

Tema 3

Grafos y árboles

(Apuntes elaborados por Juan Manuel Urbano Blanco)

Contents

1	Introducción	1
2	Conceptos básicos	6
3	Secuencias de grados	12
4	Caminos en un grafo	17
5	Grafos dirigidos	22
6	Tipos de caminos en un grafo	26
7	Coloración de grafos	34
8	Árboles y grafos bipartidos	36
9	Grafos planos	39
9.1	Una pequeña historia con historia	47
10	El algoritmo de Huffman	48

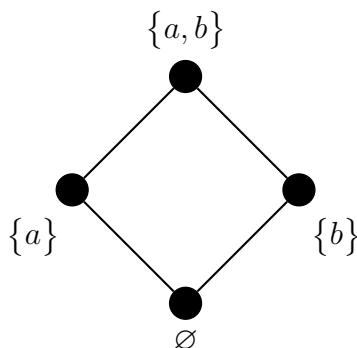
1 Introducción

Intuitivamente, un grafo viene dado por un conjunto V de objetos, llamados **vértices**, **puntos o nodos**, y otro conjunto L de objetos, llamados **lados o segmentos** (e incluso **flechas**, cuando el grafo sea dirigido) que unen nodos pertenecientes a V .

Los grafos proporcionan una herramienta para **representar** una o mas **relaciones binarias definidas sobre un conjunto de objetos**. Veamos algunos ejemplos.

1. El alumno ya conoce de la asignatura ALEM el *diagrama de Hasse* de un conjunto ordenado (A, \leq) . En este caso $V := A$ y L es el conjunto de pares ordenados $(a_i, a_j) \in A \times A$ tales que $a_i < a_j$ y no existe ningún elemento $a_k \in A$ verificando que $a_i < a_k < a_j$.

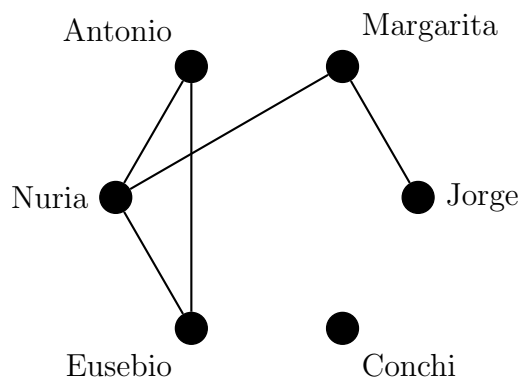
Por ejemplo, el diagrama de Hasse del conjunto ordenado $(P(\{a, b\}), \subseteq)$ es:



El diagrama de Hasse es un tipo concreto de grafo que nos permite representar (de la forma más económica posible) conjuntos ordenados. Aquí $V = P(\{a, b\})$ y $L = \{(\emptyset, \{a\}), (\emptyset, \{b\}), (\{a\}, \{a, b\}), (\{b\}, \{a, b\})\}$.

2. En un conjunto de personas, X , consideramos la relación binaria $a R b$ que indica que la persona a conoce a la persona b . Es razonable suponer que R verifica la propiedad simétrica.

Una situación concreta podría ser aquella en la cual las personas son Antonio, Margarita, Jorge, Conchi, Eusebio, y Nuria, y las conocencias vienen dadas por el diagrama siguiente:



Por tanto tenemos un grafo en el que

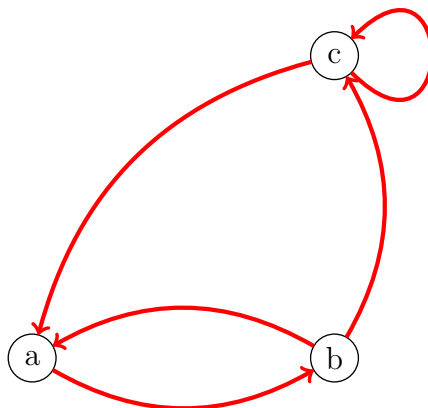
$$V = \{\text{Antonio, Margarita, Jorge, Conchi, Eusebio, Nuria}\}$$

y

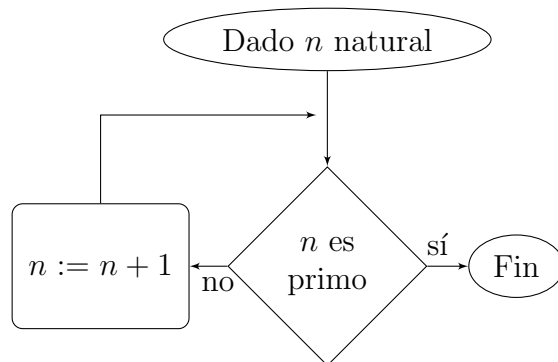
$$L = \{\{a, b\} : a, b \in X, a \neq b, a \text{ y } b \text{ se conocen}\}.$$

3. En un mapa turístico de una ciudad señalamos los edificios, monumentos y lugares más emblemáticos, así como las líneas de autobuses o de metro y las calles con el sentido de circulación para los vehículos. En este ejemplo los puntos son los edificios

y monumentos, y los lados son las líneas de autobuses, de metro y las calles. Nótese que pueden haber calles donde se permita la circulación en ambos sentidos y también pueden haber calles donde la circulación sea de sentido único. Por tanto en este ejemplo los segmentos o lados que unen a los vértices tendrán un sentido asignado.



4. El *diagrama de flujo* de un algoritmo es un grafo que representa a dicho algoritmo:



Todos estos ejemplos nos llevan a dar las siguientes definiciones.

Un **grafo dirigido u orientado** viene dado por dos conjuntos V y L , llamados **conjunto de vértices** y **conjunto de lados**, respectivamente, y por dos aplicaciones

$$\text{ini} : L \rightarrow V, \quad \text{fin} : L \rightarrow V,$$

denominadas **aplicaciones de incidencia**. Si $\ell \in L$, decimos que $\text{ini}(\ell)$ y $\text{fin}(\ell)$ son, respectivamente, el vértice inicial y el vértice final del lado ℓ .

Si tenemos un grafo dirigido G , lo usual es representarlo gráficamente dibujando un punto (o pequeño círculo) por cada elemento de V y a continuación para cada elemento $\ell \in L$, dibujamos un segmento orientado o flecha etiquetado con ℓ que va del punto $\text{ini}(\ell)$ al punto $\text{fin}(\ell)$.

Ejemplo 1. Sea el grafo dirigido G , donde:

- $V = \{v_1, v_2, v_3\}$,
- $L = \{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8\}$,
- y las aplicaciones de incidencia $\text{ini}, \text{fin} : L \rightarrow V$ vienen dadas por:

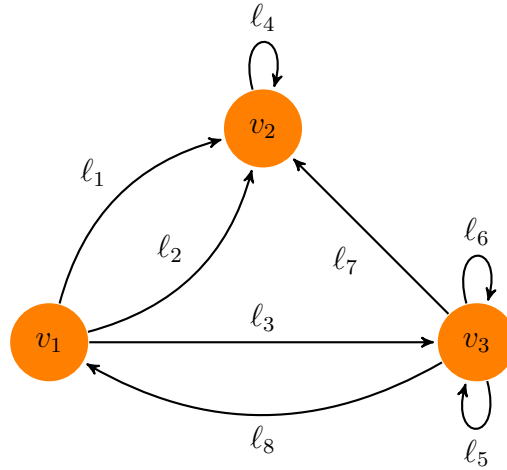
$$\text{ini}(\ell_1) = v_1, \text{ini}(\ell_2) = v_1, \text{ini}(\ell_3) = v_1, \text{ini}(\ell_4) = v_2,$$

$$\text{fin}(\ell_1) = v_2, \text{fin}(\ell_2) = v_2, \text{fin}(\ell_3) = v_3, \text{fin}(\ell_4) = v_2,$$

$$\text{ini}(\ell_5) = v_3, \text{ini}(\ell_6) = v_3, \text{ini}(\ell_7) = v_3, \text{ini}(\ell_8) = v_3$$

$$\text{fin}(\ell_5) = v_3, \text{fin}(\ell_6) = v_3, \text{fin}(\ell_7) = v_2, \text{fin}(\ell_8) = v_1.$$

Representamos gráficamente este grafo de la forma siguiente:



□

También es posible definir un grafo dirigido G dando los conjuntos V y L de vértices y lados, respectivamente, y una aplicación

$$\varphi : L \rightarrow V \times V$$

que se denominada aplicación de incidencia, donde

$$\varphi(\ell) = (\text{ini}(\ell), \text{fin}(\ell)).$$

Escribimos en tal caso $G = (V, L, \varphi)$ para denotar al grafo dirigido G .

Para el grafo del ejemplo anterior, tenemos que

$$\varphi(\ell_1) = (v_1, v_2), \varphi(\ell_2) = (v_1, v_2), \varphi(\ell_3) = (v_1, v_3), \varphi(\ell_4) = (v_2, v_2),$$

$$\varphi(\ell_5) = (v_3, v_3), \varphi(\ell_6) = (v_3, v_3), \varphi(\ell_7) = (v_3, v_2), \varphi(\ell_8) = (v_3, v_1).$$

Un **grafo no dirigido** o **multigrafo** viene dado por dos conjuntos V y L , llamados conjunto de vértices y conjunto de lados, respectivamente, y una aplicación

$$f : L \rightarrow \left\{ \{v\} : v \in V \right\} \cup \left\{ \{u, v\} : u, v \in V, \text{ con } u \neq v \right\},$$

llamada aplicación de incidencia.

Si tenemos un multigrafo G , podemos representarlo gráficamente dibujando un punto por cada elemento de V , y a continuación para cada elemento $\ell \in L$, dibujamos un segmento (no orientado) etiquetado con ℓ que une los puntos correspondientes a u y v si $f(\ell) = \{u, v\}$, o bien une el punto correspondiente a v consigo mismo, si $f(\ell) = \{v\}$.

Ejemplo 2. Sea el multigrafo G dado por:

- $V = \{v_1, v_2, v_3, v_4, v_5\}$,
- $L = \{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6, \ell_7, \ell_8, \ell_9, \ell_{10}, \ell_{11}\}$,
- y la aplicación de incidencia f , viene definida por:

$$f(\ell_1) = f(\ell_2) = \{v_1\},$$

$$f(\ell_3) = f(\ell_4) = f(\ell_5) = \{v_1, v_2\},$$

$$f(\ell_6) = f(\ell_7) = \{v_1, v_3\},$$

$$f(\ell_8) = \{v_2\},$$

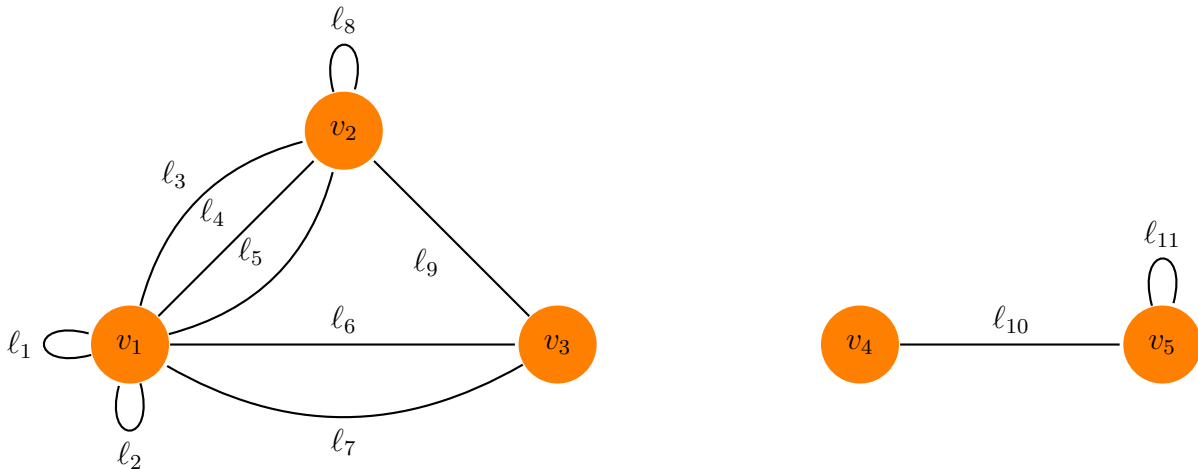
$$f(\ell_9) = \{v_2, v_3\},$$

$$f(\ell_{10}) = \{v_4, v_5\},$$

$$f(\ell_{11}) = \{v_5\}.$$

□

Representamos gráficamente este grafo de la forma siguiente:

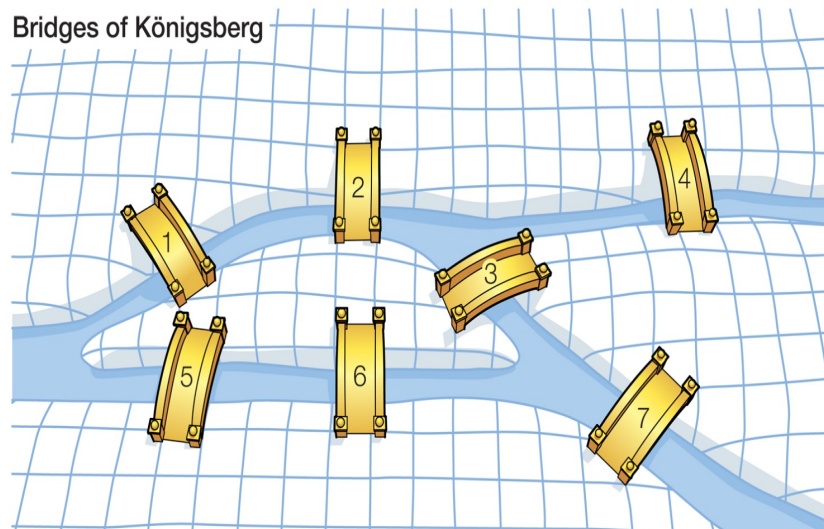


□

Dado un grafo G , ya sea dirigido o no dirigido, denotaremos por $V(G)$ el conjunto de los vértices de G y por $L(G)$ el conjunto de los lados de G .

A una representación gráfica de un grafo G , también la llamaremos grafo, aunque no hemos de olvidar que realmente el grafo es el objeto formal dado por los conjuntos $V(G)$, $L(G)$ y las aplicaciones de incidencia.

La teoría de grafos se originó en el siglo XVIII con el llamado **problema de los puentes de Königsberg**. En esta ciudad, antaño perteneciente a Prusia oriental y actualmente llamada Kaliningrado y perteneciente a Rusia, se encuentra el río Pregel, el cual dejaba dos islotes que se comunicaban con las orillas con un total de siete puentes.



A alguien se le ocurrió el siguiente problema. ¿Es posible comenzar en alguna orilla o islote, y recorrer todos los puentes pasando por cada puente una sola vez? Muchos lugareños y turistas intentaban in situ resolver el problema sin éxito. Fue **Leonhard Euler** (Basilea, Suiza, 15 de abril de 1707 - San Petersburgo, Rusia, 18 de septiembre de 1783) quien resolvió este problema como veremos más adelante.

2 Conceptos básicos

Sea G un grafo no dirigido. Si $\ell \in L(G)$ y $f(\ell) = \{u\}$, se dice que ℓ es un **autolazo** de G . Si $\ell_1, \ell_2 \in L(G)$, distintos, y $f(\ell_1) = f(\ell_2) = \{v, w\}$, con $v \neq w$, se dice que ℓ_1 y ℓ_2 son **lados paralelos** de G .

En el grafo del Ejemplo 2, los autolazos son ℓ_1, ℓ_2, ℓ_8 y ℓ_{11} . Por otra parte, ℓ_3, ℓ_4 y ℓ_5 son lados paralelos entre sí. También son lados paralelos ℓ_6 y ℓ_7 .

A partir de ahora en este tema, y salvo que se diga otra cosa, la palabra **grafo** se referirá a un grafo no dirigido, sin autolazos ni lados paralelos. Por tanto la definición de grafo que vamos a adoptar es la siguiente.

