Bounds on distances for spanning trees of graphs

Mr Mthobisi Luca Ntuli

March 8, 2018

Bounds on distances for spanning trees of graphs

by

Mr Mthobisi Luca Ntuli

Submitted in partial fulfilment of the academic requirements for the degree of Masters in Science in the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal

As the candidate's supervisor, I have approved this dissertation for submission

Dr MJ Morgan	Jane Morgan	1 Dec 2017
Name	Signature	Date
Prof S Mukwembi		4/12/17
Name	Signature	Date

Declaration

The research done in this dissertation is original work and has not been previously submitted to any other institution. Where reference to the work of other researchers was made, it has been duly acknowledged.

Mr ML Ntuli

To

Mphemba Legacy

Acknowledgments

I would like to thank my supervisors, Dr MJ Morgan and Prof S Mukwembi. It has been a great privilege working with them. I would also like to thank the College of Agriculture, Engineering and Science for the bursary they offered me for the past two years.

Abstract

In graph theory, there are several techniques known in literature for constructing spanning trees. Some of these techniques yield spanning trees with many leaves. We will use these constructed spanning trees to bound several distance parameters.

The cardinality of the vertex set of graph G is called the *order*, $n(G)$ or n. The cardinality of the edge set of graph G is called the *size*, $m(G)$ or m. The *minimum degree* of G, $\delta(G)$ or δ, is the minimum degree among the degrees of the vertices of G. A spanning tree T of a graph G is a subgraph that is a tree which includes all the vertices of G. The distance $d(u, v)$ between two vertices u and v is the length of a shortest $u - v$ path of G. The *eccentricity*, $ec\,(v)$, of a vertex $v \in V(G)$ is the maximum distance from it to any other vertex in G. The *diameter*, $diam(G)$ or d, is the maximum eccentricity amongst all vertices of G. The *radius*, $rad(G)$, is the minimum eccentricity among all vertices of G. The *average distance* of a graph G, $\mu\,(G)$, is the expected distance between a randomly chosen pair of distinct vertices.

We investigate how each constructed spanning tree can be used to bound diameter, radius or average distance in terms of order, size and minimum degree. The techniques to be considered include the radius-preserving spanning trees by Erdős et al, the Ding et al technique, and the Dankelmann and Entringer technique. Finally, we use the Kleitman and West dead leaves technique to construct spanning trees with many leaves for various values of the minimum degree $\delta \geq k$ (for $k = 3, 4$ and $k > 4$) and order n. We then use the leaf number to bound diameter.

Contents

Chapter 1

Introduction and Preliminaries

The purpose of this chapter is to define the most important terms that will be used in this dissertation and to present motivation for our study. Terms not defined in this chapter will be defined in subsequent chapters.

1.1 General Graph Theory terminology and definitions

A graph G is a finite non-empty set of objects, $V(G)$, called *vertices* (the singular is vertex), together with a set of unordered pairs of distinct vertices called *edges*, $E(G)$. The cardinality of the vertex set of graph G is called the *order*, $n(G)$ or n. The cardinality of the edge set of graph G is called the *size*, $m(G)$ or m. The edge $e = uv$ is said to *join* the vertices u and v. If $e = uv$ is an edge of G, then u and v are *adjacent vertices*, while u and e are *incident*, as are v and e. A *complete graph*, K_n, of order n, is a graph in which every two distinct vertices are adjacent. The *complement* of a graph G, $\overline{G}$, is the graph with the same vertex set as G, and where distinct vertices u and v are adjacent in $\overline{G}$ if and only if they are not adjacent in G. The *degree* of v, $d(v)$, is the number of edges of G incident with v. The *Handshaking Lemma* states that in any graph, the sum of all the vertex degrees is equal to twice the number of edges. A *degree sequence* is a list of the degrees of vertices of a graph in non increasing order. The *irregularity index of* G, $\eta(G)$ or η, is the number of distinct terms in the degree sequence of G. A graph G is called k-regular if the degree of every vertex of G is k. The *minimum degree* of G, $\delta(G)$ or δ, is the minimum degree among the degrees of the vertices of G. The *maximum degree* of G, $\Delta(G)$ or Δ is the maximum degree among the degrees of the vertices of G. An *end vertex* is a vertex of a graph that has exactly one edge incident to it. A $u - v$ *walk* in a graph G is a finite, alternating sequence of vertices and edges that begins with the vertex u and ends with the vertex v and in which each edge of the sequence joins the vertex that precedes it in the sequence to the vertex that follows it in the sequence. A walk in which all the edges are distinct is a *trail*. A $u-v$ walk is *closed* if $u = v$ and *open* otherwise. If all the vertices and edges of a walk are distinct then that trail is a *path*, P. A *cycle*, C_n, is a closed path. A *subgraph* H of a graph G, $H \leq G$, is a graph whose vertices belong to $V(G)$ and the set of edges belong to $E(G)$. If W is a non-empty subset of vertices of graph G, then the subgraph $G(W)$ of G induced by W is the graph having vertex set W and whose edge set consists of all those edges of G incident with two vertices in W. A subgraph H of G is called an *induced subgraph* of G if $H = G(W)$ for some subset W of $V(G)$. Two graphs G and H are *isomorphic*, $G \simeq H$ if H can be obtained from G by relabeling the vertices that is, there exists a one-to-one ($f : V(G) \to V(H)$) correspondence between vertices of G and those of H such that an edge joins any pair of vertices in G if and only if an edge joins the corresponding pair of vertices in H. Such a function f is called an *isomorphism* from G to H. A *triangle-free graph* is a graph in which no three vertices form a triangle of edges. A graph is called C_4-*free* if it contains no cycle of length four as an induced subgraph. A graph is called *connected* if given any two vertices v_i, v_j, there is a path from v_i to v_j. A *component* of a graph G is a maximal connected subgraph of G. An *edge cut* of G is an edge whose deletion increases the number of components. A *tree* T is a connected graph which contains no cycles. A *forest* is a graph that has no cycles (each component of a forest is a tree). A well known property of a tree is that a tree T of order n has size $m(T)=$ $n - 1$. A *spanning tree* T of a graph G is a subgraph that is a tree which includes all the vertices of G. A *leaf* of a tree is an end vertex of T. The *leaf number*, $L(G)$ or L, is the maximum number of end vertices contained in a spanning tree of G.

The *distance* $d(u, v)$ between two vertices u and v is the minimum length of the $u - v$ paths of G. This distance function $d(u, v)$ is a *metric*, that is, it satisfies the following fundamental properties, for all $u, v, w \in V(G)$:

(i) $d(u, v) \geq 0$ and $d(u, v) = 0$ if and only if $u = v$;

(ii) $d(u, v)= d(v, u)$ (symmetry property);

(iii) $d(u, v) \leq d(u, w) + d(w, v)$ (triangle inequality).

We denote this distance by $d(u, v)$, and in situations where clarity of context is important, we may write $d_G(u, v)$. The *open neighbourhood*, $N_G(v)$, of a vertex v of G is the set $\{x \in V : d(x, v) = 1\}$. The *closed neighbourhood*, $N_G[v]$, of a vertex v of

G is the set $N_G[v] = N_G(v) \cup \{v\}$. The *eccentricity*, $ec(v)$, of a vertex $v \in V(G)$ is the maximum distance from it to any other vertex in G. The *diameter*, $diam(G)$ or d, is the maximum eccentricity amongst all vertices of G. The *radius*, $rad(G)$, is the minimum eccentricity among all vertices of G. The *average eccentricity*, $avec(G)$, is the mean of the eccentricities of vertices in G. A vertex v is a *central vertex* of G if $ec(v) = rad(G)$. The *average distance* of a graph G, $\mu(G)$, is the expected distance between a randomly chosen pair of distinct vertices. A *weighted graph* is a graph G in which each edge e is assigned a positive real number, called the *weight of* e, denoted by $w(e)$.

1.2 Overall approach and Motivation

In this dissertation, we will study several techniques for constructing a spanning tree, with the aim of seeing how these spanning trees can be used to bound various distance parameters. In particular, the techniques to be considered are the radius-preserving spanning tree by Erdős et al [10], the spanning tree constructed by Ding et al [8], the spanning tree constructed by Dankelmann and Entringer [7], and the dead leaves spanning tree construction by Kleitman and West [13]. Thereafter, for each of these techniques, we will derive various distance-related bounds on these spanning trees. The distance-based upper or lower bounds to be considered will be radius, diameter and average distance. These bounds will be functions of graph invariants such as order, size and minimum degree.

Graph theory plays an important role in solving biological networks, network communication and computer problems (for example, metabolic and gene regulation networks in each cell, a city road system, computer processes or a telephonic exchange) [19, 21]. Using the properties of spanning trees, problems such as data congestion, cost of devices (i.e., software and hardware) and performance of the network can be solved [17]. A spanning tree can be used in an optical fibre network system to bring us internet, cable TV and telephone services [18].

We study the distance parameters because of their importance in solving network or communication problems. For example, diameter plays a significant role in analyzing communication networks. In such networks, the time delay or signal degradation for sending a message from one point to another is often proportional to the distance between the two points. The diameter can be used to indicate the worst-case performance in this scenario [5].

The radius also plays an important role in road systems. For example, a municipality might want to place an emergency facility like a hospital in the city. Here the primary interest would be the distance to the emergency facility and a location furthest away. The municipality would want to place the emergency facility where the response time or distance is minimum. Then the radius of a graph would be the minimum response time or distance from the emergency facility to a location furthest away.

The average distance has been studied by a number of authors, [4, 6, 7, 14, 16]. For example the average distance of a graph can be used as the average travel time between any two randomly chosen locations in the city, and in architecture as a tool for evaluation of floor plans. Here each room corresponds to a vertex, and two vertices are adjacent if it is possible to move directly between the corresponding rooms [16].

Chapter 2

Bounds on distances in terms of Order, Size and Minimum degree

In this chapter, we consider a distance-preserving spanning tree, and find bounds on distance parameters in terms of order and size, as well as bounds in terms of order and minimum degree. In each case, we investigate how the constructed spanning tree can be used to bound radius, diameter or average distance.

2.1 Distance-preserving spanning tree

A spanning tree T of a connected graph G is said to be *distance-preserving* from a vertex v in G if $d_T(u,v) = d_G(u,v)$ for every vertex $u \in V(G)$. To find a distance-preserving spanning tree, we use the *Breadth First Search* (BFS). This is an algorithm for traversing or searching tree or graph data structures. It starts at the tree *root* (an arbitrary vertex v) and explores the neighbour vertices first, before moving to the next level neighbours. This algorithm guarantees that we will get a spanning tree which is distance-preserving from the root vertex v.

We present two well known properties of graphs.

Theorem 2.1. *Let G be a connected graph. Then $rad(G) \leq diam(G) \leq 2\,rad(G)$.*

Proof. The inequality $rad(G) \leq diam(G)$ follows directly from the definitions of radius and diameter.
To prove the second inequality, let $u, v \in V(G)$ be a diametral pair of vertices, that is, $d(u,v) = diam(G)$. Furthermore, let $w \in V(G)$ be a central vertex. Therefore $ec\,(w) = rad(G)$. Observe that since $ec\,(w)$ is the maximum distance from w to all

other vertices, then $ec\,(w) \geq d(u,w)$ for every vertex $u \in V(G)$.
By the triangle inequality and the symmetry property,

$$\begin{aligned}
\text{LHS} &= diam(G) \\
&= d(u,v) \\
&\leq d(u,w) + d(w,v) \\
&= d(u,w) + d(v,w) \\
&\leq ec\,(w) + ec\,(w) \\
&= 2\,ec\,(w) \\
&= 2\,rad(G) \\
&= \text{RHS}.
\end{aligned}$$

Thus, we have shown that $diam(G) \leq 2\,rad(G)$, and the proof is complete.

□

Theorem 2.2. *For every vertex v of a connected graph G, there exists a spanning tree T of G that is distance-preserving from v.*

Proof. Let κ be the eccentricity of vertex v in G. For $1 \leq i \leq \kappa$, consider the distance layers

$$N_i(v)=\{u \in V(G) \,|\, d(u,v) \;=\; i\}$$

We want to construct a spanning tree T that is distance-preserving from v. To do this we are going to use the BFS algorithm. Since G is connected it follows that every vertex $u \neq v$ belongs to N_i for some $0 \leq i \leq \kappa$. Let vertex v be a root vertex. By connectivity every vertex $u \in N_i$ is adjacent to at least one vertex in the set $N_{i-1}(v)$ and possibly to vertices in sets $N_i(v)$ and $N_{i+1}(v)$. We begin the construction of tree T by joining v to all neighbour vertices in set $N_1(v)$. To further construct T we do it step by step for $1 \leq i \leq \kappa$.
Step 1: Level $N_i(v)$
For each $u \in N_i(v)$, delete all edges except for one edge that joins u to vertex of $N_{i-1}(v)$ or $N_{i+1}(v)$. Also remove every other edge that joins vertices in this set $N_i(v)$ (or $N_{i+1}(v)$) such that no two vertices in the same set are joined.
Step 2: Level $N_{i+1}(v)$
To proceed to the next level neighbour $N_{i+1}(v)$, we choose any vertex u in the previous level $N_i(v)$ and we join it to all neighbour vertices in set $N_{i+1}(v)$. We then use step 1 to make sure that no vertices in set $N_{i+1}(v)$ are joined. We repeat this

process for each vertex in set $N_i(v)$ until we get T. See for example Figure 2.1. The BFS creates a $u - v$ path for each $u \neq v$ in G to produce T and $V(G) = V(T)$. Hence, T spans G. Therefore, T is connected, and is distance-preserving from v.

We show that T is a tree. We only need to prove that T is acyclic. By contradiction, assume T has a cycle C. Let w be a vertex of C whose distance from v is maximum. Furthermore, let w_1 and w_2 be the vertices adjacent to w on C. Assume that $w \in N_k$. Since w_1 and w_2 are on C, w_i $(i = 1, 2)$ may belong to set N_k or N_{k-1}. If $w_1 \in N_k$ (or $w_2 \in N_k$), this means that w and w_1 (or w_2) are joined. This contradicts our construction of T, which requires that no two verticies in set N_i are joined. Now we look at the case where w_i $(i = 1, 2)$ belong to the set N_{k-1}. We earlier assumed that both w_1 and w_2 are adjacent to w on C. This means that w_1 and w_2 in set N_{k-1} are joined to the same vertex w in set N_k. This contradicts our construction of T which requires that only one edge in T joins a vertex in $N_i(v)$ to a vertex in $N_{i-1}(v)$. Therefore, T is acyclic and hence is a tree. □

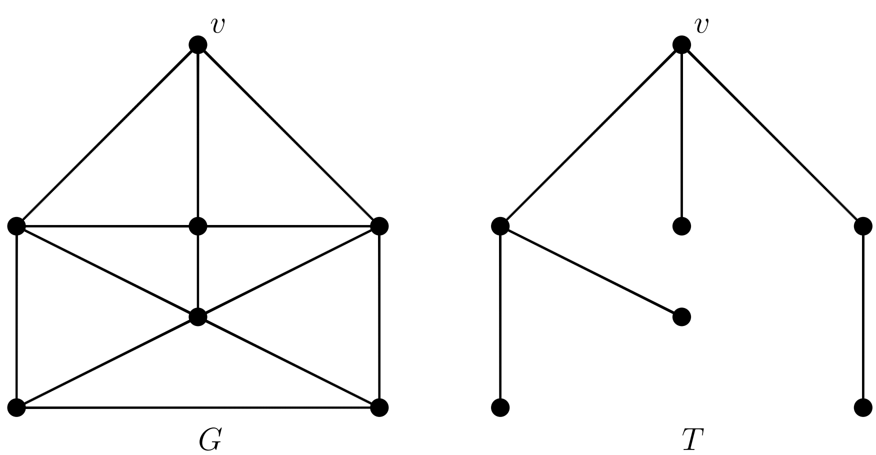


Figure 2.1: A connected graph G and a spanning tree T that is distance-preserving from v.

The following corollary shows that for every connected graph, it is possible to form a radius-preserving spanning tree.

Corollary 2.3. *Every connected graph G has a spanning tree T with rad(G) = rad(T).*

Proof. Let $v \in V(G)$. We use the method of Breadth First Search to construct a distance-preserving spanning tree T of G from a root vertex v. Since $d_G(v, u) \leq d_T(v, u)$ for all u in $V(G)$, we conclude that $\forall\, x \in V(G)$, $ec_G(x) \leq ec_T(x)$. But since we are constructing a distance-preserving spanning tree from v, we have

$d_G(v,u) = d_T(v,u)$ for all u in $V(G)$, and hence, $ec_T(v) = ec_G(v)$. Let v be a central vertex of G. Then v must also be a central vertex of T. We therefore have

$$rad\,(T) \geq rad\,(G) = ec_G(v) = ec_T(v) \geq rad\,(T).$$

Equality holds, and hence

$$rad\,(G) = rad\,(T).$$

$\square$

In the following we are going to use the properties of a distance-preserving spanning tree from Theorem 2.2 ($d_G(v,u) = d_T(v,u)$) and Corollary 2.3 ($rad(G) = rad(T)$) to prove Theorem 2.4.

Theorem 2.4. *(Erdős et al [10]) Let G be a connected graph of order n and minimum degree $\delta \geq 2$. Then*

$$rad(G) \leq \frac{3(n-3)}{2(\delta+1)} + 5.$$

Proof. Let v be a root vertex of G such that the eccentricity of v, $ec(v) = rad(G) = r$. Let $N_i(v) = \{u \in V(G) \,|\, d_G(v,u) = i\}$ for $0 \leq i \leq r$. Now we generate a distance-preserving spanning tree.

For each vertex $u \in N_i$, join it to a vertex $\acute{u} \in N_{i-1}$ by means of an edge such that $u\acute{u} \in E(G)$ $(1 \leq i \leq r)$. We do this for each $u \in N_i$ until we reach the root vertex v. Note that we have generated a spanning tree T which is distance-preserving from v. Therefore, by Theorem 2.2 we have $d_G(v,u) = d_T(v,u)$, $\forall y \in V(G) = V(T)$.

