Math 5707 Spring 2017 (Darij Grinberg): midterm 1
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.

0.2. Exercise 1: a from-dominating set from a Hamiltonian path

Exercise 1. Let $D = (V, A)$ be a digraph. A *from-dominating set* of D shall mean a subset S of V such that for each vertex $v \in V \setminus S$, there exists at least one arc $uv \in A$ with $u \in S$.

Assume that D has a Hamiltonian path. Prove that D has a from-dominating set of size $\leq \dfrac{|V|+1}{2}$.

Solution sketch to Exercise 1. We have assumed that D has a Hamiltonian path. Fix such a path, and denote it by $(v_1, v_2, \ldots, v_n)$.

Each vertex of D appears exactly once in the path $(v_1, v_2, \ldots, v_n)$ (since $(v_1, v_2, \ldots, v_n)$ is a Hamiltonian path). In other words, each element of V appears exactly once in the path $(v_1, v_2, \ldots, v_n)$. Hence, $|V| = n$ and $V = \{v_1, v_2, \ldots, v_n\}$.

Define a subset S of V by

$$S = \{v_i \mid i \in \{1, 2, \ldots, n\} \text{ is odd}\} = \{v_1, v_3, v_5, \ldots\}.$$

Then,[1] $|S| = \left\lfloor \dfrac{n+1}{2} \right\rfloor \leq \dfrac{n+1}{2} = \dfrac{|V|+1}{2}$ (since $n = |V|$). Hence, the set S has size $\leq \dfrac{|V|+1}{2}$. It thus merely remains to prove that S is from-dominating.

According to the definition of "from-dominating", this means proving that for each vertex $v \in V \setminus S$, there exists at least one arc $uv \in A$ with $u \in S$. So let us prove this now.

Let $v \in V \setminus S$. Thus,

$$\begin{aligned} v \in V \setminus S &= \{v_1, v_2, \ldots, v_n\} \setminus \{v_1, v_3, v_5, \ldots\} \\ &\qquad (\text{since } V = \{v_1, v_2, \ldots, v_n\} \text{ and } S = \{v_1, v_3, v_5, \ldots\}) \\ &= \{v_2, v_4, v_6, \ldots\} = \{v_i \mid i \in \{1, 2, \ldots, n\} \text{ is even}\}. \end{aligned}$$

In other words, $v = v_i$ for some even $i \in \{1, 2, \ldots, n\}$. Consider this i. The integer i is even and positive; hence, the integer $i - 1$ is odd and positive. Hence, $i - 1 \in \{1, 2, \ldots, n-1\}$ (since $i - 1 < i \leq n$), so that $i - 1 \in \{1, 2, \ldots, n\}$. Thus, $v_{i-1} \in \{v_1, v_3, v_5, \ldots\}$ (since $i - 1$ is odd). Hence, $v_{i-1} \in \{v_1, v_3, v_5, \ldots\} = S$.

Furthermore, recall that $(v_1, v_2, \ldots, v_n)$ is a path in D. Thus, $v_j v_{j+1}$ is an arc of D for each $j \in \{1, 2, \ldots, n-1\}$. Applying this to $j = i - 1$, we conclude that $v_{i-1} v_i$ is an arc of D. In other words, $v_{i-1} v_i \in A$. Since $v_i = v$, this rewrites as $v_{i-1} v \in A$. Hence, there exists at least one arc $uv \in A$ with $u \in S$ (namely, the arc $v_{i-1} v$ with $u = v_{i-1}$). This is exactly what we wanted to prove. Hence, we have shown that S is from-dominating. □

0.3. Exercise 2: Hamiltonian paths of a line graph

Exercise 2. Let $G = (V, E)$ be a simple graph. The *line graph* $L(G)$ is defined as the simple graph (E, F), where

$$F = \{\{e_1, e_2\} \in \mathcal{P}_2(E) \mid e_1 \cap e_2 \neq \varnothing\}.$$

(In other words, $L(G)$ is the graph whose **vertices** are the **edges** of G, and in which two vertices e_1 and e_2 are adjacent if and only if the edges e_1 and e_2 of G share a common vertex.)

Assume that $|V| > 1$.

(a) If G has a Hamiltonian path, then prove that $L(G)$ has a Hamiltonian path.

(b) If G has a Eulerian walk, then prove that $L(G)$ has a Hamiltonian path.

Hints to Exercise 2. **(a)** Let $(v_0, v_1, \ldots, v_n)$ be a Hamiltonian path in G. Let $e_1, e_2, \ldots, e_n$ be the edges of this path (so that $e_i = v_{i-1} v_i$ for each i). Then, $\mathbf{p} = (e_1, e_2, \ldots, e_n)$ is a path in $L(G)$. This path is not necessarily a Hamiltonian path; but we can turn it into a Hamiltonian path by the following procedure:

[1] Here, we are using the following notation: If x is a real number, then $\lfloor x \rfloor$ denotes the largest integer that is smaller or equal to x. For example, $\lfloor 2.8 \rfloor = 2$ and $\lfloor 3 \rfloor = 3$ and $\lfloor -1.6 \rfloor = -2$.

- Insert all edges of G that contain v_0 and are not already in the path $\mathbf{p}$ into $\mathbf{p}$, placing them at the beginning of $\mathbf{p}$ (just before e_1).
- Insert all edges of G that contain v_1 and are not already in the path $\mathbf{p}$ into $\mathbf{p}$, placing them between e_1 and e_2.
- Insert all edges of G that contain v_2 and are not already in the path $\mathbf{p}$ into $\mathbf{p}$, placing them between e_2 and e_3.
- And so on, until all edges of G have been inserted.

It is easy to check that the result is a Hamiltonian path in $L(G)$.

(b) Let $e_1, e_2, \ldots, e_m$ be the edges of an Eulerian walk in G. Then, $(e_1, e_2, \ldots, e_m)$ is a Hamiltonian path in $L(G)$.

[*Remark:* It is also true that if G has an Eulerian circuit, then $L(G)$ has an Eulerian circuit. To prove this, show that $L(G)$ is connected and that each vertex of $L(G)$ has even degree.] □

0.4. Exercise 3: a condition for a digraph to have a cycle

Exercise 3. Let $D = (V, A)$ be a digraph with $|V| > 0$. Assume that each vertex $v \in V$ satisfies $\deg^- v > 0$. Prove that D has at least one cycle.

(Keep in mind that a length-1 circuit (v, v) counts as a cycle when A contains the loop (v, v).)

Hints to Exercise 3. Fix a longest path $(v_0, v_1, \ldots, v_k)$ in D. There is at least one arc with target v_0 (since $\deg^-(v_0) > 0$). Let v_{-1} be the source of this arc. Then, $(v_{-1}, v_0, v_1, \ldots, v_k)$ is a walk, but not a path (since $(v_0, v_1, \ldots, v_k)$ is a longest path). Hence, $v_{-1} = v_i$ for some $i \in \{0, 1, \ldots, k\}$. Fix the **smallest** such i. Then, $(v_{-1}, v_0, \ldots, v_i)$ is a cycle.

[*Remark:* This is very similar to Lemma 0.2 in the solutions to homework set 2.] □

0.5. Exercise 4: a condition for a multigraph to have a cycle

Recall that the degree $\deg v$ of a vertex v of a multigraph G is defined as the number of all edges of G containing v.

Exercise 4. Let G be a multigraph with at least one edge. Assume that each vertex of G has even degree. Prove that G has a cycle.

Hints to Exercise 4. Fix a longest path $(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$ in G. There is at least one edge containing v_0 (namely, e_1). Thus, the number $\deg(v_0)$ is positive. This number $\deg(v_0)$ is furthermore even (since each vertex of G has even degree). Hence, this number $\deg(v_0)$ is ≥ 2 (because it is even and positive). In other words, there are at least two edges containing v_0. Thus, there is at least one edge containing v_0 that is distinct from e_1. Denote this edge by e_0, and let v_{-1} be its other endpoint. Then, $(v_{-1}, e_0, v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$ is a walk, but not a path (since $(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k)$ is a longest path). Hence, $v_{-1} = v_i$ for some $i \in \{0, 1, \ldots, k\}$. Fix the **smallest** such i. Then, $(v_{-1}, e_0, v_0, e_1, v_1, \ldots, e_i, v_i)$ is a cycle.

[*Remark:* This is (so to speak) the undirected version of Lemma 0.2 in the solutions to homework set 2.] □

0.6. Exercise 5: a coloring where neighbors shun equal colors

Exercise 5. Let $k \in \mathbb{N}$. Let $p_1, p_2, \ldots, p_k$ be k nonnegative real numbers such that $p_1 + p_2 + \cdots + p_k \geq 1$.

Let $G = (V, E)$ be a simple graph. A *k-coloring* of G shall mean a map $f : V \to \{1, 2, \ldots, k\}$.

Prove that there exists a k-coloring f of G with the following property: For each vertex $v \in V$, at most $p_{f(v)} \deg v$ neighbors of v have the same color as v. Here, the *color* of a vertex $w \in V$ (under the k-coloring f) means the value $f(w)$.

Hints to Exercise 5. This is a generalization of Exercise 1 on homework set 0. Indeed, you obtain the latter exercise if you set $k = 2$, $p_1 = 1/2$ and $p_2 = 1/2$.

