

UNIVERSIDAD DE SANTIAGO DE CHILE

FACULTAD DE CIENCIA
DEPARTAMENTO DE MATEMÁTICA Y CIENCIA
DE LA COMPUTACIÓN



1 Grafos cuasi-aleatorios y el lema de regularidad de 2 Szemerédi

3 Felipe Andrés Sánchez Erazo

4 Profesor(a) Guía:
Hiệp Hàn

5 Trabajo de titulación para optar al
título profesional de Ingeniero
Matemático

6 Santiago - Chile

7 2024

8 © Felipe Andrés Sánchez Erazo.

9 Se autoriza la reproducción parcial o total de esta obra con fines académicos, por cual-
10 quier forma, medio o procedimiento, siempre y cuando se incluya la cita bibliográfica
11 del documento.

UNIVERSIDAD DE SANTIAGO DE CHILE

FACULTAD DE CIENCIA
DEPARTAMENTO DE MATEMÁTICA Y CIENCIA
DE LA COMPUTACIÓN



12 Grafos cuasi-aleatorios y el lema de regularidad de 13 Szemerédi

14 Felipe Andrés Sánchez Erazo

15 Trabajo de Titulación presentado a la Facultad de Ciencia, en cumplimiento parcial
16 de los requisitos exigidos para optar al título de **Ingeniero Matemático**.

17 Este trabajo de Graduación fue presentado bajo la supervisión del profesor guía
18 Dr. Hiệp Hàn, del Departamento de Matemática y Ciencia de la Computación de la
19 Facultad de Ciencia.

20

Dr. Hiệp Hàn
Profesor Guía

21

Dr. Marcos Kiwi Krauskopf
Profesor Informante

Pablo Marín Álvarez
Director del Departamento

Dr. Sebastián Barbieri Lemp
Profesor Informante

A mi abuelo, Sergio Sánchez

24 Agradecimientos

25 Primero quiero agradecer a mi madre, por ser una figura y pilar fundamental en
26 mi vida, que me ha brindado una comprensión y cariño que no se puede medir. A
27 mi padre, por su apoyo completamente incondicional en todas las decisiones que he
28 tomado, y por despertar mi curiosidad por los saberes de este mundo. A mi hermano,
29 mi compañero de viaje, mi cómplice, gracias por traer alegrías a cada lugar que vas.
30 Te amo y admiro mucho, y aunque este logro es personal, espero que solo sea uno de
31 muchos en nuestro andar.

32 También agradecer a los otros miembros de mi familia, quienes en todo momento
33 me dieron palabras de aliento, y que de algún u otro modo, fueron parte de mi
34 motivación en la elaboración de esta tesis.

35 Quiero dar una mención especial y profunda gratitud a mi amada, Lourdes. Sin
36 su constante apoyo, los días habrían sido más grises en el desarrollo de este trabajo.
37 Gracias por tu admirable empatía al entender mis problemas y tristezas, incluso
38 cuando yo no les hallaba. Te agradezco por hacer que mis momentos de alegrías
39 tengan muchos más colores.

40 Mil gracias a aquellas personas que fueron parte de mi etapa universitaria y de
41 la vida, hicieron que la travesía estuviese llena de momentos bellos. En particular,
42 agradecer a Rodrigo, infinito amigo, quien ha estado en todos los momentos más
43 importantes de mi vida, y esta vez no fue la excepción. Cristian, por mostrarme lo
44 lindas que pueden ser las coincidencias en la vida y la amistad duradera que a veces
45 traen consigo. Catalina, por el Apoyo, Risas, Afecto y Motivación en la recta final de

46 este trabajo. Ailine, por nuestras conversaciones terapéuticas y todas las veces que
47 me dijiste “Amigo, tú puedes”. Camilo, por mostrarme que la amistad puede ir más
48 allá que el ser compañeros de trabajo, por tu dedicación, y por ser un referente para
49 mí como profesional.

50 Doy gracias a mis amistades de la música por mantener viva esta gran pasión
51 que compartimos. En especial, a mis cofrades de la Tuna Mayor de Ciencia de la
52 Universidad de Santiago de Chile. ¡Aúpa Tuna!

53 Quiero agradecer a mi profesor guía Hiệp Hàn por su buena disposición, paciencia,
54 y su inmensa comprensión y calidad humana. Aprecio mucho las conversaciones de
55 la vida que tuvimos tomando un café. Fue realmente un honor trabajar con usted.

56 Finalmente, agradezco al proyecto FONDECYT 1231599 por financiar esta tesis.

57 *Adiós, adiós, adiós, ciudad de mi querer, donde con ilusión mi carrera*
58 *estudié. Adiós, mi universidad, cuyo reloj no volveré a escuchar...*

Índice general

59		
60	Agradecimientos	IV
61	Índice de figuras	VII
62	Introducción	1
63	1. Preliminares	3
64	1.1. Teoría de grafos	4
65	1.2. Álgebra lineal y teoría espectral de grafos	9
66	1.3. Grafos aleatorios	18
67	2. Cuasi-aleatoriedad	21
68	2.1. Teorema de Chung, Graham y Wilson	22
69	2.2. Aspectos adicionales del teorema de CGW	35
70	3. Lema de regularidad de Szemerédi	42
71	3.1. Demostración por incremento de energía	47
72	3.2. Demostración espectral	57
73	3.3. Aplicaciones	62
74	3.3.1. Eliminación de triángulos	65

75	3.3.2. Emparejamiento inducido	67
76	3.3.3. Teorema de Roth	70
77	Bibliografía	75

Índice de figuras

78		
79	1.1. Ejemplo de un grafo	4
80	1.2. Ejemplo de un grafo 3-partito	7
81	1.3. Ejemplos de un grafo especial $K_{3,5}$, K_6 y 3-regular	7
82	1.4. Ejemplo de una caminata y un ciclo en un grafo	8
83	1.5. Ejemplo de una copia etiquetada	9
84	1.6. Ejemplo de la matriz de adyacencia de un grafo	10
85	2.1. Esquema de un contraejemplo	36
86	3.1. Diagrama de flujo de la demostración del teorema de Szemerédi . . .	43
87	3.2. Ilustración del resultado del lema de regularidad de Szemerédi	44
88	3.3. Ejemplo de un par ε -regular	45
89	3.4. Ilustración del procedimiento de refinamiento simultaneo	53
90	3.5. Ejemplo de un emparejamiento y un emparejamiento inducido	68
91	3.6. Esquema construcción grafo en demostración del teorema de Roth . .	71
92	3.7. Ejemplo de emparejamiento generado por una clase de equivalencia .	73
93	3.8. Esquema de esquina formada en el teorema de Ajtai-Szemerédi	74

Introducción

Esta tesis tiene lugar en la *combinatoria extremal*, que en la actualidad, es una de las ramas más populares en la matemática discreta. Una figura destacada del área es el matemático húngaro Endre Szemerédi, cuyo logro más conocido es su demostración de lo que hoy se conoce como *teorema de Szemerédi* (1975), pero que antes era una famosa conjetura con décadas de antigüedad propuesta por Erdős y Turán. Sus aportes fueron tales que, en 2012, fue nombrado ganador del premio Abel con la siguiente cita:

“Por sus contribuciones fundamentales a las matemáticas discretas y la informática teórica, y en reconocimiento al profundo y duradero impacto de sus aportaciones sobre la teoría aditiva de números y la teoría ergódica.”

Uno de los ingredientes en la demostración del teorema de Szemerédi es el hoy denominado *lema de regularidad*, el cual se ha singularizado como un potente instrumento con múltiples aplicaciones y ha sido analizado desde diversas perspectivas. Antes del lema de regularidad, los grafos se consideraban como objetos carentes de estructura, después de todo, al especificar un grafo se decide para cada par de vértices si unirlos con una arista, y no hay restricción en absoluto sobre la decisión. No obstante, Szemerédi notó que con la aleatoriedad se puede dar una descripción estructural útil a cualquier grafo completamente arbitrario. De manera muy general, el lema dice lo siguiente: dado cualquier grafo, existe una forma de dividir sus vértices en un número finito de conjuntos de tal manera que si se observan las aristas que

unen a dos de esos conjuntos, parecen haber sido elegidas al azar. En otras palabras, cada grafo está compuesto por un número acotado de “grafos similares a aleatorios”. Esto nos indica que se puede dar una buena descripción de un grafo utilizando una cantidad pequeña de datos, puesto que una vez se han encontrado los conjuntos que proporciona el lema, solo hay que decir aproximadamente cuántas aristas hay entre cada par de esos conjuntos, y sabemos de antemano que estarán distribuidas de manera que parezcan aleatorias.

La tesis se divide esencialmente en tres partes. En la primera de ellas, se establecen las bases para las dos siguientes, abordando conceptos elementales de la teoría clásica y espectral de grafos, álgebra lineal, y la noción de un grafo aleatorio. La segunda, fundamenta la frase “grafos similares a aleatorios” con la idea de *cuasi-aleatoriedad*, es decir, el estudio de estructuras discretas deterministas que tienen propiedades con comportamientos esperados según su contraparte aleatoria. En particular, se presenta y demuestra rigurosamente el célebre teorema de Chung, Graham y Wilson, el cual entrega una extensa lista de propiedades equivalentes a la noción de cuasi-aleatoriedad. Por último, en la tercera parte, se exponen dos demostraciones del lema de regularidad de Szemerédi; la versión clásica vía *incremento de energía*, y otra propuesta por Terence Tao usando un enfoque espectral. Seguido de esto, se estudia cómo utilizar el lema de regularidad en dos resultados particulares, los cuales prueban de manera independiente un caso particular del teorema de Szemerédi, el teorema de Roth.

Capítulo 1

Preliminares

Este capítulo proporciona una introducción concisa de los conceptos y terminologías que se utilizarán en esta tesis. La sección 1.1 repasa las nociones más básicas de la teoría de grafos, otorgando una línea de base para el desarrollo del documento. En la sección 1.2 se repasan algunos conceptos y resultados clásicos del álgebra lineal para abordar las propiedades necesarias de la teoría espectral de grafos. Por último, la sección 1.3 contextualiza y motiva el contenido de la sección 2.

En muchos de los resultados de esta tesis, la *desigualdad de Cauchy-Schwarz* (DCS) es un argumento fundamental en sus demostraciones. En particular, se emplearán dos variantes que se enunciarán a continuación. Primero, recuerde que la DCS establece que todo $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$ satisfacen

$$\sum_{i=1}^k a_i^2 \sum_{i=1}^k b_i^2 \geq \left(\sum_{i=1}^k a_i b_i \right)^2. \quad (1.1)$$

Entonces, si $\mathbf{b} = (1, \dots, 1)$, se obtiene la primera variante:

$$\sum_{i=1}^k a_i^2 \geq \frac{1}{k} \left(\sum_{i=1}^k a_i \right)^2. \quad (1.2)$$

Adicionalmente, considerando los reales $\alpha_1, \dots, \alpha_k > 0$ y $\beta_1, \dots, \beta_k \geq 0$, defina $a_i = \sqrt{\alpha_i}$ y $b_i = \frac{\beta_i}{\sqrt{\alpha_i}}$ para conseguir la segunda variante:

$$\sum_{i=1}^k \frac{\beta_i^2}{\alpha_i} \geq \frac{\left(\sum_{i=1}^k \beta_i\right)^2}{\sum_{i=1}^k \alpha_i}. \quad (1.3)$$

152 Por otro lado, será usual utilizar la notación asintótica para destacar la intuición
 153 de algunos resultados. Por esto, se define la notación considerando $f, g \neq 0$ como
 154 funciones de n :

155 ■ Si $\lim_{n \rightarrow \infty} f(n)/g(n) \rightarrow 0$, se dice que $f = o(g)$.

156 ■ Si $\lim_{n \rightarrow \infty} f(n)/g(n) < \infty$, se dice que $f = O(g)$.

157 1.1. Teoría de grafos

158 Se denota al conjunto de los primeros n naturales por $[n] := \{1, 2, \dots, n\}$. También,
 159 si S es un conjunto finito y r es un entero positivo, se establece $\binom{S}{r}$ como el conjunto
 160 de todos los subconjuntos de r elementos de S .

161 Un **grafo** es un par $G = (V, E)$, donde V representa el conjunto de **vértices**, y
 162 $E \subseteq \binom{V}{2}$ el conjunto de **aristas**. Dado un grafo G , se escribe $V(G)$ como su conjunto
 163 de vértices, $E(G)$ como su conjunto de aristas, y $e_G := |E(G)|$ como la cantidad de
 164 aristas presentes en el grafo.

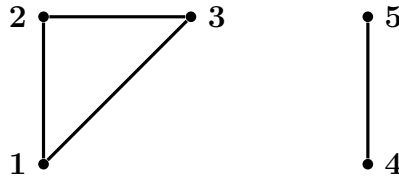


Figura 1.1: Ejemplo de un grafo con conjunto de vértices $V = \{1, 2, 3, 4, 5\}$ y conjunto de aristas $E = \{12, 23, 13, 45\}$.

165 Dado un grafo cualquiera $G = (V, E)$ y $u, v \in V$, se dirá que u es **adyacente**
 166 a v (o viceversa) si y solamente si $uv \in E$. Si $X, Y \subset V$ son dos subconjuntos no
 167 necesariamente disjuntos, se define el conjunto de tuplas que forman una arista en
 168 G de la siguiente manera:

$$e(X, Y) := \left| \{ (x, y) \in X \times Y : xy \in E \} \right|. \quad (1.4)$$

169 Cuando $X \cap Y = \emptyset$, $e(X, Y)$ cuenta el número de aristas entre X e Y , y cuando
 170 $X \cap Y \neq \emptyset$, $e(X, Y)$ realiza un doble conteo sobre las aristas que se encuentran en
 171 $X \cap Y$. Se entenderá por **vecindad** de $u \in V$ como el conjunto de todos los vértices
 172 adyacentes a u , es decir,

$$N(u) := \{v \in V(G) : uv \in E(G)\}. \quad (1.5)$$

173 Si $\mathbb{1}_X$ denota la función indicatriz de un conjunto X , se define el **grado** de un
 174 vértice $u \in V$ con respecto a algún subconjunto de vértices $Y \subseteq V$ de la siguiente
 175 manera:

$$\deg(u; Y) := \sum_{v \in Y} \mathbb{1}_E(uv) = |N(u) \cap Y|.$$

176 En particular, cuando $Y = V$,

$$\deg(u) = \sum_{v \in V} \mathbb{1}_E(uv) = |N(u)|.$$

177 Una propiedad elemental en teoría de grafos es la relación que guarda la suma del
 178 grado de todos los vértices y la cantidad de aristas de un grafo.

179 **Proposición 1.** *Dado un grafo $G = (V, E)$, entonces*

$$\sum_{u \in V} \deg(u) = 2e_G. \quad (1.6)$$

180 *Demostración.* Cada arista $uv \in E$ será contada dos veces en la suma, una contri-
 181 bución por u , y otra por v . □

182 En algunas ocasiones estaremos interesados en la cantidad de vecinos que com-
 183 parten dos vértices del grafo $G = (V, E)$. Entonces, se define el **cogrado** de un par
 184 de vértices $u, v \in V$ no necesariamente diferentes mediante:

$$\text{codeg}(u, v) = \sum_{w \in V} \mathbb{1}_E(wu) \mathbb{1}_E(wv) = |N(u) \cap N(v)|.$$

185 Mostraremos que existe una relación intrínseca entre los conceptos de grado y
186 cogrado, la cual será de utilidad en la sección 2.

187 **Proposición 2.** Sea $G = (V, E)$ un grafo e $Y \subset V$ un subconjunto de vértices,
188 entonces

$$\sum_{u \in V} \deg(u; Y)^2 = \sum_{v \in Y} \sum_{v' \in Y} \text{codeg}(v, v').$$

189 *Demostración.* Utilizando las respectivas definiciones de grado y cogrado, el resulta-
190 do se obtiene de seguir el siguiente cálculo:

$$\begin{aligned} \sum_{u \in V} \deg(u; Y)^2 &= \sum_{u \in V} \sum_{v \in Y} \sum_{v' \in Y} \mathbb{1}_E(uv) \mathbb{1}_E(uv') \\ &= \sum_{v \in Y} \sum_{v' \in Y} \sum_{u \in V} \mathbb{1}_E(vu) \mathbb{1}_E(v'u) \\ &= \sum_{v \in Y} \sum_{v' \in Y} \text{codeg}(v, v'). \end{aligned}$$

191 □

192 Observe que en particular, cuando $Y = V$, se satisface

$$\sum_{u \in V} \deg(u)^2 = \sum_{u \in V} \sum_{v \in V} \text{codeg}(u, v). \quad (1.7)$$

193 A continuación, se enuncian algunos grafos especiales que son contemplados en
194 esta tesis. Diremos que un grafo $G = (V, E)$ es k -**partito** si V se puede dividir en
195 k subconjuntos disjuntos V_1, V_2, \dots, V_k tales que si $uv \in E$ entonces $u \in V_i$ y $v \in V_j$,
196 con $i \neq j$. En particular, a un grafo 2-partito lo llamaremos **bipartito**.

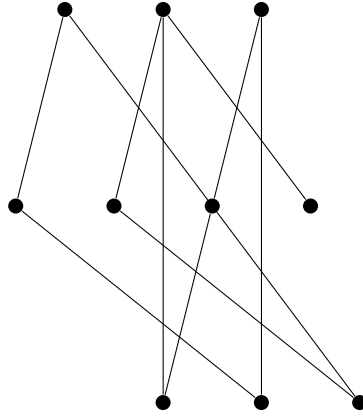
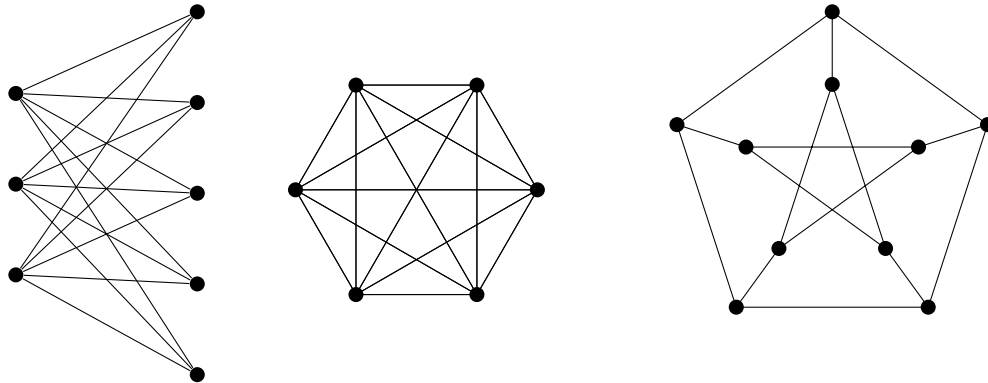


Figura 1.2: Ejemplo de un grafo 3-partito.

197 Un **grafo completo** de n vértices, denotado por K_n , es un grafo en el cual todos
 198 sus vértices son adyacentes entre ellos, es decir, todo par de vértices en el grafo posee
 199 una arista que los conecta. Similarmente, se denota por $K_{n,m}$ al **grafo bipartito**
 200 **completo** con n y m elementos en sus respectivos conjuntos de vértices. Observe
 201 que la cantidad de aristas en los grafos anteriores son exactamente $e_{K_n} = \binom{n}{2}$ y
 202 $e_{K_{n,m}} = n \cdot m$. Por último, un grafo d -**regular** es aquel que presenta todos sus
 203 vértices con grado d .

Figura 1.3: Ejemplo de los grafos especiales $K_{3,5}$, K_6 y 3-regular.

204 Otro término relevante en este trabajo son las diferentes nociones de rutas que
 205 se pueden encontrar siguiendo una determinada secuencia de aristas en un grafo.
 206 Suponga que el grafo G posee $n \geq k$ vértices, entonces se definen los siguientes
 207 conceptos:

- 208 ■ Una **caminata**, es una secuencia de vértices no necesariamente distintos
 209 v_0, v_1, \dots, v_k tales que $v_{i-1}v_i \in E(G)$ para todo $i \in [k]$. Si $v_0 = v_k$, se dice que
 210 es una **caminata cerrada**. El **largo** de una caminata está determinado por
 211 la cantidad de aristas que esta posea.
- 212 ■ Un **ciclo**, es una caminata con $k \geq 2$ vértices únicos a excepción de v_k , que
 213 coincide con v_0 . Se denotará por C_k al ciclo de largo k .

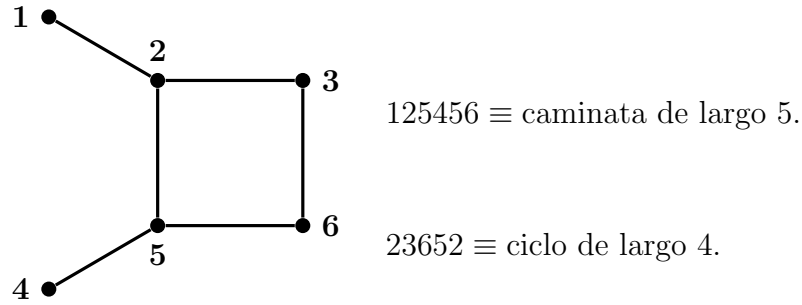


Figura 1.4: Ejemplo de una caminata y un ciclo.

214 Por otro lado, para estudiar posteriormente relaciones estructurales entre grafos, se
 215 define un **isomorfismo** entre los grafos H y G como una biyección $f : V(H) \rightarrow V(G)$
 216 tal que $uv \in E(H)$ si y solamente si $f(u)f(v) \in E(G)$. Si existe tal biyección, diremos
 217 que H y G son isomorfos.

