



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 de las siguientes secciones. 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 1.4.

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  $f(n)/g(n) \rightarrow \infty$  cuando  $n \rightarrow \infty$ , se dice que  $f = \omega(g)$  y  $g = o(f)$ . Esto significa que  $f$  crece más rápido que  $g$  cuando  $n$  es lo suficientemente grande.
- Si existe una constante  $C > 0$  tal que  $|f(n)| \leq C|g(n)|$  para todo  $n$ , se dice que  $f = O(g)$  y  $g = \Omega(f)$ . Esto indica que  $f$  está acotada superiormente por un múltiplo constante de  $g$ .

En el caso de que  $C$  dependa de otro parámetro, se indica utilizando un subíndice para especificar esta dependencia. Por ejemplo,  $f = O_\ell(g)$ .

### 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** (o *nodos*, o *puntos*), y  $E \subseteq \binom{V}{2}$  el conjunto de **aristas** (o *líneas*). Dado un grafo  $G$ , se escribe  $V(G)$  como su conjunto de

41 vértices,  $E(G)$  como su conjunto de aristas, y  $e_G := |E(G)|$  como la cantidad de aristas presentes  
 42 en el grafo.

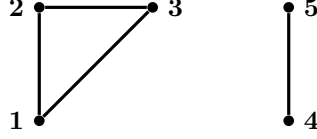


Figura 1: Ejemplo de un grafo con conjunto de vértices  $V = \{1, 2, 3, 4, 5\}$  y conjunto de aristas  $E = \{\{1, 2\}, \{2, 3\}, \{1, 3\}, \{4, 5\}\}$ .

43 Dado un grafo cualquiera  $G = (V, E)$ , se define su **complemento** como un nuevo grafo  $\overline{G} =$   
 44  $(V, \binom{V}{2} \setminus E)$ . Es decir, dos vértices están conectados por una arista en  $G$  si y solo si no lo están  
 45 en  $\overline{G}$ . Para  $u, v \in V$ , se dirá que  $u$  es **adyacente** a  $v$  (o viceversa) si y solamente si  $\{u, v\} \in E$ .  
 46 Si  $X, Y \subset V$  son dos subconjuntos no necesariamente disjuntos, se define el conjunto de tuplas que  
 47 forman una arista en  $G$  de la siguiente manera:

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

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

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

51 Se define el **grado** de un vértice  $u \in V$  con respecto a algún subconjunto de vértices  $Y \subseteq V$  de  
 52 la siguiente manera:

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

53 En particular, cuando  $Y = V$ ,

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

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

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

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

57 *Demostración.* Cada arista  $\{u, v\} \in E$  será contada dos veces en la suma, una contribución por  $u$ ,  
 58 y otra por  $v$ . □

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

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

Mostraremos que existe una relaci on intr ınseca entre los conceptos de grado y cogrado, cual ser a de utilidad en la secci on 14.

**Proposici n 2.** Sea  $G = (V, E)$  un grafo e  $Y \subset V$  un subconjunto de v ertices, entonces

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

*Demostraci n.* Utilizando las respectivas definiciones de grado y cogrado, el resultado se obtiene de seguir el siguiente c alculo:

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

□

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

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

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

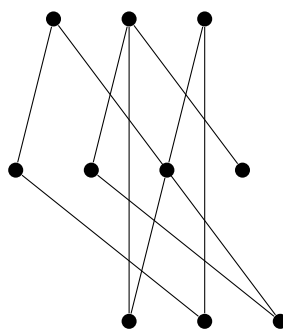


Figura 2: Ejemplo de un grafo 3-partito.

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

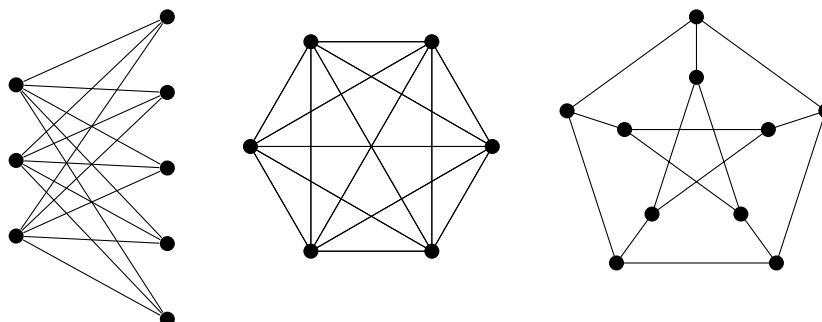


Figura 3: Ejemplo de los grafos especiales  $K_{3,5}$ ,  $K_6$  y 3-regular.

Otro concepto relevante en este trabajo son las diferentes nociones de rutas que se pueden encontrar siguiendo una secuencia de determinadas aristas de 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  $k \geq 2$  en la cual todos sus vértices  $v_i$  son distintos salvo el último, que coincide con el primero, es decir,  $v_k = v_0$ . Se denotará por  $C_k$  al ciclo de largo  $k$ .

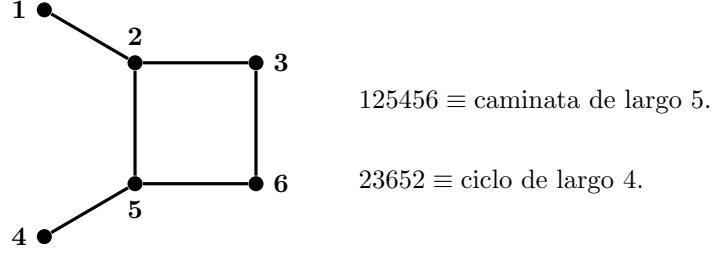


Figura 4: Ejemplo de una caminata y un ciclo.

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

90 Finalmente, se entenderá como una **copia etiquetada** de un grafo  $H$  en  $G$  a una aplicación  
91 inyectiva  $f : V(H) \rightarrow V(G)$  tal que  $\{f(u)f(v)\} \in E(G)$  cada vez que  $\{u, v\} \in E(H)$ . Se denota por  
92  $\binom{G}{H}$  al conjunto de copias etiquetadas de  $H$  en  $G$ .

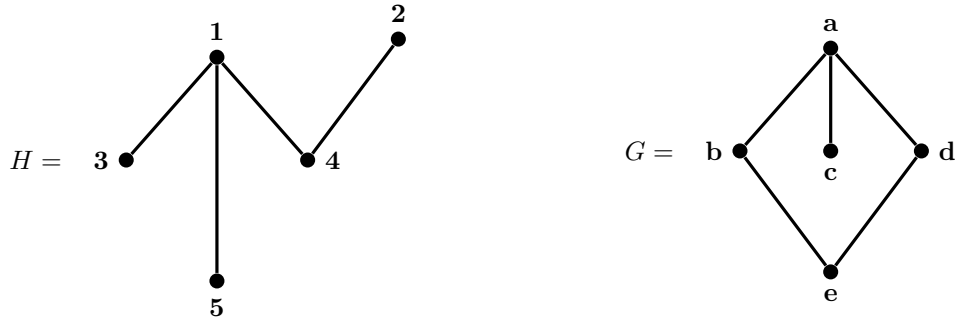


Figura 5: Ejemplo de una copia etiquetada de  $H$  en  $G$  mediante la función  $f : V(H) \rightarrow V(G)$  definida por  $f(1) = a$ ,  $f(2) = e$ ,  $f(3) = c$ ,  $f(4) = b$  y  $f(5) = d$ .

## 93 2.2. Álgebra lineal y teoría espectral de grafos

94 Se define  $\mathcal{M}_{n \times m}(\mathbb{R})$  como el conjunto de matrices reales de  $n$  filas y  $m$  columnas. Entonces,  
95 se denota por  $\mathbb{1}_X$  a la función indicatriz del conjunto  $X$ ,  $\mathbf{1} \in \mathcal{M}_{n \times 1}(\mathbb{R})$  al vector de 1-entradas,  
96  $J \in \mathcal{M}_{n \times n}(\mathbb{R})$  a la matriz de 1-entradas,  $I_n \in \mathcal{M}_{n \times n}(\mathbb{R})$  a la matriz identidad, y  $\mathbf{e}_i \in \mathcal{M}_{n \times 1}(\mathbb{R})$   
97 como el vector de la base canónica de  $\mathbb{R}^n$  con entrada 1 en la posición  $i$ . Además,  $\|\cdot\|$  y  $\langle \cdot, \cdot \rangle$   
98 representarán en todo momento la norma y producto interno usuales de  $\mathbb{R}^n$  ( $\mathbb{C}^n$ , según corresponda)  
99 respectivamente.

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

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

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

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

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

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

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

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

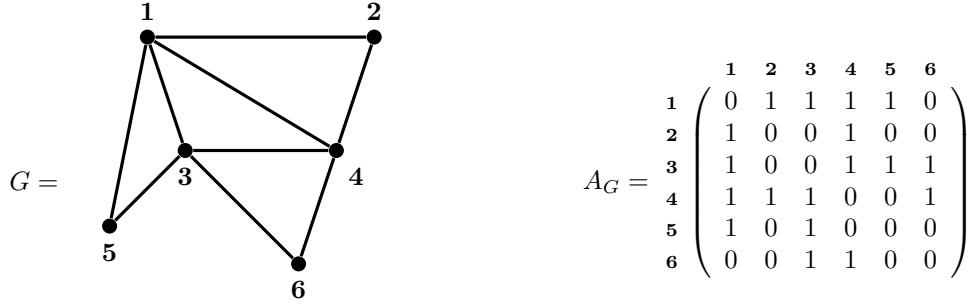


Figura 6: Ejemplo de representación mediante la matriz de adyacencia de un grafo.

112 Observe que la representación anterior resulta siempre en una matriz simétrica, es decir,  $A_G^T =$   
113  $A_G$ . Además, en todo grafo  $G = ([n], E)$  con matriz de adyacencia  $A$ , se puede obtener un vector  
114 con los grados de cada vértice del grafo aplicando el operador  $A$  al vector de 1-entradas:

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

115 .  
116 Otro aspecto interesante de la matriz de adyacencia que será de utilidad en la sección 4, es  
117 que nos permite reescribir la ecuación (4) en función de ella. Para ver esto, considere la matriz de  
118 adyacencia  $A$  del grafo  $G = ([n], E)$ , y los vértices  $i, j \in [n]$ . Luego, por la definición de  $A$ ,

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



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

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

□

Como consecuencia de la proposición anterior, en cualquier grafo  $G = ([n], E)$  con matriz de adyacencia  $A$ , se puede representar la cantidad total de caminatas cerradas de largo  $t$  en el grafo por medio de la traza,  $\text{Tr}(A^t) = \sum_{i=1}^n a_{ii}^{(t)}$ . Con esto, note que  $\text{Tr}(A^2) = 2e_G$ .

Por otro lado, para introducir algunos aspectos de la teoría espectral de grafos, recuerde que el 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  $\lambda \in \mathbb{C}$  si  $A\mathbf{v} = \lambda\mathbf{v}$ . Esto significa que  $\lambda$  es un valor propio si y solo si  $\lambda I_n - A$  es una matriz singular. Así, los valores propios vienen dados por las raíces del polinomio característico  $\det(xI_n - A)$ . En este trabajo, cuando se haga referencia a los valores y vectores propios de un grafo  $G$ , siempre será con respecto a su matriz de adyacencia  $A$ .

Por ejemplo, si  $G$  es un grafo  $d$ -regular, entonces con la igualdad (8) se puede deducir que  $d$  es el valor propio asociado al vector propio normalizado de 1-entradas de la matriz de adyacencia  $A_G$ .

