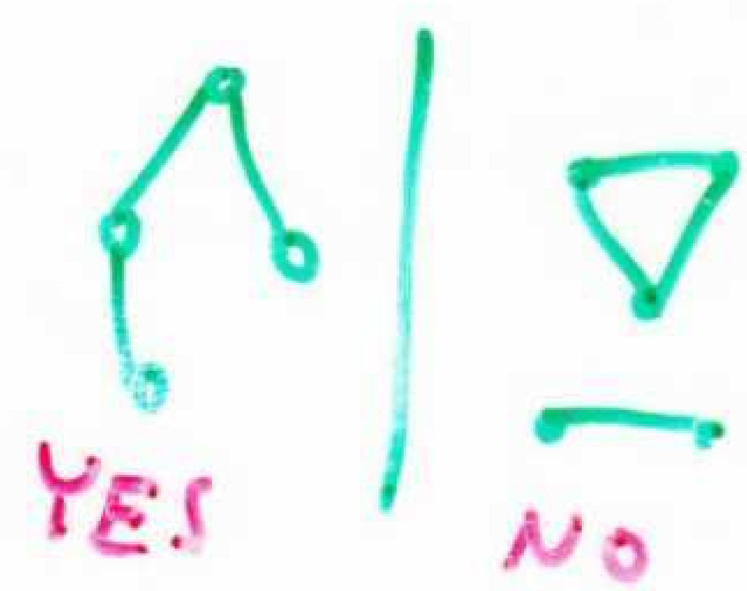
Graph Theory

A graph consists of a set of <u>vertices</u> V and a set of edges, or <u>ordered</u> pairs.

The reason graphs are so important is that they represent the set of binary relations $V \times V$.

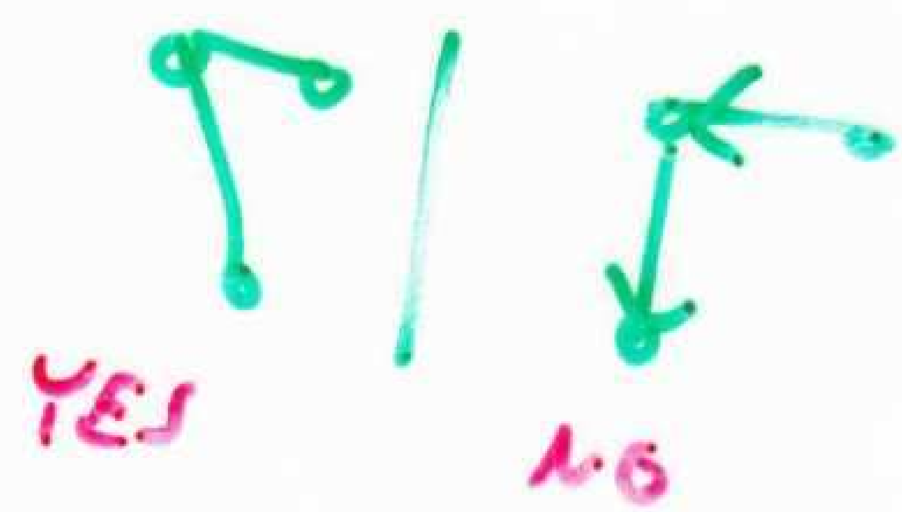
Graphs come in many different flavors

connected - there is a path between any pair of vertices.



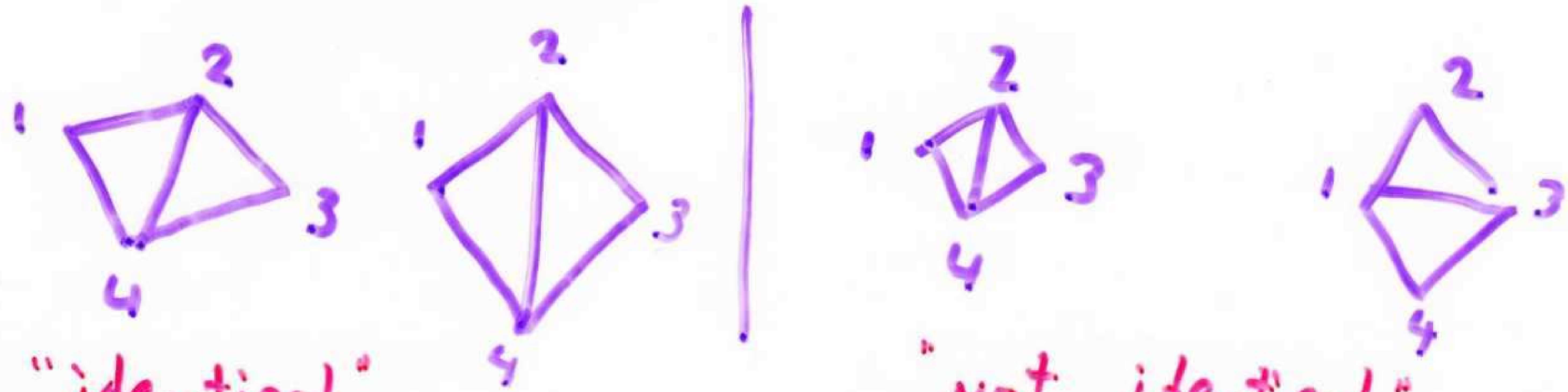
Simple - there are no self-loops and only one edge per pair of vertices.