218 Finalmente, se define una **copia etiquetada** de un grafo H en G , como la aplica-
 219 ción inyectiva $f : V(H) \rightarrow V(G)$ tal que $f(u)f(v) \in E(G)$ cada vez que $uv \in E(H)$.
 220 En otras palabras, es un mapeo de los vértices de H a los de G que preserva las
 221 aristas. Se denotará por $\binom{G}{H}$ al conjunto de copias etiquetadas de H en G .

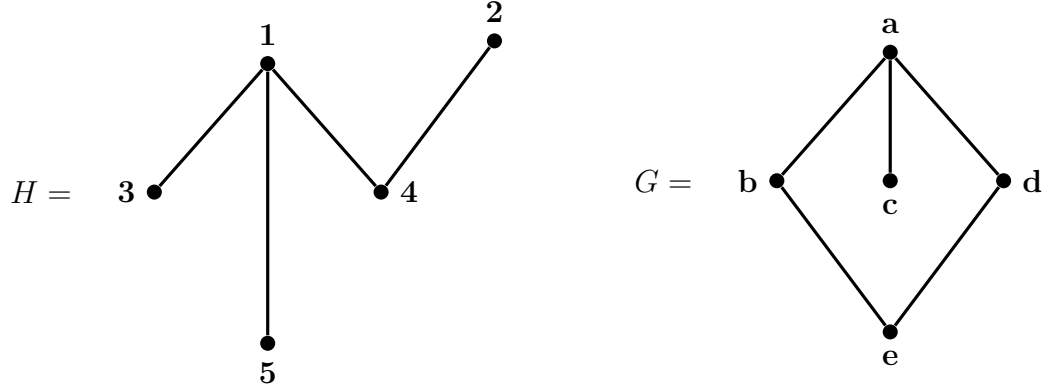


Figura 1.5: Ejemplo de una copia etiquetada de H en G mediante la función $f : V(H) \rightarrow V(G)$ definida por $f(1) = a$, $f(2) = e$, $f(3) = c$, $f(4) = b$ y $f(5) = d$.

1.2. Álgebra lineal y teoría espectral de grafos

Se define $\mathcal{M}_{n \times m}(\mathbb{R})$ como el conjunto de matrices reales de n filas y m columnas, y denotaremos A^T a la matriz traspuesta de $A \in \mathcal{M}_{n \times m}(\mathbb{R})$. También, representaremos por $\mathbf{1} \in \mathcal{M}_{n \times 1}(\mathbb{R})$ al vector de solo 1-entradas, $J \in \mathcal{M}_{n \times n}(\mathbb{R})$ a la matriz de solo 1-entradas, $I_n \in \mathcal{M}_{n \times n}(\mathbb{R})$ a la matriz identidad, y $\mathbf{e}_i \in \mathcal{M}_{n \times 1}(\mathbb{R})$ como el vector de la base canónica de \mathbb{R}^n con entrada 1 en la posición i . Además, $\|\cdot\|$ y $\langle \cdot, \cdot \rangle$ representarán en todo momento la norma y producto interno usuales de \mathbb{R}^n (\mathbb{C}^n , según corresponda) respectivamente.

Considerando una matriz cuadrada $A \in \mathcal{M}_{n \times n}(\mathbb{R})$, se define la **traza** de A como la suma de los elementos de su diagonal principal. Esto es,

$$\text{Tr}(A) = a_{11} + a_{22} + \dots + a_{nn}.$$

Si $A, B \in \mathcal{M}_{n \times n}(\mathbb{R})$, entonces la traza resulta invariante bajo el orden de multiplicación de dichas matrices. En efecto,

$$\text{Tr}(AB) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} b_{ji} \right) = \sum_{j=1}^n \left(\sum_{i=1}^n a_{ij} b_{ji} \right) = \text{Tr}(BA).$$

Una manera muy útil de representar un grafo es mediante una matriz cuadrada binaria, en la que sus filas y columnas representarán todos los vértices del grafo. La matriz adopta el valor 1 en las coordenadas en que sus respectivos vértices forman una arista, y 0 cuando no. Bajo esta representación, se consigue una visión clara y eficiente entre las relaciones de los vértices del grafo y se goza de las propiedades que ofrece el álgebra lineal.

Definición 3. Dado un grafo G sobre n vértices, se define su **matriz de adyacencia** $A_G \in \mathcal{M}_{n \times n}(\mathbb{R})$ de la siguiente manera:

$$a_{ij} = \begin{cases} 1 & \text{si } ij \in E(G), \\ 0 & \text{en otro caso.} \end{cases}$$

Cuando el contexto sea claro, la matriz de adyacencia será representada simplemente por A .

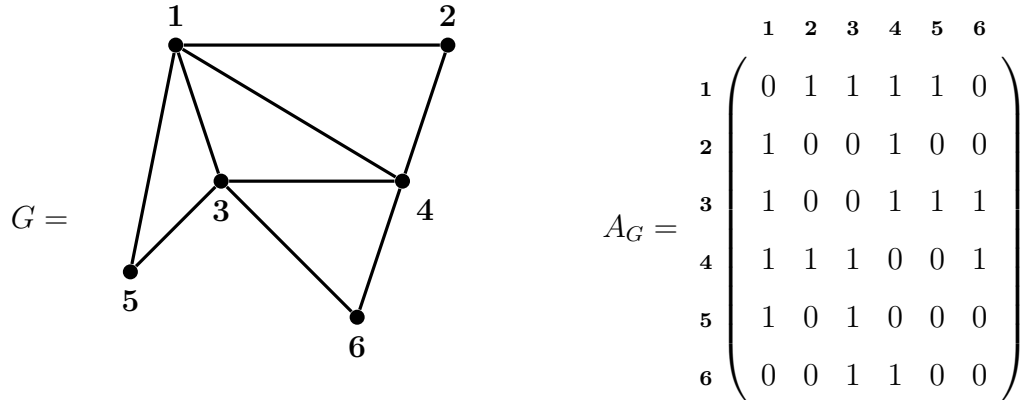


Figura 1.6: Ejemplo de la representación de un grafo mediante la matriz de adyacencia.

Observe que la construcción anterior resulta siempre en una matriz simétrica, es decir, $A_G^T = A_G$. Además, a partir de todo grafo $G = ([n], E)$ con matriz de

246 adyacencia A , se puede obtener un vector con los grados de todos los vértices del
 247 grafo aplicando el operador A al vector de 1-entradas:

$$A\mathbf{1} = \begin{pmatrix} \deg(1) \\ \vdots \\ \deg(n) \end{pmatrix} \quad (1.8)$$

248 .

249 Otro aspecto interesante de la matriz de adyacencia que será de utilidad en la
 250 sección 3, es que nos permite reescribir la ecuación (1.4) en función de ella. Para
 251 ver esto, considere la matriz de adyacencia A del grafo $G = ([n], E)$, y los vértices
 252 $i, j \in [n]$. Luego, según la definición 3,

$$e(\{i\}, \{j\}) = \mathbf{e}_i^T A \mathbf{e}_j = a_{ij}.$$

253 Y así, por linealidad, se extiende el resultado anterior sobre cualquier conjunto
 254 $X, Y \subset [n]$.

$$e(X, Y) = \sum_{i \in X} \sum_{j \in Y} \mathbf{e}_i^T A \mathbf{e}_j = \mathbf{v}_X^T A \mathbf{v}_Y. \quad (1.9)$$

255 En la ecuación anterior, y desde ahora en adelante, el vector $\mathbf{v}_X = \sum_{i \in X} \mathbf{e}_i$
 256 representa el vector indicador del subconjunto de vértices $X \subset [n]$ de algún grafo
 257 $G = ([n], E)$.

258 Es importante destacar que la matriz de adyacencia no solo describe las conexiones
 259 entre cada par de vértices en un grafo, sino que también revela la cantidad exacta
 260 de caminatas que existen entre dos vértices de un largo determinado. En específico,
 261 la posición ij de la t -ésima potencia de la matriz de adyacencia de un grafo guarda
 262 la cantidad de caminatas de largo t entre los vértices i y j .

263 **Proposición 4.** *Sea A la matriz de adyacencia del grafo $G = ([n], E)$. La (i, j) -*
 264 *ésima entrada $a_{ij}^{(t)}$ de A^t , cuenta la cantidad de caminatas de largo t que comienzan*
 265 *y terminan en los vértices i y j respectivamente.*

266 *Demostración.* Cuando $t = 1$, existe una caminata de largo 1 que conecta los vértices
 267 i y j si y solamente si $a_{ij}^{(1)} = 1$. Ahora, asuma que el lema se cumple para algún $t > 1$
 268 fijo. Note que cualquier caminata de largo $t + 1$ entre i y j contiene una caminata
 269 de largo t desde i hasta un vecino de j , digamos k . Entonces si $k \in N(j)$, por la
 270 asunción del lema, el número de caminatas de largo t entre i y k es $a_{ik}^{(t)}$. Por lo tanto,
 271 el número total de caminatas de largo $t + 1$ desde i hasta j es exactamente

$$\sum_{k \in V} a_{ik}^{(t)} \mathbb{1}_{N(j)}(k) = \sum_{\ell=1}^n a_{i\ell}^{(t)} a_{\ell j} = a_{ij}^{(t+1)}.$$

272

□

273 Como consecuencia de la proposición anterior, en cualquier grafo $G = ([n], E)$ con
 274 matriz de adyacencia A , se puede representar la cantidad total de caminatas cerradas
 275 de largo t en el grafo por medio de la traza, $\text{Tr}(A^t) = \sum_{i=1}^n a_{ii}^{(t)}$. Con esto, note que
 276 $\text{Tr}(A^2) = 2e_G$. Con esencialmente la misma demostración, se enuncia el siguiente
 277 corolario que nos será de utilidad más adelante.

278 **Corolario 5.** Sea la matriz cuadrada $F = (f_{ij}) \in \mathcal{M}_{n \times n}(\mathbb{R})$, entonces $\text{Tr}(F^2) =$
 279 $\sum_{(i,j) \in [n]^2} f_{ij}^2$.

280 Por otro lado, para introducir algunos aspectos de la teoría espectral de grafos,
 281 recuerde que el vector no nulo $\mathbf{v} \in \mathbb{R}^n$ es un **vector propio** de alguna matriz
 282 $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ con **valor propio** $\lambda \in \mathbb{C}$ si $A\mathbf{v} = \lambda\mathbf{v}$. Esto significa que λ es un
 283 valor propio si y solo si $\lambda I_n - A$ es una matriz singular. Así, los valores propios
 284 vienen dados por las raíces del polinomio característico $\det(xI_n - A)$. En este trabajo,
 285 cuando se haga referencia a los valores y vectores propios de un grafo G , siempre será
 286 con respecto a su matriz de adyacencia A . Por ejemplo, si G es un grafo d -regular,
 287 entonces con la igualdad (1.8) se puede deducir que d es el valor propio asociado al
 288 vector propio normalizado de 1-entradas de la matriz de adyacencia A_G .

289 **Proposición 6.** Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz real simétrica, entonces todos sus
 290 valores propios son reales. Además, si dos vectores propios están asociados a distintos

291 valores propios, entonces éstos son ortogonales. Más aún, el conjunto de todos los
 292 vectores propios define una base ortonormal de \mathbb{R}^n .

293 *Demostración.* Se comienza probando que los valores propios de A son reales. Sea
 294 λ un valor propio de A y $\mathbf{x} \neq 0$ su correspondiente vector propio. Tomando su
 295 conjugado (denotado por \bar{z} al complejo conjugado de $z \in \mathbb{C}$), se obtiene paralelamente
 296 que

$$\begin{array}{ccc} A\mathbf{x} = \lambda\mathbf{x} & & A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}} \\ \Downarrow & & \Downarrow \\ \bar{\mathbf{x}}^T A\mathbf{x} = \lambda\|\mathbf{x}\|^2 & & \mathbf{x}^T A\bar{\mathbf{x}} = \bar{\lambda}\|\mathbf{x}\|^2. \end{array}$$

297 Además, como A es simétrica,

$$\bar{\mathbf{x}}^T A\mathbf{x} = (A\mathbf{x})^T \bar{\mathbf{x}} = \mathbf{x}^T A\bar{\mathbf{x}}.$$

298 Así, ya que $\mathbf{x} \neq 0$, debe ocurrir que $\lambda = \bar{\lambda}$, permitiendo concluir que todos los
 299 valores propios de A son números reales.

300 Por otro lado, considere $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vectores propios distintos de A asociados a los
 301 valores propios $\lambda, \mu \in \mathbb{R} \setminus \{0\}$ respectivamente. Calculamos como sigue,

$$\lambda\langle \mathbf{u}, \mathbf{v} \rangle = \langle \lambda\mathbf{u}, \mathbf{v} \rangle = \langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A^T \mathbf{v} \rangle = \langle \mathbf{u}, A\mathbf{v} \rangle = \langle \mathbf{u}, \mu\mathbf{v} \rangle = \mu\langle \mathbf{u}, \mathbf{v} \rangle.$$

302 De esta manera, $\lambda\langle \mathbf{u}, \mathbf{v} \rangle = \mu\langle \mathbf{u}, \mathbf{v} \rangle$ si y solamente si $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Ya probada
 303 la ortogonalidad de los vectores propios de A , defina $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ como el
 304 conjunto de vectores propios normalizados de A para probar que \mathcal{B} constituye una
 305 base ortonormal de \mathbb{R}^n . Para esto, sean $c_1, \dots, c_n \in \mathbb{R}$ tales que

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_n\mathbf{u}_n = 0.$$

Entonces, para cualquier $i \in [n]$, multiplicando por la izquierda la igualdad anterior por \mathbf{u}_i^T ,

$$\mathbf{u}_i^T(c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n) = c_i\mathbf{u}_i^T\mathbf{u}_i = c_i = 0.$$

Así, queda demostrado que \mathcal{B} es una base ortonormal de \mathbb{R}^n . \square

A continuación, se enunciará sin demostración uno de los teoremas más importantes del álgebra lineal, y se estudiarán algunas de sus consecuencias bajo el contexto de la teoría de grafos.

Teorema 7. (*Teorema espectral*) Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz real simétrica. Entonces existen matrices P ortogonal y D diagonal tales que

$$A = PDP^T = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T. \quad (1.10)$$

En donde la matriz diagonal D está compuesta por los valores propios $\lambda_i \in \mathbb{R}$ de A , y las columnas de P son los vectores propios ortonormales $\mathbf{v}_i \in \mathbb{R}^n$ de A .

Con el teorema anterior es posible representar la matriz de adyacencia de un grafo por medio de dos matrices que dependen únicamente de sus valores y vectores propios para obtener propiedades desde una perspectiva espectral.

En primera instancia, observe que la descomposición espectral (1.10) permite trabajar eficientemente con las potencias de una matriz real simétrica, puesto a que el problema se reduce a calcular la respectiva potencia de la matriz diagonal de valores propios. Para visualizar este hecho, primero observe cómo se comporta el cuadrado de una matriz simétrica $A \in \mathcal{M}_{n \times n}(\mathbb{R})$:

$$A^2 = (PDP^T)(PDP^T) = PD(P^TP)DP^T = PD^2P^T.$$

Luego, de manera inductiva se obtiene que $A^k = PD^kP^T$. Esta propiedad resulta altamente útil de cara al cálculo de caminatas de largo k entre dos vértices de un

326 grafo. Más aún, la Proposición 8 y el Corolario 9 mostrarán que el número de ca-
 327 minatas cerradas en un grafo queda totalmente determinado por los valores propios
 328 del mismo.

329 **Proposición 8.** *La traza de toda matriz simétrica $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ es igual a la suma*
 330 *de sus valores propios.*

331 *Demostración.* Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz simétrica, $\lambda_1, \dots, \lambda_n$ sus valores propios,
 332 y $\mathbf{v}_1, \dots, \mathbf{v}_n$ sus vectores propios. Se escribe la traza estratégicamente de la siguiente
 333 manera:

$$\text{Tr}(A) = \sum_{i=1}^n \mathbf{e}_i^T A \mathbf{e}_i.$$

334 Con esto, se concluye utilizando la descomposición espectral.

$$\begin{aligned} \text{Tr}(A) &= \sum_{i=1}^n \mathbf{e}_i^T \left(\sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T \right) \mathbf{e}_i \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \mathbf{e}_i^T \mathbf{v}_j \mathbf{v}_j^T \mathbf{e}_i \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \langle \mathbf{e}_i, \mathbf{v}_j \rangle \langle \mathbf{v}_j, \mathbf{e}_i \rangle \\ &= \sum_{j=1}^n \lambda_j \|\mathbf{v}_j\|^2 \\ &= \sum_{j=1}^n \lambda_j. \end{aligned}$$

335

□

336 El siguiente corolario extiende el resultado anterior sobre cualquier potencia de
 337 una matriz real simétrica.

338 **Corolario 9.** *Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz simétrica y $\lambda_1, \dots, \lambda_n$ sus valores pro-*
 339 *prios, entonces se cumple $\text{Tr}(A^k) = \sum_{i=1}^n \lambda_i^k$.*

340 *Demostración.* El resultado sigue de utilizar que la traza de la multiplicación de dos
 341 matrices es invariante bajo el orden de la multiplicación,

$$\text{Tr}(A^k) = \text{Tr}([PDP^T]^k) = \text{Tr}(P[D^kP^T]) = \text{Tr}([D^kP^T]P) = \text{Tr}(D^k) = \sum_{i=1}^n \lambda_i^k.$$

342

□

343 De esta manera, la cantidad de caminatas cerradas de largo k entre dos vértices
 344 de un grafo se simplifica a solo calcular la suma de la k -ésima potencia de todos sus
 345 valores propios. Más adelante, en la sección 2, esta propiedad será de utilidad debido
 346 a que entrega una buena aproximación de la cantidad de copias etiquetadas de ciclos
 347 de largo k que existen en un grafo $G = ([n], E)$. En particular, si A es la matriz de
 348 adyacencia de G y $\lambda_1, \dots, \lambda_n$ sus valores propios,

$$\left| \binom{G}{C_k} \right| = \text{Tr}(A^k) + o(n^k) = \sum_{i=1}^n \lambda_i^k + o(n^k) \quad (1.11)$$

349 Finalmente, es posible caracterizar cada valor propio de una matriz real simétrica
 350 por medio del teorema de *Courant-Fischer*. Se enuncia sin demostración.

351 **Teorema 10.** (*Teorema de Courant-Fischer*) Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz real
 352 simétrica, cuyos valores propios son $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, y $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ sus vectores
 353 propios. Entonces,

(i)

$$\lambda_k = \inf_{\substack{\mathbf{x} \perp \{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

(ii)

$$\lambda_k = \sup_{\substack{\mathbf{x} \perp \{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\} \\ \mathbf{x} \neq \mathbf{0}}} \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

354 **Corolario 11.** Sea λ_1 el valor propio más grande de la matriz simétrica $A \in$
 355 $\mathcal{M}_{n \times n}(\mathbb{R})$, entonces

$$\lambda_1 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

356 *Demostración.* Si \mathbf{v}_1 un vector propio de A correspondiente a λ_1 , entonces

$$\lambda_1 = \frac{\|A\mathbf{v}_1\|}{\|\mathbf{v}_1\|} \leq \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

357 Por otro lado, observando que el valor propio más grande de A^2 es λ_1^2 , se concluye
 358 para cualquier $\mathbf{x} \in \mathbb{R}^n$ que

$$\|A\mathbf{x}\|^2 = \langle A\mathbf{x}, A\mathbf{x} \rangle = \langle \mathbf{x}, A^T A\mathbf{x} \rangle = \langle \mathbf{x}, A^2\mathbf{x} \rangle \leq \lambda_1^n \langle \mathbf{x}, \mathbf{x} \rangle = \lambda_1^2 \|\mathbf{x}\|^2.$$

359

□

360 Usualmente, el primer valor propio de todo grafo juega un papel protagónico.
 361 Para los fines de estas tesis, es necesario establecer una cota inferior del primer valor
 362 propio.

363 **Proposición 12.** *El primer valor propio de la matriz de adyacencia de un grafo es*
 364 *al menos el promedio de los grados. En particular, cuando el grafo es d -regular, el*
 365 *primer valor propio coincide con d .*

366 *Demostración.* Considerando A como la matriz de adyacencia del grafo $G = ([n], E)$,
 367 se desarrolla en función del Teorema 10:

$$\lambda_1 = \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{x} \neq \mathbf{0}}} \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} \geq \frac{\langle \mathbf{1}, A\mathbf{1} \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle} = \frac{2e_G}{n} = \frac{\sum_{v \in V(G)} \deg(v)}{n}.$$

368 Adicionalmente, con apoyo de la igualdad (1.8) y usando la cota anterior, se
 369 concluye que $\lambda_1 = d$ cada vez que G es un grafo d -regular. □

1.3. Grafos aleatorios

El objetivo de la presente sección es proporcionar un breve contexto teórico para abordar más adelante la noción y propiedades de los grafos *cuasi-aleatorios*.

Intuitivamente, se podría pensar en un *grafo aleatorio* de n vértices como el resultado de seleccionar aleatoriamente un subconjunto de aristas de K_n . En 1959, P. Erdős y A. Rényi [7] (también E. Gilbert [11]) proponen dicha selección de la siguiente manera: comenzando con un grafo sin aristas $G = ([n], \emptyset)$, decidir sobre cada par de vértices de G si agregar una arista con una probabilidad p establecida. En cada repetición del proceso anterior se genera un nuevo grafo de n vértices, que contribuye a la creación del espacio de probabilidad conocido como $G(n, p)$, y se denomina modelo binomial. Entonces, considerando \mathcal{G}^n como el conjunto de todos los grafos de n vértices, se define formalmente.

