WITS UNIVERSITY

School Of Mathematics

Master Of Science

Dissertation

Distance measures, independence number and chromatic number

Supervisor:
Prof. Simon Mukwembi
Co-Supervisor:
Prof. Betsie Jonck

Student:
Letlhogonolo Moholane

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Declaration

I, Letlhogonolo Moholane, declare that this dissertation titled "Distance measures, independence number and chromatic number" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a masters degree at this University.
- Where any part of this dissertation has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- I have acknowledged all main sources of help.
- Where the dissertation is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signature of candidate: L.D Moholane. Date: 02 April 2023

Abstract

There are numerous parameters in graph theory. In this dissertation, we pay a special attention to average distance, independence number, average eccentricity, order and the chromatic number of a graph.

In 1975, Doyle and Graver proved an upper bound on the average distance with respect to the order of the graph. This gave rise to studies that focus on upper and lower bounds on average distance in terms of other graph parameters. Approximately, three decades after Doyle and Graver proved their result, Dankelmann, Goddard, and Swart in 2004 produced a study that gave an upper bound on average eccentricity in terms of minimum degree and order of the graph, initiating studies that focus on giving bounds on average eccentricity with respect to other known graph parameters.

In this dissertation, we investigate bounds on average eccentricity and on average distance. We give upper bounds on average eccentricity in terms of independence number of the graph and order of the graph. Then, we present bounds on average eccentricity when order and chromatic number of the graph are prescribed. The second part of the dissertation is dedicated to presenting upper bounds on average distance with respect to independence number and order of the graph, and again, in terms of chromatic number and order of the graph.

Acknowledgement

My deepest gratitudes to my supervisors, Prof. Simon Mukwembi and Prof. Betsie Jonck. Their support and guidance throughout writing and compiling this dissertation is much appreciated. For without their guidance and support, this dissertation would not have been completed in such a timely manner. Above all, I appreciate the patience they had with me thoughout this process.

Contents

List of Symbols and Abbreviations

Symbol	**Description**
$d_G(u, v)$	The distance between vertices u and v in graph G.
$e_G(v)$	The eccentricity of a vertex v in graph G.
$\gamma(G)$	The domination number of graph G.
$\sigma_G(v)$	The total distance of a vertex v in the graph G.
$\sigma(G)$	The transmission of graph G.
$\mu(G)$	Average distance of graph G.
$\gamma_c(G)$	The connected domination number of graph G.
$\alpha(G)$	The independence number of graph G.
$\chi(G)$	The chromatic number of graph G.
$\zeta(G)$	Total eccentricity of graph G.
$W(G)$	Total distance of graph G (also referred to as the Weiner index of graph G).

Abbreviations

$rad(G)$	Radius of graph G.
$diam(G)$	Diameter of graph G.
$avec(G)$	Average eccentricity of graph G.
$deg_G(x)$	The degree of vertex x in graph G.

Special Graphs

P_q	A path of order q.
K_q	A complete graph of order q.
$B_{n-b,b}$	A tree consisting only of a path $P = v_1, v_2, \ldots, v_{n-b}$ with $\lfloor \frac{b}{2} \rfloor$ end vertices attached to v_1 and $\lceil \frac{b}{2} \rceil$ end vertices attached to v_{n-b}.
$H_{s,c}$	A c-chromatic graph of order s consisting of a path P of order $s - c$ with end vertices x and y, and a complete graph K_c, with x joined to a vertex v of the c-clique, K_c.

$T(s,c)$	A family of complete c-partite graphs with order s and maximum number of edges, without $(c+1)$-cliques for $1 \leq c \leq s-1$.
$\mathcal{T}_{s,c}$	A family of trees T: T is a combination of some tree T' of order s with c end vertices attached to some vertices of T'.

Chapter 1

Introduction and Preliminaries

1.1 Introduction

The concept of distance in graph theory is one of the building blocks for most graph invariants, such as average eccentricity and average distance of a graph, which will be our key focus areas in this study. The graph invariants, average eccentricity and average distance, will be studied in relation to the independence number and the chromatic number. The independence number deals with subsets of vertices of a graph such that no two vertices are adjacent to one another in that subset, and the chromatic number deals with the coloring of vertices of a graph such that no two adjacent vertices are colored with the same color. In particular, in this dissertation, we will give bounds on average eccentricity and average distance when the independence number and the chromatic number are known.

In this chapter, we give basic terminology that we will use throughout this dissertation. In Section 1.2, we give a definition of a graph, and give some of the elementary definitions that relate to graphs. In the subsequent section, Section 1.3, we give definitions that are related to distances in a graph. In the last two sections, we define independence number, and chromatic number of a graph. Terminology that is given in this introduction can be found in [2], [8] and [21].

1.2 Graph theory terminology

A *graph* $G = (V, E)$ consists of two sets, a non-empty set V of elements called *vertices* and a set E (with a possibility of being empty) of 2-element subsets of V called *edges*. In some cases, the set of vertices of G is denoted by $V(G)$ and the set of edges of G is denoted by $E(G)$. If $E(G) = \emptyset$, then we say that graph G is *empty*. The *order* of graph G is defined as the number of vertices in G while the number of edges in G is the *size* of G. We will denote by n and m the order and size of a graph, respectively.

Example 1.2.1. Consider a graph G of order $n = 5$ and size $m = 6$, with vertex set

$$V(G) = \{v_1, v_2, v_3, v_4, v_5\}$$

and edge set

$$E(G) = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_2, v_5\}, \{v_3, v_4\}, \{v_4, v_5\}\}$$

or simply

$$E(G) = \{v_1v_2, v_1v_3, v_2v_3, v_2v_5, v_3v_4, v_4v_5\}.$$

We use a diagram shown in Figure 1.1, to represent graph G. Vertices of graph G are represented by dots while edges are represented by lines joining the corresponding vertices.

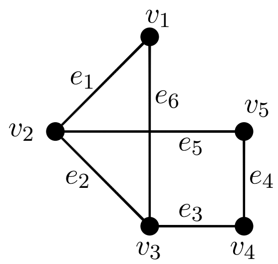


Figure 1.1: Graph G

If two vertices in G are joined by an edge then we say that those vertices are *adjacent* to one another or they are *neighbours*. If u and v are adjacent vertices in G, then we say that u and v are *incident* with the edge uv. A set consisting only of vertices that are adjacent to a vertex v is called an *open neighbourhood* of v, or just a neighbourhood of v. We use $N(v)$ to denote the neighbourhood of v, while the *closed neighbourhood* of v is the set $N[v] = N(v) \cup \{v\}$. A neighbour w of v is called a private neighbour of v with respect to a set $S \subseteq V(G)$ if no other vertex of S is adjacent to w. The *degree* of a vertex v in G is the number of vertices adjacent to v. The smallest degree among vertices of G is called the *minimum degree* of G, while the largest degree among vertices of G is called the *maximum degree* of G, and is denoted by $\Delta(G)$. If every two distinct vertices in G are adjacent, then G is called a *complete* graph. A complete graph of order n is denoted by K_n.

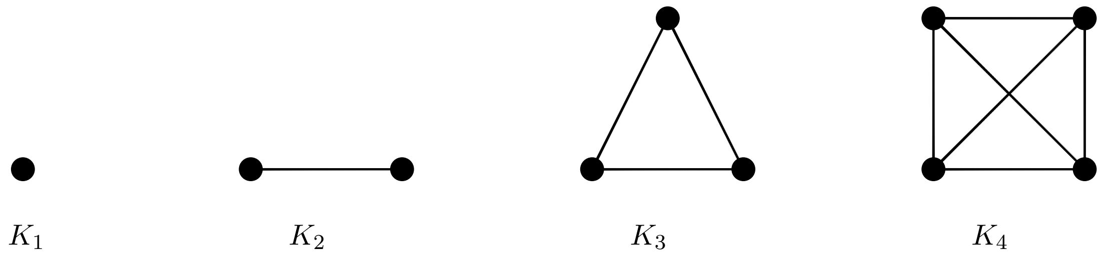


Figure 1.2: K_n graphs for $n \leq 4$.

Let G be a graph. A graph H is a *subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. Suppose H is a subgraph of G, we say that H is a *spanning subgraph* of G if $V(H) = V(G)$. The graph H is said to be a *proper subgraph* of G if either $V(H) \subset V(G)$ or $E(H) \subset E(G)$. A complete subgraph of graph G is called a *clique*. A clique of order k is called a *k-clique*. Let S be a nonempty subset of $V(G)$. The *subgraph* $G[S]$ of G induced by S has vertex set S (i.e., $V(G[S]) = S$) and two vertices u and v in S are adjacent in $G[S]$ if and only if u and v are adjacent in G. A subgraph H of a graph G is called an *induced subgraph* of G if there is a nonempty set S of $V(G)$ such that $H = G[S]$.

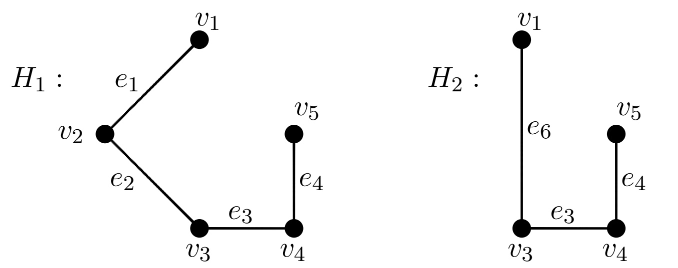
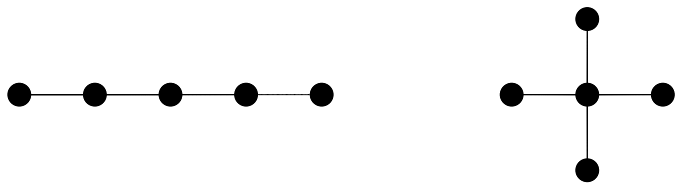


Figure 1.3: Examples of subgraphs of graph G given in Figure 1.1

Note that, in Figure 1.3, H_1 is a spanning subgraph of G since $V(H_1) = V(G)$ and $E(H_1) = \{e_1, e_2, e_3, e_4\} \subset E(G)$. H_2 is a subgraph of G since $V(H_2) = \{v_1, v_3, v_4, v_5\} \subset V(G)$ and $E(H_2) = \{e_6, e_3, e_4\} \subset E(G)$. $H_2 = G[V(H_2)]$ thus H_2 is an induced subgraph of G. $H_1 \neq G[V(H_1)]$ since v_1 and v_3 are not adjacent in H_1 and also, v_2 and v_5 are not adjacent in H_1.

We define a *u-v walk* W within a graph G as an alternating sequence of vertices and edges, with W explicitly given by the sequence $u = v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k = v$ and $e_i = v_{i-1}v_i$ for $i = 1, 2, \ldots, k$. The number of edges we encounter within a u-v walk in a graph is the *length* of that u-v walk. If no vertices are repeated within a walk in G then we call that walk a *path*. A walk in a graph is a *closed walk* if the initial vertex and the last vertex are the same otherwise we say that it is an *open walk*. A closed path in graph G is called a *cycle* in G. If G contains a u-v walk, then we say that vertices u and v are *connected* in G. Also, we say that G is *connected* if for each vertex u and any vertex v in G, there is a u-v walk in G, otherwise, we say that G is *disconnencted*. A connected subgraph H of G is a *component* of G if it is not a proper subgraph of a connected subgraph of G. Let G be a graph with vertex set $V(G)$ and $x \in V(G)$. We denote the graph with vertex set $V(G) - \{x\}$ by $G - x$. If for any vertex $x \in V(G)$, $G - x$ has more components than G, then x is said to be a *cut-vertex* of G. A connected graph, with no cycles, is called a *tree*.



P_4 $K_{1,4}$

Figure 1.4: Examples of trees

Example 1.2.2. The following walks in Figure 1.1, $v_1v_2v_5$, $v_1v_2v_3v_4v_5$, and $v_1v_3v_4v_5$, are v_1-v_5 walks. These walks are also paths, and so, they can also be referred to as v_1-v_5 paths. The following paths are cycles in Figure 1.1, $v_2v_3v_4v_5v_2$, $v_1v_2v_3$, and $v_1v_3v_4v_5v_2v_1$.

Let G and H be graphs. We say that G and H are *isomorphic* if there exists a bijective function $\phi : V(G) \to V(H)$ called an *isomorphism* such that two vertices u and v are adjacent in G if and only if $\phi(u)$ and $\phi(v)$ are adjacent in H. If G and H are isomorphic, then we write $G \cong H$.

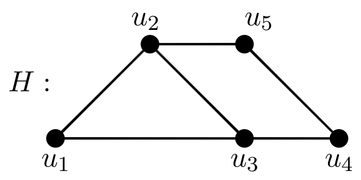


Figure 1.5: Graph H

Graph G and graph H that are in Figure 1.1 and Figure 1.5, respectively, are isomorphic, under the isomorphism $\phi : V(G) \to V(H)$ defined as:

$$\phi : \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ u_1 & u_2 & u_3 & u_4 & u_5 \end{pmatrix}.$$

This is because, for each pair of adjacent vertices u and v in $V(G)$, there is a pair $\phi(u)$ and $\phi(v)$ of adjacent vertices in $V(H)$ and vice versa. Note that, G and H both have order 5, size 6, two vertices with degree 3 while other vertices have only degree 2. Moreover, each of the two graphs has three cycles.

Let G be a graph. A subset S of $V(G)$ is called a *dominating set* of G if every vertex not in S is adjacent to at least one vertex in S. The *domination number* $\gamma(G)$ of G is the minimum cardinality of a dominating set of G. If a dominating set S of G induces a connected subgraph in G, then we say that S is a *connected dominating set* of G, and a connected dominating set of of minimum cardinality is called the *connected domination number* of G and is denoted by $\gamma_c(G)$, or simply γ_c if G is understood. The graph G in Figure 1.1, the following sets are some of the dominating sets of graph G.

$$\{v_1, v_2, v_3, v_4\}, \{v_1, v_3, v_4\}, \{v_3, v_4\}, \{v_3, v_5\}.$$

The domination number $\gamma(G)$ of graph G in Figure 1.1 is

$$|\{v_3, v_4\}| = |\{v_3, v_5\}| = 2.$$

The following sets are some of the connected dominating sets of graph G in Figure 1.1,

$$\{v_1, v_2, v_3, v_4\}, \{v_3, v_4\}.$$

The connected domination number of G in Figure 1.1 is thus,

$$|\{v_3, v_4\}| = 2.$$

1.3 Distance concepts

For a connected graph G, we define *distance* $d_G(x, y)$ between vertex x and vertex y in G as the length of a shortest x-y path in G. A x-y path of length $d_G(x, y)$ is called a *x-y geodesic*. The *total distance of a vertex* x in G is defined as

$$\sigma_G(x) = \sum_{y \in V(G)} d_G(x, y)$$

and the *average distance* of G is defined as

$$\mu(G) = \frac{1}{n(n-1)} \sum_{x \in V(G)} \sigma_G(x) = \binom{n}{2}^{-1} \sum_{\{x,y\} \subseteq V(G)} d_G(x, y) = \frac{2}{n(n-1)} \sum_{\{x,y\} \subseteq V(G)} d_G(x, y),$$

where n is the order of G.

In some instances, we will use $d_G(x)$ to imply $\sigma_G(x)$. Also, it is tedious to work with $\mu(G)$, instead, we will use the Wiener index of G defined as

$$W(G) = \sum_{\{x,y\}\subseteq V(G)} d_G(x,y).$$

The *transmission* of G is defined as

$$\sigma(G) = \sum_{x\in V(G)} \sigma_G(x) = \sum_{x,y\in V(G)} d_G(x,y).$$

Let $S \subset V(G)$. We use $d_G(u,S)$ to denote minimum distance between $u \in V(G)$ and vertices of S. The *eccentricity* $e_G(v)$ of a vertex v in G is the distance between v and a vertex furthest from v in G. The greatest eccentricity among the vertices of G is called the *diameter* of G and it is denoted by $diam(G)$ while the smallest eccentricity among the vertices of G is called the *radius* and it is denoted by $rad(G)$. The *average eccentricity* of a graph G is the mean eccentricity of G and it is denoted by $avec(G)$. Thus, we have that

$$diam(G) = \max_{v\in V(G)} e_G(v), rad(G) = \min_{v\in V(G)} e_G(v) \text{ and } avec(G) = \frac{1}{n}\sum_{v\in V(G)} e_G(v).$$

A vertex v in G with eccentricity $e_G(v) = rad(G)$ is called a *central vertex* of G.

1.4 Independence concepts

A set Q is said to be *maximal* with respect to a given property P, if adding an element to Q violates the property P. Now, a set U of vertices in a graph G is said to be *independent* if no two vertices in it are adjacent in G. A maximal independent set U in G is referred to as a *vertex independence number* or simply, just *independence number* of G and is denoted by $\alpha(G)$.

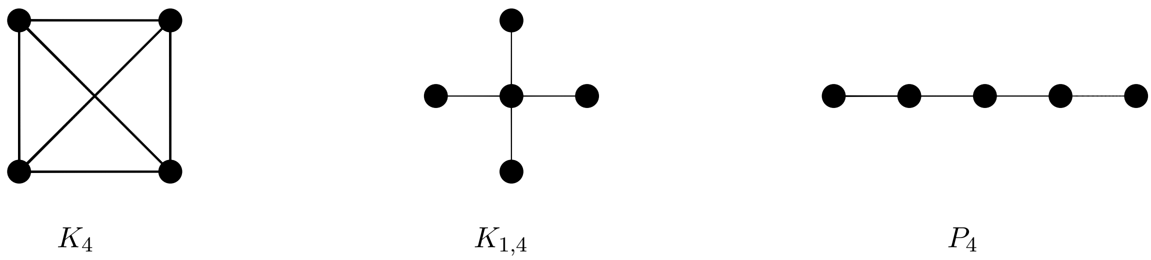


Figure 1.6: $\alpha(K_4) = 1$; $\alpha(K_{1,4}) = 4$ and $\alpha(P_5) = 3$.

1.5 Chromatic concepts

Let G be a graph. A *proper vertex coloring* of G (or just *coloring* of G) is an assignment of colors to the vertices of G, one color to each vertex, so that adjacent vertices are colored differently. Positive integers are used to represent different colors used for coloring of the graph. We can then present the colors used for coloring a graph G by a function $c : V(G) \to \mathbb{N}$ such that $c(u) \neq c(v)$ if u and v are adjacent. ($\mathbb{N}$ is a set of positive integers which represents the colors used for the coloring of graph G). If we use k or less colors for a coloring of G, then we say such coloring is a *k-coloring*. Graph G is *k-colorable* if there exists a k-coloring of G. The smallest positive integer k such that G is k-colorable is called the *chromatic number* of G and is denoted by $\chi(G)$. If G has chromatic number k then we say G is a *k-chromatic graph*.

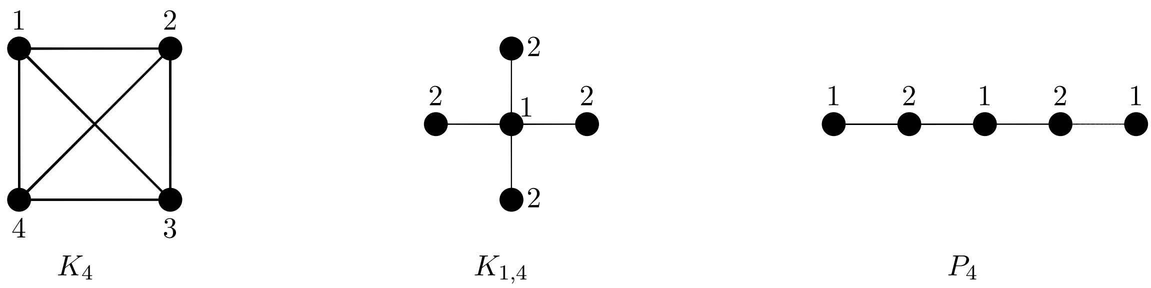


Figure 1.7: Coloring of $K_4, K_{1,4}$ and P_5 to show their chromatic numbers.

For any other undefined concepts, please see [2].

1.6 Conclusion

In this chapter, we gave definitions that we will use through out this dissertation. In Section 1.1, we introduced the reader to what we intend to study. Then, we gave general definitions of graphs in Section 1.2. Then, in the remaining three sections, Section 1.3, Section 1.4, and Section 1.5, we respectively, gave definitions that focus on distance of a graph, independence number and chromatic number. In the next chapter, we will review relevant literature for our work.

Chapter 2

Literature review

2.1 Introduction

Graph theory ideologies have been used in many science, humanities and engineering fields. In particular, the idea of distance is one of the most explored concepts in graph theory, since from it, we study how far apart one vertex is from the rest of the vertices in a graph. Most of the graph theory concepts can be linked back to distance. Average eccentricity and average distance are some of the concepts that can be linked to distance, as seen from the definitions of the two concepts. Bounds on average eccentricity and average distance have been presented before in terms of different graph parameters. We aim to show upper bounds on average eccentricity when order and independence number of a graph are known. Also, we will show upper bounds on average eccentricity when order and chromatic number are prescribed. We give bounds on average distance in terms of order and independence number, and also, in terms of order and chromatic number of the graph.

In this chapter, we will first give the motivation for studying average eccentricity and average distance, and then we will give some of the background work that has been done on the bounds on average eccentricity and average distance in terms of other graph parameters.

2.2 Motivation and Background

2.2.1 Motivation

Graphs have many parameters that help us to describe how the vertices are connected to one another. For instance, we can use degrees of vertices to describe the number of vertices a vertex is adjacent to. The graph parameters help with classifying graphs of the same order, which is convenient for application. The interest in studying bounds on average eccentricity and average distance is due to the vast application power the two concepts provide. For instance, in chemistry, the Wiener index, a variant of average distance, is used to estimate the boiling points of some of the chemical compounds [25], and in communications network, average eccentricity can be used to estimate the time a message takes to arrive [8]. These two concepts are also applied in architecture, social sciences, crystallography, cryptography, and facility location [13, 25].

2.2.2 Background

Average distance was first studied by a chemist named Harold Wiener. In 1947, Wiener noticed that, the boiling points of acyclic compounds could be calculated using the Wiener index [20, 25]. Initially, other chemists were skeptical of the result, but eventually, they accepted the concept [25]. Mathematicians began to study the Wiener index almost thirty years after chemists had started working on it. This was initially done without the aid of chemists' work on $W(G)$ [11, 25], leading to a slightly different, but equivalent definition of $W(G)$, see [11]. Doyle and Graver were the first to define average distance as a graph parameter [20], and since then, the study on average distance has grown greatly. For instance, graph theorists have studied bounds on average distance with respect to other graph parameters, like minimum degree, maximum degree, order, and size of the graph.

From eccentricity of a vertex in a graph, we can define more graph parameters. For instance, average eccentricity was introduced few decades after the introduction of average distance. In 1988, Skorobogatov and Dobryninin introduced the concept of average eccentricity [23], then, in 1990, Buckley and Harary named it eccentric mean [8, 23]. Researchers in chemistry and biology have defined other topological parameters that are closely related to the average eccentricity, one of them being eccentric connectivity index, defined as the sum of products of degree and eccentricity of each vertex in the graph, see [16]. Some other parameters that are defined using eccentricity, are first Zagreb eccentricity index and the second Zagreb eccentricity index. The first Zagreb eccentricity index is defined as the sum of the squares of eccentricities of each vertex in a graph, and the second Zagreb eccentricity index is defined as the sum of the product of eccentricities of adjacent vertices in a graph, see [22].

Considering the number of graph parameters that have been defined in graph theory, it is difficult to relate all of them manually. To combat this problem, computer programs were designed to take graph parameters, graph models and a relation, amongst others, as inputs, see [1, 10, 15]. From 1981 to 1984, the University of Belgrade, Faculty of Electrical Engineering, developed a computer program called "Graph", which is designed to help with verifying, giving and disproving conjectures in graph theory [4, 5]. In 1985, Siemion Fajtlowicz designed a computer program called Graffiti, see [15], and in 2001, Ermelinda DeLaVina developed Graffiti.pc, with an end goal of using Graffiti-like program without a middle man, see [10]. Unlike "Graph", Graffiti and Graffiti.pc give conjectures, but they do not prove or disprove those conjectures [10]. Before Ermelinda DeLaVina developed Graffiti.pc, AutoGraphix had been developed at Gerad, Canada, since 1997 [1]. AutoGraphiX or AGX, was developed to assist graph theorists in being able to find extremal graphs, and it gives conjectures with suggestions on how to prove them [1]. In the sequel, we will be mentioning, from time to time, some of the relevant contributions into literature by these computer programs.

Coming back to bounds, in [8], upper bounds on average eccentricity are given in terms of order and minimum degree of the graph. They also give extremal graphs. But, there are other graph parameters, other than, minimum degree and order, which can be related to the average eccentricity. This is what Yunfang Tang and Bo Zhou did in [23]; they gave lower bounds on average eccentricity in terms of the number of vertices and number of edges, and they gave the extremal graph as well. They then continued to give lower and upper bounds on the average eccentricity of n-vertex trees with fixed diameter, fixed number of pendent vertices and fixed matching number, respectively. Du and Ilic studied some of the conjectures created by AGX in [13]. One of the two conjectures they proved true give an upper bound on average eccentricity with respect to order and independence number, while the other conjecture gives an upper bound on average eccentricity with respect to order and chromatic number of the graph. Also, the conjecture that relates average eccentricity, order and chromatic number is similar to a theorem we are going to prove in this dissertation, and the method they used to prove the conjecture is similar to the method we used in this dissertation. Interestingly, the authors of [9] and [13] worked on the proof of the conjecture independently.

Average distance, Wiener index and the transmissions of a graph are all related (see, [18] and

[24]). In most cases, it is tedious to work with average distance, so, researchers opt to work with either Weiner index or the transmission of the graph. For instance, in [6], Dankelmann presented a sharp upper bound on the average distance of a graph in terms of order and the independence number of the graph by first proving an upper bound on the transmission of the graph in terms of the independence number and the order. He then deduced an upper bound on the average distance in terms of independence number and the order. Few years later, Dankelmann [7], presented upper bounds on average distance in terms of order and domination number . He used similar methods as in [6] to draw conclusion in [7]. Chung in [3], proved a conjecture produced by Graffiti, which states that the average distance is at most the independence number of the graph. To prove the conjecture, Chung used neither of the Wiener index nor the transmission, rather, she used only the average distance. In this dissertation, we will use Wiener index and the transmission to draw conclusions on upper bounds on average distance.
A lot of research results on average distance and average eccentricity are found in literature. In the next section, we present relevant results that are important for this study.

2.3 Survey of Important Results

2.3.1 Average distance results

In 1975, when Doyle and Graver first defined average distance as a graph parameter in [12], they gave bounds on average distance in terms of order. They used $h(v)$, a sum of tripple products of order of branches, which is defined as follows. Assume that G is a tree and $|V(G)| = n$. For $v \in V(G)$, let $m_1, m_2, m_3, \ldots, m_q$ be the orders of branches of $G - v$. Then

$$h(v) = \sum_{1 \leq i < j < k \leq q} m_i \cdot m_j \cdot m_k.$$

They then proved the following theorem which provided a basis for their result.

Theorem 2.3.1. *If G is a tree with n vertices, then*

$$\mu(G) = \frac{n+1}{3} - \frac{2}{n(n-1)} \sum_{v \in V(G)} h(v).$$

By observing that $1 = \mu(K_n) \leq \mu(G)$, for any connected graph G of order n, and by applying the fact that, for any spanning tree T of G, we have that $\mu(G) \leq \mu(T)$, to Theorem 2.3.1, they deduced the following corollary.

Corollary 2.3.2. *If G is a connected graph with $|V(G)| = n$, then*

$$1 \leq \mu(G) \leq \frac{n+1}{3},$$

with equality holding on the left if and only if G is complete and on the right if and only if G is a path.

Then in 1976, Entringer, Jackson and Snyder in [14] presented bounds on average distance in terms of size and order, and their results were given by the following theorem,

Theorem 2.3.3. *If G is a connected graph with n vertices and k edges, then*

$$n(n-1) \leq d(G) + k \leq \frac{1}{6}(n^3 + 5n - 6),$$

where $d(G) = \frac{1}{2}\sum_{u,v\in V(G)} d_G(u,v)$. To prove the lower bound of Theorem 2.3.3, $d(G)$ was expressed in terms of the size of the graph and the sum of distances of non-adjacent vertices in G, and from the obtained expression, they obtained the lower bound. They then used induction on n, followed by induction on k, to show that the upper bound holds.
Few years later, in 1984, Plesník improved the bound from Corollary 2.3.2 to 2-vertex-connected and 2-edge-connected graphs in [19]. They showed the improved bounds by presenting proofs for the following theorems:

Theorem 2.3.4. *Consider a graph G and assume that G is a 2-vertex-connected graph with n vertices. Then*

$$\sigma(G) \leq n\left\lfloor \frac{1}{4}n^2 \right\rfloor,$$

and equality holds true if and only if G is a cycle,

and

Theorem 2.3.5. *Consider a graph G. Assume G is a 2-edge-connected graph with n vertices. Then*

$$\sigma(G) \leq n\left\lfloor \frac{1}{4}n^2 \right\rfloor,$$

and equality holds true if and only if G is a cycle.

Theorem 2.3.4, followed from the following lemma,

Lemma 2.3.6. *Consider a 2-vertex-connected graph G with n vertices. Then for any vertex v we have $\sigma(v) \leq \left\lfloor \frac{1}{4}n^2 \right\rfloor$. Equality holds for every $v \in V(G)$ if and only if G is a cycle.*

To prove Lemma 2.3.6, Plesník used the idea of distance layers. Then using Lemma 2.3.6 and the fact that $\sigma(G) = \sum_{v\in V(G)} \sigma(v)$, he was able to show that Theorem 2.3.4 is true. Induction is used on n to prove Theorem 2.3.5.

In 1988, Chung showed that, average distance of a connected graph is at most the independence number of the graph, by presenting a proof for the following theorem in [3],

Theorem 2.3.7. *For every connected graph G,*

$$\mu(G) \leq \alpha(G),$$

and equality holds if and only if G is a complete graph.