Undirected - An edge (b, a) implies (a, b).



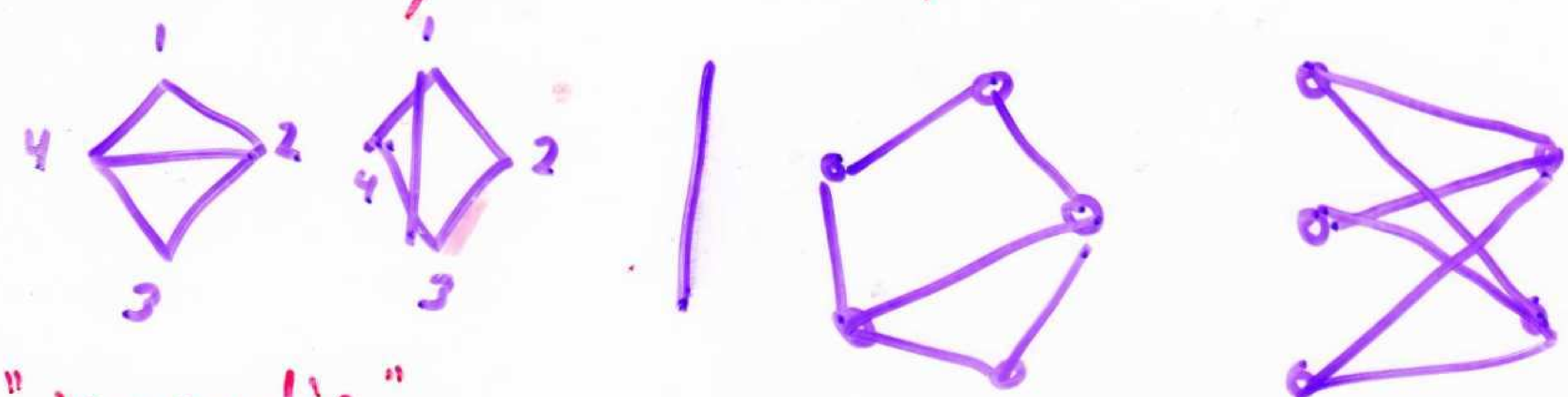
Most often, the graphs we will be interested in will be connected, simple, and undirected

Two labeled graphs are <u>identical</u> if there vertex and edge sets are equal:



"identical" "not identical"

Two labeled graphs are <u>isomorphic</u> if there exists a labeling of the vertices of the second graph such that they are identical:



"isomorphic" "not isomorphic"

Two graphs are isomorphic if for all "invariants", or functions independent of labeling, the invariants are equal.

The example on the right passes the "degree sequence" invariant but not the "girth" invariant- The length of the shortest cycle.

The most elementary invariants are the number of edges or vertices.

The **degree** d_i of a vertex is the number of edges **incident** upon it or the number of vertices **adjacent** to it. The **degree sequence** of a graph $(d_1, d_2, d_3 \ldots d_n)$ is the set of all vertex degrees.

Clearly, the degree sequence of a graph is an invariant, since two graphs of different degree sequences cannot be isomorphic.

Theorem: $\sum_{v \in V} d(v) = 2m$, where m is the number of edges

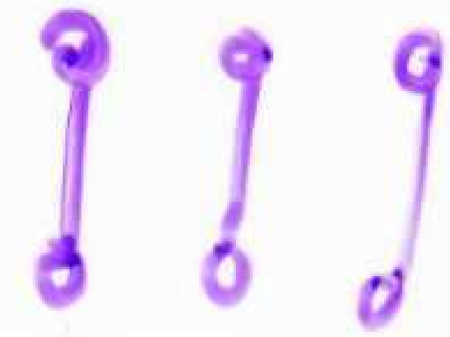
Proof: each edge contributes 1 to the degree of two vertices

Corollary: In any graph, the number of vertices of odd degree is even.

Graphic Degree Sequences

A list of numbers is called <u>graphic</u> when it can be realized as the degree sequence of some simple, undirected graph. Under what conditions is a sequence graphic?

(4, 4, 4, 4, 4) YES  Complete graph, K_5

(1, 1, 1, 1, 1, 1) YES  No such "connected" graph exists

(1, 1, 1, 1, 1, 1, 1) <u>No</u> sums to an odd number

(6, 5, 4, 3, 2, 1) <u>No</u> largest degree too large.

