ell, \epsilon)$  tales que si  $n > N(\epsilon)$  y  $\mathcal{C}$  es una familia de  $\ell$  gráficas (no dirigidas) con conjunto de vértices igual a  $[n]$ , entonces existe una partición  $\mathcal{Q} = \{Q_0, \dots, Q_q\}$  de  $[n]$ , donde  $q \leq q(\ell, \epsilon)$ , con las siguientes propiedades:

- (i)  $|Q_0| < \epsilon n$ ;
- (ii)  $|Q_1| = \dots = |Q_q| < \epsilon n$ ;
- (iii) para cada gráfica  $G \in \mathcal{C}$  existen al menos  $(1-\epsilon)\binom{q}{2}$  conjuntos  $\{i_1, i_2\}$  de elementos distintos de  $[q]$  tales que  $G[Q_{j_1}, Q_{j_2}]$  es  $\epsilon$ -regular para cada permutación  $(j_1, j_2)$  de  $\{i_1, i_2\}$ .

Otra muestra de la ventaja conceptual ganada usando probabilidad se vuelve evidente menos de una década más tarde. En 2016 Dodos, Kanellopoulos y Karageorgos [DKK16] utilizan el algoritmo de aproximación a una variable aleatoria propuesto por Tao [Tao06] para construir una martingala y con ello una sucesión de diferencias de martingala. La clave en su demostración es una desigualdad reciente debida a Ricard y Xu [RX<sup>+</sup>16] en este tipo de sucesiones. Con dicha desigualdad muestran que es posible regularizar mediante una sola partición una familia finita de variables aleatorias. De igual forma, los semianillos juegan un papel importante en su prueba, pues a través ellos generalizan la llamada *norma de corte*. La idea intrínseca en [DKK16] es que podemos ver el lema de regularidad en probabilidad como un resultado de aproximación a una familia finita de variables aleatorias mediante un solo objeto en una norma “uniforme”. La conexión entre teoría de gráficas y esta versión del lema de regularidad es clara gracias a los semianillos expuestos por Bollobás y Nikiforov [BN08].

El resultado principal en [DKK16] es cercano al de Tao [Tao06]. Lo nuevo en [DKK16] es que el resultado es para variables aleatorias en el espacio  $L^p$  con  $1 < p \leq 2$ , mientras que Tao usa propiedades de  $\sigma$ -álgebras y de operadores ortogonales para lograr un resultado en el espacio  $L^2$ . Aunque la idea de aproximarse a una variable aleatoria es esencialmente la misma en ambos artículos, una diferencia notable es la manera en la que ambos llegan a mostrar la existencia de la cota en la cardinalidad de la partición. En tanto que Tao llega a ella mediante propiedades de ortogonalidad en el espacio  $L^2$ , Dodos, Kanellopoulos y Karageorgos aseguran la cota en la cardinalidad debido a la desigualdad de Ricard y Xu [RX<sup>+</sup>16], además esta desigualdad es vital para extender el resultado al caso  $L^p$  con  $1 < p \leq 2$ .

En la versión original del lema de regularidad de Szemerédi, un punto clave es que la cardinalidad de la partición obtenida —llamémosla de nuevo  $\mathcal{P}$ — no depende del orden de la gráfica.

Desafortunadamente, el número de elementos que conforman la partición  $\mathcal{P}$  es demasiado grande, incluso la mejor prueba da una torre exponencial de números dos cuyo tamaño es cercano a  $1/\epsilon^2$ , dependiendo del número real positivo  $\epsilon$  dado; esto es, de la forma

$$2^{2^{\cdot^{\cdot^2}}}$$

Gowers muestra en 1997 que esta es una característica del lema de regularidad que no puede mejorar significativamente [Gow97].

La búsqueda por conseguir una partición del lema de regularidad con un menor número de elementos se debe a que esto facilita las aplicaciones, en particular, en ciencias de la computación. Con este objetivo, Frieze y Kannan desarrollan en 1999 una primera versión débil del lema de regularidad de Szemerédi, donde la partición obtenida en el lema reduce considerablemente su cardinalidad en comparación con su versión original [FK99]. Por supuesto, dicha partición no satisface completamente la condición de regularidad; sin embargo, es suficiente para aplicaciones en algoritmos de la teoría de gráficas, entre ellos, los que conciernen al *problema de corte máximo*.

Frieze y Kannan definen, también en [FK99], una nueva norma en el espacio de matrices llamada norma de corte, inspirada en el problema de corte máximo. Es sabido que a través de la matriz de adyacencia de una gráfica podemos inducir una métrica en el espacio de gráficas, y la norma de corte revela una nueva noción de cercanía entre dos gráficas.

Para entender la importancia de la norma de corte como una norma apropiada para gráficas de gran tamaño pensemos en lo siguiente. Consideremos una sucesión de gráficas  $(G_n)$  donde la cantidad de vértices de  $G_n$  crece conforme  $n$  tiende a infinito. Las preguntas naturales son: ¿la sucesión  $(G_n)$  converge? y, si converge, ¿a qué objeto?. En 2006 Lovász y Szegedy [LS06] describen la noción de convergencia para este tipo de sucesiones, a través de homomorfismos de gráficas, y descubren que la representación del objeto límite buscado es una función  $W : [0, 1]^2 \rightarrow [0, 1]$  que es  $(\mathcal{B}([0, 1]^2), \mathcal{B}([0, 1]))$ -medible y simétrica.<sup>7</sup> Tal función es definida como *grafón* y puede ser pensada como una generalización de una gráfica.<sup>8</sup> En un trabajo conjunto, los matemáticos Borgs, Chayes, Lovász, Sós y Vesztergombi muestran en 2008 [BCL<sup>+</sup>08] que la noción de convergencia establecida por Lovász y Szegedy [LS06] es equivalente a la convergencia con una métrica llamada *distancia de corte*. La distancia de corte es un ajuste a la métrica inducida por la norma de corte, y esta modificación se debe a un resultado de compacidad que retomaremos en breve dentro de esta introducción.

Por el momento, volvamos de nuevo al modelo de Erdős-Renyi [ER60] para mostrar un ejemplo de un grafón como objeto límite. Si  $p$  es un número fijo en el intervalo  $(0, 1)$  y consideramos la sucesión  $(G_{n,p})$ , entonces, cuando  $n$  tiende a infinito, la gráfica límite es representada por el grafón  $W : [0, 1]^2 \rightarrow [0, 1]$  con regla de correspondencia  $W(x, y) = p$  para todo  $x, y \in [0, 1]$ . Recíprocamente, dada una función  $W : [0, 1]^2 \rightarrow [0, 1]$  que es idénticamente igual a  $p \in (0, 1)$ , podemos construir una gráfica aleatoria  $G(n, p)$ , para algún entero positivo  $n$ , con conjunto de vértices igual a  $[n]$  y con aristas elegidas con probabilidad  $p$ . En este trabajo analizamos grafones, no necesariamente constantes, obtenidos como límite de una sucesión de gráficas aleatorias dada una sucesión de grados [CDS11].

Como consecuencia de los trabajos hechos en 1999 y 2006, Lovász y Szegedy exhiben de manera concreta, en 2007, una nueva perspectiva del lema de regularidad de Szemerédi, ahora desde el área de análisis [LS07]. Dentro de esta nueva visión podemos ver el lema débil de regularidad como un resultado para grafones, es decir, en su versión “continua”, mientras que la versión discreta corresponde al enunciado para gráficas. Lovász y Szegedy [LS07] retoman la generalización de Frieze y Kannan [FK99] del lema débil de regularidad para funciones Lebesgue-medibles y modifican ligeramente el enunciado para establecer un lema débil de regularidad en el espacio de grafones. El lema débil para grafones se resume en un enunciado de aproximación, en la norma de corte, a un

<sup>7</sup>Una función  $W : [0, 1]^2 \rightarrow [0, 1]$  es simétrica si, para cualesquiera  $x, y \in [0, 1]$ , se satisface  $W(x, y) = W(y, x)$ .

<sup>8</sup>La palabra grafón es el resultado de la contracción de los términos gráfica y función. Para una introducción intuitiva de este objeto límite véase [Lov12, p. 16].

grafón mediante funciones simples.

A su vez, Lovász y Szegedy [LS07] demuestran que la versión débil del lema de regularidad, concerniente a grafones, implica que el espacio de grafones dotado de la distancia de corte —bajo la respectiva generalización— es un espacio compacto (resultado anunciado en párrafos anteriores). Lovász comenta [Lov12, p. 149] que el teorema de compacidad en el espacio de grafones puede ser considerado como la forma más fuerte de regularidad.

Por otro lado, existe también en la teoría de gráficas una versión fuerte del lema de regularidad de Szemerédi, publicada por primera vez en el año 2000 por Alon, Fisher, Krivelevich y Szegedy [AFKS00]. Esta nueva versión fuerte describe aún más la estructura de los elementos de la partición obtenida, por supuesto, con el costo de incrementar la cardinalidad de la partición en una torre exponencial aún más formidable. El lema fuerte de regularidad da una partición regular  $\mathcal{P}$  y un refinamiento  $\mathcal{Q}$  “más regular” y cercano a  $\mathcal{P}$  [CF12, p. 5]. De igual manera Lovász y Szegedy [LS07] enuncian y demuestran el análogo del lema fuerte en su versión “continua” que corresponde a grafones.

Sorprendentemente, el teorema de compacidad en el espacio de grafones implica el lema fuerte de regularidad de Lovász y Szegedy [LS07]. Como es de esperarse, el lema fuerte de regularidad implica su versión débil, también de Lovász y Szegedy [LS07]. Tenemos, por tanto, la equivalencia entre los lemas fuerte y débil de regularidad.

El resultado principal de Dodos, Kanellopoulos y Karageorgos [DKK16] y el de Tao [Tao06] dan una perspectiva cercana al lema fuerte de regularidad desde el área de la probabilidad. Uno de los resultados principales demostrados en esta tesis (Teorema 1.40 a partir del resultado principal en [DKK16]) puede considerarse como una versión fuerte del lema de regularidad.

Como hemos visto a lo largo de este breve panorama, la importancia del lema de regularidad de Szemerédi se volvió evidente décadas después. Sin saberlo en el momento, un lema utilizado en la demostración del teorema de regularidad de Szemerédi influiría no solamente en la teoría de números, sino en diversas ramas de las matemáticas. El lema de regularidad de Szemerédi [Sze75b] marca el punto de partida de nuevas preguntas, y con ello nuevas rutas de investigación.

El presente trabajo está dirigido a estudiantes de licenciatura que apetecen de aplicaciones de la teoría de martingalas, o de quienes quieren profundizar un poco más en la Probabilidad después de un curso de la teoría de la medida. No obstante, con la finalidad de que el trabajo fuese autocontenido, se anexaron dos apéndices, que además dan constancia de la estricta formalidad de este trabajo en los cuatro capítulos. Las demostraciones presentadas en esta tesis se basan mayoritariamente en resultados de los apéndices A y B, propios de un curso riguroso de Probabilidad.

El objetivo del primer capítulo es presentar la prueba del lema de regularidad de Szemerédi en la teoría de gráficas de principio a fin, vía la teoría de la probabilidad. A lo largo del Capítulo 1 nos guiamos en todo momento del trabajo de Dodos, Kanellopoulos y Karageorgos [DKK16]. En las primeras tres secciones se desarrollan, con lujo de detalle, las herramientas necesarias para la demostración del lema fuerte de regularidad en su versión probabilista (Teorema 1.40), junto con algunos ejemplos. Acto seguido, comenzamos la cuarta sección del Capítulo 1 con una motivación y una heurística de la demostración del Teorema 1.40. El Teorema 1.40 es el teorema principal de esta tesis. Luego, re-escribimos el Teorema 1.40 al Corolario 1.42 para mayor similitud y menor abstracción en la relación de estos dos resultados con el lema fuerte de regularidad en su versión analista. Concluimos el primer capítulo exponiendo y ahondando en los detalles de la demostración de Tao [Tao06] del lema de regularidad de Szemerédi en su versión casi original como una consecuencia del Corolario 1.42.

El segundo capítulo está dedicado a recopilar los conceptos básicos del lenguaje de una teoría reciente conocida como la teoría de límites de gráficas. Así mismo, se muestran algunas relaciones entre grafones y gráficas. También, se analizan propiedades de la métrica de corte en el espacio de gráficas, así como la métrica análoga en el espacio de grafones. Por último, se expone el teorema

de compacidad en el espacio de grafones, consecuencia del lema débil de regularidad en su versión analista y del teorema de convergencia de martingalas acotadas (Teorema B.14).

El tercer capítulo es una continuación del primer capítulo. En él se exponen algunas de las aplicaciones teóricas del resultado principal (Teorema 1.40) mostrado en el Capítulo 1. Se aborda en la segunda sección la aplicación más importante del tercer capítulo, enunciada en la Proposición 3.7, y obtenida como consecuencia del lema fuerte de regularidad (versión probabilista) en virtud del Lema 3.6.

Finalmente, en el cuarto y último capítulo, se expone un resumen del artículo de Chatterjee, Diaconis y Sly [CDS11]. Este capítulo supone que las lectoras y los lectores ya conocen los conceptos básicos de la teoría de límites de gráficas, aunque se mencionan brevemente en la primera sección del cuarto capítulo como una introducción formal. El teorema principal de [CDS11] se enuncia en esta tesis como el Teorema 4.3, y concierne a la convergencia de una sucesión de gráficas aleatorias cada una con una sucesión de grados dada, a un grafón,  $W$ , descrito de manera explícita en términos de una cierta función  $g$ . Después se enuncia y demuestra la Proposición 4.4, que se interpreta como la versión continua del criterio de Erdős-Gallai en la teoría de gráficas tradicional. Seguido de ello, vemos un modelo de simulación de gráficas aleatorias, llamado Modelo  $\beta$ . Su importancia se debe a que subyace en la prueba del Teorema 4.3 (resultado principal de [CDS11]). Por último, pero quizá lo más detallado de este último capítulo, se enuncia y demuestra el Teorema 4.12, donde se desarrolla un algoritmo para obtener el estimador máximo verosímil,  $\hat{\beta}$ , suponiendo su existencia. La prueba del Teorema 4.12 es cercana al Teorema del punto fijo de Banach y es asequible con un primer curso de Análisis matemático.



## El lema de regularidad de Szemerédi

El presente capítulo está dedicado a la demostración del lema fuerte de regularidad en su versión probabilista, cercana a la versión desde el área de análisis. Comenzamos enunciando el concepto de semianillo, el cual será central en nuestra demostración. Después, desarrollamos propiedades y ejemplos de otras definiciones que conforman la estructura en la demostración de nuestro resultado principal. Finalmente, aplicamos todos los conceptos previos en nuestra última sección, donde demostraremos la versión más cercana, dentro de la teoría de gráficas, del lema de regularidad de Szemerédi.

### 1. Semianillos: Construcciones y propiedades

Una filtración juega un papel importante para establecer la definición de una martingala. Intuitivamente, dados dos elementos de la filtración, digamos  $\mathcal{A}$  y  $\mathcal{B}$  con  $\mathcal{A} \subset \mathcal{B}$ , la  $\sigma$ -álgebra  $\mathcal{B}$  contiene más información que la  $\sigma$ -álgebra  $\mathcal{A}$  en el sentido de que  $\mathcal{B}$  contiene más elementos como conjunto ó, en el sentido probabilista, hay más eventos en  $\mathcal{B}$ . Podemos pensar a una sucesión de semianillos como el análogo de una filtración. La ventaja de trabajar con semianillos se debe a que son conocidos como una familia estructurada de conjuntos ([DKK16, p. 3], [BN08]).

DEFINICIÓN 1.1. *Supongamos que  $\Omega$  es un conjunto no vacío y  $k$  es un entero positivo. Supongamos también que  $\mathcal{S}$  es una colección de subconjuntos de  $\Omega$ . Decimos que  $\mathcal{S}$  es un  $k$ -semianillo en  $\Omega$  si se satisfacen simultáneamente las siguientes propiedades:*

- (P1).  $\emptyset, \Omega \in \mathcal{S}$ .
- (P2). Para cualesquiera  $S, T \in \mathcal{S}$  se cumple  $S \cap T \in \mathcal{S}$ .
- (P3). Para cualesquiera  $S, T \in \mathcal{S}$  existe un entero  $\ell \in [k]$  y conjuntos  $R_1, \dots, R_\ell \in \mathcal{S}$  ajenos por pares tales que  $S \setminus T = R_1 \cup \dots \cup R_\ell$ .

Una breve ilustración del concepto de semianillo se encuentra en la Figura 1, donde suponemos que  $S$  y  $T$  son dos elementos de un  $k$ -semianillo  $\mathcal{S}$  para algún entero positivo  $k$  mayor o igual a 5. En dicha figura mostramos que podemos descomponer a los elementos  $S$  y  $T$  en conjuntos más pequeños. Esto será de utilidad cuando trabajemos con una partición y busquemos un refinamiento.

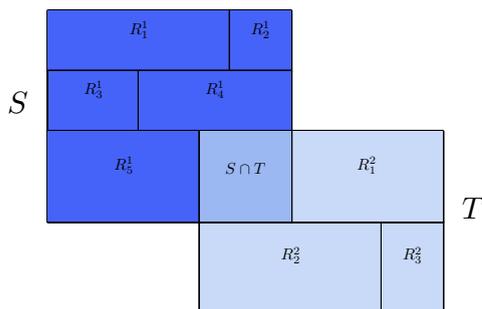


FIGURA 1. Descomposición de dos elementos de un  $k$ -semianillo para algún  $k \geq 5$ .

La siguiente observación enuncia la relación de un semianillo con el concepto de álgebra de la teoría de la medida.

OBSERVACIÓN 1.2. Una familia  $\mathcal{S}$  de subconjuntos de un conjunto no vacío,  $\Omega$ , es un 1-semianillo si y sólo si es un álgebra.

DEMOSTRACIÓN. Las propiedades (P1), (P2) y (P3) referidas son las de un 1-semianillo (Definición 1.1). Mientras que las propiedades (i), (ii) y (iii) son propiedades de un álgebra (Definición A.1).

Supongamos que  $\mathcal{S}$  es un 1-semianillo. Entonces (P1) asegura  $\Omega \in \mathcal{S}$ . Sean  $A, B \in \mathcal{S}$ . Luego, (P3) nos dice que existe  $R_1 \in \mathcal{S}$  tal que  $\Omega \setminus A = R_1$ ; es decir,  $\Omega \setminus A \in \mathcal{S}$ , por lo que  $\mathcal{S}$  es cerrada bajo complementos. A su vez,  $\Omega \setminus B \in \mathcal{S}$ . Finalmente, por la propiedad (P2) tenemos  $\Omega \setminus (A \cup B) = (\Omega \setminus A) \cap (\Omega \setminus B) \in \mathcal{S}$ . Consecuentemente, como  $\mathcal{S}$  es cerrada bajo complementos,  $A \cup B = \Omega \setminus (\Omega \setminus (A \cup B)) \in \mathcal{S}$ . Por ende,  $\mathcal{S}$  es un álgebra.

Inversamente, supongamos que  $\mathcal{S}$  es un álgebra. Por (i),  $\Omega$  pertenece a  $\mathcal{S}$ . Además, el conjunto vacío pertenece a  $\mathcal{S}$  por ser el complemento de  $\Omega$ . Sean  $S, T \in \mathcal{S}$ . Entonces  $\Omega \setminus S, \Omega \setminus T \in \mathcal{S}$  debido a (ii). Esto último implica  $\Omega \setminus (S \cap T) = (\Omega \setminus S) \cup (\Omega \setminus T) \in \mathcal{S}$  por (iii). Nuevamente, por ser  $\mathcal{S}$  cerrada bajo complementos,  $S \cap T = \Omega \setminus (\Omega \setminus (S \cap T)) \in \mathcal{S}$ ; en consecuencia,  $\mathcal{S}$  es cerrada bajo intersección. Por último, para verificar que la propiedad (P3) se satisface, basta notar que  $S \setminus T$  es igual a la intersección entre los conjuntos  $S$  y  $\Omega \setminus T$  que, como ya se hizo notar, pertenecen a  $\mathcal{S}$ . Por lo tanto, al ser  $\mathcal{S}$  cerrada bajo intersección,  $S \setminus T \in \mathcal{S}$ .  $\square$

DEFINICIÓN 1.3. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Para cada  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  definimos

$$\|X\|_{\mathcal{S}} := \sup \left\{ \left| \int_S X d\mathbb{P} \right| : S \in \mathcal{S} \right\}.$$

Llamaremos a  $\|X\|_{\mathcal{S}}$  la  $\mathcal{S}$ -norma uniforme de  $X$ . No obstante, el hecho:  $\|\cdot\|_{\mathcal{S}}$  es una norma, dependerá de una propiedad particular en el  $k$ -semianillo, como se expone a continuación.

PROPOSICIÓN 1.4. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . La función  $\|\cdot\|_{\mathcal{S}}$  es una norma en  $L^1(\Omega, \mathcal{F}, \mathbb{P})$  si y sólo si para cualquier par de variables aleatorias  $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  con  $X \neq Y$ , existe  $S \in \mathcal{S}$  con la propiedad

$$\int_S X d\mathbb{P} \neq \int_S Y d\mathbb{P}.$$

DEMOSTRACIÓN. Supongamos primeramente que  $\|\cdot\|_{\mathcal{S}}$  es una norma en  $L^1(\Omega, \mathcal{F}, \mathbb{P})$ . Sean  $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  tales que  $X \neq Y$ . Entonces,  $\|X - Y\|_{\mathcal{S}} > 0$ . Consecuentemente, existe  $S \in \mathcal{S}$  con la propiedad

$$\left| \int_S X - Y d\mathbb{P} \right| > 0.$$

De aquí se deduce

$$\int_S X d\mathbb{P} \neq \int_S Y d\mathbb{P}.$$

Recíprocamente, supongamos ahora que para cualquier par de variables aleatorias  $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  con  $X \neq Y$ , existe  $S \in \mathcal{S}$  con la propiedad

$$\int_S X d\mathbb{P} \neq \int_S Y d\mathbb{P}.$$

Tomemos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  tal que  $X \equiv 0$ . En consecuencia, para cualquier  $S \in \mathcal{S}$ ,  $X \mathbb{1}_S \equiv 0$ . Así, para cualquier  $S \in \mathcal{S}$ ,

$$\int_S X d\mathbb{P} = 0.$$

Se sigue que  $\|X\|_{\mathcal{S}} = 0$ .

Sea  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  tal que  $\|X\|_{\mathcal{S}} = 0$ . Entonces, para cualquier  $S \in \mathcal{S}$ ,

$$\left| \int_S X d\mathbb{P} \right| = 0.$$

Si  $X \neq 0$ , entonces, por hipótesis, existe  $S \in \mathcal{S}$  tal que

$$\int_S X d\mathbb{P} \neq \int_S 0 d\mathbb{P} = 0.$$

Por lo cual,

$$\left| \int_S X d\mathbb{P} \right| > 0.$$

Así,  $\|X\|_{\mathcal{S}} > 0$ . Por ende, concluimos  $X \equiv 0$ .

Sean  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $\alpha \in \mathbb{R}$ . Tenemos, por linealidad de la integral en  $L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,

$$\begin{aligned} \|\alpha X\|_{\mathcal{S}} &= \sup \left\{ \left| \int_S \alpha X d\mathbb{P} \right| : S \in \mathcal{S} \right\} \\ &= \sup \left\{ \left| \alpha \int_S X d\mathbb{P} \right| : S \in \mathcal{S} \right\} \\ &= |\alpha| \sup \left\{ \left| \int_S X d\mathbb{P} \right| : S \in \mathcal{S} \right\} = |\alpha| \|X\|_{\mathcal{S}}, \end{aligned}$$

donde en la penúltima igualdad usamos un hecho básico de cálculo para la constante  $|\alpha| \geq 0$ . Finalmente, sean  $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $S \in \mathcal{S}$  arbitrarios. De la desigualdad del triángulo, usando implícitamente para la primera igualdad que  $X \mathbb{1}_S$  y  $Y \mathbb{1}_S$  son integrables,

$$\begin{aligned} \left| \int_S (X + Y) d\mathbb{P} \right| &= \left| \int_S X d\mathbb{P} + \int_S Y d\mathbb{P} \right| \\ &\leq \left| \int_S X d\mathbb{P} \right| + \left| \int_S Y d\mathbb{P} \right| \\ &\leq \|X\|_{\mathcal{S}} + \|Y\|_{\mathcal{S}}. \end{aligned}$$

Por consiguiente, dado que  $S \in \mathcal{S}$  fue arbitrario,  $\|X + Y\|_{\mathcal{S}} \leq \|X\|_{\mathcal{S}} + \|Y\|_{\mathcal{S}}$ . De modo que la  $\mathcal{S}$ -norma uniforme es, en efecto, una norma.  $\square$

LEMA 1.5. *Supongamos que  $\Omega$  es un conjunto no vacío. Supongamos también que  $m, k_1, \dots, k_m$  son enteros positivos y defina  $k = \sum_{i=1}^m k_i$ . Si  $\mathcal{S}_i$  es un  $k_i$ -semianillo en  $\Omega$  para cada  $i \in [m]$ , entonces la familia*

$$\mathcal{S} = \left\{ \bigcap_{i=1}^m S_i \mid S_i \in \mathcal{S}_i \text{ para cada } i \in [m] \right\}$$

es un  $k$ -semianillo en  $\Omega$ .

DEMOSTRACIÓN. El caso  $m = 1$  es inmediato por definición. Supongamos  $m \geq 2$ . Note que el conjunto vacío pertenece a  $\mathcal{S}_1$ , pues  $\mathcal{S}_1$  es un  $k_1$ -semianillo en  $\Omega$ . Como la intersección de cualquier conjunto con el vacío es, de vuelta, el vacío, entonces  $\emptyset \in \mathcal{S}$ . Ahora, para cada  $i \in [m]$ , tenemos  $\Omega \in \mathcal{S}_i$ . De manera que  $\Omega = \bigcap_{i=1}^m \Omega \in \mathcal{S}$ . Esto muestra que (P1) en la Definición 1.1 se satisface. Tomemos ahora dos elementos (al menos el vacío y  $\Omega$  están) arbitrarios  $S = \bigcap_{i=1}^m S_i, T = \bigcap_{i=1}^m T_i \in \mathcal{S}$ , donde  $S_i, T_i \in \mathcal{S}_i$  para cada  $i \in [m]$ . Observe que

$$S \cap T = \left( \bigcap_{i=1}^m S_i \right) \cap \left( \bigcap_{i=1}^m T_i \right) = \left( \bigcap_{i=1}^m (S_i \cap T_i) \right).$$

Ya que  $\mathcal{S}_i$  es un  $k_i$ -semianillo, entonces para cada  $i \in [m]$  se satisface  $S_i \cap T_i \in \mathcal{S}_i$ . Consecuentemente  $S \cap T \in \mathcal{S}$ , por lo que  $\mathcal{S}$  cumple la propiedad (P2) en la Definición 1.1. Para verificar que la propiedad (P3), de la misma definición, se satisface, defina los conjuntos  $P_1 = \Omega \setminus T_1$  y  $P_j = T_1 \cap \dots \cap T_{j-1} \cap (\Omega \setminus T_j)$  siempre que  $j \in \{2, \dots, m\}$ . Los conjuntos  $P_1, \dots, P_m$  son ajenos por pares, debido a que si  $i, j \in [m]$  son tales que  $i < j$ , entonces  $i \in [j-1]$ , por lo que  $P_i \subseteq \Omega \setminus T_i$  y  $P_j \subseteq T_1 \cap \dots \cap T_{j-1} \subseteq T_i$ ; por consiguiente,  $P_i \cap P_j = \emptyset$ . Más aún,

$$(1) \quad \Omega \setminus \left( \bigcap_{i=1}^m T_i \right) = \bigcup_{j=1}^m P_j.$$

En efecto, si  $x \in \bigcup_{j=1}^m P_j$ , existe  $j_0 \in [m]$  tal que  $x \in P_{j_0}$ , por lo cual  $x \in \Omega \setminus T_{j_0}$ ; en consecuencia,  $x \in \bigcup_{i=1}^m (\Omega \setminus T_i)$ ; por lo tanto,  $x \in \Omega \setminus \left( \bigcap_{i=1}^m T_i \right)$ . Recíprocamente, si  $x \in \Omega \setminus \left( \bigcap_{i=1}^m T_i \right)$ , entonces existe  $i_0 \in [m]$  tal que  $x \in \Omega \setminus T_{i_0}$ ; por el teorema del buen orden podemos considerar, sin perder generalidad, que  $i_0$  es el mínimo número en el conjunto  $[m]$  con la propiedad  $x \in \Omega \setminus T_{i_0}$ ; de manera que  $x \in P_{i_0}$ ; por ende,  $x \in \bigcup_{j=1}^m P_j$ .

La igualdad (1) implica

$$(2) \quad S \setminus T = S \cap \left( \bigcup_{j=1}^m P_j \right) = \bigcup_{j=1}^m (S \cap P_j).$$

Sea  $j \in [m]$  arbitrario. Puesto que  $\mathcal{S}_j$  es un  $k_j$ -semianillo, existe  $\ell_j \in [k_j]$  y conjuntos ajenos por pares  $R_1^j, \dots, R_{\ell_j}^j \in \mathcal{S}_j$  tales que  $S_j \setminus T_j = R_1^j \cup \dots \cup R_{\ell_j}^j$ . Para facilitar la notación define

- (a)  $B_1 = \Omega$  y  $B_j = \bigcap_{1 \leq i < j} (S_i \cap T_i)$  si  $j \in \{2, \dots, m\}$ ,
- (b)  $C_j = \bigcap_{j < i \leq m} S_i$  si  $j \in [m-1]$  y  $C_m = \Omega$ .

Recordemos la definición de  $P_1$  y también que  $C_1 = \bigcap_{1 < i \leq m} S_i$ . Esto implica la primer igualdad de la siguiente cadena de igualdades, en tanto que las igualdades subsecuentes se deben solamente a propiedades de asociatividad y distributividad de conjuntos:

$$(3) \quad S \cap P_1 = \left( C_1 \cap S_1 \right) \cap (\Omega \setminus T_1) = C_1 \cap (S_1 \setminus T_1) = C_1 \cap \left( \bigcup_{n=1}^{\ell_1} R_n^1 \right) = \bigcup_{n=1}^{\ell_1} (B_1 \cap R_n^1 \cap C_1),$$

donde en la última igualdad usamos  $B_1 = \Omega$ . Para  $j \in \{2, \dots, m\}$  podemos descomponer a  $S$  como

$$S = \left( \bigcap_{1 \leq i < j} S_i \right) \cap S_j \cap C_j.$$

De manera que obtenemos nuevamente una cadena de igualdades para  $j \in \{2, \dots, m\}$ , donde se han colocado paréntesis de manera apropiada aprovechándonos en varias ocasiones de la asociatividad, conmutatividad y distributividad de la intersección y la unión de conjuntos,

$$(4) \quad \begin{aligned} S \cap P_j &= \left( \left( \bigcap_{1 \leq i < j} S_i \right) \cap S_j \cap C_j \right) \cap (T_1 \cap \dots \cap T_{j-1} \cap (\Omega \setminus T_j)) \\ &= \left( \left( \bigcap_{1 \leq i < j} S_i \right) \cap (T_1 \cap \dots \cap T_{j-1}) \right) \cap S_j \cap C_j \cap (\Omega \setminus T_j) \\ &= \left( \bigcap_{1 \leq i < j} (S_i \cap T_i) \right) \cap (S_j \setminus T_j) \cap C_j \\ &= B_j \cap \left( \bigcup_{n=1}^{\ell_j} R_n^j \right) \cap C_j \\ &= \bigcup_{n=1}^{\ell_j} (B_j \cap R_n^j \cap C_j). \end{aligned}$$

Total, para  $j \in [m]$ , las igualdades (3) y (4) muestran que

$$(5) \quad S \cap P_j = \bigcup_{n=1}^{\ell_j} (B_j \cap R_n^j \cap C_j).$$

Este hecho implica, al sustituir (5) en (2), que  $S \setminus T$  se expresa como sigue:

$$(6) \quad S \setminus T = \bigcup_{j=1}^m \left( \bigcup_{n=1}^{\ell_j} (B_j \cap R_n^j \cap C_j) \right).$$

Defina el producto cartesiano  $I = \bigcup_{j=1}^m (\{j\} \times [\ell_j])$ . Es evidente que su cardinalidad es justamente  $|I| = \ell_1 + \dots + \ell_m$ . Dado que  $\ell_i \leq k_i$  para  $i \in [m]$ , tenemos  $|I| \leq k_1 + \dots + k_m = k$ . Para cada par

$(j, n) \in I$  defina

$$U_n^j = B_j \cap R_n^j \cap C_j$$

y observe que  $S_i \cap T_i \in \mathcal{S}_i$  para  $i \in [j-1]$  por ser  $\mathcal{S}_i$  un  $k_i$ -semianillo en  $\Omega$ , por lo que, por construcción,  $U_n^j \in \mathcal{S}$ ; además,  $U_n^j \subseteq R_n^j$ . Entonces,  $U_n^j \cap U_m^j \subseteq R_n^j \cap R_m^j = \emptyset$  siempre que  $n \neq m$ . También note que para el caso  $j = 1$ ,  $U_n^1 \subseteq R_n^1 \subseteq S_1 \setminus T_1 \subseteq \Omega \setminus T_1 = P_1$ . A su vez, para  $j \in \{2, \dots, m\}$ ,  $U_n^j \subseteq B_j \cap R_n^j \subseteq (T_1 \cap \dots \cap T_j) \cap (S_j \setminus T_j) \subseteq P_j$ . De cualquier manera, para cualquier  $(j, n) \in I$ ,  $U_n^j \subseteq P_j$ . Esto implica  $U_n^i \cap U_n^j \subseteq P_i \cap P_j = \emptyset$  siempre que  $i$  sea distinto de  $j$ . Todo lo anterior asegura que la familia  $\{U_n^j \mid (j, n) \in I\}$  es un subconjunto de  $\mathcal{S}$  y consiste de conjuntos ajenos por pares. Más aún, por (6), tenemos

$$S \setminus T = \bigcup_{(j,n) \in I} U_n^j.$$

Por lo tanto, la familia  $\mathcal{S}$  satisface la propiedad (P3) en la Definición 1.1. Concluimos que  $\mathcal{S}$  es un  $k$ -semianillo.  $\square$

**OBSERVACIÓN 1.6.** Sean  $d, k_1, \dots, k_d$  enteros positivos y  $\Omega_1, \dots, \Omega_d$  conjuntos no vacíos. Supongamos que  $\mathcal{S}_i$  es un  $k_i$ -semianillo en  $\Omega_i$  para  $i \in [d]$ . Entonces

$$\overline{\mathcal{S}_i} := \{\Omega_1 \times \dots \times S_i \times \dots \times \Omega_d \mid S_i \in \mathcal{S}_i\}$$

es un  $k_i$ -semianillo en  $\Omega_1 \times \dots \times \Omega_d$ .

**DEMOSTRACIÓN.** Fijemos  $i \in [d]$ . Dado que  $\mathcal{S}_i$  es un  $k_i$ -semianillo en  $\Omega_i$ , entonces  $\emptyset, \Omega_i \in \mathcal{S}_i$ . De manera que  $\emptyset, \Omega_1 \times \dots \times \Omega_i \times \dots \times \Omega_d \in \overline{\mathcal{S}_i}$ .

Supongamos ahora  $\Omega_1 \times \dots \times S \times \dots \times \Omega_d, \Omega_1 \times \dots \times T \times \dots \times \Omega_d \in \overline{\mathcal{S}_i}$ . Como  $S, T \in \mathcal{S}_i$ , entonces  $S \cap T \in \mathcal{S}_i$ . Por consiguiente,

$$(\Omega_1 \times \dots \times S \times \dots \times \Omega_d) \cap (\Omega_1 \times \dots \times T \times \dots \times \Omega_d) = \Omega_1 \times \dots \times (S \cap T) \times \dots \times \Omega_d \in \overline{\mathcal{S}_i}.$$

Verifiquemos la última propiedad de un  $k_i$ -semianillo. Observe que se cumple la igualdad siguiente.

$$\begin{aligned} & (\Omega_1 \times \dots \times S \times \dots \times \Omega_d) \setminus (\Omega_1 \times \dots \times T \times \dots \times \Omega_d) = \\ (7) \quad & = ((\Omega_1 \setminus \Omega_1) \times \dots \times S \times \dots \times \Omega_d) \cup \dots \cup (\Omega_1 \times \dots \times (S \setminus T) \times \dots \times \Omega_d) \cup \\ & \quad \dots \cup (\Omega_1 \times \dots \times S \times \dots \times (\Omega_d \setminus \Omega_d)) \\ & = \Omega_1 \times \dots \times (S \setminus T) \times \dots \times \Omega_d. \end{aligned}$$

Por otro lado,  $S \setminus T = \bigcup_{j=1}^{\ell_i} R_j$  para algún  $\ell_i \in [k_i]$ , donde  $R_1, \dots, R_{\ell_i} \in \mathcal{S}_i$ . Además, para cualesquiera conjuntos  $A, B$  se satisface

$$\left( A \times \left( \bigcup_{j=1}^{\ell_i} R_j \right) \right) \times B = \left( \bigcup_{j=1}^{\ell_i} (A \times R_j) \right) \times B = \bigcup_{j=1}^{\ell_i} ((A \times R_j) \times B) = \bigcup_{j=1}^{\ell_i} (A \times R_j \times B).$$

El hecho anterior nos permite mostrar que

$$\Omega_1 \times \dots \times \left( \bigcup_{j=1}^{\ell_i} R_j \right) \times \dots \times \Omega_d = \bigcup_{j=1}^{\ell_i} (\Omega_1 \times \dots \times R_j \times \dots \times \Omega_d),$$

con  $\Omega_1 \times \dots \times R_j \times \dots \times \Omega_d \in \overline{\mathcal{S}_i}$  para  $j \in [\ell_i]$ . Consecuentemente, esta última igualdad aplicada a lo obtenido en (7) nos da

$$(\Omega_1 \times \dots \times S \times \dots \times \Omega_d) \setminus (\Omega_1 \times \dots \times T \times \dots \times \Omega_d) = \bigcup_{j=1}^{\ell_i} (\Omega_1 \times \dots \times R_j \times \dots \times \Omega_d).$$

Esto muestra que  $\overline{\mathcal{S}_i}$  es un  $k_i$ -semianillo en  $\Omega_1 \times \dots \times \Omega_d$ .  $\square$

**COROLARIO 1.7.** Lo siguiente se satisface.

(a) Sea  $\Omega$  un conjunto no vacío. Sea  $k$  un entero positivo y para cada  $i \in [k]$  sea  $\mathcal{A}_i$  un álgebra en  $\Omega$ . Entonces la familia

$$\{A_1 \cap \cdots \cap A_k \mid A_i \in \mathcal{A}_i \text{ para cada } i \in [k]\}$$

es un  $k$ -semianillo en  $\Omega$ .

(b) Sean  $d, k_1, \dots, k_d$  enteros positivos y defina  $k = \sum_{i=1}^d k_i$ . También sean  $\Omega_1, \dots, \Omega_d$  conjuntos no vacíos y para cada  $i \in [d]$  sea  $\mathcal{S}_i$  un  $k_i$ -semianillo en  $\Omega_i$ . Entonces la familia

$$\{S_1 \times \cdots \times S_d \mid S_i \in \mathcal{S}_i \text{ para cada } i \in [d]\}$$

es un  $k$ -semianillo en  $\Omega_1 \times \cdots \times \Omega_d$ .

DEMOSTRACIÓN. (a): La Observación 1.2 asegura que  $\mathcal{A}_i$  es un 1-semianillo en  $\Omega$  para toda  $i \in [k]$ . En consecuencia, la parte (a) se satisface por el Lema 1.5.

(b): La Observación 1.6 nos dice que  $\overline{\mathcal{S}_i} := \{\Omega_1 \times \cdots \times S_i \times \cdots \times \Omega_d \mid S_i \in \mathcal{S}_i\}$  es un  $k_i$ -semianillo en  $\Omega_1 \times \cdots \times \Omega_d$  para  $i \in [d]$ . También se cumple que

$$(S_1 \times \Omega_2 \times \cdots \times \Omega_d) \cap (\Omega_1 \times S_2 \times \cdots \times \Omega_d) \cap \cdots \cap (\Omega_1 \times \Omega_2 \times \cdots \times S_d) = S_1 \times S_2 \times \cdots \times S_d,$$

Así, el Lema 1.5 asegura que

$$\{S_1 \times \cdots \times S_d \mid S_i \in \mathcal{S}_i \text{ para cada } i \in [d]\}$$

es un  $k$ -semianillo en  $\Omega_1 \times \cdots \times \Omega_d$ . □

EJEMPLO 1.8. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. El conjunto

$$\mathcal{S}_{\square} := \{S \times T \mid S, T \in \mathcal{F}\}$$

es un 2-semianillo en  $\Omega \times \Omega$ .

DEMOSTRACIÓN. Dado que  $\mathcal{F}$  es una  $\sigma$ -álgebra, en particular,  $\mathcal{F}$  es un álgebra. Equivalentemente  $\mathcal{F}$  es un 1-semianillo, como se hizo notar en la Observación 1.2. La parte (b) del Corolario 1.7 asegura que  $\mathcal{S}_{\square}$  es un 2-semianillo en  $\Omega \times \Omega$ . □

La  $\mathcal{S}_{\square}$ -norma uniforme es comúnmente conocida como norma de corte.

El siguiente semianillo fue introducido por Bollobás y Nikiforov [BN08] y puede ser considerado como la versión “simétrica” del Ejemplo 1.8.

EJEMPLO 1.9. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Defina

$$\Sigma_{\square} := \{S \times T \mid S, T \in \mathcal{F}, S \cap T = \emptyset \text{ o } S = T\}.$$

Entonces  $\Sigma_{\square}$  es un 4-semianillo.

DEMOSTRACIÓN. (P1): Es evidente que el conjunto vacío es un elemento de  $\Sigma_{\square}$  ya que  $\emptyset$  pertenece a  $\mathcal{F}$ . También,  $\Omega$  pertenece a  $\mathcal{F}$ , por lo cual  $\Omega \times \Omega$  es elemento de  $\Sigma_{\square}$ .

(P2): Sean  $S \times T, U \times V \in \Sigma_{\square}$ . Luego,

$$(S \times T) \cap (U \times V) = (S \cap U) \times (T \cap V).$$

Recordemos que  $\mathcal{F}$  es una  $\sigma$ -álgebra; en particular,  $\mathcal{F}$  es un álgebra, o equivalentemente,  $\mathcal{F}$  es un 1-semianillo, por esta razón  $\mathcal{F}$  es cerrada bajo intersecciones. De modo que  $S \cap U, T \cap V \in \mathcal{F}$ .

En caso de que se cumpla al menos una de las afirmaciones  $S \cap T = \emptyset$  o  $U \cap V = \emptyset$ , podemos asegurar  $(S \cap U) \cap (T \cap V) = \emptyset$ . Entonces el producto cartesiano de  $S \cap U$  y  $T \cap V$  pertenece a  $\Sigma_{\square}$ .

En el caso  $S \cap T \neq \emptyset$  y  $U \cap V \neq \emptyset$ , se satisface  $S = T$  y  $U = V$  por definición de  $\Sigma_{\square}$ . Consecuentemente,  $S \cap U = T \cap V$ . Por ende,  $(S \cap U) \times (T \cap V) \in \Sigma_{\square}$ .

Total, en cualquier caso,

$$(S \times T) \cap (U \times V) \in \Sigma_{\square}.$$

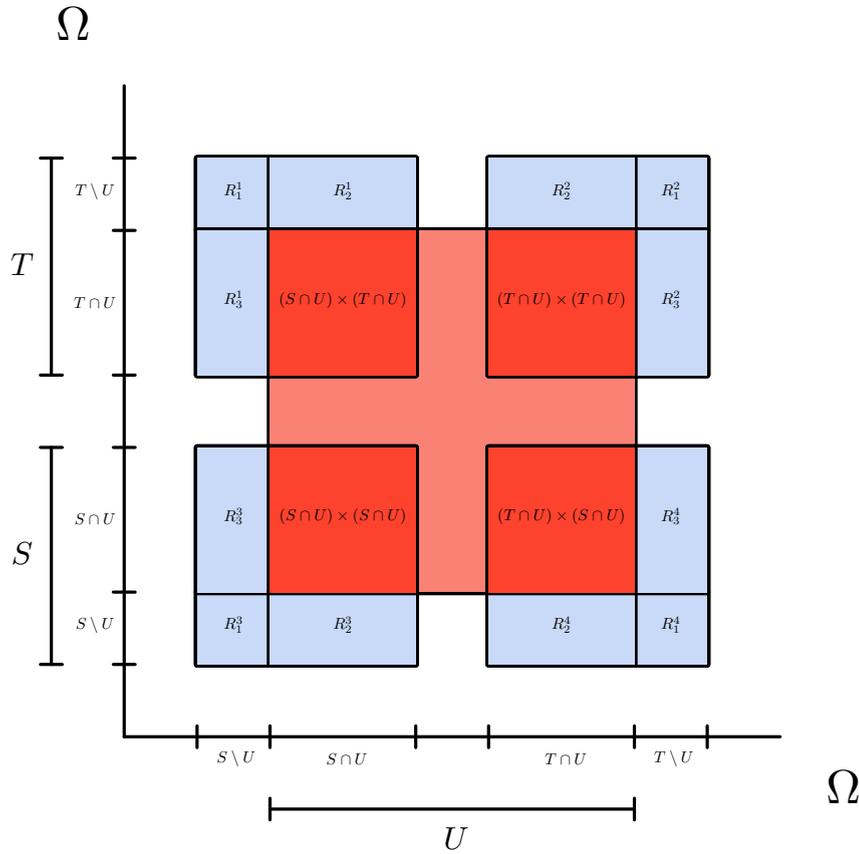


FIGURA 2. Ilustración del caso 1 ( $U$  igual a  $V$ ) en la propiedad (P3), dentro de la demostración del Ejemplo 1.9: Descomposición de los conjuntos  $(S \times T) \setminus (U \times U)$ ,  $(T \times S) \setminus (U \times U)$ ,  $(S \times S) \setminus (U \times U)$  y  $(T \times T) \setminus (U \times U)$  en tres conjuntos, cada uno, que pertenecen al 4-semianillo  $\Sigma_{\square}$ , cuando  $S$  es ajeno a  $T$ .

(P3): Sean  $S \times T, U \times V \in \Sigma_{\square}$ . Queremos ver que  $(S \times T) \setminus (U \times V)$  es igual a la unión de a lo más cuatro elementos en  $\Sigma_{\square}$  que son ajenos dos a dos.

*Caso 1.* Supongamos primero  $U = V$ . Es un ejercicio básico de conjuntos ver que la siguiente igualdad se satisface:

$$(S \times T) \setminus (U \times U) = ((S \setminus U) \times (T \cap U)) \cup ((S \setminus U) \times (T \setminus U)) \cup ((S \cap U) \times (T \setminus U)).$$

Dado que  $\mathcal{F}$  es cerrado bajo complementos e intersecciones, los conjuntos  $S \setminus U$ ,  $S \cap U$ ,  $T \setminus U$  y  $T \cap U$  pertenecen a  $\mathcal{F}$  porque  $S, T, U \in \mathcal{F}$ . Puesto que  $\Omega \setminus U$  y  $U$  son conjuntos ajenos, entonces  $S \setminus U$  y  $T \cap U$  también son conjuntos ajenos; análogamente  $S \cap U$  y  $T \setminus U$  son ajenos. Lo anterior asegura

$$(S \setminus U) \times (T \cap U), (S \cap U) \times (T \setminus U) \in \Sigma_{\square}.$$

Finalmente observemos que  $S \setminus U$  es igual al conjunto  $T \setminus U$  si  $S = T$ , y  $(S \setminus U) \cap (T \setminus U) = \emptyset$  si  $S \cap T = \emptyset$ . De cualquier forma

$$(S \setminus U) \times (T \setminus U) \in \Sigma_{\square}.$$

Esto muestra que  $(S \times T) \setminus (U \times U)$  es igual a la unión de a lo más tres conjuntos que pertenecen a  $\Sigma_{\square}$  y son ajenos dos a dos.

*Caso 2.* Supongamos ahora  $U \cap V = \emptyset$ . Mediante argumentos elementales de teoría de conjuntos es sencillo verificar, usando que  $U$  y  $V$  son ajenos, que se satisface:

$$(S \times T) \setminus (U \times V) = ((S \setminus U) \times (T \cap U)) \cup ((S \setminus U) \times (T \setminus U)) \cup ((S \cap U) \times (T \cap U)) \cup ((S \cap U) \times (T \setminus (U \cup V))).$$

Por el caso 1, resta mostrar

$$(S \cap U) \times (T \cap U), (S \cap U) \times (T \setminus (U \cup V)) \in \Sigma_{\square}.$$

Ya habíamos asegurado  $S \cap U, T \cap U \in \mathcal{F}$  en el caso pasado. Nuevamente porque  $\mathcal{F}$  es un álgebra y  $T, U, V \in \mathcal{F}$ , sabemos que se cumple  $T \setminus (U \cup V) \in \mathcal{F}$ . Por último,  $(S \cap U) \cap (T \cap U) = \emptyset$  si  $S \cap T = \emptyset$ , y  $S \cap U = T \cap U$  si  $S = T$ . En consecuencia,  $(S \cap U) \times (T \cap U) \in \Sigma_{\square}$ . Como  $T \setminus (U \cup V)$  es subconjunto de  $\Omega \setminus U$ , entonces  $T \setminus (U \cap V)$  y  $S \cap U$  son conjuntos ajenos. Por lo tanto,  $(S \cap U) \times (T \setminus (U \cup V)) \in \Sigma_{\square}$ .

Vemos que  $(S \times T) \setminus (U \times V)$  es la unión de a lo más cuatro elementos en  $\Sigma_{\square}$  que son ajenos dos a dos. Esto concluye el caso 2, y con ello la demostración de la propiedad (P3).  $\square$

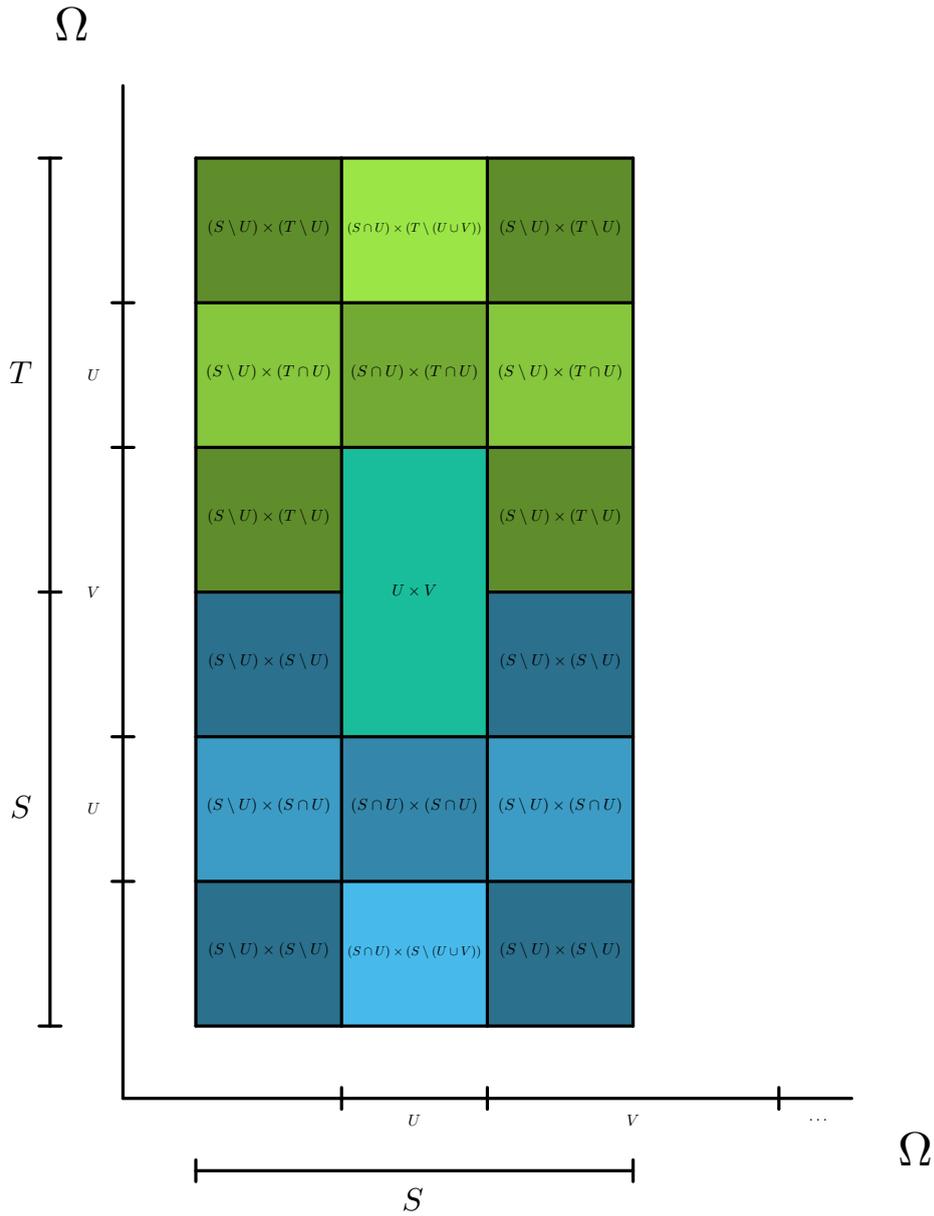


FIGURA 3. Ilustración del caso 2, propiedad (P3), en la demostración del Ejemplo 1.9: Descomposición de  $(S \times T) \setminus (U \times V)$  y  $(S \times S) \setminus (U \times V)$  en cuatro conjuntos, cada uno, que pertenecen al 4-semianillo  $\Sigma_{\square}$ , cuando  $S$  es ajeno a  $T$  y  $U$  es ajeno a  $V$ .