Theorem 2.3.7 was proven by induction on the order of the graph. They considered average distance from some vertex v, defined by

$$f(v) = \frac{1}{n-1}\sum_{u\neq v} d_G(u,v),$$

and then they related $f(v)$ to $\mu(G)$. Then, they made four claims. In the first claim, they proved that if $G - w$ is connected for a vertex w of G, then $\alpha(G) \geq \mu(G)$, by using the fact that $\alpha(G) \geq \alpha(G - w)$ and the induction hypothesis. Then, they proceeded by assuming that, every vertex w such that $f(w) \leq \mu(G)$ is a cut-vertex. Then, they claimed that, if, for some vertex w, $f(w) \leq \mu(G) - 1$, then $\alpha(G) > \mu(G)$. They proved the second claim by considering neighbors of w, which they assumed to be $w_1, w_2, \ldots, w_k$, and they have that, for $1 \leq i \leq k$, $f(w_i) \leq f(w) + 1 \leq \mu(G)$, giving that w_i's are cut-vertices as well. Then, they concluded the proof for the second claim by showing the claim holds for two cases, when $k = 2$ and for the case when $k \geq 3$. The third claim they proved is given as follows, if $diam(G) \geq 2\lfloor\mu(G)\rfloor - 1$, then $\alpha(G) > \mu(G)$. To prove the third claim, they used distance layers L_i. They considered the maximum independent set L of the L_i, and they further considered two independent subsets of L, where one set is contained in L if i is even and the other set is contained in L if i is odd. Using the two subsets of L, they showed that the claim holds. To prove the fourth claim, which states that, if $f(w) > \mu(G) - 1$, then $\alpha(G) > \mu(G)$, they first took w to be a vertex such that, $\max\{d_G(w, v) : v \in G\}$ is minimized among all w such that $f(w) \leq \mu(G)$, and they assume that, the number of vertices v with $d_G(w, v) = \max\{d_G(w, v) : v \in G\}$ is as small as possible. Then, they considered one component of $G - w$, denoted by A, with vertex u such that $d_G(u, w) = \max\{d_G(w, v) : v \in G\} \geq \lfloor\mu(G)\rfloor$. Then they used the third claim, together with A to bound $d_G(w, x)$ in terms of $\lfloor\mu(G)\rfloor$, for any $x \in G - A - w$. Then, they chose a vertex z that lies on the shortest path $P(u, w)$ joining u and w such that $d_G(z, u) = \lfloor\mu(G)\rfloor - 1$. Then, to conclude the fourth claim, they considered two sub-cases. A case when there is a vertex v such that $d_G(z, u) = \lfloor\mu(G)\rfloor + 1$ and the case when $d_G(v, z) \leq \lfloor\mu(G)\rfloor$ for any v.

In 1994, Dankelmann presented upper bounds on average distance with respect to order and independence number of the graph in [6]. The extremal graph in his results, Theorem 2.3.8, is defined as follows. For positive integers n, k with $2 \leq k \leq \frac{1}{2}n$, let $G_{n,k}$ be the graph obtained from a path P_{2k-2} with end vertices v_1, v_2 and two disjoint complete graphs G_1, G_2 of order

$$n_1 = \left\lfloor \frac{n}{2} \right\rfloor - k + 1, \;\; n_2 = \left\lceil \frac{n}{2} \right\rceil - k + 1$$

by joining v_i with each vertex of G_i for $i = 1, 2$. While for positive integers n, k with $\frac{1}{2}n \leq k \leq n - 1$, $G_{n,k}$ is the graph obtained from a path $P_{2n-2k-1}$ with end vertices v_1, v_2 and two disjoint empty graphs G_1, G_2 of order

$$n_1 = k - \left\lfloor \frac{n-1}{2} \right\rfloor, \;\; n_2 = k - \left\lceil \frac{n-1}{2} \right\rceil$$

by joining v_i with each vertex of G_i for $i = 1, 2$. In particular, Dankelmann proved the following theorem.

Theorem 2.3.8. *Let n and α be two integers satisfying $2 \leq \alpha \leq n - 1$. Then*

$$\mu(G) \leq \mu(G_{n,\alpha}).$$

Equality is true if and only if G is isomorphic to $G_{n,\alpha}$.

To prove Theorem 2.3.8, construction and induction methods were used to draw a conclusion. Define $\rho(n, \alpha)$ to be the class of all connected graphs of order n and independence number α. The following lemmas were used as well to prove Theorem 2.3.8.

Lemma 2.3.9. *For graph $G \in \rho(n, 2)$, and $n \geq 3$, we have that*

$$\sigma(G) \leq \sigma(G_{n,2}).$$

Equality is possible if and only if G is isomorphic to $G_{n,2}$.

Lemma 2.3.10. *(a) For integers n, k, $2 \leq k \leq \frac{1}{2}n$, let $H'_{n,k}$ denote the connected graph of order n, which consist of a path P_{2k-1} with an end vertex v, a complete graph K_{n-2k+1}, and an edge joining v to exactly one vertex of the k-clique.*

(b) For integers n, k, $\frac{1}{2}n < k \leq n - 1$, let $H'_{n,k}$ denote the connected graph of order n, which consists of a path P_{2n-2k} with an end vertex v, an empty graph of order $2k - n$ and edges joining

v to the empty graph.
Let $H \in \rho(n,k)$ and $x \in V(H)$. Then we have

$$\sigma_H(x) \leq (2k-1)(n-k).$$

Equality holds if and only if H is isomorphic to $H'_{n,k}$ and x is the unique end vertex of H.

After having related average distance with order and independence number, in 1997, Dankelmann presented upper bounds on average distance with respect to order and domination number, by proving Theorem 2.3.11 and Theorem 2.3.15 in [7]. Theorem 2.3.11 is stated as follows.

Theorem 2.3.11. *For a connected graph G with n vertices and domination number $\gamma \leq \frac{n}{3}$, we have that*

$$\mu(G) \leq \begin{cases} \frac{n+1}{3} - \frac{(n-3\gamma)(n-3\gamma+2)(2n+3\gamma-7)}{6n(n-1)} & if \quad n-\gamma \text{ is even,} \\ \\ \frac{n+1}{3} - \frac{(n-3\gamma)(n-3\gamma+2)(2n+3\gamma-7)-9(\gamma-1)}{6n(n-1)} & if \quad n-\gamma \text{ is odd.} \end{cases}$$

Equality holds if and only if $G = G_{n,\gamma}$.

To prove Theorem 2.3.11, they inducted on n and γ using Lemma 2.3.12, Lemma 2.3.13 and Lemma 2.3.14 to draw conclusions. The forementioned lemmas are stated as follows.

Lemma 2.3.12. *Let H be a graph. The following two statements are equivalent.*

1. *$\gamma(H-e) > \gamma(H)$ for each edge $e \in E(H)$.*
2. *H is the union of vertex disjoint stars.*

Lemma 2.3.13. 1. *Let G be a connected graph and $v \in V(G)$ a cut-vertex. Then there is a vertex $w \in N(v)$ with $\sigma(w) > \sigma(v)$.*

2. *If G is a tree and v is neither an end vertex nor adjacent to an end vertex, then there is a vertex w, adjacent to an end vertex, with $\sigma(w) > \sigma(v)$.*

In Lemma 2.3.14, the extremal graph is defined as follows.

i. For positive integers n, γ with $1 \leq \gamma \leq \frac{n}{3}$ let $H_{n,\gamma}$ be the graph consisting of a path $P_{3\gamma-1} = v_1v_2\ldots v_{3\gamma-1}$ and independent vertices $w_1,\ldots,w_{n+1-3\gamma}$ that are joined with $v_{3\gamma-1}$.

ii. For positive integers n, γ with $\frac{n}{3} \leq \gamma \leq \frac{n}{2}$ let $H_{n,\gamma}$ be the graph obtained from a path $P_{2n-3\gamma+1} = v_1v_2\ldots v_{2n-3\gamma+1}$ and independent vertices $w_{3n-6\gamma+3},\ldots,w_{2n-3\gamma+1}$ by joining v_i and w_i for $3n-6\gamma+3 \leq i \leq 2n-3\gamma+1$.

Sketches of the described extremal graphs are provided in [7]. We now state Lemma 2.3.14,

Lemma 2.3.14. *Consider a tree G with n vertices and domination number γ. For each vertex $v \in V(G)$, we have*

$$\sigma(v) \leq \begin{cases} (3\gamma-1)(n-\frac{3}{2}\gamma) & if \quad \gamma \leq \frac{n}{3}, \\ \\ (2n-3\gamma+1)^2 - \frac{1}{2}(3n-6\gamma+3)(3n-6\gamma+2) & if \quad \gamma > \frac{n}{3}. \end{cases}$$

Equality is attained if and only if $G = H_{n,\gamma}$ and $v = v_1$.

Theorem 2.3.15 is stated as follows,

Theorem 2.3.15. *For a connected graph G with n vertices and domination number $\gamma \geq \frac{n}{3}$, we have*

$$\mu(G) \leq \begin{cases} \frac{n+1}{3} - \frac{(3\gamma-n)(3\gamma-n-2)(5n-6\gamma-4)}{3n(n-1)} & if \quad n-\gamma \text{ is even,} \\ \frac{n+1}{3} - \frac{(3\gamma-n-1)(3\gamma-n-3)(5n-6\gamma-2)+6(2n-3\gamma-1)}{3n(n-1)} & if \quad n-\gamma \text{ is odd.} \end{cases}$$

with equality if and only if $G = G_{n,\gamma}$.

They presented a proof of Theorem 2.3.15, using induction on n and Lemma 2.3.13.

In the same year, 1997, Kouider and Winkler presented the upper bounds on average distance in terms of order and minimum degree of the graph in [17]. They presented a proof of the following theorem.

Theorem 2.3.16. *Let G be a graph with n vertices and minimum degree δ. Then*

$$\mu(G) \leq \frac{n}{\delta+1} + 2.$$

To prove Theorem 2.3.16, they first considered a path $P(u, v)$ given by $u = x_0 x_1 \ldots x_{d_G(u,v)} = v$, followed by considering neighborhoods of vertices on $P(u, v)$, and selecting three disjoint sets A_1, A_2, and A_3 using vertices of $P(u, v)$. By using the cardinalities of the forementioned sets, A_1, A_2, and A_3, they defined a "score" function, and showed that, the score function is at most

$$\frac{n(n-1)(n-2)}{6}.$$

They proceeded to calculate the lower bound on the score function in terms of minimum degree, $|A_1|$, $|A_2|$, $|A_3|$ and average distance. Now, using the bounds of the score function, they were able to show that Theorem 2.3.16 holds.

2.3.2 Average eccentricity results

In 2004, Dankelmann, Goddard, and Swart initiated studies of bounds of average eccentricity by presenting upper bounds on average eccentricity in terms of order and minimum degree in [8]. They needed the following definition to prove their results:

Definition 2.3.17. Let G be a graph without multiple edges. Let $V(G) = \{v_1, v_2, \ldots, v_n\}$. A pair (G, w) is called a *weighted graph* if $w : V(G) \times V(G) \to \mathbb{R}$ is a function satisfying: $w(v_j, v_i) = w(v_i, v_j)$; $w(v_i, v_j) \geq 0$ and $w(v_i, v_j) > 0$ if and only if $v_i v_j$ is an edge in G, for all $v_i, v_j \in V(G)$.

Then, for a weighted graph G with weight function $c : V(G) \to \mathbb{R}$, they defined the *eccentricity of G with respect to c* as

$$EX_c(G) = \sum_{x \in V(G)} c(x) e_G(x)$$

and the *average eccentricity of G with respect to c* as

$$avec_c(G) = \frac{EX_c(G)}{\sum_{x \in V(G)} c(x)}.$$

Being concerned about upper bounds on average eccentricity in terms of order and minimum degree, they proved the following lemma.

Lemma 2.3.18. *Let G be a weighted graph with weight function c such that $c(v) \geq 1$ for every vertex v, and $N = \left\lceil \sum_{v \in V(G)} c(v) \right\rceil$. Then*

$$avec_c(G) \leq avec(P_N).$$

They used Lemma 2.3.18 to show upper bound on average eccentricity of a connected graph with order n and minimum degree δ is as stated in Theorem 2.3.19.

Theorem 2.3.19. *Consider a connected graph G with n vertices and minimum degree δ. Then*

$$avec(G) \leq \frac{9n}{4\delta+4} + \frac{15}{4}. \tag{2.1}$$

Excluding the additive constant, we have that the inequality is attainable.

To prove Theorem 2.3.19, they first constructed a maximal set A such that every vertex v in A is at least distance 3 from all other vertices of A. Afterwards, they considered a forest subgraph T_1 of G, such that, T_1 has vertex set made up of vertices in A and any vertex adjacent to A in G, denoted by $N[A]$. Using the construction of A and T_1, they constructed a subtree T_2 of G that is obtained by joining components of T_1 using $|A|-1$ edges in G that join the components through vertices in $N[A]-A$. By considering a spanning tree T of G with edge set $E(T_2) \cup \{uu' : u \in V(G) - N[A], u' \in N[A]\}$, they showed that

$$avec(T) \leq \frac{9n}{4\delta+4} + \frac{15}{4},$$

by also considering a weight function $c : V(T) \to \mathbb{R}$, given by

$$c(u) = |\{x \in V(T)\,|x_A = u\}|,$$

where $u \in V(T)$, with x_A being a vertex closest to x in A. Then, they argued that

$$avec(T) \leq avec_c(T) + 2.$$

Then, they constructed a tree U using A and T to obtain

$$avec_c(T) \leq 3avec_c(U) + 1. \tag{2.2}$$

By using Lemma 2.3.18, (2.2), with $N = \left\lceil \frac{n}{\delta+1} \right\rceil$, they concluded that

$$avec(T) \leq \frac{9n}{4\delta+4} + \frac{15}{4}.$$

To show that the inequality is best possible in (2.1), they constructed a graph that achieves equality in (2.1) for large n.
Eight years after Dankelmann, Goddard and Swart had presented their results in [8], Tang and Zhou followed with a study that gives bounds on average eccentricity in terms of order and size in [23]. Firstly, they proved a proposition that gives the lower bound on average eccentricity of a connected graph with order n in terms of the number of vertices with degree $n-1$ and order n, and it is stated as follows,

Proposition 2.3.20. *Let G be an n-vertex connected graph, and k the number of vertices of degree $n-1$ in G, where $0 \leq k \leq n$. Then*

$$avec(G) \geq 2 - \frac{k}{n}$$

with equality if and only if all of the vertices of degree less than $n-1$ have eccentricity two.

To prove Proposition 2.3.20, Tang and Zhou provided the reader with a note, mentioning that, there are k vertices with eccentricity one and $n-k$ vertices with eccentricity at least two. So, the result will then follow. They followed the above proposition with a definition. They defined the set $G(n,m)$ as follows. Let $G \vee H$ be the graph formed from vertex-disjoint graphs G and H by adding edges between each vertex in G and each vertex in H. Then, $G(n,m)$ denotes the set of graphs $K_a \vee H$ with n vertices and m edges, where $a = \left\lfloor \frac{2n-1-\sqrt{(2n-1)^2-8m}}{2} \right\rfloor$.
Then, they used Proposition 2.3.20 to prove Proposition 2.3.21, which gives the lower bound on average eccentricity in terms of order and size,

Proposition 2.3.21. *Let G be an n-vertex connected graph with m edges, where $n-1 \leq m < \binom{n}{2}$. Let $a = \left\lfloor \frac{2n-1-\sqrt{(2n-1)^2-8m}}{2} \right\rfloor$. Then*

$$avec(G) \geq 2 - \frac{a}{n}$$

with equality if and only if $G \in G(n,m)$.

To prove Proposition 2.3.21, Tang and Zhou started by denoting the number of vertices with degree $n-1$ by k, where $0 \leq k \leq n-1$. They showed that Proposition 2.3.21 is true by using Proposition 2.3.20, and considering the cases when $k = 0$ and $k \geq 1$, together with the fact that $2m \geq k(n-1) + k(n-k)$.
Then, Tang and Zhou proceeded to give proof for Proposition 2.3.22, which gives an upper bound on average eccentricity in terms of order and size. The extremal graph used in Proposition 2.3.22 is defined as follows. Let $K_n - ke$ be a graph formed by deleting k independent edges from the complete graph K_n, where $k = 1, 2, \ldots, \lfloor \frac{n}{2} \rfloor$.

Proposition 2.3.22. *For a connected graph G with $n \geq 2$ vertices and size m, we have that*

$$avec(G) \leq n - \frac{2m}{n}. \tag{2.3}$$

Equality is possible if and only if $G = K_n - ke$ for $k = 0, 1, \ldots, \lfloor \frac{n}{2} \rfloor$, or $G = P_4$.

To prove Proposition 2.3.22, Tang and Zhou used order of distance layers, $|N_i(x)|$, for $i = 1, 2, \ldots, e_G(x)$. From the fact that

$$n - 1 = deg_G(x) + \sum_{i=2}^{e_G(x)} |N_i(x)|,$$

they showed that $deg_G(x) + e_G(x) \leq n$. This gives,

$$avec(G) = \frac{1}{n} \sum_{x \in V(G)} e_G(x) \leq n - \frac{2m}{n}. \tag{2.3}$$

To show the conditions for equality in (2.3), they first noted that, either $e_G(x) = 1$ with $deg_G(x) = 1$ or $e_G(x) \geq 2$ with $|N_2(x)| = |N_3(x)| = \cdots = |N_{e_G(x)}(x)| = 1$ for every $x \in V(G)$. Then, they showed that equality holds in the proposition for the case $e_G(x) = 1$, and for the case $e_G(x) \geq 2$. They assumed that, it is easy to show that $avec(G) = n - \frac{2m}{n}$ for $G = K_n - ke$ with $k = 0, 1, \ldots, \lfloor \frac{n}{2} \rfloor$ or $G = P_4$, thus omitting the proof.
In the same year, 2012, that Tang and Zhou presented their results, in [23], Ilić presented a proof for a conjecture that relates average eccentricity to independence number and order, in [16]. Ilić used the following lemma in the proof of the conjecture,

Lemma 2.3.23. *Let T be an arbitrary tree on n vertices, not isomorphic to a path P_n. Then there is a pendent vertex v such that for each $u \in V(T)$ it holds*

$$e_T(u) = e_{T-v}(u).$$

The conjecture was obtained from AutoGraphix. The extremal case is given by the graph $B(n, \Delta)$ consisting of a star $S_{\Delta+1}$ and a path $P_{n-\Delta+1}$ with one of the end vertices of $P_{n-\Delta+1}$ attached to one of the pendent vertices of $S_{\Delta+1}$. The conjecture is stated as follows,

Conjecture 2.3.24. *For every $n \geq 4$ it holds*

$$avec(G) + \alpha(G) \leq \begin{cases} \frac{3n^2-2n-1}{4n} + \frac{n+1}{2} & \text{if } n \text{ is odd,} \\ \frac{3n^2-4n-4}{4n} + \frac{n+2}{2} & \text{if } n \text{ is even} \end{cases}$$

with equality if and only if G is isomorphic to P_n for odd n, and G is isomorphic to $B(n,3)$ for even n.

To prove Theorem 2.3.24, Ilić considered an extremal tree T^* with diametrical path $P = v_0v_1 \dots v_d$. Then using Lemma 2.3.23, he argued that $\alpha(B(n, n-d+1)) + avec(B(n, n-d+1))$ and $\alpha(T^*) + avec(T^*)$ are equal. Then by direct calculation of $\alpha(B(n, n-d+1)) + avec(B(n, n-d+1))$, he concluded that the theorem holds.
Then, in 2013, Du partnered with Ilić, to present proofs to more conjectures provided by the AutoGraphix computer system in [13]. Firstly, they provided a proof to the conjecture that relates average eccentricity to order and independence number. This was followed by a conjecture relating average eccentricity and chromatic number and order. The first conjecture is stated as follows,

Conjecture 2.3.25. *Let G be a graph with n vertices such that $n \geq 4$. Then $avec(G)/\alpha(G)$ is maximized by some graph composed of two cliques linked by a path.*

To prove Conjecture 2.3.25, Du and Ilić showed that, the conjecture holds for two cases. For the first case, they considered a connected graph G of order n with $\alpha(G) \geq \left\lceil \frac{n}{2} \right\rceil$, and in the second case they have that $\alpha(G) < \left\lceil \frac{n}{2} \right\rceil$.
To prove the second conjecture, they had to define an extremal graph, as follows. Let $LP(n,t)$ denote the graph obtained from a complete graph K_t and a path P_{n-t+1} with one of the end vertices of P_{n-t+1} attached to one of the K_t vertices by an edge. $LP(n,t)$ is called the lollipop graph. The second conjecture is stated as follows,

Conjecture 2.3.26. *For every $n \geq 4$ it holds*

$$avec(G) - \alpha(G) \leq \begin{cases} \frac{n}{4} - \frac{1}{2} & \text{if } n \text{ is even,} \\ \frac{n}{4} - \frac{1}{2} - \frac{3}{4n} & \text{if } n \text{ is odd} \end{cases}$$

with equality if and only if G is isomorphic to P_n for even n, and G is isomorphic to $LP(n,3)$ for odd n.

To prove Conjecture 2.3.26, Du and Ilić considered a connected graph G of order $n \geq 4$ and maximum $avec(G) - \alpha(G)$. Then, they noted that, the inequality can be obtained using some of the methods used in the proof of Theorem 2.3.25. Then, they proceeded to use contradiction to show that equality holds for when G is isomorphic to P_n with n even, and G is isomorphic to $LP(n,3)$ with odd n.
Then, with the aid of Brooks theorem, which is stated as follows: *For any connected graph G with maximum degree Δ, $\chi(G) \leq \Delta$ unless G is a clique or an odd cycle, in which case $\chi(G) = \Delta + 1$*, Du and Ilić proceeded to prove the following conjecture, which gives upper bound on $avec(G) \cdot \chi(G)$, and is stated as follows,

Conjecture 2.3.27. *Consider a graph G with order n with $n \geq 4$. Then $avec(G) \cdot \chi(G)$ is maximized by some lollipop graph.*

To prove Conjecture 2.3.27, Du and Ilić considered a connected graph G of order n with chromatic number χ, maximum degree Δ and diameter d. Then, they directly showed that $avec(G) \leq avec(LP(n,\chi))$ for $\chi = 2$, $\chi = 3$ and $\chi = n$. Then, they used Brooks theorem to show that $avec(G) \leq avec(LP(n,\chi))$ when $4 \leq \chi \leq n-1$. Using the inequality $avec(G) \leq avec(LP(n,\chi))$, they concluded that the maximum value of $avec(G) \cdot \chi$ is achieved by $LP(n,\chi)$.

2.4 Conclusion

In this chapter, we began with a brief introduction, where we informed the reader about the study fields that incorporate graph theory ideologies within their analysis. Also, we made the reader aware of how most of the graph theory concepts relate back to distance. Then, we gave motivation of studying graph parameters, and why we might want to study bounds on average eccentricity and average distance in terms of other graph parameters. We proceeded to give background work on average eccentricity and average distance. We gave a history of how the study on bounds on average eccentricity and average distance came to be, and we mentioned authors who have taken interest in studying those bounds. We also mentioned computer programs used to generate some of the hypothesis for the results that have been studied thus far. We proceeded to give survey of results that are closely related to the results of this dissertation.

Chapter 3

Upper bounds on the average eccentricity

3.1 Introduction

In this chapter, we are going to study upper bounds on average eccentricity in terms of independence number and order of the graph, and then in terms of chromatic number and order of the graph. Instead of working with $avec(G)$, for easy calculations, we will work with $\zeta(G) = \sum_{v \in V(G)} e_G(v)$ to analyse average eccentricity results, and convert back to $avec(G)$ thereafter.

Main results of this chapter were studied by Dankelmann and Mukwembi in [9]. We will employ methods of [9] to prove main results of this chapter.

3.2 Average eccentricity, order and independence number

We dedicate this section to proving results that will assist us in showing an upper bound on average eccentricity in terms of order and independence number of a graph, and we conclude the section by proving a theorem that gives an upper bound on average eccentricity in terms of order and independence number of a graph. We begin with proving a result that allows us to compute the eccentricities of vertices in a tree easily.

Lemma 3.2.1. *Consider a tree T with two vertices u and v at distance diam(T). For all $x \in V(T)$, we have*

$$e_T(x) = \max\{d_T(x, u),\, d_T(x, v)\}.$$

Proof: We proceed by contradiction. Suppose T is a tree and u and v are vertices at distance $diam(T)$ in T, with

$$e_T(x) \neq \max\{d_T(x, u),\, d_T(x, v)\}$$

for some vertex $x \in V(T)$.
Then under the supposition, we have the following claim.

Claim 3.2.2. $e_T(x) > \max\{d_T(x, u),\, d_T(x, v)\}$

Proof of Claim 3.2.2
Since we have that $e_T(x) \neq \max\{d_T(x, u),\, d_T(x, v)\}$ for vertex $x \in V(T)$, it follows that either

$$e_T(x) > \max\{d_T(x, u), d_T(x, v)\}$$

or

$$e_T(x) < \max\{d_T(x,u), d_T(x,v)\}.$$

Suppose that $e_T(x) < \max\{d_T(x,u),\, d_T(x,v)\}$, and assume without loss of generality that

$$\max\{d_T(x,u), d_T(x,v)\} = d_T(x,u)$$

and let $e_T(x) = d_T(x,y)$ for some vertex $y \in V(T)$. Then, since $e_T(x) < d_T(x,u)$ we have that

$$e_T(x) = d_T(x,y) < d_T(x,u),$$

contradicting the fact that y is a vertex furthest away from x. Therefore, we must have that $e_T(x) > \max\{d_T(x,u), d_T(x,v)\}$ and Claim 3.2.2 is proven.

Now, let $e_T(x) = d_T(x,y)$ for some vertex $y \in V(T)$. By Claim 3.2.2, we have that,

$$e_T(x) = d_T(x,y) > \max\{d_T(x,u),\, d_T(x,v)\}.$$

Let $P(u,v)$ and $P(x,y)$ denote a u-v and an x-y path in T, respectively. We consider the following cases.

Case 1: When the intersection between vertices in $V(P(u,v))$ and vertices in $V(P(x,y))$ contain at most one vertex in it.

Case 2: When the intersection between the vertices in $V(P(u,v))$ and vertices in $V(P(x,y))$ contain more than one vertex.

Proof for the first case.

Let w be a vertex that lies on $P(u,v)$ and let z be a vertex that lies on $P(x,y)$. Assume that w and z are at a minimum distance, it follows that, $w = z$ if and only if the two paths have only one vertex in common. We demonstrate the relationship of $P(u,v)$ and $P(x,y)$ in T when $V(P(u,v)) \cap V(P(x,y)) = \{z\} = \{w\}$ by Figure 3.1 below.

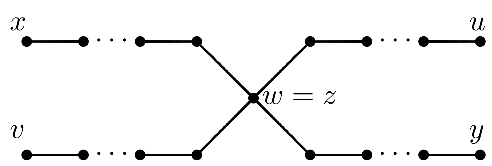


Figure 3.1: Subgraph of Tree T with paths $P(u,v)$ and $P(x,y)$ having one vertex in common.

From Figure 3.1, we see that, the distances $d_T(x,y)$, $d_T(x,u)$, and $d_T(x,v)$ can be written as follows,

$$d_T(x,y) = d_T(x,z) + d_T(z,y) = d_T(x,w) + d_T(w,y), \tag{3.1}$$
$$d_T(x,u) = d_T(x,z) + d_T(z,u) = d_T(x,w) + d_T(w,u), \tag{3.2}$$
$$d_T(x,v) = d_T(x,z) + d_T(z,v) = d_T(x,w) + d_T(w,v). \tag{3.3}$$

Since

$$d_T(x,y) > \max\{d_T(x,u), d_T(x,v)\},$$

we have

$$d_T(x, y) > d_T(x, u)$$

and

$$d_T(x, y) > d_T(x, u).$$

Hence, from (3.1) and (3.2), we get

$$\begin{aligned} d_T(x, z) + d_T(z, y) &= d_T(x, y) \\ &> d_T(x, u) \\ &= d_T(x, z) + d_T(z, u). \end{aligned}$$

Giving that, $d_T(z, y) > d_T(z, u)$. Also, from (3.1) and (3.3), we get

$$\begin{aligned} d_T(x, z) + d_T(z, y) &= d_T(x, y) \\ &> d_T(x, v) \\ &= d_T(x, z) + d_T(z, v). \end{aligned}$$

It follows that $d_T(z, y) > d_T(z, v)$. Then we have that $d_T(z, u)$, $d_T(z, v) < d_T(z, y)$. Since $w = z$, we have that $d_T(w, u)$, $d_T(w, v) < d_T(w, y)$.
This, in conjunction with the fact that $w = z$, yields

$$\begin{aligned} d_T(u, y) &= d_T(u, z) + d_T(z, y) \\ &= d_T(u, w) + d_T(w, y) \\ &> d_T(u, w) + d_T(w, v) \\ &= d_T(u, z) + d_T(z, v) \\ &= d_T(u, v) \\ &= diam(T). \end{aligned}$$

Thus, we obtain a contradiction since $diam(T) = d_T(u, v) \geq d_T(x, y)$ for all $x, y \in V(T)$. Therefore, $e_T(x) = \max\{d_T(x, u), d_T(x, v)\}$.
Proof for the second case.
Suppose that $P(u, v)$ and $P(x, y)$ share more than one vertex. For some vertices $w, z \in V(T)$, we have that $P(u, v)$ and $P(x, y)$ have a path segment $P(w, z)$ in common. Assume, without loss of generality, that w is closer to u on $P(u, v)$ and z is closer to v.

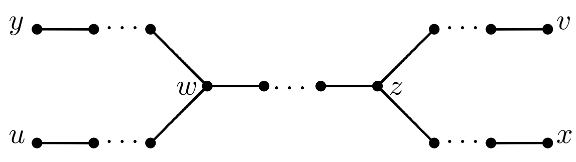


Figure 3.2: Subgraph of Tree T with paths $P(u, v)$ and $P(x, y)$ having more than one vertex in common.

Since $d_T(x, y) > \max\{d_T(u, x), d_T(x, v)\}$, we have that

$$d_T(x, y) > d_T(x, v).$$

Also,

$$z \in V(P(x,y)) \text{ and } z \in V(P(x,v)).$$

So,

$$d_T(x,y) = d_T(x,z) + d_T(z,y),$$

and, since z is closer to v on $P(u,v)$, we have

$$d_T(x,v) = d_T(x,z) + d_T(z,v).$$

Therefore,

$$\begin{aligned} d_T(x,z) + d_T(z,y) &= d_T(x,y) \\ &> d_T(x,v) \\ &= d_T(x,z) + d_T(z,v). \end{aligned}$$

It follows that $d_T(z,y) > d_T(z,v)$, and so

$$\begin{aligned} d_T(u,y) &= d_T(u,z) + d_T(z,y) \\ &> d_T(u,z) + d_T(z,v) \\ &= d_T(u,v) \\ &= diam(T). \end{aligned}$$

We have a contradiction, since $diam(T) = d_T(u,v) \geq d_T(x,y)$ for all $x, y \in V(T)$. Therefore, $e_T(x) = \max\{d_T(x,u), d_T(x,v)\}$, and this completes the proof of the lemma. □

Note 3.2.3. We note that:

- If $x \in \mathbb{N}$, then $\lfloor x \rfloor = x$.
- If $x \in \mathbb{N}$ and $y \in \mathbb{R}^+$, then $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$.