There are two different characterizations of when a degree sequence is graphic, one recursive due to Havel & Hakimi, the other by Erdos & Gallai.

Theorem: A degree sequence of an even integer $(d_1, d_2, d_3, \ldots, d_N)$ where $N > d_1 \geq d_2 \ldots \geq d_N \geq 0$ is graphical if and only if

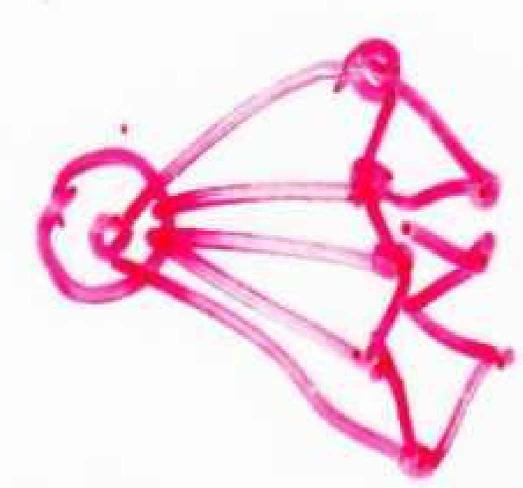
$$(d_2-1, d_3-1, \ldots d_{d_1+1}-1, d_{d_1+2}, \ldots d_N)$$

is graphical.

Proof: Clearly, any sequence satisfying this

if

condition is graphical - construct it on $N-1$ verties and connect the N^{th} to the d_1 highest degree verties:



(B → A)

only if

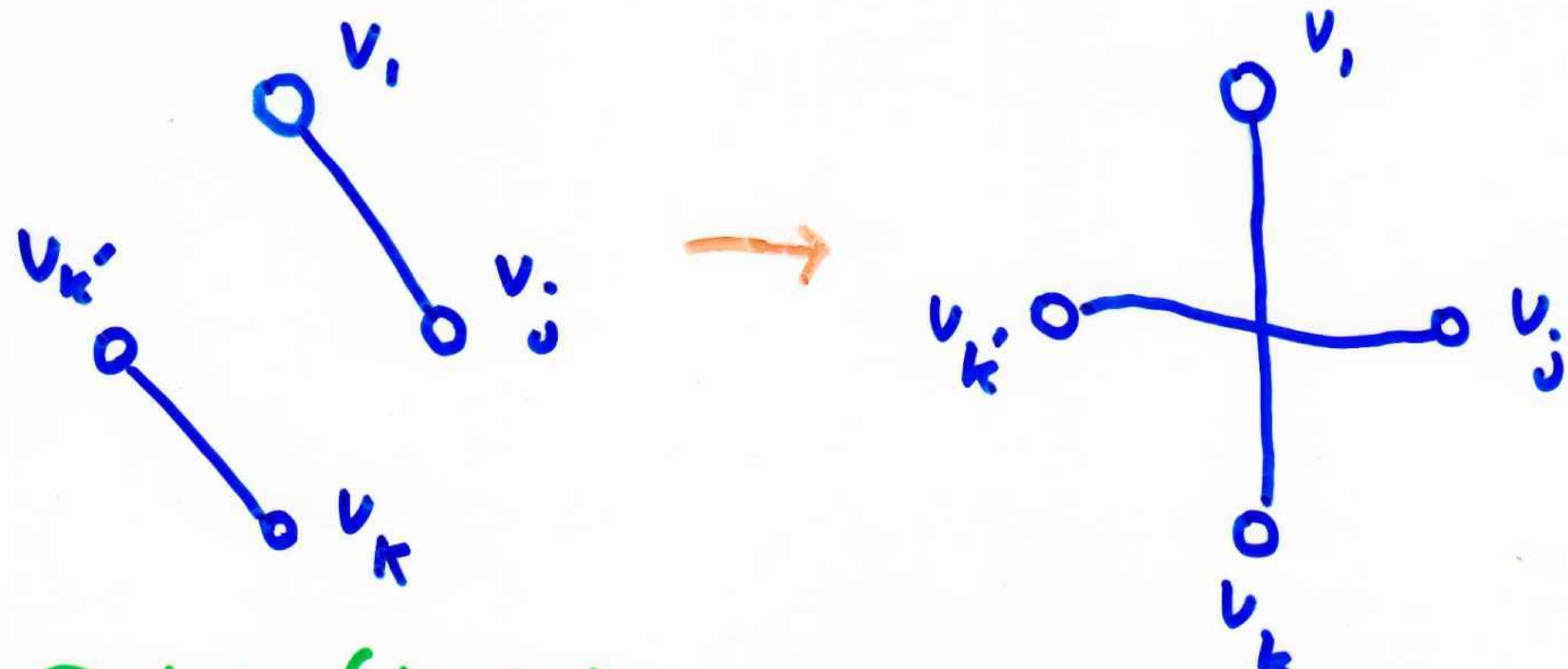
Suppose you have a graph with some degree sequence which is not constructed like this. We can construct a graph with the same sequence with the highest degree vertex connected to high degree verties by edge exchange operation:

(A → B)

Edge Interchange Operation

Suppose v_1 is adjacent to v_j but not v_k, where $d(v_k) > d(v_j)$.