LEMA 1.10. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Supongamos también  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ . Entonces se satisface lo siguiente:

- (a) Tenemos  $\|X\|_{\mathcal{S}} \leq \|X\|_1$ .
- (b) Si  $\mathcal{G}$  es una  $\sigma$ -álgebra en  $\Omega$  con  $\mathcal{G} \subseteq \mathcal{S}$ , entonces  $\|\mathbb{E}(X|\mathcal{G})\|_{\mathcal{S}} \leq \|X\|_{\mathcal{S}}$ .
- (c) Si  $\mathcal{S}$  es una  $\sigma$ -álgebra, entonces  $\|X\|_{\mathcal{S}} \leq \|\mathbb{E}(X|\mathcal{S})\|_1 \leq 2\|X\|_{\mathcal{S}}$ .

DEMOSTRACIÓN. (a): Sea  $S \in \mathcal{S}$  fijo. Dado que  $X$  es integrable y  $S \in \mathcal{F}$ , sabemos que  $X\mathbb{1}_S$  también es integrable (Proposición A.34). Consecuentemente, por la desigualdad de Jensen (Teorema A.43),

$$\left| \int_S X d\mathbb{P} \right| := \left| \int X \mathbb{1}_S d\mathbb{P} \right| \leq \int |X| \mathbb{1}_S d\mathbb{P} = \int_S |X| d\mathbb{P}.$$

Puesto que  $S \subseteq \Omega$  y  $|X| \in L^+(\Omega, \mathcal{F})$ , en virtud de la Proposición A.33,

$$\int_S |X| d\mathbb{P} \leq \int |X| d\mathbb{P} = \|X\|_1.$$

Entonces, siguiendo la cadena de desigualdades,

$$\left| \int_S X d\mathbb{P} \right| \leq \|X\|_1.$$

Observe que la desigualdad anterior es válida para cada  $S \in \mathcal{S}$ . De modo que

$$\|X\|_{\mathcal{S}} := \sup \left\{ \left| \int_S X d\mathbb{P} \right| : S \in \mathcal{S} \right\} \leq \|X\|_1.$$

(b): Fijemos una  $\sigma$ -álgebra  $\mathcal{G}$  en  $\Omega$  con  $\mathcal{G} \subseteq \mathcal{S}$ . La pre-imagen de  $[0, \infty)$  bajo la variable aleatoria  $\mathbb{E}(X|\mathcal{G})$  es  $P := \{\omega \in \Omega \mid \mathbb{E}(X|\mathcal{G})(\omega) \geq 0\}$ . Debido a que  $\mathbb{E}(X|\mathcal{G})$  es  $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$ -medible y  $[0, \infty) \in \mathcal{B}(\mathbb{R})$ , entonces  $P \in \mathcal{G}$ . Defina  $N := \Omega \setminus P$ . Como  $\mathcal{G}$  es cerrada bajo complementos, por ser una  $\sigma$ -álgebra,  $N \in \mathcal{G} \subseteq \mathcal{S}$ . Fijemos también  $S \in \mathcal{S}$  arbitrario. Puesto que  $\mathbb{E}(X|\mathcal{G})$  es una variable aleatoria integrable por definición, y como  $S, P, N \in \mathcal{F}$  porque  $\mathcal{G} \subseteq \mathcal{S} \subseteq \mathcal{F}$ , entonces  $\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap P}$  y  $\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap N}$  también son integrables. Más aún,  $\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap N} \leq 0$  y  $\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap P} \geq 0$  por definición de  $P$  y  $N$ . Así, en uso de la Proposición A.35,

$$(8) \quad \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \leq 0 \leq \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}.$$

De la primera desigualdad, al sumar una integral finita,

$$\begin{aligned} \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} + \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} &\leq \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \\ &\leq \max \left\{ \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}, - \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right\}. \end{aligned}$$

En tanto que de la segunda desigualdad de (8),

$$\begin{aligned} \min \left\{ - \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}, \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right\} &\leq \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \\ &\leq \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} + \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}. \end{aligned}$$

Además,

$$\min \left\{ - \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}, \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right\} = - \max \left\{ \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}, - \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right\}.$$

Y en virtud del Teorema A.32, pues  $S \cap N$  y  $S \cap P$  son ajenos ya que  $P$  y  $N$  lo son,

$$\int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} + \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} = \int_S \mathbb{E}(X|\mathcal{G}) d\mathbb{P}.$$

Por consiguiente,

$$(9) \quad \left| \int_S \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right| \leq \max \left\{ \int_{S \cap P} \mathbb{E}(X|\mathcal{G}) d\mathbb{P}, - \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) d\mathbb{P} \right\}.$$

Notemos también que  $\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap P} \leq \mathbb{E}(X|\mathcal{G})\mathbb{1}_P$  y  $-\mathbb{E}(X|\mathcal{G})\mathbb{1}_{S \cap N} \leq -\mathbb{E}(X|\mathcal{G})\mathbb{1}_N$ . Entonces, nuevamente por la Proposición A.35,

$$\int_{S \cap P} \mathbb{E}(X|\mathcal{G})d\mathbb{P} \leq \int_P \mathbb{E}(X|\mathcal{G})d\mathbb{P} \text{ y } \int_{S \cap N} -\mathbb{E}(X|\mathcal{G})d\mathbb{P} \leq \int_N -\mathbb{E}(X|\mathcal{G})d\mathbb{P}.$$

De manera que

$$\max \left\{ \int_{S \cap P} \mathbb{E}(X|\mathcal{G}), - \int_{S \cap N} \mathbb{E}(X|\mathcal{G}) \right\} \leq \max \left\{ \int_P \mathbb{E}(X|\mathcal{G}), - \int_N \mathbb{E}(X|\mathcal{G}) \right\}.$$

Como último paso notemos que

$$\max \left\{ \int_P \mathbb{E}(X|\mathcal{G}), - \int_N \mathbb{E}(X|\mathcal{G}) \right\} = \max \left\{ \int_P X d\mathbb{P}, - \int_N X d\mathbb{P} \right\},$$

porque  $P, N \in \mathcal{G}$ , donde usamos una propiedad importante de la definición de esperanza condicional dada una  $\sigma$ -álgebra (Definición B.4, (iii)). Por lo tanto, aplicando este último hecho a (9),

$$\left| \int_S \mathbb{E}(X|\mathcal{G})d\mathbb{P} \right| \leq \max \left\{ \int_P X d\mathbb{P}, - \int_N X d\mathbb{P} \right\}.$$

A su vez,  $P, N \in \mathcal{S}$ , así que por definición de la  $\mathcal{S}$ -norma uniforme,

$$\left| \int_S \mathbb{E}(X|\mathcal{G})d\mathbb{P} \right| \leq \max \left\{ \int_P X d\mathbb{P}, - \int_N X d\mathbb{P} \right\} \leq \|X\|_{\mathcal{S}}.$$

Debido a la arbitrariedad de  $S \in \mathcal{S}$ , la desigualdad anterior es válida para todo  $S \in \mathcal{S}$ . Tomando supremos de ambos lados de la desigualdad podemos concluir con el resultado.

(c): Supongamos que  $\mathcal{S}$  es una  $\sigma$ -álgebra. De una propiedad de la variable aleatoria  $\mathbb{E}(X|\mathcal{S})$  (Definición B.4, (iii)),

$$\left| \int_S X d\mathbb{P} \right| = \left| \int_S \mathbb{E}(X|\mathcal{S})d\mathbb{P} \right|,$$

para todo  $S \in \mathcal{S}$ . Fijemos  $S \in \mathcal{S}$ . Puesto que  $\mathbb{E}(X|\mathcal{S})\mathbb{1}_S$  pertenece a  $L^1(\Omega, \mathcal{F}, \mathbb{P})$ , entonces, por la desigualdad de Jensen,

$$\left| \int_S X d\mathbb{P} \right| \leq \int_S |\mathbb{E}(X|\mathcal{S})| d\mathbb{P}.$$

De lo anterior, ya que  $S \subseteq \Omega$ , la Proposición A.33 asegura que

$$\left| \int_S X d\mathbb{P} \right| \leq \int |\mathbb{E}(X|\mathcal{S})| d\mathbb{P} =: \|\mathbb{E}(X|\mathcal{S})\|_1.$$

Hemos mostrado entonces que  $\|\mathbb{E}(X|\mathcal{S})\|_1$  es una cota superior para  $|\int_S X d\mathbb{P}|$  con  $S \in \mathcal{S}$ . De aquí es inmediato ver que

$$\|X\|_{\mathcal{S}} \leq \|\mathbb{E}(X|\mathcal{S})\|_1.$$

De la misma manera como escribimos en incisos anteriores, defina  $P := \{\omega \in \Omega \mid \mathbb{E}(X|\mathcal{S}) \geq 0\}$  y  $N = \Omega \setminus P$ . Dado que  $P \cap N = \emptyset$  y  $P \cup N = \Omega$ ,

$$\int |\mathbb{E}(X|\mathcal{S})| d\mathbb{P} = \int_P |\mathbb{E}(X|\mathcal{S})| d\mathbb{P} + \int_N |\mathbb{E}(X|\mathcal{S})| d\mathbb{P} = \int_P \mathbb{E}(X|\mathcal{S})d\mathbb{P} - \int_N \mathbb{E}(X|\mathcal{S})d\mathbb{P}.$$

De modo que

$$\|\mathbb{E}(X|\mathcal{S})\|_1 \leq 2 \cdot \max \left\{ \int_P \mathbb{E}(X|\mathcal{S})d\mathbb{P}, - \int_N \mathbb{E}(X|\mathcal{S})d\mathbb{P} \right\}.$$

Finalmente, ya que  $P, N \in \mathcal{S}$ ,

$$2 \cdot \max \left\{ \int_P \mathbb{E}(X|\mathcal{S})d\mathbb{P}, - \int_N \mathbb{E}(X|\mathcal{S})d\mathbb{P} \right\} = 2 \cdot \max \left\{ \int_P X d\mathbb{P}, - \int_N X d\mathbb{P} \right\} \leq 2\|X\|_{\mathcal{S}}.$$

Por ende,

$$\|\mathbb{E}(X|\mathcal{S})\|_1 \leq 2\|X\|_{\mathcal{S}}.$$

□

## 2. Particiones y refinamientos

DEFINICIÓN 1.11. Decimos que una familia de conjuntos no vacíos,  $\mathcal{P}$ , es una **partición de  $\Omega$**  si satisfacen simultáneamente las siguientes propiedades:

- (a) Los conjuntos que forman  $\mathcal{P}$  son ajenos dos a dos, es decir,  $A, B \in \mathcal{P}$  y  $A \neq B$  implica  $A \cap B = \emptyset$ .
- (b) La unión de  $\mathcal{P}$  es  $\Omega$ , es decir,  $\Omega = \bigcup \mathcal{P}$ .

Si además  $\mathcal{P}$  tiene cardinalidad finita, decimos que  $\mathcal{P}$  es una **partición finita de  $\Omega$** .

Una ventaja de trabajar con particiones en espacios de medida es que podemos describir la  $\sigma$ -álgebra generada por una partición de manera explícita.

PROPOSICIÓN 1.12. Supongamos que  $\mathcal{P}$  es una partición finita de un conjunto no vacío  $\Omega$ . Entonces

$$\sigma(\mathcal{P}) = \left\{ \bigcup \mathcal{R} \mid \mathcal{R} \subseteq \mathcal{P} \right\}.$$

DEMOSTRACIÓN. Supongamos que  $\mathcal{P} = \{A_1, \dots, A_n\}$  es una partición de  $\Omega$  tal que  $|\mathcal{P}| = n$  para algún entero positivo  $n$ . Definamos

$$\mathcal{F} := \left\{ \bigcup \mathcal{R} \mid \mathcal{R} \subseteq \mathcal{P} \right\}.$$

Mostraremos primero  $\mathcal{F} \subseteq \sigma(\mathcal{P})$ . Sea  $\mathcal{R}$  un subconjunto de  $\mathcal{P}$ . Recordemos que  $\mathcal{P} \subseteq \sigma(\mathcal{P})$ , por lo que  $\mathcal{R} \subseteq \sigma(\mathcal{P})$ . Además  $\mathcal{R}$  tiene un número finito de elementos ya que  $\mathcal{P}$  tiene cardinalidad finita. Consecuentemente,  $\bigcup \mathcal{R} \in \sigma(\mathcal{P})$ . Por lo tanto,  $\mathcal{F} \subseteq \sigma(\mathcal{P})$ .

Observemos que  $\mathcal{P} \subseteq \mathcal{F}$ . En efecto, para  $A \in \mathcal{P}$ , se cumplen  $\{A\} \subseteq \mathcal{P}$  y  $A = \bigcup \{A\}$ . Entonces  $A \in \mathcal{F}$ . Por ende,  $\mathcal{P} \subseteq \mathcal{F}$ .

El siguiente paso será mostrar que  $\mathcal{F}$  es una  $\sigma$ -álgebra. Notemos que  $\mathcal{P}$  es subconjunto de sí mismo y, debido a que  $\mathcal{P}$  es una partición,  $\Omega = \bigcup \mathcal{P}$ . En consecuencia,  $\Omega \in \mathcal{F}$ . Ahora supongamos  $A \in \mathcal{F}$ . Queremos ver que se cumple  $\Omega \setminus A \in \mathcal{F}$ . Por definición de  $\mathcal{F}$ , el conjunto  $A$  es de la forma  $A = \bigcup \mathcal{R}$  para algún  $\mathcal{R} \subseteq \mathcal{P}$ . Si  $\mathcal{R} = \emptyset$ , entonces  $A = \emptyset$ , y ya vimos que el complemento de  $A$ ,  $\Omega = \Omega \setminus \emptyset$ , pertenece a  $\mathcal{F}$ . Supongamos, pues, que  $\mathcal{R}$  es un subconjunto no vacío de  $\mathcal{P}$ . De manera que podemos describir a  $\mathcal{R}$  como  $\mathcal{R} = \{A_{i_1}, \dots, A_{i_m}\}$  con  $|\mathcal{R}| = m$  para  $\{i_1, \dots, i_m\} \subseteq [n]$  y  $m \in \mathbb{Z}_{>0}$  con  $m \leq n$ . Si  $m$  es igual a  $n$ , entonces  $\mathcal{P} = \mathcal{R}$ , por lo que  $A = \Omega$ ; consecuentemente el complemento de  $A$ ,  $\emptyset$ , pertenece a  $\mathcal{F}$  evidentemente. Entonces supongamos  $0 < |\mathcal{R}| < n$  y escribamos  $\mathcal{P} \setminus \mathcal{R} = \{A_{i_{m+1}}, \dots, A_{i_n}\}$ . Por consiguiente,

$$\Omega \setminus \bigcup \mathcal{R} = \Omega \setminus \left( \bigcup_{j=1}^m A_{i_j} \right) = \bigcup_{j=m+1}^n A_{i_j}$$

y es claro que  $\bigcup_{j=m+1}^n A_{i_j} \in \mathcal{F}$ . En cualquier caso, para todo  $A \in \mathcal{F}$ , se satisface  $\Omega \setminus A \in \mathcal{F}$ . Ahora supongamos  $A, B \in \mathcal{F}$ . Para terminar de mostrar la afirmación “ $\mathcal{F}$  es una  $\sigma$ -álgebra” consideremos una colección  $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$ ; entonces, para cada  $n \in \mathbb{N}$ ,  $A_n$  es de la forma  $A_n = \bigcup \mathcal{R}_n$  para algún  $\mathcal{R}_n \subseteq \mathcal{P}$ . Consecuentemente,

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left( \bigcup \mathcal{R}_n \right) = \bigcup \left( \bigcup_{n=1}^{\infty} \mathcal{R}_n \right).$$

Notando que  $\bigcup_{n=1}^{\infty} \mathcal{R}_n \subseteq \mathcal{P}$ , concluimos  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ . Por lo tanto,  $\mathcal{F}$  es una  $\sigma$ -álgebra. Los hechos  $\mathcal{P}$  es subconjunto de  $\mathcal{F}$  y  $\mathcal{F}$  es una  $\sigma$ -álgebra aseguran  $\sigma(\mathcal{P}) \subseteq \mathcal{F}$  (Proposición A.3). Y recordemos, de la primera parte de la demostración,  $\mathcal{F} \subseteq \sigma(\mathcal{P})$ . Esto muestra  $\mathcal{F} = \sigma(\mathcal{P})$ .  $\square$

En Cálculo se busca aproximar el área de una región mediante particiones cada vez más finas, y es ahí donde se introduce por primera vez la noción de refinamiento. Nos gustaría refinar una partición, teniendo más conjuntos, para que nuestra variable aleatoria “esperanza condicional” sea

muy parecida (cerca en norma) a una variable aleatoria dada. Necesitaremos entonces una noción abstracta de refinamiento de una partición. La siguiente definición es debida a Bollobás y Nikiforov [BN08].

**DEFINICIÓN 1.13.** Sean  $\Omega$  un conjunto no vacío, y  $\mathcal{P}$  y  $\mathcal{Q}$  dos particiones de  $\Omega$ . Decimos que  $\mathcal{Q}$  es un **refinamiento de  $\mathcal{P}$**  si  $\mathcal{Q}$  es una partición de  $\Omega$  y para todo  $A \in \mathcal{P}$ ,  $A = \bigcup \mathcal{R}$  para algún  $\mathcal{R} \subseteq \mathcal{Q}$ .

En ocasiones resulta tedioso verificar cuándo una partición es un refinamiento de otra partición. No obstante, podremos identificar, dadas dos particiones, cuándo una partición es un refinamiento de la otra mediante la siguiente equivalencia.

**OBSERVACIÓN 1.14.** Sea  $\Omega$  un conjunto no vacío. Si  $\mathcal{P}$  y  $\mathcal{Q}$  son dos particiones finitas de  $\Omega$ , entonces  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$  si y sólo si  $\sigma(\mathcal{P}) \subseteq \sigma(\mathcal{Q})$ .

**DEMOSTRACIÓN.** Sean  $\mathcal{P}$  y  $\mathcal{Q}$  son dos particiones finitas de un conjunto no vacío  $\Omega$ . Supongamos que  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ . Sea  $A \in \mathcal{P}$ . Entonces existe  $\mathcal{R}$  un subconjunto de  $\mathcal{Q}$  con la propiedad  $\bigcup \mathcal{R} = A$ . Como  $\mathcal{Q} \subseteq \sigma(\mathcal{Q})$ , entonces también  $\mathcal{R} \subseteq \sigma(\mathcal{Q})$ . Luego,  $\mathcal{R}$  tiene cardinalidad finita, porque  $\mathcal{Q}$  es una partición finita y  $\mathcal{R} \subseteq \mathcal{Q}$ . De modo que  $\bigcup \mathcal{R} \in \sigma(\mathcal{Q})$ . Por lo tanto,  $A \in \sigma(\mathcal{Q})$ . En consecuencia,  $\mathcal{P} \subseteq \sigma(\mathcal{Q})$ . Por consiguiente,  $\sigma(\mathcal{P}) \subseteq \sigma(\mathcal{Q})$  (Proposición A.3). Recíprocamente, supongamos  $\sigma(\mathcal{P}) \subseteq \sigma(\mathcal{Q})$ . Sea  $A \in \mathcal{P}$ . Dado que  $\mathcal{P} \subseteq \sigma(\mathcal{P}) \subseteq \sigma(\mathcal{Q})$ , entonces  $A \in \sigma(\mathcal{Q})$ . La Proposición 1.12 asegura que existe  $\mathcal{R} \subseteq \mathcal{Q}$  con la propiedad  $A = \bigcup \mathcal{R}$ . Por lo tanto,  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ .  $\square$

**DEFINICIÓN 1.15.** Sean  $\mathcal{P}$  y  $\mathcal{Q}$  dos particiones de  $\Omega$ . Definimos  $\mathcal{P} \cap \mathcal{Q}$  como la partición consistente de todas las intersecciones no vacías  $P \cap Q$ , con  $P \in \mathcal{P}$  y  $Q \in \mathcal{Q}$ . En símbolos,

$$\mathcal{P} \cap \mathcal{Q} := \{P \cap Q \mid P \in \mathcal{P}, Q \in \mathcal{Q}, \text{ y } P \cap Q \neq \emptyset\}.$$

**OBSERVACIÓN 1.16.** Dadas dos particiones  $\mathcal{P}$  y  $\mathcal{Q}$  de  $\Omega$ ,  $\mathcal{P} \cap \mathcal{Q}$  es, en efecto, una partición de  $\Omega$ .

**DEMOSTRACIÓN.** Por definición, el conjunto  $\mathcal{P} \cap \mathcal{Q}$  es una familia de conjuntos no vacíos. Sean  $A, B \in \mathcal{P} \cap \mathcal{Q}$  tales que  $A \neq B$ . Entonces  $A = P_1 \cap Q_1$  y  $B = P_2 \cap Q_2$  para  $P_1, P_2 \in \mathcal{P}$  y  $Q_1, Q_2 \in \mathcal{Q}$ . Observe que  $P_1 = P_2$  y  $Q_1 = Q_2$  implica  $A = B$ , por lo que, necesariamente,  $P_1 \neq P_2$  ó  $Q_1 \neq Q_2$ . Si  $P_1 \neq P_2$ , entonces  $A \cap B \subseteq P_1 \cap P_2 = \emptyset$ . Análogamente, si  $Q_1 \neq Q_2$ , entonces  $A \cap B \subseteq Q_1 \cap Q_2 = \emptyset$ . En cualquier caso,  $A \neq B$  implica  $A \cap B = \emptyset$ . Por último,

$$\bigcup (\mathcal{P} \cap \mathcal{Q}) = \left( \bigcup \mathcal{P} \right) \cap \left( \bigcup \mathcal{Q} \right) = \Omega.$$

Por lo tanto,  $\mathcal{P} \cap \mathcal{Q}$  es una partición de  $\Omega$ .  $\square$

Un argumento similar muestra que lo anterior es válido para un número finito de intersecciones.

**OBSERVACIÓN 1.17.** Supongamos que  $k$  es un entero positivo y que  $\mathcal{P}_1, \dots, \mathcal{P}_k$  son particiones de  $\Omega$ . Entonces

$$\bigcap_{i \in [k]} \mathcal{P}_i := \{P_1 \cap \dots \cap P_k \mid P_i \in \mathcal{P}_i \text{ para cada } i \in [k] \text{ y } P_1 \cap \dots \cap P_k \neq \emptyset\}.$$

es una partición de  $\Omega$ .

Para culminar con esta sección de particiones en espacios de probabilidad enunciamos la siguiente definición.

**DEFINICIÓN 1.18.** Sea  $\mathcal{P}$  una partición finita de un conjunto no vacío  $\Omega$ . Definimos

$$\mathcal{P}^2 := \{P_{i_1} \times P_{i_2} \mid P_{i_1}, P_{i_2} \in \mathcal{P}\}.$$

**OBSERVACIÓN 1.19.** Sea  $\Omega$  un conjunto no vacío. Dada una partición  $\mathcal{P}$  de  $\Omega$ ,  $\mathcal{P}^2$  es una partición de  $\Omega \times \Omega$ .

DEMOSTRACIÓN. La familia de conjuntos  $\mathcal{P}^2$  consiste de conjuntos no vacíos porque  $\mathcal{P}$  es una familia de conjuntos no vacíos al ser una partición. Sean  $P_1 \times P_2, Q_1 \times Q_2 \in \mathcal{P}^2$ . Queremos ver que  $P_1 \times P_2 \neq Q_1 \times Q_2$  implica  $(P_1 \times P_2) \cap (Q_1 \times Q_2) = \emptyset$ . Argumentando por contrapositiva, supongamos  $(P_1 \times P_2) \cap (Q_1 \times Q_2) \neq \emptyset$ . Recordemos que se cumple la igualdad

$$(P_1 \times P_2) \cap (Q_1 \times Q_2) = (P_1 \cap Q_1) \times (P_2 \cap Q_2).$$

Entonces, necesariamente,  $P_1 \cap Q_1 \neq \emptyset$  y  $P_2 \cap Q_2 \neq \emptyset$ . Ya que  $\mathcal{P}$  es una partición, tenemos como consecuencia  $P_1 = Q_1$  y  $P_2 = Q_2$ . Por ende,  $P_1 \times P_2 = Q_1 \times Q_2$ , como queríamos demostrar.

Para verificar la última propiedad de una partición, tomemos  $(x, y) \in \Omega \times \Omega$  arbitrario. Como  $\mathcal{P}$  es una partición, existen  $P, Q \in \mathcal{P}$  tales que  $x \in P$  y  $y \in Q$ . Por consiguiente,  $(x, y) \in P \times Q$ . Se sigue que  $\Omega \times \Omega \subseteq \bigcup \mathcal{P}^2$ . La otra contención es clara. Por lo tanto,  $\Omega \times \Omega = \bigcup \mathcal{P}^2$ .

Así,  $\mathcal{P}^2$  es una partición de  $\Omega \times \Omega$ .  $\square$

### 3. Funciones de crecimiento

DEFINICIÓN 1.20. Una función  $F : \mathbb{N} \rightarrow \mathbb{R}$  es **creciente** si para cualesquiera  $n, m \in \mathbb{N}$  con  $n < m$  se satisface  $F(n) < F(m)$ .

DEFINICIÓN 1.21. Decimos que una función  $F : \mathbb{N} \rightarrow \mathbb{R}$  es una **función de crecimiento** si cumple que

- (i)  $F$  es creciente, y
- (ii)  $F(n) \geq n + 1$  para cada  $n \in \mathbb{N}$ .

NOTACIÓN 1.22. Una función  $F : \mathbb{N} \rightarrow \mathbb{N}$  también será llamada una **función de crecimiento** si satisface las propiedades (i) y (ii) en la Definición 1.21.

EJEMPLO 1.23. Sea  $0 < \epsilon \leq 1$ . La función  $F : \mathbb{N} \rightarrow \mathbb{N}$  con regla de correspondencia

$$F(n) := \left\lceil \frac{(1 + \epsilon)^2 (2^{2n+2})^2}{\epsilon^3} \right\rceil,$$

es una función de crecimiento.

DEMOSTRACIÓN. Es fácil ver que  $F$  es una función creciente. En efecto, si  $n, m \in \mathbb{N}$  son tales que  $n < m$ , entonces  $2^{2n+2} < 2^{2m+2}$ , lo cual implica  $F(n) \leq F(m)$ . Supongamos  $F(n) = F(m)$ . Entonces

$$F(m) - 1 < \frac{(1 + \epsilon)^2 (2^{2m+2})^2}{\epsilon^3} \leq F(m)$$

y

$$F(m) - 1 < \frac{(1 + \epsilon)^2 (2^{2n+2})^2}{\epsilon^3} \leq F(m).$$

Por ende,

$$\left| \frac{(1 + \epsilon)^2 (2^{2m+2})^2}{\epsilon^3} - \frac{(1 + \epsilon)^2 (2^{2n+2})^2}{\epsilon^3} \right| < 1.$$

En consecuencia,

$$|(2^{2m+2})^2 - (2^{2n+2})^2| < \epsilon^3 / (1 + \epsilon)^2 \leq 1,$$

lo cual no puede ocurrir. Entonces se satisface  $F(n) < F(m)$ . Por consiguiente,  $F$  es una función creciente.

Mostremos que se cumple  $F(n) \geq n + 1$  para todo  $n \in \mathbb{N}$ . La prueba es por inducción sobre  $n$ . En el caso base,  $n = 0$ , tenemos

$$F(0) \geq \frac{(1 + \epsilon)^2 (2^2)^2}{\epsilon^3} \geq 2^4 \geq 1.$$

Fijemos un número natural  $n$  mayor o igual a uno y supongamos  $F(n) \geq n + 1$ . Como  $F$  es creciente,  $F(n + 1) > F(n) \geq n + 1$ . Por ende,  $F(n + 1) \geq n + 2$ . El principio de inducción asegura que  $F(n) \geq n + 1$  para todo  $n \in \mathbb{N}$ . Esto prueba que  $F$  es una función de crecimiento.  $\square$

EJEMPLO 1.24. Sean  $k$  un entero positivo y  $F : \mathbb{N} \rightarrow \mathbb{N}$  una función de crecimiento. La función  $F' : \mathbb{N} \rightarrow \mathbb{N}$  definida como  $F'(n) = F((k + 1)^n)$  para  $n \in \mathbb{N}$  es una función de crecimiento.

DEMOSTRACIÓN. Primero veamos que  $F'$  es creciente. Dado que  $1 \leq k+1$  implica  $(k+1)^n \leq (k+1)^m$  para  $n, m \in \mathbb{N}$  tales que  $n \leq m$ , entonces  $F'(n) \leq F'(m)$  porque  $F$  es una función creciente. Verifiquemos ahora la segunda propiedad de una función de crecimiento. En el caso  $n = 0$  note que  $F'(0) = F(1) \geq 2 \geq 1$ . Luego, para  $n \in \mathbb{N}$  con  $n \geq 1$  tenemos, por el teorema del binomio de Newton,

$$(k+1)^n = \binom{n}{0}k^n + \cdots + \binom{n}{n}1^n \geq n.$$

Por consiguiente,  $F'((k+1)^n) \geq F'(n) \geq n+1$ . Es decir,  $F'(n) \geq n+1$ . Concluimos que  $F'$  es una función de crecimiento.  $\square$

NOTACIÓN 1.25. *Supongamos que  $F : \mathbb{N} \rightarrow \mathbb{N}$  es una función de crecimiento. Para cada  $\ell \in \mathbb{N}$  denotamos por  $F^{\circ\ell} : \mathbb{N} \rightarrow \mathbb{N}$  a la  $\ell$ -ésima composición de  $F$ , definida recursivamente por  $F^{\circ 0}(n) = n$  y  $F^{\circ(\ell+1)}(n) = F(F^{\circ\ell}(n))$  para  $n \in \mathbb{N}$ .*

OBSERVACIÓN 1.26. *Supongamos que  $F : \mathbb{N} \rightarrow \mathbb{N}$  es una función de crecimiento. Entonces  $F^{\circ\ell}(0) < F(F^{\circ\ell}(0))$  para todo  $\ell \in \mathbb{N}$ .*

DEMOSTRACIÓN. Sea  $F : \mathbb{N} \rightarrow \mathbb{N}$  una función de crecimiento. La demostración será por inducción sobre  $\ell \in \mathbb{N}$ . Sabemos,  $0 < 1 \leq F(0)$ . Ahora supongamos que para  $\ell \in \mathbb{N}$  se satisface  $F^{\circ\ell}(0) < F(F^{\circ\ell}(0))$ . Veamos que la afirmación es cierta para  $\ell+1$ . En efecto, como  $F$  es creciente,

$$F^{\circ(\ell+1)}(0) = F(F^{\circ\ell}(0)) < F(F(F^{\circ\ell}(0))) = F(F^{\circ(\ell+1)}(0)).$$

$\square$

EJEMPLO 1.27. *Supongamos que  $F$  es una función de crecimiento. Afirmamos que  $H : \mathbb{N} \rightarrow \mathbb{N}$  definida mediante la regla  $H(n) = F^{\circ(n+2)}(0)$  es una función de crecimiento.*

DEMOSTRACIÓN. Sea  $F : \mathbb{N} \rightarrow \mathbb{N}$  una función de crecimiento. La Observación 1.26 asegura  $H(i) < H(j)$  siempre que  $i < j$  con  $i, j \in \mathbb{N}$ ; esto es,  $H$  es una función creciente. La segunda propiedad de una función de crecimiento (Definición 1.21, (ii)) se verifica por inducción sobre  $n \in \mathbb{N}$ . Para la base de inducción,

$$H(0) = F^{\circ 2}(0) \geq F(0) \geq 1.$$

Luego, suponga que  $H(n) \geq n+1$  para algún  $n \in \mathbb{N}$ . Entonces, ya que  $F$  es una función de crecimiento,

$$H(n+1) = F^{\circ(n+3)}(0) = F(H(n)) \geq F(n+1) \geq n+2.$$

El principio de inducción nos permite deducir que  $H$  es una función de crecimiento.  $\square$

#### 4. Prueba del lema fuerte de regularidad vía martingalas

En esta sección exponemos la prueba del Teorema 1.40 concerniente al resultado principal de Dodos, Kanellopoulos y Karageorgos [DKK16, p. 8]. Este resultado afirma que existe una sola partición y un solo refinamiento que “regularizan” simultáneamente una familia finita de variables aleatorias que pertenecen al espacio  $L^p$  con  $1 < p \leq 2$ . Tal como se anunció en la introducción de esta tesis, dicho resultado contiene al de Terence Tao [Tao06, p. 15]. Mediante una sucesión de diferencias de martingala es posible extender el lema fuerte de regularidad a cualquier espacio  $L^p$  con  $1 < p \leq 2$ . No obstante, con el fin de extraer la esencia del resultado principal, trabajaremos en el espacio  $L^2$  con una sola variable aleatoria. En esta simplificación nos ahorramos un argumento combinatorio acerca del principio de las casillas.

**4.1. Motivación.** Cuando hablamos de la esperanza de cierta variable aleatoria podemos pensar en un promedio de los valores que toma nuestra variable cuando un experimento bajo las mismas condiciones se repite muchas veces.<sup>1</sup> De manera que una forma de aproximarnos, en promedio, a una variable aleatoria es con su esperanza. No obstante, la esperanza de una variable aleatoria no siempre es suficientemente cercana a la variable aleatoria. ¿Qué podemos hacer para que nuestra aproximación sea mejor?. El concepto de esperanza condicional dado un evento generaliza la noción de esperanza al restringirnos a un promedio de una variable aleatoria en un evento particular. De modo que dividir el espacio muestral en eventos más pequeños (en el sentido de medida) y calcular la esperanza relativa a cada evento parece ser una mejor idea para aproximar.

El concepto de esperanza condicional dada una  $\sigma$ -álgebra, digamos  $\mathcal{G}$ , es también una noción de aproximación en promedio a una variable aleatoria integrable  $X$  en el conjunto de eventos  $\mathcal{G}$  con otra variable aleatoria, denotada  $\mathbb{E}(X|\mathcal{G})$ . En ocasiones  $\mathbb{E}(X|\mathcal{G})$  puede ser una variable aleatoria más sencilla de describir que  $X$ .

Por ejemplo, dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  y una variable aleatoria  $X$ , integrable claro, consideremos una partición  $\mathcal{P}$  de nuestro espacio muestral  $\Omega$  consistente de eventos de probabilidad positiva y tal que  $\mathcal{P} \subseteq \mathcal{F}$ . Podemos aproximarnos a  $X$  en cada  $A \in \mathcal{P}$  al calcular la esperanza condicional de  $X$  dado el evento  $A$ , es decir,  $\mathbb{E}(X|A)$ . Por otro lado, al considerar la esperanza condicional de  $X$  dada la  $\sigma$ -álgebra generada por  $\mathcal{P}$ , a quien denotaremos por  $\mathbb{E}(X|\mathcal{P})$ ,<sup>2</sup> podemos decir explícitamente cómo se comporta la variable aleatoria  $\mathbb{E}(X|\mathcal{P})$  en cada evento  $A \in \mathcal{P}$ , ya que sabemos

$$\mathbb{E}(X|\mathcal{P}) \equiv \mathbb{E}(X|A) \text{ en cada } A \in \mathcal{P},$$

como asegura el Ejemplo B.6. Por consiguiente,  $\mathbb{E}(X|\mathcal{P})$  es ni más ni menos que una función simple en el sentido casi seguro.

Parece entonces que el problema de aproximarnos a una variable aleatoria se reduce a encontrar una partición adecuada  $\mathcal{P}$  de  $\Omega$  de tal suerte que  $\mathbb{E}(X|\mathcal{P})$  sea cercana a  $X$ . La idea subyacente en nuestra prueba del lema de regularidad de Szemerédi es aproximar una variable aleatoria en el espacio  $L^2$  mediante el siguiente algoritmo.

Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , tomemos cualquier  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  y fijemos  $0 < \epsilon \leq 1$ . Comenzamos con la partición trivial  $\{\Omega\}$ . Definamos  $\mathcal{F}_0 := \sigma(\{\Omega\}) = \{\emptyset, \Omega\}$  y  $X_0 := \mathbb{E}(X|\mathcal{F}_0)$ , que en este caso resulta ser una función constante, a saber,  $\mathbb{E}(X)$  (Ejemplo B.5). La norma  $\|\cdot\|_1$  dictará, en este ejemplo, la cercanía entre  $X_0$  y  $X$ . Si se cumple  $\|X - X_0\|_1 \leq \epsilon$ , entonces acabamos. De otra manera, elegimos un evento  $A_0 \in \mathcal{F}$  con la propiedad

$$\frac{\epsilon}{2} < \left| \int_{A_0} X - X_0 d\mathbb{P} \right|.$$

Como siguiente paso definimos  $\mathcal{F}_1 := \sigma(\mathcal{F}_0 \cup \{A_0\})$  y  $X_1 := \mathbb{E}(X|\mathcal{F}_1)$ . Note que agregamos más información a nuestra  $\sigma$ -álgebra original,  $\mathcal{F}_0$ , y por propiedades de esperanza condicional dada una  $\sigma$ -álgebra (Definición B.4, (iii))

$$\frac{\epsilon}{2} < \left| \int_{A_0} X_1 - X_0 d\mathbb{P} \right| \leq \|X_1 - X_0\|_1.$$

Observe que de alguna manera garantizamos que nos alejamos de nuestra aproximación anterior  $X_0$ . Si  $\|X - X_1\|_1 \leq \epsilon$ , detenemos nuestro proceso. En otro caso, realizamos una elección de un evento  $A_1 \in \mathcal{F}$  tal que

$$\frac{\epsilon}{2} < \left| \int_{A_1} X_1 - X_0 d\mathbb{P} \right|$$

y continuamos recursivamente.

La pregunta es si este algoritmo terminará eventualmente en un número finito de pasos. Notemos que siguiendo este proceso obtenemos una filtración  $(\mathcal{F}_n)$ , cuyos elementos son  $\sigma$ -álgebras finitas,

<sup>1</sup>Así lo afirma Miguel Ángel García Álvarez en [Álv15, p. 316] y detalla algunas ideas erróneas del concepto de esperanza.

<sup>2</sup>La lectora, o el lector, no debe tener inconveniente con esta notación, pues, dadas dos variables aleatorias integrables  $X, Y$ , es común denotar  $\mathbb{E}(X|Y)$ , en lugar de  $\mathbb{E}(X|\sigma(Y))$ .

y una sucesión de variables aleatorias  $(X_n)$  con la propiedad  $\|X_n - X_{n-1}\|_1 > \epsilon/2$  para todo entero positivo  $n$ . Más aún, la sucesión  $(X_n)$  es una martingala adaptada a la filtración  $(\mathcal{F}_n)$ , cuya sucesión de diferencias de martingala está lejos de cero en la norma  $\|\cdot\|_1$ . Las sucesiones de diferencias de martingala son estructuras ampliamente estudiadas; en particular, hay una desigualdad debida a Ricard y Xu [RX<sup>+</sup>16] que asegura que nuestro algoritmo terminará en a lo más  $\lfloor 4\|X\|_2\epsilon^{-2} \rfloor + 1$  pasos.

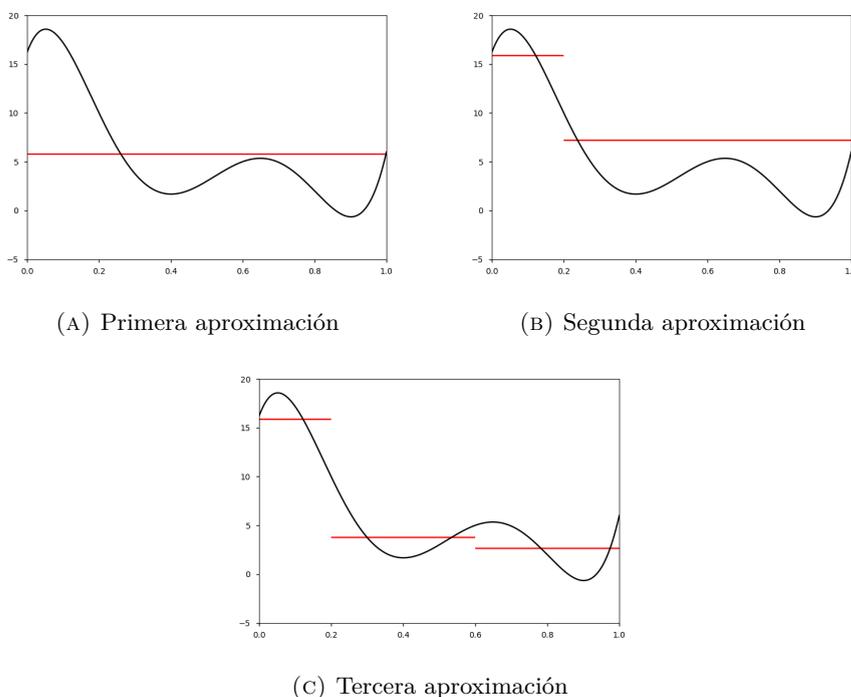


FIGURA 4. Simulación de una aproximación a un polinomio vía esperanza condicional.

Para afianzar aún más la idea de este algoritmo y esclarecer la visualización, permítanme exhibir en un ejemplo concreto (y determinista) los tres primeros pasos mencionados. Consideremos el espacio de probabilidad  $([0, 1], \mathcal{B}([0, 1]), \lambda)$ , donde convenimos en que  $\lambda$  es la restricción de la medida de Lebesgue a  $\mathcal{B}([0, 1])$ . Es claro que una función continua satisface la definición de una variable aleatoria; en particular, consideremos un polinomio de grado 5,  $X : [0, 1] \rightarrow \mathbb{R}$  dado por  $X(\omega) = a_5\omega^5 + \dots + a_1\omega + a_0$ , para  $\omega \in [0, 1]$ ; donde  $a_0, \dots, a_5 \in \mathbb{R}$  son algunas constantes fijas.<sup>3</sup>

El primer paso es considerar  $\mathcal{F}_0 := \{\emptyset, \Omega\}$ , con  $\Omega = [0, 1]$ , y  $X_0 := \mathbb{E}(X|\mathcal{F}_0) \equiv \mathbb{E}(X)$ , como lo muestra la Figura 4 (a). Luego, elegimos el evento  $A_0 = [0, 1/5]$ . En nuestro algoritmo no es difícil ver que  $\mathcal{F}_1 := \sigma(\mathcal{F}_0 \cup \{A_0\})$  es igual a  $\{\emptyset, A_0, \Omega \setminus A_0, \Omega\}$ , que a su vez coincide con la  $\sigma$ -álgebra generada por la partición  $\{A_0, \Omega \setminus A_0\}$  de  $\Omega$ . De modo que sabemos el comportamiento de  $X_1 := \mathbb{E}(X|\mathcal{F}_1)$  en  $\Omega = [0, 1]$ , a saber,

$$X_1(\omega) = \begin{cases} \mathbb{E}(X|A_0) & \text{si } \omega \in A_0, \\ \mathbb{E}(X|\Omega \setminus A_0) & \text{si } \omega \in \Omega \setminus A_0. \end{cases}$$

Posteriormente, elegimos el intervalo  $A_1 := [1/5, 3/5]$ . No es complicado demostrar (Observación 1.28) que  $\mathcal{F}_2 := \sigma(\mathcal{F}_1 \cup \{A_1\})$  es precisamente la  $\sigma$ -álgebra generada por la partición  $\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\}$  y, por ende, esto nos permite calcular el comportamiento de  $\mathbb{E}(X|\mathcal{F}_2)$  en los eventos  $A_0, A_1$  y  $\Omega \setminus (A_0 \cup A_1)$  como se muestra en la Figura 4 (c).

OBSERVACIÓN 1.28.  $\mathcal{F}_2 := \sigma(\mathcal{F}_1 \cup \{A_1\})$  es igual a  $\sigma(\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\})$ .

<sup>3</sup>El grado del polinomio es irrelevante en nuestro ejemplo, así como las constantes  $a_0, \dots, a_5$ .

DEMOSTRACIÓN. Ya sabemos  $\mathcal{F}_1 = \{\emptyset, A_0, \Omega \setminus A_0, \Omega\}$ . De modo que

$$\mathcal{F}_1 \cup \{A_1\} \subseteq \sigma(\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\}).$$

Consecuentemente,  $\sigma(\mathcal{F}_1 \cup \{A_1\}) \subseteq \sigma(\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\})$ . La otra contención se muestra de manera similar, es decir, notemos que  $\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\} \subseteq \sigma(\mathcal{F}_1 \cup \{A_1\})$  y, por tanto,  $\sigma(\{A_0, A_1, \Omega \setminus (A_0 \cup A_1)\}) \subseteq \sigma(\mathcal{F}_1 \cup \{A_1\})$ .  $\square$

**4.2. Tres lemas previos.** En esta subsección se encuentra la prueba, con todo rigor, de tres lemas previos al resultado principal. Se expone también la relación de estos lemas con teoría de gráficas y la demostración de Tao [Tao06].

LEMA 1.29 (Primer lema). *Fijemos  $k$  un entero positivo y un número real  $0 < \delta \leq 1$ . Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\Sigma$  un  $k$ -semianillo en  $\Omega$  con  $\Sigma \subseteq \mathcal{F}$ ,  $\mathcal{Q}$  una partición finita de  $\Omega$  subconjunto de  $\Sigma$  y  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  con la propiedad  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma} > \delta$ . Entonces existe un refinamiento de  $\mathcal{Q}$ ,  $\mathcal{R}$ , que también es un subconjunto de  $\Sigma$ , es tal que  $|\mathcal{R}| \leq |\mathcal{Q}|(k+1)$  y satisface*

$$\|\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})\|_2 > \delta.$$

DEMOSTRACIÓN. Supongamos que  $\mathcal{Q} = \{A_1, \dots, A_n\}$  es una partición finita de  $\Omega$  tal que  $|\mathcal{Q}| = n$  para algún entero positivo  $n$  y  $\mathcal{Q} \subseteq \Sigma$ . La hipótesis  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma} > \delta$  garantiza la existencia de un conjunto  $S \in \Sigma$  que cumple

$$(10) \quad \left| \int_S (X - \mathbb{E}(X|\mathcal{Q})) d\mathbb{P} \right| > \delta.$$

Notemos que la integral anterior es distinta de cero, por lo que  $S$  no es un conjunto nulo, en particular,  $S \neq \emptyset$ . De modo que, al ser  $\mathcal{Q}$  una partición, existe un subconjunto no vacío  $I$  de  $[n]$  tal que, para cada  $i \in I$ ,  $S \cap A_i \neq \emptyset$ . Sin perder generalidad podemos considerar a  $I$  como el subconjunto de  $[n]$  de cardinalidad máxima con la propiedad anterior. La propiedad (P3) en la Definición 1.1 asegura que para cada  $i \in I$ , existen un entero  $\ell_i \in [k]$  y conjuntos  $R_1^i, \dots, R_{\ell_i}^i \in \Sigma$  ajenos por pares tales que  $A_i \setminus S = R_1^i \cup \dots \cup R_{\ell_i}^i$ . Definamos

$$\mathcal{R} := \{S \cap A_i \mid i \in I\} \cup \{A_i \mid i \in [n] \setminus I\} \cup \left( \bigcup_{i \in I} \{R_j^i \mid j \in [\ell_i]\} \right).$$

Para ver que  $\mathcal{R}$  es una partición de  $\Omega$  note que cada  $A_i \neq \emptyset$  porque  $\mathcal{Q}$  es una partición. Luego,  $S \cap A_i \neq \emptyset$  para cada  $i \in I$  por construcción de  $I$ . A su vez, cada elemento de la forma  $R_j^i$  es no vacío. Es decir,  $\mathcal{R}$  es una familia de conjuntos no vacíos. Los elementos de  $\mathcal{R}$  son ajenos dos a dos, pues, al ser los elementos de  $\mathcal{Q}$  ajenos por pares, tenemos  $(S \cap A_i) \cap A_j = \emptyset$  para  $i \in I, j \in [n] \setminus I$ . Notemos también que para  $i \in I$  y  $j \in [\ell_i]$  se cumple  $R_j^i \subseteq A_i \setminus S$ , por lo que  $R_j^i \cap (A_i \cap S) = \emptyset$ ; además, si  $k \in [n]$  es distinto de  $i$ ,  $R_j^i \cap A_k = \emptyset$  y  $R_j^i \cap (A_k \cap S) = \emptyset$ . Consecuentemente, los elementos de  $\mathcal{R}$  son ajenos dos a dos (Definición 1.11, (a)). Para verificar que la unión de  $\mathcal{R}$  es  $\Omega$  (Definición 1.11, (b)), basta recordar que  $\bigcup_{j=1}^{\ell_i} R_j^i = A_i \setminus S$  para  $i \in I$ , de manera que, como  $\bigcup \mathcal{Q} = \Omega$  y

$$A_i = (A_i \setminus S) \cup (A_i \cap S) \subseteq \bigcup \mathcal{R},$$

tenemos  $\bigcup \mathcal{R} = \Omega$ . Por ende,  $\mathcal{R}$  es una partición de  $\Omega$ .

Ahora, notemos que  $A_i \in \mathcal{R} \subseteq \sigma(\mathcal{R})$  siempre que  $i \in [n] \setminus I$ . Para  $i \in I$  podemos descomponer  $A_i$  como la unión de los dos conjuntos  $(S \cap A_i)$  y  $A_i \setminus S = (\bigcup_{j=1}^{\ell_i} R_j^i)$ , los cuales pertenecen a  $\sigma(\mathcal{R})$ . Es decir,  $A_i \in \sigma(\mathcal{R})$  para  $i \in I$ . Por lo tanto,  $\mathcal{Q} \subseteq \sigma(\mathcal{R})$ . Consecuentemente,  $\sigma(\mathcal{Q}) \subseteq \sigma(\mathcal{R})$ ; equivalentemente,  $\mathcal{R}$  es un refinamiento de  $\mathcal{Q}$  por la Observación 1.14.