Definición 13. (*Modelo binomial*) Sea $p \in (0, 1)$. Se define $G(n, p)$ como el espacio de probabilidad $(\mathcal{G}^n, \mathcal{P}(\mathcal{G}^n), \mathbb{P})$, con

$$\mathbb{P}(\{G\}) = p^{e_G}(1-p)^{\binom{n}{2}-e_G}, \quad \forall G \sim G(n, p).$$

Diremos que $\mathcal{P}_n \subset \mathcal{G}^n$ es una propiedad de grafos si es cerrada bajo isomorfismos de grafos. Más aún, G satisface la propiedad \mathcal{P}_n **con alta probabilidad** si $\mathbb{P}(\mathcal{P}_n) \rightarrow 1$ cuando $n \rightarrow \infty$. Dicho esto, se probará que $G(n, p)$ posee una distribución de aristas en el siguiente sentido:

Proposición 14. Sea $\varepsilon > 0$ y $p \in (0, 1)$. Si $G \sim G(n, p)$, entonces satisface con alta probabilidad la siguiente propiedad:

$$\mathcal{P}_n^{p, \varepsilon} := \{G \in \mathcal{G}^n : |e(A, B) - p|A||B|| \leq \varepsilon n^2, \quad \forall A, B \subset V(G)\}.$$

Para dar prueba a la proposición anterior es necesario utilizar la desigualdad de Chernov. Existiendo diversas formas de expresar tal desigualdad, en esta tesis se

392 utiliza el resultado para el caso en que cada variable aleatoria solo toma los valores
 393 0 o 1, como se plantea en [15], en la ecuación (2.12) de la observación 2.5.

394 **Teorema 15.** (*Desigualdad de Chernov*) Sean X_1, \dots, X_N variables aleatorias inde-
 395 pendientes tales que $X_i = 1$ con probabilidad p , y $X_i = 0$ con probabilidad $1 - p$.
 396 Entonces, si $X = \sum_{i=1}^N X_i$, se satisface

$$\mathbb{P}(|X - \mathbb{E}[X]| > t) \leq 2 \exp\left(-\frac{2t^2}{N}\right), \quad \forall t \geq 0.$$

397 Con esto, damos paso a la demostración prometida.

398 *Demostración Proposición 14.* Dado $p \in (0, 1)$ y $\varepsilon > 0$, considere $G \sim G(n, p)$ y
 399 $A, B \subset V(G)$. Defina la variable aleatoria $X = e(A, B) = \sum_{a \in A} \sum_{b \in B} X_{ab}$, en donde

$$X_{ab} = \begin{cases} 1 & \text{si } ab \in E(G), \\ 0 & \text{en otro caso.} \end{cases}$$

400 Para utilizar la cota de Chernov más adelante, se separa la variable aleatoria X
 401 en sumas de variables aleatorias independientes. Vale decir $X = X_1 + X_2$, en donde

$$X_1 = 2 \sum_{ab \in \binom{A \cap B}{2}} X_{ab}, \text{ y } X_2 = \sum_{\substack{a \in A, b \in B \\ a \neq b \\ ab \notin \binom{A \cap B}{2}}} X_{ab}.$$

402 Al calcular la esperanza de X_1 y X_2 se obtiene el siguiente resultado:

$$\mathbb{E}[X_1] = \mu_1 = 2p \binom{|A \cap B|}{2}, \text{ y } \mathbb{E}[X_2] = \mu_2 = p \left(|A||B| - |A \cap B| - 2 \binom{|A \cap B|}{2} \right).$$

403 Notando ahora que $|A||B| \leq n^2$, se utiliza la desigualdad de Chernov con $t = \frac{\varepsilon}{3}n^2$
 404 sobre $i \in \{1, 2\}$ para obtener lo siguiente:

$$\begin{aligned}
\mathcal{P}\left(|X_i - \mu_i| > \frac{\varepsilon}{3}n^2\right) &\leq 2 \exp\left(-\frac{2\left(\frac{\varepsilon}{3}n^2\right)^2}{|A||B|}\right) \\
&\leq 2 \exp\left(-\frac{2}{9}\varepsilon^2 n^2\right).
\end{aligned} \tag{1.12}$$

405 Luego, si ocurre simultáneamente que $|X_1 - \mu_1| \leq \frac{\varepsilon}{3}n^2$ y $|X_2 - \mu_2| \leq \frac{\varepsilon}{3}n^2$, entonces

$$\left|X_1 + X_2 - (\mu_1 + \mu_2)\right| \leq \frac{2}{3}\varepsilon n^2, \quad \forall A, B \in V(G).$$

406 Y así, como $\mu_1 + \mu_2 = p(|A||B| - |A \cap B|) = p|A||B| \pm \varepsilon n$, se tendrá que todo
 407 $A, B \subset V(G)$ satisface $\left|X - p|A||B|\right| \leq \varepsilon n^2$.

408 Por lo anterior, se concluye utilizando la cota de la unión de la siguiente manera:

$$\begin{aligned}
1 - \mathbb{P}(\mathcal{P}_n^{p,\varepsilon}) &= \mathbb{P}\left(\exists A, B \subset V(G) : \left|X - p|A||B|\right| > \varepsilon n^2\right) \\
&\leq \mathbb{P}\left(\exists A, B \subset V(G) : |X_1 - \mu_1| > \frac{\varepsilon}{3}n^2 \vee |X_2 - \mu_2| > \varepsilon n^2\right) \\
&\leq \mathbb{P}\left(\exists A, B \subset V(G) : |X_1 - \mu_1| > \frac{\varepsilon}{3}n^2\right) + \mathbb{P}\left(|X_2 - \mu_2| > \frac{\varepsilon}{3}n^2\right) \\
&\leq \sum_{\substack{A \subset V(G) \\ B \subset V(G)}} \mathbb{P}\left(|X_1 - \mu_1| > \frac{\varepsilon}{3}n^2\right) + \sum_{\substack{A \subset V(G) \\ B \subset V(G)}} \mathbb{P}\left(|X_2 - \mu_2| > \frac{\varepsilon}{3}n^2\right) \\
&\stackrel{(1.12)}{\leq} 2^{2n+1} \exp\left(-\frac{2}{9}\varepsilon^2 n^2\right) \xrightarrow{n \rightarrow \infty} 0.
\end{aligned}$$

410 Capítulo 2

411 Cuasi-aleatoriedad

412 Trabajar con estructuras discretas aleatorias brinda una amplia gama de pro-
413 piedades ideales y deseables, lo que las convierte en piezas fundamenteles tanto en
414 matemáticas como en ciencias de la computación. Por ejemplo, el modelo de grafo
415 aleatorio binomial goza de una distribución uniforme de aristas, buenas propieda-
416 des y es robusto. El tema ahora es cómo capturar las propiedades esenciales de la
417 aleatoriedad dentro de un marco *determinista*. Esta idea condujo a la noción de
418 cuasi-aleatoriedad, que en la actualidad, es un tópico central en las matemáticas
419 discretas. En términos generales, las propiedades cuasi-aleatorias son características
420 deterministas que son propias de objetos realmente aleatorios. Aunque la noción
421 de cuasi-aleatoriedad es interesante por sí misma, su estudio ha revelado profundas
422 conexiones entre varias ramas de la matemática y ciencias de la computación, encon-
423 trando aplicaciones en teoría de grafos, teoría de números, teoría ergódica, geometría,
424 y algoritmos y complejidad.

425 Como se verá a detalle más adelante en la sección 3, una de las razones principales
426 por las cuales el estudio de la cuasi-aleatoriedad no se limita a un área específica,
427 es el hecho de que existe un teorema de partición que permite la aproximación de
428 cualquier objeto discreto por otros cuasi-aleatorios. Con esto, nos referimos al célebre
429 lema de regularidad de Szemerédi, que establece que todo grafo se puede aproximar
430 mediante un número finito de grafos cuasi-aleatorios, permitiendo la conexión entre

431 un grafo arbitrario y los cuasi-aleatorios.

432 El estudio sistemático de los grafos cuasi-aleatorios fue iniciado por V. Rödl [20] y
 433 A. Thomason [27], y su punto inicial es la siguiente noción de *distribución uniforme*
 434 *de aristas* para definir la cuasi-aleatoriedad de un grafo.

435 **Definición 16.** Sea $p \in (0, 1)$ y $(G_n)_{n \rightarrow \infty}$ una secuencia de grafos, en donde cada
 436 G_n posee n vértices. Entonces el grafo G_n es **cuasi-aleatorio** si en todo par de
 437 subconjuntos $X, Y \subset V(G_n)$ se encuentra una distribución de aristas similar, es
 438 decir,

$$e(X, Y) = p|X||Y| + o(n^2). \quad (2.1)$$

439 En otras palabras, la distribución uniforme de aristas establece que, hasta el
 440 término de error $o(n^2)$, cualquier par de subconjuntos de vértices poseen tantas
 441 aristas como se esperaría de un grafo aleatorio $G(n, p)$. Es importante destacar que
 442 esta propiedad no solo se cumple con alta probabilidad en un grafo aleatorio $G(n, p)$,
 443 sino que también se considera como una de sus características distintivas.

444 2.1. Teorema de Chung, Graham y Wilson

445 Una contribución revolucionaria en la teoría de grafos cuasi-aleatorios fue en 1989
 446 por Fan Chung, Ronald Graham y Richard M. Wilson [4]. Ellos presentaron una
 447 extensa lista de propiedades superficialmente diferentes entre sí, y demostraron que
 448 todas son equivalentes al concepto de cuasi-aleatoriedad entendido en la Definición
 449 2.

450 En la presente sección se enuncia el teorema de Chung, Graham y Wilson junto
 451 a una demostración formal.

452 **Teorema 17.** (Chung, Graham y Wilson) Sea $p \in (0, 1)$ fijo. Para cualquier secuen-
 453 cia de grafos $(G_n)_{n \rightarrow \infty}$ con $|V(G_n)| = n$ vértices y $e_{G_n} = (p + o(1))\binom{n}{2}$ aristas, las
 454 siguientes propiedades son equivalentes:

455 DISC_p : Para todo $X, Y \subseteq (G_n)$,

$$\left| e(X, Y) - p|X||Y| \right| = o(n^2).$$

456 DISC'_p : Para todo $X \subseteq V(G_n)$,

$$\left| e(X) - p\binom{|X|}{2} \right| = o(n^2).$$

457 COUNT_p : Para cada grafo H , la cantidad de copias etiquetadas de H en G_n está
458 dada por

$$\left| \binom{G_n}{H} \right| = (p^{e(H)} + o(1)) n^{(H)}.$$

459 $\text{COUNT}_{C_4, p}$: La cantidad de copias etiquetadas de ciclos de largo 4 es

$$\left| \binom{G_n}{C_4} \right| = (p^4 + o(1)) n^4.$$

CODEG_p :

$$\sum_{u, v \in V(G_n)} \left| \text{codeg}(u, v) - p^2 n \right| = o(n^3).$$

460 EIG_p : Si $\lambda_1 \geq \dots \geq \lambda_n$ son los valores propios de la matriz de adyacencia de
461 G_n , entonces

$$\lambda_1 = pn + o(n) \quad , \quad \max_{i \neq 1} |\lambda_i| = o(n).$$

462 Para una comprensión e intuición inicial de cada propiedad del Teorema 17, se ha
463 utilizado notación asintótica en su enunciado. Sin embargo, con dicha formulación no
464 queda del todo claro las dependencias cuantificadas de los errores en las implicancias
465 para cada par de propiedades. Entonces, se replantean equivalentemente las propie-
466 dades con una versión cuantitativa, asociando algún parámetro de error ε en todo
467 grafo específico G con un conjunto de vértices suficientemente grande. Por ejemplo,

bajo los supuestos del Teorema 17, asuma que la sucesión de grafos $(G_n)_{n \rightarrow \infty}$ satisface DISC_p , y luego, la versión equivalente establece que para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que el grafo G sobre $n \geq n_0$ vértices satisface

$$\text{DISC}_p(\varepsilon) : \quad e(X, Y) = p|X||Y| \pm \varepsilon n^2, \quad \forall X, Y \subseteq V(G).$$

De manera general, diremos que una secuencia de grafos $(G_n)_{n \rightarrow \infty}$ con $|V(G_n)| = n$ satisface la propiedad P_{x_1, \dots, x_k} ¹ si para cada elección de $\varepsilon > 0$, existe algún $n_0 \in \mathbb{N}$ tal que el grafo G con $n \geq n_0$ vértices satisface $P_{x_1, \dots, x_k}(\varepsilon)$. Más aún, se dirá que la propiedad Q_{y_1, \dots, y_ℓ} implica la propiedad P_{x_1, \dots, x_k} si y solamente si $P_{x_1, \dots, x_k}(\varepsilon)$ implica $Q_{y_1, \dots, y_\ell}(\delta)$. Es decir, para todo $\varepsilon > 0$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G con $n \geq n_0$ vértices cumple con $Q_{y_1, \dots, y_\ell}(\delta)$ cada vez que satisfaga la propiedad $P_{x_1, \dots, x_k}(\varepsilon)$. Se desarrollará la demostración formal del Teorema 17 utilizando notación ε - δ , mostrando que cada par de propiedades P_{x_1, \dots, x_k} y Q_{y_1, \dots, y_ℓ} son equivalentes entre sí con un cambio polinomial en el error, esto es, $P_{x_1, \dots, x_k}(\varepsilon) \Rightarrow Q_{y_1, \dots, y_\ell}(C\varepsilon^c)$ para algún par de constantes $C, c > 0$.

Demostración Teorema de Chung, Graham y Wilson

La demostración del Teorema fue descompuesta en ocho proposiciones, las cuales mostrarán la equivalencia entre todas las propiedades conforme al siguiente esquema:

$$\begin{array}{ccccc}
 \text{DISC}'_p & \xRightarrow{\text{Prop. 20.}} & \text{COUNT}_p & \xRightarrow{\text{Prop. 21.}} & \text{COUNT}_{C4,p} & \xLeftrightarrow{\text{Prop. 24. y 25.}} & \text{EIG}_p \\
 \Downarrow \text{Prop. 18 y 19.} & & & & \Downarrow \text{Prop. 22.} & & \\
 \text{DISC}_p & & \xLeftrightarrow{\text{Prop. 23.}} & & \text{CODEG}_p & &
 \end{array} \tag{2.2}$$

Con esto en mente, damos paso a la demostración de cada proposición considerada en el esquema (2.2).

¹Los parámetros x_1, \dots, x_k pueden ser de distinta naturaleza, dependiendo de la propiedad simbolizada. En las propiedades del Teorema 17 se utiliza $k = 1$ con $x_1 = p$ salvo en la propiedad $\text{COUNT}_{C4,p}$, en donde $k = 2$.

Proposición 18. Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre $n \geq n_0$ vértices satisface $\text{DISC}'_p(\varepsilon)$ cada vez que cumpla con $\text{DISC}_p(\delta)$. En particular,

$$\text{DISC}_p \Rightarrow \text{DISC}'_p.$$

Demostración. Dado $\varepsilon > 0$ y $p \in (0, 1)$, se elige $\delta < \frac{\varepsilon}{2}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Entonces, considerando el grafo G con $n \geq n_0$ vértices que satisface $\text{DISC}_p(\delta)$ y $X \subset V(G)$, se utiliza la propiedad $\text{DISC}_p(\delta)$ para obtener el resultado de la siguiente manera:

$$e(X) = p \frac{|X|^2}{2} \pm \delta n^2 = p \binom{|X|}{2} \pm 2\delta n^2.$$

Las igualdades anteriores consideran, por definición, $e(X, X) = 2e(X)$, y la aproximación $\binom{|X|}{2} = \frac{|X|^2}{2} \pm \delta n^2$. \square

Proposición 19. Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre $n \geq n_0$ vértices satisface $\text{DISC}_p(\varepsilon)$ cada vez que cumpla con $\text{DISC}'_p(\delta)$. En particular,

$$\text{DISC}'_p \Rightarrow \text{DISC}_p.$$

Demostración. Dado $\varepsilon > 0$ y $p \in (0, 1)$, se elige $\delta < \frac{\varepsilon}{4}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Considere también el grafo G sobre $n \geq n_0$ vértices que satisface $\text{DISC}'_p(\delta)$.

En primera instancia, se lleva el conteo de aristas que existen entre pares de subconjuntos de vértices a un conteo equivalente mediante la combinación aditiva de las aristas que se encuentran en un subconjunto único de vértices. Es decir, para $X, Y \subset V(G)$,

$$e(X, Y) = e(X \cup Y) + e(X \cap Y) - e(X \setminus Y) - e(Y \setminus X). \quad (2.3)$$

504 Observe que con esta configuración, el conteo de las aristas entre X e Y es doble
 505 cuando los vértices que componen las aristas pertenecen a $X \cap Y$. Luego, se utiliza
 506 la propiedad $\text{DISC}'_p(\delta)$ sobre la identidad (2.3) para conseguir el resultado.

$$\begin{aligned} e(X, Y) &= p \left(\binom{|X \cup Y|}{2} + \binom{|X \cap Y|}{2} - \binom{|X \setminus Y|}{2} - \binom{|Y \setminus X|}{2} \right) \pm 4\delta n^2 \\ &= p|X||Y| \pm 4\delta n^2 \\ &= p|X||Y| \pm \varepsilon n^2. \end{aligned}$$

507

□

508 **Proposición 20.** Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que
 509 el grafo G sobre $n \geq n_0$ vértices satisface $\text{COUNT}_p(\varepsilon)$ cada vez que cumpla con
 510 $\text{DISC}'_p(\delta)$. En otras palabras,

$$\text{DISC}'_p \Rightarrow \text{COUNT}_p.$$

511 *Demostración.* Sea $\varepsilon > 0$, $p \in (0, 1)$ y H un grafo sobre ℓ vértices, elegimos $\delta < \frac{\varepsilon}{6\ell^2}$
 512 y $n_0 \in \mathbb{N}$ suficientemente grande. Considere también el grafo $G = (V, E)$ con $n \geq n_0$
 513 vértices que satisface la propiedad $\text{COUNT}_p(\varepsilon)$.

514 Dado cualquier grafo F con ℓ vértices y $e_F \geq 1$ aristas, razonamos por inducción
 515 sobre su cantidad de aristas para probar que

$$\left| \binom{G}{F} \right| = p^{e_F} n^\ell \pm 4e_F \delta n^\ell. \quad (2.4)$$

516 Una vez probada la ecuación (2.4), el resultado seguirá de tomar $F = H$ y la
 517 elección de δ para conseguir las siguientes desigualdades:

$$4e_F \delta n^\ell \leq 4 \binom{\ell}{2} \delta n^\ell \leq 4\delta \left(\frac{\ell^2}{2} + \delta \ell^2 \right) n^\ell \leq 6\delta \ell^2 n^\ell < \varepsilon n^\ell.$$

518 Entonces, cuando $e_F = 1$, $\left| \binom{G}{F} \right|$ es el número de pares ordenados de vértices de G

519 que forman una arista junto a cualquier combinación de $\ell - 2$ vértices para completar
 520 una copia de F . Es decir,

$$\left| \binom{G}{F} \right| = 2e_G(n-2)(n-3) \cdots (n-\ell+1).$$

521 Luego, si aplicamos la propiedad $\text{DISC}'_p(\delta)$ sobre V , se obtiene que la cantidad
 522 de aristas es

$$e_G = \frac{pn(n-1)}{2} \pm \delta n^2.$$

523 Así, con $\left| \binom{G}{F} \right| = pn^\ell \pm 4\delta n^\ell$, se prueba el caso inicial de la inducción. Ahora, sea
 524 F un grafo con $e_F > 1$ aristas y asuma que se satisface la ecuación (2.4) en cual-
 525 quier grafo con una cantidad de aristas menor que e_F . Para desarrollar la inducción,
 526 suponga que $ij \in E(F)$ y considere la siguiente notación:

- 527 i) F^- corresponde al grafo producido por eliminar la arista ij de F .
- 528 ii) F^* es el resultado de eliminar los vértices de la arista ij en F .