Un **grafo** G viene dado por dos conjuntos V y L , siendo V un conjunto cuyos elementos se denominan vértices y L un conjunto cuyos elementos son subconjuntos de V de cardinal 2. Los elementos de L se denominan lados. Escribiremos $G = (V, L)$ para indicar un grafo G cuyo conjunto de vértices es V y cuyo conjunto de lados es L .

Dos **vértices** $u, v \in V(G)$ son **adyacentes** si $\{u, v\} \in L(G)$. En tal caso escribimos $u \sim v$. Dos **lados** distintos $\ell_1, \ell_2 \in L(G)$ son **adyacentes** si tienen un vértice en común, es decir, $\ell_1 \cap \ell_2 = \{v\}$. Un **lado** $\ell \in L(V)$ y un **vértice** $v \in V(L)$ son **incidentes** si $v \in \ell$.

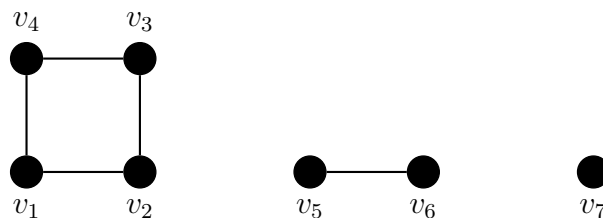
El **grado de un vértice** $v \in V(G)$ es el número de vértices de G que son adyacentes con v . Equivalentemente, el número de lados de G que son incidentes con v . Designaremos por $\text{gr}_G(v)$ el grado del vértice v de G . Un **grafo** G se dice **regular** de grado d , si todo vértice de G tiene grado d .

Si $V(G) = \{v_1, \dots, v_n\}$, diremos que $\text{gr}_G(v_1), \dots, \text{gr}_G(v_n)$ es una **secuencia de grados** del grafo G .

Ejemplo 3. Sea el grafo $G = (V, L)$, donde

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\} \quad \text{y} \quad L = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_4\}, \{v_4, v_1\}, \{v_5, v_6\}\}.$$

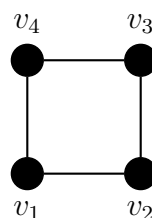
Representamos G de la forma siguiente:



Entonces vemos que los vértices v_2 y v_3 son adyacentes mientras que los vértices v_2 y v_4 no lo son. El grado de cada uno de los vértices v_1, v_2, v_3 y v_4 es 2, el grado de v_5 y v_6 es 1, y el grado de v_7 es 0. Por tanto G no es un grafo regular.

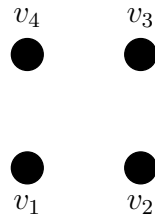
Una secuencia de grados de G es 2, 2, 2, 2, 1, 1, 0. □

Ejemplo 4. El grafo G dado por el dibujo siguiente, es regular de grado 2:



□

Ejemplo 5. El grafo G dado por el dibujo siguiente, es regular de grado 0:



La secuencia de grados de G es 0, 0, 0, 0.

□

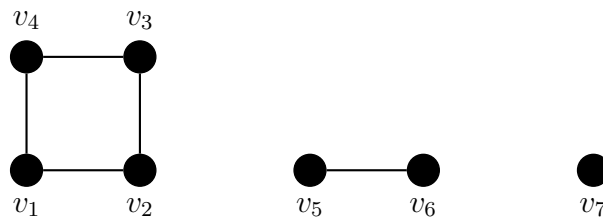
Proposición 1 (Teorema de Euler). En todo grafo G se verifica que

$$\sum_{v \in V(G)} \text{gr}_G(v) = 2 \cdot |L(G)|.$$

□

La justificación de esta propiedad es que cada lado de G aporta dos unidades a la suma total de grados de G .

Ejemplo 6. Ilustramos el teorema de Euler para el grafo G siguiente:



Por una parte,

$$\sum_{v \in V(G)} \text{gr}_G(v) = 2 + 2 + 2 + 2 + 1 + 1 + 0 = 10.$$

Como $|L(G)| = 5$, por otra parte tenemos que

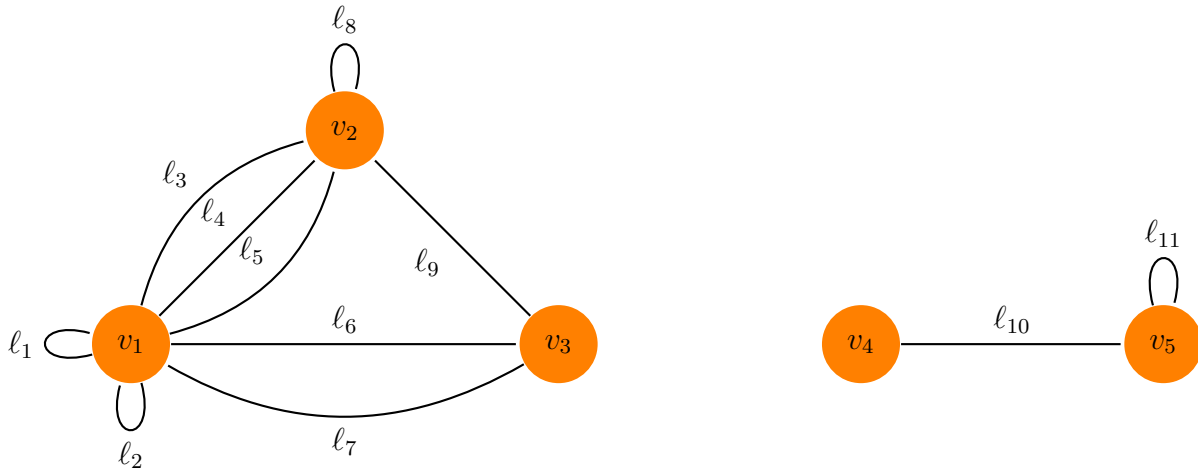
$$2 \cdot |L(G)| = 2 \cdot 5 = 10.$$

□

El Teorema de Euler puede ser extendido al caso de multigrafos. Para ello hemos de tener en cuenta que **cada autolazo aporta dos unidades al grado** del vértice correspondiente.

Para ilustrar esta idea, consideremos de nuevo el grafo del Ejemplo 2 en el que

$$|V(G)| = 5 \quad \text{y} \quad |L(G)| = 11.$$



Tenemos los siguientes grados:

$$\text{gr}(v_1) = 9, \quad \text{gr}(v_2) = 6, \quad \text{gr}(v_3) = 3, \quad \text{gr}(v_4) = 1, \quad \text{gr}(v_5) = 3.$$

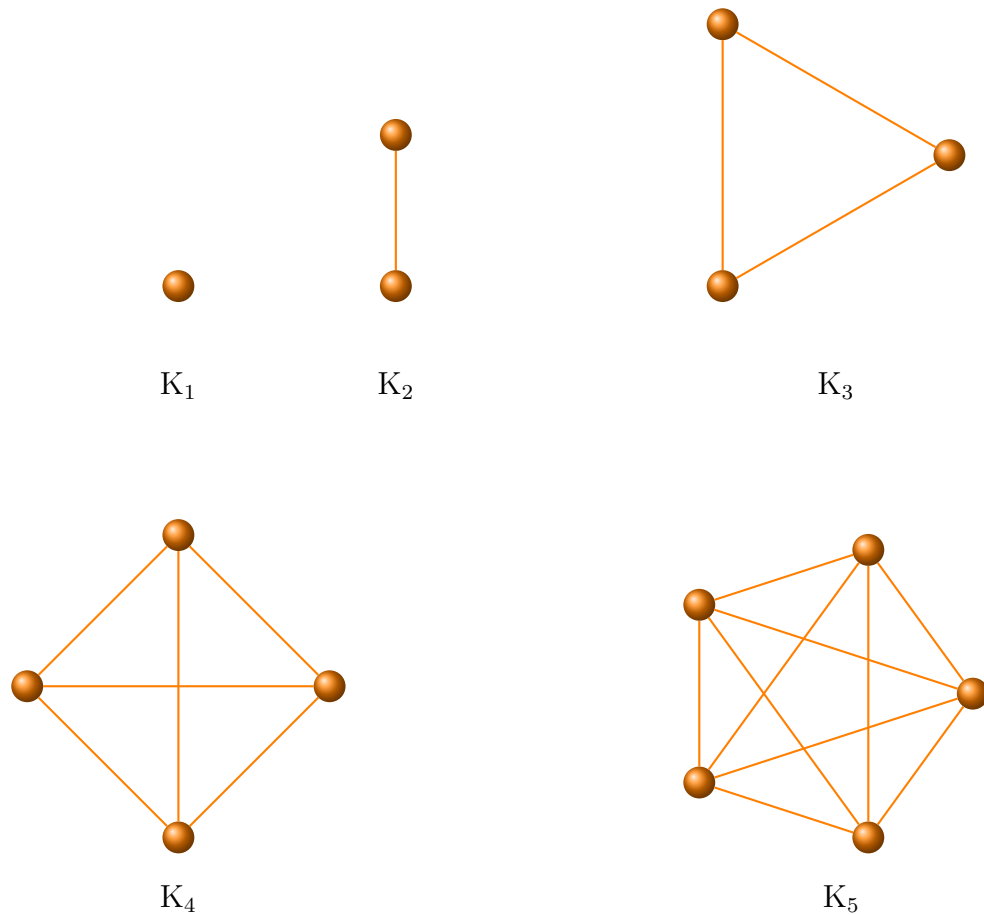
Entonces se cumple que

$$\sum_{v \in V(G)} \text{gr}_G(v) = 9 + 6 + 3 + 1 + 3 = 22$$

que es precisamente igual a

$$2 \cdot |L(G)| = 2 \cdot 11.$$

Un grafo G se dice **completo**, cuando cualesquiera dos vértices de G , distintos, son adyacentes. Es costumbre escribir K_n para representar un grafo completo de n vértices. Los primeros grafos completos, en los que no hemos escrito los nombres de los vértices por no ser ello relevante en este momento, son:



La siguiente propiedad es una consecuencia inmediata de las definiciones anteriores.

Proposición 2. K_n es un grafo regular de grado $n - 1$. Además

$$|L(K_n)| = \binom{n}{2}.$$

□

El **complemento o complementario** de un grafo G , es un nuevo grafo denotado por \overline{G} y definido como sigue:

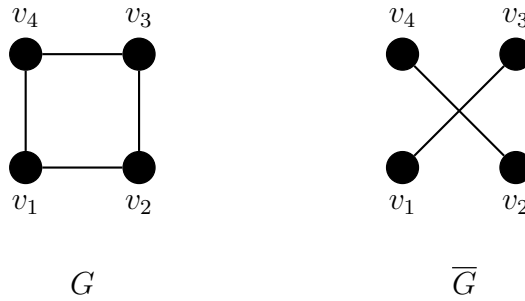
$$V(\overline{G}) = V(G),$$

$$L(\overline{G}) = \left\{ \{u, v\} : u, v \in V(G), u \neq v, \{u, v\} \notin L(G) \right\}.$$

Es decir, \overline{G} es un nuevo grafo cuyos vértices son los mismos de G y cuyos lados son aquellos que le faltan a G para ser un grafo completo. Claramente

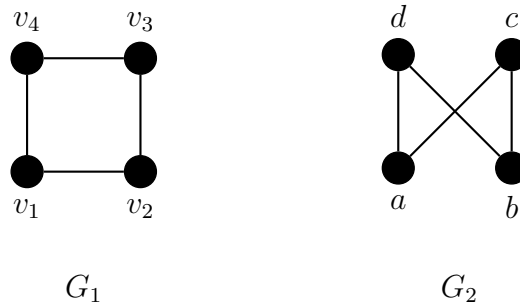
$$\overline{\overline{G}} = G.$$

Ejemplo 7. Los dos grafos siguientes son complementarios:

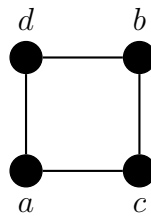


□

Consideremos los grafos siguientes.



Claramente, el grafo G_2 también lo podemos dibujar de la manera siguiente,



con lo cual vemos que el grafo G_2 puede ser dibujado de modo que tenga la misma forma que el grafo G_1 . Ésta observación nos permite dar la siguiente definición.

Dos **grafos** G_1 y G_2 son **isomorfos**, si existe una aplicación biyectiva

$$\varphi : V(G_1) \rightarrow V(G_2)$$

tal que para cualesquiera $u, v \in V(G_1)$ se cumple que

$$\{u, v\} \in L(G_1) \quad \text{si y sólo si} \quad \{\varphi(u), \varphi(v)\} \in L(G_2).$$

En tal caso se dice que φ es un **isomorfismo de G_1 en G_2** . Es inmediato probar que la aplicación inversa φ^{-1} es un isomorfismo de G_2 en G_1 . Escribiremos $G_1 \cong G_2$ para indicar que G_1 y G_2 son isomorfos.

Así, los dos grafos G_1 y G_2 que hemos considerado anteriormente, son isomorfos y la aplicación $\varphi : V(G_1) \rightarrow V(G_2)$ dada por

$$\varphi(v_1) = a, \quad \varphi(v_2) = c, \quad \varphi(v_3) = b, \quad \varphi(v_4) = d,$$

es un isomorfismo de G_1 en G_2 . La aplicación $\sigma : V(G_1) \rightarrow V(G_2)$ dada por

$$\sigma(v_1) = d, \quad \sigma(v_2) = a, \quad \sigma(v_3) = c, \quad \sigma(v_4) = b,$$

es otro isomorfismo de G_1 en G_2 .

Intuitivamente, dos grafos son isomorfos cuando ambos tienen una misma representación gráfica, salvo, como es lógico, los nombres con los que etiquetamos los nodos en uno y en otro grafo. Cada isomorfismo indica una forma de “superponer” un grafo sobre el otro.

Si $\varphi : V(G_1) \rightarrow V(G_2)$ es un isomorfismo del grafo G_1 en G_2 , entonces es obvio que para cualquier $v \in V(G_1)$ se verifica que

$$\text{gr}_{G_1}(v) = \text{gr}_{G_2}(\varphi(v)).$$

El **problema del isomorfismo de grafos**, consiste en dados dos grafos, decidir si ambos son o no son isomorfos.

Proposición 3. Si los grafos G_1 y G_2 son isomorfos, entonces:

1. $|V(G_1)| = |V(G_2)|$.
2. $|L(G_1)| = |L(G_2)|$.
3. Para todo $d \in \mathbb{N}$, el número de vértices en G_1 con grado d es igual al número de vértices en G_2 con grado d .

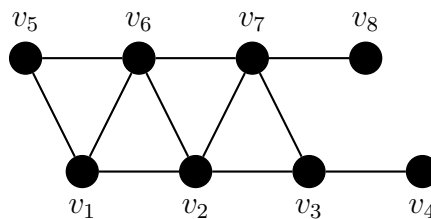
□

Esta propiedad nos da tres **condiciones necesarias** para que dos grafos puedan ser isomorfos. Sin embargo, ni tan siquiera las tres juntas son suficientes para que dos grafos sean isomorfos, tal y como veremos en los ejercicios.

3 Secuencias de grados

La **secuencia de grados de un grafo** es aquella formada por los grados de sus vértices ordenados de manera decreciente. Tal y como ya se estableció anteriormente, si tenemos una lista con los grados de los vértices en un grafo y no nos preocupamos de ordenar (de mayor a menor) los valores que aparecen en dicha lista, decimos que tenemos una secuencia de grados del grafo.

Ejemplo 8. Para el grafo siguiente,



la secuencia de grados es:

$$4, 4, 4, 3, 3, 2, 1, 1.$$

Si escribimos

$$1, 2, 4, 3, 4, 3, 1, 4,$$

tenemos una secuencia de grados del grafo dado. □

Como consecuencia del comentario que sigue a la Proposición 3, dos grafos pueden tener la misma secuencia de grados y no ser isomorfos.

