**Math 5707 Spring 2017 (Darij Grinberg): homework set 3**
**Solution sketches (DRAFT).**

# Contents



## 0.1. Reminders

See the lecture notes and also the handwritten notes for relevant material. See also the solutions to homework set 2 for various conventions and notations that are in use here.

If $G = (V, E, \phi)$ is a multigraph, and if $v \in V$ and $e \in E$, then the edge $e$ is said to be *incident* to the vertex $v$ (in the multigraph $G$) if and only if $v \in \phi(e)$ (in other words, if and only if $v$ is an endpoint of $e$).

## 0.2. Exercise 1: Centers of trees lie on longest paths

If $v$ is a vertex of a simple graph $G = (V, E)$, then the *eccentricity* of $v$ is defined to be $\max\{d(v, u) \mid u \in V\}$ (where $d(v, u)$ is the distance between $v$ and $u$, as usual). A *center* of a simple graph $G$ means a vertex whose eccentricity is minimum (among the eccentricities of all vertices).

**Exercise 1.** Let $T$ be a tree. Let $(v_0, v_1, \ldots, v_k)$ be a longest path of $T$. Prove that each center of $T$ belongs to this path (i.e., is one of the vertices $v_0, v_1, \ldots, v_k$).

In preparation for the solution of this exercise, we cite a result from lecture 10:

**Proposition 0.1.** Let $T$ be a tree with $\geq 3$ vertices. Let $L$ be the set of all leaves of $T$. Let $T \setminus L$ be the multigraph obtained from $T$ by removing the vertices in $L$ and all edges incident to them.

The eccentricity of a vertex $v$ of a graph $G$ will be denoted by $\operatorname{ecc}_G v$.

**(a)** The graph $T \setminus L$ is a tree.

**(b)** Each vertex $v$ of $T \setminus L$ satisfies $\operatorname{ecc}_T v = \operatorname{ecc}_{T \setminus L} v + 1$.

**(c)** Each $v \in L$ satisfies $\operatorname{ecc}_T v = \operatorname{ecc}_T w + 1$, where $w$ is the unique neighbor of $v$ in $T$.

**(d)** The centers of $T$ are precisely the centers of $T \setminus L$.

We contrast this with the following simple fact:

**Proposition 0.2.** Let $T$ be a tree with $\geq 3$ vertices. Let $L$ be the set of all leaves of $T$. Let $T \setminus L$ be the multigraph obtained from $T$ by removing the vertices in $L$ and all edges incident to them.

Let $(v_0, v_1, \ldots, v_k)$ be a longest path of $T$. Then, $(v_1, v_2, \ldots, v_{k-1})$ is a longest path of $T \setminus L$.

*Proof of Proposition 0.2 (sketched).* For each $i \in \{1, 2, \ldots, k-1\}$, the vertex $v_i$ is a vertex of $T \setminus L$ [1]. Hence, $(v_1, v_2, \ldots, v_{k-1})$ is a path of $T \setminus L$. It remains to prove that it is a **longest** path.

Indeed, assume the contrary. Hence, there exists a longest path $(w_1, w_2, \ldots, w_m)$ of $T \setminus L$ whose length $m-1$ is greater than the length $k-2$ of the path $(v_1, v_2, \ldots, v_{k-1})$. Consider such a path. From $m-1 > k-2$, we obtain $m > k-1$. Hence, $m \geq k$.

The vertex $w_1$ must be a leaf of $T \setminus L$ (since otherwise, it would have a neighbor distinct from $w_2$, which would then allow us to extend the path $(w_1, w_2, \ldots, w_m)$ by attaching this neighbor to its front; but this would contradict the fact that this path $(w_1, w_2, \ldots, w_m)$ is a longest path of $T \setminus L$). But the vertex $w_1$ cannot be a leaf of $T$ (since in this case, it would belong to $L$, and hence could not be a vertex of $T \setminus L$). Hence, the vertex $w_1$ has at least one more neighbor in $T$ than it has in $T \setminus L$ (because it is a leaf of $T \setminus L$ but not of $T$). Therefore, the vertex $w_1$ has at least one neighbor in $T$ that is not a vertex of $T \setminus L$. Fix such a neighbor, and denote it by $w_0$. This neighbor $w_0$ must lie in $L$ (since it is not a vertex of $T \setminus L$).

Similarly, the vertex $w_m$ has at least one neighbor in $T$ that is not a vertex of $T \setminus L$. Fix such a neighbor, and denote it by $w_{m+1}$. This neighbor $w_{m+1}$ must lie in $L$ (since it is not a vertex of $T \setminus L$).

Now, $(w_0, w_1, \ldots, w_{m+1})$ is a walk in $T$. Furthermore, all the $m+2$ vertices of this walk are distinct[2]. Hence, this walk is a path. This path has length $m+1 > m \geq k$, and thus is longer than the path $(v_0, v_1, \ldots, v_k)$. But this is absurd, since the

[1] *Proof.* Let $i \in \{1, 2, \ldots, k-1\}$. Then, $v_{i-1}v_i$ and $v_i v_{i+1}$ are two distinct edges of $T$ (since $(v_0, v_1, \ldots, v_k)$ is a path of $T$). Hence, the vertex $v_i$ of $T$ belongs to at least two distinct edges (namely, $v_{i-1}v_i$ and $v_i v_{i+1}$), and thus is not a leaf of $T$. In other words, $v_i$ is not an element of $L$. Hence, $v_i$ is a vertex of $T \setminus L$.

[2] *Proof.* First, we notice that the $m$ vertices $w_1, w_2, \ldots, w_m$ are distinct (since $(w_1, w_2, \ldots, w_m)$ is a path of $T \setminus L$). Second, we observe that the two vertices $w_0$ and $w_{m+1}$ are distinct from the $m$

latter path is a longest path of $T$. Hence, we have found a contradiction, which is precisely what we wanted to find. The proof is thus complete. $\square$

*Hints to Exercise 1.* Proceed by strong induction on $|\operatorname{V}(T)|$. Thus, fix $N \in \mathbb{N}$, and assume (as the induction hypothesis) that Exercise 1 has already been solved for all trees $T$ satisfying $|\operatorname{V}(T)| < N$. Now, fix a tree $T$ satisfying $|\operatorname{V}(T)| = N$. We must prove that Exercise 1 holds for this particular tree $T$.

If $|\operatorname{V}(T)| < 3$, then this is obvious (because in this case, each vertex of $T$ lies on the longest path $(v_0, v_1, \ldots, v_k)$). Thus, we WLOG assume that $|\operatorname{V}(T)| \geq 3$.

Let $L$ be the set of all leaves of $T$. Thus, $|L| \geq 2$ (since $T$ has at least two leaves (since $T$ is a tree with at least 2 vertices)).

Let $T \setminus L$ be the multigraph obtained from $T$ by removing the vertices in $L$ and all edges incident to them.

Proposition 0.1 **(a)** shows that the graph $T \setminus L$ is a tree. Proposition 0.1 **(d)** shows that the centers of $T$ are precisely the centers of $T \setminus L$.

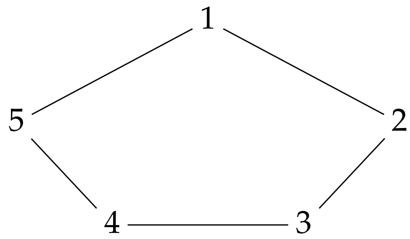
Let $(v_0, v_1, \ldots, v_k)$ be a longest path of $T$. Proposition 0.2 shows that $(v_1, v_2, \ldots, v_{k-1})$ is a longest path of $T \setminus L$.

Clearly, $\operatorname{V}(T \setminus L) = \operatorname{V}(T) \setminus L$, so that $|\operatorname{V}(T \setminus L)| = |\operatorname{V}(T) \setminus L| = \underbrace{|\operatorname{V}(T)|}_{=N} - \underbrace{|L|}_{\geq 2 > 0} < N$. Hence, the induction hypothesis shows that Exercise 1 holds for $T \setminus L$ instead of $T$. Thus, we can apply Exercise 1 to $T \setminus L$ and $(v_1, v_2, \ldots, v_{k-1})$ instead of $T$ and $(v_0, v_1, \ldots, v_k)$. We thus conclude that each center of $T \setminus L$ belongs to the path $(v_1, v_2, \ldots, v_{k-1})$. Since the centers of $T$ are precisely the centers of $T \setminus L$, we can rewrite this as follows: Each center of $T$ belongs to the path $(v_1, v_2, \ldots, v_{k-1})$. Thus, each center of $T$ belongs to the path $(v_0, v_1, \ldots, v_k)$ as well (since the latter path contains the former path). In other words, Exercise 1 holds for our particular tree $T$. This completes the induction; thus, Exercise 1 is solved. $\square$

vertices $w_1, w_2, \ldots, w_m$ (since the former lie in $L$, whereas the latter are vertices of $T \setminus L$). Thus, it only remains to prove that the vertices $w_0$ and $w_{m+1}$ are distinct. But this is easy: If they were not, then the walk $(w_0, w_1, \ldots, w_{m+1})$ would be a cycle; but this would contradict the fact that $T$ has no cycles (since $T$ is a tree). Thus, we have shown that all the $m+2$ vertices of the walk $(w_0, w_1, \ldots, w_{m+1})$ are distinct.

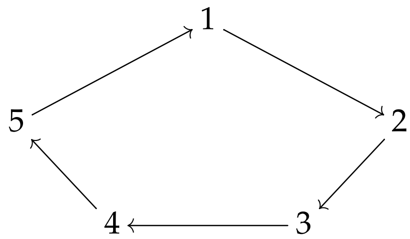
## 0.3. Exercise 2: Counting spanning trees in some cases

**Exercise 2. (a)** Consider the cycle graph $C_n$ for some $n \geq 2$. Its vertices are $1, 2, \ldots, n$, and its edges are $12, 23, \ldots, (n-1)\,n, n1$. (Here is how it looks for $n = 5$:




) Find the number of spanning trees of $C_n$.

**(b)** Consider the directed cycle graph $\overrightarrow{C}_n$ for some $n \geq 2$. It is a digraph; its vertices are $1, 2, \ldots, n$, and its arcs are $12, 23, \ldots, (n-1)\,n, n1$. (Here is how it looks for $n = 5$:




) Find the number of oriented spanning trees of $\overrightarrow{C}_n$ with root 1.

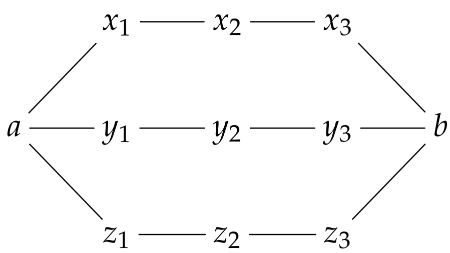
**(c)** Fix $m \geq 1$. Let $G$ be the simple graph with $3m + 2$ vertices

$$a, b, x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_m, z_1, z_2, \ldots, z_m$$

and the following $3m + 3$ edges:

$$\begin{aligned} & ax_1, ay_1, az_1, \\ & x_i x_{i+1}, y_i y_{i+1}, z_i z_{i+1} \qquad \text{for all } i \in \{1, 2, \ldots, m-1\}, \\ & x_m b, y_m b, z_m b. \end{aligned}$$

(Thus, the graph consists of two vertices $a$ and $b$ connected by three paths, each of length $m + 1$, with no overlaps between the paths except for their starting and ending points. Here is a picture for $m = 3$:




) Compute the number of spanning trees of $G$.

[To argue why your number is correct, a sketch of the argument in 1-2 sentences should be enough; a fully rigorous proof is not required.]

*Hints to Exercise 2.* **(a)** The number is $n$. Indeed, any graph obtained from $C_n$ by removing a single edge is a spanning tree of $C_n$.

[*Proof:* Recall that a tree with $n$ vertices must have exactly $n-1$ edges. Thus, a spanning subgraph of $C_n$ can be a tree only if it has $n-1$ edges, i.e., only if it is obtained from $C_n$ by removing a single edge. It only remains to check that any subgraph obtained from $C_n$ by removing a single edge is indeed a spanning tree. But this is easy, since all such subgraphs are isomorphic to the path graph $P_n$.]

**(b)** The number is 1. Indeed, the only oriented spanning tree of $\overrightarrow{C}_n$ with root 1 is the subdigraph of $\overrightarrow{C}_n$ obtained by removing the arc 12.

[*Proof:* How can an oriented spanning tree of $\overrightarrow{C}_n$ with root 1 look like? It must have no arc with source 1 (since 1 is its root), so it must not contain the arc 12.

But it must contain a walk from $i$ to 1 for each vertex $i \in \{2, 3, \ldots, n\}$. Thus, it must contain at least one arc with source $i$ for each $i \in \{2, 3, \ldots, n\}$ (because otherwise, we would have no way to get out of $i$, and this would render a walk from $i$ to 1 impossible). This arc clearly must be $i(i+1)$, where we set $n+1 = 1$ (because the only arc with source $i$ in $\overrightarrow{C}_n$ is the arc $i(i+1)$).

Hence, our oriented spanning tree must not contain the arc 12, but it must contain the arc $i(i+1)$ for each $i \in \{2, 3, \ldots, n\}$. This uniquely defines this oriented spanning tree. Conversely, it is trivial that the subdigraph of $\overrightarrow{C}_n$ obtained by removing the arc 12 is indeed an oriented spanning tree.]

**(c)** The number is $3(m+1)^2$. Indeed, let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ be the three paths $(a, x_1, x_2, \ldots, x_m, b)$, $(a, y_1, y_2, \ldots, y_m, b)$ and $(a, z_1, z_2, \ldots, z_m, b)$. Then, the spanning trees of $G$ are the subgraphs of $G$ obtained

- either by removing an edge from $\mathbf{x}$ and an edge from $\mathbf{y}$ (there are $(m+1)^2$ ways to do that);
- or by removing an edge from $\mathbf{x}$ and an edge from $\mathbf{z}$ (there are $(m+1)^2$ ways to do that);
- or by removing an edge from $\mathbf{y}$ and an edge from $\mathbf{z}$ (there are $(m+1)^2$ ways to do that).

[*Proof:* The graph $G$ has $3m+2$ vertices. Hence, any spanning tree of $G$ must have $(3m+2) - 1 = 3m+1$ edges. This means that any spanning tree of $G$ can be obtained from $G$ by removing two edges (since $G$ has $3m+3$ edges). But not each pair of edges yields a spanning tree when removed. Which ones do, and which ones do not?

- If we remove two edges from $\mathbf{x}$, then the subgraph is not connected (indeed, at least one vertex on the path $\mathbf{x}$ lies between these two edges, and this vertex is disconnected from $a$ in this subgraph), and thus not a tree. The same problem happens if we remove two edges from $\mathbf{y}$ or two edges from $\mathbf{z}$.

- If we remove an edge from **x** and an edge from **y**, then the subgraph is connected (because any vertex is still connected to at least one of $a$ and $b$, but $a$ and $b$ are also still connected to each other via the undamaged path **z**), and thus is a spanning tree of $G$ (since it is connected and has the "right" number of edges). The same happens if we remove an edge from **x** and an edge from **z**, or if we remove an edge from **y** and an edge from **z**.

There are no other cases. Thus, tallying these possibilities, we obtain the characterization of spanning trees given above, and thus there are $(m+1)^2 + (m+1)^2 + (m+1)^2 = 3(m+1)^2$ spanning trees.] □

[*Remark:* I guess that parts **(a)** and **(c)** of Exercise 2 can also be solved using the Matrix-Tree Theorem. But the solutions given above are definitely easier!]

## 0.4. Exercise 3: The number of connected components is supermodular

### 0.4.1. Statement of the problem

We first recall how the connected components of a multigraph were defined:

**Definition 0.3.** Let $G = (V, E, \phi)$ be a multigraph.
**(a)** We define a binary relation $\simeq_G$ on the set $V$ as follows: For two vertices $u$ and $v$ in $V$, we set $u \simeq_G v$ if and only if there exists a walk from $u$ to $v$ in $G$.
**(b)** The binary relation $\simeq_G$ is an equivalence relation on $V$. Its equivalence classes are called the *connected components* of $G$.

**Definition 0.4.** If $G$ is a multigraph, then $\operatorname{conn} G$ shall denote the number of connected components of $G$. (We have previously called this number $b_0(G)$ in lecture notes. Note that it equals 0 when $G$ has no vertices, and 1 if $G$ is connected.)

**Exercise 3.** Let $(V, H, \phi)$ be a multigraph. Let $E$ and $F$ be two subsets of $H$.
**(a)** Prove that

$$\begin{aligned} &\operatorname{conn}(V, E, \phi\mid_E) + \operatorname{conn}(V, F, \phi\mid_F) \\ &\leq \operatorname{conn}(V, E \cup F, \phi\mid_{E\cup F}) + \operatorname{conn}(V, E \cap F, \phi\mid_{E\cap F}). \end{aligned} \tag{1}$$

**(b)** Give an example where the inequality (1) does **not** become an equality.

### 0.4.2. Hints

**Remark 0.5.** The following two hints are helpful for solving Exercise 3 **(a)**:

- Feel free to restrict yourself to the case of a simple graph; in this case, $E$ and $F$ are two subsets of $\mathcal{P}_2(V)$, and you have to show that

$$\operatorname{conn}(V,E) + \operatorname{conn}(V,F) \leq \operatorname{conn}(V, E \cup F) + \operatorname{conn}(V, E \cap F).$$

  This isn't any easier than the general case, but saves you the hassle of carrying the map $\phi$ around.

- Also, feel free to take inspiration from the proof of the classical fact that $\dim X + \dim Y = \dim(X+Y) + \dim(X \cap Y)$ when $X$ and $Y$ are two subspaces of a finite-dimensional vector space $U$. That proof relies on choosing a basis of $X \cap Y$ and extending it to bases of $X$ and $Y$, then merging the extended bases to a basis of $X + Y$. A "basis" of a multigraph $G$ is a spanning forest: a spanning subgraph that is a forest and has the same number of connected components as $G$. More precisely, it is the set of the edges of a spanning forest.

  Actually, the second solution to Exercise 3 sketched below follows this idea (of imitating the proof of $\dim X + \dim Y = \dim(X+Y) + \dim(X \cap Y)$), whereas the third solution uses the identity $\dim X + \dim Y = \dim(X+Y) + \dim(X \cap Y)$ itself.

### 0.4.3. First solution

The following solution to Exercise 3 is probably the most conventional one. It is rather long due to the fact that certain properties of connected components, while being obvious to the eye and easy to explain with some handwaving, are painstakingly difficult to rigorously formulate. Despite its length, a few details are left to the reader (but they should be easy to fill in).

We prepare for the solution of Exercise 3 with a definition and two simple lemmas:

**Definition 0.6.** Let $G = (V, E, \phi)$ be a multigraph. Let $e$ be an edge of $G$. Then, $G - e$ will denote the multigraph obained from $G$ by removing the edge $e$. (Formally speaking, $G - e$ is the multigraph $\left(V, E \setminus \{e\}, \phi\mid_{E \setminus \{e\}}\right)$.)

Similar notations are used for simple graphs, for digraphs, and for multidigraphs.

**Lemma 0.7.** Let $G = (V, E, \phi)$ be a multigraph. Let $e$ be an edge of $G$. Let $u \in V$ and $v \in V$ be such that $u \simeq_G v$. Assume that we do not have $u \simeq_{G-e} v$. Then, there exist $x \in V$ and $y \in V$ such that $\phi(e) = \{x, y\}$ and $u \simeq_{G-e} x$ and $v \simeq_{G-e} y$.

*Proof of Lemma 0.7.* We do not have $u \simeq_{G-e} v$. In other words, there exists no walk from $u$ to $v$ in $G - e$.