To solve Exercise 5, we can generalize the solution of Exercise 1 on homework set 0. Four changes need to be made:

1. We need to deal with the cases $k \leq 1$ separately.
2. In the algorithm, we need to explain how precisely to "flip" the color $f(v)$ of the vertex v. (Indeed, for $k > 2$, there is more than one possibility.)
3. We need to change the definition of "enmity".
4. We need to prove that the enmity cannot keep decreasing forever.

I leave change 1 to the reader (the cases $k \leq 1$ are essentially trivial).

As for change 2: We have some $v \in V$ such that more than $p_{f(v)} \deg v$ among the neighbors of v have the same color of v. Then, there exists **some** $i \in \{1, 2, \ldots, k\}$ such that **at most** $p_i \deg v$ among the neighbors of v have the color i (because otherwise, by summing over all i, we conclude that v has more than $\sum_{i=1}^{k} p_i \deg v = \underbrace{(p_1 + p_2 + \ldots + p_k)}_{\geq 1} \deg v \geq \deg v$ neighbors in total; but this is absurd)[2]. Fix such an i, and replace the color $f(v)$ of v by this i.

What about change 3? We formerly defined the enmity of a coloring f to be the number of f-monochromatic edges. This definition no longer works. However,

[2] You can even find an i such that **fewer than** $p_i \deg v$ among the neighbors of v have the color i. But that's not necessary for us.

what works is the following subtler definition: For each $i \in \{1, 2, \ldots, k\}$, an edge e of G is said to be (f, i)-*chromatic* if the two endpoints of e both have color i in the k-coloring f. The *enmity* of a k-coloring f is now defined as the sum

$$\sum_{i=1}^{k} \frac{1}{p_i} \left(\text{the number of all } (f, i)\text{-chromatic edges}\right).$$

This definition requires a minor technical fix: It only works if all p_i are positive (i.e., nonzero). Fortunately, we can WLOG assume that all p_i are positive (indeed, if some p_i is zero, then we can discard this p_i, and correspondingly agree to never use the color i in our coloring). Proving that the enmity decreases throughout the algorithm is rather easy (a straightforward modification of the original argument).

Change 4 is simple but subtle; it is really easy to miss its importance. In the solution of Exercise 1 on homework set 0, we just argued that the enmity of a 2-coloring cannot keep decreasing forever because it is a nonnegative integer. However, in our more general setting, the enmity no longer is a nonnegative integer, and thus one could imagine it decreasing indefinitely (e.g., from 1 to $1/2$, then on to $1/3$, then on to $1/4$, etc.). So we need a new argument. Fortunately, there is an easy one: There are only finitely many k-colorings of G (namely, $k^{|V|}$ many). Hence, there are only finitely many values that the enmity of a k-coloring of G can take. Hence, the enmity cannot keep decreasing forever (because that would force it to take infinitely many different values along the way). So we are done. $\square$

0.7. Exercise 6: adding edges to get an Eulerian circuit

Exercise 6. Let G be a connected multigraph. Let m be the number of vertices of G that have odd degree. Prove that we can add $m/2$ new edges to G in such a way that the resulting multigraph will have an Eulerian circuit. (It is allowed to add an edge even if there is already an edge between the same two vertices.)

Hints to Exercise 6. Proposition 2.5.8 in the lecture notes says that the number of vertices of G having odd degree is even[3]. In other words, m is even. Let $v_1, v_2, \ldots, v_m$ be the m vertices of G that have odd degree.

Now, let us add $m/2$ new edges $e_1, e_2, \ldots, e_{m/2}$ to G, where each e_i has $\phi(e_i) = \{v_{2i-1}, v_{2i}\}$. (This is well-defined, since m is even.) The resulting multigraph is clearly connected (since G was connected), and has the property that each of its vertices has even degree[4]. Hence, by the Euler-Hierholzer theorem, this new graph has an Eulerian circuit. $\square$

[3] To be fully precise, Proposition 2.5.8 in the lecture notes only states this fact for simple graphs, not for multigraphs. But the proof for multigraphs is almost the same. (The only difference is that "$v \in e$" must be replaced by "$v \in \phi(e)$".)

[4] *Proof.* Let us see how the degrees of the vertices of our multigraph have been affected by adding the $m/2$ new edges $e_1, e_2, \ldots, e_{m/2}$:

- The degrees of the vertices $v_1, v_2, \ldots, v_m$ have been incremented by 1 when we added our $m/2$ new edges. This caused these degrees to become even (because back in the original

0.8. Exercise 7: how large can the perimeter of a triangle on a graph be?

0.8.1. Distances in a graph

If u and v are two vertices of a simple graph G, then $d(u,v)$ denotes the *distance* between u and 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 ∞.

We observe the following simple facts:

Lemma 0.1. Let u and v be two vertices of a connected simple graph $G = (V, E)$. Then, $d(u,v) \leq |V| - 1$.

Proof of Lemma 0.1. The simple graph G is connected. Hence, there exists a walk from u to v in G. Let k be the length of this walk. Therefore, there exists a walk from u to v in G having length $\leq k$ (namely, the walk we have just constructed). Hence, Corollary 2.8.10 in the lecture notes shows that there exists a path from u to v having length $\leq k$. Let $(u_0, u_1, \ldots, u_g)$ be this path. Then, the vertices $u_0, u_1, \ldots, u_g$ are pairwise distinct (since $(u_0, u_1, \ldots, u_g)$ is a path). Hence, $|\{u_0, u_1, \ldots, u_g\}| = g + 1$. But from $\{u_0, u_1, \ldots, u_g\} \subseteq V$, we obtain $|\{u_0, u_1, \ldots, u_g\}| \leq |V|$. Thus, $g + 1 = |\{u_0, u_1, \ldots, u_g\}| \leq |V|$. Hence, $g \leq |V| - 1$.

Now, there exists a path from u to v having length g (namely, the path $(u_0, u_1, \ldots, u_g)$). Hence, the minimum length of a path from u to v is $\leq g$. But this minimum length is $d(u,v)$ (by the definition of $d(u,v)$). Hence, we obtain $d(u,v) \leq g \leq |V| - 1$. This proves Lemma 0.1. □

Lemma 0.1 shows that if u and v are two vertices of a connected simple graph G, then $d(u,v)$ is an actual integer (as opposed to ∞).

Lemma 0.2. Let u and v be two vertices of a simple graph G. Let $k \in \mathbb{N}$. If there exists a walk from u to v in G having length k, then $d(u,v) \leq k$.

Proof of Lemma 0.2. We assumed that there exists a walk from u to v in G having length k. Therefore, Corollary 2.8.10 in the lecture notes shows that there exists a path from u to v having length $\leq k$. Therefore, the minimum length of a path from u to v is $\leq k$. But this minimum length is $d(u,v)$ (by the definition of $d(u,v)$). Hence, we obtain $d(u,v) \leq k$. This proves Lemma 0.2. □

multigraph G, they were odd).

- The degrees of all other vertices have not changed when we added our $m/2$ new edges (because none of these new edges contains any of the other vertices). Hence, these degrees have remained even (because they were even in the original multigraph G).

Thus, in the resulting multigraph, the degrees of **all** vertices have become even.

Lemma 0.3. Let $G = (V, E)$ be a simple graph.

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

(b) Each $u \in V$ and $v \in V$ satisfy $d(u, v) = d(v, u)$.

(c) 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 ∞ to be greater than any integer, and we understand $\infty + m$ to be ∞ whenever $m \in \mathbb{Z}$.)

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

Proof of Lemma 0.3 (sketched). Parts **(a)**, **(b)** and **(d)** of Lemma 0.3 are easy to check, and thus left to the reader.

(c) We WLOG assume that none of the two distances $d(u, v)$ and $d(v, w)$ is ∞ (since otherwise, Lemma 0.3 **(c)** holds for obvious reasons).

If there was no path from u to v, then $d(u, v)$ would be ∞ (by the definition of $d(u, v)$), which would contradict the fact that none of the two distances $d(u, v)$ and $d(v, w)$ is ∞. Hence, there must exist at least one path from u to v. Thus, $d(u, v)$ is the minimum length of a path from u to v (by the definition of $d(u, v)$). Hence, there exists a path from u to v having length $d(u, v)$. Fix such a path, and denote it by $\mathbf{p} = (p_0, p_1, \ldots, p_g)$. Hence, (the length of the path $\mathbf{p}$) $= g$. Therefore, $g =$ (the length of the path $\mathbf{p}$) $= d(u, v)$ (since $\mathbf{p}$ has length $d(u, v)$).

We have shown that there exists a path from u to v having length $d(u, v)$. Similarly, we can show that there exists a path from v to w having length $d(v, w)$. Fix such a path, and denote it by $\mathbf{q} = (q_0, q_1, \ldots, q_h)$. Hence, (the length of the path $\mathbf{q}$) $= h$. Therefore, $h =$ (the length of the path $\mathbf{q}$) $= d(v, w)$ (since $\mathbf{h}$ has length $d(v, w)$).

Now, $(p_0, p_1, \ldots, p_g)$ is a path from u to v. Hence, $p_0 = u$ and $p_g = v$. Also, $(q_0, q_1, \ldots, q_h)$ is a path from v to w. Hence, $q_0 = v$ and $q_h = w$.