There <u>must</u> exist a vertex v_k' which is adjacent to v_k but not v_j



Delete (v_1, v_j)
Delete (v_k', v_k)
Add (v_1, v_k)
Add (v_k', v_j)

The degree sequence is preserved, but v_1 is adjacent to a higher degree vertex than before. Repeat until the condition of the theorem is true. ∎

Example: (5, 5, 3, 3, 2, 2, 2)

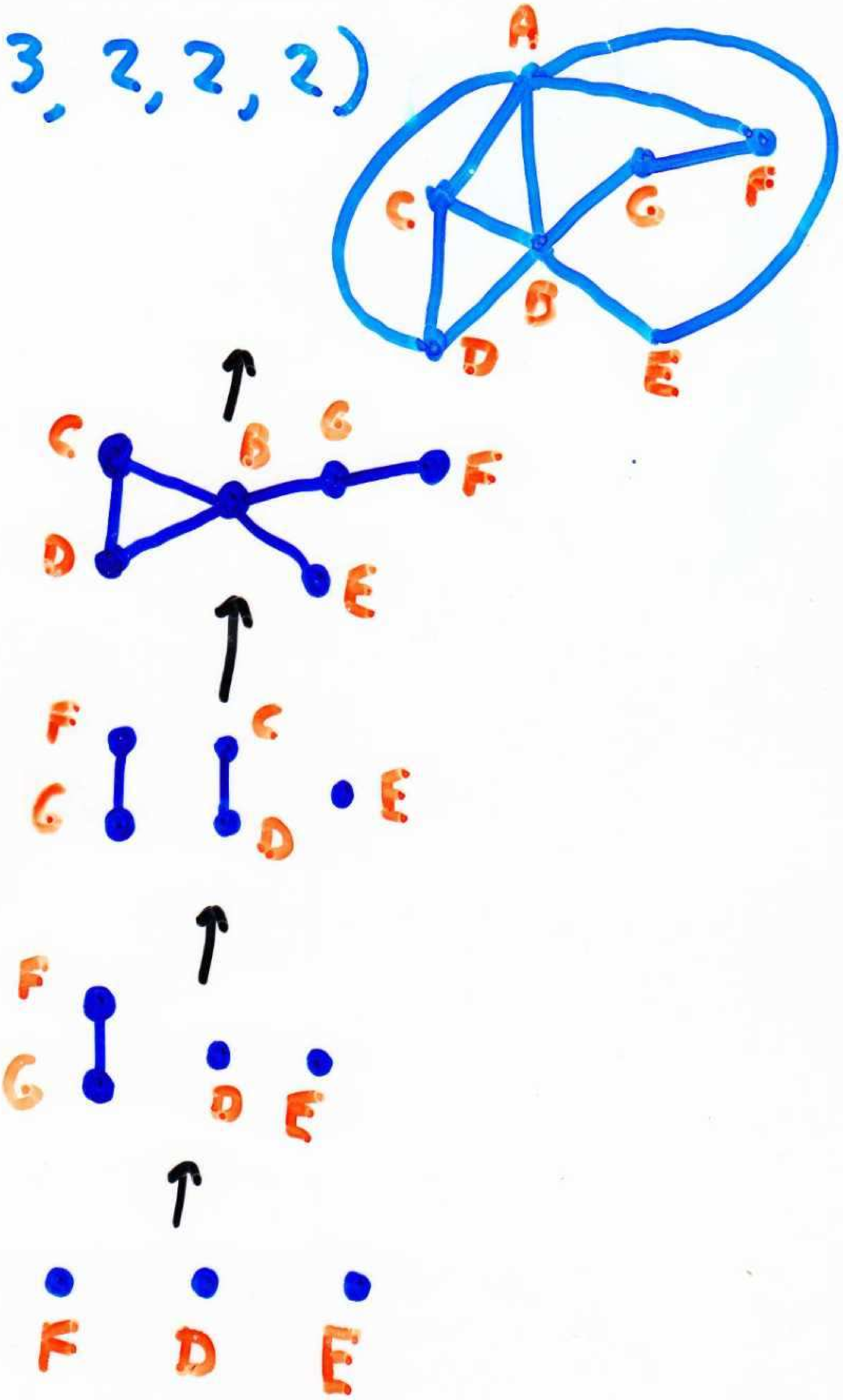
A B C D E F G
(5, 5, 3, 3, 2, 2, 2)

B C D G E F
(4, 2, 2, 2, 1, 1)

C D G F E
(1, 1, 1, 1, 0)

