



1

2

3

Grafos cuasi-aleatorios y lema de regularidad de Szemerédi

4

5

Estudiante:
Felipe Sánchez Erazo

6

Profesor Guía:

7

Dr. Hiệp Hàn

8

9

Tesis para optar al título de Ingeniero Matemático de la Universidad de Santiago de Chile

10

11

Departamento de Matemática y Ciencia de la computación
Universidad de Santiago de Chile

A mi abuelo, Sergio Sánchez.

1. Introducción

2. Preliminares

Este capítulo proporciona una introducción concisa de los conceptos y terminologías que se utilizarán en esta tesis. La sección 2.1 da un paseo por las nociones más básicas de la teoría de grafos, otorgando una línea de base para el desarrollo del documento. En la sección 2.2 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 2.3 contextualiza y motiva el contenido de la sección 3.

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

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

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

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

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:

$$\sum_{i=1}^k \frac{\beta_i^2}{\alpha_i} \geq \frac{\left(\sum_{i=1}^k \beta_i \right)^2}{\sum_{i=1}^k \alpha_i}. \quad (3)$$

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

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

2.1. Teoría de grafos

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

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

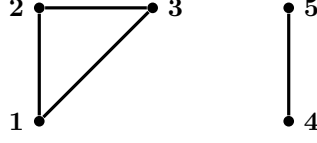


Figura 1: Ejemplo de un grafo con conjunto de vértices $V = \{1, 2, 3, 4, 5\}$ y conjunto de aristas $E = \{12, 23, 13, 45\}$.

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

$$e(X, Y) := \left| \{(x, y) \in X \times Y : xy \in E\} \right|. \quad (4)$$

41 Cuando $X \cap Y = \emptyset$, $e(X, Y)$ cuenta el número de aristas entre X e Y , y cuando $X \cap Y \neq \emptyset$,
 42 $e(X, Y)$ realiza un doble conteo sobre las aristas que se encuentran en $X \cap Y$. Se entenderá por
 43 **vecindad** de $u \in V$ como el conjunto de todos los vértices adyacentes a u , es decir,

$$N(u) := \{v \in V(G) : uv \in E(G)\}. \quad (5)$$

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

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

46 En particular, cuando $Y = V$,

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

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

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

$$\sum_{u \in V} \deg(u) = 2e_G. \quad (6)$$

50 *Demostración.* Cada arista $uv \in E$ será contada dos veces en la suma, una contribución por u , y
 51 otra por v . \square

52 En algunas ocasiones estaremos interesados en la cantidad de vecinos que comparten dos vértices
 53 del grafo $G = (V, E)$. Entonces, se define el **cogrado** de un par de vértices $u, v \in V$ no necesaria-
 54 mente diferentes mediante:

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

55 Mostraremos que existe una relación intrínseca entre los conceptos de grado y cogrado, cual será
 56 de utilidad en la sección 3.

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

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

58 *Demostración.* Utilizando las respectivas definiciones de grado y cogrado, el resultado se obtiene
59 de seguir el siguiente cálculo:

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

60 □

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

$$\sum_{u \in V} \deg(u)^2 = \sum_{u \in V} \sum_{v \in V} \text{codeg}(u, v). \quad (7)$$

62 A continuación, se enuncian algunos grafos especiales que son contemplados en esta tesis. Di-
63 remos que un grafo $G = (V, E)$ es **k -partito** si V se puede dividir en k subconjuntos disjuntos
64 V_1, V_2, \dots, V_k tales que si $uv \in E$ entonces $u \in V_i$ y $v \in V_j$, con $i \neq j$. En particular, a un grafo
65 2-partito lo llamaremos **bipartito**.

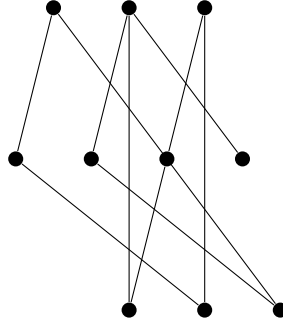


Figura 2: Ejemplo de un grafo 3-partito.

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

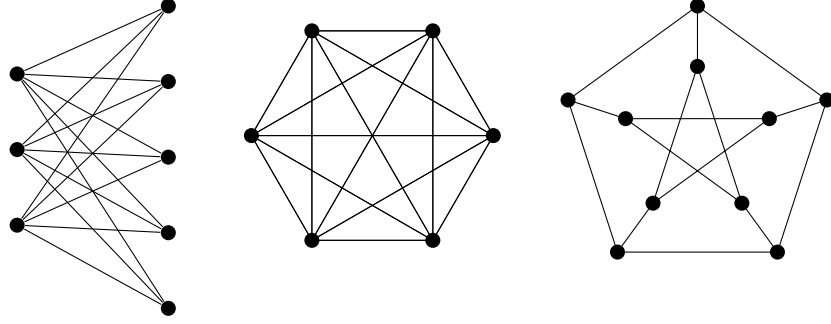


Figura 3: Ejemplo de los grafos especiales $K_{3,5}$, K_6 y 3-regular.

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

- Una **caminata**, es una secuencia de vértices no necesariamente distintos 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 es una **caminata cerrada**. El **largo** de una caminata está determinado por la cantidad de aristas que ésta posea.
- Un **ciclo**, es una caminata con $k \geq 2$ vértices únicos a excepción de v_k , que coincide con v_0 . Se denotará por C_k al ciclo de largo k .

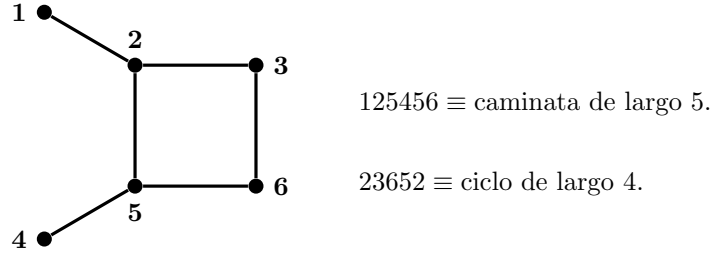


Figura 4: Ejemplo de una caminata y un ciclo.

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

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

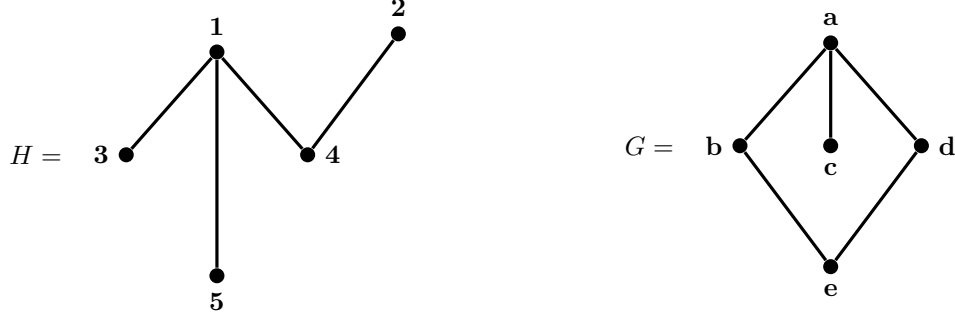


Figura 5: Ejemplo de una copia etiquetada de H en G mediante la función $f : V(H) \rightarrow V(G)$ definida por $f(1) = a$, $f(2) = e$, $f(3) = c$, $f(4) = b$ y $f(5) = d$.

2.2. 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 como 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 sus elementos de la 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 gozan 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 .

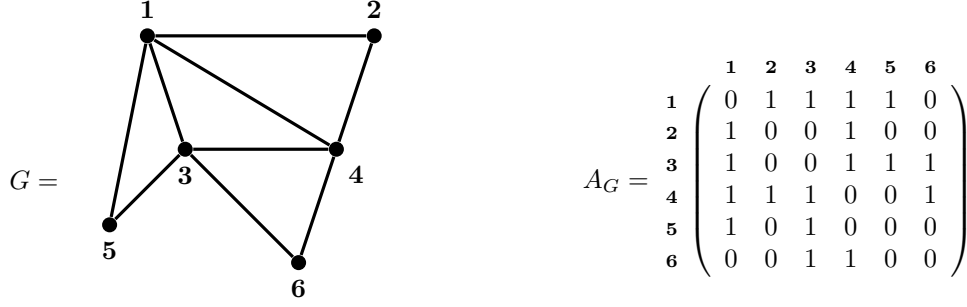


Figura 6: Ejemplo de representación de un grafo mediante la matriz de adyacencia.

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

$$A\mathbf{1} = \begin{pmatrix} \deg(1) \\ \vdots \\ \deg(n) \end{pmatrix} \quad (8)$$

Otro aspecto interesante de la matriz de adyacencia que será de utilidad en la sección 4, es que nos permite reescribir la ecuación (4) en función de ella. Para ver esto, considere la matriz de adyacencia A del grafo $G = ([n], E)$, y los vértices $i, j \in [n]$. Luego, según la definición 3,

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

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

$$e(X, Y) = \sum_{i \in X} \sum_{j \in Y} \mathbf{e}_i^T A \mathbf{e}_j = \mathbf{v}_X^T A \mathbf{v}_Y. \quad (9)$$

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

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

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

Demostración. Cuando $t = 1$, existe una caminata de largo 1 que conecta los vértices i y j si y solamente si $a_{ij}^{(1)} = 1$. Ahora, asuma que el lema se cumple para algún $t > 1$ fijo. Note que cualquier

126 caminata de largo $t + 1$ entre i y j contiene una caminata de largo t desde i hasta un vecino de j ,
 127 digamos k . Entonces si $k \in N(j)$, por la asunción del lema, el número de caminatas de largo t entre
 128 i y k es $a_{ik}^{(t)}$. Por lo tanto, el número total de caminatas de largo $t + 1$ desde i hasta j es exactamente

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

129

□

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

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

135 Por otro lado, para introducir algunos aspectos de la teoría espectral de grafos, recuerde que el
 136 vector no nulo $\mathbf{v} \in \mathbb{R}^n$ es un **vector propio** de alguna matriz $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ con **valor propio**
 137 $\lambda \in \mathbb{C}$ si $A\mathbf{v} = \lambda\mathbf{v}$. Esto significa que λ es un valor propio si y solo si $\lambda I_n - A$ es una matriz singular.
 138 Así, los valores propios vienen dados por las raíces del polinomio característico $\det(xI_n - A)$. En
 139 este trabajo, cuando se haga referencia a los valores y vectores propios de un grafo G , siempre será
 140 con respecto a su matriz de adyacencia A . Por ejemplo, si G es un grafo d -regular, entonces con
 141 la igualdad (8) se puede deducir que d es el valor propio asociado al vector propio normalizado de
 142 1-entradas de la matriz de adyacencia A_G .