En esta sección vamos a dar respuesta a la siguiente pregunta.

Dada una secuencia de números naturales, ¿cómo sabemos si ésta es una secuencia de grados de algún grafo?, y en caso afirmativo, ¿cómo encontramos al menos un grafo con dicha secuencia de grados?

Una **secuencia** (no necesariamente decreciente) de números naturales se dice que es **gráfica**, si es una secuencia de grados de algún grafo.

Ejemplo 9.

1. La secuencia $0, 0, \dots, 0$ formada por n ceros, es gráfica.
2. La secuencia $1, 1, \dots, 1$ formada por $2n$ unos, es gráfica.
3. La secuencia $1, 1, \dots, 1$ formada por $2n + 1$ unos, no es gráfica, pues ya sabemos que en cualquier grafo el número de vértices de grado impar, ha de ser par.
4. La secuencia $4, 3, 3, 2, 2, 1$ tampoco es gráfica por la misma razón que la anterior.
5. La secuencia $2, 2, 2, 2, 2$ es gráfica.
6. La secuencia $5, 4, 3, 2, 2$ no es gráfica, pues en un grafo de cinco vértices, el grado de cada vértice es a lo sumo 4.

□

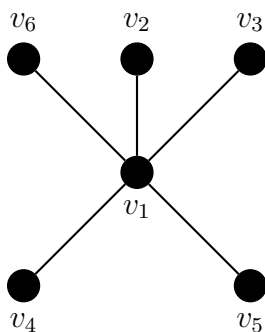
En vista de los ejemplos anteriores, deducimos que si d_1, d_2, \dots, d_n es una secuencia gráfica, entonces es necesario que

$$\sum_{j=1}^n d_j \text{ es par y } 0 \leq d_j \leq n-1 \text{ para todo } j = 1, \dots, n.$$

Sin embargo estas dos condiciones no son suficientes como pone de manifiesto el ejemplo siguiente.

Ejemplo 10. La secuencia 5, 4, 4, 2, 2, 1 de longitud 6 no es gráfica aún cuando cada valor está comprendido entre 0 y $6-1$, y la suma $5+4+4+2+2+1=18$ es par.

Supongamos que la secuencia sí fuese gráfica. Entonces existiría un vértice v_1 de grado 5, el cual por tanto habría de ser adyacente con todos los demás vértices v_2, v_3, v_4, v_5, v_6 del grafo. Dibujamos esta situación:



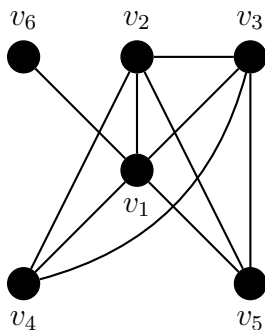
Además en G ha de haber exactamente un vértice de grado 1, y dos vértices de grado 4. Supongamos por ejemplo que

$$\text{gr}_G(v_6) = 1 \quad \text{y} \quad \text{gr}_G(v_2) = \text{gr}_G(v_3) = 4,$$

y por tanto

$$\text{gr}_G(v_4) = 2 \quad \text{y} \quad \text{gr}_G(v_5) = 2.$$

En base a esta información, ampliamos el grafo anterior, añadiendo los lados que faltan. Resulta:



Ésto nos lleva a que $\text{gr}_G(v_4) = \text{gr}_G(v_5) = 3$, lo que contradice las hipótesis anteriores. Por tanto la secuencia dada no es gráfica. \square

Necesitamos un método sistemático para dar respuesta a las preguntas anteriores. Éste vendrá dado por el siguiente teorema y el algoritmo resultante.

Teorema 1 (de **Havel–Hakimi**). Sea $s : d_1, d_2, \dots, d_n$ una secuencia decreciente de números naturales. Entonces s es gráfica si y sólo si la secuencia

$$s' : d_2 - 1, d_3 - 1, \dots, d_{d_1+1} - 1, d_{d_1+2}, \dots, d_n$$

es gráfica. \square

Nótese que en el teorema anterior, la secuencia s' se obtiene a partir de s suprimiendo el elemento que aparece más a la izquierda, y a continuación a cada uno de los primeros d_1 elementos por la izquierda se le resta 1.

Aplicamos este resultado para resolver algunos ejemplos.

Ejemplo 11. Veamos de nuevo que la secuencia 5, 4, 4, 2, 2, 1 no es gráfica.

Tenemos las siguientes secuencias.

$$\begin{aligned} 5, \overbrace{4, 4, 2, 2, 1} &\Rightarrow 3, 3, 1, 1, 0 \\ 3, \overbrace{3, 1, 1, 0} &\Rightarrow 2, 0, 0, 0 \end{aligned}$$

Como la secuencia última obtenida, 2, 0, 0, 0, no es gráfica, por el Teorema de Havel-Hakimi deducimos que la secuencia inicial tampoco lo es. \square

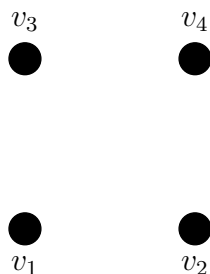
Ejemplo 12. Comprobamos que la secuencia 5, 4, 4, 3, 2, 1, 1 es gráfica.

Tenemos las siguientes secuencias.

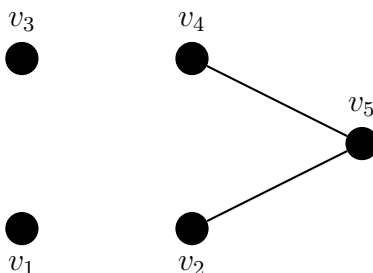
$$\begin{aligned} 5, \overbrace{4, 4, 3, 2, 1} &\Rightarrow 3, 3, 2, 1, 0, 1 \\ 3, \overbrace{3, 2, 1} &\Rightarrow 2, 1, 0, 1, 0 \\ 2, \overbrace{1, 1} &\Rightarrow 0, 0, 0, 0 \end{aligned}$$

Como la secuencia última obtenida 0, 0, 0, 0 es gráfica, por el Teorema de Havel-Hakimi tenemos que la secuencia dada también lo es.

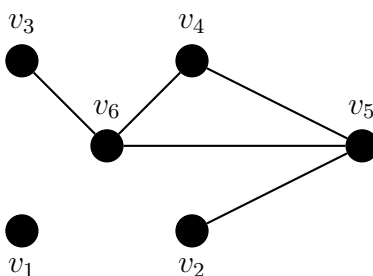
A partir de esta información es además inmediato construir un grafo concreto cuya secuencia de grados es la secuencia inicial. Para ello, partimos de la última secuencia obtenida, 0, 0, 0, 0, para la cual dibujamos el grafo siguiente G_0 :



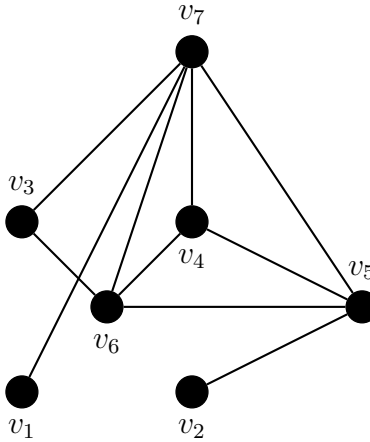
Como la secuencia 0, 0, 0, 0 procede de la secuencia 2, 1, 1, 0, 0, añadimos al grafo anterior un nuevo vértice v_5 que será adyacente con dos vértices que en G_0 tengan grado cero. Resulta el grafo G_1 :



Ahora, la secuencia 2, 1, 1, 0, 0 la hemos obtenido tras ordenar la secuencia 2, 1, 0, 1, 0, la cual procede de la secuencia 3, 3, 2, 1, 1, 0. Por tanto, añadimos a G_1 un vértice nuevo v_6 que será adyacente con tres vértices de G_1 . Concretamente, con uno de grado 2, con otro de grado 1, y con un tercer vértice de grado 0. Obtenemos así el grafo G_2 :



Por último, consideramos la secuencia 3, 3, 2, 1, 1, 0 que proviene de ordenar la secuencia 3, 3, 2, 1, 0, 1, y ésta a su vez de la secuencia 5, 4, 4, 3, 2, 1, 1. Por tanto, añadimos a G_2 un nuevo vértice v_7 que será adyacente con cinco vértices de G_2 , cuyos grados respectivos eran 3, 3, 2, 1, 0. Resulta finalmente el grafo G_3 , cuya secuencia de grados es la secuencia inicial:



□

A continuación damos un algoritmo que implementa el método estudiado en los ejemplos anteriores.

Algoritmo (de **Havel-Hakimi**).

ENTRADA: Una secuencia decreciente s de números naturales.

SALIDA: SÍ, en caso de que s sea gráfica, y NO, en caso contrario.

1. $s' := s$.
2. Si en s' aparece algún elemento mayor que $\ell - 1$, donde ℓ es la longitud de s' , devolver NO y terminar.
3. Si s' está formada sólo por ceros, devolver SÍ y terminar.
4. Si s' contiene algún número negativo, devolver NO y terminar.
5. Reordenar s' , si es necesario, de modo que ésta sea decreciente.
6. Eliminar el primer elemento d_1 en s' y restarle 1 a cada uno de los d_1 elementos siguientes. Llamar a la secuencia resultante s' e ir al paso (2). □

4 Caminos en un grafo

Un **camino de longitud k** en un grafo G , es una secuencia formada por k lados

$$\ell_1, \ell_2, \dots, \ell_k$$

de G tales que ℓ_j y ℓ_{j+1} son adyacentes para todo $j = 1, \dots, k - 1$.

Ya que estamos considerando grafos sin lados paralelos, podemos dar un camino de longitud k simplemente escribiendo los vértices por los que va pasando el camino. De esta

forma, si $\ell_j = \{v_{j-1}, v_j\}$ para $j = 1, \dots, k$, escribiremos el camino referido simplemente como

$$v_0, v_1, \dots, v_{k-1}, v_k,$$

y diremos que es un camino de v_0 a v_k .

Si $u, v \in V(G)$ son dos vértices distintos, se dice que u está **conectado** con v , si existe en G al menos un camino de u a v . En tal caso, cualquier camino de longitud mínima que conecta a u con v se denomina una **geodésica** de u a v .

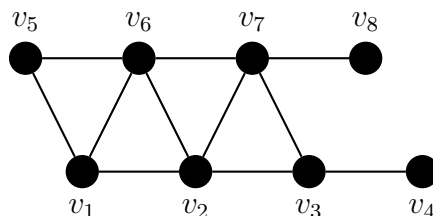
La **distancia entre dos vértices conectados** u y v de G , denotada por $d(u, v)$, es el mínimo de las longitudes de los caminos que conectan a u con v , es decir, la longitud de cualquier geodésica de u a v .

Ya que estamos considerando grafos no dirigidos, es inmediato que $d(u, v) = d(v, u)$.

Por definición, todo vértice $v \in V(G)$ está conectado consigo mismo mediante el camino vacío, que es aquel en el que no hay ningún lado y cuya longitud es 0. Entonces $d(v, v) = 0$.

Un **grafo** G se dice que es **conexo** cuando dos vértices cualesquiera de G están conectados.

Ejemplo 13. El grafo siguiente es conexo:



La secuencia de vértices

$$v_1, v_5, v_6, v_7, v_3, v_7, v_8$$

es un camino de longitud 6 del vértice v_1 al vértice v_8 .

Entre los vértices v_1 y v_8 hay exactamente dos geodésicas que son:

$$v_1, v_6, v_7, v_8 \quad \text{y} \quad v_1, v_2, v_7, v_8.$$

Por tanto, la distancia entre v_1 y v_8 es igual a 3, es decir, $d(v_1, v_8) = 3$.

Nótese que

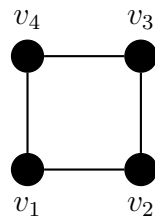
$$v_2, v_3, v_7, v_2 \quad \text{y} \quad v_2, v_3, v_7, v_2, v_3, v_7, v_2$$

son ambos caminos de v_2 a v_2 de longitudes 3 y 6, respectivamente. \square

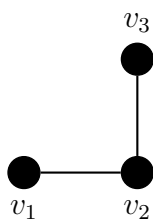
Un grafo $G_1 = (V_1, L_1)$ se dice que es un **subgrafo** de un grafo $G_2 = (V_2, L_2)$ si $V_1 \subseteq V_2$ y $L_1 \subseteq L_2$. Escribiremos en tal caso $G_1 \subseteq G_2$.

De esta forma hemos definido una relación de orden en la clase de todos los grafos.

Ejemplo 14. Sea el grafo G siguiente:



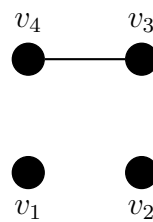
Cada uno de los grafos siguientes es un subgrafo de G :



G_1



G_2

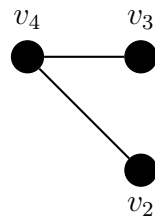


G_3



G_4

Sin embargo, el grafo siguiente, no es un subgrafo de G :



pues el lado $\{v_2, v_4\}$ no aparece en G . □

Una **componente conexa** de un grafo G es un subgrafo conexo maximal de G , es decir, es un subgrafo H de G tal que:

1. H es conexo.
2. Si H' es un subgrafo conexo de G tal que $H \subseteq H'$, entonces $H = H'$.

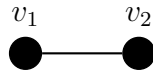
Claramente un grafo G es conexo si y sólo si G consta de una única componente conexa.

Ejemplo 15. Consideramos el grafo G siguiente:



Los siguientes hechos son ciertos para G :

- En G hay tres componentes conexas, por lo que G no es un grafo conexo.
- Los vértices v_1 y v_5 no están conectados.
- El grafo siguiente, es un subgrafo de G , pero no es una componente conexa de G :



- El grafo siguiente es una componente conexa de G :



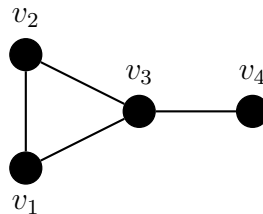
□

Dado un grafo $G = (V, L)$, con $V = \{v_1, \dots, v_n\}$, la **matriz de adyacencia** de G , denotada por $A(G)$, es una matriz $n \times n$ con coeficientes en \mathbb{R} definida como sigue. Si escribimos $A(G) = (a_{ij})$, entonces

$$a_{ij} = \begin{cases} 0 & \text{si } i = j \\ 0 & \text{si } i \neq j \text{ y } \{v_i, v_j\} \notin L(G) \\ 1 & \text{si } i \neq j \text{ y } \{v_i, v_j\} \in L(G) \end{cases}$$

Es inmediato que $A(G)$ es simétrica y el número de unos en la fila (columna) i -ésima de $A(G)$ es igual a $\text{gr}_G(v_i)$ para cada $i = 1, \dots, n$. Nótese además que la matriz de adyacencia depende del orden en el que escribimos los vértices del grafo G , en nuestro caso, v_1, \dots, v_n .

Ejemplo 16. Sea G el grafo siguiente:



La matriz de adyacencia de G , para la ordenación v_1, v_2, v_3, v_4 de sus vértices, es

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Si consideramos la ordenación v_3, v_2, v_4, v_1 de sus vértices, la matriz de adyacencia es

$$B = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

□

La matriz de adyacencia de un grafo G proporciona una forma de representar a G aunque de manera muy redundante. Piense en un grafo con 1000 vértices y un sólo lado. El interés en el estudio de la matriz de adyacencia de un grafo es principalmente teórico como pone de manifiesto la propiedad siguiente.

Proposición 4 (Teorema del número de caminos). Sea un grafo $G = (V, L)$, con $V = \{v_1, \dots, v_n\}$. Entonces para todo $k \in \mathbb{N}$, el número de caminos en G de longitud k desde v_i a v_j viene dado por el elemento que aparece en la fila i -ésima y en la columna j -ésima de la matriz $A(G)^k$.