G F D E
(1, 1, 0, 0)

F D E
(0, 0, 0)



Note that subtraction permutes the order of the partition and edge flips

When does a partition define the degree sequence of a tree?

iff $2q = \sum_{i=1}^{N} d_i$, $d_i > 0$, $q = N-1$

Why?

if $2(N-1) = \sum_{i=1}^{N} d_i$, then $d_N = 1$

(if $2q = \sum d_i \rightarrow$ tree)

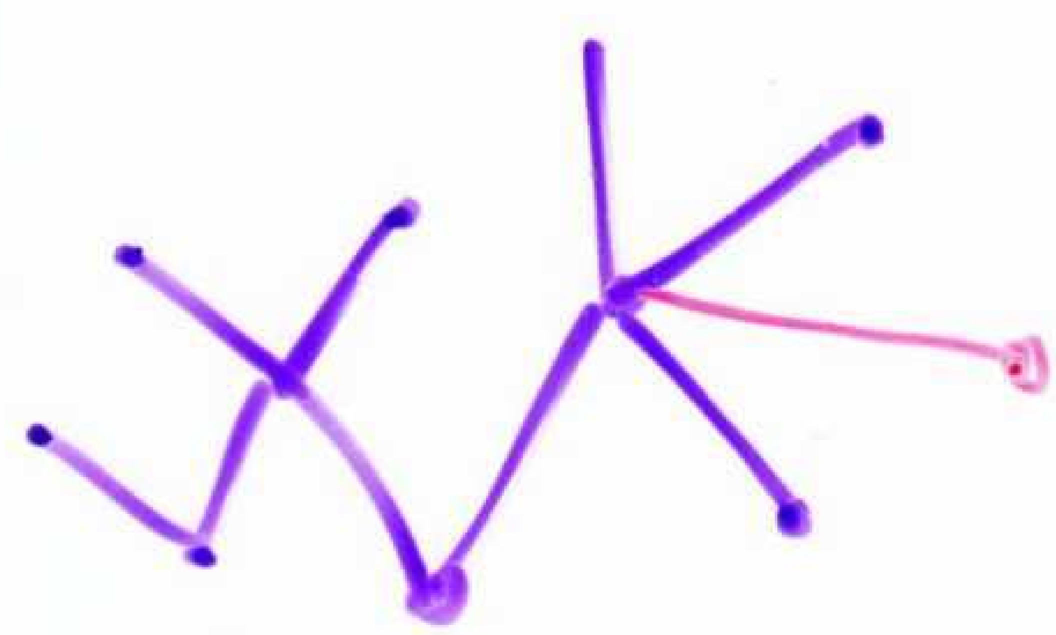
Proof: by induction

if $N=2$  ✓

assume true for $N-1$

then $(d_1 - 1, d_2 \ldots d_{N-1})$ defines

a tree:



add back the missing edge to create a tree!

(if tree $\rightarrow 2q = \sum d_i$)

EASY

The Erdos-Gallai Condition

An alternate condition for graphical degree sequences does not rely on recursion:

<u>Theorem</u>: Let $(d_1, d_2, \dots d_N)$ be an <u>even</u> sequence listed in non-increasing order. The sequence is graphical if and only if for each $1 \le r \le N-1$,

$$\sum_{i=1}^{r} d_i \le r(r-1) + \sum_{i=r+1}^{N} \min\{r, d_i\}$$

<u>Example</u>: $(\overset{1}{5}, \overset{2}{5}, \overset{3}{3}, \overset{4}{3}, \overset{5}{2}, \overset{6}{2}, \overset{7}{2})$

$$5 \le 0 + (1+1+1+1+1+1)$$
$$10 \le 2 + (2+2+2+2+2)$$
$$13 \le 6 + (3+2+2+2)$$
$$16 \le 12 + (2+2+2)$$
$$18 \le 20 + (2+2)$$
$$20 \le 30 + (2)$$

✓

GRAPHIC

$(8,7,6,5,4,3,2,2,1)$

$$8 \le 0 + (1+1+1+1+1+1+1+1)$$
$$15 \le 2 + 2\cdot 6 + 1$$
$$21 \not\le 6 + 3\cdot 3 + 2\cdot 2 + 1$$

NOT GRAPHIC

$$\sum_{i=1}^{r} d_i \leq r(r-1) + \sum_{i=r+1}^{p} \min\{r, d_i\}$$

Must prove two things for "if and only if"

1. Any graphic sequence satisfies this condition

→ Observe that for any set of r vertices the sum of the degrees are at most r(r-1) among themselves + that other vertices are each connected to at most $\min\{r, d_i\}$ members of the set.

2. Any sequence which satisfies this condition is graphic.

Our proof of part 2 is by induction induction on the number of vertices.