For an arbitrary vertex z, let $P_z = T\,(v,z)$ be the $v-z$ path in T. Define $N_{\leq j} = \bigcup_{0\leq i\leq j} N_i$, and $N_{\geq j} = \bigcup_{j\leq i\leq r} N_i$. Fix a vertex $z \in N_r$. See Figure 2.2. We say that a vertex $w \in V(G)$ is *related to* z if there exists a vertex $\acute{z} \in P_z \cap N_{\geq 5}$ and a vertex $\acute{w} \in P_w \cap N_{\geq 5}$ such that $d_G(\acute{z},\acute{w}) \leq 2$.
We show that there exists a vertex far from v that is not related to z.

Lemma 2.5. *There exists a vertex $w \in N_{\geq r-5}$ which is not related to z.*

Suppose to the contrary that every vertex $w \in N_{\geq r-5}$ is related to z. Let a be the only vertex in P_w which belongs to N_5. So $d_T(v,a) = 5$. Then, for any $y \in N_{\leq r-6}$,

$d_T(v, y) \leq r - 6$.

$$\begin{aligned} d_T(a, y) &\leq d_T(a, v) + d_T(v, y) \\ &\leq 5 + r - 6 \\ &= r - 1 \end{aligned}$$

This tell us that the distance between a and any vertex $y \in N_{\leq r-6}$ is at most $r - 1$.
Note: $d_T(v, \acute{z}) = d_T(v, a) + d_T(a, \acute{z})$, $d_T(v, w) = d_T(v, \acute{w}) + d_T(\acute{w}, w)$ and $d_T(v, w) \leq r$.
Let $w \in N_{\geq r-5}$ be an arbitrary vertex. Since w and a are related, there exists a vertex $\acute{w} \in P_w \cap N_{\geq 5}$ and vertex $\acute{a} \in P_a \cap N_{\geq 5}$ such that $d_T(\acute{w}, \acute{a}) \leq 2$.

$$\begin{aligned} d_T(a, w) &\leq d_T(a, \acute{z}) + d_T(\acute{z}, \acute{w}) + d_T(\acute{w}, w) \\ &\leq (d_T(v, \acute{z}) - d_T(v, a)) + 2 + (d_T(v, w) - d_T(v, \acute{w})) \\ &\leq (d_T(v, \acute{z}) - 5) + 2 + (r - d_T(v, \acute{w})) \\ &= r - 3 + d_T(v, \acute{z}) - d_T(v, \acute{w}). \end{aligned}$$

By the triangle inequality $d_T(v, \acute{z}) \leq d_T(\acute{z}, \acute{w}) + d_T(v, \acute{w})$ we get,

$$\begin{aligned} d_T(a, w) &\leq r - 3 + d_T(\acute{z}, \acute{w}) \\ &\leq r - 1. \end{aligned}$$

We have shown that $d_T(a, y) \leq r - 1$ for all $y \in V(G) = V(T)$ in the set $N_{\leq r-6}$. We have furthermore shown that $d_T(a, w) \leq r - 1$ for all $w \in V(G) = V(T)$ in the set $N_{\geq r-5}$. This means that $\forall y \in V(G) = V(T)$, $ec(a) \leq r - 1$ which contradicts the assumption that the $ec(v) = r$. Therefore it cannot be true that every vertex $w \in N_{\geq r-5}$ is related to z. Hence the lemma is shown.
If $w \in N_{\geq r-5}$ is not related to z, then we cannot find two vertices $\acute{z} \in P_z \cap N_{\geq 5}$ and $\acute{w} \in P_w \cap N_{\geq 5}$ such that $d_G(\acute{z}, \acute{w}) \leq 2$. For any i, let $\acute{N}_i$ and $\acute{\acute{N}}_i$ denote a set of all vertices in N_i whose distance is at most 1 from any vertex q in $P_w \cap N_{\geq 5}$ and $P_z \cap N_{\geq 5}$ respectively. Therefore $\acute{N}_i = \{q \in V(G) : d_T(q, P_z \cap N_{\geq 5}) \leq 1\}$ and $\acute{\acute{N}}_i = \{q \in V(G) : d_T(q, P_w \cap N_{\geq 5}) \leq 1\}$. Since z and w are not related,

$$\left(\cup_{i=4}^{r} \acute{N}_i \right) \cap \left(\cup_{i=4}^{r} \acute{\acute{N}}_i \right) = \emptyset.$$

By the definition of minimum degree,

$$\begin{aligned} |\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}| \geq \delta + 1 \text{ for all } 5 \leq i \leq r, \\ |\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}| \geq \delta + 1 \text{ for all } 5 \leq i \leq s, \end{aligned} \tag{2.1}$$

where

$$N = d_T(v,w) \geq r-5. \tag{2.2}$$

We show that $\displaystyle\sum_{i=5}^{r}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) \geq \sum_{i=4}^{r}|\acute{N}_i|$,

$$\begin{aligned}
\text{LHS} &= \sum_{i=5}^{r}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|)\\
&= (|\acute{N}_4| + |\acute{N}_5| + |\acute{N}_6| + \ldots + |\acute{N}_{r-1}|) + (|\acute{N}_5| + |\acute{N}_6| + |\acute{N}_7| + \ldots + |\acute{N}_{r-1}| + |\acute{N}_r|)\\
&\quad + (|\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}| + |\acute{N}_r| + |\acute{N}_{r+1}|)\\
&= |\acute{N}_4| + 2\,|\acute{N}_5| + 3\,(|\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|) + 2\,|\acute{N}_r| + |\acute{N}_{r+1}|.
\end{aligned}$$

Note:
(i) $|\acute{N}_4| + 2\,|\acute{N}_5| + 2\,|\acute{N}_r| \geq |\acute{N}_4| + |\acute{N}_5| + |\acute{N}_r|$
(ii) $3\,(|\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|) \geq |\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|$
(iii) $|\acute{N}_{r+1}| = 0$, since the $max_{y\in V(G)}\, d_G(v,y) = rad(G) = r$.
Therefore,

$$\begin{aligned}
\text{LHS} &= |\acute{N}_4| + 2\,|\acute{N}_5| + 3\,(|\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|) + 2\,|\acute{N}_r| && (2.3)\\
&\geq |\acute{N}_4| + |\acute{N}_5| + |\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}| + |\acute{N}_r|. && (2.4)
\end{aligned}$$

We have,

$$\text{RHS} = \sum_{i=4}^{r}|\acute{N}_i| = |\acute{N}_4| + |\acute{N}_5| + |\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}| + |\acute{N}_r|. \tag{2.5}$$

Therefore, by inequality (2.4) and equation (2.5), LHS $\geq$ RHS.
Similarly, it can be shown that $\displaystyle\sum_{i=5}^{s}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|) \geq \sum_{i=4}^{s+1}|\acute{\acute{N}}_i|$.
We know that $n = n_T = n_G$, then we have

$$n \geq |N_{\leq 3}| + \sum_{i=4}^{r}|\acute{N}_i| + \sum_{i=4}^{s+1}|\acute{\acute{N}}_i|. \tag{2.6}$$

Consider $|N_{\leq 3}|$, then we have $|N_{\leq 3}| = |N_3| + |N_2| + |N_1| + |N_0|$. By inequalities (2.1), $|N_3| + |N_2| + |N_1| \geq \delta + 1$ and $|N_0| = 1$. Therefore,

$$|N_{\leq 3}| \geq \delta + 2. \tag{2.7}$$

We want to show:

$\sum_{i=4}^{r} |\acute{N}_i| \geq \left\{ \sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) + 1 \right\}$ or equivalently we need to show that $\sum_{i=4}^{r} |\acute{N}_i| - \sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) \geq 1$. Let LHS$= \sum_{i=4}^{r} |\acute{N}_i|$. Using equation (2.5) the LHS $= \sum_{i=4}^{r} |\acute{N}_i| = |\acute{N}_4| + |\acute{N}_5| + |\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}| + |\acute{N}_r|$.
Consider the second term $\left(\sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}| \right)$ on the right hand side.
Then by equation (2.3),

$$\begin{aligned}\sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) &= \frac{1}{3}(|\acute{N}_4| + 2\,|\acute{N}_5| + 3\,(|\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|) + 2\,|\acute{N}_r|) \\ &= \frac{1}{3}|\acute{N}_4| + \frac{2}{3}\,|\acute{N}_5| + \frac{2}{3}\,|\acute{N}_r| + |\acute{N}_6| + |\acute{N}_7| + |\acute{N}_8| + \ldots + |\acute{N}_{r-1}|.\end{aligned}$$

So now we are going to show that LHS$-$RHS≥ 1.
We have LHS$-$RHS $= \frac{2}{3}|\acute{N}_4| + \frac{1}{3}|\acute{N}_5| + \frac{1}{3}|\acute{N}_r|$. Since $|\acute{N}_4| \geq 1$, $|\acute{N}_5| \geq 1$ and $|\acute{N}_r| \geq 1$, we get that $\frac{2}{3}|\acute{N}_4| + \frac{1}{3}|\acute{N}_5| + \frac{1}{3}|\acute{N}_r| \geq \frac{2}{3} + \frac{1}{3} + \frac{1}{3} \geq 1$. Therefore,

$$\begin{aligned}\text{LHS} - \text{RHS} &\geq 1 \\ \text{LHS} &\geq \text{RHS} + 1 \\ &\geq \left\{ \sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) + 1 \right\}.\end{aligned} \tag{2.8}$$

Also, we want to show that,

$$\sum_{i=4}^{s+1} |\acute{\acute{N}}_i| - \sum_{i=5}^{s} \frac{1}{3}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|) \geq 1.$$

Let LHS$= \sum_{i=4}^{s+1} |\acute{\acute{N}}_i|$ and RHS$= \sum_{i=5}^{s} \frac{1}{3}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|)$.
So, the LHS $= \sum_{i=4}^{s+1} |\acute{\acute{N}}_i| = |\acute{\acute{N}}_4| + |\acute{\acute{N}}_5| + |\acute{\acute{N}}_6| + |\acute{\acute{N}}_7| + |\acute{\acute{N}}_8| + \ldots + |\acute{\acute{N}}_{s-1}| + |\acute{\acute{N}}_s| + |\acute{\acute{N}}_{s+1}|$.
Note

$$\sum_{i=5}^{s} \frac{1}{3}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|) = \frac{1}{3}|\acute{\acute{N}}_4| + \frac{2}{3}\,|\acute{\acute{N}}_5| + |\acute{\acute{N}}_6| + |\acute{\acute{N}}_7| + |\acute{\acute{N}}_8| + \ldots + |\acute{\acute{N}}_{s-1}| + \frac{2}{3}\,|\acute{\acute{N}}_s| + \frac{1}{3}|\acute{\acute{N}}_{s+1}|). \tag{2.9}$$

We show that LHS$-$RHS≥ 1.
We have LHS$-$RHS $= \frac{2}{3}|\acute{\acute{N}}_4| + \frac{1}{3}\,|\acute{\acute{N}}_5| + \frac{1}{3}\,|\acute{\acute{N}}_s| + \frac{2}{3}|\acute{\acute{N}}_{s+1}|$. Since $|\acute{\acute{N}}_4| \geq 1$, $|\acute{\acute{N}}_5| \geq 1$ $|\acute{\acute{N}}_s| \geq 1$ and $|\acute{\acute{N}}_{s+1}| \geq 1$, we get that $\frac{2}{3}|\acute{\acute{N}}_4|+\frac{1}{3}\,|\acute{\acute{N}}_5|+\frac{1}{3}\,|\acute{\acute{N}}_s|+\frac{2}{3}|\acute{\acute{N}}_{s+1}| \geq \frac{2}{3}+\frac{1}{3}+\frac{1}{3}+\frac{2}{3} \geq 1$. Therefore,

$$\begin{aligned}
\text{LHS} - \text{RHS} &\geq 1 \\
\text{LHS} &\geq \text{RHS} + 1 \\
&\geq \left\{ \sum_{i=5}^{s} \frac{1}{3}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|) + 1 \right\}.
\end{aligned} \tag{2.10}$$

Substitute inequalities (2.1), (2.7), (2.8) and (2.10) into (2.6) and using the observation that $\sum_{i=a}^{r}(1) = r - a + 1$ we get,

$$\begin{aligned}
n &\geq |N_{\leq 3}| + \sum_{i=4}^{r} |\acute{N}_i| + \sum_{i=4}^{s+1} |\acute{\acute{N}}_i| \\
&\geq \delta + 2 + \left\{ \sum_{i=5}^{r} \frac{1}{3}(|\acute{N}_{i-1}| + |\acute{N}_i| + |\acute{N}_{i+1}|) + 1 \right\} + \left\{ \sum_{i=5}^{s} \frac{1}{3}(|\acute{\acute{N}}_{i-1}| + |\acute{\acute{N}}_i| + |\acute{\acute{N}}_{i+1}|) + 1 \right\} \\
&\geq \delta + 4 + \frac{1}{3}(r-4)(\delta+1) + \frac{1}{3}(s-4)(\delta+1).
\end{aligned}$$

From inequality (2.2) we have $s \geq r - 5$,

$$\begin{aligned}
\frac{1}{3}(s-4) &\geq \frac{1}{3}(r-9) \\
\frac{1}{3}(s-4)(\delta+1) &\geq \frac{1}{3}(r-9)(\delta+1).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\delta + 4 + \frac{1}{3}(r-4)(\delta+1) + \frac{1}{3}(s-4)(\delta+1) &\geq (\delta+1) + 3 + \frac{1}{3}(r-4)(\delta+1) + \frac{1}{3}(r-9)(\delta+1) \\
&\geq \frac{1}{3}(\delta+1)(r-4+r-9+3) + 3 \\
&\geq \frac{1}{3}(2r-10)(\delta+1) + 3.
\end{aligned}$$

Rearranging we get,

$$rad(G) = rad(T) \leq \frac{3(n-3)}{2(\delta+1)} + 5, \text{ by Corollary 2.3.}$$

We have shown that there exists a vertex far from v that is not related to z. Then the proof of Theorem 2.4 is complete. $\square$

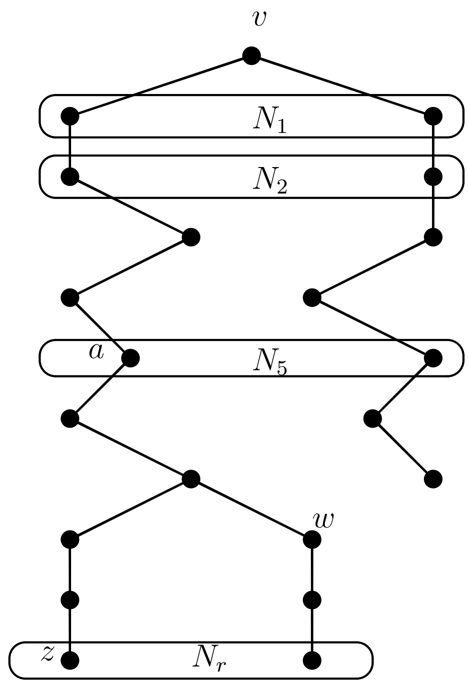


Figure 2.2: Spanning tree T illustrating path P_z and path P_w.

In the next section, we consider bounds in terms of order and size.

2.2 Bounds in terms of Order and Size

In this section we will use the Ding et al [8] spanning tree construction method to find an upper bound on diameter in terms of order and size. The Ali et al [1] paper uses the Ding et al method to derive this bound on diameter.

2.2.1 Upper bound on Diameter

We are going to prove Theorem 2.6 to illustrate how Ding et al constructed a spanning tree. We are not going to prove Theorem 2.7, but we will use it later to help derive a bound on diameter on a constructed spanning tree.

For the next theorem we will need the following definition. For any integer $n > t \geq 2$, define

$$f(n,t) = \begin{cases} n + \frac{1}{2}(t^2 - 4) & \text{if } n = t+2 \text{ and } t \text{ is even} \\ n + \frac{1}{2}(t^2 - 5) & \text{if } n = t+2 \text{ and } t \text{ is odd} \\ n + \frac{1}{2}(t^2 - t - 2) & \text{if } n = t+1 \text{ or } n \geq t+3. \end{cases}$$

Theorem 2.6. *(Ding et al [8]) For all integers $n > t \geq 2$ there is a connected graph with n vertices and $f(n,t)$ edges in which every spanning tree has $\leq t$ leaves.*

Theorem 2.7. *(Ding et al [8]) For all integers $n > t \geq 2$ every connected graph with n vertices and $> f(n,t)$ edges has a spanning tree with $> t$ leaves.*

The following observation will help prove part of Theorem 2.6.

Observation 1. *A graph and its complement cannot both be disconnected.*

Proof. Let G be a disconnected graph and $\overline{G}$ be its complement. Consider two vertices u and v in both G and $\overline{G}$. If u and v are not adjacent in G, then they must be adjacent in $\overline{G}$. Hence there exists a $u-v$ path in $\overline{G}$. If u and v are adjacent in G, then u and v must belong to the same component of G. Now let w be some vertex in another component of G. Since w is in a different component in G this implies that the edges uw and vw do not exist in G but exist in $\overline{G}$. Therefore in $\overline{G}$ there exists a $u-w-v$ path. Hence there exists a path between any two vertices of $\overline{G}$, then $\overline{G}$ is connected. □

The following is a proof of Theorem 2.6.