143 **Proposición 6.** *Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz real simétrica, entonces todos sus valores propios*
 144 *son reales. Además, si dos vectores propios están asociados a distintos valores propios, entonces éstos*
 145 *son ortogonales. Más aún, el conjunto de todos los vectores propios define una base ortonormal de*
 146 \mathbb{R}^n .

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

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

150 Además, como A es simétrica,

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

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

153 Por otro lado, considere $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ vectores propios distintos de A asociados a los valores propios
 154 $\lambda, \mu \in \mathbb{R} \setminus \{0\}$ respectivamente. Calculamos como sigue,

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

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 la ortogonalidad de los vectores propios de A , defina $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ como el conjunto de vectores propios normalizados de A para probar que \mathcal{B} constituye una base ortonormal de \mathbb{R}^n . Para esto, sean $c_1, \dots, c_n \in \mathbb{R}$ tales que

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

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

$$A = PDP^T = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^T. \quad (10)$$

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 (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 como se comporta el cuadrado de una matriz simétrica $A \in \mathcal{M}_{n \times n}(\mathbb{R})$:

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

Luego, de manera inductiva se obtiene que $A^k = PD^k P^T$. Esta propiedad resulta altamente útil de cara al cálculo de caminatas de largo k entre dos vértices de un grafo. Más aún, la Proposición 8 y el Corolario 9 mostrarán que el número de caminatas cerradas en un grafo queda totalmente determinado por los valores propios del mismo.

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

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

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

182

Con esto, se concluye utilizando la descomposición espectral.

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

183

□

184

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

185

186

187

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

188

189

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

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

190

□

191

192

193

194

195

De esta manera, la cantidad de caminatas cerradas de largo k entre dos vértices de un grafo se simplifica a solo calcular la suma de la k -ésima potencia de todos sus valores propios. Más adelante, en la sección 3, esta propiedad será de utilidad debido a que entrega una buena aproximación de la cantidad de copias etiquetadas de ciclos de largo k que existen en un grafo $G = ([n], E)$. En particular, si A es la matriz de adyacencia de G y $\lambda_1, \dots, \lambda_n$ sus valores propios,

$$\left| \binom{G}{C_k} \right| = \text{Tr}(A^k) + o(n^k) = \sum_{i=1}^n \lambda_i^k + o(n^k) \quad (11)$$

196

197

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

198

199

Teorema 10. (*Teorema de Courant-Fischer*) Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz real 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 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)

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

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

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

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

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

202 Por otro lado, observando que el valor propio más grande de A^2 es λ_1^2 , se concluye para cualquier
203 $\mathbf{x} \in \mathbb{R}^n$ que

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

204

□

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

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

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

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

212 Adicionalmente, con apoyo de la igualdad (8) y usando la cota anterior, se concluye que $\lambda_1 = d$
213 cada vez que G es un grafo d -regular. □

214 2.3. Grafos aleatorios

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

217 Intuitivamente, se podría pensar en un *grafo aleatorio* de n vértices como el resultado de se-
218 leccionar aleatoriamente un subconjunto de aristas de K_n . En 1959, Erdős-Rényi y Edgar Gilbert*
219 [referencia](#) * proponen dicha selección de la siguiente manera: comenzando con un grafo sin aristas

220 $G = ([n], \emptyset)$, decidir sobre cada par de vértices de G si agregar una arista con una probabilidad
 221 p establecida. En cada repetición del proceso anterior se genera un nuevo grafo de n vértices, que
 222 contribuye a la creación del espacio de probabilidad conocido como $G(n, p)$, y se denomina modelo
 223 binomial. Entonces, considerando \mathcal{G}^n como el conjunto de todos los grafos de n vértices, se define
 224 formalmente.

225 **Definición 13.** (*Modelo binomial*) Sea $p \in (0, 1)$. Se define $G(n, p)$ como el espacio de probabilidad
 226 $(\mathcal{G}^n, \mathcal{P}(\mathcal{G}^n), \mathbb{P})$, con

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

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

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

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

232 Para dar prueba a la proposición anterior es necesario utilizar la desigualdad de Chernov. Exis-
 233 tiendo diversas formas de expresar tal desigualdad, en esta tesis se utiliza el resultado para el caso
 234 en que cada variable aleatoria solo toma los valores 0 o 1, como se plantea en [* referencia *](#) en la
 235 ecuación (2.12) de la observación 2.5.

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

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

239 Con esto, damos paso a la demostración prometida.

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

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

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

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

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

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

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

$$\begin{aligned} \mathcal{P} \left(|X_i - \mu_i| > \frac{\varepsilon}{3}n^2 \right) &\leq 2 \exp \left(-\frac{2 \left(\frac{\varepsilon}{3}n^2 \right)^2}{|A||B|} \right) \\ &\leq 2 \exp \left(-\frac{2}{9}\varepsilon^2 n^2 \right). \end{aligned} \quad (12)$$

247 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

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

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

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

$$\begin{aligned} 1 - \mathbb{P}(\mathcal{P}_n^{p,\varepsilon}) &= \mathbb{P} \left(\exists A, B \subset V(G) : \left| X - p|A||B| \right| > \varepsilon n^2 \right) \\ &\leq \mathbb{P} \left(\exists A, B \subset V(G) : |X_1 - \mu_1| > \frac{\varepsilon}{3}n^2 \vee |X_2 - \mu_2| > \varepsilon n^2 \right) \\ &\leq \mathbb{P} \left(\exists A, B \subset V(G) : |X_1 - \mu_1| > \frac{\varepsilon}{3}n^2 \right) + \mathbb{P} \left(|X_2 - \mu_2| > \frac{\varepsilon}{3}n^2 \right) \\ &\leq \sum_{\substack{A \subset V(G) \\ B \subset V(G)}} \mathbb{P} \left(|X_1 - \mu_1| > \frac{\varepsilon}{3}n^2 \right) + \sum_{\substack{A \subset V(G) \\ B \subset V(G)}} \mathbb{P} \left(|X_2 - \mu_2| > \frac{\varepsilon}{3}n^2 \right) \\ &\stackrel{(12)}{\leq} 2^{2n+1} \exp \left(-\frac{2}{9}\varepsilon^2 n^2 \right) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

251

□

252 3. Cuasi-aleatoriedad

253 Trabajar con estructuras discretas aleatorias brinda una amplia gama de propiedades ideales
 254 y deseables, lo que las convierte en piezas fundamentales tanto en matemáticas como en ciencias
 255 de la computación. Por ejemplo, el modelo de grafo aleatorio binomial goza de una distribución
 256 uniforme de aristas, buenas propiedades y es robusto. La cuestión ahora es cómo capturar las
 257 propiedades esenciales de la aleatoriedad dentro de un marco *determinista*. Esta idea condujo a la
 258 noción de cuasi-aleatoriedad, que en la actualidad, es un tópico central en las matemáticas discretas.
 259 En términos generales, las propiedades cuasi-aleatorias son características deterministas que son
 260 propias de objetos realmente aleatorios. Aunque la noción de cuasi-aleatoriedad es interesante por
 261 sí misma, su estudio ha revelado profundas conexiones entre varias ramas de la matemática y

ciencias de la computación, encontrando aplicaciones en teoría de grafos, teoría de números, teoría ergódica, geometría, y algoritmos y complejidad.

Como se verá a detalle más adelante en la sección 4, una de las razones principales por las cuales el estudio de la cuasi-aleatoriedad no se limita a un área específica, es el hecho de que existe un teorema de partición que permite la aproximación de cualquier objeto discreto por otros cuasi-aleatorios. Con esto, nos referimos al célebre lema de regularidad de Szemerédi, que establece que todo grafo se puede aproximar mediante un número finito de grafos cuasi-aleatorios, permitiendo la conexión entre un grafo arbitrario y los cuasi-aleatorios.

El estudio sistemático de los grafos cuasi-aleatorios fue iniciado por Rödl *referencia* y Thomason *referencia*, y su punto inicial es la siguiente noción de *distribución uniforme de aristas* para definir la cuasi-aleatoriedad de un grafo. *Buscar año...*

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

$$e(X, Y) = p|X||Y| + o(n^2). \quad (13)$$

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

3.1. Teorema de Chung, Graham y Wilson

Una contribución revolucionaria en la teoría de grafos cuasi-aleatorios fue en 1989 por Fan Chung, Ronald Graham y Richard M. Wilson *referencia*. Ellos presentaron una extensa lista de propiedades superficialmente diferentes entre sí, y demostraron que todas son equivalentes al concepto de cuasi-aleatoriedad entendido en la Definición 3.