Ejemplo 17. Consideramos de nuevo el grafo G del Ejemplo 16 cuya matriz de adyacencia es

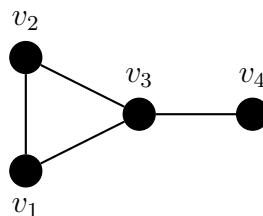
$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Calculamos las primeras potencias de A :

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 3 & 2 & 4 & 1 \\ 4 & 4 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{pmatrix},$$

$$A^4 = \begin{pmatrix} 7 & 6 & 6 & 4 \\ 6 & 7 & 6 & 4 \\ 6 & 6 & 11 & 2 \\ 4 & 4 & 2 & 3 \end{pmatrix}, \quad A^5 = \begin{pmatrix} 12 & 13 & 17 & 6 \\ 13 & 12 & 17 & 6 \\ 17 & 17 & 14 & 11 \\ 6 & 6 & 11 & 2 \end{pmatrix}.$$

Una representación de G es:



Obtenemos la siguiente información:

1. Los valores que aparecen la diagonal principal de A^2 son los grados de los vértices.

2. En la fila 2, columna 3 de A^3 aparece el valor 4, lo que nos indica que hay exactamente 4 caminos de longitud 3 desde v_2 hasta v_3 . Éstos son:

v_2, v_3, v_2, v_3
 v_2, v_3, v_4, v_3
 v_2, v_3, v_1, v_3
 v_2, v_1, v_2, v_3

3. El número de caminos de longitud 5 que empiezan y acaban en el vértice v_4 es 2. Éstos son:

$v_4, v_3, v_1, v_2, v_3, v_4$
 $v_4, v_3, v_2, v_1, v_3, v_4$

□

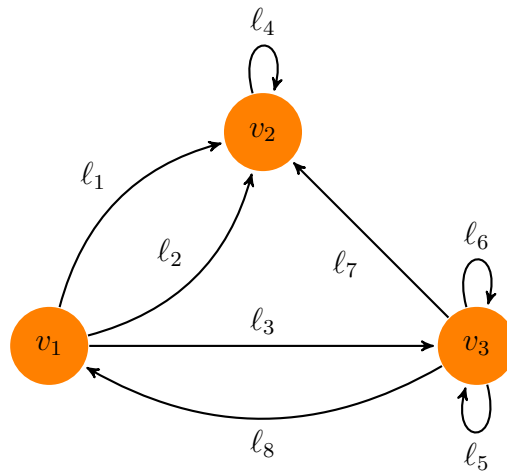
5 Grafos dirigidos

Comentamos en esta sección algunos aspectos que hemos estudiado para (multi)grafos y que son extensibles a grafos dirigidos.

Comenzamos viendo la versión del **Teorema de Euler para grafos dirigidos**. Antes necesitamos las siguientes definiciones.

Si G es un grafo dirigido y $v \in V(G)$, se define el **grado de salida** de v como el número de flechas que salen de v , y el **grado de entrada** de v como el número de flechas que llegan a v . Ambas cantidades son denotadas respectivamente por $\text{gr}_G^+(v)$ y $\text{gr}_G^-(v)$. Nótese que cada autolazo dirigido aporta una unidad al grado de salida y otra al grado de entrada del vértice correspondiente.

Consideremos de nuevo el grafo dirigido del Ejemplo 1.



Se tiene que:

$$\begin{aligned} \text{gr}_G^+(v_1) &= 3, & \text{gr}_G^+(v_2) &= 1, & \text{gr}_G^+(v_3) &= 4, \\ \text{gr}_G^-(v_1) &= 1, & \text{gr}_G^-(v_2) &= 4, & \text{gr}_G^-(v_3) &= 3. \end{aligned}$$

Proposición 5 (Teorema de Euler para grafos dirigidos). En todo grafo dirigido G se verifica que

$$\sum_{v \in V(G)} \text{gr}_G^+(v) = |L(G)| = \sum_{v \in V(G)} \text{gr}_G^-(v).$$

□

Si aplicamos este teorema al grafo dirigido anterior, obtenemos

$$3 + 1 + 4 = 8 = 1 + 4 + 3.$$

También se define el concepto de **isomorfismo de grafos dirigidos**. Dos grafos dirigidos $G_1 = (V_1, L_1, \varphi_1)$ y $G_2 = (V_2, L_2, \varphi_2)$ son isomorfos si existen aplicaciones biyectivas

$$\theta : V_1 \rightarrow V_2 \quad \text{y} \quad \sigma : L_1 \rightarrow L_2$$

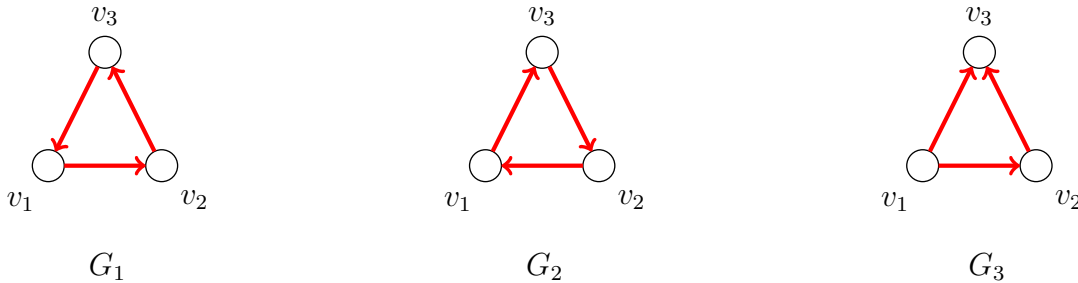
tales que

$$\varphi_1(\ell) = (u, v) \quad \text{si y sólo si} \quad \varphi_2(\sigma(\ell)) = (\theta(u), \theta(v)),$$

para todo $\ell \in L_1$ y $u, v \in V_1$. En tal caso se escribe $G_1 \cong G_2$.

Intuitivamente ésto significa que ambos grafos se pueden dibujar de la misma forma, teniendo en cuenta siempre las direcciones de los lados.

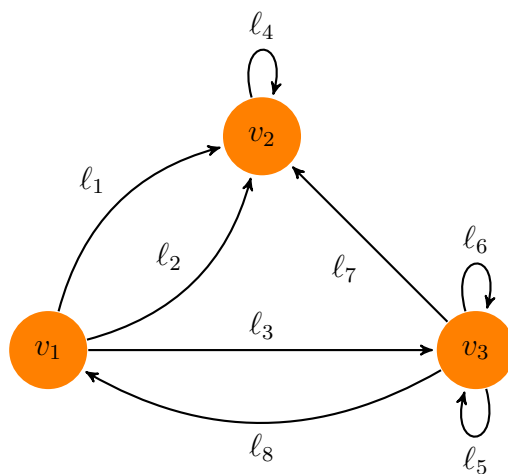
Ejemplo 18. Dados los grafos dirigidos



entonces G_1 y G_2 son isomorfos, pero G_3 no es isomorfo a ninguno de los anteriores. □

El concepto de camino se generaliza intuitivamente al caso de grafos dirigidos imponiendo además que en todo momento se siga el sentido asignado a las flechas. La distancia de un vértice v_i a otro v_j es el menor número de flechas que hay que recorrer para ir de v_i a v_j .

Ejemplo 19. Consideramos de nuevo el grafo siguiente:



Entonces:

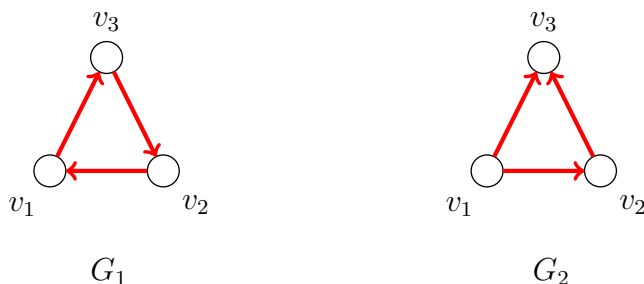
- La secuencia de lados ℓ_1, ℓ_4, ℓ_4 es un camino de longitud 3 del vértice v_1 al vértice v_2 .
- $d(v_1, v_2) = 1$.
- El vértice v_1 está conectado con el vértice v_2 , pero v_2 no está conectado con v_1 , ya que no hay ningún camino de v_2 a v_1 .

□

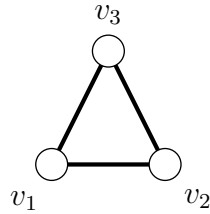
El grafo asociado de un grafo dirigido G , es el grafo no dirigido que resulta de suprimir los sentidos de las flechas que hay en G .

Un grafo dirigido G se dice que es **conexo**, cuando su grafo asociado es conexo. Un grafo dirigido G se dice **fuertemente conexo**, si para cualesquiera $v_i, v_j \in V(G)$, existe algún camino dirigido de v_i a v_j .

Ejemplo 20. Sean los grafos dirigidos



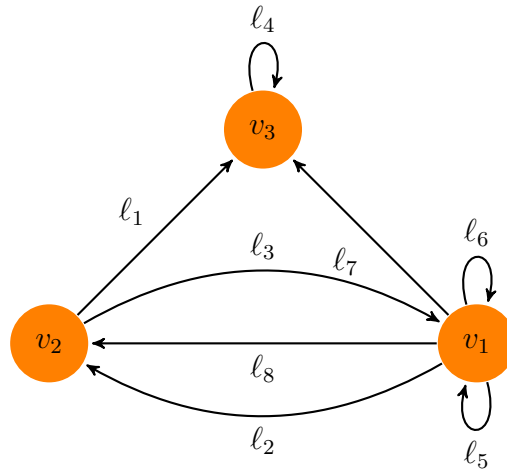
G_1 y G_2 tienen el mismo grafo asociado que es



El grafo G_1 es fuertemente conexo. El grafo G_2 no es fuertemente conexo ya que no existe ningún camino del vértice v_3 al v_1 . \square

Para grafos dirigidos también se define la matriz de adyacencia, que en general no será simétrica. Por supuesto la matriz depende del orden que establezcamos para los vértices.

Ejemplo 21. Para el grafo siguiente,



si ordenamos los vértices en la forma usual, v_1, v_2, v_3 , la matriz de adyacencia es

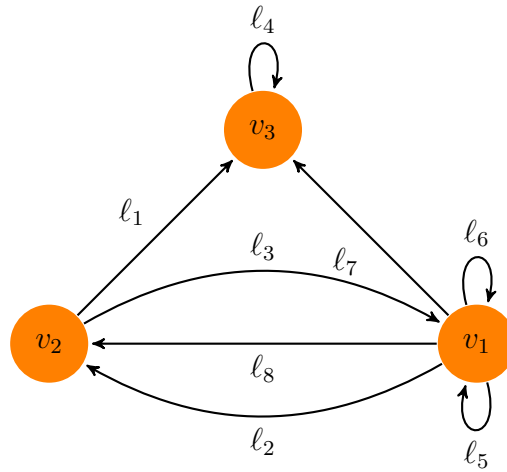
$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

\square

Además tenemos el correspondiente Teorema del número de caminos.

Proposición 6 (Teorema del número de caminos para grafos dirigidos). Sea $G = (V, L, \text{ini}, \text{fin})$ un grafo dirigido con $V = \{v_1, \dots, v_n\}$. Entonces para todo $k \in \mathbb{N}$, el número de caminos dirigidos en G de longitud k desde v_i hasta v_j viene dado por el elemento que aparece en la fila i -ésima y en la columna j -ésima de la matriz $A(G)^k$. \square

Ejemplo 22. Consideramos de nuevo el grafo del Ejemplo 21,



y su matriz de adyacencia

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Las primeras potencias de A son:

$$A^2 = \begin{pmatrix} 6 & 4 & 5 \\ 2 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 16 & 12 & 15 \\ 6 & 4 & 6 \\ 0 & 0 & 1 \end{pmatrix}, \quad A^4 = \begin{pmatrix} 44 & 32 & 43 \\ 16 & 12 & 16 \\ 0 & 0 & 1 \end{pmatrix}$$

Por ejemplo vemos que hay:

- 6 caminos de longitud 2 del vértice v_1 al vértice v_1 .
- 6 caminos de longitud 3 del vértice v_2 al vértice v_1 .
- 43 caminos de longitud 4 del vértice v_1 al vértice v_3 .

□

6 Tipos de caminos en un grafo

Referido a un grafo G de cualquier tipo, un **camino** se dice que es **abierto** cuando el nodo de inicio es diferente del nodo de llegada, y se dice que es **cerrado (ó un lazo)** en caso contrario.

Nos restringimos ahora al caso de multigrafos G .

Para un multigrafo G , un **camino o recorrido de Euler** es aquel que pasa por cada lado de G una sólo vez. Si es cerrado, se denomina un **lazo o camino cerrado de Euler**.

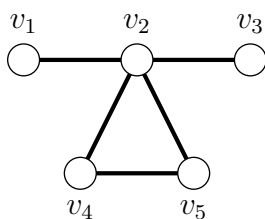
Para que en G pueda existir algún camino de Euler, es necesario que G sea conexo.

Teorema 2. Sea G un multigrafo conexo. Entonces:

1. En G hay un lazo de Euler si y sólo si el grado de cada vértice de G es par.
2. En G hay un camino de Euler entre los vértices u y v , con $u \neq v$, si y sólo si u y v son los únicos vértices de G de grado impar.

□

Ejemplo 23. El grafo de la figura siguiente es conexo, y además todos sus vértices tienen grado par, excepto, v_1 y v_3 que tienen grado impar.



Por tanto hay un camino de Euler del vértice v_1 al vértice v_3 :

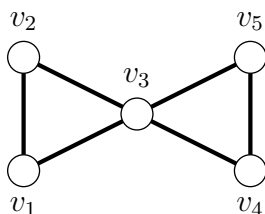
$$v_1, v_2, v_4, v_5, v_2, v_3.$$

Otra posibilidad es

$$v_1, v_2, v_5, v_4, v_2, v_3.$$

□

Ejemplo 24. El grafo de la figura siguiente es conexo y todos sus vértices tienen grado par, por lo cual hay un lazo de Euler ó camino cerrado de Euler.



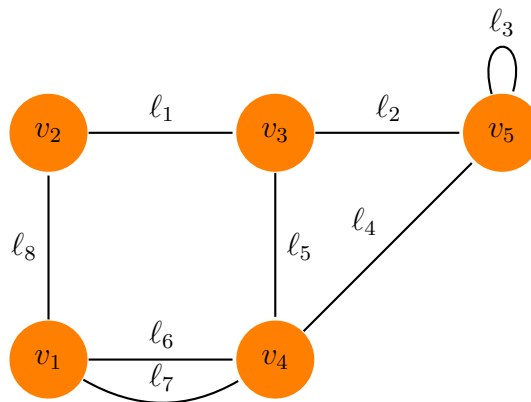
Por ejemplo, un lazo de Euler es:

$$v_3, v_4, v_5, v_3, v_2, v_1, v_3.$$

□

Recordemos que en multigrafo cada autolazo aporta dos unidades al grado del vértice correspondiente.

Ejemplo 25. El multigrafo de la figura es conexo, y v_1 y v_3 son los únicos vértices de grado impar.