Afirmamos que la cardinalidad de  $\mathcal{R}$  está acotada por  $|\mathcal{Q}|(k+1)$ . En efecto, al ser  $\{S \cap A_i \mid i \in I\}$  y  $\{A_i \mid i \in [n] \setminus I\}$  conjuntos ajenos, tenemos que la cardinalidad de  $\{S \cap A_i \mid i \in I\} \cup \{A_i \mid i \in [n] \setminus I\}$  es  $|\mathcal{Q}|$ . Por otro lado, para  $i \in I$  sucede que  $\ell_i \leq k$ . Es decir, el conjunto  $\{R_j^i \mid j \in [\ell_i]\}$  contiene a lo más  $k$  elementos. Además, como  $I \subseteq [n]$ ,  $|I| \leq |\mathcal{Q}|$ . En consecuencia,

$$\left| \bigcup_{j \in I} \{R_k^j \mid k \in [\ell_j]\} \right| \leq |\mathcal{Q}|k.$$

Entonces  $|\mathcal{R}| \leq |\mathcal{Q}| + |\mathcal{Q}|k$ , como queríamos mostrar. Por otro lado, también se satisface que  $\mathcal{R} \subseteq \Sigma$ , pues  $\mathcal{Q} \subseteq \Sigma$  por hipótesis,  $\{S \cap A_i \mid i \in I\} \subseteq \Sigma$  y  $\bigcup_{j \in I} \{R_k^j \mid k \in [\ell_j]\} \subseteq \Sigma$  por la segunda y tercera propiedad de un  $k$ -semianillo, respectivamente.

Observemos que  $S = (\bigcup_{i=1}^n A_i) \cap S = \bigcup_{i \in I} (A_i \cap S)$ . Por consiguiente,  $S \in \sigma(\mathcal{R})$ . Se sigue, por una propiedad de la variable aleatoria  $\mathbb{E}(X|\mathcal{R})$  (Definición B.4, (iii)),

$$\int_S \mathbb{E}(X|\mathcal{R}) d\mathbb{P} = \int_S X d\mathbb{P}.$$

Aplicando este último hecho a (10), la desigualdad de Jensen y la propiedad de monotonía de las normas (Proposición A.50), obtenemos lo siguiente:

$$\delta < \left| \int_S (\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})) d\mathbb{P} \right| \leq \|\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})\|_1 \leq \|\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})\|_2.$$

Así,  $\mathcal{R}$  es el refinamiento buscado.  $\square$

Para exhibir la relación entre nuestros tres lemas previos al resultado principal (Lema 1.29, Lema 1.37 y Lema 1.38) y el lema de regularidad de Szemerédi [Sze75a], seguiremos la generalización abstracta de Bollobás y Nikiforov [BN08] del lema de regularidad, pues sus definiciones son cercanas a las utilizadas en la demostración original del lema de Szemerédi. Primero recordemos la definición, en probabilidad, de densidad entre dos eventos, análoga al concepto de densidad de aristas de una gráfica. La siguiente definición es debida a Bollobás y Nikiforov [BN08, p. 222].

DEFINICIÓN 1.30. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Para cada  $A, V \in \mathcal{F}$  definimos

$$d(A, V) := \frac{\mathbb{P}(A \cap V)}{\mathbb{P}(V)},$$

si  $\mathbb{P}(V) > 0$  y  $d(A, V) := 0$  si  $\mathbb{P}(V) = 0$ .

El concepto de *índice* de una partición del conjunto de vértices de una gráfica, introducido por Szemerédi [Sze75b], es clave en la demostración original del lema de regularidad. Posteriormente este concepto fue generalizado por Bollobás y Nikiforov [BN08, p. 231], expuesto a continuación.

DEFINICIÓN 1.31. Fijemos un entero positivo  $k$ . Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\mathcal{S} \subseteq \mathcal{F}$  un  $k$ -semianillo,  $\mathcal{P} \subseteq \mathcal{S}$  una partición finita de  $\Omega$  y  $A \in \mathcal{F}$ . Definimos el *índice* de  $\mathcal{P}$  con respecto a  $A$  como

$$\text{ind}_A \mathcal{P} := \sum_{P \in \mathcal{P}} \mathbb{P}(P) \cdot (d(A, P))^2.$$

Una observación relevante en la demostración original del lema de regularidad es que el índice de una partición siempre está acotado.

OBSERVACIÓN 1.32. Fijemos un entero positivo  $k$ . Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\mathcal{S} \subseteq \mathcal{F}$  un  $k$ -semianillo,  $\mathcal{P} \subseteq \mathcal{S}$  una partición finita de  $\Omega$  y  $A \in \mathcal{F}$ . Entonces

$$\text{ind}_A \mathcal{P} \leq 1.$$

DEMOSTRACIÓN. Supongamos que se satisfacen las hipótesis del teorema. Entonces

$$\text{ind}_A \mathcal{P} = \sum_{P \in \mathcal{P}} \mathbb{P}(P) (d(A, P))^2 \leq \sum_{\substack{P \in \mathcal{P} \\ \mathbb{P}(P) > 0}} \frac{\mathbb{P}(A \cap P) \mathbb{P}(P)}{\mathbb{P}(P)} = \mathbb{P}(A) \leq 1,$$

donde usamos que el hecho  $d(A, P) \leq 1$  implica  $(d(A, P))^2 \leq d(A, P)$  para todo  $P \in \mathcal{P}$ .  $\square$

El argumento estándar en la prueba del lema de regularidad de Szemerédi es que, dada una partición  $\mathcal{P}$  de un conjunto no vacío  $\Omega$ , el índice de cualquier refinamiento,  $\mathcal{Q}$ , de  $\mathcal{P}$  siempre incrementa o es igual. La generalización, de gráficas a espacios de probabilidad, se encuentra en el siguiente lema, también debido a Bollobás y Nikiforov [BN08, p. 231].

LEMA 1.33. Sean  $k$  un entero positivo,  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\mathcal{S}$  un  $k$ -semianillo con  $\mathcal{S} \subseteq \mathcal{F}$ . Supongamos que  $\mathcal{P}$  y  $\mathcal{Q}$  son dos particiones finitas de  $\Omega$  tales que  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ . Entonces para cada  $A \in \mathcal{F}$  se cumple  $\text{índice}_A \mathcal{Q} \geq \text{índice}_A \mathcal{P}$ .

DEMOSTRACIÓN. La demostración puede consultarse en [BN08, p. 232].  $\square$

En términos informales, el siguiente lema nos dice que la ausencia de regularidad implica la existencia de un refinamiento con un incremento significativo en su índice.

LEMA 1.34. Fijemos un entero positivo  $k$ . Sea  $0 < \epsilon < 1$  y  $\mathcal{S}$  un  $k$ -semianillo. Si  $\mathcal{P}$  es una partición finita de  $\Omega$  tal que  $\mathcal{P} \subseteq \mathcal{S}$  y  $A \in \mathcal{F}$  no es  $\epsilon$ -regular en  $\mathcal{P}$ , entonces existe un refinamiento finito  $\mathcal{Q}$  de  $\mathcal{P}$  con las propiedades  $\mathcal{Q} \subseteq \mathcal{S}$  y

$$\text{índice}_A \mathcal{Q} \geq \text{índice}_A \mathcal{P} + \epsilon^4.$$

DEMOSTRACIÓN. De igual manera, la prueba puede consultarse en [BN08, p. 234].  $\square$

Un esbozo, informal, de la demostración original del lema de regularidad de Szemerédi es aprovechar el incremento de índices del Lema 1.34 y utilizar la cota del índice de una partición (Observación 1.32) para asegurar que no podemos continuar refinando una partición indefinidamente y tener, al mismo tiempo, ausencia de regularidad.

Para poder “regularizar” una variable aleatoria en el espacio  $L^2$ , Tao [Tao06] introduce una versión análoga al concepto de *índice* de una partición mediante  $\sigma$ -álgebras, y llama *energía* a esta nueva versión.

DEFINICIÓN 1.35. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Dada una variable aleatoria  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ , para cualquier  $\sigma$ -álgebra  $\mathcal{G} \subseteq \mathcal{F}$ , definimos la *energía* de  $\mathcal{G}$ ,  $\mathcal{E}(\mathcal{G})$ , como

$$\mathcal{E}(\mathcal{G}) := \|\mathbb{E}(X|\mathcal{G})\|_2^2.$$

El siguiente ejemplo exhibe la relación entre el índice de una partición  $\mathcal{P}$  del conjunto de vértices de una gráfica y la energía de la  $\sigma$ -álgebra generada por  $\mathcal{P}$ ,  $\sigma(\mathcal{P})$ .

EJEMPLO 1.36. Fijemos un entero positivo  $n$ . Consideremos el espacio de probabilidad discreto  $([n]^2, \mathcal{P}([n]^2), \mathbb{P})$ . Sea  $G = ([n], E)$  una gráfica y  $\mathcal{P}$  una partición de  $[n]$ . Entonces

$$\mathcal{E}(\sigma(\mathcal{P}^2)) := \|\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2)\|_2^2 = \text{índice}_E(\mathcal{P}^2).$$

DEMOSTRACIÓN. Sean  $n, p$  dos enteros positivos,  $G = ([n], E)$  una gráfica y  $\mathcal{P} = \{P_1, \dots, P_p\}$  una partición de  $[n]$  de cardinalidad  $p$ .

Al considerar el espacio de probabilidad discreto  $([n]^2, \mathcal{P}([n]^2), \mathbb{P})$ , cualquier subconjunto de  $[n]^2$  es un evento en nuestro espacio de probabilidad; en particular,  $E \in \mathcal{P}([n]^2)$ . Ya sabemos que  $\mathcal{P}^2$  es, en efecto, una partición de  $[n]^2$  (Observación 1.19). Además,

$$(11) \quad \mathbb{E}(\mathbb{1}_E|\mathcal{P}^2) \equiv \mathbb{E}(\mathbb{1}_E|P_{i_1} \times P_{i_2}) \text{ en cada } P_{i_1} \times P_{i_2} \in \mathcal{P}^2 \text{ con } i_1, i_2 \in [p],$$

como asegura el Ejemplo B.6.<sup>4</sup> Luego, para cualesquiera  $i_1, i_2 \in [p]$ ,

$$(12) \quad \begin{aligned} \mathbb{E}(\mathbb{1}_E|P_{i_1} \times P_{i_2}) &= \frac{1}{\mathbb{P}(P_{i_1} \times P_{i_2})} \mathbb{E}(\mathbb{1}_E \cdot \mathbb{1}_{P_{i_1} \times P_{i_2}}) \\ &= \frac{1}{\mathbb{P}(P_{i_1} \times P_{i_2})} \mathbb{P}(E \cap (P_{i_1} \times P_{i_2})) =: d(E, P_{i_1} \times P_{i_2}). \end{aligned}$$

Por un lado,

$$\|\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2)\|_2^2 = \int (\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2))^2 d\mathbb{P} = \sum_{i_1, i_2 \in [p]} \int_{P_{i_1} \times P_{i_2}} (\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2))^2 d\mathbb{P},$$

<sup>4</sup>Nuestra partición  $\mathcal{P}$  consiste de eventos de probabilidad positiva porque una partición es una familia de conjuntos no vacíos y  $\mathbb{P}$  es, en este caso, una medida de probabilidad uniforme.

donde en la última igualdad usamos que  $\mathcal{P}^2$  es una partición de  $[n]^2$ , junto con el Teorema A.32. Por otro lado, dado que  $\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2)$  es constante en cada  $P_{i_1} \times P_{i_2} \in \mathcal{P}^2$  por (11), y por la igualdad en (12), respectivamente,

$$\begin{aligned} \|\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2)\|_2^2 &= \sum_{i_1, i_2 \in [p]} \mathbb{P}(P_{i_1} \times P_{i_2}) (\mathbb{E}(\mathbb{1}_E|P_{i_1} \times P_{i_2}))^2 \\ &= \sum_{i_1, i_2 \in [p]} \mathbb{P}(P_{i_1} \times P_{i_2}) (d(E, P_{i_1} \times P_{i_2}))^2. \end{aligned}$$

Por lo tanto,

$$\|\mathbb{E}(\mathbb{1}_E|\mathcal{P}^2)\|_2^2 = \text{ind}_E(\mathcal{P}^2).$$

□

Observemos que el primer lema previo a nuestro teorema principal (Lema 1.29) es el análogo, en probabilidad, al argumento de incremento de energía de un refinamiento de una partición (Lema 1.33). Es decir, el Lema 1.29 asegura la existencia de un refinamiento  $\mathcal{R}$  con un incremento de energía significativo. Nuestro siguiente lema es similar al Lema 1.29, salvo que esta vez tenemos la opción de detenernos en algún momento (parte (b) del Lema 1.37), y esta parte es mostrada mediante martingalas, en términos exactos, la parte (b) del Lema 1.37 se debe a la desigualdad de Ricard y Xu [RX<sup>+</sup>16] que concierne a una sucesión de diferencias de martingala.

LEMA 1.37 (Segundo lema). *Fijemos un entero positivo  $k$  y números reales  $0 < \delta, \sigma \leq 1$ . Denotemos*

$$(13) \quad n = \left\lceil \frac{\sigma^2}{\delta^2} \right\rceil.$$

Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $(\Sigma_i)$  una sucesión creciente de  $k$ -semianillos en  $\Omega$  con  $\Sigma_i \subseteq \mathcal{F}$  para cada  $i \in \mathbb{N}$ . Además, sean  $m \in \mathbb{N}$  y  $\mathcal{P}$  una partición de  $\Omega$  con  $\mathcal{P} \subseteq \Sigma_m$  y  $|\mathcal{P}| \leq (k+1)^m$ . Entonces para cada variable aleatoria  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  existe  $j \in \{m, \dots, m+n\}$  y un refinamiento  $\mathcal{Q}$  de  $\mathcal{P}$  con  $\mathcal{Q} \subseteq \Sigma_j$  y  $|\mathcal{Q}| \leq (k+1)^j$ , y es tal que

- (a)  $\|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 > \sigma$ , o
- (b)  $\|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 \leq \sigma$  y  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma_{j+1}} \leq \delta$ .

Terence Tao prueba el Lema 1.37 con un *argumento de incremento de energía* [Tao06, p. 19], distinto a nuestro argumento con sucesiones de diferencias de martingala. A su vez, el argumento de incremento de energía es análogo al argumento de incremento de índice de una partición escrito originalmente por Szemerédi [Sze75a, p. 4]. En el fondo, lo que nos dice la parte (a) del Lema 1.37 es que un refinamiento de  $\mathcal{P}$  implica un incremento de energía, mientras que la parte (b) asegura que si  $\mathcal{Q}$  es un refinamiento finito de  $\mathcal{P}$  con la propiedad de que  $\mathbb{E}(X|\mathcal{Q})$  es cercana a  $\mathbb{E}(X|\mathcal{P})$ , entonces automáticamente  $\mathbb{E}(X|\mathcal{Q})$  está cerca de  $X$ . La cardinalidad de  $\mathcal{Q}$  depende directamente de qué tan cerca queremos estar de  $X$  y qué tan cerca queremos que estén  $\mathbb{E}(X|\mathcal{Q})$  y  $\mathbb{E}(X|\mathcal{P})$ .

DEMOSTRACIÓN DEL LEMA 1.37. Supongamos que la parte (a) no se satisface. Esto es equivalente a decir que

- H(1) para cada  $j \in \{m, \dots, m+n\}$ , para cada refinamiento  $\mathcal{Q}$  de  $\mathcal{P}$ , con  $\mathcal{Q} \subseteq \Sigma_j$ , y  $|\mathcal{Q}| \leq (k+1)^j$ , sucede que  $\|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 \leq \sigma$ .

Usaremos H(1) para mostrar que la parte (b) se satisface. Argumentando por contradicción, supongamos que (b) no se satisface. Sean  $j \in \{m, \dots, m+n\}$  arbitraria y  $\mathcal{Q}$  un refinamiento de  $\mathcal{P}$  con  $\mathcal{Q} \subseteq \Sigma_j$  y  $|\mathcal{Q}| \leq (k+1)^j$ . Observemos que la hipótesis H(1) y el supuesto de que (b) no se satisface implican que  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma_{j+1}} > \delta$ . Puesto que  $(\Sigma_i)$  es una sucesión creciente,  $\mathcal{Q} \subseteq \Sigma_{j+1}$ . Por el lema previo existe  $\mathcal{R}$  un refinamiento de  $\mathcal{Q}$  con  $\mathcal{R} \subseteq \Sigma_{j+1}$  y  $|\mathcal{R}| \leq |\mathcal{Q}|(k+1) \leq (k+1)^{j+1}$  y es tal que  $\|\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})\|_2 > \delta$ . Por la arbitrariedad de  $j \in \{m, \dots, m+n\}$  obtenemos que

H(2) para cada  $j \in \{m, \dots, m+n\}$ , para cada refinamiento  $\mathcal{Q}$  de  $\mathcal{P}$  con  $\mathcal{Q} \subseteq \Sigma_j$  y  $|\mathcal{Q}| \leq (k+1)^j$  existe un refinamiento  $\mathcal{R}$  de  $\mathcal{Q}$  con  $\mathcal{R} \subseteq \Sigma_{j+1}$  y  $|\mathcal{R}| \leq (k+1)^{j+1}$ , que satisface  $\|\mathbb{E}(X|\mathcal{R}) - \mathbb{E}(X|\mathcal{Q})\|_2 > \delta$ .

A manera de ilustrar el siguiente paso recursivo apliquemos la hipótesis (H2) al caso  $j = m$ . Ya que  $\mathcal{P}$  es un refinamiento de  $\mathcal{P}$  con  $\mathcal{P} \subseteq \Sigma_m$  y  $|\mathcal{P}| \leq (k+1)^m$ , H(2) nos dice que existe un refinamiento de  $\mathcal{P}$ , al que ahora denotamos por  $\mathcal{P}_1$ , que cumple que  $\mathcal{P}_1 \subseteq \Sigma_{m+1}$  y  $|\mathcal{P}_1| \leq (k+1)^{m+1}$ ; a su vez se satisface que  $\|\mathbb{E}(X|\mathcal{P}_1) - \mathbb{E}(X|\mathcal{P})\|_2 > \delta$ .

Es así como, recursivamente y usando la hipótesis H(2), obtenemos una sucesión finita de particiones de  $\Omega$  comenzando con la partición  $\mathcal{P}_0 = \mathcal{P}$  tales que para cada  $i \in [n]$  tenemos que

- (P1)  $\mathcal{P}_i$  es un refinamiento de  $\mathcal{P}_{i-1}$ ,
- (P2)  $\mathcal{P}_i \subseteq \Sigma_{m+i}$  y  $|\mathcal{P}_i| \leq (k+1)^{m+i}$ , y
- (P3)  $\|\mathbb{E}(X|\mathcal{P}_i) - \mathbb{E}(X|\mathcal{P}_{i-1})\|_2 > \delta$ .

Se sigue, en particular, que  $(\sigma(\mathcal{P}_i))_{i=1}^n$  es una sucesión creciente de sub- $\sigma$ -álgebras finitas de  $\mathcal{F}$ .

Defina la variable  $Z := X - \mathbb{E}(X|\mathcal{P})$ . Puesto que  $Z$  es una combinación lineal de variables aleatorias en  $L^2(\Omega, \mathcal{F}, \mathbb{P})$ ,  $Z$  pertenece al espacio vectorial  $L^2(\Omega, \mathcal{F}, \mathbb{P})$ . Más aún, por la monotonía de las normas (Proposición A.50), sucede que  $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ . Consecuentemente, la sucesión finita  $\mathbb{E}(Z|\mathcal{P}_0), \dots, \mathbb{E}(Z|\mathcal{P}_n)$  es una martingala finita, pues es una martingala de la esperanza condicional (Ejemplo B.13) ya que  $(\sigma(\mathcal{P}_i))_{i=1}^n$  es una sucesión creciente de sub- $\sigma$ -álgebras. Para la martingala construida considere su sucesión de diferencias de martingala  $(d_i)_{i=0}^n$ . De manera que, por linealidad de la esperanza condicional (Proposición B.7) y dado que  $\mathcal{P}_0 \subseteq \mathcal{P}_i$  para todo  $i \in [n]$ , en virtud de la Proposición B.8,

$$\begin{aligned} d_i &= \mathbb{E}(Z|\mathcal{P}_i) - \mathbb{E}(Z|\mathcal{P}_{i-1}) \\ &= (\mathbb{E}(X|\mathcal{P}_i) - \mathbb{E}(X|\mathcal{P}_0)) - (\mathbb{E}(X|\mathcal{P}_{i-1}) - \mathbb{E}(X|\mathcal{P}_0)) \\ &= \mathbb{E}(X|\mathcal{P}_i) - \mathbb{E}(X|\mathcal{P}_{i-1}), \end{aligned}$$

para cada  $i \in [n]$ . También, notemos que  $d_0 := \mathbb{E}(Z|\mathcal{P}_0) = 0$ . A su vez, de la propiedad (P3) expuesta en esta demostración,  $\|d_i\|_2 > \delta$  para cualquier  $i \in [n]$ . Esto último implica:

$$n\delta^2 = \sum_{i=1}^n \delta^2 < \sum_{i=1}^n \|d_i\|_2^2.$$

De aquí obtenemos

$$n^{1/2}\delta < \left( \sum_{i=0}^n \|d_i\|_2^2 \right)^{1/2}.$$

Por otro lado, de la definición de  $n$ ,

$$\sigma \leq n^{1/2}\delta.$$

De manera que, juntando las dos desigualdades previas,

$$\sigma < \left( \sum_{i=0}^n \|d_i\|_2^2 \right)^{1/2}.$$

Finalmente, para concluir la cadena de desigualdades, hacemos uso de nuestra desigualdad principal (Proposición B.16) en lo anterior para mostrar que

$$(14) \quad \sigma < \left\| \sum_{i=0}^n d_i \right\|_2 = \|\mathbb{E}(X|\mathcal{P}_n) - \mathbb{E}(X|\mathcal{P}_0)\|_2.$$

Sin embargo, por las propiedades (P1) y (P2) de esta demostración, vemos que  $\mathcal{P}_n$  es un refinamiento de  $\mathcal{P}$  con  $\mathcal{P}_n \subseteq \Sigma_{m+n}$  y  $|\mathcal{P}_n| \leq (k+1)^{m+n}$ . Por lo tanto, por (H1),  $\|\mathbb{E}(X|\mathcal{P}_n) - \mathbb{E}(X|\mathcal{P}_0)\|_2 \leq \sigma$ , lo que contradice (14). Esto concluye con la demostración.  $\square$

LEMA 1.38 (Tercer lema). *Fijemos un entero positivo  $k$ , un número real  $0 < \sigma \leq 1$ , y  $H : \mathbb{N} \rightarrow \mathbb{R}$  una función de crecimiento. Sea  $L = \lceil \sigma^{-2} \rceil$ . Definimos recursivamente una sucesión  $(n_i)$  mediante la regla*

$$(15) \quad \begin{cases} n_0 = 0, \\ n_{i+1} = n_i + \lceil \sigma^2 H(n_i)^2 \rceil. \end{cases}$$

Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $(\Sigma_i)$  una sucesión creciente de  $k$ -semianillos en  $\Omega$  con  $\Sigma_i \subseteq \mathcal{F}$  para cada  $i \in \mathbb{N}$ . Finalmente, sea  $X$  una variable aleatoria en  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  tal que  $\|X\|_2 \leq 1$ . Entonces existen dos enteros  $j \in \{0, \dots, L-1\}$ ,  $J \in \{n_j, \dots, n_{j+1}\}$  y dos particiones,  $\mathcal{P}$  y  $\mathcal{Q}$ , de  $\Omega$  con las siguientes propiedades:

- (i)  $\mathcal{P} \subseteq \Sigma_{n_j}$  y  $\mathcal{Q} \subseteq \Sigma_J$ ,
- (ii)  $|\mathcal{P}| \leq (k+1)^{n_j}$  y  $|\mathcal{Q}| \leq (k+1)^J$ ,
- (iii)  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ ,
- (iv)  $\|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 \leq \sigma$  y  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma_{J+1}} \leq 1/H(n_j)$ .

La demostración del Lema 1.38 es similar a la demostración del Lema 1.37, procediendo por contradicción. Concretamente, aplicamos el Lema 1.37 recursivamente para obtener una martingala finita, posteriormente obtenemos una sucesión de diferencias de martingala y aplicamos nuestra desigualdad principal (Proposición B.16) para llegar a la contradicción buscada.

DEMOSTRACIÓN DEL LEMA 1.38. Suponga que el lema es falso. Es decir, una vez establecidas las hipótesis del presente lema, para una variable aleatoria con las condiciones ya mencionadas, para cualesquiera enteros  $j \in \{0, \dots, L-1\}$ ,  $J \in \{n_j, \dots, n_{j+1}\}$  y particiones  $\mathcal{P}$  y  $\mathcal{Q}$ , no se cumple al menos una de las condiciones (i), (ii), (iii) o (iv). Sabemos que, en particular, el lema no se cumple para los números naturales  $j = 0$  y  $J_0 = n_0 = 0$  (en este caso coinciden). Luego, puesto que  $H$  es una función de crecimiento,  $1 \leq H(n_0)$ , por lo que, para  $\delta_0 := H(n_0)^{-1}$ , se cumple  $0 < \delta_0 \leq 1$ . Hasta el momento los números  $k \in \mathbb{Z}_{>0}$ ,  $0 < \delta_0, \sigma \leq 1$  están dados. Defina también

$$\mu_0 = \left\lceil \frac{\sigma^2}{\delta_0^2} \right\rceil.$$

Observe que contamos con las hipótesis necesarias para poder aplicar el Lema 1.37. Para aplicarlo, considere nuevamente el número natural  $n_0$  y la partición trivial  $\mathcal{P}_0 = \{\Omega\}$  contenida en  $\Sigma_0$ , con  $|\mathcal{P}_0| \leq (k+1)^0$ .

Entonces, dado que  $n_0 + \mu_0 = n_1$ , el Lema 1.37 asegura la existencia de un número  $J_1 \in \{n_0, \dots, n_1\}$  y  $\mathcal{P}_1$  un refinamiento de  $\mathcal{P}_0$ , con  $\mathcal{P}_1 \subseteq \Sigma_{J_1}$  y cardinalidad acotada por  $|\mathcal{P}_1| \leq (k+1)^{J_1}$ , que satisface

- (a)  $\|\mathbb{E}(X|\mathcal{P}_1) - \mathbb{E}(X|\mathcal{P}_0)\|_2 > \sigma$ , o
- (b)  $\|\mathbb{E}(X|\mathcal{P}_1) - \mathbb{E}(X|\mathcal{P}_0)\|_2 \leq \sigma$  y  $\|X - \mathbb{E}(X|\mathcal{P}_1)\|_{\Sigma_{J_1+1}} \leq \delta_0$ .

Hemos mostrado que para  $j = 0$ ,  $J_1 \in \{n_0, \dots, n_1\}$ , las particiones  $\mathcal{P}_0$  y  $\mathcal{P}_1$  satisfacen (i), (ii) y (iii) del presente lema. Como estamos asumiendo que el lema es falso, deducimos que (iv) no se satisface. Entonces la condición (b) no se satisface. En consecuencia, (a) se satisface.

Antes de proceder mediante un argumento recursivo, hacemos hincapié sobre el siguiente paso. Note que  $J_1 \leq n_1$  y que, al ser  $(\Sigma_i)$  una sucesión creciente de  $k$ -semianillos,  $\mathcal{P}_1 \subseteq \Sigma_{n_1}$  y  $|\mathcal{P}_1| \leq (k+1)^{n_1}$ . Consecuentemente, un razonamiento análogo nos da un número natural  $J_2 \in \{n_1, \dots, n_2\}$ ,  $\mathcal{P}_2$  un refinamiento de  $\mathcal{P}_1$  con  $\mathcal{P}_2 \subseteq \Sigma_{J_2}$  y  $|\mathcal{P}_2| \leq (k+1)^{J_2}$  tal que

$$(16) \quad \|\mathbb{E}(X|\mathcal{P}_2) - \mathbb{E}(X|\mathcal{P}_1)\|_{\Sigma_{J_2+1}} > \sigma.$$

Así, recursivamente y usando el Lema 1.37 seleccionamos una sucesión finita de números  $J_0, \dots, J_L \in \mathbb{N}$ , una sucesión finita  $\mathcal{P}_0, \dots, \mathcal{P}_L$  de particiones en  $\Omega$  tales que para cada  $i \in [L]$  tenemos:

1.  $J_i \in \{n_{i-1}, \dots, n_i\}$ ,
2.  $\mathcal{P}_i$  es un refinamiento de  $\mathcal{P}_{i-1}$ ,

3.  $\mathcal{P}_i \subseteq \Sigma_{J_i}$  con  $|\mathcal{P}_i| \leq (k+1)^{J_i}$ , y
4.  $\|\mathbb{E}(X|\mathcal{P}_i) - \mathbb{E}(X|\mathcal{P}_{i-1})\|_2 > \sigma$ .

De manera similar a la demostración del Lema 1.37 observemos que  $(\sigma(\mathcal{P}_i))_{i=0}^L$  es una sucesión creciente de sub- $\sigma$ -álgebras finitas de  $\mathcal{F}$ .

Es un hecho que la sucesión finita  $\mathbb{E}(X|\mathcal{P}_0), \dots, \mathbb{E}(X|\mathcal{P}_L)$  es una martingala, comúnmente conocida como martingala de la esperanza condicional (Ejemplo B.13). Consideremos la sucesión finita de diferencias de martingala asociada  $(d_i)_{i=0}^L$ .

Por la propiedad 4, se satisface  $\|d_i\|_2 > \sigma$  para cualquier  $i \in [L]$ . De modo que

$$\sigma^2 L < \sum_{i=0}^L \|d_i\|_2^2.$$

Esto implica una desigualdad relevante, a saber,

$$(17) \quad \sigma L^{1/2} < \left( \sum_{i=0}^L \|d_i\|_2^2 \right)^{1/2}.$$

Por otro lado, de la definición de  $L$ ,  $\sigma^{-2} \leq L$ . En consecuencia,

$$(18) \quad 1 \leq \sigma L^{1/2}.$$

De esta última desigualdad junto con (17),

$$1 < \left( \sum_{i=0}^L \|d_i\|_2^2 \right)^{1/2}.$$

Y de la Proposición B.16,

$$1 < \left( \sum_{i=0}^L \|d_i\|_2^2 \right)^{1/2} \leq \left\| \sum_{i=0}^L d_i \right\|_2 = \|\mathbb{E}(X|\mathcal{P}_L)\|_2.$$

Por último, como la esperanza condicional es una contracción en  $L^2$  (Teorema B.9) y por el hecho  $\|X\|_2 \leq 1$ , deducimos

$$1 < \|\mathbb{E}(X|\mathcal{P}_L)\|_2 \leq \|X\|_2 \leq 1.$$

Contradicción que concluye con la demostración.  $\square$

Notemos que en el Lema 1.38 describimos el exponente en la cota de la cardinalidad de la partición obtenida. Tal exponente depende de una función de crecimiento,  $F$ , del error de aproximación,  $\sigma$ , y de una sucesión de números naturales,  $(n_i)$ . Es por ello que es relevante detenernos en definir este exponente como  $Reg(\sigma, F)$ . En palabras,  $Reg(\sigma, F)$  dicta en a lo más cuántos pasos tendremos una aproximación, menor a  $\sigma$ , a una variable aleatoria en  $L^2$ . Esta aproximación depende de qué tan rápido queremos aproximarnos con nuestra función  $F$  y con qué precisión, dada por  $\sigma$ .

NOTACIÓN 1.39. Para cada  $0 < \sigma \leq 1$  y cada función de crecimiento  $F : \mathbb{N} \rightarrow \mathbb{N}$ , considere la sucesión  $(n_i)$  definida como

$$\begin{cases} n_0 = 0, \\ n_{i+1} = n_i + \lceil \sigma^2 F^{\circ(n_i+2)}(0)^2 \rceil. \end{cases}$$

Con la finalidad de establecer una notación más sencilla, denotemos por  $r$  al número natural

$$(19) \quad r = \left\lceil \frac{1}{\sigma^2} \right\rceil - 1.$$

Definimos

$$Reg(\sigma, F) := F^{\circ n_r}(0).$$

### 4.3. Resultado principal.

TEOREMA 1.40. *Fijemos un entero positivo  $k$ , y un número real  $0 < \sigma \leq 1$ . Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $(\mathcal{S}_i)$  una sucesión creciente de  $k$ -semianillos en  $\Omega$  con  $\mathcal{S}_i \subseteq \mathcal{F}$  para cada  $i \in \mathbb{N}$ . Sean, también,  $F : \mathbb{N} \rightarrow \mathbb{N}$  una función de crecimiento y  $X$  una variable aleatoria en  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  con la propiedad  $\|X\|_2 \leq 1$ . Entonces existen*

- (T1) *un número natural  $N \in \mathbb{N}$  que satisface  $N \leq \text{Reg}(\sigma, F)$ ,*
- (T2) *una partición  $\mathcal{P}$  de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{S}_N$  y  $|\mathcal{P}| \leq (k+1)^N$ , y*
- (T3) *un refinamiento finito  $\mathcal{Q}$  de  $\mathcal{P}$ , con  $\mathcal{Q} \subseteq \mathcal{S}_i$  para alguna  $i \geq N$*

que cumple que, al definir  $X = X_{\text{est}} + X_{\text{err}} + X_{\text{unf}}$ , donde

$$(20) \quad X_{\text{est}} = \mathbb{E}(X|\mathcal{P}), \quad X_{\text{err}} = \mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P}) \quad \text{y} \quad X_{\text{unf}} = X - \mathbb{E}(X|\mathcal{Q}),$$

tenemos las estimaciones

$$(21) \quad \|X_{\text{err}}\|_2 \leq \sigma \quad \text{y} \quad \|X_{\text{unf}}\|_{\mathcal{S}_i} \leq \frac{1}{F(i)},$$

para cada  $i \in \{0, \dots, F(N)\}$ .

DEMOSTRACIÓN. Supongamos que se satisfacen las hipótesis del presente teorema. Definamos, para cada  $i \in \mathbb{N}$ ,

$$m_i := F^{\circ i}(0).$$

La Observación 1.26 muestra que  $F^{\circ i}(0) < F(F^{\circ i}(0))$  para cualquier  $i \in \mathbb{N}$ , por lo que  $(m_i)$  es una sucesión creciente. Definamos también una sucesión de semianillos  $(\Sigma_i)$  como

$$\Sigma_i := \mathcal{S}_{m_i},$$

para cada  $i \in \mathbb{N}$ . Dado que  $(m_i)$  y  $(\mathcal{S}_i)$  son sucesiones crecientes de números reales y de  $k$ -semianillos, respectivamente, entonces  $(\Sigma_i)$  es una sucesión creciente de  $k$ -semianillos de  $\Omega$ , con la propiedad  $\Sigma_i \subseteq \mathcal{F}$  para todo número natural  $i$ .

Consideremos la función  $H : \mathbb{N} \rightarrow \mathbb{N}$  definida mediante la regla  $H(n) = F^{\circ(n+2)}(0)$ . El Ejemplo 1.27 afirma que  $H$  es una función de crecimiento. Dados los números ya fijos  $k, \sigma$ , la función de crecimiento  $H$  y la sucesión  $(\Sigma_i)$ , considere el entero  $L = \lceil \sigma^{-2} \rceil$  y la sucesión  $(n_i)$  definida como

$$\begin{cases} n_0 = 0, \\ n_{i+1} = n_i + \lceil \sigma^2 H(n_i)^2 \rceil. \end{cases}$$

Observemos que las condiciones están establecidas para aplicar el Lema 1.38 a lo anterior. En consecuencia, existen números enteros  $j \in \{0, \dots, L-1\}$ ,  $J \in \{n_j, \dots, n_{j+1}\}$ , y particiones finitas,  $\mathcal{P}$  y  $\mathcal{Q}$ , de  $\Omega$  que cumplen las siguientes condiciones:

- (i)  $\mathcal{P} \subseteq \Sigma_{n_j}$ ,  $\mathcal{Q} \subseteq \Sigma_J$ ,
- (ii)  $|\mathcal{P}| \leq (k+1)^{n_j}$ ,  $|\mathcal{Q}| \leq (k+1)^J$ ,
- (iii)  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ ,
- (iv)  $\|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 \leq \sigma$  y  $\|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma_{J+1}} \leq 1/H(n_j)$ .

Denotemos

$$(22) \quad N := m_{n_j} = F^{\circ n_j}(0),$$

por lo que  $\Sigma_{n_j} = \mathcal{S}_N$ . Como la notación lo sugiere, afirmamos que  $N, \mathcal{P}$  y  $\mathcal{Q}$  son los buscados. Debido a que  $j \leq L-1$  y  $(n_i)$  es una sucesión creciente,  $n_j \leq n_{L-1} = n_r$ , con  $r = \lceil \sigma^{-2} \rceil - 1$ . Puesto que  $(m_i)$  es una sucesión creciente,

$$N = F^{\circ n_j}(0) \leq F^{\circ n_r}(0) =: \text{Reg}(\sigma, F).$$

Además, ya que  $F$  es una función de crecimiento,

$$n_j \leq n_j + 1 \leq F^{\circ n_j}(0) = N.$$

Esto implica, de la parte (ii),  $|\mathcal{P}| \leq (k+1)^N$  y, de la parte (i) junto con la observación  $\Sigma_{n_j} = \mathcal{S}_N$ ,  $\mathcal{P} \subseteq \mathcal{S}_N$ . Consecuentemente,  $N$  y  $\mathcal{P}$  satisfacen las condiciones (T1) y (T2) en el enunciado del presente teorema. Finalmente, por la parte (iii), ya sabemos que  $\mathcal{Q}$  es un refinamiento de  $\mathcal{P}$ . Además,  $\mathcal{Q} \subseteq \mathcal{S}_i$  para algún número natural  $i$  mayor o igual a  $N$ , a saber, para el número natural  $m_J$ , ya que  $\mathcal{Q} \subseteq \Sigma_J = \mathcal{S}_{m_J}$  y  $N \leq m_J$  porque  $n_j \leq J$ .

Sea  $X = X_{\text{est}} + X_{\text{err}} + X_{\text{unf}}$ , donde

$$X_{\text{est}} = \mathbb{E}(X|\mathcal{P}), X_{\text{err}} = \mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P}) \text{ y } X_{\text{unf}} = X - \mathbb{E}(X|\mathcal{Q}).$$

De la parte (iv) en esta demostración,

$$(23) \quad \|X_{\text{err}}\|_2 = \|\mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P})\|_2 \leq \sigma.$$

Para acotar la norma de  $X_{\text{unf}}$ , notemos que  $n_j + 1 \leq J + 1$ . Por consiguiente,

$$F(N) = F(F^{\circ n_j}(0)) = m_{n_j+1} \leq m_{J+1}.$$

De modo que

$$\mathcal{S}_{F(N)} \subseteq \mathcal{S}_{m_{J+1}} = \Sigma_{J+1}.$$

Por lo tanto, también de la parte (iv) en esta demostración, para cada número natural  $i \in \{0, \dots, F(N)\}$ , tenemos

$$\|X_{\text{unf}}\|_{\mathcal{S}_i} = \|X - \mathbb{E}(X|\mathcal{Q})\|_{\mathcal{S}_i} \leq \|X - \mathbb{E}(X|\mathcal{Q})\|_{\Sigma_{J+1}} \leq \frac{1}{H(n_j)} = \frac{1}{F(F(N))} \leq \frac{1}{F(i)}.$$

□

Convengamos en la siguiente notación.

NOTACIÓN 1.41. *Para cada entero positivo  $k$ , cada  $0 < \sigma \leq 1$  y cada función de crecimiento  $F : \mathbb{N} \rightarrow \mathbb{N}$  definimos*

$$(24) \quad \text{Reg}'(k, \sigma, F) := (k+1)^{\text{Reg}(\sigma, F')},$$

donde  $F' : \mathbb{N} \rightarrow \mathbb{N}$  está definida por la regla  $F'(n) = F((k+1)^n)$  para cada  $n \in \mathbb{N}$ . El Ejemplo 1.24 asegura que  $F'$  es, en efecto, una función de crecimiento.

Necesitaremos una ligera modificación del Teorema 1.40 para demostrar el lema de regularidad de Szemerédi.

COROLARIO 1.42 (Lema fuerte de regularidad, versión probabilista). *Supongamos que tenemos un entero positivo  $k$ , un número real  $0 < \sigma \leq 1$  y una función de crecimiento  $F : \mathbb{N} \rightarrow \mathbb{N}$ . Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Finalmente, sea  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  con  $\|X\|_2 \leq 1$ . Entonces existen*

- (C1) un entero positivo  $M$  con  $M \leq \text{Reg}'(k, \sigma, F)$ ,
- (C2) una partición  $\mathcal{P}$  de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{S}$  y  $|\mathcal{P}| = M$ , y
- (C3) un refinamiento finito  $\mathcal{Q}$  de  $\mathcal{P}$  con  $\mathcal{Q} \subseteq \mathcal{S}$

con la propiedad de que, escribiendo  $X = X_{\text{est}} + X_{\text{err}} + X_{\text{unf}}$ , donde

$$X_{\text{est}} = \mathbb{E}(X|\mathcal{P}), X_{\text{err}} = \mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P}) \text{ y } X_{\text{unf}} = X - \mathbb{E}(X|\mathcal{Q}),$$

tenemos las estimaciones

$$\|X_{\text{err}}\|_2 \leq \sigma \quad \text{y} \quad \|X_{\text{unf}}\|_{\mathcal{S}} \leq \frac{1}{F(M)}.$$

DEMOSTRACIÓN. Supongamos que las hipótesis del Corolario están dadas. Considere la sucesión creciente, y constante, de  $k$ -semianillos  $(\mathcal{S})$  en  $\Omega$ . Aplicamos el Teorema 1.40 a las hipótesis establecidas, salvo que, en este caso, consideramos la función de crecimiento  $F' : \mathbb{N} \rightarrow \mathbb{N}$  con regla  $F'(n) = F((k+1)^n)$ . Entonces, dicho teorema asegura la existencia de

- (T1) un número natural  $N \in \mathbb{N}$  tal que  $N \leq \text{Reg}(\sigma, F')$ ,
- (T2) una partición  $\mathcal{P}$  de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{S}$  y  $|\mathcal{P}| \leq (k+1)^N$ , y

(T3) un refinamiento finito  $\mathcal{Q}$  de  $\mathcal{P}$ , con  $\mathcal{Q} \subseteq \mathcal{S}$

que cumple que, al definir  $X = X_{\text{est}} + X_{\text{err}} + X_{\text{unf}}$ , donde

$$(25) \quad X_{\text{est}} = \mathbb{E}(X|\mathcal{P}), \quad X_{\text{err}} = \mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P}) \quad \text{y} \quad X_{\text{unf}} = X - \mathbb{E}(X|\mathcal{Q}),$$

tenemos las estimaciones

$$(26) \quad \|X_{\text{err}}\|_2 \leq \sigma \quad \text{y} \quad \|X_{\text{unf}}\|_{\mathcal{S}} \leq \frac{1}{F'(i)},$$

para cada  $i \in \{0, \dots, F'(N)\}$ .

De la parte (T1) tenemos, siguiendo la definición en (24),

$$(k+1)^N \leq \text{Reg}'(k, \sigma, F).$$

Luego, por (T2),  $|\mathcal{P}| = M$  para alguna  $M \in \mathbb{N}$  acotada superiormente por  $\text{Reg}'(k, \sigma, F)$ , lo que muestra las afirmaciones (C1) y (C2) del presente Corolario. Finalmente, como  $M \leq (k+1)^N$  y  $F$  es, en particular, una función creciente, se cumple  $F(M) \leq F((k+1)^N)$ ; es decir,  $F(M) \leq F'(N)$ . Además,  $N \in \{0, \dots, F'(N)\}$  debido a que  $F'$  es una función de crecimiento. En consecuencia, de (26),

$$\|X_{\text{unf}}\|_{\mathcal{S}} \leq \frac{1}{F'(N)} \leq \frac{1}{F(M)},$$

como queríamos demostrar.  $\square$

## 5. Prueba del lema de regularidad en su versión de la teoría de gráficas

Volvamos ahora a la teoría de gráficas. Para formular el lema de regularidad de Szemerédi, como ya se mencionó en la introducción, necesitamos de algunas definiciones.

DEFINICIÓN 1.43. *Supongamos que  $G = (V, E)$  es una gráfica. Sean  $S$  y  $T$  dos subconjuntos de  $V$ . Definimos*

$$e_G(S, T) := |\{uv \in E \mid u \in S, v \in T\}|.$$

DEFINICIÓN 1.44. *Supongamos que  $G = (V, E)$  es una gráfica. Sean  $S, T$  subconjuntos de  $V$ . Definimos la **densidad de aristas entre  $S$  y  $T$**  como*

$$d_G(S, T) = \frac{e_G(S, T)}{|S||T|}.$$

De igual manera recordemos la definición de regularidad de una gráfica.

DEFINICIÓN 1.45 (Condición de regularidad). *Sean  $G$  una gráfica bipartita con bipartición  $\{V_1, V_2\}$  y  $\epsilon > 0$ . Decimos que  $G$  es  $\epsilon$ -regular si para cualesquiera  $A_1 \subseteq V_1$  y  $A_2 \subseteq V_2$  con la propiedad  $|A_1| > \epsilon|V_1|$  y  $|A_2| > \epsilon|V_2|$  se satisface*

$$|d_G(A_1, A_2) - d_G(V_1, V_2)| \leq \epsilon.$$

No obstante, Tao [Tao06] y Lovász [Lov12] utilizan una definición esencialmente equivalente, la de homogeneidad de una gráfica.

DEFINICIÓN 1.46. *Sean  $G$  una gráfica bipartita con bipartición  $\{V_1, V_2\}$  y  $\epsilon > 0$ . Decimos que  $G$  es  $\epsilon$ -homogénea si para cualesquiera  $A_1 \subseteq V_1$  y  $A_2 \subseteq V_2$  se satisface*

$$|e_G(A_1, A_2) - d_G(V_1, V_2) |A_1| |A_2|| \leq \epsilon |V_1| |V_2|,$$

Veamos que las condiciones de ser homogénea o regular en una gráfica son esencialmente equivalentes.

PROPOSICIÓN 1.47. *Sean  $G = (V, E)$  una gráfica bipartita con bipartición  $\{V_1, V_2\}$  y  $0 < \epsilon \leq 1$ . Entonces*

(a) *Si  $G$  es  $\epsilon$ -regular, entonces  $G$  es  $\epsilon$ -homogénea.*

(b) Si  $G$  es  $\epsilon^3$ -homogénea, entonces  $G$  es  $\epsilon$ -regular.

DEMOSTRACIÓN. (a): Supongamos que  $G$  es  $\epsilon$ -regular. Sean  $A_1 \subseteq V_1$  y  $A_2 \subseteq V_2$ . Si  $|A_1| > \epsilon|V_1|$  y  $|A_2| > \epsilon|V_2|$ , entonces, por hipótesis,

$$|e_G(A_1, A_2) - d_G(V_1, V_2) |A_1| |A_2|| \leq \epsilon |A_1| |A_2| \leq \epsilon |V_1| |V_2|,$$

pues  $|A_1| \leq |V_1|$  y  $|A_2| \leq |V_2|$ .

Supongamos ahora que se satisface al menos una de las condiciones:  $|A_1| \leq \epsilon|V_1|$  o  $|A_2| \leq \epsilon|V_2|$ . Como  $|A_1| \leq |V_1|$  y  $|A_2| \leq |V_2|$ , entonces  $|A_1| |A_2| \leq \epsilon |V_1| |V_2|$ . De modo que

$$d_G(V_1, V_2) |A_1| |A_2| \leq \epsilon \cdot e_G(V_1, V_2).$$

Luego, puesto que  $0 < \epsilon \leq 1$ ,  $\epsilon \cdot e_G(A_1, A_2) \leq e_G(A_1, A_2)$  implica  $-e_G(A_1, A_2) \leq -\epsilon \cdot e_G(A_1, A_2)$ . Por un lado,

$$\begin{aligned} d_G(V_1, V_2) |A_1| |A_2| - e_G(A_1, A_2) &\leq \epsilon \cdot e_G(V_1, V_2) - \epsilon \cdot e_G(A_1, A_2) \\ &= \epsilon(e_G(V_1, V_2) - e_G(A_1, A_2)) \\ &\leq \epsilon \cdot e_G(V_1, V_2) \\ &\leq \epsilon |V_1| |V_2|. \end{aligned}$$

Por otro lado,

$$\begin{aligned} e_G(A_1, A_2) - d_G(V_1, V_2) |A_1| |A_2| &= \frac{e_G(A_1, A_2) |V_1| |V_2| - e_G(V_1, V_2) |A_1| |A_2|}{|V_1| |V_2|} \\ &\leq \frac{e_G(A_1, A_2) (|V_1| |V_2| - |A_1| |A_2|)}{|V_1| |V_2|} \\ &\leq \frac{\epsilon |V_1| |V_2| |V_1| |V_2|}{|V_1| |V_2|} \\ &= \epsilon |V_1| |V_2|. \end{aligned}$$

Por lo tanto,

$$|e_G(A_1, A_2) - d_G(V_1, V_2) |A_1| |A_2|| \leq \epsilon |V_1| |V_2|.$$

Es decir,  $G$  es  $\epsilon$ -homogénea.

(b): Supongamos que  $G$  es  $\epsilon^3$ -homogénea. Sean  $A_1 \subseteq V_1$  y  $A_2 \subseteq V_2$  tales que  $\epsilon|V_1| < |A_1|$  y  $\epsilon|V_2| < |A_2|$ . Entonces

$$\begin{aligned} |e_G(A_1, A_2) - d_G(V_1, V_2) |A_1| |A_2|| &\leq \epsilon^3 |V_1| |V_2| = \epsilon \cdot \epsilon |V_1| \cdot \epsilon |V_2| \\ &\leq \epsilon |A_1| |A_2|, \end{aligned}$$

como queríamos demostrar.  $\square$

Haremos uso del concepto de particiones equitables.

DEFINICIÓN 1.48. Sean  $\epsilon > 0$ ,  $p$  un entero positivo y  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Supongamos que  $\mathcal{P} = \{P_0, \dots, P_p\}$  es una partición finita, de cardinalidad  $p + 1$ , de  $\Omega$  tal que  $\mathcal{P} \subseteq \mathcal{F}$ . Decimos que  $\mathcal{P}$  es  $\epsilon$ -**equitable** si  $\mathbb{P}(P_0) \leq \epsilon$  y  $\mathbb{P}(P_1) = \dots = \mathbb{P}(P_p) \leq \epsilon$ .

NOTACIÓN 1.49. Sea  $G = (V, E)$  una gráfica. Supongamos que  $V_1, V_2 \subseteq V$  son subconjuntos ajenos. Denotamos

$$G[V_1, V_2] := (V_1 \cup V_2, E \cap (V_1 \times V_2))$$

a la gráfica bipartita inducida por  $V_1, V_2$ .

El Corolario 1.42 implica algunas versiones del lema de regularidad de Szemerédi. Entre ellas, se encuentra el lema para gráficas bipartitas demostrado por Terence Tao en 2006 en el artículo [Tao06], donde Tao utiliza argumentos probabilísticos para su demostración. En esta tesis demostraremos la siguiente versión del lema de regularidad de Szemerédi que atañe a cualquier gráfica, cercana a la versión original [Sze75a]. Nuestra demostración está basada en la demostración de Terence Tao [Tao06] del lema de regularidad para gráficas bipartitas a partir del lema fuerte de regularidad en su versión probabilista.

TEOREMA 1.50 (Lema de regularidad de Szemerédi, versión casi original). *Para cada  $0 < \epsilon \leq 1$  existe un número natural  $N(\epsilon)$  con la propiedad de que si  $G = (V, E)$  es una gráfica de orden  $n \in \mathbb{Z}_{\geq N(\epsilon)}$ , entonces existe una partición  $\mathcal{T} = \{V_0, V_1, \dots, V_q\}$  de  $V$ , con  $1/\epsilon \leq q \leq N(\epsilon)$ , que satisface simultáneamente las siguientes propiedades:*

- (i) (Conjunto excepcional)  $|V_0| < \epsilon n$ .
- (ii) (Partición uniforme)  $|V_1| = \dots = |V_q|$ .
- (iii) (Regularidad) Cualesquiera  $i, j \in [q]$  con  $i \neq j$ , salvo una cantidad  $\epsilon q^2$  de pares  $(i, j)$ , satisfacen que la gráfica bipartita  $G[V_i, V_j]$  es  $\epsilon$ -homogénea.