The lists $(p_0, p_1, \ldots, p_g)$ and $(q_0, q_1, \ldots, q_h)$ are paths, and therefore walks. The ending point p_g of the first of these two walks is the starting point q_0 of the second (because $p_g = v = q_0$). Hence, we can combine these two walks to a walk $(p_0, p_1, \ldots, p_g, q_1, q_2, \ldots, q_h) = (p_0, p_1, \ldots, p_{g-1}, q_0, q_1, \ldots, q_h)$. This latter walk has length $g + h$, and is a walk from u to w (since $p_0 = u$ and $q_h = w$). Thus, there exists a walk from u to w having length $g + h$ (namely, the walk that we have just constructed). Hence, Lemma 0.2 (applied to w and $g + h$ instead of v and k) shows that $d(u, w) \leq g + h = d(u, v) + d(v, w)$ (since $g = d(u, v)$ and $h = d(v, w)$). In other words, $d(u, v) + d(v, w) \geq d(u, w)$. This proves Lemma 0.3 **(c)**. $\square$

Lemma 0.3 **(c)** is known as the *triangle inequality* for distances on a graph. Of course, this is due to its similarity to the well-known triangle inequality in Euclidean geometry. In fact, this similarity goes deeper: Lemma 0.1 shows that if $G = (V, E)$ is a connected simple graph, then the definition of the distance $d(u, v)$ for each pair $(u, v) \in V \times V$ gives rise to a well-defined map $d : V \times V \to \mathbb{N}$. Lemma 0.3 shows that this map d is a metric. We shall not use this in the following, but it is a useful fact to keep in one's mind.

0.8.2. Statement of the exercise

Exercise 7. Let a, b and c be three vertices of a connected simple graph $G = (V, E)$. Prove that $d(b, c) + d(c, a) + d(a, b) \leq 2|V| - 2$.

0.8.3. First solution

Solution sketch to Exercise 7. The following solution was found by Jiali Huang and Nicholas Rancourt.

Fix some path $\mathbf{z} = (z_0, z_1, \ldots, z_g)$ from a to b having minimum length. Then, the length of $\mathbf{z}$ is $d(a, b)$ (since $d(a, b)$ is defined as the minimum length of a path from a to b). Hence,

$$d(a, b) = (\text{the length of the path } \mathbf{z}) = g \qquad (1)$$

(since $\mathbf{z} = (z_0, z_1, \ldots, z_g)$).

Since $(z_0, z_1, \ldots, z_g)$ is a path from a to b, we have $z_0 = a$ and $z_g = b$.

Since $(z_0, z_1, \ldots, z_g)$ is a path, the $g + 1$ vertices $z_0, z_1, \ldots, z_g$ are distinct. Hence, $|\{z_0, z_1, \ldots, z_g\}| = g + 1$.

Now, pick an element $i \in \{0, 1, \ldots, g\}$ for which the number $d(c, z_i)$ is minimum. Hence,

$$d(c, z_i) \leq d(c, z_j) \qquad \text{for each } j \in \{0, 1, \ldots, g\}. \qquad (2)$$

Fix some path $\mathbf{t} = (t_0, t_1, \ldots, t_h)$ from c to z_i having minimum length. Then, the length of $\mathbf{t}$ is $d(c, z_i)$ (since $d(c, z_i)$ is defined as the minimum length of a path from c to z_i). Hence,

$$d(c, z_i) = (\text{the length of the path } \mathbf{t}) = h$$

(since $\mathbf{t} = (t_0, t_1, \ldots, t_h)$).

Since $(t_0, t_1, \ldots, t_h)$ is a path from c to z_i, we have $t_0 = c$ and $t_h = z_i$.

Since $(t_0, t_1, \ldots, t_h)$ is a path, the $h + 1$ vertices $t_0, t_1, \ldots, t_h$ are distinct. In particular, the h vertices $t_0, t_1, \ldots, t_{h-1}$ are distinct. Hence, $|\{t_0, t_1, \ldots, t_{h-1}\}| = h$.

Now, there is a walk from c to a in G having length $h + i$ [5]. Hence, Lemma 0.2 (applied to $u = c$, $v = a$ and $k = h + i$) shows that

$$d(c, a) \leq h + i. \qquad (3)$$

On the other hand, there is a walk from b to c in G having length $(g - i) + h$ [6]. Hence, Lemma 0.2 (applied to $u = b$, $v = c$ and $k = (g - i) + h$) shows that

$$d(b, c) \leq (g - i) + h. \qquad (4)$$

[5] *Proof.* We know that $(z_0, z_1, \ldots, z_g)$ is a path, thus a walk. Hence, $(z_0, z_1, \ldots, z_i)$ is a walk as well. Therefore, $(z_i, z_{i-1}, \ldots, z_0)$ is a walk (being the reversal of the walk $(z_0, z_1, \ldots, z_i)$). On the other hand, $(t_0, t_1, \ldots, t_h)$ is a walk. Since the ending point of the walk $(t_0, t_1, \ldots, t_h)$ is the starting point of the walk $(z_i, z_{i-1}, \ldots, z_0)$ (because $t_h = z_i$), we can combine these two walks, obtaining a new walk $(t_0, t_1, \ldots, t_{h-1}, z_i, z_{i-1}, \ldots, z_0)$. This new walk is a walk from c to a (since $t_0 = c$ and $z_0 = a$) and has length $h + i$. Hence, there is a walk from c to a in G having length $h + i$ (namely, the walk that we have just constructed).

[6] *Proof.* We know that $(z_0, z_1, \ldots, z_g)$ is a path, thus a walk. Hence, $(z_i, z_{i+1}, \ldots, z_g)$ is a walk as well. Therefore, $(z_g, z_{g-1}, \ldots, z_i)$ is a walk (being the reversal of the walk $(z_i, z_{i+1}, \ldots, z_g)$). On the other hand, $(t_0, t_1, \ldots, t_h)$ is a walk. Hence, $(t_h, t_{h-1}, \ldots, t_0)$ is a walk as well (being the

It is easy to see that the sets $\{t_0, t_1, \ldots, t_{h-1}\}$ and $\{z_0, z_1, \ldots, z_g\}$ are disjoint[7]. Hence, the sum of the sizes of these sets equals the size of their union. In other words,

$$\left|\{t_0, t_1, \ldots, t_{h-1}\}\right| + \left|\{z_0, z_1, \ldots, z_g\}\right| = \left|\underbrace{\{t_0, t_1, \ldots, t_{h-1}\} \cup \{z_0, z_1, \ldots, z_g\}}_{\subseteq V}\right| \leq |V|.$$

Since $\left|\{z_0, z_1, \ldots, z_g\}\right| = g+1$ and $\left|\{t_0, t_1, \ldots, t_{h-1}\}\right| = h$, this rewrites as $h + (g+1) \leq |V|$. Hence, $(g+h)+1 = h + (g+1) \leq |V|$, so that $g + h \leq |V| - 1$.

Adding together the two inequalities (4) and (3) and the equation (1), we obtain

$$d(b,c) + d(c,a) + d(a,b) \leq ((g-i)+h) + (h+i) + g = 2\left(\underbrace{g+h}_{\leq |V|-1}\right) \leq 2(|V|-1) = 2|V| - 2.$$

This solves the exercise. □

0.8.4. Second solution

Hints to a second solution of Exercise 7. The following solution is how I originally solved the exercise.

Let $\ell(\mathbf{w})$ denote the length of a walk $\mathbf{w}$.

Fix

reversal of the walk $(t_0, t_1, \ldots, t_h)$). Since the ending point of the walk $(z_g, z_{g-1}, \ldots, z_i)$ is the starting point of the walk $(t_h, t_{h-1}, \ldots, t_0)$ (because $z_i = t_h$), we can combine these two walks, obtaining a new walk $(z_g, z_{g-1}, \ldots, z_i, t_{h-1}, t_{h-2}, \ldots, t_0)$. This new walk is a walk from b to c (since $z_g = b$ and $t_0 = c$) and has length $(g-i)+h$. Hence, there is a walk from b to c in G having length $(g-i)+h$ (namely, the walk that we have just constructed).

[7] *Proof.* Let $v \in \{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\}$. We shall derive a contradiction.

We have $v \in \{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\} \subseteq \{t_0, t_1, \ldots, t_{h-1}\}$. Hence, $v = t_p$ for some $p \in \{0, 1, \ldots, h-1\}$. Consider this p.

We have $v \in \{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\} \subseteq \{z_0, z_1, \ldots, z_g\}$. Hence, $v = z_j$ for some $j \in \{0, 1, \ldots, g\}$. Consider this j.

Recall that $(t_0, t_1, \ldots, t_h)$ is a walk. Hence, $(t_0, t_1, \ldots, t_p)$ is a walk as well. This walk $(t_0, t_1, \ldots, t_p)$ is a walk from c to v (since $t_0 = c$ and $t_p = v$) and has length p. Hence, there is a walk from c to v in G having length p (namely, the walk $(t_0, t_1, \ldots, t_p)$). Consequently, Lemma 0.2 (applied to $u = c$ and $k = p$) shows that $d(c,v) \leq p \leq h-1$ (since $p \in \{0, 1, \ldots, h-1\}$).

But (2) yields $d(c, z_i) \leq d(c, z_j) = d(c, v)$ (since $z_j = v$). Thus, $d(c,v) \geq d(c, z_i) = h > h-1$. This contradicts $d(c,v) \leq h-1$.