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

Teorema 17. (Chung, Graham y Wilson) Sea $p \in (0, 1)$ fijo. Para cualquier secuencia 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 siguientes propiedades son equivalentes:

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

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

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

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

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

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

294 $\text{COUNT}_{C4,p}$: La cantidad de copias etiquetadas de ciclos de largo 4 es

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

295 CODEG_p :

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

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

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

297 Para una comprensión e intuición inicial de cada propiedad del Teorema 17, se ha utilizado
 298 notación asintótica en su enunciado. Sin embargo, con dicha formulación no queda del todo claro
 299 las dependencias cuantificadas de los errores en las implicancias cada par de propiedades. Enton-
 300 ces, se replantean equivalentemente las propiedades con una versión cuantitativa, asociando algún
 301 parámetro de error ε en todo grafo específico G con un conjunto de vértices suficientemente grande.
 302 Por ejemplo, bajo los supuestos del Teorema 17, asuma que la sucesión de grafos $(G_n)_{n \rightarrow \infty}$ satisface
 303 DISC_p , y luego, la versión equivalente establece que para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que el grafo
 304 G sobre $n \geq n_0$ vértices satisface

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

305 De manera general, diremos que una secuencia de grafos $(G_n)_{n \rightarrow \infty}$ con $|V(G_n)| = n$ satisface
 306 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
 307 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
 308 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 $\varepsilon > 0$,
 309 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
 310 satisfaga la propiedad $P_{x_1, \dots, x_k}(\varepsilon)$. Se desarrollará la demostración formal del Teorema 17 utilizando
 311 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
 312 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
 313 constantes $C, c > 0$.

314 Demostración Teorema de Chung, Graham y Wilson

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

$$\begin{array}{ccccc} \text{DISC}'_p & \xRightarrow{\text{Prop. 20.}} & \text{COUNT}_p & \xRightarrow{\text{Prop. 21.}} & \text{COUNT}_{C4,p} & \xRightarrow{\text{Prop. 24. y 25.}} & \text{EIG}_p \\ \updownarrow \text{Prop. 18 y 19.} & & & & \downarrow \text{Prop. 22.} & & \\ \text{DISC}_p & & \xleftarrow{\text{Prop. 23.}} & & \text{CODEG}_p. & & \end{array} \quad (14)$$

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 $k = 2$.

317 Con esto en mente, damos paso a la demostración de cada proposición considerada en el esquema
318 (14).

319 **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*
320 *$n \geq n_0$ vértices satisface $\text{DISC}'_p(\varepsilon)$ cada vez que cumpla con $\text{DISC}_p(\delta)$. En particular,*

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

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

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

324 Las igualdades anteriores consideran $e(X, X) = 2e(X)$, por definición, y la aproximación $\binom{|X|}{2} =$
325 $\frac{|X|^2}{2} \pm \delta n^2$. □

326 **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*
327 *$n \geq n_0$ vértices satisface $\text{DISC}_p(\varepsilon)$ cada vez que cumpla con $\text{DISC}'_p(\delta)$. En particular,*

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

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

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

$$e(X, Y) = e(X \cup Y) + e(X \cap Y) - e(X \setminus Y) - e(Y \setminus X). \quad (15)$$

333 Observe que con esta configuración, el conteo de las aristas entre X e Y es doble cuando los
334 vértices que componen las aristas pertenecen a $X \cap Y$. Luego, se utiliza la propiedad $\text{DISC}'_p(\delta)$
335 sobre la identidad (15) para conseguir el resultado.

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

336 □

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

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

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

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

$$\left| \binom{G}{F} \right| = p^{e_F} n^\ell \pm 4e_F \delta n^\ell. \quad (16)$$

344 Una vez probada la ecuación (16), el resultado seguirá de tomar $F = H$ y la elección de δ para
 345 conseguir las siguientes desigualdades:

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

346 Entonces, cuando $e_F = 1$, $\left| \binom{G}{F} \right|$ es el número de pares ordenados de vértices de G que forman
 347 una arista junto a cualquier combinación de $\ell - 2$ vértices para completar una copia de F . Es decir,

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

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

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

349 Así, con $\left| \binom{G}{F} \right| = pn^\ell \pm 4\delta n^\ell$, se prueba el caso inicial de la inducción. Ahora, sea F un grafo
 350 con $e_F > 1$ aristas y asuma que se satisface la ecuación (16) en cualquier grafo con una cantidad de
 351 aristas menor que e_F . Para desarrollar la inducción, suponga que $ij \in E(F)$ y considere la siguiente
 352 notación:

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

355 Sea T^- una copia etiquetada de F^- en G , es decir, T^- se corresponde una aplicación inyectiva
 356 $f : V(F^-) \rightarrow V(T^-) \subseteq V$ tal que $f(u)f(v) \in E(T^-)$ cada vez que $uv \in E(F^-)$. Entonces,
 357 considerando $e_{T^-} := f(i)f(j)$, se escribe la cantidad de copias etiquetadas de F en G de manera
 358 conveniente para utilizar la hipótesis inductiva como se muestra a continuación:

$$\begin{aligned}
\left| \binom{G}{F} \right| &= \sum_{T^- \in \binom{G}{F^-}} \mathbb{1}_E(e_{T^-}) \\
&= \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) + p - p] \\
&= \sum_{T^- \in \binom{G}{F^-}} p + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
&= p \left| \binom{G}{F^-} \right| + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
&\stackrel{(16)}{=} p^{e_F} n^\ell + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \pm 4(e_F - 1)\delta n^\ell.
\end{aligned} \tag{17}$$

En este punto, es suficiente probar que el segundo sumando de la desigualdad (17) es pequeño. Para esto, considere T^* una copia de F^* , y denote por F_i^* y F_j^* a los grafos resultantes de eliminar de F^- los vértices j e i respectivamente. Con esto, defina los siguientes conjuntos:

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

Los conjuntos anteriores, por construcción, son tales que para cada tupla $(a, b) \in A_i^{T^*} \times A_j^{T^*}$ añadida a T^* se obtiene una copia de F^- . Así, reescribiendo el segundo sumando de la igualdad (17) convenientemente y utilizando la propiedad $\text{DISC}'_p(\delta)$,

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

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

367 **Proposición 21.** Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G sobre
 368 $n \geq n_0$ vértices satisface $\text{COUNT}_{C_4, p}(\varepsilon)$ cada vez que cumpla con $\text{COUNT}_p(\delta)$. En otras palabras,

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

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

370 **Proposición 22.** Para todo $\varepsilon > 0$ y $p \in (0, 1)$, existe $\delta > 0$ y $n_0 \in \mathbb{N}$ tales que el grafo G
 371 sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{CODEG}_p(\varepsilon)$ cada vez que cumpla con
 372 $\text{COUNT}_{C_4, p}(\delta)$. En particular,

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

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

375 La clave de esta demostración radica en encontrar una buena cota para $\sum_{u, v \in V(G)} \text{codeg}(u, v)$
 376 y $\sum_{u, v \in V(G)} \text{codeg}(u, v)^2$. Para esto, será necesario la utilización apropiada de la desigualdad de
 377 Cauchy-Schwarz vista en (2). Por un lado, con la relación entre el grado y el cogrado (7) se obtiene
 378 la primera de las cotas:

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

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

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

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

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

381

□

382 **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
383 $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{DISC}_p(\varepsilon)$ cada vez que cumpla con $\text{CODEG}_p(\delta)$.
384 En particular,

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

385 *Demostración.* Dado $\varepsilon > 0$, $p \in (0, 1)$, seleccionamos $\delta < \frac{\varepsilon^4}{81}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Sea
386 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 $\text{CODEG}_p(\delta)$.

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

$$\begin{aligned}
\sum_{x \in V(G)} \left| \deg(x) - pn \right| &\stackrel{\text{DCS}}{\leq} n^{1/2} \left(\sum_{x \in V(G)} (\deg(x) - pn)^2 \right)^{1/2} \\
&= n^{1/2} \left(\sum_{x \in V(G)} \deg(x)^2 - 2pn \sum_{x \in V(G)} \deg(x) + p^2 n^3 \right)^{1/2} \\
&\stackrel{(7)}{=} n^{1/2} \left(\left(\sum_{u,v \in V(G)} \text{codeg}(u,v) - p^2 n \right) - 4pne_G + 2p^2 n^3 \right)^{1/2} \\
&\leq n^{1/2} \left(\left(\sum_{u,v \in V(G)} \left| \text{codeg}(u,v) - p^2 n \right| \right) + 4pn \left(\delta n^2 - \frac{pn^2}{2} \right) + 2p^2 n^3 \right)^{1/2} \\
&\leq n^{1/2} (2p^2 n^3 - 2p^2 n^3 + 4p\delta n^3 + \delta n^3)^{1/2} \\
&< 3\delta^{1/2} n^2.
\end{aligned}$$

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

$$\left| e(X, Y) - p|X||Y| \right| = \left| \sum_{x \in X} (\deg(x; Y) - p|Y|) \right| \stackrel{DCS}{\leq} n^{1/2} \left(\sum_{x \in X} (\deg(x; Y) - p|Y|)^2 \right)^{1/2}. \quad (18)$$

391 En la desigualdad anterior se ha conseguido que el argumento de la suma sea siempre no negativo,
 392 lo que permite extender su dominio de X a $V(G)$. De esta manera, usando a la cota proveniente de
 393 la conentración de los grados en los vértices de G , se prueba el resultado continuando desde (18):

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

394 □

395 **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
 396 $n \geq n_0$ 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 $\text{EIG}_p(\delta)$.
 397 En particular,

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

398 *Demostración.* Dado $\varepsilon > 0$ y $p \in (0, 1)$, se elige $\delta < \frac{\varepsilon}{20}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Conside-
 399 ramos 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 $\text{EIG}_p(\delta)$,
 400 $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ como la matriz de adyacencia de G , y $\lambda_1 \geq \dots \geq \lambda_n$ los valores propios de A .