But $u \simeq_G v$. Thus, there exists a walk from $u$ to $v$ in $G$. Hence, there exists a path from $u$ to $v$ in $G$. Fix such a path. Write this path in the form $(p_0, e_1, p_1, e_2, p_2, \ldots, e_k, p_k)$ with $p_0 = u$ and $p_k = v$. This path must contain the edge $e$ (since otherwise, it would be a path in $G - e$, thus also a walk in $G - e$; but this would contradict the fact that there exists no walk from $u$ to $v$ in $G - e$). In other words, $e_i = e$ for some $i \in \{1, 2, \ldots, k\}$. Consider this $i$.

The edges $e_1, e_2, \ldots, e_k$ are the edges of the path $(p_0, e_1, p_1, e_2, p_2, \ldots, e_k, p_k)$, and thus are distinct (since the edges of a path are always distinct). Hence, none of the edges $e_1, e_2, \ldots, e_{i-1}, e_{i+1}, e_{i+2}, \ldots, e_k$ equals $e_i$. Since $e_i = e$, this rewrites as follows: None of the edges $e_1, e_2, \ldots, e_{i-1}, e_{i+1}, e_{i+2}, \ldots, e_k$ equals $e$. Thus, all of the edges $e_1, e_2, \ldots, e_{i-1}, e_{i+1}, e_{i+2}, \ldots, e_k$ are edges of the multigraph $G - e$. Hence, the two subwalks[3] $(p_0, e_1, p_1, e_2, p_2, \ldots, e_{i-1}, p_{i-1})$ and $(p_i, e_{i+1}, p_{i+1}, e_{i+2}, p_{i+2}, \ldots, e_k, p_k)$ of the path $(p_0, e_1, p_1, e_2, p_2, \ldots, e_k, p_k)$ are walks in $G - e$.

Now, there exists a walk from $p_0$ to $p_{i-1}$ in $G - e$ (namely, the walk $(p_0, e_1, p_1, e_2, p_2, \ldots, e_{i-1}, p_{i-1})$). In other words, $p_0 \simeq_{G-e} p_{i-1}$. Since $p_0 = u$, this rewrites as $u \simeq_{G-e} p_{i-1}$.

Also, there exists a walk from $p_i$ to $p_k$ in $G - e$ (namely, the walk $(p_i, e_{i+1}, p_{i+1}, e_{i+2}, p_{i+2}, \ldots, e_k, p_k)$). In other words, $p_i \simeq_{G-e} p_k$. Since $\simeq_{G-e}$ is an equivalence relation, this shows that $p_k \simeq_{G-e} p_i$. Since $p_k = v$, this rewrites as $v \simeq_{G-e} p_i$.

We have $\phi(e_i) = \{p_{i-1}, p_i\}$ (since $(p_0, e_1, p_1, e_2, p_2, \ldots, e_k, p_k)$ is a path). Therefore, $\{p_{i-1}, p_i\} = \phi(e_i) = \phi(e)$ (since $e_i = e$). Hence, $\phi(e) = \{p_{i-1}, p_i\}$. Thus, there exist $x \in V$ and $y \in V$ such that $\phi(e) = \{x, y\}$ and $u \simeq_{G-e} x$ and $v \simeq_{G-e} y$ (namely, $x = p_{i-1}$ and $y = p_i$). This proves Lemma 0.7. □

**Lemma 0.8.** Let $G = (V, E, \phi)$ be a multigraph. Let $e$ be an edge of $G$. Let us use the Iverson bracket notation.

Then,
$$\operatorname{conn}(G - e) = \operatorname{conn} G + [e \text{ belongs to no cycle of } G].$$

*Proof of Lemma 0.8 (sketched).* Consider the two relations $\simeq_G$ and $\simeq_{G-e}$. (Recall that two vertices $u$ and $v$ of $G$ satisfy $u \simeq_G v$ if and only if there exists a walk from $u$ to $v$ in $G$. The relation $\simeq_{G-e}$ is defined similarly, but using $G - e$ instead of $G$.)

The connected components of $G$ are the equivalence classes of the relation $\simeq_G$. The connected components of $G - e$ are the equivalence classes of the relation $\simeq_{G-e}$.

Every two vertices $u \in V$ and $v \in V$ satisfying $u \simeq_{G-e} v$ satisfy $u \simeq_G v$ [4].

We shall use the notation $\operatorname{conncomp}_H w$ for the connected component of a multigraph $H$ containing a given vertex $w$. Thus, for each vertex $v \in V$, we have a

[3] Here, a *subwalk* of a walk $(w_0, f_1, w_1, f_2, w_2, \ldots, f_m, w_m)$ means a list of the form $(w_I, f_{I+1}, w_{I+1}, f_{I+2}, w_{I+2}, \ldots, f_J, w_J)$ for two elements $I$ and $J$ of $\{0, 1, \ldots, m\}$ satisfying $I \leq J$. Such a list is always a walk.

[4] *Proof.* Let $u \in V$ and $v \in V$ be two vertices satisfying $u \simeq_{G-e} v$. We must show that $u \simeq_G v$.

We know that $u \simeq_{G-e} v$. In other words, there exists a walk from $u$ to $v$ in $G - e$. This walk is clearly also a walk from $u$ to $v$ in $G$ (since $G - e$ is a submultigraph of $G$). Hence, there exists a walk from $u$ to $v$ in $G$. In other words, $u \simeq_G v$.

connected component $\operatorname{conncomp}_{G-e} v$ and a connected component $\operatorname{conncomp}_G v$. These two connected components satisfy

$$\operatorname{conncomp}_{G-e} v \subseteq \operatorname{conncomp}_G v$$

(because all vertices $u \in \operatorname{conncomp}_{G-e} v$ satisfy $u \simeq_{G-e} v$, thus $u \simeq_G v$, thus $u \in \operatorname{conncomp}_G v$); but the reverse inclusion might not hold. Hence, each connected component of $G$ is a union of a nonzero number (possibly just one, but possibly more) of connected components of $G - e$.

Write the set $\phi(e) \in \mathcal{P}_2(V)$ in the form $\phi(e) = \{a, b\}$.

We are in one of the following two cases:

*Case 1:* The edge $e$ belongs to no cycle of $G$.

*Case 2:* The edge $e$ belongs to at least one cycle of $G$.

We shall treat these two cases separately:

- Let us first consider Case 1. In this case, the edge $e$ belongs to no cycle of $G$. Then, we do not have $a \simeq_{G-e} b$ [5]. Hence, $\operatorname{conncomp}_{G-e} a \neq \operatorname{conncomp}_{G-e} b$. But we do have $a \simeq_G b$ (because the edge $e$ provides a walk $(a, e, b)$ from $a$ to $b$ in $G$).

  Now, recall that each connected component of $G$ is a union of a nonzero number (possibly just one, but possibly more) of connected components of $G - e$. In other words, the connected components of $G$ are obtained by merging **some** of the connected components of $G - e$. Which connected components get merged? On the one hand, we know that the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ of $G - e$ get merged in $G$ (since $a \simeq_G b$); and these two components were indeed distinct in $G - e$ (since $\operatorname{conncomp}_{G-e} a \neq \operatorname{conncomp}_{G-e} b$). On the other hand, we know that these are the **only** two connected components that get merged[6]. Altogether, we thus see that only two connected components are merged when passing

[5]*Proof.* Assume the contrary. Thus, we have $a \simeq_{G-e} b$. In other words, there exists a walk from $a$ to $b$ in $G - e$. Hence, there exists a path from $a$ to $b$ in $G - e$. Combining this path with the edge $e$, we obtain a cycle of $G$ that contains the edge $e$. Thus, the edge $e$ belongs to at least one cycle of $G$ (namely, to the cycle we have just constructed). This contradicts the fact that the edge $e$ belongs to no cycle of $G$.

[6]*Proof.* Let $P$ and $Q$ be two distinct connected components of $G - e$ that get merged in $G$ (possibly together with other connected components). We must show that $P$ and $Q$ are the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ (in some order).

We know that $P$ and $Q$ are two connected components of $G - e$. Hence, $P = \operatorname{conncomp}_{G-e} u$ and $Q = \operatorname{conncomp}_{G-e} v$ for some $u \in V$ and $v \in V$. Consider these $u$ and $v$. Clearly, $u \in \operatorname{conncomp}_{G-e} u = P$ and $v \in \operatorname{conncomp}_{G-e} v = Q$.

Since $P$ and $Q$ are distinct, we have $P \neq Q$, so that $\operatorname{conncomp}_{G-e} u = P \neq Q = \operatorname{conncomp}_{G-e} v$. In other words, we do not have $u \simeq_{G-e} v$. But the connected components $P$ and $Q$ get merged in $G$ (possibly together with other connected components). The resulting connected component of $G$ contains both $P$ and $Q$ as subsets, and therefore contains both $u$ and $v$ as elements (because $u \in P$ and $v \in Q$). Hence, $u$ and $v$ lie in the same connected component of $G$ (namely, in the connected component we have just mentioned). In other words, $u \simeq_G v$. Hence, Lemma 0.7 shows that there exist $x \in V$ and $y \in V$ such that $\phi(e) = \{x, y\}$ and $u \simeq_{G-e} x$

from $G-e$ to $G$ (namely, the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$). Hence, the number of connected components of $G$ equals the number of connected components of $G-e$ minus 1. In other words, $\operatorname{conn} G = \operatorname{conn}(G-e) - 1$. But recall that $e$ belongs to no cycle of $G$. Hence, $[e \text{ belongs to no cycle of } G] = 1$. Thus,

$$\underbrace{\operatorname{conn} G}_{=\operatorname{conn}(G-e)-1} + \underbrace{[e \text{ belongs to no cycle of } G]}_{=1} = (\operatorname{conn}(G-e) - 1) + 1 = \operatorname{conn}(G-e).$$

Hence, Lemma 0.8 is proven in Case 1.

- Let us now consider Case 2. In this case, the edge $e$ belongs to at least one cycle of $G$. Fix such a cycle, and write it in the form $(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$, with $v_k = v_0$. Thus, $e$ is one of the edges $e_1, e_2, \ldots, e_k$ (since $e$ belongs to this cycle). We WLOG assume that $e = e_1$ (since otherwise, we can achieve $e = e_1$ by rotating the cycle). The edges $e_1, e_2, \ldots, e_k$ are the edges of the cycle $(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$, and thus are distinct (since the edges of a cycle are always distinct). Thus, in particular, the edges $e_2, e_3, \ldots, e_k$ are all different from $e_1$. Since $e = e_1$, this rewrites as follows: The edges $e_2, e_3, \ldots, e_k$ are all different from $e$. Hence, $e_2, e_3, \ldots, e_k$ are edges of the multigraph $G-e$. Thus, $(v_1, e_2, v_2, e_3, v_3, \ldots, e_k, v_k)$ is a walk in $G-e$. This is clearly a walk from $v_1$ to $v_0$ (since $v_1 = v_1$ and $v_k = v_0$). Hence, there exists a walk from $v_1$ to $v_0$ in $G-e$. In other words, $v_1 \simeq_{G-e} v_0$. Thus, $v_0 \simeq_{G-e} v_1$ (since $\simeq_{G-e}$ is an equivalence relation).

  Every two vertices $u \in V$ and $v \in V$ satisfying $u \simeq_G v$ satisfy $u \simeq_{G-e} v$ [7]. Conversely, every two vertices $u \in V$ and $v \in V$ satisfying $u \simeq_{G-e} v$ satisfy

and $v \simeq_{G-e} y$. Consider these $x$ and $y$. We have $P = \operatorname{conncomp}_{G-e} u = \operatorname{conncomp}_{G-e} x$ (since $u \simeq_{G-e} x$) and $Q = \operatorname{conncomp}_{G-e} v = \operatorname{conncomp}_{G-e} y$ (since $v \simeq_{G-e} y$). Hence, $\operatorname{conncomp}_{G-e} x = P \neq Q = \operatorname{conncomp}_{G-e} y$. Therefore, $x \neq y$. But $\{x, y\} = \phi(e) = \{a, b\}$. Hence, $x \in \{x, y\} = \{a, b\}$ and $y \in \{x, y\} = \{a, b\}$. Thus, $x$ and $y$ are two elements of $\{a, b\}$. Since $x \neq y$, we can hence conclude that $x$ and $y$ are two **distinct** elements of $\{a, b\}$. Thus, we have either ($x = a$ and $y = b$) or ($x = b$ and $y = a$). But each of these two options quickly leads us to our desired conclusion (namely, to the conclusion that $P$ and $Q$ are the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ (in some order)):

– If ($x = a$ and $y = b$), then we have $P = \operatorname{conncomp}_{G-e} x = \operatorname{conncomp}_{G-e} a$ (since $x = a$) and $Q = \operatorname{conncomp}_{G-e} y = \operatorname{conncomp}_{G-e} b$ (since $y = b$), and therefore we conclude that $P$ and $Q$ are the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ (in some order).

– If ($x = b$ and $y = a$), then we have $P = \operatorname{conncomp}_{G-e} x = \operatorname{conncomp}_{G-e} b$ (since $x = b$) and $Q = \operatorname{conncomp}_{G-e} y = \operatorname{conncomp}_{G-e} a$ (since $y = a$), and therefore we conclude that $P$ and $Q$ are the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ (in some order).

Thus, we have shown that $P$ and $Q$ are the two connected components $\operatorname{conncomp}_{G-e} a$ and $\operatorname{conncomp}_{G-e} b$ (in some order). This is what we wanted to prove.

[7] *Proof.* Let $u \in V$ and $v \in V$ be two vertices satisfying $u \simeq_G v$. We must show that $u \simeq_{G-e} v$.

Indeed, assume the contrary. Thus, we do not have $u \simeq_{G-e} v$. Lemma 0.7 thus shows that there exist $x \in V$ and $y \in V$ such that $\phi(e) = \{x, y\}$ and $u \simeq_{G-e} x$ and $v \simeq_{G-e} y$. Consider these

$u \simeq_G v$ [8]. Combining the preceding two sentences, we conclude that for any two vertices $u \in V$ and $v \in V$, we have $u \simeq_{G-e} v$ if and only if we have $u \simeq_G v$. In other words, the equivalence relations $\simeq_{G-e}$ and $\simeq_G$ on the set $V$ are identical. Hence, the connected components of $G - e$ (being the equivalence classes of the relation $\simeq_{G-e}$) are precisely the connected components of $G$ (which are the equivalence classes of the relation $\simeq_G$). Therefore, the number of the connected components of $G - e$ equals the number of the connected components of $G$. In other words, $\operatorname{conn} G = \operatorname{conn}(G - e)$. But recall that $e$ belongs to at least one cycle of $G$. Hence, $[e \text{ belongs to no cycle of } G] = 0$. Thus,

$$\underbrace{\operatorname{conn} G}_{=\operatorname{conn}(G-e)} + \underbrace{[e \text{ belongs to no cycle of } G]}_{=0} = \operatorname{conn}(G-e) + 0 = \operatorname{conn}(G-e).$$

Hence, Lemma 0.8 is proven in Case 2.

We thus have proven Lemma 0.8 in each of the two Cases 1 and 2. This completes its proof. □

A further lemma that I shall use has nothing to do with graphs; it is a simple (but important) property of sums of numbers:

**Lemma 0.9** (telescope principle for sums)**.** Let $k \in \mathbb{N}$. Let $r_0, r_1, \ldots, r_k$ be $k+1$ integers. Then,

$$\sum_{i=1}^{k} (r_i - r_{i-1}) = r_k - r_0.$$

Of course, the $r_0, r_1, \ldots, r_k$ in Lemma 0.9 can just as well be rational numbers or real numbers or complex numbers or elements of any abelian group (if you know what this means).

---

$x$ and $y$.

From $v \simeq_{G-e} y$, we obtain $y \simeq_{G-e} v$ (since $\simeq_{G-e}$ is an equivalence relation). If we had $x \simeq_{G-e} y$, then (using the fact that $\simeq_{G-e}$ is an equivalence relation) we would obtain $u \simeq_{G-e} x \simeq_{G-e} y \simeq_{G-e} v$, which would contradict the fact that we do not have $u \simeq_{G-e} v$. Hence, we cannot have $x \simeq_{G-e} y$. Consequently, we cannot have $x = y$. Thus, $x$ and $y$ are distinct.

We have $\phi(e_1) = \{v_0, v_1\}$ (since $(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$ is a cycle in $G$). Hence, $\{x, y\} = \phi(e) = \{v_0, v_1\}$. Since $x$ and $y$ are distinct, this yields that we must have either $(x = v_0 \text{ and } y = v_1)$ or $(x = v_1 \text{ and } y = v_0)$. But the first of these two options cannot happen (because if we had $(x = v_0 \text{ and } y = v_1)$, then the we would have $x = v_0 \simeq_{G-e} v_1 = y$, which would contradict the fact that we cannot have $x \simeq_{G-e} y$). Hence, the second of these two options must be the case. In other words, we have $(x = v_1 \text{ and } y = v_0)$. Thus, $x = v_1 \simeq_{G-e} v_0 = y$. This contradicts the fact that we cannot have $x \simeq_{G-e} y$. This contradiction proves that our assumption was wrong. Hence, $u \simeq_{G-e} v$ is proven.

[8] This was proven above.

*Proof of Lemma 0.9.* If $k=0$, then Lemma 0.9 holds (because both sides of the equality in question equal 0 in this case). Hence, for the rest of this proof, we WLOG assume that we don't have $k=0$. Thus, $k \geq 1$.

Now,

$$\sum_{i=1}^{k}\left(r_i - r_{i-1}\right) = \sum_{i=1}^{k} r_i - \sum_{i=1}^{k} r_{i-1} = \underbrace{\sum_{i=1}^{k} r_i}_{\substack{=\sum_{i=1}^{k-1} r_i + r_k \\ \text{(since } k\geq 1\text{)}}} - \underbrace{\sum_{i=0}^{k-1} r_i}_{\substack{=r_0 + \sum_{i=1}^{k-1} r_i \\ \text{(since } k\geq 1\text{)}}}$$

(here, we have substituted $i$ for $i-1$ in the second sum)

$$= \left(\sum_{i=1}^{k-1} r_i + r_k\right) - \left(r_0 + \sum_{i=1}^{k-1} r_i\right) = r_k - r_0.$$

This proves Lemma 0.9. □

*First solution to Exercise 3 (sketched).* **(a)** Whenever $K$ is a subset of $H$, we shall use the notation $c(K)$ for the nonnegative integer $\operatorname{conn}(V, K, \phi \mid_K)$. Using this notation, we can rewrite the inequality (1) (which we must prove) as follows:

$$c(E) + c(F) \leq c(E \cup F) + c(E \cap F). \tag{2}$$

We shall use the Iverson bracket notation. We observe that each subset $K$ of $H$ and each $f \in K$ satisfy

$$c(K \setminus \{f\}) - c(K) = [f \text{ belongs to no cycle of } (V, K, \phi \mid_K)] \tag{3}$$

[9].