El siguiente lema será útil en la demostración del Teorema 1.50.

LEMA 1.51. *Sean  $0 < \epsilon \leq 1$ ,  $F$  una función de crecimiento y  $n$  un entero positivo con la propiedad*

$$n > \frac{(1 + \epsilon)2^{2^{\text{Reg}'\left(4, \frac{\epsilon^{3/2}}{8(1+\epsilon)}, F\right)}}}{\epsilon} =: N(\epsilon).$$

*Consideremos el espacio de probabilidad discreto  $([n], \mathcal{P}([n]), \mathbb{P})$ . Supongamos que  $\mathcal{P}$  es una partición finita de  $[n]^2$  de cardinalidad  $M \leq \text{Reg}'\left(4, \frac{\epsilon^{3/2}}{8(1+\epsilon)}, F\right)$  y es tal que  $\mathcal{P} \subseteq \Sigma_{\square}$ , donde*

$$\Sigma_{\square} := \{S \times T \mid S \subseteq [n], T \subseteq [n] \text{ con } S \cap T = \emptyset \text{ o } S = T\}.$$

*Entonces existen un entero positivo  $q$  y  $\mathcal{T} = \{V_0, V_1, \dots, V_q\}$  una partición  $\epsilon$ -equitable de  $[n]$ , de cardinalidad  $|\mathcal{T}| = q + 1$ , con  $2^{2^M}/\epsilon \leq q \leq N(\epsilon)$ . Además,  $\mathcal{T}$  cumple*

$$|V_0| \leq \epsilon n, \quad |V_1| = \dots = |V_q| = \left\lfloor \frac{\epsilon n}{(1 + \epsilon)2^{2^M}} \right\rfloor,$$

*y, para cualesquiera  $i, j \in [q] \cup \{0\}$ ,  $V_i \times V_j \subseteq P$  para algún  $P \in \mathcal{P}$ .*

La demostración del Lema 1.51 puede encontrarse enseguida de la demostración del Teorema 1.50.

DEMOSTRACIÓN DEL TEOREMA 1.50. Sea  $0 < \epsilon \leq 1$ . Definamos una función  $F : \mathbb{N} \rightarrow \mathbb{N}$  mediante la regla

$$(27) \quad F(n) := \left\lceil \frac{(1 + \epsilon)^2 (2^{2n+2})^2}{\epsilon^3} \right\rceil,$$

para  $n \in \mathbb{N}$ . El Ejemplo 1.23 asegura que  $F$  es una función de crecimiento. Definamos

$$N(\epsilon) := \frac{(1 + \epsilon)2^{2^{\text{Reg}'\left(4, \frac{\epsilon^{3/2}}{8(1+\epsilon)}, F\right)}}}{\epsilon}.$$

Sean  $n$  un entero positivo con la propiedad  $n > N(\epsilon)$  y  $G = (V, E)$  una gráfica de orden  $n$ . Sin perder generalidad, supongamos que el conjunto de vértices de  $G$  es  $V = [n]$ . Aplicamos el Corolario 1.42 a la función de crecimiento  $F : \mathbb{N} \rightarrow \mathbb{N}$  dada por (27),  $\sigma = \frac{\epsilon^{3/2}}{8(1+\epsilon)}$ , el espacio de probabilidad discreto  $([n]^2, \mathcal{P}([n]^2), \mathbb{P})$ , el 4-semianillo

$$\Sigma_{\square} := \{S \times T \mid S \subseteq [n], T \subseteq [n] \text{ con } S \cap T = \emptyset \text{ o } S = T\}$$

y la variable aleatoria  $X = \mathbb{1}_E$ . Entonces el Corolario 1.42 afirma que existen

- (a) un entero positivo  $M$  con  $M \leq \text{Reg}'\left(4, \frac{\epsilon^{3/2}}{8(1+\epsilon)}, F\right)$ ;
- (b) una partición  $\mathcal{P}$  de  $[n]^2$  con  $\mathcal{P} \subseteq \Sigma_{\square}$  y  $|\mathcal{P}| = M$ ; y
- (c) un refinamiento finito,  $\mathcal{Q}$ , de  $\mathcal{P}$ , con  $\mathcal{Q} \subseteq \Sigma_{\square}$ , y tales que satisfacen las estimaciones

$$(28) \quad \|\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})\|_2 \leq \frac{\epsilon^{3/2}}{8(1 + \epsilon)}$$

y

$$(29) \quad \|\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E|\mathcal{Q})\|_{\Sigma_{\square}} \leq \frac{1}{F(M)} \leq \frac{\epsilon^3}{(1+\epsilon)^2(2^{2M+2})^2},$$

donde

$$\|\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E|\mathcal{Q})\|_{\Sigma_{\square}} := \sup \left\{ \left| \int_{S \times T} \mathbb{1}_E - \mathbb{E}(\mathbb{1}_E|\mathcal{Q}) d\mathbb{P} \right| : S \times T \in \Sigma_{\square} \right\}.$$

Sea  $\mathcal{P} = \{P_1, \dots, P_M\}$  la partición de  $[n]^2$  del inciso (b) y notemos la desigualdad que satisface  $M$  en el inciso (a). Entonces el Lema 1.51 garantiza la existencia de un entero positivo  $q$  y una partición  $\epsilon$ -equitable,  $\mathcal{T} = \{V_0, V_1, \dots, V_q\}$ , de  $[n]$ , de cardinalidad  $|\mathcal{T}| = q + 1$ . Más aún,

$$|V_0| \leq \epsilon n, \quad |V_1| = \dots = |V_q| = \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor,$$

con  $J := 2^{2M}/\epsilon$  y  $J \leq q$ .

Ahora, fijemos  $j_1, j_2 \in [q]$  con  $j_1 \neq j_2$ . Consideremos la gráfica bipartita inducida  $G[V_{j_1}, V_{j_2}] := (V_{j_1} \cup V_{j_2}, E \cap (V_{j_1} \times V_{j_2}))$ . Primero mostraremos que  $G[V_{j_1}, V_{j_2}]$  es  $\epsilon$ -homogénea; esto es, queremos mostrar

$$|e_G(A_1, A_2) - d_G(V_{j_1}, V_{j_2})|A_1||A_2|| \leq \epsilon|V_{j_1}||V_{j_2}|,$$

para cualesquiera  $A_1 \subseteq V_{j_1}$ ,  $A_2 \subseteq V_{j_2}$ . Es decir, buscamos que se cumpla

$$\left| |E \cap (A_1 \times A_2)| - \frac{|E \cap (V_{j_1} \times V_{j_2})|}{|V_{j_1} \times V_{j_2}|} |A_1 \times A_2| \right| \leq \epsilon \frac{n^2}{(1+\epsilon)^2 J^2},$$

para cualesquiera  $A_1 \subseteq V_{j_1}$  y  $A_2 \subseteq V_{j_2}$ . Para ello será suficiente encontrar una cantidad  $d$  independiente de  $A_1, A_2$  (pero que depende de  $E, V_{j_1}, V_{j_2}$ ) tal que

$$\left| |E \cap (A_1 \times A_2)| - d|A_1 \times A_2| \right| \leq \epsilon \frac{n^2}{(1+\epsilon)^2 J^2}.$$

Dividiendo entre  $|V \times V| = n^2$  podemos re-escribir lo anterior como

$$\left| \frac{|E \cap (A_1 \times A_2)|}{|V \times V|} - d \frac{|A_1 \times A_2|}{|V \times V|} \right| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2}.$$

Equivalentemente,

$$|\mathbb{P}(E \cap (A_1 \times A_2)) - d \cdot \mathbb{P}(A_1 \times A_2)| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2},$$

pues  $\mathbb{P}$  es la medida de probabilidad uniforme. Escribiendo la última desigualdad en términos de esperanza se obtiene

$$|\mathbb{E}(\mathbb{1}_{E \cap (A_1 \times A_2)}) - d \cdot \mathbb{E}(\mathbb{1}_{A_1 \times A_2})| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2}.$$

Recordemos que  $\mathbb{1}_{E \cap (A_1 \times A_2)} = \mathbb{1}_E \mathbb{1}_{A_1 \times A_2}$ , por ende,

$$|\mathbb{E}(\mathbb{1}_E \mathbb{1}_{A_1 \times A_2}) - d \cdot \mathbb{E}(\mathbb{1}_{A_1 \times A_2})| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2}.$$

Finalmente, por linealidad de la esperanza,

$$|\mathbb{E}((\mathbb{1}_E - d) \mathbb{1}_{A_1 \times A_2})| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2}.$$

Observemos que  $A_1 \times A_2$  está contenido en un sólo elemento de  $\mathcal{P}$ . Entonces podemos tomar  $d := \mathbb{E}(\mathbb{1}_E|\mathcal{P})$  en este elemento. Queremos mostrar entonces

$$|\mathbb{E}((\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E|\mathcal{P})) \mathbb{1}_{A_1 \times A_2})| \leq \frac{\epsilon}{(1+\epsilon)^2 J^2}.$$

*Demostración de la afirmación  $G[V_{j_1}, V_{j_2}]$  es  $\epsilon$ -homogénea:*

Observemos que  $A_1 \cap A_2 = \emptyset$ , por lo que  $A_1 \times A_2 \in \Sigma_{\square}$ . De (29) tenemos

$$|\mathbb{E}((\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E|\mathcal{Q})) \mathbb{1}_{A_1 \times A_2})| \leq \frac{\epsilon^3}{(1+\epsilon)^2(2^{2M+2})^2} \leq \frac{\epsilon}{8(1+\epsilon)^2 J^2}.$$

En consecuencia, para probar que  $G[V_{j_1}, V_{j_2}]$  es  $\epsilon$ -homogénea, es suficiente probar, por la desigualdad del triángulo y la desigualdad de Jensen,

$$\mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})| \mathbb{1}_{V_{j_1} \times V_{j_2}} \right) \leq \frac{\epsilon}{8(1+\epsilon)^2 J^2}.$$

Recordemos que de (28) tenemos

$$(30) \quad \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \right) = \|\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})\|_2^2 \leq \frac{\epsilon^3}{8^2(1+\epsilon)^2}.$$

Afirmamos que a lo más para  $\epsilon q^2$  pares de índices  $(i_1, i_2)$  se satisface

$$(31) \quad \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \mathbb{1}_{V_{i_1} \times V_{i_2}} \right) > \frac{\epsilon^2}{8^2(1+\epsilon)^2 J^2}.$$

Recordemos, al ser  $\mathcal{I}^2 = \{V_{i_1} \times V_{i_2} \mid i_1, i_2 \in [q] \cup \{0\}\}$  una partición de  $[n]^2$ ,

$$\sum_{i_1, i_2=0}^q \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \mathbb{1}_{V_{i_1} \times V_{i_2}} \right) = \mathbb{E} \left( (\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P}))^2 \right).$$

Consecuentemente, si para más de  $\epsilon q^2$  pares de índices  $(i_1, i_2)$  se satisface (31),

$$\mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \right) > \epsilon q^2 \cdot \frac{\epsilon^2}{8^2(1+\epsilon)^2 J^2} \geq \frac{\epsilon^3}{8^2(1+\epsilon)^2},$$

lo que contradice (30). Por lo tanto, para cualquier par  $(i_1, i_2)$ , salvo para  $\epsilon q^2$  pares de índices  $(i_1, i_2)$ , se cumple

$$(32) \quad \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \mathbb{1}_{V_{i_1} \times V_{i_2}} \right) \leq \frac{\epsilon^2}{8^2(1+\epsilon)^2 J^2}.$$

Supongamos que  $(j_1, j_2)$  es un par de índices que satisface la desigualdad (32). Observemos, además, que se cumple

$$\mathbb{E}(\mathbb{1}_{V_{j_1} \times V_{j_2}}) = \frac{|V_{j_1}| |V_{j_2}|}{n^2} \leq \frac{1}{(1+\epsilon)^2 J^2}.$$

De la desigualdad de Cauchy-Schwartz (Proposición A.49),

$$\begin{aligned} \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})| \mathbb{1}_{V_{j_1} \times V_{j_2}} \right) &\leq \sqrt{\mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})|^2 \mathbb{1}_{V_{j_1} \times V_{j_2}} \right) \mathbb{E}(\mathbb{1}_{V_{j_1} \times V_{j_2}})} \\ &\leq \sqrt{\frac{\epsilon^2}{8^2(1+\epsilon)^2 J^2} \cdot \frac{1}{(1+\epsilon)^2 J^2}} \\ &= \frac{\epsilon}{8(1+\epsilon)^2 J^2}. \end{aligned}$$

Por consiguiente,

$$\mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})| \mathbb{1}_{V_{j_1} \times V_{j_2}} \right) \leq \frac{\epsilon}{8(1+\epsilon)^2 J^2}.$$

Por último, de la desigualdad del triángulo, la desigualdad de Jensen (Teorema A.43), y dado que  $A_1 \subseteq V_{j_1}$  y  $A_2 \subseteq V_{j_2}$ ,

$$\begin{aligned} |\mathbb{E}((\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E | \mathcal{P})) \mathbb{1}_{A_1 \times A_2})| &\leq |\mathbb{E}((\mathbb{1}_E - \mathbb{E}(\mathbb{1}_E | \mathcal{Q})) \mathbb{1}_{A_1 \times A_2})| + \mathbb{E} \left( |\mathbb{E}(\mathbb{1}_E | \mathcal{Q}) - \mathbb{E}(\mathbb{1}_E | \mathcal{P})| \mathbb{1}_{V_{j_1} \times V_{j_2}} \right) \\ &\leq \frac{\epsilon}{8(1+\epsilon)^2 J^2} + \frac{\epsilon}{8(1+\epsilon)^2 J^2} \\ &= \frac{\epsilon}{4(1+\epsilon)^2 J^2}. \end{aligned}$$

Concluimos que

$$\left| \frac{|E \cap (A_1 \times A_2)|}{|V \times V|} - \frac{|E \cap (V_{j_1} \times V_{j_2})|}{|V_{j_1} \times V_{j_2}|} \cdot \frac{|A_1 \times A_2|}{|V \times V|} \right| \leq \frac{\epsilon}{4(1+\epsilon)^2 J^2}.$$

En otras palabras,

$$\left| |E \cap (A_1 \times A_2)| - \frac{|E \cap (V_{j_1} \times V_{j_2})|}{|V_{j_1} \times V_{j_2}|} \cdot |A_1 \times A_2| \right| \leq \frac{\epsilon}{4} \left( \frac{n}{(1+\epsilon)J} \right)^2.$$

Definamos

$$\alpha := \frac{n}{(1+\epsilon)J} - \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor = \frac{n}{(1+\epsilon)J} - |V_{j_1}|.$$

Observemos  $0 \leq \alpha < 1$ . De modo que

$$\left| |E \cap (A_1 \times A_2)| - \frac{|E \cap (V_{j_1} \times V_{j_2})|}{|V_{j_1} \times V_{j_2}|} \cdot |A_1 \times A_2| \right| \leq \frac{\epsilon}{4} (|V_{j_1}| + \alpha)^2.$$

Consecuentemente, como  $|V_{j_1}| + \alpha < 2|V_{j_1}|$ ,

$$\left| |E \cap (A_1 \times A_2)| - \frac{|E \cap (V_{j_1} \times V_{j_2})|}{|V_{j_1} \times V_{j_2}|} \cdot |A_1 \times A_2| \right| \leq \epsilon |V_{j_1}|^2 = \epsilon |V_{j_1}| |V_{j_2}|.$$

Por lo tanto,  $G[V_{j_1}, V_{j_2}]$  es  $\epsilon$ -homogénea.  $\square$

DEMOSTRACIÓN DEL LEMA 1.51. Fijemos un número real  $0 < \epsilon \leq 1$ , una función de crecimiento  $F : \mathbb{N} \rightarrow \mathbb{N}$  y un entero positivo  $n > N(\epsilon)$ . Supongamos que  $M$  es un entero positivo menor o igual a  $\text{Reg}'\left(4, \frac{\epsilon^{3/2}}{8(1+\epsilon)}, F\right)$ ,  $\mathcal{P} = \{P_1, \dots, P_M\}$  es una partición de  $[n]^2$  de cardinalidad  $M$  y es tal que  $\mathcal{P} \subseteq \Sigma_{\square}$ . Para cada  $i \in [M]$ ,  $P_i$  es de la forma  $P_i = S_{i_1} \times S_{i_2}$ , donde se cumple  $S_{i_1}, S_{i_2} \subseteq [n]$ , y  $S_{i_1} \cap S_{i_2} = \emptyset$  o  $S_{i_1} = S_{i_2}$ . Para cada  $i \in [M]$  y  $j \in [2]$ , definamos  $\mathcal{P}_{i_j} = \{S_{i_j}, [n] \setminus S_{i_j}\}$  si  $S_{i_j} \neq [n]$  y  $\mathcal{P}_{i_j} = \{S_{i_j}\}$  si  $S_{i_j} = [n]$ . Notemos que  $\mathcal{P}_{i_j}$  es una partición de  $[n]$ . De manera que, por la Observación 1.17, podemos definir una partición,  $\mathcal{R}$ , de  $[n]$  como

$$\mathcal{R} := \bigcap_{i \in [M], j \in [2]} \mathcal{P}_{i_j}.$$

Definamos también  $r := |\mathcal{R}|$ , y notemos  $r \leq 2^{2M}$ . Describamos a  $\mathcal{R}$  como  $\mathcal{R} = \{R_1, \dots, R_r\}$ . Es un hecho que podemos refinar a la partición  $\mathcal{R}$  en conjuntos de tamaño  $\lfloor \frac{n}{(1+\epsilon)J} \rfloor$  con  $J := 2^{2M}/\epsilon$ . Para precisar esta afirmación consideremos

$$\mathcal{A} := \left\{ R_i \in \mathcal{R} : |R_i| \geq \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor \right\} \text{ y } \mathcal{B} := \mathcal{R} \setminus \mathcal{A}.$$

Para ver que  $\mathcal{A}$  es un conjunto no vacío, supongamos que para todo  $R_i \in \mathcal{R}$ , con  $i \in [r]$ , se cumple

$$|R_i| < \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor.$$

Consecuentemente, como  $\mathcal{R}$  es una partición de  $[n]$ ,

$$n = \sum_{i=1}^r |R_i| < r \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor \leq r \cdot \frac{n\epsilon}{(1+\epsilon)2^{2M}} \leq \frac{\epsilon}{1+\epsilon} \cdot n \leq n,$$

un absurdo. Por lo tanto,  $\mathcal{A}$  es un conjunto no vacío.

Para cada  $i \in [r]$  y  $R_i \in \mathcal{R}$  existen únicos  $q_i, \ell_i \in \mathbb{Z}_{\geq 0}$  con la propiedad

$$|R_i| = \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor q_i + \ell_i, \text{ con } 0 \leq \ell_i < \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor.$$

Denotemos, para cada  $i \in [r]$ ,  $|R_i| = r_i$  y  $R_i = \{n_1^i, \dots, n_{r_i}^i\}$ , donde, para  $k, j \in [r_i]$ , suponemos  $n_k^i < n_j^i$  si  $k < j$ . Refinaremos, para cada  $i \in [r]$ , el elemento  $R_i \in \mathcal{R}$  como sigue. Definamos, para cada  $i \in [r]$  y para cada  $j \in [q_i]$ ,

$$R_j^i := \{n_{(j-1)\lfloor \frac{n}{(1+\epsilon)J} \rfloor + 1}^i, \dots, n_{j\lfloor \frac{n}{(1+\epsilon)J} \rfloor}^i\} \text{ y } R_0^i = R_i \setminus \bigcup_{j=1}^{q_i} R_j^i.$$

Notemos que  $R_0^i$  puede ser el conjunto vacío y que su cardinalidad es  $|R_0^i| = \ell_i$ . Finalmente definimos

$$V_0 := \bigcup_{i=1}^r R_0^i,$$

que también puede ser un conjunto vacío. Además, tenemos como una consecuencia inmediata,

$$|V_0| = \sum_{i=1}^r \ell_i < r \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor \leq \frac{\epsilon}{1+\epsilon} n \leq \epsilon n.$$

Consecuentemente, al definir  $q := \sum_{i=1}^r q_i$ , tenemos que  $\mathcal{S} = \{V_0\} \cup \{R_j^i \mid i \in [r], j \in [q_i]\}$  es la partición buscada, donde enfatizamos nuevamente que el conjunto  $V_0$  puede ser el conjunto vacío. Por último, verifiquemos que se satisface  $J = 2^{2^M}/\epsilon \leq q$ . La prueba de esta última afirmación es por contradicción. Supongamos  $q < J$ . Entonces, como  $|V_0| \leq \frac{\epsilon}{1+\epsilon}n$ ,

$$n \leq q \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor + \frac{\epsilon}{1+\epsilon}n < J \left\lfloor \frac{n}{(1+\epsilon)J} \right\rfloor + \frac{\epsilon}{1+\epsilon}n \leq \frac{n}{1+\epsilon} + \frac{\epsilon}{1+\epsilon}n = n,$$

una contradicción. Por lo tanto,  $J \leq q$ . □

## Elementos de la teoría de límites de gráficas

Este segundo capítulo está dedicado a mostrar los conceptos básicos de la teoría de límites de gráficas. Esta exposición está basada mayoritariamente en el libro “Large networks and graph limits” de László Lovász [Lov12]. Aquí se exponen también las distintas nociones de métricas en el espacio de gráficas; entre ellas, la que más destaca en el presente trabajo es la métrica de corte. Además, desarrollamos la noción de un grafón como un objeto límite de una sucesión de gráficas con número de vértices creciente, y su relación con gráficas de la teoría tradicional. Por último, se da un esbozo de la demostración del teorema de compacidad en el espacio de grafones, donde citamos el lema de regularidad de Szemerédi, en su versión débil del área de análisis, como una herramienta dentro de la prueba.

Aprovecho este espacio para agradecer a Pablo Hernández Hernández y a Santiago Guzmán Pro por permitirme discutir con ellos algunas secciones del libro “Large networks and graph limits”.

### 1. Conceptos básicos

DEFINICIÓN 2.1. Una **gráfica**  $G$  es un par ordenado  $(V, E)$  consistente de un conjunto finito y no vacío,  $V$ , al que llamaremos conjunto de vértices de  $G$ , y un conjunto  $E$ , llamado el conjunto de aristas de  $G$ , cuyos elementos son pares no ordenados  $(u, v)$  con  $u, v \in V$ .

NOTACIÓN 2.2. Denotaremos a una arista  $(u, v) \in E(G)$  por  $uv$  o, equivalentemente, por  $vu$ . Para enfatizar que  $V$  y  $E$  son el conjunto de vértices y aristas de una gráfica  $G$ , respectivamente, escribiremos  $V$  como  $V(G)$  y  $E$  como  $E(G)$ .

Un lazo es un par  $(v, v) \in E(G)$  para  $v \in V(G)$ ; en otras palabras, un lazo es una arista que une a un vértice consigo mismo.

Una gráfica  $G$  también es llamada una *gráfica con posibles lazos*, pues puede suceder  $vv \in E(G)$  para  $v \in V(G)$ .

DEFINICIÓN 2.3. Una gráfica con posibles lazos  $G$  es una **gráfica completa con lazos** si satisface  $(u, v) \in E(G)$  para cualesquiera  $u, v \in V(G)$ .

DEFINICIÓN 2.4. Supongamos que  $G$  es una gráfica. El **orden** de  $G$ , denotado por  $v(G)$ , es el número de vértices de  $G$ . Es decir,  $v(G) = |V(G)|$ .

DEFINICIÓN 2.5. Supongamos que  $G$  es una gráfica. El **tamaño** de  $G$ , denotado por  $e(G)$ , es el número de aristas de  $G$ . Esto es,  $e(G) = |E(G)|$ .

DEFINICIÓN 2.6. Una **gráfica simple**  $G$  es una gráfica con la propiedad  $(v, v) \notin E(G)$  para todo  $v \in V(G)$ .

Notemos que una gráfica simple es una gráfica sin lazos.

DEFINICIÓN 2.7. Una gráfica simple  $G$  es una **gráfica completa** si satisface que cualesquiera dos vértices distintos son adyacentes; es decir, para todo par  $u, v \in V(G)$  con  $u \neq v$ , se cumple  $uv \in E(G)$ .

NOTACIÓN 2.8. Sea  $n$  un entero positivo. Denotamos a una gráfica completa de orden  $n$  por  $K_n$ .

### 1.1. Gráficas con peso.

DEFINICIÓN 2.9. *Fijemos un entero positivo  $n$ . Una cuarteta  $H = (V(H), E(H), \alpha, \beta)$  es una **gráfica con pesos de orden  $n$**  si el par  $(V(H), E(H))$  es una gráfica con posibles lazos de orden  $n$ ; y  $\alpha$  y  $\beta$  son dos funciones  $\alpha : V \rightarrow \mathbb{R}^+$  y  $\beta : E \rightarrow \mathbb{R}$ .*

NOTACIÓN 2.10. *Sean  $n$  un entero positivo y  $H = (V(H), E(H), \alpha, \beta)$  una gráfica con pesos de orden  $n$ , donde  $V(H) = \{v_1, \dots, v_n\}$ . Denotamos*

$$\alpha_i(H) := \alpha(v_i) \quad \text{y} \quad \beta_{ij}(H) := \beta(v_i v_j),$$

para cualesquiera  $i, j \in [n]$ .

NOTACIÓN 2.11. *Sean  $n$  un entero positivo y  $H = (V(H), E(H), \alpha, \beta)$  una gráfica con pesos de orden  $n$ . Definimos*

$$\alpha_H := \sum_{i=1}^n \alpha_i(H).$$

Sean  $n$  un entero positivo y  $H' = (V(H'), E(H'), \alpha', \beta')$  una gráfica con pesos de orden  $n$  con conjunto de vértices  $V(H') = \{v_1, \dots, v_n\}$ . Consideremos una nueva gráfica con pesos  $H = (V(H), E(H), \alpha, \beta)$  definida como sigue. El conjunto de vértices de  $H$  es  $V(H) := V(H')$  y el conjunto de aristas es  $E(H) := V(H') \times V(H')$ . La función  $\alpha : V \rightarrow \mathbb{R}^+$  es  $\alpha := \alpha'$  y la función  $\beta$  tiene regla de correspondencia

$$\beta(v_i v_j) = \begin{cases} \beta'(v_i v_j) & \text{si } v_i v_j \in E(H') \\ 0 & \text{si } v_i v_j \notin E(H'). \end{cases}$$

De modo que, a partir de cualquier gráfica con pesos  $H' = (V(H'), E(H'), \alpha', \beta')$ , siempre podemos inducir una gráfica con pesos  $H = (V(H), E(H), \alpha, \beta)$  con la propiedad de que el par  $(V(H), E(H))$  es una gráfica completa con lazos. Es decir, en  $H'$ , las aristas faltantes en  $(V(H'), E(H'))$  (los pares  $uv \in (V(H') \times V(H')) \setminus E(H')$ ) pueden ser agregadas con peso cero.

La observación previa nos permite asumir que, en toda gráfica con pesos

$$H = (V(H), E(H), \alpha, \beta),$$

la gráfica  $(V(H), E(H))$  es una gráfica completa con lazos. Consecuentemente, una gráfica con pesos queda determinada por un vector  $a = (\alpha_1(H), \dots, \alpha_n(H))$  cuyas entradas son los pesos de los vértices y una matriz  $B = (\beta_{ij}(H))$  que tiene por entradas los pesos de las aristas. Además, notemos que para cualquier par  $i, j \in [n]$  se tiene la igualdad  $\beta_{ij}(H) = \beta_{ji}(H)$ , pues la arista  $v_i v_j \in E(H)$  es igual a la arista  $v_j v_i$ , así,  $B$  es una matriz simétrica con entradas reales. En otras palabras,  $H$  está completamente descrita por su orden  $v(H) = n$ , un vector real positivo  $a = (\alpha_1(H), \dots, \alpha_n(H)) \in \mathbb{R}^n$  de pesos de vértices y una matriz real simétrica  $B = (b_{ij}) \in \mathbb{R}^{n \times n}$  de pesos de aristas. Por esta razón, en adelante denotaremos a una gráfica con pesos,  $H$ , por  $H(a, B)$ .

### 1.2. Gráficas etiquetadas.

DEFINICIÓN 2.12. *Sean  $k$  y  $n$  dos enteros positivos que cumplen  $k \leq n$ . Supongamos que  $G$  es una gráfica de orden  $n$ . Decimos que una función inyectiva  $\varphi : [k] \rightarrow V(G)$  es un **etiquetamiento de  $G$** .*

DEFINICIÓN 2.13. *Sean  $k$  y  $n$  dos enteros positivos que cumplen  $k \leq n$ . Una **gráfica simple  $k$ -etiquetada** es una terna  $(V, E, \varphi)$  tal que el par ordenado  $(V, E)$  es una gráfica de orden  $n$  y  $\varphi : [k] \rightarrow V$  es un etiquetamiento de  $G$ .*

DEFINICIÓN 2.14. *Sean  $n$  un entero positivo,  $G$  una gráfica de orden  $n$  y  $\varphi : [n] \rightarrow V(G)$  un etiquetamiento de  $G$ . Definimos la gráfica  $G^\varphi$  como el par ordenado  $([n], E^\varphi(G))$ , donde el conjunto de aristas es*

$$E^\varphi(G) := \{\varphi^{-1}(v_i)\varphi^{-1}(v_j) \mid v_i v_j \in E(G)\}.$$

### 1.3. Operaciones con gráficas.

DEFINICIÓN 2.15. Supongamos que tenemos una gráfica simple  $G = (V, E)$  y fijemos un entero positivo  $m$ . Para cada  $i \in [m]$  considere los conjuntos  $V_i := \{i\} \times V$ . Definimos la  **$m$ -gráfica autoreplicada de  $G$** , denotada por  $G(m)$ , como la gráfica con conjunto de vértices  $V^m := \bigcup_{i=1}^m V_i$  y conjunto de aristas  $E^m := \{((i, j_1), (k, j_2)) \mid (j_1, j_2) \in E, i, k \in [m]\}$ .

OBSERVACIÓN 2.16. Dada una gráfica  $G$  y un entero positivo  $m$ , hacemos hincapié en que las aristas de la  $m$ -gráfica autoreplicada de  $G$ ,  $G(m)$ , son también pares no ordenados.

En términos informales, dada una gráfica  $G = (V(G), E(G))$  y un entero positivo  $m$ , la  $m$ -gráfica autoreplicada de  $G$  es la gráfica resultante de reemplazar cada vértice  $v \in V(G)$  por  $m$  copias de  $v$ . Las copias de dos vértices en  $G(m)$  son adyacentes si los vértices originales en  $G$  lo son.

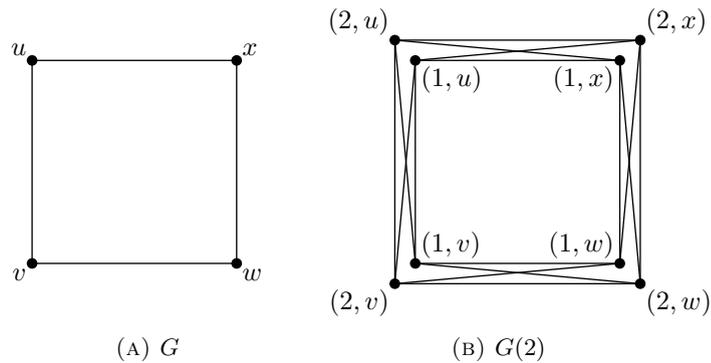


FIGURA 1. Ejemplo de una gráfica  $G$  y la 2-gráfica autoreplicada de  $G$ .

De este modo, dada una gráfica  $G$  y un entero positivo  $m$ , la  $m$ -gráfica autoreplicada<sup>1</sup> de  $G$ ,  $G(m)$ , tiene como número de vértices  $v(G(m)) = m \cdot v(G)$ . Además, por cada arista  $uv \in E(G)$  reemplazamos el vértice  $u$  por  $m$  copias de  $u$  y  $v$  por  $m$  copias de  $v$ , digamos,  $\{u_1, \dots, u_m\}$  y  $\{v_1, \dots, v_m\}$ . Luego, fijando  $i \in [m]$  arbitrario,  $u_i$  es adyacente a  $v_j$  para toda  $j \in [m]$ , lo que resulta en  $m$  aristas de  $G(m)$ . Sumando las aristas por cada  $i \in [m]$  obtenemos que  $e(G(m))$  es igual a  $m^2 \cdot e(G)$ . Hemos probado la siguiente observación.

OBSERVACIÓN 2.17. Supongamos que  $G$  es una gráfica con número de vértices  $v(G)$  y que  $m$  es un entero positivo. Entonces la  $m$ -gráfica autoreplicada de  $G$ ,  $G(m)$ , tiene  $m \cdot v(G)$  vértices y  $m^2 \cdot e(G)$  aristas.

DEFINICIÓN 2.18. Una sucesión de gráficas  $(G_n)$  es **densa** si existen  $n_0 \in \mathbb{N}$  y una constante  $c > 0$  tales que para todo  $n > n_0$  se satisface  $|E_n| \geq c|V_n|^2$ , donde  $G_n = (V_n, E_n)$ .

En la literatura de sucesiones de gráficas el término “denso” se refiere al orden de crecimiento de número de aristas de las gráficas, en lugar de la densidad de la sucesión en el sentido topológico.

En una gráfica  $G = (V(G), E(G))$ , de orden  $n$ , el número de posibles aristas es  $e(G) = n(n-1)/2$ , que es asintóticamente equivalente a  $n^2/2$ . Entonces estamos pidiendo que para un número natural  $n$  suficientemente grande,  $E$  sea, en proporción, mayor o igual que  $n^2/2$  (el número de aristas en una gráfica completa de orden  $n$ ).

Un ejemplo trivial de una sucesión de gráficas densa es la sucesión de gráficas que son completas,  $(K_n)$ .

<sup>1</sup>En inglés:  **$m$ -blowup**.

**1.4. Matriz de adyacencia.** Aunque la representación visual de una gráfica resulta útil, hay ocasiones en las que se necesita de una representación más algebraica. Por ejemplo, la matriz de adyacencia de una gráfica. Dicha matriz será de utilidad para definir la distancia entre dos gráficas.

DEFINICIÓN 2.19. Supongamos que  $G = (V, E)$  es una gráfica de orden  $n$  con conjunto de vértices  $V = \{v_1, \dots, v_n\}$ . La **matriz de adyacencia** de  $G$  es una matriz  $A_G = (a_{ij})$  de tamaño  $n \times n$  definida como

$$a_{ij} = \begin{cases} 1 & \text{si } v_i v_j \in E \\ 0 & \text{si } v_i v_j \notin E, \end{cases}$$

para  $i, j \in [n]$ .

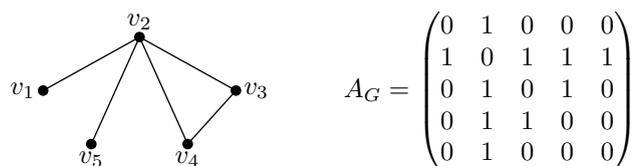


FIGURA 2. Una gráfica y su matriz de adyacencia

## 2. Densidad de homomorfismos

DEFINICIÓN 2.20. Supongamos que  $G$  y  $H$  son dos gráficas simples. Un **homomorfismo de  $G$  en  $H$**  es una función  $f : V(G) \rightarrow V(H)$  que satisface  $f(u)f(v) \in E(H)$  siempre que  $uv \in E(G)$ .

De manera que un homomorfismo es una función que preserva adyacencias.

NOTACIÓN 2.21. Definamos al conjunto de homomorfismos de  $G$  en  $H$  como  $\text{Hom}(G, H)$ . En símbolos,

$$\text{Hom}(G, H) = \{f : V(G) \rightarrow V(H) \mid f \text{ es un homomorfismo}\}.$$

La notación estándar es denotar a la cardinalidad de  $\text{Hom}(G, H)$  por  $\text{hom}(G, H)$ .

Dado cualquier par de conjuntos finitos  $A$  y  $B$ , recordemos que el conjunto de todas las funciones de  $A$  en  $B$  es denotado por  $B^A$ , existe y tiene cardinalidad  $|B|^{|A|}$ . Con esto en mente, el conjunto de todas las funciones de  $V(G)$  en  $V(H)$ , denotado por  $V(H)^{V(G)}$ , tiene cardinalidad  $v(H)^{v(G)}$ . Así, en un espacio de probabilidad discreto  $(\Omega, \mathcal{F}, \mathbb{P})$ , donde el espacio muestral es  $\Omega = V(H)^{V(G)}$ , la probabilidad de escoger una función al azar  $f \in V(H)^{V(G)}$  y ésta sea un homomorfismo es precisamente:

$$\frac{\text{hom}(G, H)}{v(H)^{v(G)}}.$$

Lo anterior motiva la siguiente definición.

DEFINICIÓN 2.22. Sean  $k$  y  $n$  dos enteros positivos. Sean, también,  $G$  y  $H$  gráficas simples de ordenes  $k$  y  $n$ , respectivamente. Definimos

$$t(G, H) = \frac{\text{hom}(G, H)}{n^k}$$

como la **densidad de homomorfismos de  $G$  en  $H$** .

Consideremos nuevamente dos gráficas simples  $G$  y  $H$ . Decimos que  $f : V(G) \rightarrow V(H)$  es un **isomorfismo** si  $f$  es un homomorfismo y  $f$  es biyectiva.

### 3. Algunas relaciones entre grafones y gráficas

DEFINICIÓN 2.23. Una función  $W : [0, 1]^2 \rightarrow \mathbb{R}$  es **simétrica** si para cualesquiera  $x, y \in [0, 1]$  se cumple  $W(x, y) = W(y, x)$ .

DEFINICIÓN 2.24. Una función  $W : [0, 1]^2 \rightarrow \mathbb{R}$  es un **núcleo** si es acotada,  $(\mathcal{B}([0, 1]^2), \mathcal{B}(\mathbb{R}))$ -medible y simétrica. Denotamos por  $\mathcal{W}$  al espacio de núcleos.

Una primera consecuencia es que todo núcleo es integrable.

PROPOSICIÓN 2.25. Consideremos el espacio de probabilidad  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ . Cualquier núcleo  $W \in \mathcal{W}$  pertenece a  $L^1([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ .

DEMOSTRACIÓN. Sea  $W \in \mathcal{W}$ . Entonces  $W$  está acotado; esto es, existe  $M \in \mathbb{R}$  tal que  $|W| \leq M$ . Por consiguiente, por la propiedad de monotonía (parte (i) de la Proposición A.33),

$$\int_{[0, 1]^2} |W| d\lambda^2 \leq \int_{[0, 1]^2} M d\lambda^2 = \lambda^2([0, 1]^2) M = M < \infty.$$

Por lo tanto,  $W \in L^1([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ .  $\square$

DEFINICIÓN 2.26. Una función  $W : [0, 1]^2 \rightarrow \mathbb{R}$  es un **grafón** si  $W$  es un núcleo y además, para cualesquiera  $x, y \in [0, 1]$ , cumple  $0 \leq W(x, y) \leq 1$ . Denotamos por  $\mathcal{W}_0$  al espacio de grafones.

El nombre grafón viene de una contracción de los términos gráfica y función en una sola palabra. La idea detrás de este nombre se debe a que los grafones son, en cierto sentido que se verá más adelante, una generalización de las gráficas.

DEFINICIÓN 2.27. Un núcleo  $W \in \mathcal{W}$  es una **función escalonada** si existen un entero positivo  $k$  y una partición finita  $\mathcal{Q} = \{Q_1, \dots, Q_k\}$  de  $[0, 1]$ , con cardinalidad  $k$ , consistente de conjuntos medibles, esto es,  $\mathcal{Q} \subseteq \mathcal{B}([0, 1])$ , con la propiedad de que, para todo par  $i, j \in [k]$ , existe  $c_{ij} \in \mathbb{R}$  tal que para todo  $(x, y) \in Q_i \times Q_j$  se cumple  $W(x, y) = c_{ij}$ .

NOTACIÓN 2.28. Supongamos que  $W \in \mathcal{W}$  es una función escalonada. Sea  $\mathcal{Q} = \{Q_1, \dots, Q_k\}$  la partición referida en la Definición 2.27. Llamamos a los elementos de  $\mathcal{Q}$  los **pasos** de  $W$ .

Los núcleos son una generalización de las gráficas en el siguiente sentido. Supongamos que  $H = (V(H), E(H), \alpha, \beta)$  es una gráfica con pesos de orden  $n$  con conjunto de vértices  $V(H) = \{v_1, \dots, v_n\}$ . A partir de  $H$  podemos construir una función escalonada como sigue. Primero dividimos el intervalo  $[0, 1]$  en  $n$  intervalos  $J_1, \dots, J_n$  de longitud  $\lambda(J_i) = \alpha_i / \alpha_H$  para cada  $i \in [n]$ . Defina  $W_H : [0, 1]^2 \rightarrow \mathbb{R}$  mediante la regla  $W_H(x, y) = \beta_{ij}(H)$  para  $(x, y) \in J_i \times J_j$  con  $i, j \in [n]$ . Note que  $W_H$  es simétrica debido al hecho  $\beta_{ij}(H) = \beta_{ji}(H)$ . Inversamente, cualquier función escalonada corresponde a una gráfica con peso. En efecto, supongamos que  $U \in \mathcal{W}$  es una función escalonada con pasos  $\mathcal{P} = \{P_1, \dots, P_n\}$ . Definamos una gráfica  $H$  como sigue. Primero consideremos el conjunto de vértices  $V(H) := \{v_1, \dots, v_n\}$  con  $v(H) = n$ . El conjunto de aristas será el conjunto  $V(H) \times V(H)$ , donde la arista  $v_i v_j$  tiene peso  $U(x, y)$  para  $(x, y) \in P_i \times P_j$ . No habrá inconveniente con esto último, pues, al ser una función escalonada,  $U$  es constante en cada  $P_i \times P_j$  con  $i, j \in [n]$ .

Una propiedad de las gráficas que también tiene su análogo a grafones es que podemos definir una función de grado, donde consideramos el intervalo  $[0, 1]$  como un “continuo” del conjunto de vértices.

Un grafón cuya imagen es  $\{0, 1\}$  puede ser considerado como una gráfica en el conjunto de vértices  $[0, 1]$ .

DEFINICIÓN 2.29. Supongamos que  $W \in \mathcal{W}$  es un núcleo y tomemos  $x \in [0, 1]$ . Definimos la **función de grado** como

$$d_W(x) = \int_0^1 W(x, y) dy.$$

#### 4. Métrica de corte y otras distancias en el espacio de gráficas

Nos interesa definir la distancia entre dos gráficas arbitrarias y no sólo eso, sino que además esta distancia refleje similitud en sus estructuras. Para ello hay que recordar las siguientes definiciones de normas.

Algunas normas en el espacio  $\mathbb{R}^{n \times n}$ , de matrices de tamaño  $n \times n$  con entradas reales, son las siguientes. Sea  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ . Una primer norma es

$$\|A\|_1 = \frac{1}{n^2} \sum_{i,j=1}^n |a_{ij}|,$$

que recibe el nombre de norma  $\ell_1$ .

Otra norma es

$$\|A\|_2 = \left( \frac{1}{n^2} \sum_{i,j=1}^n (a_{ij})^2 \right)^{1/2},$$

llamada norma  $\ell_2$  o norma de Frobenius.

Finalmente la norma  $\ell_\infty$

$$\|A\|_\infty = \max_{i,j} |a_{ij}|.$$

Notemos la normalización de las normas  $\ell_1$  y  $\ell_2$ , cuando  $A$  es una matriz de adyacencia estas normas están entre cero y uno, como lo muestra la siguiente observación.

**OBSERVACIÓN 2.30.** *Fijemos un entero positivo  $n$ . Sea  $A = (a_{ij})$  la matriz de adyacencia de una gráfica simple de orden  $n$ . Entonces  $0 \leq \|A\|_1 \leq 1$  y  $0 \leq \|A\|_2 \leq 1$ .*

**DEMOSTRACIÓN.** Por ser  $A$  una matriz de adyacencia, tenemos  $0 \leq a_{ij} \leq 1$  para  $i, j \in [n]$ . Entonces,  $0 \leq |a_{ij}| \leq 1$ . De modo que

$$0 \leq \frac{1}{n^2} \sum_{i,j=1}^n |a_{ij}| \leq \frac{1}{n^2} \cdot n^2 = 1.$$

Por lo tanto,  $0 \leq \|A\|_1 \leq 1$ . De lo anterior también tenemos que  $0 \leq a_{ij}^2 \leq 1$  para  $i, j \in [n]$ , de donde obtenemos

$$0 \leq \frac{1}{n^2} \sum_{i,j=1}^n a_{ij}^2 \leq \frac{1}{n^2} \cdot n^2 = 1.$$

Por ende,  $0 \leq \|A\|_2 \leq 1$ . □

Otra norma para el espacio  $\mathbb{R}^{n \times n}$  es una un poco menos estándar llamada *norma de corte*, introducida en 1999 por Alan Frieze y Ravi Kannan. Está definida por

$$\|A\|_\square = \frac{1}{n^2} \max_{S,T \subseteq [n]} \left| \sum_{i \in S, j \in T} a_{ij} \right|.$$

**OBSERVACIÓN 2.31.** *Sea  $A \in \mathbb{R}^{n \times n}$ . Entonces,  $\|A\|_\square \leq \|A\|_1$ .*

**DEMOSTRACIÓN.** Suponga  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ . Primero veamos la primera desigualdad. Para  $S, T$  subconjuntos arbitrarios de  $[n]$ ,

$$\left| \sum_{i \in S, j \in T} a_{ij} \right| \leq \sum_{i \in S, j \in T} |a_{ij}| \leq \sum_{i,j=1}^n |a_{ij}|,$$

donde usamos la desigualdad del triángulo y además el hecho de que  $|a_{ij}|$  es positivo para  $i, j \in [n]$ . Como  $S, T$  fueron arbitrarios y  $\sum_{i,j=1}^n |a_{ij}|$  es constante, entonces

$$\max_{S, T \subseteq [n]} \left| \sum_{i \in S, j \in T} a_{ij} \right| \leq \sum_{i,j=1}^n |a_{ij}|.$$

Esto último implica  $\|A\|_{\square} \leq \|A\|_1$ .  $\square$

Queremos encontrar una manera de definir la distancia entre dos gráficas  $G$  y  $G'$ . Aquí es donde la matriz de adyacencia juega un papel fundamental. Es bien sabido que una norma definida en un espacio vectorial induce una métrica. Así que una manera de definir la distancia entre dos gráficas es asociar a cada gráfica su correspondiente matriz de adyacencia, en este caso  $A_G$  y  $A_{G'}$ . Notemos que para esto es necesario pedir que las matrices  $A_G$  y  $A_{G'}$  sean del mismo tamaño. En cuyo caso, el orden de  $G$  y el orden de  $G'$  coinciden. Analicemos entonces el caso en el que  $G$  y  $G'$  tienen el mismo orden. Además, para un primer análisis será necesario pedir no sólo que su orden coincida, sino que el conjunto de vértices de  $G$  y  $G'$  sea el mismo. Posteriormente se analizará el caso en el que los conjuntos de vértices son distintos.

**4.1. Dos gráficas con igual conjunto de vértices.** Fijemos un entero positivo  $n$ . A lo largo de esta sección suponemos que  $G = (V(G), E(G))$  y  $G' = (V(G'), E(G'))$  son dos gráficas de orden  $n$  con la propiedad  $V(G) = V(G')$ . Para facilitar nuestro análisis, supongamos  $V(G) = V(G') = [n]$ . Sean  $A_G$  y  $A_{G'}$  las matrices de adyacencia de  $G$  y  $G'$ , respectivamente.

DEFINICIÓN 2.32. Definimos la **distancia**  $\ell_1$  como

$$d_1(G, G') := \|A_G - A_{G'}\|_1.$$

Es posible darle una interpretación a  $d_1$  en términos de la teoría de gráficas con la siguiente proposición.

PROPOSICIÓN 2.33. Para cualesquiera dos gráficas  $G$  y  $G'$  con la propiedad  $V(G) = V(G')$  se cumple

$$d_1(G, G') = \frac{|E(G) \Delta E(G')|}{n^2}.$$

DEMOSTRACIÓN. Mostraremos la igualdad buscada mediante un argumento probabilista. Como  $G$  y  $G'$  tienen el mismo conjunto de vértices, podemos denotar a dicho conjunto como  $V = [n]$ . Considere el espacio de probabilidad discreto conformado por el espacio muestral  $[n]^2$ , la  $\sigma$ -álgebra  $\mathcal{P}([n]^2)$  y  $\mathbb{P}$  la medida de probabilidad uniforme. Notemos que, en particular, tanto  $E(G)$  como  $E(G')$  son eventos en nuestro espacio de probabilidad. Entonces es posible considerar a las variables aleatorias  $\mathbb{1}_{E(G)}, \mathbb{1}_{E(G')} : [n]^2 \rightarrow \mathbb{R}$ . De este modo, podemos ver que las matrices de adyacencia de  $A_G = (a_{ij})$  y  $A_{G'} = (b_{ij})$  de  $G$  y  $G'$ , respectivamente, pueden ser vistas como una función indicadora en el sentido de que para cualquier  $ij \in [n]^2$  se tiene

$$\mathbb{1}_{E(G)}(ij) = a_{ij} \quad \text{y} \quad \mathbb{1}_{E(G')}(ij) = b_{ij}.$$

Así,

$$(33) \quad \|A_G - A_{G'}\|_1 = \frac{1}{n^2} \sum_{i,j=1}^n |a_{ij} - b_{ij}| = \frac{1}{n^2} \sum_{i,j=1}^n |\mathbb{1}_{E(G)}(ij) - \mathbb{1}_{E(G')}(ij)|.$$

Recordemos que en un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , para cualesquiera  $A, B \in \mathcal{F}$  se cumple  $|\mathbb{1}_A - \mathbb{1}_B| = \mathbb{1}_{A \Delta B}$ . En particular, para  $E(G), E(G') \in \mathcal{P}([n]^2)$ , la igualdad  $|\mathbb{1}_{E(G)} - \mathbb{1}_{E(G')}| = \mathbb{1}_{E(G) \Delta E(G')}$  se satisface. Consecuentemente, aplicando esta afirmación a lo ya obtenido en (33),

$$\|A_G - A_{G'}\|_1 = \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{1}_{E(G) \Delta E(G')}(ij) = \frac{1}{n^2} |E(G) \Delta E(G')|.$$

Por lo tanto,  $d_1(G, G') := \|A_G - A_{G'}\|_1 = \frac{|E(G) \Delta E(G')|}{n^2}$ .  $\square$

La demostración de la proposición anterior muestra, a su vez, que al considerar el espacio de probabilidad discreto  $([n]^2, \mathcal{P}([n]^2), \mathbb{P})$  donde  $\mathbb{P}$  es la medida de probabilidad uniforme, obtenemos  $\mathbb{P}(E(G) \Delta E(G')) = d_1(G, G')$ . De manera que podemos interpretar a la distancia  $\ell_1$  entre dos gráficas  $G$  y  $G'$  con el mismo conjunto de vértices como la probabilidad de elegir al azar un par  $ij \in [n]^2$  y éste pertenezca exactamente a una gráfica.

Otra manera de interpretar a  $d_1(G, G')$  es la fracción de pares de vértices que tenemos que agregar en ambas gráficas para obtener una misma gráfica.