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

$$\left| \binom{G}{C_4} \right| = \text{Tr}(A^4) \pm \delta n^4 = \sum_{i=1}^n \lambda_i^4 \pm \delta n^4 = \lambda_1^4 + \sum_{i=2}^n \lambda_i^4 \pm \delta n^4. \quad (19)$$

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

$$\sum_{i=2}^n \lambda_i^4 \leq \max_{i \neq 1} \lambda_i^2 \sum_{i=1}^n \lambda_i^2 \leq \delta n^2 \text{Tr}(A^2) \leq \delta n^2 (pn^2 + 2\delta n^2) \leq 3\delta n^4. \quad (20)$$

405 Finalmente, se concluirá tras usar la propiedad $\text{EIG}_p(\delta)$ sobre el primer valor propio y la cota
 406 mostrada en (20). Entonces, continuando desde la ecuación (19),

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

407 □

408 **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
 409 $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas satisface $\text{EIG}_p(\varepsilon)$ cada vez que cumpla la propiedad
 410 $\text{COUNT}_{C_4,p}(\delta)$. Es decir,*

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

411 *Demostración.* Sea $\varepsilon > 0$ y $p \in (0, 1)$, escogemos $\delta < \frac{\varepsilon^4}{4}$ y $n_0 \in \mathbb{N}$ suficientemente grande. Sea
 412 también G un grafo sobre $n \geq n_0$ vértices y $e_G = \frac{pn^2}{2} \pm \delta n^2$ aristas que satisface la propiedad
 413 $\text{COUNT}_{C_4,p}(\delta)$, $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ la matriz de adyacencia de G , y $\lambda_1 \geq \dots \geq \lambda_n$ los valores propios
 414 de A .

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

$$\lambda_1 \geq \frac{\sum_{x \in V(G)} \deg(x)}{n} = \frac{2e_G}{n} = \frac{2}{n} \left(\frac{pn^2}{2} \pm \delta n^2 \right) \geq pn - 2\delta n. \quad (21)$$

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

$$\lambda_1^4 \leq \sum_{i=1}^n \lambda_i^4 = \text{Tr}(A^4) = \left| \binom{G}{C_4} \right| \pm \delta n^4 \leq p^4 n^4 + 2\delta n^4. \quad (22)$$

418 La desigualdad (22) implica que $\lambda_1 \leq pn + (2\delta)^{1/4}n$, y en combinación con la cota vista en (21),
 419 se obtiene que $\lambda_1 = pn \pm (2\delta)^{1/4}n$. Por último, observe por las cotas vistas anteriormente que

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

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

3.2. Aspectos adicionales del teorema de CGW

La noción inicial presentada de un grafo cuasi-aleatorio por distribución de aristas según la Definición 3 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, en primer lugar se 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.

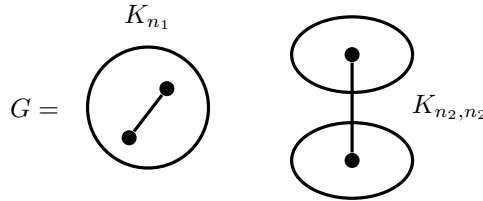


Figura 7: Esquema de la configuración del grafo usado como contraejemplo. Aquí, $\bullet-\bullet$ representa las aristas permitidas dentro del grafo G .

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:

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

443 Bajo esta configuración, se encontrará el parámetro $p \in (0, 1)$ de manera tal que el grafo G
444 posea la cantidad esperada de aristas y copias etiquetadas de C_3 según lo haría un grafo aleatorio
445 $G(n, p)$. Para ello, se plantea el siguiente sistema de ecuaciones:

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

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

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

448 Como también,

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

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

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

454 De manera más general, es posible extender la propiedad $\text{COUNT}_{C_4, p}$ a $\text{COUNT}_{C_{2t}, p}$ con $t \geq 2$.
455 Es decir,

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

456 Se expone un bosquejo de la demostración.

457 **Proposición 27.** Sea $p \in (0, 1)$ y $(G_n)_{n \rightarrow \infty}$ una secuencia de grafos con $|V(G_n)| = n$ vértices y
458 $e_{G_n} = (p + o(1)) \binom{n}{2}$ aristas, entonces las propiedades $\text{COUNT}_{C_{2t}, p}$ y EIG_p son equivalentes.

459 *Demostración.* Este resultado es una consecuencia directa de las demostraciones de la Proposición
460 24 y 25 tras el siguiente par de observaciones. En primer lugar, la cantidad de copias etiquetadas
461 caminatas cerradas de largo $2t$ que no son C_{2t} en G_n están dentro de un error $O(n^{2t-1})$, es decir,

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

462 También, se debe modificar la cota presentada en la ecuación (20) como sigue:

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

463 Con estas observaciones el resultado queda demostrado. \square

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

467 Finalmente, se explora un caso siempre interesante de estudio, ya que simplifica varios cálculos
 468 y surge de manera recurrente en la vida cotidiana: un grafo d -regular. En nuestro contexto, se verá
 469 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
 470 con $\text{EIG}_{\frac{d}{n}}$. Dicha equivalencia nace como una consecuencia del siguiente teorema.

471 **Teorema 28.** (*Expander Mixing Lemma*) Sea $G = ([n], E)$ un grafo d -regular, y $d = \lambda_1 \geq \lambda_2 \geq$
 472 $\dots \geq \lambda_n$ los valores propios asociados a la matriz de adyacencia A de G . Si se denota:

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

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

$$\left| e(X, Y) - \frac{d}{n} |X| |Y| \right| \leq \lambda \sqrt{|X| |Y| \left(1 - \frac{|X|}{n} \right) \left(1 - \frac{|Y|}{n} \right)}. \quad (23)$$

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

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

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

477 Conforme a la ecuación (9), para todo $X, Y \subset [n]$, se cumple la siguiente igualdad:

$$e(X, Y) = \mathbf{v}_X^T A \mathbf{v}_Y = \mathbf{v}_X^T A_1 \mathbf{v}_Y + \mathbf{v}_X^T \Delta \mathbf{v}_Y. \quad (24)$$

478 De la ecuación anterior se espera que el primer sumando sea el término principal, mientras que
 479 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} .
 480 Es decir,

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

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

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

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

483 Ahora, se estudian los sumandos de la igualdad (24) por separado. Por un lado,

$$\begin{aligned}
\mathbf{v}_X^T A_1 \mathbf{v}_Y &= \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i \right)^T (\lambda_1 \mathbf{v}_1 \mathbf{v}_1^T) \left(\sum_{j=1}^n \beta_j \mathbf{v}_j \right) \\
&= \lambda_1 \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i^T \right) (\mathbf{v}_1 \mathbf{v}_1^T) \left(\sum_{j=1}^n \beta_j \mathbf{v}_j \right) \\
&= \lambda_1 \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (\mathbf{v}_i^T \mathbf{v}_1) (\mathbf{v}_1^T \mathbf{v}_j) \\
&= \lambda_1 \alpha_1 \beta_1.
\end{aligned} \tag{25}$$

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

$$\mathbf{v}_X^T \Delta \mathbf{v}_Y = \left(\sum_{i=1}^n \alpha_i \mathbf{v}_i \right)^T \left(\sum_{j=2}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^T \right) \left(\sum_{k=1}^n \beta_k \mathbf{v}_k \right) = \sum_{i=1}^n \lambda_i \alpha_i \beta_i. \tag{26}$$

485 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 vector propio
486 respectivamente de A . En consecuencia,

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

487

488 Así, la ecuación (25) resulta en $\mathbf{v}_X^T A_1 \mathbf{v}_Y = \frac{d}{n} |X| |Y|$.

489 Para el término de error, recordando la definición de λ , se desarrolla el valor absoluto de la
490 ecuación (26) usando la desigualdad de Cauchy-Schwarz.

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

491 Finalmente, el resultado se obtiene directamente de tomar el valor absoluto de la ecuación (24)
492 de la siguiente manera:

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

493

□

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

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

498 Finalmente, en un grafo d -regular, la equivalencia entre las propiedades $\text{EIG}_{\frac{d}{n}}$ y $\text{DISC}_{\frac{d}{n}}$ se
499 completa por el camino de implicancias ya demostradas según el esquema (14).

500 4. Lema de regularidad de Szemerédi

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

505 Erdős y Turán *referencia* conjeturaron en 1936 que todo conjunto de números enteros su-
506 ficientemente grande posee una progresión aritmética de longitud arbitraria, digamos k . En 1953,

507 Klaus Roth [* referencia *](#) da el primer resultado positivo a la conjetura para el caso en que $k = 3$
508 utilizando análisis de Fourier, dando paso al teorema de Roth. Más adelante, en 1969, Szemerédi [*
509 referencia *](#) extendió el teorema de Roth a progresiones aritméticas de largo 4 vía métodos combi-
510 natorios. Seis años después, en 1975, Endre Szemerédi demuestra la conjetura de Erdős-Turán para
511 progresiones aritméticas de longitud arbitraria, lo que se conoce actualmente como el teorema de
512 Szemerédi.

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

516 La demostración que propone Szemerédi es considerada una obra maestra del razonamiento
517 combinatorio. El siguiente esquema entrega una noción de la complejidad de la idea.