529 Sea T^- una copia etiquetada de F^- en G , es decir, T^- se corresponde una apli-
 530 cación inyectiva $f : V(F^-) \rightarrow V(T^-) \subseteq V$ tal que $f(u)f(v) \in E(T^-)$ cada vez
 531 que $uv \in E(F^-)$. Entonces, considerando $e_{T^-} := f(i)f(j)$, se escribe la cantidad
 532 de copias etiquetadas de F en G de manera conveniente para utilizar la hipótesis
 533 inductiva como se muestra a continuación:

$$\begin{aligned}
| \binom{G}{F} | &= \sum_{T^- \in \binom{G}{F^-}} \mathbb{1}_E(e_{T^-}) \\
&= \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) + p - p] \\
&= \sum_{T^- \in \binom{G}{F^-}} p + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
&= p | \binom{G}{F^-} | + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
&\stackrel{(2.4)}{=} p^{e_F} n^\ell + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \pm 4(e_F - 1)\delta n^\ell. \tag{2.5}
\end{aligned}$$

534 En este punto, es suficiente probar que el segundo sumando de la igualdad (2.5) es
535 pequeño. Para esto, considere T^* una copia de F^* , y denote por F_i^* y F_j^* a los grafos
536 resultantes de eliminar de F^- los vértices j e i respectivamente. Con esto, defina los
537 siguientes conjuntos:

$$\begin{aligned}
A_i^{T^*} &:= \{v \in V : T^* \text{ con } v \text{ forma una copia de } F_i^*\} \\
A_j^{T^*} &:= \{v \in V : T^* \text{ con } v \text{ forma una copia de } F_j^*\}.
\end{aligned}$$

539 Los conjuntos anteriores, por construcción, son tales que para cada tupla $(a, b) \in$
540 $A_i^{T^*} \times A_j^{T^*}$ añadida a T^* se obtiene una copia de F^- . Así, reescribiendo el segundo
541 sumando de la igualdad (2.5) convenientemente y utilizando la propiedad $\text{DISC}'_p(\delta)$,

$$\begin{aligned}
\left| \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \right| &= \left| \sum_{T^* \in \binom{G}{F^*}} \sum_{f \in A_i^{T^*} \times A_j^{T^*}} [\mathbb{1}_E(f) - p] \right| \\
&\leq \sum_{T^* \in \binom{G}{F^*}} \left| \sum_{f \in A_i^{T^*} \times A_j^{T^*}} [\mathbb{1}_E(f) - p] \right| \\
&= \sum_{T^* \in \binom{G}{F^*}} |e(A_i^{T^*}, A_j^{T^*}) - p| |A_i^{T^*}| |A_j^{T^*}| \\
&\leq \sum_{T^* \in \binom{G}{F^*}} \delta n^2 \\
&\leq 4\delta n^\ell.
\end{aligned}$$

542 De esta manera, tomando la elección de δ y $F = H$ se obtiene el resultado. \square

543 **Proposición 21.** Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que
544 el grafo G sobre $n \geq n_0$ vértices satisface $\text{COUNT}_{C_4, p}(\varepsilon)$ cada vez que cumpla con
545 $\text{COUNT}_p(\delta)$. En otras palabras,

$$\text{COUNT}_p \Rightarrow \text{COUNT}_{C_4, p}.$$

546 *Demostración.* Se trata de un caso particular de COUNT_p , en donde $H = C_4$ y
547 $\delta < \varepsilon$. \square

548 **Proposición 22.** Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el
549 grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{CODEG}_p(\varepsilon)$ cada
550 vez que cumpla con $\text{COUNT}_{C_4, p}(\delta)$. En particular,

$$\text{COUNT}_{C_4, p} \Rightarrow \text{CODEG}_p.$$

551 *Demostración.* Dado $\varepsilon > 0$ y $p \in (0, 1)$, elegimos $\delta < \frac{\varepsilon^2}{16}$ y $n_0 \in \mathbb{N}$ suficientemente
552 grande. También considere el grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas
553 que satisface $\text{COUNT}_{C_4, p}(\delta)$.

La clave de esta demostración radica en encontrar una buena cota para $\sum_{u,v \in V(G)} \text{codeg}(u, v)$ y $\sum_{u,v \in V(G)} \text{codeg}(u, v)^2$. Para esto, será necesario la utilización apropiada de la desigualdad de Cauchy-Schwarz vista en (1.2). Por un lado, con la relación entre el grado y el cogrado (1.7) se obtiene la primera de las cotas:

$$\begin{aligned}
 \sum_{u,v \in V(G)} \text{codeg}(u, v) &= \sum_{x \in V(G)} \deg(x)^2 \\
 &\stackrel{\text{DCS}}{\geq} \frac{1}{n} \left(\sum_{x \in V(G)} \deg(x) \right)^2 \\
 &= \frac{4e_G^2}{n} \\
 &\geq \frac{4}{n} \left(\frac{pn^2}{2} - \delta n^2 \right)^2 \\
 &\geq p^2 n^3 - 4\delta n^3.
 \end{aligned}$$

Por otro lado, usando $\text{COUNT}_{C_4, p}(\delta)$,

$$\sum_{u,v \in V(G)} \text{codeg}(u, v)^2 = \left| \binom{G}{C_4} \right| \pm \delta n^4 \leq p^4 n^4 + 2\delta n^4.$$

Así, con las cotas anteriores, se obtiene el resultado de la siguiente manera:

$$\begin{aligned}
 \sum_{u,v \in V(G)} \left| \text{codeg}(u, v) - p^2 n \right| &\stackrel{\text{DCS}}{\leq} n \left(\sum_{u,v \in V(G)} (\text{codeg}(u, v) - p^2 n)^2 \right)^{1/2} \\
 &= n \left(\sum_{u,v \in V(G)} \text{codeg}(u, v)^2 - 2p^2 n \sum_{u,v \in V(G)} \text{codeg}(u, v) + \sum_{u,v \in V(G)} p^4 n^2 \right)^{1/2} \\
 &\leq n \left(p^4 n^4 + 2\delta n^4 + 2p^2 n(4\delta n^3 - p^2 n^3) + p^4 n^4 \right)^{1/2} \\
 &= n((2 + 8p^2)\delta n^4)^{1/2} \\
 &\leq 4\delta^{1/2} n^3.
 \end{aligned}$$

Proposición 23. Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{DISC}_p(\varepsilon)$ cada vez que cumpla con $\text{CODEG}_p(\delta)$. En particular,

$$\text{CODEG}_p \Rightarrow \text{DISC}_p.$$

Demostración. Dado $\varepsilon > 0$, $p \in (0, 1)$, seleccionamos $\delta < \frac{\varepsilon^4}{81}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Sea G un grafo de $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas que satisface la propiedad $\text{CODEG}_p(\delta)$.

En primera instancia, note que la propiedad $\text{CODEG}_p(\delta)$ induce una concentración en los grados de los vértices de G . En efecto,

$$\begin{aligned} \sum_{x \in V(G)} \left| \deg(x) - pn \right| &\stackrel{\text{DCS}}{\leq} n^{1/2} \left(\sum_{x \in V(G)} (\deg(x) - pn)^2 \right)^{1/2} \\ &= n^{1/2} \left(\sum_{x \in V(G)} \deg(x)^2 - 2pn \sum_{x \in V(G)} \deg(x) + p^2 n^3 \right)^{1/2} \\ &\stackrel{(1.7)}{=} n^{1/2} \left(\left(\sum_{u, v \in V(G)} \text{codeg}(u, v) - p^2 n \right) - 4pne_G + 2p^2 n^3 \right)^{1/2} \\ &\leq n^{1/2} \left(\left(\sum_{u, v \in V(G)} \left| \text{codeg}(u, v) - p^2 n \right| \right) + 4pn \left(\delta n^2 - \frac{pn^2}{2} \right) + 2p^2 n^3 \right)^{1/2} \\ &\leq n^{1/2} (2p^2 n^3 - 2p^2 n^3 + 4p\delta n^3 + \delta n^3)^{1/2} \\ &< 3\delta^{1/2} n^2. \end{aligned}$$

Luego, para todo $X, Y \in V(G)$, se reescribe la expresión de la propiedad DISC_p de forma conveniente para posteriormente utilizar la desigualdad de Cauchy-Schwarz.

$$\begin{aligned}
\left| e(X, Y) - p|X||Y| \right| &= \left| \sum_{x \in X} (\deg(x; Y) - p|Y|) \right| \\
&\stackrel{DCS}{\leq} n^{1/2} \left(\sum_{x \in X} (\deg(x; Y) - p|Y|)^2 \right)^{1/2}. \quad (2.6)
\end{aligned}$$

571 En la desigualdad anterior se ha conseguido que el argumento de la suma sea
572 siempre no negativo, lo que permite extender su dominio de X a $V(G)$. De esta
573 manera, usando a la cota proveniente de la concentración de los grados en los vértices
574 de G , se prueba el resultado continuando desde (2.6):

$$\begin{aligned}
\left| e(X, Y) - p|X||Y| \right| &\leq n^{1/2} \left(\sum_{x \in V(G)} (\deg(x; Y) - p|Y|)^2 \right)^{1/2} \\
&= n^{1/2} \left(\sum_{x \in V(G)} \deg(x; Y)^2 - 2p|Y| \sum_{x \in V(G)} \deg(x; Y) + \sum_{x \in V(G)} p^2|Y|^2 \right)^{1/2} \\
&= n^{1/2} \left(2p^2n|Y|^2 - p^2n|Y|^2 + 2p|Y||Y|pn - 2p|Y||Y|pn \right. \\
&\quad \left. + \sum_{y, y' \in Y} \text{codeg}(y, y') - 2p|Y| \sum_{y \in Y} \deg(y) \right)^{1/2} \\
&= n^{1/2} \left(\sum_{y, y' \in Y} (\text{codeg}(y, y') - p^2n) - 2p|Y| \sum_{y \in Y} (\deg(y) - pn) \right)^{1/2} \\
&\leq n^{1/2} \left(\left| \sum_{y, y' \in Y} (\text{codeg}(y, y') - p^2n) \right| + \left| 2p|Y| \sum_{y \in Y} (\deg(y) - pn) \right| \right)^{1/2} \\
&\leq n^{1/2} \left(\sum_{u, v \in V(G)} \left| \text{codeg}(u, v) - p^2n \right| + 2p|Y| \sum_{x \in V(G)} \left| \deg(x) - pn \right| \right)^{1/2} \\
&\leq n^{1/2} (\delta n^3 + 6p\delta^{1/2}n^3)^{1/2} \\
&< 3\delta^{1/4}n^2.
\end{aligned}$$

Proposición 24. Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{COUNT}_{C_4, p}(\varepsilon)$ cada vez que cumpla con $\text{EIG}_p(\delta)$. En particular,

$$\text{EIG}_p \Rightarrow \text{COUNT}_{C_4, p}.$$

Demostración. Dado $\varepsilon > 0$ y $p \in (0, 1)$, se elige $\delta < \frac{\varepsilon}{20}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Consideramos el grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas que satisface la propiedad $\text{EIG}_p(\delta)$, $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ como la matriz de adyacencia de G , y $\lambda_1 \geq \dots \geq \lambda_n$ los valores propios de A .

Como vimos en la ecuación (1.11), la cantidad de copias etiquetadas de caminatas cerradas de largo 4 que no son C_4 en G , se encuentran dentro de un error de a lo más δn^4 con respecto al número de copias etiquetadas de C_4 en G . Con esto, junto al Lema 4 y el Corolario 9 se obtiene lo siguiente:

$$\left| \binom{G}{C_4} \right| = \text{Tr}(A^4) \pm \delta n^4 = \sum_{i=1}^n \lambda_i^4 \pm \delta n^4 = \lambda_1^4 + \sum_{i=2}^n \lambda_i^4 \pm \delta n^4. \quad (2.7)$$

Luego, recordando que $\text{Tr}(A^2) = 2e_G$, y usando $\text{EIG}_p(\delta)$,

$$\sum_{i=2}^n \lambda_i^4 \leq \max_{i \neq 1} \lambda_i^2 \sum_{i=1}^n \lambda_i^2 \leq \delta n^2 \text{Tr}(A^2) \leq \delta n^2 (pn^2 + 2\delta n^2) \leq 3\delta n^4. \quad (2.8)$$

Finalmente, se concluirá tras usar la propiedad $\text{EIG}_p(\delta)$ sobre el primer valor propio y la cota mostrada en (2.8). Entonces, continuando desde la ecuación (2.7),

$$\left| \binom{G}{C_4} \right| = \lambda_1^4 + \sum_{i=2}^n \lambda_i^4 \pm \delta n^4 \leq p^4 n^4 + 20\delta n^4.$$

590

□

Proposición 25. Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{EIG}_p(\varepsilon)$ cada vez que cumpla la propiedad $\text{COUNT}_{C_4, p}(\delta)$. Es decir,

$$\text{COUNT}_{C_4,p} \Rightarrow \text{EIG}_p.$$

594 *Demostración.* Sea $\varepsilon > 0$ y $p \in (0, 1)$, escogemos $\delta < \frac{\varepsilon^4}{4}$ y $n_0 \in \mathbb{N}$ suficientemente
 595 grande. Sea también G un grafo sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas que
 596 satisface la propiedad $\text{COUNT}_{C_4,p}(\delta)$, $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ la matriz de adyacencia de G ,
 597 y $\lambda_1 \geq \dots \geq \lambda_n$ los valores propios de A .

598 En lo que respecta al primer valor propio, sabemos por un lado que este es al
 599 menos el promedio de los grados gracias al Lema 12. Es decir,

$$\lambda_1 \geq \frac{\sum_{x \in V(G)} \deg(x)}{n} = \frac{2e_G}{n} = \frac{2}{n} \left(\frac{pn^2}{2} \pm \delta n^2 \right) \geq pn - 2\delta n. \quad (2.9)$$

600 Por otro lado, mediante el Lema 4, el Corolario 9 y la propiedad $\text{COUNT}_{C_4,p}(\delta)$,

$$\lambda_1^4 \leq \sum_{i=1}^n \lambda_i^4 = \text{Tr}(A^4) = \left| \binom{G}{C_4} \right| \pm \delta n^4 \leq p^4 n^4 + 2\delta n. \quad (2.10)$$

601 La desigualdad (2.10) implica que $\lambda_1 \leq pn + (2\delta)^{1/4}n$, y en combinación con la
 602 cota vista en (2.9), se obtiene que $\lambda_1 = pn \pm (2\delta)^{1/4}n$. Por último, observe por las
 603 cotas vistas anteriormente que

$$\begin{aligned} \max_{i \neq 1} |\lambda_i|^4 &\leq \sum_{i=2}^n \lambda_i^4 + \lambda_1^4 - \lambda_1^4 \\ &= \text{Tr}(A^4) - \lambda_1^4 \\ &\leq p^4 n^4 + 2\delta n^2 - p^4 n^4 + 2\delta n^4 \\ &= 4\delta n^4. \end{aligned}$$

604 Así, se logra probar el resultado determinando que $\max_{i \neq 1} |\lambda_i| \leq (4\delta)^{1/4}n$. \square

2.2. Aspectos adicionales del teorema de CGW

La noción inicial presentada de un grafo cuasi-aleatorio por distribución de aristas según la Definición 2 contempla verificar que si todo par de subconjuntos de vértices del grafo satisfacen la condición DISC_p para determinar la cuasi-aleatoriedad. En otras palabras, se requiere comprobar un número exponencial de subconjuntos. Por esto, resulta sorprendente que tal propiedad sea equivalente a todas las otras (salvo DISC'_p), debido a que se verifican de manera polinomial. Otro aspecto interesante es que la propiedad más débil $\text{COUNT}_{C_4,p}$, que solo requiere que la condición de conteo sea verdadera para el ciclo C_4 , sea suficientemente sólida para implicar la afirmación de conteo de la propiedad COUNT_p ; que dice que el número de copias etiquetadas de cualquier grafo F de tamaño fijo en $G = ([n], E)$ es aproximadamente el esperado de los grafos aleatorios $G(n, p)$.

A continuación mostraremos que no es suficiente que la condición de conteo sea verdadera para ciclos de largo inferior a 4 para determinar la cuasi-aleatoriedad de un grafo. Para ver esto, se realiza la construcción de un contraejemplo de un grafo que posee la cantidad de copias etiquetadas esperadas de C_3 , pero que no cumple con las condiciones para ser cuasi-aleatorio.

Proposición 26. *Existe un grafo $G = ([n], E)$ con $(\frac{1}{3})^3 n^3 + o(n^3)$ copias etiquetadas de C_3 , pero que no es cuasi-aleatorio.*

Demostración. La idea de la construcción consiste en la combinación de dos grafos, uno con una cantidad mayor que la esperada en un grafo aleatorio $G(n, p)$ de copias etiquetadas de C_3 , y otro con una cantidad menor. Consideramos entonces independientemente los grafos completos K_{n_1} y K_{n_2, n_2} tales que su unión disjunta forma el grafo $G = K_{n_1} \cup K_{n_2, n_2}$ con $n_1 + 2n_2 = n$ vértices.

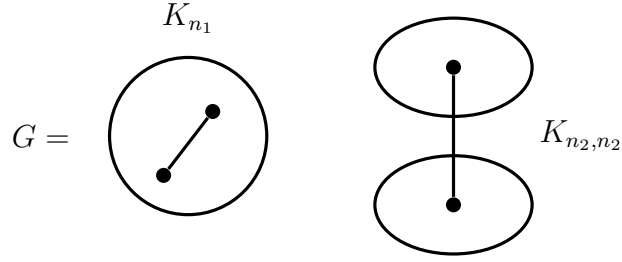


Figura 2.1: Esquema de la configuración del grafo usado como contraejemplo. Aquí, $\bullet\text{---}\bullet$ representa las aristas permitidas dentro del grafo G .

En K_{n_1} y K_{n_2, n_2} , observe que la cantidad de sus aristas y copias etiquetadas de C_3 son las siguientes:

$$\begin{aligned} e_{K_{n_1}} &\approx \frac{n_1^2}{2} & , & \quad \left| \binom{K_{n_1}}{C_3} \right| \approx n_1^3 \quad , \\ e_{K_{n_2, n_2}} &\approx \frac{(n - n_1)^2}{4} & , & \quad \left| \binom{K_{n_2, n_2}}{C_3} \right| = 0. \end{aligned}$$

Bajo esta configuración, se encontrará el parámetro $p \in (0, 1)$ de manera tal que el grafo G posea la cantidad esperada de aristas y copias etiquetadas de C_3 según lo haría un grafo aleatorio $G(n, p)$. Para ello, se plantea el siguiente sistema de ecuaciones:

$$\begin{cases} p \frac{n^2}{2} = \frac{n_1^2}{2} + \frac{(n - n_1)^2}{4}, \\ p^3 n^3 = n_1^3. \end{cases}$$

Resolviendo el sistema anterior, se obtiene que $p = \frac{1}{3}$ y $n_1 = n_2 = \frac{n}{3}$. Esta construcción, en efecto, presenta

$$e_G = \binom{\frac{n}{3}}{2} + \frac{n^2}{9} = \frac{1}{3} \binom{n}{2} + o(n^2),$$

Como también,

$$\left| \binom{G}{C_3} \right| = \left(\frac{n}{3} \right)^3 + o(n^3) = \left(\frac{1}{3} \right)^3 n^3 + o(n^3).$$

638 Sin embargo, el grafo G no es cuasi-aleatorio debido a que no existen aristas entre
 639 K_{n_1} y K_{n_2, n_2} ni dentro de los conjuntos de vértices que conforman a K_{n_2, n_2} . \square

640 Lo expuesto se enfoca en el caso muy particular en el que $p = \frac{1}{3}$, pero es importante
 641 destacar la técnica utilizada. En específico, la interpolación de dos grafos arbitrarios
 642 con una cantidad esperada menor y mayor de copias etiquetadas de C_3 según $G(n, p)$
 643 produce un nuevo contraejemplo.

644 Más aún, es posible extender la propiedad $\text{COUNT}_{C_4, p}$ a $\text{COUNT}_{C_{2t}, p}$ con $t \geq 2$.
 645 Es decir,

$$\text{COUNT}_{C_{2t}, p} : \left| \binom{G_n}{C_{2t}} \right| = (p^{2t} + o(1)) n^{2t}, \quad \forall t \geq 2.$$

646 Se expone un bosquejo de la demostración.

647 **Proposición 27.** Sea $p \in (0, 1)$ y $(G_n)_{n \rightarrow \infty}$ una secuencia de grafos con $|V(G_n)| = n$
 648 vértices y $e_{G_n} = (p + o(1)) \binom{n}{2}$ aristas, entonces las propiedades $\text{COUNT}_{C_{2t}, p}$ y EIG_p
 649 son equivalentes.

650 *Demostración.* Este resultado es una consecuencia directa de las demostraciones de
 651 la Proposición 24 y 25 tras el siguiente par de observaciones. En primer lugar, la
 652 cantidad de copias etiquetadas caminatas cerradas de largo $2t$ que no son C_{2t} en G_n
 653 están dentro de un error $O(n^{2t-1})$, es decir,

$$\text{Tr}(A^{2t}) = \sum_{i=1}^n \lambda_i^{2t} = \left| \binom{G_n}{C_{2t}} \right| + O(n^{2t-1}).$$

654 También, se debe modificar la cota presentada en la ecuación (2.8) como sigue:

$$\sum_{i=2}^n \lambda_i^{2t} \leq \max_{i \neq 1} \lambda_i^{2(t-1)} \sum_{i=1}^n \lambda_i^2 = \max_{i \neq 1} \lambda_i^{2t-2} \text{Tr}(A^2).$$

655 Con estas observaciones el resultado queda demostrado. \square

Es importante destacar que la demostración anterior es funcional gracias a que los ciclos de largo par preservan una contribución positiva en la suma de cada uno de los valores propios de G , eliminando la posibilidad de cancelaciones entre ellos.

Finalmente, se explora un caso siempre interesante de estudio, ya que simplifica varios cálculos y surge de manera recurrente en la vida cotidiana: un grafo d -regular. En nuestro contexto, se verá que toda secuencia $(G_n)_{n \rightarrow \infty}$ de grafos d -regular satisface la propiedad $\text{DISC}_{\frac{d}{n}}$ si y solo si cumple con $\text{EIG}_{\frac{d}{n}}$. Dicha equivalencia nace como una consecuencia del siguiente teorema.

Teorema 28. (*Expander Mixing Lemma*) Sea $G = ([n], E)$ un grafo d -regular, y $d = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ los valores propios asociados a la matriz de adyacencia A de G . Si se denota:

$$\lambda = \max_{i \neq 1} |\lambda_i|.$$

Entonces, para cada $X, Y \subset [n]$,

$$\left| e(X, Y) - \frac{d}{n} |X| |Y| \right| \leq \lambda \sqrt{|X| |Y| \left(1 - \frac{|X|}{n} \right) \left(1 - \frac{|Y|}{n} \right)}. \quad (2.11)$$

Demostración. Sea $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ la base ortonormal de \mathbb{R}^n compuesta por los vectores propios de A . Utilizando la descomposición espectral, denotamos

$$A_1 = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T \quad \text{y} \quad \Delta = \sum_{i=2}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T,$$

de manera que $A = A_1 + \Delta$.