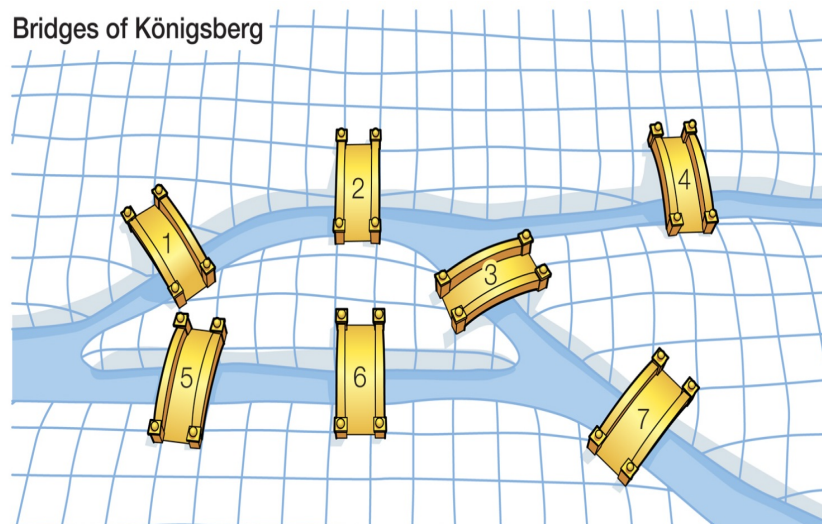
□

Por tanto existe un camino de Euler del vértice v_1 al v_3 :

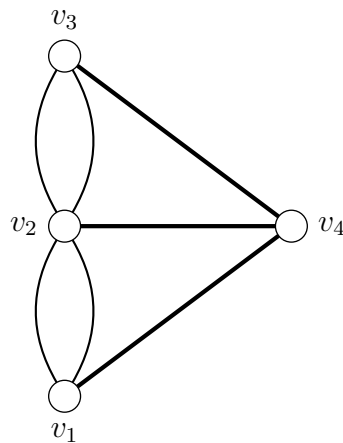
$$\ell_7, \ell_6, \ell_8, \ell_1, \ell_2, \ell_3, \ell_4, \ell_5.$$

□

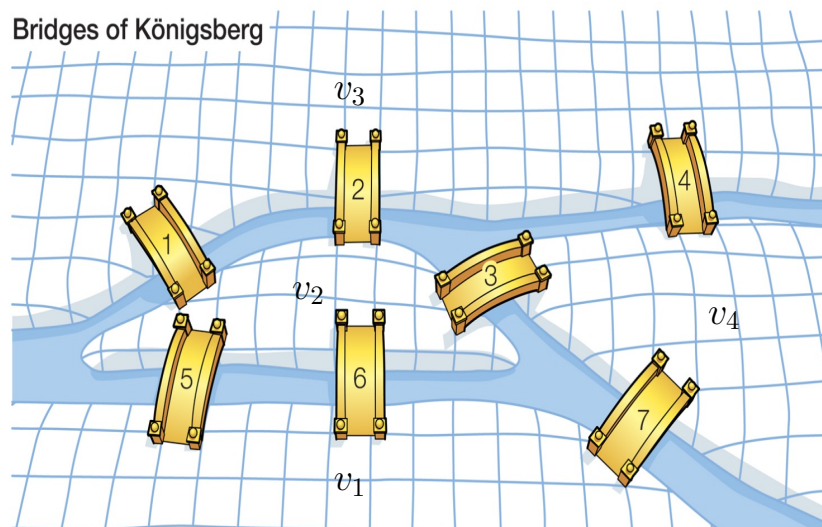
Como ya hemos comentado al principio del tema, la teoría de grafos se originó en el siglo XVIII con el llamado **problema de los puentes de Königsberg**, el cual fué resuelto por Leonhard Euler. Recordemos de nuevo el dibujo.



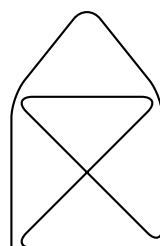
Euler asoció al dibujo el siguiente grafo.



Especificamos los nodos sobre el dibujo.



Ejemplo 26. Un problema clásico relacionado con los recorridos de Euler en un grafo es el problema de dibujar un sobre abierto de un sólo trazo. Vea el dibujo siguiente.



Como se aprecia, el problema tiene solución ya que la figura del sobre se puede interpretar como un grafo conexo en el que todos los vértices tienen grado par, excepto los dos inferiores que tienen grado impar. \square

Para un grafo dirigido G , un **camino o recorrido de Euler** es aquel que pasa por

cada flecha de G una sólo vez. Si es cerrado, se denomina un **lazo o camino cerrado de Euler**.

También es posible dar una versión del Teorema 2 para grafos dirigidos usando la noción de grado de salida y grado de entrada de un vértice. Concretamente ha de ocurrir que el grafo sea conexo y para cada vértice, dependiendo del tipo de camino, se verifique una condición que relaciona el grado de entrada con el grado de salida.

Proposición 7. Sea G un grafo dirigido conexo. Entonces:

1. En G hay un lazo de Euler si y sólo si para todo $v \in V(G)$,

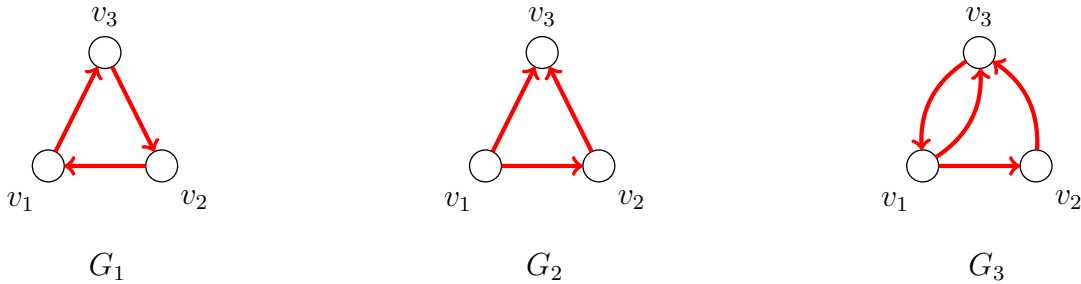
$$\text{gr}_G^+(v) = \text{gr}_G^-(v).$$

2. En G hay un camino de Euler de v_i a v_j , con $i \neq j$, si y sólo si:

- $\text{gr}_G^+(v_i) = 1 + \text{gr}_G^-(v_i)$,
- $\text{gr}_G^-(v_j) = 1 + \text{gr}_G^+(v_j)$,
- Para todo $v \in V(G) \setminus \{v_i, v_j\}$, $\text{gr}_G^-(v) = \text{gr}_G^+(v)$.

□

Ejemplo 27. Sean los grafos dirigidos



Tanto G_1 , G_2 y G_3 son conexos. Para G_1 se verifica que $\text{gr}_G^+(v) = \text{gr}_G^-(v)$ para cualquier vértice v , por lo que el apartado (1.) de la Proposición 7 nos dice que en G_1 hay un lazo dirigido de Euler.

Sin embargo, para G_2 no se verifica ni la condición del apartado (1.) ni (todas) las condiciones del apartado (2.) de la Proposición 7. Por tanto en G_2 no hay lazos dirigidos de Euler.

Analicemos lo que ocurre en G_3 . En primer lugar observamos que

$$\text{gr}_G^+(v_2) = \text{gr}_G^-(v_2).$$

Por otra parte

$$\text{gr}_G^+(v_1) = 1 + \text{gr}_G^-(v_1) \quad \text{y} \quad \text{gr}_G^-(v_3) = 1 + \text{gr}_G^+(v_3).$$

Por el apartado (3.) de la Proposición 7, en G_3 existe un camino (abierto) de Euler del vértice v_1 al v_3 . \square

En lo que queda de sección, consideramos de nuevo grafos no dirigidos. En un grafo G , un **ciclo** es un camino cerrado

$$v_0, v_1, \dots, v_{k-1}, v_0,$$

tal que los vértices v_0, v_1, \dots, v_{k-1} son distintos entre sí.

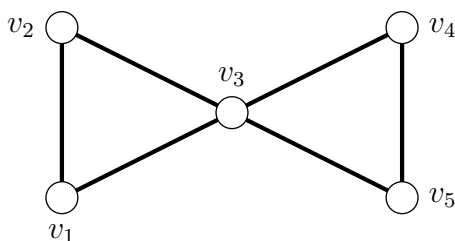
Ejemplo 28. En el grafo de la figura, la secuencia de vértices

$$v_1, v_2, v_3, v_1$$

es un ciclo, pero

$$v_1, v_2, v_3, v_4, v_5, v_3, v_1$$

no lo es.



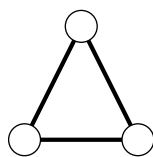
\square

Para todo entero $n \geq 3$, definimos el grafo C_n como sigue:

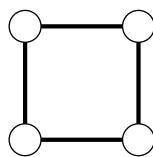
$$V(C_n) = \{1, 2, \dots, n\} \quad \text{y} \quad L(C_n) = \{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}, \{n, 1\}\}.$$

Diremos que C_n , así como cualquier otro grafo isomorfo a él, es un **grafo ciclo de n vértices**.

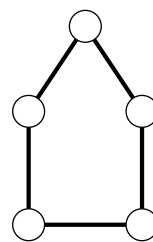
Mostramos los primeros grafos ciclo.



C_3



C_4



C_5

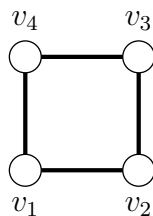
Un **camino abierto** en un grafo G se denomina **de Hamilton**, si pasa por cada vértice de G una sola vez.

Un **ciclo de Hamilton** es un camino cerrado que pasa por cada vértice una sólo vez, salvo el inicial y el final que son iguales.

Englobamos a los caminos abiertos de Hamilton y a los ciclos de Hamilton bajo la designación de **recorridos de Hamilton**.

Evidentemente, en cualquier grafo ciclo C_n , existe un ciclo de Hamilton. Es más, también existe un camino abierto de Hamilton.

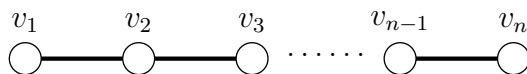
Por ejemplo, para el grafo C_4 ,



la secuencia de vértices v_1, v_2, v_3, v_4, v_1 es un ciclo de Hamilton, y la secuencia de vértices v_1, v_2, v_3, v_4 es un camino abierto de Hamilton.

Para todo entero $n \geq 1$, definimos el grafo P_n como sigue:

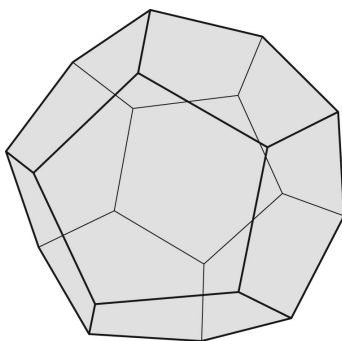
$$V(P_n) = \{v_1, v_2, \dots, v_n\} \quad \text{y} \quad L(P_n) = \{\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{n-1}, v_n\}\}.$$



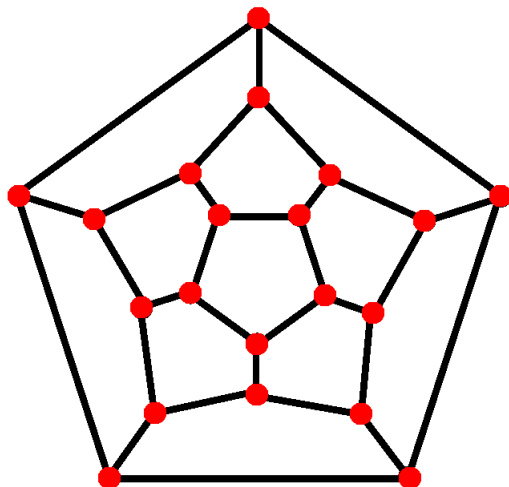
Diremos que P_n , así como cualquier otro grafo isomorfo con él, es un **camino de n vértices**. Para todo $n \geq 2$, es evidente que la secuencia de vértices $v_1, v_2, v_3, \dots, v_{n-1}, v_n$ en P_n es un camino abierto de Hamilton.

El problema de encontrar caminos de Hamilton en un grafo se originó a raíz de un juego propuesto por **William Hamilton** (4 de agosto de 1805 - 2 de septiembre de 1865).

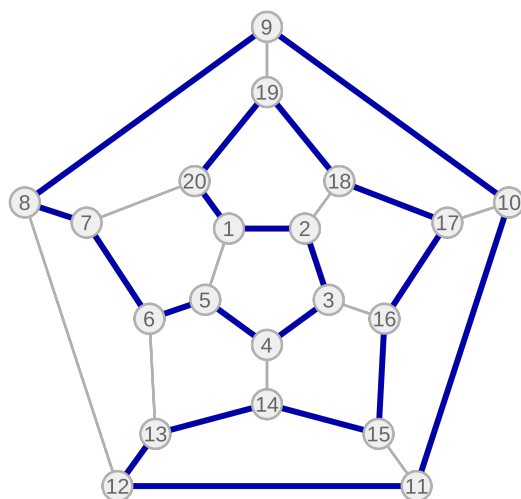
Éste consistía en: *Dado un **dodecaedro**, es decir un poliedro formado por 12 caras que son pentágonos, se trata de partir de uno de sus vértices, viajar a través de las aristas y recorrer cada uno de los vértices restantes una sólo vez, terminando en el vértice inicial.*



Realmente el problema equivale a encontrar un ciclo de Hamilton en el grafo siguiente.



Una solución a este problema nada trivial, es la siguiente:



El problema de decidir si en un grafo hay o no un recorrido de Hamilton no se resuelve tan elegantemente como el problema de los recorridos de Euler que acabamos de ver. Sólo se conocen condiciones suficientes, pero no necesarias, para la existencia de ciclos de Hamilton. Mencionamos a título ilustrativo una de ellas.

Proposición 8. Sea G un grafo tal que $|V(G)| = n \geq 3$ y para todo $v \in V(G)$ se verifica $\text{gr}_G(v) \geq n/2$. Entonces G tiene un ciclo de Hamilton.

Sin embargo existen grafos G en los que hay ciclos de Hamilton pero no se verifica la condición de la proposición anterior. Por ejemplo, C_n cuando $n \geq 5$.

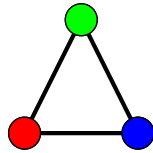
7 Coloración de grafos

Una **coloración** (de los vértices) de un grafo $G = (V, L)$ es una aplicación $f : V \rightarrow E$, donde E es un conjunto finito llamado conjunto de colores, tal que si $\{u, v\} \in L$, entonces $f(u) \neq f(v)$.

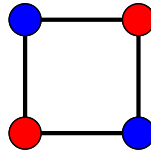
Es decir, a cada vértice de G se le asigna un color (elemento del conjunto E) de modo que dos vértices adyacentes reciban colores diferentes.

El **número cromático** de un grafo G es el menor cardinal que puede tener un conjunto de colores E de forma que se pueda definir una coloración en G . Dicha cantidad se denota por $\chi(G)$.

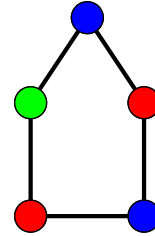
Ejemplo 29. Supongamos el conjunto de colores $E = \{\text{rojo}, \text{azul}, \text{verde}\}$. Mostramos sendas coloraciones para los tres primeros grafos ciclo.



$$\chi(C_3) = 3$$



$$\chi(C_4) = 2$$



$$\chi(C_5) = 3$$

Generalizamos ésto en el ejemplo siguiente.

Ejemplo 30.

- Para $n \geq 1$, $\chi(K_n) = n$.
- Para $n \geq 3$, $\chi(C_n) = \begin{cases} 2 & \text{si } n \text{ es par} \\ 3 & \text{si } n \text{ es impar} \end{cases}$
- Para $n \geq 1$, $\chi(P_n) = \begin{cases} 1 & \text{si } n = 1 \\ 2 & \text{si } n \geq 2 \end{cases}$

□

La siguiente propiedad nos proporciona una cota superior para el número cromático de un grafo.

Proposición 1. Sea G un grafo y sea $\delta(G)$ el máximo de los grados de los vértices de G . Entonces

$$\chi(G) \leq \delta(G) + 1.$$

□

La siguiente propiedad es consecuencia inmediata de las definiciones.

Proposición 2.