Now, forget that we fixed v. Thus, we have obtained a contradiction for each $v \in \{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\}$. Hence, there exists no $v \in \{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\}$. Thus, $\{t_0, t_1, \ldots, t_{h-1}\} \cap \{z_0, z_1, \ldots, z_g\} = \varnothing$. In other words, the sets $\{t_0, t_1, \ldots, t_{h-1}\}$ and $\{z_0, z_1, \ldots, z_g\}$ are disjoint.

- a shortest path $\mathbf{x}$ from b to c;
- a shortest path $\mathbf{y}$ from c to a.
- a shortest path $\mathbf{z}$ from a to b.

Let X be the set of all vertices of $\mathbf{x}$. Similarly define Y and Z. Thus,

$$d(b,c) = \ell(\mathbf{x}) = |X| - 1 \qquad \text{and} \tag{5}$$
$$d(c,a) = \ell(\mathbf{y}) = |Y| - 1 \qquad \text{and} \tag{6}$$
$$d(a,b) = \ell(\mathbf{z}) = |Z| - 1. \tag{7}$$

Now, we claim that $|X \cap Y \cap Z| \leq 1$. Indeed, assume the contrary. Then, $|X \cap Y \cap Z| \geq 2$. Hence, there exist two distinct vertices p and q in $X \cap Y \cap Z$. Consider these p and q.

Both vertices p and q belong to $X \cap Y \cap Z \subseteq X$, thus appear on the path $\mathbf{x}$. We WLOG assume that p appears before q on this path (i.e., the path $\mathbf{x}$ has the form $(\ldots, p, \ldots, q, \ldots)$, where some of the $\ldots$ may be empty). (This WLOG assumption is legitimate, since we can switch p with q.)

But the vertices p and q also appear on the path $\mathbf{y}$ (since they belong to $X \cap Y \cap Z \subseteq Y$). The vertex p must appear after q on this path[8]. In other words, the vertex q appears before p on the path $\mathbf{y}$.

[8] *Proof.* Assume the contrary. Thus, p appears before q on the path $\mathbf{y}$.

Now, let us split the path $\mathbf{x}$ into three parts: Namely,

- let $\mathbf{x}_1$ be the part between b and p;
- let $\mathbf{x}_2$ be the part between p and q;
- let $\mathbf{x}_3$ be the part between q and c.

(This is possible because p appears before q in $\mathbf{x}$.) Then, $\ell(\mathbf{x}) = \ell(\mathbf{x}_1) + \ell(\mathbf{x}_2) + \ell(\mathbf{x}_3)$.

Next, let us split the path $\mathbf{y}$ into three parts: Namely,

- let $\mathbf{y}_1$ be the part between c and p;
- let $\mathbf{y}_2$ be the part between p and q;
- let $\mathbf{y}_3$ be the part between q and a.

(This is possible because p appears before q in $\mathbf{y}$.) Then, $\ell(\mathbf{y}) = \ell(\mathbf{y}_1) + \ell(\mathbf{y}_2) + \ell(\mathbf{y}_3)$.

The path $\mathbf{x}_2$ connects p and q, and thus has length > 0 (since p and q are distinct). Hence, $\ell(\mathbf{x}_2) > 0$. Thus,

$$\ell(\mathbf{x}) = \ell(\mathbf{x}_1) + \underbrace{\ell(\mathbf{x}_2)}_{>0} + \ell(\mathbf{x}_3) > \ell(\mathbf{x}_1) + \ell(\mathbf{x}_3),$$

so that

$$\ell(\mathbf{x}_1) + \ell(\mathbf{x}_3) < \ell(\mathbf{x}) = d(b,c). \tag{8}$$

Similarly,

$$\ell(\mathbf{y}_1) + \ell(\mathbf{y}_3) < d(c,a). \tag{9}$$

The same reasoning (but applied to b, c, a, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{x}$, q and p instead of a, b, c, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, p and q) now shows that the vertex p appears before q on the path $\mathbf{z}$ (because the vertex q appears before p on the path $\mathbf{y}$). The same reasoning (but applied to b, c, a, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{x}$ instead of a, b, c, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$) therefore shows that the vertex p appears before q on the path $\mathbf{y}$. But this contradicts the fact that the vertex q appears before p on the path $\mathbf{y}$.

This contradiction shows that our assumption was wrong. Hence, $|X \cap Y \cap Z| \leq 1$ is proven. From this, we can easily obtain $|X| + |Y| + |Z| \leq 2|X \cup Y \cup Z| + 1$ [9]. Now, adding the equalities (5), (6) and (7) together, we obtain

$$d(b,c) + d(c,a) + d(a,b) = (|X| - 1) + (|Y| - 1) + (|Z| - 1) = \underbrace{|X| + |Y| + |Z|}_{\leq 2|X \cup Y \cup Z| + 1} - 3$$

$$\leq 2 \left| \underbrace{X \cup Y \cup Z}_{\subseteq V} \right| + 1 - 3 \leq 2|V| + 1 - 3 = 2|V| - 2.$$

□

0.8.5. Third solution

Let me finally sketch a third solution of the exercise. The idea of the below solution is due to Sasha Pevzner, although I am restating it in slightly different terms. First, let me generalize the problem as follows:

Combining the reversal of the walk $\mathbf{x}_3$ with the walk $\mathbf{y}_3$, we obtain a walk from c to a (via q). This walk has length $\ell(\mathbf{x}_3) + \ell(\mathbf{y}_3)$. Hence, there exists a walk from c to a having length $\ell(\mathbf{x}_3) + \ell(\mathbf{y}_3)$. Consequently, Lemma 0.2 (applied to $u = c$, $v = a$ and $k = \ell(\mathbf{x}_3) + \ell(\mathbf{y}_3)$) shows that

$$d(c,a) \leq \ell(\mathbf{x}_3) + \ell(\mathbf{y}_3). \tag{10}$$

Combining the walk $\mathbf{x}_1$ with the reversal of the walk $\mathbf{y}_1$, we obtain a walk from b to c (via p). This walk has length $\ell(\mathbf{x}_1) + \ell(\mathbf{y}_1)$. Hence, there exists a walk from b to c having length $\ell(\mathbf{x}_1) + \ell(\mathbf{y}_1)$. Consequently, Lemma 0.2 (applied to $u = b$, $v = c$ and $k = \ell(\mathbf{x}_1) + \ell(\mathbf{y}_1)$) shows that

$$d(b,c) \leq \ell(\mathbf{x}_1) + \ell(\mathbf{y}_1). \tag{11}$$

Adding together the inequalities (10) and (11), we obtain

$$\begin{aligned} d(c,a) + d(b,c) &\leq (\ell(\mathbf{x}_3) + \ell(\mathbf{y}_3)) + (\ell(\mathbf{x}_1) + \ell(\mathbf{y}_1)) \\ &= \underbrace{\ell(\mathbf{y}_1) + \ell(\mathbf{y}_3)}_{\substack{< d(c,a) \\ \text{(by (9))}}} + \underbrace{\ell(\mathbf{x}_1) + \ell(\mathbf{x}_3)}_{\substack{< d(b,c) \\ \text{(by (8))}}} < d(c,a) + d(b,c). \end{aligned}$$

This is absurd. Hence, we have found a contradiction, qed.

[9] *Proof.* The sum $|X| + |Y| + |Z|$ counts the elements of $X \cup Y \cup Z$, but it counts some of them twice and some thrice: Namely, an element is counted thrice if it belongs to $X \cap Y \cap Z$; otherwise, it is counted at most twice. Since $|X \cap Y \cap Z| \leq 1$, we know that at most one element is counted thrice. All other elements are counted at most twice. Hence, the total sum $|X| + |Y| + |Z|$ is at most $3 + 2(|X \cup Y \cup Z| - 1) = 2|X \cup Y \cup Z| + 1$. Qed.

Theorem 0.4. Let $k \in \mathbb{N}$ be odd. Let $a_1, a_2, \ldots, a_k$ be k vertices of a connected simple graph G. Let us set $a_{k+1} = a_1$. Then,

$$\sum_{i=1}^{k} d(a_i, a_{i+1}) \leq (k-1)(|\operatorname{V}(G)| - 1).$$

Before we start proving this, let us notice that Theorem 0.4 does not hold for k even, and the whole question of maximizing $\sum_{i=1}^{k} d(a_i, a_{i+1})$ for k even is rather trivial[10]. This makes Theorem 0.4 (which answers the question of maximizing $\sum_{i=1}^{k} d(a_i, a_{i+1})$ for k odd[11]) all the more interesting.

Exercise 7 is the particular case of Theorem 0.4 for $k = 3$.

For the proof of Theorem 0.4, we shall need a slightly more advanced notation for the distance between two vertices in a graph: Namely, if u and v are two vertices of a simple graph G, then the distance $d(u, v)$ will often be denoted by $d_G(u, v)$ (in order to stress the dependence on G). This allows us to unambiguously speak of distances between u and v even when there are several different graphs containing u and v as vertices.

Sasha's argument begins by reducing the problem to the situation in which the graph is a tree. Thus, we will need to prove the following lemma, which of course is a particular case of Theorem 0.4:

Lemma 0.5. Let $k \in \mathbb{N}$ be odd. Let $a_1, a_2, \ldots, a_k$ be k vertices of a tree G. Let us set $a_{k+1} = a_1$. Then,

$$\sum_{i=1}^{k} d(a_i, a_{i+1}) \leq (k-1)(|\operatorname{V}(G)| - 1).$$

(Here, we regard a tree as a simple graph.)