Conforme a la ecuación (1.9), para todo $X, Y \subset [n]$, se cumple la siguiente igualdad:

$$e(X, Y) = \mathbf{v}_X^T A \mathbf{v}_Y = \mathbf{v}_X^T A_1 \mathbf{v}_Y + \mathbf{v}_X^T \Delta \mathbf{v}_Y. \quad (2.12)$$

De la ecuación anterior se espera que el primer sumando sea el término principal,
 mientras que el segundo el factor de error. Para ver esto, se representan los vectores
 \mathbf{v}_X y \mathbf{v}_Y según la base \mathcal{B} . Es decir,

$$\mathbf{v}_X = \sum_{i=1}^n \alpha_i \mathbf{v}_i \quad \text{y} \quad \mathbf{v}_Y = \sum_{i=1}^n \beta_i \mathbf{v}_i,$$

donde $\alpha_i = \mathbf{v}_X^T \mathbf{v}_i$ y $\beta_i = \mathbf{v}_Y^T \mathbf{v}_i$. Con esto, se calcula:

$$\begin{aligned} \|\alpha_i\|^2 &= \sum_{i=1}^n \alpha_i^2 \\ &= \sum_{i=1}^n \langle \mathbf{v}_X, \mathbf{v}_i \rangle^2 \\ &= \sum_{i=1}^n \langle \sum_{j \in X} \mathbf{e}_j, \mathbf{v}_i \rangle^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle \mathbf{e}_j, \mathbf{v}_i \rangle^2 \mathbb{1}_X(i) \\ &= \sum_{i=1}^n \|\mathbf{v}_i\|^2 \mathbb{1}_X(i) \\ &= |X|. \end{aligned}$$

Análogamente, $\|\beta_i\|^2 = \sum_{i=1}^n \beta_i^2 = |Y|$.

Ahora, se estudian los sumandos de la igualdad (2.12) por separado. Por un lado,

$$\begin{aligned} \mathbf{v}_X^T A_1 \mathbf{v}_Y &= \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i \right)^T (\lambda_1 \mathbf{v}_1 \mathbf{v}_1^T) \left(\sum_{j=1}^n \beta_j \mathbf{v}_j \right) \\ &= \lambda_1 \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i^T \right) (\mathbf{v}_1 \mathbf{v}_1^T) \left(\sum_{j=1}^n \beta_j \mathbf{v}_j \right) \\ &= \lambda_1 \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (\mathbf{v}_i^T \mathbf{v}_1) (\mathbf{v}_1^T \mathbf{v}_j) \\ &= \lambda_1 \alpha_1 \beta_1. \end{aligned} \tag{2.13}$$

679 Por otro lado, de la misma manera que el cálculo anterior,

$$\mathbf{v}_X^T \Delta \mathbf{v}_Y = \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i \right)^T \left(\sum_{j=2}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T \right) \left(\sum_{k=1}^n \beta_k \mathbf{v}_k \right) = \sum_{i=1}^n \lambda_i \alpha_i \beta_i. \quad (2.14)$$

680 Luego, dado que G es un grafo d -regular, $\lambda_1 = d$ y $\mathbf{v}_1 = \frac{1}{\sqrt{n}}(1, \dots, 1)^T$ son valor y
681 vector propio respectivamente de A . En consecuencia,

$$\alpha_1 = \frac{|X|}{\sqrt{n}} \quad \text{y} \quad \beta_1 = \frac{|Y|}{\sqrt{n}}.$$

682

683 Así, la ecuación (2.13) resulta en $\mathbf{v}_X^T A_1 \mathbf{v}_Y = \frac{d}{n} |X| |Y|$.

684 Para el término de error, recordando la definición de λ , se desarrolla el valor
685 absoluto de la ecuación (2.14) usando la desigualdad de Cauchy-Schwarz.

$$\begin{aligned} |\mathbf{v}_X^T \Delta \mathbf{v}_Y| &= \left| \sum_{i=2}^n \lambda_i \alpha_i \beta_i \right| \\ &\leq \lambda \left| \sum_{i=2}^n \alpha_i \beta_i \right| \\ &\stackrel{\text{DCS}}{\leq} \lambda \sqrt{\sum_{i=2}^n \alpha_i^2 \sum_{i=2}^n \beta_i^2} \\ &= \lambda \sqrt{(\|\alpha_i\|^2 - \alpha_1^2) (\|\beta_i\|^2 - \beta_1^2)} \\ &= \lambda \sqrt{|X| |Y| \left(1 - \frac{|X|}{n}\right) \left(1 - \frac{|Y|}{n}\right)}. \end{aligned}$$

686 Finalmente, el resultado se obtiene directamente de tomar el valor absoluto de la
687 ecuación (2.12) de la siguiente manera:

$$|e(X, Y) - \mathbf{v}_X^T A_1 \mathbf{v}_Y| = |\mathbf{v}_X^T \Delta \mathbf{v}_Y|.$$

688

□

El teorema anterior permite asegurar que todo grafo d -regular $G = ([n], E)$ con un conjunto de vértices suficientemente grande que satisface la propiedad $\text{EIG}_{\frac{d}{n}}(\delta)$, también cumple con $\text{DISC}_{\frac{d}{n}}(\varepsilon)$. En efecto, para todo $\varepsilon > 0$ y $X, Y \subset [n]$, elija $n_0 \in \mathbb{N}$ suficientemente grande y $\delta < \frac{\varepsilon n}{\sqrt{|X||Y|}}$. Entonces, si G satisface la propiedad $\text{EIG}_{\frac{d}{n}}(\delta)$, por el Teorema 28:

$$\begin{aligned} \left| e(X, Y) - \frac{d}{n}|X||Y| \right| &\leq \lambda \sqrt{|X||Y| \left(1 - \frac{|X|}{n}\right) \left(1 - \frac{|Y|}{n}\right)} \\ &< \delta n \sqrt{|X||Y|} \\ &< \varepsilon n^2. \end{aligned}$$

Finalmente, en un grafo d -regular, la equivalencia entre las propiedades $\text{EIG}_{\frac{d}{n}}$ y $\text{DISC}_{\frac{d}{n}}$ se completa por el camino de implicancias ya demostradas según el esquema (2.2).

697 Capítulo 3

698 Lema de regularidad de Szemerédi

699 El lema de regularidad de Szemerédi ha mostrado ser un resultado muy poderoso
700 e importante en la teoría extremal de grafos. Sus aplicaciones no solo se restringen
701 a la teoría de grafos, sino que también las tiene en teoría de números combinatorios,
702 geometría discreta y ciencias de la computación. A continuación, un poco de historia
703 de su origen.

704 P. Erdős y P. Turán [8] conjeturaron en 1936 que todo conjunto de números en-
705 teros suficientemente grande posee una progresión aritmética de longitud arbitraria,
706 digamos k . En 1953, K. Roth [21] da el primer resultado positivo a la conjetura para
707 el caso en que $k = 3$ utilizando análisis de Fourier, dando paso al teorema de Roth.
708 Más adelante, en 1969, E. Szemerédi [24] extiende el teorema de Roth a progresiones
709 aritméticas de largo 4 vía métodos combinatorios. Seis años después, en 1975, Sze-
710 merédi [25] demuestra la conjetura de Erdős-Turán para progresiones aritméticas de
711 longitud arbitraria, lo que se conoce actualmente como el teorema de Szemerédi.

712 **Teorema 29.** (*Teorema de Szemerédi*) Para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que si
713 el conjunto $S \subset [n]$ posee $|S| > \varepsilon n$ elementos, entonces S contiene una progresión
714 aritmética de largo k no trivial cada vez que $n \geq n_0$.

715 La demostración que propone Szemerédi es considerada una obra maestra del ra-
716 zonamiento combinatorio. El siguiente esquema entrega una noción de la complejidad

717 de la idea.

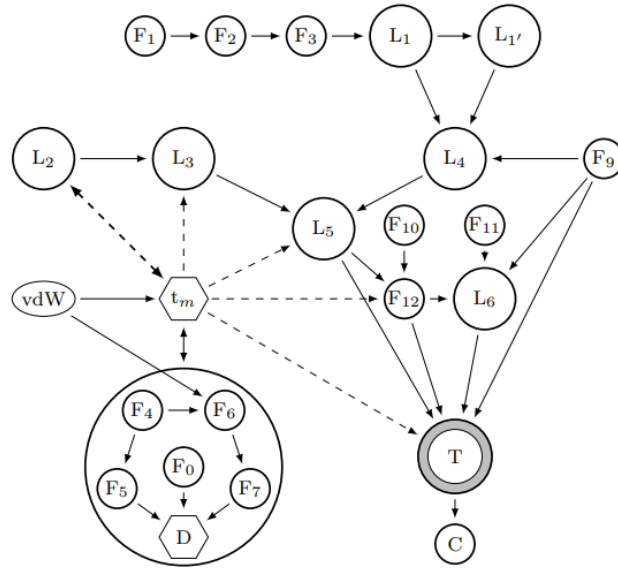


Figura 3.1: El esquema representa el diagrama de flujo utilizado en la demostración del teorema de Szemerédi ([25], página 202). Algunos de los símbolos involucrados significan lo siguiente: $F_k \equiv$ Hecho k , $L_k \equiv$ Lema k , $T \equiv$ Teorema, $C \equiv$ Corolario, $D \equiv$ Definiciones varias, etc.

718 Probablemente la dificultad de tal demostración motivó a otros matemáticos en
 719 encontrar nuevas formas de probar el teorema. Poco después, en 1977, H. Furstenberg
 720 [10] obtiene una demostración fundamentada en la teoría ergódica. También, en 2001,
 721 T. Gowers [13] entrega una tercera demostración por medio de análisis de Fourier,
 722 extendiendo el camino iniciado por Roth para progresiones aritméticas de largo 3.

723 En esta tesis, nos enfocaremos en uno de los elementos que compone la demos-
 724 tración entregada por Szemerédi del Teorema 29, el hoy conocido como *lema de*
 725 *regularidad de Szemerédi*. A grandes rasgos, el lema dice que el conjunto de vértices
 726 de todo grafo puede ser particionado en una cantidad finita de partes que mues-
 727 tran comportamientos *regulares* (o cuasi-aleatorios) entre la mayoría de los pares
 728 de partes. Este hecho permite entender cualquier grafo con menos información, y se
 729 aprovechan cada una de las propiedades equivalentes vistas en el Teorema 17.

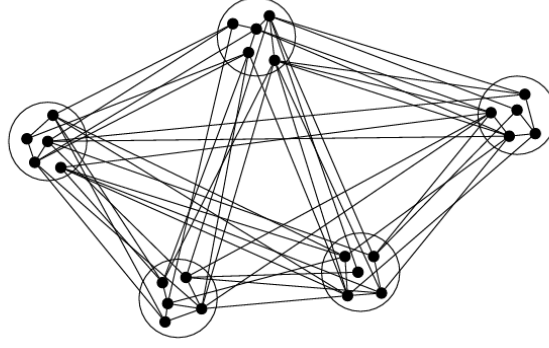


Figura 3.2: Ejemplo ilustrativo con cinco partes del resultado del lema de regularidad de Szemerédi sobre un grafo.

Hasta enunciar formalmente el lema de regularidad, se definirán los conceptos necesarios para su buena comprensión. Dado un grafo G y los subconjuntos de vértices $X, Y \subset V(G)$, se define la **densidad de aristas** entre X e Y de la siguiente manera:

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}. \quad (3.1)$$

Diremos que $\mathcal{P} = \{X_1, X_2, \dots, X_k\}$ es una **partición** del conjunto X si:

- i) $\bigcup_{i=1}^k X_i = X$.
- ii) $X_i \cap X_j = \emptyset$ para todo $i, j \in [k]$.

Cuando $|X_1| \leq |X_2| \leq \dots \leq |X_k| = |X_1| + 1$, llamaremos a \mathcal{P} como una **equipartición**. En particular, cada parte posee $\lceil |X|/k \rceil$ o $\lfloor |X|/k \rfloor$ elementos.

También, es necesario conocer en qué sentido los pares de partes entregados por el lema son regulares.

Definición 30. Sea G un grafo y $X, Y \subset V(G)$ subconjuntos no necesariamente disjuntos. Diremos que (X, Y) es un **par ε -regular** en G si para todo $A \subset X$ y $B \subset Y$ con $|A| \geq \varepsilon|X|$ y $|B| \geq \varepsilon|Y|$, se cumple

$$\left| d(A, B) - d(X, Y) \right| \leq \varepsilon.$$

743 Cuando (X, Y) no es un par ε -regular, entonces la irregularidad es evidenciada
 744 por algún $A \subseteq X$ y $B \subseteq Y$ que satisfacen $|A| \geq \varepsilon|X|$ y $|B| \geq \varepsilon|Y|$, pero $\left|d(A, B) - \right.$
 745 $\left.d(X, Y)\right| > \varepsilon$.

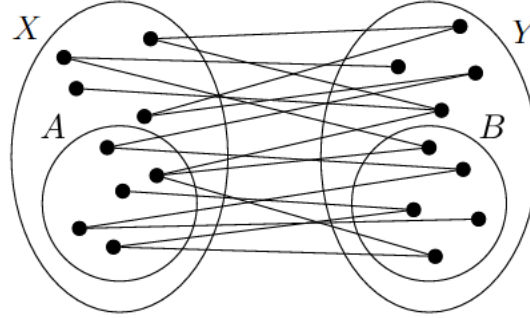


Figura 3.3: Ejemplo del par (X, Y) ε -regular, cuando $X \cap Y = \emptyset$.

Las densidades $d(X, Y)$ y $d(A, B)$ no varían demasiado.

746 Notaremos que la noción de un par ε -regular es, de hecho, una analogía de la
 747 propiedad $\text{DISC}_p(\varepsilon)$ para grafos bipartitos. Es decir, si G es tal que $V(G) = U \cup W$
 748 y $p \in (0, 1)$, se cumple

$$\left|e(X, Y) - p|X||Y|\right| = o(|U||W|), \quad \forall X \subset U, \quad \forall Y \subset W. \quad (3.2)$$

749 En efecto, si (U, W) es un par ε -regular, entonces todo $A \subset U$ y $B \subset W$ tales que
 750 $|A| \geq \varepsilon|U|$ y $|B| \geq \varepsilon|W|$ satisfacen

$$e(A, B) = d(U, W)|A||B| \pm \varepsilon|A||B| = d(U, W)|A||B| \pm \varepsilon|U||W|.$$

751 Ahora bien, si al menos uno de los subconjuntos de la condición de un par ε -regular
 752 no es suficientemente grande, digamos $|A| < \varepsilon|X|$, entonces

$$e(A, B) \leq |A||B| \leq \varepsilon|U||W| < d(U, W)|A||B| + \varepsilon|U||W|.$$

753 Luego, notando que $e(A, B) > d(U, W)|A||B| - \varepsilon|U||W|$, se obtiene la analogía
 754 planteada con $p = d(U, W)$.

755 Por último, debemos saber la noción de regularidad en una partición del conjunto
756 de vértices de un grafo.

757 **Definición 31.** Dado un grafo G , una partición $\mathcal{P} = \{V_1, \dots, V_k\}$ del conjunto de
758 vértices $V(G)$ es una **partición ε -regular** si

$$\sum_{\substack{(i,j) \in [k]^2 \\ (V_i, V_j) \text{ no } \varepsilon\text{-regular}}} |V_i||V_j| \leq \varepsilon |V(G)|^2.$$

759 Es decir, todos los pares de subconjuntos de vértices en la partición son ε -regular
760 salvo una fracción ε de pares de vértices.

761 Note que si una partición ε -regular de k partes es en particular una equipartición,
762 entonces a lo más εk^2 pares de elementos de la partición no son ε -regular.

763 Ya con todo lo requerido, se enuncia formalmente el lema de regularidad.

764 **Teorema 32.** (Lema de regularidad de Szemerédi) Para todo $\varepsilon > 0$, existe un entero
765 $M = M(\varepsilon)$ tal que todo grafo admite una partición ε -regular de a lo más M partes.

766 Otra forma de encontrar el resultado es cuando todas las partes de la partición
767 poseen aproximadamente el mismo tamaño.

768 **Teorema 33.** (Regularidad de Szemerédi - Equipartición) Para todo $\varepsilon > 0$ y $m_0 \in \mathbb{N}$,
769 existe un entero $M = M(\varepsilon)$ tal que todo grafo admite una equipartición ε -regular de
770 su conjunto de vértices de k partes, con $m_0 \leq k \leq M$.

771 Lo poderoso del lema de regularidad es que la cota de partes que entrega no
772 depende del tamaño del grafo, sino que únicamente del parámetro ε . Esto se debe
773 a que en grafos más grandes, el tamaño de las partes podría ser considerablemente
774 mayor.

775 En la subsección 3.1 se demostrará rigurosamente el Teorema 32 desde una mirada
776 clásica, y se expondrá la manera de probar el Teorema 33. Más adelante, en la
777 sección 3.2, nos limitaremos a mostrar una forma novedosa e ingeniosa de demostrar

778 el Teorema 32 desde una perspectiva espectral. Finalmente, la subsección 3.3 tiene
 779 el objetivo de demostrar de dos maneras diferentes el teorema de Roth utilizando el
 780 lema de regularidad de Szemerédi.

781 3.1. Demostración por incremento de energía

782 Se empleará una técnica llamada *argumento de incremento de energía*, la cual
 783 para todo grafo G , funciona algorítmicamente de la siguiente manera:

- 784 1. Comenzar con la partición trivial de $V(G)$, i.e, $\mathcal{P} = \{V(G)\}$.
- 785 2. Mientras la partición actual \mathcal{P} no es ε -regular:
 - 786 (a) Para cada par (V_i, V_j) no ε -regular, encontrar los subconjuntos $A^{ij} \subset V_i$
 787 y $A^{ji} \subset V_j$ que evidencian la irregularidad del par.
 - 788 (b) Refinar \mathcal{P} utilizando simultáneamente los conjuntos A^{ij} y A^{ji} encontrados
 789 de cada par (V_i, V_j) no ε -regular para obtener \mathcal{Q} .
 - 790 (c) Actualizar \mathcal{P} con \mathcal{Q} .

791 Siendo \mathcal{P} y \mathcal{Q} dos particiones de un mismo conjunto de vértices, diremos que \mathcal{Q}
 792 **refina** a \mathcal{P} si cada parte de \mathcal{Q} está contenida en una parte de \mathcal{P} . En lo que resta
 793 de esta subsección, mostraremos que el algoritmo tiene un fin y que entrega una
 794 partición ε -regular en un número de iteraciones que solo depende de ε .

795 **Definición 34.** (*Energía*) Sea G un grafo sobre n vértices y $X, Y \subset V(G)$. Se define
 796 en primer lugar

$$q(X, Y) := \frac{|X||Y|}{n^2} d(X, Y)^2 = \frac{e(X, Y)^2}{n^2 |X||Y|}.$$

797 Luego, para particiones $\mathcal{P}_X = \{X_1, \dots, X_k\}$ de X y $\mathcal{P}_Y = \{Y_1, \dots, Y_\ell\}$ de Y , se
 798 define

$$q(\mathcal{P}_X, \mathcal{P}_Y) := \sum_{i=1}^k \sum_{j=1}^{\ell} q(X_i, Y_j).$$

799 Finalmente, para una partición $\mathcal{P} = \{V_1, \dots, V_k\}$, se define su **energía** mediante

$$q(\mathcal{P}) := \sum_{i=1}^k \sum_{j=1}^k q(V_i, V_j) = \sum_{i=1}^k \sum_{j=1}^k \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2.$$

800 Observe que en toda partición \mathcal{P} de $V(G)$, siempre se tendrá que $0 \leq q(\mathcal{P}) \leq 1$.

801 En efecto,

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i=1}^k \sum_{j=1}^k \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 \\ &\leq \frac{1}{n^2} \sum_{i=1}^k |V_i| \sum_{j=1}^k |V_j| \\ &= 1. \end{aligned}$$

802 La última observación es crucial en la demostración del Teorema 32, puesto que los
 803 Lemas 35, 36 y 37 nos asegurarán que la energía de una partición nunca decrece bajo
 804 refinamiento. Por consecuencia, el algoritmo de la técnica *argumento de incremento*
 805 *de energía* tendrá un fin, y entregará una partición ε -regular. Dicho esto, procedemos
 806 a enunciar y demostrar cada uno de los lemas mencionados para probar clásicamente
 807 el Teorema 32.

808 El primero de los lemas, afirma que la energía de una partición no decrece bajo
 809 cualquier refinamiento arbitrario.

810 **Lema 35.** Sea G un grafo, $X, Y \subset V(G)$, y \mathcal{P}_X y \mathcal{P}_Y particiones de X e Y respecti-
 811 vamente, entonces $q(\mathcal{P}_X, \mathcal{P}_Y) \geq q(X, Y)$. Además, si \mathcal{P} y \mathcal{P}' son dos particiones de
 812 vértices de G , entonces $q(\mathcal{P}) \leq q(\mathcal{P}')$ cada vez que \mathcal{P}' refina a \mathcal{P} .

813 *Demostración.* Considere un grafo G sobre n vértices, los conjuntos $X, Y \subset V(G)$,
 814 y las particiones $\mathcal{P}_X = \{X_1, \dots, X_k\}$ y $\mathcal{P}_Y = \{Y_1, \dots, Y_\ell\}$ de X e Y respectivamente.