**Proposición 5.** Sea  $A \in \mathcal{M}_{n \times n}(\mathbb{R})$  una matriz real simétrica, entonces todos sus valores propios son reales. Además, si dos vectores propios están asociados a distintos valores propios, entonces éstos son ortogonales. Más aún, el conjunto de todos los vectores propios definen una base ortonormal de  $\mathbb{R}^n$ .

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

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

Además, como  $A$  es simétrica,

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

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

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

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

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  $\mathbb{R}^n$ . □

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 6.** (*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  $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ :

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

178 Luego, de manera inductiva se obtiene que  $A^k = PD^kP^T$ . Esta propiedad resulta altamente útil  
 179 de cara al cálculo de caminatas de largo  $k$  entre dos vértices de un grafo. Más aún, la Proposición  
 180 7 y el Corolario 8 mostrarán que el número de caminatas cerradas en un grafo queda totalmente  
 181 determinado por los valores propios del mismo.

182 **Proposición 7.** *La traza de toda matriz simétrica  $A \in \mathcal{M}_{n \times n}(\mathbb{R})$  es igual a la suma de sus*  
 183 *autovalores.*

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

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

186 Con esto, se concluye utilizando la descomposición espectral.

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

187

□

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

190 **Corolario 8.** *Sea  $A \in \mathcal{M}_{n \times n}(\mathbb{R})$  una matriz simétrica y  $\lambda_1, \dots, \lambda_n$  sus autovalores, entonces se*  
 191 *cumple  $\text{Tr}(A^k) = \sum_{i=1}^n \lambda_i^k$ .*

192 *Demostración.* El resultado sigue de utilizar que la traza de la multiplicación de dos matrices es  
 193 invariante bajo orden del producto,

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

194

□

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 14, 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,

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

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

**Teorema 9.** (*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)

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

(ii)

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

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

**Proposición 10.** El primer autovalor de la matriz de adyacencia de un grafo es al menos el promedio de los grados. En particular, cuando el grafo es  $d$ -regular, el primer autovalor coincide con  $d$ .

*Demostración.* Considerando  $A$  como la matriz de adyacencia del grafo  $G = ([n], E)$ , se desarrolla en función del Teorema 9:

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

Adicionalmente, con apoyo de la igualdad (8), usando la cota anterior se concluye que si  $G$  es  $d$ -regular, entonces  $\lambda_1 = d$ .  $\square$

### 2.3. Grafos aleatorios

El objetivo de la presente sección es proporcionar un breve contexto teórico para abordar más adelante la noción y propiedades de los grafos *cuasi-aleatorios*. Para ello, se asumirán los conceptos básicos asociados a la teoría de la probabilidad.

De manera intuitiva, se podría pensar en primera instancia en un *grafo aleatorio* de  $n$  vértices como el resultado de seleccionar aleatoriamente un subconjunto de aristas de  $K_n$ . La manera de seleccionar dicho subconjunto de aristas determina el tipo de modelo de grafo aleatorio. El modelo

más popular, propuesto por Edgar Gilbert en 1959 [\\* referencia \\*](#), contempla una estructura que se desarrolla algorítmicamente agregando una arista a la vez, en la que comenzando con el grafo vacío  $\overline{K}_n$ , se decide iterativamente si existe una arista entre cada par de vértices del grafo con probabilidad independiente  $p$ .

Cada repetición del proceso anterior genera un nuevo grafo de  $n$  vértices. El conjunto de todos esos grafos constituye un espacio de probabilidad, al que denotaremos por  $G(n, p)$  y llamaremos modelo binomial. Entonces, considerando  $\mathcal{G}^n$  como el conjunto de todos los grafos de  $n$  vértices, se define formalmente.

**Definición 11.** (*Modelo binomial*) Sea  $p \in (0, 1)$ . Se define  $G(n, p)$  como el espacio de probabilidad  $(\mathcal{G}^n, \mathcal{P}(\mathcal{G}^n), \mathbb{P})$ , con

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

Se adopta el modelo binomial como referencia para trabajar más adelante con el concepto de un grafo *cuasi-aleatorio*. La mayoría de la literatura sobre grafos aleatorios dedica su contenido a los casos en que  $p = p(n) \rightarrow 0$  cuando  $n \rightarrow \infty$ . En esta tesis, se considera  $p \in (0, 1)$  constante.

A continuación, se presenta sin demostración una de las cotas más importantes de la teoría de la probabilidad, la *desigualdad de Chernov*. Existen diferentes formas de expresar la desigualdad de Chernov, y cada una varía levemente en su hipótesis. En esta tesis se estudia el resultado para el caso en que cada variable aleatoria solo toma valores 0 o 1, como se plantea en [\\* referencia \\*](#) en la ecuación (2.12) de la observación 2.5.

**Teorema 12.** (*Desigualdad de Chernov*) Sean  $X_1, \dots, X_N$  variables aleatorias independientes tales 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 satisface

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

Por otro lado, diremos que  $\mathcal{P}_n \subset \mathcal{G}^n$  es una propiedad de grafos si es cerrada bajo isomorfismos de grafos. Más aún,  $G$  satisface la propiedad  $\mathcal{P}_n$  **con alta probabilidad** si  $\mathbb{P}(G \in \mathcal{P}_n) \rightarrow 1$  cuando  $n \rightarrow \infty$ . Dicho esto, se probará que todo grafo en  $G(n, p)$  posee casi seguramente una distribución uniforme de aristas a lo largo de todo el grafo.

**Proposición 13.** Sea  $\varepsilon > 0$  y  $p \in (0, 1)$ . Si  $G \sim G(n, p)$ , entonces satisface con alta probabilidad la siguiente propiedad:

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

*Demostración.* Al pendiente de resolver el cálculo de la esperanza □

Sea  $p \in (0, 1)$ ,  $G \in G(n, p)$  y  $X, Y \subset V(G)$ . Entonces la cantidad de aristas esperadas entre  $X$  e  $Y$  es  $p|X||Y|$ . Enumere las aristas del grafo bipartito completo  $K_{|X|, |Y|}$  desde 1 hasta  $|X||Y|$ , y para cada arista  $1 \leq i \leq |X||Y|$ , considere  $Z_i$  como la siguiente variable aleatoria indicadora:

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

251

252 Así, la variable aleatoria  $Z = \sum_{i=1}^{|X||Y|} Z_i$  cuenta el número total de aristas que existen entre  $X$   
 253 e  $Y$ . Finalmente,

$$\mathbb{E}[Z] = \sum_{i=1}^{|X||Y|} \mathbb{E}[Z_i] = \sum_{i=1}^{|X||Y|} \mathbb{P}(Z_i = 1) = \sum_{i=1}^{|X||Y|} p = p|X||Y|.$$

### 254 3. Cuasi-aleatoriedad

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

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

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

275 **Definición 14.** Sea  $p \in (0, 1)$  y  $(G_n)_{n \rightarrow \infty}$  una secuencia de grafos, en donde cada  $G_n$  posee  $n$   
 276 vértices. Entonces el grafo  $G_n$  es **cuasi-aleatorio** si en todo par de subconjuntos  $X, Y \subset V(G_n)$   
 277 se encuentra una distribución de aristas similar, es decir,

$$e(X, Y) = p|X||Y| + o(n^2). \quad (12)$$

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

#### 283 3.1. Teorema de Chung, Graham y Wilson

284 Una contribución revolucionaria en la teoría de grafos cuasi-aleatorios fue en 1989 por Fan  
 285 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 14.

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

**Teorema 15.** *(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:*

$\text{DISC}_p$ : Para todo  $X, Y \subseteq V(G_n)$ ,

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

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

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

$\text{COUNT}_p$ : Para cada grafo  $H$ , la cantidad de copias etiquetadas de  $H$  en  $G_n$  está dada por

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

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

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

$\text{CODEG}_p$ :

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

$\text{EIG}_p$ : Si  $\lambda_1 \geq \dots \geq \lambda_n$  son los autovalores de la matriz de adyacencia de  $G_n$ , entonces

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

Para una comprensión e intuición inicial de cada propiedad del Teorema 15, se ha utilizado notación asintótica en su enunciado. Sin embargo, con dicha formulación no queda del todo claro las dependencias cuantificadas de los errores en las implicancias entre dos propiedades. Entonces, se replantea equivalentemente cada propiedad con una versión cuantitativa asociando algún parámetro de error  $\varepsilon$  para todo grafo específico  $G$  con un conjunto de vértices suficientemente grande. Por ejemplo, bajo los supuestos del Teorema 15, asuma que la sucesión de grafos  $(G_n)_{n \rightarrow \infty}$  satisface  $\text{DISC}_p$ . Luego, la versión equivalente establece que para todo  $\varepsilon > 0$ , existe  $n_0 \in \mathbb{N}$  tal que el grafo  $G$  sobre  $n \geq n_0$  vértices satisface

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

De manera general, diremos que una secuencia de grafos  $(G_n)_{n \rightarrow \infty}$  con  $|V(G_n)| = n$  satisface la propiedad  $P_{x_1, \dots, x_k}$ <sup>1</sup> si para cada elección de  $\varepsilon > 0$ , existe algún  $n_0 \in \mathbb{N}$  tal que el grafo  $G$  con  $n \geq n_0$  vértices satisface  $P_{x_1, \dots, x_k}(\varepsilon)$ . Más aún, se dirá que la propiedad  $Q_{y_1, \dots, y_\ell}$  implica la propiedad  $P_{x_1, \dots, x_k}$  si y solamente si  $P_{x_1, \dots, x_k}(\varepsilon)$  implica  $Q_{y_1, \dots, y_\ell}$ . Es decir, para todo  $\varepsilon > 0$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  con  $n \geq n_0$  vértices cumple con  $Q_{y_1, \dots, y_\ell}(\delta)$  cada vez que satisfaga la propiedad  $P_{x_1, \dots, x_k}(\varepsilon)$ .

Se desarrollará la demostración formal del teorema de Chung, Graham y Wilson utilizando notación  $\varepsilon$ - $\delta$ , mostrando que cada par de propiedades  $P_{x_1, \dots, x_k}$  y  $Q_{y_1, \dots, y_\ell}$  son equivalentes entre sí con un cambio polinomial en el error. Esto es,  $P_{x_1, \dots, x_k}(\varepsilon) \Rightarrow Q_{y_1, \dots, y_\ell}(C\varepsilon^c)$  para algún par de constantes  $C, c > 0$ .

## Demostración Teorema de Chung, Graham y Wilson

La demostración del Teorema 15 fue descompuesta en ocho proposiciones, las cuales mostrarán la equivalencia entre todas las propiedades conforme al siguiente esquema:

$$\begin{array}{ccccc}
 \text{DISC}'_p & \xRightarrow{\text{Prop. 18.}} & \text{COUNT}_p & \xRightarrow{\text{Prop. 19.}} & \text{COUNT}_{C4,p} & \xRightarrow{\text{Prop. 22. y 23.}} & \text{EIG}_p \\
 \Downarrow \text{Prop. 16 y 17.} & & & & \Downarrow \text{Prop. 20.} & & \\
 \text{DISC}_p & & \xleftarrow{\text{Prop. 21.}} & & \text{CODEG}_p & & 
 \end{array} \quad (13)$$

Así, damos paso a la demostración de cada proposición considerada en el esquema 13.