Proposition 3.2.4. *Consider a tree T with n vertices and diameter d. Then*

$$\zeta(T) \leq \left\lfloor dn - \frac{1}{4}d^2 + \frac{1}{4} \right\rfloor.$$

Proof: Let $P = v_0, v_1, \ldots, v_d$ be a diametral path of T. By Lemma 3.2.1, we have,

$$\begin{aligned} e_T(v_i) &= \max\{d_T(v_0, v_i), d_T(v_i, v_d)\} \\ &= \max\{i, d-i\} \\ &= \begin{cases} d-i & \text{if} \quad i \leq \lfloor \frac{d}{2} \rfloor, \\ i & \text{if} \quad i > \lfloor \frac{d}{2} \rfloor. \end{cases} \end{aligned}$$

Summing the eccentricities for all of $v_i \in V(P)$, we get the following:

$$\sum_{v_i \in V(P)} e_T(v_i) = \sum_{i=0}^{d} e_T(v_i)$$

$$
\begin{aligned}
&= \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} e_T(v_i) + \sum_{i=\lfloor \frac{d}{2} \rfloor+1}^{d} e_T(v_i) \\
&= \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} (d-i) + \sum_{i=\lfloor \frac{d}{2} \rfloor+1}^{d} i \\
&= \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} d - \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} i + \sum_{i=\lfloor \frac{d}{2} \rfloor+1}^{d} i \\
&= \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} d - \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} i + \sum_{i=0}^{d} i - \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} i \\
&= \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} d - 2\sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} i + \sum_{i=0}^{d} i \\
&= \Big(\Big\lfloor \frac{d}{2} \Big\rfloor + 1\Big)d - 2\Big[\frac{\big\lfloor \frac{d}{2} \big\rfloor \big(\big\lfloor \frac{d}{2} \big\rfloor + 1\big)}{2}\Big] + \frac{d(d+1)}{2} \\
&= \Big\lfloor \frac{d}{2} \Big\rfloor d + d - \Big\lfloor \frac{d}{2} \Big\rfloor \Big(\Big\lfloor \frac{d}{2} \Big\rfloor + 1\Big) + \frac{d(d+1)}{2} \\
&= \Big\lfloor \frac{d}{2} \Big\rfloor \Big(d - 1 - \Big\lfloor \frac{d}{2} \Big\rfloor\Big) + \frac{d^2}{2} + \frac{d}{2} + d. \\
&= \Big\lfloor \frac{d}{2} \Big\rfloor \Big(d - 1 - \Big\lfloor \frac{d}{2} \Big\rfloor\Big) + \frac{d^2}{2} + \frac{3d}{2}.
\end{aligned}
$$

If d is even, then,

$$
\begin{aligned}
\sum_{v_i \in V(P)} e_T(v_i) &= \Big\lfloor \frac{d}{2} \Big\rfloor \Big(d - 1 - \Big\lfloor \frac{d}{2} \Big\rfloor\Big) + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{d}{2}\Big(d - 1 - \frac{d}{2}\Big) + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{d}{2}\Big(\frac{d}{2} - 1\Big) + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{d^2}{4} - \frac{d}{2} + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{3d^2}{4} + d \\
&= \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} \Big\rfloor.
\end{aligned}
$$

If d is odd, then,

$$
\begin{aligned}
\sum_{v_i \in V(P)} e_T(v_i) &= \Big\lfloor \frac{d}{2} \Big\rfloor \Big(d - 1 - \Big\lfloor \frac{d}{2} \Big\rfloor\Big) + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{d-1}{2}\Big(d - 1 - \frac{d-1}{2}\Big) + \frac{d^2}{2} + \frac{3d}{2} \\
&= \frac{3d^2}{4} + d + \frac{1}{4} \\
&= \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} \Big\rfloor.
\end{aligned}
$$

Thus, since d is either odd or even, we have that

$$\sum_{v_i \in V(P)} e_T(v_i) = \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} \Big\rfloor. \tag{3.4}$$

Let $x \in V(T) - V(P)$. Since $diam(T) = d$, we have, $e_T(x) \leq d$ and so

$$\sum_{x \in V(T)-V(P)} e_T(x) \leq \sum_{x \in V(T)-V(P)} d.$$

Since T has order n, and the order of the set $V(P)$ is $d+1$, we have that,

$$|V(T) - V(P)| = n - (d+1) = n - d - 1.$$

It follows that if we sum eccentricities for verticies in $V(T) - V(P)$, we get,

$$\begin{aligned}\sum_{x \in V(T)-V(P)} e_T(x) &\leq \sum_{x \in V(T)-V(P)} d \\ &= |V(T) - V(P)|(d) \\ &= (n-d-1)d. \end{aligned} \tag{3.5}$$

Since $(n-d-1)d \in \mathbb{N}$ and $\left(\frac{3d^2}{4} + d + \frac{1}{4}\right) \in \mathbb{R}^+$, by (3.4), (3.5) and Note 3.2.3, we have that,

$$\begin{aligned}\zeta(T) &= \sum_{x \in V(T)} e_T(x) \\ &= \sum_{x \in V(P) \cup [V(T)-V(P)]} e_T(x) \\ &= \sum_{x \in V(P)} e_T(x) + \sum_{x \in V(T)-V(P)} e_T(x) \\ &\leq \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} \Big\rfloor + (n-d-1)d \\ &= \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} \Big\rfloor + \Big\lfloor (n-d-1)d \Big\rfloor \\ &= \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} + (n-d-1)d \Big\rfloor \\ &= \Big\lfloor \frac{3d^2}{4} + d + \frac{1}{4} + dn - d^2 - d \Big\rfloor \\ &= \Big\lfloor dn + \frac{3d^2}{4} - d^2 + d - d + \frac{1}{4} \Big\rfloor \\ &= \Big\lfloor dn - \frac{1}{4}d^2 + \frac{1}{4} \Big\rfloor, \end{aligned}$$

as desired. □

Lemma 3.2.5. *Consider a connected graph G with n vertices. Take T to be a connected dominating subgraph of G with vertex set S. We have that*

$$\zeta(G) \leq \zeta(T) + |S| + (|V(G)| - |S|)(diam(T) + 2).$$

Proof: We will consider the contribution made to $\zeta(G)$ by vertices in S and those in $V(G) - S$ seperately.

Claim 3.2.6. *Let $x \in S$. Then $e_G(x) \leq e_T(x) + 1$.*

Proof of Claim 3.2.6: Let $x \in S$. Then by the definition of eccentricity, $e_G(x) = \max_{y \in V(G)} \{d_G(x, y)\}$. On one hand, if $y \in S$, then since T is connected, there is an $x - y$ path in T and we have,

$$d_G(x, y) \leq d_T(x, y) < d_T(x, y) + 1.$$

So, we have that $e_G(x) < e_T(x) + 1$.
On the other hand, if y is not in S, then since T is a dominating subgraph of G, y is adjacent to some vertex y' on T. Hence

$$d_G(x, y) \leq d_G(x, y') + d_G(y, y') \leq d_T(x, y') + 1.$$

Therefore,

$$\begin{aligned} e_G(x) &= \max_{y \in V(G)} \{d_G(x, y)\} \\ &\leq \max_{y' \in V(G)} \{d_G(x, y') + d_G(y', y)\} \\ &= \max_{y' \in V(G)} \{d_G(x, y') + 1\} \\ &\leq \max_{y' \in V(T)} \{d_T(x, y') + 1\} \\ &\leq e_T(x) + 1, \end{aligned}$$

and the claim is proven. We now consider vertices in $V(G) - S$.

Claim 3.2.7. *Let $x \in V(G) - S$. Then $e_G(x) \leq diam(T) + 2$.*

Proof of Claim 3.2.7
Let $x \in V(G) - S$ and assume that $e_G(x) = d_G(x, y)$ for some $y \in V(G)$.
Since T is a connected dominating subgraph of G, let x' and y' be vertices on T that are adjacent to x and y, respectively. Then

$$\begin{aligned} e_G(x) &= d_G(x, y) \\ &\leq d_G(x, x') + d_G(x', y') + d_G(y', y) \\ &\leq 1 + d_T(x', y') + 1 \\ &\leq diam(T) + 2, \end{aligned}$$

as claimed. It follows from Claim 3.2.6 and Claim 3.2.7, that

$$\begin{aligned} \zeta(G) &= \sum_{x \in V(G)} e_G(x) \\ &= \sum_{x \in S} e_G(x) + \sum_{x \in V(G) - S} e_G(x) \\ &\leq \sum_{x \in S} [e_T(x) + 1] + \sum_{x \in V(G) - S} [diam(T) + 2] \\ &= \sum_{x \in S} e_T(x) + \sum_{x \in S} 1 + [diam(T) + 2] \sum_{x \in V(G) - S} 1 \end{aligned}$$

$$= \zeta(T) + |S| + (diam(T) + 2)(V(G) - |S|).$$

This completes the proof of the lemma. □

The following algorithm gives a way of constructing a sequence $A_1 \subset A_2 \subset A_3 \subset \ldots$ of independent sets, with any A_i, $i = 1, 2, 3, \ldots$, being contained within vertices of some subtree T of a connected graph G.

Algorithm 3.2.8. Let G be a connected graph of order n. We construct a sequence $A_1 \subset A_2 \subset A_3 \subset \ldots$ of independent sets, together with a sequence of subtrees of G, given by $T_1 \leq T_2 \leq T_3 \leq \ldots$, with $A_i \subseteq V(T_i)$ for $i = 1, 2, \ldots$. The forementioned sequences are constructed as follows: Take some vertex v of G, and let $A_1 = \{v\}$. Set $T_1 \leq G$ to be a tree with order one, with $V(T_1) = \{v\}$. Now, assume we have obtained A_{i-1}, as well as the tree T_{i-1}. If there exists a vertex x in G with $d_G(x, A_{i-1}) = 2$, let $A_i = A_{i-1} \cup \{x\}$.
To obtain T_i, let $xya, a \in A_{i-1}$, be an x-A_{i-1} shortest path in G and define $T_i = xya \cup T_{i-1}$.

Lemma 3.2.9. *Let G be a connected graph with subtrees $T_1 \leq T_2 \leq T_3 \ldots$ and let $A_1 \subset A_2 \subset A_3 \subset \ldots$ be a sequence of independent sets such that $A_i \subseteq V(T_i)$ for $i = 1, 2, 3, \ldots$ in G. Then $|V(T_i)| \leq |V(T_{i-1})| + 2$.*

Proof: Assume that $T_i = xya \cup T_{i-1}$, where $a \in A_{i-1}$ and $d_G(x, A_{i-1}) = 2$. Since $A_{i-1} \subseteq V(T_{i-1})$, we have $a \in V(T_{i-1})$.
It follows that

$$\begin{aligned} |V(T_i)| &= |V(xya \cup T_{i-1})| \\ &\leq |V(T_{i-1})| + |\{x, y\}| \\ &= |V(T_{i-1})| + 2, \end{aligned}$$

as desired. □

Remark 3.2.10. At some point, the options of having a vertex x in G such that $d_G(x, A_i) = 2$ for some i gets exhausted in Algorithm 3.2.8. So, for some set A_k, we have that A_k is maximally independent. It follows that every vertex in $V(G) - A_k$ is adjacent to some vertex in A_k. Let k be the smallest integer such that every vertex not in A_k is adjacent to some vertex in A_k. Then, $V(T_k)$ is a connected dominating set of G.

We illustrate the construction of T_k by way of an example.

Example 3.2.11. Suppose we have a graph G of order $n = 10$ as drawn below.

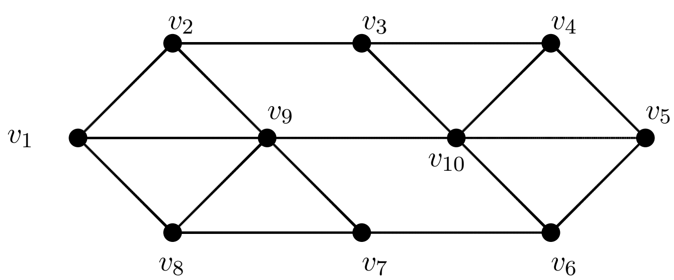


Figure 3.3: Graph G

- Firstly, we choose A_1 and T_1 in G. Let $A_1 = \{v_1\}$, then T_1 consists only of vertex v_1.
- Now, we find A_2 and T_2 in G. We have the set $A_{2-1} = A_1$ and the tree $T_{2-1} = T_1$. We see that in G, $\{v_3, v_7, v_{10}\}$ is the set of all vertices that are a distance 2 from A_1. So, from the set $\{v_3, v_7, v_{10}\}$, we let $x = v_7$, then we have that $A_2 = \{v_1, v_7\}$.
  There are two shortest v_7-A_1 paths, $v_1v_8v_7$ and $v_1v_9v_7$ in G. We choose $v_1v_9v_7$ to be the v_7-A_1 shortest path.
  So, we have that $T_2 = v_1v_9v_7 \cup T_1 = v_1v_9v_7 \cup v_1 = v_1v_9v_7$.

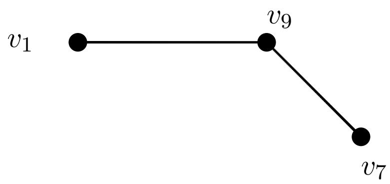


Figure 3.4: T_2

- We find A_3 and T_3 in G. We have the set $A_{3-1} = A_2$ and the tree $T_{3-1} = T_2$. We see that, in G, $\{v_3, v_5, v_{10}\}$ is the set of all vertices that are a distance 2 from the set A_2. So, from the set $\{v_3, v_5, v_{10}\}$, let $x = v_{10}$. Then we have that $A_3 = \{v_1, v_7, v_{10}\}$.
  There are two shortest v_{10}-A_2 paths, $v_7v_9v_{10}$ and $v_1v_9v_{10}$. We choose $v_7v_9v_{10}$ to be the v_7-A_2 shortest path.
  So, we have that $T_3 = v_1v_9v_{10} \cup T_2$.

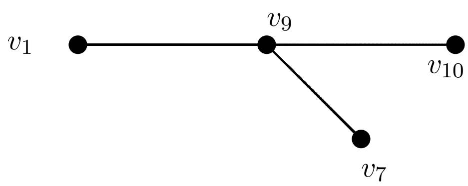


Figure 3.5: T_3

- We look to find the set A_4 and tree T_4 in G. We have the set $A_{4-1} = A_3$ and tree $T_{4-1} = T_3$. Observe that, there are no longer any vertices at distance 2 from A_3. All vertices not in A_3 are within distance 1 of A_3. Hence A_3 is maximal, and we see that T_3 is a connected dominating tree of G.

Proposition 3.2.12. *Let G be a connected graph of order n. For any integer k, the independence sets obtained by Algorithm 3.2.8 have order $|A_k| = k$.*

Proof By construction of A_k, we have, $A_k = A_{k-1} \cup \{x\}$ with x some vertex in G such that $d_G(x, A_{k-1}) = 2$. The proposition follows inductively from the fact that $|A_1| = 1$.

Lemma 3.2.13. *Let G be a connected graph of order n and connected domination number $\gamma_c(G)$. Let T_k be a dominating subtree of G obtained using Algorithm 3.2.8. Then $\gamma_c(G) \leq 2\alpha - 1$.*

Proof: Suppose the assumption is true. By Lemma 3.2.9, $|V(T_i)| \leq |V(T_{i-1})| + 2$ for $i = 2, 3, \ldots, k$.
Therefore,

$$\begin{aligned} |V(T_k)| &\leq |V(T_{k-1})| + 2 \\ &\leq (|V(T_{k-2})| + 2) + 2 \\ &= |V(T_{k-2})| + 2(2) \\ &\leq ((|V(T_{k-3})| + 2) + 2) + 2 \\ &= |V(T_{k-3})| + 2(3) \\ &\vdots \\ &\leq |V(T_3)| + 2(k-3) \\ &\leq |V(T_2)| + 2(k-2) \\ &\leq |V(T_1)| + 2(k-1) \\ &\leq 1 + 2(k-1) \\ &= 2k - 1. \end{aligned} \tag{3.6}$$

Since T_k is a dominating tree of G, we have that,

$$\gamma_c(G) \leq |V(T_k)| \leq 2k - 1. \tag{3.7}$$

By the construction in Algorithm 3.2.8., $d_G(x, y) \geq 2$ for any $x, y \in A_k$. Therefore, A_k is an independent set of G. So it follows that $|A_k| \leq \alpha(G)$, but $k = |A_k|$, so from (3.7), we have,

$$\gamma_c(G) \leq 2k - 1 \leq 2\alpha - 1.$$

This proves Lemma 3.2.13. □

Lemma 3.2.14. *Let G be a connected graph with independence number α. Let $T = v_1, v_2, \ldots, v_{2\alpha-1}$ be an induced dominating path of G of order $\gamma_c(G)$ and let T be one of the dominating trees of G with minimum diameter. There exist private neighbours u_1 of v_1 and $u_{2\alpha-1}$ of $v_{2\alpha-1}$ in T, u_1 and $u_{2\alpha-1}$ are adjacent in G.*

Proof: Suppose the assumption holds. Since T is a minimal dominating tree, the trees $T - v_1$ and $T - v_{2\alpha-1}$ are not dominating trees of G. Hence there exist private neighbours u_1 of v_1 and $u_{2\alpha-1}$ of $v_{2\alpha-1}$ in T.

We now show that u_1 and $u_{2\alpha-1}$ are adjacent in G.

Suppose that u_1 and $u_{2\alpha-1}$ are not adjacent in G.
Then since T is an induced path of G, with u_1 being a private neighbor of v_1 in T and G and $u_{2\alpha-1}$ being a private neighbor of $v_{2\alpha-1}$ in T and G, we have that $u_1v_1 \in E(G)$, $u_1v_1 \in E(T)$, $u_{2\alpha-1}v_{2\alpha-1} \in E(G)$ and $u_{2\alpha-1}v_{2\alpha-1} \in E(T)$.
So $P = u_1, v_1, v_2, \ldots, v_{2\alpha-1}, u_{2\alpha-1}$ is also an induced path of G, and so $\{u_1, v_2, v_4, \ldots, v_{2\alpha-2}, u_{2\alpha-1}\}$ is an independent set of order $\alpha + 1$.
But by the construction, we have that $A_k \subseteq V(T)$ is a maximally independent set that has at most order α.
So, we have a contradiction. Hence $u_1u_{2\alpha-1} \in E(G)$. □

Lemma 3.2.15. *Consider a connected graph G with n vertices and independence number 2. Then $diam(G) \leq 3$.*

Proof: Suppose G is a connected graph of order n with independence number $\alpha = 2$ and $diam(G) > 3$.
Let $\{x, y\} \subseteq V(G)$ be a set of vertices at distance $diam(G)$ in G.
Since $diam(G) > 3$, there is a vertex $u \in V(G)$ on an x-y path in G such that $d_G(u, \{x, y\}) \geq 2$.
Thus, the set $\{u, x, y\}$ is independent, contradicting the fact that $\alpha = 2$.
So we have that $diam(G) \leq 3$. □

Theorem 3.2.16. *Consider a connected graph G with n vertices and independence number α when $\alpha \leq \frac{n}{2}$. We have that*

$$\zeta(G) \leq (2\alpha - 1)n - \alpha^2 + \alpha,$$

and the bound is sharp.

Proof: Let G be a connected graph. We begin by constructing a dominating tree T_k of G:

Suppose a subtree $T_i \leq G$ and the set A_i have been obtained using Algorithm 3.2.8. We have, by Lemma 3.2.9, that

$$|V(T_i)| \leq |V(T_{i-1})| + 2.$$

Remark 3.2.10 gives that, there is a maximally independent set A_k of G obtained from Algorithm 3.2.8. Let k be the smallest integer such that every vertex not in A_k is adjacent to some vertex in A_k. Then, $V(T_k)$ is a connected dominating set of G.

Set T to be a dominating tree of G of order $\gamma_c(G)$ and let T be one of the dominating trees of G with minimum diameter. From Lemma 3.2.13, we have $\gamma_c(G) \leq 2\alpha - 1$. We go through three cases of $\gamma_c(G)$ and T to prove the remainder of the theorem, and the cases are,

1. $\gamma_c(G) \leq 2\alpha - 2$,
2. $\gamma_c(G) = 2\alpha - 1$, and T is not a path,
3. $\gamma_c(G) = 2\alpha - 1$, and T is a path.

Case 1: $\gamma_c(G) \leq 2\alpha - 2$.
We know that, the diameter of any graph G is maximised if G is a path, and so, we have,

$$diam(T) \leq |V(T)| - 1 = \gamma_c(G) - 1. \tag{3.8}$$

Since T is a dominating tree of G, we have that, (3.8) in conjunction with Lemma 3.2.5, gives that,

$$\begin{aligned}
\zeta(G) &\leq \zeta(T) + |V(T)| + (|V(G)| - |V(T)|)[diam(T) + 2] \\
&\leq \zeta(T) + \gamma_c(G) + (n - \gamma_c(G))[(\gamma_c(G) - 1) + 2] \\
&= \zeta(T) + \gamma_c(G) + n(\gamma_c(G) - 1) + 2n - \gamma_c(G)(\gamma_c(G) - 1) - 2\gamma_c(G) \\
&= \zeta(T) + \gamma_c(G) + \gamma_c(G)n - n + 2n - \gamma_c(G)^2 + \gamma_c(G) - 2\gamma_c(G) \\
&= \zeta(T) + \gamma_c(G)n + n - \gamma_c(G)^2.
\end{aligned} \tag{3.9}$$

Using Proposition 3.2.4, we get that,

$$\zeta(T) \leq \left\lfloor diam(T)\gamma_c(G) - \frac{1}{4}diam(T)^2 + \frac{1}{4} \right\rfloor.$$

The function $t(diam(T)) := diam(T)\gamma_c(G) - \frac{1}{4}diam(T)^2 + \frac{1}{4}$ has a local maxima at $diam(T) = 2\gamma_c(G)$, and it is increasing in the interval $0 \leq diam(T) < 2\gamma_c(G)$.
Since by (3.8), $diam(T) \leq \gamma_c(G) - 1 < 2\gamma_c(G)$, we have that t is increasing in the interval $0 \leq diam(T) \leq \gamma_c(G) - 1$ and it is maximised for $diam(T) = \gamma_c(G) - 1$, giving

$$\begin{aligned}
\zeta(T) &\leq \left\lfloor t(\gamma_c(G) - 1) \right\rfloor \\
&= \left\lfloor (\gamma_c(G) - 1)\gamma_c(G) - \frac{1}{4}(\gamma_c(G) - 1)^2 + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \gamma_c(G)^2 - \gamma_c(G) - \frac{1}{4}(\gamma_c(G)^2 - 2\gamma_c(G) + 1) + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \gamma_c(G)^2 - \gamma_c(G) - \frac{1}{4}\gamma_c(G)^2 + \frac{1}{2}\gamma_c(G) - \frac{1}{4} + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \frac{3}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) \right\rfloor \\
&\leq \frac{3}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G).
\end{aligned}$$

We use the above inequality to simplify (3.9) further to get,

$$\begin{aligned}
\zeta(G) &\leq \frac{3}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) + (\gamma_c(G) + 1)n - \gamma_c(G)^2 \\
&= \frac{3}{4}\gamma_c(G)^2 - \gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) + (\gamma_c(G) + 1)n \\
&= -\frac{1}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) + (\gamma_c(G) + 1)n.
\end{aligned}$$

A simple differentiation shows that $f(\gamma_c(G)) = -\frac{1}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) + (\gamma_c(G) + 1)n$ is increasing in $\gamma_c(G)$. Subject to $\gamma_c(G) \leq 2\alpha - 2$, f is maximised for $\gamma_c(G) = 2\alpha - 2$ and we obtain

$$\begin{aligned}
\zeta(G) &\leq (\gamma_c(G) + 1)n - \frac{1}{4}\gamma_c(G)^2 - \frac{1}{2}\gamma_c(G) \\
&\leq -\frac{1}{4}(2(\alpha - 1))^2 - \frac{1}{2}(2(\alpha - 1)) + (2\alpha - 2 + 1)n \\
&= -(\alpha^2 - 2\alpha + 1) - \alpha + 1 + (2\alpha - 1)n \\
&= (2\alpha - 1)n - \alpha^2 + 2\alpha - 1 - \alpha + 1 \\
&= (2\alpha - 1)n - \alpha^2 + \alpha,
\end{aligned}$$

as desired.

Case 2: We consider $\gamma_c(G) = 2\alpha - 1$ with T not a path. Let $diam(T) = d$. We have established that T has order $2\alpha - 1$ and T is not a path. It follows that

$$d \leq |V(T)| - 2 = (2\alpha - 1) - 2 = 2\alpha - 3. \tag{3.10}$$

By Proposition 3.2.4, we have

$$\begin{aligned}
\zeta(T) &\leq \left\lfloor diam(T)\gamma_c(G) - \frac{1}{4}diam(T)^2 + \frac{1}{4} \right\rfloor \\
&= \left\lfloor d(2\alpha - 1) - \frac{1}{4}d^2 + \frac{1}{4} \right\rfloor
\end{aligned}$$

$$\leq d(2\alpha - 1) - \frac{1}{4}d^2 + \frac{1}{4}.$$

This, together with Lemma 3.2.5, gives

$$\begin{aligned}
\zeta(G) &\leq \zeta(T) + |V(T)| + (|V(G)| - |V(T)|)[diam(T) + 2] \\
&= \zeta(T) + \gamma_c(G) + (n - \gamma_c(G))(d + 2) \\
&= \zeta(T) + (2\alpha - 1) + (n - (2\alpha - 1))(d + 2) \\
&\leq d(2\alpha - 1) - \frac{1}{4}d^2 + \frac{1}{4} + (2\alpha - 1) + (n - (2\alpha - 1))(d + 2) \\
&= 2d\alpha - d - \frac{1}{4}d^2 + \frac{1}{4} + 2\alpha - 1 + dn + 2n - 2d\alpha + d - 4\alpha + 2 \\
&= 2d\alpha - 2d\alpha - d + d - \frac{1}{4}d^2 + dn + 2n - 4\alpha + 2\alpha + 2 - 1 + \frac{1}{4} \\
&= -\frac{1}{4}d^2 + dn + 2n - 2\alpha + \frac{5}{4}.
\end{aligned}$$

For a real-valued function $f : \mathbb{R} \to \mathbb{R}$, given by $f(d) := -\frac{1}{4}d^2 + dn + 2n - 2\alpha + \frac{5}{4}$, we have that f has a local maximum at $d = 2n$, and it is increasing in the interval where $0 \leq d < 2n$.
Since $\alpha \leq \frac{n}{2}$, we have that $2\alpha \leq n$, giving that $2\alpha - 3 \leq n - 3 < 2n$. From (3.10), we have that $0 \leq d \leq 2\alpha - 3$. But f is increasing in the interval $0 \leq d \leq 2\alpha - 3$ and it is maximised for $d = 2\alpha - 3$, giving

$$\begin{aligned}
\zeta(G) &\leq f(2\alpha - 3) \\
&= -\frac{1}{4}(2\alpha - 3)^2 + (2\alpha - 3)n + 2n - 2\alpha + \frac{5}{4} \\
&= -\frac{1}{4}((2\alpha)^2 - 12\alpha + 9) + 2\alpha n - 3n + 2n - 2\alpha + \frac{5}{4} \\
&= -\alpha^2 + 3\alpha + 2\alpha n - n - 2\alpha + \frac{5}{4} - \frac{9}{4} \\
&= -\alpha^2 + \alpha + (2\alpha - 1)n - 1 \\
&< (2\alpha - 1)n - \alpha^2 + \alpha,
\end{aligned}$$

as desired.

Case 3: $\gamma_c(G) = 2\alpha - 1$ and T is a path.
Let $T = v_1, v_2, \ldots, v_{2\alpha-1}$. We first show that the bound in the theorem holds if T is not an induced path of G. If $v_1v_{2\alpha-1} \in E(G)$, then the cycle $C = v_1, v_2, \ldots, v_{2\alpha-1}, v_1$ has diameter at most $\alpha - 1$ and it dominates G. It follows, by Lemma 3.2.5, that

$$\zeta(G) \leq \zeta(C) + |V(C)| + (|V(G)| - |V(C)|)[diam(C) + 2].$$

Since every vertex x on cycle C of order $2\alpha - 1$ has eccentricity

$$e_C(x) = \left\lfloor \frac{2\alpha - 1}{2} \right\rfloor = \alpha - \left\lceil \frac{1}{2} \right\rceil = \alpha - 1,$$

we have

$$\begin{aligned}
\zeta(G) &\leq \zeta(C) + |V(C)| + (|V(G)| - |V(C)|)[diam(C) + 2] \\
&\leq \sum_{x \in V(C)} e_C(x) + (2\alpha - 1) + (n - (2\alpha - 1))(\alpha - 1 + 2)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{2\alpha-1} (\alpha-1) + (2\alpha-1) + (n-(2\alpha-1))(\alpha+1) \\
&= (2\alpha-1)(\alpha-1) + (2\alpha-1) + n(\alpha+1) - (2\alpha-1)(\alpha+1) \\
&= (2\alpha-1)\Big(\alpha-1+1-(\alpha+1)\Big) + n(\alpha+1) \\
&= n(\alpha+1) - (2\alpha-1) \\
&= n(\alpha+1) - 2\alpha+1.
\end{aligned}$$

Now, we relate $n(\alpha+1)-2\alpha+1$ to $(2\alpha-1)n-\alpha^2+\alpha$, we see that

$$\begin{aligned}
(2\alpha-1)n-\alpha^2+\alpha-\Big((\alpha+1)n-2\alpha+1\Big) &= (2\alpha-1-\alpha-1)n-\alpha^2+3\alpha-1 \\
&= (\alpha-2)n-\alpha^2+3\alpha-1.
\end{aligned}$$

But, since G has cycle C in it, we have that $\alpha \geq 2$. Let $q(\alpha) = (\alpha-2)n-\alpha^2+3\alpha-1$. Then, we see that q is increasing in the interval $2 \leq \alpha \leq \frac{n}{2}$. So, we have that

$$q(\alpha) \geq q(2) = (2-2)n - 2^2 + 3(2) - 1 = 1.$$

Giving that,

$$\begin{aligned}
(2\alpha-1)n-\alpha^2+\alpha-\Big((\alpha+1)n-2\alpha+1\Big) &= (\alpha-2)n-\alpha^2+3\alpha-1 \\
&\geq 1 \\
&> 0.
\end{aligned}$$

Thus, we obtain that

$$(2\alpha-1)n-\alpha^2+\alpha > n(\alpha+1)-2\alpha+1,$$

giving that

$$\zeta(G) \leq n(\alpha+1)-2\alpha+1 < (2\alpha-1)n-\alpha^2+\alpha.$$

So, we have that, the theorem holds when T is not an induced path of G, and $v_1v_{2\alpha-1} \in E(G)$. Now assume that for some $i, j \in \{1, 2, \ldots, 2\alpha-1\}$, where $j-i \geq 2$ and $v_iv_j \neq v_1v_{2\alpha-1}$, we have $v_iv_j \in E(G)$.