815 En primera instancia, se utiliza la desigualdad (1.3) proveniente de Cauchy-Schwarz
 816 para probar que $q(\mathcal{P}_X, \mathcal{P}_Y) \geq q(X, Y)$. Para esto, se desarrolla como sigue:

$$\begin{aligned}
 q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^k \sum_{j=1}^{\ell} q(X_i, Y_j) \\
 &= \frac{1}{n^2} \sum_{i=1}^k \sum_{j=1}^{\ell} \frac{e(X_i, Y_j)^2}{|X_i||Y_j|} \\
 &\stackrel{(1.3)}{\geq} \frac{1}{n^2} \frac{\left(\sum_{i=1}^k \sum_{j=1}^{\ell} e(X_i, Y_j) \right)^2}{\sum_{i=1}^k \sum_{j=1}^{\ell} |X_i||Y_j|} \\
 &= \frac{1}{n^2} \frac{e(X, Y)^2}{\left(\sum_{i=1}^k |X_i| \right) \left(\sum_{j=1}^{\ell} |Y_j| \right)} \\
 &= \frac{e(X, Y)^2}{n^2 |X||Y|} \\
 &= q(X, Y).
 \end{aligned} \tag{3.3}$$

817 Sea ahora la partición $\mathcal{P} = \{V_1, \dots, V_k\}$ de $V(G)$ y $\mathcal{P}' = \{\mathcal{P}'_{V_1}, \dots, \mathcal{P}'_{V_k}\}$ un refina-
 818 miento de \mathcal{P} . Entonces, se utiliza el resultado probado previamente para completar
 819 el resultado:

$$q(\mathcal{P}) = \sum_{i=1}^k \sum_{j=1}^k q(V_i, V_j) \stackrel{(3.3)}{\leq} \sum_{i=1}^k \sum_{j=1}^k q(\mathcal{P}'_{V_i}, \mathcal{P}'_{V_j}).$$

820

□

821 Ahora, veremos que refinar un par (X, Y) no ε -regular de un grafo G , mediante
 822 los subconjuntos que evidencian su irregularidad, provoca un aumento estricto en la
 823 energía.

824 **Lema 36.** Sea $\varepsilon > 0$, G un grafo de n vértices y $X, Y \subset V(G)$. Si (X, Y) no es un
 825 par ε -regular, existen particiones $\mathcal{P}_X = \{X_1, X_2\}$ de X y $\mathcal{P}_Y = \{Y_1, Y_2\}$ de Y tales
 826 que

$$q(\mathcal{P}_X, \mathcal{P}_Y) > q(X, Y) + \varepsilon^4 \frac{|X||Y|}{n^2}.$$

827 *Demostración.* Dado $\varepsilon > 0$, considere el grafo G sobre n vértices y $X, Y \subset V(G)$
 828 subconjuntos tales que el par (X, Y) no es ε -regular. Entonces, existen los subcon-
 829 juntos $X_1 \subset X$ e $Y_1 \subset Y$ que evidencian la irregularidad del par (X, Y) , y son tales
 830 que

$$|X_1| \geq \varepsilon|X| \text{ y } |Y_1| \geq \varepsilon|Y|. \quad (3.4)$$

831 Se define adicionalmente los conjuntos $X_2 := X \setminus X_1$, $Y_2 := Y \setminus Y_1$, y $\eta :=$
 832 $d(X_1, Y_1) - d(X, Y)$, el cual por definición de par ε -regular, satisface

$$|\eta| > \varepsilon. \quad (3.5)$$

833 Por un lado, observe la siguiente descomposición,

$$\begin{aligned} e(X, Y) &= e(X_1, Y) + e(X_2, Y) \\ &= e(X_1, Y_1) + e(X_1, Y_2) + e(X_2, Y_1) + e(X_2, Y_2). \end{aligned}$$

834 De esta manera,

$$\sum_{i+j>2} e(X_i, Y_j) = e(X, Y) - e(X_1, Y_1). \quad (3.6)$$

835 Por otro lado, se tiene que,

$$\begin{aligned} |X||Y| &= (|X_1| + |X_2|)(|Y_1| + |Y_2|) \\ &= |X_1||Y_1| + |X_1||Y_2| + |X_2||Y_1| + |X_2||Y_2|. \end{aligned}$$

836 Así,

$$\sum_{i+j>2} |X_i||Y_j| = |X||Y| - |X_1||Y_1|. \quad (3.7)$$

837 Ahora, definiendo las particiones $\mathcal{P}_X = \{X_1, X_2\}$ de X y $\mathcal{P}_Y = \{Y_1, Y_2\}$ de Y ,
 838 desarrollamos,

$$\begin{aligned}
 q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^2 \sum_{j=1}^2 q(X_i, Y_j) \\
 &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{e(X_i, Y_j)^2}{n^2 |X_i| |Y_j|} \\
 &= \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \sum_{i+j>2} \frac{e(X_i, Y_j)^2}{|X_i| |Y_j|} \right) \\
 &\stackrel{(1.3)}{\geq} \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{\left(\sum_{i+j>2} e(X_i, Y_j) \right)^2}{\sum_{i+j>2} |X_i| |Y_j|} \right) \\
 &\stackrel{(3.6)}{=} \stackrel{(3.7)}{=} \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X| |Y| - |X_1| |Y_1|} \right). \quad (3.8)
 \end{aligned}$$

839 Luego, por definición, note que

$$e(X_1, Y_1) = \frac{|X_1| |Y_1| e(X, Y)}{|X| |Y|} + \eta |X_1| |Y_1|. \quad (3.9)$$

840 Con esto, se continúa el cálculo desde la desigualdad (3.8) como sigue:

$$\begin{aligned}
 n^2 q(\mathcal{P}_X, \mathcal{P}_Y) &\geq \frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X| |Y| - |X_1| |Y_1|} \\
 &\stackrel{(3.9)}{=} \frac{1}{|X_1| |Y_1|} \left(\frac{|X_1| |Y_1| e(X, Y)}{|X| |Y|} + \eta |X_1| |Y_1| \right)^2 \\
 &\quad + \frac{1}{|X| |Y| - |X_1| |Y_1|} \left(\frac{|X| |Y| - |X_1| |Y_1|}{|X| |Y|} e(X, Y) - \eta |X_1| |Y_1| \right)^2 \\
 &= \frac{|X_1| |Y_1|}{|X|^2 |Y|^2} e(X, Y)^2 + 2 \frac{|X_1| |Y_1|}{|X| |Y|} \eta e(X, Y) + \eta^2 |X_1| |Y_1| \\
 &\quad + \frac{|X| |Y| - |X_1| |Y_1|}{|X|^2 |Y|^2} e(X, Y)^2 - 2 \frac{|X_1| |Y_1|}{|X| |Y|} \eta e(X, Y) + \frac{\eta^2 |X_1|^2 |Y_1|^2}{|X| |Y| - |X_1| |Y_1|} \\
 &= \frac{e(X, Y)^2}{|X| |Y|} + \eta^2 |X_1| |Y_1| \left(1 + \frac{|X_1| |Y_1|}{|X| |Y| - |X_1| |Y_1|} \right) \\
 &\geq \frac{e(X, Y)^2}{|X| |Y|} + \eta^2 |X_1| |Y_1|. \quad (3.10)
 \end{aligned}$$

Finalmente, utilizando las cotas (3.4) y (3.5), podemos concluir desde la desigualdad (3.10),

$$\begin{aligned} q(\mathcal{P}_X, \mathcal{P}_Y) &= \frac{e(X, Y)^2}{n^2 |X| |Y|} + \eta^2 \frac{|X_1| |Y_1|}{n^2} \\ &= q(X, Y) + \eta^2 \frac{|X_1| |Y_1|}{n^2} \\ &> q(X, Y) + \varepsilon^4 \frac{|X| |Y|}{n^2}. \end{aligned}$$

□

Vimos que particionar cualquier par de conjuntos no ε -regular por medio de sus subconjuntos que evidencian la irregularidad produce un aumento en la energía. Entonces, haciendo alusión al paso 2(b) del algoritmo de la técnica *argumento de incremento de energía*, se mostrará que refinar simultáneamente todos los pares de conjuntos no ε -regular de un grafo produce un aumento estricto de al menos ε^5 en la energía.

Lema 37. Sea $\varepsilon > 0$, un grafo G y una partición $\mathcal{P} = \{V_1, \dots, V_k\}$ no ε -regular de $V(G)$. Entonces existe un refinamiento \mathcal{Q} de \mathcal{P} , en el que cada V_i se particiona en a lo más 2^k partes y es tal que

$$q(\mathcal{Q}) > q(\mathcal{P}) + \varepsilon^5.$$

Demostración. Sea $\varepsilon > 0$ y $\mathcal{P} = \{V_1, \dots, V_k\}$ una partición no ε -regular del conjunto de n vértices de un grafo G . Sabemos que para todos los $(i, j) \in [k]^2$ tales que el par (V_i, V_j) no es ε -regular, existen los subconjuntos $A^{ij} \subset V_i$ y $A^{ji} \subset V_j$ testigos de su irregularidad. Observe que en cada V_i se podrían encontrar a lo más k conjuntos no vacíos A^{ij} que evidencian la irregularidad de los pares (V_i, V_j) no ε -regular. Considere ahora la partición $\mathcal{Q} = \{Q_1, \dots, Q_k\}$ que refina a \mathcal{P} , en la que cada Q_i es una partición resultante de dividir el conjunto V_i según la intersección de todos los subconjuntos no vacíos A^{ij} que atestiguan la irregularidad de los pares (V_i, V_j) no ε -regular (ver Figura 3.4). En consecuencia, $|Q_i| \leq 2^k$.

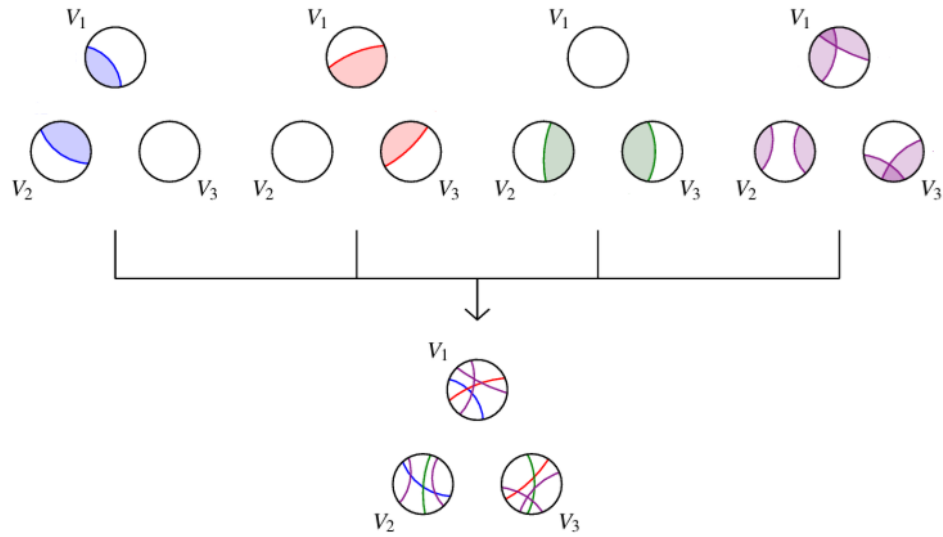


Figura 3.4: Ejemplo del refinamiento simultaneo por todos los subconjuntos que evidencian la irregularidad usando tres conjuntos de vértices ([28], página 59).

862 Para simplicidad en la notación, se define $\Theta := \{(i, j) \in [k]^2 : (V_i, V_j) \text{ es } \varepsilon\text{-regular}\}$.

863 Luego, como la partición \mathcal{P} no es ε -regular, se cumple la desigualdad

$$\sum_{(i,j) \notin \Theta} \frac{|V_i||V_j|}{n^2} > \varepsilon. \quad (3.11)$$

864 Así, junto a los lemas probados previamente, se da prueba al resultado de la

865 siguiente manera:

$$\begin{aligned}
q(\mathcal{Q}) &= \sum_{(i,j) \in [k]^2} q(\mathcal{Q}_i, \mathcal{Q}_j) \\
&= \sum_{(i,j) \in \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) + \sum_{(i,j) \notin \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) \\
&\stackrel{\text{Lema 35}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} q(\{A^{ij}, V_i \setminus A^{ij}\}, \{A^{ji}, V_j \setminus A^{ji}\}) \\
&\stackrel{\text{Lema 36}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \left(q(V_i, V_j) + \varepsilon^4 \frac{|V_i||V_j|}{n^2} \right) \\
&= \sum_{(i,j) \in [k]^2} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \varepsilon^4 \frac{|V_i||V_j|}{n^2} \\
&\stackrel{(3.11)}{\geq} q(\mathcal{P}) + \varepsilon^5.
\end{aligned}$$

866 * Cambiar por $>$ en la última línea y donde dice lema 33, cuando lo cambio se
867 me descuadra :c * □

868 Este último Lema culmina lo que se necesita para dar prueba formal del lema de
869 regularidad de Szemerédi mediante el argumento de incremento de energía.

870 *Demostración del Teorema 32.* Dado $\varepsilon > 0$ y un grafo G , elegimos inicialmente la
871 partición trivial del conjunto de vértices $\mathcal{P} = \{V(G)\}$. Ahora, iterativamente (actua-
872 lizando \mathcal{P}), aplicaremos el Lema 37 cada vez que la partición actual no sea ε -regular.
873 Observe que por cada aplicación del Lema 37 se consigue un aumento de al menos
874 ε^5 en la energía, y como la energía de toda partición está acotada superiormente
875 por 1, el proceso iterativo terminará luego de a lo más ε^{-5} pasos. El resultado será
876 necesariamente una partición ε -regular debido a la cota de la energía.

877 Para una partición no ε -regular con k elementos, el Lema 37 encuentra un refina-
878 miento de a lo más $k2^k$ partes. Dicho refinamiento será producido en cada iteración
879 del algoritmo de *argumento de incremento de energía*, y la cantidad de partes pro-
880 ducidas las acotaremos crudamente en cada paso por $k2^k < 2^{2^k}$. Comenzando con la
881 partición trivial de una parte, ejemplificaremos con las tres primeras iteraciones del

882 algoritmo para mostrar la cantidad de partes producidas en cada paso tras aplicar
 883 el Lema 37.

$$\begin{aligned}
 &1^{\text{ra}} \text{ Iteración: } 1 \rightarrow 2 < 2^2 \quad \text{partes.} \\
 &2^{\text{da}} \text{ Iteración: } 2^2 \rightarrow (2^2) 2^{(2^2)} < 2^{2^{2^2}} \quad \text{partes.} \\
 &3^{\text{ra}} \text{ Iteración: } 2^{2^{2^2}} \rightarrow (2^{2^{2^2}}) 2^{(2^{2^{2^2}})} < 2^{2^{2^{2^{2^2}}}} \quad \text{partes.}
 \end{aligned}$$

884 Finalmente, como el algoritmo debe terminar luego de a lo más ε^{-5} iteraciones,
 885 la cantidad de partes al final del proceso será

$$M(\varepsilon) \leq 2^{2^{\dots^2}} \left. \vphantom{2^{2^{\dots^2}}} \right\} \text{Altura } 2\varepsilon^{-5}.$$

886

□

887 Desde ahora en adelante, vamos a definir y considerar una *torre de altura* k de la
 888 siguiente manera:

$$\text{torre}(k) := 2^{2^{\dots^2}} \left. \vphantom{2^{2^{\dots^2}}} \right\} \text{Altura } k.$$

889 Durante la demostración del Teorema 32 se utilizó una cota que podría parecer
 890 exagerada para encontrar la cantidad de partes que devuelve el algoritmo implemen-
 891 tado, por sobre todo, considerando lo rápido que crece a medida que ε se hace más
 892 pequeño. Sorprendentemente, en 1997, T. Gowers [12] prueba que tal límite inferior
 893 de partes es necesario. Más precisamente, mostró que es posible encontrar una cons-
 894 tante $c > 0$ tal que para todo suficientemente pequeño $\varepsilon > 0$, existe un grafo sin
 895 partición ε -regular siempre que posea una cantidad menor que $\text{torre}(\lceil \varepsilon^{-c} \rceil)$ partes
 896 (ver G. Moshkovitz y A. Shapira [19] para una demostración corta).

897 Finalmente, se expone la forma de probar el Teorema 33. La idea de la demos-
 898 tración consiste en modificar el algoritmo de la técnica de argumento de incremento
 899 de energía, de manera que en cada iteración del refinamiento se logre obtener una

900 equipartición. Este procedimiento conservará el incremento de energía en cada paso
 901 y terminará con una equipartición del conjunto de vértices de un grafo cualquiera.
 902 Entonces, para todo grafo G , la modificación del algoritmo es la siguiente:

- 903 1. Comenzar con una equipartición inicial arbitraria \mathcal{P} de $V(G)$ con m_0 partes.
- 904 2. Mientras la partición actual \mathcal{P} no es ε -regular:
 - 905 (a) Para cada par (V_i, V_j) no ε -regular, encontrar los subconjuntos $A^{ij} \subset V_i$
 906 y $A^{ji} \subset V_j$ que evidencian la irregularidad de los pares.
 - 907 (b) Refinar \mathcal{P} usando simultáneamente los conjuntos A^{ij} y A^{ji} para obtener
 908 la partición \mathcal{Q} , cual divide cada parte de \mathcal{P} en a lo más $2^{|\mathcal{P}|}$ partes.
 - 909 (c) Modificar la partición \mathcal{Q} refinando, si es posible, cada uno de sus elementos
 910 para formar partes iguales de tamaño $|V(G)|/m$, dada alguna elección
 911 apropiada del entero $m = m(|\mathcal{Q}|, \varepsilon)$. Luego, los elementos de \mathcal{Q} que no
 912 fueron refinados previamente a causa de su bajo tamaño y los conjuntos
 913 de vértices residuales del refinamiento anterior, deben ser combinados y
 914 posteriormente dividir el resultado en partes iguales de tamaño $|V(G)|/m$.
 - 915 (d) Actualizar \mathcal{P} con la modificación de \mathcal{Q} .

916 El algoritmo anterior obtiene una equipartición del conjunto de vértices del grafo
 917 G . En lo que respecta a la energía del proceso, el paso 2(b) conserva un aumento de al
 918 menos ε^5 en cada iteración. El paso 2(c) podría ocasionar una baja en la energía, sin
 919 embargo, no debería ser significativa con una elección de m suficientemente grande.
 920 En resumidas cuentas, el proceso anterior aumenta la energía en cada iteración en al
 921 menos $\varepsilon^5/2$, logrando terminar luego de a lo más $2\varepsilon^{-5}$ pasos con una equipartición
 922 de a lo más $\text{torre}(\lceil \varepsilon^{-5} \rceil)$ partes.

3.2. Demostración espectral

En 2012, T. Tao [26] publica en su blog una prueba del lema de regularidad de Szemerédi usando la descomposición espectral de la matriz de adyacencia. La idea original de la demostración proviene de los autores A. Frieze y R. Kannan [9], a quienes Tao les da el crédito en su publicación. Más adelante, en 2013, S. Cioba y R. Martin [5] escribieron la demostración con más detalles. La prueba que se expone en esta sección está basada esencialmente en la publicación de Cioba y Martin.

Demostración espectral del Teorema 3.2. Dado $\varepsilon > 0$, consideramos la función $F : \mathbb{N} \rightarrow \mathbb{N}$ definida por

$$F(\ell) = \left\lceil \frac{1}{\varepsilon^6} \left(\frac{2\ell^2}{\varepsilon^2} \right)^{4\ell} \right\rceil.$$

Denotaremos por $F^{(i)}$ a la i -ésima composición de F con ella misma, y escogemos $n_0 \in \mathbb{N}$ suficientemente grande.

Sea $G = ([n], E)$ un grafo con $n \geq n_0$ vértices, y A su matriz de adyacencia. Ordenamos los valores propios $|\lambda_1| \geq \dots \geq |\lambda_n|$ de A de manera decreciente y consideramos $\mathbf{u}_1, \dots, \mathbf{u}_n$ los vectores propios correspondientes, que forman una base ortonormal de \mathbb{R}^n .

Por la Proposición 4 y el Corolario 9, se satisface

$$\text{Tr}(T) = \sum_{i=1}^n \lambda_i^2 = 2e_G \leq n^2. \quad (3.12)$$

De lo anterior, al notar que $i\lambda_i^2 \leq \sum_{j=1}^i \lambda_j^2 \leq n^2$, se encuentra la cota

$$\lambda_i \leq \frac{n}{\sqrt{i}}, \quad \forall i \in [n]. \quad (3.13)$$

Consideramos también los intervalos $I_1, \dots, I_{\lceil 1/\varepsilon^3 \rceil} \subset [n]$ definidos por

■ $I_1 = \{1, 2, \dots, F^{(1)}(1) - 1\}$ y

942 ■ $I_k = \{F^{(k-1)}(1), \dots, F^k(1) - 1\}$ para todo $k = 2, \dots, I_{\lceil 1/\varepsilon^3 \rceil}$.

943 Con esta construcción, debe existir un natural $1 \leq L \leq \lceil 1/\varepsilon^3 \rceil$ que cumple con

$$\sum_{L \leq j < F(L)} \lambda_j^2 \leq \varepsilon^3 n^2, \quad (3.14)$$

944 porque de lo contrario, se obtiene

$$\sum_{i=1}^n \lambda_i^2 \geq \sum_{k=1}^{\lceil 1/\varepsilon^3 \rceil} \sum_{i \in I_k} \lambda_i^2 > \lceil 1/\varepsilon^3 \rceil \cdot \varepsilon^3 n^2 > n^2,$$

945 que contradice la desigualdad (3.12).