Let us first see why proving this lemma is sufficient:

Proof of Theorem 0.4 using Lemma 0.5 (sketched). Assume that Lemma 0.5 is already proven.

[10] To wit, it is obvious that $\sum_{i=1}^{k} d(a_i, a_{i+1}) \leq k(|\operatorname{V}(G)| - 1)$ (since Lemma 0.1 shows that each of the k numbers $d(a_i, a_{i+1})$ is $\leq |\operatorname{V}(G)| - 1$). When k is even, this bound cannot be improved, because it is attained whenever G is the path graph P_n, the vertices $a_1, a_3, a_5, \ldots$ all equal to one endpoint of this path, and the vertices $a_2, a_4, a_6, \ldots$ all equal to the other endpoint.

[11] Indeed, it is not hard to see that equality can be obtained in the inequality in Theorem 0.4; hence, it really maximizes $\sum_{i=1}^{k} d(a_i, a_{i+1})$.

The graph G is connected, and thus has a spanning tree T. Fix such a T. Clearly, $\operatorname{V}(T) = \operatorname{V}(G)$. Being a tree, T is of course still connected.

But T is a tree. Therefore, Lemma 0.5 (applied to T instead of G) shows that $\sum\limits_{i=1}^{k} d_T(a_i, a_{i+1}) \leq (k-1)(|\operatorname{V}(T)| - 1) = (k-1)(|\operatorname{V}(G)| - 1)$ (since $\operatorname{V}(T) = \operatorname{V}(G)$).

Since T is a subgraph of G, each walk in T is a walk in G. Hence, $d_G(u, v) \leq d_T(u, v)$ for any two vertices u and v of G. Thus,

$$\sum_{i=1}^{k} \underbrace{d_G(a_i, a_{i+1})}_{\leq d_T(a_i, a_{i+1})} \leq \sum_{i=1}^{k} d_T(a_i, a_{i+1}) \leq (k-1)(|\operatorname{V}(G)| - 1).$$

Hence, Theorem 0.4 is proven. □

It thus remains to prove Lemma 0.5.

Proof of Lemma 0.5 (sketched). We proceed by induction on $|\operatorname{V}(G)|$.

The *induction base* is the case when $|\operatorname{V}(G)| = 1$. This case is easy (indeed, G has only one vertex in this case, so that all k vertices $a_1, a_2, \ldots, a_k$ must be identical, and therefore their distances $d(a_i, a_{i+1})$ are all 0).

Now, to the *induction step*: Fix an integer $N > 1$. Assume (as the induction hypothesis) that Lemma 0.5 is proven in the case when $|\operatorname{V}(G)| = N - 1$. We must now prove Lemma 0.5 in the case when $|\operatorname{V}(G)| = N$.

Thus, let us consider the situation of Lemma 0.5 under the assumption that $|\operatorname{V}(G)| = N$. We need to prove that

$$\sum_{i=1}^{k} d_G(a_i, a_{i+1}) \leq (k-1)(|\operatorname{V}(G)| - 1). \tag{12}$$

The graph G is a tree with more than 1 vertex (since $|\operatorname{V}(G)| = N > 1$). Hence, G has at least one leaf. Pick such a leaf, and denote it by v. Let v' be the only neighbor of v in G.

Let G' be the subgraph of G obtained from G by removing the vertex v and the unique edge with endpoint v (that is, the edge vv'). Thus, G' is a tree with $|\operatorname{V}(G) \setminus \{v\}| = N - 1$ vertices. Hence, $|\operatorname{V}(G')| = N - 1$.

For each vertex u of G, we define a vertex $\overline{u}$ of G' as follows:

$$\overline{u} = \begin{cases} u, & \text{if } u \neq v; \\ v', & \text{if } u = v \end{cases}.$$

Then, all of $\overline{a_1}, \overline{a_2}, \ldots, \overline{a_k}$ are vertices of G'. Hence, (by the induction hypothesis) we can apply Lemma 0.5 to G' and $\overline{a_i}$ instead of G and a_i. We thus obtain

$$\sum_{i=1}^{k} d_{G'}(\overline{a_i}, \overline{a_{i+1}}) \leq (k-1)\left(\left|\operatorname{V}(G')\right| - 1\right). \tag{13}$$

Recall that v is a leaf of the tree G. Hence, any path in G that neither starts nor ends at v must be a path in G' as well (because otherwise, it would have to traverse v, but this would entail that v is contained in at least two edges, which would contradict the fact that v is a leaf). Thus,

$$d_G(x,y) = d_{G'}(x,y) \tag{14}$$

for any two vertices x and y of G'.

Let us, however, notice that

$$d_G(x,y) = d_{G'}(\overline{x},\overline{y}) + [\text{exactly one of } x \text{ and } y \text{ equals } v] \tag{15}$$

for any two vertices x and y of G, where we are using the Iverson bracket notation[12].

[12] *Proof of* (15). Let x and y be two vertices of G. We must prove (15). If none of x and y equals v, then (15) follows from (14) (because in this case, we have $\overline{x} = x$ and $\overline{y} = y$). If both x and y equal v, then (15) holds as well (since in this case, we have $x = y$). Hence, it remains to prove (15) in the case when exactly one of x and y equals v. Thus, consider this case. WLOG assume that $x = v$ and $y \neq v$ (since otherwise, we can switch x with y). Thus, $[\text{exactly one of } x \text{ and } y \text{ equals } v] = 1$. From $x = v$, we obtain $\overline{x} = v'$ (by the definition of $\overline{x}$). From $y \neq v$, we obtain $\overline{y} = y$ (by the definition of $\overline{y}$). Recall that v' is the only neighbor of v. Hence, vv' is the only edge that contains v. In other words, vv' is the only edge that contains x (since $x = v$). Recall that G is a tree. Hence, there is only one path from x to y. The length of this path must therefore be $d_G(x,y)$. This path has length > 0 (since $y \neq v = x$) and therefore begins with an edge that contains x. Hence, it begins with the edge vv' (because the only edge that contains x is vv'). After traversing this edge, the path must proceed from v' to y, which requires $d_G(v',y)$ edges (again because G is a tree, and thus there is exactly one path from v' to y). Hence, the length of this path is $1 + d_G(v',y)$. Since we already know that the length of this path is $d_G(x,y)$, we thus obtain

$$d_G(x,y) = 1 + \underbrace{d_G(v',y)}_{\substack{= d_{G'}(v',y) \\ \text{(by (14), applied to } v' \text{ instead of } x \text{ (since } v' \text{ and } y \text{ are vertices of } G'))}} = 1 + d_{G'}(v',y)$$

$$= d_{G'}\Bigl(\underbrace{v'}_{=\overline{x}}, \underbrace{y}_{=\overline{y}}\Bigr) + 1 = d_{G'}(\overline{x},\overline{y}) + 1 = d_{G'}(\overline{x},\overline{y}) + [\text{exactly one of } x \text{ and } y \text{ equals } v]$$

(since $1 = [\text{exactly one of } x \text{ and } y \text{ equals } v]$). This proves (15).

Hence,

$$\begin{aligned}
&\sum_{i=1}^{k} \underbrace{d_G\left(a_i, a_{i+1}\right)}_{\substack{=d_{G'}\left(\overline{a_i}, \overline{a_{i+1}}\right)+[\text{exactly one of } a_i \text{ and } a_{i+1} \text{ equals } v] \\ \text{(by (15), applied to } x=a_i \text{ and } y=a_{i+1})}} \\
&= \sum_{i=1}^{k} \left(d_{G'}\left(\overline{a_i}, \overline{a_{i+1}}\right) + [\text{exactly one of } a_i \text{ and } a_{i+1} \text{ equals } v]\right) \\
&= \underbrace{\sum_{i=1}^{k} d_{G'}\left(\overline{a_i}, \overline{a_{i+1}}\right)}_{\substack{\leq (k-1)\left(\left|\operatorname{V}(G')\right|-1\right) \\ \text{(by (13))}}} + \sum_{i=1}^{k} \underbrace{[\text{exactly one of } a_i \text{ and } a_{i+1} \text{ equals } v]}_{\leq 1} \\
&\leq (k-1)\left(\left|\operatorname{V}\left(G'\right)\right| - 1\right) + \underbrace{\sum_{i=1}^{k} 1}_{=k} = (k-1)\left(\underbrace{\left|\operatorname{V}\left(G'\right)\right|}_{=N-1} - 1\right) + k \\
&= (k-1)\left((N-1)-1\right) + k = (k-1)(N-1) + 1. \qquad (16)
\end{aligned}$$

But recall that our goal is to prove (12). We assume the contrary (for the sake of contradiction). Hence, we have

$$\sum_{i=1}^{k} d_G\left(a_i, a_{i+1}\right) > (k-1)\left(\underbrace{\left|\operatorname{V}(G)\right|}_{=N} - 1\right) = (k-1)(N-1).$$

Since both sides of this inequality are integers, we thus obtain

$$\sum_{i=1}^{k} d_G\left(a_i, a_{i+1}\right) \geq (k-1)(N-1) + 1.$$

Combining this with (16), we see that the inequality (16) must be an equality. Hence, each inequality that was used in the derivation of (16) must also be an equality[13]. In particular, the inequality

$$[\text{exactly one of } a_i \text{ and } a_{i+1} \text{ equals } v] \leq 1$$

must be an equality for each $i \in \{1, 2, \ldots, k\}$ (because all of these inequalities were used in the derivation of (16)). In other words, the following claim holds:

Claim 1: For each $i \in \{1, 2, \ldots, k\}$, exactly one of a_i and a_{i+1} equals v.