1. Si $G_1 \cong G_2$, entonces $\chi(G_1) = \chi(G_2)$.
2. Si G_1 es un subgrafo de G_2 , entonces $\chi(G_1) \leq \chi(G_2)$.
3. Si G_1, \dots, G_r son las componentes conexas de un grafo G , entonces

$$\chi(G) = \max\{\chi(G_1), \dots, \chi(G_r)\}.$$

□

Ejemplo 31. Si G_1 y G_2 son las componentes conexas de un grafo G , y se verifica que $G_1 \cong P_4$ y $G_2 \cong C_3$, entonces

$$\chi(G) = \max\{2, 3\} = 3.$$

□

Un concepto relacionado con el número cromático de un grafo es el de polinomio cromático de un grafo.

Si x representa una variable en \mathbb{N} , el **polinomio cromático de un grafo** G , denotado por $p(G, x)$, es el número de formas de colorear los vértices de G usando a lo sumo x colores.

Ejemplo 32.

- $p(\overline{K_n}, x) = x^n$
- $p(K_n, x) = x \cdot (x - 1) \cdots (x - n + 1)$
- $p(P_n, x) = x \cdot (x - 1)^{n-1}$
- $p(C_3, x) = x \cdot (x - 1) \cdot (x - 2)$

□

Proposición 3. Si G_1, \dots, G_r son las componentes conexas de un grafo G , entonces

$$p(G, x) = p(G_1, x) \cdots p(G_r, x).$$

□

Ejemplo 33. Si G_1 y G_2 son las componentes conexas de un grafo G , con $G_1 \cong P_4$ y $G_2 \cong C_3$, entonces

$$p(G, x) = p(G_1, x) \cdot p(G_2, x) = x \cdot (x-1)^3 \cdot x \cdot (x-1) \cdot (x-2).$$

□

Surge de forma natural la siguiente cuestión.

Para un grafo G , ¿qué relación hay entre $\chi(G)$ y $p(G, x)$?

La respuesta es bien sencilla:

$\chi(G)$ es el menor número natural d tal que $p(G, d) > 0$.

Ejemplo 34. Sabemos que $p(K_4, x) = x \cdot (x-1) \cdot (x-2) \cdot (x-3)$. Resulta entonces que

$$p(K_4, 1) = p(K_4, 2) = p(K_4, 3) = 0, \quad \text{y} \quad p(K_4, 4) = 4! = 24 > 0.$$

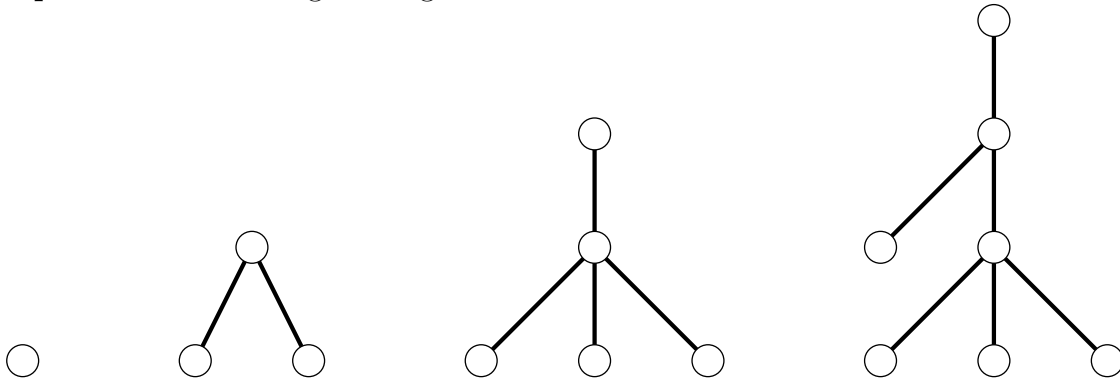
Por consiguiente $\chi(K_4) = 4$.

□

8 Árboles y grafos bipartidos

Un **árbol** es un grafo conexo que no contiene ciclos.

Ejemplo 35. Los cuatro grafos siguientes son árboles.

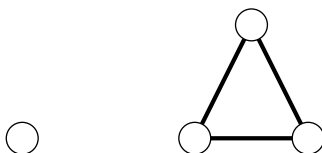


□

Proposición 9. Si G es un árbol, entonces $|L(G)| = |V(G)| - 1$.

□

Comentario 1. Si G es un grafo y se verifica $|L(G)| = |V(G)| - 1$, ésto no implica que G sea un árbol. Por ejemplo, el grafo siguiente no es conexo y verifica que $|L(G)| = 3 = |V(G)| - 1$:



□

La siguiente propiedad caracteriza el concepto de árbol.

Proposición 10. Sea G un grafo. Son equivalentes:

1. G es un árbol.
2. Dos nodos cualesquiera de G están conectados por un único camino en el que no se repiten vértices.
3. G es conexo y se verifica $|L(G)| = |V(G)| - 1$.

□

El concepto de árbol se generaliza mediante el concepto de bosque.

Un **bosque** es un grafo que no contiene ciclos, es decir, un grafo cuyas componentes conexas son árboles.

La Proposición 9 se generaliza de la manera siguiente.

Proposición 11. Si G es un bosque, entonces $|L(G)| = |V(G)| - c$, siendo c el número de componentes conexas de G . □

A su vez el concepto de bosque se generaliza mediante el de grafo bipartido.

Un **grafo** $G = (V, L)$ se dice que es **bipartido**, si es posible descomponer el conjunto $V(G)$ como

$$V(G) = V_1 \cup V_2$$

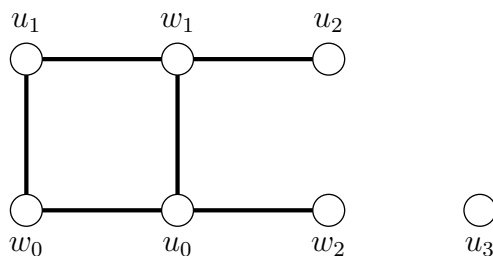
tal que

$$V_1 \cap V_2 = \emptyset,$$

y si $\{u, v\} \in L(G)$, entonces $u \in V_1$ y $v \in V_2$, o bien $u \in V_2$ y $v \in V_1$.

Es decir, todo lado de G (en caso de que lo haya) es incidente con un vértice de V_1 y con otro de V_2 .

Ejemplo 36. Consideremos el grafo G siguiente:



Entonces el conjunto de vértices

$$V = \{u_0, u_1, u_2, u_3, w_0, w_1, w_2\}$$

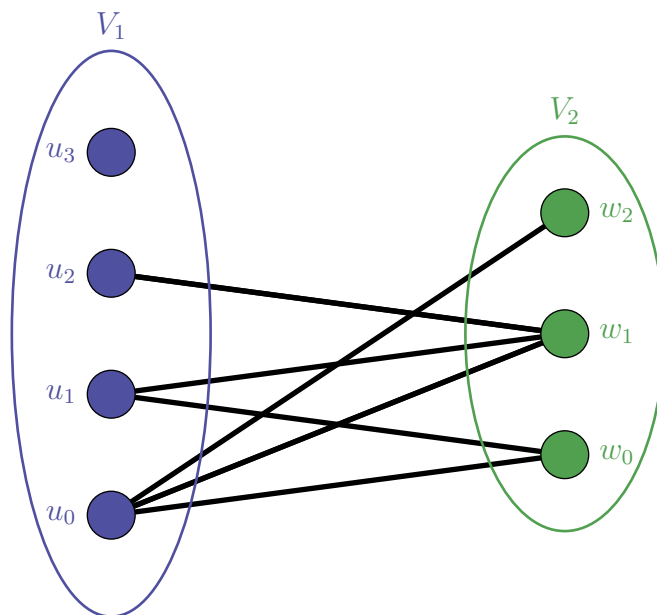
lo podemos particionar como

$$V = V_1 \cup V_2,$$

donde

$$V_1 = \{u_0, u_1, u_2, u_3\} \quad \text{y} \quad V_2 = \{w_0, w_1, w_2\}.$$

Además ningún vértice de V_1 es adyacente con ningún vértice de V_2 , por lo que G es un grafo bipartido. Se suele hacer la representación siguiente de G .

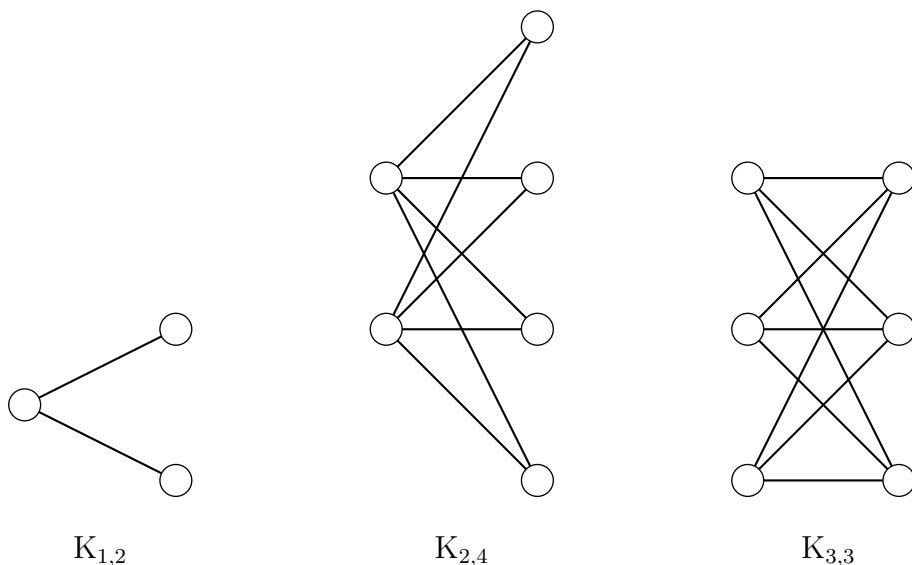


Nótese que el vértice u_3 , al ser un vértice aislado, lo podíamos haber puesto en el conjunto V_2 en vez de en V_1 . \square

Comentario 2. Un grafo que consta de un único vértice v es un grafo bipartido, donde podemos tomar $V_1 = \{v\}$ y $V_2 = \{\}$, o bien, $V_1 = \{\}$ y $V_2 = \{v\}$. \square

Si en un grafo bipartido G además para todo $u \in V_1$ y para todo $v \in V_2$ se verifica que $\{u, v\} \in L(G)$, se dice que G es **bipartido-completo**. Cualquier grafo bipartido-completo tal que $|V_1| = m$ y $|V_2| = n$ se suele denotar por $K_{m,n}$.

Ejemplo 37. Los siguientes grafos son bipartido-completos.



□

Los grafos bipartidos se usan para representar relaciones entre dos conjuntos disjuntos. Por ejemplo, en una oficina de empleo hay por una parte personas demandantes de empleo, y por otra hay empleos los cuales pueden requerir determinadas cualificaciones para ser desempeñados.

La siguiente propiedad caracteriza a los grafos bipartidos.

Proposición 12. Para un grafo G , las siguientes afirmaciones son equivalentes:

1. G es bipartido.
2. En G no hay ciclos de longitud impar.
3. $\chi(G) \leq 2$.

□

Vea de nuevo el grafo del Ejemplo 36, y entenderá mejor por qué es bipartido.

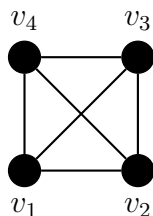
Otro ejemplo viene dado por el grafo ciclo C_n , con $n \geq 3$. Claramente C_n es bipartido, si y sólo si n es par.

En un bosque no hay ciclos, ni de longitud par ni impar, por lo cual todo bosque es un grafo bipartido.

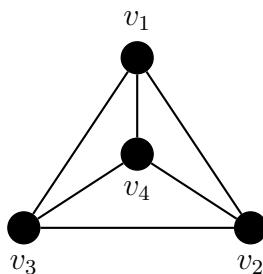
9 Grafos planos

Recordemos que un grafo es un objeto formal (definido por ciertos conjuntos y aplicaciones de incidencia) y que por abuso del lenguaje también llamamos grafo a cualquiera de sus representaciones gráficas.

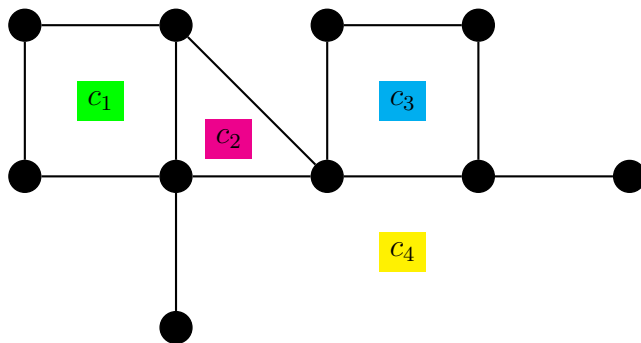
Un **grafo** G es **plano** si puede ser representado en el plano sin que se crucen dos lados. En tal caso decimos que tenemos una **representación plana** de G . Por ejemplo el grafo K_4 es plano. Aquí lo tenemos representado con cruces de sus lados.



Pero también es posible representarlo sin cruces de sus lados, por lo cual es un grafo plano.



Toda representación plana de un grafo plano G divide el plano en un número finito de regiones o **caras**. La única cara de área infinita que aparece se denomina **cara exterior**. En el ejemplo siguiente tenemos una representación plana de un grafo plano en la que aparecen 4 caras.



Comentario 3. Nótese que cualquier bosque es un grafo plano y en cualquier representación plana suya aparece sólo una cara (la cara exterior). Recíprocamente, si tenemos un grafo plano G tal que en cualquier representación plana suya aparece sólo una cara, entonces G es un bosque. \square

Teorema 3. Sea G un grafo plano conexo. Supongamos que en una representación plana

de G aparecen c caras. Entonces se verifica

$$|V(G)| - |L(G)| + c = 2.$$

□

Consideremos de nuevo el grafo plano anterior. Dicho grafo es conexo, $|V(G)| = 10$, $|L(G)| = 12$ y $c = 4$. Entonces se verifica:

$$10 - 12 + 4 = 2$$

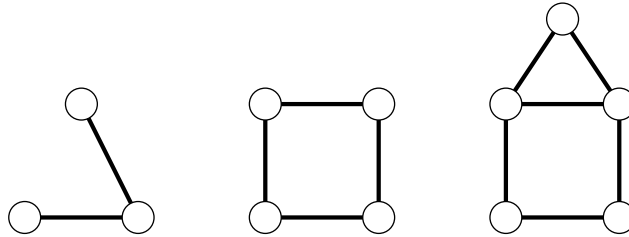
El teorema anterior se puede generalizar para grafos planos no conexos como sigue.

Teorema 4. Sea G un grafo plano formado por τ componentes conexas. Supongamos que en una representación plana de G aparecen c caras. Entonces

$$|V(G)| - |L(G)| + c = 1 + \tau.$$

□

Ejemplo 38. Sea el grafo siguiente que consta de tres componentes conexas:



Entonces $|V(G)| = 12$, $|L(G)| = 12$ y $c = 4$, y se verifica:

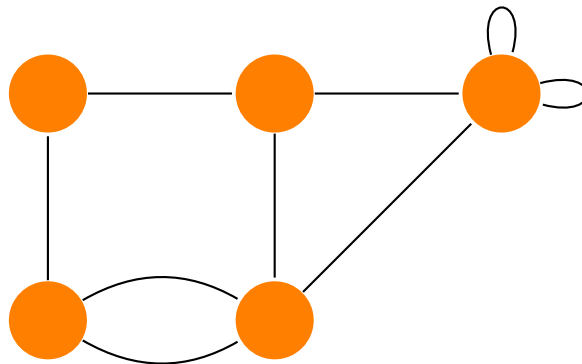
$$12 - 12 + 4 = 4 = 1 + \tau$$

de donde $\tau = 3$ que coincide con el número de componentes conexas.