$N = 2$ $(1,1)$, $(0,0)$ (good) ✓

$(2,0)$ (bad)

Assuming $d_1 \cdots d_N$ satisfy the terms of our condition, we can construct our graph as we did before - connecting the largest degree vertex to the other largest vertices:

if m & n are the largest integers such that:

$$d_{m+1} = \ldots d_{d_1+1} = \cdots = d_N,$$

$$e_i \begin{cases} d_{i+1} - 1 & \text{for } i = 1 \text{ to } m-1 \text{ + } N-1-(d_1-m) \text{ to } N-1 \\ d_{i+1} & \text{otherwise} \end{cases}$$

We must now show that if $d_1, d_2 \ldots d_N$ satisfied the condition, so does $e_1, e_2 \ldots e_{N-1}$.

We will proceed by contradiction.

Let h be the first index which violates the balance.

 From Harary's Graph Theory

The proof of the sufficiency is by induction on p. Clearly the result holds for sequences of two parts. Assume that it holds for sequences of p parts, and let $d_1, d_2, \cdots, d_{p+1}$ be a sequence satisfying the hypotheses of the theorem.

Let m and n be the smallest and largest integers such that

$$d_{m+1} = \cdots = d_{d_1+1} = \cdots = d_n.$$

Form a new sequence of p terms by letting

$$e_i = \begin{cases} d_{i+1} - 1 & \text{for } i = 1 \text{ to } m - 1 \text{ and } n - 1 - (d_1 - m) \text{ to } n - 1, \\ d_{i+1} & \text{otherwise.} \end{cases}$$

If the hypotheses of the theorem hold for the new sequence $e_1, \cdots, e_p$, then by the induction hypothesis, there will be a graph with the numbers e_i as degrees. A graph having the given degree sequence d_i will be formed by adding a new point of degree d_1 adjacent to points of degrees corresponding to those terms e_i which were obtained by subtracting 1 from terms d_{i+1} as above.

Assume h is the first condition that fails.

Clearly $p > e_1 \geq e_2 \geq \cdots \geq e_p$. Suppose that condition (6.1) does not hold and let h be the least value of r for which it does not. Then

$$\sum_{i=1}^{h} e_i > h(h-1) + \sum_{i=h+1}^{p} \min\{h, e_i\}. \tag{6.2}$$

But the following inequalities do hold:

original graph satisfies

$$\sum_{i=1}^{h+1} d_i \leq h(h+1) + \sum_{i=h+2}^{p+1} \min\{h+1, d_i\}, \tag{6.3}$$

before h satisfies

$$\sum_{i=1}^{h-1} e_i \leq (h-1)(h-2) + \sum_{i=h}^{p} \min\{h-1, e_i\}, \tag{6.4}$$

$$\sum_{i=1}^{h-2} e_i \leq (h-2)(h-3) + \sum_{i=h-1}^{p} \min\{h-2, e_i\}. \tag{6.5}$$

Let s denote the number of values of $i \leq h$ for which $e_i = d_{i+1} - 1$. Then (6.3)–(6.5) when combined with (6.2) yield

6.2+6.3

$$d_1 + s < 2h + \sum_{i=h+1}^{p} (\min\{h+1, d_{i+1}\} - \min\{h, e_i\}), \tag{6.6}$$

6.2+6.4

$$e_h > 2(h-1) - \min\{h-1, e_h\} + \sum_{i=h+1}^{p} (\min\{h, e_i\} - \min\{h-1, e_i\}), \tag{6.7}$$

$e_h \geq h$ is implied

6.2+6.5

$$e_{h-1} + e_h > 4h - 6 - \min\{h-2, e_{h-1}\} - \min\{h-2, e_h\} + \sum_{i=h+1}^{p} (\min\{h, e_i\} - \min\{h-2, e_i\}). \tag{6.8}$$

Note that $e_h \geq h$ since otherwise inequality (6.7) gives a contradiction. Let a, b, and c denote the number of values of $i > h$ for which $e_i > h$, $e_i = h$, and $e_i < h$, respectively. Furthermore, let a', b', and c' denote the numbers of these for which $e_i = d_{i+1} - 1$. Then

$$d_1 = s + a' + b' + c'. \tag{6.9}$$

by definition

The inequalities (6.6)–(6.8) now become

$$d_1 + s < 2h + a + b' + c', \tag{6.10}$$

$$e_h \geq h + a + b, \tag{6.11}$$

$$e_{h-1} + e_h \geq 2h - 1 + \sum_{i=h+1}^{p} (\min\{h, e_i\} - \min\{h - 2, e_i\}). \tag{6.12}$$

There are now several cases to consider.

You can verify that this case analysis is complete, and always leads to a contradiction

CASE 1. $c' = 0$. Since $d_1 \geq e_h$, we have from (6.11),

$$h + a + b \leq d_1.$$

But a combination of (6.9) and (6.10) gives

$$2d_1 < 2h + a + a' + 2b',$$

which is a contradiction.