Proof. We are going to prove this theorem using two cases.
Case 1:
If $n = t+2$, let H be a graph with n vertices and $\lceil \frac{1}{2}n \rceil$ edges, in which every vertex has degree ≥ 1. Note that H is disconnected. Let P_n be a path of order n, so $H \simeq \frac{n}{2} P_2$ if n is even, and $H \simeq \frac{n-3}{2} P_2 \cup P_3$ if n is odd. Let G be its complement, then by Observation 1, G is connected.
Now we show that $|E(G)| = f(n,t)$, for t even and for t odd.
For t even:

$$\begin{aligned}
\text{LHS} &= |E(G)| \\
&= \binom{n}{2} - \left\lceil \frac{n}{2} \right\rceil \\
&= \frac{n(n-1)}{2} - \frac{n}{2} \qquad\qquad \text{substituting } n = t+2 \\
&= \frac{(t+2)(t+2-1)}{2} - \frac{n}{2}
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{n}{2} + \frac{1}{2}(t^2 + 3t + 2) \\
&= -\frac{n}{2} + \frac{1}{2}(t^2 + 3(n-2) + 2) \quad \text{for } t = n-2 \\
&= -\frac{n}{2} + \frac{1}{2}(t^2 + 3n - 4) \\
&= n + \frac{1}{2}(t^2 - 4).
\end{aligned}
$$

For t odd:

$$
\begin{aligned}
\text{LHS} &= |E(G)| \\
&= \binom{n}{2} - \left\lceil \frac{n}{2} \right\rceil \\
&= \frac{n(n-1)}{2} - \frac{n+1}{2} && \text{substituting } n = t+2 \\
&= \frac{(t+2)(t+2-1)}{2} - \frac{n+1}{2} \\
&= -\frac{n+1}{2} + \frac{1}{2}(t^2 + 3t + 2) \\
&= -\frac{n+1}{2} + \frac{1}{2}(t^2 + 3(n-2) + 2) && \text{for } t = n-2 \\
&= -\frac{n+1}{2} + \frac{1}{2}(t^2 + 3n - 4) \\
&= n + \frac{1}{2}(t^2 - 5).
\end{aligned}
$$

Now we show that every spanning tree T of G has at most t leaves. Since $t = n-2$, we will show that T has maximum $n-2$ leaves. In both even and odd cases each vertex of G has degree $(n-2)$ except for the central vertex in P_3 in the odd case. The degree of the central vertex in P_3 is $(n-3)$. We first look at the even case. Since every vertex of the even case has degree $(n-2)$, let w be a vertex in G. The vertex w is adjacent to $(n-2)$ vertices, so every possible spanning tree T is constructed such that w is adjacent to all $(n-2)$ vertices. There is a vertex u which is adjacent to w in G but not adjacent to w in H. Since G is connected, u must be connected to some vertex x which is adjacent to w. The vertex w is adjacent to $(n-2)$ vertices which gives $(n-2)$ leaves. But the addition of the ux edge decreases the number of leaves by one. We note that u in T is adjacent to x only. Therefore, u is a leaf, hence the number of leaves of T is,

$$(n-2) + 1 - 1 \;=\; n-2 \;=\; t.$$

Now we look at the odd case. Graph G has vertices of degree $(n-2)$ and a vertex of degree $(n-3)$. If we choose any vertex say w of degree $(n-2)$, this follows exactly the even case in which we showed that every spanning tree T has at most t leaves. Suppose w is the vertex of degree $(n-3)$. By applying the same logic applied to w in the even case we get that the maximum number of leaves is

$$(n-3)+1-1 = n-3 = t-1.$$

Thus, if w has a degree of $(n-3)$ we get $(t-1)$ leaves which are less than t leaves. Therefore, we can conclude that every spanning tree T of G has at most t leaves.

Case 2:
Consider $n = t+1$ or $n \geq t+3$. Let G be obtained from graph K_{t+1} by replacing some edge $e = xy$ by a path P with $n-t$ edges between vertices x and y, the ends of e. By definition a complete graph is connected. So, replacing e by the path P to form graph G does not effect the connectivity.
We show that $|E(G)|= f(n,t)$. By the construction of G, we have:

$$\begin{aligned}
\text{LHS} &= |E(G)| \\
&= \binom{t+1}{2} - 1 + (n-t) \\
&= \frac{(t+1)(t+1-1)}{2} - 1 + (n-t) \\
&= \frac{t^2+t}{2} - 1 + (n-t) \\
&= n + \frac{1}{2}(t^2 - t - 2) \\
&= \text{RHS}.
\end{aligned}$$

We know that every vertex of a complete graph K_n has degree $n-1$. For our case we have a complete graph K_{t+1} with every vertex having degree $(t+1)-1$. Note that a tree of a complete graph K_n has at most $n-1$ leaves. Hence $n-1$ is the maximum number of leaves of a spanning tree of K_n. The order of G is equal to $(t+1)-1+n-t = n$, which satisfies the hypothesis of the theorem.
Now we show that every spanning tree T of G has at most t leaves. We will show that t is the maximum number of leaves every spanning tree T can possibly have. In G we have, $\acute{K}_{t+1}$, a complete graph with one edge, e, missing and P. To construct T of G with maximum t leaves, we first look at $\acute{K}_{t+1}$ and we see that every spanning tree T will have a maximum number of leaves equal to $(t+1)-1 = t$. Then we add the path P reducing by one the number of leaves of T. But since we are constructing

a spanning tree we remove one edge from P which allows at most one extra leaf in T. Then the maximum number of leaves of T is,

$$t-1+1=t.$$

Therefore, every spanning tree has at most t leaves. Our proof is complete. $\square$

We will present Theorem 2.9, an upper bound on diameter in terms of order and size. We will show that this bound is close to sharp on the diameter except for a small difference of 2 when compared to the diameter of a path-complete graph.
To help prove Theorem 2.9 we are going to use Theorem 2.8 (stated without proof) which uses the Ding et al approach to construct a spanning tree.

Theorem 2.8. *(Ali et al [1]) Let G be a connected graph of order n and size m. If $m \geq n+\frac{1}{2}t(t-1)$, then G has a spanning tree with more than t leaves.*

A *path-complete* graph, $PK_{n,m}$, is a graph obtained by taking one copy of a path, P, and one copy of a complete graph K_a, and joining one end vertex of P to one or more vertices of K_a. See example Figure 2.3.
The path-complete graph, for $n-1 \leq m \leq \binom{n}{2}$, has diameter

$$diam(PK_{n,m}) = n + \left\lfloor \frac{1}{2} - \left(\sqrt{2m-2n+\frac{17}{4}}\right) \right\rfloor.$$

This equality is attained if an end vertex of the path is joined to all vertices of the complete graph except for one vertex.

It can be seen in Figure 2.3 that the diameter of $PK_{9,22}$ is 4. Applying the values $(n=9, m=22)$ into the formula we find:

$$\begin{aligned} diam(PK_{9,22}) &= 9 + \left\lfloor \frac{1}{2} - \left(\sqrt{2(22)-2(9)+\frac{17}{4}}\right) \right\rfloor \\ &= 9 + \lfloor -5 \rfloor \\ &= 9-5 \\ &= 4. \end{aligned}$$

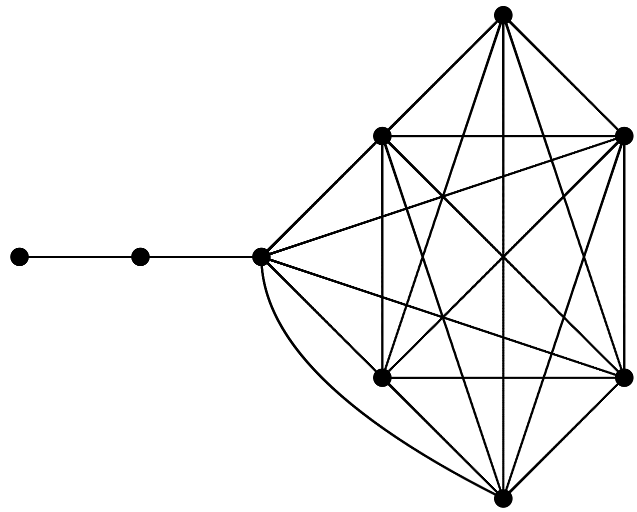


Figure 2.3: A path-complete graph $PK_{9,22}$ with diameter = 4.

For the following proof, note that since $x-1 < \lfloor x \rfloor \leq x$, we have that

$$x-1 < \lfloor x \rfloor \leq x$$
$$x-1 < \lfloor x \rfloor,$$

and hence
$$-x+1 > -\lfloor x \rfloor. \tag{2.11}$$

Theorem 2.9. *(Ali et al [1]) Let G be a connected graph of order n and size m, $m \geq n$. Then G has a spanning tree T of diameter at most*

$$n + \frac{1}{2} - \sqrt{2m - 2n + \frac{1}{4}}.$$

Proof. Let t be the largest integer such that $m \geq n + \dfrac{1}{2}t(t-1)$. We prove our theorem in two cases.

Case 1: $n \neq t+2$.

By Theorem 2.8, G has a spanning tree T with at least $t+1$ leaves. Let d be the diameter of T. The number of leaves of T, $L(T)$, is at most the difference between the order and the number of internal vertices on a diametral path of T. We know that there are $(d-1)$ internal vertices on the diametral path of T. Therefore, we have $L(T) \leq n-(d-1)$. By Theorem 2.8, $L(T) \geq t+1$. This yield

$$t+1 \leq L(T) \leq n-d+1.$$

$$d \leq n - L(T) + 1 \leq n - t. \tag{2.12}$$

Since t is the largest integer such that $m \geq n + \frac{1}{2}t(t-1)$:

$$t^2 - t + 2m - 2n \leq 0.$$

$$\text{Using the quadratic formula,} \quad t_{1,2} = \frac{-(-1) \pm \sqrt{(-1)^2 - 4(1)(2m-2n)}}{2}$$
$$= \frac{1}{2} \pm \sqrt{2m - 2n + \frac{1}{4}}$$
$$\text{For } t \text{ is the largest integer,} \quad t \geq \left\lfloor \frac{1}{2} + \sqrt{2m - 2n + \frac{1}{4}} \right\rfloor.$$

We take this inequality and substitute it back into inequality (2.12).

$$d \leq n - t \leq n - \left\lfloor \frac{1}{2} + \sqrt{2m - 2n + \frac{1}{4}} \right\rfloor$$
$$d \leq n - \left\lfloor \frac{1}{2} + \sqrt{2m - 2n + \frac{1}{4}} \right\rfloor.$$

By observation (2.11), and substituting $x = \frac{1}{2} + \sqrt{2m - 2n + \frac{1}{4}}$, we conclude that

$$d \leq n + \frac{1}{2} - \sqrt{2m - 2n + \frac{1}{4}},$$

as desired.

Case 2: $n = t + 2$.

Therefore $t = n - 2$, then $m \geq n + \frac{1}{2}t(t-1) = \frac{1}{2}n\,(n\,-\,3) + 3$.

Claim 1. *G has a vertex of degree at least $(n-2)$.*

We prove this claim by contradiction. Assume that G does not have a vertex of degree at least $(n-2)$. So, the largest degree of every vertex in G is $(n-3)$.
By the Handshaking Lemma

$$\begin{aligned} m &= \frac{1}{2} \sum_{v \in V(G)} d(v) \\ &\leq \frac{1}{2} n\,(n-3) \\ &< \frac{1}{2} n\,(n\,-\,3) + 3. \end{aligned}$$

This is a contradiction to our choice of t ($m \geq \frac{1}{2}n\,(n\,-\,3) + 3$), and the claim is shown.

Claim 2. *G has a spanning tree of diameter at most 3.*

We show that G has a spanning tree T' of diameter at most 3. We have shown in Claim 1 that G has some vertex, say w, of degree at least $(n-2)$, so the vertex w is adjacent to at least $(n-2)$ vertices. Since any vertex in graph G of order n can have a maximum of $(n-1)$ degree, we are only going to consider the vertex w to have degree $(n-2)$ or $(n-1)$. If w has degree $(n-2)$, there must exist some vertex u in G not adjacent to vertex w, but adjacent to some vertex x which is adjacent to w. Applying the BFS on G we are able to construct a spanning tree T' with diameter 3. On the other hand, if w has degree $(n-1)$, applying the BFS on G we get a spanning tree T' of diameter 2. We have shown that in G we can get a spanning tree T' with diameter 2 or 3. Therefore

$$d_{T'} \leq 3.$$

Our claim is shown.

Claim 3. $n + \dfrac{1}{2} - \sqrt{2m - 2n + \dfrac{1}{4}} = 3.$

We show that $3 \geq n + \dfrac{1}{2} - \sqrt{2m - 2n + \dfrac{1}{4}}$, given $m \geq \frac{1}{2}\, n\,(n-3) + 3$.

$$\begin{aligned} \text{So,} \qquad\qquad 2m &\geq n(n-3) + 6 \\ 2m &\geq n^2 - 3n + 6 \\ 2m - 2n + \frac{1}{4} &\geq n^2 - 5n + 6 + \frac{1}{4}. \end{aligned} \tag{2.13}$$

We would like to introduce a square root on both sides of (2.13.)
Since $m \geq n$, we have

$$2m - 2n \geq 0$$
$$2m - 2n + \frac{1}{4} \geq \frac{1}{4}.$$

The hypothesis of Theorem 2.9 ($m \geq n$) implies that $n \geq 3$. Thus, $n - \frac{5}{2} \geq 0$. But $n^2 - 5\,n + 6 + \frac{1}{4} = (n - \frac{5}{2})^2$. Hence, we can introduce the square roots on both sides of inequality (2.13.).

$$
\begin{aligned}
\sqrt{2m-2n+\frac{1}{4}} &\geq \sqrt{n^2-5n+\frac{25}{4}} \\
\sqrt{2m-2n+\frac{1}{4}} &\geq \sqrt{\left(n-\frac{5}{2}\right)^2} \\
-\sqrt{2m-2n+\frac{1}{4}} &\leq -\sqrt{\left(n-\frac{5}{2}\right)^2} \\
-\sqrt{2m-2n+\frac{1}{4}} &\leq -\left|n-\frac{5}{2}\right| \\
\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}} &\leq -n+\frac{5}{2}+\frac{1}{2} \\
n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}} &\leq 3,
\end{aligned}
$$

as required.

However, if $3 > n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}}$, we get $2 \geq n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}}$ which implies

$$
\begin{aligned}
2 &\geq n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}} \\
2m-2n+\frac{1}{4} &\geq \left(n-\frac{3}{2}\right)^2 \\
2m &\geq n^2-n+2 \\
m &\geq \frac{n^2-n}{2}+1 \\
m &\geq \binom{n}{2}+1.
\end{aligned}
$$

This is a contradiction since $m \leq \binom{n}{2}$ the number of edges in a complete graph K_n.

Therefore $3 \not> n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}}$. Hence $3 = n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}}$.

Our claim is shown.

By Claims 2 and 3,

$$
\begin{aligned}
d_{T'} &\leq 3 = n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}} \\
d_{T'} &\leq n+\frac{1}{2}-\sqrt{2m-2n+\frac{1}{4}}.
\end{aligned}
$$

Case 2 is shown, with $T = T'$ and the proof of Theorem 2.9 is complete.

$\square$

Theorem 2.9 has shown that there is a spanning tree T for which $d_T \leq n + \frac{1}{2} - \sqrt{2m - 2n + \frac{1}{4}}$. This bound is sharp up to a small additive constant. Note that the diameter of the path-complete graph, $PK_{n,m}$ and the bound of Theorem 2.9, differ by at most 2.

The Ali et al paper [1] uses the result of Theorem 2.9 to derive an almost sharp upper bound on the average eccentricity of graphs of given order and size. This is stated here without proof.

Theorem 2.10. *(Ali et al [1]) Let G be a connected graph of order n and size m, $m \geq n$. Then*

$$avec\,(G) \leq \frac{3}{4}n - \frac{m}{2n} - \frac{1}{2}\sqrt{2m - 2n + \frac{1}{4}} + \frac{1}{4p}\sqrt{2m - 2n + \frac{1}{4}} + \frac{1}{8p} + \frac{3}{4}.$$

In the next section, we investigate upper bounds on average distance and diameter in terms of order and minimum degree.

2.3 Bounds in terms of Order and Minimum degree

We consider how a constructed spanning tree by the methods of Dankelmann and Entringer [7] and Mukwembi [15] can be used to bound average distance and diameter respectively.

2.3.1 Upper bound on Average Distance

We now present an upper bound on the average distance for graphs in terms of order and minimum degree.

Definition 1. *The kth power of G, denoted by G^k, is a graph with the same vertex set as G, in which two vertices $u \neq v \in V(G)$ are adjacent if $d_G(u, v) \leq k$.*

Definition 2. *A k packing A of G is a subset of $V(G)$ such that for all $u, v \in A$, $d(u, v) > k$. A k packing set A is maximal if A is a set with the largest cardinality.*

Definition 3. *Let $A \subset V(G)$, then $d_G(x, A)$, the distance between a vertex x and A, is defined as $min_{v \in A}\, d_G(x, v)$.*

The next definition is motivated as follows. Let G be a graph with vertices that act as a host of facilities. And let $\mu(G)$ be the expected distance between any pair of randomly chosen vertices or facilities. Note that every vertex hosts exactly one facility.

Now assume that some vertices host more than one facility. Furthermore, let the distance between the two facilities located on the same host be zero.

Let $c(x)$ be the number of facilities located in vertex x. Hence $N = \sum\limits_{x \in V(G)} c(x)$ is the total number of facilities. Then the expected distance between two randomly selected distinct facilities equals $\binom{N}{2}^{-1} \sum\limits_{x,y \in V(G)} c(x)c(y)d(x, y)$.

Definition 4. *For a weighted graph G with weight function $c : V(G) \to Z$ define the distance of G with respect to c by*

$$\sigma_c(G) = \textstyle\sum_{x,y \in V(G)} c(x)c(y)d(x, y),$$

and the average distance of G with respect to c by

$$\mu_c(G) = \binom{N}{2}^{-1} \sigma_c(G),$$

where $N = \sum\limits_{v \in V(G)} c(x)$ is the total weight of the vertices in G.

We are going to use Lemma 2.11 (stated without proof) to help prove Theorem 2.12.

Lemma 2.11. *Let G be a weighted graph with weight function c, and let k, N be positive integers, N a multiple of k such that $c(v) \geq k$ for every vertex v of G and $\sum\limits_{v \in V(G)} c(v) \leq N$. Then*

$$\mu_c(G) \leq \frac{N-k}{N-1} \frac{N+k}{3k}.$$

Equality holds if and only if G is a path and $c(v) = k$ for every $v \in V(G)$.

Theorem 2.12. *(Dankelmann and Entringer [7]) Let G be a connected graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \leq \frac{n}{\delta + 1} + 5.$$

Apart from the additive constant, this inequality is best possible.

Proof. We first construct a set $A \subset V(G)$ using the following procedure. Choose an arbitrary vertex v of G and let $A = \{v\}$. If there is a vertex x such that $d(x, A) = 3$, add x to A. Add all such vertices x such that $d(x, A) = 3$ until every vertex not in A is within distance 2 of A. The resulting A is a maximal 2 packing.
Then we note that $G(A)$, the subgraph induced by A is not connected, it is a collection of vertices of degree zero. But by definition of the kth power, for $k = 3$, $G^3(A)$ is connected.