[9] *Proof of* (3). Let $K$ be a subset of $H$. Let $f \in K$. Thus, $f$ is an edge of the multigraph $(V, K, \phi \mid_K)$. Hence, Lemma 0.8 (applied to $(V, K, \phi \mid_K)$ and $f$ instead of $G$ and $e$) yields

$$\operatorname{conn}((V, K, \phi \mid_K) - f) = \operatorname{conn}(V, K, \phi \mid_K) + [f \text{ belongs to no cycle of } (V, K, \phi \mid_K)].$$

Since $(V, K, \phi \mid_K) - f = \left(V, K \setminus \{f\}, \phi \mid_{K\setminus\{f\}}\right)$, this rewrites as

$$\operatorname{conn}\left(V, K \setminus \{f\}, \phi \mid_{K\setminus\{f\}}\right) = \operatorname{conn}(V, K, \phi \mid_K) + [f \text{ belongs to no cycle of } (V, K, \phi \mid_K)]. \tag{4}$$

Now, the definition of $c(K)$ yields $c(K) = \operatorname{conn}(V, K, \phi \mid_K)$, whereas the definition of $c(K \setminus \{f\})$ yields $c(K \setminus \{f\}) = \operatorname{conn}\left(V, K \setminus \{f\}, \phi \mid_{K\setminus\{f\}}\right)$. Subtracting the former equality from the latter, we obtain

$$c(K \setminus \{f\}) - c(K) = \operatorname{conn}\left(V, K \setminus \{f\}, \phi \mid_{K\setminus\{f\}}\right) - \operatorname{conn}(V, K, \phi \mid_K)$$
$$= [f \text{ belongs to no cycle of } (V, K, \phi \mid_K)]$$

(by (4)). This proves (3).

If $K$ and $L$ are two subsets of $H$ satisfying $K \subseteq L$, and if $f$ is an element of $K$, then

$$c(K \setminus \{f\}) - c(K) \geq c(L \setminus \{f\}) - c(L) \tag{5}$$

[10].

Let $(f_1, f_2, \ldots, f_k)$ be a list of all elements of $F \setminus E$ (with no element occurring twice). Thus, the elements $f_1, f_2, \ldots, f_k$ are distinct, and satisfy $F \setminus E = \{f_1, f_2, \ldots, f_k\}$.

Let $F' = E \cup F$. Thus, $F' \subseteq H$ (since $E$ and $F$ are subsets of $H$) and $F \subseteq E \cup F = F'$.

We have

$$\begin{aligned} & c(F \setminus \{f_1, f_2, \ldots, f_i\}) - c(F \setminus \{f_1, f_2, \ldots, f_{i-1}\}) \\ & \geq c(F' \setminus \{f_1, f_2, \ldots, f_i\}) - c(F' \setminus \{f_1, f_2, \ldots, f_{i-1}\}) \end{aligned} \tag{7}$$

for each $i \in \{1, 2, \ldots, k\}$ [11].

---

[10] *Proof of* (5). Let $K$ and $L$ be two subsets of $H$ satisfying $K \subseteq L$, and let $f$ be an element of $K$. We need to prove the inequality (5).

We have $f \in K$, so that $f \in K \subseteq L$. Hence, from (3) (applied to $L$ instead of $K$), we obtain

$$c(L \setminus \{f\}) - c(L) = [f \text{ belongs to no cycle of } (V, L, \phi|_L)] \leq 1 \tag{6}$$

(since the truth value of any statement is $\leq 1$). Now, if $f$ belongs to no cycle of $(V, K, \phi|_K)$, then

$$\begin{aligned} c(K \setminus \{f\}) - c(K) &= [f \text{ belongs to no cycle of } (V, K, \phi|_K)] \qquad (\text{by (3)}) \\ &= 1 \qquad (\text{since } f \text{ belongs to no cycle of } (V, K, \phi|_K)) \\ &\geq c(L \setminus \{f\}) - c(L) \qquad (\text{by (6)}). \end{aligned}$$

Thus, if $f$ belongs to no cycle of $(V, K, \phi|_K)$, then (5) is proven. Hence, for the rest of this proof, we WLOG assume that $f$ belongs to at least one cycle of $(V, K, \phi|_K)$. In other words, there exists a cycle $c$ of $(V, K, \phi|_K)$ such that $f$ belongs to $c$. Consider this $c$. But $(V, K, \phi|_K)$ is a sub-multigraph of $(V, L, \phi|_L)$ (since $K \subseteq L$). Hence, $c$ is a cycle of $(V, L, \phi|_L)$ (because $c$ is a cycle of $(V, K, \phi|_K)$). Therefore, $f$ belongs to at least one cycle of $(V, L, \phi|_L)$ (namely, to $c$). Now,

$$\begin{aligned} c(K \setminus \{f\}) - c(K) &= [f \text{ belongs to no cycle of } (V, K, \phi|_K)] \qquad (\text{by (3)}) \\ &= 0 \qquad (\text{since } f \text{ belongs to at least one cycle of } (V, K, \phi|_K)). \end{aligned}$$

Comparing this with

$$\begin{aligned} c(L \setminus \{f\}) - c(L) &= [f \text{ belongs to no cycle of } (V, L, \phi|_L)] \\ &= 0 \qquad (\text{since } f \text{ belongs to at least one cycle of } (V, L, \phi|_L)), \end{aligned}$$

we obtain $c(K \setminus \{f\}) - c(K) = c(L \setminus \{f\}) - c(L)$. Hence, $c(K \setminus \{f\}) - c(K) \geq c(L \setminus \{f\}) - c(L)$. This proves (5).

[11] *Proof of* (7). Fix $i \in \{1, 2, \ldots, k\}$. We have $\underbrace{F}_{\subseteq F'} \setminus \{f_1, f_2, \ldots, f_{i-1}\} \subseteq F' \setminus \{f_1, f_2, \ldots, f_{i-1}\}$. Furthermore, $f_i \in \{f_1, f_2, \ldots, f_k\} = F \setminus E \subseteq F$. Combining this with $f_i \notin \{f_1, f_2, \ldots, f_{i-1}\}$ (since $f_1, f_2, \ldots, f_k$ are distinct), we obtain $f_i \in F \setminus \{f_1, f_2, \ldots, f_{i-1}\}$. Hence, (5) (applied to

But

$$\begin{aligned}
&\sum_{i=1}^{k}\left(c\left(F\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\right)\\
&=c\left(F\setminus\underbrace{\left\{f_1,f_2,\ldots,f_k\right\}}_{=F\setminus E}\right)-c\left(F\setminus\underbrace{\left\{f_1,f_2,\ldots,f_0\right\}}_{=\varnothing}\right)\\
&\qquad\left(\text{by Lemma 0.9, applied to } r_i=c\left(F\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)\right)\\
&=c\left(\underbrace{F\setminus\left(F\setminus E\right)}_{=E\cap F}\right)-c\left(\underbrace{F\setminus\varnothing}_{=F}\right)=c\left(E\cap F\right)-c\left(F\right).
\end{aligned} \tag{8}$$

The same argument (but with $F$ replaced by $F'$) shows that

$$\begin{aligned}
&\sum_{i=1}^{k}\left(c\left(F'\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\right)\\
&=c\left(E\cap F'\right)-c\left(F'\right).
\end{aligned} \tag{9}$$

But $E\subseteq E\cup F=F'$, so that $E\cap F'=E$.
Now, (8) shows that

$$\begin{aligned}
&c\left(E\cap F\right)-c\left(F\right)\\
&=\sum_{i=1}^{k}\underbrace{\left(c\left(F\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\right)}_{\substack{\geq c\left(F'\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\\ \text{(by (7))}}}\\
&\geq\sum_{i=1}^{k}\left(c\left(F'\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\right)\\
&=c\left(\underbrace{E\cap F'}_{=E}\right)-c\left(\underbrace{F'}_{=E\cup F}\right)\qquad\text{(by (9))}\\
&=c\left(E\right)-c\left(E\cup F\right).
\end{aligned}$$

---

$K=F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}$, $L=F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}$ and $f=f_i$) yields

$$\begin{aligned}
&c\left(\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\setminus\left\{f_i\right\}\right)-c\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\\
&\geq c\left(\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\setminus\left\{f_i\right\}\right)-c\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right).
\end{aligned}$$

Since $\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\setminus\left\{f_i\right\}=F\setminus\left\{f_1,f_2,\ldots,f_i\right\}$ and $\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\setminus\left\{f_i\right\}=F'\setminus\left\{f_1,f_2,\ldots,f_i\right\}$, this rewrites as

$$c\left(F\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right)\geq c\left(F'\setminus\left\{f_1,f_2,\ldots,f_i\right\}\right)-c\left(F'\setminus\left\{f_1,f_2,\ldots,f_{i-1}\right\}\right).$$

This proves (7).

In other words, $c(E) + c(F) \leq c(E \cup F) + c(E \cap F)$. This proves (2). As we know, this yields (1) (since (2) is just a rewritten version of (1)). Hence, Exercise 3 **(a)** is solved.

**(b)** There are many possible examples.

For example, let $(V, H, \phi) = (\{1, 2, 3\}, \{12, 13, 23\}, \mathrm{id})$ (this is just the complete graph $K_3$ on the three vertices $1, 2, 3$, regarded as a multigraph), and set $E = \{12, 23\}$ and $F = \{23, 13\}$. In this case, $\operatorname{conn}(V, E, \phi \mid_E) = 1$, $\operatorname{conn}(V, F, \phi \mid_F) = 1$, $\operatorname{conn}(V, E \cup F, \phi \mid_{E \cup F}) = 1$ and $\operatorname{conn}(V, E \cap F, \phi \mid_{E \cap F}) = 2$, and thus the inequality (1) becomes $1 + 1 \leq 1 + 2$. $\square$

### 0.4.4. Second solution

Now, we shall outline a second solution of Exercise 3, following the hint.

**Definition 0.10.** Let $G = (V, E, \phi)$ be a multigraph.

**(a)** A subset $X$ of $E$ is said to be *connective* if the connected components of $(V, X, \phi \mid_X)$ are precisely the connected components of $G$. (Equivalently: A subset $X$ of $E$ is connective if and only if every two vertices that are connected by a walk in $G$ are also connected by a walk that only uses edges from $X$.)

**(b)** A subset $X$ of $E$ is said to be *independent* if the multigraph $(V, X, \phi \mid_X)$ has no cycles (i.e., if no cycle of $G$ has all its edges belong to $X$).

**(c)** A *basis* of $G$ shall mean a subset $X$ of $E$ that is both connective and independent.

**(d)** When $X$ is a basis of a multigraph $G = (V, E, \phi)$, the sub-multigraph $(V, X, \phi \mid_X)$ is called a *spanning forest* of $G$.

There is an analogy between multigraphs and vector spaces. Under this analogy, a connective subset corresponds to a spanning subset; an independent subset corresponds to a linearly independent subset; a basis corresponds to a basis.

**Proposition 0.11.** Let $G$ be a multigraph. Then, a basis of $G$ exists.

*Proof of Proposition 0.11.* For each connected component $C$ of $G$, consider a spanning tree of the induced sub-multigraph $G[C]$ of $G$ [12]. Let $X_C$ be the set of all edges of this spanning tree.

Let $X$ be the union of these sets $X_C$ over all connected components $C$ of $G$. Then, $X$ is connective. (Indeed, any two vertices lying in one and the same connected component of $G$ must also be connected by the spanning tree of this component.) Also, $X$ is independent. (Indeed, the sub-multigraph $(V, X, \phi \mid_X)$ has no cycles, because each spanning tree separately has no cycles, and because we cannot "jump" from one spanning tree to the other since they are in different connected components.) Hence, $X$ is a basis of $G$. $\square$

[12] Induced sub-multigraphs of $G$ are defined as follows: Write $G$ in the form $G = (V, E, \phi)$. If $S$ is a subset of $V$, then the induced sub-multigraph $G[S]$ is defined to be the sub-multigraph $(S, E_S, \phi \mid_{E_S})$ of $G$, where $E_S$ is the set of all edges $e \in E$ such that $\phi(e) \subseteq S$ (in other words, such that both endpoints of $e$ lie in $S$).

**Proposition 0.12.** Let $G$ be a forest. Then, $|\operatorname{E}(G)| = |\operatorname{V}(G)| - \operatorname{conn} G$.

Proposition 0.12 is precisely Corollary 20 from lecture 9 (with the only difference being that we denoted $\operatorname{conn} G$ by $b_0(G)$ in lecture 9), so we omit its proof here.

**Proposition 0.13.** Let $G = (V, E, \phi)$ be a multigraph. Let $X$ be a basis of $G$. Then, $|X| = |V| - \operatorname{conn} G$.

*Proof of Proposition 0.13 (sketched).* The set $X$ is a basis of $G$. In other words, $X$ is a subset of $E$ that is both connective and independent.

Since $X$ is connective, the connected components of $(V, X, \phi\mid_X)$ are precisely the connected components of $G$. Hence, the number of the former connected components equals the number of the latter connected components. In other words, $\operatorname{conn}(V, X, \phi\mid_X) = \operatorname{conn} G$.

But $X$ is independent. In other words, the multigraph $(V, X, \phi\mid_X)$ has no cycles. In other words, this multigraph is a forest. Hence, Proposition 0.12 (applied to $(V, X, \phi\mid_X)$ instead of $G$) shows that $|\operatorname{E}((V, X, \phi\mid_X))| = |\operatorname{V}((V, X, \phi\mid_X))| - \operatorname{conn}(V, X, \phi\mid_X)$. Since $\operatorname{E}((V, X, \phi\mid_X)) = X$, $\operatorname{V}((V, X, \phi\mid_X)) = V$ and $\operatorname{conn}(V, X, \phi\mid_X) = \operatorname{conn} G$, this rewrites as $|X| = |V| - \operatorname{conn} G$. □

Compare Proposition 0.13 to the well-known fact from linear algebra that any two bases of a vector space have the same size (the dimension of this vector space).

**Proposition 0.14.** Let $G = (V, E, \phi)$ be a multigraph. Let $X$ be a subset of $E$. Assume that for each edge $e \in E \setminus X$, there exists a path that uses only edges from $X$, and that connects the two endpoints of $e$ (that is, the starting point and the ending point of this path are the two endpoints of $e$). Then, the subset $X$ of $E$ is connective.

*Proof of Proposition 0.14.* For each edge $e \in E \setminus X$, we fix some path that uses only edges from $X$, and that connects the two endpoints of $e$ [13]. We shall refer to this path as the *$X$-detour for $e$*.

Now, let $u$ and $v$ be two vertices that are connected by a walk in $G$. Fix such a walk. Some of the edges of this walk may belong to $E \setminus X$. But if we replace all these edges by their $X$-detours, we obtain a (possibly longer) walk that uses only edges from $X$. Thus, $u$ and $v$ are connected by a walk that uses only edges from $X$.

We thus have proven that every two vertices that are connected by a walk in $G$ are also connected by a walk that only uses edges from $X$. In other words, the subset $X$ of $E$ is connective. This proves Proposition 0.14. □

**Lemma 0.15.** Let $G = (V, E, \phi)$ be a multigraph. Let $X$ be an independent subset of $E$. Assume that $X$ is not connective. Then, there exists an edge $e \in E \setminus X$ such that the subset $X \cup \{e\}$ is independent.

[13] The existence of such a path is guaranteed by the hypothesis of Proposition 0.14.

*Proof of Lemma 0.15.* Assume the contrary. Thus, for each edge $e \in E \setminus X$, the subset $X \cup \{e\}$ is not independent.

Now, for each edge $e \in E \setminus X$, there exists a path that uses only edges from $X$, and that connects the two endpoints of $e$ (that is, the starting point and the ending point of this path are the two endpoints of $e$) [14]. Hence, Proposition 0.14 shows that the subset $X$ of $E$ is connective. This contradicts the hypothesis that $X$ is not connective. This contradiction completes the proof. □

**Proposition 0.16.** Let $G = (V, E, \phi)$ be a multigraph. Let $Y$ be an independent subset of $E$. Then, there exists a basis of $G$ that contains $Y$ as a subset.

*Proof of Proposition 0.16 (sketched).* We construct a sequence $(Y_0, Y_1, \ldots, Y_k)$ of independent subsets of $E$ by the following algorithm:

- Set $Y_0 = Y$ and $i = 0$.
- **While** there exists an edge $e \in E \setminus Y_i$ such that the set $Y_i \cup \{e\}$ is still independent, **do** the following:
  - Pick such an $e$, and set $Y_{i+1} = Y_i \cup \{e\}$. Then, set $i = i + 1$.
- Set $k = i$.

This algorithm must terminate. (Indeed, each subset $Y_{i+1}$ constructed during the algorithm has a larger size than the previous subset $Y_i$, but a subset of $E$ cannot have size larger than $|E|$, so we cannot build an infinite sequence $(Y_0, Y_1, Y_2, \ldots)$ of subsets of $E$ where each subset has larger size than the previous one.) Thus, the subset $Y_k$ built at the end of the algorithm has the following property: It is an independent subset of $E$, but there exists no edge $e \in E \setminus Y_k$ such that the set $Y_k \cup \{e\}$ is still independent.

The algorithm guarantees that $Y_i \subseteq Y_{i+1}$ for each $i$ for which $Y_{i+1}$ has been constructed. Thus, $Y_0 \subseteq Y_1 \subseteq \cdots \subseteq Y_k$, so that $Y_0 \subseteq Y_k$ and thus $Y = Y_0 \subseteq Y_k$. Consequently, $Y_k$ contains $Y$ as a subset.

It the subset $Y_k$ of $E$ was not connective, then Lemma 0.15 (applied to $X = Y_k$) would show that there exists an edge $e \in E \setminus Y_k$ such that the subset $Y_k \cup \{e\}$ is independent. This would contradict the fact that there exists no edge $e \in E \setminus Y_k$ such that the set $Y_k \cup \{e\}$ is still independent. Hence, the subset $Y_k$ is connective. Thus, $Y_k$ is both connective and independent. In other words, $Y_k$ is a basis of $G$. Hence, there exists a basis of $G$ that contains $Y$ as a subset (namely, $Y_k$). This proves Proposition 0.16. □

[14] *Proof.* Let $e \in E \setminus X$ be an edge. We know (from the previous paragraph) that the subset $X \cup \{e\}$ is not independent. In other words, there exists a cycle of $G$ all of whose edges belong to $X \cup \{e\}$. Fix such a cycle. Clearly, not all of the edges of this cycle belong to $X$ (because the subset $X$ is independent, and thus there exists no cycle of $G$ all of whose edges belong to $X$). Hence, at least one edge of this cycle must be $e$. Therefore, **exactly one** edge of this cycle is $e$ (since the edges of a cycle are always distinct). If we remove the edge $e$ from this cycle, we thus obtain a path that uses only edges from $X$, and that connects the two endpoints of $e$. Thus, there exists a path that uses only edges from $X$, and that connects the two endpoints of $e$.

Compare Proposition 0.16 to the well-known fact that a linearly independent subset of a vector space can always be extended to a basis.

Notice that we can also use Proposition 0.16 to prove Proposition 0.11 again (namely, by applying Proposition 0.16 to $Y = \varnothing$, exploiting the obvious fact that $\varnothing$ is independent).

**Proposition 0.17.** Let $G = (V, E, \phi)$ be a multigraph. Then, $\operatorname{conn} G \geq |V| - |E|$.

Proposition 0.17 is precisely Proposition 14 from lecture 9 (with the only difference being that we denoted $\operatorname{conn} G$ by $b_0(G)$ in lecture 9), so we omit its proof here.

**Proposition 0.18.** Let $G = (V, E, \phi)$ be a multigraph. Let $X$ be a connective subset of $E$. Then, $|X| \geq |V| - \operatorname{conn} G$.

*Proof of Proposition 0.18 (sketched).* Since $X$ is connective, the connected components of $(V, X, \phi\mid_X)$ are precisely the connected components of $G$. Hence, the number of the former connected components equals the number of the latter connected components. In other words, $\operatorname{conn}(V, X, \phi\mid_X) = \operatorname{conn} G$.