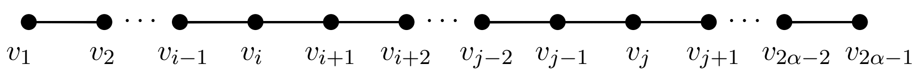


Figure 3.6: Graph of path T.

Since T is a dominating tree of G, we have that the tree T' obtained from T by removing the edge v_iv_{i+1} or v_jv_{j-1} and adding the edge v_iv_j is a dominating tree of G of order $2\alpha-1$ that is not a path.

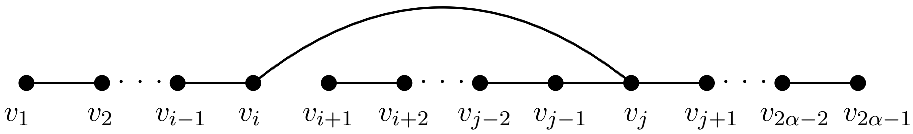


Figure 3.7: Graph T' with the edge v_iv_{i+1} removed.

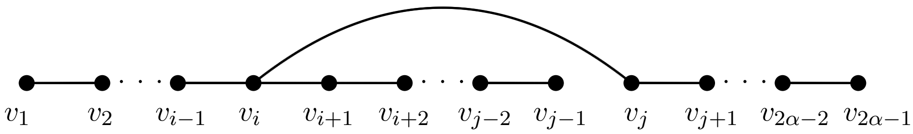


Figure 3.8: Graph T' with the edge $v_{j-1}v_j$ removed.

By Lemma 3.2.1, we have,

$$\zeta(G) \leq \zeta(T') + (2\alpha - 1) + (n - (2\alpha - 1))(diam(T') + 2).$$

Since T' is a tree that is not a path, we have by Case 2 that

$$\zeta(G) < (2\alpha - 1)n - \alpha^2 + \alpha.$$

So, we have that, the theorem holds when T is not an induced path of G.

Now, assume that $T = v_1, v_2, \ldots, v_{2\alpha-1}$ is an induced path of G. Since T is a minimal dominating tree, the trees $T - v_1$ and $T - v_{2\alpha-1}$ are not dominating trees of G. Hence there exist private neighbours u_1 of v_1 and $u_{2\alpha-1}$ of $v_{2\alpha-1}$ in T.

It follows from by Lemma 3.2.14 that $V(T) \cup \{u_1, u_{2\alpha-1}\}$ induces a dominating cycle C_T of length $2\alpha + 1$.
Since $diam(C_T) = \alpha$, we conclude by Lemma 3.2.5 that

$$\begin{aligned}
\zeta(G) &\leq \zeta(C_T) + |V(C_T)| + (|V(G)| - |V(C_T)|)[diam(C_T) + 2] \\
&= \sum_{x \in V(C_T)} e_{C_T}(x) + (2\alpha + 1) + (n - (2\alpha + 1))(\alpha + 2) \\
&\leq (2\alpha + 1)\alpha + (2\alpha + 1) + (n - 2\alpha - 1)(\alpha + 2) \\
&= (2\alpha + 1)(\alpha + 1) + (\alpha + 2)n - (2\alpha + 1)(\alpha + 2) \\
&= (2\alpha + 1 - 2\alpha - 1)(\alpha + 1) + (\alpha + 2)n - (2\alpha + 1) \\
&= (\alpha + 2)n - 2\alpha - 1.
\end{aligned}$$

If $\alpha \geq 3$, then we have that

$$\begin{aligned}
(2\alpha - 1)n - \alpha^2 + \alpha - \Big((\alpha + 2)n - 2\alpha - 1\Big) &= (2\alpha - 1 - \alpha - 2)n - \alpha^2 + 3\alpha + 1 \\
&= (\alpha - 3)n - \alpha^2 + 3\alpha + 1.
\end{aligned}$$

By the assumption, we have that $3 \leq \alpha \leq \frac{n}{2}$. Let $s(\alpha) := (\alpha - 3)n - \alpha^2 + 3\alpha + 1$. Then s is increasing in the interval $3 \leq \alpha \leq \frac{n}{2}$, giving that

$$s(\alpha) \geq s(3) = (3 - 3)n - 3^2 + 3(3) + 1 = 1.$$

Thus, we have that

$$(2\alpha - 1)n - \alpha^2 + \alpha > (\alpha + 2)n - 2\alpha - 1,$$

giving that

$$\zeta(G) < (2\alpha - 1)n - \alpha^2 + \alpha$$

as desired.

We now show that the theorem holds for the case when $\alpha = 2$.

Let $\alpha = 2$. Lemma 3.2.15 gives that $diam(G) \leq 3$. We consider cases when $diam(G) \leq 2$ and $diam(G) = 3$.

If $diam(G) \leq 2$, then for every vertex $x \in V(G)$, $e_G(x) \leq 2$ giving that

$$\zeta(G) = \sum_{x \in V(G)} e_G(x) \leq \sum_{x \in V(G)} 2 = 2n.$$

But since $\alpha = 2$, we have that

$$\begin{aligned}(2\alpha - 1)n - \alpha^2 + \alpha &= (2(2) - 1)n - 2^2 + 2 \\ &= 3n - 2. \end{aligned} \tag{3.11}$$

Thus, we have that

$$\zeta(G) \leq 2n < 3n - 2 = (2\alpha - 1)n - \alpha^2 + \alpha.$$

Now assume that $diam(G) = 3 = d_G(v_0, v_3)$, say. Let v_0, v_1, v_2, v_3 be a shortest path between these vertices.
Since $\alpha = 2$, we have that no three vertices of G are independent. Thus we have that

$$N[v_0] \cup N[v_3] = V(G),$$

and for all $x, y \in N[v_0]$ $(N[v_3])$, $d_G(x, y) = 1$, giving that $N[v_0]$ $(N[v_3])$ induces a complete graph. Also, we have that $v_1 \in N[v_0]$ and $v_2 \in N[v_3]$, and the two cliques are joined by the edge v_1v_2, giving that $d_G(v_1, x) \leq 2$ and $d_G(v_2) \leq 2$ for all $x \in V(G)$ and so $e_G(v_1) = 2$ and $e_G(v_2) = 2$. The remaining vertices have eccentricity at most 3, so we have that

$$\begin{aligned}\zeta(G) &= \sum_{x \in V(G)} e_G(x) \\ &= e_G(v_1) + e_G(v_2) + \sum_{x \in V(G) - \{v_1, v_2\}} e_G(x) \\ &\leq 2 + 2 + 3(n - 2) \\ &= 3n - 2.\end{aligned}$$

Since $\alpha = 2$ in this case, we have by (3.11) that

$$\zeta(G) \leq 3n - 2 = (2\alpha - 1)n - \alpha^2 + \alpha.$$

To show that the bound is sharp, consider the graph $G_{n,\alpha}$ obtained by taking a path $P_{2\alpha-2}$ of order $2\alpha - 2$, with end vertices u and v, joining u to every vertex of a complete graph K_a, and joining v to every vertex of a complete graph K_c, where

$$a + c = n - 2\alpha + 2, \; a, c \geq 1. \tag{3.12}$$

We see that, $G_{n,\alpha}$ has order

$$2\alpha - 2 + a + c = 2\alpha - 2 + n - 2\alpha + 2 = n.$$

Also, since u is joined to every vertex of K_a, we have that, the graph with the vertex set $V(K_a) \cup \{u\}$ is a complete graph of order $a+1$ in $G_{n,\alpha}$. Similarly, we have that $V(K_c) \cup \{v\}$ is a complete graph of order $c+1$ in $G_{n,\alpha}$. The remaining $2\alpha - 4$ vertices of the path $P_{2\alpha-2}$ are not adjacent to any of the K_a or K_c vertices. So, the independence number of $G_{n,\alpha}$ is

$$1 + 1 + \frac{2\alpha - 4}{2} = 2 + \alpha - 2 = \alpha.$$

We calculate $\zeta(G_{n,\alpha})$.
Let $P_{2\alpha} = v_0, v_1, \ldots, v_{2\alpha-1}$ be a diametral path in $G_{n,\alpha}$. Then by Lemma 3.2.1, we have that $e_{G_{n,\alpha}}(v_i) = (2\alpha - 1) - i$ for $i \leq \lfloor \frac{2\alpha-1}{2} \rfloor$ and $e_{G_{n,\alpha}}(v_i) = i$ for $i > \lfloor \frac{2\alpha-1}{2} \rfloor$.
We have

$$\begin{aligned}
\zeta(P_{2\alpha}) &= \sum_{i=0}^{\lfloor \frac{2\alpha-1}{2} \rfloor} \Big((2\alpha - 1) - i\Big) + \sum_{i=\lfloor \frac{2\alpha-1}{2} \rfloor + 1}^{2\alpha-1} i \\
&= \sum_{i=0}^{\alpha-1} \Big((2\alpha - 1) - i\Big) + \sum_{i=\alpha-1+1}^{2\alpha-1} i \\
&= \sum_{i=0}^{\alpha-1} \Big((2\alpha - 1) - i\Big) + \sum_{i=0}^{2\alpha-1} i - \sum_{i=0}^{\alpha-1} i \\
&= \sum_{i=0}^{\alpha-1} (2\alpha - 1) - \sum_{i=0}^{\alpha-1} i + \sum_{i=0}^{2\alpha-1} i - \sum_{i=0}^{\alpha-1} i \\
&= (2\alpha - 1)(\alpha - 1 + 1) - \frac{(\alpha - 1)\alpha}{2} + \frac{(2\alpha - 1)2\alpha}{2} - \frac{(\alpha - 1)\alpha}{2} \\
&= (2\alpha - 1)\alpha - (\alpha - 1)\alpha + (2\alpha - 1)\alpha \\
&= \alpha(2\alpha - 1 - \alpha + 1 + 2\alpha - 1) \\
&= \alpha(3\alpha - 1).
\end{aligned}$$

$P_{2\alpha}$ in $G_{n,\alpha}$ has length $2\alpha - 3 + 1 + 1 = 2\alpha - 1$ and the $2\alpha - 2$ of the vertices are those that make up the subgraph $P_{2\alpha-2}$ in $G_{n,\alpha}$ and the other two vertices are in the vertices of the K_a and K_c cliques of $G_{n,\alpha}$.

We note that

$$\begin{aligned}
\zeta(G_{n,\alpha}) = &\sum_{x \in V(G_{n,\alpha}) - (V(K_a) \cup V(K_c))} e_{G_{n,\alpha}}(x) + \sum_{x \in V(G_{n,\alpha}) - (V(P_{2\alpha-2}) \cup V(K_a))} e_{G_{n,\alpha}}(x) \\
&+ \sum_{x \in V(G_{n,\alpha}) - (V(P_{2\alpha-2}) \cup (V(K_b))} e_{G_{n,\alpha}}(x).
\end{aligned}$$

We see that

- $V(G_{n,\alpha}) - (V(K_a) \cup V(K_c))$ is the set of vertices for the path $P_{2\alpha-2}$ in $G_{n,\alpha}$.
- $V(G_{n,\alpha}) - (V(P_{2\alpha-2}) \cup V(K_a))$ is the set of vertices for the clique K_c in $G_{n,\alpha}$.
- $V(G_{n,\alpha}) - (V(P_{2\alpha-2}) \cup V(K_c))$ is the set of vertices for the clique K_a in $G_{n,\alpha}$.

We have,

$$\sum_{x\in V(G_{n,\alpha})-(V(K_a)\cup V(K_c))} e_{G_{n,\alpha}}(x) = \zeta(P_{2\alpha}) - 2(2\alpha-1)$$
$$= \alpha(3\alpha-1) - 2(2\alpha-1).$$

$$\sum_{x\in V(G_{n,\alpha})-(V(P_{2\alpha-2})\cup V(K_c))} e_{G_{n,\alpha}}(x) = a(2\alpha-1),$$

and

$$\sum_{x\in V(G_{n,\alpha})-(V(P_{2\alpha-2})\cup V(K_a))} e_{G_{n,\alpha}}(x) = c(2\alpha-1).$$

Hence,

$$\begin{aligned}\zeta(G_{n,\alpha}) &= \Big(\alpha(3\alpha-1) - 2(2\alpha-1)\Big) + a(2\alpha-1) + c(2\alpha-1).\\ &= \alpha(3\alpha-1) - 2(2\alpha-1) + (a+c)(2\alpha-1).\end{aligned}$$

By (3.12), we have that

$$\begin{aligned}\zeta(G_{n,\alpha}) &= \alpha(3\alpha-1) - 2(2\alpha-1) + (n-2\alpha+2)(2\alpha-1)\\ &= \alpha(3\alpha-1) - 2(2\alpha-1) + (2\alpha-1)n - 2\alpha(2\alpha-1) + 2(2\alpha-1)\\ &= \alpha(3\alpha-1-4\alpha+2) + (2\alpha-1)n\\ &= \alpha(1-\alpha) + (2\alpha-1)n\\ &= (2\alpha-1)n - \alpha^2 + \alpha.\end{aligned}$$

□

To analyse Theorem 3.2.19, we need the following definition:

Definition 3.2.17. (Eccentric sequence)
If G is a connected graph with vertex set $\{v_1, v_2, \ldots, v_n\}$, and if v_i has eccentricity e_i, then the sequence $e_1, e_2, \ldots, e_n$ is the eccentric sequence of G.

We assume the eccentric sequence of G is ordered as a nondecreasing sequence, so e_1 is the radius of G, and e_n is the diameter of G.

Lemma 3.2.18. *Let G be a connected graph. The eccentricities of adjacent vertices in G differ by not more than 1.*

Proof: Let u and v be two adjacent vertices in G. Assume that $e_G(v) = d_G(v,x)$ and $e_G(u) = d_G(u,y)$ for some $x, y \in V(G)$. Assume, without loss of generality, that $e_G(v) \geq e_G(u)$. It suffices to prove that $e_G(v) - e_G(u) \leq 1$.
Note that by the triangle inequality we have that

$$d_G(v,x) \leq d_G(v,u) + d_G(u,x) \tag{3.13}$$

and also

$$d_G(u,x) \leq e_G(u), \tag{3.14}$$

so we have,

$$e_G(v) - e_G(u) = d_G(v,x) - e_G(u)$$

$$\begin{aligned} &\leq d_G(u,v) + d_G(u,x) - e_G(u) && \text{(by 3.13)} \\ &\leq d_G(u,v) + e_G(u) - e_G(u) && \text{(by 3.14)} \\ &= 1, \end{aligned}$$

and Lemma 3.2.18 is proven. □

Theorem 3.2.19. *Consider a connected graph G of order n. We have that, for any integer k with $rad(G) < k < diam(G)$ there exist at least two vertices in G of eccentricity k.*

Proof. Let G be a graph with eccentric sequence $e_1, e_2, \ldots, e_n$. Since we assume that the eccentric sequence $e_1, e_2, \ldots, e_n$ is in a nondecreasing order, we have by Lemma 3.2.18 that

$$e_{i-1} \leq e_i \leq e_{i-1} + 1 \text{ for } i = 2, 3, \ldots, n.$$

So we have that a path from a vertex of eccentricity e_1 to a vertex of eccentricity e_n contains vertices of each eccentricity between e_1 to e_n. It follows that, there exists a vertex u of eccentricity

$$e_i = e_G(u)$$

such that

$$e_1 < e_i \leq e_n \text{ for some } i = 2, 3, \ldots, n-1.$$

Let v be a vertex at distance e_i from u. Then

$$e_G(v) \geq e_i.$$

Let w be a central vertex, so

$$e_G(w) = e_1 < e_i.$$

So we have that

$$e_G(w) < e_i \leq e_G(v).$$

By Lemma 3.2.18, we have that there is a vertex v' of eccentricity e_i in the v-w path of length $d_G(v,w)$. We show that $u \neq v'$. Since by the assumption we have that $d_G(u,v) = e_i$, and since v' is in the v-w path of length $d_G(v,w)$, and since $e_G(w) = e_1$ we have that $d_G(v,w) \leq e_1$, so we have $d_G(v,v') \leq d_G(v,w) \leq e_1 < e_i$. Hence, $d_G(u,v) = e_i > e_1 \geq d_G(v',v)$. It follows that $u \neq v'$.
Thus, we have proved that the theorem holds. □

Lemma 3.2.20. *Let G be a connected graph of order n and independence number at least α and, among those, a graph of maximum average eccentricity and having smallest size. Then G is a tree.*

Proof: Suppose to the contrary, that G is not a tree. Then G has at least one cycle in it. Say cycle C_G. Let u and v be two adjacent vertices in C_G such that, the graph $G - uv$ is a connected spanning graph of G and we have that

$$\sum_{x \in V(G-uv)} e_{G-uv}(x) \geq \sum_{x \in V(G)} e_G(x), \text{ and } \alpha(G-uv) \geq \alpha(G).$$

So, if G is not a tree, we see that, we can find a connected graph $G - uv$ of less edges than G contradicting our choice of G. So, we must have that G is a tree. □

Lemma 3.2.21. *Let T be a tree. Then $2rad(T) - 1 \leq diam(T) \leq 2rad(T)$.*

Proof: Suppose T has diameter d and let $P_d = v_0, v_1, v_2, \ldots, v_d$ be a diametral path in T. By Lemma 3.2.1, we have that, for any vertex $x \in V(T) - V(P_d)$,

$$e_T(x) = \max\{d_T(v_0, x), d_T(x, v_d)\}.$$

Let x be any vertex in T and not on P_d. Assume without loss of generality that

$$e_T(x) = d_T(v_0, x).$$

Since T is a tree, we have that, there is only one v_0-x path in T. So, for some $v_i \in V(P_d)$, we have that

$$d_T(v_0, x) = d_T(v_0, v_i) + d_T(v_i, x) > d_T(v_0, v_i).$$

Also, since

$$e_T(x) = d_T(v_0, x),$$

we have that

$$d_T(v_0, x) \geq d_T(x, v_d).$$

So, since there is only one x-v_d path in T, we have that,

$$d_T(x, v_d) = d_T(x, v_i) + d_T(v_i, v_d) > d_T(v_i, v_d).$$

Also, since

$$d_T(v_0, x) \geq d_T(x, v_d),$$

we have that,

$$\begin{aligned} 0 &\geq d_T(x, v_d) - d_T(v_0, x) \\ &= d_T(x, v_i) + d_T(v_i, v_d) - [d_T(v_0, v_i) + d_T(v_i, x)] \\ &= d_T(v_i, v_d) - d_T(v_0, v_i), \end{aligned}$$

giving that

$$d_T(v_0, v_i) \geq d_T(v_i, v_d),$$

by Lemma 3.2.1, we have that

$$e_T(v_i) = d_T(v_0, v_i).$$

So, we have that, for every $x \in V(T) - V(P_d)$, $e_T(x) > e_T(v_i)$ for some $v_i \in V(P_d)$.
So, a vertex that has the smallest eccentricity in T will be in the path P_d. But we know that for any path Q we have that $rad(Q) = \lfloor \frac{|V(Q)|}{2} \rfloor$.

So we must have that $rad(T) = rad(P_d) = \left\lfloor \frac{|V(P_d)|}{2} \right\rfloor = \begin{cases} \frac{|V(P_d)|}{2} & \text{if } |V(P_d)| \text{ is even} \\ \frac{|V(P_d)|-1}{2} & \text{if } |V(P_d)| \text{ is odd} \end{cases}$

But, since P_d is a path, we have that $diam(P_d) = |V(P_d)| - 1$, which gives that $|V(P_d)| = diam(P_d) + 1$.
So, for $|V(P_d)|$ even, we have

$$rad(T) = rad(P_d) = \frac{|V(P_d)|}{2} = \frac{diam(P_d) + 1}{2} = \frac{diam(T) + 1}{2}.$$

Giving that $2rad(T) = diam(T) + 1$, thus $2rad(T) - 1 = diam(T)$.

For $|V(P_d)|$ odd, we have

$$rad(T) = rad(P_d) = \frac{|V(P_d)| - 1}{2} = \frac{diam(P_d) + 1 - 1}{2} = \frac{diam(P_d)}{2} = \frac{diam(T)}{2}. \tag{3.15}$$

Giving that $2rad(T) = diam(T)$.
So, we have that $2rad(T) - 1 \leq diam(T) \leq 2rad(T)$. Thus, Lemma 3.2.21 is proven. □

Lemma 3.2.22. *Let G be a tree of order n and independence number α. We have that*

$$rad(G) \leq n - \alpha.$$

Proof:
Consider a shortest path P between two diametral vertices u and v. Let S be a maximum independent set of G. Any two adjacent vertices cannot both be in S. The path P guarantees that consecutive vertices on it are adjacent. So, any two consecutive vertices on P cannot both be in S. Let $V(P_s)$ be an independent set of G such that $V(P_s) \subset V(P)$. Since P is a path, we have the following.

- If $|V(P)|$ is odd, then we have that at most, $|V(P_s)| = \frac{|V(P)|+1}{2}$.
- If $|V(P)|$ is even, then we have that at most, $|V(P_s)| = \frac{|V(P)|}{2}$.

We now study the order of $V(P) - S$.

Case 1 If $V(P_s) \subseteq S$, then by the first bullet above, we have that, at most, $\frac{|V(P)|+1}{2}$ vertices of P will be in S. So,

$$|V(P) \cap S| \leq \frac{|V(P)|+1}{2},$$

and we have,

$$\begin{aligned} |V(P) - S| &= |V(P)| - |V(P) \cap S| \\ &\geq |V(P)| - \frac{|V(P)|+1}{2} \\ &= \frac{|V(P)|-1}{2}. \end{aligned}$$

Case 2 If $V(P_s) \nsubseteq S$, then there is at least one vertex $x \in V(P_s)$ such that $x \notin S$.
Since at most, $|V(P_s)| = \frac{|V(P)|+1}{2}$, we have that at most

$$|V(P) \cap S| = \frac{|V(P)|+1}{2} - 1 < \frac{|V(P)|+1}{2}.$$

So, we have that,

$$\begin{aligned} |V(P) - S| &= |V(P)| - |V(P) \cap S| \\ &> |V(P)| - \frac{|V(P)|+1}{2} \\ &= \frac{|V(P)|-1}{2}. \end{aligned}$$

So, Case 1 and Case 2 gives that

$$|V(P) - S| \geq \frac{|V(P)|-1}{2}.$$

Since P is a path, we have that

$$diam(G) = d_G(u,v) = d_P(u,v) = diam(P) = |V(P)| - 1,$$

We have that,

$$V(G) - S = [V(P) - S] \cup [V(G) - (V(P) \cup S)],$$

giving,

$$|V(G) - S| = |V(P) - S| + |V(G) - V(P) \cup S| \geq |V(P) - S| \geq \frac{|V(P)| - 1}{2} = \frac{diam(G)}{2}.$$

But since $|V(G)| = n$, $|S| = \alpha$ and $|V(G) \cap S| = \alpha$, we have that $|V(G) - S| = n - \alpha$. Thus, we have that

$$\frac{diam(G)}{2} \leq |V(G) - S| = n - \alpha. \tag{3.16}$$

But, since G is a tree, and the bound in (3.16) is obtained from path P of odd $|V(P)|$, we have by (3.15) that $\frac{diam(G)}{2} = rad(G)$, and so, we have that

$$rad(G) = \frac{diam(G)}{2} \leq |V(G) - S| = n - \alpha. \tag{3.17}$$

as desired. □

Theorem 3.2.23. *Consider a connected graph G with n vertices and independence number α, where $\alpha > \frac{n}{2}$. We have that*

$$\zeta(G) \leq n^2 - \alpha^2,$$

and the bound is sharp.

Proof:
Since $\zeta(G) \leq n^2 - \alpha^2$, we have that the bound is decreasing as α increases. So we show that $\zeta(G) \leq n^2 - \alpha^2$ holds for all connected graphs of order n with independence number at least α (this will be sufficient to show that the theorem holds). Let G be a connected graph of order n and independence number at least α and, among those, a graph of maximum average eccentricity and having smallest size. By Lemma 3.2.20 , we have that G is a tree.

Consider the eccentric sequence $e_1, e_2, \ldots, e_n$ of G. By (3.16) and (3.17), we have that

$$rad(G) = e_1 \leq n - \alpha$$

and

$$diam(G) \leq 2(n - \alpha),$$

and Theorem 3.2.19 gives that, there are at least two vertices with eccentricity k such that $e_1 < k < e_n$ appearing in the eccentric sequence of G. Subject to the above mentioned conditions, $\sum_{i=1}^{n} e_i$ is maximised for the sequence

$$(n-\alpha), (n-\alpha+1)^{(2)}, (n-\alpha+2)^{(2)}, (n-\alpha+3)^{(2)}, \ldots, (2n-2\alpha-1)^{(2)}, (2n-2\alpha)^{(2\alpha-n+1)} \tag{3.18}$$

Note 3.2.24. Note that, in the above sequence, the superscript on the eccentricities indicates the number of vertices in G that have that eccentricity.

The sequence in (3.18) has $n - \alpha$ terms, and there are two vertices of eccentricity $n - \alpha + i$ for $i = 1, 2, 3, \ldots, n - \alpha - 1$ in G and one vertex of eccentricity $n - \alpha$. Since in total, G has n vertices, we have that $n - 2(n - \alpha - 1) - 1 = 2\alpha - n + 1$ vertices have not been accounted for, thus, to have $\sum_{i=1}^{n} e_i$ maximised, we will need the $2\alpha - n + 1$ vertices at maximum eccentricity, which we find to be $2n - 2\alpha$.

So, we have that,

$$\zeta(G) \leq (n - \alpha) + 2(n - \alpha + 1) + 2(n - \alpha + 2) + 2(n - \alpha + 2)$$

$$
\begin{aligned}
&+\cdots+2(2n-2\alpha-1)+(2\alpha-n+1)(2n-2\alpha)\\
&=(n-\alpha)+\sum_{i=1}^{n-\alpha-1}2(n-\alpha+i)+2(2\alpha-n+1)(n-\alpha)\\
&=(n-\alpha)+2(2\alpha-n+1)(n-\alpha)+2\sum_{i=1}^{n-\alpha-1}(n-\alpha)+2\sum_{i=1}^{n-\alpha-1}i\\
&=(n-\alpha)[1+2(2\alpha-n+1)]+2(n-\alpha)(n-\alpha-1)+2\left(\frac{(n-\alpha-1)(n-\alpha)}{2}\right)\\
&=(n-\alpha)[4\alpha-2n+3]+2(n-\alpha)(n-\alpha-1)+(n-\alpha-1)(n-\alpha)\\
&=(n-\alpha)[4\alpha-2n+3]+3(n-\alpha)(n-\alpha-1)\\
&=(n-\alpha)[4\alpha-2n+3+3n-3\alpha-3]\\
&=(n-\alpha)[\alpha+n]\\
&=n\alpha+n^2-\alpha^2-n\alpha\\
&=n^2-\alpha^2,
\end{aligned}
$$

as desired.

To see that the above bound is sharp, consider the tree T obtained from a path of order $2n-2\alpha$ by appending $2\alpha-n$ end vertices to one end of the path. See Figure 3.9. We will denote the vertices in the path of order $2n-2\alpha$ by $P_{2n-2\alpha}$, and the attached vertices will be denoted by $V_{attached}$.
Let $v_1, v_2, v_3, \ldots, v_{2n-2\alpha-2}, v_{2n-2\alpha-1}, v_{2n-2\alpha}$ be the vertices of the path $P_{2n-2\alpha}$, and let

$$V_{attached}=\{u_1,u_2,u_3,\ldots,u_{2\alpha-n-2},u_{2\alpha-n-1},u_{2\alpha-n}\}$$

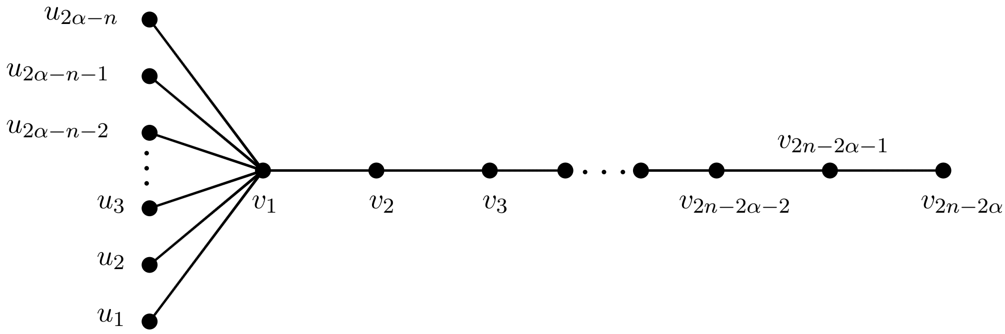


Figure 3.9: Tree T.

Tree T has order

$$2n-2\alpha+(2\alpha-n)=2n-n=n.$$

We see that, for any $x, y \in V_{attached}$, we have that $d_T(x,y)=2$, giving that the $2\alpha-n$ vertices attached to $P_{2n-2\alpha}$ form an independent set of order $2\alpha-n$. Also, vertices in $P_{2n-2\alpha}$ give two independent sets that have order $\frac{2n-2\alpha}{2}=n-\alpha$ at most, and one of these independent sets will have vertices that are not adjacent to any of the $V_{attached}$ vertices. Denote this independent set by $A_{2n-2\alpha}$. So, we have that, the biggest independent set in T is given by $A_{2n-2\alpha}\cup V_{attached}$, thus, we have that, the independence number of T is $\alpha(T)=|A_{2n-2\alpha}|+|V_{attached}|=n-\alpha+(2\alpha-n)=\alpha$.

So, we see that T has order n and independence number α. We calculate $\zeta(T)$.