CASE 2. $c' > 0$ and $d_{h+1} > h$. This means that $d_{i+1} = e_i + 1$ whenever $d_{i+1} > h$. Therefore since $d_{h+1} > h$, $s = h$ and $a = a'$. But the inequalities (6.10) and (6.9) imply that

$$d_1 + h < 2h + a' + b' + c' = d_1 + h,$$

a contradiction.

CASE 3. $c' > 1$ and $d_{h+1} = h$. Under these circumstances, $e_h = h$ and $a = b = 0$, so $d_1 = s + c'$. Furthermore, since $e_h = d_{h+1}$, $e_i = h - 1$ for at least c' values of $i > h$. Hence inequality (6.12) implies

$$e_{h-1} \geq h - 1 + c' > h$$

so that $e_{h-1} = d_h - 1$. Therefore $s = h - 1$, and

$$d_1 = h - 1 + c' \leq e_{h-1} < d_h,$$

a contradiction.

CASE 4. $c' = 1$ and $d_{h+1} = h$. Again, $e_h = h$, $a = b = 0$, and $d_1 = s + c'$. Since $s \leq h - 1$, $d_1 = h$. But this implies $s = 0$ and $d_1 = 1$, so all $d_i = 1$. Thus (6.1) is obviously satisfied, which is a contradiction.

$d_{h+1} > h$

Since $e_h \geq h$ and $d_{h+1} \geq e_h$, we see that d_{h+1} cannot be less than h. Thus all possible cases have been considered and the proof is complete.

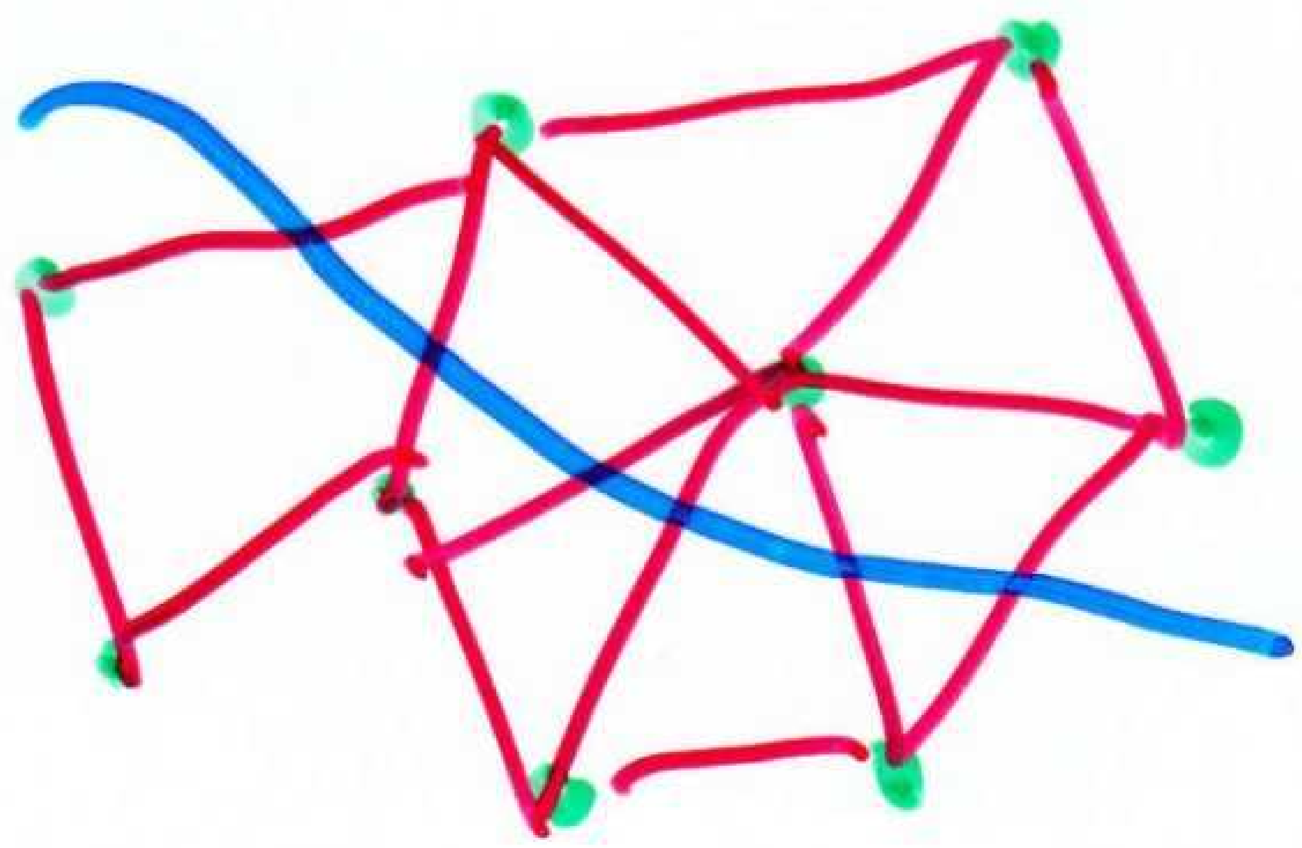
We started talking about degree sequences as an invariant under isomorphism.