But Proposition 0.17 (applied to $(V, X, \phi\mid_X)$, $X$ and $\phi\mid_X$ instead of $G$, $E$ and $\phi$) shows that $\operatorname{conn}(V, X, \phi\mid_X) \geq |V| - |X|$. Thus, $|X| \geq |V| - \operatorname{conn}(V, X, \phi\mid_X) = \operatorname{conn} G$. This proves Proposition 0.18. □

*Second solution to Exercise 3 (sketched).* **(a)** Proposition 0.11 (applied to $(V, E \cap F, \phi\mid_{E\cap F})$ instead of $G$) shows that a basis of the multigraph $(V, E \cap F, \phi\mid_{E\cap F})$ exists. Fix such a basis, and denote it by $Y$. Thus, $Y$ is an independent subset of $E \cap F$.

Proposition 0.16 (applied to $(V, E, \phi\mid_E)$ and $\phi\mid_E$ instead of $G$ and $\phi$) shows that there exists a basis of $(V, E, \phi\mid_E)$ that contains $Y$ as a subset. Fix such a basis, and denote it by $P$. Thus, $P$ is a subset of $E$ that is both independent and connective (with respect to the multigraph $(V, E, \phi\mid_E)$).

Proposition 0.16 (applied to $(V, F, \phi\mid_F)$, $F$ and $\phi\mid_F$ instead of $G$, $E$ and $\phi$) shows that there exists a basis of $(V, F, \phi\mid_F)$ that contains $Y$ as a subset. Fix such a basis, and denote it by $Q$. Thus, $Q$ is a subset of $F$ that is both independent and connective (with respect to the multigraph $(V, F, \phi\mid_F)$).

From $Y \subseteq P$ and $Y \subseteq Q$, we obtain $Y \subseteq P \cap Q$.

It is not necessarily true that $P \cup Q$ is a basis of the multigraph $(V, E \cup F, \phi\mid_{E\cup F})$. However, it is not hard to see that $P \cup Q$ is a connective subset of $E \cup F$ (with respect to this multigraph). Indeed, for each edge $e \in (E \cup F) \setminus (P \cup Q)$, there exists a path that uses only edges from $P \cup Q$, and that connects the two endpoints of $e$ [15]. Thus, Proposition 0.14 (applied to $(V, E \cup F, \phi\mid_{E\cup F})$, $E \cup F$, $\phi\mid_{E\cup F}$ and

[15] *Proof.* Let $e \in (E \cup F) \setminus (P \cup Q)$ be an edge. Then, $e \in (E \cup F) \setminus (P \cup Q) \subseteq E \cup F$. Hence, either $e \in E$ or $e \in F$ (or both). We WLOG assume that $e \in E$ (since otherwise, an analogous argument works). Then, the two endpoints of $e$ lie in the same connected component of the multigraph $(V, E, \phi\mid_E)$ (because they are adjacent in this multigraph). Hence, these two endpoints are connected by a path using only edges from $P$ (since $P$ is a connective subset of $E$ with respect

$P \cup Q$ instead of $G$, $E$, $\phi$ and $X$) shows that the subset $P \cup Q$ of $E \cup F$ is connective with respect to the multigraph $(V, E \cup F, \phi\mid_{E\cup F})$. Hence, Proposition 0.18 (applied to $(V, E \cup F, \phi\mid_{E\cup F})$, $E \cup F$, $\phi\mid_{E\cup F}$ and $P \cup Q$ instead of $G$, $E$, $\phi$ and $X$) shows that $|P \cup Q| \geq |V| - \operatorname{conn}(V, E \cup F, \phi\mid_{E\cup F})$. Hence,

$$\operatorname{conn}(V, E \cup F, \phi\mid_{E\cup F}) \geq |V| - |P \cup Q|. \tag{10}$$

But $Y$ is a basis of $(V, E \cap F, \phi\mid_{E\cap F})$. Hence, Proposition 0.13 (applied to $(V, E \cap F, \phi\mid_{E\cap F})$, $E \cap F$, $\phi\mid_{E\cap F}$ and $Y$ instead of $G$, $E$, $\phi$ and $X$) shows that $|Y| = |V| - \operatorname{conn}(V, E \cap F, \phi\mid_{E\cap F})$. Hence,

$$\operatorname{conn}(V, E \cap F, \phi\mid_{E\cap F}) = |V| - \underbrace{|Y|}_{\substack{\leq |P\cap Q| \\ (\text{since } Y \subseteq P \cap Q)}} \geq |V| - |P \cap Q|.$$

Adding this inequality to (10), we obtain

$$\begin{aligned}\operatorname{conn}(V, E \cup F, \phi\mid_{E\cup F}) + \operatorname{conn}(V, E \cap F, \phi\mid_{E\cap F}) &\geq (|V| - |P \cup Q|) + (|V| - |P \cap Q|) \\ &= 2|V| - \underbrace{(|P \cup Q| + |P \cap Q|)}_{=|P|+|Q|} \\ &= 2|V| - (|P| + |Q|).\end{aligned}$$

On the other hand, $P$ is a basis of $(V, E, \phi\mid_E)$. Hence, Proposition 0.13 (applied to $(V, E, \phi\mid_E)$, $\phi\mid_E$ and $P$ instead of $G$, $\phi$ and $X$) shows that $|P| = |V| - \operatorname{conn}(V, E, \phi\mid_E)$. Hence,

$$\operatorname{conn}(V, E, \phi\mid_E) = |V| - |P|.$$

The same reasoning (but applied to $F$ and $Q$ instead of $E$ and $P$) shows that

$$\operatorname{conn}(V, F, \phi\mid_F) = |V| - |Q|.$$

Thus,

$$\begin{aligned}&\operatorname{conn}(V, E \cup F, \phi\mid_{E\cup F}) + \operatorname{conn}(V, E \cap F, \phi\mid_{E\cap F}) \\ &\geq 2|V| - (|P| + |Q|) = \underbrace{|V| - |P|}_{=\operatorname{conn}(V,E,\phi\mid_E)} + \underbrace{|V| - |Q|}_{=\operatorname{conn}(V,F,\phi\mid_F)} \\ &= \operatorname{conn}(V, E, \phi\mid_E) + \operatorname{conn}(V, F, \phi\mid_F).\end{aligned}$$

This solves Exercise 3 **(a)**.

**(b)** See the First solution above. □

---

to the the multigraph $(V, E, \phi\mid_E)$). This path thus uses only edges from $P \cup Q$ (since the edges from $P$ clearly are edges from $P \cup Q$), and connects the two endpoints of $e$. Hence, there exists a path that uses only edges from $P \cup Q$, and that connects the two endpoints of $e$.

### 0.4.5. Third solution

The third solution of Exercise 3 (which will be roughly outlined) uses linear algebra. Let us first introduce some notations.

**Definition 0.19.** Let $n \in \mathbb{N}$. Then, $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ will denote the standard basis of the $\mathbb{Q}$-vector space $\mathbb{Q}^n$. This means that $\mathbf{e}_i$ is the column vector whose $i$-th coordinate is 1, and whose all other coordinates are 0.

(Instead of $\mathbb{Q}$-vector spaces, we can just as well use $\mathbb{R}$-vector spaces or $\mathbb{C}$-vector spaces[16]. I have chosen $\mathbb{Q}$ merely because rational numbers feel more concrete to me.)

The crux of the third solution is the following neat result from linear algebra:

**Proposition 0.20.** Let $G = (V, E, \phi)$ be a multigraph, where $V = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. For each edge $e$ of $G$, define a vector $v_e \in \mathbb{Q}^n$ by picking $i \in V$ and $j \in V$ such that $\phi(e) = \{i, j\}$, and setting $v_e = \mathbf{e}_i - \mathbf{e}_j$. (We are free to choose which of the two endpoints of $e$ is to become $i$ and which is to become $j$ here.)

Then,

$$\operatorname{conn} G = |V| - \dim(\operatorname{span}(\{v_e \mid e \in E\})). \tag{11}$$

(More precisely, $\operatorname{span}(\{v_e \mid e \in E\})$ is the $\mathbb{Q}$-vector subspace of $\mathbb{Q}^n$ that consists of all vectors $\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{Q}^n$ satisfying

$$\sum_{i \in C} p_i = 0 \text{ for each connected component } C \text{ of } G.$$

Thus, it is the solution set of a system of $\operatorname{conn} G$ many linear equations.)

(Notice that the equality (11) appears in the literature in various guises. For example, [Quinla17, Theorem 1.3.5] is a restatement of (11) in terms of matrices.)

*Hints to a proof of Proposition 0.20.* Let $\mathcal{P}$ denote the $\mathbb{Q}$-vector subspace of $\mathbb{Q}^n$ that consists of all vectors $\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{Q}^n$ satisfying

$$\sum_{i \in C} p_i = 0 \text{ for each connected component } C \text{ of } G. \tag{12}$$

[16] Or vector spaces over any field – if you know what this means.

For each connected component $C$ of $G$, fix a vertex of $C$ (chosen arbitrarily), and call it the *root* of $C$. A vertex of $G$ is said to be a *root* if and only if it is the root of its connected component.

The *depth* of a vertex $v \in V$ shall be defined as the distance from $v$ to the root of the connected component of $v$. This depth is a nonnegative integer, and it equals 0 if and only if $v$ itself is a root.

Let $\mathcal{S}$ denote the $\mathbb{Q}$-vector subspace $\operatorname{span}(\{v_e \mid e \in E\})$ of $\mathbb{Q}^n$.

Let $\mathcal{R}$ be the $\mathbb{Q}$-vector subspace $\operatorname{span}(\{\mathbf{e}_v \mid v \in V \text{ is a root}\})$ of $\mathbb{Q}^n$. In other words, $\mathcal{R}$ is the set of all vectors $\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{Q}^n$ satisfying

$$p_i = 0 \text{ for each } i \in V \text{ that is not a root.} \tag{13}$$

Clearly,

$$\dim \mathcal{R} = (\text{the number of roots}) = \operatorname{conn} G.$$

Now, it is not hard to see that

$$\mathbf{e}_i \in \mathcal{S} + \mathcal{R} \qquad \text{for all } i \in V. \tag{14}$$

Indeed, this is easily proven by induction[17]. Since the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ form a basis of $\mathbb{Q}^n$, this shows that $\mathbb{Q}^n = \mathcal{S} + \mathcal{R}$.

Also, $\mathcal{S} \subseteq \mathcal{P}$ (to check this, show that all the generators $v_e$ of $\mathcal{S}$ lie in $\mathcal{P}$). Furthermore, $\mathcal{P} \cap \mathcal{R} = 0$ [18]. Now, from $\mathcal{S} \subseteq \mathcal{P}$, we obtain $\mathcal{S} \cap \mathcal{R} \subseteq \mathcal{P} \cap \mathcal{R} = 0$. Combined with $\mathbb{Q}^n = \mathcal{S} + \mathcal{R}$, this yields

$$\mathbb{Q}^n = \mathcal{S} \oplus \mathcal{R}.$$

[17] In (slightly) more detail:

We proceed by induction over the depth of $i$.

The *induction base* is the case when the depth of $i$ is 0. This case is easy (because if the depth of $i$ is 0, then $i$ is a root, whence $\mathbf{e}_i \in \mathcal{R} \subseteq \mathcal{S} + \mathcal{R}$).

The *induction step* requires us to prove (14) for all $i$ of depth $k+1$, assuming that (14) holds for all $i$ of depth $k$. This can be argued as follows: Fix an $i \in V$ of depth $k+1$. Then, there exists a neighbor $j \in V$ of $i$ that has depth $k$ (since the depth is the distance from the root). Fix such a neighbor, and let $e$ be an edge connecting $i$ to this neighbor. Then, $\mathbf{e}_j \in \mathcal{S} + \mathcal{R}$ by the induction hypothesis (since $j$ has depth $k$).

But the edge $e$ connects $i$ with $j$. Hence, either $v_e = \mathbf{e}_i - \mathbf{e}_j$ or $v_e = \mathbf{e}_j - \mathbf{e}_i$ (depending on the way we defined $v_e$). Thus, in either case, $\mathbf{e}_i - \mathbf{e}_j \in \{v_e, -v_e\} \subseteq \mathcal{S}$, so that $\mathbf{e}_i \in \mathcal{S} + \mathbf{e}_j \subseteq \mathcal{S} + \mathcal{R}$ (since $\mathbf{e}_j \in \mathcal{S} + \mathcal{R}$). This completes the induction step.

[18] *Proof.* Let $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$ be a vector in $\mathcal{P} \cap \mathcal{R}$. Then, this vector must satisfy both equations (12) and (13) (since it lies in both $\mathcal{P}$ and $\mathcal{R}$).

Now, let $j \in V$ be a vertex. We want to show that $p_j = 0$. Indeed, if $j$ is not a root, then this follows from (13) (applied to $i = j$). So let us WLOG assume that $j$ is a root. Let $C$ be the connected component containing $j$. Then, the only root in $C$ is $j$. Hence, all vertices $i \in C$ except

Taking dimensions, we find $\dim(\mathbb{Q}^n) = \dim \mathcal{S} + \dim \mathcal{R}$. Hence,

$$\dim \mathcal{S} = \underbrace{\dim(\mathbb{Q}^n)}_{=n} - \underbrace{\dim \mathcal{R}}_{=\operatorname{conn} G} = n - \operatorname{conn} G.$$

Hence, $\operatorname{conn} G = n - \dim \mathcal{S} = |V| - \dim \mathcal{S}$ (since $n = |V|$). This proves (11).

Let us now prove the "More precisely" statement in Proposition 0.20. Indeed, this statement simply claims that $\mathcal{S} = \mathcal{P}$. To prove it, we assume the contrary. Thus, $\mathcal{S}$ is a proper subset of $\mathcal{P}$ (because we know that $\mathcal{S} \subseteq \mathcal{P}$). Hence, $\dim \mathcal{S} < \dim \mathcal{P}$. But $\mathcal{P} \cap \mathcal{R} = 0$ yields that the sum $\mathcal{P} + \mathcal{R}$ is a direct sum. Hence,

$$\dim(\mathcal{P} + \mathcal{R}) = \underbrace{\dim \mathcal{P}}_{>\dim \mathcal{S}} + \dim \mathcal{R} > \dim \mathcal{S} + \dim \mathcal{R} = \dim(\mathbb{Q}^n).$$

This contradicts the fact that $\dim(\mathcal{P} + \mathcal{R}) \leq \dim(\mathbb{Q}^n)$ (which is a trivial consequence of the fact that $\mathcal{P} + \mathcal{R}$ is a subspace of $\mathbb{Q}^n$). This contradiction shows that our assumption was wrong, and so $\mathcal{S} = \mathcal{P}$ is proven. Finally, the proof of Proposition 0.20 is complete. □

This allows solving Exercise 3 as follows:

*Hints to a third solution of Exercise 3.* WLOG assume that $V = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$. For each edge $e \in H$, define a vector $v_e \in \mathbb{Q}^n$ by picking $i \in V$ and $j \in V$ such that $\phi(e) = \{i, j\}$, and setting $v_e = \mathbf{e}_i - \mathbf{e}_j$.

Define two $\mathbb{Q}$-vector subspaces $X$ and $Y$ of $\mathbb{Q}^n$ by

$$X = \operatorname{span}(\{v_e \mid e \in E\}) \qquad \text{and} \qquad Y = \operatorname{span}(\{v_e \mid e \in F\}).$$

Then, notice that

$$X + Y = \operatorname{span}(\{v_e \mid e \in E \cup F\})$$

and

$$X \cap Y \supseteq \operatorname{span}(\{v_e \mid e \in E \cap F\}) \tag{15}$$

(this is not an equality, just an inclusion). But a classical fact from linear algebra says that $\dim X + \dim Y = \dim(X + Y) + \dim(X \cap Y)$. Substitute the above expressions for $X$, $Y$, $X + Y$ and $X \cap Y$ into this equality (thus turning it into an inequality, since (15) is merely an inclusion). Finally, rewrite the dimensions of the spans using Proposition 0.20. The result is precisely the claim of Exercise 3. □

## 0.5. Exercise 4: forcing a Hamiltonian cycle on a tree

for $j$ are not roots. Thus, all vertices $i \in C$ except for $j$ satisfy $p_i = 0$ (by (13)). Hence, $\sum_{i \in C} p_i = p_j$. Therefore, the equality (12) simplifies to $p_j = 0$.

Now, forget that we fixed $j$. We thus have proven that $p_j = 0$ for each $j \in V$. In other words, all coordinates of our vector $\mathbf{p}$ are 0. In other words, $\mathbf{p} = 0$.

We thus have shown that any vector $\mathbf{p} \in \mathcal{P} \cap \mathcal{R}$ satisfies $\mathbf{p} = 0$. In other words, $\mathcal{P} \cap \mathcal{R} = 0$.

**Exercise 4.** Let $T$ be a tree having more than 1 vertex. Let $L$ be the set of leaves of $T$. Prove that it is possible to add $|L| - 1$ new edges to $T$ in such a way that the resulting multigraph has a Hamiltonian cycle.

Exercise 4 is taken from [Wang17, Lemma 3.1, second inequality sign]. Rather than solve this exercise directly, we shall prove a stronger fact. First, we define a notation:

**Definition 0.21.** Let $T$ be a tree. Then, $\operatorname{Leaves}(T)$ shall mean the set of all leaves of $T$.

The following fact is simple:

**Proposition 0.22.** Let $T$ be a tree such that $|\operatorname{V}(T)| > 2$. Let $v$ be a leaf of $T$. Let $T'$ denote the multigraph obtained from $T$ by removing this leaf $v$ and the unique edge that contains $v$. Let $u$ be the unique neighbor of $v$ in $T$.

**(a)** The multigraph $T'$ is a tree again.

**(b)** We have $u \notin \operatorname{Leaves}(T)$.

**(c)** If $u$ is a leaf of $T'$, then $\operatorname{Leaves}(T') = (\operatorname{Leaves}(T) \setminus \{v\}) \cup \{u\}$.

**(d)** If $u$ is not a leaf of $T'$, then $\operatorname{Leaves}(T') = \operatorname{Leaves}(T) \setminus \{v\}$.

Unfortunately, the length of a proof of Proposition 0.22 is far out of proportion to its simplicity. We recommend the reader to prove it themselves instead of reading the below argument.

*Proof of Proposition 0.22 (sketched).* The vertex $v$ is a leaf of $T$. In other words, $\deg_T v = 1$. Hence, there is a unique edge of $T$ that contains $v$. Let $e$ be this edge. The other endpoint of this edge $e$ (besides $v$) is $u$ (since $u$ is the unique neighbor of $v$ in $T$). Thus, the two endpoints of this edge $e$ are $u$ and $v$.

Recall that $e$ is the unique edge of $T$ that contains $v$. Hence, $e$ is the unique edge that gets removed from $T$ in the construction of $T'$. In other words, $e$ is the only edge of $T$ that is not an edge of $T'$.

Observe that $v$ is not a vertex of $T'$ (since the multigraph $T'$ is obtained from $T$ by removing the vertex $v$).

Also, the multigraph $T'$ is obtained from $T$ by removing the vertex $v$ (and an edge). Thus, the vertex set of $T'$ is obtained from the vertex set of $T$ by removing the vertex $v$. In other words, $\operatorname{V}(T') = \operatorname{V}(T) \setminus \{v\}$. Since $v \in \operatorname{V}(T)$, we thus obtain $|\operatorname{V}(T')| = |\operatorname{V}(T)| - 1 > 1$ (since $|\operatorname{V}(T)| > 2$).