[13]This is not to be taken fully literally. For example, we can derive the inequality $2 \cdot 0 \leq 0$ by multiplying the two inequalities $2 \leq 0$ and $0 \leq 0$, and it is not true that each of the latter two inequalities must be an equality, even though the former inequality is an equality. However, the derivation of the inequality (16) involved only **addition** (not multiplication) of other inequalities; and therefore, equality in (16) forces equality in each of the inequalities that were added.

Now, let us assume that $a_1 = v$. Then, Claim 1 (applied to $i = 1$) shows that exactly one of a_1 and a_2 equals v. Hence, $a_2 \neq v$ (since $a_1 = v$). But Claim 2 (applied to $i = 2$) shows that exactly one of a_2 and a_3 equals v. Hence, $a_3 = v$ (since $a_2 \neq v$). We can continue this line of reasoning, thus showing that $a_i = v$ for each odd $i \in \{1, 2, \ldots, k+1\}$ and that $a_i \neq v$ for each even $i \in \{1, 2, \ldots, k+1\}$. In particular, this shows that $a_{k+1} \neq v$ (since $k+1$ is even (since k is odd)). But this contradicts $a_{k+1} = a_1 = v$.

Thus, we have obtained a contradiction under the assumption that $a_1 = v$. But we could similarly have obtained a contradiction under the assumption that $a_1 \neq v$ (indeed, the very same argument would have worked, except that the roles of "being equal to v" and "being not equal to v" would be interchanged). Thus, we always have a contradiction. This completes the proof of Lemma 0.5. □

0.8.6. The generalization derived from the problem

As already mentioned, Exercise 7 can be obtained as a particular case of Theorem 0.4 for $k = 3$. Conversely, we can also prove Theorem 0.4 using a fairly simple induction argument that relies on Exercise 7 and Lemma 0.3. Unlike Sasha Pevzner's proof above, this does not give a new solution to Exercise 7, but might still be of interest.

Proof of Theorem 0.4 using Exercise 7 and Lemma 0.3. Write the graph G in the form $G = (V, E)$. Thus, $\operatorname{V}(G) = V$ and $\operatorname{E}(G) = E$. Set $n = |V|$. Also, set $n' = n - 1$.

We shall first show that for each $j \in \mathbb{N}$ satisfying $2j \leq k$, we have

$$\sum_{i=1}^{2j} d(a_i, a_{i+1}) \leq 2jn' - d\left(a_{2j+1}, a_1\right). \tag{17}$$

Proof of (17)*:* We shall prove (17) by induction over j.

Induction base: The inequality (17) holds for $j = 0$ [14]. Hence, the induction base is complete.

Induction step: Fix a positive $J \in \mathbb{N}$ satisfying $2J \leq k$. Assume that (17) holds for $j = J - 1$. We must then show that (17) holds for $j = J$.

Since J is a positive integer, we have $J - 1 \in \mathbb{N}$. Moreover, $2(J-1) \leq 2J \leq k$. Hence, (17) is applicable to $j = J - 1$. Since we have assumed that (17) holds for

[14] *Proof.* We have $d\left(\underbrace{a_{2\cdot 0+1}}_{=a_1}, a_1\right) = d(a_1, a_1) = 0$ (by Lemma 0.3 **(a)**), so that $\underbrace{2 \cdot 0n'}_{=0} - \underbrace{d(a_1, a_{2\cdot 0+1})}_{=0} = 0 - 0 = 0$. Now, $\sum_{i=1}^{2\cdot 0} d(a_i, a_{i+1}) = (\text{empty sum}) = 0 \leq 0 = 2 \cdot 0n' - d(a_{2\cdot 0+1}, a_1)$. In other words, the inequality (17) holds for $j = 0$.

$j = J - 1$, we thus obtain

$$\sum_{i=1}^{2(J-1)} d(a_i, a_{i+1}) \leq 2(J-1)n' - d\left(a_{2(J-1)+1}, a_1\right)$$
$$= 2(J-1)n' - d\left(a_{2J-1}, a_1\right) \qquad (18)$$

(since $2(J-1)+1 = 2J-1$).

Lemma 0.3 **(c)** (applied to $u = a_{2J-1}$, $v = a_1$ and $w = a_{2J}$) yields $d\left(a_{2J-1}, a_1\right) + d\left(a_1, a_{2J}\right) \geq d\left(a_{2J-1}, a_{2J}\right)$. Hence,

$$d\left(a_{2J-1}, a_{2J}\right) \leq d\left(a_{2J-1}, a_1\right) + d\left(a_1, a_{2J}\right).$$

But Exercise 7 (applied to $a = a_{2J}$, $b = a_{2J+1}$ and $c = a_1$) yields

$$d\left(a_{2J+1}, a_1\right) + d\left(a_1, a_{2J}\right) + d\left(a_{2J}, a_{2J+1}\right) \leq 2\underbrace{|V|}_{=n} - 2 = 2n - 2 = 2\underbrace{(n-1)}_{=n'} = 2n'.$$

Subtracting $d\left(a_{2J+1}, a_1\right)$ from both sides of this inequality, we obtain

$$d\left(a_1, a_{2J}\right) + d\left(a_{2J}, a_{2J+1}\right) \leq 2n' - d\left(a_{2J+1}, a_1\right).$$

Now,

$$\begin{aligned}
\sum_{i=1}^{2J} d(a_i, a_{i+1}) &= \sum_{i=1}^{2J-2} d(a_i, a_{i+1}) + d\left(a_{2J-1}, a_{2J}\right) + d\left(a_{2J}, a_{2J+1}\right) \\
&= \underbrace{\sum_{i=1}^{2(J-1)} d(a_i, a_{i+1})}_{\substack{\leq 2(J-1)n' - d\left(a_{2J-1}, a_1\right) \\ \text{(by (18))}}} + \underbrace{d\left(a_{2J-1}, a_{2J}\right)}_{\leq d\left(a_{2J-1}, a_1\right) + d\left(a_1, a_{2J}\right)} + d\left(a_{2J}, a_{2J+1}\right) \\
&\qquad (\text{since } 2J-2 = 2(J-1)) \\
&\leq 2(J-1)n' - d\left(a_1, a_{2J-1}\right) + d\left(a_1, a_{2J-1}\right) + d\left(a_1, a_{2J}\right) + d\left(a_{2J}, a_{2J+1}\right) \\
&= 2(J-1)n' + \underbrace{d\left(a_1, a_{2J}\right) + d\left(a_{2J}, a_{2J+1}\right)}_{\leq 2n' - d\left(a_{2J+1}, a_1\right)} \\
&\leq \underbrace{2(J-1)n' + 2n'}_{=2Jn'} - d\left(a_{2J+1}, a_1\right) = 2Jn' - d\left(a_{2J+1}, a_1\right).
\end{aligned}$$

In other words, (17) holds for $j = J$. This completes the induction step. Thus, (17) is proven by induction. □

We know that k is an odd nonnegative integer. Hence, there exists some $j \in \mathbb{N}$ satisfying $k = 2j+1$. Consider this j. From $2j+1 = k$, we obtain $a_{(2j+1)+1} = a_{k+1} = a_1$. Also, $2j = k-1$ (since $k = 2j+1$) and $n' = \underbrace{n}_{=|V|} - 1 = \left|\underbrace{V}_{=\operatorname{V}(G)}\right| - 1 = |\operatorname{V}(G)| - 1$.

From $k = 2j+1$, we obtain

$$\sum_{i=1}^{k} d(a_i, a_{i+1}) = \sum_{i=1}^{2j+1} d(a_i, a_{i+1}) = \underbrace{\sum_{i=1}^{2j} d(a_i, a_{i+1})}_{\substack{\leq 2jn' - d(a_{2j+1}, a_1) \\ \text{(by (17))}}} + d\left(a_{2j+1}, \underbrace{a_{(2j+1)+1}}_{=a_1}\right)$$
$$\leq 2jn' - d(a_{2j+1}, a_1) + d(a_{2j+1}, a_1) = \underbrace{2j}_{=k-1} \underbrace{n'}_{=|V(G)|-1}$$
$$= (k-1)(|V(G)|-1).$$

This proves Theorem 0.4. □

0.9. Exercise 8: two steps away from a forest

We prepare for Exercise 8 by showing several lemmas.

Lemma 0.6. Let H be a multigraph such that $|E(H)| \geq |V(H)|$. Assume that H has at least one vertex. Then, H contains a cycle.

Proof of Lemma 0.6 (sketched). Assume the contrary. Hence, H contains no cycle, i.e., is a forest. Therefore, Corollary 20 from lecture 9 shows that[15] $|E(H)| = |V(H)| - b_0(H) < |V(H)|$ (since $b_0(H) > 0$), which contradicts $|E(H)| \geq |V(H)|$. Hence, our assumption was false, qed. □

In the following, we shall use the Iverson bracket notation. In other words, for any logical statement $\mathcal{A}$, we set $[\mathcal{A}] = \begin{cases} 1, & \text{if } \mathcal{A} \text{ is true;} \\ 0, & \text{if } \mathcal{A} \text{ is false} \end{cases}$.