EJEMPLO 2.34. Consideremos  $G = ([8], E(G))$  la gráfica completa de orden 8, también denotada por  $K_8$ . Definamos  $G'$  como la gráfica con conjunto de vértices  $[8]$  y conjunto de aristas  $E(G') := E(G) \setminus \{(2, 3), (5, 6), (7, 8)\}$ . De manera que  $E(G')$  es subconjunto de  $E(G)$ . Es fácil ver que  $|E(G) \Delta E(G')|$  es igual a 3. Entonces, por la Proposición 2.33,

$$d_1(G, G') = \frac{|E(G) \Delta E(G')|}{8^2} = \frac{3}{64}.$$

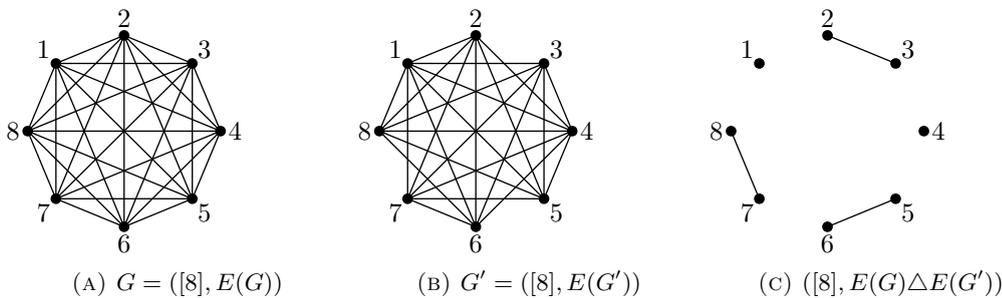


FIGURA 3. La distancia  $d_1$  entre dos gráficas.

No obstante, la distancia  $\ell_1$  no resulta ser adecuada en algunas ocasiones, como en el siguiente ejemplo.

EJEMPLO 2.35. Definamos las gráficas  $G = ([8], E(G))$  y  $G' = ([8], E(G'))$  con conjuntos de aristas  $E(G) = \{(1, 6), (2, 5), (3, 8), (4, 7)\}$  y  $E(G') = \{(1, 4), (2, 7), (3, 6), (5, 8)\}$ . La Figura 4 nos dice que en el fondo las gráficas  $G$  y  $G'$  son iguales, sin embargo, dada nuestra definición de la distancia  $d_1$ , la distancia entre  $G$  y  $G'$  no es cero. Notemos que  $E(G)$  y  $E(G')$  son conjuntos ajenos, de manera que

$$d_1(G, G') = \frac{|E(G) \Delta E(G')|}{8^2} = \frac{|E(G) \cup E(G')|}{8^2} = \frac{8}{8^2} = \frac{1}{8}.$$

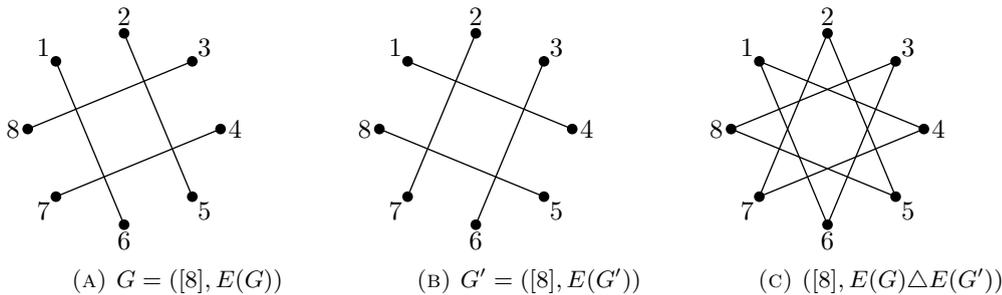


FIGURA 4. La distancia  $d_1$  no es realmente efectiva en el espacio de gráficas en el sentido de que dos gráficas esencialmente iguales tienen distancia positiva.

Por otra parte, la norma de corte induce una distancia, llamada la *métrica de corte*.

DEFINICIÓN 2.36. Sean  $G = (V(G), E(G))$  y  $G' = (V(G'), E(G'))$  dos gráficas tales que  $V(G) = V(G')$ . Definimos la **distancia de corte** como

$$d_{\square}(G, G') := \|A_G - A_{G'}\|_{\square}.$$

DEFINICIÓN 2.37. Sean  $G = (V, E)$  una gráfica y  $S, T \subseteq V$ . Definimos

$$e_G(S, T) := |\{(u, v) \mid u \in S, v \in T\}|.$$

PROPOSICIÓN 2.38. Para cualesquiera dos gráficas  $G = (V(G), E(G))$  y  $G' = (V(G'), E(G'))$  con  $V(G) = V(G')$  se satisface

$$d_{\square}(G, G') = \max_{S, T \subseteq V(G)} \frac{|e_G(S, T) - e_{G'}(S, T)|}{n^2}.$$

DEMOSTRACIÓN. Sean  $A_G = (a_{ij})$  y  $A_{G'} = (b_{ij})$  las matrices de adyacencia de  $G$  y  $G'$ , respectivamente. Observemos que, de la definición de una matriz de adyacencia, para  $S, T \subseteq V(G)$  se tiene

$$(34) \quad e_G(S, T) = \sum_{i \in S, j \in T} a_{ij} \quad \text{y} \quad e_{G'}(S, T) = \sum_{i \in S, j \in T} b_{ij}.$$

Entonces

$$d_{\square}(G, G') = \frac{1}{n^2} \max_{S, T \subseteq V(G)} \left| \sum_{i \in S, j \in T} (a_{i,j} - b_{i,j}) \right| = \frac{1}{n^2} \max_{S, T \subseteq V(G)} \left| \sum_{i \in S, j \in T} a_{i,j} - \sum_{i \in S, j \in T} b_{i,j} \right|.$$

De esta última igualdad podemos concluir por (34).  $\square$

Para un entero positivo  $n$ , debido a que para toda  $A \in \mathbb{R}^{n \times n}$  se cumple  $\|A\|_{\square} \leq \|A\|_1$  (Observación 2.31), entonces tenemos  $d_{\square}(G, G') \leq d_1(G, G')$  para cualesquiera dos gráficas  $G$  y  $G'$  de orden  $n$ .

**4.2. Dos gráficas con el mismo número de vértices.** Fijemos un entero positivo  $n$ . A lo largo de esta sección suponemos que  $G = (V(G), E(G))$  y  $G' = (V(G'), E(G'))$  son dos gráficas de orden  $n$ . Un siguiente paso en la definición de una métrica entre dos gráficas es suponer que son del mismo orden, esto es, las gráficas  $G$  y  $G'$  tienen la propiedad  $v(G) = v(G')$ . Enfatizamos que es posible el caso  $V(G) \neq V(G')$ .

DEFINICIÓN 2.39. Sean  $n$  un entero positivo,  $G$  y  $G'$  dos gráficas de orden  $n$ . Definimos la distancia entre ellas por

$$\widehat{d}_{\square}(G, G') := \min\{d_{\square}(G^{\varphi}, G'^{\psi}) \mid \varphi : [n] \rightarrow V(G), \psi : [n] \rightarrow V(G') \\ \text{son etiquetamientos de } G \text{ y } G', \text{ respectivamente}\}.$$

EJEMPLO 2.40. Consideremos las gráficas  $G$  y  $G'$  de orden 5 con conjuntos de vértices  $V(G) = \{v_1, \dots, v_5\}$  y  $V(G') = \{u_1, \dots, u_5\}$  con la propiedad  $V(G) \neq V(G')$ . Y con conjuntos de aristas

$$E(G) = \{v_1v_2, v_2v_3, v_3v_4, v_1v_5\}$$

y

$$E(G') = \{u_1u_2, u_4u_5, u_3u_4, u_1u_5\}.$$

Las gráficas  $G$  y  $G'$  están representadas en la Figura 5.

En base a la Figura 5 no es difícil ver que las gráficas  $G$  y  $G'$  son esencialmente la misma gráfica y, por ende, la distancia entre  $G$  y  $G'$  debería ser cero. En efecto, consideremos el etiquetamiento de  $G$   $\varphi : [5] \rightarrow V(G)$  dado por  $\varphi(i) = v_i$  para cada  $i \in [5]$ . Y  $\psi : [5] \rightarrow V(G')$  el etiquetamiento de  $G'$  con regla de correspondencia

$$\psi(1) = u_4, \psi(2) = u_5, \psi(3) = u_1, \psi(4) = u_2, \psi(5) = u_3.$$

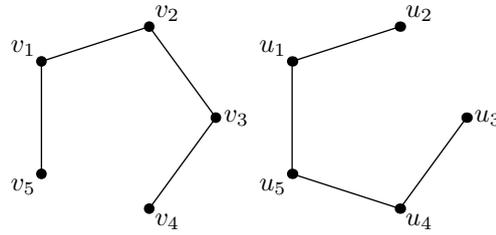
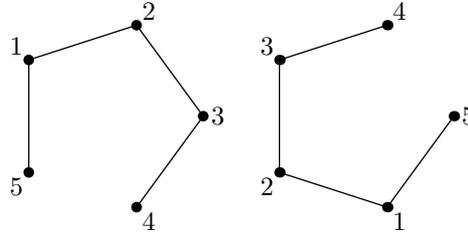


FIGURA 5. Dos gráficas del mismo orden esencialmente iguales.

FIGURA 6.  $G^\varphi$  y  $G'^\psi$ 

Entonces los conjuntos de aristas de  $G_1^\varphi$  y  $G_2^\psi$  son

$$E^\varphi(G) = \{(1, 2), (2, 3), (3, 4), (1, 5)\}$$

y

$$E^\psi(G') = \{(3, 4), (1, 2), (5, 1), (3, 2)\}.$$

Por consiguiente,  $G^\varphi$  es igual a la gráfica  $G'^\psi$ , pues comparten el mismo conjunto de vértices y coinciden también sus conjuntos de aristas. Visualmente podemos confirmar este hecho en la Figura 6. Consecuentemente,  $0 \leq \widehat{\delta}_\square(G, G') \leq \delta_\square(G^\varphi, G'^\psi) = 0$ . Por lo tanto,  $\widehat{\delta}_\square(G, G') = 0$ .

**4.3. Dos gráficas arbitrarias.** Fijemos dos enteros positivos  $n$  y  $n'$ . Supongamos que  $G = (V, E)$  y  $G' = (V', E')$  son dos gráficas con conjuntos de vértices  $V = [n]$  y  $V' = [n']$ . Dado que  $n$  y  $n'$  son arbitrarios, para poder considerar la distancia entre  $G$  y  $G'$  necesitamos que ambas gráficas tengan la misma cantidad de vértices. De la Observación 2.17 sabemos que  $G(n')$  tiene  $n' \cdot v(G) = n' \cdot n$  vértices, mientras que  $G'(n)$  tiene  $n \cdot v(G') = n \cdot n'$  vértices, es decir,  $G(n')$  y  $G'(n)$  son del mismo orden. Más generalmente, para cada entero positivo  $k$ ,  $G(kn')$  y  $G'(kn)$  son del mismo orden. Por ende, podemos establecer la definición de la distancia entre  $G$  y  $G'$  a través de gráficas autoreplicadas.

**DEFINICIÓN 2.41.** Sean  $n$  y  $n'$  dos enteros positivos. Supongamos que  $G$  y  $G'$  dos gráficas con conjuntos de vértices  $[n]$  y  $[n']$ , respectivamente. Definimos la distancia  $\delta_\square$  entre  $G$  y  $G'$  como

$$\delta_\square(G, G') = \lim_{k \rightarrow \infty} \widehat{\delta}_\square(G(kn'), G'(kn)).$$

## 5. Norma de corte y distancia de corte para núcleos

Ahora definiremos una métrica en el espacio de núcleos. Tal y como se observó en la introducción, los grafones son una generalización de las gráficas.

### 5.1. Norma de corte.

**DEFINICIÓN 2.42.** Definimos la **norma de corte** en el espacio de núcleos  $\mathcal{W}$  como:

$$\|W\|_\square := \sup \left\{ \left| \int_{S \times T} W d\lambda^2 \right| : S, T \in \mathcal{B}([0, 1]) \right\}.$$

OBSERVACIÓN 2.43. Consideremos el espacio de probabilidad  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ . La norma de corte en  $\mathcal{W}$  es la  $\mathcal{S}_\square$ -norma uniforme (Definición 1.3) al considerar el 2-semianillo en  $[0, 1]^2$  (Ejemplo 1.8)

$$\mathcal{S}_\square := \{S \times T \mid S, T \in \mathcal{B}([0, 1])\}.$$

PROPOSICIÓN 2.44. La norma de corte es, en efecto, una norma.

DEMOSTRACIÓN. Es similar a la demostración en la Proposición 1.4. Sean  $S, T \in \mathcal{B}([0, 1])$  y  $W \in \mathcal{W}$  tal que  $W = 0$   $\lambda^2$ -casi donde sea. En consecuencia,  $\int_{S \times T} W d\lambda^2 = 0$ . Se sigue que  $\|W\|_\square = 0$ .

Recíprocamente, sea  $W \in \mathcal{W}$  tal que  $\|W\|_\square = 0$ . De manera que

$$\left| \int_{S \times T} W d\lambda^2 \right| = 0$$

para cualesquiera  $S, T \in \mathcal{B}([0, 1])$ ; en particular, para cualquier  $a_1, b_1, a_2, b_2 \in \mathbb{R}$  con  $0 \leq a_1 \leq b_1 \leq 1$  y  $0 \leq a_2 \leq b_2 \leq 1$ ,

$$\int_{[a_1, b_1] \times [a_2, b_2]} W d\lambda^2 = 0.$$

Defina  $\Pi_2 := \{[a, b] \mid a, b \in \mathbb{R}^2, 0 \leq a \leq b \leq 1\} \cup \{\emptyset\}$  (Notación A.7). Las igualdades en las dos integrales previas nos dicen que

$$\int_A W d\lambda^2 = 0$$

para cualquier  $A \in \Pi_2$  (la primera igualdad es para el caso  $A = \emptyset$ ). Sabemos que  $\Pi_2$  es un  $\pi$ -sistema que cumple  $\mathcal{B}([0, 1]^2) = \sigma(\Pi_2)$  (Corolario A.15). Por lo cual, considerando el espacio de medida  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ , dado que  $W$  es integrable (Proposición 2.25) y medible por ser un núcleo, entonces la Proposición A.37 asegura que  $W = 0$   $\lambda^2$ -casi donde sea.

Fijemos ahora  $W \in \mathcal{W}$  arbitrario y  $\alpha \in \mathbb{R}$ . Tenemos, por linealidad de la integral,

$$\begin{aligned} \|\alpha W\|_\square &= \sup \left\{ \left| \int_{S \times T} \alpha W d\lambda^2 \right| : S, T \in \mathcal{B}([0, 1]) \right\} \\ &= \sup \left\{ \left| \alpha \int_{S \times T} W d\lambda^2 \right| : S, T \in \mathcal{B}([0, 1]) \right\} \\ &= |\alpha| \sup \left\{ \left| \int_{S \times T} W d\lambda^2 \right| : S, T \in \mathcal{B}([0, 1]) \right\} = |\alpha| \|W\|_\square, \end{aligned}$$

donde en la penúltima igualdad usamos un hecho básico de cálculo para la constante  $|\alpha| \geq 0$ .

Por último, sean  $W, U \in \mathcal{W}$  y  $S, T \in \mathcal{B}([0, 1])$  arbitrarios. De la desigualdad del triángulo, usando implícitamente para la primera igualdad que  $W \mathbb{1}_{S \times T}$  y  $U \mathbb{1}_{S \times T}$  son integrables,

$$\begin{aligned} \left| \int_{S \times T} (W + U) d\lambda^2 \right| &= \left| \int_{S \times T} W d\lambda^2 + \int_{S \times T} U d\lambda^2 \right| \\ &\leq \left| \int_{S \times T} W d\lambda^2 \right| + \left| \int_{S \times T} U d\lambda^2 \right| \\ &\leq \|W\|_\square + \|U\|_\square. \end{aligned}$$

Por consiguiente, dado que  $S, T \in \mathcal{B}([0, 1])$  son arbitrarios,  $\|W + U\|_\square \leq \|W\|_\square + \|U\|_\square$ . De modo que la norma de corte es, en efecto, una norma.  $\square$

Por consiguiente, es posible dar una métrica al espacio de núcleos  $\mathcal{W}$ .

DEFINICIÓN 2.45. Definimos la **métrica de corte** en  $\mathcal{W}$  como

$$d_\square(U, W) = \|U - W\|_\square.$$

**5.2. Distancia de corte para núcleos no etiquetados.** Los núcleos definidos en un conjunto fijo  $[0, 1]$  corresponden a gráficas etiquetadas. Así como en gráficas, introducimos una versión “no etiquetada” de la norma de corte, encontrando la mejor “cubierta” de los conjuntos subyacentes.

**DEFINICIÓN 2.46.** Sean  $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$  y  $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  dos espacios de probabilidad, y  $\varphi : \Omega_1 \rightarrow \Omega_2$  una función. Decimos que  $\varphi : (\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \rightarrow (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  **preserva la medida** si  $\varphi$  es  $(\mathcal{F}_1, \mathcal{F}_2)$ -medible y además  $\mathbb{P}_1(\varphi^{-1}(A)) = \mathbb{P}_2(A)$  para todo  $A \in \mathcal{F}_2$ .

**EJEMPLO 2.47.** Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , la función identidad  $Id : \Omega \rightarrow \Omega$ , esto es, dada por  $Id(\omega) = \omega$ , es una función que preserva la medida.

**DEMOSTRACIÓN.** Sea  $A \in \mathcal{F}$ . Por definición  $Id^{-1}(A)$  es igual a  $A$ . Consecuentemente,  $Id^{-1}(A)$  pertenece a  $\mathcal{F}$  y  $\mathbb{P}(Id^{-1}(A)) = \mathbb{P}(A)$ . Por lo tanto,  $Id : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\Omega, \mathcal{F}, \mathbb{P})$  preserva la medida.  $\square$

El nombre es un poco engañoso porque es  $\varphi^{-1}$  quien preserva la medida en lugar de  $\varphi$ . Una función que preserva la medida no necesariamente es biyectiva.

**DEFINICIÓN 2.48.** Sean  $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$  y  $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  dos espacios de probabilidad, y  $\varphi : \Omega_1 \rightarrow \Omega_2$  una función. Supongamos que  $\varphi : (\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \rightarrow (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$  preserva la medida. Decimos que  $\varphi$  es **invertible** si  $\varphi$  es biyectiva y  $\varphi^{-1} : (\Omega_2, \mathcal{F}_2, \mathbb{P}_2) \rightarrow (\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$  preserva la medida.

Es importante ser precisos en los espacios de probabilidad con los que se trabaja cuando consideramos una función que preserva la medida. No obstante, para el conjunto  $[0, 1]$ , en este capítulo siempre consideraremos el espacio de probabilidad  $([0, 1], \mathcal{B}([0, 1]), \lambda)$ . Es por ello que se establece la siguiente notación.

**NOTACIÓN 2.49.** Supongamos que tenemos una función  $\varphi : [0, 1] \rightarrow [0, 1]$ . En lugar de la extensa notación (y en este caso particular poco útil)  $\varphi : ([0, 1], \mathcal{B}([0, 1]), \lambda) \rightarrow ([0, 1], \mathcal{B}([0, 1]), \lambda)$  preserva la medida, denotamos  $\varphi : [0, 1] \rightarrow [0, 1]$  preserva la medida o simplemente  $\varphi$  preserva la medida.

Aclarada esta notación, definamos el conjunto

$$S_{[0,1]} := \{\varphi : [0, 1] \rightarrow [0, 1] \mid \varphi \text{ preserva la medida y es invertible}\}.$$

**PROPOSICIÓN 2.50.** Considere el espacio de medida  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ . Tomemos  $\varphi \in S_{[0,1]}$  arbitraria y definamos  $\phi : [0, 1]^2 \rightarrow [0, 1]^2$  mediante la regla  $\phi(x, y) = (\varphi(x), \varphi(y))$ . Entonces  $\phi$  es  $(\mathcal{B}([0, 1]^2), \mathcal{B}([0, 1]^2))$ -medible y su medida inducida  $\lambda^2 \circ \phi^{-1}$  (ver Proposición A.44) satisface  $\lambda^2 \circ \phi^{-1} = \lambda^2$ .

**DEMOSTRACIÓN.** Veamos primero que  $\phi$  es medible. En virtud del Corolario A.15 sabemos que  $\Pi_2 = \{[a, b] \mid a, b \in \mathbb{R}^2, 0 \leq a \leq b \leq 1\} \cup \{\emptyset\}$  es un  $\pi$ -sistema que cumple  $\mathcal{B}([0, 1]^2) = \sigma(\Pi_2)$ . De manera que, por la Proposición A.22, bastará mostrar  $\phi^{-1}(A) \in \mathcal{B}([0, 1]^2)$  para todo  $A \in \Pi_2$ . Para  $\emptyset \in \Pi_2$ ,  $\phi^{-1}(\emptyset) = \emptyset \in \mathcal{B}([0, 1]^2)$ . Ahora, fijemos  $A_0 \in \Pi_2 \setminus \{\emptyset\}$  un elemento arbitrario. Por lo cual,  $A_0 = [a_1, b_1] \times [a_2, b_2]$  para  $a_1, a_2, b_1, b_2 \in \mathbb{R}$  con  $a_1 \leq b_1$  y  $a_2 \leq b_2$ . Sencillamente de la definición, la pre-imagen de  $A_0$  bajo  $\phi$  es

$$(35) \quad \phi^{-1}([a_1, b_1] \times [a_2, b_2]) = \{(x, y) \in [0, 1]^2 \mid (\varphi(x), \varphi(y)) \in [a_1, b_1] \times [a_2, b_2]\}.$$

No es difícil ver que se cumple la igualdad

$$(36) \quad \{(x, y) \in [0, 1]^2 \mid (\varphi(x), \varphi(y)) \in [a_1, b_1] \times [a_2, b_2]\} = \varphi^{-1}([a_1, b_1]) \times \varphi^{-1}([a_2, b_2]).$$

Como  $[a_1, b_1] \in \mathcal{B}([0, 1])$  y  $\varphi$  es medible, entonces  $\varphi^{-1}([a_1, b_1]) \in \mathcal{B}([0, 1])$ . De igual manera,  $\varphi^{-1}([a_2, b_2]) \in \mathcal{B}([0, 1])$ . Consecuentemente,  $\varphi^{-1}([a_1, b_1]) \times \varphi^{-1}([a_2, b_2]) \in \bigotimes_{i=1}^2 \mathcal{B}([0, 1])$ . Puesto que  $\bigotimes_{i=1}^2 \mathcal{B}([0, 1]) = \mathcal{B}([0, 1]^2)$  (Proposición A.5), entonces  $\varphi^{-1}([a_1, b_1]) \times \varphi^{-1}([a_2, b_2]) \in \mathcal{B}([0, 1]^2)$ . De la cadena de igualdades compuesta por (35) y (36),

$$(37) \quad \phi^{-1}([a_1, b_1] \times [a_2, b_2]) = \varphi^{-1}([a_1, b_1]) \times \varphi^{-1}([a_2, b_2]).$$

Por lo tanto,  $\phi^{-1}([a_1, b_1] \times [a_2, b_2]) \in \mathcal{B}([0, 1]^2)$ . La arbitrariedad de  $A_0$  nos permite deducir que  $\phi^{-1}(A) \in \mathcal{B}([0, 1]^2)$  para cualquier  $A \in \Pi_2$ , como queríamos demostrar. Se sigue que  $\phi$  es  $(\mathcal{B}([0, 1]^2), \mathcal{B}([0, 1]^2))$ -medible.

Ahora veamos  $\lambda^2 \circ \phi^{-1} = \lambda^2$ . Para ello usaremos la Proposición A.16. La igualdad de las medidas es evidente en el caso  $\emptyset \in \Pi_2$ , pues  $\phi^{-1}(\emptyset) = \emptyset$ , así que  $\lambda^2 \circ \phi^{-1}(\emptyset) = \lambda^2(\emptyset)$ . En el caso  $A \in \Pi_2 \setminus \{\emptyset\}$ , podemos considerar  $A_0$  como antes, esto es,  $A_0$  es de la forma  $A_0 = [a_1, b_1] \times [a_2, b_2]$  y por (37),

$$\lambda^2(\phi^{-1}([a_1, b_1] \times [a_2, b_2])) = \lambda^2(\varphi^{-1}([a_1, b_1]) \times \varphi^{-1}([a_2, b_2])) = \lambda(\varphi^{-1}([a_1, b_1])) \times \lambda(\varphi^{-1}([a_2, b_2])).$$

Como  $\varphi$  es una función que preserva la medida,  $\lambda(\varphi^{-1}([a_i, b_i])) = \lambda([a_i, b_i])$  para  $i = 1, 2$ . Por ende,

$$(38) \quad \lambda^2(\phi^{-1}([a_1, b_1] \times [a_2, b_2])) = \lambda([a_1, b_1]) \times \lambda([a_2, b_2]) = \lambda^2([a_1, b_1] \times [a_2, b_2]).$$

Puesto que  $A_0$  fue un elemento arbitrario, esto muestra que para cualquier  $A \in \Pi_2$ ,  $\lambda^2 \circ \phi^{-1}(A) = \lambda^2(A)$ . Resta ver que  $\lambda^2 \circ \phi^{-1}([0, 1]^2) = \lambda^2([0, 1]^2) < \infty$ . La igualdad  $\lambda^2 \circ \phi^{-1}([0, 1]^2) = \lambda^2([0, 1]^2)$  se obtiene de (38) en el caso particular  $a_1 = a_2 = 0$ ,  $b_1 = b_2 = 1$ . Por otro lado,  $\lambda^2([0, 1]^2) = 1$ . En consecuencia,  $\lambda^2 \circ \phi^{-1}([0, 1]^2) = \lambda^2([0, 1]^2) = 1 < \infty$ . Por lo tanto, la Proposición A.16 garantiza la igualdad  $\lambda^2 \circ \phi^{-1} = \lambda^2$ .  $\square$

NOTACIÓN 2.51. Dado un núcleo  $W : [0, 1]^2 \rightarrow \mathbb{R}$  y una función  $\varphi \in S_{[0,1]}$  definimos  $W^\varphi : [0, 1]^2 \rightarrow \mathbb{R}$  como  $W^\varphi(x, y) := W(\varphi(x), \varphi(y))$ .

OBSERVACIÓN 2.52. Para cada  $W \in \mathcal{W}$  y  $\varphi \in S_{[0,1]}$ , la función  $W^\varphi$  pertenece a  $\mathcal{W}$ .

DEMOSTRACIÓN. Fijemos un núcleo  $W \in \mathcal{W}$  y una función  $\varphi \in S_{[0,1]}$ . Defina  $\phi : [0, 1]^2 \rightarrow [0, 1]^2$  mediante la regla  $\phi(x, y) = (\varphi(x), \varphi(y))$ . Por la Proposición 2.50  $\phi$  es  $(\mathcal{B}([0, 1]^2), \mathcal{B}([0, 1]^2))$ -medible. Notemos la igualdad

$$W^\varphi = W \circ \phi.$$

Dado que composición de funciones medibles es medible (Proposición A.24) tenemos que  $W^\varphi$  es  $(\mathcal{B}([0, 1]^2), \mathcal{B}(\mathbb{R}))$ -medible.

Para cada  $x, y \in [0, 1]$ , como  $W$  es simétrica,

$$W^\varphi(x, y) = W(\varphi(x), \varphi(y)) = W(\varphi(y), \varphi(x)) = W^\varphi(y, x).$$

Por ende,  $W^\varphi$  es simétrica.

Finalmente,  $W^\varphi$  está acotada porque  $W$  es una función acotada. Por lo tanto,  $W^\varphi$  es un núcleo, como queríamos demostrar.  $\square$

DEFINICIÓN 2.53. Definimos la **distancia de corte** para  $U, W \in \mathcal{W}$  como

$$\delta_{\square}(U, W) := \inf_{\varphi \in S_{[0,1]}} d_{\square}(U, W^\varphi).$$

Fijemos dos núcleos  $U, W \in \mathcal{W}$ . Dado que, para cada  $\varphi \in S_{[0,1]}$ , sabemos que  $W^\varphi$  es un núcleo, ciertamente es válido considerar  $d_{\square}(U, W^\varphi)$ . Como  $d_{\square}(U, W^\varphi)$  está acotada inferiormente por cero y la función identidad pertenece a  $S_{[0,1]}$ , entonces podemos considerar el ínfimo sobre  $d_{\square}(U, W^\varphi)$ . Esto es, la distancia de corte tiene sentido. Ahora bien, el ínfimo de  $d_{\square}(U, W^\varphi)$  con respecto a cada elemento  $\varphi \in S_{[0,1]}$  debería coincidir al considerar el ínfimo  $d_{\square}(U^\varphi, W)$  con  $\varphi \in S_{[0,1]}$ . Esta igualdad correspondería al análogo en gráficas; es decir, dadas dos gráficas con el mismo número de vértices, la distancia entre ellas es el resultado de minimizar las distancias de entre todos los posibles etiquetamientos. En núcleos,  $U^\varphi$  corresponde a un etiquetamiento de  $U$  para cada  $\varphi \in S_{[0,1]}$ , de igual manera con  $W^\varphi$ . Buscamos que la distancia de corte no dependa sobre qué grafón estamos "etiquetando". Es decir, queremos probar

$$(39) \quad \inf_{\varphi \in S_{[0,1]}} d_{\square}(U^\varphi, W) = \inf_{\varphi \in S_{[0,1]}} d_{\square}(U, W^\varphi).$$

Trabajar con funciones que preservan la medida hace que la igualdad buscada sea posible. Antes de mostrar la igualdad (39) serán necesarios algunos resultados técnicos. Primero, veamos la siguiente proposición.

PROPOSICIÓN 2.54. Sea  $W \in \mathcal{W}$  y  $\varphi \in S_{[0,1]}$ . Entonces

$$(W^\varphi)^{\varphi^{-1}} = W = (W^{\varphi^{-1}})^\varphi.$$

DEMOSTRACIÓN. El resultado se sigue de una sencilla cadena de igualdades. Para  $(x, y) \in [0, 1]^2$ ,

$$(W^\varphi)^{\varphi^{-1}}(x, y) = W^\varphi(\varphi^{-1}(x), \varphi^{-1}(y)) = W(\varphi(\varphi^{-1}(x)), \varphi(\varphi^{-1}(y))) = W(x, y).$$

Por lo tanto,  $(W^\varphi)^{\varphi^{-1}} = W$ . Análogamente,  $(W^{\varphi^{-1}})^\varphi = W$ .  $\square$

Como consecuencia de la Proposición 2.50 y el teorema de cambio de variable (Teorema A.45) tenemos el siguiente corolario.

COROLARIO 2.55. Sean  $W : [0, 1]^2 \rightarrow \mathbb{R}$  un núcleo y  $\varphi \in S_{[0,1]}$ . Entonces para cualesquiera  $S, T \in \mathcal{B}([0, 1])$  se cumple

$$\int_{S \times T} W^\varphi d\lambda^2 = \int_{\varphi(S) \times \varphi(T)} W d\lambda^2.$$

Para mostrar el corolario será útil la siguiente observación.

OBSERVACIÓN 2.56. Sean  $W$  un núcleo y  $\varphi \in S_{[0,1]}$ . Para cada  $S, T \in \mathcal{B}([0, 1])$  se cumple

$$W^\varphi \cdot \left( \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi = \left( W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi.$$

Procedamos con la demostración del Corolario 2.55 y posteriormente con la demostración de la Observación 2.56.

DEMOSTRACIÓN DEL COROLARIO 2.55. Sean  $W$  un núcleo y  $\varphi \in S_{[0,1]}$ . Fijemos  $S, T \in \mathcal{B}([0, 1])$ . Entonces  $S \times T \in \mathcal{B}([0, 1]^2)$  (Proposición A.5). También ya vimos que  $W^\varphi$  es un núcleo, en particular, una función medible, por lo que es válido considerar la integral de  $W^\varphi$  en  $S \times T$ . Además por la Proposición 2.54 y la Observación 2.56 tenemos, respectivamente,

$$\int_{S \times T} W^\varphi d\lambda^2 = \int W^\varphi \cdot \left( \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi d\lambda^2 = \int \left( W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi d\lambda^2.$$

Al definir  $\phi : [0, 1]^2 \rightarrow [0, 1]^2$  mediante la regla  $\phi(x, y) = (\varphi(x), \varphi(y))$  sabemos  $\lambda^2 \circ \phi^{-1} = \lambda^2$  (Proposición 2.50). Por consiguiente, en virtud del teorema de cambio de variable (Teorema A.45),

$$\int_{S \times T} W^\varphi d\lambda^2 = \int \left( W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi d\lambda^2 = \int W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}} d\lambda^2.$$

Es evidente que  $(\varphi^{-1}(x), \varphi^{-1}(y)) \in S \times T$  implica  $(x, y) \in \varphi(S) \times \varphi(T)$ . Recíprocamente, como  $\varphi$  es biyectiva,  $(x, y) \in \varphi(S) \times \varphi(T)$  implica  $(\varphi^{-1}(x), \varphi^{-1}(y)) \in S \times T$ . Por ende,

$$\mathbb{1}_{S \times T}^{\varphi^{-1}} = \mathbb{1}_{\varphi(S) \times \varphi(T)}.$$

Consecuentemente, de nuestra última integral,

$$\int_{S \times T} W^\varphi d\lambda^2 = \int W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}} d\lambda^2 = \int_{\varphi(S) \times \varphi(T)} W d\lambda^2.$$

$\square$

DEMOSTRACIÓN DE LA OBSERVACIÓN 2.56. Fijemos un núcleo  $W \in \mathcal{W}$ , una función  $\varphi \in S_{[0,1]}$  y dos elementos  $S, T \in \mathcal{B}([0, 1])$ . Entonces  $S \times T \in \mathcal{B}([0, 1]^2)$ . Para cada  $(x, y) \in [0, 1]^2$  se cumple

$$\left( W^\varphi \cdot \left( \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi \right)(x, y) = W^\varphi(x, y) \cdot \left( \mathbb{1}_{S \times T}^{\varphi^{-1}} \right)^\varphi(x, y) = W(\varphi(x), \varphi(y)) \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}}(\varphi(x), \varphi(y)).$$

A su vez,

$$\left(W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}}\right)^{\varphi}(x, y) = \left(W \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}}\right)(\varphi(x), \varphi(y)) = W(\varphi(x), \varphi(y)) \cdot \mathbb{1}_{S \times T}^{\varphi^{-1}}(\varphi(x), \varphi(y)).$$

□

Estamos preparados para demostrar la igualdad (39).

PROPOSICIÓN 2.57. *Para todo  $U, W \in \mathcal{W}$  se satisface*

$$\delta_{\square}(U, W) := \inf_{\varphi \in S_{[0,1]}} d_{\square}(U, W^{\varphi}) = \inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W).$$

DEMOSTRACIÓN. Sean  $U, W \in \mathcal{W}$  y  $\varphi \in S_{[0,1]}$ . Fijemos  $S, T \in \mathcal{B}([0, 1])$  arbitrarios. Recordemos  $U = (U^{\varphi})^{\varphi^{-1}}$  y, puesto que  $\varphi^{-1}$  también pertenece a  $S_{[0,1]}$  y  $U^{\varphi}$  es un núcleo, por el Corolario 2.55 se cumple

$$\int_{S \times T} U d\lambda^2 = \int_{S \times T} (U^{\varphi})^{\varphi^{-1}} d\lambda^2 = \int_{\varphi^{-1}(S) \times \varphi^{-1}(T)} U^{\varphi} d\lambda^2.$$

De la misma manera,

$$\int_{S \times T} W^{\varphi^{-1}} d\lambda^2 = \int_{\varphi^{-1}(S) \times \varphi^{-1}(T)} W d\lambda^2.$$

Entonces, dado que todo núcleo es integrable (Proposición 2.25), podemos aplicar la propiedad de linealidad de la integral (Proposición A.36) para obtener la igualdad

$$\int_{S \times T} U - W^{\varphi^{-1}} d\lambda^2 = \int_{\varphi^{-1}(S) \times \varphi^{-1}(T)} U^{\varphi} - W d\lambda^2.$$

Observemos que  $\varphi^{-1}(S), \varphi^{-1}(T) \in \mathcal{B}([0, 1])$  porque  $\varphi$  es medible. En consecuencia,

$$\left| \int_{S \times T} U - W^{\varphi^{-1}} d\lambda^2 \right| \leq \sup \left\{ \left| \int_{S \times T} U^{\varphi} - W d\lambda^2 \right| : S, T \in \mathcal{B}([0, 1]) \right\} = d_{\square}(U^{\varphi}, W).$$

Notemos que  $d_{\square}(U, W)$  es constante. Por ende, tomando supremo sobre cada  $S, T \in \mathcal{B}([0, 1])$  en ambos lados de la desigualdad anterior,

$$(40) \quad d_{\square}(U, W^{\varphi^{-1}}) \leq d_{\square}(U^{\varphi}, W).$$

Esto implica

$$\delta_{\square}(U, W) = \inf_{\varphi^{-1} \in S_{[0,1]}} d_{\square}(U, W^{\varphi^{-1}}) \leq d_{\square}(U^{\varphi}, W).$$

En consecuencia, al tomar el ínfimo en ambos lados de la desigualdad anterior sobre cada  $\varphi \in S_{[0,1]}$  y como  $\delta_{\square}(U, W)$  es constante, llegamos a

$$(41) \quad \delta_{\square}(U, W) \leq \inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W).$$

Consideremos ahora los núcleos  $U^{\varphi^{-1}}, W^{\varphi}$ . De igual manera a como mostramos (40) tenemos

$$d_{\square}(U^{\varphi^{-1}}, (W^{\varphi})^{\varphi^{-1}}) \leq d_{\square}((U^{\varphi^{-1}})^{\varphi}, W^{\varphi}).$$

Es decir,  $d_{\square}(U^{\varphi^{-1}}, W) \leq d_{\square}(U, W^{\varphi})$ , en uso de la Proposición 2.54. Por consiguiente,

$$\inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W) = \inf_{\varphi^{-1} \in S_{[0,1]}} d_{\square}(U^{\varphi^{-1}}, W) \leq d_{\square}(U, W^{\varphi}).$$

De manera que encontramos una cota inferior para  $d_{\square}(U, W^{\varphi})$ . Debido a la arbitrariedad de  $\varphi$  se sigue que

$$(42) \quad \inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W) \leq \delta_{\square}(U, W).$$

Total, las desigualdades (41) y (42) muestran

$$\inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W) = \inf_{\varphi \in S_{[0,1]}} d_{\square}(U, W^{\varphi}).$$

Esto concluye con la demostración. □

Desafortunadamente, la distancia de corte en el espacio de núcleos solamente es una pseudo-métrica.

PROPOSICIÓN 2.58. *La distancia de corte  $\delta_{\square}$  es una pseudométrica.*

DEMOSTRACIÓN. Sean  $U, V, W \in \mathcal{W}$  fijos.

- (a) Ya vimos que siempre se cumple  $0 \leq \delta_{\square}(U, U)$ . Por otro lado,  $Id : [0, 1] \rightarrow [0, 1]$ , dada por  $Id(x) = (x)$ , preserva la medida (Ejemplo 2.47) y es invertible. Esto implica  $Id \in S_{[0,1]}$ . De manera que

$$\delta_{\square}(U, U) \leq d_{\square}(U, U^{Id}) = 0.$$

Por lo tanto,  $\delta_{\square}(U, U) = 0$ .

- (b) (Simetría) Notemos que  $d_{\square}(U^{\varphi}, W)$  es igual a  $d_{\square}(W, U^{\varphi})$  para cada  $\varphi \in S_{[0,1]}$  porque  $d_{\square}$  es una métrica. Entonces la propiedad de simetría es inmediata de la Proposición 2.57 mediante la siguiente cadena de igualdades:

$$\delta_{\square}(U, W) = \inf_{\varphi \in S_{[0,1]}} d_{\square}(U^{\varphi}, W) = \inf_{\varphi \in S_{[0,1]}} d_{\square}(W, U^{\varphi}) = \delta_{\square}(W, U).$$

- (c) (Desigualdad del triángulo) Para  $\varphi \in S_{[0,1]}$ , como  $d_{\square}$  es una métrica,  $d_{\square}$  satisface la desigualdad del triángulo:

$$d_{\square}(U, W^{\varphi}) \leq d_{\square}(U, V^{\varphi}) + d_{\square}(V, V^{\varphi}) + d_{\square}(V, W^{\varphi}).$$

Entonces

$$\delta_{\square}(U, W) \leq \inf_{\varphi \in S_{[0,1]}} (d_{\square}(U, V^{\varphi}) + d_{\square}(V, V^{\varphi}) + d_{\square}(V, W^{\varphi})).$$

Los conjuntos  $\{d_{\square}(U, V^{\varphi}) \mid \varphi \in S_{[0,1]}\}$ ,  $\{d_{\square}(V, V^{\varphi}) \mid \varphi \in S_{[0,1]}\}$  y  $\{d_{\square}(V, W^{\varphi}) \mid \varphi \in S_{[0,1]}\}$  son no vacíos ( $Id$  pertenece a  $S_{[0,1]}$ ) y están acotados inferiormente por cero. Por consiguiente,

$$\delta_{\square}(U, W) \leq \inf_{\varphi \in S_{[0,1]}} d_{\square}(U, V^{\varphi}) + \inf_{\varphi \in S_{[0,1]}} d_{\square}(V, V^{\varphi}) + \inf_{\varphi \in S_{[0,1]}} d_{\square}(V, W^{\varphi}).$$

Así, como  $\delta_{\square}(V, V) = 0$  por la parte (a) de la presente demostración, se cumple

$$\delta_{\square}(U, W) \leq \delta_{\square}(U, V) + \delta_{\square}(V, W).$$

□

DEFINICIÓN 2.59. Sean  $U, W \in \mathcal{W}$ . Definimos la relación  $\sim$  como:  $U \sim W$  si y sólo si  $\delta_{\square}(U, W) = 0$ .

OBSERVACIÓN 2.60.  $\sim$  es una relación de equivalencia.

DEMOSTRACIÓN. Sean  $U, V, W \in \mathcal{W}$ .

- Ya sabemos  $\delta_{\square}(W, W) = 0$ , por lo que  $W \sim W$ .
- Supongamos  $U \sim W$ . Entonces  $\delta_{\square}(U, W) = 0$ . Ya que  $\delta_{\square}$  es simétrica, se cumple  $\delta_{\square}(W, U) = 0$ . Consecuentemente  $W \sim U$ .
- Supongamos ahora  $U \sim V$  y  $V \sim W$ . En consecuencia, por la desigualdad del triángulo,  $\delta_{\square}(U, W) \leq \delta_{\square}(U, V) + \delta_{\square}(V, W) = 0$ . Por lo tanto,  $\delta_{\square}(U, W) = 0$ , esto es,  $U \sim W$ .

En conclusión,  $\sim$  es una relación de equivalencia. □

Por lo tanto, para cada núcleo  $W \in \mathcal{W}$ , podemos considerar la clase de equivalencia de  $W$  módulo  $\sim$ , es decir,

$$[W]_{\sim} = \{U \in \mathcal{W} \mid U \sim W\}.$$

DEFINICIÓN 2.61. Definimos  $\widetilde{\mathcal{W}} = \{[W]_{\sim} \mid W \in \mathcal{W}\}$ .

NOTACIÓN 2.62. Como es usual en estos casos, haremos un abuso notacional, pues denotaremos a un elemento  $[W]_{\sim} \in \widetilde{\mathcal{W}}$  simplemente por el núcleo representante  $W$ . Esto es, decimos que dos elementos  $U, W \in \widetilde{\mathcal{W}}$  pertenecen a  $\mathcal{W}$ , y son iguales si y solamente si se satisface  $\delta_{\square}(U, W) = 0$ .

Las complicaciones derivadas de usar la norma de corte en el espacio de núcleos y después minimizar sobre funciones que preservan la medida se justifican por el hecho importante de que la métrica  $\delta_{\square}$  en  $\widetilde{\mathcal{W}}$  define un espacio métrico compacto en el espacio  $\widetilde{\mathcal{W}}$  [Lov12, p. 132].

## 6. El lema de regularidad como un teorema de compacidad en el espacio de grafones

Nuestra aproximación a un núcleo será con el llamado “operador paso”.

DEFINICIÓN 2.63 (El operador paso). Supongamos que  $p$  es un entero positivo. Sea  $W \in \mathcal{W}$  y sea  $\mathcal{P} = \{S_1, \dots, S_p\}$  una partición de  $[0, 1]$  de cardinalidad  $p$  y con la propiedad  $\mathcal{P} \subseteq \mathcal{B}([0, 1])$ . Definimos la función  $W_{\mathcal{P}} : [0, 1]^2 \rightarrow \mathbb{R}$  como

$$W_{\mathcal{P}}(x, y) = \frac{1}{\lambda(S_i)\lambda(S_j)} \int_{S_i \times S_j} W(x, y) dx dy,$$

para  $x \in S_i, y \in S_j$  y  $\lambda(S_i), \lambda(S_j) > 0$ . con  $i, j \in [p]$ . En el caso  $\lambda(S_i) = 0$  o  $\lambda(S_j) = 0$ , para  $i, j \in [p], x \in S_i, y \in S_j$ , definimos  $W_{\mathcal{P}}(x, y) = 0$ .

NOTACIÓN 2.64. Dado un núcleo  $W \in \mathcal{W}$  y una partición  $\mathcal{P}$  del intervalo  $[0, 1]$ , el operador paso de  $W$ ,  $W_{\mathcal{P}}$ , es llamado un **escalonamiento** de  $W$ . En términos probabilísticos, al considerar el espacio de probabilidad  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$  y la partición  $\mathcal{P}^2$  de  $[0, 1]^2$ ,  $W_{\mathcal{P}}$  es precisamente  $\mathbb{E}(W | \mathcal{P}^2)$  (Ejemplo B.6).

**6.1. Lema débil de regularidad.** En esta sección enunciamos el lema débil de regularidad en la versión del área de análisis, después usaremos dicho lema dentro de la demostración del teorema de compacidad en el espacio de grafones.

Un hecho ampliamente conocido es que cada núcleo  $W \in \mathcal{W}$  puede ser aproximado, en la norma  $L^1$ , por funciones simples. La siguiente proposición asegura este hecho [Lov12, p. 145].

PROPOSICIÓN 2.65. Sea  $(\mathcal{P}_n)$  una sucesión de particiones de  $[0, 1]$  con la propiedad de que, para cada  $n \in \mathbb{N}$ ,  $\mathcal{P}_n \subseteq \mathcal{B}([0, 1])$  y tales que cada par de puntos es separado por todos salvo un número finito de particiones  $\mathcal{P}_n$ . Entonces, para cada  $W \in \mathcal{W}$ ,  $W_{\mathcal{P}_n} \rightarrow W$  casi donde sea.

El lema débil de regularidad para núcleos, probado por Frieze y Kannan en 1999 [FK99, p. 217], es un enunciado acerca de aproximación a una función medible mediante funciones simples en la norma de corte.

LEMA 2.66 (Lema débil de regularidad para núcleos). Para cada núcleo  $W \in \mathcal{W}$  y cada entero positivo  $k$  existe una función escalonada  $U \in \mathcal{W}$  con  $k$  pasos (Notación 2.28) tal que

$$\|W - U\|_{\square} \leq \frac{2}{\sqrt{\log k}} \|W\|_2.$$

Entonces cada núcleo puede ser aproximado por funciones simples. La Proposición 2.65 asegura algo similar sobre aproximación en la norma  $L^1$ . Dado que, para  $W \in \mathcal{W}$ , se cumple la desigualdad  $\|W\|_{\square} \leq \|W\|_1$ , aproximar en  $L^1$  parece un resultado más fuerte. Sin embargo, el error en la aproximación en la norma  $L^1$  depende no sólo del número de pasos de una función escalonada (Notación 2.28), sino de  $W$ . El hecho crucial sobre el Lema 2.66 es que el error tiende a cero conforme  $k$  tiende a infinito, uniformemente en  $W$  [Lov12, p. 145].

Una función escalonada aproximándose a un grafón  $W \in \mathcal{W}$  en el Lema 2.66 usualmente no es un escalonamiento de  $W$  (Notación 2.64). No obstante, tenemos el siguiente lema:

LEMA 2.67. Sea  $W \in \mathcal{W}_0$ , sea  $U \in \mathcal{W}$  una función escalonada, y sea  $\mathcal{P}$  una partición de  $[0, 1]$ , donde los elementos de  $\mathcal{P}$  son los pasos de  $U$  (Notación 2.28). Entonces

$$\|W - W_{\mathcal{P}}\|_{\square} \leq 2\|W - U\|_{\square}.$$

Los lemas 2.66 y 2.67 implican el siguiente corolario.

COROLARIO 2.68. Para cada grafón  $W \in \mathcal{W}_0$  y cada entero positivo  $k$  existe una partición  $\mathcal{P}$  de  $[0, 1]$  en a lo más  $k$  conjuntos con medida positiva tales que

$$\|W - W_{\mathcal{P}}\|_{\square} \leq \frac{2}{\sqrt{\log k}}.$$

DEFINICIÓN 2.69. Sean  $\epsilon > 0$  y  $W \in \mathcal{W}$ . Una partición  $\mathcal{P}$  de  $[0, 1]$  que cumple

$$\|W - W_{\mathcal{P}}\|_{\square} \leq \epsilon$$

es llamada una **partición débil regular** de  $W$  con error  $\epsilon$ .

Usaremos la siguiente versión del lema débil de regularidad (C. Borgs, J.T. Chayes y Y. Zhao).

TEOREMA 2.70 (Lema débil de regularidad). Para cualquier  $\epsilon > 0$  y cualquier grafón  $W : [0, 1]^2 \rightarrow \mathbb{R}$  existe una partición  $\mathcal{P}$  de  $[0, 1]$  que satisface  $\mathcal{P} \subseteq \mathcal{B}([0, 1])$ , cumple  $|\mathcal{P}| \leq 4^{1/\epsilon^2}$  y

$$\|W - W_{\mathcal{P}}\|_{\square} \leq \epsilon.$$

LEMA 2.71 (Incremento de energía en  $L^2$ ). Sea  $W \in \mathcal{W}_0$  un grafón y  $\mathcal{P}$  una partición del intervalo  $[0, 1]$ , tal que  $\|W - W_{\mathcal{P}}\|_{\square} > \epsilon$ . Existe un refinamiento  $\mathcal{Q}$  de  $\mathcal{P}$  con la propiedad

$$\|W_{\mathcal{Q}}\|_2^2 > \|W_{\mathcal{P}}\|_2^2 + \epsilon^2.$$

TEOREMA 2.72. Sean  $\epsilon > 0$ ,  $W \in \mathcal{W}_0$  un grafón y  $\mathcal{P}_0$  una partición de  $[0, 1]$ . Existe un refinamiento  $\mathcal{Q}$  de  $\mathcal{P}_0$  tal que  $|\mathcal{Q}| \leq 4^{1/\epsilon^2} |\mathcal{P}_0|$  que cumple

$$\|W - W_{\mathcal{Q}}\|_{\square} \leq \epsilon.$$

Definamos  $\widetilde{\mathcal{W}}_0$  como el espacio de grafones módulo la relación de equivalencia  $W \sim U$  si  $\delta_{\square}(W, U) = 0$ . Esto es,

$$\widetilde{\mathcal{W}}_0 = \{[W]_{\sim} \mid W \in \mathcal{W}_0\}.$$

Notemos que  $(\widetilde{\mathcal{W}}_0, \delta_{\square})$  es un espacio métrico. El lema débil de regularidad, en su versión analista, es una herramienta fundamental en la demostración del siguiente teorema de compacidad, el cual puede pensarse como la versión más fuerte de regularidad [Lov12, p. 149].

TEOREMA 2.73. El espacio métrico  $(\widetilde{\mathcal{W}}_0, \delta_{\square})$  es compacto.

La prueba del teorema de compacidad en el espacio de grafones (Teorema 2.73) expuesta en el presente trabajo es una prueba no muy formal, puesto que analizar sus argumentos con lujo de detalle requeriría de otra tesis. Fueron consultadas las notas de Zhao [Zha, p. 104] para la demostración.