**Proposición 16.** *Para todo  $\varepsilon > 0$  y  $p \in (0, 1)$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  sobre  $n \geq n_0$  vértices satisface  $\text{DISC}'_p(\varepsilon)$  cada vez que cumpla con  $\text{DISC}_p(\delta)$ . En particular,*

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

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

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

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

**Proposición 17.** *Para todo  $\varepsilon > 0$  y  $p \in (0, 1)$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  sobre  $n \geq n_0$  vértices satisface  $\text{DISC}_p(\varepsilon)$  cada vez que cumpla con  $\text{DISC}'_p(\delta)$ . En particular,*

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

---

<sup>1</sup>Los parámetros  $x_1, \dots, x_k$  pueden ser de distinta naturaleza, dependiendo de la propiedad simbolizada. En las propiedades del Teorema 15 se utiliza  $k = 1$  con  $x_1 = p$  salvo en la propiedad  $\text{COUNT}_{C4,p}$ , en donde  $k = 2$ .



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

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

$$e(X, Y) = e(X \cup Y) + e(X \cap Y) - e(X \setminus Y) - e(Y \setminus X). \quad (14)$$

335 Observe que con esta configuración, el conteo de las aristas entre  $X$  e  $Y$  es doble cuando los  
 336 vértices que componen las aristas pertenecen a  $X \cap Y$ . Luego, se utiliza la propiedad  $\text{DISC}'_p(\delta)$   
 337 sobre la identidad (14) para conseguir el resultado.

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

338 □

339 **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  
 340  $n \geq n_0$  vértices satisface  $\text{COUNT}_p(\varepsilon)$  cada vez que cumpla con  $\text{DISC}'_p(\delta)$ . En otras palabras,

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

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

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

$$\left| \binom{G}{F} \right| = p^{e_F} n^\ell \pm 4e_F \delta n^\ell. \quad (15)$$

346 Una vez probada la ecuación (15), el resultado seguirá de tomar  $F = H$  y la elección de  $\delta$  para  
 347 conseguir las siguientes desigualdades:

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

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

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

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

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

351 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  
 352 con  $e_F > 1$  aristas y asuma que se satisface la ecuación (15) en cualquier grafo con una cantidad  
 353 de aristas menor que  $e_F$ . Para desarrollar la inducción, suponga que  $\{i, j\} \in E(F)$  y considere la  
 354 siguiente notación:

- 355 i)  $F^-$  corresponde es el grafo producido por eliminar la arista  $\{i, j\}$  de  $F$ .
- 356 ii)  $F^*$  es el resultado de eliminar los vértices de la arista  $\{i, j\}$  en  $F$ .

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

$$\begin{aligned}
 \left| \binom{G}{F} \right| &= \sum_{T^- \in \binom{G}{F^-}} \mathbb{1}_E(e_{T^-}) \\
 &= \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) + p - p] \\
 &= \sum_{T^- \in \binom{G}{F^-}} p + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
 &= p \left| \binom{G}{F^-} \right| + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \\
 &\stackrel{(15)}{=} p^{e_F} n^\ell + \sum_{T^- \in \binom{G}{F^-}} [\mathbb{1}_E(e_{T^-}) - p] \pm 4(e_F - 1)\delta n^\ell.
 \end{aligned} \tag{16}$$

361 Ahora, es suficiente probar que el segundo sumando de la desigualdad (16) es pequeño. Para  
 362 esto, considere  $T^*$  una copia de  $F^*$ , y denote por  $F_i^*$  y  $F_j^*$  a los grafos resultantes de eliminar de  
 363  $F^-$  los vértices  $j$  e  $i$  respectivamente. Defina los siguientes conjuntos:

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

365 Los conjuntos anteriores, por construcción, son tales que para cada tupla  $(a, b) \in A_i^{T^*} \times A_j^{T^*}$   
 366 añadida a  $T^*$  se obtiene una copia de  $F^-$ . Así, reescribiendo el segundo sumando de la igualdad  
 367 (16) 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  $\delta$  y  $F = H$  se obtiene el resultado.  $\square$

**Proposición 19.** Para todo  $\varepsilon > 0$  y  $p \in (0, 1)$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  sobre  $n \geq n_0$  vértices satisface  $\text{COUNT}_{C_4, p}(\varepsilon)$  cada vez que cumpla con  $\text{COUNT}_p(\delta)$ . En otras palabras,

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

*Demostración.* Se trata de un caso particular de  $\text{COUNT}_p$ , en donde  $H = C_4$  y  $\delta < \varepsilon$ .  $\square$

**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  $n \geq n_0$  vértices y  $e_G = \frac{pn^2}{2} \pm \delta n^2$  aristas satisface  $\text{CODEG}_p(\varepsilon)$  cada vez que cumpla con  $\text{COUNT}_{C_4, p}(\delta)$ . En particular,

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

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

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

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

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

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

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

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

383 □

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

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

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

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

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

391 Luego, para todo  $X, Y \in V(G)$ , se reescribe la expresión de la propiedad  $\text{DISC}_p$  de forma  
392 conveniente para posteriormente utilizar la desigualdad de Cauchy-Schwarz.

$$\left| e(X, Y) - p|X||Y| \right| = \left| \sum_{x \in X} (\deg(x; Y) - p|Y|) \right| \stackrel{\text{DCS}}{\leq} n^{1/2} \left( \sum_{x \in X} (\deg(x; Y) - p|Y|)^2 \right)^{1/2}. \quad (17)$$

393 En la desigualdad anterior se ha conseguido que el argumento de la suma sea siempre no negativo,  
394 lo que permite extender su dominio de  $X$  a  $V(G)$ . De esta manera, usando a la cota proveniente de  
395 la concentración de los grados en los vértices de  $G$ , se prueba el resultado continuando desde (17):

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

396

□

397 **Proposición 22.** Para todo  $\varepsilon > 0$  y  $p \in (0, 1)$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  sobre  
398  $n \geq n_0$  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  $\text{EIG}_p(\delta)$ .  
399 En particular,

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

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

403 Recuerde que la cantidad de copias etiquetadas de caminatas cerradas de largo 4, que no son  $C_4$ ,  
404 en  $G$  se encuentran dentro de un error de a lo más  $\delta n^4$  con respecto al número de copias etiquetadas  
405 de  $C_4$  en  $G$ . Con esto, junto al Lema 4 y el Corolario 8 se obtiene lo siguiente:

$$\left| \binom{G}{C_4} \right| = \text{Tr}(A^4) \pm \delta n^4 = \sum_{i=1}^n \lambda_i^4 \pm \delta n^4 = \lambda_1^4 + \sum_{i=2}^n \lambda_i^4 \pm \delta n^4. \quad (18)$$

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

$$\sum_{i=2}^n \lambda_i^4 \leq \max_{i \neq 1} \lambda_i^2 \sum_{i=1}^n \lambda_i^2 \leq \delta n^2 \text{Tr}(A^2) \leq \delta n^2 (pn^2 + 2\delta n^2) \leq 3\delta n^4. \quad (19)$$

Finalmente, se concluirá tras usar la propiedad  $\text{EIG}_p(\delta)$  sobre el primer autovalor y la cota mostrada en (19). Entonces, continuando desde la ecuación (18),

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

□

**Proposición 23.** Para todo  $\varepsilon > 0$  y  $p \in (0, 1)$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tales que el grafo  $G$  sobre  $n \geq n_0$  vértices y  $e_G = \frac{pn^2}{2} \pm \delta n^2$  aristas satisface  $\text{EIG}_p(\varepsilon)$  cada vez que cumpla la propiedad  $\text{COUNT}_{C_4, p}(\delta)$ . Es decir,

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

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

En lo que respecta al primer autovalor, sabemos por un lado que éste es al menos el promedio de los grados gracias al Lema 10. Es decir,

$$\lambda_1 \geq \frac{\sum_{x \in V(G)} \deg(x)}{n} = \frac{2e_G}{n} = \frac{2}{n} \left( \frac{pn^2}{2} \pm \delta n^2 \right) \geq pn - 2\delta n. \quad (20)$$

Por otro lado, mediante el Lema 4, el Corolario 8 y la propiedad  $\text{COUNT}_{C_4, p}(\delta)$ ,

$$\lambda_1^4 \leq \sum_{i=1}^n \lambda_i^4 = \text{Tr}(A^4) = \left| \binom{G}{C_4} \right| \pm \delta n^4 \leq p^4 n^4 + 2\delta n^4. \quad (21)$$

La desigualdad (21) implica que  $\lambda_1 \leq pn + (2\delta)^{1/4}n$ , y en combinación con la cota vista en (20), 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  $\max_{i \neq 1} |\lambda_i| \leq (4\delta)^{1/4}n$ . □

### 3.2. Aspectos adicionales del teorema de CGW

La noción inicial presentada de un grafo cuasi-aleatorio por distribución de aristas según la Definición 14 contempla verificar si todo par de subconjuntos de vértices del grafo satisfacen la condición  $\text{DISC}_p$  para determinar la cuasi-aleatoriedad. En otras palabras, se requiere verificar un número exponencial de subconjuntos. Por esto, resulta sorprendente que tal propiedad sea equivalente a todas las otras (salvo  $\text{DISC}'_p$ ), 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  $\text{COUNT}_p$ ; que dice que el número etiquetadas de cualquier grafo  $F$  de tamaño fijo en  $G = ([n], E)$  es aproximadamente el esperado de los grafos aleatorios  $G(n, p)$ .

\* **Cambiar el enfoque de este párrafo.** \* Una pregunta natural al observar la propiedad  $\text{COUNT}_{C_4,p}$  es: ¿Podemos debilitar la condición para un conteo esperado de copias etiquetadas de  $K_3$ ? La respuesta es no, de hecho, la propiedad puede ser extendida a  $\text{COUNT}_{C_{2t},p}$  con  $t \geq 2$ . Es decir, para el grafo  $G$ ,

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

Se expone un bosquejo de la demostración de la extensión de  $\text{COUNT}_{C_4,p}$  a la propiedad  $\text{COUNT}_{C_{2t},p}$ .

**Proposición 24.** Sea  $p \in (0, 1)$  y  $(G_n)_{n \rightarrow \infty}$  una secuencia de grafos con  $|V(G_n)| = n$  vértices y  $e_{G_n} = (p + o(1)) \binom{n}{2}$  aristas, entonces las propiedades  $\text{COUNT}_{C_{2t},p}$  y  $\text{EIG}_p$  son equivalentes.

*Demostración.* Este resultado es una consecuencia directa de las demostraciones de la Proposición 22 y 23 tras notar el siguiente par de observaciones. En primer lugar, la cantidad de copias etiquetadas caminatas cerradas de largo  $2t$  que no son  $C_{2t}$  en  $G_n$  están dentro de un error  $O(n^{2t-1})$ , es decir,

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

También, se debe modificar la cota presentada en la ecuación (19) como sigue:

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

Con estas observaciones el resultado queda demostrado.  $\square$

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



454 A continuación, se expone la construcción de un contraejemplo de un grafo que posee la cantidad  
 455 de copias etiquetadas esperadas de  $K_3$ , pero no cumple las condiciones para ser casi-aleatorio. La  
 456 idea de la construcción consiste en la combinación de dos grafos, uno con una cantidad mayor que  
 457 la esperada de copias etiquetadas de  $K_3$ , y otro con una cantidad menor. Consideramos entonces  
 458 independientemente los grafos completos  $K_{n_1}$  y  $K_{n_2, n_2}$  tales que su unión disjunta forma el grafo  
 459  $G = K_{n_1} \cup K_{n_2, n_2}$  con  $n_1 + 2n_2 = n$  vértices.

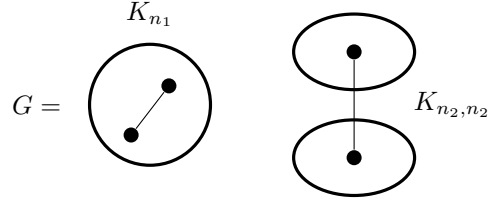


Figura 7: Esquema de la configuración del grafo usado como contraejemplo.  
 Aquí,  $\bullet\text{---}\bullet$  representa las aristas permitidas dentro del grafo  $G$ .