Let us introduce one more notation:

- If A and B are two sets, then $A \triangle B$ shall denote the *symmetric difference* of A and B; this is defined as the set $(A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$. In other words, $A \triangle B$ is the set of all elements that belong to **exactly one** of A and B. (This is the operation on sets that corresponds to the logical operation XOR.)

Lemma 0.7. Let $G = (V, E, \phi)$ be a multigraph. Let A and B be two subsets of E. Let $v \in V$. Then,

$$\deg_{(V, A \triangle B, \phi|_{A \triangle B})} v = \deg_{(V, A, \phi|_A)} v + \deg_{(V, B, \phi|_B)} v - 2 \deg_{(V, A \cap B, \phi|_{A \cap B})} v.$$

[15] Recall that $b_0(H)$ denotes the number of connected components of H. Since H has at least one connected component (because H has at least one vertex), we have $b_0(H) > 0$.

Proof of Lemma 0.7. Each edge $e \in E$ satisfies

$$[e \in A \triangle B] = [e \in A] + [e \in B] - 2[e \in A \cap B] \tag{19}$$

[16].

But for each subset F of E, we have

$$\deg_{(V,F,\phi|_F)} v = \sum_{e \in E;\ v \in \phi(e)} [e \in F] \tag{20}$$

[17]. Applying this to $F = A \triangle B$, we obtain

$$\begin{aligned}
\deg_{(V,A\triangle B,\phi|_{A\triangle B})} v &= \sum_{e \in E;\ v \in \phi(e)} \underbrace{[e \in A \triangle B]}_{\substack{=[e\in A]+[e\in B]-2[e\in A\cap B] \\ \text{(by (19))}}} \\
&= \sum_{e \in E;\ v \in \phi(e)} ([e \in A] + [e \in B] - 2[e \in A \cap B]) \\
&= \sum_{e \in E;\ v \in \phi(e)} [e \in A] + \sum_{e \in E;\ v \in \phi(e)} [e \in B] - 2 \sum_{e \in E;\ v \in \phi(e)} [e \in A \cap B].
\end{aligned}$$

---

[16] *Proof of* (19). Let $e \in E$. We must prove the equality (19). We are in one of the following four cases:

- *Case 1:* We have $e \in A$ and $e \in B$.
- *Case 2:* We have $e \in A$ but not $e \in B$.
- *Case 3:* We have $e \in B$ but not $e \in A$.
- *Case 4:* We have neither $e \in A$ nor $e \in B$.

Let us prove (19) in Case 1. In this case, we have $e \in A$ and $e \in B$. Thus, $e \in A \cap B$, so that $e \notin (A \cup B) \setminus (A \cap B) = A \triangle B$. Thus, $[e \in A \triangle B] = 0$. Also, $[e \in A \cap B] = 1$ (since $e \in A \cap B$) and $[e \in A] = 1$ (since $e \in A$) and $[e \in B] = 1$ (since $e \in B$). Thus,

$$\underbrace{[e \in A]}_{=1} + \underbrace{[e \in B]}_{=1} - 2\underbrace{[e \in A \cap B]}_{=1} = 1 + 1 - 2 \cdot 1 = 0 = [e \in A \triangle B].$$

Hence, (19) is proven in Case 1. Similarly, (19) can be proven in Case 2, in Case 3, and in Case 4. Therefore, (19) always holds. This completes the proof of (19).

[17] *Proof of* (20). Let F be a subset of E. The definition of $\deg_{(V,F,\phi|_F)} v$ shows that $\deg_{(V,F,\phi|_F)} v$ is the number of edges of the multigraph $(V, F, \phi|_F)$ that contain v. Since the edges of the multigraph $(V, F, \phi|_F)$ are the elements of F, this rewrites as follows: $\deg_{(V,F,\phi|_F)} v$ is the number of elements of F that contain v. In other words, $\deg_{(V,F,\phi|_F)} v$ is the number of all $e \in F$ that satisfy $v \in (\phi|_F)(e)$. In other words, $\deg_{(V,F,\phi|_F)} v$ is the number of all $e \in F$ that satisfy $v \in \phi(e)$ (because obviously, we have $(\phi|_F)(e) = \phi(e)$ for each $e \in F$). In other words,

$$\deg_{(V,F,\phi|_F)} v = (\text{the number of all } e \in F \text{ that satisfy } v \in \phi(e)). \tag{21}$$

Comparing this with

$$\underbrace{\deg_{(V,A,\phi|_A)} v}_{\substack{=\sum_{e\in E;\ v\in\phi(e)} [e\in A] \\ \text{(by (20), applied to } F=A)}} + \underbrace{\deg_{(V,B,\phi|_B)} v}_{\substack{=\sum_{e\in E;\ v\in\phi(e)} [e\in B] \\ \text{(by (20), applied to } F=B)}} - 2 \underbrace{\deg_{(V,A\cap B,\phi|_{A\cap B})} v}_{\substack{=\sum_{e\in E;\ v\in\phi(e)} [e\in A\cap B] \\ \text{(by (20), applied to } F=A\cap B)}}$$
$$= \sum_{e\in E;\ v\in\phi(e)} [e\in A] + \sum_{e\in E;\ v\in\phi(e)} [e\in B] - 2 \sum_{e\in E;\ v\in\phi(e)} [e\in A\cap B],$$

we obtain

$$\deg_{(V,A\triangle B,\phi|_{A\triangle B})} v = \deg_{(V,A,\phi|_A)} v + \deg_{(V,B,\phi|_B)} v - 2\deg_{(V,A\cap B,\phi|_{A\cap B})} v.$$

This proves Lemma 0.7. □

Lemma 0.8. Let $G = (V, E, \phi)$ be a multigraph. Let $\mathbf{a}$ and $\mathbf{b}$ be two cycles of G. Let A be the set of the edges of $\mathbf{a}$. Let B be the set of the edges of $\mathbf{b}$.

(a) For each vertex $v \in V$, the number $\deg_{(V,A\triangle B,\phi|_{A\triangle B})} v$ is even.

(b) If $A \neq B$, then the multigraph $(V, A \triangle B, \phi\mid_{A\triangle B})$ has a cycle.

(c) Let $\mathbf{c}$ be a cycle of the multigraph $(V, A \triangle B, \phi\mid_{A\triangle B})$. Then, at least one of the cycles $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ has length $\leq \dfrac{2|E|}{3}$.

Proof of Lemma 0.8. **(a)** Let $v \in V$. Then, $\deg_{(V,A,\phi|_A)} v \equiv 0 \mod 2$ [18]. Similarly,

---

But

$$\sum_{e\in E;\ v\in\phi(e)} [e\in F] = \sum_{e\in E;\ v\in\phi(e);\ e\in F} \underbrace{[e\in F]}_{\substack{=1 \\ \text{(since } e\in F)}} + \sum_{e\in E;\ v\in\phi(e);\ \text{not } e\in F} \underbrace{[e\in F]}_{\substack{=0 \\ \text{(since we do not have } e\in F)}}$$
$$\text{(since each } e\in E \text{ either satisfies } e\in F \text{ or does not)}$$
$$= \sum_{e\in E;\ v\in\phi(e);\ e\in F} 1 + \underbrace{\sum_{e\in E;\ v\in\phi(e);\ \text{not } e\in F} 0}_{=0}$$
$$= \sum_{e\in E;\ v\in\phi(e);\ e\in F} 1$$
$$= (\text{the number of all } e\in E \text{ that satisfy } v\in\phi(e) \text{ and } e\in F)\cdot 1$$
$$= (\text{the number of all } e\in E \text{ that satisfy } v\in\phi(e) \text{ and } e\in F)$$
$$= (\text{the number of all } e\in F \text{ that satisfy } v\in\phi(e))$$

(because the $e \in E$ that satisfy $e \in F$ are precisely the $e \in F$). Comparing this with (21), we obtain

$$\deg_{(V,F,\phi|_F)} v = \sum_{e\in E;\ v\in\phi(e)} [e\in F].$$

This proves (20).

[18] *Proof.* We are in one of the following two cases:

$\deg_{(V,B,\phi|_B)} v \equiv 0 \mod 2$. Now, Lemma 0.7 yields

$$\begin{aligned}\deg_{\left(V,A\triangle B,\phi|_{A\triangle B}\right)} v &= \underbrace{\deg_{(V,A,\phi|_A)} v}_{\equiv 0 \mod 2} + \underbrace{\deg_{(V,B,\phi|_B)} v}_{\equiv 0 \mod 2} - \underbrace{2}_{\equiv 0 \mod 2} \deg_{(V,A\cap B,\phi|_{A\cap B})} v \\ &\equiv 0+0-0 = 0 \mod 2.\end{aligned}$$

In other words, the number $\deg_{\left(V,A\triangle B,\phi|_{A\triangle B}\right)} v$ is even. This proves Lemma 0.8 **(a)**.

(b) It is well-known that if X and Y are two sets satisfying $X \neq Y$, then $X \triangle Y \neq \varnothing$. Applying this to $X = A$ and $Y = B$, we obtain $A \triangle B \neq \varnothing$. Hence, the multigraph $\left(V, A \triangle B, \phi\mid_{A\triangle B}\right)$ has at least one edge. Furthermore, each vertex of this multigraph has even degree (because of Lemma 0.8 **(a)**). Thus, Exercise 4 (applied to $\left(V, A \triangle B, \phi\mid_{A\triangle B}\right)$ instead of G) shows that this multigraph $\left(V, A \triangle B, \phi\mid_{A\triangle B}\right)$ has a cycle. This proves Lemma 0.8 **(b)**.