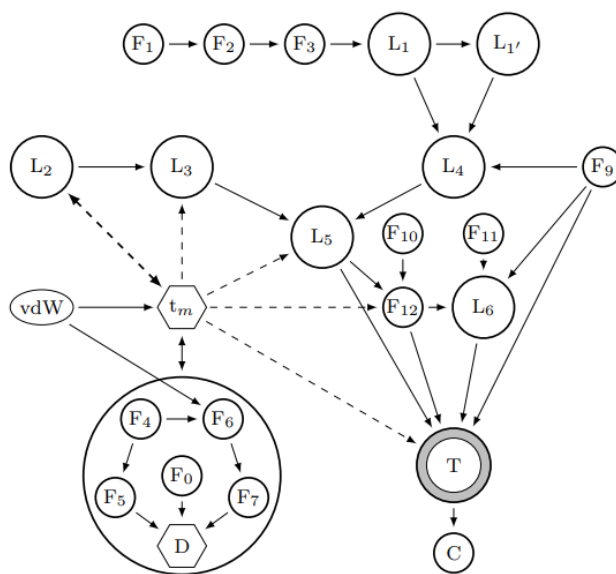


Figura 8: El esquema representa el diagrama de flujo de la demostración del teorema de Szemerédi ([* referencia *](#), página 202). Algunos de los símbolos involucrados significan lo siguiente: $F_k \equiv$ Hecho k , $L_k \equiv$ Lema k , $T \equiv$ Teorema, $C \equiv$ Corolario, $D \equiv$ Definiciones varias, etc.

518 Probablemente la complejidad de tal demostración motivó a otros matemáticos en encontrar
519 nuevas formas de probar el teorema. Poco después, en 1977, Hillel Furstenberg [* referencia *](#) obtiene
520 una demostración fundamentada en la teoría ergódica. También, en 2001, T. Gowers [* referencia *](#)
521 entrega una tercera demostración por medio de análisis de Fourier, extendiendo el camino iniciado
522 por Roth para progresiones aritméticas de largo 3.

523 En esta tesis, nos enfocaremos en uno de los elementos que compone la demostración entregada
524 por Szemerédi del Teorema 29, el hoy conocido como *lema de regularidad de Szemerédi*. A grandes
525 rasgos, el lema dice que el conjunto de vértices de todo grafo puede ser particionado en una cantidad
526 finita de partes que muestran comportamientos *regulares* (o cuasi-aleatorios) entre la mayoría de

los pares de partes. Este hecho permite entender cualquier grafo con menos información, y se aprovechan cada una de las propiedades equivalentes vistas en el Teorema 17.

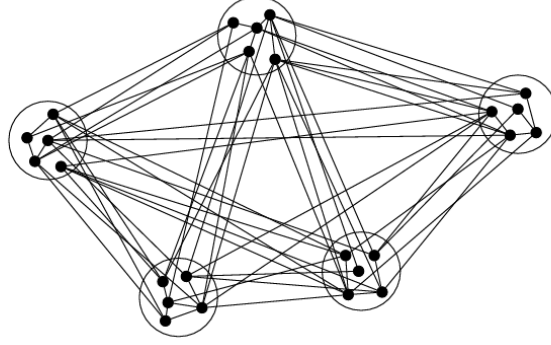


Figura 9: Ejemplo ilustrativo con cinco partes del resultado del lema de regularidad de Szemerédi sobre un grafo.

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:

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}. \quad (27)$$

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

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

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

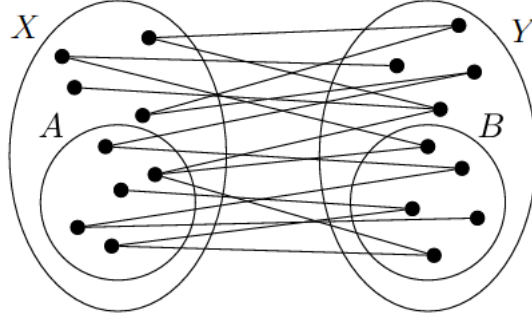


Figura 10: Ejemplo del par (X, Y) ε -regular, cuando $X \cap Y = \emptyset$. Las densidades $d(X, Y)$ y $d(A, B)$ no varían demasiado.

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

$$\left| e(X, Y) - p|X||Y| \right| = o(|U||W|), \quad \forall X \subset U, \forall Y \subset W. \quad (28)$$

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

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

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

$$d(U, W)|A||B| - \varepsilon|U||W| < 0 \leq e(A, B) \leq |A||B| \leq \varepsilon|U||W| < d(U, W)|A||B| + \varepsilon|U||W|.$$

De esta manera, tomando $p = d(U, W)$, se obtiene la analogía planteada.

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

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

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

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

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

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

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

562 Otra forma de encontrar el resultado, es cuando todas las partes de la partición poseen aproxi-
563 madamente el mismo tamaño.

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

567 Lo poderoso del lema de regularidad es que la cota de partes que entrega es independiente del
568 tamaño del grafo, y solo depende del parámetro ε . Esto ya que en grafo más grandes, el tamaño de
569 las partes podrían ser considerablemente más grandes.

570 En la subsección 4.1 se demostrará rigurosamente el Teorema 32 desde una mirada clásica, y
571 se expondrá la manera de probar el Teorema 33. Más adelante, en la sección 4.2, nos limitaremos
572 a mostrar una forma novedosa e ingeniosa de demostrar el Teorema 32 desde una perspectiva
573 espectral. Finalmente, la subsección 4.3 tiene el objetivo de demostrar de dos maneras diferentes el
574 teorema de Roth utilizando el lema de regularidad de Szemerédi.

575 4.1. Demostración por incremento de energía

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

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

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

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

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

591 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 define

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

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

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

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

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

594 La última observación es crucial en la demostración del Teorema 32, puesto que los Lemas
595 35, 36 y 37 nos asegurarán que la energía de una partición nunca decrece bajo refinamiento. Por
596 consecuencia, el algoritmo de la técnica *argumento de incremento de energía* tendrá un fin, y
597 entregará una partición ε -regular. Dicho esto, procedemos a enunciar y demostrar cada uno de los
598 lemas mencionados para probar clásicamente el Teorema 32.

599 El priero de los lemas, afirma que la energía de una partición no decrece bajo cualquier refina-
600 miento arbitrario.

601 **Lema 35.** Sea G un grafo, $X, Y \subset V(G)$, y \mathcal{P}_X y \mathcal{P}_Y particiones de X e Y respectivamente,
602 entonces $q(\mathcal{P}_X, \mathcal{P}_Y) \geq q(X, Y)$. Además, si \mathcal{P} y \mathcal{P}' son dos particiones de vértices de G , entonces
603 $q(\mathcal{P}) \leq q(\mathcal{P}')$ cada vez que \mathcal{P}' refina a \mathcal{P} .

604 *Demostración.* Considere un grafo G sobre n vértices, los conjuntos $X, Y \subset V(G)$, y las particiones
605 $\mathcal{P}_X = \{X_1, \dots, X_k\}$ y $\mathcal{P}_Y = \{Y_1, \dots, Y_\ell\}$ de X e Y respectivamente. En primera instancia, se utiliza
606 la desigualdad (3) proveniente de Cauchy-Schwarz para probar que $q(\mathcal{P}_X, \mathcal{P}_Y) \geq q(X, Y)$. Para
607 esto, se desarrolla como sigue:

$$\begin{aligned} q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^k \sum_{j=1}^{\ell} q(X_i, Y_j) \\ &= \frac{1}{n^2} \sum_{i=1}^k \sum_{j=1}^{\ell} \frac{e(X_i, Y_j)^2}{|X_i||Y_j|} \\ &\stackrel{(3)}{\geq} \frac{1}{n^2} \frac{\left(\sum_{i=1}^k \sum_{j=1}^{\ell} e(X_i, Y_j) \right)^2}{\sum_{i=1}^k \sum_{j=1}^{\ell} |X_i||Y_j|} \\ &= \frac{1}{n^2} \frac{e(X, Y)^2}{\left(\sum_{i=1}^k |X_i| \right) \left(\sum_{j=1}^{\ell} |Y_j| \right)} \\ &= \frac{e(X, Y)^2}{n^2 |X||Y|} \\ &= q(X, Y). \end{aligned} \tag{29}$$

608 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 refinamiento de \mathcal{P} .
 609 Entonces, se utiliza el resultado probado previamente para completar el resultado:

$$q(\mathcal{P}) = \sum_{i=1}^k \sum_{j=1}^k q(V_i, V_j) \stackrel{(29)}{\leq} \sum_{i=1}^k \sum_{j=1}^k q(\mathcal{P}'_{V_i}, \mathcal{P}'_{V_j}).$$

610 □

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

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

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

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

$$|X_1| \geq \varepsilon|X| \quad \text{y} \quad |Y_1| \geq \varepsilon|Y|. \quad (30)$$

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

$$|\eta| > \varepsilon. \quad (31)$$

620 Por un lado, observe la siguiente descomposición,

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

621 De esta manera,

$$\sum_{i+j>2} e(X_i, Y_j) = e(X, Y) - e(X_1, Y_1). \quad (32)$$

622 Por otro lado, se tiene que,

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

623 Así,

$$\sum_{i+j>2} |X_i||Y_j| = |X||Y| - |X_1||Y_1|. \quad (33)$$

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

$$\begin{aligned}
q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^2 \sum_{j=1}^2 q(X_i, Y_j) \\
&= \sum_{i=1}^2 \sum_{j=1}^2 \frac{e(X_i, Y_j)^2}{n^2 |X_i| |Y_j|} \\
&= \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \sum_{i+j>2} \frac{e(X_i, Y_j)^2}{|X_i| |Y_j|} \right) \\
&\stackrel{(3)}{\geq} \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{\left(\sum_{i+j>2} e(X_i, Y_j) \right)^2}{\sum_{i+j>2} |X_i| |Y_j|} \right) \\
&\stackrel{(32) \text{ y } (33)}{=} \frac{1}{n^2} \left(\frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X| |Y| - |X_1| |Y_1|} \right). \tag{34}
\end{aligned}$$