460 Sobre  $K_{n_1}$  y  $K_{n_2, n_2}$ , observe que la cantidad de sus aristas y copias etiquetadas de  $K_3$  son:

$$\begin{aligned} e_{K_{n_1}} &\approx \frac{n_1^2}{2} & , & \quad \left| \binom{K_{n_1}}{K_3} \right| \approx n_1^3 \quad , \\ e_{K_{n_2, n_2}} &\approx \frac{(n - n_1)^2}{4} & , & \quad \left| \binom{K_{n_2, n_2}}{K_3} \right| = 0. \end{aligned}$$

461

462 Ahora, se encontrará el parámetro  $p \in (0, 1)$  de manera tal que el grafo  $G$  posea la cantidad  
 463 esperada de aristas y copias etiquetadas de  $K_3$ , acorde a un grafo aleatorio binomial con densidad  
 464  $p$ . Para ello, se plantea el siguiente sistema de ecuaciones:

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

465

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

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

468

469 Como también,

$$\left| \binom{G}{K_3} \right| = \left( \frac{n}{3} \right)^3 + o(n^3) = \left( \frac{1}{3} \right)^3 n^3 + o(n^3).$$

470

471 Sin embargo, el grafo  $G$  no es casi-aleatorio debido a que no existen aristas entre  $K_{n_1}$  y  $K_{n_2, n_2}$   
 472 ni dentro de los conjuntos que conforman a  $K_{n_2, n_2}$ .

473 Finalmente, se explora un caso siempre interesante de estudio, ya que simplifica varios cálculos  
 474 y surge de manera recurrente en la vida cotidiana: un grafo  $d$ -regular. En nuestro contexto, se verá  
 475 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  
 476 con  $\text{EIG}_{\frac{d}{n}}$ . Dicha equivalencia nace como una consecuencia del siguiente teorema.

477 **Teorema 25.** (*Expander Mixing Lemma*) Sea  $G = ([n], E)$  un grafo  $d$ -regular, y  $d = \lambda_1 \geq \lambda_2 \geq$   
 478  $\dots \geq \lambda_n$  los valores propios asociados a la matriz de adyacencia  $A$  de  $G$ . Se denota:

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

479

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

$$\left| e(X, Y) - \frac{d}{n} |X| |Y| \right| \leq \lambda \sqrt{|X| |Y| \left( 1 - \frac{|X|}{n} \right) \left( 1 - \frac{|Y|}{n} \right)}. \quad (22)$$

481 *Demostración.* Sea  $\mathcal{B} = \{v_1, \dots, v_n\}$  la base ortonormal de  $\mathbb{R}^n$  compuesta por los vectores propios  
 482 de  $A$ . Utilizando la descomposición espectral, se denota

$$A_1 = \lambda_1 v_1 v_1^T \text{ y } \Delta = \sum_{i=2}^n \lambda_i v_i v_i^T,$$

483

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

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

$$e(X, Y) = v_X^T A v_Y = v_X^T A_1 v_Y + v_X^T \Delta v_Y. \quad (23)$$

486

487 De la ecuación anterior se espera que el primer sumando sea el término principal, mientras que  
 488 el segundo el factor de error. Para esto, se representan los vectores  $v_X$  y  $v_Y$  según la base  $\mathcal{B}$ . Es  
 489 decir,

$$v_X = \sum_{i=1}^n \alpha_i v_i \text{ y } v_Y = \sum_{i=1}^n \beta_i v_i,$$

490

491 donde  $\alpha_i = v_X^T v_i$  y  $\beta_i = v_Y^T v_i$ . Con esto, se calcula:

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

492

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

494 Ahora, se estudian los sumandos de la igualdad (23) por separado. Por un lado,

$$\begin{aligned}
v_X^T A_1 v_Y &= \left( \sum_{i=1}^n \alpha_i v_i \right)^T (\lambda_1 v_1 v_1^T) \left( \sum_{j=1}^n \beta_j v_j \right) \\
&= \lambda_1 \left( \sum_{i=1}^n \alpha_i v_i^T \right) (v_1 v_1^T) \left( \sum_{j=1}^n \beta_j v_j \right) \\
&= \lambda_1 \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j (v_i^T v_1) (v_1^T v_j) \\
&= \lambda_1 \alpha_1 \beta_1.
\end{aligned} \tag{24}$$

495

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

$$v_X^T \Delta v_Y = \left( \sum_{i=1}^n \alpha_i v_i \right)^T \left( \sum_{j=2}^n \lambda_j v_j v_j^T \right) \left( \sum_{k=1}^n \beta_k v_k \right) = \sum_{i=1}^n \lambda_i \alpha_i \beta_i. \tag{25}$$

497

498 Luego, dado que  $G$  es un grafo  $d$ -regular,  $\lambda_1 = d$  y  $v_1 = \frac{1}{\sqrt{n}}(1, \dots, 1)^T$  son valor y vector propio  
499 respectivamente de  $A$ . En consecuencia,

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

500

501 Así, la ecuación (24) resulta en  $v_X^T A_1 v_Y = \frac{d}{n} |X||Y|$ .

502 Para el término de error, recordando la definición de  $\lambda$ , se desarrolla el valor absoluto de la  
 503 ecuación (25) usando la desigualdad de Cauchy-Schwarz.

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

504

505 Finalmente, el resultado se obtiene directamente de tomar el valor absoluto de la ecuación (23)  
 506 de la siguiente manera:

$$|e(X, Y) - v_X^T A_1 v_Y| = |v_X^T \Delta v_Y|.$$

507

□

508 El teorema anterior permite asegurar que todo grafo  $d$ -regular  $G = ([n], E)$  con un conjunto de  
 509 vértices suficientemente grande que satisface la propiedad  $\text{EIG}_{\frac{d}{n}}(\delta)$ , también cumple con  $\text{DISC}_{\frac{d}{n}}(\varepsilon)$ .  
 510 En efecto, para todo  $\varepsilon > 0$  y  $X, Y \subset [n]$ , elija  $n_0 \in \mathbb{N}$  suficientemente grande y  $\delta < \frac{\varepsilon n}{\sqrt{|X||Y|}}$ .  
 511 Entonces, si  $G$  satisface la propiedad  $\text{EIG}_{\frac{d}{n}}(\delta)$ , por el Teorema 25:

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

512

513 Finalmente, en un grafo  $d$ -regular, la equivalencia entre las propiedades  $\text{EIG}_{\frac{d}{n}}$  y  $\text{DISC}_{\frac{d}{n}}$  se  
 514 completa por el camino de implicancias ya demostradas según el esquema (13).

## 4. Lema de regularidad de Szemerédi

\* Aquí debo ingresar una breve descripción del lema de regularidad de Szemerédi y la fuerza que toma al combinarlo con el teorema de Chung-Graham-Wilson y comentar la aplicación que se mostrará. (Quizás también hablar de las dos demostraciones de este lema, usual y espectral.) \*

### 4.1. Enunciado y demostración

Se tratará el concepto de regularidad de una manera un poco diferente a como es tradicionalmente conocida. En particular, se permite intersección entre pares de subconjuntos de vértices de un grafo en las futuras definiciones.

**Definición 26.** Sea  $G$  un grafo y  $X, Y \subset V(G)$  subconjuntos no necesariamente disjuntos. Diremos que  $(X, Y)$  es un **par  $\varepsilon$ -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  $\varepsilon$ -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  $\left| d(A, B) - d(X, Y) \right| > \varepsilon$ .

Notaremos que la noción de un par  $\varepsilon$ -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

$$\left| e(X, Y) - p|X||Y| \right| = o(|U||W|), \quad \forall X \subset U, \forall Y \subset W. \quad (26)$$

En efecto, si  $(U, W)$  es un par  $\varepsilon$ -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  $\varepsilon$ -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 otro lado, con el espíritu del Teorema 15, es posible expresar un resultado análogo a la propiedad  $\text{COUNT}_p(\varepsilon)$  utilizando el concepto de par  $\varepsilon$ -regular. Dicho resultado, para  $H = K_3$  es conocido como el lema de conteo de triángulos.

542 **Lema 27.** (*Lema de conteo de triángulos*) Sea  $\varepsilon > 0$ ,  $G = (V, E)$  un grafo, y los conjuntos no  
 543 necesariamente disjuntos  $X, Y, Z \subset V$  tales que los pares  $(X, Y)$ ,  $(Y, Z)$  y  $(X, Z)$  son  $\varepsilon$ -regular.  
 544 Entonces,

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

545 *Demostración.* Se realizará un proceso inductivo similar al visto en la demostración de la Propo-  
 546 sición 18 sobre la cantidad de aristas del grafo  $K_3 = ([3], \{12, 23, 13\})$ . Cuando el grafo no posea  
 547 aristas, entonces

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

548

549 También, recordando que la condición de un par  $\varepsilon$ -regular es equivalente a  $\text{BI} - \text{DISC}_p(\varepsilon)$  para  
 550 algún  $p \in (0, 1)$ , cuando el grafo presenta una arista,

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

551

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

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

553

554 Defina  $e^- = \varphi(1)\varphi(3)$ , y  $T^-$  como el grafo correspondido a una copia etiquetada del grafo  
 555  $([3], \{12, 23\})$  en  $G$  bajo la aplicación inyectiva  $\varphi : [3] \rightarrow V(T^-) \subset V$ . Con esto, se desarrolla  
 556 inductivamente como sigue:

$$\begin{aligned} |\{(x, y, z) \in X \times Y \times Z : xy, yz, xz \in E\}| &= \sum_{T^-} [\mathbb{1}_E(e^-) + d(X, Z) - d(X, Z)] \\ &= d(X, Y)d(Y, Z)d(X, Z)|X||Y||Z| \\ &\quad + \sum_{T^-} (\mathbb{1}_E(e^-) - d(X, Z)) \pm 2\varepsilon|X||Y||Z|. \end{aligned} \quad (27)$$

557

558 En este punto, nos falta probar que el segundo sumando de la igualdad (27) se corresponde  
 559 con un factor de error, para esto, sea  $T^*$  una copia del grafo singleton  $\{2\}$  en  $G$ , y considere los  
 560 siguientes conjuntos:

$$A_1^{T^*} = \{x \in X : T^* \text{ con } x \text{ forma una copia de } (\{1, 2\}, \{12\}) \text{ en } G\}.$$

$$A_3^{T^*} = \{z \in Z : T^* \text{ con } z \text{ forma una copia de } (\{2, 3\}, \{23\}) \text{ en } G\}.$$

De esta manera, dada la equivalencia de la condición del par  $(X, Z)$   $\varepsilon$ -regular con versión bipartita de la propiedad  $\text{DISC}_{d(X, Z)}(\varepsilon)$  vista en (26), 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 (27) se prueba lo prometido.  $\square$

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

Bajo el mismo planteamiento de la inducción vista en la demostración del Lema 27 (y Proposición 18), es posible generalizar el resultado para contar apropiadamente cualquier grafo  $H$ . Se enuncia sin demostración.

**Lema 28.** (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  $\varepsilon$ -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

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

Ya conociendo el concepto de regularidad entre pares de subconjuntos de vértices, estudiamos la regularidad en una partición del conjunto de vértices del grafo.