Let us first show that $T'$ is a tree. This is an easy fact (and was done in class), but we briefly recall the argument for the sake of completeness: The multigraph $T'$ is connected[19]

[19] *Proof.* Let $p$ and $q$ be two vertices of $T'$. Then, $p$ and $q$ are distinct from $v$ (since $v$ is not a vertex of $T'$). Since $T$ is connected (because $T$ is a tree), there exists a walk from $p$ to $q$ in $T$. Hence, there exists a path from $p$ to $q$ in $T$. The starting point and the ending point of this path are both distinct from $v$ (since these two vertices are $p$ and $q$, and we know that $p$ and $q$ are distinct from $v$). All intermediate vertices of this path (i.e., all its vertices other than the starting point and the ending point) must be distinct from $v$ as well (since $v$ is a leaf, but an intermediate vertex of a

and is a forest[20]. Thus, $T'$ is a tree. This proves Proposition 0.22 **(a)**.

Each vertex $q$ of $T'$ satisfies

$$\deg_{T'} q \geq 1 \qquad (16)$$

[21].

The vertex $u$ is not a leaf of $T$ [22]. In other words, $u \notin \operatorname{Leaves}(T)$. This proves Proposition 0.22 **(b)**.

Next, let us notice that each $q \in \operatorname{Leaves}(T) \setminus \{v\}$ satisfies $q \in \operatorname{Leaves}(T')$ [23]. In other words,

$$\operatorname{Leaves}(T) \setminus \{v\} \subseteq \operatorname{Leaves}(T'). \qquad (17)$$

On the other hand, each $q \in \operatorname{Leaves}(T') \setminus \{u\}$ satisfies $q \in \operatorname{Leaves}(T) \setminus \{v\}$ [24]. In

---

path can never be a leaf). Thus, **all** vertices of this path are distinct from $v$. Hence, all vertices of this path are vertices of $T'$. Thus, this path is a path in $T'$. Consequently, there is a path from $p$ to $q$ in $T'$ (namely, the path that we have just constructed).

Thus, we have shown that if $p$ and $q$ are two vertices of $T'$, then there is a path from $p$ to $q$ in $T'$. Hence, $T'$ is connected (since $|\operatorname{V}(T')| > 1 > 0$).

[20] *Proof.* The multigraph $T$ has no cycles (since it is a tree). Thus, the multigraph $T'$ has no cycles either (since any cycle of $T'$ would be a cycle of $T$). In other words, $T'$ is a forest.

[21] *Proof of* (16). Let $q$ be a vertex of $T'$. Recall that $|\operatorname{V}(T')| > 1$, so that $|\operatorname{V}(T')| \geq 2$. Hence, there exist at least two vertices of $T'$. Thus, there exists at least one vertex of $T'$ distinct from $q$. Fix such a vertex, and denote it by $p$.

The multigraph $T'$ is connected. Thus, there exists a walk from $q$ to $p$ in $T'$. This walk has length $> 0$ (since $p$ is distinct from $q$), and thus has at least one edge. Hence, there exists an edge containing $q$ in $T'$ (namely, the very first edge of this walk). In other words, $\deg_{T'} q \geq 1$. This proves (16).

[22] *Proof.* Assume the contrary. Thus, $u$ is a leaf of $T$. In other words, $\deg_T u = 1$. In other words, there exists a unique edge of $T$ containing $u$. This unique edge must be $e$ (since the edge $e$ contains $u$ (because the two endpoints of this edge $e$ are $u$ and $v$)). Thus, $e$ is the only edge of $T$ containing $u$.

But $u \neq v$ (since $u$ is a neighbor of $v$ in $T$). Hence, $u \in \operatorname{V}(T) \setminus \{v\} = \operatorname{V}(T')$. In other words, $u$ is a vertex of $T'$. Therefore, (16) (applied to $q = u$) shows that $\deg_{T'} u \geq 1$. In other words, the number of edges of $T'$ containing $u$ is $\geq 1$. Hence, there exists at least one edge of $T'$ containing $u$. Fix such an edge, and denote it by $f$.

Now, $f$ is an edge of $T'$ containing $u$. But $T'$ is a sub-multigraph of $T$. Thus, $f$ is an edge of $T$ (since $f$ is an edge of $T'$). Hence, $f$ is an edge of $T$ containing $u$. Since $e$ is the only edge of $T$ containing $u$, we thus conclude that $f = e$. But $e$ is not an edge of $T'$ (since $e$ is the unique edge that gets removed from $T$ in the construction of $T'$). In other words, $f$ is not an edge of $T'$ (since $f = e$). This contradicts the fact that $f$ is an edge of $T'$. This contradiction proves that our assumption was false; qed.

[23] *Proof.* Let $q \in \operatorname{Leaves}(T) \setminus \{v\}$. Thus, $q \in \operatorname{Leaves}(T)$ and $q \neq v$. Since $q \neq v$, we conclude that $q$ is a vertex of $T'$.

We know that $q \in \operatorname{Leaves}(T)$. In other words, $q$ is a leaf of $T$. In other words, $\deg_T q = 1$.

But $T'$ is a submultigraph of $T$. Hence, $\deg_{T'} q \leq \deg_T q = 1$. Combining this with (16), we obtain $\deg_{T'} q = 1$. In other words, $q$ is a leaf of $T'$. In other words, $q \in \operatorname{Leaves}(T')$.

[24] *Proof.* Let $q \in \operatorname{Leaves}(T') \setminus \{u\}$. Thus, $q \in \operatorname{Leaves}(T')$ and $q \neq u$.

We know that $q \in \operatorname{Leaves}(T')$. In other words, $q$ is a leaf of $T'$. In other words, $\deg_{T'} q = 1$. Hence, there is only one edge of $T'$ that contains $q$.

Recall that $T'$ is a submultigraph of $T$. Thus, $\deg_T q \geq \deg_{T'} q = 1$.

Since $q$ is a vertex of $T'$, we have $q \neq v$ (since the multigraph $T'$ is obtained from $T$ by removing the vertex $v$).

Assume (for the sake of contradiction) that $\deg_T q \neq 1$. Thus, $\deg_T q > 1$ (since $\deg_T q \geq 1$), other words,

$$\text{Leaves}\left(T'\right) \setminus \{u\} \subseteq \text{Leaves}\left(T\right) \setminus \{v\} \qquad (18)$$

**(c)** Assume that $u$ is a leaf of $T'$. In other words, $u \in \text{Leaves}\left(T'\right)$. Hence,

$$\text{Leaves}\left(T'\right) = \underbrace{\left(\text{Leaves}\left(T'\right) \setminus \{u\}\right)}_{\substack{\subseteq \text{Leaves}(T)\setminus\{v\} \\ \text{(by (18))}}} \cup \{u\} \subseteq \left(\text{Leaves}\left(T\right) \setminus \{v\}\right) \cup \{u\}.$$

Combining this with

$$\underbrace{\left(\text{Leaves}\left(T\right) \setminus \{v\}\right)}_{\substack{\subseteq \text{Leaves}(T') \\ \text{(by (17))}}} \cup \underbrace{\{u\}}_{\substack{\subseteq \text{Leaves}(T') \\ \text{(since } u \in \text{Leaves}(T'))}} \subseteq \text{Leaves}\left(T'\right) \cup \text{Leaves}\left(T'\right) = \text{Leaves}\left(T'\right),$$

we obtain $\text{Leaves}\left(T'\right) = \left(\text{Leaves}\left(T\right) \setminus \{v\}\right) \cup \{u\}$. This proves Proposition 0.22 **(c)**.

**(d)** Assume that $u$ is not a leaf of $T'$. In other words, $u \notin \text{Leaves}\left(T'\right)$. Hence,

$$\text{Leaves}\left(T'\right) = \text{Leaves}\left(T'\right) \setminus \{u\} \subseteq \text{Leaves}\left(T\right) \setminus \{v\} \qquad \text{(by (18))}.$$

Combining this with (17), we obtain $\text{Leaves}\left(T'\right) = \text{Leaves}\left(T\right) \setminus \{v\}$. This proves Proposition 0.22 **(d)**. □

Also, here is a particularly trivial fact:

**Proposition 0.23.** Let $T$ be a tree such that $\left|\text{V}\left(T\right)\right| \leq 2$. Let $v$ and $w$ be two distinct leaves of $T$. Then, the vertices $v$ and $w$ are adjacent, and we have $\text{V}\left(T\right) = \text{Leaves}\left(T\right) = \{v, w\}$.

*Proof of Proposition 0.23.* We have $\text{Leaves}\left(T\right) \subseteq \text{V}\left(T\right)$ (since each leaf of $T$ is a vertex of $T$). Thus, $\left|\text{Leaves}\left(T\right)\right| \leq \left|\text{V}\left(T\right)\right| \leq 2$.

On the other hand, $v$ and $w$ are two distinct leaves of $T$. Thus, $T$ has at least two distinct leaves. In other words, $\left|\text{Leaves}\left(T\right)\right| \geq 2$. Combined with $\left|\text{Leaves}\left(T\right)\right| \leq 2$, this yields $\left|\text{Leaves}\left(T\right)\right| = 2$. Hence, $2 = \left|\text{Leaves}\left(T\right)\right| \leq \left|\text{V}\left(T\right)\right|$. Combining this with $\left|\text{V}\left(T\right)\right| \leq 2$, we find $\left|\text{V}\left(T\right)\right| = 2$. In other words, $\text{V}\left(T\right)$ is a 2-element set. Hence, the only 2-element subset of the set $\text{V}\left(T\right)$ is this set $\text{V}\left(T\right)$ itself.

Now, $v$ and $w$ are two distinct leaves of $T$. In other words, $v$ and $w$ are two distinct elements of $\text{Leaves}\left(T\right)$. Hence, $\{v, w\} \subseteq \text{Leaves}\left(T\right) \subseteq \text{V}\left(T\right)$. Thus, $\{v, w\}$ is a 2-element subset of the set $\text{V}\left(T\right)$ (in fact, $\{v, w\}$ is a 2-element set since $v$ and $w$ are distinct). Therefore, $\{v, w\}$ is the set $\text{V}\left(T\right)$ itself (since the only 2-element subset of the set $\text{V}\left(T\right)$ is this

so that $\deg_T q \geq 2$ (since $\deg_T q$ is an integer). In other words, there are at least 2 edges of $T$ that contain $q$. This yields that there are more edges of $T$ that contain $q$ than there are edges of $T'$ that contain $q$ (because there is only one edge of $T'$ that contains $q$). Hence, there exists an edge of $T$ that contains $q$ but that is not an edge of $T'$. This edge must be $e$ (since $e$ is the only edge of $T$ that is not an edge of $T'$). Thus, the edge $e$ contains $q$. In other words, $q$ is one of the two endpoints of $e$. In other words, $q$ is one of $u$ and $v$ (since the two endpoints of $e$ are $u$ and $v$). Since $q \neq u$, we thus obtain $q = v$. This contradicts $q \neq v$. This contradiction proves that our assumption was wrong. Hence, we cannot have $\deg_T q \neq 1$.

In other words, we have $\deg_T q = 1$. In other words, $q$ is a leaf of $T$. In other words, $q \in \text{Leaves}\left(T\right)$. Combining this with $q \neq v$, we obtain $q \in \text{Leaves}\left(T\right) \setminus \{v\}$.

set $\operatorname{V}(T)$ itself). In other words, $\{v,w\} = \operatorname{V}(T)$. Combining $\{v,w\} \subseteq \operatorname{Leaves}(T)$ with $\operatorname{Leaves}(T) \subseteq \operatorname{V}(T) = \{v,w\}$, we find $\{v,w\} = \operatorname{Leaves}(T)$. Altogether, we thus know that $\operatorname{V}(T) = \{v,w\} = \operatorname{Leaves}(T)$, so that $\operatorname{V}(T) = \operatorname{Leaves}(T) = \{v,w\}$.

The vertex $v$ of $T$ is a leaf of $T$. In other words, $v$ is a vertex of $T$ such that $\deg v = 1$. In other words, $v$ is a vertex of $T$ having exactly one neighbor. Let $q$ be this neighbor.

Since $q \in \operatorname{V}(T)$ (since $q$ is a vertex of $T$) and $q \neq v$ (because $q$ is a neighbor of $v$), we have $q \in \underbrace{\operatorname{V}(T)}_{=\{v,w\}} \setminus \{v\} = \{v,w\} \setminus \{v\} \subseteq \{w\}$, so that $q = w$. But the vertices $v$ and $q$ are adjacent (since $q$ is a neighbor of $v$). Since $q = w$, this rewrites as follows: The vertices $v$ and $w$ are adjacent. This completes the proof of Proposition 0.23. □

Next, we introduce some simple notations:

**Definition 0.24.** For any $k \in \mathbb{N}$, we let $[k]$ denote the set $\{1, 2, \ldots, k\}$.

**Definition 0.25.** Let $V$ be a finite set. A *listing* of $V$ shall mean a list of elements of $V$ such that each element of $V$ appears exactly once in this list.

For example, the set $\{1, 4, 6\}$ has exactly 6 listings; two of them are $(1, 4, 6)$ and $(4, 1, 6)$.

**Definition 0.26.** Let $G$ be a multigraph. Let $p$ and $q$ be two vertices of $G$. Then, we write $p \operatorname{nad}_G q$ if and only if $p$ is not adjacent to $q$ in $G$.

The following fact is obvious:

**Proposition 0.27.** Let $V$ be a finite set. Let $v \in V$. Let $(v_1, v_2, \ldots, v_{n-1})$ be a listing of the set $V \setminus \{v\}$. Then, $(v_1, v_2, \ldots, v_{n-1}, v)$ is a listing of the set $V$.

Now comes a theorem which (once proven) will quickly yield the claim of Exercise 4:

**Theorem 0.28.** Let $n \in \mathbb{N}$. Let $T$ be a tree such that $|\operatorname{V}(T)| = n$. Let $v$ and $w$ be two distinct leaves of $T$. Then, there exists a listing $(v_1, v_2, \ldots, v_n)$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_n = v$ and

$$|\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

*Proof of Theorem 0.28.* We shall prove Theorem 0.28 by induction over $n$:

*Induction base:* Theorem 0.28 holds in the case when $n \leq 2$.

[*Proof.* Assume that $n \leq 2$. Thus, $|\operatorname{V}(T)| = n \leq 2$. Proposition 0.23 shows that the vertices $v$ and $w$ are adjacent, and that we have $\operatorname{V}(T) = \operatorname{Leaves}(T) = \{v,w\}$. From $\operatorname{Leaves}(T) = \{v,w\}$, we obtain $|\operatorname{Leaves}(T)| = |\{v,w\}| = 2$ (since $v$ and $w$ are distinct). From $\operatorname{V}(T) = \{v,w\}$, we obtain $|\operatorname{V}(T)| = |\{v,w\}| = 2$, whence $n = |\operatorname{V}(T)| = 2$.

But the list $(w, v)$ is a listing of the set $\{v, w\}$ (since $v$ and $w$ are distinct), i.e., is a listing of the set $\operatorname{V}(T)$ (since $\{v, w\} = \operatorname{V}(T)$). Denote this listing by $(v_1, v_2, \ldots, v_n)$. (This is well-defined, since the length of this listing is $2 = n$.) Thus, $(v_1, v_2, \ldots, v_n) = (w, v)$; hence, $v_1 = w$ and $v_n = v$. Moreover, there exists no $i \in [n-1]$ satisfying $v_i \operatorname{nad}_T v_{i+1}$ [25]. Hence, $\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\} = \varnothing$, so that

$$|\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| = 0 \leq 0 = \underbrace{2}_{=|\operatorname{Leaves}(T)|} - 2 = |\operatorname{Leaves}(T)| - 2.$$

Hence, we have constructed a listing $(v_1, v_2, \ldots, v_n)$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_n = v$ and

$$|\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

This proves that such a listing exists. In other words, Theorem 0.28 holds. We thus have proven Theorem 0.28 in the case when $n \leq 2$.]

This completes the induction base.

*Induction step:* Let $N > 2$ be an integer. Assume (as the induction hypothesis) that Theorem 0.28 holds in the case when $n = N - 1$. We must then prove that Theorem 0.28 holds in the case when $n = N$.

We have assumed that Theorem 0.28 holds in the case when $n = N - 1$. In other words, the following fact holds:

> *Fact 1:* Let $T$ be a tree such that $|\operatorname{V}(T)| = N - 1$. Let $v$ and $w$ be two distinct leaves of $T$. Then, there exists a listing $(v_1, v_2, \ldots, v_{N-1})$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_{N-1} = v$ and
>
> $$|\{i \in [(N-1) - 1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

Now, let us prove that Theorem 0.28 holds in the case when $n = N$.

Let $T$ be a tree such that $|\operatorname{V}(T)| = N$. Let $v$ and $w$ be two distinct leaves of $T$.

We want to prove that there exists a listing $(v_1, v_2, \ldots, v_N)$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_N = v$ and

$$|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

Such a listing will be called a *helpful listing*. Thus, we want to prove that there exists a helpful listing.

We have $|\operatorname{V}(T)| = N > 2$.

[25] *Proof.* Assume the contrary. Thus, there exists some $i \in [n-1]$ satisfying $v_i \operatorname{nad}_T v_{i+1}$. Consider this $i$.

We have $n = 2$, thus $n - 1 = 1$, thus $[n-1] = [1] = \{1\}$. Hence, $i \in [n-1] = \{1\}$, so that $i = 1$. Hence, $v_i = v_1 = w$. Moreover, from $i = 1$, we obtain $i + 1 = 2 = n$, so that $v_{i+1} = v_n = v$. Now, recall that $v_i \operatorname{nad}_T v_{i+1}$. In light of $v_i = w$ and $v_{i+1} = v$, this rewrites as $w \operatorname{nad}_T v$. In other words, the vertex $w$ is not adjacent to $v$ in $T$. This contradicts the fact that $w$ is adjacent to $v$ (since the vertices $v$ and $w$ are adjacent). This contradiction completes our proof.

Recall that $v$ is a leaf of $T$. Let $T'$ denote the multigraph obtained from $T$ by removing this leaf $v$ and the unique edge that contains $v$. Let $u$ be the unique neighbor of $v$ in $T$. Proposition 0.22 **(a)** shows that the multigraph $T'$ is a tree again. Proposition 0.22 **(b)** shows that $u \notin \operatorname{Leaves}(T)$, and thus $u \notin \operatorname{Leaves}(T) \setminus \{v\}$. Notice that $v \in \operatorname{Leaves}(T)$ (since $v$ is a leaf of $T$) and $w \in \operatorname{Leaves}(T)$ (since $w$ is a leaf of $T$). Thus, $w \neq u$ (because otherwise, we would have $w = u \notin \operatorname{Leaves}(T)$, which would contradict $w \in \operatorname{Leaves}(T)$). In other words, $u$ and $w$ are distinct. Also, $w \neq v$ (since $v$ and $w$ are distinct).

We know that $u$ is a neighbor of $v$ in $T$. Hence, there exists an edge of $T$ having endpoints $u$ and $v$. Let us denote this edge by $e$.

We have $N \geq 3$ (because $N$ is an integer and satisfies $N > 2$). The multigraph $T'$ was obtained by removing the vertex $v$ and one edge from $T$. Hence, the vertices of $T'$ are exactly the vertices of $T$ other than $v$. In other words, $\operatorname{V}(T') = \operatorname{V}(T) \setminus \{v\}$. Thus,

$$\left| \underbrace{\operatorname{V}(T')}_{=\operatorname{V}(T)\setminus\{v\}} \right| = \left|\operatorname{V}(T) \setminus \{v\}\right| = \underbrace{\left|\operatorname{V}(T)\right|}_{=N} - 1 \qquad (\text{since } v \in \operatorname{V}(T))$$
$$= N - 1 \geq 2 \qquad (\text{since } N \geq 3).$$

We are in one of the following two cases:

*Case 1:* The vertex $u$ is a leaf of $T'$.

*Case 2:* The vertex $u$ is not a leaf of $T'$.