We construct a forest $T_1 \leq G$. T_1 consists of all the vertices in A and their neighbours. To construct T_2 from T_1, we form a larger subgraph of G such that we join two neighbours of two different elements of A. Observe that we must join $|A| - 1$ edges of G to get a tree T_2.
We now construct a spanning tree T of G.
Case 1: $V(T_2) = V(G)$.
Then $T_2 = T$ is a spanning tree of G, and case 1 is complete.
Case 2: $V(T_2) \subset V(G)$.
We choose any vertex u of $V(G)$ that is not in T_2. Since G is connected, u is adjacent to some vertex w in T_2 by the edge uw such that the distance from u to some element in A is at most 2. To form T we take T_2 and add the edge uw such that $E(T) = E(T_2) \cup \{uw\}$. Repeat this procedure until all vertices of G are in T. Hence T is a spanning tree of G.

We now prove that

$$\mu(T) \leq \frac{n}{\delta + 1} + 5.$$

We consider the vertex set of T. For every vertex x in T find the unique vertex u in A which is closest to it and assign $x_A = u$. We define a weight function $c : V(T) \to Z$ by $c(u) = |\{x \in V(T) | x_A = u\}|$. Note

$$c(u) = 0 \,\text{if}\, u \notin A, \text{since all}\, x_A \in A. \tag{2.14}$$

Furthermore $c(u) \geq \delta + 1$ for all $u \in A$, since $d(u) \geq \delta$ and if $u \in A$ then $u_A = u$. Since A is a maximal 2 packing, the number of facilities $c(u)$ (let us say the number of facilities hosted by a vertex is equivalent to the weight of each vertex) will not move with a distance exceeding 2 if we moved them closer to A. Consider two vertices not in A. If they are moved closer to A their distance will change up to 4.

Hence the average of the distances between all facilities can change by up to four, since each distance between two facilities can change by up to four. Hence

$$\mu(T) \leq \mu_c(T) + 4. \tag{2.15}$$

We now turn our focus to $T^3(A)$. Recall that by (2.14), only the vertices in A have non zero weight. $T^3(A)$ is connected. We note that the distance between any two vertices of A in T is exactly $3l$ ($l \geq 1$), and this distance has been reduced to l in $T^3(A)$. This implies,

$$\mu_c(T) \leq 3\mu_c(T^3(A)). \tag{2.16}$$

We apply Lemma 2.11, taking $k = \delta + 1$. Recall $c(u) \geq \delta + 1$ for each $u \in A$. We follow the same idea as in Lemma 2.11. We choose an integer N to be the least multiple of k for which $N \geq n$, then

$$N \leq n + \delta. \tag{2.17}$$

$$\begin{aligned}
\mu_c(T^3(A)) &\leq \frac{N-(\delta+1)}{N-1} \cdot \frac{N+(\delta+1)}{3(\delta+1)} \\
&= \frac{N-\delta-1}{N-1} \cdot \frac{N+\delta+1}{3(\delta+1)} \\
&= \frac{N^2-1-2\delta-\delta^2}{3(N-1)(\delta+1)} \\
&= \frac{(N-1)(N+1)}{3(N-1)(\delta+1)} - \frac{\delta^2+2\delta}{3(N-1)(\delta+1)} \\
&= \frac{N+1}{3(\delta+1)} - \frac{\delta^2+2\delta}{3(N-1)(\delta+1)} \qquad (2.18) \\
&\leq \frac{N+1}{3(\delta+1)}. \qquad (2.19)
\end{aligned}$$

Inequality (2.18) is less than or equal to inequality (2.19) since at $\delta = 0$ inequality (2.18) is equal to inequality (2.19). But for $\delta > 0$ inequality (2.18) decreases to be less than inequality (2.19).

Now we consider inequalities (2.15), (2.16), (2.19) and also (2.17). Then

$$\begin{aligned}
\mu(T) &\leq \mu_c(T) + 4 \\
&\leq 3\mu_c(T^3(A)) + 4 \\
&\leq 3\frac{N+1}{3(\delta+1)} + 4 \\
&\leq \frac{(n+\delta)+1}{(\delta+1)} + 4 \\
&= \frac{n}{(\delta+1)} + \frac{\delta+1}{(\delta+1)} + 4 \\
&= \frac{n}{(\delta+1)} + 5,
\end{aligned}$$

as desired.

Finally, we show that the bound of Theorem 2.12 is best possible except for a small additive constant. For given integers n, δ, k with $n = k\,(\delta+1)$, let $G_{n,\delta}$ be a graph obtained from the union of disjoint copies $(G_1, G_2, G_3 \ldots, G_k)$ of the complete graph $K_{\delta+1}$, removing an edge from each copy except the end copies, then connecting copies by means of the ends of the removed edges. For example, see Figure 2.4 for $k = 4$. As shown in [14], $G_{n,\delta}$ has order n, minimum degree δ and

$$\mu(G_{n,\delta}) > \frac{n}{\delta+1}.$$

Hence, every spanning tree T of $G_{n,\delta}$ has an average distance greater than $\frac{n}{\delta+1}$ and so the bound is best possible apart from the value of an additive constant. $\square$

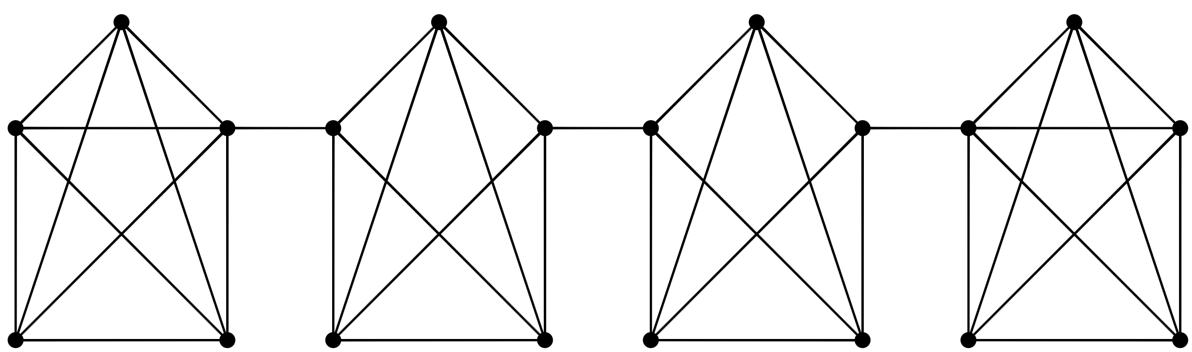


Figure 2.4: Graph $G_{20,4}$.

.

We will illustrate the process of the above proof with an example.

Example 1. *Consider the graph G given in Figure 2.5, with $n = 23$. The 6 vertices (u, d, i, o, m, t) form the set A, which is a maximal 2 packing of G.*

We construct a forest $T_1 \leq G$. Recall that T_1 consists of all the vertices in A and their neighbours.

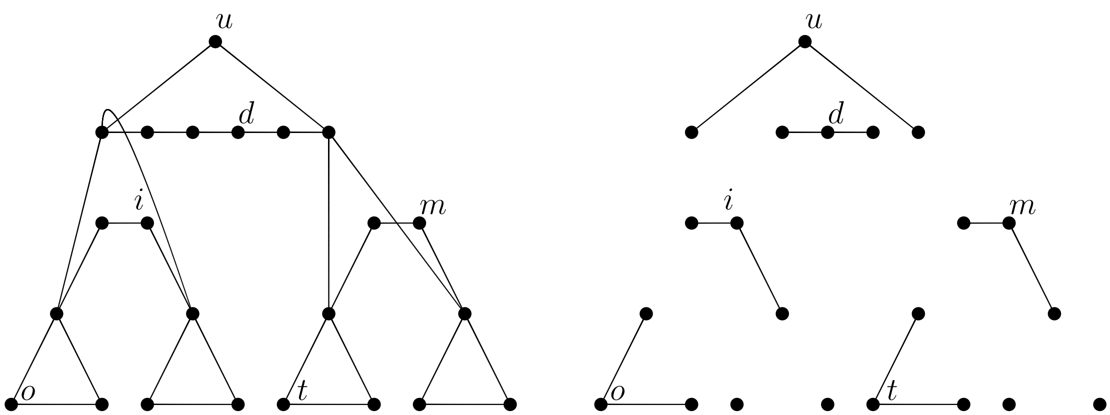


Figure 2.5: Graph G with $n = 23$ and Forest T_1.

To construct T_2 from T_1, we form a larger subgraph of G such that we join two neighbours of two different elements of A. Observe that we must join $|A| - 1 = 6 - 1 = 5$ edges of G to get a tree T_2. We then construct a spanning tree T from T_2 by adding the missing vertices $\{b, q, r, v, w\}$ and joining each of them to a vertex in T_2 using an edge already in G. The resulting graph is a spanning tree T. Both graphs are given in Figure 2.6.

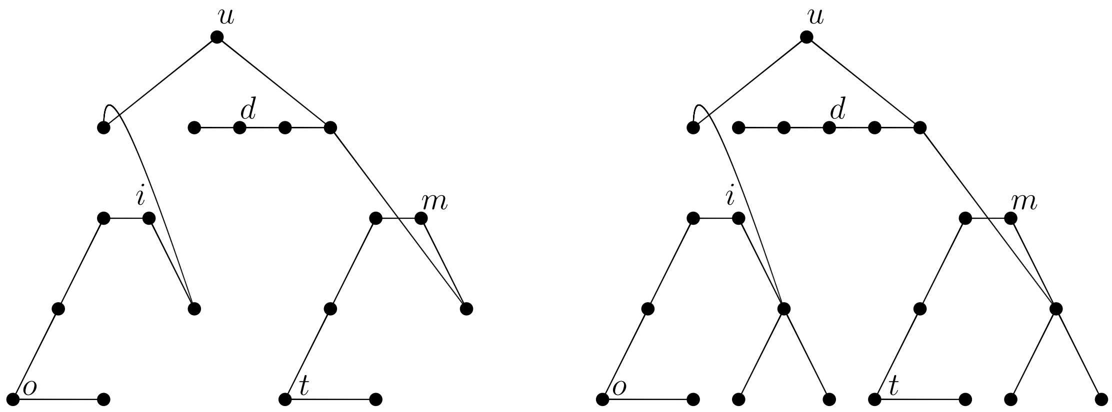


Figure 2.6: Tree T_2 and Spanning tree T.

Hence from a connected graph G, we have obtained a spanning tree T. Using the spanning tree T, we add the weights using the weight function $c : V(T) \to Z$ by $c(u) = |\{x \in V(T)|x_A = u\}|$ to get the weighted graph T. See Figure 2.7.
Note that only the vertices in A have non zero weight and the rest of the vertices have zero weight. Now using only the vertices in A we form graph $T^3[A]$.

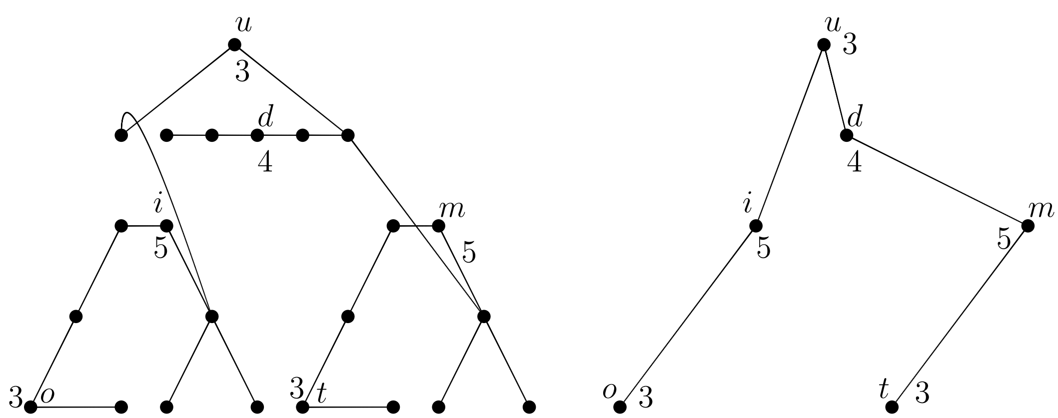


Figure 2.7: Weighted graph T and graph $T^3[A]$.

The example has illustrated how to form a spanning tree T from G.

The Dankelmann and Entringer paper [7] uses the same proof technique of Theorem 2.12 to prove Theorem 2.13 for a triangle-free graph. Similarly, Theorem 2.14

for a C_4-free graph uses the Dankelmann and Entringer [7] technique of constructing a spanning tree to find an upper bound on the average distance of the constructed spanning tree in terms of order and minimum degree. Both theorems are stated here without proof.

Theorem 2.13. *Let G be a connected triangle-free graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \leq \frac{2}{3}\frac{n}{\delta} + \frac{25}{3}.$$

Apart from the additive constant, this inequality is best possible.

Theorem 2.14. *(i) Let G be a connected C_4-free graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \leq \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor \delta/2 \rfloor + 1} + \frac{29}{3}.$$

(ii) There exists an infinite number of C_4-free graphs with n vertices and minimum degree δ such that, for every spanning tree T of G,

$$\mu(T) \leq \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor \delta/2 \rfloor + 2} + \frac{29}{3} + O(1).$$

Furthermore, the paper gives us a corollary for each of the theorems in terms of an upper bound of the average distance of G. Observe that $\mu(G) \leq \mu(T)$ if T is a spanning tree of G. The corollaries of the theorems are as follows respectively,

Corollary 2.15. *Let G be a connected triangle-free graph with n vertices and minimum degree δ. Then*

$$\mu(G) \leq \frac{2}{3}\frac{n}{\delta} + \frac{25}{3}.$$

Apart from the additive constant, this inequality is best possible.

Corollary 2.16. *Let G be a connected C_4-free graph with n vertices and minimum degree δ. Then*

$$\mu(G) \leq \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor \delta/2 \rfloor + 1} + \frac{29}{3}.$$

We note that Kouider and Winkler [14] proved that every connected graph with n vertices and minimum degree δ has, $\mu(G) \leq \dfrac{n}{\delta + 1} + 2$.

2.3.2 Upper bound on Diameter

We now present a bound by Mukwembi [15] (Theorem 2.18) on the diameter of a graph in terms of order, minimum degree and the irregularity index. We are going to use the Dankelmann and Entringer [7] technique of constructing a spanning tree to find an upper bound on the diameter of the constructed spanning tree in terms of order, minimum degree and the irregularity index of a graph.
We begin by proving a proposition on a bound of the diameter of a graph with a prescribed irregularity index.

Proposition 2.1. *Let G be a connected graph of order n. The diameter of G satisfies the inequality*

$$diam(G) \leq n - \eta + 1,$$

where η is the irregularity index of G. Moreover, this inequality is sharp.

Proof. We first show that $\Delta \leq n - diam(G) + 1$. Consider a vertex v with maximum degree, $d(v) = \Delta$ and the fact that every vertex can be adjacent to at most three consecutive vertices on a diametral path P. Then,
Case 1: $v \notin P$.
v is adjacent to at most 3 consecutive vertices of P.

$$\begin{aligned} n &\geq (diam(G) + 1) + \Delta + 1 - 3 \\ &\geq diam(G) - 1 + \Delta \\ \Delta &\leq n - diam(G) + 1. \end{aligned}$$

Case 2: $v \in P$.

$$\begin{aligned} n &\geq (diam(G) + 1) + \Delta - 2 \\ &\geq diam(G) - 1 + \Delta \\ \Delta &\leq n - diam(G) + 1. \end{aligned}$$

Hence we have shown that $\Delta \leq n - diam(G) + 1$.
Now by the definitions of Δ, η and the degree sequence we get that $\eta \leq \Delta$, therefore $\eta \leq n - diam(G) + 1$. Hence

$$diam(G) \leq n - \eta + 1,$$

as desired.
The following example graph illustrates that the inequality $diam(G) \leq n - \eta + 1$ is sharp.

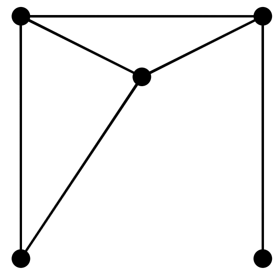


Figure 2.8: Graph $G_{5,3.}$

It can be seen in Figure 2.8 that the diameter of the graph is 3.
Applying the values $(n = 5, \eta = 3)$ into the inequality we get

$$\begin{aligned} 3 = diam(G) &\leq n - \eta + 1 \\ &\leq 5 - 3 + 1 \\ &\leq 3. \end{aligned}$$

□

In 1989 Erdős et al [10] introduced the following bounds in terms of order and minimum degree.

Theorem 2.17. *(Erdős et al [10]) Let G be a connected graph of order n and minimum degree $\delta \geq 2$. Then*

$$(i) \qquad diam\,(G) \leq \left\lfloor \frac{3n}{\delta + 1} \right\rfloor - 1. \tag{2.20}$$

$$(ii) \qquad rad\,(G) \leq \frac{3(n-3)}{2(\delta + 1)} + 5.$$

Furthermore, (i) and (ii) are tight apart from the exact value of the additive constant, and for every $\delta > 5$ equality can hold in (i) for infinitely many values of n.

We have already proven in Section 2.1 the bound on radius. Similarly, the same approach used to prove the bound on radius can be used to prove the bound on diameter.
The next theorem by Mukwembi [15] has a stronger bound on diameter compared to bound (2.20) on diameter given by Erdős et al. This is true if the irregularity index is prescribed on the Mukwembi bound. Observe that when the irregularity index,

$\eta = 1$, the right hand side of the Mukwembi bound $\left(d \leq \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1}\right)$ is equal to the right hand side of (2.20). Furthermore, for $\eta > 1$ the value of the right hand side of $\left(d \leq \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1}\right)$ is smaller compared to that of (2.20). Therefore, the bound given by Mukwembi is stronger than the bound given by Erdős et al except for $\eta = 1$.

Theorem 2.18. *(Mukwembi [15]) Let G be a connected graph of order n, minimum degree δ and diameter $d, d \neq 3, 4$. Then the inequality*

$$d \leq \frac{3(n-\eta)}{\delta+1} - 1 + \frac{3}{\delta+1}, \tag{2.21}$$

where η is the irregularity index of G, holds. Moreover, this inequality is essentially tight.