□

Además la fórmula del teorema anterior sigue siendo válida para el caso de multigrafos.

Ejemplo 39. Comprobamos la identidad del Teorema 4 para el multigrafo siguiente.



Aquí $|V(G)| = 5$, $|L(G)| = 9$, $c = 6$ y $\tau = 1$. Entonces

$$5 - 9 + 6 = 2 = 1 + \tau.$$

□

El Teorema 4 tiene también como consecuencia que el número de caras que resultan en cualquier representación plana de un grafo plano G , es siempre la misma, y por tanto no depende de la habilidad del dibujante.

Ahora nos planteamos la siguiente pregunta. *¿Cuál es el mayor número de lados que puede tener un grafo plano con n vértices?*

La siguiente proposición nos permite acotar el número de lados en un grafo plano, en función de su número de vértices, y por tanto responde a la pregunta anterior.

Proposición 4. Si G es un grafo plano con al menos tres vértices, entonces

$$|L(G)| \leq 3 \cdot [|V(G)| - 2].$$

□

Un **grafo plano** se dice **maximal** si para cualesquiera $u, v \in V(G)$ tales que $u \neq v$, si $\ell = \{u, v\} \notin L(G)$, entonces el grafo $G' = (V(G), L(G) \cup \{\ell\})$ no es plano. Es decir, si hay dos vértices no adyacentes y los unimos por un lado, el grafo que resulta no es plano.

Nótese que en una representación plana de un grafo plano maximal, con al menos tres vértices, cada una de las caras ha de ser un triángulo.

Incluso podemos afinar un poco más el resultado en la proposición anterior.

Proposición 5. Si G es un grafo plano con al menos tres vértices y en el que no hay ciclos de longitud 3, entonces

$$|L(G)| \leq 2 \cdot [|V(G)| - 2].$$

□

Las Proposiciones 4 y 5 son principalmente útiles para averiguar si un grafo dado no es plano.

Ejemplo 40. ¿Qué grafos completos son planos?

Ya hemos visto a lo largo de diversos ejemplos que los grafos K_1 , K_2 , K_3 y K_4 son planos.

Si intentamos obtener una representación plana de K_5 vemos que no se puede. Sin embargo ésto no es suficiente para asegurar que K_5 no es plano, salvo que hagamos un

análisis exhaustivo de todas las posibilidades que resultan. Es más fácil utilizar el siguiente razonamiento de tipo aritmético. Sabemos que

$$|V(K_5)| = 5 \quad \text{y} \quad |L(K_5)| = \binom{5}{2} = 10.$$

Por la Proposición 4, si K_5 fuese plano, entonces tendría que ocurrir que

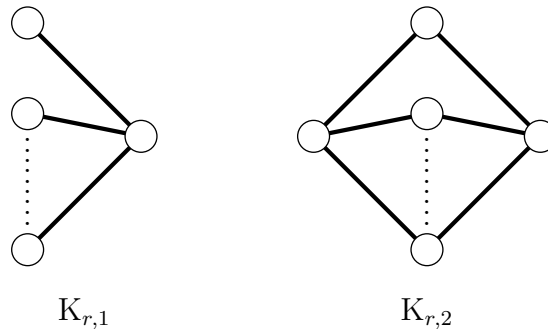
$$10 \leq 3 \cdot (5 - 2) = 9,$$

lo cual es falso. Por tanto K_5 no es plano. Como consecuencia, K_n tampoco es plano para $n \geq 6$. Así hemos probado que K_n es plano si y sólo si $n \in \{1, 2, 3, 4\}$.

Además, si en K_5 suprimimos un lado, pero no los dos vértices incidentes con él, resulta un grafo plano maximal, pues si le añadimos de nuevo el lado que hemos suprimido, obtenemos K_5 que no es plano. \square

Ejemplo 41. ¿Qué grafos bipartido-completos son planos?

Recurriendo a los dibujos es inmediato comprobar que $K_{r,1}$ y $K_{r,2}$ son planos para todo $r \geq 1$.



¿Es $K_{3,3}$ plano? Si intentamos obtener una representación plana del mismo, vemos que no se puede. Sin embargo, al igual que antes, ésto no es suficiente para asegurar que $K_{3,3}$ no es plano, salgo que analicemos todas las posibilidades de manera completa. En lugar de ello, de nuevo trabajamos con la aritmética. Tenemos que

$$|V(K_{3,3})| = 6 \quad \text{y} \quad |L(K_{3,3})| = 9.$$

Por la Proposición 4, si $K_{3,3}$ fuese plano, entonces tendría que ocurrir que

$$9 \leq 3 \cdot (6 - 2) = 12,$$

lo cual no supone ninguna contradicción. Es decir, la Proposición 4 no aporta ninguna información sobre la planaridad o no planaridad de $K_{3,3}$.

Pero sabemos que $K_{3,3}$ es un grafo bipartido y por tanto no contiene ciclos de longitud impar, y en particular tampoco contiene ciclos de longitud 3. Por tanto aplicamos la Proposición 5, según la cual si $K_{3,3}$ fuese plano, entonces

$$9 \leq 2 \cdot (6 - 2) = 8,$$

lo cual no es cierto. Ésto significa que $K_{3,3}$ no es plano.

Como consecuencia, ni $K_{r,3}$ ni $K_{3,s}$ son planos para $r, s \geq 3$.

Así hemos probado que $K_{r,s}$ es plano si y sólo si $\min\{r, s\} \leq 2$. \square

Los grafos no planos K_5 y $K_{3,3}$ juegan un papel muy importante en la caracterización de cuándo un grafo dado es plano.

En 1930 el matemático polaco Kazimierz Kuratowski publicó un trabajo en el que establecía que cualquier grafo no plano contiene un subgrafo que está relacionado con K_5 ó con $K_{3,3}$. Damos a continuación las definiciones precisas para poder enunciar el Teorema de Kuratowski.

Una **subdivisión de un lado** $\ell = \{u, v\}$ de un grafo G consiste en eliminar el lado ℓ de G , y añadir un nuevo vértice w así como los lados $\{u, w\}$ y $\{w, v\}$. Intuitivamente lo que estamos haciendo es insertar un nuevo vértice w “en medio” del lado ℓ .

Es decir, definimos un nuevo grafo \tilde{G} tal que

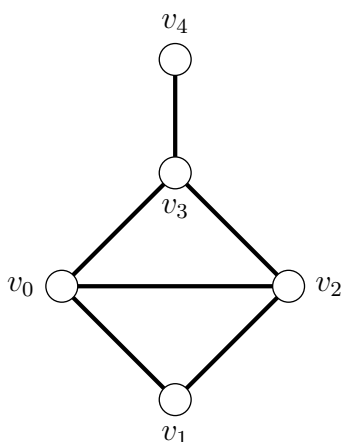
$$V(\tilde{G}) = V(G) \cup \{w\} \quad \text{y} \quad L(\tilde{G}) = (L(G) \setminus \{\ell\}) \cup \{\{u, w\}, \{w, v\}\}.$$

Dos grafos G_1 y G_2 son **homeomorfos**, si existe un grafo H tal que:

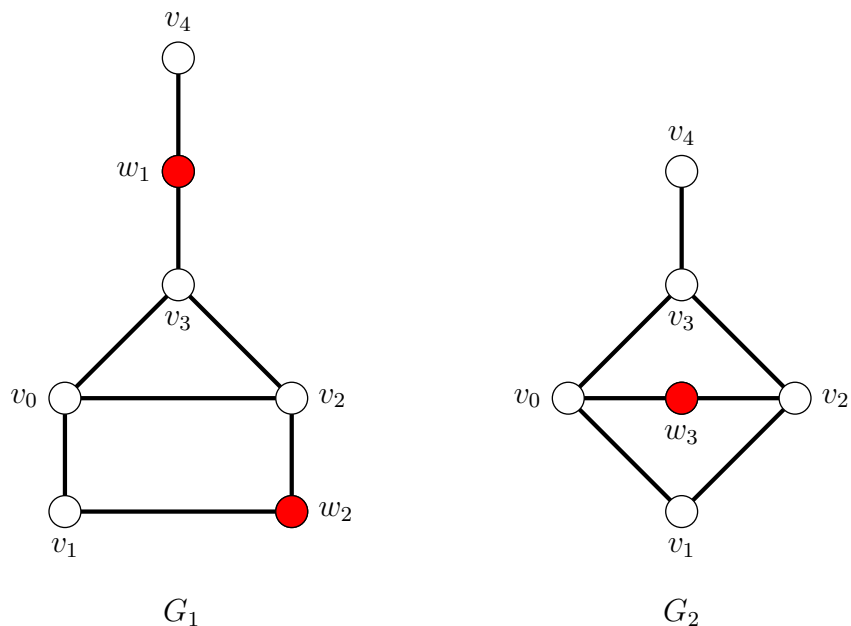
1. G_1 es isomorfo a H ó es obtenible a partir de H aplicando un número finito de subdivisiones de lados.
2. G_2 es isomorfo a H ó es obtenible a partir de H aplicando un número finito de subdivisiones de lados.

Claramente, “ser homeomorfo a” define una relación de equivalencia sobre grafos.

Ejemplo 42. Sea el grafo siguiente H :



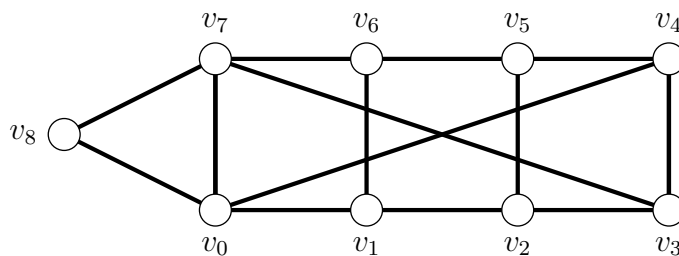
Cada uno de los grafos siguientes G_1 y G_2 se obtienen aplicando ciertas subdivisiones de lados sobre el grafo H . Por tanto, G_1 y G_2 son homeomorfos.



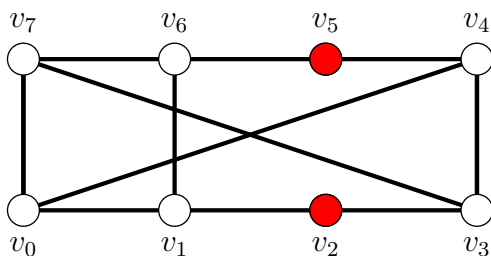
□

Teorema 5 (Kuratowski (1930)). Un grafo G es plano si y sólo si ninguno de sus subgrafos es homeomorfo a K_5 ni a $K_{3,3}$. □

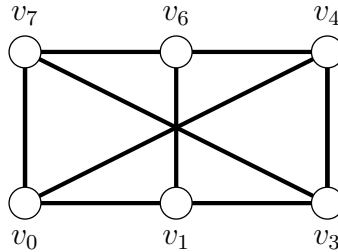
Ejemplo 43. Consideramos el grafo G ,



Queremos saber si G es o no plano. El grafo siguiente, que denotamos por H , es un subgrafo de G .



H ha sido obtenido mediante dos subdivisiones de lados aplicadas al grafo que presentamos a continuación:



Como este grafo es isomorfo al grafo $K_{3,3}$ (compruébelo), concluimos que H no es plano, y por tanto G tampoco lo es. \square

Existe otra forma equivalente de enfocar el problema de la planaridad de un grafo.

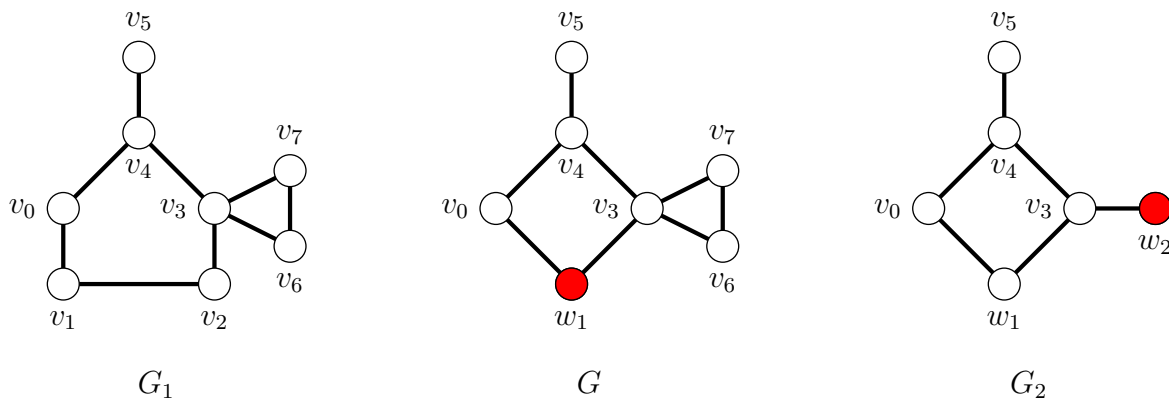
Una **contracción elemental sobre un lado** $\{u, v\}$ de un grafo $G = (V, L)$ consiste en aplicar los pasos siguientes:

- Eliminar el lado $\{u, v\}$ del conjunto L .
- Reemplazar cada aparición de los vértices u y v en lados de L por un nuevo vértice w .
- Quitar los vértices u y v del conjunto V .
- Añadir w a V .

Desde un punto de vista intuitivo, estamos fusionando los vértices u y v de G , eliminando el lado $\{u, v\}$, y el vértice resultante lo llamamos w .

Un grafo G_1 es **contraíble** al grafo G_2 , si G_2 puede ser obtenido a partir de G_1 mediante una secuencia de contracciones elementales.

Ejemplo 44. El grafo G_1 se contrae al grafo G_2 aplicando dos contracciones elementales.



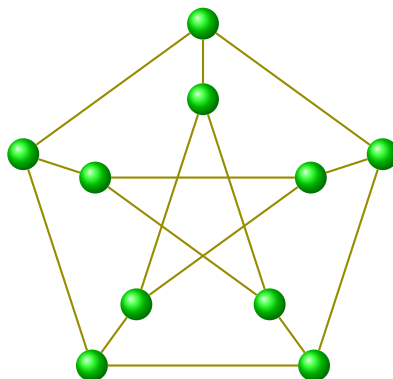
En primer lugar hemos aplicado una contracción elemental sobre el lado $\{v_1, v_2\}$ de G_1 para obtener el grafo G . A continuación hemos aplicado una contracción elemental sobre el lado $\{v_6, v_7\}$ de G para obtener el grafo G_2 . \square

Teorema 6 (Wagner (1937), Harary y Tutte (1965)). Un grafo G es plano si y sólo si ninguno de sus subgrafos es contraíble a K_5 ni a $K_{3,3}$. \square

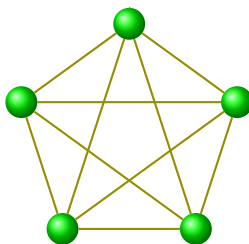
Comentario 4. El Teorema de Wagner-Harary-Tutte nos dice que un grafo G no es plano, si y sólo si existe algún subgrafo de G que es contraíble a K_5 ó a $K_{3,3}$. \square

Hay ejemplos en los que es más fácil concluir que el grafo dado no es plano utilizando el Teorema de Kuratowski, y otros en los que es preferible aplicar el Teorema de Wagner-Harary-Tutte.

Ejemplo 45. El grafo siguiente, G , se conoce como *grafo de Petersen*.



Aplicamos el Teorema de Wagner-Harary-Tutte tomando como subgrafo el propio grafo G . Vemos que G es contraíble mediante cinco contracciones elementales de lados, a un grafo isomorfo a K_5 .



Por tanto, G no es plano. \square

9.1 Una pequeña historia con historia

Otro problema de coloración de grafos es el **problema de colorear las caras de un grafo plano**, de modo que si dos caras son adyacentes (es decir, tienen algún lado común), entonces ambas reciban colores diferentes.