946 Ahora, usando L , separamos la matriz A en tres matrices simétricas:

$$A = S + F + Q,$$

947 donde la matriz S se interpretará como la componente *estructural*,

$$S = \sum_{i < L} \lambda_i \mathbf{u}_i \mathbf{u}_i^T,$$

948 la matriz F como la componente de *error*,

$$F = \sum_{L \leq i < F(L)} \lambda_i \mathbf{u}_i \mathbf{u}_i^T,$$

949 y la matriz Q como la componente *cuasi-aleatoria*,

$$Q = \sum_{i \geq F(L)} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

950 Usaremos los vectores propios $\mathbf{u}_1, \dots, \mathbf{u}_{L-1}$ de S para definir una partición de
 951 $V(G)$ como mostraremos a continuación. Consideramos el intervalo de \mathbb{R} de longitud

952 $2\sqrt{(L/\varepsilon)} \cdot n^{-1/2}$ centrado en el origen y lo particionamos en $t = 2(L/\varepsilon)^2$ subinter-
 953 valos J_1, \dots, J_t de longitud $(\varepsilon/L)^{3/2}n^{-1/2}$ cada uno. Luego, clasificamos los vértices
 954 $v \in V(G)$ según su valor $\mathbf{u}_i(v)$ de la siguiente manera:

$$V_i^j = \{v \in V(G) : \mathbf{u}_i(v) \in J_j\}, \quad 1 \leq j \leq t.$$

955 Con esto, tomamos el refinamiento de todos estos conjuntos $\{V_i^j \neq \emptyset : i \in$
 956 $[L-1], j \in [t]\}$ para obtener los conjuntos V_0, V_1, \dots, V_M , en donde $M \leq (\frac{2L^2}{\varepsilon^2})^L$.
 957 El resultado anterior considera un conjunto excepcional de vértices V_0 que está de-
 958 finido como sigue:

$$V_0 = \left\{ v \in V(G) : |\mathbf{u}_i(v)| > \sqrt{\frac{L}{\varepsilon}} n^{-1/2} \text{ para algún } i \in [L-1] \right\}.$$

959 Mostraremos que el conjunto excepcional V_0 es suficientemente pequeño. En efecto,
 960 observando que

$$L-1 = \sum_{i=1}^{L-1} \|\mathbf{u}_i\|^2 \geq \sum_{i=1}^{L-1} \sum_{v \in V(G)} \mathbf{u}_i(v)^2 \geq |V_0| \left(\frac{L}{\varepsilon n} \right),$$

961 se determina que $|V_0| < \varepsilon n$.

962 Probaremos que la partición construida del conjunto de vértices del grafo $\mathcal{P} =$
 963 $\{V_0, V_1, \dots, V_M\}$ es ε -regular. Comenzamos identificando los pares excepcionales. Para
 964 esto, sea $F = (f_{xy})$ y defina

$$\Sigma_F = \left\{ (i, j) : \sum_{(x,y) \in V_i \times V_j} f_{xy}^2 > \varepsilon |V_i| |V_j| \right\}$$

965 Entonces, por la definición de F y el Corolario 5, tenemos que

$$\begin{aligned}
\varepsilon^3 n^2 &\geq \sum_{L \leq i < F(L)} \lambda_i^2 = \text{Tr}(F^2) = \sum_{(x,y) \in V(G)^2} f_{xy}^2 \\
&\geq \sum_{(i,j) \in \Sigma_F} \sum_{(x,y) \in V_i \times V_j} f_{xy}^2 > \varepsilon \sum_{(i,j) \in \Sigma_F} |V_i| |V_j|.
\end{aligned}$$

966 y por consecuencia

$$\varepsilon^2 n^2 \geq \sum_{(i,j) \in \Sigma_F} |V_i| |V_j|. \quad (3.15)$$

967 Además, sea

$$\Sigma_Q = \left\{ (i, j) : \min\{|V_i|, |V_j|\} < \frac{\varepsilon}{M} n \right\} \cup \left\{ (i, j) : i = 0 \text{ o } j = 0 \right\},$$

968 y observe que

$$\sum_{(i,j) \in \Sigma_Q} |V_i| |V_j| \leq 2M \cdot \frac{\varepsilon}{M} n + 2|V_0|n < 4\varepsilon n.$$

969 Ahora, sea $(i, j) \notin \Sigma_F \cup \Sigma_Q$, y $d_{ij} = d(V_i, V_j)$ la densidad del par (V_i, V_j) . Entonces,
 970 dado los subconjuntos $X \subset V_i$ e $Y \subset V_j$, note la siguiente descomposición:

$$\begin{aligned}
\left| e(X, Y) - d_{ij} |X| |Y| \right| &= \left| \mathbf{v}_X^T A \mathbf{v}_Y - d_{ij} |X| |Y| \right| \\
&\leq \left| \mathbf{v}_X^T S \mathbf{v}_Y - d_{ij} |X| |Y| \right| + \left| \mathbf{v}_X^T F \mathbf{v}_Y \right| + \left| \mathbf{v}_X^T Q \mathbf{v}_Y \right|. \quad (3.16)
\end{aligned}$$

971 En este punto, el objetivo es encontrar cotas para cada uno de los sumandos
 972 anteriores.

973 Para comenzar, por la definición de Σ_F y la desigualdad de Cauchy-Schwarz, se
 974 obtiene la primera de las cotas de la siguiente manera:

$$\begin{aligned}
|\mathbf{v}_X^T F \mathbf{v}_Y|^2 &= \left| \sum_{(x,y) \in X \times Y} f_{xy} \right|^2 \stackrel{\text{DCS}}{\leq} \left(\sum_{(x,y) \in X \times Y} f_{xy}^2 \right) |X||Y| \\
&\leq \varepsilon^2 |V_i| |V_j| |X| |Y| \leq \varepsilon^2 |V_i|^2 |V_j|^2.
\end{aligned} \tag{3.17}$$

975 Para la próxima cota, debemos observar por la construcción de Q y el Teorema
976 10 que

$$|\mathbf{v}_X^T Q \mathbf{v}_Y| \stackrel{\text{DCS}}{\leq} \|\mathbf{v}_X\| \|Q \mathbf{v}_Y\| \leq \|\mathbf{v}_X\| \|\mathbf{v}_Y\| \frac{n}{\sqrt{F(L)}} = \sqrt{|X||Y|} \frac{n}{\sqrt{F(L)}} \leq \frac{n^2}{\sqrt{F(L)}}.$$

977 Además, como $M \leq (\frac{2L^2}{\varepsilon^2})^L$, concluimos de la elección de $F(\cdot)$ que $F(L) \geq$
978 $\frac{1}{\varepsilon^6} \left(\frac{2L^2}{\varepsilon^2} \right)^{4L} \geq \frac{1}{\varepsilon^6} M^4$. Y así, cuando $(i, j) \neq \Sigma_Q$, se tiene

$$|\mathbf{v}_X^T Q \mathbf{v}_Y| \leq \frac{n^2}{\sqrt{F(L)}} \leq \frac{M^2 |V_i| |V_j|}{\varepsilon^2 \sqrt{F(L)}} \leq \varepsilon |V_i| |V_j|. \tag{3.18}$$

979 Por último, para la tercera cota, analizamos $S = (s_{xy})$. Sean s_{ab} y s_{cd} los valores
980 mínimo y máximo de todos los s_{xy} sobre $(u, v) \in V_i \times V_j$. Entonces,

$$\begin{aligned}
s_{cd} - s_{ab} &= \sum_{i < L} \lambda_i \mathbf{u}_i(c) \mathbf{u}_i(d) - \lambda_i \mathbf{u}_i(a) \mathbf{u}_i(b) \\
&\leq \sum_{i < L} |\lambda_i| \left| \mathbf{u}_i(c) \mathbf{u}_i(d) - \mathbf{u}_i(a) \mathbf{u}_i(d) + \mathbf{u}_i(a) \mathbf{u}_i(d) - \mathbf{u}_i(a) \mathbf{u}_i(b) \right| \\
&\leq n \sum_{i < L} \left| \mathbf{u}_i(d) (\mathbf{u}_i(c) - \mathbf{u}_i(a)) + \mathbf{u}_i(a) (\mathbf{u}_i(d) - \mathbf{u}_i(b)) \right| \\
&\leq n \sum_{i < L} |\mathbf{u}_i(b)| \left| \mathbf{u}_i(a) - \mathbf{u}_i(c) \right| + |\mathbf{u}_i(c)| \left| \mathbf{u}_i(b) - \mathbf{u}_i(d) \right| \\
&\leq Ln \cdot 4 \cdot \sqrt{\frac{L}{\varepsilon n}} \cdot \frac{\varepsilon}{L} \sqrt{\frac{\varepsilon}{Ln}} \\
&= 4\varepsilon.
\end{aligned}$$

981 Ahora bien, como d_{ij} es el promedio de S sobre $V_i \times V_j$, tenemos que $s_{ab} \leq d_{ij} \leq s_{cd}$,
982 y por ende, $|s_{xy} - d_{ij}| \leq s_{cd} - s_{ab}$ para cada $(u, v) \in V_i \times V_j$. Como resultado,

$$\left| \mathbf{v}_X^T S \mathbf{v}_Y - d_{ij} |V_i| |V_j| \right| \leq \sum_{(x,y) \in X \times Y} |s_{xy} - d_{ij}| \leq (s_{cd} - s_{ab}) |X| |Y| \leq 4\epsilon |X| |Y|. \quad (3.19)$$

Utilizando las desigualdades (3.17), (3.18) y (3.19) en la expresión enunciada en (3.16) se concluye la demostración del teorema. \square

3.3. Aplicaciones

Usualmente las aplicaciones del lema de regularidad de Szemerédi son abordadas con el *método de regularidad* que se describe en los siguientes pasos:

1. Obtener una **partición** del conjunto de vértices de un grafo con el lema de regularidad.
2. **Limpiar** las aristas del que tengan un “mal comportamiento” según el problema. Generalmente, se eliminan las aristas entre los pares de partes que presentan:
 - (i) Irregularidad.
 - (ii) Baja densidad.
 - (iii) Al menos una de las partes demasiado pequeña.
3. **Contar** un determinado patrón en el grafo limpio.

Para el último paso se utilizará un resultado análogo a la propiedad $\text{COUNT}_p(\epsilon)$ del Teorema 17, pero con el concepto de par ϵ -regular. Las aplicaciones que se estudiarán en esta tesis solo necesitan el caso en que $H = K_3$, el cual es conocido como lema de conteo de triángulos.

1001 **Lema 38.** (*Lema de conteo de triángulos*) Sea $\varepsilon > 0$, $G = (V, E)$ un grafo, y los
 1002 conjuntos no necesariamente disjuntos $X, Y, Z \subset V$ tales que los pares $(X, Y), (Y, Z)$
 1003 y (X, Z) son ε -regular. Entonces,

$$\left| \{(x, y, z) \in X \times Y \times Z : xy, xz, yz \in E\} \right| = d(X, Y)d(X, Z)d(Y, Z)|X||Y||Z| \\ \pm 3\varepsilon|X||Y||Z|.$$

1004 *Demostración.* Se realizará un proceso inductivo similar al visto en la demostración
 1005 de la Proposición 20 sobre la cantidad de aristas del grafo $K_3 = ([3], \{12, 23, 13\})$.
 1006 Cuando el grafo no posee aristas, entonces

$$\left| \{(x, y, z) \in X \times Y \times Z : xy, xz, yz \notin E\} \right| = |X||Y||Z|.$$

1007 También, como vimos en (3.2), recordamos que la condición de un par ε -regular
 1008 es equivalente a la propiedad $\text{DISC}_p(\varepsilon)$ en grafos bipartitos para algún $p \in (0, 1)$.
 1009 Entonces, cuando el grafo presenta solo una arista,

$$\left| \{(x, y, z) \in X \times Y \times Z : xy \in E\} \right| = (d(X, Y)|X||Y| \pm \varepsilon|X||Y|)|Z|.$$

1010 Ahora, se plantea la hipótesis inductiva de la siguiente manera:

$$\left| \{(x, y, z) \in X, Y, Z : xy, yz \in E\} \right| = d(X, Y)d(Y, Z)|X||Y||Z| \pm 2\varepsilon|X||Y||Z|.$$

1011 Defina $e^- = \varphi(1)\varphi(3)$, y T^- como el grafo correspondido a una copia etiquetada
 1012 del grafo $([3], \{12, 23\})$ en G bajo la aplicación inyectiva $\varphi : [3] \rightarrow V(T^-) \subset V$. Con
 1013 esto, se desarrolla inductivamente como sigue:

$$\begin{aligned}
\left| \{(x, y, z) \in X \times Y \times Z : xy, yz, xz \in E\} \right| &= \sum_{T^-} [\mathbb{1}_E(e^-) + d(X, Z) - d(X, Z)] \\
&= d(X, Y)d(Y, Z)d(X, Z)|X||Y||Z| + \sum_{T^-} (\mathbb{1}_E(e^-) - d(X, Z)) \pm 2\varepsilon|X||Y||Z|.
\end{aligned} \tag{3.20}$$

1014 En este punto, nos falta probar que el segundo sumando de la igualdad (3.20) se
1015 corresponde con un factor de error, para esto, sea T^* una copia del grafo singleton
1016 $\{2\}$ en G , y considere los siguientes conjuntos:

$$\begin{aligned}
1017 \quad A_1^{T^*} &= \{x \in X : T^* \text{ con } x \text{ forma una copia de } (\{1, 2\}, \{12\}) \text{ en } G\}. \\
A_3^{T^*} &= \{z \in Z : T^* \text{ con } z \text{ forma una copia de } (\{2, 3\}, \{23\}) \text{ en } G\}.
\end{aligned}$$

1018 De esta manera, dada la equivalencia de la condición del par (X, Z) ε -regular con
1019 versión bipartita de la propiedad $\text{DISC}_{d(X, Z)}(\varepsilon)$ vista en (3.2), se consigue la siguiente
1020 desigualdad:

$$\begin{aligned}
\left| \sum_{T^-} (\mathbb{1}_E(e^-)) - d(X, Z) \right| &\leq \sum_{T^*} \left| \sum_{f \in A_1^{T^*} \times A_3^{T^*}} (\mathbb{1}_E(f) - d(X, Z)) \right| \\
&= \sum_{T^*} |e(A_1^{T^*}, A_3^{T^*}) - d(X, Z)| A_1^{T^*} || A_3^{T^*} || \\
&\leq \sum_{T^*} \varepsilon |X| |Z| \\
&\leq \varepsilon |X| |Y| |Z|.
\end{aligned}$$

1021 Finalmente, aplicando la última desigualdad en la ecuación (3.20) se prueba lo
1022 prometido. \square

1023 En la demostración anterior solo fue necesario utilizar que los pares (X, Y) y
1024 (X, Z) son ε -regular, por lo que es interesante destacar que uno de los pares de

1025 conjuntos de vértices podría no ser necesariamente un par ε -regular para que el lema
1026 de conteo de triángulos funcione correctamente.

1027 Bajo el mismo planteamiento de la inducción vista en la demostración del Lema
1028 38 (y Proposición 20), es posible generalizar el resultado para contar apropiadamente
1029 cualquier grafo H . Se enuncia sin demostración.

1030 **Lema 39.** (*Lema de conteo de grafos*) Sea $\varepsilon > 0$, H un grafo sobre k vértices, y G
1031 un grafo de n vértices con los subconjuntos disjuntos $V_1, \dots, V_k \subset V(G)$ tales que los
1032 pares (V_i, V_j) son ε -regular siempre que $ij \in E(H)$. Entonces, la cantidad de tuplas
1033 $(v_1, \dots, v_k) \in V_1 \times \dots \times V_k$ tales que $v_i v_j \in E(G)$ cada vez que $ij \in E(H)$ es

$$\left(\prod_{ij \in E(H)} d(V_i, V_j) \right) \left(\prod_{\ell=1}^k |V_\ell| \right) \pm e_H \cdot \varepsilon \prod_{\ell=1}^k |V_\ell|.$$

1034 En las siguientes subsecciones se discutirán dos aplicaciones del método de regu-
1035 laridad para entregar dos demostraciones alternativas al teorema de Roth.

1036 3.3.1. Eliminación de triángulos

1037 El lema de eliminación de triángulos fue probado por los autores I. Ruzsa y E. Sze-
1038 merédi [22] en 1976, y es una de las primeras aplicaciones del método de regularidad.
1039 La intuición del lema dice que todo grafo con *pocos* triángulos se puede convertir en
1040 un grafo libre de triángulos eliminando *pocas* aristas. Formalmente,

1041 **Teorema 40.** (*Lema de eliminación de triángulos*) Para todo $\varepsilon > 0$, existe $\delta > 0$ y
1042 $n_0 \in \mathbb{N}$ tal que todo grafo sobre $n \geq n_0$ vértices con a lo más δn^3 triángulos se puede
1043 hacer libre de triángulos eliminando a lo más εn^2 aristas.

1044 *Demostración.* Dado $\varepsilon > 0$, elija $\varepsilon_r = \frac{1}{4} \left(\frac{\varepsilon}{3}\right)^3$ y utilice el Teorema 32 con tal elec-
1045 ción para obtener la constante $M = M(\varepsilon_r)$. Considere además $\delta = \frac{1}{2} \frac{\varepsilon_r^4}{M^3}$ y $n_0 \in \mathbb{N}$
1046 suficientemente grande, de manera tal que el grafo $G = (V, E)$ con $n \geq n_0$ vértices
1047 posee a lo más δn^3 triángulos. Luego, nuevamente por el Teorema 32, se asegura la
1048 existencia de una partición ε_r -regular $\mathcal{P} = \{V_1, \dots, V_k\}$, con $k \leq M$.

1049 Para limpiar el grafo, para cada $(i, j) \in [k]^2$, se eliminan todas las aristas entre V_i
 1050 y V_j cuando

1051 (a) (V_i, V_j) no es un par ε_r -regular,

1052 (b) $d(V_i, V_j) < (4\varepsilon_r)^{1/3}$, o

1053 (c) $\min\{|V_i||V_j|\} < \frac{n}{k}\varepsilon_r$.

1054 De esta manera, como la partición es ε_r -regular, las aristas removidas por la
 1055 condición (a) son a lo más

$$\sum_{\substack{(i,j) \in [k]^2 \\ (V_i, V_j) \text{ no } \varepsilon_r\text{-regular}}} |V_i||V_j| \leq \varepsilon_r n^2.$$

1056 Las aristas eliminadas en los conjuntos de baja densidad por la condición (b) son
 1057 a lo más

$$\sum_{\substack{(i,j) \in [k]^2 \\ d(V_i, V_j) < (4\varepsilon_r)^{1/3}}} d(V_i, V_j)|V_i||V_j| < (4\varepsilon_r)^{1/3} \sum_{(i,j) \in [k]^2} |V_i||V_j| = (4\varepsilon_r)^{1/3} n^2.$$

1058 Por último, debido a que cada vértice de G puede ser adyacente con a lo más $\frac{n}{k}\varepsilon_r$
 1059 vértices en a lo más k subconjuntos demasiado pequeños, las aristas removidas por
 1060 (c) son a lo más

$$k \cdot \frac{n}{k}\varepsilon_r \cdot n = \varepsilon_r n^2.$$

1061 En total, en la limpieza, se eliminan a lo más εn^2 aristas.

1062 Ahora, nos falta probar que el grafo limpio $G' = (V, E')$ es libre de triángulos.
 1063 Para esto, con nuestra elección de δ , buscaremos formular la siguiente contradicción:
 1064 si existe un triángulo en el grafo limpio G' , el lema de conteo de triángulos asegura
 1065 que en realidad existen más de δn^3 triángulos. No obstante, como el grafo original

1066 posee a lo más δn^3 triángulos, se podrá concluir que el grafo G' es libre de triángulos
 1067 eliminando a lo más εn^2 aristas.

1068 Dicho esto, estudiamos la cantidad de triángulos en G' . Dada la eliminación de
 1069 aristas según la condición (a), cada par (V_i, V_j) es ε -regular, y por ende se satisface
 1070 la hipótesis del Lema 38. Entonces, gracias a la ausencia de las aristas que cumplían
 1071 con las condiciones (b) y (c),

$$\begin{aligned}
 |\{(x, y, z) \in V_i \times V_j \times V_\ell : xy, yz, xz \in E'\}| &\geq d(V_i, V_j)d(V_i, V_\ell)d(V_j, V_\ell)|V_i||V_j||V_\ell| \\
 &\quad - 3\varepsilon_r|V_i||V_j||V_\ell| \\
 &\geq \varepsilon_r|V_i||V_j||V_\ell| \\
 &\geq \frac{\varepsilon^4 n^3}{k^3} \\
 &> \delta n^3.
 \end{aligned}$$

1072 Así, de la contradicción anterior, se determina que el grafo G es libre de triángulos
 1073 eliminando a lo más εn^2 aristas. □

1074 Otra forma de entender el Teorema 40 es de la siguiente manera: si se necesitan
 1075 eliminar al menos εn^2 para hacer de G libre de triángulos, entonces G contiene al
 1076 menos δn^3 triángulos.

1077 3.3.2. Emparejamiento inducido

1078 Dado un grafo $G = (V, E)$, un subconjunto $R \subset E$ es un **emparejamiento** en
 1079 G si no existe un par de aristas en R que compartan algún vértice. Diremos que R
 1080 es un **emparejamiento inducido** si es un emparejamiento y no existen un par de
 1081 aristas en R que estén conectadas por una arista de G , es decir, no existen aristas
 1082 en G entre cada par de vértices de R .