(c) Assume the contrary. Thus, none of the cycles **a**, **b** and **c** has length $\leq \dfrac{2|E|}{3}$. In other words, each of the cycles **a**, **b** and **c** has length $> \dfrac{2|E|}{3}$.

Let C be the set of edges of **c**. Then, $C \subseteq A \triangle B$ (since **c** is a cycle of the multigraph $\left(V, A \triangle B, \phi\mid_{A\triangle B}\right)$). Hence,

$$\begin{aligned}|C| &\leq \left|\underbrace{A \triangle B}_{=(A\cup B)\setminus(A\cap B)}\right| = |(A\cup B)\setminus(A\cap B)| \\ &= |A\cup B| - \underbrace{|A\cap B|}_{=|A|+|B|-|A\cup B|} \qquad (\text{since } A\cap B \subseteq A \subseteq A\cup B) \\ &= |A\cup B| - (|A|+|B|-|A\cup B|) = 2\left|\underbrace{A\cup B}_{\subseteq E}\right| - |A| - |B| \leq 2|E| - |A| - |B|.\end{aligned}$$

- *Case 1:* The vertex v does not lie on the cycle **a**.
- *Case 2:* The vertex v lies on the cycle **a**.

Let us consider Case 1 first. In this case, the vertex v does not lie on the cycle **a**. Hence, no edge of **a** contains v. In other words, no edge in A contains v (since the edges in A are precisely the edges of **a**). In other words, $\deg_{(V,A,\phi|_A)} v = 0$. Hence, $\deg_{(V,A,\phi|_A)} v = 0 \equiv 0 \mod 2$. Thus, $\deg_{(V,A,\phi|_A)} v \equiv 0 \mod 2$ is proven in Case 1.

Let us next consider Case 2. In this case, the vertex v lies on the cycle **a**. Hence, exactly two edges of **a** contain v (because a vertex on a cycle is entered exactly once by the cycle, and exited exactly once by the cycle). These two edges are distinct (since the edges of a cycle are always distinct). Thus, exactly two distinct edges of **a** contain v. In other words, exactly two edges in A contain v (since the edges in A are precisely the edges of **a**). In other words, $\deg_{(V,A,\phi|_A)} v = 2$. Hence, $\deg_{(V,A,\phi|_A)} v = 2 \equiv 0 \mod 2$. Thus, $\deg_{(V,A,\phi|_A)} v \equiv 0 \mod 2$ is proven in Case 2.

We have now proven $\deg_{(V,A,\phi|_A)} v \equiv 0 \mod 2$ in each of the two Cases 1 and 2. Hence, $\deg_{(V,A,\phi|_A)} v \equiv 0 \mod 2$ always holds.

Adding $|A| + |B|$ to both sides of this inequality, we obtain $|A| + |B| + |C| \leq 2|E|$.

But

$$\begin{aligned} |A| &= (\text{the number of all elements of } A) = (\text{the number of all edges of } \mathbf{a}) \\ &\qquad (\text{since the elements of } A \text{ are the edges of } \mathbf{a}, \text{ and since these edges are all distinct}) \\ &= (\text{the length of } \mathbf{a}) > \frac{2|E|}{3} \end{aligned}$$

(since the cycle $\mathbf{a}$ has length $> \frac{2|E|}{3}$). Similarly, $|B| > \frac{2|E|}{3}$ and $|C| > \frac{2|E|}{3}$. Adding these three inequalities together, we obtain

$$|A| + |B| + |C| > \frac{2|E|}{3} + \frac{2|E|}{3} + \frac{2|E|}{3} = 2|E|.$$

This contradicts $|A| + |B| + |C| \leq 2|E|$. This contradiction shows that our assumption was wrong. Hence, Lemma 0.8 **(c)** is proven. □

Lemma 0.9. Let $G = (V, E, \phi)$ be a multigraph such that $|E| > |V|$. Let $n = |V|$. Then, G has a cycle of length $\leq \frac{2n+2}{3}$.

Proof of Lemma 0.9 (sketched). Recall that $|E| > |V|$. Thus, $|E| \geq |V| + 1$. Hence, we can WLOG assume that $|E| = |V| + 1$ (since otherwise, we can keep deleting edges from G until $|E| = |V| + 1$ holds; if we can find a cycle of length $\leq \frac{2n+2}{3}$ after that, then we clearly also get such a cycle in the original graph).

Thus, $|E| = \underbrace{|V|}_{=n} + 1 = n + 1$.

The multigraph G has at least one edge (since $|E| > |V| \geq 0$), and therefore has at least one vertex. Also, $|\mathrm{E}(G)| = |E| \geq |V| = |\mathrm{V}(G)|$. Hence, Lemma 0.6 (applied to $H = G$) yields that G contains a cycle. Fix such a cycle, and denote it by $\mathbf{a}$. Fix an edge a of $\mathbf{a}$.

Removing the edge a from G yields a new multigraph G', which has one fewer edge than G. Thus, this new graph G' satisfies $|\mathrm{E}(G')| = |E| - 1 \geq |V|$ (since $|E| > |V|$). This rewrites as $|\mathrm{E}(G')| \geq |\mathrm{V}(G')|$ (since $V = \mathrm{V}(G')$). Consequently, Lemma 0.6 (applied to $H = G'$) yields that G' contains a cycle. Fix such a cycle, and denote it by $\mathbf{b}$. Of course, $\mathbf{b}$ is a cycle of G as well (since G' is a subgraph of G).

Let A be the set of the edges of $\mathbf{a}$, and let B be the set of edges of $\mathbf{b}$. Then, $A \neq B$ (since $a \in A$ but $a \notin B$). Hence, Lemma 0.8 **(b)** shows that the multigraph $(V, A \triangle B, \phi \mid_{A \triangle B})$ has a cycle. Fix such a cycle, and denote it by $\mathbf{c}$. Clearly, $\mathbf{c}$ is a cycle of G (since $(V, A \triangle B, \phi \mid_{A \triangle B})$ is a subgraph of G). Lemma 0.8 **(c)** shows that at least one of the cycles $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ has length $\leq \frac{2|E|}{3}$. Since $2\underbrace{|E|}_{=n+1} = 2(n+1) =$

$2n+2$, this rewrites as follows: At least one of the cycles **a**, **b** and **c** has length $\leq \frac{2n+2}{3}$. Hence, G has a cycle of length $\leq \frac{2n+2}{3}$ (since all of **a**, **b** and **c** are cycles of G). This proves Lemma 0.9. □

Exercise 8. Let $G = (V, E)$ be a simple graph such that $|E| > |V|$. Prove that G has a cycle of length $\leq \frac{2n+2}{3}$, where $n = |V|$.

Solution sketch to Exercise 8. Recall that each simple graph can be viewed as a multigraph in an obvious way. Thus, we can view G as a multigraph this way. Hence, Exercise 8 follows immediately from Lemma 0.9. □

[*Remarks:*

- A statement fairly close to Exercise 8 appears in [BoCaDu13, Lemma 6].
- The bound $\frac{2n+2}{3}$ cannot be improved. For a graph G which achieves this bound, see Exercise 2 **(c)** on homework set 3. A slight modification of this construction allows us to find a graph G with n vertices achieving the bound $\left\lfloor \frac{2n+2}{3} \right\rfloor$ for each $n > 3$.
- We have generalized Exercise 8 to Lemma 0.9 by replacing the simple graph G by a multigraph. However, this generalization does not add any significant new power to the statement, because each multigraph with no two parallel edges (= two distinct edges e_1 and e_2 satisfying $\phi(e_1) = \phi(e_2)$) can be regarded as a simple graph (as long as we are willing to forget the names of the edges, which in our case is harmless), whereas Lemma 0.9 holds obviously for a multigraph with two parallel edges (in fact, two parallel edges form a cycle of length 2). Nevertheless, I find the generalization worth stating, since it appears to me that the setting of multigraphs is more natural for this exercise.
- Given two integers $k \geq 0$ and $n > 1$, we can define $\rho(n,k)$ to be the smallest integer such that each multigraph G satisfying $|\mathrm{E}(G)| \geq |\mathrm{V}(G)| + k$ and $|\mathrm{V}(G)| = n$ must have a cycle of length $\leq \rho(n,k)$. Can we compute this $\rho(n,k)$, or at least find a good upper bound on it? From Lemma 0.6 (and the example of the cycle graph C_n), we can easily obtain $\rho(n,0) = n$. From Lemma 0.9 (and an example of a graph G achieving the bound), we obtain $\rho(n,1) = \left\lfloor \frac{2n+2}{3} \right\rfloor$ for each $n > 3$. What can we say about $\rho(n,2)$? The notion of cages seems relevant (even though not directly applicable).

]

References

[BoCaDu13] Prosenjit Bose, Paz Carmi, Stephane Durocher, *Bounding the locality of distributed routing algorithms*, Distrib. Comput. (2013) 26, pp. 39–58.
A preprint can be found at https://pdfs.semanticscholar.org/424b/649b86c848cde7072c4a3500ecb40119e442.pdf