**Definición 29.** Dado un grafo  $G$ , una partición  $\mathcal{P} = \{V_1, \dots, V_k\}$  del conjunto de vértices  $V(G)$  es una **partición  $\varepsilon$ -regular** si

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

581

582 *Es decir, todos los pares de subconjuntos de vértices en la partición son  $\varepsilon$ -regular salvo una*  
 583 *fracción  $\varepsilon$  de pares de vértices.*

584 Note que si una partición  $\varepsilon$ -regular de  $k$  partes es en particular una equipartición, entonces a lo  
 585 más  $\varepsilon k^2$  pares de elementos de la partición no son  $\varepsilon$ -regular.

586 Ya con todo lo necesario, se introduce el lema de regularidad de Szemerédi. Intuitivamente, el  
 587 lema permite particionar el conjunto de vértices de todo grafo en una cantidad finita de partes,  
 588 satisfaciendo que la mayoría de sus pares de partes son  $\varepsilon$ -regular. Enunciamos el célebre lema, y se  
 589 dará prueba formal más adelante.

590 **Teorema 30.** *(Lema de regularidad de Szemerédi) Para todo  $\varepsilon > 0$ , existe un entero  $M = M(\varepsilon)$*   
 591 *tal que todo grafo admite una partición  $\varepsilon$ -regular de a lo más  $M$  partes.*

592 Para dar prueba a este teorema, se utilizará una técnica llamada *argumento de incremento de*  
 593 *energía*. Para todo grafo  $G$ , la técnica funciona algorítmicamente de la siguiente manera:

- 594 1. Comenzar con la partición trivial de  $V(G)$ , i.e,  $\mathcal{P} = \{V(G)\}$ .
- 595 2. Mientras la partición actual  $\mathcal{P}$  no es  $\varepsilon$ -regular:
- 596 (a) Para cada par  $(V_i, V_j)$  no  $\varepsilon$ -regular, encontrar los subconjuntos  $A^{ij} \subset V_i$  y  $A^{ji} \subset V_j$  que  
 597 evidencian la irregularidad de cada par.
- 598 (b) Refinar  $\mathcal{P}$  utilizando simultáneamente los conjuntos  $A^{ij}$  y  $A^{ji}$  encontrados de cada par  
 599  $(V_i, V_j)$  no  $\varepsilon$ -regular para obtener  $\mathcal{Q}$ .
- 600 (c) Actualizar  $\mathcal{P}$  con  $\mathcal{Q}$ .

601 Siendo  $\mathcal{P}$  y  $\mathcal{Q}$  dos particiones de un mismo conjunto de vértices, diremos que  $\mathcal{Q}$  refina a  $\mathcal{P}$  si  
 602 cada parte de  $\mathcal{Q}$  está contenida en una parte de  $\mathcal{P}$ . En lo que resta de esta sección mostraremos  
 603 que el algoritmo tiene un fin, y que entrega una partición  $\varepsilon$ -regular en un número de iteraciones  
 604 que solo depende de  $\varepsilon$ .

605 **Definición 31.** *(Energía) Sea  $G$  un grafo sobre  $n$  vértices y  $X, Y \subset V(G)$ . Se define en primer*  
 606 *lugar*

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

607

608 *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*

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

609



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

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

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

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

612

613 La última observación es crucial en la demostración del Teorema 30, puesto que los Lemas 32,  
614 33 y 34 nos asegurarán que la energía de una partición nunca decrece bajo refinamiento. Por conse-  
615 cuencia, el algoritmo de la técnica *argumento de incremento de energía* tendrá un fin, entregando  
616 una partición  $\varepsilon$ -regular.

617 El siguiente lema muestra que la energía no disminuye al particionar o refinar arbitrariamente  
618 un conjunto o partición respectivamente.

619 **Lema 32.** Sea  $G$  un grafo,  $X, Y \subset V(G)$ ,  $\mathcal{P}_X$  y  $\mathcal{P}_Y$  particiones de  $X$  e  $Y$  respectivamente, entonces  
620  $q(\mathcal{P}_X, \mathcal{P}_Y) \geq q(X, Y)$ . Además, dadas dos particiones de vértices  $\mathcal{P}$  y  $\mathcal{P}'$  de  $G$ ,  $q(\mathcal{P}) \leq q(\mathcal{P}')$  cada  
621 vez que  $\mathcal{P}'$  refina a  $\mathcal{P}$ .

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

$$\begin{aligned}
q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^k \sum_{j=1}^{\ell} q(X_i, Y_j) \\
&= \frac{1}{n^2} \sum_{i=1}^k \sum_{j=1}^{\ell} \frac{e(X_i, Y_j)^2}{|X_i||Y_j|} \\
&\stackrel{(3)}{\geq} \frac{1}{n^2} \frac{\left( \sum_{i=1}^k \sum_{j=1}^{\ell} e(X_i, Y_j) \right)^2}{\sum_{i=1}^k \sum_{j=1}^{\ell} |X_i||Y_j|} \\
&= \frac{1}{n^2} \frac{e(X, Y)^2}{\left( \sum_{i=1}^k |X_i| \right) \left( \sum_{j=1}^{\ell} |Y_j| \right)} \\
&= \frac{e(X, Y)^2}{n^2 |X||Y|} \\
&= q(X, Y).
\end{aligned} \tag{28}$$

625

626 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  $\mathcal{P}$ .  
627 Entonces, se utiliza el resultado probado previamente para completar el resultado:

$$q(\mathcal{P}) = \sum_{i=1}^k \sum_{j=1}^k q(V_i, V_j) \stackrel{(28)}{\leq} \sum_{i=1}^k \sum_{j=1}^k q(\mathcal{P}'_{V_i}, \mathcal{P}'_{V_j}).$$

628

□

629 Ahora, veremos que refinar un par  $(X, Y)$  no  $\varepsilon$ -regular de un grafo  $G$ , mediante los subconjuntos  
630 que evidencian su irregularidad, provoca un aumento estricto en la energía.

631 **Lema 33.** Sea  $\varepsilon > 0$ ,  $G$  un grafo de  $n$  vértices y  $X, Y \subset V(G)$ . Si  $(X, Y)$  no es un par  $\varepsilon$ -regular,  
632 existen particiones  $\mathcal{P}_X = \{X_1, X_2\}$  de  $X$  y  $\mathcal{P}_Y = \{Y_1, Y_2\}$  de  $Y$  tales que

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

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

$$|X_1| \geq \varepsilon |X| \quad \text{y} \quad |Y_1| \geq \varepsilon |Y|. \tag{29}$$

636

637 Se define adicionalmente los conjuntos  $X_2 := X \setminus X_1$ ,  $Y_2 := Y \setminus Y_1$ , y  $\eta := d(X_1, Y_1) - d(X, Y)$ ,  
638 cual por definición de par  $\varepsilon$ -regular, satisface\* Aquí quedé \*

$$|\eta| > \varepsilon. \quad (30)$$

639 Por un lado, observe la siguiente descomposición,

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

640 De esta manera,

$$\sum_{i+j>2} e(X_i, Y_j) = e(X, Y) - e(X_1, Y_1). \quad (31)$$

641 Por otro lado, se tiene que,

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

642 Así,

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

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

$$\begin{aligned} q(\mathcal{P}_X, \mathcal{P}_Y) &= \sum_{i=1}^2 \sum_{j=1}^2 q(X_i, Y_j) \\ &= \sum_{i=1}^2 \sum_{j=1}^2 \frac{e(X_i, Y_j)^2}{n^2 |X_i||Y_j|} \\ &= \frac{1}{n^2} \left( \frac{e(X_1, Y_1)^2}{|X_1||Y_1|} + \sum_{i+j>2} \frac{e(X_i, Y_j)^2}{|X_i||Y_j|} \right) \\ &\stackrel{(3)}{\geq} \frac{1}{n^2} \left( \frac{e(X_1, Y_1)^2}{|X_1||Y_1|} + \frac{\left( \sum_{i+j>2} e(X_i, Y_j) \right)^2}{\sum_{i+j>2} |X_i||Y_j|} \right) \\ &\stackrel{(31) \text{ y } (32)}{=} \frac{1}{n^2} \left( \frac{e(X_1, Y_1)^2}{|X_1||Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X||Y| - |X_1||Y_1|} \right). \end{aligned} \quad (33)$$

644 Luego, por definición, note que

$$e(X_1, Y_1) = \frac{|X_1||Y_1|e(X, Y)}{|X||Y|} + \eta |X_1||Y_1|. \quad (34)$$

645 Con esto, se continúa el cálculo desde la desigualdad (33) como sigue, \* Me podré saltar un  
 646 espacio más pequeño abajo? \*

$$\begin{aligned}
 n^2 q(\mathcal{P}_X, \mathcal{P}_Y) &\geq \frac{e(X_1, Y_1)^2}{|X_1||Y_1|} + \frac{(e(X, Y) - e(X_1, Y_1))^2}{|X||Y| - |X_1||Y_1|} \\
 &\stackrel{(34)}{=} \frac{1}{|X_1||Y_1|} \left( \frac{|X_1||Y_1|e(X, Y)}{|X||Y|} + \eta |X_1||Y_1| \right)^2 \\
 &\quad + \frac{1}{|X||Y| - |X_1||Y_1|} \left( \frac{|X||Y| - |X_1||Y_1|}{|X||Y|} e(X, Y) - \eta |X_1||Y_1| \right)^2 \\
 &= \frac{|X_1||Y_1|}{|X|^2|Y|^2} e(X, Y)^2 + 2 \frac{|X_1||Y_1|}{|X||Y|} \eta e(X, Y) + \eta^2 |X_1||Y_1| \\
 &\quad + \frac{|X||Y| - |X_1||Y_1|}{|X|^2|Y|^2} e(X, Y)^2 - 2 \frac{|X_1||Y_1|}{|X||Y|} \eta e(X, Y) + \frac{\eta^2 |X_1|^2 |Y_1|^2}{|X||Y| - |X_1||Y_1|} \\
 &= \frac{e(X, Y)^2}{|X||Y|} + \eta^2 |X_1||Y_1| \left( 1 + \frac{|X_1||Y_1|}{|X||Y| - |X_1||Y_1|} \right) \\
 &\geq \frac{e(X, Y)^2}{|X||Y|} + \eta^2 |X_1||Y_1|. \tag{35}
 \end{aligned}$$

649 Finalmente, utilizando las cotas (29) y (30), podemos concluir desde la desigualdad (35),

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

650

□

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

656 **Lema 34.** Sea  $\varepsilon > 0$ , un grafo  $G$  y una partición  $\mathcal{P} = \{V_1, \dots, V_k\}$  no  $\varepsilon$ -regular de  $V(G)$ . Entonces  
 657 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

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

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

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

$$\sum_{(i,j) \notin \Theta} \frac{|V_i||V_j|}{n^2} > \varepsilon. \quad (36)$$

668 Así, junto a los lemas probados previamente, damos prueba al resultado de la siguiente manera,

$$\begin{aligned} q(\mathcal{Q}) &= \sum_{(i,j) \in [k]^2} q(\mathcal{Q}_i, \mathcal{Q}_j) \\ &= \sum_{(i,j) \in \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) + \sum_{(i,j) \notin \Theta} q(\mathcal{Q}_i, \mathcal{Q}_j) \\ &\stackrel{\text{Lema 32}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} q(\{A^{ij}, V_i \setminus A^{ij}\}, \{A^{ji}, V_j \setminus A^{ji}\}) \\ &\stackrel{\text{Lema 33}}{\geq} \sum_{(i,j) \in \Theta} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \left( q(V_i, V_j) + \varepsilon^4 \frac{|V_i||V_j|}{n^2} \right) \\ &= \sum_{(i,j) \in [k]^2} q(V_i, V_j) + \sum_{(i,j) \notin \Theta} \varepsilon^4 \frac{|V_i||V_j|}{n^2} \\ &\stackrel{(36)}{\geq} q(\mathcal{P}) + \varepsilon^5. \end{aligned}$$

669 \* Cambiar por  $>$  en la última línea y donde dice lema 5, cuando lo cambio se me descuadra :c  
670 \* □

671 Ya con todo lo necesario, damos prueba formal al Teorema 30 mediante la técnica de *argumento*  
672 *de incremento de energía*.