DEMOSTRACIÓN DEL TEOREMA 2.73. Dado que estamos trabajando en un espacio métrico, basta probar que toda sucesión de grafones  $(W_n)$  tiene una subsucesión convergente a algún grafón. Fijemos entonces una sucesión de grafones  $(W_n)$ . Para cada  $n \in \mathbb{N}$  aplicamos el lema débil de regularidad (Teorema 2.72) repetidamente para obtener una sucesión de particiones

$$(43) \quad \mathcal{P}_{n,1}, \mathcal{P}_{n,2}, \mathcal{P}_{n,3}, \dots$$

tal que, para cualesquiera  $n, k \in \mathbb{N}$ ,

- (a)  $\mathcal{P}_{n,k+1}$  refina  $\mathcal{P}_{n,k}$ ;
- (b)  $|\mathcal{P}_{n,k}| = m_k$ , donde  $m_k$  depende solamente de  $k$ ; y
- (c)  $\|W_n - W_{n,k}\|_{\square} \leq 1/k$ , donde  $W_{n,k} = (W_n)_{\mathcal{P}_{n,k}}$ .

El lema débil de regularidad solamente garantiza que  $|\mathcal{P}_{n,k}| \leq m_k$ , pero si permitimos conjuntos de medida cero como elementos de la partición  $\mathcal{P}_{n,k}$ , podemos alcanzar la igualdad.

Inicialmente, cada elemento de una partición  $\mathcal{P}_{n,k}$  puede ser un conjunto medible arbitrario. Sin embargo, para cada  $n$ , podemos aplicar una función invertible que preserva la medida a  $W_{n,1}$  y a  $\mathcal{P}_{n,1}$  de tal forma que  $\mathcal{P}_{n,1}$  es una partición de  $[0, 1]$  en intervalos. Para cada  $k \geq 2$ , asumiendo que  $\mathcal{P}_{n,k-1}$  es una partición de  $[0, 1]$  en intervalos, podemos aplicar una función invertible que preserva la medida a  $W_{n,k}$  y  $\mathcal{P}_{n,k}$  de tal forma que  $\mathcal{P}_{n,k}$  es una partición de  $[0, 1]$  en intervalos, y refina a  $\mathcal{P}_{n,k-1}$ . Por inducción, tenemos entonces que  $\mathcal{P}_{n,k}$  consiste de intervalos para toda  $n, k$ . Las propiedades (a), (b) aún se cumplen. En tanto que la propiedad (c) puede dejar de satisfacerse, y tampoco puede ser cierto que  $W_{n,k} = (W_n)_{\mathcal{P}_{n,k}}$ , aún sabemos que  $\delta_{\square}(W_n, W_{n,k}) \leq 1/k$  para toda  $n, k$ . Esto será suficiente para nuestro fin.

Ahora, la parte crucial de la prueba es un argumento de diagonalización en numerables pasos. Comenzamos con la sucesión  $W_1, W_2, \dots$  y pasaremos a una subsucesión repetidamente. En el paso  $k$  seleccionamos una subsucesión  $W_{n_1}, W_{n_2}, \dots$  tal que:

1. los extremos de los intervalos de  $\mathcal{P}_{n_i,k}$  convergen individualmente cuando  $i \rightarrow \infty$ .
2.  $W_{n_i,k}$  converge puntualmente casi donde sea a un grafón  $U_k$  cuando  $i \rightarrow \infty$ .

Existe una subsucesión que satisface (1) dado que, para cada  $n, k \in \mathbb{N}$ , la partición  $\mathcal{P}_{n,k}$  tiene exactamente  $m_k$  elementos, y cada elemento tiene una longitud dentro del intervalo  $[0, 1]$ . Así, consideremos una subsucesión  $(W_{a_i})_{i=1}^{\infty}$  que satisface (1). Cada  $W_{a_i,k}$  puede ser naturalmente identificada con una función  $f_{a_i,k} : [m_k]^2 \rightarrow [0, 1]$ . El espacio de tales funciones está acotado, así que hay una subsucesión  $(f_{n_i})_{i=1}^{\infty}$  de  $(f_{a_i})_{i=1}^{\infty}$  que converge a alguna  $f : [m_k]^2 \rightarrow [0, 1]$ . Ahora  $f$  corresponde a un grafón  $U_k$  el cual es el límite de la sucesión  $(W_{n_i})_{i=1}^{\infty}$ . Por ende, (2) también se satisface.

Para concluir el paso  $k$ , la subsucesión es re-etiquetada como  $W_1, W_2, \dots$  e ignoramos los términos descartados. Las particiones correspondientes también son re-etiquetadas. Sin perder generalidad en el paso  $k$  pasamos a una subsucesión que contiene a  $W_1, \dots, W_k$ . Entonces, el resultado final de los pasos  $k = 1, 2, \dots$  es una sucesión infinita con la propiedad de que  $(W_{n,k})_{n=1}^{\infty}$  converge puntualmente casi donde sea a  $U_k$  para toda  $k$ .

Similarmente,  $(\mathcal{P}_{n,k})_{n=1}^{\infty}$  converge a un intervalo partición  $\mathcal{P}_k$  para toda  $k$ .

De la propiedad (a), cada partición  $\mathcal{P}_{n,k+1}$  refina a  $\mathcal{P}_{n,k}$ , lo cual implica

$$W_{n,k} = (W_{n,k+1})_{\mathcal{P}_{n,k}}.$$

Tomando  $n \rightarrow \infty$ , se sigue que  $U_k = (U_{k+1})_{\mathcal{P}_k}$ . Ahora, cada  $U_k$  puede pensarse como una variable aleatoria en el espacio de probabilidad  $([0, 1]^2, \mathcal{B}([0, 1]^2), \lambda^2)$ . Desde este punto de vista, las igualdades  $U_k = (U_{k+1})_{\mathcal{P}_{n,k}}$  implican que la sucesión  $U_1, U_2, \dots$  es una martingala.

El rango de cada  $U_k$  está contenido en el intervalo  $[0, 1]$ , así que la martingala es acotada. Por el teorema de convergencia de martingalas (Teorema B.14) existe un grafón  $U$  tal que  $U_k \rightarrow U$  puntualmente casi donde sea cuando  $k \rightarrow \infty$ .

Recordemos que nuestro objetivo era encontrar una subsucesión de  $W_1, W_2, \dots$  convergente a un grafón bajo  $\delta_{\square}$ . Hemos llegado a una subsucesión por el argumento de diagonalización de arriba, y afirmamos que tal subsucesión converge a  $U$  bajo  $\delta_{\square}$ . Es decir, queremos mostrar que  $\delta_{\square}(W_n, U) \rightarrow 0$  cuando  $n \rightarrow \infty$ . Esto se sigue por un "argumento estándar de 3-epsilons". Sea  $\epsilon > 0$ . Entonces existe  $k > 3/\epsilon$  con la propiedad  $\|U - U_k\|_1 < \epsilon/3$ , por convergencia puntual y el teorema de convergencia dominada. Dado que  $W_{n,k} \rightarrow U_k$  puntualmente casi donde sea (y por otra aplicación del teorema de convergencia dominada), existe  $n_0 \in \mathbb{N}$  tal que  $\|U_k - W_{n,k}\|_1 < \epsilon/3$  para toda  $n > n_0$ . Finalmente, dado que  $k > 3/\epsilon$ , sabemos que  $\delta_{\square}(W_n, W_{n,k}) < \epsilon/3$  para toda  $n$ . Concluimos que

$$\begin{aligned} \delta_{\square}(U, W_n) &\leq \delta_{\square}(U, U_k) + \delta_{\square}(U_k, W_{n,k}) + \delta_{\square}(W_{n,k}, W_n) \\ &\leq \|U - U_k\|_1 + \|U_k - W_{n,k}\|_1 + \delta_{\square}(W_{n,k}, W_n) \\ &\leq \epsilon. \end{aligned}$$

En la segunda desigualdad usamos

$$\delta_{\square}(W_1, W_2) \leq \|W_1 - W_2\|_{\square} \leq \|W_1 - W_2\|_1,$$

para grafones  $W_1, W_2$ .

□

## Aplicaciones del lema fuerte de regularidad

Aquí se exponen las primeras consecuencias del resultado principal (Teorema 1.40) de esta tesis. Primero se introduce la noción de eventos uniformes y su relación con la  $\mathcal{S}$ -norma uniforme (Definición 1.3), introducida en el Capítulo 1. La segunda sección está dedicada a la aplicación más importante de este capítulo, enunciada en la Proposición 3.7, y obtenida como consecuencia del lema fuerte de regularidad (versión probabilista) en virtud del Lema 3.6. La Proposición 3.7 brinda una partición finita del espacio muestral consistente de una subcolección de eventos uniformes.<sup>1</sup> En la última sección se enuncia el lema fuerte de regularidad, en su versión analista, para grafones, como una consecuencia del resultado principal en el Capítulo 1 (Teorema 1.40) a través del Corolario 1.42.<sup>2</sup>

### 1. Eventos uniformes

**DEFINICIÓN 3.1.** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Supongamos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $0 < \eta \leq 1$  y  $S \in \mathcal{S}$ . Decimos que el evento  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme si, para cada  $T \subseteq S$  con  $T \in \mathcal{S}$ , tenemos

$$\left| \int_T (X - \mathbb{E}(X|S)) d\mathbb{P} \right| \leq \eta \cdot \mathbb{P}(S).$$

Más aún, para cada  $\mathcal{C} \subseteq \mathcal{S}$  definimos  $\text{Unf}(\mathcal{C}, X, \eta) = \{C \in \mathcal{C} \mid C \text{ es } (X, \mathcal{S}, \eta)\text{-uniforme}\}$ .

En los ejemplos 3.2 y 3.3 suponemos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . A su vez, suponemos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $0 < \eta \leq 1$  y  $S \in \mathcal{S}$ .

**EJEMPLO 3.2.** Si  $S$  es un evento con probabilidad cero, entonces  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme.

**DEMOSTRACIÓN.** Supongamos un evento  $S \in \mathcal{S}$  tal que  $\mathbb{P}(S) = 0$ . Sea  $T \in \mathcal{S}$  con la propiedad  $T \subseteq S$ . Como  $\mathbb{P}(S) = 0$ , entonces, por monotonía, la probabilidad del evento  $T$  también es cero. Consecuentemente,

$$\int_T (X - \mathbb{E}(X|S)) d\mathbb{P} = 0.$$

De manera que si  $S \in \mathcal{S}$  tiene probabilidad cero, entonces el evento  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme para cada  $0 < \eta \leq 1$ .  $\square$

**EJEMPLO 3.3.** Si  $X$  es constante en  $S$ , entonces  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme.

**DEMOSTRACIÓN.** Supongamos que  $X$  es igual a una constante  $c \in \mathbb{R}$  en el evento  $S$ . Entonces la esperanza condicional de  $X$  dado  $S$ ,  $\mathbb{E}(X|S)$ , es igual a  $c$  (Proposición B.2), por lo que, para  $T \subseteq S$ ,

$$\int_T (X - \mathbb{E}(X|S)) d\mathbb{P} = \int_T (c - c) d\mathbb{P} = 0.$$

Así, si  $X$  es constante en  $S$ , el evento  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme para toda  $0 < \eta \leq 1$ .  $\square$

<sup>1</sup>Véase [DKK16, p. 16] para una aplicación de la Proposición 3.7 a hipercubos.

<sup>2</sup>Véase [DKK16, p. 17] para una demostración del lema fuerte de regularidad para grafones en el espacio  $L^p$ .

El siguiente lema muestra que el concepto de  $(X, \mathcal{S}, \eta)$ -uniformidad está relacionado con la definición de la  $\mathcal{S}$ -norma uniforme (Definición 1.3).

LEMA 3.4. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Supongamos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $0 < \eta \leq 1$  y un evento  $S \in \mathcal{S}$  con la propiedad  $\mathbb{P}(S) > 0$ . Entonces,  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme si y sólo si la variable aleatoria  $X - \mathbb{E}(X|S) \in L^1(\Omega, \mathcal{F}, \mathbb{P}(\cdot | S))$  satisface que su  $\mathcal{S}$ -norma uniforme es menor o igual a  $\eta$ .

DEMOSTRACIÓN. Supongamos que el evento  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme. Para facilitar la notación, denotemos a la medida de probabilidad condicional de  $\mathbb{P}$  relativa a  $S$  como  $\nu = \mathbb{P}(\cdot | S)$ . Entonces, de la desigualdad del triángulo,

$$\begin{aligned} \int |X - \mathbb{E}(X|S)| d\nu &\leq \int |X| d\nu + \int |\mathbb{E}(X|S)| d\nu \\ &= \int |X| d\nu + |\mathbb{E}(X|S)| \\ &\leq \int |X| d\mathbb{P} + |\mathbb{E}(X|S)|. \end{aligned}$$

La última desigualdad se debe al Ejemplo A.42, pues  $|X|$  es de vuelta una variable aleatoria positiva (pertenecer a  $L^+(\Omega, \mathcal{F})$ ) por la Proposición A.24. Luego, puesto que  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ , entonces  $\int |X| d\mathbb{P}$  y  $|\mathbb{E}(X|S)|$  son números reales. Consecuentemente,  $\int |X - \mathbb{E}(X|S)| d\nu < \infty$ . Además, si consideramos a  $\mathbb{E}(X|S)$  como una función constante, entonces, en virtud de las Proposiciones A.25 y A.26,  $X - \mathbb{E}(X|S)$  es medible. Por ende, se cumple  $X - \mathbb{E}(X|S) \in L^1(\Omega, \mathcal{F}, \nu)$ . Ahora, tomemos  $T \in \mathcal{S}$  arbitrario. Al ser  $\mathcal{S}$  un  $k$ -semianillo, se satisface  $S \cap T \in \mathcal{S}$ ; además, notemos que  $S \cap T$  es subconjunto de  $S$ . Por consiguiente, debido a que  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme,

$$\left| \int_{S \cap T} (X - \mathbb{E}(X|S)) d\mathbb{P} \right| \leq \eta \cdot \mathbb{P}(S).$$

Equivalentemente,

$$\left| \frac{1}{\mathbb{P}(S)} \int_{S \cap T} (X - \mathbb{E}(X|S)) d\mathbb{P} \right| \leq \eta.$$

Del Ejemplo A.39,

$$\left| \int_T (X - \mathbb{E}(X|S)) d\nu \right| = \left| \frac{1}{\mathbb{P}(S)} \int_{S \cap T} (X - \mathbb{E}(X|S)) d\mathbb{P} \right| \leq \eta.$$

La arbitrariedad de  $T$  nos permite deducir

$$\|X - \mathbb{E}(X|S)\|_{\mathcal{S}} := \sup \left\{ \left| \int_T (X - \mathbb{E}(X|S)) d\nu \right| : T \in \mathcal{S} \right\} \leq \eta.$$

Recíprocamente, supongamos que la variable aleatoria  $X - \mathbb{E}(X|S) \in L^1(\Omega, \mathcal{F}, \nu)$  satisface que su  $\mathcal{S}$ -norma uniforme es menor o igual a  $\eta$ . Sea  $R$  cualquier subconjunto, fijo, de  $S$  con  $R \in \mathcal{S}$ . Entonces, por hipótesis,

$$\left| \int_R (X - \mathbb{E}(X|S)) d\nu \right| \leq \|X - \mathbb{E}(X|S)\|_{\mathcal{S}} \leq \eta.$$

Nuevamente, haciendo uso del Ejemplo A.39,

$$(44) \quad \left| \frac{1}{\mathbb{P}(S)} \int_{S \cap R} (X - \mathbb{E}(X|S)) d\mathbb{P} \right| = \left| \int_R (X - \mathbb{E}(X|S)) d\nu \right| \leq \eta.$$

Por el hecho de que  $R$  es subconjunto de  $S$ ,  $S \cap R = R$ . Por lo tanto, de (44),

$$\left| \int_R (X - \mathbb{E}(X|S)) d\mathbb{P} \right| \leq \eta \cdot \mathbb{P}(S),$$

mostrando así que  $S$  es  $(X, \mathcal{S}, \eta)$ -uniforme.  $\square$

**COROLARIO 3.5.** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . Supongamos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $0 < \eta \leq 1$ . El espacio muestral,  $\Omega$ , es  $(X, \mathcal{S}, \eta)$ -uniforme si y sólo si la variable aleatoria  $X - \mathbb{E}(X) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  satisface  $\|X - \mathbb{E}(X)\|_{\mathcal{S}} \leq \eta$ .

**DEMOSTRACIÓN.** Debido a que  $\Omega$  es el espacio muestral, sabemos  $\mathbb{P}(\Omega) = 1$ . Así, por definición de probabilidad condicional (Ejemplo A.39), es fácil ver que la medida de probabilidad  $\mathbb{P}(\cdot | \Omega)$  coincide con la medida  $\mathbb{P}$ , es decir, tenemos la igualdad  $\mathbb{P}(\cdot | \Omega) = \mathbb{P}$ . Por esta razón, los espacios  $L^1(\Omega, \mathcal{F}, \mathbb{P}(\cdot | \Omega))$  y  $L^1(\Omega, \mathcal{F}, \mathbb{P})$  son iguales. También, de la definición de esperanza condicional dado un evento (Definición B.1), tenemos  $\mathbb{E}(X|\Omega) = \mathbb{E}(X)$ . Por lo tanto, del Lema 3.4,  $\Omega$  es  $(X, \mathcal{S}, \eta)$ -uniforme si y sólo si  $\|X - \mathbb{E}(X)\|_{\mathcal{S}} \leq \eta$ .  $\square$

## 2. Regularidad en Probabilidad

**LEMA 3.6.** Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $k$  un entero positivo y  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$ . También, sean  $\mathcal{P}$  una partición finita de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{F}$ ,  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $0 < \eta \leq 1$ . Supongamos que  $X$  admite una descomposición  $X = X_{est} + X_{err} + X_{unf}$  en variables aleatorias integrables con la propiedad de que  $X_{est}$  es constante en cada  $S \in \mathcal{P}$  y las funciones  $X_{err}$  y  $X_{unf}$  obedecen las estimaciones  $\|X_{err}\|_1 \leq \eta^2/8$  y  $\|X_{unf}\|_{\mathcal{S}} \leq (\eta^2/8)|\mathcal{P}|^{-1}$ . Entonces tenemos

$$\sum_{S \in \mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) \leq \eta.$$

**DEMOSTRACIÓN.** Supongamos  $S \in \mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)$ . Por definición, podemos seleccionar  $T \subseteq S$  con  $T \in \mathcal{S}$  tal que

$$(45) \quad \eta \cdot \mathbb{P}(S) < \left| \int_T (X - \mathbb{E}(X|S)) d\mathbb{P} \right|.$$

La función  $X_{est}$  es constante en  $S$  y por lo tanto, del Ejemplo 3.3,

$$\left| \int_T (X_{est} - \mathbb{E}(X_{est}|S)) d\mathbb{P} \right| = 0.$$

Por consiguiente, de (45) y la desigualdad del triángulo,

$$(46) \quad \begin{aligned} \eta \cdot \mathbb{P}(S) &< \left| \int_T (X_{est} + X_{err} + X_{unf} - \mathbb{E}(X_{est}|S) - \mathbb{E}(X_{err}|S) - \mathbb{E}(X_{unf}|S)) d\mathbb{P} \right| \\ &\leq \left| \int_T (X_{err} - \mathbb{E}(X_{err}|S)) d\mathbb{P} \right| + \left| \int_T (X_{unf} - \mathbb{E}(X_{unf}|S)) d\mathbb{P} \right|. \end{aligned}$$

Por un lado, usando la desigualdad de Jensen y la desigualdad del triángulo,

$$(47) \quad \left| \int_T (X_{err} - \mathbb{E}(X_{err}|S)) d\mathbb{P} \right| \leq \int_T |X_{err}| d\mathbb{P} + \int_T |\mathbb{E}(X_{err}|S)| d\mathbb{P}.$$

Continuando con la desigualdad (47), la Proposición A.33 asegura que

$$(48) \quad \begin{aligned} \int_T |X_{err}| d\mathbb{P} + \int_T |\mathbb{E}(X_{err}|S)| d\mathbb{P} &\leq \int_S |X_{err}| d\mathbb{P} + \int_S |\mathbb{E}(X_{err}|S)| d\mathbb{P} \\ &= \mathbb{E}(|X_{err}| | S) \mathbb{P}(S) + |\mathbb{E}(X_{err}|S)| \mathbb{P}(S). \end{aligned}$$

Y recordemos, de la Proposición B.3,  $|\mathbb{E}(X_{err}|S)| \leq \mathbb{E}(|X_{err}| | S)$ . En consecuencia, por las desigualdades (47) y (48),

$$(49) \quad \left| \int_T (X_{err} - \mathbb{E}(X_{err}|S)) d\mathbb{P} \right| \leq 2\mathbb{E}(|X_{err}| | S) \mathbb{P}(S).$$

Por otro lado, de la bien conocida desigualdad del triángulo y por definición de la  $\mathcal{S}$ -norma uniforme,

$$(50) \quad \left| \int_T (X_{\text{unf}} - \mathbb{E}(X_{\text{unf}}|S)) d\mathbb{P} \right| \leq \left| \int_T X_{\text{unf}} d\mathbb{P} \right| + \left| \int_T \mathbb{E}(X_{\text{unf}}|S) d\mathbb{P} \right| \\ \leq \|X_{\text{unf}}\|_{\mathcal{S}} + \mathbb{P}(T) |\mathbb{E}(X_{\text{unf}}|S)|.$$

La propiedad de monotonía de la medida nos dice que  $\mathbb{P}(T) \leq \mathbb{P}(S)$ , por lo cual  $\mathbb{P}(T)/\mathbb{P}(S) \leq 1$ . La afirmación anterior es consistente, pues ya vimos que si  $S$  tiene probabilidad cero, entonces  $S$  pertenece a  $\text{Unf}(\mathcal{P}, X, \eta)$  (Ejemplo 3.2), por lo que podemos deducir que  $\mathbb{P}(S) > 0$ . Así, tenemos lo siguiente:

$$(51) \quad \mathbb{P}(T) |\mathbb{E}(X_{\text{unf}}|S)| = \frac{\mathbb{P}(T)}{\mathbb{P}(S)} \left| \int_S X_{\text{unf}} d\mathbb{P} \right| \leq \left| \int_S X_{\text{unf}} d\mathbb{P} \right| \leq \|X_{\text{unf}}\|_{\mathcal{S}}.$$

Atando cabos, de las desigualdades (50) y (51),

$$(52) \quad \left| \int_T (X_{\text{unf}} - \mathbb{E}(X_{\text{unf}}|S)) d\mathbb{P} \right| \leq 2\|X_{\text{unf}}\|_{\mathcal{S}}.$$

De lo que hasta ahora hemos desarrollado, la desigualdad que hay que tener en cuenta es aquella que es el resultado de las desigualdades (46), (49) y (52), es decir,

$$(53) \quad \eta \cdot \mathbb{P}(S) < 2\mathbb{E}(|X_{\text{err}}| | S) \mathbb{P}(S) + 2\|X_{\text{unf}}\|_{\mathcal{S}}.$$

La arbitrariedad de  $S$  nos permite afirmar que lo anterior se satisface para todo elemento de  $\mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)$ .

Definamos los conjuntos

$$\mathcal{A} = \{S \in \mathcal{P} \mid \mathbb{E}(|X_{\text{err}}| | S) \geq \eta/4\} \text{ y } \mathcal{B} = \{S \in \mathcal{P} \mid \mathbb{P}(S) \leq 4\eta^{-1}\|X_{\text{unf}}\|_{\mathcal{S}}\}.$$

Afirmamos que  $\mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta) \subseteq \mathcal{A} \cup \mathcal{B}$ . En efecto, argumentando por contradicción, suponga que existe un elemento en la intersección  $(\mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)) \cap (\mathcal{P} \setminus (\mathcal{A} \cup \mathcal{B}))$ , a quien, para facilitar la notación, denotaremos nuevamente por  $S$ . En consecuencia,  $S$  cumple (53). También

$$\mathbb{E}(|X_{\text{err}}| | S) < \frac{\eta}{4} \text{ y } \mathbb{P}(S) > \frac{4\|X_{\text{unf}}\|_{\mathcal{S}}}{\eta}.$$

Equivalentemente,

$$2\mathbb{E}(|X_{\text{err}}| | S) \mathbb{P}(S) < \frac{\eta \mathbb{P}(S)}{2} \text{ y } \frac{\eta \mathbb{P}(S)}{2} > 2\|X_{\text{unf}}\|_{\mathcal{S}}.$$

Sustituyendo este último hecho en (53),

$$\eta \mathbb{P}(S) < \frac{\eta \mathbb{P}(S)}{2} + \frac{\eta \mathbb{P}(S)}{2},$$

una contradicción. Esto muestra que  $\mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta) \subseteq \mathcal{A} \cup \mathcal{B}$ . Ahora bien, para  $A \in \mathcal{A}$ ,  $\mathbb{E}(|X_{\text{err}}| | A) \geq \eta/4$  implica, de la definición de esperanza condicional dado un evento  $A$  (Definición B.1),

$$\mathbb{P}(A) \leq \frac{4}{\eta} \int_A |X_{\text{err}}| d\mathbb{P}.$$

Consecuentemente,

$$\sum_{A \in \mathcal{A}} \mathbb{P}(A) \leq \frac{4}{\eta} \sum_{A \in \mathcal{A}} \int_A |X_{\text{err}}| d\mathbb{P} \leq \frac{4}{\eta} \sum_{A \in \mathcal{P}} \int_A |X_{\text{err}}| d\mathbb{P} = \frac{4}{\eta} \int |X_{\text{err}}| d\mathbb{P},$$

donde usamos el Teorema A.32, pues  $\mathcal{P}$  consiste de conjuntos ajenos por pares al ser una partición (finita) de  $\Omega$ , además,  $\bigcup \mathcal{P} = \Omega$ . De modo que, como  $\|X_{\text{err}}\|_1$  es una cantidad acotada por  $\eta^2/8$ ,

$$(54) \quad \sum_{A \in \mathcal{A}} \mathbb{P}(A) \leq \frac{4}{\eta} \|X_{\text{err}}\|_1 \leq \frac{\eta}{2}.$$

A su vez,  $\mathcal{B}$  es subconjunto de  $\mathcal{P}$  y  $\|X_{\text{unf}}\|_{\mathcal{S}}$  también es una cantidad finita. Por lo tanto,

$$(55) \quad \sum_{B \in \mathcal{B}} \mathbb{P}(B) \leq \frac{4\|X_{\text{unf}}\|_{\mathcal{S}}}{\eta} |\mathcal{B}| \leq \frac{4\|X_{\text{unf}}\|_{\mathcal{S}}}{\eta} |\mathcal{P}| \leq \frac{\eta}{2}.$$

Finalmente, dado que  $\mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta) \subseteq \mathcal{A} \cup \mathcal{B}$ , por (54) y (55),

$$(56) \quad \sum_{S \in \mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) \leq \sum_{A \in \mathcal{A}} \mathbb{P}(S) + \sum_{B \in \mathcal{B}} \mathbb{P}(S) \leq \eta.$$

□

El Lema 3.6 nos permitirá reducir la siguiente proposición al lema fuerte de regularidad en su versión probabilista (Corolario 1.42).

**PROPOSICIÓN 3.7.** *Para cada entero positivo  $k$  y cada  $0 < \eta \leq 1$  existe un entero positivo  $U(k, \eta)$  con la siguiente propiedad. Si  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad,  $\mathcal{S}$  es un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$  y  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  con  $\|X\|_2 \leq 1$ , entonces existe un entero positivo  $M \leq U(k, \eta)$  y una partición  $\mathcal{P}$  de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{S}$  y  $|\mathcal{P}| = M$ , y satisface*

$$\sum_{S \in \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) \geq 1 - \eta.$$

**DEMOSTRACIÓN.** Fijemos un entero positivo  $k$  y un número real  $0 < \eta \leq 1$ . Definamos  $\sigma = \eta^2/8$  y  $F : \mathbb{N} \rightarrow \mathbb{N}$  mediante la regla  $F(n) = \lceil (n/\sigma) + 1 \rceil = \lceil (8n/\eta^2) + 1 \rceil$  para cada  $n \in \mathbb{N}$ . Para ver que  $F$  es creciente tomemos  $n, m \in \mathbb{N}$  con la propiedad  $n < m$ . Entonces

$$\frac{n}{\sigma} + 1 < \frac{m}{\sigma} + 1 \leq F(m).$$

Por consiguiente,  $F(n) \leq F(m)$ . Supongamos que se satisface  $F(n) = F(m)$ . Consecuentemente,

$$F(m) - 1 < \frac{m}{\sigma} + 1 \leq F(m) \quad \text{y} \quad F(m) - 1 < \frac{n}{\sigma} + 1 \leq F(m).$$

En consecuencia,

$$\left| \frac{8m - 8n}{\eta^2} \right| = \left| \frac{m - n}{\sigma} \right| < 1.$$

Así,  $|8m - 8n| < \eta^2 \leq 1$ , lo cual no puede suceder. Por ende,  $F(n) < F(m)$ ; es decir,  $F$  es una función creciente. Por otro lado,  $\sigma \leq 1$  implica  $n \leq n\sigma^{-1}$  para cualquier  $n \in \mathbb{N}$ . Así, para cada  $n \in \mathbb{N}$ , se cumple  $n + 1 \leq F(n)$ . Por tanto,  $F$  es una función de crecimiento. Definamos

$$U(k, \eta) = \text{Reg}'(k, \sigma, F).$$

Afirmamos que  $U(k, \eta)$  es el entero positivo buscado. Supongamos que  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $\mathcal{S}$  un  $k$ -semianillo en  $\Omega$  con  $\mathcal{S} \subseteq \mathcal{F}$  y  $X$  una variable aleatoria que pertenece a  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  con  $\|X\|_2 \leq 1$ . Por el Corolario 1.42, existe  $M$  un entero positivo menor o igual a  $U(k, \eta)$ , una partición  $\mathcal{P}$  de  $\Omega$  con  $\mathcal{P} \subseteq \mathcal{S}$  y  $|\mathcal{P}| = M$ , y un refinamiento finito  $\mathcal{Q}$  de  $\mathcal{P}$  con  $\mathcal{Q} \subseteq \mathcal{S}$  tal que, para

$$(57) \quad X_{\text{est}} = \mathbb{E}(X|\mathcal{P}), \quad X_{\text{err}} = \mathbb{E}(X|\mathcal{Q}) - \mathbb{E}(X|\mathcal{P}) \quad \text{y} \quad X_{\text{unf}} = X - \mathbb{E}(X|\mathcal{Q}),$$

tenemos las estimaciones  $\|X_{\text{err}}\|_2 \leq \sigma$  y  $\|X_{\text{unf}}\|_{\mathcal{S}} \leq 1/F(M)$ . Por monotonía de las normas  $L^p$  (Proposición A.50),  $\|X_{\text{err}}\|_1 \leq \sigma$ . A su vez, notemos que

$$\frac{8M}{\eta^2} \leq \frac{8M}{\eta^2} + 1 \leq F(M),$$

por lo cual,

$$\frac{1}{F(M)} \leq \frac{\eta^2}{8M}.$$

Por consiguiente, recordando que  $|\mathcal{P}| = M$ , tenemos  $\|X_{\text{unf}}\|_{\mathcal{S}} \leq (\eta^2/8) |\mathcal{P}|^{-1}$ . Como última observación,  $X_{\text{est}}$  es constante en cada  $S \in \mathcal{P}$  porque está definida como la esperanza condicional de  $X$  dada la  $\sigma$ -álgebra finita generada por la partición  $\mathcal{P}$  (Ejemplo B.6) y  $X_{\text{est}}$  es integrable por definición de esperanza condicional. A su vez,  $X_{\text{err}}$  y  $X_{\text{unf}}$  son integrables por ser el resultado de una combinación lineal de variables aleatorias integrables, pues  $L^1(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio vectorial.

Entonces,  $X$  admite una descomposición en variables aleatorias integrables. De este modo, el Lema 3.6 asegura que

$$(58) \quad \sum_{S \in \mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) \leq \eta.$$

Y, ya que  $\mathcal{P}$  es una partición,

$$\sum_{S \in \mathcal{P}} \mathbb{P}(S) = \mathbb{P}\left(\bigcup \mathcal{P}\right) = \mathbb{P}(\Omega) = 1.$$

Entonces,

$$\sum_{S \in \mathcal{P} \setminus \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) + \sum_{S \in \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) = 1.$$

Despejando lo anterior y usando la desigualdad (58) tenemos:

$$\sum_{S \in \text{Unf}(\mathcal{P}, X, \eta)} \mathbb{P}(S) \geq 1 - \eta.$$

Para poder despejar, en este último paso usamos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de medida finita.  $\square$

### 3. Lema fuerte de regularidad para grafones

DEFINICIÓN 3.8. *Supongamos que  $\Omega$  es un conjunto no vacío. Decimos que una función  $W : \Omega \times \Omega \rightarrow \mathbb{R}$  es **simétrica** si, para cualesquiera  $x, y \in \Omega$ ,  $W$  satisface  $W(x, y) = W(y, x)$ .*

DEFINICIÓN 3.9. *Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Un **núcleo** es una función  $W : \Omega \times \Omega \rightarrow \mathbb{R}$  que es acotada, simétrica y  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -medible.*

DEFINICIÓN 3.10. *Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Un **grafón** es un núcleo  $W : \Omega \times \Omega \rightarrow \mathbb{R}$  tal que  $0 \leq W(x, y) \leq 1$  para cualesquiera  $x, y \in \Omega$ .*

Otra consecuencia del Corolario 1.42 es el lema fuerte de regularidad para grafones, cuya demostración puede encontrarse en “Szemerédi’s regularity lemma via martingales” [DKK16].

COROLARIO 3.11 (Lema fuerte de regularidad para grafones). *Para cada  $0 < \epsilon \leq 1$  y cada función positiva  $h : \mathbb{N} \rightarrow \mathbb{R}$  existe un entero positivo  $s(\epsilon, h)$  con la siguiente propiedad. Si  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad y  $W : \Omega \times \Omega \rightarrow \mathbb{R}$  es un grafón con  $\|W\|_2 \leq 1$ , entonces existe una partición  $\mathcal{R}$  de  $\Omega$  con  $\mathcal{R} \subseteq \mathcal{F}$  y  $|\mathcal{R}| \leq s(\epsilon, h)$ , y  $U : \Omega \times \Omega \rightarrow \mathbb{R}$  un grafón tal que  $\|W - U\|_2 \leq \epsilon$  y  $\|U - U_{\mathcal{R}}\|_{\square} \leq h(|\mathcal{R}|)$ .*

## Gráficas con sucesión de grados dada

Este capítulo expone brevemente el artículo “Random graphs with a given degree sequence” [CDS11]. Comenzamos haciendo un recuento de algunos elementos de la teoría de límites de gráficas para, posteriormente, introducir la noción de un *límite de escalamiento*, representado por una función  $f : [0, 1] \rightarrow \mathbb{R}$ , correspondiente al límite de una sucesión de sucesiones de grados de gráficas,  $(\mathbf{d}^n)$ . El resultado principal de [CDS11] se enuncia en esta tesis como el Teorema 4.3 y concierne a una sucesión de gráficas aleatorias, con una sucesión de sucesiones de grados dada, a un grafón,  $W$ , descrito de manera explícita en términos de una cierta función  $g$ .

En seguida se enuncia y demuestra la Proposición 4.4, que caracteriza a un cierto tipo de límites de escalamiento y se interpreta como una versión continua del criterio de Erdős-Gallai en la teoría de gráficas tradicional; su demostración brinda, al mismo tiempo, un algoritmo para obtener una sucesión de grados de una gráfica. En virtud del algoritmo obtenido en la demostración de la Proposición 4.4, es posible dar una sucesión de grados de una gráfica mediante funciones continuas y acotadas. Por esta razón, la tercera sección de este cuarto capítulo muestra una simulación de gráficas aleatorias en base al Modelo de configuración.

Otro modelo de simulación de gráficas aleatorias relevante en este trabajo es el llamado Modelo  $\beta$ . Dicho modelo influye dentro de la prueba del Teorema 4.3 (resultado principal de [CDS11]). El Modelo  $\beta$  dicta las probabilidades para construir gráficas aleatorias. Cuando  $\beta$  es desconocido, el estimador máximo verosímil,  $\hat{\beta}$ , lleva a un algoritmo que permite encontrar  $g$ , y con ello, el grafón límite,  $W$ , del Teorema 4.3. El Teorema 4.9 asegura la existencia de un estimador  $\hat{\beta}$  con probabilidad alta. A su vez, el Teorema 4.10, garantiza una caracterización de todas las posibles sucesiones de grados esperadas del Modelo  $\beta$  conforme  $\beta$  varía en  $\mathbb{R}^n$ . En la última sección desarrollamos con rigor la demostración de un teorema (Teorema 4.12) sobre convergencia al estimador máximo verosímil,  $\hat{\beta}$ , suponiendo su existencia. La demostración del Teorema 4.12 es cercana a la prueba del teorema del punto fijo de Banach.

### 1. Introducción

Las gráficas de un orden suficientemente grande a menudo son estudiadas a través de su sucesión de grados. En el artículo “Random graphs from a given degree sequence” [CDS11] se estudian gráficas que son uniformemente escogidas con una sucesión de grados dada. Se ha mostrado que bajo ciertas condiciones tales gráficas tienen límites de gráficas en el sentido de Lovász y Szegedy [LS06] con límites identificables por una función conocida como grafón. Estudiaremos ecuaciones particulares de un modelo exponencial que tienen a las sucesiones de grados como estadística suficiente. En [CDS11] se muestra que el estimador máximo verosímil de los parámetros es único y consistente con alta probabilidad. Entonces, para un entero positivo  $n$ ,  $n$  parámetros pueden ser estimados consistentemente basados en una muestra de tamaño uno.

Supongamos que  $G$  es una gráfica de  $n$  vértices y sean  $d_1, \dots, d_n$  los grados de los vértices de  $G$ . Correspondientemente, la distribución de grados de  $G$  es la función de distribución acumulada  $F$  con soporte en  $[0, 1]$ , definida como

$$F(x) = \frac{|\{i : d_i \leq nx\}|}{n}.$$

En otras palabras, si un vértice es escogido uniformemente al azar, entonces el grado de ese vértice, dividido por  $n$ , es una variable aleatoria con función de distribución acumulada  $F$ .

Uno de los objetivos de [CDS11] es dar una descripción bastante precisa de la estructura de gráficas aleatorias densas vía la noción de gráficas límite introducida recientemente por Lovász y Szegedy [LS06] y desarrollada por Borgs et al. Esto da, en particular, una forma de escribir fórmulas exactas para el número esperado de subgráficas de un tipo dado sin simulación.

**1.1. Elementos de la teoría límite de gráficas.** Supongamos que  $(G_n)$  es una sucesión de gráficas simples tales que el número de vértices tiende a infinito. Para dos gráficas  $G$  y  $H$ , recordemos la densidad de homomorfismos de  $H$  en  $G$  definida como

$$t(H, G) = \frac{\text{hom}(H, G)}{v(G)^{v(H)}}.$$

En un espacio de probabilidad discreto  $(V(G)^{V(H)}, \mathcal{F}, \mathbb{P})$ , la densidad de homomorfismos de  $H$  en  $G$  es precisamente la probabilidad de que al elegir al azar una función  $f : V(H) \rightarrow V(G)$  ésta sea un homomorfismo.

En [LS06], se analizan los límites de la sucesión  $(G_n)$  mediante densidades de homomorfismos. Es decir, para cada gráfica  $H$ , las densidades de homomorfismo  $t(H, G_n)$  tienden a un límite  $t(H)$  conforme  $n \rightarrow \infty$ .

El resultado principal de Lovász y Szegedy [LS06] es que, en efecto, hay un “objeto límite” en el espacio de grafones  $\mathcal{W}_0$ . Inversamente, todo grafón es un objeto límite de una sucesión de gráficas apropiada. Este objeto límite determina todos los límites de densidades de subgráficas: si  $H$  es una gráfica simple con  $V(H) = [k]$ , entonces

$$\int_{[0,1]^k} \prod_{(i,j) \in E(H)} W(x_i, x_j) dx_1 \dots dx_k.$$

Por ejemplo, para la gráfica de Erdős-Renyi  $G_{n,p}$  si  $p$  está fijo y  $n \rightarrow \infty$ , entonces la gráfica límite es representada por la función que es idénticamente igual a  $p$  en  $[0, 1]^2$ .

## 2. Límites de escalamiento

**DEFINICIÓN 4.1.** *Supongamos que, para cada entero positivo  $n$ , una sucesión de grados  $\mathbf{d}^n = (d_1^n, \dots, d_n^n)$  está dada. Sin pérdida de generalidad supongamos  $d_1^n \geq d_2^n \geq \dots \geq d_n^n$ . Decimos que la sucesión  $(\mathbf{d}^n)$  tiene un **límite de escalamiento** si existe una función  $f : [0, 1] \rightarrow \mathbb{R}$  no creciente tal que*

$$(59) \quad \lim_{n \rightarrow \infty} \left( \left| \frac{d_1^n}{n} - f(0) \right| + \left| \frac{d_n^n}{n} - f(1) \right| + \frac{1}{n} \sum_{i=1}^n \left| \frac{d_i^n}{n} - f\left(\frac{i}{n}\right) \right| \right) = 0.$$

Notemos que la convergencia en este sentido puede ser enunciada equivalentemente en términos de convergencia de distribuciones de grados:  $d_1^n/n \rightarrow f(0)$ ,  $d_n^n/n \rightarrow f(1)$  y  $D_n/n \rightarrow f(U)$  en distribución, donde  $D_n$  es algún  $d_i^n$  escogido al azar y  $U$  se distribuye uniformemente en  $[0, 1]$ . La idea de controlar  $d_1^n$  y  $d_n^n$  viene de la necesidad de eliminar vértices atípicos que son adyacentes a demasiados o pocos vértices [CDS11, p. 1403]. Dado que  $d_i^n$  es decreciente en  $i$ , los vértices de grado atípico pueden ser eliminados simplemente controlando  $d_1^n$  y  $d_n^n$ .

**DEFINICIÓN 4.2.** *Definamos*

$$D^\downarrow[0, 1] := \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ es no creciente y } f \text{ es continua por la derecha en } (0, 1)\}.$$

Para cada  $n \in \mathbb{N}$ , sea  $G_n$  una gráfica aleatoria escogida uniformemente en el conjunto de gráficas simples con sucesión de grados  $\mathbf{d}^n$ . Supongamos que  $f$  es el límite de escalamiento de la

sucesión  $(\mathbf{d}^n)$  como en (59). Dotamos al conjunto de límites de escalamiento (es decir,  $D^\downarrow[0, 1]$ ) con la topología inducida por una modificación de la norma  $L^1$  dada por

$$\|f\|_{1'} := |f(0)| + |f(1)| + \int_0^1 |f(x)| dx.$$

La elección de esta norma es debido a la necesidad de hacerla compatible con nuestra noción de convergencia de sucesiones de grados.

No todas las funciones pueden ser límites de escalamiento de sucesiones de grados. Denotemos

$$\mathcal{F} := \{f \in D^\downarrow[0, 1] \mid \text{existe una sucesión de grados } (\mathbf{d}^n) \text{ tal que } f \text{ es límite de escalamiento de } (\mathbf{d}^n)\}.$$

El resultado principal en el artículo “Random graphs with a given degree sequence” [CDS11] es el siguiente:

**TEOREMA 4.3.** *Sea  $(G_n)$  una sucesión de gráficas aleatorias tal que, para cada  $n \in \mathbb{N}$ ,  $G_n$  es escogida uniformemente en el conjunto de gráficas simples con sucesión de grados  $\mathbf{d}^n$ . Supongamos que  $f$  es el límite de escalamiento de la sucesión  $(\mathbf{d}^n)$  como en (59). Supongamos también que  $f$  pertenece al interior topológico del conjunto  $\mathcal{F}$ . Entonces existe una única función  $g : [0, 1] \rightarrow \mathbb{R}$  en  $D^\downarrow[0, 1]$  tal que la función*

$$W(x, y) := \frac{e^{g(x)+g(y)}}{1 + e^{g(x)+g(y)}}$$

satisface, para toda  $x \in [0, 1]$ ,

$$f(x) = \int_0^1 W(x, y) dy.$$

En esta situación, la sucesión  $(G_n)$  converge casi seguramente al grafón  $W$ .

La demostración del Teorema 4.3 puede ser consultada en [CDS11]. El Teorema 4.3 puede ser útil sólo si damos una manera simple de verificar cuándo  $f$  pertenece al interior de  $\mathcal{F}$ . Si  $f$  es el límite de escalamiento de una sucesión de grados, entonces  $f \in \mathcal{F}$ . La pregunta es ¿cuándo  $f$  pertenece al interior de  $\mathcal{F}$ ? El siguiente resultado da una condición equivalente que es sencilla de verificar.

**PROPOSICIÓN 4.4** (Caracterización del interior de  $\mathcal{F}$ ). *Una función  $f : [0, 1] \rightarrow [0, 1]$  en  $D^\downarrow[0, 1]$  pertenece al interior de  $\mathcal{F}$  si y sólo si:*

- (i) *existen dos constantes  $c_1 > 0$  y  $c_2 < 1$  tales que  $c_1 \leq f(x) \leq c_2$  para todo  $x \in [0, 1]$ .*
- (ii) *para cada  $x \in (0, 1]$ ,*

$$\int_x^1 \min(f(y), x) dy + x^2 - \int_0^x f(y) dy > 0.$$

Demostremos la Proposición 4.4 después de las siguientes observaciones.

Observaciones:

1. La condición (ii) en el resultado previo es una versión continua del bien conocido criterio de Erdős-Gallai: Supongamos que  $d_1 \geq d_2 \geq \dots \geq d_n$  son enteros no negativos. El criterio de Erdős-Gallai nos dice que  $d_1, \dots, d_n$  puede ser la sucesión de grados de una gráfica simple de  $n$  vértices si y sólo si  $\sum_{i=1}^n d_i$  es par y para cada  $k \in [n]$ ,

$$\sum_{i=1}^k d_i \leq k(k-1) + \sum_{i=k+1}^n \min(d_i, k).$$

2. Como un ejemplo, consideremos el límite de la gráfica de Erdős-Renyi  $G(n, p)$  conforme  $n \rightarrow \infty$ . Entonces  $f(x) = p$  para todo  $x$ . La condición (ii) se vuelve  $(1-x) \min(p, x) + x^2 - px > 0$  para toda  $x$ . Considerando los casos  $x \geq p$  y  $x < p$  es posible verificar que las gráficas de Erdős-Renyi pertenecen al interior de  $\mathcal{F}$  para cualquier  $p$  fija,  $0 < p < 1$ .

DEFINICIÓN 4.5. Sea  $f \in D^\downarrow[0, 1]$ . Definimos  $G_f : [0, 1] \rightarrow \mathbb{R}$  como

$$G_f(x) := \int_x^1 \min(f(y), x) dy + x^2 - \int_0^x f(y) dy.$$

OBSERVACIÓN 4.6. Para cualesquiera  $f, f' \in D^\downarrow[0, 1]$  se satisface

$$\sup_{x \in [0, 1]} |G_f(x) - G_{f'}(x)| \leq \|f - f'\|_{1'}.$$

DEMOSTRACIÓN. Sean  $f, f' \in D^\downarrow[0, 1]$  y  $x \in [0, 1]$ . Por un lado, para todo  $y \in [x, 1]$ ,

$$(60) \quad \begin{aligned} \min(f(y), x) - \min(f'(y), x) &= \frac{f(y) + x - |f(y) - x|}{2} - \frac{f'(y) + x - |f'(y) - x|}{2} \\ &\leq \frac{f(y) - f'(y) - |f(y) - f'(y)|}{2} \leq f(y) - f'(y). \end{aligned}$$

Por otro lado,

$$\begin{aligned} |G_f(x) - G_{f'}(x)| &= \left| \int_x^1 \min(f(y), x) dy - \int_0^x f(y) dy - \int_x^1 \min(f'(y), x) dy + \int_0^x f'(y) dy \right| \\ &\leq \left| \int_x^1 \min(f(y), x) - \min(f'(y), x) dy \right| + \left| \int_0^x f(y) - f'(y) dy \right| \\ &\leq \int_x^1 |\min(f(y), x) - \min(f'(y), x)| dy + \int_0^x |f(y) - f'(y)| dy. \end{aligned}$$

Así, puesto que para cada  $y \in [x, 1]$  se cumple (60),

$$\begin{aligned} |G_f(x) - G_{f'}(x)| &\leq \int_x^1 |f(y) - f'(y)| dy + \int_0^x |f(y) - f'(y)| dy = \int_0^1 |f(y) - f'(y)| dy \\ &\leq \|f - f'\|_{1'}. \end{aligned}$$

Por lo tanto, como  $\|f - f'\|_{1'}$  es constante, al tomar supremo sobre todo  $x \in [0, 1]$  en ambos lados de la desigualdad obtenemos el resultado deseado.  $\square$

DEMOSTRACIÓN DE LA PROPOSICIÓN 4.4. La prueba de que una función  $f : [0, 1] \rightarrow [0, 1]$  en  $D^\downarrow[0, 1]$  pertenece al interior topológico de  $\mathcal{F}$  puede encontrarse en [CDS11]. Nos enfocaremos a demostrar el resultado inverso. Sea  $f \in D^\downarrow[0, 1]$  tal que, para toda  $x \in [0, 1]$ , se satisface  $0 < c_1 < f(x) < c_2 < 1$  y, para toda  $x \in (0, 1]$ ,  $G_f(x) > 0$ .

Probaremos primero  $f \in \mathcal{F}$ . Fijemos  $n \in \mathbb{N}$ . Definamos  $d_1^n = \lfloor nf(0) \rfloor$  y, para  $i = 2, \dots, n$ ,

$$(61) \quad d_i^n = \left\lfloor nf\left(\frac{i}{n}\right) \right\rfloor.$$

Dado que  $f$  es no creciente, tenemos  $d_1^n \geq d_2^n \geq \dots \geq d_n^n$ . Podemos incrementar  $d_1^n$  por 1 (ó incrementar cualquier otro  $d_i^n$  para  $i \in [n]$ ) si es necesario, de tal forma que  $\sum_{i=1}^n d_i^n$  sea un número par.

En el caso en que no se incremente  $d_1^n$  por una unidad,  $d_1^n \leq nf(0) < n$ , por lo que  $0 \leq f(0) - d_1^n/n < 1/n$ . Para el caso en que incrementemos  $d_1^n$  en una unidad,  $nf(0) \leq d_1^n \leq n$ . Total, en cualquier caso siempre podemos asegurar  $|f(0) - d_1^n/n| \leq 1/n$ . Análogamente, para cada  $i \in \{2, \dots, n\}$ , se cumple  $|f(i/n) - d_i^n/n| \leq 1/n$ , por lo que

$$\frac{1}{n} \sum_{i=2}^n \left| \frac{d_i^n}{n} - f\left(\frac{i}{n}\right) \right| \leq \frac{n-1}{n^2} \leq \frac{1}{n}.$$

Como  $f$  es no creciente,  $n > nf(0) \geq nf(1/n)$ , entonces  $n > d_1^n \geq nf(1/n)$  ó  $n > nf(1/n) \geq d_1^n$ ; en cualquier caso,  $|d_1^n/n - f(1/n)| \leq 1/n$ . Por lo tanto, con esta construcción obtenemos

$$(62) \quad \left| \frac{d_1^n}{n} - f(0) \right| + \left| \frac{d_n^n}{n} - f(1) \right| + \frac{1}{n} \sum_{i=1}^n \left| \frac{d_i^n}{n} - f\left(\frac{i}{n}\right) \right| \leq \frac{4}{n}.$$

Para cada  $n \in \mathbb{N}$  definamos al vector  $\mathbf{d}^n = (d_1^n, \dots, d_n^n)$  como en (61). La desigualdad (62) asegura que  $\mathbf{d}^n$  converge al límite de escalamiento  $f$  cuando  $n \rightarrow \infty$ . Falta mostrar que para cualquier  $n$  suficientemente grande,  $\mathbf{d}^n$  es, en efecto, una sucesión de grados de alguna gráfica.