Let us treat these cases separately:

- Let us first consider Case 1. In this case, the vertex $u$ is a leaf of $T'$. Hence, Proposition 0.22 **(c)** shows that $\operatorname{Leaves}(T') = (\operatorname{Leaves}(T) \setminus \{v\}) \cup \{u\}$. Hence,

$$\begin{aligned} \left|\operatorname{Leaves}(T')\right| &= \left|(\operatorname{Leaves}(T) \setminus \{v\}) \cup \{u\}\right| \\ &= \underbrace{\left|\operatorname{Leaves}(T) \setminus \{v\}\right|}_{\substack{=\left|\operatorname{Leaves}(T)\right|-1 \\ (\text{since } v \in \operatorname{Leaves}(T))}} + 1 \qquad (\text{since } u \notin \operatorname{Leaves}(T) \setminus \{v\}) \\ &= \left(\left|\operatorname{Leaves}(T)\right| - 1\right) + 1 = \left|\operatorname{Leaves}(T)\right|. \end{aligned}$$

  But $w \in \operatorname{Leaves}(T) \setminus \{v\}$ (since $w \in \operatorname{Leaves}(T)$ and $w \neq v$) and thus $w \in \operatorname{Leaves}(T) \setminus \{v\} \subseteq (\operatorname{Leaves}(T) \setminus \{v\}) \cup \{u\} = \operatorname{Leaves}(T')$. Hence, $w$ is a leaf of $T'$. Hence, $u$ and $w$ are two distinct leaves of $T'$ (since $u$ and $w$ are leaves of $T'$, and are distinct). Thus, Fact 1 (applied to $T'$ and $u$ instead of $T$ and $v$) shows that there exists a listing $(v_1, v_2, \ldots, v_{N-1})$ of $\operatorname{V}(T')$ such that $v_1 = w$ and $v_{N-1} = u$ and

$$\left|\left\{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\right\}\right| \leq \left|\operatorname{Leaves}(T')\right| - 2. \qquad (19)$$

  Consider this listing.

We know that $(v_1, v_2, \ldots, v_{N-1})$ is a listing of the set $\operatorname{V}(T') = \operatorname{V}(T) \setminus \{v\}$. Hence, Proposition 0.27 (applied to $V = \operatorname{V}(T)$ and $n = N$) shows that $(v_1, v_2, \ldots, v_{N-1}, v)$ is a listing of the set $\operatorname{V}(T)$.

Let us extend the $(N-1)$-tuple $(v_1, v_2, \ldots, v_{N-1})$ to an $N$-tuple $(v_1, v_2, \ldots, v_N)$ by setting $v_N = v$. Thus, $(v_1, v_2, \ldots, v_N) = (v_1, v_2, \ldots, v_{N-1}, v)$ is a listing of the set $\operatorname{V}(T)$ (as we have just seen). Furthermore, $v_1 = w$ and $v_N = v$. Next, we claim that

$$\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \subseteq \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}. \qquad (20)$$

[*Proof of* (20): Let $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$ be arbitrary. Thus, $j$ is an element of $[N-1]$ satisfying $v_j \operatorname{nad}_T v_{j+1}$. We have $j \neq N-1$ [26]. Combined with $j \in [N-1]$, this yields $j \in [N-1] \setminus \{N-1\} = [(N-1)-1]$. Hence, both $v_j$ and $v_{j+1}$ are entries of the list $(v_1, v_2, \ldots, v_{N-1})$, and therefore are elements of $\operatorname{V}(T')$ (since this list $(v_1, v_2, \ldots, v_{N-1})$ is a listing of $\operatorname{V}(T')$). But we have $v_j \operatorname{nad}_T v_{j+1}$. In other words, the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T$. If the vertex $v_j$ was adjacent to $v_{j+1}$ in $T'$, then it would also be adjacent to $v_{j+1}$ in $T$ (since $T'$ is a sub-multigraph of $T$), which would contradict the fact that the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T$. Hence, the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T'$. In other words, we have $v_j \operatorname{nad}_{T'} v_{j+1}$.

Now, we have shown that $j$ is an element of $[(N-1)-1]$, and that this element $j$ satisfies $v_j \operatorname{nad}_{T'} v_{j+1}$. Hence, $j \in \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}$.

Now, forget that we fixed $j$. We thus have proven that $j \in \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}$ for each $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$. In other words,

$$\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \subseteq \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}.$$

This proves (20).]

From (20), we obtain

$$\begin{aligned} |\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| &\leq |\{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}| \\ &\leq \underbrace{|\operatorname{Leaves}(T')|}_{=|\operatorname{Leaves}(T)|} - 2 \qquad \text{(by (19))} \\ &= |\operatorname{Leaves}(T)| - 2. \end{aligned}$$

Thus, we have shown that $(v_1, v_2, \ldots, v_N)$ is a listing of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_N = v$ and

$$|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

[26] *Proof.* Assume the contrary. Thus, $j = N-1$. Hence, $v_j = v_{N-1} = u$. Also, from $j = N-1$, we obtain $j+1 = N$, so that $v_{j+1} = v_N = v$. Now, $v_j \operatorname{nad}_T v_{j+1}$ rewrites as $u \operatorname{nad}_T v$ (since $v_j = u$ and $v_{j+1} = v$). In other words, the vertex $u$ is not adjacent to $v$ in $T$. In other words, $u$ is not a neighbor of $v$ in $T$. This contradicts the fact that $u$ is a neighbor of $v$ in $T$. This contradiction proves that our assumption was wrong, qed.

In other words, $(v_1, v_2, \ldots, v_N)$ is a helpful listing. Hence, there exists a helpful listing in Case 1.

- Let us now consider Case 2. In this case, the vertex $u$ is not a leaf of $T'$. Hence, Proposition 0.22 **(d)** shows that $\operatorname{Leaves}(T') = \operatorname{Leaves}(T) \setminus \{v\}$. Hence,

$$\left|\operatorname{Leaves}(T')\right| = \left|\operatorname{Leaves}(T) \setminus \{v\}\right| = \left|\operatorname{Leaves}(T)\right| - 1$$

(since $v \in \operatorname{Leaves}(T)$).

But $w \in \operatorname{Leaves}(T) \setminus \{v\}$ (since $w \in \operatorname{Leaves}(T)$ and $w \neq v$) and thus $w \in \operatorname{Leaves}(T) \setminus \{v\} = \operatorname{Leaves}(T')$. Hence, $w$ is a leaf of $T'$.

The tree $T'$ has at least two vertices (since $|\operatorname{V}(T')| \geq 2$). It is known that any tree with at least two vertices must have at least two leaves. Since $T'$ is a tree with at least two vertices, we thus conclude that $T'$ has at least two leaves. Hence, $T'$ has at least one leaf distinct from $w$. Pick such a leaf, and denote it by $p$. Hence, $p$ and $w$ are two distinct leaves of $T'$ (since $p$ and $w$ are leaves of $T'$, and since $p$ is distinct from $w$). Thus, Fact 1 (applied to $T'$ and $p$ instead of $T$ and $v$) shows that there exists a listing $(v_1, v_2, \ldots, v_{N-1})$ of $\operatorname{V}(T')$ such that $v_1 = w$ and $v_{N-1} = p$ and

$$\left|\{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}\right| \leq \left|\operatorname{Leaves}(T')\right| - 2. \tag{21}$$

Consider this listing.

We know that $(v_1, v_2, \ldots, v_{N-1})$ is a listing of the set $\operatorname{V}(T') = \operatorname{V}(T) \setminus \{v\}$. Hence, Proposition 0.27 (applied to $V = \operatorname{V}(T)$ and $n = N$) shows that $(v_1, v_2, \ldots, v_{N-1}, v)$ is a listing of the set $\operatorname{V}(T)$.

Let us extend the $(N-1)$-tuple $(v_1, v_2, \ldots, v_{N-1})$ to an $N$-tuple $(v_1, v_2, \ldots, v_N)$ by setting $v_N = v$. Thus, $(v_1, v_2, \ldots, v_N) = (v_1, v_2, \ldots, v_{N-1}, v)$ is a listing of the set $\operatorname{V}(T)$ (as we have just seen). Furthermore, $v_1 = w$ and $v_N = v$. Next, we claim that

$$\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\} \subseteq \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}. \tag{22}$$

[*Proof of* (22)*:* Let $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\}$ be arbitrary. Thus, $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$ and $j \notin \{N-1\}$. From $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$, we see that $j$ is an element of $[N-1]$ satisfying $v_j \operatorname{nad}_T v_{j+1}$. From $j \notin \{N-1\}$, we obtain $j \neq N-1$. Combined with $j \in [N-1]$, this yields $j \in [N-1] \setminus \{N-1\} = [(N-1)-1]$. Hence, both $v_j$ and $v_{j+1}$ are entries of the list $(v_1, v_2, \ldots, v_{N-1})$, and therefore are elements of $\operatorname{V}(T')$ (since this list $(v_1, v_2, \ldots, v_{N-1})$ is a listing of $\operatorname{V}(T')$). But we have $v_j \operatorname{nad}_T v_{j+1}$. In other words, the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T$. If the vertex $v_j$ was adjacent to $v_{j+1}$ in $T'$, then it would also be adjacent to $v_{j+1}$ in $T$ (since $T'$ is a sub-multigraph of $T$), which would contradict the fact that the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T$. Hence, the vertex $v_j$ is not adjacent to $v_{j+1}$ in $T'$. In other words, we have $v_j \operatorname{nad}_{T'} v_{j+1}$.

Now, we have shown that $j$ is an element of $[(N-1)-1]$, and that this element $j$ satisfies $v_j \operatorname{nad}_{T'} v_{j+1}$. Hence, $j \in \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}$.

Now, forget that we fixed $j$. We thus have proven that $j \in \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}$ for each $j \in \{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\}$. In other words,

$$\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\} \subseteq \{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}.$$

This proves (22).]

Now,

$$\begin{aligned} &|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\}| \\ &\geq |\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| - \underbrace{|\{N-1\}|}_{=1} \\ &= |\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| - 1. \end{aligned}$$

Hence,

$$\begin{aligned} &|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| - 1 \\ &\leq |\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\} \setminus \{N-1\}| \\ &\leq |\{i \in [(N-1)-1] \mid v_i \operatorname{nad}_{T'} v_{i+1}\}| \qquad \text{(by (22))} \\ &\leq \underbrace{|\operatorname{Leaves}(T')|}_{=|\operatorname{Leaves}(T)|-1} - 2 \qquad \text{(by (19))} \\ &= (|\operatorname{Leaves}(T)| - 1) - 2 = |\operatorname{Leaves}(T)| - 3. \end{aligned}$$

Adding 1 to both sides of this inequality, we obtain

$$|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

Thus, we have shown that $(v_1, v_2, \ldots, v_N)$ is a listing of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_N = v$ and

$$|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

In other words, $(v_1, v_2, \ldots, v_N)$ is a helpful listing. Hence, there exists a helpful listing in Case 2.

Thus, in each of the two Cases 1 and 2, we have shown that there exists a helpful listing. Hence, there is always a helpful listing (since the two cases cover all possibilities). In other words, there always exists a listing $(v_1, v_2, \ldots, v_N)$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_N = v$ and

$$|\{i \in [N-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2.$$

Hence, we have proven that Theorem 0.28 holds in the case when $n = N$. This completes the induction step. Thus, Theorem 0.28 is proven. □

*Solution to Exercise 4 (sketched).* Let $n = |\operatorname{V}(T)|$. The tree $T$ has more than 1 vertex. In other words, $|\operatorname{V}(T)| > 1$. Hence, $n = |\operatorname{V}(T)| > 1$. Thus, $n \geq 2$ (since $n$ is an integer).

The tree $T$ has at least two vertices (since $|\operatorname{V}(T)| = n \geq 2$). It is known that any tree with at least two vertices must have at least two leaves. Since $T$ is a tree with at least two vertices, we thus conclude that $T$ has at least two leaves. In other words, there exist two distinct leaves of $T$. Pick two such leaves, and denote them by $v$ and $w$. Theorem 0.28 shows that there exists a listing $(v_1, v_2, \ldots, v_n)$ of $\operatorname{V}(T)$ such that $v_1 = w$ and $v_n = v$ and

$$|\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq |\operatorname{Leaves}(T)| - 2. \tag{23}$$

Consider such a listing.

The $n$ elements $v_1, v_2, \ldots, v_n$ are distinct elements of $\operatorname{V}(T)$ (since $(v_1, v_2, \ldots, v_n)$ is a listing of $\operatorname{V}(T)$). Hence, $v_n \neq v_1$ (since $n \geq 2$).

Define a subset $S$ of $[n-1]$ by $S = \{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$. Thus,

$$|S| = |\{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}| \leq \left|\underbrace{\operatorname{Leaves}(T)}_{=L}\right| - 2 \qquad (\text{by (23)})$$
$$= |L| - 2.$$

Hence,

$$|L| - 1 - \underbrace{|S|}_{\leq |L|-2} \geq |L| - 1 - (|L| - 2) = 1.$$

Now, let $G$ be the multigraph obtained from $T$ by the following procedure:

- **Step 1:** For each $i \in S$, add an edge with endpoints $v_i$ and $v_{i+1}$. (This is possible, because $v_i$ and $v_{i+1}$ are distinct vertices of $T$ (since $v_1, v_2, \ldots, v_n$ are distinct).)
- **Step 2:** Add $|L| - 1 - |S|$ many edges with endpoints $v_n$ and $v_1$. (This is possible, since $v_n \neq v_1$, and because $|L| - 1 - |S| \geq 1 \geq 0$.)

Let us state some simple observations about this multigraph $G$:

> *Observation 1:* The multigraph $G$ is obtained from $T$ by adding $|L| - 1$ new edges to $T$.

*Proof of Observation 1:* The multigraph $G$ is obtained from $T$ by the above procedure, which has two steps. Both steps consist in adding edges to the multigraph. In Step 1, exactly $|S|$ many edges are added (because one edge is added for each $i \in S$). In Step 2, exactly $|L| - 1 - |S|$ further edges are added. Thus, altogether, $|S| + (|L| - 1 - |S|) = |L| - 1$ many edges are added. Hence, the multigraph $G$ is obtained from $T$ by adding $|L| - 1$ new edges to $T$. This proves Observation 1. $\square$

*Observation 2:* The list $(v_1, v_2, \ldots, v_n)$ is a listing of $\operatorname{V}(G)$.

*Proof of Observation 2:* The multigraph $G$ is obtained from $T$ by adding some edges; but no vertices are ever added. Hence, the vertices of $G$ are precisely the vertices of $T$. In other words, $\operatorname{V}(G) = \operatorname{V}(T)$.

But recall that the list $(v_1, v_2, \ldots, v_n)$ is a listing of $\operatorname{V}(T)$. Since $\operatorname{V}(G) = \operatorname{V}(T)$, this rewrites as follows: The list $(v_1, v_2, \ldots, v_n)$ is a listing of $\operatorname{V}(G)$. This proves Observation 2. □

Now, let us set $v_{n+1} = v_1$. Thus, $n+1$ vertices $v_1, v_2, \ldots, v_{n+1}$ of $G$ are defined.

*Observation 3:* Let $j \in [n]$. Then, there exists an edge of $G$ whose endpoints are $v_j$ and $v_{j+1}$.

*Proof of Observation 3:* We are in one of the following three cases:

- *Case 1:* We have $j = n$.
- *Case 2:* We have $j \in S$.
- *Case 3:* We have neither $j = n$ nor $j \in S$.

We shall consider these cases separately:

- Let us first consider Case 1. In this case, we have $j = n$. Thus, $v_j = v_n$ and $v_{j+1} = v_{n+1} = v_1$. Now, recall the procedure that we used to define the multigraph $G$. This procedure consisted of two steps. In Step 2, we have added $|L| - 1 - |S|$ many edges with endpoints $v_n$ and $v_1$. Since $|L| - 1 - |S| \geq 1$, this shows that we have added **at least one** edge with endpoints $v_n$ and $v_1$. Hence, the multigraph $G$ has at least one edge whose endpoints are $v_n$ and $v_1$. Since $v_j = v_n$ and $v_{j+1} = v_1$, this rewrites as follows: The multigraph $G$ has at least one edge whose endpoints are $v_j$ and $v_{j+1}$. In other words, there exists an edge of $G$ whose endpoints are $v_j$ and $v_{j+1}$. Hence, Observation 3 is proven in Case 1.

- Let us now consider Case 2. In this case, we have $j \in S$. Now, recall the procedure that we used to define the multigraph $G$. This procedure consisted of two steps. In Step 1, we have added an edge with endpoints $v_i$ and $v_{i+1}$ for each $i \in S$. Thus, in particular, we have added an edge with endpoints $v_j$ and $v_{j+1}$ (since $j \in S$). Hence, the multigraph $G$ has at least one edge whose endpoints are $v_j$ and $v_{j+1}$. In other words, there exists an edge of $G$ whose endpoints are $v_j$ and $v_{j+1}$. Hence, Observation 3 is proven in Case 2.

- Let us now consider Case 3. In this case, we have neither $j = n$ nor $j \in S$. Since we have $j \in [n]$ and $j \neq n$ (since we do not have $j = n$), we must have $j \in [n] \setminus \{n\} = [n-1]$. But the multigraph $G$ was obtained from $T$ by adding edges. Thus, $T$ is a sub-multigraph of $G$. Hence, each edge of $T$ is an edge of $G$. Now, the vertices $v_j$ and $v_{j+1}$ of $T$ are adjacent[27]. In other words, there exists an edge of $T$ whose endpoints are $v_j$ and $v_{j+1}$. Therefore, there exists an edge of $G$ whose endpoints are $v_j$ and $v_{j+1}$ (since each edge of $T$ is an edge of $G$). Hence, Observation 3 is proven in Case 3.

We thus have proven Observation 3 in all three cases. □

The list $(v_1, v_2, \ldots, v_n)$ is a listing of $\operatorname{V}(G)$ (by Observation 2). Thus, this list $(v_1, v_2, \ldots, v_n)$ contains each element of $\operatorname{V}(G)$ exactly once. In other words, this list $(v_1, v_2, \ldots, v_n)$ contains each vertex of $G$ exactly once. In other words, each vertex of $G$ appears exactly once among the vertices $v_1, v_2, \ldots, v_n$.

Now, we define $n$ edges $e_1, e_2, \ldots, e_n$ of $G$ as follows: For each $j \in [n]$, we pick some edge of $G$ whose endpoints are $v_j$ and $v_{j+1}$ (such an edge exists because of Observation 3), and denote this edge by $e_j$. Thus, $(v_1, e_1, v_2, e_2, v_3, \ldots, v_n, e_n, v_{n+1})$ is a walk in $G$. This walk is a circuit (since $v_{n+1} = v_1$) and therefore a cycle (since the elements $v_1, v_2, \ldots, v_n$ are distinct). The non-ultimate vertices of this cycle are $v_1, v_2, \ldots, v_n$. Hence, each vertex of $G$ appears exactly once among the non-ultimate vertices of this cycle (since each vertex of $G$ appears exactly once among the vertices $v_1, v_2, \ldots, v_n$). In other words, this cycle is a Hamiltonian cycle of $G$. Thus, the multigraph $G$ has a Hamiltonian cycle.

The multigraph $G$ is obtained by adding $|L| - 1$ new edges to $T$ (by Observation 1), and has a Hamiltonian cycle. Thus, it is possible to add $|L| - 1$ new edges to $T$ in such a way that the resulting multigraph has a Hamiltonian cycle. This solves the exercise. □

## 0.6. Exercise 5: the maximum perimeter of a triangle on a digraph

### 0.6.1. Distances in a digraph

If $u$ and $v$ are two vertices of a digraph $D$, then $d(u, v)$ denotes the *distance* from $u$ to $v$. This is defined to be the minimum length of a path from $u$ to $v$ if such a path exists; otherwise it is defined to be the symbol $\infty$. Notice that $d(u, v)$ is not usually the same as $d(v, u)$ (unlike for simple graphs).