625 Luego, por definici3n, note que

$$e(X_1, Y_1) = \frac{|X_1| |Y_1| e(X, Y)}{|X| |Y|} + \eta |X_1| |Y_1|. \tag{35}$$

626 Con esto, se contin3a el c3lculo desde la desigualdad (34) como sigue:

$$\begin{aligned}
n^2 q(\mathcal{P}_X, \mathcal{P}_Y) &\geq \frac{e(X_1, Y_1)^2}{|X_1| |Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X| |Y| - |X_1| |Y_1|} \\
&\stackrel{(35)}{=} \frac{1}{|X_1| |Y_1|} \left(\frac{|X_1| |Y_1| e(X, Y)}{|X| |Y|} + \eta |X_1| |Y_1| \right)^2 \\
&\quad + \frac{1}{|X| |Y| - |X_1| |Y_1|} \left(\frac{|X| |Y| - |X_1| |Y_1|}{|X| |Y|} e(X, Y) - \eta |X_1| |Y_1| \right)^2 \\
&= \frac{|X_1| |Y_1|}{|X|^2 |Y|^2} e(X, Y)^2 + 2 \frac{|X_1| |Y_1|}{|X| |Y|} \eta e(X, Y) + \eta^2 |X_1| |Y_1| \\
&\quad + \frac{|X| |Y| - |X_1| |Y_1|}{|X|^2 |Y|^2} e(X, Y)^2 - 2 \frac{|X_1| |Y_1|}{|X| |Y|} \eta e(X, Y) + \frac{\eta^2 |X_1|^2 |Y_1|^2}{|X| |Y| - |X_1| |Y_1|} \\
&= \frac{e(X, Y)^2}{|X| |Y|} + \eta^2 |X_1| |Y_1| \left(1 + \frac{|X_1| |Y_1|}{|X| |Y| - |X_1| |Y_1|} \right) \\
&\geq \frac{e(X, Y)^2}{|X| |Y|} + \eta^2 |X_1| |Y_1|. \tag{36}
\end{aligned}$$

627 Finalmente, utilizando las cotas (30) y (31), podemos concluir desde la desigualdad (36),

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

628

□

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

634 **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*
635 *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*

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

636 *Demostración.* Sea $\varepsilon > 0$ y $\mathcal{P} = \{V_1, \dots, V_k\}$ una partición no ε -regular del conjunto de n vértices
637 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,
638 existen los subconjuntos $A^{ij} \subset V_i$ y $A^{ji} \subset V_j$ testigos de su irregularidad. Observe que en cada
639 V_i se podrían encontrar a lo más k conjuntos no vacíos A^{ij} que evidencian la irregularidad de los
640 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
641 que cada Q_i es una partición resultante de dividir el conjunto V_i según la intersección de todos los
642 subconjuntos no vacíos A^{ij} que atestiguan la irregularidad de los pares (V_i, V_j) no ε -regular (ver
643 Figura 11). En consecuencia, $|Q_i| \leq 2^k$.

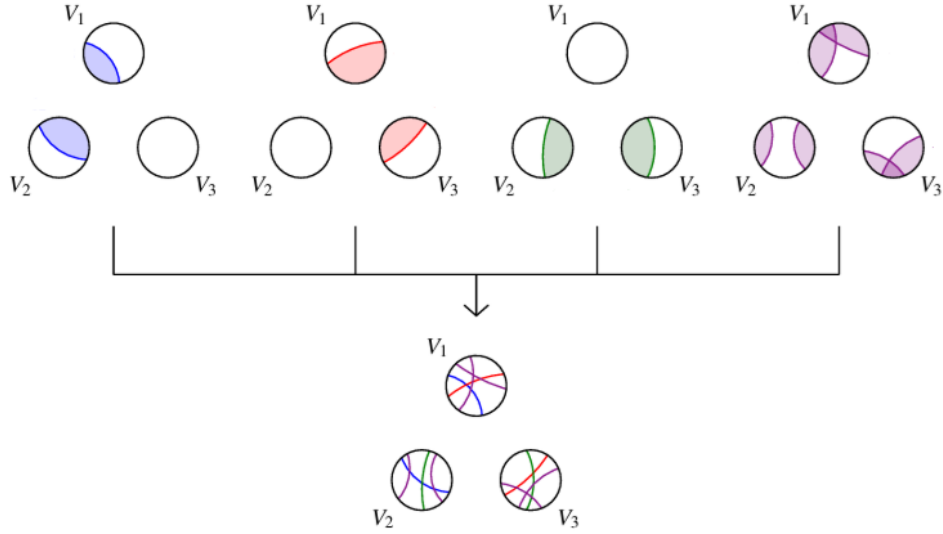


Figura 11: Ejemplo del refinamiento simultaneo por todos los subconjuntos que evidencian la irregularidad usando tres conjuntos de vértices [* referencia *](#).

644 Para simplicidad en la notación, se define $\Theta := \{(i, j) \in [k]^2 : (V_i, V_j) \text{ es } \varepsilon\text{-regular}\}$. Luego, como
645 la partición \mathcal{P} no es ε -regular, se cumple la desigualdad

$$\sum_{(i,j) \notin \Theta} \frac{|V_i||V_j|}{n^2} > \varepsilon. \quad (37)$$

646 Así, junto a los lemas probados previamente, se da prueba al resultado de la siguiente manera:

$$\begin{aligned} q(\mathcal{Q}) &= \sum_{(i,j) \in [k]^2} q(\mathcal{Q}_i, \mathcal{Q}_j) \\ &= \sum_{(i,j) \in \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) + \sum_{(i,j) \notin \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) \\ &\stackrel{\text{Lema 35}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} q(\{A^{ij}, V_i \setminus A^{ij}\}, \{A^{ji}, V_j \setminus A^{ji}\}) \\ &\stackrel{\text{Lema 36}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \left(q(V_i, V_j) + \varepsilon^4 \frac{|V_i||V_j|}{n^2} \right) \\ &= \sum_{(i,j) \in [k]^2} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \varepsilon^4 \frac{|V_i||V_j|}{n^2} \\ &\stackrel{(37)}{\geq} q(\mathcal{P}) + \varepsilon^5. \end{aligned}$$

647 * Cambiar por $>$ en la última línea y donde dice lema 33, cuando lo cambio se me descuadra :c
648 *

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

651 *Demostración del Teorema 32.* Dado $\varepsilon > 0$ y un grafo G , elegimos inicialmente la partición trivial
652 del conjunto de vértices $\mathcal{P} = \{V(G)\}$. Ahora, iterativamente (actualizando \mathcal{P}), aplicaremos el Lema
653 37 cada vez que la partición actual no sea ε -regular. Observe que por cada aplicación del Lema 37 se
654 consigue un aumento de al menos ε^5 en la energía, y como la energía de toda partición está acotada
655 superiormente por 1, el proceso iterativo terminará luego de a lo más ε^{-5} pasos. El resultado será
656 necesariamente una partición ε -regular debido a la cota de la energía.

657 Para una partición no ε -regular con k elementos, el Lema 37 encuentra un refinamiento de a lo
658 más $k2^k$ partes. Dicho refinamiento será producido en cada iteración del algoritmo de *argumento*
659 *de incremento de energía*, y la cantidad de partes producidas las acotaremos crudamente en cada
660 paso por $k2^k < 2^{2^k}$. Comenzando con la partición trivial de una parte, ejemplificaremos con las
661 tres primeras iteraciones del algoritmo para mostrar la cantidad de partes producidas en cada paso
662 tras aplicar el Lema 37.

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

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

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

665

666 Desde ahora en adelante, vamos a definir y consirar una *torre de altura k* de la siguiente manera:

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

667 Durante la demostración del Teorema 32 se utilizó una cota que podría parecer exagerada
668 para encontrar la cantidad de partes que devuelve el algoritmo implementado, por sobre todo,
669 considerando lo rápido que crece a medida que ε se hace más pequeño. Sorprendentemente, en 1997,
670 T. Gowers * referencia * prueba que tal límite inferior de partes es necesario. Más precisamente,
671 mostró que es posible encontrar una constante $c > 0$ tal que para todo suficientemente pequeño
672 $\varepsilon > 0$, existe un grafo sin partición ε -regular siempre que posea una cantidad menor que $\text{torre}(\lceil \varepsilon^{-c} \rceil)$
673 partes (ver Moshkovitz y Shapira * referencia * (2016) para una demostración corta).

674 Finalmente, se expone la forma de probar el Teorema 33. La idea de la demostración consiste
675 en modificar el algoritmo de la técnica de argumento de incremento de energía, de manera que en
676 cada iteración del refinamiento se logre obtener una equipartición. Este procedimiento conservará

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

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

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

4.2. Demostración espectral

En 2012, Terence Tao [* referencia *](#) 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 Frieze y Kannan [* referencia *](#), a quienes Tao les da el crédito en su publicación. Más adelante, en 2013, Cioba y Martin [* referencia *](#) 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 32. 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

$$\text{Tr}(T) = \sum_{i=1}^n \lambda_i^2 = 2e_G \leq n^2. \quad (38)$$

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

$$\lambda_i \leq \frac{n}{\sqrt{i}}, \quad \forall i \in [n]. \quad (39)$$

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

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

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

$$\sum_{L \leq j < F(L)} \lambda_j^2 \leq \varepsilon^3 n^2, \quad (40)$$

porque de lo contrario, se obtiene

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

que contradice la desigualdad (38).