All vertices $x \in V_{attached}$ have eccentricity $2n-2\alpha$ in T. Also, a diametral path in T has order $2n-2\alpha+1$. We denote a diametral path by $P_{2n-2\alpha+1}$, and we note that in $V(P_{2n-2\alpha+1})$, there is one vertex that is in the set $V_{attached}$, giving that we will be counting it for the second time.
Let $v_0, v_1, v_2, \ldots, v_{2n-2\alpha}$ be vertices in $P_{2n-2\alpha+1}$. By Lemma 3.2.1, we have that,

$$\begin{aligned}\sum_{x\in V(P_{2n-2\alpha+1})} e_G(x) &= \sum_{i=0}^{\lfloor\frac{2n-2\alpha}{2}\rfloor}(2n-2\alpha-i) + \sum_{i=\lfloor\frac{2n-2\alpha}{2}\rfloor+1}^{2n-2\alpha} i \\ &= \sum_{i=0}^{n-\alpha}(2n-2\alpha-i) + \sum_{i=n-\alpha+1}^{2n-2\alpha} i. \end{aligned} \tag{3.19}$$

We have that,

$$\begin{aligned}\sum_{i=0}^{n-\alpha}(2n-2\alpha-i) &= \sum_{i=0}^{n-\alpha}(2n-2\alpha) - \sum_{i=0}^{n-\alpha} i \\ &= (2n-2\alpha)(n-\alpha+1) - \frac{(n-\alpha)(n-\alpha+1)}{2} \\ &= 2(n-\alpha)(n-\alpha+1) - \frac{(n-\alpha)(n-\alpha+1)}{2}. \end{aligned} \tag{3.20}$$

And,

$$\begin{aligned}\sum_{i=n-\alpha+1}^{2n-2\alpha} i &= \sum_{i=0}^{2n-2\alpha} i - \sum_{i=0}^{n-\alpha} i \\ &= \frac{(2n-2\alpha)(2n-2\alpha+1)}{2} - \frac{(n-\alpha)(n-\alpha+1)}{2} \\ &= (n-\alpha)(2n-2\alpha+1) - \frac{(n-\alpha)(n-\alpha+1)}{2}. \end{aligned} \tag{3.21}$$

Plugging (3.20) and (3.21) into (3.19), we get,

$$\begin{aligned}\sum_{x\in V(P_{2n-2\alpha+1})} e_G(x) &= 2(n-\alpha)(n-\alpha+1) - \frac{(n-\alpha)(n-\alpha+1)}{2} \\ &\quad + (n-\alpha)(2n-2\alpha+1) - \frac{(n-\alpha)(n-\alpha+1)}{2} \\ &= (n-\alpha)\Bigg(2n-2\alpha+2+2n-2\alpha+1\Bigg) - 2\Bigg(\frac{(n-\alpha)(n-\alpha+1)}{2}\Bigg) \\ &= (n-\alpha)(4n-4\alpha+3) - (n-\alpha)(n-\alpha+1) \\ &= (n-\alpha)\Bigg(4n-4\alpha+3-n+\alpha-1\Bigg) \\ &= (n-\alpha)(3n-3\alpha+2). \end{aligned}$$

So, we have that, (keeping in mind that we need to substract one of the eccentricity of the $2\alpha-n$ vertices because of the double counting.)

$$\zeta(T) = \sum_{x\in V_{attached}} e_G(x) + \sum_{x\in V(P_{2n-2\alpha+1})} e_G(x) - 1(2n-2\alpha)$$

$$
\begin{aligned}
&= (2\alpha - n)(2n - 2\alpha) + (n - \alpha)(3n - 3\alpha + 2) - 1(2n - 2\alpha) \\
&= (n - \alpha)\Bigg(4\alpha - 2n + 3n - 3\alpha + 2 - 2\Bigg) \\
&= (n - \alpha)(\alpha + n) \\
&= n^2 - \alpha^2.
\end{aligned}
$$

□

Since $avec(G) = \frac{1}{n}\zeta(G)$, we are able to express Theorem 3.2.16, and Theorem 3.2.23, in terms of $avec(G)$ as follows: By Theorem 3.2.16, we have that, for a connected graph G of order n and independence number α such that $\alpha \leq \frac{n}{2}$, the average eccentricity of G is at most $(2\alpha - 1) - \frac{(\alpha^2 - \alpha)}{n}$. While Theorem 3.2.23 gives that for a connected graph G of order n and independence number α such that $\alpha < \frac{n}{2}$, the average eccentricity of G is at most $n - \frac{\alpha^2}{n}$.

3.3 Average eccentricity, order and chromatic number

This section is dedicated to proving upper bound on average eccentricity in terms of order and chromatic number of the graph. We start by providing proofs (for the sake of completeness) to results that will assist us in proving main theorem of this section, Theorem 3.3.19. To be able to analyse some of the results of this section, we will need the following definitions:

Definition 3.3.1. A connected graph G of order $n > 1$, without cut-vertices is called *nonseperable*.

Definition 3.3.2. Let G be a connected graph with n vertices, where n is greater than one and let C be a subgraph of G with no cut-vertices. The subgraph C is said to be a *block* of G if C is not a proper subgraph of any nonseperable subgraphs of G.

Definition 3.3.3. Let G be a connected graph and let $S \subset V(G)$. If $G - S$ is disconnected, then S is called a *vertex-cut* of G. The smallest cardinality of vertex-cut in G is called the *vertex connectivity* of G and we denote it by $\kappa(G)$. Let k be an integer greater than one. If $\kappa(G) \geq k$, then we say that G is *k-connected*.

Definition 3.3.4. A graph G is called *color critical* if $\chi(H) < \chi(G)$ for every proper subgraph H of G. If G is k-chromatic and color-critical graph, then G is called *k-critical*.

Definition 3.3.5. (The union of G and H)

Let G and H be graphs, with H not a subgraph of G and vice-versa. The *union* $G \cup H$ of G and H, is the graph given by

$$V(G \cup H) = V(G) \cup V(H) \text{ and } E(G \cup H) = E(G) \cup E(H).$$

Definition 3.3.6. (The join of G and H)

For two verterx-disjoint graphs G and H, the *join* $G + H$ of G and H has vertex set

$$V(G + H) = V(G) \cup V(H)$$

and edge set

$$E(G + H) = E(G) \cup E(H) \cup \{uv \,:\, u \in V(G), v \in V(H)\}.$$

Results of this section begin with the following proposition:

Proposition 3.3.7. *For graphs $G_1, G_2, \ldots, G_k$ and $G = G_1 \cup G_2 \cup \cdots \cup G_k$,*

$$\chi(G) = \max\{\chi(G_i) : 1 \leq i \leq k\}.$$

Proof

Let $G = G_1 \cup G_2 \cup \cdots \cup G_k$ and let $c : V(G) \to \mathbb{N}$ be a coloring map in G such that for any $u, v \in V(G)$, if $d_G(u, v) = 1$, then $c(u) \neq c(v)$.
Without loss of generality, assume that $\chi(G_k) \geq \chi(G_i)$ for any $i \neq k$. Since for all $v_i \in V(G_i)$ and $v_j \in V(G_j)$, where $i \neq j$, we have that $d_G(v_i, v_j) = \infty$, we can use colors of G_k to color any G_i for $i \neq k$ in G. Therefore,

$$\chi(G) = \chi(G_k) = \max\{\chi(G_i) : 1 \leq i \leq k\}.$$

□

From Proposition 3.3.7, we get the following corollary.

Corollary 3.3.8. *If G is a graph with components $G_1, G_2, \ldots, G_k$, then*

$$\chi(G) = \max\{\chi(G_i) : 1 \leq i \leq k\}.$$

Proof:

Let G be a graph with components $G_1, G_2, \ldots, G_k$. Then, we have that the components of G are vertex-disjoint, giving that $G = G_1 \cup G_2 \cup \cdots \cup G_k$. So, by Proposition 3.3.7, we have that

$$\chi(G) = \max\{\chi(G_i) : 1 \leq i \leq k\}.$$

□

Theorem 3.3.9. *If a connected graph G has cut-vertices, then G has at least two end-blocks.*

Proof:

Let G be a graph with cut-vertices. Let G' be a graph with vertex set the blocks of G where uv is an edge in G' if the corresponding blocks share a cut-vertex of u and v. Then G' is a tree. But any tree contains at least two end vertices. Therefore, G' has at least two end vertices and so G has at least two end blocks. □

Proposition 3.3.10. *Let G be a connected graph with order greater than one, and blocks $B_1, B_2, \ldots, B_k$. We have that,*

$$\chi(G) = \max\{\chi(B_i) \, : \, 1 \leq i \leq k\}.$$

Proof:

We prove the proposition by induction on the number of blocks of G. Let G be a non-trivial connected graph with blocks $B_1, B_2, \ldots, B_k$.
If $G = B_1$, then $\chi(G) = \chi(B_1)$, therefore, proposition holds if $k = 1$.
Induction base: For $k = 2$; G has two blocks B_1 and B_2.
Let u be a cut-vertex connecting B_1 and B_2 and assume without loss of generality that

$$\chi(B_2) \geq \chi(B_1). \tag{3.22}$$

Note that for $x \in B_1$ and $y \in B_2$, $x, y \neq u$ we have $d_G(x, y) > 1$. Hence by (3.22), we have that colours used to colour vertices in B_2 can colour vertices in $B_1 - \{u\}$. So, $\chi(G) = \chi(B_2) =$

$\max\{\chi(B_i) : 1 \leq i \leq 2\}$. Thus, the proposition is true for $k = 2$.
Assume that the theorem holds for a graph with $k-1$ blocks. Consider graph G. Since G has cut-vertices, by Theorem 3.3.9, we have that G has end blocks. Let B_k be an end block in G. Let G' be the graph obtained from G by removing block B_k except for the cut-vertex of B_k, say u. Then G' has $k-1$ blocks, $B_1, B_2, \ldots, B_{k-1}$. Assume without loss of generality that

$$\chi(B_{k-1}) \geq \chi(B_i), \text{ for any } i \neq k-1.$$

so that $\chi(G') = \chi(B_{k-1})$.
If

$$\chi(B_{k-1}) \geq \chi(B_k),$$

then colours used to colour vertices of B_{k-1} can be used to colour $B_k - \{u\}$ and so $\chi(G) = \chi(B_{k-1})$. If

$$\chi(B_k) \geq \chi(B_{k-1}),$$

then since $\chi(G') = \chi(B_{k-1})$ and vertices in B_k are coloured using $\chi(B_k)$ colours, we have $\chi(G) = \chi(B_k)$ and the theorem is proven. □

Proposition 3.3.11. *If H is a k-critical graph with $k \geq 2$, then H is connected.*

Proof:
Suppose not. Then H is not connected. Thus, H has two or more components. Let $H_1, H_2, \ldots, H_k$ be components in H. By Corollary 3.3.8, we have that $\chi(H) = \max\{\chi(H_i) : 1 \leq i \leq k\}$. Without loss of generality, let $\chi(H) = \chi(H_k)$. Since H_k is a proper subgraph of H, we have that H is not k-critical, which contradicts the assumption.
So, we must have that H is connected. □

Proposition 3.3.12. *If H is a k-critical graph with $k \geq 2$, then H is 2-connected.*

Proof: Suppose not. Since H is k-critical, we have by Proposition 3.3.11 that H is connected. Thus, we have that H has a cut-vertex. Let $B_1, B_2, \ldots, B_k$ be blocks in H.
By Proposition 3.3.10, we have that $\chi(H) = \max\{\chi(B_i) : 1 \leq i \leq k\}$. Let without loss of generality, $\chi(H) = \chi(B_k)$. Since B_k is a proper subgraph of H, we get that H is not k-critical, which contradicts the assumption. So, we must have that H is 2-connected. □

Proposition 3.3.13. *Let G be a k-chromatic graph with $k \geq 2$. Then G contains a k-critical subgraph.*

Proof: Amongst all k-chromatic subgraphs of G, let H be a subgraph with minimum size having no isolated vertices.
We claim that every proper subgraph F of H satisfies $\chi(F) < \chi(H)$.
To see that the claim is true, suppose to the contrary that F is a proper subgraph of H with $\chi(F) \geq \chi(H)$.
Therefore, since F is a proper subgraph of H, we have $\chi(F) = \chi(H)$, and so F is k-chromatic since H is k-chromatic. But since H has no isolated vertices, we have that size of F is less than the size of H. This contradicts our choice of H. Hence, $\chi(F) < \chi(H)$.
So by our claim, we have that H is k-critical. □

Algorithm 3.3.14. The Greedy Coloring Algorithm
Suppose vertices of graph G are listed in the order $v_1, v_2, \ldots, v_n$. Assume vertex v_1 is assigned the color 1. Then when the vertices $v_1, v_2, \ldots, v_j$ have been assigned colors, where $1 \leq j < n$, the vertex v_{j+1} is assigned the smallest color that is not assigned to any neighbor of v_{j+1} belonging to the set $\{v_1, v_2, \ldots, v_j\}$.

Lemma 3.3.15. *Let G be a connected graph that is not complete. Let H be a complete subgraph of G, then $\Delta(G) \geq k$.*

Proof: Since H is a subgraph of G, and every vertex $x \in V(H) = V(K_k)$ has degree $deg_H(x) = k - 1$, and G is not complete, we have that, there is at least one vertex $v \in V(G) - V(H)$.
But G is connected, so, for some $u \in V(H)$, we have that $d_G(u, v) = 1$, giving that $deg_G(u) \geq k - 1 + 1 = k$.
Since $\Delta(G) \geq deg_G(u)$, we have that $\Delta(G) \geq k$, as desired. □

Lemma 3.3.16. *Let G be a connected graph that is not an odd cycle. Let H be a subgraph of G and let H be an odd cycle. Then $\Delta(G) \geq 3$.*

Proof: Since H is a subgraph of G, and every vertex $x \in V(H)$ has degree $deg_H(x) = 2$, and since G is connected and not an odd cycle, we have that, there is at least one vertex $v \in V(G) - V(H)$ that is adjacent to some $u \in V(H)$. Hence, $deg_G(u) \geq 2 + 1 = 3$. So, we have that $\Delta(G) \geq 3$ as desired. □

Corollary 3.3.17. *From Lemma 3.3.15 and Lemma 3.3.16, we have that $\chi(G) \leq \Delta(G)$ when G is neither an odd cycle nor a complete graph, but H is either a complete graph or an odd cycle. Giving that the theorem is true for H being either a complete graph or an odd cycle.*

Theorem 3.3.18. ***(Brooks' Theorem)***
If G is a connected graph that is not an odd cycle or a complete graph, then

$$\chi(G) \leq \Delta(G).$$

Proof: Let $\chi(G) = k \geq 2$ and let H be a k-critical subgraph of G (This follows by Proposition 3.3.13). Then, we have that H is k-chromatic, so $\chi(H) = k = \chi(G)$ and by Proposition 3.3.12, we have that H is 2-connected. Since H is a subgraph of G, we have that $\Delta(H) \leq \Delta(G)$. If H is a complete graph (notice that $H = K_k$ is this case) or H is an odd cycle, then $G \neq H$ since G is neither an odd cycle nor a complete graph.

By Lemma 3.3.15, if $H = H_k$, then $\chi(G) = \chi(H) = k \leq \Delta(G)$ and by Lemma 3.3.16, if H is an odd cycle, then $\Delta(G) \geq 3 = k = \chi(H) = \chi(G)$. So theorem holds for these cases. We show that the theorem holds when H is a k-critical subgraph that is not an odd cycle or a complete graph.

Suppose H is a k-critical subgraph that is neither an odd cycle nor a complete graph.

We show that we can assume that $k \geq 4$. Suppose $k \leq 3$. Since H is 2-connected, $deg_H(u) \geq 2$ for all $u \in V(H)$. If $deg_H(u) = 2$ for all $u \in V(H)$, then $H = G$ is a cycle and so by the hypothesis $H = G$ is an even cycle. It follows that $k = 2 \leq \Delta(H) = 2$. If on the other hand there is a vertex u on H with $deg_H(u) \geq 3$, then $k \leq 3 \leq deg_H(u) \leq \Delta(H)$ and the theorem holds. So, we can assume that $k \geq 4$.

Assume that H has n vertices. Since $\chi(G) = k \geq 4$ and H is not complete, we have that $n > k$ and so $n \geq 5$. But H is 2-connected. So either two vertices are required to be removed from H to obtain a disconnected graph or we will need to remove a minimum of three vertices to disconnect H. Thus, we have that H is either 3-connected or H has connectivity 2.

Case 1. Suppose H is $3 -$ *connected*. By our assumption, we have that H is not complete. It follows that, at least two vertices u and w in H are at a distance 2 from one another.

Let (u, v, w) be a $u - w$ geodesic in H. We have that H is 3-connected, so removing two vertices will not disconnect H, thus, $H - u - w$ is connected. Suppose $H - u - w$ is made of vertices $v = u_1, u_2, \dots, u_{n-2}$, such that each vertex $u_i\,(2 \le i \le n-2)$ is adjacent to some vertex preceding it. Let $u_{n-1} = u$ and $u_n = w$. Consequently, for each set

$$U_j = \{u_1, u_2, \dots, u_j\}, 1 \le j \le n,$$

the induced subgraph $H[U_j]$ is connected. Therefore, $H[U_n] = H[V(H)] = H$ is connected. We use Algorithm 3.3.14 on H with respect to the reverse ordering

$$w = u_n, u = u_{n-1}, u_{n-2}, \dots, u_2, u_1 = v \tag{3.23}$$

of the vertices of H. But $d_H(u, w) = 2$, so we assign color 1 to u and w. Also, the smallest color in $\{1, 2, \dots, \Delta(H)\}$ that was not used to color any neighbours of $u_i(2 \le i \le n-2)$ that preceded u_i in the sequence (3.23), is assigned to u_i. Since we apply the greedy coloring theorem in reverse order in (3.23), by the assumption, we have that each vertex u_i has at least one neighbor following it in the sequence (3.23), thus, u_i has at most $\Delta(H) - 1$ neighbors that come before it in the sequence. It follows, by the greedy coloring theorem, that there is a color available for u_i. Since there u-w geodesic in H of distance 2 given by (u, v, w), we have that the vertex $u_1 = v$ is adjacent to two vertices colored 1 (namely $w = u_n$ and $u = u_{n-1}$) and so at most $\Delta(H) - 1$ colors are assigned to the neighbors for v. So we have that $\chi(H) \le \Delta(H) - 1 + 1 = \Delta(H)$. Hence it follows that

$$\chi(G) = \chi(H) \le \Delta(H) \le \Delta(G). \tag{3.24}$$

Case 2. $\kappa(H) = 2$. We claim that H contains a vertex x such that $2 < deg_H x < n - 1$. Suppose that this is not the case. Then every vertex of H has degree 2 or $n-1$. Because $\chi(H) \ge 4$, it follows that H cannot contain only vertices of degree 2; and because H is not complete, H cannot contain only vertices of degree $n - 1$.

If H contains vertices of both degrees 2 and $n - 1$ only, then either

$$H = K_1 + \Big(\frac{n-1}{2}\Big)K_2 \text{ or } H = K_{1,1,n-2}.$$

Note that $H = K_1 + \big(\frac{n-1}{2}\big)K_2$ has $\frac{n-1}{2} + 1$ partite sets, with $\frac{n-1}{2}$ sets having order 2. Denote these partite sets by D_i, $1 \le i \le \frac{n-1}{2}$. We can use the same two colors to color vertices in D_i for any i, while the remaining one vertex can be colored using a different color. So, in this case $\chi(H) = 3$. But, for any $u \in D_i \subset V(H)$, for some $1 \le i \le \frac{n-1}{2}$, $\chi(H - u) = 3$, so H is not critical.

Note that $H = K_{1,1,n-2}$ has three partite sets, with two sets having order one, and the other one having order $n - 2$. The vertices in the set of order $n - 2$ can be colored using only one color, and the two remaining vertices require two different colors. So, $\chi(H) = 3$, but for any vertex u in the partite set of order $n-2$, $\chi(H-u) = 3$, so H is not critical. So, we obtain a contradiction, since we assume H is k-critical. Thus, as claimed, H contains a vertex x such that $2 < deg_H x < n - 1$.

Since $\kappa(H) = 2$, we have that, either x is in a vertex-cut of the smallest cardinality, or it is not, giving that $\kappa(H - x) = 1$ or $\kappa(H - x) = 2$. Suppose $\kappa(H - x) = 2$. It follows that x is not in any minimum vertex-cut of H, giving that every neighbor of x has degree at least two. Thus, we have that H contains a vertex y with $d_H(x, y) = 2$. Continuing as we did in Case 1, where $u = x$ and $w = y$, we have that there is a coloring of H with at most $\Delta(H)$ colors and so once again we have (3.24), that is, $\chi(G) \le \Delta(G)$.

We may assume that $\kappa(H - x) = 1$. Thus $H - x$ has a cut-vertex, and so, $H - x$ contains end blocks B_1 and B_2, containing cut-vertices x_1 and x_2, respectively, of $H - x$. Since H

is 2-connected, there exist vertices $y_1 \in V(B_1) - \{x_1\}$ and $y_2 \in V(B_2) - \{x_2\}$ such that x is adjacent to both y_1 and y_2. Proceeding as in Case 1 with $u = y_1$ and $w = y_2$, we obtain a coloring of H with at most $\Delta(H)$ colors, giving us (3.24) and so $\chi(G) \leq \Delta(G)$. □

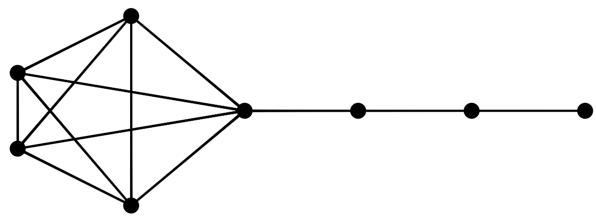
Theorem 3.3.19. *Consider a connected graph G with n vertices and chromatic number χ, where $2 \leq \chi \leq n-1$. We have that*

$$\zeta(G) \leq \left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor,$$

with equality holding true if G is the graph obtained from combining the complete graph K_χ and a path $P_{n-\chi}$ with an edge that joins an end vertex of $P_{n-\chi}$ to one of the vertices of K_χ.

We illustrate the described graph in the last statement of the theorem by a way of example.

Example 3.3.20. If $n = 8$ and $\chi = 5$, then the described graph is,



Proof of Theorem 3.3.19:

We note the following,

$$\begin{aligned}\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor &= \left\lfloor \frac{1}{4}\Big(3n^2 - \chi^2 - 2\chi n\Big) + \frac{1}{2}\Big(n + \chi\Big) \right\rfloor \\ &= \left\lfloor \frac{1}{4}\Big(3n + \chi\Big)\Big(n - \chi\Big) + \frac{1}{2}\Big(n + \chi\Big) \right\rfloor.\end{aligned}$$

If G is an odd cycle, then $\chi = 3$ and n is odd, so, $n-1$ is even, thus, for $\chi = 3$ and n odd, we have,

$$\begin{aligned}\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor &= \left\lfloor \frac{3}{4}n^2 - \frac{1}{4}(3)^2 - \frac{1}{2}(3)n + \frac{1}{2}n + \frac{1}{2}(3) \right\rfloor \\ &= \left\lfloor \frac{3n^2}{4} - \frac{3n}{2} + \frac{n}{2} - \frac{9}{4} + \frac{3}{2} \right\rfloor \\ &= \left\lfloor \frac{3n^2}{4} - n - \frac{3}{4} \right\rfloor \\ &= \left\lfloor \frac{3(n-1+1)^2}{4} - n - \frac{3}{4} \right\rfloor \\ &= \left\lfloor \frac{3}{4}\Big((n-1)^2 + 2(n-1) + 1\Big) - n - \frac{3}{4} \right\rfloor\end{aligned}$$

$$\begin{aligned}
&= \left\lfloor \frac{3}{4}\Big((n-1)^2 + 2(n-1) + 1\Big) - n - \frac{3}{4} \right\rfloor \\
&= \left\lfloor \frac{3(n-1)^2}{4} + \frac{3(n-1)}{2} - n + \frac{3}{4} - \frac{3}{4} \right\rfloor \\
&= \left\lfloor \frac{3(n-1)^2}{4} + \frac{3(n-1)}{2} - n \right\rfloor \\
&= \frac{3(n-1)^2}{4} + \frac{3(n-1)}{2} - n \\
&= \frac{3(n-1)}{2}\Big(\frac{(n-1)}{2} + 1\Big) - n \\
&= \frac{3(n-1)}{2}\Big(\frac{n+1}{2}\Big) - n \\
&= \frac{3(n-1)(n+1)}{4} - n.
\end{aligned}$$

If G is an odd cycle, then,

$$\begin{aligned}
\zeta(G) &= \sum_{x \in V(G)} e_G(x) \\
&= \sum_{i=1}^{n} \frac{n-1}{2} \\
&= \frac{n(n-1)}{2}.
\end{aligned}$$

Now let $f(n) = \frac{n(n-1)}{2}$ and $g(n) = \frac{3(n-1)(n+1)}{4} - n$. Simple differentiation shows that $f(n)$ and $g(n)$ are increasing functions for $n \geq 3$ and $g(n) \geq f(n)$ for $n \geq 3$. Thus, for G an odd cycle, the theorem holds.

Let G be a complete graph. Then $\chi = n$, so,

$$\begin{aligned}
\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor &= \left\lfloor \frac{3}{4}n^2 - \frac{1}{4}n^2 - \frac{1}{2}(n)n + \frac{1}{2}n + \frac{1}{2}n \right\rfloor \\
&= \left\lfloor \frac{1}{2}n^2 - \frac{1}{2}n^2 + n \right\rfloor \\
&= n,
\end{aligned}$$

and

$$\begin{aligned}
\zeta(G) &= \sum_{x \in V(G)} e_G(x) \\
&= \sum_{i=1}^{n} 1 \\
&= n.
\end{aligned}$$

Thus, the theorem holds when G is a complete graph.

Assume that G is neither an odd cycle, nor complete. Since G has chromatic number χ and is connected, it follows by Theorem 3.3.18, that $\Delta(G) \geq \chi$. It follows that G contains a vertex with degree χ or more, giving that G has a spanning tree T of maximum degree at least χ. Thus,

$$diam(T) \leq n + 1 - \chi.$$

By Proposition 3.2.4, and since $f(d) = nd - \frac{1}{4}x^2 + \frac{1}{4}$ is increasing, we have,

$$\begin{aligned}
\zeta(G) \leq \zeta(T) &\leq \left\lfloor diam(T)n - \frac{1}{4}diam(T)^2 + \frac{1}{4} \right\rfloor \\
&\leq \left\lfloor (n+1-\chi)n - \frac{1}{4}(n+1-\chi)^2 + \frac{1}{4} \right\rfloor \\
&= \left\lfloor (n+1-\chi)\left[n - \frac{n+1-\chi}{4}\right] + \frac{1}{4} \right\rfloor \\
&= \left\lfloor (n+1-\chi)\left[\frac{3n-1+\chi}{4}\right] + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \frac{1}{4}(n-\chi+1)(3n+\chi-1) + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \frac{1}{4}\Big((n-\chi)(3n+\chi) - (n-\chi) + (3n+\chi) - 1\Big) + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \frac{1}{4}\Big((n-\chi)(3n+\chi) - n + \chi + 3n + \chi\Big) - \frac{1}{4} + \frac{1}{4} \right\rfloor \\
&= \left\lfloor \frac{1}{4}\Big((n-\chi)(3n+\chi) + 2n + 2\chi\Big) \right\rfloor \\
&= \left\lfloor \frac{1}{4}(n-\chi)(3n+\chi) + \frac{n+\chi}{2} \right\rfloor \\
&= \left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor,
\end{aligned}$$

as desired.
We now verify the second statement of the theorem.