We observe the following simple facts:[28]

[27] *Proof.* Assume the contrary. Thus, the vertices $v_j$ and $v_{j+1}$ of $T$ are not adjacent. In other words, $v_j \operatorname{nad}_T v_{j+1}$. Now, we know that $j$ is an element of $[n-1]$ satisfying $v_j \operatorname{nad}_T v_{j+1}$. In other words, $j \in \{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$. In light of $S = \{i \in [n-1] \mid v_i \operatorname{nad}_T v_{i+1}\}$, this rewrites as $j \in S$. This contradicts the fact that we don't have $j \in S$. This contradiction shows that our assumption was false.

[28] The following facts are analogues of the facts proven in Section 0.8.1 of the solution to midterm 1. They are proven in the same way as the latter facts (of course, simple graphs must be replaced by digraphs).

**Lemma 0.29.** Let $u$ and $v$ be two vertices of a strongly connected digraph $D = (V, A)$. Then, $d(u, v) \leq |V| - 1$.

Lemma 0.29 shows that if $u$ and $v$ are two vertices of a strongly connected digraph $D$, then $d(u, v)$ is an actual integer (as opposed to $\infty$).

**Lemma 0.30.** Let $u$ and $v$ be two vertices of a digraph $D$. Let $k \in \mathbb{N}$. If there exists a walk from $u$ to $v$ in $D$ having length $k$, then $d(u, v) \leq k$.

**Lemma 0.31.** Let $D = (V, A)$ be a digraph.

**(a)** Each $u \in V$ satisfies $d(u, u) = 0$.

**(b)** Each $u \in V$, $v \in V$ and $w \in V$ satisfy $d(u, v) + d(v, w) \geq d(u, w)$. (This inequality has to be interpreted appropriately when one of the distances is infinite: For example, we understand $\infty$ to be greater than any integer, and we understand $\infty + m$ to be $\infty$ whenever $m \in \mathbb{Z}$.)

**(c)** If $u \in V$ and $v \in V$ satisfy $d(u, v) = 0$, then $u = v$.

### 0.6.2. The exercise

**Exercise 5.** Let $a$, $b$ and $c$ be three vertices of a strongly connected digraph $D = (V, A)$ such that $|V| \geq 4$.

**(a)** Prove that $d(b, c) + d(c, a) + d(a, b) \leq 3|V| - 4$.

**(b)** For each $n \geq 5$, construct an example in which $|V| = n$ and $d(b, c) + d(c, a) + d(a, b) = 3|V| - 4$. (No proof required for the example.)

We shall solve Exercise 5 after proving some lemmas:

**Lemma 0.32.** Let $a$, $b$ and $c$ be three vertices of a digraph $D = (V, A)$. Let $(y_0, y_1, \ldots, y_j)$ be a walk from $a$ to $c$. Let $(z_0, z_1, \ldots, z_k)$ be a walk from $c$ to $b$. Set $Y = \{y_0, y_1, \ldots, y_{j-1}\}$ and $Z = \{z_0, z_1, \ldots, z_{k-1}\}$.

Assume that $Y \cap Z \neq \varnothing$. Then, $d(a, b) < j + k$.

*Proof of Lemma 0.32.* Recall that $(y_0, y_1, \ldots, y_j)$ is a walk from $a$ to $c$. Hence, $y_0 = a$ and $y_j = c$.

Also, $(z_0, z_1, \ldots, z_k)$ is a walk from $c$ to $b$. Thus, $z_0 = c$ and $z_k = b$.

We have $Y \cap Z \neq \varnothing$. Hence, there exists some $v \in Y \cap Z$. Consider this $v$.

We have $v \in Y = \{y_0, y_1, \ldots, y_{j-1}\}$. In other words, $v = y_g$ for some $g \in \{0, 1, \ldots, j-1\}$. Consider this $g$.

We have $v \in Z = \{z_0, z_1, \ldots, z_{k-1}\}$. In other words, $v = z_h$ for some $h \in \{0, 1, \ldots, k-1\}$. Consider this $h$.

From $g \in \{0, 1, \ldots, j-1\}$, we obtain $g \leq j - 1 < j$.

Consider the subwalk[29] $(y_0, y_1, \ldots, y_g)$ of the walk $(y_0, y_1, \ldots, y_j)$. This subwalk $(y_0, y_1, \ldots, y_g)$ is a walk from $a$ to $v$ (since $y_0 = a$ and $y_g = v$).

[29] Here, a *subwalk* of a walk $(w_0, w_1, \ldots, w_m)$ means a list of the form $(w_I, w_{I+1}, \ldots, w_J)$ for two elements $I$ and $J$ of $\{0, 1, \ldots, m\}$ satisfying $I \leq J$. Such a list is always a walk.

Consider the subwalk $(z_h, z_{h+1}, \ldots, z_k)$ of the walk $(z_0, z_1, \ldots, z_k)$. This subwalk $(z_h, z_{h+1}, \ldots, z_k)$ is a walk from $v$ to $b$ (since $z_h = v$ and $z_k = b$).

Now, the ending point of the walk $(y_0, y_1, \ldots, y_g)$ is the starting point of the walk $(z_h, z_{h+1}, \ldots, z_k)$ (since $y_g = v = z_h$). Hence, these two walks can be combined to one walk

$$(y_0, y_1, \ldots, y_{g-1}, z_h, z_{h+1}, \ldots, z_k) = (y_0, y_1, \ldots, y_g, z_{h+1}, z_{h+2}, \ldots, z_k).$$

This resulting walk is a walk from $a$ to $b$ (since $y_0 = a$ and $z_k = b$), and has length $g + 1 + (k - h - 1)$. Hence, there exists a walk from $a$ to $b$ in $D$ having length $g + 1 + (k - h - 1)$ (namely, the walk we have just constructed). Hence, Lemma 0.30 (applied to $a$, $b$ and $g + 1 + (k - h - 1)$ instead of $u$, $v$ and $k$) shows that

$$d(a, b) \leq g + 1 + (k - h - 1) = \underbrace{g}_{<j} + k - \underbrace{h}_{\geq 0} < j + k - 0 = j + k.$$

This proves Lemma 0.32. □

**Lemma 0.33.** Let $a$, $b$ and $c$ be three vertices of a strongly connected digraph $D = (V, A)$.

Assume that we have

$$d(b, c) = d(b, a) + d(a, c) \qquad \text{and}$$
$$d(c, a) = d(c, b) + d(b, a) \qquad \text{and}$$
$$d(a, b) = d(a, c) + d(c, b).$$

Then,

$$d(c, b) + d(a, c) + d(b, a) \leq |V|.$$

*Proof of Lemma 0.33.* There exists a walk from $c$ to $b$ in $D$ [30]. Hence, there exists a path from $c$ to $b$ in $D$. Hence, there exists a path from $c$ to $b$ in $D$ having minimum length. Fix such a path, and denote it by $(x_0, x_1, \ldots, x_i)$. Define a subset $X$ of $V$ by $X = \{x_0, x_1, \ldots, x_{i-1}\}$.

There exists a walk from $a$ to $c$ in $D$ [31]. Hence, there exists a path from $a$ to $c$ in $D$. Hence, there exists a path from $a$ to $c$ in $D$ having minimum length. Fix such a path, and denote it by $(y_0, y_1, \ldots, y_j)$. Define a subset $Y$ of $V$ by $Y = \{y_0, y_1, \ldots, y_{j-1}\}$.

There exists a walk from $b$ to $a$ in $D$ [32]. Hence, there exists a path from $b$ to $a$ in $D$. Hence, there exists a path from $b$ to $a$ in $D$ having minimum length. Fix such a path, and denote it by $(z_0, z_1, \ldots, z_k)$. Define a subset $Z$ of $V$ by $Z = \{z_0, z_1, \ldots, z_{k-1}\}$.

[30] since $D$ is strongly connected.
[31] since $D$ is strongly connected.
[32] since $D$ is strongly connected.

There exists a path from $c$ to $b$ (as we know). Hence, $d(c,b)$ is defined as the minimum length of a path from $c$ to $b$. Thus,

$$\begin{aligned} d(c,b) &= (\text{the minimum length of a path from } c \text{ to } b) \\ &= (\text{the length of the path } \{x_0, x_1, \ldots, x_{i-1}\}) \\ &\qquad (\text{since } \{x_0, x_1, \ldots, x_{i-1}\} \text{ is a path from } c \text{ to } b \text{ having minimum length}) \\ &= i. \end{aligned}$$

The same argument (applied to $a$, $c$, $j$ and $(y_0, y_1, \ldots, y_j)$ instead of $c$, $b$, $i$ and $(x_0, x_1, \ldots, x_i)$) shows that $d(a,c) = j$. The same argument (applied to $b$, $a$, $k$ and $(z_0, z_1, \ldots, z_k)$ instead of $a$, $c$, $j$ and $(y_0, y_1, \ldots, y_j)$) shows that $d(b,a) = k$.

We have $|X| = i$ [33]. Similarly, $|Y| = j$ and $|Z| = k$.

Now, using Lemma 0.32, we can easily see that

$$Y \cap Z = \varnothing \tag{24}$$

[34]. Similarly,

$$Z \cap X = \varnothing \tag{25}$$

and

$$X \cap Y = \varnothing. \tag{26}$$

The equalities (24), (25) and (26) (combined) show that the sets $X$, $Y$ and $Z$ are disjoint. Hence, the size of the union of these sets equals the sum of their sizes. In other words, $|X \cup Y \cup Z| = |X| + |Y| + |Z|$. But $X \cup Y \cup Z \subseteq V$ (since $X$, $Y$ and $Z$ are subsets of $V$), and thus $|X \cup Y \cup Z| \leq |V|$. Hence,

$$|V| \geq |X \cup Y \cup Z| = \underbrace{|X|}_{=i=d(c,b)} + \underbrace{|Y|}_{=j=d(a,c)} + \underbrace{|Z|}_{=k=d(b,a)} = d(c,b) + d(a,c) + d(b,a).$$

This proves Lemma 0.33. □

**Lemma 0.34.** Let $a$, $b$ and $c$ be three vertices of a strongly connected digraph $D = (V, A)$.

Assume that $d(a,b) = |V| - 1$. Then, $d(a,b) = d(a,c) + d(c,b)$.

[33] *Proof.* The list $(x_0, x_1, \ldots, x_i)$ is a path. Hence, the vertices $x_0, x_1, \ldots, x_i$ are the vertices of a path, and therefore are distinct (since the vertices of any path are distinct). Thus, the $i$ vertices $x_0, x_1, \ldots, x_{i-1}$ are distinct as well. Hence, $|\{x_0, x_1, \ldots, x_{i-1}\}| = i$. Since $\{x_0, x_1, \ldots, x_{i-1}\} = X$, this rewrites as $|X| = i$.

[34] *Proof of* (24). Assume the contrary. Thus, $Y \cap Z \neq \varnothing$. Also, the list $(y_0, y_1, \ldots, y_j)$ is a walk from $a$ to $c$ (since it is a path from $a$ to $c$). Furthermore, the list $(z_0, z_1, \ldots, z_k)$ is a walk from $b$ to $a$ (since it is a path from $b$ to $a$). Thus, Lemma 0.32 (applied to $b$, $c$, $a$, $k$, $j$, $(z_0, z_1, \ldots, z_k)$, $(y_0, y_1, \ldots, y_j)$, $Z$ and $Y$ instead of $a$, $b$, $c$, $j$, $k$, $(y_0, y_1, \ldots, y_j)$, $(z_0, z_1, \ldots, z_k)$, $Y$ and $Z$) shows that $d(b,c) < \underbrace{k}_{=d(b,a)} + \underbrace{j}_{=d(a,c)} = d(b,a) + d(a,c)$. This contradicts $d(b,c) = d(b,a) + d(a,c)$. This contradiction shows that our assumption was wrong. This proves (24).

*Proof of Lemma 0.34.* We have $d(a,b) = |V| - 1 \neq \infty$. Hence, there exists a path from $a$ to $b$. Thus, $d(a,b)$ is defined as the minimum length of a path from $a$ to $b$. Hence, there exists a path from $a$ to $b$ having length $d(a,b)$. Fix such a path, and denote it by $(p_0, p_1, \ldots, p_k)$.

Thus, $(p_0, p_1, \ldots, p_k)$ is a path of length $d(a,b)$. Hence,

$$d(a,b) = (\text{the length of the path } (p_0, p_1, \ldots, p_k)) = k.$$

Hence, $k = d(a,b) = |V| - 1$, so that $k + 1 = |V|$.

But the $k+1$ elements $p_0, p_1, \ldots, p_k$ are the vertices of a path (namely, of the path $(p_0, p_1, \ldots, p_k)$), and thus are distinct (since the vertices of a path are always distinct). Hence, $|\{p_0, p_1, \ldots, p_k\}| = k + 1 = |V|$. Clearly, $\{p_0, p_1, \ldots, p_k\}$ is a subset of the finite set $V$.

Now, recall the following simple fact: If $S$ is a finite set, and if $T$ is a subset of $S$ satisfying $|T| = |S|$, then $T = S$. Applying this fact to $S = V$ and $T = \{p_0, p_1, \ldots, p_k\}$, we obtain $\{p_0, p_1, \ldots, p_k\} = V$ (since $\{p_0, p_1, \ldots, p_k\}$ is a subset of the finite set $V$ satisfying $|\{p_0, p_1, \ldots, p_k\}| = |V|$). Hence, $c \in V = \{p_0, p_1, \ldots, p_k\}$. In other words, $c = p_i$ for some $i \in \{0, 1, \ldots, k\}$. Consider this $i$.

Clearly, the list $(p_0, p_1, \ldots, p_k)$ is a walk (since it is a path). Consider the subwalk[35] $(p_0, p_1, \ldots, p_i)$ of the walk $(p_0, p_1, \ldots, p_k)$. This subwalk $(p_0, p_1, \ldots, p_i)$ is a walk from $a$ to $c$ (since $p_0 = a$ and $p_i = c$) and has length $i$. Thus, there exists a walk from $a$ to $c$ having length $i$ (namely, the subwalk we have just constructed). Therefore, Lemma 0.30 (applied to $a$, $c$ and $i$ instead of $u$, $v$ and $k$) shows that $d(a,c) \leq i$.

Consider the subwalk $(p_i, p_{i+1}, \ldots, p_k)$ of the walk $(p_0, p_1, \ldots, p_k)$. This subwalk $(p_i, p_{i+1}, \ldots, p_k)$ is a walk from $c$ to $b$ (since $p_i = c$ and $p_k = b$) and has length $k - i$. Thus, there exists a walk from $c$ to $b$ having length $k - i$ (namely, the subwalk we have just constructed). Therefore, Lemma 0.30 (applied to $c$, $b$ and $k - i$ instead of $u$, $v$ and $k$) shows that $d(c,b) \leq k - i$.

Lemma 0.31 **(b)** (applied to $a$, $c$ and $b$ instead of $u$, $v$ and $w$) shows that

$$d(a,c) + d(c,b) \geq d(a,b).$$

Combining this with the inequality

$$\underbrace{d(a,c)}_{\leq i} + \underbrace{d(c,b)}_{\leq k-i} \leq i + (k - i) = k = d(a,b),$$

we obtain $d(a,c) + d(c,b) = d(a,b)$. This proves Lemma 0.34. □

*Solution to Exercise 5 (sketched).* **(a)** Assume the contrary. Thus, $d(b,c) + d(c,a) + d(a,b) > 3|V| - 4$. Since $d(b,c) + d(c,a) + d(a,b)$ and $3|V| - 4$ are integers, this shows that

$$d(b,c) + d(c,a) + d(a,b) \geq (3|V| - 4) + 1 = 3|V| - 3. \tag{27}$$

[35] Here, a *subwalk* of a walk $(w_0, w_1, \ldots, w_m)$ means a list of the form $(w_I, w_{I+1}, \ldots, w_J)$ for two elements $I$ and $J$ of $\{0, 1, \ldots, m\}$ satisfying $I \leq J$. Such a list is always a walk.

Lemma 0.29 (applied to $u = b$ and $v = c$) yields $d(b,c) \leq |V| - 1$. Lemma 0.29 (applied to $u = c$ and $v = a$) yields $d(c,a) \leq |V| - 1$. Lemma 0.29 (applied to $u = a$ and $v = b$) yields $d(a,b) \leq |V| - 1$.

Now, subtracting $d(c,a) + d(a,b)$ from both sides of the inequality (27), we obtain

$$d(b,c) \geq 3|V| - 3 - \left( \underbrace{d(c,a)}_{\leq |V|-1} + \underbrace{d(a,b)}_{\leq |V|-1} \right) \geq 3|V| - 3 - \left( (|V|-1) + (|V|-1) \right) = |V| - 1.$$

Combining this with $d(b,c) \leq |V| - 1$, we obtain $d(b,c) = |V| - 1$. Similarly, $d(c,a) = |V| - 1$ and $d(a,b) = |V| - 1$.

Now, Lemma 0.34 yields

$$d(a,b) = d(a,c) + d(c,b). \tag{28}$$

Furthermore, Lemma 0.34 (applied to $b$, $c$ and $a$ instead of $a$, $b$ and $c$) yields

$$d(b,c) = d(b,a) + d(a,c). \tag{29}$$

Also, Lemma 0.34 (applied to $c$, $a$ and $b$ instead of $a$, $b$ and $c$) yields

$$d(c,a) = d(c,b) + d(b,a). \tag{30}$$

Hence, Lemma 0.33 yields

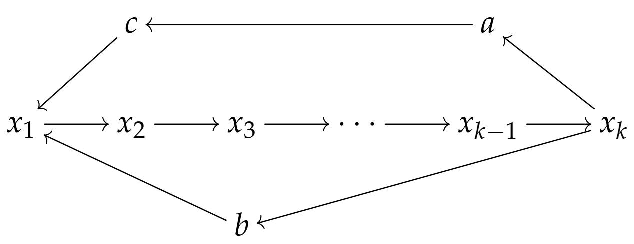
$$d(c,b) + d(a,c) + d(b,a) \leq |V|.$$

Now, adding together the three equalities (29), (30) and (28), we find

$$\begin{aligned} & d(b,c) + d(c,a) + d(a,b) \\ & = (d(b,a) + d(a,c)) + (d(c,b) + d(b,a)) + (d(a,c) + d(c,b)) \\ & = 2 \underbrace{(d(c,b) + d(a,c) + d(b,a))}_{\leq |V|} \leq 2|V| < 3|V| - 3 \end{aligned}$$

(since $(3|V| - 3) - 2|V| = \underbrace{|V|}_{\geq 4} - 3 \geq 4 - 3 = 1 > 0$). This contradicts (27). This contradiction proves that our assumption was false. Hence, Exercise 5 **(a)** is solved.

**(b)** One possible example is the following digraph:




where $k = n - 3$. In this digraph, we have $d(b,c) = n - 1$ and $d(c,a) = n - 2$ and $d(a,b) = n - 1$. □

## 0.7. Exercise 6: cycles of length divisible by 3, and proper 3-colorings

Recall that a *k-coloring* of a simple graph $G = (V, E)$ means a map $f : V \to \{1, 2, \ldots, k\}$. Such a $k$-coloring $f$ is said to be *proper* if no two adjacent vertices $u$ and $v$ have the same color (i.e., satisfy $f(u) = f(v)$).