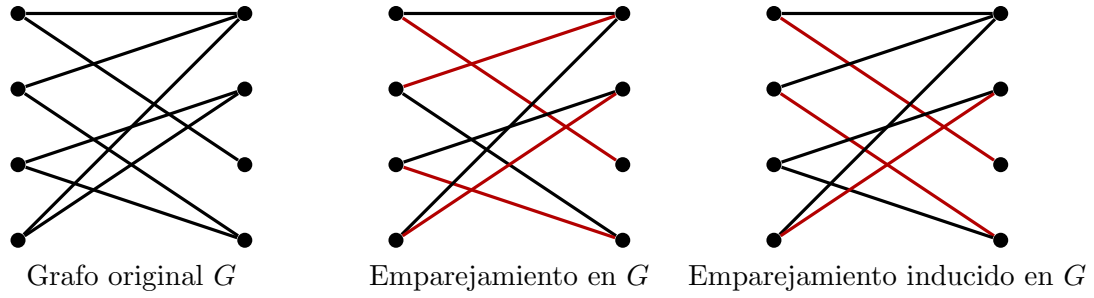


Figura 3.5: Ejemplo de un emparejamiento y un emparejamiento inducido.

La presente aplicación del método de regularidad responde a la pregunta: ¿Cuántas aristas puede tener un grafo que es la unión de emparejamientos inducidos? * No sabría si aquí agregar un comentario referente a que la intuición dice que deben ser “pocas” aristas, pero la prueba no es trivial *

Teorema 41. (Emparejamiento inducido) Para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que todo grafo $G = (V, E)$ de $n \geq n_0$ vértices que está compuesto por la unión de n emparejamientos inducidos, posee a lo más εn^2 aristas.

Demostración. Dado $\varepsilon > 0$, aplique el Teorema 32 con $\varepsilon_r = \frac{\varepsilon}{10}$ para obtener la constante $M(\varepsilon_r)$. Considere $n_0 \in \mathbb{N}$ suficientemente grande, y asuma que el grafo $G = (V, E)$ con $n \geq n_0$ vértices y compuesto por n emparejamientos inducidos satisface $e_G > \varepsilon n$. Nuevamente, por el Teorema 32, se asegura la existencia de la partición $\mathcal{P} = \{V_1, \dots, V_k\}$ con $k \leq M(\varepsilon)$ partes que es ε_r -regular.

Para cada $(i, j) \in [k]^2$ se eliminan todas las aristas entre los conjuntos V_i y V_j cuando estos presenten irregularidad, densidad menor que $2\varepsilon_r$, o al menos uno de los conjuntos es menor que $\frac{n}{k}\varepsilon_r$. En total, el proceso de limpieza remueve a lo más $4\varepsilon_r n^2$ aristas de G para obtener un nuevo grafo G' . En consecuencia,

$$e'_G \geq e_G - 4\varepsilon_r n^2 > \varepsilon n^2 - \frac{4}{10}\varepsilon n^2 > \frac{\varepsilon}{2}n^2.$$

Ahora, observe que debe existir un emparejamiento inducido R en G' con al menos $\frac{\varepsilon}{2}n$ aristas (y al menos εn vértices). De no ser así, todos los emparejamientos tendrán

1101 a lo más $\frac{\varepsilon}{2}n$ aristas, por lo que $e'_G < \frac{\varepsilon}{2}n^2$.

1102 Denotando por $V(R)$ al conjunto de vértices que componen las aristas de R , se
 1103 define $U_i := V_i \cap V(R)$ como el subconjunto de vértices de R que comparte elementos
 1104 con V_i , y $U := \bigcup_{i \in [k]} \{U_i : |U_i| \geq \varepsilon_r |V_i|\}$. Es decir, U es la unión de todos los conjuntos
 1105 $U_i \subset V(R)$ que comparten una fracción suficientemente grande de vértices con V_i .
 1106 Note que podemos obtener el conjunto U removiendo a lo más $\varepsilon_r n = \frac{\varepsilon}{10}n$ vértices de
 1107 $V(R)$, pues

$$\sum_{i \in [k]} |U_i| < \sum_{i \in [k]} \varepsilon_r |V_i| = \frac{\varepsilon}{10}n.$$

1108 De esta manera, recordando que $|V(R)| \geq \varepsilon n$, se determina que $|U| > \varepsilon n - \frac{\varepsilon}{10}n =$
 1109 $\frac{9}{10}\varepsilon n$. Además, como también $|R| \geq \frac{\varepsilon}{2}n$, debe existir al menos un vértice en U que sea
 1110 parte de una arista en R . Luego, dada la limpieza de G , dicha arista debe pertenecer
 1111 a algún par $U_t \times U_\ell$ que satisfacen $|U_k| \geq \varepsilon_r |V_k|$ y $|U_\ell| \geq \varepsilon_r |V_\ell|$, y son tales que su
 1112 correspondiente par (V_t, V_ℓ) es ε_r -regular con densidad $d(V_t, V_\ell) \geq 2\varepsilon_r$. Entonces, por
 1113 regularidad,

$$d(U_t, U_\ell) = d(V_t, V_\ell) \pm \varepsilon_r \geq 2\varepsilon_r - \varepsilon_r = \varepsilon_r. \quad (3.21)$$

1114 Ahora, como R es un emparejamiento inducido, todo par de subconjuntos $A, B \subset$
 1115 $V(M)$ debe satisfacer

$$e(A, B) \leq \min\{|A|, |B|\}.$$

1116 Sin embargo, la desigualdad (3.21) implica que

$$\begin{aligned}
e(U_t, U_\ell) &= d(U_t, U_\ell)|U_t||U_\ell| \\
&\geq |U_t||U_\ell|\varepsilon_r \\
&\geq |U_t||V_\ell|\varepsilon_r^2 \\
&\geq |U_t|\frac{n}{k}\varepsilon_r^3 \\
&> |U_t|.
\end{aligned}$$

La desigualdad anterior nos dice que existe una arista entre U_k y U_ℓ que no pertenece a R , por lo que se contradice la hipótesis de que R es un emparejamiento inducido. \square

Otra mirada del Teorema 41 es la siguiente: si G posee al menos εn^2 aristas, entonces G tiene al menos un emparejamiento no inducido.

3.3.3. Teorema de Roth

Como hemos visto en el comienzo de la sección 3, el teorema de Roth es un caso particular del teorema de Szemerédi, el cual en un principio fue demostrado utilizando análisis de Fourier.

Teorema 42. *(Teorema de Roth) Para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que si el conjunto $S \subset [n]$ posee $|S| \geq \varepsilon n$ elementos, entonces S contiene una progresión aritmética no trivial de largo 3.*

En esta sección se entregarán dos demostraciones del teorema de Roth por medio del lema de regularidad de Szemerédi. Como el enunciado del Teorema 42 alude a la teoría de números, en ambas pruebas, la idea es traducir el problema al lenguaje de la teoría de grafos con la construcción de un grafo apropiado. La primera demostración se fundamenta en el Teorema 40.

Primera demostración del Teorema 42. Sea $\varepsilon > 0$ y el conjunto $S \subset [n]$ con $|S| \geq \varepsilon n$ elementos. Escogemos $n_0 \in \mathbb{N}$ suficientemente grande.

Una manera natural de construir un grafo de n vértices utilizando el conjunto S , es definir una arista entre los vértices i y j si y solamente si $|i - j| \in S$. Sin embargo, con n elementos no podemos describir una progresión aritmética de largo 3 por medio de relaciones entre las aristas. Por esto, agregamos un poco de asimetría a la construcción. Para asegurar que las sumas siempre tengan la posibilidad de estar en S , consideramos el grafo 3-partito $G = (V, E)$ con partición de vértices $V = V_1 \cup V_2 \cup V_3$, en donde $V_1 = [n]$, $V_2 = [2n]$ y $V_3 = [3n]$ son tales que $6n \geq n_0$. Las aristas de G son definidas de la siguiente manera:

1. Existe una arista desde $i \in V_1$ hasta $j \in V_2$ si y solamente si $j - i \in S$.
2. Existe una arista desde $j \in V_2$ hasta $k \in V_3$ si y solamente si $k - j \in S$.
3. Existe una arista desde $i \in V_1$ hasta $k \in V_3$ si y solamente si $\frac{k-i}{2} \in S$.

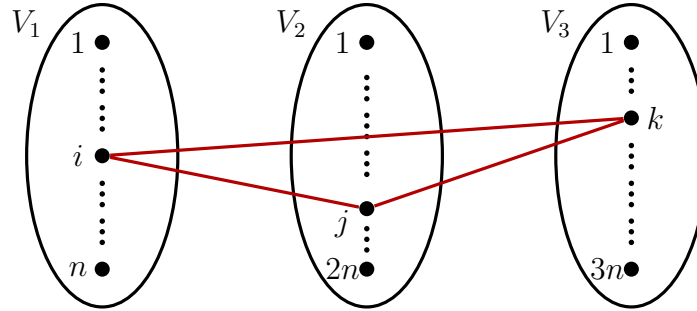


Figura 3.6: Esquema de relación entre un triángulo en G y una progresión aritmética de largo 3 en S .

Nótese que la tupla $(i, j, k) \in V_1 \times V_2 \times V_3$ define un triángulo en G si y solamente si $j - i \in S$, $k - j \in S$ y $\frac{k-i}{2} \in S$, o bien, $\{j - i, \frac{k-i}{2}, k - j\}$ es una progresión aritmética de largo 3 en S con diferencia $\frac{k-2j+i}{2}$.

Diremos que un triángulo $(i, j, k) \in V_1 \times V_2 \times V_3$ es trivial en G si para algún $s \in S$ se satisface que $j - i = \frac{k-i}{2} = k - j = s$. Entonces, observando que cada triángulo trivial se puede identificar con el par $(i, s) \in V_1 \times S$, la cantidad de estos es exactamente $n|S| \geq \varepsilon n^2$. Además, por construcción, no existen triángulos triviales que compartan una arista, por lo que no se pueden eliminar dos de ellos removiendo

1155 solo una arista. Por consecuencia, se tienen que eliminar al menos $\varepsilon n^2 = \frac{\varepsilon}{36}(6n)^2$
 1156 aristas para hacer de G libre de triángulos. Luego, utilizando el lema de eliminación
 1157 de triángulos con $\frac{\varepsilon}{36}$, recibimos $\delta\left(\frac{\varepsilon}{36}\right)$, y aseguramos que existen al menos $\delta(6n)^3 =$
 1158 $216\delta n^3$ triángulos en G . De esta manera, existen al menos $216\delta n^3 - n^2$ triángulos no
 1159 triviales. En conclusión, como $216\delta n^3 > n^2$, debe existir una progresión aritmética
 1160 no trivial de largo 3 en S .

1161

□

1162 Para la segunda demostración del teorema de Roth, será necesario un resultado
 1163 intermedio proporcionado por los autores M. Ajtai y E. Szemerédi [1]. Para su prueba,
 1164 se utiliza el Teorema 41.

1165 **Teorema 43.** (*Teorema de la esquina, Ajtai-Szemerédi*) Para todo $\varepsilon > 0$, existe
 1166 $n_0 \in \mathbb{N}$ tal que siempre que $n \geq n_0$, todo subconjunto $S \subset [n]^2$ con $|S| \geq \varepsilon n^2$ posee
 1167 elementos de la forma $\{(a, b), (a + d, b), (a, b + d)\}$ para algún $a, b, d \in \mathbb{N}$, con $d \neq 0$.

1168 *Demostración.* Sea $\varepsilon > 0$. Considere $n_0 \in \mathbb{N}$ suficientemente grande y $S \subset [n]^2$ con
 1169 al menos εn elementos, y $n \geq n_0$.

1170 Vamos a construir un grafo bipartito $G = (U \cup W, E)$ con conjunto de vértices
 1171 $U = \{u_1, \dots, u_n\}$ y $W = \{w_1, \dots, w_n\}$ definiendo las aristas de la siguiente manera:

$$u_i w_j \in E \iff (i, j) \in S.$$

1172 Ahora, interpretando a $[n]^2$ como una grilla bidimensional, se define una relación
 1173 entre pares de aristas de G de manera que se preserve cierta noción de distancia en
 1174 la grilla. Esto es:

$$u_i w_j \sim u_k w_\ell \iff i + j = k + \ell = q.$$

1175 Observe que para cada $2 \leq q \leq 2n$ se define un emparejamiento en G debido a
 1176 que no existen aristas que compartan un vértice, por lo que las clases de equivalencia

1177 (cada una asociada a algún q) de la relación forman una partición de emparejamientos
 1178 de E . En efecto, suponga que las aristas que pertenecen a la misma clase $u_i w_j$ y $u_k w_j$
 1179 comparten el vértice w_j . Entonces, como $i + j = k + j$, se determina que $u_i = u_k$ y
 1180 se concluye que $u_i w_j$ y $u_k w_j$ son la misma arista.

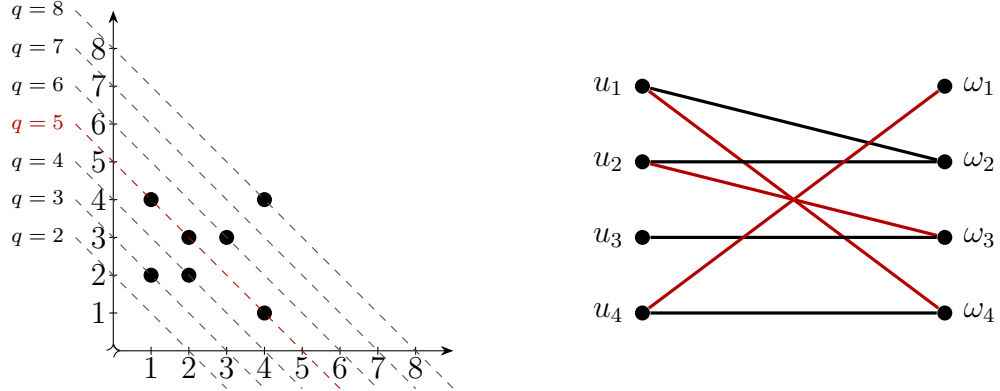


Figura 3.7: Ejemplo del emparejamiento generado por la
 clase de equivalencia asociada a $q = 5$, y $S =$
 $\{(1, 2), (1, 4), (2, 2), (2, 3), (3, 3), (4, 1), (4, 4)\}$.

1181 Luego, como $e_G = |S| \geq \varepsilon n^2 = \frac{\varepsilon}{4}(2n)^2$, el teorema de emparejamiento inducido
 1182 asegura que existe al menos un emparejamiento no inducido. Esto significa que en
 1183 al menos un emparejamiento que contiene las aristas con la relación $u_i w_j \sim u_k w_\ell$
 1184 puede existir el trío de aristas $u_i w_j$, $u_k w_\ell$ y $u_i w_\ell$. Así, para algún $d \in \mathbb{N}$, (i, j) , (k, ℓ)
 1185 y (i, ℓ) son elementos de S que satisfacen

$$k - i = j - \ell = d.$$

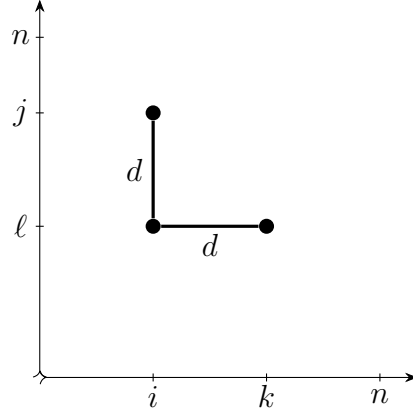


Figura 3.8: Esquema ilustrativo de la esquina formada en la demostración

1186 Finalmente, la demostración del teorema se consigue tomando $(i, \ell) = (a, b)$ para
 1187 obtener $j = b + d$ y $k = a + d$. □

1188 El resultado anterior otorga lo necesario para la segunda demostración (J. Solymosi [23]) del teorema de Roth utilizando el lema de regularidad de Szemerédi.
 1189

1190 *Segunda demostración Teorema 42.* Dado $\varepsilon > 0$, escogemos $n_0 \in \mathbb{N}$ suficientemente
 1191 grande. Para $n \geq n_0$, sea $S \subset [n]$ un conjunto que posee al menos εn elementos. Se
 1192 define el siguiente conjunto:

$$B = \{(x, y) \in [2n]^2 : x - y \in S\},$$

1193 Observe que cada $s \in S$ da lugar a exactamente n elementos en B con $x - y = s$,
 1194 permitiendo determinar que $|B| = n|S| \geq \varepsilon n^2$. Luego, el Teorema 43 asegura la
 1195 existencia de elementos de la forma $\{(s, b), (s, b+d), (s+d, b)\}$ en B . Por consecuencia,
 1196 se encuentra una progresión aritmética de largo 3 no trivial en A tomando $x = s - b$,
 1197 e $y = d$. □

Bibliografía

- [1] Miklós Ajtai y Endre Szemerédi. “Sets of lattice points that form no squares”. En: *Stud. Sci. Math. Hungar* 9.1975 (1974), págs. 9-11.
- [2] Béla Bollobás. *Random graphs*. Springer, 1998.
- [3] Béla Bollobás. *Modern graph theory*. Vol. 184. Springer Science & Business Media, 2013.
- [4] Fan R. K. Chung, Ronald L. Graham y Richard M. Wilson. “Quasi-random graphs”. En: *Combinatorica* 9 (1989), págs. 345-362.
- [5] S CIOABA y RR MARTIN. “TAO’S SPECTRAL PROOF OF THE SZEMERÉDI REGULARITY LEMMA”. En: (2013).
- [6] Reinhard Diestel. “Graph theory. fifth. vol. 173”. En: *Graduate Texts in Mathematics*. Springer, Berlin (2018).
- [7] P Erdős y A Rényi. “On random graphs I”. En: *Publ. math. debrecen* 6.290-297 (1959), pág. 18.
- [8] Paul Erdős y Paul Turán. “On some sequences of integers”. En: *Journal of the London Mathematical Society* 1.4 (1936), págs. 261-264.
- [9] Alan Frieze y Ravi Kannan. “The regularity lemma and approximation schemes for dense problems”. En: *Proceedings of 37th conference on foundations of computer science*. IEEE. 1996, págs. 12-20.
- [10] Harry Furstenberg. “Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions”. En: *Journal d’Analyse Mathématique* 31.1 (1977), págs. 204-256.

- [11] Edgar N Gilbert. “Random graphs”. En: *The Annals of Mathematical Statistics* 30.4 (1959), págs. 1141-1144.
- [12] William T Gowers. “Lower bounds of tower type for Szemerédi’s uniformity lemma”. En: *Geometric & Functional Analysis GAFA* 7.2 (1997), págs. 322-337.
- [13] William T Gowers. “A new proof of Szemerédi’s theorem”. En: *Geometric & Functional Analysis GAFA* 11.3 (2001), págs. 465-588.
- [14] WT Gowers. “The work of Endre Szemerédi”. En: *Online* (2012). URL: <http://www.abelprize.no/c54147/binfil/download.php>.
- [15] Svante Janson, Tomasz Łuczak y Andrzej Ruciński. *Random graphs*. John Wiley & Sons, 2011.
- [16] János Komlós, Ali Shokoufandeh, Miklós Simonovits y Endre Szemerédi. “The regularity lemma and its applications in graph theory”. En: *Summer school on theoretical aspects of computer science* (2000), págs. 84-112.
- [17] János Komlós y Miklós Simonovits. *Szemerédi’s Regularity Lemma and its applications in graph theory*. 1995.
- [18] Michael Krivelevich y Benny Sudakov. “Pseudo-random graphs”. En: *More sets, graphs and numbers: A Salute to Vera Sos and András Hajnal*. Springer, 2006, págs. 199-262.
- [19] Guy Moshkovitz y Asaf Shapira. “A short proof of Gowers’ lower bound for the regularity lemma”. En: *Combinatorica* 36.2 (2016), págs. 187-194.
- [20] Vojtěch Rödl. “On universality of graphs with uniformly distributed edges”. En: *Discrete Mathematics* 59.1-2 (1986), págs. 125-134.
- [21] Klaus F Roth. “On certain sets of integers”. En: *J. London Math. Soc* 28.104-109 (1953), pág. 3.
- [22] Imre Z Ruzsa y Endre Szemerédi. “Triple systems with no six points carrying three triangles”. En: *Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bol-yai* 18.939-945 (1978), pág. 2.

- 1247 [23] József Solymosi. “Note on a generalization of Roth’s theorem”. En: *Discrete and Computational Geometry: The Goodman-Pollack Festschrift*. Springer, 1248 2003, págs. 825-827.
- 1250 [24] Endre Szemerédi. “On sets of integers containing no four elements in arithmetic 1251 progression”. En: *Acta Mathematica Hungarica* 20.1-2 (1969), págs. 89-104.
- 1252 [25] Endre Szemerédi. “On sets of integers containing no k elements in arithmetic 1253 progression”. En: *Acta Arith* 27.199-245 (1975), pág. 2.
- 1254 [26] Terence Tao. *The spectral proof of the Szemerédi regularity lemma*. [Consultado 1255 el 31 de marzo de 2024]. 2012. URL: <https://terrytao.wordpress.com/2012/12/03/the-spectral-proof-of-the-szemerédi-regularity-lemma/>. 1256
- 1257 [27] Andrew Thomason. “Pseudo-random graphs”. En: *North-Holland Mathematics Studies*. Vol. 144. Elsevier, 1987, págs. 307-331. 1258
- 1259 [28] Yufei Zhao. *Graph theory and additive combinatorics: exploring structure and 1260 randomness*. Cambridge University Press, 2023.