So, if both n and χ are either even or odd, then each of the following sums is even, $3n + \chi$, $n - \chi$, $n + \chi$.
So, $\frac{1}{4}(3n+\chi)(n-\chi)$ is an integer, since $3n + \chi = 2k$ for some $k \in \mathbb{Z}$ and $n - \chi = 2t$ for some $t \in \mathbb{Z}$, we have, $\frac{1}{4}(3n+\chi)(n-\chi) = \frac{1}{4}(2k)(2t) = kt \in \mathbb{Z}$.
$\frac{1}{2}(n+\chi) = \frac{1}{2}(2q) = q$ for some $q \in \mathbb{Z}$.
So, for when both n and χ are either even or odd, we have

$$\begin{aligned}
\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor &= \left\lfloor \frac{1}{4}\Big(3n+\chi\Big)\Big(n-\chi\Big) + \frac{1}{2}\Big(n+\chi\Big) \right\rfloor \\
&= \frac{1}{4}\Big(3n+\chi\Big)\Big(n-\chi\Big) + \frac{1}{2}\Big(n+\chi\Big). \qquad (3.25)
\end{aligned}$$

If either n or χ is even and the other one is odd, then, each of the following sums is odd, $3n + \chi$, $n - \chi$ and $n + \chi$.
So, subtracting one from each of the above sums, we get an even sum. So, the following sums are even,

$$3n + \chi - 1,\ n - \chi - 1 \text{ and } n + \chi - 1. \qquad (3.26)$$

So,

$$\begin{aligned}
\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor &= \left\lfloor \frac{1}{4}\Big(3n+\chi\Big)\Big(n-\chi\Big) + \frac{1}{2}\Big(n+\chi\Big) \right\rfloor \\
&= \left\lfloor \frac{1}{4}\Big(3n+\chi-1+1\Big)\Big(n-\chi-1+1\Big) + \frac{1}{2}\Big(n+\chi-1+1\Big) \right\rfloor
\end{aligned}$$

$$
\begin{aligned}
&= \Bigg\lfloor \frac{1}{4}\Big[\Big(3n+\chi-1\Big)\Big(n-\chi-1\Big)+\Big(3n+\chi-1\Big)+\Big(n-\chi-1\Big)+1\Big] \\
&\quad + \frac{1}{2}\Big(n+\chi-1+1\Big)\Bigg\rfloor \\
&= \Bigg\lfloor \frac{1}{4}\Big[\Big(3n+\chi-1\Big)\Big(n-\chi-1\Big)+4n-2+1\Big]+\frac{1}{2}\Big(n+\chi-1+1\Big)\Bigg\rfloor \\
&= \Bigg\lfloor \frac{(3n+\chi-1)(n-\chi-1)}{4}+\frac{4n-1}{4}+\frac{n+\chi-1+1}{2}\Bigg\rfloor \\
&= \Bigg\lfloor \frac{(3n+\chi-1)(n-\chi-1)}{4}+\frac{4n}{4}-\frac{1}{4}+\frac{n+\chi-1}{2}+\frac{1}{2}\Bigg\rfloor \\
&= \Bigg\lfloor \frac{(3n+\chi-1)(n-\chi-1)}{4}+n+\frac{n+\chi-1}{2}-\frac{1}{4}+\frac{1}{2}\Bigg\rfloor \\
&= \Bigg\lfloor \frac{(3n+\chi-1)(n-\chi-1)}{4}+n+\frac{n+\chi-1}{2}+\frac{1}{4}\Bigg\rfloor
\end{aligned}
$$

But, by (3.26), we have, $\frac{(3n+\chi-1)(n-\chi-1)}{4}$ and $\frac{n+\chi-1}{2}$ are integers. Therefore,

$$
\begin{aligned}
\Bigg\lfloor \frac{3}{4}n^2-\frac{1}{4}\chi^2-\frac{1}{2}\chi n+\frac{1}{2}n+\frac{1}{2}\chi\Bigg\rfloor &= \Bigg\lfloor \frac{(3n+\chi-1)(n-\chi-1)}{4}+n+\frac{n+\chi-1}{2}+\frac{1}{4}\Bigg\rfloor \\
&= \frac{(3n+\chi-1)(n-\chi-1)}{4}+n+\frac{n+\chi-1}{2} \\
&= \frac{(3n+\chi)(n-\chi)-(3n+\chi)-(n-\chi)+1}{4}+n+\frac{n+\chi-1}{2} \\
&= \frac{(3n+\chi)(n-\chi)-4n+1}{4}+n+\frac{n+\chi-1}{2} \\
&= \frac{(3n+\chi)(n-\chi)}{4}-\frac{4n}{4}+n+\frac{1}{4}+\frac{n+\chi}{2}-\frac{1}{2} \\
&= \frac{(3n+\chi)(n-\chi)}{4}-n+n+\frac{n+\chi}{2}+\frac{1}{4}-\frac{1}{2} \\
&= \frac{(3n+\chi)(n-\chi)}{4}+\frac{n+\chi}{2}-\frac{1}{4} \\
&= \frac{1}{4}\Big((3n+\chi)(n-\chi)+2(n+\chi)-1\Big). \qquad (3.27)
\end{aligned}
$$

Let $T_{K_\chi P_{n-\chi}}$ be the described graph in the statement, with K_χ being the complete graph of order χ in $T_{K_\chi P_{n-\chi}}$ and $P_{n-\chi}$ being the path of order $n-\chi$ in $T_{K_\chi P_{n-\chi}}$ and $V(K_\chi)\cap V(P_{n-\chi})=\emptyset$ in $T_{K_\chi P_{n-\chi}}$. Let the diametral path be denoted by $P_{n-\chi+2}$ in $T_{K_\chi P_{n-\chi}}$. $\chi-1$ vertices in K_χ have eccentricity $n-\chi+1$, but the one vertex K_χ that is adjacent to the end vertex of $P_{n-\chi}$ has eccentricity $n-\chi$ and it lies on the diametral path. So, we only count twice one vertex in K_χ that has eccentricity $n-\chi+1$. To solve the double counting, we will disregard one vertex in K_χ when calculating $\sum_{x\in V(K_\chi)} e_{T_{K_\chi P_{n-\chi}}}(x)$. So, let u be some vertex in $V(K_\chi)$ such that $e_{T_{K_\chi P_{n-\chi}}}(u)=(n-\chi+1)$ and let v be a vertex in K_χ such that $e_{T_{K_\chi P_{n-\chi}}}(v)=(n-\chi)$. Then,

$$
\zeta(T_{K_\chi P_{n-\chi}}) = \sum_{x\in V(K_\chi)-\{u,v\}} e_{T_{K_\chi P_{n-\chi}}}(x) + \sum_{x\in V(P_{n-\chi+2})} e_{T_{K_\chi P_{n-\chi+2}}}(x)
$$

$$
\begin{aligned}
&= \sum_{i=1}^{\chi-2}(n-\chi+1) + \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} (n-\chi+1-i) + \sum_{i=\left\lfloor \frac{n-\chi+1}{2} \right\rfloor+1}^{n-\chi+1} i \\
&= (n-\chi+1)\sum_{i=1}^{\chi-2} 1 + \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} (n-\chi+1) - \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} i + \sum_{i=0}^{n-\chi+1} i - \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} i \\
&= (n-\chi+1)\sum_{i=1}^{\chi-2} 1 + (n-\chi+1) \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} 1 - 2 \sum_{i=0}^{\left\lfloor \frac{n-\chi+1}{2} \right\rfloor} i + \sum_{i=0}^{n-\chi+1} i \\
&= (n-\chi+1)(\chi-2) + (n-\chi+1)\Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big) - 2\left(\frac{\Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor\Big)\Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big)}{2}\right) \\
&\quad + \frac{(n-\chi+1)(n-\chi+1+1)}{2} \\
&= (n-\chi+1)(\chi-2+\frac{n-\chi+1+1}{2}) + (n-\chi+1)\Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big) \\
&\quad - \Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor\Big)\Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big) \\
&= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big)\Big(n-\chi+1-\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor\Big).
\end{aligned}
$$

If χ and n are both even or both odd, then $n-\chi$ is even.
So, $\frac{n-\chi}{2}$ is an integer, thus,

$$
\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor = \Big\lfloor \frac{n-\chi}{2} + \frac{1}{2} \Big\rfloor = \frac{n-\chi}{2}.
$$

So, when χ and n are both even or odd, we have,

$$
\begin{aligned}
\zeta(T_{K_\chi P_{n-\chi}}) &= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor + 1\Big)\Big(n-\chi+1-\Big\lfloor \frac{n-\chi+1}{2} \Big\rfloor\Big) \\
&= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\frac{n-\chi}{2} + 1\Big)\Big(n-\chi+1-\frac{n-\chi}{2}\Big) \\
&= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\frac{n-\chi+2}{2}\Big)\Big(\frac{n-\chi+2}{2}\Big) \\
&= \frac{1}{2}\Big(n^2 + \chi n - 2n - \chi n - \chi^2 + 2\chi + n + \chi - 2\Big) \\
&\quad + \frac{1}{4}\Big(n^2 - \chi n + 2n - \chi n + \chi^2 - 2\chi + 2n - 2\chi + 4\Big) \\
&= \frac{1}{2}\Big(n^2 - n - \chi^2 + 3\chi - 2\Big) + \frac{1}{4}\Big(n^2 - 2\chi n + 4n + \chi^2 - 4\chi + 4\Big) \\
&= \frac{n^2}{2} - \frac{n}{2} - \frac{\chi^2}{2} + \frac{3\chi}{2} - 1 + \frac{n^2}{4} - \frac{2\chi n}{4} + \frac{4n}{4} + \frac{\chi^2}{4} - \frac{4\chi}{4} + \frac{4}{4} \\
&= \frac{n^2}{2} - \frac{n}{2} - \frac{\chi^2}{2} + \frac{3\chi}{2} - 1 + \frac{n^2}{4} - \frac{\chi n}{2} + n + \frac{\chi^2}{4} - 4\chi + 1 \\
&= \frac{n^2}{2} + \frac{n^2}{4} - \frac{n}{2} + n - \frac{\chi^2}{2} + \frac{\chi^2}{4} + \frac{3\chi}{2} - \chi - \frac{\chi n}{2} - 1 + 1 \\
&= \frac{3n^2}{4} + \frac{n}{2} - \frac{\chi^2}{4} + \frac{\chi}{2} - \frac{\chi n}{2}
\end{aligned}
$$

$$= \frac{3n^2}{4} - \frac{\chi n}{2} - \frac{\chi^2}{4} + \frac{n}{2} + \frac{\chi}{2}$$
$$= \frac{1}{4}\Big(3n+\chi\Big)\Big(n-\chi\Big) + \frac{1}{2}\Big(n+\chi\Big). \tag{3.28}$$

If χ or n is even and the other is odd, then $n-\chi$ is odd, thus, $n-\chi+1$ is even. So, $\frac{n-\chi+1}{2}$ is an integer. So, when χ or n is even and the other is odd, we have,

$$\begin{aligned}
\zeta(T_{K_\chi P_{n-\chi}}) &= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\Big\lfloor\frac{n-\chi+1}{2}\Big\rfloor+1\Big)\Big(n-\chi+1-\Big\lfloor\frac{n-\chi+1}{2}\Big\rfloor\Big)\\
&= (n-\chi+1)\Big(\frac{n+\chi-2}{2}\Big) + \Big(\frac{n-\chi+1}{2}+1\Big)\Big(n-\chi+1-\frac{n-\chi+1}{2}\Big)\\
&= \Big(\frac{n-\chi+1}{2}\Big)(n+\chi-2) + \Big(\frac{n-\chi+3}{2}\Big)\Big(\frac{n-\chi+1}{2}\Big)\\
&= \frac{n-\chi+1}{2}\Big(n+\chi-2+\frac{n-\chi+3}{2}\Big)\\
&= \frac{n-\chi+1}{2}\Big(\frac{3n+\chi-1}{2}\Big)\\
&= \frac{1}{4}(n-\chi+1)(3n+\chi-1)\\
&= \frac{1}{4}(3n^2+\chi n-n-3\chi n-\chi^2+\chi+3n+\chi-1)\\
&= \frac{1}{4}(3n^2-2\chi n-\chi^2+2n+2\chi-1)\\
&= \frac{1}{4}\Big((3n+\chi)(n-\chi)+2(n+\chi)-1\Big).
\end{aligned} \tag{3.29}$$

So, the second statement of the theorem is true, since (3.25) is equal to (3.28), and (3.27) is equal to (3.29). □

By the previous theorem and the fact that for any connected graph G, $avec(G) = \frac{1}{n}\zeta(G)$, we have that, for any connected graph G of order n and chromatic number χ, the average eccentricity of G is at most $\frac{1}{n}\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor$.

3.4 Conclusion

We divided this chapter into four sections. We gave a brief introduction in the first section on what we studied in this chapter. In the second section, we were concerned with proving upper bounds on average eccentricity in terms of order of the graph and independence number of the graph, while in the third section we proved results on upper bounds on average eccentricity in terms of order of the graph and chromatic number of the graph. To present our results in the second section, we split the results into two theorems. The first theorem is given by Theorem 3.2.16 and is stated as: "Let G be a connected graph of order n and independence number α, where $\alpha \leq \frac{n}{2}$. Then

$$\zeta(G) \leq (2\alpha-1)n - \alpha^2 + \alpha,$$

and this bound is sharp." The second theorem is given by Theorem 3.2.23, and is stated as "Let G be a connected graph of order n and independence number α, where $\alpha > \frac{n}{2}$. Then

$$\zeta(G) \leq n^2 - \alpha^2,$$

and the bound is sharp." After proving the aforementioned theorems, we presented results of Section 3.3. In Section 3.3, our results are given by Theorem 3.3.19, which is stated as follows:

"Let G be a connected graph of order n and chromatic number χ, where
$2 \leq \chi \leq n-1$. Then

$$\zeta(G) \leq \left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor.$$

Equality in the above bound is attained for the graph obtained from the union of a complete graph of order χ and a path of order $n-\chi$ by adding an edge joining an end vertex of the path to a vertex of the complete graph." In the next chapter, we study average distance in relation to independence number.

Chapter 4

Upper bound on the average distance in terms of order and independence number

4.1 Introduction

In this chapter, we are going to present results on upper bounds on average distance in terms of order and independence number of the graph. The results of this chapter are from a 2012 study by Mukwembi, presented in [18]. We use methods used in [18] to prove our results. To analyse some results of this chapter, we need the following:

Notation 4.1.1. We use $\mathcal{T}_{n,b}$ to denote the set of trees T; T results from taking some tree T' of order n, and then attaching b end vertices to some vertices of T'.

Notation 4.1.2. We use $B_{n-b,b}$ to denote the tree made of a path $P = v_1, v_2, \ldots, v_{n-b}$ with $\lfloor \frac{b}{2} \rfloor$ end vertices attached to v_1 and $\lceil \frac{b}{2} \rceil$ end vertices attached to v_{n-b}.

Notation 4.1.3. Let $a = \lceil \frac{n-2\alpha+2}{2} \rceil$ and $b = \lfloor \frac{n-2\alpha+2}{2} \rfloor$. We use $G_{n,\alpha}$ to denote the graph obtained by taking the path $P_{2\alpha-2}$, with end vertices p and q and joining p to each of the vertices of a complete graph K_a, while q is joined to each of the vertices of the complete graph K_b, with $V(K_a) \cap V(K_b) = \emptyset$.

4.2 Average distance, order and independence number

We begin our results with the following proposition:

Proposition 4.2.1. *For any connected graph G with a vertex v such that $G - v$ is connected, we have, for all $u, w \in V(G) - \{v\}$, $d_{G-v}(u,w) \geq d_G(u,w)$.*

Proof: Let G be a connected graph and v a vertex of G such that $G - v$ is connected. Then, for any $u, w \in V(G)$, a $u - w$ shortest path P in $G - v$ is also a $u - w$ path in G and so $d_G(u,w)$ is at most the length of P, i.e., $d_G(u,w) \leq d_{G-v}(u,w)$. □

Lemma 4.2.2. *Let G be a connected graph and v a vertex of G such that $G - v$ is connected. Then*

$$W(G) \leq W(G - v) + d_G(v).$$

Proof: Let G be a connected graph and v a vertex of G such that $G-v$ is connected. Then, applying Proposition 4.2.1, we get

$$\begin{aligned} W(G) &= \sum_{\{u,w\}\subset V(G)} d_G(u,w) \\ &= \sum_{\{u,w\}\subset V(G)-\{v\}} d_G(u,w) + \sum_{u\in V(G)} d_G(u,v) \\ &= \sum_{\{u,w\}\subset V(G-v)} d_G(u,w) + \sum_{u\in V(G)-\{v\}} d_G(u,v) \\ &\leq \sum_{\{u,w\}\subset V(G-v)} d_{G-v}(u,w) + \sum_{u\in V(G)-\{v\}} d_G(u,v) \\ &= W(G-v) + d_G(v). \end{aligned}$$

□

The following Proposition is folklore. We include the proof for completeness.

Proposition 4.2.3. *Consider a connected graph G with n vertices. We have that*

$$W(G) \leq \frac{n(n-1)(n+1)}{6}$$

and equality holds true if and only if G is a path.

Proof: We use induction on n to prove the proposition.
If $n=2$, then $G=K_2$, so $W(K_2)=1$ and $\frac{2(3)(1)}{6}=1$, thus, Proposition 4.2.3 is true for $n=2$. Assume that $n>2$. Choose a vertex v such that $G-v$ is connected. Assume that

$$W(G-v) \leq \frac{(n-1)(n-1+1)(n-1-1)}{6} = \frac{n(n-1)(n-2)}{6}.$$

Note that,

$$\begin{aligned} d_G(v) &\leq 1+2+3+\cdots+(n-1) \\ &= \frac{n(n-1)}{2}. \end{aligned}$$

By Lemma 4.2.2,

$$\begin{aligned} W(G) &\leq W(G-v) + d_G(v) \\ &\leq \frac{n(n-1)(n-2)}{6} + \frac{n(n-1)}{2} \\ &= \frac{n(n-1)}{2}\Big(\frac{n-2}{3}+1\Big) \\ &= \frac{n(n-1)}{2}\Big(\frac{n+1}{3}\Big) \\ &= \frac{n(n-1)(n+1)}{6}, \end{aligned}$$

as desired.
It can easily be verified that equality in the bound holds if and only if G is a path. □

Lemma 4.2.4. *For any $T \in \mathcal{T}_{n,b}$, we have that,*

$$W(T) \leq \frac{1}{6}n(n-1)(n+1) + \frac{1}{2}b(n+1)\left(n+\frac{1}{2}b\right) + \frac{1}{2}b^2 - b$$

and equality holds true if and only if $T = B_{n,b}$ and b is even.

Proof: Let T be a tree of maximum Wiener index in $\mathcal{T}_{n,b}$. Assume that T is obtained by attaching b end vertices to some tree H of order n. Let $B = V(T) - V(H)$. Then,

$$\begin{aligned} W(T) &= \sum_{\{x,y\}\subseteq V(T)} d_T(x,y) \\ &= \sum_{\{x,y\}\subseteq V(H)} d_T(x,y) + \sum_{x\in B, y\in V(H)} d_T(x,y) + \sum_{\{x,y\}\subseteq B} d_T(x,y) \\ &= \sum_{\{x,y\}\subseteq V(H)} d_H(x,y) + \sum_{x\in B, y\in V(H)} d_T(x,y) + \sum_{\{x,y\}\subseteq B} d_T(x,y) \\ &= W(H) + \sum_{x\in B, y\in V(H)} d_T(x,y) + \sum_{\{x,y\}\subseteq B} d_T(x,y). \end{aligned} \tag{4.1}$$

We bound terms of (4.1) independently, and use the obtained upper bounds to further simplify $W(T)$ expression above.

Since $T \in \mathcal{T}_{n,b}$ and is obtained from attaching b end vertices to some tree H, we have that H has order n and T has order $n+b$. So, from Proposition 4.2.3, we have that $W(H) \leq \frac{n(n-1)(n+1)}{6}$. By our assumption of T, we have that, for any $x \in B$, the neighbor of x is in H. But H is a tree, giving that H is connected. So the sum of the distances between x and vertices in H is at most

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

This is because $W(H)$ is maximised if H is a path, and attaching any end vertex $x \in B$ to H, we see that, for all graphs with vertices $V(H) \cup x$, the sum of distance is maximised if x is attached to one of the end vertices of path H. Since $V(H) \subseteq V(T)$, we have that

$$|B| = |V(T) - V(H)| = |V(T)| - |V(H)| = b + n - n = b.$$

So, it follows that

$$\sum_{x\in B, y\in V(H)} d_T(x,y) \leq b\frac{n(n+1)}{2}.$$

To bound $\sum_{\{x,y\}\subseteq B} d_T(x,y)$ define an integer-valued function f on the set of edges $e \in E(T)$, where $f(e) = n_1 n_2$, where n_1 is the number of vertices in $Q \cap B$ and n_2 is the number of vertices in $R \cap B$, where Q and R are the sets of vertices of the two components of $T - e$. Then

$$\sum_{\{x,y\}\subseteq B} d_T(x,y) \leq \sum_{e\in E(T)} f(e).$$

When removing any of the b edges that are incident with the vertices of B, one of the two componets of $T - e$ will have only one vertex, and the other component will have $n + b - 1$ vertices. The vertex in the component of order one is contained in B, while the other component of order $n + 1 - 1$ has $b - 1$ vertices that are in B. So, for each of the b edges incident with a vertex of B we have that $f(e) = 1(b-1) = b-1$. For each of the remaining edges $e \in E(T)$, that is, for edges that are in $e \in E(T) - B = E(H)$, we have that $f(e)$ is maximised if one component of $T - e$ has order $\lfloor \frac{n+b}{2} \rfloor$ and the other component has order $\lceil \frac{n+b}{2} \rceil$. So, one component has $\lfloor \frac{b}{2} \rfloor$ vertices that are in B and the other one has $\lceil \frac{b}{2} \rceil$ vertices that are in B, and so we have $f(e) \leq \lfloor \frac{b}{2} \rfloor \lceil \frac{b}{2} \rceil \leq \frac{b^2}{4}$.

Since T is a tree of order $n + b$ and H is a tree of order n, we have that the size of H is $|E(H)| = n - 1$ and since $E(H) \subseteq E(T)$, we have that $|E(T) - E(H)| = |E(T)| - |E(H)| = n + b - 1 - (n - 1) = b$, we get,

$$\sum_{\{x,y\}\subseteq B} d_T(x,y) \leq \sum_{e\in E(T)} f(e)$$

$$
\begin{aligned}
&= \sum_{e \in E(H)} f(e) + \sum_{e \in E(T)-E(H)} f(e) \\
&\leq (n-1)\frac{b^2}{4} + b(b-1).
\end{aligned}
$$

Adding the three inequalities gives,

$$
\begin{aligned}
W(T) &= W(H) + \sum_{x \in B, y \in V(H)} d_T(x,y) + \sum_{\{x,y\} \subseteq B} d_T(x,y) \\
&\leq \frac{n(n-1)(n+1)}{6} + b\frac{n(n+1)}{2} + (n-1)\frac{b^2}{4} + b(b-1) \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{nb^2}{4} - \frac{b^2}{4} + b^2 - b \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{nb^2}{4} + \frac{3b^2}{4} - b \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{nb^2}{4} + \frac{(1+2)b^2}{4} - b \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{nb^2}{4} + \frac{b^2}{4} + \frac{b^2}{2} - b \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{b^2}{4}(n+1) + \frac{b^2}{2} - b \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{1}{2}b(n+1)\left[n + \frac{1}{2}b\right] + \frac{b^2}{2} - b,
\end{aligned}
$$

as desired. Suppose that

$$
W(T) = \frac{1}{6}n(n-1)(n+1) + \frac{1}{2}b(n+1)\left(n + \frac{1}{2}b\right) + \frac{1}{2}b^2 - b. \tag{4.2}
$$

Then,

$$
\begin{aligned}
W(T) &= \sum_{\{x,y\} \subseteq V(T)} d_T(x,y) \\
&= W(H) + \sum_{x \in B, y \in V(H)} d_T(x,y) + \sum_{\{x,y\} \subseteq B} d_T(x,y).
\end{aligned}
$$

Each of the three sums are maximised, so,

$$
W(H) = \frac{n(n-1)(n+1)}{6},
$$

$$
\sum_{x \in B, y \in V(H)} d_T(x,y) = b\frac{n(n+1)}{2} \tag{4.3}
$$

and for b even, with e an edge in H,

$$
f(e) = \lfloor \frac{b}{2} \rfloor \lceil \frac{b}{2} \rceil = \frac{b^2}{4}. \tag{4.4}
$$

So, by Proposition 4.2.3, since $W(H) = \frac{n(n-1)(n+1)}{6}$, H is a path of order n in T. (4.3) implies that, each of the b vertices that are attached to H are at a distance n from one of the vertices of H, since H is a path of order n, the attached b vertices must be adjacent to any of the end vertices of path H. Since H is a path, and the b vertices are attached to the end vertices of H, (4.4) gives that, one of the components of $T-e$ will always have $\lfloor \frac{b}{2} \rfloor = \frac{b}{2}$ vertices of the b attached vertices and the other component will have $\lceil \frac{b}{2} \rceil = \frac{b}{2}$. So, equality in (4.2) gives that $T = B_{n,b}$ for b even.

Suppose that $T = B_{n,b}$ with b even. Let $P = v_1, v_2, \ldots, v_n$ be the path in $B_{n,b}$ that we attach the b end vertices on, and let T_β be the set of end vertices of order $\left\lfloor \frac{b}{2} \right\rfloor$ attached to v_1 and T_γ be the set of end vertices of order $\left\lceil \frac{b}{2} \right\rceil$ attached to v_n. Then,

$$\begin{aligned} W(T) = &\sum_{\{x,y\}\subseteq T_\beta} d_T(x,y) + \sum_{\{x,y\}\subseteq T_\gamma} d_T(x,y) + \sum_{\{x,y\}\subseteq P} d_T(x,y) + \sum_{x\in T_\gamma, y\in T_\beta} d_T(x,y) \\ &+ \sum_{x\in T_\beta, y\in P} d_T(x,y) + \sum_{x\in P, y\in T_\gamma} d_T(x,y). \end{aligned} \tag{4.5}$$

We calculate each of the six terms in (4.5).
The vertices that are in the set T_γ are a distance 2 from each other, and is the same case for vertices that are in T_β. So,

$$\begin{aligned} \sum_{\{x,y\}\subseteq T_\beta} d_T(x,y) &= 2\Big(\left\lfloor \frac{b}{2} \right\rfloor (\left\lfloor \frac{b}{2} \right\rfloor - 1) - \frac{1}{2}(\left\lfloor \frac{b}{2} \right\rfloor - 1)\left\lfloor \frac{b}{2} \right\rfloor\Big) \\ &= 2\Big(\frac{1}{2}(\left\lfloor \frac{b}{2} \right\rfloor - 1)\left\lfloor \frac{b}{2} \right\rfloor\Big) \\ &= (\left\lfloor \frac{b}{2} \right\rfloor - 1)\left\lfloor \frac{b}{2} \right\rfloor, \end{aligned} \tag{4.6}$$

$$\begin{aligned} \sum_{\{x,y\}\subseteq T_\gamma} d_T(x,y) &= 2\Big(\frac{1}{2}(\left\lceil \frac{b}{2} \right\rceil - 1)\left\lceil \frac{b}{2} \right\rceil\Big) \\ &= (\left\lceil \frac{b}{2} \right\rceil - 1)\left\lceil \frac{b}{2} \right\rceil. \end{aligned} \tag{4.7}$$

Since P is a path of order n in T, by Proposition 4.2.3, we have,

$$\sum_{\{x,y\}\subseteq V(P)} d_T(x,y) = \frac{n(n-1)(n+1)}{6}. \tag{4.8}$$

The vertices that are in T_γ are at distance $n+1$ from the vertices that are in T_β, so,

$$\sum_{x\in T_\gamma, y\in T_\beta} d_T(x,y) = \left\lceil \frac{b}{2} \right\rceil \left\lfloor \frac{b}{2} \right\rfloor (n+1). \tag{4.9}$$

Each vertex in T_γ and T_β forms a path of order $n+1$ with P in T. So,

$$\begin{aligned} \sum_{x\in T_\beta, y\in V(P)} d_T(x,y) &= \left\lfloor \frac{b}{2} \right\rfloor (1+2+\cdots+n) \\ &= \left\lfloor \frac{b}{2} \right\rfloor \frac{n(n+1)}{2}. \end{aligned} \tag{4.10}$$

Similarly,

$$\sum_{x\in T_\gamma, y\in V(P)} d_T(x,y) = \left\lceil \frac{b}{2} \right\rceil (1+2+\cdots+n)$$

$$= \left\lceil \frac{b}{2} \right\rceil \frac{n(n+1)}{2}. \tag{4.11}$$

Plugging (4.6), (4.7), (4.8), (4.9), (4.10) and (4.11) into (4.5), we get,

$$\begin{aligned}
W(T) &= \Big(\left\lfloor \frac{b}{2} \right\rfloor - 1\Big)\left\lfloor \frac{b}{2} \right\rfloor + \Big(\left\lceil \frac{b}{2} \right\rceil - 1\Big)\left\lceil \frac{b}{2} \right\rceil + \frac{n(n-1)(n+1)}{6} + \left\lceil \frac{b}{2} \right\rceil \left\lfloor \frac{b}{2} \right\rfloor (n+1) \\
&\quad + \left\lfloor \frac{b}{2} \right\rfloor \frac{n(n+1)}{2} + \left\lceil \frac{b}{2} \right\rceil \frac{n(n+1)}{2} \\
&= \frac{1}{6}n(n-1)(n+1) + \Big(\left\lfloor \frac{b}{2} \right\rfloor - 1\Big)\left\lfloor \frac{b}{2} \right\rfloor + \Big(\left\lceil \frac{b}{2} \right\rceil - 1\Big)\left\lceil \frac{b}{2} \right\rceil + \left\lceil \frac{b}{2} \right\rceil \left\lfloor \frac{b}{2} \right\rfloor (n+1) \\
&\quad + \frac{n(n+1)}{2}\Big(\left\lfloor \frac{b}{2} \right\rfloor + \left\lceil \frac{b}{2} \right\rceil\Big) \\
&= \frac{1}{6}n(n-1)(n+1) + \Big(\left\lfloor \frac{b}{2} \right\rfloor - 1\Big)\left\lfloor \frac{b}{2} \right\rfloor + \Big(\left\lceil \frac{b}{2} \right\rceil - 1\Big)\left\lceil \frac{b}{2} \right\rceil + \left\lceil \frac{b}{2} \right\rceil \left\lfloor \frac{b}{2} \right\rfloor (n+1) + \frac{n(n+1)}{2}(b) \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \Big(\left\lfloor \frac{b}{2} \right\rfloor - 1\Big)\left\lfloor \frac{b}{2} \right\rfloor + \Big(\left\lceil \frac{b}{2} \right\rceil - 1\Big)\left\lceil \frac{b}{2} \right\rceil + \left\lceil \frac{b}{2} \right\rceil \left\lfloor \frac{b}{2} \right\rfloor (n+1).
\end{aligned} \tag{4.12}$$

For b even, $W(T)$ becomes,

$$\begin{aligned}
W(T) &= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \Big(\frac{b}{2} - 1\Big)\frac{b}{2} + \Big(\frac{b}{2} - 1\Big)\frac{b}{2} + \Big(\frac{b}{2}\Big)\Big(\frac{b}{2}\Big)(n+1) \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{nb(n+1)}{2} + \frac{b^2}{4}(n+1) + 2\Big[\Big(\frac{b}{2} - 1\Big)\frac{b}{2}\Big] \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{1}{2}b(n+1)\Big[n + \frac{b}{2}\Big] + b\Big(\frac{b}{2} - 1\Big) \\
&= \frac{1}{6}n(n-1)(n+1) + \frac{1}{2}b(n+1)\Big[n + \frac{b}{2}\Big] + \frac{b^2}{2} - b.
\end{aligned}$$

Thus, (4.2) is true if and only if $T = B_{n,b}$ for b even. □

Theorem 4.2.5. *Consider a connected graph G with n vertices and independence number α. We have that G has a spanning tree T with*

$$\mu(T) \leq \begin{cases} \alpha + 2, & \text{if } \alpha < \frac{n+1}{2}, \\ \frac{2}{3}\alpha, & \text{if } \alpha \geq \frac{n+1}{2}. \end{cases}$$

Proof: We show that the theorem holds for two cases. The case when $\alpha \geq \frac{n+1}{2}$ or simply, when $n \leq 2\alpha - 1$ and the case when $\alpha < \frac{n+1}{2}$ giving that $n > 2\alpha - 1$.

Case 1: Assume that $n \leq 2\alpha - 1$. For any spanning tree T of G, we have by Proposition 4.2.3 that $W(T) \leq \frac{n(n-1)(n+1)}{6}$. Hence, for any spanning tree T of G,

$$\begin{aligned}
\mu(T) &= \binom{n}{2}^{-1} W(T) \\
&\leq \binom{n}{2}^{-1} \frac{n(n-1)(n+1)}{6} \\
&= \frac{n+1}{3}
\end{aligned}$$

$$\leq \frac{(2\alpha-1)+1}{3}$$
$$= \frac{2}{3}\alpha,$$

as desired.

Case 2: Assume that $n > 2\alpha - 1$. We use the method that was used in Theorem 3.1 of [9], and in particular, we construct a sequence $A_1 \subset A_2 \subset A_3 \subset \ldots$ of independent sets, and a sequence $T_1 \leq T_2 \leq T_3 \leq \ldots$ of subtrees of G such that $A_i \subseteq V(T_i)$ for $i = 1, 2, \ldots$ using Algorithm 3.2.8.

Suppose a subtree $T_i \leq G$ and the set A_i have been obtained using Algorithm 3.2.8.

Let k be the smallest integer such that every vertex not in A_k is within distance one of A_k. Since $A_k \subset V(T_k)$, every vertex x not in T_k is a neighbour of some vertex x' in T_k. Let T be a spanning tree of G with $E(T) = E(T_k) \cup \{xx' \,|\, x \in V(G) - V(T_k)\}$. We show that $\mu(T) \leq \alpha + 2$.