Dado que  $f$  está acotada y es no creciente,

$$\lim_{n \rightarrow \infty} \int_0^1 |f(x) - f(\lceil nx \rceil/n)| dx = 0,$$

así que para  $k \in [n]$

$$\begin{aligned} & \left| \frac{\sum_{i=k+1}^n \min(d_i^n, k) + k(k-1) - \sum_{i=1}^k d_i^n}{n^2} - G_f(k/n) \right| \leq \\ & \leq \left| \sum_{i=k+1}^n \frac{1}{n} \min\left(\frac{d_i^n}{n}, \frac{k}{n}\right) - \int_{k/n}^1 \min\left(f(y), \frac{k}{n}\right) dy \right| + \frac{k}{n^2} + \left| \sum_{i=1}^k \frac{d_i^n}{n^2} - \int_0^{k/n} f(y) dy \right| \\ & \leq \left| \sum_{i=k+1}^n \frac{1}{n} \min\left(f\left(\frac{i}{n}\right), \frac{k}{n}\right) - \int_{k/n}^1 \min\left(f(y), \frac{k}{n}\right) dy \right| + \frac{k}{n^2} \\ & + \left| \sum_{i=2}^k \frac{1}{n} f\left(\frac{i}{n}\right) + \frac{1}{n} f(0) - \int_0^{k/n} f(y) dy \right| \\ & \leq \epsilon(n), \end{aligned}$$

donde  $\epsilon(n) \rightarrow 0$  conforme  $n \rightarrow \infty$ . Entonces existe una sucesión de enteros  $(k_0(n))$ , donde  $k_0(n)/n \rightarrow 0$  conforme  $n \rightarrow \infty$ , tales que, para todo  $k \geq k_0(n)$ , se satisface

$$\sum_{i=k+1}^n \min(d_i^n, k) + k(k-1) - \sum_{i=1}^k d_i^n > 0.$$

Además, existen  $c'_1 < 1$  y  $c'_2 > 0$  tales que para  $n$  suficientemente grande, se cumple  $c'_2 \leq d_i^n/n \leq c'_1$  para toda  $i \in [n]$ . Supongamos que para  $n$  suficientemente grande se cumple  $k_0(n)/n < c'_2$  y  $(1 - c'_1)n - k_0(n) > 0$ . Consecuentemente, si  $k \leq k_0(n)$ , tenemos

$$\begin{aligned} & \sum_{i=k+1}^n \min(d_i^n, k) + k(k-1) - \sum_{i=1}^k d_i^n \\ & \geq \sum_{i=k+1}^n \min(c'_2 n, k) + k(k-1) - \sum_{i=1}^k n c'_1 \\ & = (n-k)k + k(k-1) - c'_1 n k \\ & = ((1 - c'_1)n - k)k + k(k-1) > 0. \end{aligned}$$

En consecuencia, para  $n$  suficientemente grande, tenemos que para toda  $k \in [n]$ ,

$$\sum_{i=k+1}^n \min(d_i^n, k) + k(k-1) - \sum_{i=1}^k d_i^n > 0.$$

El criterio de Erdős-Gallai garantiza que  $(d_1^n, \dots, d_n^n)$  es una sucesión de grados válida. Es así como hemos probado que si  $f$  satisface las condiciones (i) y (ii) de la Proposición 4.4, entonces  $f$  debe pertenecer a  $\mathcal{F}$ . Ahora mostraremos que  $f$  está en el interior de  $\mathcal{F}$ .

Notemos que  $G_f$  es una función continua que es positiva en  $(0, 1]$ . Más aún, la Observación 4.6 garantiza que

$$\sup_{x \in [0, 1]} |G_f(x) - G_{f'}(x)| \leq \|f - f'\|_{1'}$$

para cualesquiera  $f, f' \in D^\downarrow[0, 1]$ ; así que si  $f_n \rightarrow f$  en  $\|\cdot\|_{1'}$ , entonces  $G_{f_n} \rightarrow G_f$  en la norma del supremo. Consecuentemente, para cualquier  $\epsilon > 0$  existe  $\delta > 0$  tal que  $\|h - f\|_{1'} < \delta$  implica  $G_h(x) > 0$  para toda  $x \in [\epsilon, 1]$ . También,  $c_1 - \delta \leq h(x) \leq c_2 + \delta$  para toda  $x \in [0, 1]$ . Escogiendo

$\delta, \epsilon$  lo suficientemente pequeñas, podemos asegurarnos que se cumple  $c_1 - \delta > \epsilon$  y  $1 - \epsilon - \delta - c_2 > 0$ . Fijemos tales  $\epsilon, \delta$  y  $h$ . Entonces para cualquier  $x \in (0, \epsilon)$ , tenemos

$$\begin{aligned} G_h(x) &\geq \int_x^1 \min(c_1 - \delta, x) dy + x^2 - \int_0^x (c_2 + \delta) dy \\ &= (1 - x)x + x^2 - (c_2 + \delta)x \\ &\geq (1 - \epsilon - \delta - c_2)x + x^2 > 0. \end{aligned}$$

Y también tenemos  $G_h(x) > 0$  para  $x \in [\epsilon, 1]$  por la elección de  $\delta$ . Es decir, hemos probado que existe  $\delta > 0$  con la propiedad de que si  $\|h - f'\|_{1'} < \delta$ , entonces  $G_h(x) > 0$  para toda  $x \in (0, 1]$ . Si escogemos  $\delta$  lo suficientemente pequeña, podemos asegurar que el rango de  $h$  no contiene a cero ni a uno.  $\square$

La Proposición 4.4 nos da un algoritmo para obtener sucesiones de grados. Recordemos

$$D^\downarrow[0, 1] := \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ es no creciente y } f \text{ es continua por la derecha en } (0, 1)\}.$$

Tomamos  $f \in D^\downarrow[0, 1]$ . Definamos, para cada  $n \in \mathbb{N}$ ,  $d_1^n = \lfloor nf(0) \rfloor$  y

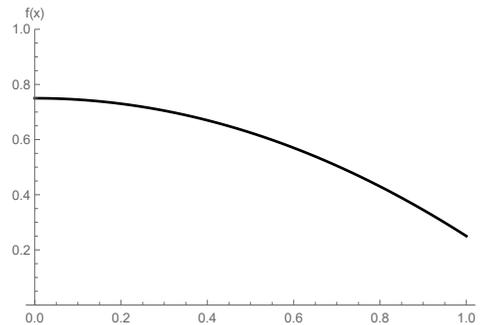
$$(63) \quad d_i^n = \left\lfloor n f\left(\frac{i}{n}\right) \right\rfloor, \text{ para } i = 2, \dots, n.$$

Podemos incrementar  $d_1^n$  por 1 (ó incrementar cualquier otro  $d_i^n$  para  $i \in [n]$ ) si es necesario, de tal forma que  $\sum_{i=1}^n d_i^n$  sea un número par.

### 3. Simulación mediante el Modelo de configuración

En los siguientes dos ejemplos se simularon gráficas con una sucesión de grados dada por una función no creciente. El código utilizado para nuestro fin se encuentra en el Apéndice C del presente trabajo.

EJEMPLO 4.7. Considere la función  $f : [0, 1] \rightarrow [0, 1]$  dada por  $f(x) = -x^2/2 + 3/4$ .



Para  $n = 10, 20, 30, 40$  construimos una sucesión de grados como en (63), denotada por  $\mathbf{d}^n = (d_1^n, \dots, d_n^n)$ . Una vez obtenida la sucesión de grados  $\mathbf{d}^n$ , simulamos una gráfica aleatoria con el siguiente código basado en el Modelo de configuración [vdH16, p. 139].

EJEMPLO 4.8. Considere la función  $f : [0, 1] \rightarrow [0, 1]$  dada por  $f(x) = 1/(x + 2)$ .

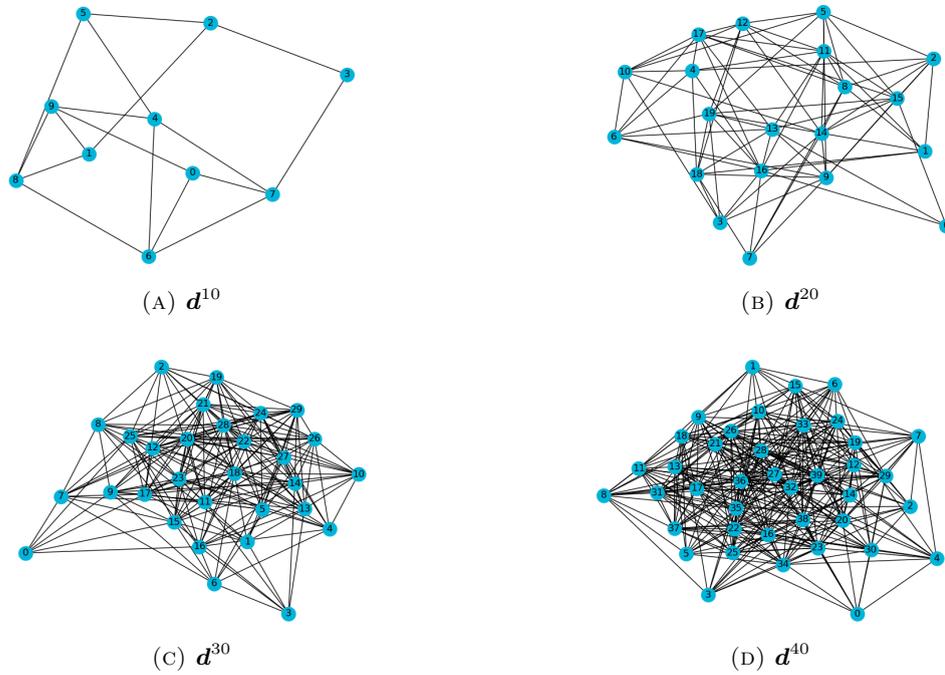
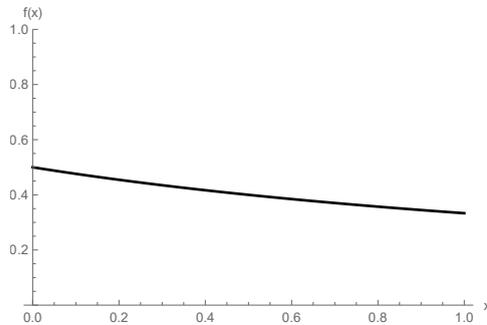


FIGURA 1. Simulación de gráficas aleatorias con una sucesión de grados dada por una función cóncava.



A diferencia del Ejemplo 4.7, en esta ocasión  $f$  es una función convexa. Notemos también que la densidad de aristas es ligeramente menor al ejemplo anterior.

La siguiente pregunta natural es si uno puede fácilmente describir la función  $g$  del Teorema 4.3 para una función  $f$  dada. Esto se vuelve un problema central en todo el análisis. De hecho, para probar el Teorema 4.3 analizamos un modelo estadístico relacionado. Resulta que obtener el estimador máximo verosímil en el modelo nos lleva a un algoritmo para encontrar  $g$ , lo que produce una prueba del Teorema 4.3. Analizaremos el modelo estadístico.

#### 4. Estadísticas con sucesiones de grados.

**4.1. Modelo  $\beta$ .** Dado un vector  $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n$  y un entero positivo  $n$ , sea  $\mathbb{P}_\beta$  la ley de una gráfica aleatoria no dirigida con conjunto de vértices  $[n]$  definida como sigue: para cada  $i, j \in [n]$  tales que  $i \neq j$ , ponemos una arista entre los vértices  $i$  y  $j$  con probabilidad

$$p_{ij} := \frac{e^{\beta_i + \beta_j}}{1 + e^{\beta_i + \beta_j}}$$

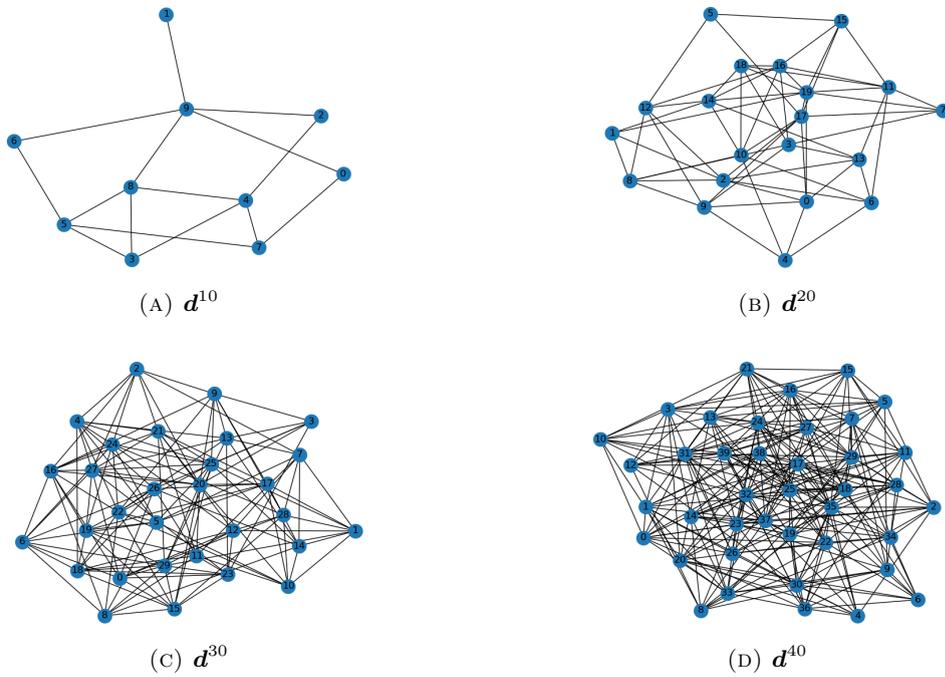


FIGURA 2. Simulación de gráficas aleatorias con una sucesión de grados dada por una función convexa.

e independientemente de todas las demás aristas. Entonces, si  $G$  es una gráfica con sucesión de grados  $d_1, \dots, d_n$ , la probabilidad de observar  $G$  bajo  $\mathbb{P}_\beta$  es

$$\frac{e^{\sum_i \beta_i d_i}}{\prod_{i < j} (1 + e^{\beta_i + \beta_j})}.$$

Este modelo de gráficas aleatorias es llamado Modelo  $\beta$ .

Supongamos que una gráfica aleatoria es generada del Modelo  $\beta$ , donde  $\beta \in \mathbb{R}^n$  es desconocido. ¿Es posible estimar  $\beta$  de una  $G$  observada? El estimador máximo verosímil,  $\hat{\beta}$ , de  $\beta$  debe satisfacer el sistema de ecuaciones

$$(64) \quad d_i = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{e^{\hat{\beta}_i + \hat{\beta}_j}}{1 + e^{\hat{\beta}_i + \hat{\beta}_j}}, \text{ para } i = 1, \dots, n,$$

donde  $d_1, \dots, d_n$  son los grados de los vértices de la gráfica observada  $G$ .

El siguiente teorema muestra que bajo ciertas condiciones en  $\beta$ , hay una alta probabilidad de que exista el estimador máximo verosímil, sea único y estime  $\beta$  con precisión uniforme en todas sus coordenadas.

**TEOREMA 4.9.** *Supongamos que  $G$  es una gráfica aleatoria dada por la medida  $\mathbb{P}_\beta$  y sea  $d_1, \dots, d_n$  la sucesión de grados de  $G$ . Sea  $L := \max_{1 \leq i \leq n} |\beta_i|$ . Entonces existe una constante  $C(L)$  que depende sólo de  $L$  tal que con probabilidad al menos  $1 - C(L)n^{-2}$ , existe una única solución  $\hat{\beta}$  a las ecuaciones (64), que satisface*

$$\max_{1 \leq i \leq n} |\hat{\beta}_i - \beta_i| \leq C(L) \sqrt{\frac{\log n}{n}}.$$

DEMOSTRACIÓN. La demostración de este teorema puede consultarse en [CDS11, p. 1414]  $\square$

El siguiente teorema caracteriza todas las posibles sucesiones de grados esperadas del modelo  $\beta$  conforme  $\beta$  varía en  $\mathbb{R}^n$ . Una buena característica es que no se deja fuera ninguna sucesión de grados.

**TEOREMA 4.10.** *Denotemos por  $\mathcal{R}$  al conjunto de todas las sucesiones de grados esperadas de gráficas aleatorias bajo la ley  $\mathbb{P}_\beta$  conforme  $\beta$  varía sobre  $\mathbb{R}^n$ . Denotemos por  $\mathcal{D}$  al conjunto de todas las sucesiones de grados posibles de gráficas no dirigidas de  $n$  vértices. Entonces*

$$\text{conv}(\mathcal{D}) = \overline{\mathcal{R}},$$

donde  $\text{conv}(\mathcal{D})$  denota la envolvente convexa de  $\mathcal{D}$  y  $\overline{\mathcal{R}}$  la cerradura topológica de  $\mathcal{R}$ .

La prueba del Teorema 4.10 es mostrada en [CDS11, p. 1413].

Por último describamos un algoritmo rápido para obtener el estimador máximo verosímil,  $\hat{\beta}$ , del Modelo  $\beta$  cuando  $\beta$  es desconocido, asumiendo su existencia.

## 5. Convergencia al estimador máximo verosímil del Modelo $\beta$

En esta sección demostramos con rigor el Teorema 4.12 que nos da un algoritmo para obtener el estimador máximo verosímil del Modelo  $\beta$ .

Recordemos que la norma  $L^\infty$  de un vector  $x = (x_1, \dots, x_n)$  está definida como

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|.$$

**NOTACIÓN 4.11.** Sean  $i, j \in [n]$  con  $i \neq j$ . Definimos  $r_{ij} : \mathbb{R}^n \rightarrow \mathbb{R}$  como

$$(65) \quad r_{ij}(x) := \frac{1}{e^{-x_j} + e^{x_i}}.$$

Dada una gráfica aleatoria con sucesión de grados  $(d_1, \dots, d_n)$ , definamos, para cada  $i \in [n]$ , la función  $\varphi_i : \mathbb{R}^n \rightarrow \mathbb{R}$  mediante la regla

$$\varphi_i(x) := \log(d_i) - \log \left( \sum_{\substack{j=1 \\ j \neq i}}^n r_{ij}(x) \right).$$

Definamos también  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  como la función dada por

$$(66) \quad \varphi(x) = (\varphi_1(x), \dots, \varphi_n(x)).$$

Notemos que si  $x$  es un punto fijo de  $\varphi$ , entonces para  $i \in [n]$ ,

$$\log(d_i) = x_i + \log \left( \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{e^{-x_j} + e^{x_i}} \right).$$

Equivalentemente,

$$d_i = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{e^{x_i}}{e^{-x_j} + e^{x_i}} = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{e^{x_i + x_j}}{1 + e^{x_i + x_j}}.$$

De manera que los puntos fijos de  $\varphi$  son precisamente las soluciones de (64).

**TEOREMA 4.12.** *Sea  $n$  un entero positivo. Dada una gráfica aleatoria con sucesión de grados  $(d_1, \dots, d_n)$ , definamos  $\varphi$  como en (66). Supongamos que las ecuaciones del estimador máximo verosímil (64) tienen una solución  $\hat{\beta}$ . Entonces  $\hat{\beta}$  es un punto fijo de la función  $\varphi$ . Comenzando desde cualquier  $x_0 \in \mathbb{R}^n$ , defina  $x_{k+1} = \varphi(x_k)$  para  $k = 0, 1, 2, \dots$ . Entonces  $x_k$  converge a  $\hat{\beta}$*

en la norma  $L^\infty$  a una velocidad geométrica donde la tasa depende sólo de  $(\|\hat{\beta}\|_\infty, \|x_0\|_\infty)$ . En particular,  $\hat{\beta}$  debe ser la única solución de (64). Más aún,

$$\|x_0 - \hat{\beta}\|_\infty \leq C\|x_0 - x_1\|_\infty,$$

donde  $C$  es una función constante del par  $(\|\hat{\beta}\|_\infty, \|x_0\|_\infty)$ . Por el contrario, si las ecuaciones del estimador máximo verosímil (64) no tienen una solución, entonces la sucesión  $(x_k)$  debe tener una subsucesión divergente.

Antes de proceder con todo rigor con la demostración del Teorema 4.12 serán necesarios un par de detalles técnicos. Comenzamos enunciando una sencilla observación.

**OBSERVACIÓN 4.13.** Para cualesquiera números reales  $a, b, c, d$  que satisfagan  $a \leq b < 0 < c \leq d$  se cumple  $ad \leq bc$ .

**DEMOSTRACIÓN.** Sean  $a, b, c, d \in \mathbb{R}$  tales que  $a \leq b < 0 < c \leq d$ . Entonces, como  $b$  es un número negativo,  $a/b \geq 1$  y  $c/d \leq 1$ . Por ende,  $c/d \leq a/b$ . Esto implica  $c \leq (ad)/b$ . Por lo tanto, nuevamente porque  $b$  es un número negativo,  $bc \geq ad$ .  $\square$

Recordemos que para una matriz  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ , la norma  $\|\cdot\|_\infty$  se define como

$$\|A\|_\infty := \max_{\|x\|_\infty \leq 1} \|Ax\|_\infty.$$

Será útil para la demostración del Teorema 4.12 la siguiente equivalencia para la norma  $\|\cdot\|_\infty$  en el espacio  $\mathbb{R}^{n \times n}$ .

**PROPOSICIÓN 4.14.** Sea  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ . Entonces

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

**DEMOSTRACIÓN.** Veamos

$$\max_{\|x\|_\infty \leq 1} \|Ax\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Sea  $x \in \mathbb{R}^n$  tal que  $\|x\|_\infty \leq 1$ . Por definición de  $\|\cdot\|_\infty$ ,

$$\|Ax\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij}x_j \right|.$$

Fijemos  $i \in [n]$ . Luego, para cada  $j \in [n]$ ,  $|x_j| \leq \|x\|_\infty \leq 1$  implica  $|a_{ij}| |x_j| \leq |a_{ij}|$ . Entonces

$$\left| \sum_{j=1}^n a_{ij}x_j \right| \leq \sum_{j=1}^n |a_{ij}x_j| \leq \sum_{j=1}^n |a_{ij}|.$$

Por consiguiente,

$$\|Ax\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Se sigue que

$$\|A\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Ahora mostraremos la desigualdad restante. Fijemos  $i \in [n]$  arbitrario. Mostraremos

$$\sum_{j=1}^n |a_{ij}| \leq \|A\|_\infty.$$

Consideremos, para cada  $j \in [n]$ ,  $y_j = \text{sgn}(a_{ij}) \cdot 1$  si  $a_{ij} \neq 0$  y  $y_j = 0$  si  $a_{ij} = 0$ . Definamos  $y \in \mathbb{R}^n$  como  $y = (y_1, \dots, y_n)$ . Notemos  $\|y\|_\infty \leq 1$ .

Puesto que se satisface

$$\|Ay\|_\infty \geq \left| \sum_{j=1}^n a_{ij} y_j \right| = \left| \sum_{j=1}^n |a_{ij}| \right| = \sum_{j=1}^n |a_{ij}|,$$

entonces

$$\|A\|_\infty \geq \|Ay\|_\infty \geq \sum_{j=1}^n |a_{ij}|.$$

Dado que  $i \in [n]$  fue arbitrario, tenemos

$$\|A\|_\infty \geq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

desigualdad que termina con la demostración.  $\square$

Para facilitar la notación definamos un tipo particular de matrices en  $\mathbb{R}^{n \times n}$ ; posteriormente veremos algunas propiedades.

DEFINICIÓN 4.15. Sea  $\delta > 0$ . Diremos que una matriz  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  pertenece a la clase  $\mathcal{L}_n(\delta)$  si  $\|A\|_\infty \leq 1$  y para cada  $i, j \in [n]$  con  $i \neq j$  se satisface

$$a_{ii} \geq \delta \text{ y } a_{ij} \leq -\frac{\delta}{n-1}.$$

LEMA 4.16. Sea  $\delta > 0$ . Si  $A, B \in \mathcal{L}_n(\delta)$ , entonces

$$\|AB\|_\infty \leq 1 - \frac{2(n-2)\delta^2}{n-1}.$$

DEMOSTRACIÓN. Fijemos  $i, k \in [n]$  tales que  $i \neq k$ . Sean  $A = (a_{ij}), B = (b_{ij}) \in \mathcal{L}_n(\delta)$ . Por definición,

$$b_{ik} \leq -\frac{\delta}{n-1} < 0 < \delta \leq a_{ii}.$$

Consecuentemente,  $a_{ii}b_{ik} \leq -\frac{\delta^2}{n-1}$  (Observación 4.13). La desigualdad  $a_{ik}b_{kk} \leq -\frac{\delta^2}{n-1}$  es análoga. Por ende,

$$(67) \quad a_{ii}b_{ik} + a_{ik}b_{kk} \leq -\frac{2\delta^2}{n-1}.$$

Sea  $j \in [n] \setminus \{i, k\}$  arbitrario. Directamente de la Definición 4.15, al multiplicar dos desigualdades, tenemos

$$a_{ij}b_{jk} \geq \frac{\delta^2}{(n-1)^2}.$$

Dado que  $j$  fue arbitraria,

$$(68) \quad \sum_{\substack{j=1 \\ j \neq i, k}}^n a_{ij}b_{jk} \geq \frac{(n-2)\delta^2}{(n-1)^2}.$$

Definamos  $x = \sum_{j \in [n] \setminus \{i, k\}} a_{ij}b_{jk}$  y  $y = -(a_{ii}b_{ik} + a_{ik}b_{kk})$ . La desigualdad (68) implica

$$(69) \quad -2x \leq \frac{-2(n-2)\delta^2}{(n-1)^2}.$$

De (67) tenemos una desigualdad para  $y$ , a saber,

$$y \geq \frac{2\delta^2}{n-1} \geq \frac{2\delta^2(n-2)}{(n-1)^2},$$

Además, de esto último vemos que se cumple  $y > 0$ , lo que implica  $y \leq 2y$ . De modo que

$$(70) \quad -2y \leq -y \leq -\frac{2\delta^2(n-2)}{(n-1)^2}.$$

En cualquier caso, (69) y (70) muestran que

$$(71) \quad -2 \min(x, y) \leq -\frac{2\delta^2(n-2)}{(n-1)^2}.$$

Es un hecho básico de cálculo que para cualquier par  $x, y$  de números reales positivos se cumple  $|x - y| = |x| + |y| - 2 \min(x, y)$ . Así, con nuestras definiciones previas de los números  $x, y$ ,

$$\left| \sum_{j=1}^n a_{ij} b_{jk} \right| = \left| \sum_{\substack{j=1 \\ j \neq i, k}}^n a_{ij} b_{jk} \right| + |(a_{ii} b_{ik} + a_{ik} b_{kk})| - 2 \min(x, y).$$

De la desigualdad del triángulo,

$$(72) \quad \left| \sum_{j=1}^n a_{ij} b_{jk} \right| \leq \sum_{j=1}^n |a_{ij} b_{jk}| - 2 \min \left( \sum_{\substack{j=1 \\ j \neq i, k}}^n a_{ij} b_{jk}, -(a_{ii} b_{ik} + a_{ik} b_{kk}) \right) \\ \leq \sum_{j=1}^n |a_{ij} b_{jk}| - \frac{2\delta^2(n-2)}{(n-1)^2},$$

donde en la última desigualdad usamos lo estimado en (71).

Nuestra equivalencia en la Proposición 4.14 para la norma  $\|\cdot\|_\infty$  en el espacio  $\mathbb{R}^{n \times n}$  nos dice que

$$\|AB\|_\infty = \max_{1 \leq i \leq n} \sum_{k=1}^n \left| \sum_{j=1}^n a_{ij} b_{jk} \right|.$$

De esta forma, por (72),

$$(73) \quad \|AB\|_\infty \leq \max_{1 \leq i \leq n} \left( \sum_{k=1}^n \left( \sum_{j=1}^n |a_{ij} b_{jk}| - \frac{2\delta^2(n-2)}{(n-1)^2} \right) \right) \\ = \max_{1 \leq i \leq n} \left( \sum_{k=1}^n \left( \sum_{j=1}^n |a_{ij} b_{jk}| \right) \right) - \frac{2n\delta^2(n-2)}{(n-1)^2}.$$

Para afinar los últimos detalles observemos que  $-n \leq -(n-1)$  implica

$$(74) \quad -\frac{2n(n-2)\delta^2}{(n-1)^2} \leq -\frac{2(n-1)(n-2)\delta^2}{(n-1)^2} = -\frac{2(n-2)\delta^2}{(n-1)}.$$

Y debido a la hipótesis  $\|A\|_\infty, \|B\|_\infty \leq 1$ ,

$$(75) \quad \max_{1 \leq i \leq n} \left( \sum_{k=1}^n \left( \sum_{j=1}^n |a_{ij} b_{jk}| \right) \right) = \max_{1 \leq i \leq n} \left( \sum_{j=1}^n |a_{ij}| \left( \sum_{k=1}^n |b_{jk}| \right) \right) \\ \leq \max_{1 \leq i \leq n} \left( \sum_{j=1}^n |a_{ij}| \cdot \|B\|_\infty \right) \\ \leq \max_{1 \leq i \leq n} \left( \sum_{j=1}^n |a_{ij}| \cdot 1 \right) \leq 1.$$

Por último sustituimos (74) y (75) en (73) para mostrar

$$\|AB\|_\infty \leq 1 - \frac{2(n-2)\delta^2}{(n-1)}.$$

□

Veremos que el Lema 4.16 es clave para la demostración del Teorema 4.12.

Por último, enunciaremos un lema algo técnico, cuya demostración se puede encontrar después de la demostración del Teorema 4.12.

LEMA 4.17. Sea  $(d_1, \dots, d_n)$  una sucesión de grados de una gráfica. Definamos, para cada  $i \in [n]$ ,

$$\varphi_i(x) := \log(d_i) - \log \left( \sum_{\substack{j=1 \\ j \neq i}}^n r_{ij}(x) \right).$$

Definamos también  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  como la función dada por

$$(76) \quad \varphi(x) = (\varphi_1(x), \dots, \varphi_n(x)).$$

Sean  $K \in \mathbb{R}_{\geq 0}$ ,  $i, j \in [n]$  con  $i \neq j$ . Entonces, para todo  $x \in \mathbb{R}^n$ , se satisface

$$\frac{\partial \varphi_i(x)}{\partial x_j} \leq 0 \quad y \quad \frac{\partial \varphi_i(x)}{\partial x_i} \geq 0.$$

Además, para cada  $x \in \mathbb{R}^n$  con la propiedad  $\|x\|_\infty \leq K$ , se cumple

$$(77) \quad \frac{e^{-4K}}{2} \leq \frac{\partial \varphi_i(x)}{\partial x_i} \leq e^{2K}$$

y

$$(78) \quad -\frac{e^{2K}}{n-1} \leq \frac{\partial \varphi_i(x)}{\partial x_j} \leq -\frac{e^{-4K}}{2(n-1)}.$$

Procedamos entonces con la esperada demostración del Teorema 4.12.

DEMOSTRACIÓN DEL TEOREMA 4.12. Fijemos  $x, y \in \mathbb{R}^n$  arbitrarios. Sea  $J(x, y)$  la matriz cuya entrada  $J_{ij}(x, y)$  es

$$J_{ij}(x, y) = \int_0^1 \frac{\partial \varphi_i}{\partial x_j}(tx + (1-t)y) dt.$$

La  $i$ -ésima entrada del vector  $J(x, y)(x - y)$  es igual a  $\sum_{j=1}^n J_{ij}(x, y)(x_j - y_j)$ . Equivalentemente, por linealidad de la integral, regla de la cadena y el teorema fundamental del cálculo,

$$\begin{aligned} \sum_{j=1}^n J_{ij}(x, y)(x_j - y_j) &= \int_0^1 \sum_{j=1}^n \frac{\partial \varphi_i}{\partial x_j}(tx + (1-t)y)(x_j - y_j) dt \\ &= \int_0^1 \frac{d}{dt} \varphi_i(tx + (1-t)y) dt \\ &= \varphi_i(x) - \varphi_i(y). \end{aligned}$$

De modo que

$$\varphi(x) - \varphi(y) = J(x, y)(x - y).$$

El Lema 4.17 asegura que, para cada  $i, j \in [n]$  con  $i \neq j$ ,  $\frac{\partial \varphi_i}{\partial x_j}$  es una función menor o igual a cero y  $\frac{\partial \varphi_i}{\partial x_i}$  es mayor o igual a cero. Más aún, para cada  $i \in [n]$ ,

$$\sum_{j=1}^n \left| \frac{\partial \varphi_i}{\partial x_j} \right| = \frac{\partial \varphi_i}{\partial x_i} - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\partial \varphi_i}{\partial x_j} \equiv 1.$$

Se sigue que para cada  $i \in [n]$  y cada  $x, y$ ,  $\sum_{j=1}^n |J_{ij}(x, y)| = 1$ . Consecuentemente,  $\|J(x, y)\|_\infty = 1$  (Proposición 4.14). Definamos  $K$  como el máximo de las normas  $L^\infty$  de  $x, y, \varphi(x)$  y  $\varphi(y)$ . Observemos que, para  $t \in [0, 1]$ ,

$$\|tx + (1-t)y\|_\infty \leq t\|x\|_\infty + (1-t)\|y\|_\infty \leq tK + (1-t)K = K.$$

Por ende, el Lema 4.17 asegura también

$$\frac{e^{-4K}}{2} \leq \frac{\partial \varphi_i(tx + (1-t)y)}{\partial x_i} \quad y \quad \frac{\partial \varphi_i(tx + (1-t)y)}{\partial x_j} \leq -\frac{e^{-4K}}{2(n-1)},$$

para toda  $t \in [0, 1]$ . Integrando de 0 a 1,

$$\frac{e^{-4K}}{2} \leq J_{ii} \quad \text{y} \quad J_{ij} \leq -\frac{e^{-4K}}{2(n-1)}.$$

Y como  $\|J(x, y)\|_\infty = 1$ , vemos que, para  $\delta = \frac{1}{2}e^{-4K}$ ,  $J(x, y) \in \mathcal{L}_n(\delta)$ . Similarmente,

$$\begin{aligned} \varphi(\varphi(x)) - \varphi(\varphi(y)) &= J(\varphi(x), \varphi(y))(\varphi(x) - \varphi(y)) \\ &= J(\varphi(x), \varphi(y))J(x, y)(x - y) \end{aligned}$$

y también  $J(\varphi(x), \varphi(y)) \in \mathcal{L}_n(\delta)$ . Aplicando el Lema 4.16 tenemos

$$\|J(\varphi(x), \varphi(y))J(x, y)\|_\infty \leq 1 - \frac{2(n-2)\delta^2}{n-1}.$$

Consecuentemente,

$$(79) \quad \|\varphi(\varphi(x)) - \varphi(\varphi(y))\|_\infty \leq \left(1 - \frac{2(n-2)\delta^2}{n-1}\right) \|x - y\|_\infty.$$

Note que  $\delta$  depende de  $K$ , que a su vez depende de  $x$  y de  $y$ . Entonces podemos denotar

$$\theta(x, y) = 1 - \frac{2(n-2)\delta^2}{n-1}.$$

Observemos que  $0 \leq \theta < 1$  y  $\theta$  está uniformemente acotada por 1 en subconjuntos de  $\mathbb{R}^n \times \mathbb{R}^n$ . Más aún, dado que  $\|J(x, y)\|_\infty = 1$ , también tenemos la cota trivial

$$\|x - y\|_\infty \leq \|x - y\|_\infty.$$

Como  $x, y \in \mathbb{R}^n$  fueron arbitrarios, entonces lo anterior se satisface para todo  $x, y \in \mathbb{R}^n$ . Ahora supongamos que  $\varphi$  tiene un punto fijo  $\hat{\beta}$ . Si comenzamos con un  $x_0$  arbitrario y definimos  $x_{k+1} = \varphi(x_k)$  para cada  $k \geq 0$ , entonces, para cada  $k \in \mathbb{N}$ , tenemos

$$\|x_{k+1} - \hat{\beta}\|_\infty = \|\varphi(x_k) - \varphi(\hat{\beta})\|_\infty \leq \|x_k - \hat{\beta}\|_\infty.$$

En particular, la sucesión  $(x_k)$  permanece acotada.

Entonces existe un  $\theta \in [0, 1)$ , que depende sólo de  $\|\hat{\beta}\|_\infty$  y  $\|x_0\|_\infty$  tal que para todo  $k \geq 0$ , tenemos

$$(80) \quad \|x_{k+3} - x_{k+2}\|_\infty \leq \theta \|x_{k+1} - x_k\|_\infty$$

y

$$(81) \quad \|x_{k+2} - \hat{\beta}\|_\infty \leq \theta \|x_k - \hat{\beta}\|_\infty.$$

De la desigualdad del triángulo,

$$\|x_0 - \hat{\beta}\|_\infty \leq \sum_{j=0}^k (\|x_j - x_{j+1}\|_\infty) + \|x_{k+1} - \hat{\beta}\|_\infty.$$

De (81), vemos que  $x_k \rightarrow \hat{\beta}$  cuando  $k \rightarrow \infty$ . En consecuencia,

$$\|x_0 - \hat{\beta}\|_\infty \leq \sum_{j=0}^{\infty} \|x_j - x_{j+1}\|_\infty.$$

Y por (80),

$$\|x_0 - \hat{\beta}\|_\infty \leq \frac{1}{1-\theta} (\|x_0 - x_1\|_\infty + \|x_1 - x_2\|_\infty) \leq \frac{2}{1-\theta} \|x_0 - x_1\|_\infty.$$

Por último, notemos que si  $\hat{\beta}$  no existe, entonces la sucesión  $(x_k)$  debe tener una subsucesión convergente. De otra manera, argumentando por contrapositiva, la sucesión está acotada; esto es, existe  $K \in \mathbb{R}_{\geq 0}$  con la propiedad  $x_k \in \overline{B_K(0_{\mathbb{R}^n})}$  para toda  $k \in \mathbb{N}$ , donde  $0_{\mathbb{R}^n}$  es el vector en  $\mathbb{R}^n$  cuyas entradas son todas cero. Por ende, existe  $\theta \in [0, 1)$  que satisface (79) para cualesquiera  $x, y \in \overline{B_K(0_{\mathbb{R}^n})}$ . En consecuencia, la desigualdad (80) también se cumple. Esto implica que la sucesión  $(x_k)$  es una sucesión de Cauchy en  $\mathbb{R}^n$ . Por ende, al ser  $(\overline{B_K(0_{\mathbb{R}^n})}, \|\cdot\|_\infty)$  un espacio de Banach, pues  $\overline{B_K(0_{\mathbb{R}^n})}$  es un conjunto compacto,  $(x_k)$  converge a un punto  $x^* \in \mathbb{R}^n$ . Puesto que  $\varphi$  es continua, se tiene entonces que  $\varphi(x_k) \rightarrow \varphi(x^*)$  en  $\overline{B_K(0_{\mathbb{R}^n})}$ . Como el límite de una sucesión

es único,  $\varphi(x^*) = x^*$ , es decir,  $x^*$  es un punto fijo de  $\varphi$ . Consecuentemente,  $x^*$  es una solución de las ecuaciones en (64).  $\square$

DEMOSTRACIÓN DEL LEMA 4.17. Fijemos  $i, j \in [n]$  con  $i \neq j$ . Sea  $x$  un elemento arbitrario de  $\mathbb{R}^n$ . Recordemos la función  $r_{ij} : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $r_{ij}(x) = (e^{-x_j} + e^{x_i})^{-1}$ . Definamos

$$q_{ij}(x) := \frac{r_{ij}(x)}{\sum_{\substack{k=1 \\ k \neq i}}^n r_{ik}(x)}.$$

Antes de derivar  $\varphi_i(x)$  respecto de  $x_i$ , primero notemos

$$\sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial r_{ik}(x)}{\partial x_i} = \sum_{\substack{k=1 \\ k \neq i}}^n -\frac{e^{x_i}}{(e^{-x_k} + e^{x_i})^2} = -\sum_{\substack{k=1 \\ k \neq i}}^n \frac{e^{x_i}}{e^{-x_k} + e^{x_i}} \cdot r_{ik}(x).$$

De manera que, por regla de la cadena,

$$\begin{aligned} \frac{\partial \varphi_i(x)}{\partial x_i} &= -\frac{1}{\sum_{\substack{\ell=1 \\ \ell \neq i}}^n r_{i\ell}(x)} \cdot \left( -\sum_{\substack{k=1 \\ k \neq i}}^n \frac{e^{x_i}}{e^{-x_k} + e^{x_i}} \cdot r_{ik}(x) \right) \\ &= \sum_{\substack{k=1 \\ k \neq i}}^n \frac{e^{x_i}}{e^{-x_k} + e^{x_i}} \cdot \frac{r_{ik}(x)}{\sum_{\substack{\ell=1 \\ \ell \neq i}}^n r_{i\ell}(x)} \\ &= \sum_{\substack{k=1 \\ k \neq i}}^n \frac{e^{x_i}}{e^{-x_k} + e^{x_i}} \cdot q_{ik}(x) = e^{x_i} \sum_{\substack{k=1 \\ k \neq i}}^n r_{ik}(x) q_{ik}(x). \end{aligned}$$

Esto muestra que la función  $\frac{\partial \varphi_i}{\partial x_i}$  es mayor ó igual a cero en  $\mathbb{R}^n$ . Mientras que

$$\frac{\partial \varphi_i(x)}{\partial x_j} = -\frac{1}{\sum_{\substack{\ell=1 \\ \ell \neq i}}^n r_{i\ell}(x)} \cdot \left( \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial r_{ik}(x)}{\partial x_j} \right).$$

Puesto que  $r_{ik}(x) = (e^{-x_k} + e^{x_i})^{-1}$ , entonces  $\frac{\partial r_{ik}(x)}{\partial x_j} = 0$  para  $k \in [n] \setminus \{j\}$ . Por consiguiente,

$$\sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial r_{ik}(x)}{\partial x_j} = \frac{\partial r_{ij}(x)}{\partial x_j} = -\frac{-e^{-x_j}}{(e^{-x_j} + e^{x_i})^2} = \frac{e^{-x_j}}{e^{-x_j} + e^{x_i}} \cdot r_{ij}(x).$$

Por lo tanto,

$$(82) \quad \frac{\partial \varphi_i(x)}{\partial x_j} = -\frac{e^{-x_j}}{e^{-x_j} + e^{x_i}} \cdot \frac{r_{ij}(x)}{\sum_{\substack{\ell=1 \\ \ell \neq i}}^n r_{i\ell}(x)} = -e^{-x_j} r_{ij}(x) \cdot q_{ij}(x).$$

Se sigue que la función  $\frac{\partial \varphi_i}{\partial x_i}$  es menor ó igual a cero en  $\mathbb{R}^n$ .

Sea  $K \in \mathbb{R}_{\geq 0}$ . Supongamos  $x \in \mathbb{R}^n$  con  $\|x\|_\infty \leq K$ . Entonces  $-x_j, x_i \leq \|x\|_\infty \leq K$ . De modo que  $e^{-x_j}, e^{x_i} \leq e^K$ . En consecuencia,  $e^{-x_j} + e^{x_i} \leq 2e^K$ . Análogamente,  $2e^{-K} \leq e^{-x_j} + e^{x_i}$ . Por ende,

$$(83) \quad \frac{1}{2}e^{-K} \leq r_{ij}(x) \leq \frac{1}{2}e^K.$$

Luego,

$$(n-1)\frac{1}{2}e^{-K} \leq \sum_{\substack{k=1 \\ k \neq i}}^n r_{ik}(x) \leq (n-1)\frac{1}{2}e^K.$$

De manera que al invertir las desigualdades anteriores obtenemos

$$(84) \quad \frac{2}{n-1}e^{-K} \leq \frac{1}{\sum_{\substack{k=1 \\ k \neq i}}^n r_{ik}(x)} \leq \frac{2}{n-1}e^K.$$

En consecuencia, cuando multiplicamos las desigualdades de (83) y (84) (las cuales son no negativas, así que se preserva el orden de la desigualdad), nos queda la desigualdad

$$\frac{1}{n-1}e^{-2K} \leq \frac{r_{ij}(x)}{\sum_{\substack{k=1 \\ k \neq i}}^n r_{ik}(x)} \leq \frac{1}{n-1}e^{2K};$$

es decir,

$$(85) \quad \frac{e^{-2K}}{n-1} \leq q_{ij}(x) \leq \frac{e^{2K}}{n-1}.$$

Entonces, al multiplicar por  $-r_{ij}(x)$ ,

$$(86) \quad -r_{ij}(x) \frac{e^{2K}}{n-1} \leq -r_{ij}(x)q_{ij}(x) \leq -r_{ij}(x) \frac{e^{-2K}}{n-1}.$$

Por consiguiente, de la segunda desigualdad en (83) y la primera desigualdad en (86),

$$-\frac{e^{3K}}{2(n-1)} \leq -r_{ij}(x) \frac{e^{2K}}{n-1} \leq -r_{ij}(x)q_{ij}(x).$$

En tanto que, de la primera desigualdad en (83) y la segunda desigualdad en (86),

$$(87) \quad -r_{ij}(x)q_{ij}(x) \leq -r_{ij}(x) \frac{e^{-2K}}{n-1} \leq -\frac{e^{-3K}}{2(n-1)}.$$

Recordemos  $e^{-K} \leq e^{-x_j}$ , así que, de (87),

$$-e^{-x_j} r_{ij}(x)q_{ij}(x) \leq -e^{-K} r_{ij}(x)q_{ij}(x) \leq -e^{-K} \frac{e^{-3K}}{2(n-1)} = -\frac{e^{-4K}}{2(n-1)}.$$

Entonces, de (82),

$$\frac{\partial \varphi_i(x)}{\partial x_j} \leq -\frac{e^{-4K}}{2(n-1)}.$$

A su vez,  $e^{-x_j} r_{ij}(x)q_{ij}(x) \leq q_{ij}(x)$ , puesto que  $e^{-x_j} \leq e^{-x_j} + e^{x_i}$ . De manera que de la segunda desigualdad en (85)

$$-\frac{e^{2K}}{n-1} \leq -q_{ij}(x) \leq -e^{-x_j} r_{ij}(x)q_{ij}(x) = \frac{\partial \varphi_i(x)}{\partial x_j}.$$

Por lo tanto,

$$-\frac{e^{2K}}{n-1} \leq \frac{\partial \varphi_i(x)}{\partial x_j} \leq -\frac{e^{-4K}}{2(n-1)}.$$

Los mismos argumentos muestran que

$$-\frac{e^{2K}}{n-1} \leq -e^{x_i} r_{ij}(x)q_{ij}(x) \leq -\frac{e^{-4K}}{2(n-1)}.$$

En consecuencia,

$$\frac{e^{-4K}}{2(n-1)} \leq e^{x_i} r_{ij}(x)q_{ij}(x) \leq \frac{e^{2K}}{n-1}.$$

Entonces,

$$\frac{e^{-4K}}{2} \leq \frac{\partial \varphi_i(x)}{\partial x_i} \leq e^{2K}.$$

□

## Nociones de la teoría de la medida

DEFINICIÓN A.1. Supongamos que  $\Omega$  es un conjunto no vacío. Una familia  $\mathcal{F}$  de subconjuntos de  $\Omega$  es un **álgebra** si contiene a  $\Omega$  y es cerrada bajo complementos y uniones finitas, esto es,

- (i)  $\Omega \in \mathcal{F}$ .
- (ii) Si  $A \in \mathcal{F}$ , entonces  $\Omega \setminus A \in \mathcal{F}$ .
- (iii) Si  $A, B \in \mathcal{F}$ , entonces  $A \cup B \in \mathcal{F}$ .

DEFINICIÓN A.2. Supongamos que  $\Omega$  es un conjunto no vacío. Una familia  $\mathcal{F}$  de subconjuntos de  $\Omega$  es una  **$\sigma$ -álgebra** si  $\mathcal{F}$  es un álgebra y es cerrada bajo uniones numerables; es decir, para  $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$  tenemos  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$ .

Note que el hecho de que  $\mathcal{F}$  cumpla con las condiciones (i) y (ii) en la Definición A.1 y también sea cerrada bajo uniones numerables, implica la condición (iii) en la Definición A.1; lo cual es bien sabido y, por tanto, no será mostrado.

Si  $\Omega$  es un conjunto no vacío y  $\mathcal{F}$  es una  $\sigma$ -álgebra de subconjuntos de  $\Omega$ , el par  $(\Omega, \mathcal{F})$  usualmente es llamado un *espacio medible*. Los conjuntos que pertenecen a  $\mathcal{F}$  son llamados *conjuntos medibles* o, en un contexto probabilístico, son llamados *eventos*. Particularmente, en probabilidad, si  $\{\omega\}$  está en  $\mathcal{F}$ , decimos que  $\{\omega\}$  es un *evento elemental*.

PROPOSICIÓN A.3. Supongamos que  $\Omega$  es un conjunto no vacío. Dada una familia  $\mathcal{C}$  de subconjuntos de  $\Omega$ , existe una mínima  $\sigma$ -álgebra en  $\Omega$  que contiene a  $\mathcal{C}$  y la denotaremos por  $\sigma(\mathcal{C})$ . Es decir, sin  $\mathcal{F}$  es una  $\sigma$ -álgebra que contiene a  $\mathcal{C}$ , entonces  $\sigma(\mathcal{C}) \subseteq \mathcal{F}$ .

En adelante diremos que  $\sigma(\mathcal{C})$  es la  $\sigma$ -álgebra generada por  $\mathcal{C}$ .

Supongamos que  $(\Omega, \tau)$  es un espacio topológico. Definimos a la  *$\sigma$ -álgebra de Borel* en  $(\Omega, \tau)$  como la  $\sigma$ -álgebra generada por  $\tau$  y la denotamos por  $\mathcal{B}(\Omega, \tau)$  o simplemente  $\mathcal{B}(\Omega)$  si no se genera alguna confusión. Si  $(\Omega, d)$  es un espacio métrico,  $\mathcal{B}(\Omega)$  denota a la  $\sigma$ -álgebra generada por los abiertos que pertenecen a la topología inducida por  $d$ . En particular, para  $n \in \mathbb{N}$ , denotamos por  $\mathcal{B}(\mathbb{R}^n)$  a la  $\sigma$ -álgebra generada por la topología inducida por la métrica euclidiana.

Supongamos que  $\{\Omega_\alpha \mid \alpha \in A\}$  es una colección de conjuntos no vacíos. Definimos a  $\prod_{\alpha \in A} \Omega_\alpha$  como el conjunto de funciones  $f : A \rightarrow \bigcup_{\alpha \in A} \Omega_\alpha$  tales que  $f(\alpha) \in \Omega_\alpha$  para todo  $\alpha \in A$ . Para cada  $\alpha \in A$  definimos al  $\alpha$ -ésimo mapeo proyección,  $\pi_\alpha : \prod_{\alpha \in A} \Omega_\alpha \rightarrow \Omega_\alpha$ , por medio de la identidad  $\pi_\alpha(\omega) = \omega(\alpha)$ .

Si  $\mathcal{F}_\alpha$  es una  $\sigma$ -álgebra en  $\Omega_\alpha$  para todo  $\alpha \in A$ , definimos a la  **$\sigma$ -álgebra-producto** en  $\prod_{\alpha \in A} \Omega_\alpha$  como

$$\bigotimes_{\alpha \in A} \Sigma_\alpha := \sigma(\{\pi_\alpha^{-1}(E_\alpha) \mid E_\alpha \in \mathcal{F}_\alpha, \alpha \in A\}).$$

PROPOSICIÓN A.4. Sea  $\{(\Omega_\alpha, \mathcal{F}_\alpha) \mid \alpha \in A\}$  una colección de espacios medibles. Supongamos que  $\mathcal{C}_\alpha \subseteq \mathcal{P}(\Omega_\alpha)$  es tal que  $\mathcal{F}_\alpha = \sigma(\mathcal{C}_\alpha)$  para todo  $\alpha \in A$ . Entonces,

$$\bigotimes_{\alpha \in A} \Sigma_\alpha := \sigma(\{\pi_\alpha^{-1}(E_\alpha) \mid E_\alpha \in \mathcal{C}_\alpha, \alpha \in A\}).$$

Si además  $A$  es a lo más numerable y  $X_\alpha \in \mathcal{C}_\alpha$  para todo  $\alpha \in A$ , entonces

$$\bigotimes_{\alpha \in A} \Sigma_\alpha := \sigma \left( \left\{ \prod_{\alpha \in A} E_\alpha \mid E_\alpha \in \mathcal{C}_\alpha \right\} \right).$$

Sean  $(X_1, d_1), \dots, (X_n, d_n)$  espacios métricos. Tenemos una métrica en  $\prod_{i=1}^n X_i$  dada por

$$d(x, y) := \left( \sum_{i=1}^n d_i(x_i, y_i)^2 \right)^{1/2},$$

para  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \prod_{i=1}^n X_i$ . No es complicado ver que  $d$  es una métrica, y es llamada la *métrica producto*.