673 *Demostración del Teorema 30.* Dado  $\varepsilon > 0$  y un grafo  $G$ , elegimos inicialmente la partición trivial  
674 del conjunto de vértices  $\mathcal{P} = \{V(G)\}$ . Ahora, iterativamente (actualizando  $\mathcal{P}$ ), aplicaremos el Lema  
675 34 cada vez que la partición actual no sea  $\varepsilon$ -regular. Observe que por cada aplicación del Lema 34 se  
676 consigue un aumento de al menos  $\varepsilon^5$  en la energía, y como la energía de toda partición está acotada  
677 superiormente por 1, el proceso iterativo terminará luego de a lo más  $\varepsilon^{-5}$  pasos. El resultado será  
678 necesariamente una partición  $\varepsilon$ -regular debido a la cota de la energía.

679 Para una partición no  $\varepsilon$ -regular con  $k$  elementos, el Lema 34 encuentra un refinamiento de a lo  
680 más  $k2^k$  partes. Dicho refinamiento será producido en cada iteración del algoritmo de *argumento*  
681 *de incremento de energía*, y la cantidad de partes producidas las acotaremos crudamente en cada

682 paso por  $k2^k < 2^{2^k}$ . Comenzando con la partición trivial de una parte, ejemplificaremos con las  
683 tres primeras iteraciones del algoritmo para mostrar la cantidad de partes producidas en cada paso  
684 tras aplicar el Lema 34.

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

685 Así, como el algoritmo debe luego de a lo más  $\varepsilon^{-5}$  iteraciones, la cantidad de partes al final de  
686 proceso será

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

687 □

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

$$\text{torre}(k) := 2^{2^{\dots^2}} \left. \vphantom{2^{2^{\dots^2}}} \right\} \text{Altura } k. \quad (37)$$

689

690 \* En esta parte me gustaría dejar un comentario sobre lo grande que es la cota y el resultado  
691 que encontró Gowers en 1997 de cota inferior, pero no lo entiendo :c \*

692 Una de las peculiaridades del lema de regularidad de Szemerédi es la flexibilidad que posee su  
693 enunciado, adaptando su aplicación a diferentes contextos. Por ejemplo, si en la demostración del  
694 Teorema 30 tomamos una partición inicial arbitraria en vez de la partición trivial del conjunto de  
695 vértices del grafo, se logra obtener la siguiente variante del lema de regularidad.

696 **Teorema 35.** (*Regularidad de Szemerédi - Partición inicial arbitraria*) Para todo  $\varepsilon > 0$ , existe un  
697 entero  $M = M(\varepsilon)$  tal que todo grafo  $G$  con una partición inicial  $\mathcal{P}_0$  de  $V(G)$  admite una partición  
698  $\varepsilon$ -regular  $\mathcal{P}$  de  $V(G)$  que refina cada parte de  $\mathcal{P}_0$  en a lo más  $M$  partes.

699 Es posible fortalecer un poco más el lema de regularidad exigiendo que el resultado sea una  
700 equipartición del conjunto de vértices de un grafo  $G$ . Es decir, una partición  $\mathcal{P} = \{V_1, \dots, V_k\}$  tal  
701 que  $|V_1| \leq |V_2| \leq \dots \leq |V_k| = |V_1| + 1$ .

702 **Teorema 36.** (*Regularidad de Szemerédi - Equipartición*) Para todo  $\varepsilon > 0$  y  $m_0 \in \mathbb{N}$ , existe un  
703 entero  $M = M(\varepsilon)$  tal que todo grafo admite una equipartición  $\varepsilon$ -regular de su conjunto de vértices  
704 de  $k$  partes, con  $m_0 \leq k \leq M$ .

705 \* Comentar que cuando tenemos esta versión del teorema (que es la clásica), entonces la defini-  
706 ción de una partición  $\varepsilon$ -regular se traduce en  $\dots \leq \varepsilon k^2$ . También hablar un poco de  $m_0$ , cual ayuda  
707 a que ninguna de las partes sea demasiado grande. \*

La idea de la demostración del Teorema 36 consiste en modificar el algoritmo de la técnica de argumento de incremento de energía, de manera que en 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. 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  $\varepsilon$ -regular:
  - (a) Para cada par  $(V_i, V_j)$  no  $\varepsilon$ -regular, encontrar los subconjuntos  $A^{ij} \subset V_i$  y  $A^{ji} \subset V_j$  que evidencian la irregularidad de cada par.
  - (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  $|V(G)|/m$  utilizando 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  $\varepsilon^5$  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.<sup>2</sup> 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. Aplicaciones

Usualmente las aplicaciones del lema de regularidad de Szemerédi son desarrolladas en base a los siguientes pasos:

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

---

<sup>2\*</sup> Aquí quiero hacer un comentario/ejemplo de  $m$ . Yufei sugiere  $m = \lfloor 100|\mathcal{Q}\varepsilon^{-5}| \rfloor$ , pero tampoco lo entiendo mucho. \*

Teniendo esta fórmula en mente, damos paso a la primera aplicación del lema de regularidad, cual plantea intuitivamente que todo grafo con *pocas* triángulos puede convertirse en un grafo libre de triángulos eliminando *pocas* aristas. Formalmente,

**Teorema 37.** (*Lema de eliminación de triángulos*) Para todo  $\varepsilon > 0$ , existe  $\delta > 0$  y  $n_0 \in \mathbb{N}$  tal que todo grafo sobre  $n \geq n_0$  vértices con a lo más  $\delta n^3$  triángulos se puede hacer libre de triángulos eliminando a lo más  $\varepsilon n^2$  aristas.

*Demostración.* Dado  $\varepsilon > 0$ , elija  $\varepsilon_r = \frac{1}{4} \left(\frac{\varepsilon}{3}\right)^3$  y utilice el Teorema 30 para obtener la constante  $M = M(\varepsilon_r)$ . Considere además  $\delta = \frac{1}{2} \frac{\varepsilon_r^4}{M^3}$  y  $n_0 \in \mathbb{N}$  suficientemente grande, de manera tal que el grafo  $G = (V, E)$  con  $n \geq n_0$  vértices posee a lo más  $\varepsilon n^3$  triángulos. Luego, nuevamente por el Teorema 30, se asegura la existencia de una partición  $\varepsilon_r$ -regular  $\mathcal{P} = \{V_1, \dots, V_M\}$ .

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

(a)  $(V_i, V_j)$  no es un par  $\varepsilon_r$ -regular,

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

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

De esta manera, como la partición es  $\varepsilon_r$ -regular, las aristas removidas por la condición (a) son a lo más

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

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

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

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

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

En total, en la limpieza, se eliminan a lo más  $\varepsilon n^2$  aristas.

Ahora, nos falta probar que el grafo limpio  $G' = (V, E')$  es libre de triángulos. Para esto, observe que la condición de eliminación de aristas (a) nos asegura que cada par  $(V_i, V_j)$  es  $\varepsilon_r$ -regular, y que se satisface la hipótesis del lema de conteo de grafos. Entonces, si luego de la limpieza del grafo aún



767 existe un triángulo  $(x, y, z) \in V_i \times V_j \times V_\ell$ , el Lema 27 nos dice que incluso hay más triángulos. En  
 768 particular, gracias a la eliminación de las aristas por la condición (b) y (c),

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

769

770 Finalmente, con nuestra elección de  $\delta$ , el resultado se prueba formulando la siguiente contradic-  
 771 ción: si existe un triángulo en el grafo limpio  $G'$ , el lema de conteo de triángulos nos dice que en  
 772 realidad existen más de  $\delta n^3$  triángulos. No obstante, el grafo original posee a lo más  $\delta n^3$  triángulos,  
 773 por lo que se concluye que el grafo  $G'$  obtenido desde  $G$  es libre de triángulos removiendo a lo más  
 774  $\varepsilon n^2$  aristas.  $\square$

775 Denotaremos por  $k$ -PA a una progresión aritmética de  $k$  elementos. En particular, diremos que  
 776 un conjunto de números naturales  $A$  es libre de 3-PA si no existen los elementos  $x, x+y, x+2y \in A$ ,  
 777 con  $y \neq 0$ . Cuando  $y = 0$ , diremos que la 3-PA es trivial.

778 **Teorema 38.** (Teorema de Roth) Para todo  $\varepsilon > 0$ , existe  $n_0 \in \mathbb{N}$  tal que si el conjunto  $A \subset [n]$   
 779 posee  $|A| \geq \varepsilon n$  elementos, entonces  $A$  contiene una 3-PA no trivial cada vez que  $n \geq n_0$ .

780 *Demostración.* Sea  $\varepsilon > 0$  y el conjunto  $A \subset [n]$  con  $|A| \geq \varepsilon n$  elementos. La idea es construir  
 781 un grafo 3-partito de manera conveniente para posteriormente utilizar el lema de eliminación de  
 782 triángulos. Considere el grafo 3-partito  $G = (V, E)$  con partición de vértices  $V = V_1 \cup V_2 \cup V_3$ , en  
 783 donde  $V_1 = [n]$ ,  $V_2 = [2n]$  y  $V_3 = [3n]$ , y son disjuntos entre cada par de ellos. Así,  $G$  tiene  $6n$   
 784 vértices, y se definen las aristas de la siguiente manera:

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

788 Luego, la tupla  $(i, j, k) \in V_1 \times V_2 \times V_3$  define un triángulo en  $G$  si y solamente si  $j - i \in A$ ,  
 789  $k - j \in A$  y  $\frac{k-i}{2} \in A$ , o bien,  $\{j - i, \frac{k-i}{2}, k - j\}$  es una 3-PA en  $A$  con diferencia  $\frac{k-2j+i}{2}$ . En  
 790 específico, diremos que un triángulo  $(i, j, k) \in V_1 \times V_2 \times V_3$  es trivial en  $G$  si para algún  $a \in A$  se  
 791 satisface que  $j - i = \frac{k-i}{2} = k - j = a$ .

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

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

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

802

803 Finalmente, el resultado queda demostrado tomando  $n_0 > \max \left\{ \frac{n'_0}{6}, \frac{\varepsilon+1}{\delta} \right\}$  suficientemente gran-  
 804 de.  $\square$

805 **Definición 39.** Dado un grafo  $G = (V, E)$ , un un conjunto  $M \subseteq E$  es un **emparejamiento**  
 806 en  $G$  si no existen un par de aristas en  $M$  que compartan algún vértice. Diremos que  $M$  es un  
 807 **emparejamiento inducido** si es un emparejamiento y toda arista de  $G$  con un vértice en  $V(M)$   
 808 es una arista en  $M$ .

809 \* Usar  $k$  o  $M$  para la cantidad de partes?, aquí se me confunde con el emparejamiento, pero en  
 810 TRL y demo espectral de regularidad usé  $M$  como las partes. De momento en esta parte lo dejaré  
 811 con  $k$ . \*

812 **Teorema 40.** (Emparejamiento inducido) Para todo  $\varepsilon > 0$ , existe  $n_0 \in \mathbb{N}$  tal que todo grafo  
 813  $G = (V, E)$  de  $n \geq n_0$  vértices que está compuesto por la unión de  $n$  emparejamientos inducidos,  
 814 posee a lo más  $\varepsilon n^2$  aristas.