In Lemma 3.2.9, we showed that for every $i = 2, 3, \ldots, k$, $|V(T_i)| \leq |V(T_{i-1})| + 2$, and in Lemma 3.2.13, (3.6), we showed that $|V(T_k)| \leq |V(T_1)| + 2(k-1) = 2k - 1$. Now, denote $|V(T_k)|$ by t. Then $t \leq 2k - 1$. Since, by Proposition 3.2.12, A_k is an independent set, with $k = |A_k| \leq \alpha$. We have,

$$t \leq 2k - 1 \leq 2\alpha - 1. \tag{4.13}$$

We have by the construction that $T \in \mathcal{T}_{t,n-t}$, and since T is obtained by attaching $|V(G) - V(T_k)| = |V(G)| - |V(T_k)| = n - t$ end vertices to a tree T_k, it follows, by Lemma 4.2.4, that

$$W(T) \leq \frac{1}{6}t(t-1)(t+1) + \frac{1}{2}(n-t)(t+1)\left(t + \frac{1}{2}(n-t)\right) + \frac{1}{2}(n-t)^2 - (n-t)$$
$$= \frac{1}{6}t(t-1)(t+1) + \frac{1}{2}(n-t)(t+1)\frac{1}{2}(n+t) + \frac{1}{2}(n-t)^2 - (n-t)$$
$$= \frac{1}{6}t(t-1)(t+1) + \frac{1}{4}(n-t)(t+1)(n+t) + \frac{1}{2}(n-t)^2 - (n-t).$$

Let $f(t) := \frac{1}{6}t(t-1)(t+1) + \frac{1}{4}(n-t)(t+1)(n+t) + \frac{1}{2}(n-t)^2 - (n-t)$. Then

$$f'(t) = \frac{1}{6}(t-1)(t+1) + \frac{1}{6}t(t+1) + \frac{1}{6}t(t-1) - \frac{1}{4}(t+1)(n+t) + \frac{1}{4}(n-t)(n+t) +$$
$$\frac{1}{4}(n-t)(t+1) - (n-t) + 1$$
$$= -\frac{1}{4}t^2 + \frac{1}{2}t + \frac{1}{4}n^2 - n + \frac{5}{6}.$$

Note that $f'(t) = 0$ if and only if $t = 1 \pm n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}}$. Since $\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}} > 0$ for $n \geq 1$, we have, $t = 1 + n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}} > 1$ for $n \geq 1$ and $t = 1 - n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}} < 0$ for $n \geq 1$.
Note that $f''(t) = -\frac{1}{2}t + \frac{1}{2}$.
Observe that $f''(t) = 0$ if and only if $t = 1$.
For $t < 1$, $f''(t) > 0$ and for $t > 1$, $f''(t) < 0$. So, $f(t)$ is concave up for $t < 1$ and $f(t)$ is concave down for $t > 1$, and the inflection point is at $t = 1$. So, $f(t)$

is increasing in the interval $1 - n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}} \leq t \leq 1 + n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}}$, and since $1 - n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}} < 0 < n - 1 < 1 + n\sqrt{1 - \frac{4}{n} + \frac{13}{3n^2}}$, we have that $f(t)$ is increasing in t on the interval $[0, n-1]$. Since, by (4.13), $t \leq 2\alpha - 1 \leq n - 1$, f is maximised for $t = 2\alpha - 1$ to give

$$\begin{aligned}
W(T) &\leq \frac{1}{6}t(t-1)(t+1) + \frac{1}{4}(n-t)(t+1)(n+t) + \frac{1}{2}(n-t)^2 - (n-t) \\
&\leq \frac{1}{6}(2\alpha-1)((2\alpha-1)-1)((2\alpha-1)+1) + \frac{1}{4}(n-(2\alpha-1))((2\alpha-1)+1)(n+(2\alpha-1)) + \\
&\quad \frac{1}{2}(n-(2\alpha-1))^2 - (n-(2\alpha-1)) \\
&= \frac{1}{6}\Big[-4\alpha^3 + 12\alpha^2 + \alpha(3n^2 - 12n + 1) + 3n^2 - 3\Big]
\end{aligned}$$

We divide $W(T)$ by $\binom{n}{2}$ to get an upper bound on $\mu(T)$.

$$\begin{aligned}
\mu(T) &= \binom{n}{2}^{-1} W(T) \\
&\leq \frac{2}{n(n-1)}\Big(\frac{1}{6}\Big[-4\alpha^3 + 12\alpha^2 + \alpha(3n^2 - 12n + 1) + 3n^2 - 3\Big]\Big) \\
&= \alpha + 1 + \frac{1}{n} - \frac{1}{n(n-1)}\Big(4\alpha^2(\frac{1}{3}\alpha - 1) + \frac{1}{3}\alpha(9n+1)\Big).
\end{aligned}$$

Note that $\alpha^2(\frac{1}{3}\alpha - 1) + \frac{1}{3}\alpha(9n+1) > 0$ for $n \geq 1$ and $\alpha \geq 1$, and $0 < \frac{1}{n} \leq 1$ for $n \geq 1$. So,

$$\begin{aligned}
\mu(T) &= \binom{n}{2}^{-1} W(T) \\
&\leq \alpha + 1 + \frac{1}{n} - \frac{1}{n(n-1)}\Big(4\alpha^2(\frac{1}{3}\alpha - 1) + \frac{1}{3}\alpha(9n+1)\Big) \\
&< \alpha + 1 + \frac{1}{n} \\
&\leq \alpha + 1 + 1 \\
&= \alpha + 2.
\end{aligned}$$

and the theorem is proven. □

Remark 4.2.6. In the next corollary, we use the extremal graphs $G_{n,\alpha}$ and $T_{n,\alpha} = B_{2n-2\alpha-1,2\alpha-n+1}$ to show that Theorem 4.2.5 is asymptotically sharp. By asymptotically sharp we mean that, for n large enough and α fixed, there exists a graph G with n vertices, independence number α, such that every spanning tree of G satisfies

$$\mu(T) \geq \begin{cases} \alpha + 2, & \text{if } \alpha < \frac{n+1}{2}, \\ \frac{2}{3}\alpha, & \text{if } \alpha \geq \frac{n+1}{2}. \end{cases}$$

Corollary 4.2.7. *Consider a connected graph G with n vertices, average distance μ and independence number α. We have that,*

$$\mu(G) \leq \begin{cases} \alpha + 2, & \textit{if } \alpha < \frac{n+1}{2}, \\ \frac{2}{3}\alpha, & \textit{if } \alpha \geq \frac{n+1}{2}. \end{cases}$$

The bound is asymptotically sharp for sufficiently large n and fixed α.

Proof: From Theorem 4.2.5, we have that G has a spanning tree T such that

$$\mu(T) \leq \begin{cases} \alpha + 2, & \text{if} \quad \alpha < \frac{n+1}{2}, \\ \frac{2}{3}\alpha, & \text{if} \quad \alpha \geq \frac{n+1}{2}. \end{cases}$$

Let T be a spanning tree of G such that the above inequality holds. Since T is a spanning tree of G, we have, $d_G(x,y) \leq d_T(x,y)$ for any $x, y \in V(G)$, and so $W(G) \leq W(T)$. This implies that, $\mu(G) \leq \mu(T)$. Thus,

$$\mu(G) \leq \mu(T) \leq \begin{cases} \alpha + 2, & \text{if} \quad \alpha < \frac{n+1}{2}, \\ \frac{2}{3}\alpha, & \text{if} \quad \alpha \geq \frac{n+1}{2}. \end{cases}$$

To show that the bound is assymptotically sharp, we consider the graphs (i) $G_{n,\alpha}$ and (ii) $T_{n,\alpha}$.
(i) Let α be a fixed positive integer, and n, $2\alpha - 1 < n$, a sufficiently large integer. Then,

$$\begin{aligned} W(G_{n,\alpha}) = & \sum_{\{x,y\} \subseteq V(P_{2\alpha-2})} d_{G_{n,\alpha}}(x,y) + \sum_{x \in V(K_{\lfloor \frac{n-2\alpha+2}{2} \rfloor}), y \in V(K_{\lceil \frac{n-2\alpha+2}{2} \rceil})} d_{G_{n,\alpha}}(x,y) \\ & + \sum_{\{x,y\} \subseteq V(K_{\lceil \frac{n-2\alpha+2}{2} \rceil})} d_{G_{n,\alpha}}(x,y) + \sum_{\{x,y\} \subseteq V(K_{\lfloor \frac{n-2\alpha+2}{2} \rfloor})} d_{G_{n,\alpha}}(x,y) \\ & + \sum_{x \in V(K_{\lceil \frac{n-2\alpha+2}{2} \rceil}), y \in V(P_{2\alpha-2})} d_{G_{n,\alpha}}(x,y) + \sum_{x \in V(P_{2\alpha-2}), y \in V(K_{\lfloor \frac{n-2\alpha+2}{2} \rfloor})} d_{G_{n,\alpha}}(x,y). \end{aligned} \tag{4.14}$$

From Proposition 4.2.3, we have,

$$\begin{aligned} \sum_{\{x,y\} \subseteq V(P_{2\alpha-2})} d_{G_{n,\alpha}}(x,y) &= \frac{(2\alpha-2)(2\alpha-1)(2\alpha-3)}{6} \\ &= \frac{2(\alpha-1)(2\alpha-1)(2\alpha-3)}{6} \\ &= \frac{(\alpha-1)(2\alpha-1)(2\alpha-3)}{3}. \end{aligned} \tag{4.15}$$

Further,

$$\sum_{x \in V(K_{\lfloor \frac{n-2\alpha+2}{2} \rfloor}), y \in V(K_{\lceil \frac{n-2\alpha+2}{2} \rceil})} d_{G_{n,\alpha}}(x,y) = \lfloor \frac{n-2\alpha+2}{2} \rfloor \lceil \frac{n-2\alpha+2}{2} \rceil (2\alpha-1), \tag{4.16}$$

and

$$\sum_{\{x,y\} \subseteq V(K_{\lfloor \frac{n-2\alpha+2}{2} \rfloor})} d_{G_{n,\alpha}}(x,y) = \frac{(\lfloor \frac{n-2\alpha+2}{2} \rfloor - 1) \lfloor \frac{n-2\alpha+2}{2} \rfloor}{2}. \tag{4.17}$$

Similarly,

$$\sum_{\{x,y\} \subseteq V(K_{\lceil \frac{n-2\alpha+2}{2} \rceil})} d_{G_{n,\alpha}}(x,y) = \frac{(\lceil \frac{n-2\alpha+2}{2} \rceil - 1) \lceil \frac{n-2\alpha+2}{2} \rceil}{2}. \tag{4.18}$$

Also,

$$\sum_{x\in V(P_{2\alpha-2}),y\in V(K_{\lfloor\frac{n-2\alpha+2}{2}\rfloor})} d_{G_{n,\alpha}}(x,y) = \lfloor\frac{n-2\alpha+2}{2}\rfloor[1+2+\cdots+(2\alpha-2)]$$

$$= \lfloor\frac{n-2\alpha+2}{2}\rfloor\frac{(2\alpha-2)(2\alpha-1)}{2}$$

$$= \lfloor\frac{n-2\alpha+2}{2}\rfloor\frac{2(\alpha-1)(2\alpha-1)}{2}$$

$$= \lfloor\frac{n-2\alpha+2}{2}\rfloor(\alpha-1)(2\alpha-1), \qquad (4.19)$$

and

$$\sum_{x\in V(K_{\lceil\frac{n-2\alpha+2}{2}\rceil}),y\in V(P_{2\alpha-2})} d_{G_{n,\alpha}}(x,y) = \lceil\frac{n-2\alpha+2}{2}\rceil[1+2+\cdots+(2\alpha-2)]$$

$$= \lceil\frac{n-2\alpha+2}{2}\rceil(\alpha-1)(2\alpha-1). \qquad (4.20)$$

Plugging (4.15), (4.16), (4.17), (4.18), (4.19), (4.20), into (4.14), we get,

$$W(G_{n,\alpha}) = \frac{(\alpha-1)(2\alpha-1)(2\alpha-3)}{3} + \lfloor\frac{n-2\alpha+2}{2}\rfloor\lceil\frac{n-2\alpha+2}{2}\rceil(2\alpha-1)$$
$$+ \frac{(\lfloor\frac{n-2\alpha+2}{2}\rfloor-1)\lfloor\frac{n-2\alpha+2}{2}\rfloor}{2} + \frac{(\lceil\frac{n-2\alpha+2}{2}\rceil-1)\lceil\frac{n-2\alpha+2}{2}\rceil}{2}$$
$$+ \lfloor\frac{n-2\alpha+2}{2}\rfloor(\alpha-1)(2\alpha-1) + \lceil\frac{n-2\alpha+2}{2}\rceil(\alpha-1)(2\alpha-1).$$

This, after simplification, gives

$$W(G_{n,\alpha}) = (\alpha-1)(2\alpha-1)\Big[n-\frac{4\alpha}{3}+1\Big] + \lfloor\frac{n-2\alpha+2}{2}\rfloor\lceil\frac{n-2\alpha+2}{2}\rceil(2\alpha-1)$$
$$+ \frac{1}{2}\Big(\lfloor\frac{n-2\alpha+2}{2}\rfloor^2 + \lceil\frac{n-2\alpha+2}{2}\rceil^2\Big) - \frac{1}{2}(n-2\alpha+2).$$

By the definition of floor functions and ceiling functions, we have that,

$$\lfloor\frac{n-2\alpha+2}{2}\rfloor \le \frac{n-2\alpha+2}{2} < \lfloor\frac{n-2\alpha+2}{2}\rfloor + 1, \qquad (4.21)$$

and

$$\lceil\frac{n-2\alpha+2}{2}\rceil - 1 < \frac{n-2\alpha+2}{2} \le \lceil\frac{n-2\alpha+2}{2}\rceil. \qquad (4.22)$$

From (4.21), we have,

$$\frac{n-2\alpha+2}{2} < \lfloor\frac{n-2\alpha+2}{2}\rfloor + 1.$$

This implies that,

$$\lfloor\frac{n-2\alpha+2}{2}\rfloor > \frac{n-2\alpha+2}{2} - 1.$$

From (4.22), we have,

$$\lceil\frac{n-2\alpha+2}{2}\rceil \ge \frac{n-2\alpha+2}{2}.$$

Using the inequalities that follow from (4.21) and (4.22) to simplify $W(G_{n,\alpha})$, we see that, the terms that have floor and ceiling functions give the following inequalities,

$$\lfloor\frac{n-2\alpha+2}{2}\rfloor\lceil\frac{n-2\alpha+2}{2}\rceil(2\alpha-1) > \Big(\frac{n-2\alpha+2}{2}\Big)\Big(\frac{n-2\alpha+2}{2}-1\Big)(2\alpha-1)$$

$$= \Big[\Big(\frac{n-2\alpha+2}{2}\Big)^2 - \frac{n-2\alpha+2}{2}\Big](2\alpha-1),$$

and

$$\begin{aligned}
\lfloor\frac{n-2\alpha+2}{2}\rfloor^2 + \lceil\frac{n-2\alpha+2}{2}\rceil^2 &> \Big(\frac{n-2\alpha+2}{2}\Big)^2 + \Big(\frac{n-2\alpha+2}{2}-1\Big)^2 \\
&= \Big(\frac{n-2\alpha+2}{2}\Big)^2 + \Big(\frac{n-2\alpha+2}{2}\Big)^2 - (n-2\alpha+2)+1 \\
&= 2\Big(\frac{n-2\alpha+2}{2}\Big)^2 - (n-2\alpha+2)+1.
\end{aligned}$$

Using the above inequalities in $W(G_{n,\alpha})$, we see that,

$$\begin{aligned}
W(G_{n,\alpha}) &= (\alpha-1)(2\alpha-1)\Big[n-\frac{4\alpha}{3}+1\Big] + \lfloor\frac{n-2\alpha+2}{2}\rfloor\lceil\frac{n-2\alpha+2}{2}\rceil(2\alpha-1) \\
&\quad + \frac{1}{2}\Big(\lfloor\frac{n-2\alpha+2}{2}\rfloor^2 + \lceil\frac{n-2\alpha+2}{2}\rceil^2\Big) - \frac{1}{2}(n-2\alpha+2) \\
&> (\alpha-1)(2\alpha-1)\Big[n-\frac{4\alpha}{3}+1\Big] + \Big[\Big(\frac{n-2\alpha+2}{2}\Big)^2 - \frac{n-2\alpha+2}{2}\Big](2\alpha-1) \\
&\quad + \frac{1}{2}\Big[2\Big(\frac{n-2\alpha+2}{2}\Big)^2 - (n-2\alpha+2)+1\Big] - \frac{1}{2}(n-2\alpha+2) \\
&= \Big(\frac{n-2\alpha+2}{2}\Big)^2(2\alpha-1) + \Big(\frac{n-2\alpha+2}{2}\Big)^2 + (\alpha-1)(2\alpha-1)\Big[n-\frac{4\alpha}{3}+1\Big] \\
&\quad - \Big(\frac{n-2\alpha+2}{2}\Big)(2\alpha-1) - \frac{1}{2}(n-2\alpha+2) + \frac{1}{2} - \frac{1}{2}(n-2\alpha+2) \\
&= \Big(\frac{n-(2\alpha-2)}{2}\Big)^2(2\alpha) + (2\alpha-1)\Big[(\alpha-1)\Big(n-\frac{4\alpha}{3}+1\Big) - \frac{n-2\alpha+2}{2}\Big] \\
&\quad - (n-2\alpha+2) + \frac{1}{2} \\
&= \frac{1}{4}(n^2 - 2(2\alpha-2)n + (2\alpha-2)^2)(2\alpha) + (2\alpha-1)\Big[(\alpha-1)\Big(n-\frac{4\alpha}{3}+1\Big) - \frac{n-2\alpha+2}{2}\Big] \\
&\quad - (n-2\alpha+2) + \frac{1}{2} \\
&= \frac{1}{2}\alpha n^2 + (2\alpha-1)\Big[(\alpha-1)\Big(n-\frac{4\alpha}{3}+1\Big) - \frac{n-2\alpha+2}{2}\Big] - 2(\alpha-1)\alpha n \\
&\quad - (n-2\alpha+2) + \frac{1}{2} + 2(\alpha-1)^2\alpha \\
&= \frac{1}{2}\alpha n^2 + \Big[(2\alpha-1)(\alpha-1) - 2(\alpha-1)\alpha - \frac{2\alpha+1}{2}\Big]n \\
&\quad + (\alpha-1)\Big[(2\alpha-1)\Big(1-\frac{4\alpha}{3}\Big) + (2\alpha-1) + 2(\alpha-1)\alpha + 2\Big] + \frac{1}{2}. \\
&= \frac{1}{2}\alpha n^2 + \Big[(\alpha-1)\Big((2\alpha-1)-2\alpha\Big) - \alpha - \frac{1}{2}\Big]n \\
&\quad + (\alpha-1)\Big[(2\alpha-1)\Big(1-\frac{4\alpha}{3}+1\Big) + 2(\alpha-1)\alpha + 2\Big] + \frac{1}{2}. \\
&= \frac{1}{2}\alpha n^2 + \Big[-(\alpha-1) - \alpha - \frac{1}{2}\Big]n + (\alpha-1)\Big[(2\alpha-1)\Big(2-\frac{4\alpha}{3}\Big) + 2(\alpha-1)\alpha + 2\Big] + \frac{1}{2}. \\
&= \frac{1}{2}\alpha n^2 + \Big(\frac{1}{2}-2\alpha\Big)n + (\alpha-1)\Big[(2\alpha-1)\Big(2-\frac{4\alpha}{3}\Big) + 2(\alpha-1)\alpha + 2\Big] + \frac{1}{2}. \\
&= \frac{1}{2}\alpha n^2 + \Big(\frac{1}{2}-2\alpha\Big)n + \frac{2(\alpha-1)}{3}\Big[5\alpha-\alpha^2\Big] + \frac{1}{2}.
\end{aligned}$$

Thus, the average distance of $G_{n,\alpha}$ is,

$$\mu(G_{n,\alpha}) = \binom{n}{2}^{-1} W(G_{n,\alpha})$$

$$
\begin{aligned}
&> \frac{2}{n(n-1)}\Big(\frac{1}{2}\alpha n^2 + \Big(\frac{1}{2} - 2\alpha\Big)n + \frac{2(\alpha-1)}{3}\Big[5\alpha - \alpha^2\Big] + \frac{1}{2}\Big) \\
&= \frac{\alpha n}{n-1} + \frac{2}{n-1}\Big[\frac{1}{2} - 2\alpha\Big] + \frac{2}{n(n-1)}\Big[\frac{2(\alpha-1)}{3}\Big[5\alpha - \alpha^2\Big] + \frac{1}{2}\Big] \\
&> \alpha + \frac{2}{n}\Big[\frac{1}{2} - 2\alpha\Big] + \frac{2}{n^2}\Big[\frac{2(\alpha-1)}{3}\Big[5\alpha - \alpha^2\Big] + \frac{1}{2}\Big]
\end{aligned}
$$

So, we see that $\mu(G_{n,\alpha}) > \alpha + C$ for some constant C, and thus, we have, for any spanning tree T of $G_{n,\alpha}$, $\mu(T) \geq \mu(G_{n,\alpha}) > \alpha + C$.

(ii) Let α and n be positive integers such that $2\alpha - 1 \geq n$, α fixed. Then, for $n = 2\alpha - 1$, in $T_{n,\alpha}$, $P = v_1, v_2, \ldots, v_{2(2\alpha-1)-2\alpha-1} = v_1, v_2, \ldots, v_{2\alpha-3}$, and we attach $\lfloor \frac{2\alpha-(2\alpha-1)+1}{2} \rfloor = \lfloor \frac{2}{2} \rfloor = 1$ vertices to v_1 and $\lceil \frac{2\alpha-(2\alpha-1)+1}{2} \rceil = \lceil \frac{2}{2} \rceil = 1$ vertices to $v_{2\alpha-3}$. So, for $n = 2\alpha - 1$, $T_{n,\alpha}$ is a path of order $2\alpha - 1$, and we have, by Proposition 4.2.3, for $n = 2\alpha - 1$,

$$
\begin{aligned}
W(T_{n,\alpha}) &= \frac{n(n-1)(n+1)}{6} \\
&= \frac{(2\alpha-1)(2\alpha-2)(2\alpha)}{6} \\
&= \frac{2(2\alpha-1)(\alpha-1)(\alpha)}{3}.
\end{aligned}
$$

And so,

$$
\begin{aligned}
\mu(T_{n,\alpha}) &= \binom{n}{2}^{-1} \frac{2(2\alpha-1)(\alpha-1)(\alpha)}{3} \\
&= \frac{2}{n(n-1)} \frac{2(2\alpha-1)(\alpha-1)(\alpha)}{3} \\
&= \frac{2}{(2\alpha-1)(2\alpha-2)} \frac{2(2\alpha-1)(\alpha-1)(\alpha)}{3} \\
&= \frac{2}{2(2\alpha-1)(\alpha-1)} \frac{2(2\alpha-1)(\alpha-1)(\alpha)}{3} \\
&= \frac{2}{3}\alpha.
\end{aligned}
$$

Now, suppose $n < 2\alpha - 1$. We have that $T_{n,\alpha} = B_{2n-2\alpha-1,2\alpha-n+1}$. So, using (4.12), we have that, n from (4.12) shifts to become $2n - 2\alpha - 1$ in this case and b shifts to become $2\alpha - n + 1$ in this case, as we have that $T = B_{n,b}$ in (4.12). Solving for $W(T_{n,\alpha})$, we get,

$$
\begin{aligned}
W(T_{n,\alpha}) &= \frac{1}{6}(2n-2\alpha-1)(2n-2\alpha-2)(2n-2\alpha) + \frac{(2n-2\alpha-1)(2\alpha-n+1)(2n-2\alpha)}{2} \\
&\quad + \Big[\Big\lfloor \frac{2\alpha-n+1}{2} \Big\rfloor - 1\Big]\Big\lfloor \frac{2\alpha-n+1}{2} \Big\rfloor + \Big[\Big\lceil \frac{2\alpha-n+1}{2} \Big\rceil - 1\Big]\Big\lceil \frac{2\alpha-n+1}{2} \Big\rceil \\
&\quad + \Big\lceil \frac{2\alpha-n+1}{2} \Big\rceil \Big\lfloor \frac{2\alpha-n+1}{2} \Big\rfloor (2n-2\alpha) \\
&> \frac{1}{6}(2n-2\alpha-1)(2n-2\alpha-2)(2n-2\alpha) + \frac{(2n-2\alpha-1)(2\alpha-n+1)(2n-2\alpha)}{2} \\
&\quad + \Big[\frac{2\alpha-n+1}{2} - 2\Big]\Big[\frac{2\alpha-n+1}{2} - 1\Big] + \Big[\frac{2\alpha-n+1}{2} - 1\Big]\Big[\frac{2\alpha-n+1}{2}\Big] \\
&\quad + \Big[\frac{2\alpha-n+1}{2}\Big]\Big[\frac{2\alpha-n+1}{2} - 1\Big](2n-2\alpha) \\
&= \frac{2}{3}(2n-2\alpha-1)(n-\alpha-1)(n-\alpha) + (2n-2\alpha-1)(2\alpha-n+1)(n-\alpha)
\end{aligned}
$$

$$
\begin{aligned}
&+\Big[\frac{2\alpha-n+1}{2}-1\Big]\Big[\frac{2\alpha-n+1}{2}-2+\frac{2\alpha-n+1}{2}\Big]\\
&+\Big[\frac{2\alpha-n+1}{2}-1\Big](2\alpha-n+1)(n-\alpha)\\
&=(2n-2\alpha-1)(n-\alpha)\Big[\frac{2}{3}(n-\alpha-1)+2\alpha-n+1\Big]\\
&+\frac{1}{2}(2\alpha-n-1)(2\alpha-n-1)+\frac{1}{2}(2\alpha-n-1)\Big[(3\alpha+1)n-n^2-2\alpha^2-\alpha\Big]\\
&=\frac{1}{3}(2n^2-2n\alpha-n-2\alpha n+2\alpha^2+\alpha)\Big[2n-2\alpha-2+6\alpha-3n+3\Big]\\
&+\frac{1}{2}(2\alpha-n-1)\Big[(3\alpha+1)n-n^2-2\alpha^2-\alpha+2\alpha-n-1\Big]\\
&=\frac{1}{3}\Big[2n^2-(4\alpha+1)n+(2\alpha^2+\alpha)\Big]\Big[4\alpha-n+1\Big]\\
&+\frac{1}{2}((2\alpha-1)-n)\Big[3\alpha n-n^2+(-2\alpha^2+\alpha-1)\Big]\\
&=\frac{1}{3}\Big[-2n^3+(4\alpha+1)n^2-(2\alpha^2+\alpha)n+2n^2(4\alpha+1)-(4\alpha+1)(4\alpha+1)n\\
&+(2\alpha^2+\alpha)(4\alpha+1)\Big]+\frac{1}{2}\Big[-3\alpha n^2+n^3-(-2\alpha^2+\alpha-1)n+3\alpha(2\alpha-1)n\\
&-n^2(2\alpha-1)+(-2\alpha^2+\alpha-1)(2\alpha-1)\Big]\\
&=\frac{1}{3}\Big[-2n^3+3(4\alpha+1)n^2-[(2\alpha^2+\alpha)+(4\alpha+1)^2]n+(2\alpha^2+\alpha)(4\alpha+1)\Big]\\
&+\frac{1}{2}\Big[-(3\alpha+(2\alpha-1))n^2+n^3+[(2\alpha^2-\alpha+1)+3\alpha(2\alpha-1)]n\\
&+(-2\alpha^2+\alpha-1)(2\alpha-1)\Big]\\
&=\Big[\frac{1}{2}-\frac{2}{3}\Big]n^3+\Big[\frac{3(4\alpha+1)}{3}-\frac{5\alpha-1}{2}\Big]n^2\\
&+\Big[\frac{(2\alpha^2-\alpha+1)+3\alpha(2\alpha-1)}{2}-\frac{(2\alpha^2+\alpha)+(4\alpha+1)^2}{3}\Big]n\\
&+\frac{1}{3}(2\alpha^2+\alpha)(4\alpha+1)+\frac{1}{2}(\alpha-2\alpha^2-1)(2\alpha-1)\\
&=-\frac{1}{6}n^3+\Big[\frac{3\alpha}{2}+\frac{3}{2}\Big]n^2+\Big[\frac{8\alpha^2-4\alpha+1}{2}-\frac{(18\alpha^2+9\alpha+1)}{3}\Big]n\\
&+\frac{1}{3}(2\alpha^2+\alpha)(4\alpha+1)+\frac{1}{2}(\alpha-2\alpha^2-1)(2\alpha-1)\\
&=-\frac{1}{6}n^3+\frac{3}{2}(\alpha+1)n^2+\Big[\frac{-12\alpha^2-30\alpha+1}{6}\Big]n+\frac{1}{3}(2\alpha^2+\alpha)(4\alpha+1)\\
&+\frac{1}{2}(\alpha-2\alpha^2-1)(2\alpha-1).
\end{aligned}
$$

So, for sufficiently large n, we have,

$$
\begin{aligned}
\mu(T_{n,\alpha})&=\frac{2}{n(n-1)}W(T_{n,\alpha})\\
&>\frac{2}{n(n-1)}\Big[-\frac{1}{6}n^3+\frac{3}{2}(\alpha+1)n^2+\Big[\frac{-12\alpha^2-30\alpha+1}{6}\Big]n\\
&+\frac{1}{3}(2\alpha^2+\alpha)(4\alpha+1)+\frac{1}{2}(\alpha-2\alpha^2-1)(2\alpha-1)\Big]
\end{aligned}
$$

$$\begin{aligned}
&> \frac{2}{n^2}\Big[-\frac{1}{6}n^3+\frac{3}{2}(\alpha+1)n^2+\Big[\frac{-12\alpha^2-30\alpha+1}{6}\Big]n+\frac{1}{3}(2\alpha^2+\alpha)(4\alpha+1)\\
&\quad+\frac{1}{2}(\alpha-2\alpha^2-1)(2\alpha-1)\Big]\\
&= -\frac{1}{3}n+3(\alpha+1)+\Big[\frac{-12\alpha^2-30\alpha+1}{3n}\Big]+\frac{2}{3n^2}(2\alpha^2+\alpha)(4\alpha+1)\\
&\quad+\frac{1}{n^2}(\alpha-2\alpha^2-1)(2\alpha-1)\Big]\\
&= -\frac{1}{3}n+3(\alpha+1).
\end{aligned}$$

Since $n < 2\alpha - 1$, we have that, $-n > 1 - 2\alpha$, give that,

$$\begin{aligned}
\mu(T_{n,\alpha}) &> -\frac{1}{3}n+3\alpha+3\\
&> \frac{1}{3}(1-2\alpha)+3\alpha+3\\
&= \frac{7\alpha}{3}+\frac{10}{3}\\
&> \frac{2}{3}\alpha.
\end{aligned}$$

□

4.3 Conclusion

In this chapter, we were concerned with proving one of the results that give upper bounds on average distance in terms of order of the graph and independence number of the graph. The main result of this chapter, given by Corollary 4.2.7 and stated as
"Let G be a connected graph of order n with average distance μ and independence number α. Then

$$\mu(G) \le \begin{cases} \alpha+2, & \text{if } \alpha < \frac{n+1}{2},\\ \frac{2}{3}\alpha, & \text{if } \alpha \ge \frac{n+1}{2}, \end{cases}$$

and this bound, for sufficiently large n and fixed α, is asymptotically sharp." follows directly from Theorem 4.2.5, which is stated as
"Let G be a connected graph of order n and independence number α. Then G has a spanning tree T with

$$\mu(T) \le \begin{cases} \alpha+2, & \text{if } \alpha < \frac{n+1}{2},\\ \frac{2}{3}\alpha, & \text{if } \alpha \ge \frac{n+1}{2}." \end{cases}$$

In the next chapter, we turn to studying bounds on average distance in terms of chromatic number.