There exist much "better" invariants to compute to test whether two graphs are isomorphic - all pairs shortest paths, number of k length cycles through each vertex, etc.

Although these work well in practice, no polynomial algorithm is know and determining whether one exists is a famous open problem!

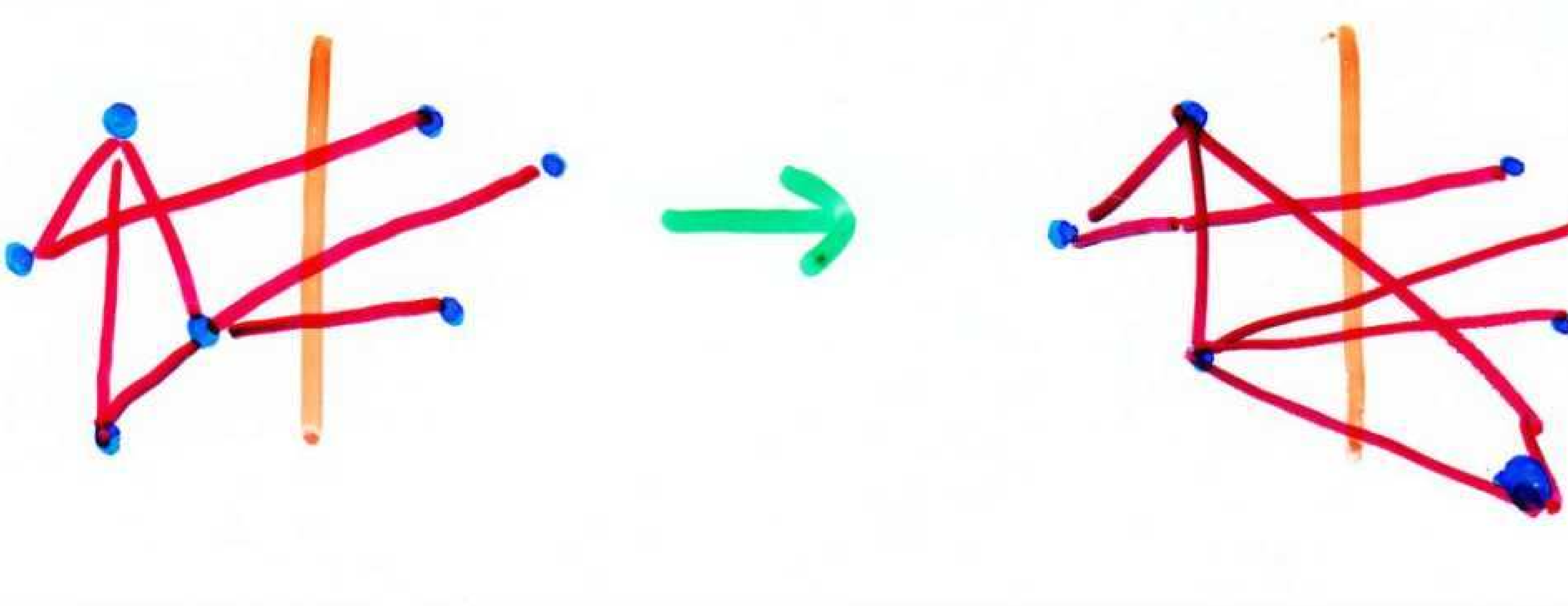
Finding a Large Cutset

A cutset of a graph is a set of edges each of which is needed to partition the vertices into two components (subsets of vertices A & B, with no edge (x, y) $x \in A$ & $y \in B$)



Every graph has a cutset with at least half the edges in the graph!

Proof: Arbitarily partition the vertices into sets A & B. If there is a vertex connected to more vertices on its side than the other, move it to the other side – repeat until no such vertices exist.



At the end, at least half of each vertices degree must be part of the cutset.

$$\sum_{i=1}^{V} d_i = 2E \rightarrow \sum_{i=1}^{N} d_i^c \geq \sum_{i=1}^{N} d_i/2 \geq E$$

Since each cut degree of 2 is one cut edge, the cut $\geq E/2$

The problem of finding the largest cut is NP-complete, but a random assignment of vertices to partitions should find a ½ approximation

$$\text{Prob}(\text{edge}(i,j) \text{ is cut}) = \begin{array}{c} \text{Prob}(i \in A) \cdot (j \in B) \\ \text{or} \\ \text{Prob}(i \in B) + (j \in A) \end{array}$$

$$= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

Thus each edge has a ½ chance of getting cut on a random assignment!