815 *Demostración.* Dado  $\varepsilon > 0$ , aplique el Teorema 30 con  $\varepsilon_r = \frac{\varepsilon}{10}$  para obtener la constante  $M(\varepsilon_r)$ .  
 816 Considere  $n_0 \in \mathbb{N}$  suficientemente grande, y asuma que el grafo  $G = (V, E)$  con  $n \geq n_0$  vértices y  
 817 compueston por  $n$  emparejamientos inducidos satisface  $e_G > \varepsilon n$ . Nuevamente, por el Teorema 30,  
 818 se asegura la existencia de la partición  $\mathcal{P} = \{V_1, \dots, V_k\}$  con  $k \leq M(\varepsilon)$  partes que es  $\varepsilon_r$ -regular.

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

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

823

824 Ahora, observe que debe existir un emparejamiento inducido  $M$  en  $G'$  con al menos  $\frac{\varepsilon}{2}n$  aristas  
 825 (y al menos  $\varepsilon n$  vértices). De no ser así, todos los emparejamientos tendrán a lo más  $\frac{\varepsilon}{2}n$  aristas, por  
 826 lo que  $e'_G < \frac{\varepsilon}{2}n^2$ .

827 Se define  $U_i := V_i \cap V(M)$  como el subconjunto de vértices de  $M$  que comparte elementos con  
 828  $V_i$ , 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  $U_i \subset V(M)$  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  $V(M)$ , pues

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

831

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

$$d(U_t, U_\ell) = d(V_t, V_\ell) \pm \varepsilon_r \geq 2\varepsilon_r - \varepsilon_r = \varepsilon_r. \quad (38)$$

837

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

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

840

Sin embargo, la desigualdad (38) implica que

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

842

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

\* Comentar que el siguiente teorema será utilizado para demostrar alternativamente el Teorema de Roth. \*

**Teorema 41.** (Ajtai-Szemerédi) Para todo  $\varepsilon > 0$ , existe  $n_0 \in \mathbb{N}$  tal que siempre que  $n \geq n_0$ , todo subconjunto  $S \subset [n]^2$  con  $|S| \geq \varepsilon n^2$  posee elementos de la forma  $\{(a, b), (a + d, b), (a, b + d)\}$  para algún  $a, b, d \in \mathbb{N}$ , con  $d \neq 0$ .

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

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

853

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

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

857

858 *\* Dibujito con 2 ejemplos de  $q$ .* \* Observe que para cada  $2 \leq q \leq 2n$  se define un emparejamiento  
859 en  $G$  debido a que no existen aristas que compartan un vértice, por lo que las clases de equivalencia  
860 (cada una asociada a algún  $q$ ) de la relación forman una partición de emparejamientos de  $E$ . En  
861 efecto, suponga que las aristas que pertenecen a la misma clase  $u_i w_j$  y  $u_k w_j$  comparten el vértice  
862  $w_j$ . 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  
863 misma arista.

864 Luego, como  $e_G = |S| \geq \varepsilon n^2$ , el Teorema 40 asegura que existe al menos un emparejamiento no  
865 inducido. Esto significa que en un emparejamiento que contiene las aristas con la relación  $u_i w_j \sim$   
866  $u_k w_\ell$  puede existir el trío de aristas  $u_i w_j$ ,  $u_k w_\ell$  y  $u_i w_\ell$ . Así, para algún  $d \in \mathbb{N}$ ,  $(i, j)$ ,  $(k, \ell)$  y  $(i, \ell)$   
867 elementos de  $S$  que satisfacen

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

868

869 Finalmente, el resultado se consigue tomando  $(i, \ell) = (a, b)$  para obtener  $j = b + d$  y  $k = a + d$ .  
870 *\* Poner dibujito de la esquina \** □

871 *\* Comentar que el Teorema de la esquina nos entrega otro camino para demostrar el Teorema*  
872 *de Roth. \**

873 *Segunda demostración Teorema 38.* Dado  $\varepsilon > 0$ , escogemos  $n_0 \in \mathbb{N}$  suficientemente grande. Para  
874  $n \geq n_0$ , sea  $A \subset [n]$  un conjunto que posee al menos  $\varepsilon n$  elementos. Se define el siguiente conjunto:

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

875

876 Observe que cada  $a \in A$  da lugar a exactamente  $n$  elementos en  $B$  con  $x - y = a$ , permitiendo  
877 determinar que  $|B| = n|A| \geq \varepsilon n^2$ . Luego, el Teorema 41 asegura la existencia de elementos de la  
878 forma  $\{(a, b), (a, b + d), (a + d, b)\}$  en  $B$ . Por consecuencia, se encuentra una 3-PA no trivial en  $A$   
879 tomando  $x = a - b$ , e  $y = d$ . □

\* Explicar que ahora vamos a demostrar con teoría espectral el lema de regularidad de Szemerédi.  
Comentar también que esta versión la realizó Terence Tao. \*

*Demostración espectral Teorema 30.* Sea  $\varepsilon > 0$ ,  $G = ([n], E)$  un grafo y  $T$  su matriz de adyacencia. Consideramos además  $\{u_1, \dots, u_n\}$  la base ortonormal de  $\mathbb{R}^n$  formada por los vectores propios de  $T$ , y  $|\lambda_1| \geq \dots \geq |\lambda_n|$  los valores propios de  $T$  ordenados de manera decreciente.

Por la Proposición 4 y el Corolario 8, se satisface

$$\text{Tr}(T) = \sum_{i=1}^n \lambda_i^2 = 2e_G \leq n^2. \quad (39)$$

De esta manera, al notar que  $i\lambda_i^2 \leq \sum_{j=1}^i \lambda_j^2 \leq n^2$ , es posible acotar cada valor propio de la siguiente manera:

$$\lambda_i \leq \frac{n}{\sqrt{i}}, \quad \forall i \in [n]. \quad (40)$$

Al final de esta demostración se entregará una función  $f : \mathbb{N} \rightarrow \mathbb{N}$  que depende únicamente de  $\varepsilon$  y que satisface  $f(i) > i$ . Denotando por  $f^{(k)}$  a la  $k$ -ésima composición de  $f$  con ella misma, consideramos una partición de  $[n]$  en intervalos de la forma  $[f^{(k-1)}(1), f^{(k)}(1)]$ , para  $k \in \{1, \dots, \frac{1}{\varepsilon^3}\}$ . Con esta construcción, debe existir un natural  $\ell = f^{(k-1)}(1)$  que cumple con

$$\sum_{\ell \leq j < f(\ell)} |\lambda_j|^2 \leq \varepsilon^3 n^2. \quad (41)$$

De lo contrario, la suma de  $|\lambda_j|^2$  sobre todos los intervalos definidos es estrictamente mayor que  $\varepsilon^3 n^2$ . Así, como son  $\frac{1}{\varepsilon^3}$  intervalos, se contradice la desigualdad (39), pues

$$\sum_{j=1}^n |\lambda_j|^2 > \frac{1}{\varepsilon^3} \cdot \varepsilon^3 n^2 = n^2.$$

Ahora, usando el natural  $\ell$ , separamos la matriz  $T$  en tres partes. En específico,

$$T = T_1 + T_2 + T_3.$$

Se interpretará  $T_1$  como la componente *estructural*,

$$T_1 := \sum_{i < \ell} \lambda u_i u_i^T,$$

902  $T_2$  como la componente de *error*,

$$T_2 := \sum_{\ell \leq i < f(\ell)} \lambda_i u_i u_i^T,$$

903

904 y  $T_3$  como la componente *casi-aleatoria*,

$$T_3 := \sum_{i \geq f(\ell)} \lambda_i u_i u_i^T.$$

905

906 Pensamos cada vector propio de  $T$  como una función  $u_i : [n] \rightarrow \mathbb{R}$ . En otras palabras, todo  
907 vector propio asigna un *peso* a cada vértice de  $G$ .

908 **Analizamos  $T_1$ .** La idea es particionar el conjunto de vértices  $[n]$  de manera tal que  $T_1$  es  
909 aproximadamente constante en la mayoría de las partes. Veremos que el número de partes será  
910  $O_{\ell, \varepsilon}(1)$ , es decir, un valor constante que depende solo de  $\ell$  y  $\varepsilon$ .

911 Para cada  $i \in [\ell - 1]$  ordenamos de manera creciente los vértices de  $G$  según la asignación de *pesos*  
912 que otorga  $u_i(\cdot)$ . En primera instancia, se agrupa en un conjunto excepcional a aquellos vértices que  
913 presenten un *peso* demasiado grande en magnitud. Dicho conjunto se define de la siguiente manera:

$$V_0^i := \left\{ k \in [n] : |u_i(k)| > \sqrt{\frac{\ell}{\varepsilon}} n^{-1/2} \right\}.$$

914

915 Dado que  $\|u_i\| = 1$ , cada  $V_0^i$  no puede tener muchos elementos. En efecto, al observar que

$$|V_0^i| \left( \sqrt{\frac{\ell}{\varepsilon}} n^{-1/2} \right)^2 < \sum_{k=1}^n u_i(k)^2 = \|u_i\|^2 = 1,$$

916

917 es posible determinar que  $|V_0^i| < \frac{\varepsilon}{\ell} n$ .

918 Aquellos vértices que no están en  $V_0^i$ , serán agrupados particionando la recta de largo  $2\sqrt{\frac{\ell}{\varepsilon}} n^{-1/2}$   
919 en subintervalos de tamaño a lo más  $\left( \frac{\varepsilon^{3/2}}{\ell^{3/2}} \right) n^{-1/2}$ . Esta configuración provoca gráficamente el  
920 siguiente esquema para cada  $u_i(\cdot)$ .

921 \* Poner dibujito... \*

922 Por consecuencia, para  $i \in [\ell - 1]$ , la cantidad de partes que genera cada  $u_i(\cdot)$  son a lo más

$$\frac{2\sqrt{\frac{\ell}{\varepsilon}} n^{-1/2}}{\frac{\varepsilon^{3/2}}{\ell^{3/2}} n^{-1/2}} = \frac{2\ell^2}{\varepsilon^2} = O_{\ell, \varepsilon}(1).$$

923

924 Para conseguir la partici3n deseada de  $[n]$ , por un lado, se toma la uni3n de todos los conjuntos  
 925 excepcionales  $V_0^i$  para dar lugar al conjunto  $V_0$  de tama3o a lo m3s  $(\ell - 1) \cdot \frac{\varepsilon n}{\ell} < \varepsilon n$ . Por otro lado,  
 926 combine las particiones generadas por los  $\ell - 1$  primeros vectores propios mediante un refinamiento  
 927 usual. As3, se consigue una partici3n del conjunto de v3rtices de  $G$  de la forma  $[n] = V_0 \cup V_1 \cup \dots \cup V_M$ .  
 928 Dada la construcci3n, la cantidad de partes que se obtienen son

$$M(\varepsilon) \leq \left( \frac{2\ell^2}{\varepsilon^2} \right)^\ell \quad (42)$$