PROPOSICIÓN A.5. Sean  $(X_1, d_1), \dots, (X_n, d_n)$  espacios métricos separables y consideremos a  $\prod_{i=1}^n X_i$  equipado con la métrica producto. Entonces  $\mathcal{B}(X) = \bigotimes_{i=1}^n \mathcal{B}(X_i)$ .

DEFINICIÓN A.6. Decimos que  $\Pi$  es un  $\pi$ -sistema si para todo  $A, B \in \Pi$  tenemos  $A \cap B \in \Pi$ .

NOTACIÓN A.7. Para  $a = (a_1, \dots, a_n)$  y  $b = (b_1, \dots, b_n)$  en  $\mathbb{R}^n$ , decimos que  $a < b$  ( $\leq$ ) si y sólo si  $a_i < b_i$  ( $a_i \leq b_i$ ) para  $i \in [n]$ . Si  $a < b$ , denotaremos al rectángulo abierto generado por  $a$  y  $b$  como

$$(a, b) := \{x \in \mathbb{R}^n \mid a < x < b\} = (a_1, b_1) \times \dots \times (a_n, b_n).$$

De manera análoga definimos al rectángulo  $[a, b]$ .

PROPOSICIÓN A.8. El conjunto  $\Pi_1 := \{[a, b] \mid a, b \in \mathbb{R}^2, a \leq b\} \cup \{\emptyset\}$  es un  $\pi$ -sistema.

PROPOSICIÓN A.9. La igualdad  $\mathcal{B}(\mathbb{R}^2) = \sigma(\Pi_1)$  se satisface.

OBSERVACIÓN A.10. Para cualesquiera  $a, b \in \mathbb{R}$  con  $a \leq b$  tenemos  $[a, b] \cap [0, 1] = [c, d]$  para  $c, d \in \mathbb{R}$  con  $0 \leq c \leq d \leq 1$ , ó  $[a, b] \cap [0, 1] = \emptyset$ .

DEMOSTRACIÓN. Fijemos dos elementos  $a, b \in \mathbb{R}$  arbitrarios con la propiedad  $a \leq b$ . Entonces,

$$[a, b] \cap [0, 1] = \begin{cases} \emptyset & \text{si } b < 0 \text{ ó } 1 < a, \\ [0, b] & \text{si } a < 0 \leq b < 1, \\ [a, b] & \text{si } 0 \leq a \leq b < 1, \\ [a, 1] & \text{si } 0 \leq a \leq 1 \leq b. \end{cases}$$

□

Sea  $\Omega$  un conjunto no vacío y  $\mathcal{C}$  una familia no vacía de subconjuntos de  $\Omega$ , para cualquier subconjunto no vacío  $E \subseteq \Omega$ , definimos

$$\mathcal{C} \cap E := \{A \cap E \mid A \in \mathcal{C}\}.$$

OBSERVACIÓN A.11. Considere  $\Pi_1 = \{[a, b] \mid a, b \in \mathbb{R}^2, a \leq b\} \cup \{\emptyset\}$  una familia de subconjuntos de  $\mathbb{R}^2$  y definamos  $\Pi_2 := \{[a, b] \mid a, b \in \mathbb{R}^2, 0 \leq a \leq b \leq 1\} \cup \{\emptyset\}$ . La igualdad  $\Pi_1 \cap [0, 1]^2 = \Pi_2$  se satisface.

DEMOSTRACIÓN. Primero mostraremos  $\Pi_1 \cap [0, 1]^2 \subseteq \Pi_2$ . Es evidente el caso  $\emptyset \cap [0, 1]^2 \in \Pi_1 \cap [0, 1]^2$  ya que  $\emptyset \cap [0, 1]^2 = \emptyset$  y  $\emptyset \in \Pi_2$ . Tomemos, pues,  $[a, b] \in \Pi_1 \setminus \{\emptyset\}$ . Note que  $[a, b]$  es de la forma  $[a, b] = [a_1, b_1] \times [a_2, b_2]$  para  $a_1, a_2, b_1, b_2 \in \mathbb{R}$ . Consecuentemente,

$$\begin{aligned} [a, b] \cap [0, 1]^2 &= ([a_1, b_1] \times [a_2, b_2]) \cap ([0, 1] \times [0, 1]) \\ &= ([a_1, b_1] \cap [0, 1]) \times ([a_2, b_2] \cap [0, 1]). \end{aligned}$$

Así, por la Observación A.10,  $[a, b] \cap [0, 1]^2 = \emptyset$ , ó  $[a, b] \cap [0, 1]^2 = [c_1, d_1] \times [c_2, d_2]$  para  $c_1, c_2, d_1, d_2 \in \mathbb{R}$  con  $0 \leq c_1 \leq d_1 \leq 1$  y  $0 \leq c_2 \leq d_2 \leq 1$ . En cualquier caso,  $[a, b] \cap [0, 1]^2 \in \Pi_2$ . Por lo tanto,  $\Pi_1 \cap [0, 1]^2 \subseteq \Pi_2$ . Recíprocamente, ya sabemos que  $\emptyset \in \Pi_1 \cap [0, 1]^2$ , por lo que será suficiente mostrar que  $[a, b] \in \Pi_2$  implica  $[a, b] \in \Pi_1 \cap [0, 1]^2$  para  $a, b \in \mathbb{R}^2$  tales que  $0 \leq a \leq b \leq 1$ . Basta ver que, por las condiciones de  $a$  y  $b$ ,  $[a, b] = [a, b] \cap [0, 1]^2$ . Finalmente,  $[a, b] \in \Pi_1 \setminus \{\emptyset\}$  puesto que  $a, b \in \mathbb{R}^2$  y, en particular,  $a \leq b$ . Por ende,  $[a, b] = [a, b] \cap [0, 1]^2 \in \Pi_1 \cap [0, 1]^2$ . Por lo tanto,  $\Pi_2 \subseteq \Pi_1 \cap [0, 1]^2$ . Esto concluye con la demostración. □

OBSERVACIÓN A.12.  $\Pi_2 = \{[a, b] \mid a, b \in \mathbb{R}^2, 0 \leq a \leq b \leq 1\} \cup \{\emptyset\}$  es un  $\pi$ -sistema.

DEMOSTRACIÓN. Ya sabemos  $\Pi_2 = \Pi_1 \cap [0, 1]^2$  de la Observación A.11. De modo que hay que ver que  $\Pi_1 \cap [0, 1]^2$  es un  $\pi$ -sistema para mostrar la afirmación. Note que para cualesquiera  $A \cap [0, 1]^2, B \cap [0, 1]^2 \in \Pi_1 \cap [0, 1]^2$ , con  $A, B \in \Pi_1$  tenemos

$$(A \cap [0, 1]^2) \cap (B \cap [0, 1]^2) = (A \cap B) \cap [0, 1]^2.$$

Dado que  $\Pi_1$  es un  $\pi$ -sistema,  $A \cap B \in \Pi_1$ . Por ende,  $(A \cap B) \cap [0, 1]^2 \in \Pi_1 \cap [0, 1]^2$ . En consecuencia,  $\Pi_1 \cap [0, 1]^2$  es un  $\pi$ -sistema.  $\square$

PROPOSICIÓN A.13. Sea  $(\Omega, \mathcal{F})$  un espacio medible y  $E$  un subconjunto no vacío de  $\Omega$ . Si  $\mathcal{F} = \sigma(\mathcal{C})$  para  $\mathcal{C}$  una familia de subconjuntos de  $\Omega$ , entonces  $\mathcal{F} \cap E = \sigma(\mathcal{C} \cap E)$ .

COROLARIO A.14. Sea  $\Omega$  un espacio métrico y consideremos a cualquier subconjunto no vacío,  $E \subseteq \Omega$ , como subespacio métrico. Entonces  $\mathcal{B}(\Omega) \cap E = \mathcal{B}(E)$ .

COROLARIO A.15. El conjunto  $\Pi_2 = \{[a, b] \mid a, b \in \mathbb{R}^2, 0 \leq a \leq b \leq 1\} \cup \{\emptyset\}$  es un  $\pi$ -sistema que cumple  $\mathcal{B}([0, 1]^2) = \sigma(\Pi_2)$ .

DEMOSTRACIÓN. Ya mostramos en la Observación A.12 que  $\Pi_2$  es un  $\pi$ -sistema. Por el Corolario A.14,  $\mathcal{B}([0, 1]^2) = \mathcal{B}(\mathbb{R}) \cap [0, 1]^2$ . Y en uso de la Proposición A.9,  $\mathcal{B}(\mathbb{R}) = \sigma(\Pi_1)$ . En consecuencia,  $\mathcal{B}([0, 1]^2) = \sigma(\Pi_1) \cap [0, 1]^2 = \sigma(\Pi_1 \cap [0, 1]^2)$ , donde la última igualdad se justifica por la Proposición A.13. La Observación A.11 nos permite concluir  $\mathcal{B}([0, 1]^2) = \sigma(\Pi_2)$ .  $\square$

PROPOSICIÓN A.16. Sea  $(\Omega, \mathcal{F})$  un espacio medible y  $\Pi \subseteq \mathcal{P}(\Omega)$  un  $\pi$ -sistema tal que  $\mathcal{F} = \sigma(\Pi)$ . Supongamos que  $\mu$  y  $\nu$  son medidas en  $\mathcal{F}$  tales que  $\mu(A) = \nu(A)$  para todo  $A \in \Pi$ . Si  $\mu(\Omega) = \nu(\Omega) < \infty$ , entonces  $\mu = \nu$ .

DEFINICIÓN A.17. Supongamos que  $(\Omega, \mathcal{F})$  es un espacio medible. Una función  $\mu : \mathcal{F} \rightarrow [0, \infty]$  es una **medida** si satisface simultáneamente las siguientes condiciones:

- (i)  $\mu(\emptyset) = 0$ .
- (ii) Si tenemos una colección  $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$  con la propiedad de que para cualquier par de números naturales distintos  $i, j$  se cumple que  $A_i \cap A_j = \emptyset$ , entonces

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Un espacio de medida es una terna  $(\Omega, \mathcal{F}, \mu)$  consistente de un conjunto no vacío  $\Omega$ ,  $\mathcal{F}$  una  $\sigma$ -álgebra en  $\Omega$  y una medida  $\mu : \mathcal{F} \rightarrow [0, \infty]$ .

DEFINICIÓN A.18. Supongamos que  $(\Omega, \mathcal{F}, \mu)$  es un espacio de medida. Decimos que  $\mu$  es una **medida de probabilidad** si  $\mu(\Omega) = 1$ . A menos que se indique lo contrario, denotaremos por  $\mathbb{P}$  a una medida de probabilidad en un espacio medible  $(\Omega, \mathcal{F})$ .

Un espacio de medida  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad si  $\mathbb{P}$  es una medida de probabilidad.

PROPOSICIÓN A.19. Supongamos que  $(\Omega, \mathcal{F}, \mu)$  es un espacio de medida. Tenemos las siguientes propiedades:

- (i) (**Monotonía**). Si  $A, B \in \mathcal{F}$  y  $A \subseteq B$ , entonces  $\mu(A) \leq \mu(B)$ .
- (ii) (**Subaditividad**). Si  $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$ , entonces  $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$ .

DEFINICIÓN A.20. Dado un espacio de medida  $(\Omega, \mathcal{F}, \mu)$ , decimos que un conjunto  $A \in \mathcal{F}$  es **nulo** (o  $\mu$ -nulo) si  $\mu(A) = 0$ .

La propiedad de subaditividad de la medida garantiza que toda unión numerable de conjuntos nulos es un conjunto nulo. Además, si  $A$  es un conjunto nulo, por monotonía se sigue que para todo  $B \in \mathcal{F}$  tal que  $B \subseteq A$ ,  $\mu(B) = 0$ . Cuando trabajemos con un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , nos referiremos a un conjunto nulo como un *conjunto de probabilidad cero*.

En un espacio de medida  $(\Omega, \mathcal{F}, \mu)$ , decimos que una proposición  $P(\omega)$  referente a  $\omega \in \Omega$  es cierta  $\mu$ -casi donde sea si existe  $A \in \mathcal{F}$  con  $\mu(A) = 0$  tal que  $P(\omega)$  es cierta si  $\omega \in \Omega \setminus A$ .

## 1. Espacios de probabilidad discretos

Sea  $\mathcal{F}$  la  $\sigma$ -álgebra de todos los subconjuntos de un espacio numerable (finito o infinito)  $\Omega$ , y sea  $p : \Omega \rightarrow [0, 1]$  una función no negativa. Supongamos que  $\sum_{\omega \in \Omega} p(\omega) = 1$ . Si  $A$  es un subconjunto de  $\Omega$ , definimos

$$(88) \quad \mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

Entonces  $\mathbb{P}$  está bien definida y es una medida de probabilidad. Siendo rigurosos, este hecho (a mi parecer) no es trivial, sin embargo, tampoco es complejo. La demostración requiere de nociones básicas de Cálculo y se omite para fines de esta tesis. El lector interesado en la prueba puede consultarla en [Bil95].

Nos referimos a  $(\Omega, \mathcal{F}, \mathbb{P})$  como un *espacio de probabilidad discreto* si  $\Omega$  es numerable (finito o infinito),  $\mathcal{F}$  consiste de todos los subconjuntos de  $\Omega$ , y  $\mathbb{P}$  se define como en (88).

En un espacio de probabilidad discreto, el caso particular donde  $\Omega$  es un conjunto finito con al menos un elemento y  $p$  está definida como  $p(\omega) = 1/|\Omega|$ , entonces para cualquier evento  $A$  tenemos

$$(89) \quad \mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

Dicho de otra manera, si los eventos elementales son equiprobables, la probabilidad de  $A$  es entonces el número de eventos elementales que conforman (cuya unión es)  $A$  dividido por el número total de eventos elementales. Este ejemplo nace de la teoría de probabilidad clásica. En este contexto, la medida  $\mathbb{P}$  dada por (89) recibe el nombre de *medida de probabilidad uniforme*.

## 2. Funciones medibles

**DEFINICIÓN A.21.** Sean  $(\Omega_1, \mathcal{F}_1)$  y  $(\Omega_2, \mathcal{F}_2)$  dos espacios medibles. Decimos que una función  $X : \Omega_1 \rightarrow \Omega_2$  es  $(\mathcal{F}_1, \mathcal{F}_2)$ -**medible**, o **medible** si no se presta a confusión, si y sólo si  $X^{-1}(A) \in \mathcal{F}_1$  siempre que  $A \in \mathcal{F}_2$ .

**PROPOSICIÓN A.22.** Sean  $(\Omega_1, \mathcal{F}_1)$  y  $(\Omega_2, \mathcal{F}_2)$  dos espacios medibles,  $\mathcal{C} \subseteq \mathcal{P}(\Omega_2)$  tal que  $\sigma(\mathcal{C}) = \mathcal{F}_2$ . Entonces,  $X : \Omega_1 \rightarrow \Omega_2$  es medible si y sólo si  $X^{-1}(A) \in \mathcal{F}_1$  para todo  $A \in \mathcal{C}$ .

**PROPOSICIÓN A.23.** Supongamos que  $m$  y  $n$  son enteros positivos. Si  $X : \mathbb{R}^n \rightarrow \mathbb{R}^m$  es una función continua, entonces es medible.

**PROPOSICIÓN A.24.** Sean  $(\Omega_1, \mathcal{F}_1)$ ,  $(\Omega_2, \mathcal{F}_2)$  y  $(\Omega_3, \mathcal{F}_3)$  espacios medibles. Si  $X : \Omega_1 \rightarrow \Omega_2$  y  $Y : \Omega_2 \rightarrow \Omega_3$  son funciones medibles, entonces  $Y \circ X : \Omega_1 \rightarrow \Omega_3$  es una función medible.

**PROPOSICIÓN A.25.** Sea  $(\Omega, \mathcal{F})$  un espacio medible. Fijemos  $n \in \mathbb{N}$ . Si  $X, Y : \Omega \rightarrow \mathbb{R}^n$  son medibles, entonces para cualesquiera  $a, b \in \mathbb{R}$ ,  $aX + bY$  es medible.

**PROPOSICIÓN A.26.** Sea  $(\Omega, \mathcal{F})$  un espacio medible. Si  $X : \Omega \rightarrow \mathbb{R}$  es constante, entonces  $X$  es medible.

**PROPOSICIÓN A.27.** Supongamos que  $(\Omega, \mathcal{F})$  es un espacio medible. Si  $X, Y : \Omega \rightarrow \mathbb{R}$  son medibles, entonces  $XY$  es medible.

DEFINICIÓN A.28. Sea  $(\Omega, \mathcal{F})$  un espacio medible. Para  $E \subseteq \Omega$ ,  $\mathbb{1}_E : \Omega \rightarrow \mathbb{R}$  dada por

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{si } x \in \Omega \setminus E \end{cases}$$

es llamada la **función indicadora de  $E$** .

*Nota:* La función indicadora de  $E$  también recibe el nombre de **función característica de  $E$** , en cuyo caso se denota como  $\chi_E$ .

PROPOSICIÓN A.29. Sea  $(\Omega, \mathcal{F})$  un espacio medible. Si  $A, B \subseteq \Omega$ , entonces

$$\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A \cap B}.$$

PROPOSICIÓN A.30. Sea  $(\Omega, \mathcal{F})$  un espacio medible y  $E$  un subconjunto de  $\Omega$ . La función indicadora de  $E$ ,  $\mathbb{1}_E$ , es medible si y sólo si  $E \in \mathcal{F}$ .

No podíamos concluir esta sección sin enunciar la definición de una variable aleatoria.

DEFINICIÓN A.31. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Una **variable aleatoria** es una función  $X : \Omega \rightarrow \mathbb{R}$  que es  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -medible.

### 3. Propiedades de la integral

Consideremos un espacio de medida  $(\Omega, \mathcal{F}, \mu)$ . Definimos

$$L^+(\Omega, \mathcal{F}) = \{X : \Omega \rightarrow \mathbb{R} \mid X \text{ es } (\mathcal{F}, \mathcal{B}(\mathbb{R}))\text{-medible y } X \geq 0\}.$$

TEOREMA A.32. *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida. Supongamos que tenemos una colección  $\{A_n \mid n \in \mathbb{N}\} \subseteq \mathcal{F}$  de conjuntos ajenos por pares. Si  $X \in L^1(\Omega, \mathcal{F}, \mu) \cup L^+(\Omega, \mathcal{F})$ , entonces*

$$\int_{\cup_n A_n} X d\mu = \sum_n \int_{A_n} X d\mu.$$

PROPOSICIÓN A.33. *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida. Supongamos  $X, Y \in L^+(\Omega, \mathcal{F})$ .*

- (i) *Si  $X \leq Y$ , entonces  $\int X d\mu \leq \int Y d\mu$ .*
- (ii) *Si  $E, F \in \mathcal{F}$  y  $E \subseteq F$ , entonces  $\int_E X d\mu \leq \int_F X d\mu$ .*

PROPOSICIÓN A.34. *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida,  $X \in L^1(\Omega, \mathcal{F}, \mu)$  y  $E \in \mathcal{F}$ . Entonces  $X \mathbb{1}_E$  pertenece a  $L^1(\Omega, \mathcal{F}, \mu)$ .*

DEMOSTRACIÓN. Sabemos que  $\mathbb{1}_E$  es medible por la Proposición A.30. Luego,  $X \mathbb{1}_E$  es medible debido a la Proposición A.27. Además,  $|X|$  es una composición de funciones medibles, así,  $|X| \in L^+(\Omega, \mathcal{F})$ . De modo que, en uso de la Proposición A.33,

$$\int |X \mathbb{1}_E| d\mu = \int |X| \mathbb{1}_E d\mu \leq \int |X| \mathbb{1}_\Omega d\mu = \int |X| d\mu < \infty.$$

Por lo tanto,  $X \mathbb{1}_E \in L^1(\Omega, \mathcal{F}, \mu)$ . □

PROPOSICIÓN A.35. *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida. Si  $X \leq Y$  con  $X, Y \in L^1(\Omega, \mathcal{F}, \mu)$ , entonces  $\int X d\mu \leq \int Y d\mu$ .*

PROPOSICIÓN A.36 (Linealidad de la integral). *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida. Si  $X, Y \in L^1(\Omega, \mathcal{F}, \mu)$  y  $a \in \mathbb{R}$ , entonces*

1.  $\int (X + Y) d\mu = \int X d\mu + \int Y d\mu$
2.  $\int aX d\mu = a \int X d\mu$ .

PROPOSICIÓN A.37. *Sea  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida tal que  $\mu(\Omega) < \infty$  y  $\Pi$  un  $\pi$ -sistema con la propiedad  $\mathcal{F} = \sigma(\Pi)$  y  $\Omega \in \Pi$ . Supongamos que  $X : \Omega \rightarrow \mathbb{R}$  es  $(\mathcal{F}, \mathcal{B}(\mathbb{R}))$ -medible e integrable. Si para todo  $A \in \Pi$  se cumple  $\int_A X d\mu = 0$ , entonces  $X = 0$   $\mu$ -casi donde sea.*

PROPOSICIÓN A.38. *Sean  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida y  $X \in L^+(\Omega, \mathcal{F})$ . La función  $\nu : \mathcal{F} \rightarrow [0, \infty]$  dada por  $\nu(E) := \int_E X d\mu$  es una medida. Además, si  $Y : \Omega \rightarrow \mathbb{R}$  es medible,  $Y \in L^1(\Omega, \mathcal{F}, \nu)$  si y sólo si  $YX \in L^1(\Omega, \mathcal{F}, \mu)$  y en este caso*

$$\int Y d\nu = \int YX d\mu.$$

La igualdad anterior también es válida para  $Y \in L^+(\Omega, \mathcal{F})$ .

EJEMPLO A.39. *Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ , fijemos un conjunto  $A \in \mathcal{F}$  con probabilidad positiva. La función  $\mathbb{P}(\cdot | A) : \mathcal{F} \rightarrow [0, 1]$  dada por*

$$(90) \quad \mathbb{P}(B | A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

*es una medida de probabilidad sobre  $(\Omega, \mathcal{F})$ . Más aún, suponga  $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P}(\cdot | A))$ . Entonces, para cualquier  $B \in \mathcal{F}$ ,*

$$(91) \quad \int_B Y d\mathbb{P}(\cdot | A) = \frac{1}{\mathbb{P}(A)} \int_{A \cap B} Y d\mathbb{P}.$$

DEMOSTRACIÓN. Puesto que  $A$  pertenece a  $\mathcal{F}$ ,  $\mathbb{1}_A$  es medible. También, por definición,  $\mathbb{1}_A$  es mayor o igual a cero. En consecuencia,  $\mathbb{1}_A \in L^+(\Omega, \mathcal{F})$ . Por otro lado, por hipótesis,  $1/\mathbb{P}(A) > 0$ . De manera que, en virtud de la Proposición A.25,  $\frac{1}{\mathbb{P}(A)}\mathbb{1}_A$  pertenece a  $L^+(\Omega, \mathcal{F})$ . Por lo tanto, la función  $\nu : \mathcal{F} \rightarrow [0, \infty]$  dada por

$$\nu(B) = \int_B \frac{1}{\mathbb{P}(A)} \mathbb{1}_A d\mathbb{P}$$

es una medida. Desarrollando lo anterior usando la Proposición A.29,

$$\nu(B) = \int \frac{1}{\mathbb{P}(A)} \mathbb{1}_A \mathbb{1}_B d\mathbb{P} = \int \frac{1}{\mathbb{P}(A)} \mathbb{1}_{A \cap B} d\mathbb{P} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

De esto último es claro que  $\nu(\Omega) = 1$ . Por consiguiente,  $\nu$  es una medida de probabilidad. Como la notación del enunciado lo sugiere, la medida  $\nu$  es denotada por  $\mathbb{P}(\cdot | A)$ . Para mostrar la segunda afirmación, tomemos  $B \in \mathcal{F}$  arbitrario. Por la Proposición A.34,  $Y \mathbb{1}_B$  pertenece a  $L^1(\Omega, \mathcal{F}, \nu)$ . En consecuencia, podemos usar la Proposición A.38 para deducir que

$$\begin{aligned} \int Y \mathbb{1}_B d\nu &= \int Y \mathbb{1}_B \frac{1}{\mathbb{P}(A)} \mathbb{1}_A d\mathbb{P} \\ &= \frac{1}{\mathbb{P}(A)} \int Y \mathbb{1}_B \mathbb{1}_A d\mathbb{P} \\ &= \frac{1}{\mathbb{P}(A)} \int Y \mathbb{1}_{A \cap B} d\mathbb{P}. \end{aligned}$$

Esto culmina con la demostración.  $\square$

La medida  $\mathbb{P}(\cdot | A)$  del ejemplo anterior tiene relevancia en la teoría de la probabilidad y da pie a la siguiente definición.

DEFINICIÓN A.40. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $A \in \mathcal{F}$  un evento de probabilidad positiva. Decimos que la función  $\mathbb{P}(\cdot | A) : \mathcal{F} \rightarrow [0, 1]$ , con regla de correspondencia dada por (90), es la **medida de probabilidad condicional de  $\mathbb{P}$  relativa a  $A$** . Para  $B \in \mathcal{F}$  arbitrario, decimos que  $\mathbb{P}(B | A)$  es la **probabilidad condicional de  $B$  dado  $A$** , típicamente abreviado como la **probabilidad de  $B$  dado  $A$** .

Para abordar más a fondo en la intuición de esta definición véase [Bil95], o también

PROPOSICIÓN A.41. Sean  $(\Omega, \mathcal{F}, \mu)$  un espacio de medida y  $X \in L^+(\Omega, \mathcal{F})$ . Si  $\nu : \mathcal{F} \rightarrow [0, \infty]$  es una medida para la cual  $\nu(E) \leq \mu(E)$  para todo  $E \in \mathcal{F}$ , entonces  $\int X d\nu \leq \int X d\mu$ .

EJEMPLO A.42. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $A \in \mathcal{F}$  un evento de probabilidad positiva. Si  $X \in L^+(\Omega, \mathcal{F})$ , entonces  $\int X d\mathbb{P}(\cdot | A) \leq \int X d\mathbb{P}$ .

DEMOSTRACIÓN. Ya que  $\mathbb{P}(A) > 0$ , entonces  $0 < (\mathbb{P}(A))^{-1} \leq 1$ . Además, para cualquier  $B \in \mathcal{F}$ , por la propiedad de monotonía,  $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$ . Así,

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \leq \mathbb{P}(B).$$

El resultado se sigue de la Proposición A.41.  $\square$

TEOREMA A.43 (Desigualdad de Jensen). Supongamos que  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  es una función convexa, esto es,

$$\lambda\varphi(x) + (1 - \lambda)\varphi(y) \geq \varphi(\lambda x + (1 - \lambda)y)$$

para toda  $\lambda \in (0, 1)$  y  $x, y \in \mathbb{R}$ . Si  $\mu$  es una medida de probabilidad, y  $X$  y  $\varphi(X)$  son integrables, entonces

$$\varphi\left(\int X d\mu\right) \leq \int \varphi(X) d\mu.$$

PROPOSICIÓN A.44. Sean  $(\Omega_1, \mathcal{F}_1)$  y  $(\Omega_2, \mathcal{F}_2)$  espacios medibles y  $\phi : \Omega_1 \rightarrow \Omega_2$  una función  $(\mathcal{F}_1, \mathcal{F}_2)$ -medible. Si  $\mu$  es una medida en  $\mathcal{F}_1$ , entonces  $\phi$  induce una medida en  $\mathcal{F}_2$  a través de la identidad

$$\mu \circ \phi^{-1}(E) := \mu(\phi^{-1}(E)),$$

para todo  $E \in \mathcal{F}_2$ .

TEOREMA A.45 (Cambio de variable). Sea  $(\Omega_1, \mathcal{F}_1, \mu)$  un espacio de medida,  $(\Omega_2, \mathcal{F}_2)$  un espacio medible,  $\phi : \Omega_1 \rightarrow \Omega_2$  una función  $(\mathcal{F}_1, \mathcal{F}_2)$ -medible y  $X : \Omega_2 \rightarrow \mathbb{R}$  una función  $(\mathcal{F}_2, \mathcal{B}(\mathbb{R}))$ -medible. Sea  $\mu \circ \phi^{-1}$  la medida inducida por  $\phi$ . Entonces  $X \circ \phi \in L^1(\Omega_1, \mathcal{F}_1, \mu)$  si y sólo si  $X \in L^1(\Omega_2, \mathcal{F}_2, \mu \circ \phi^{-1})$  y en este caso

$$\int X \circ \phi d\mu = \int X d(\mu \circ \phi^{-1}).$$

La igualdad anterior también se cumple para  $X \in L^+(\Omega_1, \mathcal{F}_1)$ .

#### 4. Espacios $L^p$ y desigualdades

Los espacios  $L^p$  están definidos para un espacio de medida arbitrario. Debido al enfoque de esta tesis, y para no confundir al lector con la notación, trabajaremos con espacios de medida de la forma  $(\Omega, \mathcal{F}, \mathbb{P})$ . Nos referiremos a  $(\Omega, \mathcal{F}, \mathbb{P})$  como un espacio de probabilidad. Es importante hacer notar que algunas de las siguientes proposiciones pueden generalizarse a cualquier espacio de medida o, en particular, a cualquier espacio de medida finita.

Consideremos, pues, un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Definimos para  $0 < p < \infty$

$$L^p(\Omega, \mathcal{F}, \mathbb{P}) := \left\{ X : \Omega \rightarrow \mathbb{R} \mid X \text{ es medible y } \int |X|^p d\mathbb{P} < \infty \right\}.$$

Es un hecho que  $L^p(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio vectorial.

PROPOSICIÓN A.46. Supongamos que  $X$  es una variable aleatoria. Son equivalentes:

1.  $\int X d\mathbb{P} \in \mathbb{R}$ .
2.  $\int X^+ d\mathbb{P}, \int X^- d\mathbb{P} < \infty$
3.  $\int |X| d\mathbb{P} < \infty$ .

De aquí podemos deducir que una variable aleatoria  $X$  tiene esperanza finita si y sólo si  $X$  pertenece al espacio  $L^1$ .

PROPOSICIÓN A.47 (Desigualdad de Hölder). Sean  $p, q \in (1, \infty)$  tales que  $1/p + 1/q = 1$ . Si  $X \in L^p$  y  $Y \in L^q$ , entonces  $XY \in L^1$  y

$$\|XY\|_1 \leq \|X\|_p \|Y\|_q.$$

OBSERVACIÓN A.48. El caso especial  $p = q = 2$  es llamado **desigualdad de Cauchy-Schwarz** y podemos reformularlo en términos probabilistas.

PROPOSICIÓN A.49 (Desigualdad de Cauchy-Schwartz). Sean  $X, Y \in L^2$  dos variables aleatorias. Entonces la variable aleatoria  $XY$  es integrable y

$$\mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

PROPOSICIÓN A.50. Supongamos que tenemos un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Si  $0 < p \leq r \leq \infty$ , entonces  $L^r(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^p(\Omega, \mathcal{F}, \mathbb{P})$  y para toda  $X \in L^r(\Omega, \mathcal{F}, \mathbb{P})$ ,

$$\|X\|_p \leq \|X\|_r.$$

## Martingalas

### 1. Esperanza condicional

#### 1.1. Esperanza condicional dado un evento.

DEFINICIÓN B.1. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $A \in \mathcal{F}$  un evento de probabilidad positiva. Definimos la **esperanza condicional de  $X$  dado  $A$** , denotada por  $\mathbb{E}(X|A)$ , como sigue

$$\mathbb{E}(X|A) = \frac{1}{\mathbb{P}(A)} \int_A X d\mathbb{P}.$$

Si  $\mathbb{P}(A) = 0$ , por convención definimos  $\mathbb{E}(X|A) = 0$ .

Es importante resaltar que la esperanza condicional de una variable aleatoria, con esperanza finita, dado un evento resulta ser un número.

Otra observación sencilla, pero útil, se establece en la siguiente proposición.

PROPOSICIÓN B.2. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $A \in \mathcal{F}$  un evento de probabilidad positiva y  $c \in \mathbb{R}$ . Si  $X$  es igual a una constante  $c$  en  $A$ , entonces  $\mathbb{E}(X|A) = c$ .

DEMOSTRACIÓN. De la definición tenemos las siguientes igualdades:

$$\begin{aligned} \mathbb{E}(X|A) &= \frac{1}{\mathbb{P}(A)} \int_A X d\mathbb{P} \\ &= \frac{1}{\mathbb{P}(A)} \int_A c d\mathbb{P} \\ &= \frac{1}{\mathbb{P}(A)} \cdot c \cdot \mathbb{P}(A). \end{aligned}$$

Por lo tanto,  $\mathbb{E}(X|A) = c$ . □

PROPOSICIÓN B.3. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $A \in \mathcal{F}$  con probabilidad positiva. Entonces

$$|\mathbb{E}(X|A)| \leq \mathbb{E}(|X| | A).$$

DEMOSTRACIÓN. Por la Definición B.1,

$$|\mathbb{E}(X|A)| = \frac{1}{\mathbb{P}(A)} \left| \int_A X d\mathbb{P} \right| = \frac{1}{\mathbb{P}(A)} \left| \int X \mathbb{1}_A d\mathbb{P} \right|.$$

Aplicando la desigualdad de Jensen a la igualdad anterior,

$$|\mathbb{E}(X|A)| \leq \frac{1}{\mathbb{P}(A)} \int |X \mathbb{1}_A| d\mathbb{P} = \frac{1}{\mathbb{P}(A)} \int |X| \mathbb{1}_A d\mathbb{P} = \mathbb{E}(|X| | A).$$

Por ende,  $|\mathbb{E}(X|A)| \leq \mathbb{E}(|X| | A)$ . □

Buscaremos generalizar este concepto al introducir una familia más, como se verá en la siguiente sección. La esperanza condicional de una variable aleatoria dada una  $\sigma$ -álgebra es una función, como se precisa en la Definición B.4.

### 1.2. Esperanza condicional dada una $\sigma$ -álgebra.

DEFINICIÓN B.4. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X : \Omega \rightarrow \mathbb{R}$  una variable aleatoria en el espacio  $L^1(\Omega, \mathcal{F}, \mathbb{P})$  y  $\mathcal{G}$  una  $\sigma$ -álgebra tal que  $\mathcal{G} \subseteq \mathcal{F}$ . Definimos la esperanza condicional de  $X$  dado  $\mathcal{G}$  como una variable aleatoria, denotada por  $\mathbb{E}(X|\mathcal{G})$ , que satisface las siguientes tres propiedades de manera simultánea:

- (i)  $\mathbb{E}(X|\mathcal{G})$  es  $\mathcal{G}$  medible.
- (ii) Tiene esperanza finita.
- (iii) Para cualquier evento  $A \in \mathcal{G}$ ,

$$(92) \quad \int_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P} = \int_A X d\mathbb{P}.$$

La pregunta obligada que surge de la definición anterior es sobre la existencia de la esperanza condicional de una variable aleatoria dada una  $\sigma$ -álgebra. La respuesta a esta pregunta estriba en el teorema de Lebesgue-Radon-Nikodym, que suele introducirse en un curso de Teoría de la medida. Como referencia el lector puede consultar [Dur10].

Ahora bien, otra pregunta que debe surgir es acerca de la unicidad. Pues, es importante notar que en la definición anterior no se da una fórmula explícita para la esperanza condicional de una variable aleatoria, sino que, en principio, cualquier variable aleatoria que cumpla solamente con tres propiedades puede ser una esperanza condicional de una cierta variable  $X$ . Esto genera la inquietud de si la esperanza condicional de una variable aleatoria es única. Adelantamos al lector que la unicidad se obtiene en el sentido casi seguro. La demostración de esta afirmación también puede consultarse en [Dur10].

De este modo, sí podemos referirnos a la esperanza condicional de una variable aleatoria dada una  $\sigma$ -álgebra (en el sentido casi seguro).

EJEMPLO B.5. Supongamos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad. Tomemos  $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ . Entonces

$$\mathbb{E}(X|\{\emptyset, \Omega\}) \equiv \mathbb{E}(X).$$

Una generalización del ejemplo anterior se enuncia a continuación.

EJEMPLO B.6. Supongamos que  $(\Omega, \mathcal{F}, \mathbb{P})$  es un espacio de probabilidad. Sea  $\mathcal{P}$  una partición del espacio muestral  $\Omega$  consistente de eventos de probabilidad positiva y tal que  $\mathcal{P} \subseteq \mathcal{F}$ . Entonces

$$\mathbb{E}(X|\mathcal{P}) \equiv \mathbb{E}(X|A) \text{ en cada } A \in \mathcal{P}.$$

### 1.3. Propiedades.

PROPOSICIÓN B.7 (Linealidad de la esperanza condicional). Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $\mathcal{G}$  una  $\sigma$ -álgebra con la propiedad  $\mathcal{G} \subseteq \mathcal{F}$ . Supongamos que tenemos dos variables aleatorias  $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$  y fijemos un escalar  $a \in \mathbb{R}$  arbitrario. Entonces

$$\mathbb{E}(aX + Y|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + \mathbb{E}(Y|\mathcal{G}).$$

PROPOSICIÓN B.8. Sean  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X$  una variable aleatoria integrable,  $\mathcal{F}_1$  y  $\mathcal{F}_2$  dos sub- $\sigma$ -álgebras que cumplen que  $\mathcal{F}_1 \subseteq \mathcal{F}_2$ . Entonces, con probabilidad uno,

- (i)  $\mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2) = \mathbb{E}(X|\mathcal{F}_1)$ .
- (ii)  $\mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1)$ .

Una manera sencilla de recordar la proposición anterior es notar que “la  $\sigma$ -álgebra más pequeña siempre gana”.

TEOREMA B.9 (La esperanza condicional es una contracción en  $L^p$ ). Fijemos un número real  $p \geq 1$ . Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad,  $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$  y  $\mathcal{G}$  una  $\sigma$ -álgebra con la propiedad  $\mathcal{G} \subseteq \mathcal{F}$ . Entonces

$$\|\mathbb{E}(X|\mathcal{G})\|_p \leq \|X\|_p.$$

## 2. Martingalas

DEFINICIÓN B.10. Una **filtración** es una sucesión de  $\sigma$ -álgebras  $(\mathcal{F}_n)$  con la propiedad de que, para cualesquiera  $n, m \in \mathbb{N}$  con la propiedad  $n \leq m$ , se cumple  $\mathcal{F}_n \subseteq \mathcal{F}_m$ .

DEFINICIÓN B.11. Supongamos que tenemos una sucesión de variables aleatorias  $(X_n)$ . Decimos que  $(X_n)$  es **adaptada a una filtración**  $(\mathcal{F}_n)$  si, para cada  $n \in \mathbb{N}$ , la variable aleatoria  $X_n$  es  $\mathcal{F}_n$ -medible.

DEFINICIÓN B.12. Sea  $(\mathcal{F}_n)$  una filtración. Una sucesión de variables aleatorias  $(X_n)$  es una **martingala** respecto a  $(\mathcal{F}_n)$  si, para cada  $n \in \mathbb{N}$ , se satisfacen simultáneamente las siguientes condiciones:

- (i)  $X_n$  es integrable.
- (ii)  $(X_n)$  es adaptada a  $(\mathcal{F}_n)$ .
- (iii)  $\mathbb{E}(X_{n+1} | \mathcal{F}_n) = X_n$ .

EJEMPLO B.13 (Martingala de la esperanza condicional). Dada una filtración  $(\mathcal{F}_n)$  y una variable aleatoria  $X$  integrable, la sucesión de variables aleatorias  $(X_n)$  definida, para cada  $n \in \mathbb{N}$ , como:

$$X_n := \mathbb{E}(X | \mathcal{F}_n)$$

es una martingala.

DEMOSTRACIÓN. Se sigue de la Proposición B.8. □

TEOREMA B.14 (Teorema de convergencia de martingalas). Supongamos que  $(X_n)$  es una martingala respecto a alguna filtración. Si  $\sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|) < \infty$ , entonces existe una variable aleatoria  $X$  que satisface  $X_n \rightarrow X$  con probabilidad uno y  $\mathbb{E}(|X|) \leq \sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|)$ .

La demostración del teorema de convergencia de martingalas puede encontrarse en [Bil95, p. 490]

## 3. Sucesión de diferencias de martingala

DEFINICIÓN B.15. Sea  $(X_n)$  una martingala. Una **sucesión de diferencias de martingala** es una sucesión  $(d_n)$  de la forma

$$d_n := X_n - X_{n-1}$$

para  $n \in \mathbb{N}$  con  $n$  mayor o igual a uno y  $d_0 := X_0$ .

PROPOSICIÓN B.16. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad y  $1 < p \leq 2$ . Entonces para cada sucesión de diferencias de martingala  $(d_i)_{i=0}^n$  que pertenece a  $L^p(\Omega, \mathcal{F}, \mathbb{P})$  tenemos

$$(93) \quad \left( \sum_{i=0}^n \|d_i\|_p^2 \right)^{1/2} \leq \left( \frac{1}{p-1} \right)^{1/2} \cdot \left\| \sum_{i=0}^n d_i \right\|_p.$$



## Códigos

En este apéndice se encuentra el código, en lenguaje Python, utilizado para los dos ejemplos de la sección 3 del Capítulo 4. El siguiente código se basa en [Sou], donde se agregó un algoritmo para simular una sucesión de grados en base a funciones continuas y acotadas en  $[0, 1]$ .

```

import networkx as nx
import random as rnd
import matplotlib.pyplot as plt
import numpy as np
import math
from networkx import Graph, draw
from bisect import bisect_left
from itertools import repeat

#Modelo de configuración dada una sucesión de grados (MCSG).
def MCSG(k):
    """
    Esta función crea un modelo de configuración de una
    gráfica aleatoria a partir de una sucesión de grados 'k',
    donde k es un vector de enteros no negativos de tamaño
    finito.
    """
    #Primero nos aseguramos que la suma de los grados en la
    #sucesión de grados dada sea un número par.
    assert ((sum(k)%2)==0)
    #Creamos una gráfica vacía.
    G = Graph()
    #El número de vértices en nuestra gráfica es igual a la
    #suma de las entradas del vector k.
    G.add_nodes_from(range(len(k)))
    #Creamos un vector llamado 'resguardos' donde guardamos
    #ordenadamente k_i copias de 'i', donde k_i es la i-ésima
    #entrada de k.
    resguardos = [ i for i in range (len(k))
                  for _ in repeat(None,k[i])]
    #Reordenamos nuestros elementos en 'resguardos' al azar.
    rnd.shuffle(resguardos)
    #Creamos las aristas seleccionado dos elementos
    #aleatoriamente de nuestro vector 'resguardos'.
    for (i,j) in zip(resguardos[0::2], resguardos [1::2]):
        #Añadimos una arista entre los vértices
        #correspondientes.
        G.add_edge(i,j)
    return G

```

```
#Especificamos una función f, acotada y continua en [0,1].
def f(x):
    value = -1/2*x**2 + 3/4
    return value
#Creación de una sucesión de grados.
def sucesion_f(n):
    #Creamos una lista vacía.
    sucesion_grados = [0]*n
    #Declaramos el grado mas grande en la gráfica.
    sucesion_grados[n-1] = math.floor(n*f(0))
    #Desde 1 hasta (n-1).
    for i in range(1,n):
        #Declaramos los demás grados.
        sucesion_grados[n-1-i] = math.floor(n*f(i/n))
    if (sum(sucesion_grados)%2 != 0):
        #Si la suma de los grados no es par,
        #entonces aumentamos en una unidad el grado más grande.
        sucesion_grados[n-1] = sucesion_grados[n-1] + 1
    #Devuelve una sucesión de grados en orden creciente
     #(devuelve una lista).
    return sucesion_grados
```

## Símbolos

Símbolo	Descripción
$\mathbb{N}$	El conjunto $\{0, 1, 2, 3, \dots\}$ .
$\mathbb{R}^+$	Los números reales positivos.
$\mathbb{R}^{n \times n}$	Conjunto de matrices con entradas reales.
$v(G)$	El número de vértices de una gráfica $G$ .
$e(G)$	El número de aristas de una gráfica $G$ .
$A_G$	Matriz de adyacencia de una gráfica $G$ .
$[0, 1]^2$	El producto cartesiano $[0, 1] \times [0, 1]$ .
$[n]$	El conjunto $\{1, \dots, n\}$ para algún $n \in \mathbb{Z}_{>0}$ .
$[n]^2$	El producto cartesiano $[n] \times [n]$ .
$\Omega$	Un conjunto no vacío.
$\mathcal{P}$	Una partición de $\Omega$ .
$\mathcal{P}(\Omega)$	El conjunto potencia de $\Omega$ .
$\mathbb{E}(X)$	Esperanza de una variable aleatoria $X$ .
$\mathbb{E}(X \mathcal{P})$	Esperanza condicional de $X$ dado $\sigma(\mathcal{P})$ .
$(X_n)$	Una sucesión de variables aleatorias.
$F^{\circ \ell}$	$F$ compuesta consigo misma, $F \circ \dots \circ F$ , $\ell$ -veces.
$(\Sigma_i)$	Una sucesión de $k$ -semianillos.
$\mathbf{d}^n$	Una sucesión de grados de una gráfica de orden $n$ .
$(\mathbf{d}^n)$	Una sucesión de sucesiones de grados.



## Índice de figuras

1. Descomposición de dos elementos de un $k$ -semianillo para algún $k \geq 5$ .	1
2. Ilustración del caso 1 ( $U$ igual a $V$ ) en la propiedad (P3), dentro de la demostración del Ejemplo 1.9: Descomposición de los conjuntos $(S \times T) \setminus (U \times U)$ , $(T \times S) \setminus (U \times U)$ , $(S \times S) \setminus (U \times U)$ y $(T \times T) \setminus (U \times U)$ en tres conjuntos, cada uno, que pertenecen al 4-semianillo $\Sigma_{\square}$ , cuando $S$ es ajeno a $T$ .	7
3. Ilustración del caso 2, propiedad (P3), en la demostración del Ejemplo 1.9: Descomposición de $(S \times T) \setminus (U \times V)$ y $(S \times S) \setminus (U \times V)$ en cuatro conjuntos, cada uno, que pertenecen al 4-semianillo $\Sigma_{\square}$ , cuando $S$ es ajeno a $T$ y $U$ es ajeno a $V$ .	8
4. Simulación de una aproximación a un polinomio vía esperanza condicional.	16
1. Ejemplo de una gráfica $G$ y la 2-gráfica autoreplicada de $G$ .	35
2. Una gráfica y su matriz de adyacencia	36
3. La distancia $d_1$ entre dos gráficas.	40
4. La distancia $d_1$ no es realmente efectiva en el espacio de gráficas en el sentido de que dos gráficas esencialmente iguales tienen distancia positiva.	40
5. Dos gráficas del mismo orden esencialmente iguales.	42
6. $G^{\varphi}$ y $G^{\psi}$	42
1. Simulación de gráficas aleatorias con una sucesión de grados dada por una función cóncava.	65
2. Simulación de gráficas aleatorias con una sucesión de grados dada por una función convexa.	66



## Bibliografía

- [AFKS00] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.
- [Álv15] Miguel Ángel García Álvarez. *Introducción a la teoría de la probabilidad I. Primer curso*. Fondo de Cultura Económica, 2015.
- [BCL<sup>+</sup>08] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.
- [Bil95] Patrick Billingsley. *Probability and measure*. A Wiley-Interscience publication. Wiley, New York [u.a.], 3. ed edition, 1995.
- [BL96] Vitaly Bergelson and Alexander Leibman. Polynomial extensions of van der waerden’s and szemerédi’s theorems. *Journal of the American Mathematical Society*, 9(3):725–753, 1996.
- [BN08] Béla Bollobás and Vladimir Nikiforov. An abstract szemerédi regularity lemma. In *Building bridges*, pages 219–240. Springer, 2008.
- [Bol01] Béla Bollobás. *Explicit Constructions*, page 348–382. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.
- [CDS11] Sourav Chatterjee, Persi Diaconis, and Allan Sly. Random graphs with a given degree sequence. 2011.
- [CF12] David Conlon and Jacob Fox. Bounds for graph regularity and removal lemmas. *Geometric and Functional Analysis*, 22:1191–1256, 2012.
- [CGW89] Fan R. K. Chung, Ronald L. Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.
- [DKK16] Pandelis Dodos, Vassilis Kanellopoulos, and Thodoris Karageorgos. Szemerédi’s regularity lemma via martingales. *Electr. J. Comb.*, 23:P3.11, 2016.
- [Dur10] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, New York, NY, USA, 4th edition, 2010.
- [ER60] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5(1):17–60, 1960.
- [ET36] Paul Erdős and Paul Turán. On some sequences of integers. *Journal of the London Mathematical Society*, 1(4):261–264, 1936.
- [FK89] H Furstenberg and Yitzhak Katznelson. Idempotents in compact semigroups and ramsey theory. *Israel Journal of Mathematics*, 68(3):257–270, 1989.
- [FK99] Alan M. Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19:175–220, 1999.
- [Fur77] Harry Furstenberg. Ergodic behavior of diagonal measures and a theorem of szemerédi on arithmetic progressions. *Journal d’Analyse Mathématique*, 31(1):204–256, 1977.
- [Gow97] William T Gowers. Lower bounds of tower type for szemerédi’s uniformity lemma. *Geometric & Functional Analysis GAFA*, 7(2):322–337, 1997.
- [Gra91] Fan Chung Graham. Regularity lemmas for hypergraphs and quasi-randomness. *Random Struct. Algorithms*, 2:241–252, 1991.
- [KS96] János Komlós and Miklós Simonovits. Szemerédi’s regularity lemma and its applications in graph theory, 1996.
- [Lov12] László Lovász. Large networks and graph limits. In *Colloquium Publications*, 2012.
- [LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [LS07] László Lovász and Balázs Szegedy. Szemerédi’s lemma for the analyst. *GAFA Geometric And Functional Analysis*, 17(1):252–270, 2007.
- [Rot53] Klaus F Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953.
- [RX<sup>+</sup>16] Éric Ricard, Quanhua Xu, et al. A noncommutative martingale convexity inequality. *The Annals of Probability*, 44(2):867–882, 2016.
- [Sou] Raissa M. D Souza. Building a configuration model random graph. [http://mae.engr.ucdavis.edu/dsouza/Courses/289-S14/HW3b\\_soln2014.pdf](http://mae.engr.ucdavis.edu/dsouza/Courses/289-S14/HW3b_soln2014.pdf). Fecha de acceso: 7 de Junio de 2020.
- [SS91] Miklós Simonovits and Vera T. Sós. Szemerédi’s partition and quasirandomness. *Random Struct. Algorithms*, 2:1–10, 1991.
- [Sze75a] Endre Szemerédi. Regular partitions of graphs. Technical report, Stanford Univ Calif Dept of Computer Science, 1975.

- 
- [Sze75b] Endre Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith.*, 27(299-345):5, 1975.
- [Tao06] Terence Tao. Szemerédi's regularity lemma revisited. *Contrib. Discrete Math.*, 1(1):8–28, 2006.
- [TV06] Terence Tao and Van H Vu. *Additive combinatorics*, volume 105. Cambridge University Press, 2006.
- [vdH16] Remco van der Hofstad. Random graphs and complex networks. In *Cambridge Series in Statistical and Probabilistic Mathematics*, 2016.
- [Zha] Yufei Zhao. Graph theory and additive combinatorics. <https://yufeizhao.com/gtac/gtac.pdf>. Fecha de acceso: 20 de Mayo de 2020.