Proof. Let $\beta = \{d(v_1), d(v_2), \ldots, d(v_\eta)\}$ be a set of the distinct degrees of vertices from the degree sequence of G, such that $d(v_1) < d(v_2) < d(v_3) < \ldots < d(v_\eta)$. Observe that $|\beta| = \eta$. We note that $d(v_1) = \delta$ and $d(v_\eta) = \Delta$. Since $d(v_\eta) = \Delta$ we have

$$|N[v_\eta]| = \Delta + 1. \tag{2.22}$$

If v_η is adjacent to all vertices in G then $d(v_\eta)$ plus one is equal to n, thus

$$n = \Delta + 1.$$

But if v_η is adjacent to some but not all vertices in G then $n > \Delta + 1$. Therefore $n \geq \Delta + 1$ and combining this with inequality (2.22) we get

$$n \geq |N[v_\eta]| = \Delta + 1. \tag{2.23}$$

We now show that $|N[v_\eta]| \geq \delta + \eta$. We are going to find an upper bound on the number of terms in β. We note that the biggest value η can take is if β contains all the numbers between δ and Δ. But this may not always be the case: depending on the degree sequence of the graph we may get a smaller value for η. So, to get our bound on η we sum over δ to Δ (inclusive on both ends) and we get the following:

$$\begin{aligned} \eta \leq{} & \Delta - \delta + 1 \\ \eta \leq{} & \Delta + 1 - \delta \\ ={} & |N[v_\eta]| - \delta \qquad \text{by (2.22)} \\ \eta + \delta \leq{} & |N[v_\eta]|. \end{aligned} \tag{2.24}$$

Combining inequalities (2.23) and (2.24) we get

$$n \geq |N[v_\eta]| \geq \delta + \eta, \tag{2.25}$$

as required.
We are going to show that the inequality (2.21) is satisfied for $d \leq 2$ and $d \geq 5$.
Now for $d \leq 2$, we are going to use inequality (2.25) to show that the inequality (2.21) is satisfied.
RHS of (2.21) $= \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1}$.

From (2.25), $n - \eta \geq \delta$. So

$$\frac{3(n-\eta)}{\delta+1} \geq \frac{3\delta}{\delta+1}.$$

Hence,

$$\begin{aligned}\frac{3(n-\eta)}{\delta+1} - 1 + \frac{3}{\delta+1} &\geq \frac{3\delta}{\delta+1} - 1 + \frac{3}{\delta+1}\\ &= \frac{3(\delta+1)}{\delta+1} - 1\\ &= 3 - 1 = 2.\end{aligned}$$

i.e., RHS of (2.21) $= \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1} \geq 2 \geq d =$ LHS of (2.21), and the inequality (2.21) is satisfied for $d \leq 2$.
Now we look at the case $d \geq 5$. Let u and v be vertices in G such that $d_G(u,v) = d \geq 5$. Since $d \geq 5$, of the two vertices u and v label u to be the vertex which is further away from v_η. Therefore $d_G(v_\eta, u) \geq 3$. Because $d_G(v_\eta, u) \geq 3$, u and v_η can be put into a 2 packing. We now construct a maximal 2 packing set A of G using the Dankelmann and Entringer technique starting with $A = \{v_\eta, u\}$.

Claim 4. $|A| \leq \dfrac{n-\eta+1}{\delta+1}$.

We are going to use the closed neighbourhood of A, $N_G[A]$, and inequality (2.25) to prove our Claim 4. Because A is a maximal 2 packing, we observe that for any two vertices x and y in A, $N[x] \cap N[y] = \emptyset$.

$$\begin{aligned} n &\geq |\cup_{x \in A} N[x]| \\ &= |N[v_\eta]| + \sum_{x \in A - \{v_\eta\}} |N[x]| \\ &\geq \delta + \eta + \sum_{x \in A - \{v_\eta\}} (\delta + 1) \\ &= \delta + \eta + (|A| - 1)(\delta + 1) \\ &= \eta - 1 + |A|(\delta + 1). \end{aligned}$$

Rearranging, we have $\quad |A| \leq \dfrac{n - \eta + 1}{\delta + 1}$.

Our claim is shown.

Continuing with the Dankelmann and Entringer technique we construct T_1, T_2 and a spanning tree T. Then we note that $T(A)$, the subgraph induced by A is not connected; it is a collection of vertices of degree zero. But by definition of the kth power, for $k = 3$, $T^3(A)$ is connected and $diam(T^3(A)) \leq |A| - 1$. Observe that the distance between any two vertices of A in T is exactly $3l$ ($l \geq 1$), and that has been reduced to l in $T^3(A)$. Hence

$$\begin{aligned} diam(T(A)) &\leq 3\, diam(T^3(A)) \\ &\leq 3\,(|A| - 1). \end{aligned}$$

This implies that $diam(T(A)) \leq 3(|A| - 1)$.

Recall that when we formed T from T_2 we joined v to some vertex w in T_2 by the edge vw such that the distance from v to some element in A is at most 2. Therefore the distance from v to some vertex $\acute{v}$ in A is at most 2, $d_T(v, \acute{v}) \leq 2$. Now note that u and $\acute{v}$ are in A, thus $d_T(u, \acute{v}) \leq diam(T(A))$, since $diam(T(A))$ is the greatest distance in A. Hence, by the triangle inequality

$$\begin{aligned} d_T(u, v) &\leq d_T(u, \acute{v}) + d_T(\acute{v}, v) \\ &\leq diam(T(A)) + 2 \\ &\leq 3(|A| - 1) + 2 \\ &= 3|A| - 1. \\ d_T(u, v) &\leq 3|A| - 1. \end{aligned} \tag{2.26}$$

Combining Claim 4 and inequality (2.26) we get

$$d_T(u, v) \leq 3\left(\frac{n - \eta + 1}{\delta + 1}\right) - 1 = \frac{3(n - \eta)}{\delta + 1} - 1 + \frac{3}{\delta + 1}.$$

Because T is a spanning tree of G, we have

$$d = d_G(u,v) \leq d_T(u,v) \leq \frac{3(n-\eta)}{\delta+1} - 1 + \frac{3}{\delta+1}.$$

Inequality (2.21) is satisfied for $d \geq 5$.

The bound (2.21) does not hold for graphs of diameter 3 or 4. Two counterexamples of the bound (2.21) are given in Figure 2.9.

We show that apart from the value of the additive constant, this bound is sharp. The Erdős et al paper gives the following example graph (see Figure 2.10) which illustrates that the bound (2.20) is sharp.

Let $k > 1, \delta > 5$, and $V(G) = V_0 \cup V_1 \ldots \cup V_{3k-1}$, where

$$|V_i| = \begin{cases} 1 & \text{if } i \equiv 0 \text{ or } 2 \pmod 3, \\ \delta & \text{if } i = 1 \text{ or } 3k-2, \\ \delta - 1 & \text{otherwise.} \end{cases}$$

Let $v \in V_i, \acute{v} \in V_j$ be joined by an edge of G if and only if $|j - i| \leq 1$.

Observe that Figure 2.10 has diameter $= 8$. Mukwembi states that a modification of the extremal graph given in the Erdős et al paper (for example Figure 2.10) shows that the bound is sharp, apart from the value of an additive constant. To modify Figure 2.10 we removed the edge fg which joins V_1 and V_2 and the edge hi which joins V_{k-1} and V_{3k-2} respectively. Also we removed pairs of edges in V_i (for $i = 1$ and $3k-2$) such that no vertex loses more than one edge and the minimum degree δ remains the same. Doing this reduces the irregularity index η from 2 in Figure 2.10 to 1. See Figure 2.11, a modification of Figure 2.10. Note that Figure 2.11 has diameter $= 8$. The edges removed from Figure 2.10 to modify it to Figure 2.11 are, ab, cd, eo, fg, hi, jp, kn and lm.

$\square$

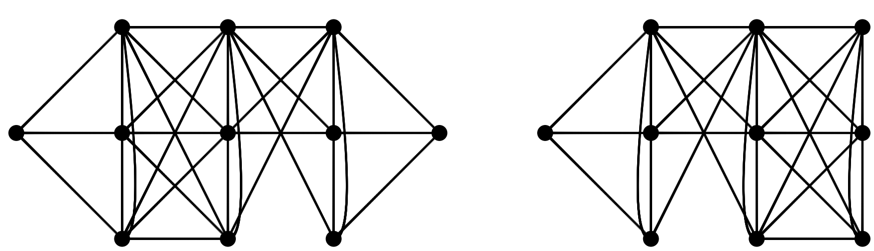


Figure 2.9: Graphs with diameter 4 and 3 illustrating two counterexamples of the bound (2.21) from Theorem 2.18.

The following theorem (stated without proof) gives a bound that hold for graphs with diameter 3 or 4.

Theorem 2.19. *(Mukwembi [15]) Let G be a connected graph of order n, minimum degree δ and diameter $d, d = 3, 4$. Then the inequality*

$$d \leq \frac{3(n-\eta)}{\delta+1} + 1 + \frac{3}{\delta+1},$$

where η is the irregularity index of G, holds. Moreover, this inequality is essentially tight.

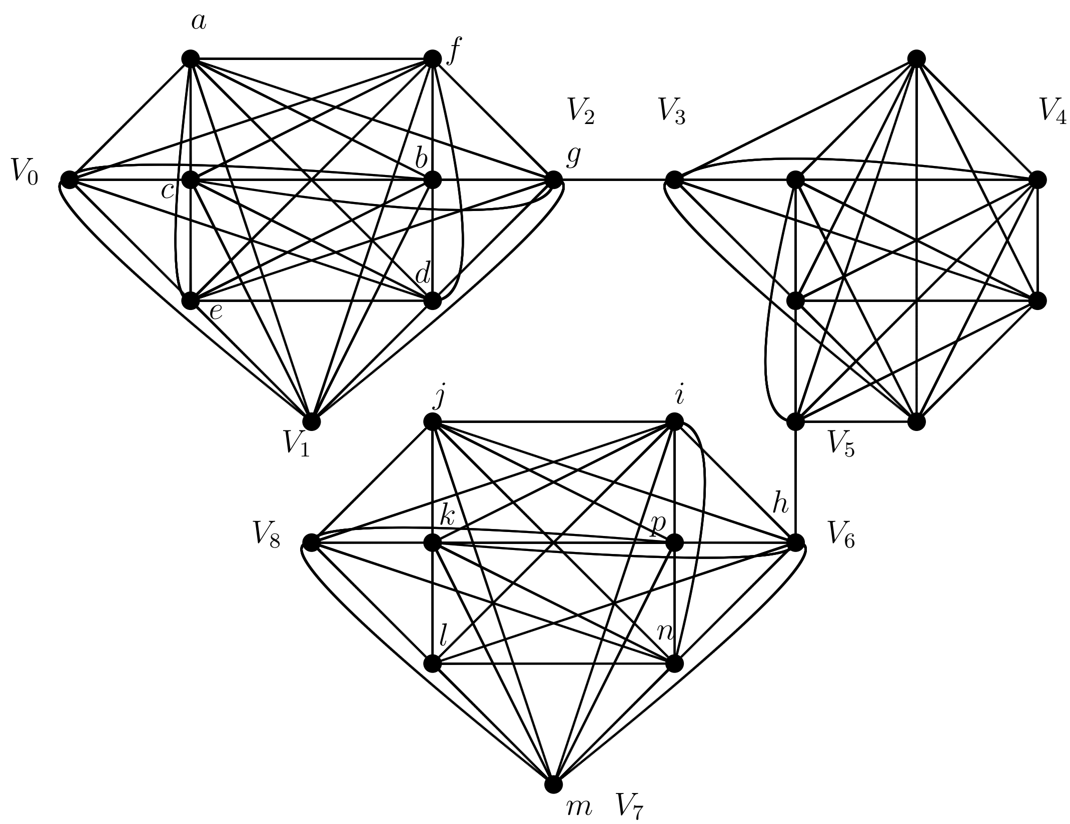


Figure 2.10: Extremal graph G with $n = 26, \delta = 7, \eta = 2, k = 3$ and diameter $= 8$.

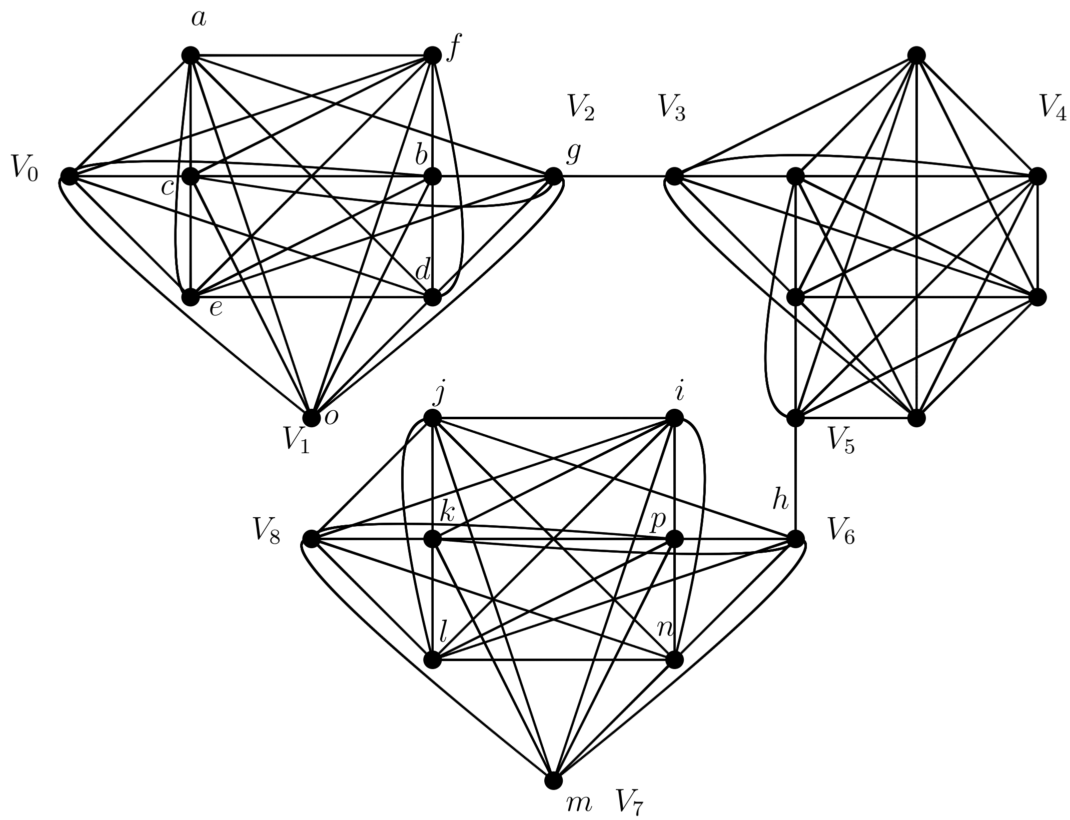


Figure 2.11: Extremal graph G with $n = 26, \delta = 7, \eta = 1, k = 3$ and diameter $= 8$.

Erdős et al in their paper proved two more theorems which give upper bounds on diameter and radius for a triangle-free graph and a C_4-free graph. The theorems are as follows.

Theorem 2.20. *Let G be a connected triangle-free graph with n vertices, and with minimum degree $\delta \geq 2$. Then*

(i)
$$diam\, G \leq \left\lceil \frac{n - \delta - 1}{2\delta} \right\rceil.$$

(ii)
$$rad\, G \leq \frac{n-2}{\delta} + 12.$$

Furthermore, (i) and (ii) are tight apart from the exact value of the additive constant, and for every $\delta \geq 2$ equality can hold in (i) for infinitely many values of n.

Theorem 2.21. *Let $\delta \geq 2$ be a fixed integer, and let G be a connected, C_4-free graph with n vertices and with minimum degree δ. Then* *(i)* $$diam\, G \leq \frac{5n}{\delta^2 - 2[\delta/2] + 1}.$$

(ii) $$rad\, G \leq \frac{5n}{2(\delta^2 - 2[\delta/2] + 1)}.$$

Furthermore, if δ is large, then these bounds are almost tight. More precisely, if $\delta + 1$ is a prime power, then there exists a graph G with the above properties and

(iii) $$diam\, G \geq \frac{5n}{\delta^2 + 3\delta + 2} - 1.$$

In summary, in this chapter we have used several methods to generate spanning tress. We used the BFS algorithm to generate a distance-preserving spanning tree from a root vertex v (see Theorem 2.2). Using the BFS algorithm we proved Corollary 2.3, which stated that every connected graph G has a spanning tree T with $rad(G) = rad(T)$. We used the Erdős et al method to construct a radius-preserving spanning tree with the bound, $rad(G) = rad(T) \leq \dfrac{3(n-3)}{2(\delta+1)} + 5$. In section 2.2 a result by Ali et al (Theorem 2.9), gives an upper bound on diameter on the constructed spanning tree $\left(d \leq n + \dfrac{1}{2} - \sqrt{2m - 2n + \dfrac{1}{4}}\right)$ in terms of order and size. The Dankelmann and Entringer (Section 2.3, Theorem 2.12) approach for constructing a spanning tree gave an upper bound on average distance $\left(\mu(T) \leq \dfrac{n}{\delta+1} + 5\right)$ in terms of order and minimum degree. In 1989 Erdős et al introduced a bound on diameter ($diam(G) \leq \left\lfloor \dfrac{3n}{\delta+1} \right\rfloor - 1$) in terms of order and minimum degree. Then Mukwembi (Theorem 2.18) in 2012 gave a stronger bound $\left(d_G \leq d_T \leq \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1}\right)$ compared to that of Erdős et al in the case when the irregularity index η is prescribed.

In the next chapter, we will consider a method which was introduced by Kleitman and West. They called it the dead leaves method. We will use this method to construct various spanning trees and use them to bound diameter.