929

930 Ahora, intuitivamente, se mostrar3 que los valores de la matriz  $T_1$  en cada bloque  $V_i \times V_j$  son  
 931 aproximadamente constante, i.e, no var3an m3s que  $o_\varepsilon(1)$ . Para esto, como se hizo con los vectores  
 932 propios, pensamos la matriz de adyacencia como una funci3n  $T : [n] \times [n] \rightarrow \mathbb{R}$  para identificar sus  
 933 entradas. De esta manera, para cada  $i, j \in [M]$ ,  $a, c \in V_i$ , y  $b, d \in V_j$ ,

$$\begin{aligned} |T_1(a, b) - T_1(c, d)| &= \left| \sum_{i < \ell} \lambda_i u_i(a) u_i(b) - \lambda_i u_i(c) u_i(d) \right| \\ &\leq \sum_{i < \ell} |\lambda_i| |u_i(a) u_i(b) - u_i(c) u_i(b) + u_i(c) u_i(b) - u_i(c) u_i(d)| \\ &\leq \sum_{i < \ell} |\lambda_i| |u_i(b) (u_i(a) - u_i(c)) + u_i(c) (u_i(b) - u_i(d))| \\ &\leq \sum_{i < \ell} n |u_i(b)| |u_i(a) - u_i(c)| + n |u_i(c)| |u_i(b) - u_i(d)| \\ &\leq \ell n \left( 2\sqrt{\frac{\ell}{\varepsilon}} n^{-1/2} \cdot \frac{\varepsilon^{3/2}}{\ell^{3/2}} n^{-1/2} + 2\sqrt{\frac{\ell}{\varepsilon}} n^{-1/2} \cdot \frac{\varepsilon^{3/2}}{\ell^{3/2}} n^{-1/2} \right) \\ &= 4\varepsilon. \end{aligned}$$

934

935 Luego, para  $i, j \in [M]$ , defina  $d_{ij}$  como el promedio de los valores del bloque  $V_i \times V_j$  en  $T_1$  y  
 936 observe que se satisface

$$|T_1(a, b) - d_{ij}| \leq 4\varepsilon, \quad \forall a \in V_i, \forall b \in V_j.$$

937

938 En efecto, como  $d_{ij}$  es un promedio, deben existir los pares  $(x_0, y_0), (x_1, y_1) \in V_i \times V_j$  tales  
 939 que  $T_1(x_0, y_0) \leq d_{ij}$  y  $T_1(x_1, y_1) \geq d_{ij}$ . Luego, si  $|T_1(a, b) - d_{ij}| > 4\varepsilon$ , entonces se encuentra una  
 940 contradicci3n al determinar que  $T_1(a, b) - T_1(x_0, y_0) > 4\varepsilon$ , o bien  $T_1(a, b) - T_1(x_1, y_1) < -4\varepsilon$ .

941 Usando lo anterior y la desigualdad triangular, para todo  $A \subset V_i$  y  $B \subset V_j$ , obtenemos la  
 942 siguiente cota.

$$\begin{aligned}
|v_A^T(T_1 - d_{ij} \mathbb{1}_{n \times n})v_B| &\leq \sum_{a \in A} \sum_{b \in B} |T_1(a, b) - d_{ij}| \\
&\leq 4\varepsilon |A| |B| \\
&\leq 4\varepsilon |V_i| |V_j|.
\end{aligned} \tag{43}$$

943

944 **Analizamos  $T_2$ .** Observe en primer lugar, por construcción,

$$\text{Tr}(T_2^2) = \sum_{\ell \leq j < f(\ell)} \lambda_j^2 \leq \varepsilon^3 n^2.$$

945

946 Adicionalmente, por la ortonormalidad de la base,

$$\begin{aligned}
\sum_{a, b \in [n]} T_2(a, b)^2 &= \sum_{a, b \in [n]} \left( \sum_{\ell \leq i < f(\ell)} \lambda_i u_i(a) u_i(b) \right)^2 \\
&= \sum_{a, b \in [n]} \sum_{\ell \leq i, j < f(\ell)} \lambda_i \lambda_j u_i(a) u_j(a) u_i(b) u_j(b) \\
&= \sum_{\ell \leq i, j < f(\ell)} \lambda_i \lambda_j \sum_{a \in [n]} u_i(a) u_j(a) \sum_{b \in [n]} u_i(b) u_j(b) \\
&= \sum_{\ell \leq i < f(\ell)} \lambda_i^2 \|u_i\|^4 \\
&= \text{Tr}(T_2^2).
\end{aligned}$$

947

948 Entonces, dada la igualdad anterior, se determina que

$$\sum_{a, b \in [n]} T_2(a, b)^2 \leq \varepsilon^3 n^2. \tag{44}$$

949

950 Ahora, defina el conjunto  $\Theta_1 \subset [M]^2$  de manera tal que todo par  $(i, j) \notin \Theta_1$  satisface

$$\sum_{a \in V_i} \sum_{b \in V_j} T_2(a, b)^2 \leq \varepsilon |V_i| |V_j|. \tag{45}$$

951

952 Más aún, para los pares  $(i, j) \in \Theta_1$ , la desigualdad (44) en particular establece que

$$\varepsilon^3 n^2 \geq \sum_{(i, j) \in \Theta_1} \sum_{a \in V_i} \sum_{b \in V_j} T_2(a, b)^2 > \varepsilon \sum_{(i, j) \in \Theta_1} |V_i| |V_j|.$$



953

954 Por consecuencia,

$$\sum_{(i,j) \in \Theta_1} |V_i| |V_j| \leq \varepsilon^2 n^2. \quad (46)$$

955

956 De esta manera, para  $(i,j) \notin \Theta_1$ ,  $A \subset V_i$  y  $B \subset V_j$ , utilizamos la desigualdad (45) y Cauchy-  
 957 Schwarz para conseguir

$$\begin{aligned} |v_A^T T_2 v_B|^2 &= \left| \sum_{a \in A} \sum_{b \in B} T_2(a, b) \right|^2 \\ &\stackrel{\text{C-S}}{\leq} \left( \sum_{a \in A} \sum_{b \in B} T_2(a, b)^2 \right) |A| |B| \\ &\leq \varepsilon^2 |V_i| |V_j| |A| |B| \\ &\leq \varepsilon^2 |V_i|^2 |V_j|^2. \end{aligned}$$

958

959 Así, se obtiene la cota asociada a  $T_2$ .

$$|v_A^T T_2 v_B| \leq \varepsilon |V_i| |V_j|. \quad (47)$$

960

961 **Analizamos  $T_3$ .** Note que el valor propio más grande en magnitud de  $T_3$  es  $\lambda_{f(\ell)}$ . Entonces,  
 962 utilizando el operador norma \* Definir... \* de la matriz  $T_3$  y el Teorema 9,

$$\frac{\|T_3 v_B\|}{\|v_B\|} \leq \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|T_3 x\|}{\|x\|} = |\lambda_{f(\ell)}| \leq \frac{n}{\sqrt{f(\ell)}}.$$

963

964 Como resultado,

$$\|T_3 v_B\| \leq \|v_B\| \frac{n}{\sqrt{f(\ell)}}.$$

965

966 Usando la desigualdad anterior junto a Cauchy-Schwarz se obtiene la siguiente cota para  $T_3$ .

$$\begin{aligned}
|v_A^T T_3 v_B| &= |\langle v_A, T_3 v_B \rangle| \\
&\stackrel{\text{C-S}}{\leq} \|v_A\| \|T_3 v_B\| \\
&\leq \|v_A\| \|v_B\| \frac{n}{\sqrt{f(\ell)}} \\
&= \sqrt{|A||B|} \frac{n}{\sqrt{f(\ell)}} \\
&\leq \frac{n^2}{\sqrt{f(\ell)}}.
\end{aligned} \tag{48}$$

967

968 Ya con el control de  $T_1$ ,  $T_2$  y  $T_3$ , nos enfocamos en estudiar  $G$  de manera global. Consideramos  
969  $\Theta \subset \{0, 1, \dots, M\}^2$  definido de la siguiente manera:

$$\Theta := \left\{ (i, j) \in \{0, 1, \dots, M\}^2 : (i, j) \in \Theta_1 \vee i = 0 \vee j = 0 \vee \min\{|V_i|, |V_j|\} \leq \frac{\varepsilon n}{M} \right\}.$$

970

971 Con esta definición, la desigualdad (46), y recordando que  $|V_0| < \varepsilon n$ ,

$$\begin{aligned}
\sum_{(i,j) \in \Theta} |V_i| |V_j| &= \sum_{(i,j) \in \Theta_1} |V_i| |V_j| + \sum_{j=0}^M |V_0| |V_j| + \sum_{i=0}^M |V_i| |V_0| + \sum_{|V_i| \leq \frac{\varepsilon n}{M}} |V_i| |V_j| + \sum_{|V_j| \leq \frac{\varepsilon n}{M}} |V_i| |V_j| \\
&\leq \sum_{(i,j) \in \Theta_1} |V_i| |V_j| + 2|V_0|n + 2 \sum_{|V_i| \leq \frac{\varepsilon n}{M}} |V_i| n \\
&\leq \varepsilon^2 n^2 + 2\varepsilon n^2 + 2M \frac{\varepsilon}{M} n^2 \\
&\leq 5\varepsilon n^2.
\end{aligned}$$

972

973 Al ver la cota anterior,  $\Theta$  se interpreta como un conjunto excepcional de pocos elementos que  
974 contiene los malos casos. Ahora bien, si  $(i, j) \notin \Theta$ , todo  $A \subset V_i$  y  $B \subset V_j$  satisfacen la desigualdad

$$\begin{aligned}
\left| e(A, B) - d_{ij} |A| |B| \right| &= \left| v_A^T (T - d_{ij} \mathbb{1}_{n \times n}) v_B \right| \\
&\leq \left| v_A^T (T_1 - d_{ij} \mathbb{1}_{n \times n}) v_B \right| + \left| v_A^T T_2 v_B \right| + \left| v_A^T T_3 v_B \right| \\
&\leq 4\varepsilon |V_i| |V_j| + \varepsilon |V_i| |V_j| + \frac{n^2}{\sqrt{f(\ell)}}.
\end{aligned} \tag{49}$$

975

976 Observando la desigualdad en (49), para  $(i, j) \notin \Theta$ , se necesita que  $\frac{n^2}{\sqrt{f(\ell)}} \leq \varepsilon|V_i||V_j|$  para  
 977 asegurar que la partición  $\{V_0, V_1, \dots, V_M\}$  de  $[n]$  es  $(6\varepsilon)$ -regular. Para esto, gracias a que  $|V_i|, |V_j| \geq$   
 978  $\frac{\varepsilon n}{M}$ , se cumple la desigualdad  $\frac{\varepsilon^2 n^2}{M^2} \leq |V_i||V_j|$ , y por consecuencia

$$\frac{n^2}{\sqrt{f(\ell)}} \leq \frac{M^2|V_i||V_j|}{\varepsilon^2 \sqrt{f(\ell)}}.$$

979

980 Finalmente, para obtener la partición  $(6\varepsilon)$ -regular del conjunto de vértices del grafo  $G$ , es sufi-  
 981 ciente asumir que  $\frac{1}{\sqrt{f(\ell)}} \leq \frac{\varepsilon^3}{M^2}$ . Así, recordando la cota vista en (42), basta elegir

$$f(x) \geq \frac{1}{\varepsilon^6} \left( \frac{2x^2}{\varepsilon^2} \right)^{4x}.$$

982

□

## 983 5. Bibliografía

- 984 [1 ] Krivelevich, M., Sudakov, B. (2006). Pseudo-random Graphs. In Bolyai Society Mathema-  
 985 tical Studies (pp. 199–262). Springer Berlin Heidelberg.
- 986 [2 ] Chung, F. R. K., Graham, R. L., Wilson, R. M. (1989). Quasi-random graphs. Combinato-  
 987 rica. An International Journal on Combinatorics and the Theory of Computing.
- 988 [3 ] Chan, T. F. N., Král', D., Noel, J. A., Pehova, Y., Sharifzadeh, M., Volec, J. (2020). Charac-  
 989 terization of quasirandom permutations by a pattern sum. Random Structures Algorithms.
- 990 [4 ] Hàn, H., Kiwi, M., Pavez-Signé, M. (2021). Quasi-random words and limits of word sequen-  
 991 ces. Journal Europeen de Combinatoire [European Journal of Combinatorics].