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

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

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

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

la matriz F como la componente de *error*,

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

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

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

Usaremos los vectores propios $\mathbf{u}_1, \dots, \mathbf{u}_{L-1}$ de S para definir una partición de $V(G)$ como mostraremos a continuación. Consideramos el intervalo de \mathbb{R} de longitud $2\sqrt{(L/\varepsilon)} \cdot n^{-1/2}$ centrado en el origen y lo particionamos en $t = 2(L/\varepsilon)^2$ subintervalos J_1, \dots, J_t de longitud $(\varepsilon/L)^{3/2} n^{-1/2}$ cada uno. Luego, clasificamos los vértices $v \in V(G)$ según su valor $\mathbf{u}_i(v)$ de la siguiente manera:

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

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

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

730 Mostraremos que el conjunto excepcional V_0 es suficientemente pequeño. En efecto, observando
 731 que

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

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

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

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

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

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

736 y por consecuencia

$$\varepsilon^2 n^2 \geq \sum_{(i,j) \in \Sigma_F} |V_i| |V_j|. \quad (41)$$

737 Además, sea

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

738 y observe que

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

739 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, dado los
 740 subconjuntos $X \subset V_i$ e $Y \subset V_j$, note la siguiente descomposición:

$$\begin{aligned} \left| e(X, Y) - d_{ij} |X| |Y| \right| &= \left| \mathbf{v}_X^T A \mathbf{v}_Y - d_{ij} |X| |Y| \right| \\ &\leq \left| \mathbf{v}_X^T S \mathbf{v}_Y - d_{ij} |X| |Y| \right| + \left| \mathbf{v}_X^T F \mathbf{v}_Y \right| + \left| \mathbf{v}_X^T Q \mathbf{v}_Y \right|. \end{aligned} \quad (42)$$

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

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

$$|\mathbf{v}_X^T F \mathbf{v}_Y|^2 = \left| \sum_{(x,y) \in X \times Y} f_{xy} \right|^2 \stackrel{\text{DCS}}{\leq} \left(\sum_{(x,y) \in X \times Y} f_{xy}^2 \right) |X||Y| \leq \varepsilon^2 |V_i| |V_j| |X||Y| \leq \varepsilon^2 |V_i|^2 |V_j|^2. \quad (43)$$

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

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

745 Además, como $M \leq (\frac{2L^2}{\varepsilon^2})^L$, concluimos de la elección de $F(\cdot)$ que $F(L) \geq \frac{1}{\varepsilon^6} \left(\frac{2L^2}{\varepsilon^2} \right)^{4L} \geq \frac{1}{\varepsilon^6} M^4$.

746 Y así, cuando $(i, j) \neq \Sigma_Q$, se tiene

$$|\mathbf{v}_X^T Q \mathbf{v}_Y| \leq \frac{n^2}{\sqrt{F(L)}} \leq \frac{M^2 |V_i| |V_j|}{\varepsilon^2 \sqrt{F(L)}} \leq \varepsilon |V_i| |V_j|. \quad (44)$$

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

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

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

$$|\mathbf{v}_X^T S \mathbf{v}_Y - d_{ij} |V_i| |V_j|| \leq \sum_{(x,y) \in X \times Y} |s_{xy} - d_{ij}| \leq (s_{cd} - s_{ab}) |X||Y| \leq 4\varepsilon |X||Y|. \quad (45)$$

751 Utilizando las desigualdades (43), (44) y (45) en la expresión enunciada en (42) se concluye la
752 demostración del teorema. \square

753 4.3. Aplicaciones

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

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

763 Para el último paso se utilizará un resultado análogo a la propiedad $\text{COUNT}_p(\varepsilon)$ del Teorema
764 17, pero con el concepto de par ε -regular. Las aplicaciones que se estudiarán en esta tesis solo
765 necesitan el caso en que $H = K_3$, cual es conocido como lema de conteo de triángulos.

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

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

769 *Demostración.* Se realizará un proceso inductivo similar al visto en la demostración de la Propo-
770 sición 20 sobre la cantidad de aristas del grafo $K_3 = ([3], \{12, 23, 13\})$. Cuando el grafo no posea
771 aristas, entonces

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

772 También, como vimos en (28), recordamos que la condición de un par ε -regular es equivalente a
773 la propiedad $\text{DISC}_p(\varepsilon)$ en grafos bipartitos para algún $p \in (0, 1)$. Entonces, cuando el grafo presenta
774 solo una arista,

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

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

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

776 Defina $e^- = \varphi(1)\varphi(3)$, y T^- como el grafo correspondido a una copia etiquetada del grafo
777 $([3], \{12, 23\})$ en G bajo la aplicación inyectiva $\varphi : [3] \rightarrow V(T^-) \subset V$. Con esto, se desarrolla
778 inductivamente como sigue:

$$\begin{aligned}
\left| \{(x, y, z) \in X \times Y \times Z : xy, yz, xz \in E\} \right| &= \sum_{T^-} [\mathbb{1}_E(e^-) + d(X, Z) - d(X, Y)] \\
&= d(X, Y)d(Y, Z)d(X, Z)|X||Y||Z| \\
&\quad + \sum_{T^-} (\mathbb{1}_E(e^-) - d(X, Z)) \pm 2\varepsilon|X||Y||Z|. \quad (46)
\end{aligned}$$

En este punto, nos falta probar que el segundo sumando de la igualdad (46) se corresponde con un factor de error, para esto, sea T^* una copia del grafo singleton $\{2\}$ en G , y considere los siguientes conjuntos:

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

De esta manera, dada la equivalencia de la condición del par (X, Z) ε -regular con versión bipartita de la propiedad $\text{DISC}_{d(X, Z)}(\varepsilon)$ vista en (28), se consigue la siguiente desigualdad:

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

Finalmente, aplicando la última desigualdad en la ecuación (46) se prueba lo prometido. \square

En la demostración anterior solo fue necesario utilizar que los pares (X, Y) y (X, Z) son ε -regular, por lo que es interesante destacar que uno de los pares de conjuntos de vértices podría no ser necesariamente un par ε -regular para el que lema de conteo de triángulos funcione correctamente.

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

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

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

En las siguientes subsecciones se discutirán dos aplicaciones del método de regularidad para entregar dos demostraciones alternativas al teorema de Roth.

798 4.3.1. Eliminación de triángulos

799 El lema de eliminación de triángulos fue probado por los autores Ruzsa y Szemerédi * [referencia](#)
800 * en 1976, y es una de las primeras aplicaciones del método de regularidad. La intuición del lema dice
801 que todo grafo con *pocas* triángulos se puede convertir en un grafo libre de triángulos eliminando
802 *pocas* aristas. Formalmente,

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

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

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

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

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

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

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

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

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

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

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

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

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

820 Ahora, nos falta probar que el grafo limpio $G' = (V, E')$ es libre de triángulos. Para esto, con
821 nuestra elección de δ , buscaremos formular la siguiente contradicción: si existe un triángulo en
822 el grafo limpio G' , el lema de conteo de triángulos asegura que en realidad existen más de δn^3
823 triángulos. No obstante, como el grafo original posee a lo más δn^3 triángulos, se podrá concluir que
824 el grafo G' es libre de triángulos eliminando a lo más εn^2 aristas.

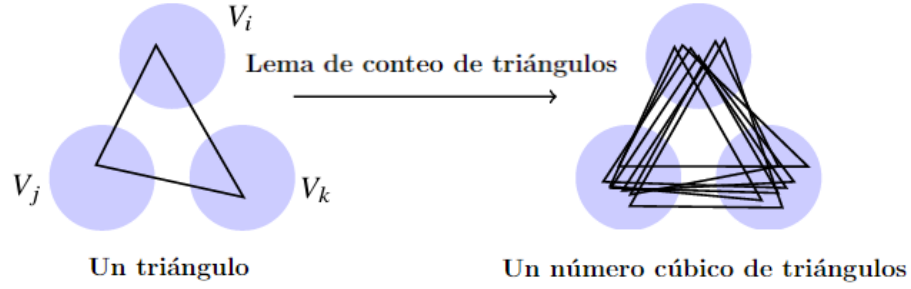


Figura 12: Esquema de la consecuencia del lema de conteo de triángulos ^{*}
referencia

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

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

828 Así, de la contradicción anterior, se determina que el grafo G es libre de triángulos eliminando
829 a lo más εn^2 aristas. \square

830 4.3.2. Emparejamiento inducido

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

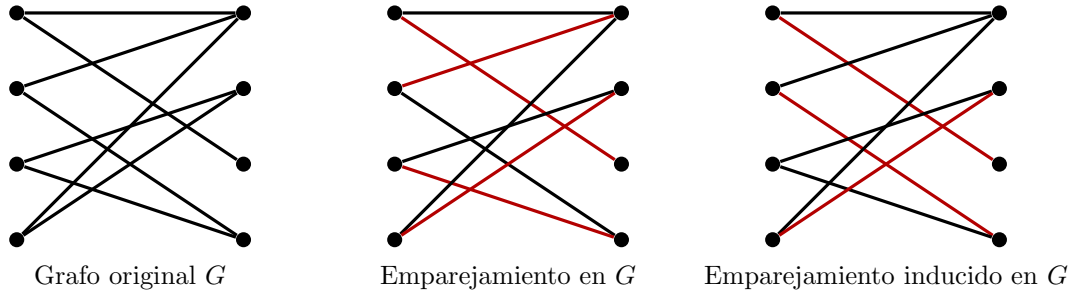


Figura 13: Ejemplo de un emparejamiento y un emparejamiento inducido.