Chapter 3

Bounds on leaf number in terms of Order and Minimum degree

In this chapter we are going to study the dead leaves method by Kleitman and West [13] for constructing a spanning tree. This method starts by finding a small tree T in G with many leaves and then growing the tree by adding several vertices to T using an iterative algorithm in such a way that the number of leaves always grows until we get the tree T to be spanning. Finding a spanning tree with the maximum number of leaves in a graph is an **NP**-complete problem [9].
If G is a cycle, we are only guaranteed to get a spanning tree with a maximum of two leaves. But we want a spanning tree with many leaves. To do this, we consider $\boldsymbol{G_{n,k}}$, the collection of connected n-vertex graphs with minimum degree at least k. Griggs and Wu [12] were one of many researchers investigating the question of finding a spanning tree with the maximum number of leaves. They also used the dead leaves method by Kleitman and West to construct a spanning tree T of G. See Section 3.3 for more details.

We will investigate bounds when $k = 3, 4$ and 5. We start with $k = 3$.

3.1 A lower bound for minimum degree $\delta \geq k = 3$

In the next theorem we look at the case $\delta \geq k = 3$. In 1981, Storer [20] conjectured that for $\delta = 3$, any 3-regular graph G with n vertices has $L(T) \geq n/4 + 2$. Kleitman and West [13] proved using the dead leaves approach that this bound is also true for any $G \in \boldsymbol{G_{n,3}}$. We will present this proof below.

We are going to use the following iterative algorithm to grow tree T to a desired spanning tree. For each iteration, we let T_i (for $i = 0, 1, 2 \ldots$) be the current tree with n' vertices and t leaves.
If x is a leaf of the current tree T_i, then the *out-degree* of x, $d'(x)$, is the number of neighbours it has in $G - T_i$. Note that if this vertex x has no neighbour(s) in the $G - T_i$ graph, then that vertex is a leaf of T_i. This particular leaf is called a dead leaf, because we cannot expand beyond it. Hence it must be a leaf in the final tree.
Initial procedure: Choose a vertex x, then we add all vertices in G adjacent to vertex x. Thus, we have our small tree T_1.
Expansion procedure: We are going to expand our tree T_i by vertex expansion sequences. We arbitrarily choose a leaf x of the current tree which is not dead i.e., $d'(x) > 0$. From x we add to T_i all $d'(x)$ vertices not in T_i and form the next tree T_{i+1}.
We say an operation is *admissible* if an expansion on a tree satisfies the augmentation inequality. And we define the *augmentation inequality* to be an inequality which determines the change in the number of leaves Δt, the number of dead leaves Δm and the number of vertices $\Delta n'$ from the current tree. Note that every time we choose a vertex and add $d'(x)$ vertices, we get a new and bigger tree T_{i+1}. We repeat this expansion using admissible operations until we get a desired spanning tree T.

Theorem 3.1. *(Kleitman and West [13]). Every $G \in \boldsymbol{G_{n,3}}$ has a spanning tree with at least $n/4 + 2$ leaves.*

Proof. We are going to construct a spanning tree T of G using the dead leaves method.
We initialize with a small tree and expand it to a spanning tree of G using the iterative algorithm, where for each step we add some number of vertices $\Delta n'$, such that the augmentation inequality $3\Delta t + \Delta m \geq \Delta n'$ is satisfied.

We consider two cases.
Case 1: G is 3-regular.
There are two sub cases to be considered.
$1a$: If every edge of G belongs to a triangle, then $G = K_4$. By applying the initial procedure of the iterative algorithm on G we get a spanning tree T with three leaves. So, $L(K_4) = 3$, and case $1a$ is proved.
$1b$: G is 3-regular but has an edge or edges which do not belong to a triangle. We are going to choose one of these edges and initialize tree T_0 from that edge. Since

G is 3-regular, the edge is incident to two vertices, each vertex having two other edges incident to it. We have now formed tree T_0. Then we expand tree T_i using the expansion procedure until we get T_i to be a spanning tree.

Case 2: G is not 3-regular.

We initialize tree T_0 from a vertex of maximum degree $\Delta \geq 4$. The tree T_0 will be inclusive of all edges incident to the vertex of maximum degree $\Delta \geq 4$. Then we expand T_i until we get a desired spanning tree.

At the end of the algorithm all the leaves of T are dead, therefore our desired spanning tree has $t = m = L$ and $n' = n$, the total number of vertices in the spanning tree T.

Now we look at the two cases and see how a tree is grown from a small tree T_0 into a spanning tree T using admissible operations. We do this until we get T_i to be a spanning tree T. We note all the leaves of the spanning tree are now dead leaves and that summing the augmentation inequalities from $T_0, T_1, \ldots, T$ we get $3(L-4) + L \geq n - 6$ for case $1b$ and $3(L - \Delta) + L \geq n - (\Delta + 1)$ for case 2. Justification of this statement is given below.

For case $1b$, we started with 4 leaves and 6 vertices and ended up with L leaves together with n vertices. So the summation of the augmentation inequality becomes $3(L-4)+L \geq n-6$. Simplifying this we get $4L \geq n+6$. For case 2, we started with maximum degree Δ leaves and $\Delta+1$ vertices and the final number of leaves and vertices is L and n respectively. Hence the summation of the augmentation inequality for case 2 is $3(L-\Delta)+L \geq n-(\Delta+1)$, which simplifies to $4L \geq n+2\Delta-1 \geq n+7$.

The following is a collection of admissible operations.

See Figure 3.1 for an illustration of the three operations.

01: If $d'(x) \geq 2$ for some current leaf x, then expanding at x yields $\Delta\, t = \Delta\, n' - 1$ and $\Delta\, m \geq 0$.

Since leaf x has at least two neighbours in $G - T_i$, the increase in number of vertices in T_{i+1} must be at least two. We show that the augmentation inequality is satisfied.

For $\Delta t = \Delta\, n' - 1$ we have

$$3\,(\Delta n' - 1) + \Delta m \geq \Delta\, n'$$

$$3\,\Delta n' + m - 3 \geq \Delta\, n'$$

$$2\,\Delta n' + m - 3 \geq 0. \tag{3.1}$$

We show that given $m \geq 0$ and $\Delta n' \geq 2$ the inequality (3.1) is satisfied.
From $m \geq 0$ we get the following

$$\begin{aligned} m &\geq 0 \\ m+1 &\geq 1. \end{aligned}$$

And from $\Delta n' \geq 2$ we get the following

$$\begin{aligned} \Delta n' &\geq 2 \\ 2\,\Delta n' &\geq 4 \\ 2\,\Delta n' - 3 &\geq 4 - 3 \\ 2\,\Delta n' - 3 &\geq 1 \\ 2\,\Delta n' - 3 + m &\geq m + 1 \\ &\geq 0 + 1 \\ &= 1 \\ &\geq 0. \end{aligned}$$

Therefore,

$$\begin{aligned} 2\,\Delta n' - 3 + m &\geq m + 1 \geq 1 \\ 2\,\Delta n' - 3 + m &\geq 1 \geq 0, \end{aligned}$$

as required.
Hence the augmentation inequality, $3\,\Delta t + \Delta m \geq \Delta\, n'$, is satisfied.

02: If $d'(x) \leq 1$ for every current leaf x and some vertex outside T_i has at least two neighbours in T, then expanding at one of them yields $\Delta\, t = 0$, $\Delta\, m \geq 1 = \Delta n'$.
We show that the augmentation inequality is satisfied..

$$\begin{aligned} \text{LHS} &= 3\,\Delta t + \Delta m \\ &= 3(0) + \Delta m \\ &= \Delta m \geq 1 = \Delta n' \\ \therefore 3\,\Delta t + \Delta m &\geq \Delta\, n' = \text{RHS}. \end{aligned}$$

Hence the augmentation inequality is satisfied.
03: If y is the only neighbour of x outside T_i and y has at least two neighbours

not in T_i, then expanding at x and then y yields $\Delta\, t = \Delta\, n' - 2 \geq 3 - 2 \geq\; 1$ and $\Delta\, m \geq 0$. When expanding at x and then y we add at least three new vertices to T_{i+1}. So $\Delta n' \geq 3$. We show that the augmentation inequality is satisfied.

$$\begin{aligned} \Delta\, n' &\geq 3 \\ 2\,\Delta\, n' &\geq 6 \\ 3\,\Delta\, n' &\geq 6 + \Delta\, n' \\ 3\,(\Delta\, n' - 2) &\geq \Delta\, n' \\ 3\,\Delta\, t &\geq \Delta\, n'. \end{aligned}$$

From $\Delta m \geq 0$ we get the following

$$\begin{aligned} \Delta\, m &\geq 0 \\ 3\,\Delta\, t + \Delta m &\geq 3\,\Delta\, t \geq \Delta\, n'. \end{aligned}$$

Therefore,

$$3\,\Delta t + \Delta m \geq \Delta\, n'.$$

Hence the augmentation inequality is satisfied.
Because $k = 3$, this implies that any vertex in $G - T_i$ has at least two neighbours in $G - T_i$ or at least two neighbours in T_i. Since operations $01 - 03$ require at least one neighbour in $G - T_i$, we are guaranteed that one of the operations will always be available until we get our desired spanning tree.

Recall that the lower bounds on the number of leaves for cases $1b$ and 2 are $4L \geq n{+}6$ and $4L \geq n + 7$ respectively. We are going to improve these bounds to $L \geq \dfrac{n}{4} + 2$ using the final iteration to get an excess of at least two dead leaves. We are going to focus on the final iteration, using the expansion procedure and one of the three operations.
Let vertex v be some vertex in tree T_{i-1}, the tree before the final iteration. The vertex v must have one neighbour (vertex w) in $G - T_{i-1}$. Because any of the three operations $01 - 03$ adds at least one vertex to T_{i-1}, we add vertex w to v and form $T_i = T$. Since this is the last iteration, vertex w must have at least two neighbours in T_{i-1}. Let u and z be the neighbours of w. Note that u and z are leaves in T_{i-1} which are not dead. Upon adding w to T_{i-1} these two vertices die. Because this is the final iteration leaf w is also dead. The final iteration has killed two leaves u and z to give us two excess dead leaves. Hence to $L \geq \dfrac{n}{4} + 2$, and our proof is complete. $\square$

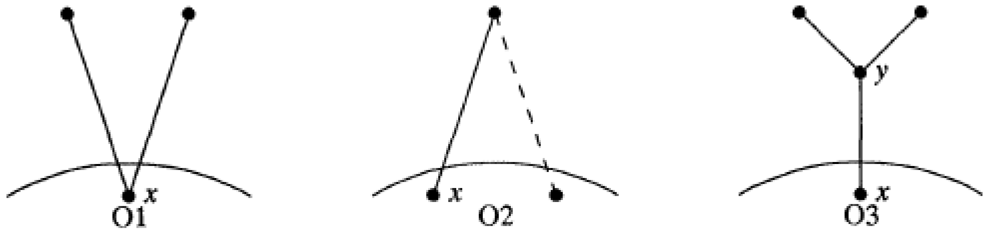
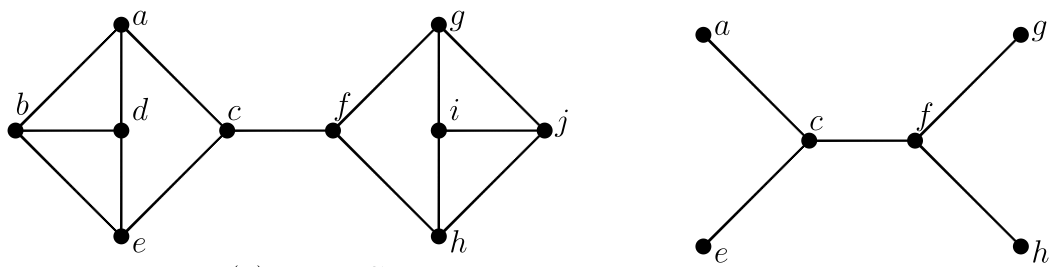


Figure 3.1: Operations used when $k = 3$.

We will illustrate the iteration algorithm and three operations of the proof with an example.

Example 2. *Consider graph G_1 for Case 1b and G_2 for Case 2.*

Case 1b:
We begin with a 3-regular graph G_1 which falls under case $1b$. We construct our initial tree T_0 for G_1 by choosing an edge, cf, which does not lie in any triangle. See Figure 3.2. Our initial tree T_0 has $n' = 6$, $t = 4$, $m = 0$.



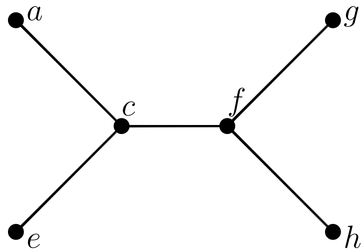
(a) Tree G_1 (b) Tree T_0

Figure 3.2: Graph G_1 and Tree T_0.

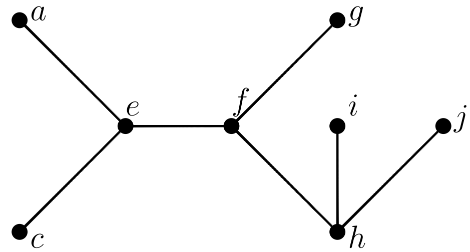
We expand T_0 to T_1 by choosing vertex h for which $d'(h) = 2$. Since $d'(h) = 2$ we expand using operation 01. See Figure 3.3. We now check if the augmentation inequality is satisfied for T_1 with $n' = 8$, $t = 5$ and $m = 3$.

$$\text{LHS} = 3\Delta t + \Delta m = 3\,(5-4) + (3-0) = 6 \geq 2 = 8 - 6 = \Delta n' = \text{ RHS}$$
$$\therefore 3\Delta t + \Delta m \geq \Delta n'.$$

The augmentation inequality is satisfied.



(a) Tree T_0

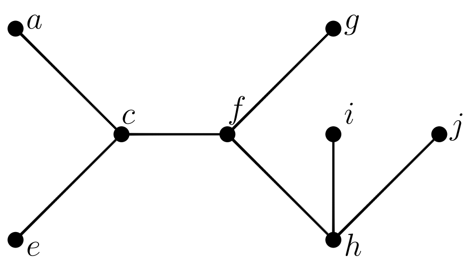


(b) Tree T_1

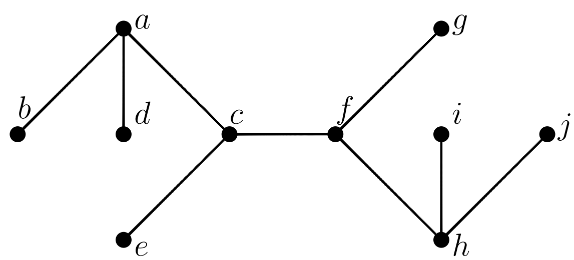
Figure 3.3: Tree T_0 and Tree T_1.

We expand tree T_1 by vertex expansion to get T_2 which is our desired spanning tree T. We choose vertex a from T_1 with 2 neighbours in $G_1 - T_1$ and expand using operation 01 to get our spanning tree T. See Figure 3.4. Since this was the final iteration, vertex e which was not dead in T_1 is now dead. So, vertices b, d, e are now dead leaves. Our spanning tree T has $n' = n = 10$ and $L = t = m = 6$. So, the augmentation inequality given below is satisfied.

$$\begin{aligned}\text{LHS} &= 3\Delta t + \Delta m = 3(6-5) + (6-3) = 6 \geq 2 = 10 - 8 = \Delta n' = \text{ RHS} \\ &\therefore 3\Delta t + \Delta m \geq \Delta n'.\end{aligned}$$



(a) Tree T_1



(b) Spanning tree T

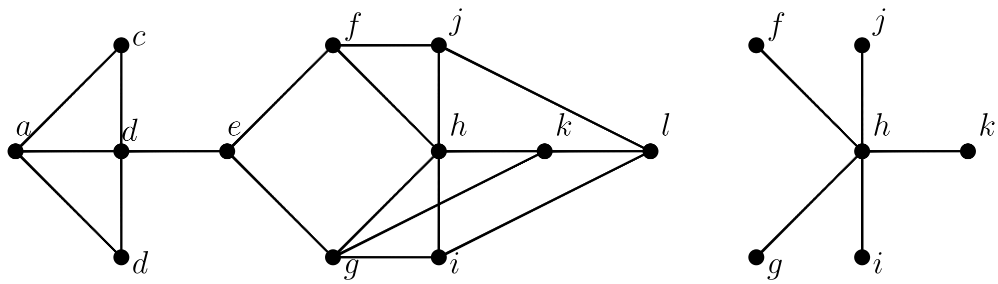
Figure 3.4: Tree T_1 and Spanning tree T.

For this example, operation 01 has been applied twice. In the last iteration 2 dead leaves are added which gives us $L(T) = 6$. And so

$$\begin{aligned}\text{LHS} = 6 = L(T) &\geq \frac{n}{4} + 2 \\ &= \frac{10}{4} + 2 \\ &= 4.5 \\ 6 &\geq 4.5 = \text{RHS}.\end{aligned}$$

Since LHS $= L(T) = 6$ and the RHS $= n/4 + 2 = 4.5$.
We started with a graph of minimum degree $k = 3$ and generated a spanning tree T with $L(T) = 6 \geq n/4 + 2 = 4.5$. This illustrates how we generate a spanning tree T with at least $n/4 + 2$ leaves.
Next, we use graph G_2 to illustrate case 2. G_2 is not 3-regular.

Case 2:
We initialize our tree T_0 from G_2 with a vertex of maximum degree $\Delta = 5$. See Figure 3.5. Our initial tree T_0 has six vertices, five leaves and zero dead leaves.
We use operation 02 to expand tree T_0 to T_1 from vertex k which has one neighbour in the graph $G_2 - T_0$. See Figure 3.6 (a). The tree T_1 has $t = 5$, $m = 3$ and $n' = 7$.
We now check if the augmentation inequality is satisfied:

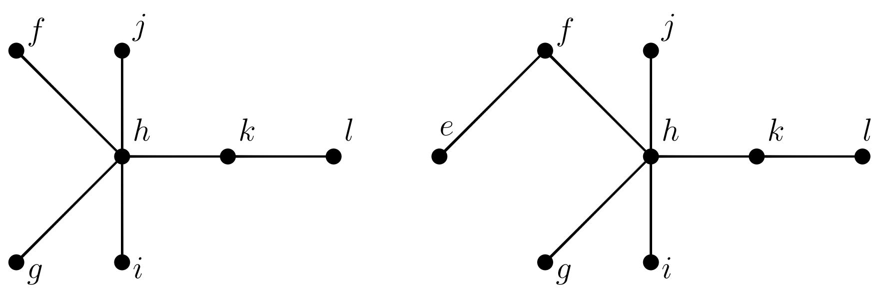


(a) Graph G_2 (b) Tree T_0

Figure 3.5: Graph G_2 and Tree T_0.

$$\text{LHS} = 3\Delta t + \Delta m = 3\,(5-5) + (3-0) = 3 \geq 1 = 7 - 6 = \Delta n' = \text{RHS}$$
$$\therefore 3\Delta t + \Delta m \geq \Delta n'.$$

The augmentation inequality is satisfied.