**Exercise 6.** We have learned that a simple graph $G$ (or multigraph $G$) has a proper 2-coloring if and only if all cycles of $G$ have even length.

**(a)** Is it true that if all cycles of a simple graph $G$ (or multigraph $G$) have length divisible by 3, then $G$ has a proper 3-coloring?

**(b)** Is it true that if a simple graph $G$ has a proper 3-coloring, then all cycles of $G$ have length divisible by 3 ?

TODO: Solution. ((a) is true, but it is harder than I thought when posing this problem. Probably a proof appears in [AGJJ09]. (b) is definitely false, as witnessed by the 4-cycle $C_4$.

Ah, there is actually a simple proof of (a)! The main step is showing that if $G$ is a simple graph whose cycles all have length divisible by 3, then at least one vertex of $G$ has degree $< 3$ (unless $G$ has 0 vertices). This fact is an olympiad problem and has been discussed in Art of Problem Solving topic #5744, where a nice proof has been given by Pascual2005. Now, using this fact, we can fix a vertex $v$ having degree $< 3$, and (by induction over the number of vertices) assume that the graph $G$ with $v$ removed already has a proper 3-coloring; we then extend this 3-coloring by assigning an appropriate color to $v$.)

## 0.8. Exercise 7: Turán's theorem via independent sets

In class, we have proven the following fact:

**Theorem 0.35.** Let $G = (V, E)$ be a simple graph. Then, $G$ has an independent set of size $\geq \dfrac{n}{1+d}$, where $n = |V|$ and $d = \dfrac{1}{n} \sum\limits_{v \in V} \deg v$. (Notice that $d$ is simply the average degree of a vertex of $G$.)

On the other hand, recall the following fact (Theorem 2.5.15 in the lecture notes) which was left unproven in class:

**Theorem 0.36** (Turán's theorem)**.** Let $r$ be a positive integer. Let $G$ be a simple graph. Let $n = |\mathrm{V}(G)|$ be the number of vertices of $G$. Assume that $|\mathrm{E}(G)| > \dfrac{r-1}{r} \cdot \dfrac{n^2}{2}$. Then, there exist $r+1$ distinct vertices of $G$ that are mutually adjacent (i.e., each two distinct ones among these $r+1$ vertices are adjacent).

**Exercise 7.** Use Theorem 0.35 to prove Theorem 0.36.

In order to solve Exercise 7, let us first rewrite Theorem 0.35 in terms of the number $|E|$ of edges of $G$:

**Corollary 0.37.** Let $G = (V, E)$ be a simple graph. Then, $G$ has an independent set of size $\geq \dfrac{n^2}{n + 2|E|}$, where $n = |V|$.

*Proof of Corollary 0.37.* From $G = (V, E)$, we obtain $\operatorname{V}(G) = V$ and $\operatorname{E}(G) = E$.

We know (from Proposition 2.5.6 in the lecture notes) that $\sum\limits_{v \in \operatorname{V}(G)} \deg v = 2 |\operatorname{E}(G)|$.

Now, set $d = \dfrac{1}{n} \sum\limits_{v \in V} \deg v$. Then,

$$d = \frac{1}{n} \sum_{v \in V} \deg v = \frac{1}{n} \underbrace{\sum_{v \in \operatorname{V}(G)} \deg v}_{= 2|\operatorname{E}(G)|} \qquad (\text{since } V = \operatorname{V}(G))$$
$$= \frac{1}{n} \cdot 2 \left| \underbrace{\operatorname{E}(G)}_{= E} \right| = \frac{1}{n} \cdot 2|E| = 2|E|/n.$$

Hence,

$$\frac{n}{1 + d} = \frac{n}{1 + 2|E|/n} = \frac{n}{(n + 2|E|)/n} = \frac{n^2}{n + 2|E|}.$$

Now, Theorem 0.35 shows that the graph $G$ has an independent set of size $\geq \dfrac{n}{1+d}$. Since $\dfrac{n}{1+d} = \dfrac{n^2}{n + 2|E|}$, this rewrites as follows: The graph $G$ has an independent set of size $\geq \dfrac{n^2}{n + 2|E|}$. This proves Corollary 0.37. $\square$

Now, we can prove Theorem 0.36 (thus solving Exercise 7):

*Proof of Theorem 0.36.* Let $V = \operatorname{V}(G)$.

Let $\overline{G}$ be the simple graph $(V, \mathcal{P}_2(V) \setminus \operatorname{E}(G))$. Note that this graph $\overline{G}$ is called the *complement graph* of $G$.

Note that $n = \left| \underbrace{\mathrm{V}(G)}_{=V} \right| = |V|$. But $\mathrm{E}(G) \subseteq \mathcal{P}_2(V)$. Thus,

$$|\mathcal{P}_2(V) \setminus \mathrm{E}(G)| = \underbrace{|\mathcal{P}_2(V)|}_{\substack{=\binom{|V|}{2}=\binom{n}{2} \\ (\text{since } |V|=n)}} - \underbrace{|\mathrm{E}(G)|}_{>\frac{r-1}{r}\cdot\frac{n^2}{2}} < \underbrace{\binom{n}{2}}_{=\frac{n(n-1)}{2}} - \frac{r-1}{r} \cdot \frac{n^2}{2} = \frac{n(n-1)}{2} - \frac{r-1}{r} \cdot \frac{n^2}{2} = \frac{n(n-r)}{2r}.$$

Hence,

$$n + 2 \underbrace{|\mathcal{P}_2(V) \setminus \mathrm{E}(G)|}_{<\frac{n(n-r)}{2r}} < n + 2 \cdot \frac{n(n-r)}{2r} = n^2/r. \tag{31}$$

Corollary 0.37 (applied to $\overline{G}$ and $\mathcal{P}_2(V) \setminus \mathrm{E}(G)$ instead of $G$ and $E$) shows that $\overline{G}$ has an independent set of size $\geq \frac{n^2}{n + 2|\mathcal{P}_2(V) \setminus \mathrm{E}(G)|}$. Fix such an independent set, and denote it by $I$.

The set $I$ has size $\geq \frac{n^2}{n + 2|\mathcal{P}_2(V) \setminus \mathrm{E}(G)|}$. In other words,

$$|I| \geq \frac{n^2}{n + 2|\mathcal{P}_2(V) \setminus \mathrm{E}(G)|} > \frac{n^2}{n^2/r} \qquad (\text{by (31)})$$
$$= r.$$

Thus, $|I| \geq r + 1$ (since both $|I|$ and $r$ are integers). Hence, there exist $r + 1$ distinct elements of $I$. Fix such $r + 1$ distinct elements, and denote them by $i_1, i_2, \ldots, i_{r+1}$.

For any two distinct elements $a$ and $b$ of $\{1, 2, \ldots, r + 1\}$, the vertices $i_a$ and $i_b$ of $G$ are adjacent[36]. In other words, the $r + 1$ vertices $i_1, i_2, \ldots, i_{r+1}$ of $G$ are mutually

[36] *Proof.* Let $a$ and $b$ be two distinct elements of $\{1, 2, \ldots, r + 1\}$. We must show that the vertices $i_a$ and $i_b$ of $G$ are adjacent.

Recall that $i_1, i_2, \ldots, i_{r+1}$ are $r + 1$ elements of $I$. Thus, $i_a \in I$ and $i_b \in I$.

Recall that the elements $i_1, i_2, \ldots, i_{r+1}$ are distinct. Thus, $i_a \neq i_b$ (since $a$ and $b$ are distinct). Hence, $i_a$ and $i_b$ are two distinct elements of $I$ (since $i_a \in I$ and $i_b \in I$).

The set $I$ is an independent set of $\overline{G}$. In other words, $I$ is a subset of $V$ having the property that no two distinct elements of $I$ are adjacent with respect to $\overline{G}$ (by the definition of an "independent set").

The elements $i_a$ and $i_b$ are two distinct elements of $I$. Hence, $i_a$ and $i_b$ are not adjacent with respect to $\overline{G}$ (since no two distinct elements of $I$ are adjacent with respect to $\overline{G}$). In other words, $\{i_a, i_b\} \notin \mathrm{E}(\overline{G})$.

adjacent. Hence, there exist $r+1$ distinct vertices of $G$ that are mutually adjacent (namely, the $r+1$ distinct vertices $i_1, i_2, \ldots, i_{r+1}$). This proves Theorem 0.36. $\square$

On the other hand, $i_a$ and $i_b$ are two distinct elements of $I$. Hence, $\{i_a, i_b\}$ is a 2-element subset of $I$. Thus, $\{i_a, i_b\} \in \mathcal{P}_2(I) \subseteq \mathcal{P}_2(V)$ (since $I \subseteq V$). Combining this with $\{i_a, i_b\} \notin \operatorname{E}(\overline{G})$, we obtain $\{i_a, i_b\} \in \mathcal{P}_2(V) \setminus \operatorname{E}(\overline{G})$. But from $\overline{G} = (V, \mathcal{P}_2(V) \setminus \operatorname{E}(G))$, we obtain $\operatorname{E}(\overline{G}) = \mathcal{P}_2(V) \setminus \operatorname{E}(G)$. Hence, $\mathcal{P}_2(V) \setminus \operatorname{E}(\overline{G}) = \mathcal{P}_2(V) \setminus (\mathcal{P}_2(V) \setminus \operatorname{E}(G)) = \operatorname{E}(G)$ (since $\operatorname{E}(G) \subseteq \mathcal{P}_2(V)$). Thus, $\{i_a, i_b\} \in \mathcal{P}_2(V) \setminus \operatorname{E}(\overline{G}) = \operatorname{E}(G)$. In other words, the vertices $i_a$ and $i_b$ of $G$ are adjacent. This completes our proof.

## 0.9. Exercise 8: bijective proofs for Vandermonde-like identities?

**Exercise 8. Extra credit:**

In this exercise, "number" means (e.g.) a real number. (Feel free to restrict yourself to positive integers if it helps you. The most general interpretation would be "element of a commutative ring", but you don't need to work in this generality.)

For any $n$ numbers $x_1, x_2, \ldots, x_n$, we define $v(x_1, x_2, \ldots, x_n)$ to be the number

$$\prod_{1 \leq i < j \leq n} (x_j - x_i) = \det\left(\left(x_j^{i-1}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right).$$

Let $x_1, x_2, \ldots, x_n$ be $n$ numbers. Let $t$ be a further number. Prove at least one of the following facts combinatorially (i.e., without using any properties of the determinant other than its definition as a sum over permutations):

**(a)** We have

$$\sum_{k=1}^{n} v(x_1, x_2, \ldots, x_{k-1}, x_k + t, x_{k+1}, x_{k+2}, \ldots, x_n) = nv(x_1, x_2, \ldots, x_n).$$

**(b)** For each $m \in \{0, 1, \ldots, n-1\}$, we have

$$\sum_{k=1}^{n} x_k^m v(x_1, x_2, \ldots, x_{k-1}, t, x_{k+1}, x_{k+2}, \ldots, x_n) = t^m v(x_1, x_2, \ldots, x_n).$$

**(c)** We have

$$\begin{aligned}
&\sum_{k=1}^{n} x_k v(x_1, x_2, \ldots, x_{k-1}, x_k + t, x_{k+1}, x_{k+2}, \ldots, x_n) \\
&= \left(\binom{n}{2} t + \sum_{k=1}^{n} x_k\right) v(x_1, x_2, \ldots, x_n).
\end{aligned}$$

*Hints to Exercise 8.* I hope that at least some parts of Exercise 8 have combinatorial proofs similar to Gessel's proof of the Vandermonde determinant (as in lecture 8). Noone solved any part of Exercise 8, though. Feel free to continue trying (the extra credit is still available, and, more importantly, a chance to get your new proof published).

Exercise 8 **(a)** can be easily proven if you know some abstract algebra (specifically, properties of polynomial rings in several variables). Let me sketch the argument. (This is modelled on a well-known proof of the Vandermonde determinant identity $\prod_{1 \leq i < j \leq n} (x_j - x_i) = \det\left(\left(x_j^{i-1}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right)$ itself, which can be found, e.g., in

[Garret10, §17].)

First of all, let $x_1, x_2, \ldots, x_n$ and $t$ no longer be numbers; but instead, let them be indeterminates. Thus, we are working in the polynomial ring $\mathbb{Z}[x_1, x_2, \ldots, x_n, t]$. Let $P$ be the polynomial $\sum\limits_{k=1}^{n} v(x_1, x_2, \ldots, x_{k-1}, x_k + t, x_{k+1}, x_{k+2}, \ldots, x_n) - nv(x_1, x_2, \ldots, x_n)$. This polynomial is homogeneous of degree $n(n-1)/2$ (check this!). But the polynomial $P$ becomes 0 whenever two of the variables $x_1, x_2, \ldots, x_n$ are equal[37]. Hence, this polynomial $P$ is divisible by $x_i - x_j$ for every pair $(i, j)$ satisfying $1 \leq i < j \leq n$ (by some basic abstract algebra). Consequently, the polynomial $P$ is divisible by the product $\prod\limits_{1 \leq i < j \leq n} (x_i - x_j)$ [38]. But this product is homogeneous of the same degree $n(n-1)/2$ as the polynomial $P$ itself (since it has $n(n-1)/2$ factors, each of them linear). Hence, the quotient $\dfrac{P}{\prod\limits_{1 \leq i < j \leq n} (x_i - x_j)}$ must be a homogeneous polynomial of degree 0, i.e., a constant. We thus merely need to show that this constant is 0. The easiest way to do so is to evaluate this polynomial at $t = 0$. The result is clearly 0. Since it is a constant, it thus is 0 whatever $t$ is. And so Exercise 8 **(a)** is proven (up to all the steps I have omitted).

Exercise 8 **(b)** can be solved in a similar manner – actually, even easier. Let $Q$ be the polynomial $\sum\limits_{k=1}^{n} x_k^m v(x_1, x_2, \ldots, x_{k-1}, t, x_{k+1}, x_{k+2}, \ldots, x_n) - t^m v(x_1, x_2, \ldots, x_n)$. Then, it is easy to see that the polynomial $Q$ vanishes whenever $t = x_i$ for any $i$. Thus, this polynomial $Q$ is divisible by all of the $n$ linear polynomials $t - x_i$ with $i \in \{1, 2, \ldots, n\}$. Hence, this polynomial $Q$ must be divisible by their product $\prod\limits_{i=1}^{n} (t - x_i)$. However, this product has degree $n$ when considered as a polynomial in $t$ (treating $x_1, x_2, \ldots, x_n$ as constants), whereas the polynomial $Q$ has degree $\leq m \leq n - 1$ (again, when considered as a polynomial in $t$). The only way a polynomial can be divisible by a polynomial of larger degree is when the former polynomial is 0. Hence, $Q$ must be 0.

Actually, Exercise 8 **(b)** is a famous formula in disguise – namely, the *Lagrange interpolation formula*. Namely, if we divide both sides of the claim by $v(x_1, x_2, \ldots, x_n)$

[37] Check this – it's not immediately obvious. The main observation to make is that the polynomial $v(x_1, x_2, \ldots, x_n)$ gets multiplied by $-1$ whenever two of the variables $x_1, x_2, \ldots, x_n$ get interchanged, and becomes 0 whenever two of the variables $x_1, x_2, \ldots, x_n$ are equal.

[38] This step is nontrivial. We are claiming that if a polynomial is divisible by $x_i - x_j$ for every pair $(i, j)$ satisfying $1 \leq i < j \leq n$, then this polynomial must also be divisible by the product $\prod\limits_{1 \leq i < j \leq n} (x_i - x_j)$. One way to prove this is by using the fact (often proven in the first semester of abstract algebra) that the polynomial ring $\mathbb{Z}[x_1, x_2, \ldots, x_n, t]$ is a unique factorization domain. There exist other arguments, but these too use at least the language of rings.

(assuming that $x_1, x_2, \ldots, x_n$ are distinct[39]), then we obtain

$$\sum_{k=1}^{n} x_k^m \frac{\prod_{i \neq k} (t - x_i)}{\prod_{i \neq k} (x_k - x_i)} = t^m.$$

More generally, if $f$ is any polynomial (in one variable) of degree $\leq n-1$, then

$$\sum_{k=1}^{n} f(x_k) \frac{\prod_{i \neq k} (t - x_i)}{\prod_{i \neq k} (x_k - x_i)} = f(t).$$

This is often stated as follows: If $x_1, x_2, \ldots, x_n$ are $n$ distinct real numbers, and $y_1, y_2, \ldots, y_n$ are $n$ further real numbers, then there is a unique polynomial $g = g(t)$ of degree $\leq n-1$ satisfying $g(x_i) = y_i$ for all $i$, and this polynomial $g$ is

$$\sum_{k=1}^{n} y_k \frac{\prod_{i \neq k} (t - x_i)}{\prod_{i \neq k} (x_k - x_i)}.$$

See, for example, `http://www.math.uconn.edu/~leykekhman/courses/MATH3795/Lectures/Lecture_14_poly_interp.pdf` .

Exercise 8 **(c)** is a harder variant of Exercise 8 **(a)**; this time the quotient of the polynomials has degree 1, which makes it harder to identify it (evaluating it at $t = 0$ is not enough). It is a step in one of the classical proofs of the hook-length formula in algebraic combinatorics – see [Fulton97, §4.3, Exercise 10] or [Uecker16, Lemma 4.13] or [GlaNg04, Lemma 2][40].

Meanwhile, elementary (but still not combinatorial) proofs also exist:

- Exercise 8 **(a)** is solved in [Grinbe16, Proposition 7.192].
- Exercise 8 **(b)** is solved in [Grinbe16, Proposition 7.194].
- Exercise 8 **(c)** is solved in [Grinbe16, Exercise 6.34].

□

# References

[AGJJ09] S. Akbari, M. Ghanbari, S. Jahanbekam, M. Jamaali, *List coloring of graphs having cycles of length divisible by a given number*, Discrete Mathematics **309**, Issue 3, 28 February 2009, pp. 613–614.

[39] This is a harmless assumption that you can make whenever you want to prove an identity like this... but proving that this is so takes a bit of abstract algebra, too.

[40] To be fully precise: [GlaNg04, Lemma 2] is obtained from Exercise 8 **(c)** by applying it to $n = m$, $t = -1$ and $x_i = z_i$. Conversely, Exercise 8 **(c)** follows from [GlaNg04, Lemma 2] (applied to $m = n$ and $z_i = -x_i/t$). Hence, the two facts are equivalent.

[Fulton97] William Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, Cambridge University Press 1997.

[Garret10] Paul Garrett, *Abstract Algebra*, ca. 2010,
http://www-users.math.umn.edu/~garrett/m/algebra/

[GlaNg04] Kenneth Glass, Chi-Keung Ng, *A Simple Proof of the Hook Length Formula*, The American Mathematical Monthly, Vol. 111, No. 8 (Oct., 2004), pp. 700–704.
https://sites.math.washington.edu/~billey/classes/581.fall.2015/articles/glass.ng.2004.pdf

[Grinbe16] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 10 January 2019.
http://www.cip.ifi.lmu.de/~grinberg/primes2015/sols.pdf

[Quinla17] Rachel Quinlan, *Advanced Linear Algebra MA500-1: Lecture Notes Semester 1 2015-2016*, version 14 March 2017.
http://www.maths.nuigalway.ie/~rquinlan/teaching/

[Uecker16] Torsten Ueckerdt and Stefan Walzer, *Lecture Notes Combinatorics*, version 19 July 2016.
http://www.math.kit.edu/iag6/lehre/co2015s/media/script.pdf

[Wang17] Jian Wang, *Maximum Linear Forests in Trees*, arXiv preprint arXiv:1701.01953v2.