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

Denotando por $V(R)$ al conjunto de vértices que componen las aristas de R , se define $U_i := V_i \cap V(R)$ como el subconjunto de vértices de R que comparte elementos 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 $U_i \subset V(R)$ que comparten una fracción suficientemente grande de vértices con V_i . Note que podemos obtener el conjunto U removiendo a lo más $\varepsilon_r n = \frac{\varepsilon}{10}n$ vértices de $V(R)$, pues

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

De esta manera, recordando que $|V(R)| \geq \varepsilon n$, se determina que $|U| > \varepsilon n - \frac{\varepsilon}{10}n = \frac{9}{10}\varepsilon n$. Además, como también $|R| \geq \frac{\varepsilon}{2}n$, debe existir al menos un vértice en U que sea parte de una arista en R .

860 Luego, dada la limpieza de G , dicha arista debe pertenecer a algún par $U_t \times U_\ell$ que satisfacen
861 $|U_k| \geq \varepsilon_r |V_k|$ y $|U_\ell| \geq \varepsilon_r |V_\ell|$, y son tales que su correspondiente par (V_t, V_ℓ) es ε_r -regular con
862 densidad $d(V_t, V_\ell) \geq 2\varepsilon_r$. Entonces, por regularidad,

$$d(U_t, U_\ell) = d(V_t, V_\ell) \pm \varepsilon_r \geq 2\varepsilon_r - \varepsilon_r = \varepsilon_r. \quad (47)$$

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

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

865 Sin embargo, la desigualdad (47) implica que

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

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

868 4.3.3. Teorema de Roth

869 Como hemos visto en el comienzo de la sección 4, el teorema de Roth es un caso particular del
870 teorema de Szemerédi, cual en un principio fue demostrado utilizando análisis de Fourier.

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

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

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

879 Una manera natural de construir un grafo de n vértices utilizando el conjunto A , es definir una
880 arista entre los vértices i y j si y solamente si $|i - j| \in A$. Sin embargo, con n elementos no podemos
881 describir una progresión aritmética de largo 3 por medio de relaciones entre las aristas. Por esto,
882 agregamos un poco de asimetría a la construcción. Para asegurar que las sumas siempre tengan
883 la posibilidad de estar en A , consideramos el grafo 3-partito $G = (V, E)$ con partición de vértices
884 $V = V_1 \cup V_2 \cup V_3$, en donde $V_1 = [n]$, $V_2 = [2n]$ y $V_3 = [3n]$ son tales que $6n \geq n_0$. Las aristas de G
885 son definidas de la siguiente manera:

- 886 1. Existe una arista desde $i \in V_1$ hasta $j \in V_2$ si y solamente si $j - i \in A$.

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

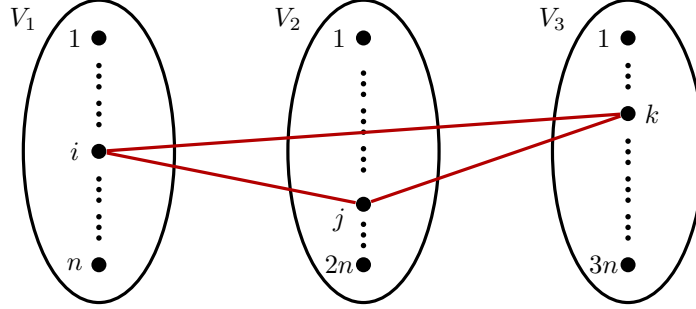


Figura 14: Esquema de relación entre un triángulo en G y una progresión aritmética de largo 3 en A .

889 * Pendiente cambiar el "no encontrar el n_0 " * Luego, la tupla $(i, j, k) \in V_1 \times V_2 \times V_3$ define un
890 triángulo en G si y solamente si $j - i \in A$, $k - j \in A$ y $\frac{k-i}{2} \in A$, o bien, $\{j - i, \frac{k-i}{2}, k - j\}$ es una
891 progresión aritmética de largo 3 en A con diferencia $\frac{k-2j+i}{2}$. También, diremos que un triángulo
892 $(i, j, k) \in V_1 \times V_2 \times V_3$ es trivial en G si para algún $a \in A$ se satisface que $j - i = \frac{k-i}{2} = k - j = a$.

893 Ahora, observando que cada triángulo trivial se puede identificar con el par $(i, a) \in V_1 \times A$, la
894 cantidad de triángulos triviales es exactamente $n|A| \geq \varepsilon n^2$. Además, por construcción, no existen
895 triángulos triviales que compartan una arista, por lo que no se puede eliminar dos triángulos triviales
896 removiendo solo una arista. En consecuencia, debemos eliminar al menos $\varepsilon n^2 = \frac{\varepsilon}{36}(6n)^2$ aristas para
897 hacer de G libre de triángulos.

898 Utilizando el lema eliminación de triángulos eligiendo $\varepsilon_0 = \frac{\varepsilon}{36}$, existen $\delta_0 > 0$ y $n'_0 \in \mathbb{N}$ tal que
899 el grafo G con $6n \geq n'_0$ vértices y a lo más $\delta_0(6n)^3$ triángulos, se convierte en libre de triángulos
900 eliminando a lo más $\varepsilon_0(6n)^2$ aristas. Entonces, estableciendo $\delta = 216\delta_0$, note que existen como
901 máximo $\delta n^3 - \varepsilon n^2$ triángulos no triviales. Sabiendo esto, aseguramos la existencia de un triángulo
902 no trivial cuando $n > \frac{\varepsilon+1}{\delta}$. En efecto,

$$n > \frac{\varepsilon+1}{\delta} \Rightarrow \delta n - \varepsilon > 1 \Rightarrow n^2(\delta n - \varepsilon) > 1.$$

903 Finalmente, el resultado queda demostrado tomando $n_0 > \max \left\{ \frac{n'_0}{6}, \frac{\varepsilon+1}{\delta} \right\}$. □

904 Para la segunda demostración del teorema de Roth, será necesario un teorema intermedio pro-
905 porcionado por los autores Ajtai y Szemerédi * referencia *. Para su prueba, se utiliza el Teorema
906 41.

907 **Teorema 43.** (Teorema de la esquina, Ajtai-Szemerédi) Para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal
908 que siempre que $n \geq n_0$, todo subconjunto $S \subset [n]^2$ con $|S| \geq \varepsilon n^2$ posee elementos de la forma
909 $\{(a, b), (a + d, b), (a, b + d)\}$ para algún $a, b, d \in \mathbb{N}$, con $d \neq 0$.

910 *Demostración.* Sea $\varepsilon > 0$, $n_0 \in \mathbb{N}$ suficientemente grande tal que $n \geq n_0$, y $S \subset [n]^2$ un subconjunto
911 con al menos εn^2 elementos. Vamos a construir un grafo bipartito $G = (U \cup W, E)$ con conjunto de
912 vértices $U = \{u_1, \dots, u_n\}$ y $W = \{w_1, \dots, w_n\}$ definiendo las aristas de la siguiente manera:

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

913 Interpretando a $[n]^2$ como una grilla bidimensional, se puede definir una relación entre pares de
 914 aristas en G en función de la distancia que abarca la suma de las coordenadas de sus respectivos
 915 pares en S . Esto es,

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

916 * Dibujito con 2 ejemplos de q . * Observe que para cada $2 \leq q \leq 2n$ se define un emparejamiento
 917 en G debido a que no existen aristas que compartan un vértice, por lo que las clases de equivalencia
 918 (cada una asociada a algún q) de la relación forman una partición de emparejamientos de E . En
 919 efecto, suponga que las aristas que pertenecen a la misma clase $u_i w_j$ y $u_k w_\ell$ comparten el vértice
 920 w_j . Entonces, como $i + j = k + j$, se determina que $u_i = u_k$ y se concluye que $u_i w_j$ y $u_k w_j$ son la
 921 misma arista.

922 Luego, como $e_G = |S| \geq \varepsilon n^2$, el Teorema 41 asegura que existe al menos un emparejamiento no
 923 inducido. Esto significa que en un emparejamiento que contiene las aristas con la relación $u_i w_j \sim$
 924 $u_k w_\ell$ 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) y (i, ℓ)
 925 son elementos de S que satisfacen

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

926 Finalmente, la demostración del teorema se consigue tomando $(i, \ell) = (a, b)$ para obtener $j = b + d$
 927 y $k = a + d$. * Poner dibujito de la esquina * \square

928 El resultado anterior otorga lo necesario para la segunda demostración del teorema de Roth
 929 utilizando el lema de regularidad de Szemerédi.

930 *Segunda demostración Teorema 42.* Dado $\varepsilon > 0$, escogemos $n_0 \in \mathbb{N}$ suficientemente grande. Para
 931 $n \geq n_0$, sea $A \subset [n]$ un conjunto que posee al menos εn elementos. Se define el siguiente conjunto:

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

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

936 5. Bibliografia

- 937 [1] Krivelevich, M., Sudakov, B. (2006). Pseudo-random Graphs. In Bolyai Society Mathema-
938 tical Studies (pp. 199–262). Springer Berlin Heidelberg.
- 939 [2] Chung, F. R. K., Graham, R. L., Wilson, R. M. (1989). Quasi-random graphs. *Combinato-*
940 *rica. An International Journal on Combinatorics and the Theory of Computing*.
- 941 [3] Chan, T. F. N., Král', D., Noel, J. A., Pehova, Y., Sharifzadeh, M., Volec, J. (2020). Charac-
942 terization of quasirandom permutations by a pattern sum. *Random Structures Algorithms*.
- 943 [4] Hàn, H., Kiwi, M., Pavez-Signé, M. (2021). Quasi-random words and limits of word sequen-
944 ces. *Journal Européen de Combinatoire [European Journal of Combinatorics]*.