(a) Graph G_2 (b) Tree T_0

Figure 3.6: Tree T_1 and Tree T_2.

Using vertex expansion on T_1 we form tree T_2 via operation 02 on vertex f. See Figure 3.6 (b). The expansion of the tree is as follows: $\Delta t = 0$, $\Delta m = \Delta n' = 1$. As done for T_1 we show that the augmentation inequality is satisfied.

$$\text{LHS} = 3\Delta t + \Delta m = \ 3\,(0) + (1) = 1 \geq 1 = \Delta n' = \ \text{RHS}$$
$$\therefore 3\Delta t + \Delta m \geq \Delta n',$$

as desired.

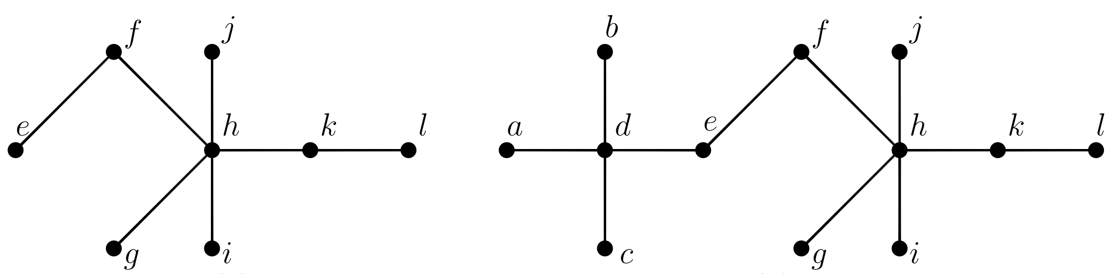
We now expand T_2 to T_3. Since vertex d is the only neighbour of e in T_2 we expand using operation 03. Furthermore, d has 3 neighbours. Since this is the final iteration leaves a, b, c die. By adding all the neighbours of vertex d we get a spanning tree T. See Figure 3.7.

$$\text{LHS} = 3\Delta t + \Delta m = \ 3\,(7-5) + (7-4) = 9 \geq 4 = 12 - 8 = \Delta n' = \ \text{RHS}$$
$$\therefore 3\Delta t + \Delta m \geq \Delta n'.$$

Hence the augmentation inequality is satisfied.

$$\text{LHS} = 3\Delta t + \Delta m = \ 3\,(5-5) + (3-0) = 3 \geq 1 = 7 - 6 = \Delta n' = \ \text{RHS}$$
$$\therefore 3\Delta t + \Delta m \geq \Delta n'.$$

The augmentation inequality is satisfied.



(a) Graph T_2 (b) Spanning tree T

Figure 3.7: Tree T_2 and Tree T_3 = the Spanning tree T.

In the last iteration 3 dead leaves are added which gives us $L(T) = 7$. And so

$$\begin{aligned}\text{LHS} = 7 = L(T) &\geq \frac{n}{4} + 2 \\ &= \frac{12}{4} + 2 \\ &= 5 \\ 7 &\geq 5 = \text{RHS}.\end{aligned}$$

We started with a graph of minimum degree $k = 3$ and generated a spanning tree T with $L(T) = 7 \geq n/4 + 2 = 5$. This illustrates how we generate a spanning tree T with at least $n/4 + 2$ leaves.
The example has illustrated how a spanning tree can be constructed using the dead leaves approach.

Before we look at the case $k = 4$ we make note of a proof by Griggs et al [11]. Using the dead leaves method, they were able to prove the following theorem.

Theorem 3.2. *(Griggs et al [11]) If G is a connected cubic graph with n vertices and contains no subgraph isomorphic to $K_4 - e$, then $L \geq \frac{1}{3}(n + 4)$.*

This bound is sharper compared to the general case $k = 3$.

Building on our case $k = 3$, we provide an overview of how the case $k = 4$ is proved.

3.2 A lower bound for minimum degree $\delta \geq k = 4$

We will use the same technique, definitions, notation and iterative algorithm as in the proof of Theorem 3.1, except now the augmentation inequality will be
$4\,\Delta t + \Delta m \geq 2\,\Delta\, n'$ and our initial tree T_0 will start with a star at a vertex with degree $k = 4$ including all its neighbours or a double star where both vertices are joined by an edge.
In this case we will require extra definitions:

Definition 5. *A principle expansion sequence is an expansion of a single leaf $x = y_0$ of T_i and then other leaves that did not belong to T_i before the initial expansion.*

Definition 6. *A principle expansion sequence is live if each expansion after y_0 introduces two new vertices to the tree.*

Definition 7. *A linear expansion sequence is a live sequence $Y = (y_0, \ldots, y_r)$ such that, for each $i \geq 1, y_{i+1}$ is one of the two leaves introduced by expanding y_i.*

The following proof is for $k = 4$.

Theorem 3.3. *(Kleitman and West [13]). Every $G \in \boldsymbol{G_{n,4}}$ has a spanning tree with at least (2n+8)/5 leaves.*

We will now give an overview of the proof.

As for the case $k = 3$, the proof starts with a small tree T_0 and expands using admissible operations until a desired spanning tree T is constructed. We now introduce new parameters. Let c_1 be the number of leaves in the initial tree, c_2 the number of leaves not counted as dead in the initial tree T_0 and c_3 the number of vertices in the initial tree. Summing the augmentation inequalities we get $4\,(L - c_1) + (L - c_2) \geq 2\,(n - c_3)$ which simplifies to $L \geq 2\,n/5 + (c_2 + 4\,c_1 - 2\,c_3)/5 = 2\,n/5 + c$, where $c = (c_2 + 4\,c_1 - 2\,c_3)/5$. The value of the additive constant will be shown later to be $c = 8/5$.

This theorem has seven admissible operations of which operations $01 - 03$ are the same as those for case $k = 3$. If operations $01 - 03$ are not available, we then proceed to use one of the operations $04 - 07$. To expand our tree using any of operations $04 - 07$, we are going to use the principle expansion sequence. When expanding we let Y be the set of vertices expanded in T_{i+1}. Operations $04 - 07$ all have a vertex y_1 which is adjacent to y_0 and involve expanding from y_1 in various different ways. See Figure 3.8 for an illustration of operations $04 - 07$.

Using the linear expansion sequence Kleitman and West were able to prove that operations $01 - 07$ will always be available until we get our desired spanning tree T. Recall:

c_1 is the number of leaves in the initial tree. Since $\delta \geq k = 4$ then $c_1 \geq 4$.

c_2 is the number of leaves not counted as dead in the initial tree. So $c_2 \geq 2$.

c_3 is the number of vertices in the initial tree. Since $k = 4$ then $c_3 = 5$.

Therefore $c = (c_2 + 4\,c_1 - 2\,c_3)/5 \geq (2 + 16 - 10)/5 = 8/5$. Hence $L \geq 2\,n/5 + 8/5$. Our overview is complete.

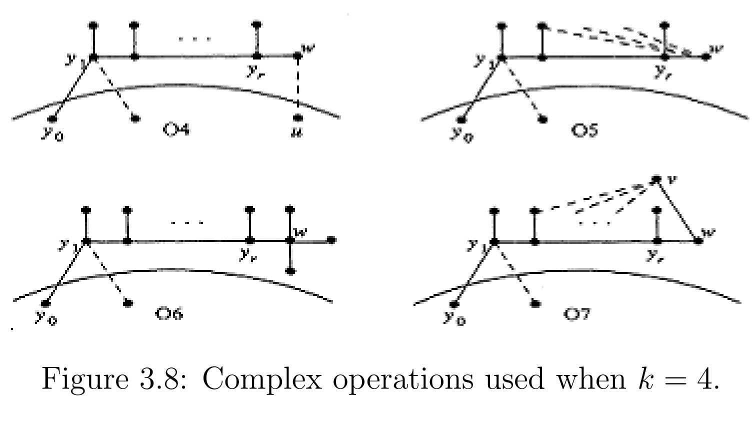


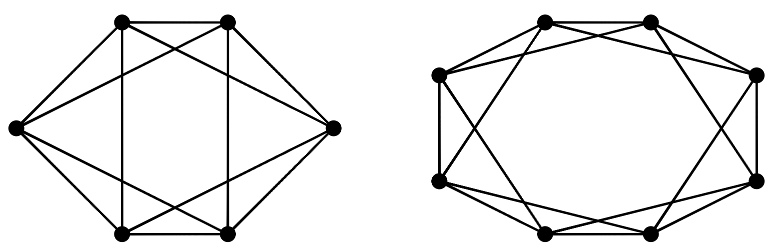
Figure 3.8: Complex operations used when $k = 4$.

Kleitman and West conjectured an improved lower bound on Theorem 3.3. The conjecture is as follows.

Conjecture 3.4. *Every $G \in \boldsymbol{G_{n,4}}$ has a spanning tree with at least $2n/5+2$ leaves, except for a 4-regular graph with every edge in a triangle.*

There are only two known examples of graphs which do not have a spanning tree T of a graph $G \in \boldsymbol{G_{n,4}}$ with at least $2n/5+2$ leaves. Both graphs have every edge belonging to a triangle. The example graphs are a 4-regular graph with 6 vertices and a 4-regular graph with eight vertices. See Figure 3.9 for the two example graphs F_1 and F_2. The conjectured bound requires F_1 and F_2 to have 5 and 6 leaves respectively. For this to be possible the spanning tree of F_1 must have a vertex of degree 5, which is impossible.

We are going to show that F_2 cannot contain a spanning tree with 6 leaves. Observe that the spanning tree must have two vertices whose neighbours cover all vertices of F_2 in one of the following ways. Either the two vertices are connected, or they are not. If they are connected, we either get four or five leaves and we note that not all vertices are covered. Therefore, the tree is not spanning. If the vertices are not connected, then this will not be a tree since a tree must be connected. Hence by the manner in which F_1 and F_2 are constructed it makes it impossible for either graph to have spanning trees attaining the bound.



(a) F_1 (b) F_2

Figure 3.9: Example graph F_1 and F_2.

Next, we illustrate that the conjectured improved bound can be attained. For our illustration we are going to use two graphs $F_3 \in \boldsymbol{G_{n,4}}$ and $F_4 \in \boldsymbol{G_{n,4}}$.

We first consider graph $F_3 \in \boldsymbol{G_{n,4}}$ to be a graph with a vertex of degree 5. We start our tree T_0 (in a similar vein to case 2 of Theorem 3.1) from that vertex of degree 5 inclusive of all the edges incident to it. We then expand our tree until we get a spanning tree T. And hence we get the following, $c_1 = 5, c_2 \geq 2$ and $c_3 = 6$. Substituting the values of c_1, c_2 and c_3 into c we get $c = (c_2 + 4\,c_1 - 2\,c_3)/5 \geq (2+20-12)/5 = 2$.

Secondly, we consider $F_2 \in \boldsymbol{G_{n,4}}$ to be 4-regular but having an edge or edges which do not belong to a triangle (this is similar to case $1b$ of Theorem 3.1). We are going to choose one of these edges and start tree T_0 from that edge and construct our spanning tree T. Since $F_2 \in \boldsymbol{G_{n,4}}$ is 4-regular, our edge is incident to two vertices, each vertex having three other edges incident to it. This yields $c_1 = 6, c_2 \geq 2$ and $c_3 = 8$. Substituting the values of c_1, c_2 and c_3 into c we get $c = (c_2 + 4\,c_1 - 2\,c_3)/5 \geq (2 + 24 - 16)/5 = 2$. In both cases we have illustrated that we can attain the improved bound $2n/5 + 2$.

We have looked at the cases for $k = 3$, $k = 4$ and we were able to construct a spanning tree with at least $n/4 + 2$ and $(2n + 8)/5$ leaves respectively. Now we look at larger values of k.

3.3 A lower bound for minimum degree $\delta \geq k > 4$

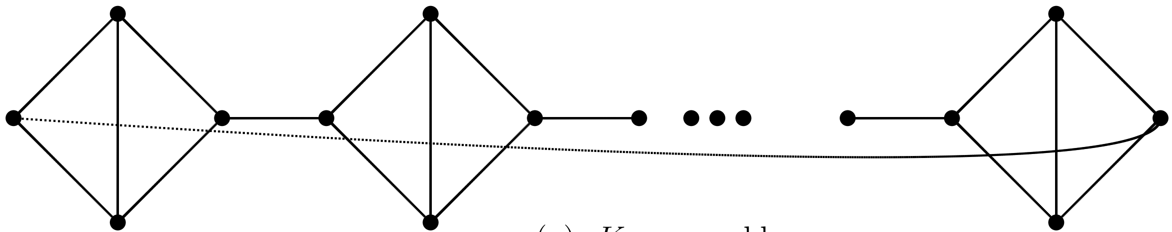
We will start with the case $k = 5$. Kleitman and West conjectured that for $k = 5$ we can construct a spanning tree with at least $n/2 + 2$ leaves. In 1992 Griggs and Wu [12] proved the Kleitman and West conjecture. That is, they proved the following theorem.

Theorem 3.5. *(Griggs and Wu [12]) If G is a connected graph with n vertices and minimum degree at least 5, then $L \geq n/2 + 2$.*

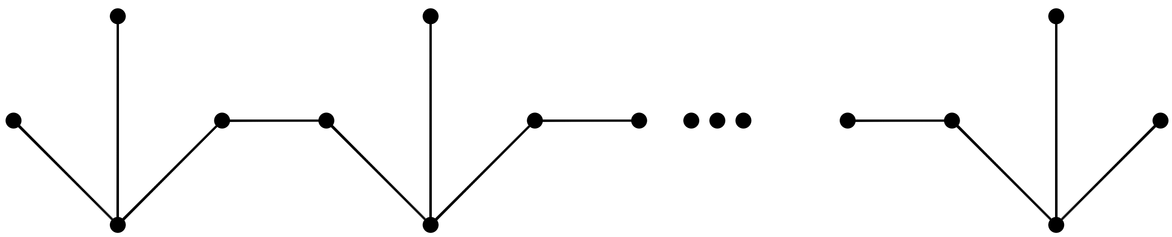
To prove this theorem Griggs and Wu used the dead leaves approach for constructing a spanning tree and used a different augmentation inequality for admissibility.
Alon et al [3] in their paper made note of a very interesting conjecture by Linial in the area of finding a spanning tree with maximum leaves. This conjecture was also mentioned earlier (1992) by Griggs and Wu in their paper. The conjecture is as follows.

Conjecture 3.6. *(Alon et al [3]) If G has n vertices and minimum degree k, then $L \geq \dfrac{(k-2)}{(k+1)} n + c_k$, where c_k depends on the value of k.*

The conjecture has been shown to hold in the the cases $k = 3, 4$ and 5 (see Theorems 3.1, 3.3 and 3.5 respectively) with $c_3 = 2, c_4 = 8/5$ and $c_5 = 2$. Griggs and Wu observed that the proofs of these theorems provide a polynomial-time algorithm to find a spanning tree which attains the lower bound on L. We can also attain this bound for a family of k-regular graphs. Figure 3.10 (a) illustrates an example of a $k = 3$-regular necklace made up of a number of beads where each bead is $K_{k+1} - e$. A corresponding spanning T with $L = n/4 + 2$ is also illustrated in Figure 3.10 (b).



(a) $K_4 - e$ necklace



(b) Spanning tree T

Figure 3.10: Illustration of a $K_4 - e$ necklace and spanning tree T with $L = n/4+2$.

The Alon et al paper states that in 1990 Alon [2] observed that Linial's conjecture was false for all large values of k.

Kleitman and West proved the following theorem for all graphs with sufficiently large minimum degree.

Theorem 3.7. *If k is sufficiently large, then there is an algorithm that constructs a spanning tree with at least [1- $b\ln k/k$]n leaves in any graph with minimum degree k, where b is any constant exceeding 2.5.*

Note that the term $b \ln k/k$ tends to zero as k approaches infinity and the coefficient $[1 - b \ln k/k]$ approaches 1. The results in Alon's paper [2] imply an upper bound on the number of leaves of a spanning tree. Indeed, there are graphs with n vertices and minimum degree k that have a spanning tree T with at most $(1 - (1 + o(1))\frac{\ln(k+1)}{k+1})n$ leaves. The term $o(1)$ tends to zero as k tends to infinity.

We present an upper bound on the leaf number of a constructed spanning tree in terms of order and minimum degree.

3.4 Upper bound on leaf number

In this section, we are going to show using two cases that the number of leaves on the spanning tree will always be at most $n - 3\lfloor n/(k+1)\rfloor + 2$.

Theorem 3.8. *(Kleitman and West [13])* $L \leq n - 3\lfloor n/(k+1)\rfloor + 2$.

Proof. We construct a graph $G_{n,k} \in \boldsymbol{G_{n,k}}$ such that no matter how hard we try, every spanning tree T that we can possibly generate from $G_{n,k} \in \boldsymbol{G_{n,k}}$ cannot have more than $n - 3\lfloor n/(k+1)\rfloor + 2$ leaves. Let $m = \lfloor n/(k+1)\rfloor$ and $r = n - m(k+1)$. Partition the vertex set $V(G)$ into sets $R_0, \ldots, R_{m-1}$, where $|R_i| = k+1$ for $i \neq 0$ and $|R_0| = k+1+r$. We choose an arbitrary pair of vertices $x_i, y_i \in R_i$. In each R_i we join every pair of distinct vertices by an edge except x_iy_i. This becomes B_i $(0 \leq i \leq m)$ blocks or components of almost complete graphs with each block having one edge x_iy_i missing. To restore the minimum degree k we add the following edges $Z = \{x_iy_{(i+1)mod\ m} : 0 \leq i < m\}$. These edges join the blocks of almost complete graphs to form one graph $G_{n,k} \in \boldsymbol{G_{n,k}}$.