Este problema a su vez es equivalente a colorear los vértices de un nuevo (multi)grafo plano definido a partir de G , llamado el **grafo dual** de G .

En 1852 Francis Guthrie plantea a su hermano Frederick, y éste a Augustus de Morgan, el problema de colorear los condados de Inglaterra, de modo que dos condados con frontera común reciban colores distintos. Desde entonces varios eminentes matemáticos intentan la demostración de la conjetura sin éxito.

Surgió así la **conjetura de los cuatro colores**, según la cual todo mapa se puede colorear usando a lo sumo cuatro colores. Esta afirmación equivale a decir que si G es un grafo plano, entonces $\chi(G) \leq 4$. De hecho el siguiente teorema era conocido.

Teorema 7. Si G es un grafo plano, entonces $\chi(G) \leq 5$. □

Según este resultado es siempre posible colorear todos los países de cualquier mapa usando a lo sumo cinco colores de forma que dos países adyacentes se pinten con colores distintos. Además, la experiencia mostraba que siempre bastaba con cuatro colores.

En 1976, Ken Appel y Wolfgang Haken probaron el resultado, con ayuda de un ordenador CRAY, para cada una de las 1.482 configuraciones posibles, empleando para ello 50 días de cálculo (menos mal que no se fué la luz). En 1996, N. Robertson, D.P. Sanders, P. Seymour y R. Thomas mejoraron la demostración también con ayuda de un ordenador (ahora sólo 633 configuraciones hubo que comprobar).

Este resultado actualmente se conoce como el **Teorema de los cuatro colores** y es el primer teorema que fué demostrado (o constatado, mejor dicho) con ayuda de una máquina. Ello creó discusiones de tipo filosófico sobre lo que ha de considerarse un teorema, y si es correcto aceptar como tal lo que ha sido comprobado por una máquina.

Para ampliar información, puede buscar en la sección de imágenes de GOOGLE la cadena de caracteres “*The four color theorem*”.

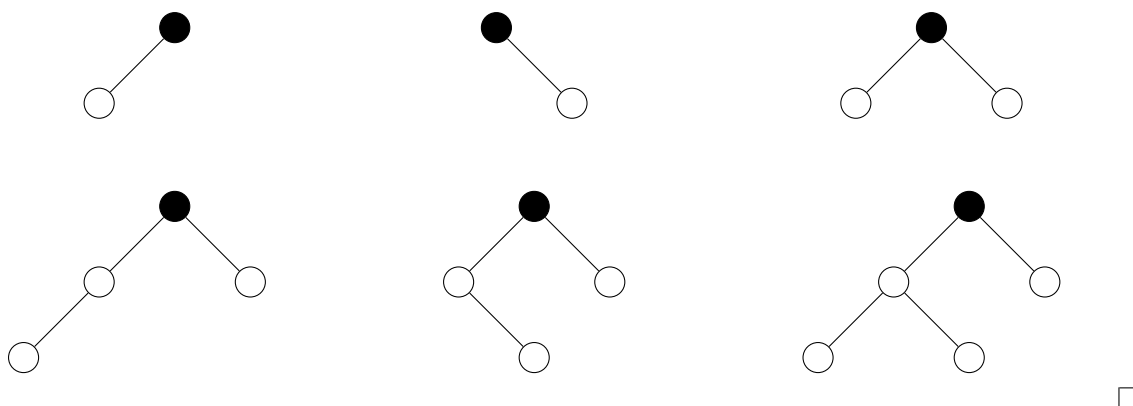
La sección siguiente se da sólo a título informativo y no será evaluada.

10 El algoritmo de Huffman

Los árboles se usan para representar información jerarquizada por niveles. Un ejemplo de ésto lo constituyen los árboles genealógicos. Hay toda una teoría al respecto. Un **árbol con raíz** es un árbol en el que se distingue ó elige un vértice que se llama raíz. A cada nodo v se le asocia un número natural llamado **nivel o profundidad**, que es la distancia de v a la raíz. La profundidad de un árbol es el máximo de las profundidades de sus nodos. En relación con los árboles con raíz, es usual hablar de padre, hijos, ascendientes y descendientes de un nodo. La raíz no tiene padre y por tanto tampoco ascendientes. Los nodos que no tienen descendientes se denominan **hojas**. Un **árbol m -ario** es un árbol con

raíz en el cual cada nodo tiene a lo sumo m descendientes. Un árbol 2-ario se denomina árbol binario. Un **árbol binario ordenado** es aquel en el cual a cada uno de los dos hijos que puede tener un nodo se le asigna de forma excluyente una posición que puede ser izquierda o derecha. Así un nodo puede no tener ningún hijo, tener el hijo izquierdo pero no tener el derecho, tener el hijo derecho pero no el izquierdo, o bien tener el hijo izquierdo y el hijo derecho.

Ejemplo 46. Mostramos algunos árboles binarios ordenados:



Si tenemos un **alfabeto** Σ , es decir, un conjunto de símbolos con los que escribimos frases, a la hora de transmitir telemáticamente un mensaje, cada símbolo perteneciente a Σ se codifica mediante una cadena binaria. Si especificamos para cada símbolo de Σ una cadena binaria, tenemos un código binario ó esquema de codificación binaria para Σ .

Ejemplo 47. Para el alfabeto $\Sigma = \{A, B, C\}$, sea el esquema de codificación binaria siguiente:

$$\begin{cases} A \longleftrightarrow 0 \\ B \longleftrightarrow 01 \\ C \longleftrightarrow 10 \end{cases}$$

Si el receptor recibe 010, ésto puede ser interpretado de dos formas distintas, lo cual no es deseable:

- Como AC, si vemos la cadena 010 como 010.
- Como BA, si vemos la cadena 010 como 010.

El problema radica en que el código (o cadena binaria) mediante el cual se codifica la letra A es un prefijo del código mediante el cual se codifica la letra B. □

Un esquema de codificación binaria para un alfabeto Σ se dice que es prefijo, si la cadena binaria asignada a cada elemento de Σ no es un prefijo de la cadena binaria asignada a cualquier otro elemento de Σ . También diremos que tenemos un **código binario prefijo** para Σ .

Ésto nos asegura que la aplicación que asigna a cada mensaje (formado a partir de símbolos de Σ) su cadena binaria correspondiente, es inyectiva, y por tanto evitará la ambigüedad a la hora de decodificar los mensajes recibidos.

Ejemplo 48. Para el alfabeto $\Sigma = \{A, B, C\}$, un código binario prefijo es

$$\begin{cases} A \longleftrightarrow 0 \\ B \longleftrightarrow 11 \\ C \longleftrightarrow 10 \end{cases}$$

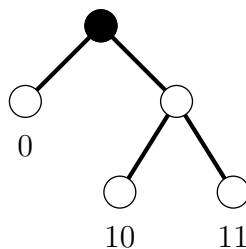
Si el receptor recibe 010, ahora claramente se refiere al mensaje AC. □

¿Qué relación existe entre los árboles binarios ordenados y los códigos binarios prefijos? Dado un alfabeto Σ formado por m símbolos, todo árbol binario ordenado con m hojas representa un código binario prefijo para Σ , y recíprocamente, todo código binario prefijo para Σ se puede representar mediante un árbol binario ordenado con m hojas.

Ejemplo 49. Para el alfabeto $\Sigma = \{A, B, C\}$ y el código binario prefijo siguiente,

$$\begin{cases} A \longleftrightarrow 0 \\ B \longleftrightarrow 11 \\ C \longleftrightarrow 10 \end{cases}$$

tenemos el árbol binario ordenado siguiente:



□

La idea es muy simple. Son las hojas del árbol las que representan a las cadenas binarias, con el siguiente criterio. Por cada nivel que descendemos hacia la izquierda, se escribe 0, y por cada nivel que descendemos hacia la derecha, se escribe 1.

En la práctica, a la hora de codificar un mensaje en binario, hay símbolos del alfabeto que se repiten más que otros. Es lógico tratar de asignarle a los símbolos que más se repiten cadenas binarias lo más cortas posible. Ello reducirá el tamaño del mensaje codificado que hay que transmitir, con el consiguiente ahorro de tiempo ó espacio.

Ejemplo 50. Supongamos un mensaje de 100 caracteres construido sobre el alfabeto $\Sigma = \{A, B, C, D, E, F\}$ con las siguientes frecuencias absolutas para cada carácter:

$$\left\{ \begin{array}{ll} A & : 25 \\ B & : 10 \\ C & : 10 \\ D & : 20 \\ E & : 15 \\ F & : 20 \end{array} \right.$$

Queremos encontrar un código binario prefijo tal que la longitud del mensaje codificado correspondiente sea lo más pequeña posible. Si elegimos como esquema de codificación binaria (sin pensar mucho) el siguiente,

$$\left\{ \begin{array}{ll} A & \longleftrightarrow 00 \\ B & \longleftrightarrow 01 \\ C & \longleftrightarrow 100 \\ D & \longleftrightarrow 1010 \\ E & \longleftrightarrow 1011 \\ F & \longleftrightarrow 11 \end{array} \right.$$

resulta que la longitud del mensaje codificado es

$$25 \cdot 2 + 10 \cdot 2 + 10 \cdot 3 + 20 \cdot 4 + 15 \cdot 4 + 20 \cdot 2 = 280.$$

□

Supongamos un árbol binario ordenado T con hojas v_1, \dots, v_k tal que cada hoja v_i tiene asignado un número real positivo α_i , llamado **peso**. Decimos que T es un **árbol binario ordenado con pesos en sus hojas**. Supongamos además que la profundidad de cada hoja v_i es ℓ_i . Se define el peso de T como

$$\omega(T) = \sum_{i=1}^k \ell_i \cdot \alpha_i.$$

Un árbol binario ordenado T con pesos $\alpha_1, \dots, \alpha_k$ en sus hojas se denomina **óptimo**, si $\omega(T)$ es mínimo entre todos los árboles binarios ordenados con k hojas y pesos $\alpha_1, \dots, \alpha_k$. Denominaremos **código binario prefijo óptimo** al asociado a un árbol binario ordenado óptimo.

Si en el Ejemplo 50 dibujamos el árbol binario ordenado correspondiente a la codificación binaria elegida y le asignamos a cada hoja la frecuencia absoluta del carácter del alfabeto Σ correspondiente a dicha hoja, obtenemos un árbol binario ordenado con pesos en sus hojas. El peso de dicho árbol es precisamente 280, valor numérico que ya ha aparecido más arriba.

De modo general, supongamos un mensaje \mathcal{M} escrito en un alfabeto Σ el cual queremos codificar mediante un código binario prefijo. Para minimizar la longitud del mensaje codificado correspondiente a \mathcal{M} , hemos de encontrar un árbol binario ordenado óptimo para los pesos dados por las frecuencias de los caracteres que aparecen en \mathcal{M} .

Precisamente el **Algoritmo de Huffman** resuelve este problema.

Huffman(L).

- Entrada: Una lista de pesos $L = [\alpha_1, \dots, \alpha_k]$ ordenados de manera creciente.
 - Salida: Un árbol binario ordenado óptimo para los pesos contenidos en L .
1. Si $k = 2$, devolver el árbol $T(L)$ de profundidad 1 cuyas hojas están etiquetadas con α_1 y α_2 .
 2. Si $k \geq 3$, sea L' la lista obtenida al insertar el elemento $\alpha_1 + \alpha_2$ en la sublista $[\alpha_3, \dots, \alpha_k]$ de modo que la lista resultante esté ordenada de manera creciente.
 3. Sea $T(L') := \text{Huffman}(L')$.
 4. Construir $T(L)$ a partir de $T(L')$ reemplazando una hoja de peso $\alpha_1 + \alpha_2$ en $T(L')$ por un subárbol de dos hojas con pesos α_1 y α_2 .

□

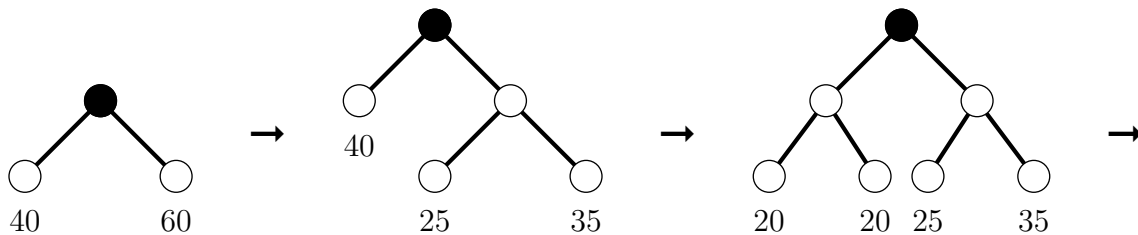
Ejemplo 51. Aplicamos el Algoritmo de Huffman a la secuencia de pesos

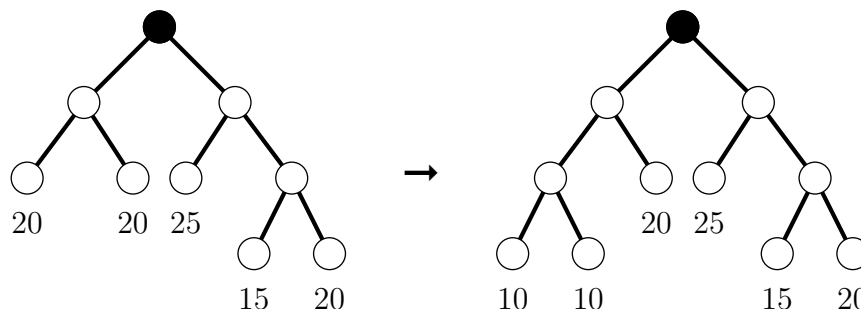
$$L = [10, 10, 15, 20, 20, 25]$$

obtenida a partir del Ejemplo 50. De hecho vamos a simplificar la mecánica del algoritmo como vemos a continuación:

$$10, 10, 15, 20, 20, 25 \rightarrow 15, \underline{20}, 20, 20, 25 \rightarrow 20, 20, 25, \underline{35} \rightarrow 25, 35, \underline{40} \rightarrow 40, \underline{60}$$

Nótese que cada elemento que insertamos en una lista, lo subrayamos. Ahora vamos hacia atrás construyendo el árbol binario ordenado óptimo para L :





Obsérvese que en el último paso tenemos tres posibilidades a la hora de construir el árbol binario ordenado óptimo.

Cualquiera de dichos árboles tiene peso 255, lo cual mejora sensiblemente el intento que hicimos en el Ejemplo 50.

Del árbol que hemos obtenido construimos un código binario prefijo óptimo. Recordemos para ello las frecuencias de los símbolos:

$$\left\{ \begin{array}{l} A : 25 \\ B : 10 \\ C : 10 \\ D : 20 \\ E : 15 \\ F : 20 \end{array} \right.$$

Resulta entonces el siguiente código binario prefijo óptimo:

$$\left\{ \begin{array}{l} A \longleftrightarrow 10 \\ B \longleftrightarrow 000 \\ C \longleftrightarrow 001 \\ D \longleftrightarrow 01 \\ E \longleftrightarrow 110 \\ F \longleftrightarrow 111 \end{array} \right.$$

□

El algoritmo de Huffman tiene múltiples aplicaciones, siendo la más popular los **algoritmos compresores de datos**. Normalmente estos algoritmos analizan el archivo a comprimir y generan una tabla de frecuencias de los caracteres que componen el archivo. A partir de las frecuencias obtenidas, aplican el Algoritmo de Huffman y obtienen un código binario prefijo óptimo con el que codifican el contenido del archivo. El mensaje codificado y el código binario prefijo óptimo son almacenados en el archivo comprimido final.

Nótese que al comprimir un archivo muy pequeño, el tamaño del archivo comprimido suele aumentar, pues apenas hay ganancia y además hay que almacenar el código binario prefijo óptimo obtenido para posteriormente llevar a cabo la descompresión.