We now construct a spanning tree T and show that any spanning tree T of $G_{n,k}$ has at most $n - 3m + 2$ leaves.
For each i $(0 \leq i \leq m-1)$ we arbitrarily choose a vertex $u \in R_i$ which has maximum degree. Then join every vertex which is adjacent to u and delete every other edge except for those edges joining vertices to u. If the vertex u is x_i or y_i $(\forall i)$, then the graph is a forest. However, if u is some other vertex, then the graph has all edges of Z missing and is not connected nor spanning. For the graph to be connected and spanning we must add all the edges of Z. We have now formed a graph which is not a tree. To generate a tree, we use the edges of Z.
Every pair of edges of Z forms an edge cut. So, the spanning tree must have at most one edge of Z missing.
Let $W = \{x_i\} \cup \{y_i\}$.
We prove that T has at most $n - 3m + 2$ leaves by using two cases.
Case 1: Suppose one edge $x_jy_{(j+1)}$ of Z is missing, for an arbitrary fixed j for $0 \leq j \leq m-1$.
Removing edge x_jy_{j+1} allows us to generate a spanning tree containing a path x_iy_i

in B_i for each i. The removal of the edge x_jy_{j+1} forces each B_i to have three vertices which are not leaves (nonleaves) except for the two blocks B_j and B_{j+1} which were joined by the edge x_jy_{j+1}. Each of these two blocks have two nonleaf vertices. Each vertex of W is not a leaf except x_j, y_{j+1}. There are $m-2$ blocks with each having 3 nonleaf vertices. Also, there are 4 nonleaves which come from the two blocks which were joined by the the edge x_jy_{j+1}. Therefore, the spanning tree T has $3(m-2)+4 = 3m-2$ nonleaf vertices.
Let H be the number of nonleaf vertices in T. Then $H \geq 3m-2$. Since T is a spanning tree of $G_{n,k}$, then $n = n_T$. Note that $n = H + L$, so $L = n - 3m - 2$. Therefore, T has

$$\begin{aligned} -H &\leq -(3m-2) \\ n - H &\leq n - (3m-2) \\ L = n - H &\leq n - (3m-2) \\ L &\leq n - 3m + 2. \end{aligned}$$

We have proven for case 1 that T has at most $n - 3m + 2$ leaves.
Case 2: No edge of Z is missing.
Since we cannot remove any of the edges of Z, to generate T we remove an edge x_ky_k in B_k for one fixed arbitrary value of k. Now all B_i blocks have a path x_iy_i in T except for the one block which had an edge x_ky_k removed. With all edges of Z in T, the vertices x_k and y_{k+1} are now nonleaf vertices. Hence the blocks which contain the vertices x_j and y_{j+1} have each gained a nonleaf vertex. The block B_i which had the edge x_ky_k removed has lost one nonleaf vertex while block B_{i-1} has gained one extra leaf due to the removal of edge x_ky_k. There are at least $(3(m-1)+2)-1 = 3m-2$ nonleaf vertices in T.
Let H be the number of nonleaf vertices. Then $H \geq 3m-2$. Using the same argument as in case 1, it follows that the number of leaves in T is $L \leq n - 3m + 2$. We have proven for case 2 that T has at most $n - 3m + 2$ leaves.
We have shown using cases 1 and 2 that we can get a spanning tree with at most $n - 3\lfloor n/(k+1) \rfloor + 2$ leaves. Hence the proof is complete. □

We will illustrate the process of the above proof with an example.

Example 3. *We consider graph $Q \in \boldsymbol{G_{27,4}}$ for cases 1 and 2. See Figure* 3.11.

Case 1: Suppose one edge x_2y_3 of Z is missing ($j = 2$).

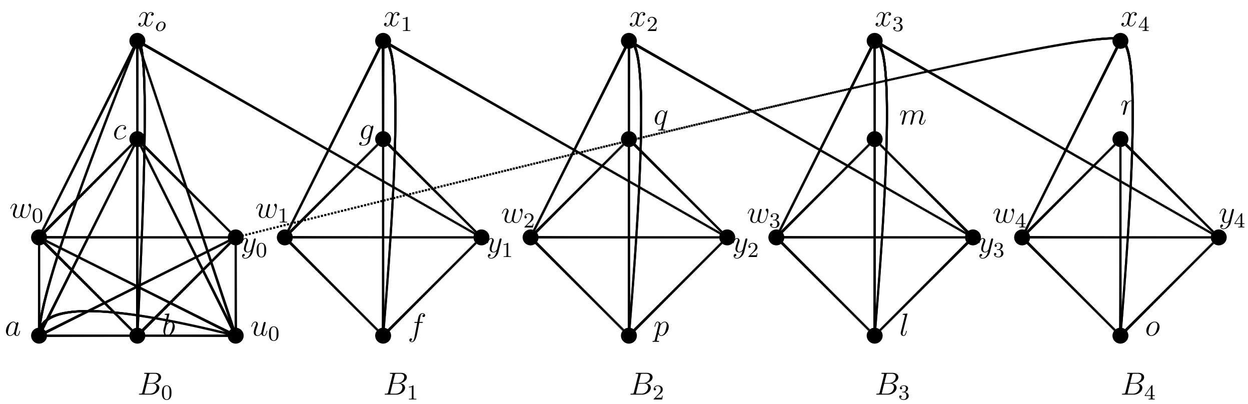


Figure 3.11: Graph $Q \in \boldsymbol{G_{27,4}}$.

We construct our spanning tree T by choosing vertex w_i for $i \neq 0$ and u_0 in B_0. Join w_i and u_0 to all neighbour vertices and add all edges of Z except for x_2y_3. See Figure 3.12. Observe that B_0, B_1 and B_4 have 3 nonleaf vertices and $4, 2, 2$ leaves respectively. Also, B_2 and B_3 have 2 nonleaves and 3 leaves each. By observation of Figure 3.12 we have generated a spanning tree T with exactly 14 leaves. We are going to show the theoretical value of L. We know that $m = \lfloor n/(k+1) \rfloor = \lfloor 27/(5) \rfloor = 5$ and $r = n - m(k+1) = 27 - 5(5) = 2$. The theoretical value of L is as follows:

$$
\begin{aligned}
L &\leq n - 3m + 2 \\
&= 27 - 3(5) + 2 \\
&\leq 14.
\end{aligned}
$$

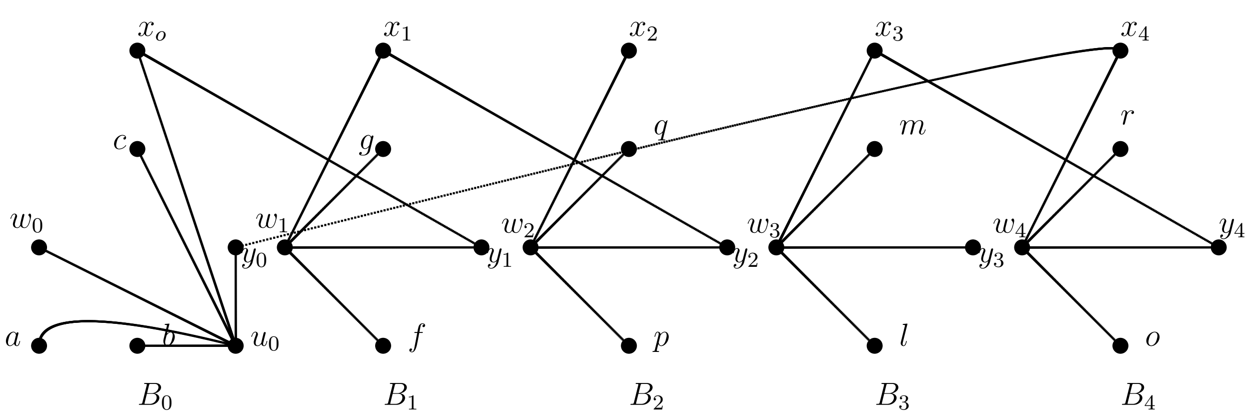


Figure 3.12: Spanning tree T

.

We have shown that we can generate a spanning tree with at most $n-3m+2$ leaves. Actually, we have shown that we can attain the exact value of $L = n - 3m + 2$.

Case 2: No edge of Z is missing.
We construct our spanning tree in a similar manner to case 1. However now instead of removing one edge of Z we remove the edge w_1y_1 $(k = 1)$ of component B_1. See Figure 3.13. Observe that B_0, B_2, B_3 and B_4 have 3 nonleaf vertices and $5, 2, 2, 2$ leaves respectively. Also, B_1 has 2 nonleaves and 2 leaves. By observation of Figure 3.13 we have generated a spanning tree T with exactly 13 leaves. The theoretical value of L is as follows:

$$\begin{aligned} L &\leq n - 3m + 2 \\ &= 27 - 3(5) + 2 \\ &\leq 14. \end{aligned}$$

We have shown theoretically that the number of leaves in Figure 3.13 must be at most 14, and observed that Figure 3.13 can have a maximum of 13 leaves which is less than 14 leaves. Hence, considering both cases 1 and 2 we have shown that the number of leaves on the spanning tree T of $G_{n,k} \in \boldsymbol{G_{n,k}}$ will always be at most $n - 3m + 2$ leaves.

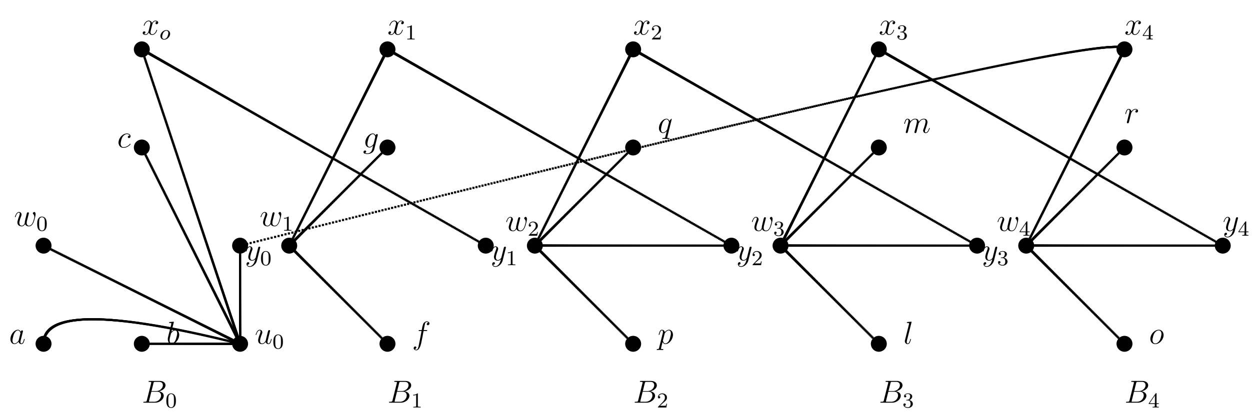


Figure 3.13: Spanning tree T.

This example shows how we can generate a spanning T of $G_{27,4}$ with at most $n - 3m + 2$ leaves.

We conclude this chapter with a link between the leaf number and the diameter of a graph. In the proof of case 1 in Theorem 2.9 we noted an upper bound on the leaf number in terms of diameter. The bound is $L \leq n - (d-1)$. Hence, we get an upper bound on the diameter for a spanning tree T as $d \leq n - L + 1$. To get our diameter in terms of order and minimum degree we use the bound of Linial's Conjecture 3.6, $L \geq \frac{(k-2)}{(k+1)}n + 2$.

$$\begin{aligned}
L &\geq \frac{(k-2)}{(k+1)}n + 2 \\
-L &\leq -\frac{(k-2)}{(k+1)}n - 2 \\
d \leq n - L + 1 &\leq n - \frac{(k-2)}{(k+1)}n - 2 + 1 \\
&= \frac{(k+1-k+2)}{(k+1)}n - 1 \\
&= \frac{3}{(k+1)}n - 1 \\
d &\leq \frac{3}{(k+1)}n - 1.
\end{aligned}$$

Recall that we earlier stated that Erdős et al [10] introduced an upper bound on diameter of a graph in terms of order and minimum degree (2.20),

$$diam(G) \leq \left\lfloor \frac{3n}{\delta + 1} \right\rfloor - 1.$$

We have now derived the related bound on diameter using leaf number.

Chapter 4

Conclusion

In this dissertation, we have investigated in detail four approaches for constructing a spanning tree, namely the radius preserving-spanning tree by Erdős et al, Ding et al's method, the Dankelmann and Entringer's method, and the dead leaves spanning tree construction method by Kleitman and West.

We used the BFS algorithm to construct a distance-preserving spanning tree from root vertex v such that $d_G(v,u) = d_T(v,u)$ for all $u \in V(G)$. From this result we were able to prove Corollary 1, which states that every connected graph G has a spanning tree T with $rad(G) = rad(T)$. We then used the results of Theorem 2.1 (($d_G(v,u) = d_T(v,u)$) and Corollary 1 to generate a radius-preserving spanning tree. Using the constructed spanning tree, we then bounded radius, $rad(G) = rad(T) \leq \dfrac{3(n-3)}{2(\delta+1)} + 5$ (Erdős et al, Theorem 2.4).

In section 2.2, Theorem 2.6 we studied the spanning tree construction approach by Ding et al. We then used the spanning tree constructed by Ding et al in Theorem 2.7 which is similar to Theorem 2.8 by Ali et al, to derive an upper bound on diameter $\left(\text{Theorem 2.9, } d \leq n + \dfrac{1}{2} - \sqrt{2m - 2n + \dfrac{1}{4}}\right)$ in terms of order and size. We concluded the chapter by studying the method for constructing spanning tree by Dankelmann and Entringer. After studying the construction method, Theorem 2.12 then bounded average distance in terms of order and minimum degree $\left(\mu(T) \leq \dfrac{n}{\delta+1} + 5\right)$. Using the same method, Mukwembi (Theorem 2.18) gave an upper bound on diameter of the constructed spanning tree in terms of order, minimum degree and irregularity index η. This bound on diameter was $d \leq \dfrac{3(n-\eta)}{\delta+1} - 1 + \dfrac{3}{\delta+1}$.

In Chapter 3, we considered $\boldsymbol{G_{n,k}}$, the collection of connected n-vertex graphs with minimum degree at least k. These graphs were looked at with the aim of finding lower and upper bounds on the number of leaves of a constructed spanning tree T of $G_{n,k} \in \boldsymbol{G_{n,k}}$. The Kleitman and West approach (dead leaves approach) was used to construct spanning trees in terms of order and minimum degree $\delta \geq k$. Kleitman and West proved that every $G_{n,k} \in \boldsymbol{G_{n,k}}$ has a spanning tree with at least $n/4+2$, $(2n+8)/5$ leaves for $k = 3, 4$ respectively. Griggs and Wu used the same method to construct a spanning tree with at least $n/2+2$ leaves for $k = 5$. In Section 3.4, we presented an upper bound on the leaf number ($L \leq n - 3\lfloor n/(k+1)\rfloor + 2$) of a constructed spanning tree in terms of order and minimum degree.
We used Conjecture 3.6 and a bound in Theorem 2.9 to derive an upper bound on the diameter of a constructed spanning tree for $k = 3, 4$ and 5. The bound is given as $d \leq \dfrac{3}{(k+1)}n - 1$.
Throughout this dissertation, we have used examples to illustrate how a spanning tree can be constructed using a particular approach.

Bibliography

[1] P. Ali, P. Dankelmann, M. J. Morgan, S. Mukwembi, H. Swart, T. Vetrík, The average eccentricity, spanning trees of plane graphs, size and order. *Util. Math.* 107 (June 2018).

[2] N. Alon, Transversal numbers of uniform hypergraphs, *Graphs Combin.* 6 (1990), 1-4.

[3] N. Alon, F. V. Fomin, G. Cutin, M. Krivelevich, S. Saurabh, Spanning directed trees with many leaves, *lecture at SIAM Discrete Math.* 23 (2009), 466-476.

[4] R.A. Beezer, J.E. Riegsecker and B.A. Smith, Using minimum degree to bound average distance. *Discrete Math.* 226 (2000), 365-371.

[5] F. R. K. Chung, The average distance and independence number, *J. Graph Theory 12* (1988), 229-235.

[6] P. Dankelmann, Average distance and domination number. *Discrete Appl. Math.* 80 (1997), 21-35.

[7] P. Dankelmann, R. Entringer, Average distance, minimum degree, and spanning trees. *J. Graph Theory 33* (2001), 1-13.

[8] D. Ding, T. Johnson, P. Seymour, Spanning trees with many leaves. *Graph Theory 37* (2001), 189-197.

[9] R. J. Douglas, NP-completeness and degree restricted spanning trees, *Discrete Math.* 105 (1992), 41-47.

[10] P. Erdős, J. Pach, R. Pollack, Z. Tuza, Radius, diameter, and minimum degree, J. *Combin. Theory Ser.* B 47 (1989), 73-79.

[11] J. R. Griggs, D. J. Kleitman, A. Shastri, Spanning trees with many leaves in cubic graphs, *J. Graph Theory 13* (1989), 669-695.

[12] J. R. Griggs, M. Wu, Spanning trees in graphs of minimum degree 4 or 5, *Discrete Math*. 104 (1992), 167-183.

[13] D. J. Kleitman, D. B. West, Spanning trees with many leaves, *lecture at SIAM Discrete Math*. Conf. San Francisco, (1988).

[14] M. Kouider, P. Winkler, Mean distance and minimum degree, *J. Graph Theory 25* (1997), 95-99.

[15] S. Mukwembi, A note on diameter and the degree sequence of a graph, *Appl. Math. Lett*. 25 (2012), 175-178.

[16] S. Mukwembi, Bounds on Distances in Graphs. Ph.D. Thesis, University of KwaZulu-Natal, Durban, 2006.

[17] S. Mukwembi, On spanning cycles, paths and trees, *Discrete Math*. 161 (2013), 2217-2222.

[18] G. Salamon, G. Wiener, On finding spanning trees with few leaves, *Inform. Process. Lett. 105* (2008) 164-169.

[19] P. J. Slater, Centrality of paths and vertices in a graph: cores and pits, theory and applications of graphs. *Wiley*, *New York* (1981), 529-542.

[20] J. A. Storer, Constructing full spanning trees for cubic graphs. *Inform. Process. Lett. 13* (1981), 8-11.

[21] S. Wuchty and P. F. Stadler, Centers of complex networks. *J. Theoret. Biol.* 223 (2003), 45-53.