Chapter 5

Upper bound on the average distance in terms of order and chromatic number

5.1 Introduction

In this chapter, we will present results on bounds on average distance in terms of order and chromatic number of the graph. The results of this chapter are from a 1989 study, presented by Tomescu and Melter, in [24]. Some of the results will be proven using methods from [24]. To analyse results of this chapter, we need the following definitions and notation:

Definition 5.1.1. A graph G is a *complete k-partite* graph if $V(G)$ can be partitioned into k disjoint subsets $V_1, V_2, \ldots, V_k$ with $uv \in E(G)$ if $u \in V_i$ and $v \in V_j$, where $1 \leq i, j \leq k$ and $i \neq j$.

Definition 5.1.2. For $n \geq k \geq 2$, $H_{n,k}$ denotes a connected k-chromatic graph of order n made of a K_k graph, and a path P of order $n-k$ with end vertices u and v, with u joined to some vertex w of K_k. We will call v a base vertex of $H_{n,k}$.

We demonstrate Definition 5.1.2 by a way of example.

Example 5.1.3. Consider a graph $H_{10,5}$. We have that $H_{10,5}$ has a complete subgraph of order 5 with a vertex w that joins a u-v path $P_{10-5} = P_5$ to K_5. We draw the graph of $H_{10,5}$ below,

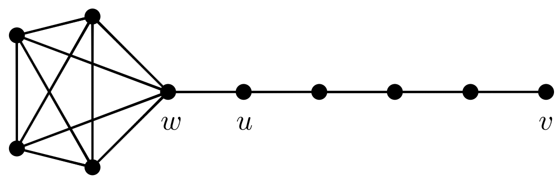


Figure 5.1: The graph of $H_{10,5}$.

Notation 5.1.4. $T(n,k)$ denotes a family of k-partite complete graphs with order n and maximum number of edges, without $(k+1)$-cliques for $1 \leq k \leq n-1$. $T(n,k)$ is constructed as follows: let T be any graph in $T(n,k)$. For a given n and k, $V(T)$ consists of k disjoint classes, with b classes having $q'+1$ vertices and $k-b$ classes containing q' vertices, where q' and b are integers such that $n = kq' + b$ with $0 \leq b < k$; two vertices in T are adjacent if and only if they belong to two different classes of $V(T)$.

We illustrate Notation 5.1.4 by a way of example:

Example 5.1.5. Suppose $G \in T(n,k)$ has order $n = 7$ and $k = 4$. Then we have that

$$n = kq' + b = 4(1) + 3 \text{ thus, } b = 3 \text{ and } q' = 1.$$

Of the $k = 4$ classes, $b = 3$ classes have order $q' + 1 = 2$, while $k - b = 4 - 3 = 1$ classes have order $q' = 1$.

Let the $b = 3$ classes be given by the sets $\{\{v_1, v_2\}, \{u_1, u_2\}, \{w_1, w_2\}\}$ and the $k - b = 1$ class be the set $\{s\}$. Then G is the graph given in Figure 5.2.

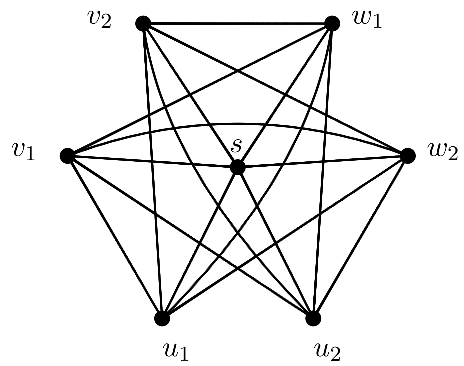


Figure 5.2: A graph in T(7,4).

5.2 Average distance, order and chromatic number

To analyse $\sigma(x)$, we will use the idea of *distance layers* defined as follows:

Definition 5.2.1. Let G be a connected graph and let x be a vertex in G. The set of vertices that are a distance i from a vertex x in G is called the *i-th distance layer* and is denoted by $N_i(x)$.

Proposition 5.2.2. *Let G be a connected graph of order n, and let x be a vertex in G. For each integer $i = 0, 1, \dots, e_G(x)$, let $N_i(x) = \{y \,|\, d_G(x,y) = i\}$ and let $k_i = |N_i(x)|$. (If vertex x is clear, we simply write k_i and N_i).*
Then

$$n = k_0 + k_1 + k_2 + \cdots + k_{e_G(x)}, \tag{5.1}$$

and

$$\sigma(x) = k_1 + 2k_2 + 3k_3 + \cdots + e_G(x)k_{e_G(x)}. \tag{5.2}$$

Proof:
Since for each vertex $y \in V(G)$, we have, $d_G(x,y) = i$ for some $i = 0, 1, \dots, e_G(x)$, we can partition vertices of G into sets of vertices that are at distance i from x, $N_i(x)$, giving that $V(G) = \cup_{i=0}^{e_G(x)} N_i(x)$. Thus, since $N_i(x) \cap N_j(x) = \emptyset$ for any $i \neq j$, we have that,

$$\begin{aligned} n &= |V(G)| \\ &= |\bigcup_{i=0}^{e_G(x)} N_i(x)| \end{aligned}$$

$$= \sum_{i=0}^{e_G(x)} |N_i(x)|$$
$$= k_0 + k_1 + k_2 + \cdots + k_{e_G(x)}.$$

Also,

$$\begin{aligned}\sigma(x) &= \sum_{y \in V(G)} d_G(x,y)\\ &= \sum_{i=0}^{e_G(x)} |N_i(x)| i\\ &= k_1 + 2k_2 + 3k_3 + \cdots + e_G(x) k_{e_G(x)},\end{aligned}$$

as desired.

Lemma 5.2.3. *Let G be a connected graph of order n and diameter d. Let $P = v_0, v_1, \ldots, v_d$ be a diametral path in G. The number of colors needed to color P in G is 2.*

Proof: Since vertices of P can be partitioned into two independent sets $A = \{v_i \,|\, i \text{ is even}, i \leq d\}$ and $B = \{v_i \,|\, i \text{ is odd}, i \leq d\}$, we have that, vertices in A can be colored with color 1, and vertices in B can be colored with color 2. Thus, Lemma 5.2.3 holds. □
The following facts and properties are going to be used to prove the main results of this chapter.

Proposition 5.2.4. *Let G be a connected graph of order n and diameter d. Then*

$$\chi(G) \leq n - d + 1.$$

Proof: Let $P = v_0, v_1, \ldots, v_d$ be a diametral path in G. Since by Lemma 5.2.3, vertices in the diametral path P can be colored with 2 colors, and the vertices in $V(G) - V(P)$ can be colored by, at most, $n - |V(P)|$ colors, we have that $\chi(G) \leq n - |V(P)| + 2 = n - (d+1) + 2 = n - d + 1$.□

We use the graph $H_{n,k}$, defined in Definition 5.1.2, to analyse the extremal graphs of the Property 5.2.5 and Property 5.2.10.

Property 5.2.5. *Let G be a connected k-chromatic graph of order n. If $x \in V(G)$ and $\chi(G - x) = k$, then*

$$\sigma(x) \leq 1 + 2 + \cdots + (n-k) + (k-1)(n-k+1) \tag{5.3}$$

and equality holds if and only if G is isomorphic to $H_{n,k}$ and $x = v$ is the base vertex of $H_{n,k}$.

Proof: Let G be a connected k-chromatic graph of order n. Suppose that $x \in V(G)$ and $\chi(G - x) = k$. Consider distance layers from x, $N_i(x)$. Since G is connected, $k_i \geq 1$ for $i = 0, 1, \ldots, e_G(x)$. Subject to this , (5.1), (5.2) is maximised for $k_1 = k_2 = \cdots = k_{e_G(x)-1} = 1$ and

$$\begin{aligned}k_{e_G(x)} &= n - (\sum_{i=1}^{e_G(x)-1} k_i + k_0)\\ &= n - (\sum_{i=1}^{e_G(x)-1} 1 + 1)\end{aligned}$$

$$= n - (e_G(x) - 1 + 1)$$
$$= n - e_G(x),$$

to give

$$\sigma(x) \leq 1 + 2 + \cdots + (e_G(x) - 1) + (n - e_G(x))e_G(x).$$

But because, by Proposition 5.2.4, $e_G(x) \leq diam(G) \leq n - k + 1$, we have,

$$\begin{aligned}\sigma(x) &\leq 1 + 2 + \cdots + (e_G(x) - 1) + (n - e_G(x))e_G(x)\\ &\leq 1 + 2 + \cdots + (n - k + 1 - 1) + [n - (n - k + 1)](n - k + 1)\\ &= 1 + 2 + \cdots + (n - k) + [k - 1](n - k + 1). \qquad (5.3)\end{aligned}$$

Now assume that there is equality in (5.3). Then every distance layer $N_i(x)$, $i = 0, 1, \ldots, e_G(x) - 1$, has 1 vertex and $k_{e_G(x)} = n - e_G(x) = n - (n - k + 1) = k - 1$.

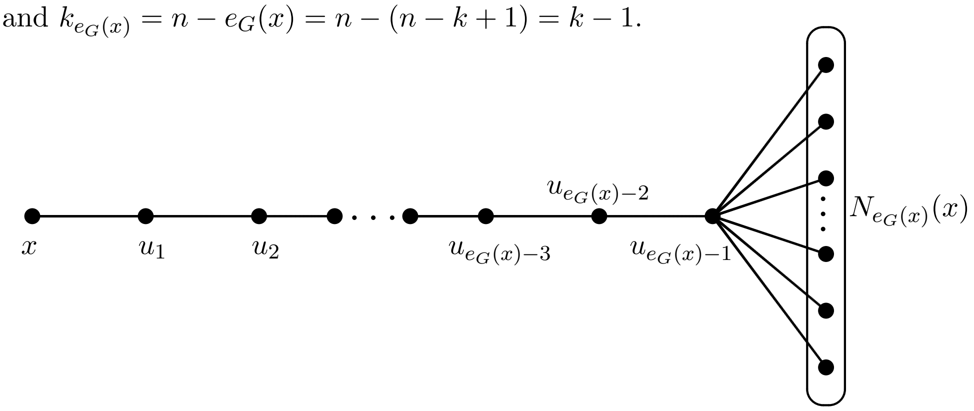


Figure 5.3: A subgraph of a graph G that is obtained when equality holds in (5.3).

By properties of distance layers, if u_i, $i = 1, 2, \ldots, e_G(x) - 1$, is the only vertex in $N_i(x)$, then $xu_1u_2 \ldots u_{e_G(x)-1}$ is a path and every vertex in $N_{e_G(x)}(x)$ is adjacent to $u_{e_G(x)-1}$.
To show that the graph is $H_{n,k}$, it is adequate to prove that any two vertices in $N_{e_G(x)}(x)$ are adjacent.
We proceed by contradiction: Suppose to the contrary that $w, z \in N_{e_G(x)}(x)$ are such that $wz \notin E(G)$.

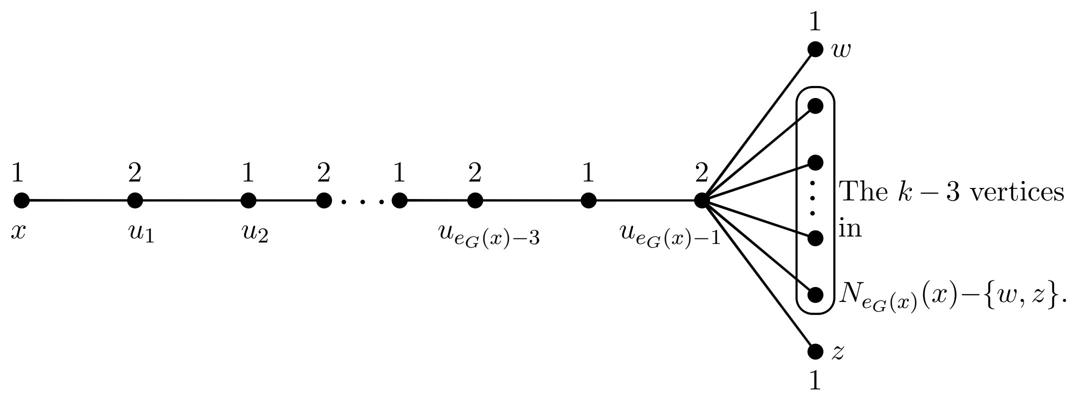


Figure 5.4: Coloring of the graph that is in Figure 5.3, for when $wz \notin E(G)$.

Then colour w and z using the same color 1, say. Using only two colors, 1 and 2, the vertices $w, z, x, u_1, u_2, \ldots, u_{e_G(x)-1}$ can be properly colored.
Since $N_{e_G(x)}$ has $k-1$ vertices, the remaining $k-3$ vertices in $N_{e_G(x)}$ can be colored using at most $k-3$ colors. Hence, $\chi(G) \leq 2+k-3 = k-1$, contradicting that $\chi(G) = k$. Thus, any two vertices in $N_{e_G(x)}(x)$ are adjacent, giving that G is isomorphic to $H_{n,k}$.

Suppose G is isomorphic to $H_{n,k}$ and $x = v$. Since by Definition 5.1.2, vertex u in the $v-u$ path, P_{n-k}, is joined to vertex w of the clique K_k, we have that v is a distance $n-k$ from w, and so, $V(P_{n-k}) \cup \{w\}$ form a $u-w$ path of order $n-k+1$ in G, and also, v is at distance $n-k+1$ from the vertices in the set $V(K_k) - \{w\}$. So,

$$\begin{aligned}\sigma(x) &= \sum_{y \in V(G)} d_G(x,y) \\ &= \sum_{y \in V(P_{n-k}) \cup w} d_G(x,y) + \sum_{y \in V(K_k) - w} d_G(x,y) \\ &= 1+2+\cdots+(n-k)+(k-1)(n-k+1).\end{aligned}$$

Thus, equality in (5.3) holds if and only if G is isomorphic to $H_{n,k}$ and $x = v$. □

Remark 5.2.6. Recall, in Definition 5.1.2, we have that $H_{n,k}$ is obtained by joining a complete graph K_k and a path P_{n-k} together, by an edge uw, where u is an end vertex in P_{n-k} and w is some vertex in K_k.

Now, suppose in $H_{n,k}$, for some vertex $x \in V(K_k)$ with $x \neq w$, we add edges to $H_{n,k}$ such that each edge joins u to any vertex $z \in V(K_k) - x$. Then we see that, the set of edges that have been added to $H_{n,k}$ can be expressed by the set $\{uz \,|\, z \in V(K_k), z \neq x\}$.

We demonstrate the above remark by a way of example.

Example 5.2.7. Consider graph $H_{10,5}$ as given in Example 5.1.3, we have that $H_{10,5}$ is the graph:

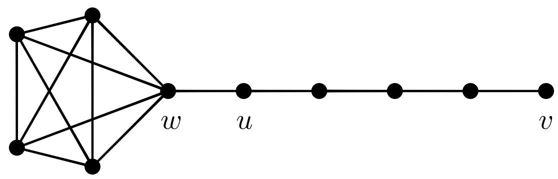


Figure 5.5: The graph of $H_{10,5}$.

Now, suppose we choose some vertex in $V(K_5) - w$ to be x, then we have the graph,

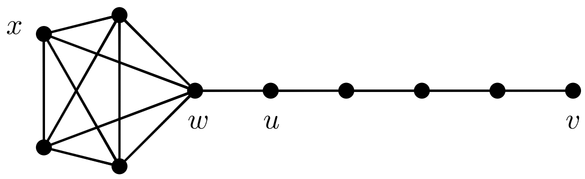
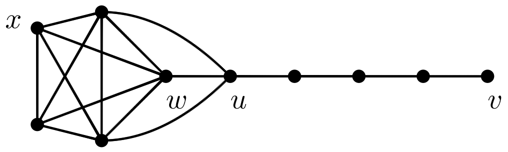
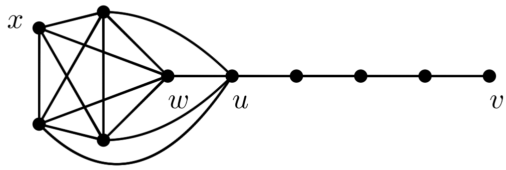


Figure 5.6: The graph of $H_{10,5}$ with a vertex x in K_5 subgraph.

Now, add two and three edges that belong to the set $\{uz \,|\, z \in V(K_5), z \neq x\}$ to $H_{10,5}$,



Adding two edges from the set $\{uz \mid z \in V(K_5), z \neq x\}$ to $H_{10,5}$.



Adding three edges from the set $\{uz \mid z \in V(K_5), z \neq x\}$ to $H_{10,5}$.

Figure 5.7: Two graphs of $H_{10,5}$ with a vertex x in K_5 clique with two and three edges in the set $\{uz \mid z \in V(K_5), z \neq x\}$ added to each graph respectively.

We see that, at most, we can add three edges from the set $\{uz \mid z \in V(K_5), z \neq x\}$ to $H_{n,k}$.

Lemma 5.2.8. *Let G be a connected k-chromatic graph of order n. Suppose that $x \in V(G)$ and $\chi(G-x) = k-1$. Then $deg_G(x) \geq k-1$.*

Proof: Suppose $x \in V(G)$ and $\chi(G-x) = k-1$ with $deg_G(x) < k-1$. Then, we have that $deg_G(x) \leq k-2$. Since $G-x$ is $k-1$ colourable and $deg_G(x) \leq k-2$, there is an extra colour, among the $k-1$ colours used to colour $G-x$, which can be used to colour x. Therefore, G is $k-1$ colourable, a contradiction. Thus, Lemma 5.2.8 holds. □

Remark 5.2.9. In the next results, Property 5.2.10, the extremal graph is given by the graphs described in Remark 5.2.6, which, in general, can be demonstrated similarly to Example 5.2.7.

Property 5.2.10. *Let G be a connected k-chromatic graph of order n. If $x \in V(G)$ and $\chi(G-x) = k-1$, then*

$$\sigma(x) \leq (k-1) + 2 + \cdots + (n-k+1) \tag{5.4}$$

and equality holds if and only if G is isomorphic to a graph obtained from $H_{n,k}$ by adding the edges from a set E_n with $x \in V(K_k)$, $x \neq w$, and $E_n \subset \{uz : z \in V(K_k), z \neq x\}$. ($E_n$ may be empty).

Proof: Let G be a connected k-chromatic graph of order n. Suppose that $x \in V(G)$ and $\chi(G-x) = k-1$.

Since G is connected, $k_i = k_i(x) \geq 1$ for $i = 0, 1, \ldots, e_G(x)$. Lemma 5.2.8 gives that $k_1 \geq k-1$, and so, subject to this, (5.1), (5.2) is maximised for $k_1 = k-1$, $k_2 = k_3 = \cdots = k_{e_G(x)-1} = 1$, and

$$\begin{aligned} k_{e_G(x)} &= n - \Big(\sum_{i=2}^{e_G(x)-1} k_i + k_0 + k_1\Big) \\ &= n - \Big(\sum_{i=2}^{e_G(x)-1} 1 + 1 + k - 1\Big) \\ &= n - (e_G(x) - 1 - 2 + 1 + 1 + k - 1) \\ &= n - e_G(x) - k + 2, \end{aligned}$$

to give

$$\sigma(x) \leq (k-1)1 + 2 + 3 + \cdots + (e_G(x) - 1) + e_G(x)(n - e_G(x) - k + 2).$$

But because $e_G(x) \leq diam(G) \leq n-k+1$, we have,

$$\sigma(x) \leq (k-1)1 + 2 + 3 + \cdots + (e_G(x) - 1) + e_G(x)(n - e_G(x) - k + 2)$$

$$\leq k-1+2+3+\cdots+(n-k+1-1)+(n-k+1)[n-(n-k+1)-k+2]$$
$$= k-1+2+3+\cdots+(n-k)+(n-k+1). \quad (5.4)$$

Now assume that equality holds in (5.4). Then every distance layer $N_i(x)$, $i = 0, 2, 3, \ldots, e_G(x)$, has 1 vertex and $k_1 = |N_1(x)| = k-1$.

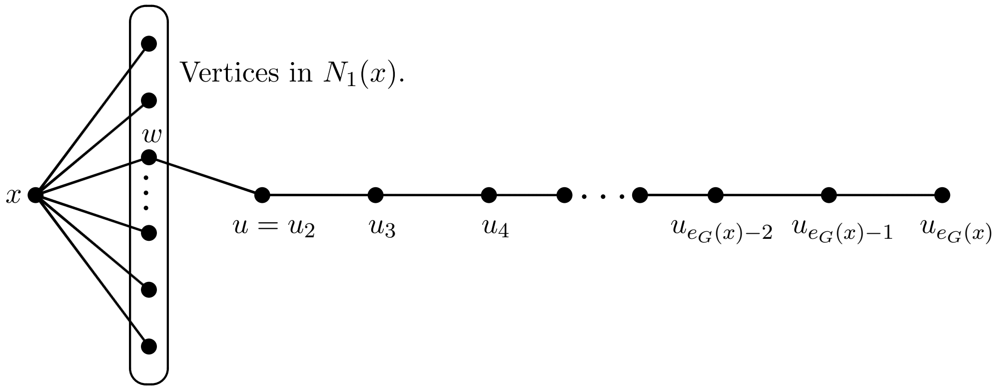


Figure 5.8: A subgraph of a graph G that is obtained when equality holds in (5.4), and x is adjacent to one vertex, w in $N_1(x)$.

By properties of distance layers, if u_i, $i = 2, \ldots, e_G(x)$, is the only vertex in $N_i(x)$, then $u_2u_3 \ldots u_{e_G(x)}$ is a path, and also, since G is connected, we have that, at least one vertex in $N_1(x)$ is adjacent to u_2.

To show that G is the graph obtained from $H_{n,k}$ by adding the edges from a set E_n with $x \in V(K_k)$, $x \neq w$, and $E_n \subset \{uz : z \in V(K_k), z \neq x\}$, it is adequate to prove that any two vertices in $N_1(x)$ are adjacent.

We proceed by contradiction: Suppose to the contrary that $y, z \in N_1(x)$ are such that $yz \notin E(G)$.

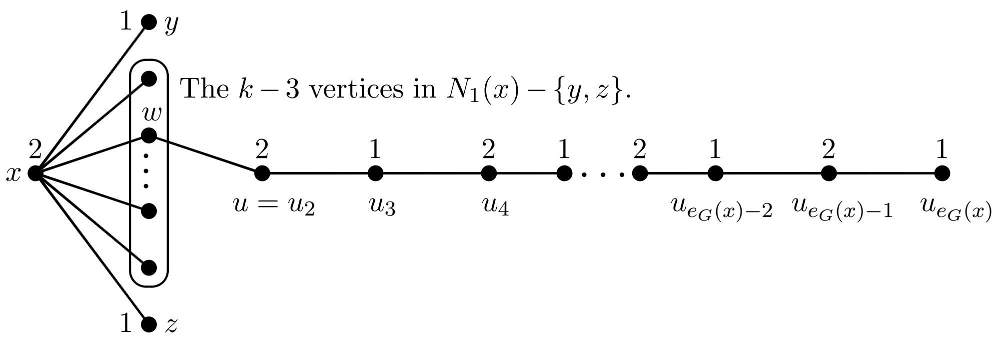


Figure 5.9: A subgraph of a graph G that is obtained when equality holds in (5.4), and u is adjacent to only w in $N_1(x)$. With $yz \notin E(G)$

Since $N_1(x)$ has $k-1$ vertices, the remaining $k-3$ vertices in $N_1(x)$ can be colored using at most $k-3$ colors. Hence, $\chi(G) \leq 2+k-3 = k-1$, contradicting that $\chi(G) = k$. Thus, any two vertices in $N_1(x)$ are adjacent, giving that G is the graph $H_{n,k}$. But, since $u_2 = u$ is adjacent to at least one vertex in $N_1(x)$ in G, we have that G is the graph obtained from $H_{n,k}$ by adding the edges from a set E_n with $x \in V(K_k)$, $x \neq w$, and $E_n \subset \{uz : z \in V(K_k), z \neq x\}$. ($E_n$ may be empty).

Suppose that G is isomorphic to a graph obtained from $H_{n,k}$ by adding the edges from a set E_n with $x \in V(K_k)$, $x \neq w$, and $E_n \subset \{uz : z \in V(K_k), z \neq x\}$. ($E_n$ may be empty). Since $x \in V(K_k)$ and $x \neq w$, we have that $deg_G(x) = k-1$, also, for $|E_n| \geq 0$, we have that $d_G(x,v) = n-k+1$, where v is an end vertex in path P_{n-k}, as given in Definition 5.1.2. Thus, we have,

$$\begin{aligned}
\sigma(x) &= \sum_{y \in V(G)} d_G(x,y) \\
&= \sum_{y \in V(P_{n-k}) \cup \{w\}} d_G(x,y) + \sum_{y \in V(K_k) - \{w\}} d_G(x,y) \\
&= 1+2+\cdots+(n-k+1)+1(k-2) \\
&= (k-2+1)+2+\cdots+n-k+1 \\
&= (k-1)+2+\cdots+n-k+1.
\end{aligned}$$

Thus, equality in (5.4) holds if and only if G is isomorphic to a graph obtained from $H_{n,k}$ by adding the edges from a set E_n with $x \in V(K_k)$, $x \neq w$, and $E_n \subset \{uz : z \in V(K_k), z \neq x\}$. ($E_n$ may be empty). □

Proposition 5.2.11. *For any connected graph G with a vertex x such that $G-x$ is connected, we have,*

$$\sigma(G) \leq \sigma(G-x) + 2\sigma(x).$$

Proof: Let G be a connected graph. Using Proposition 4.2.1, we have,

$$\begin{aligned}
\sigma(G) &= 2W(G) \\
&\leq 2\Big(W(G-x) + \sum_{v \in V(G)} d_G(v,x)\Big) \\
&= \sigma(G-x) + 2\sigma(x),
\end{aligned}$$

as desired. □

Theorem 5.2.12. *Let G be a connected graph with two or more vertices. Then G contains at least two vertices that are not cut-vertices.*

Proof. Let G be a graph with order greater than one and let P be a longest path in G. Assume that P is a x-y path. We prove, by contradiction, that x and y are not cut-vertices. Suppose x is a cut-vertex of G. It follows that $G-x$ is disconnected, giving that G has two or more components. Now, let w be the vertex adjacent to x on P and let P' be the w-y subpath of P.

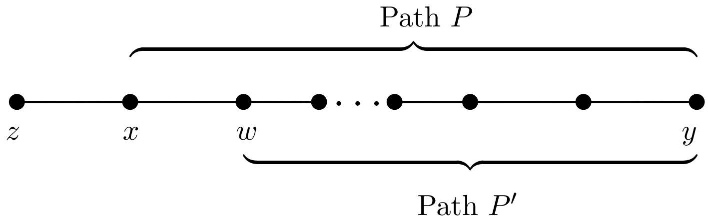


Figure 5.10: The z-y path obtained in G when we assume that x is a cut-vertex in G.

Since x is a cut-vertex, we have that P' must belong to a component, say G_1, of $G-x$. Let the other component of $G-x$ be G_2. Then G_2 contains some vertex z that is adjacent to x. Then zxP' is an $z-x$ path that is longer than P in G, which contradicts the assumption that P is a longest path in G. Similarly, y is not a cut-vertex of G. □

Lemma 5.2.13. *Consider $H_{n,k}$. Then*

$$\sigma(H_{n,k}) = \sigma(H_{n-1,k-1}) + 2[k-1+2+3+\cdots+(n-k+1)].$$

Proof: Let $x \in V(K_k)$, with $x \neq w$ in $H_{n,k}$. Then

$$\begin{aligned}\sigma(x) &= \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\\ &= k-1+2+3+\cdots+(n-k+1). \qquad (5.5)\end{aligned}$$

Now, x does not lie on any of the $u-y$ paths in $H_{n,k}$, so,

$$d_{H_{n,k}}(u,y) = d_{H_{n,k}-x}(u,y) \qquad (5.6)$$

for any vertex $u \neq x$ and vertex $y \neq x$ in $H_{n,k}$. So, using (5.5) and (5.6), we have,

$$\begin{aligned}\sigma(H_{n,k}) &= \sum_{u\in V(H_{n,k})} \sigma(u)\\
&= \sum_{u\in V(H_{n,k})} \sum_{y\in V(H_{n,k})} d_{H_{n,k}}(u,y)\\
&= \sum_{u\in V(H_{n,k})} \Big[\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(u,y) + d_{H_{n,k}}(u,x)\Big]\\
&= \sum_{u\in V(H_{n,k})} \Big(\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(u,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})-x} d_{H_{n,k}}(u,y) + d_{H_{n,k}}(x,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})-x} d_{H_{n,k}}(u,y)\Big) + \sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(x,y) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + \Big(\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(x,y) + d_{H_{n,k}}(x,x)\Big)\\
&\quad + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + \Big(\sum_{y\in V(H_{n,k})} d_{H_{n,k}}(x,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + [k-1+2+3+\cdots+(n-k+1)] \\
&\quad + [k-1+2+3+\cdots+(n-k+1)] \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + 2[k-1+2+3+\cdots+(n-k+1)] \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}-x}(u,y) + 2[k-1+2+3+\cdots+(n-k+1)] \\
&= \sigma(H_{n,k}-x) + 2[k-1+2+3+\cdots+(n-k+1)] \\
&= \sigma(H_{n-1,k-1}) + 2[k-1+2+3+\cdots+(n-k+1)],
\end{aligned}
$$

as desired. □

Lemma 5.2.14. *Consider $H_{n,k}$. Then*

$$\sigma(H_{n,k}) = \sigma(H_{n-1,k}) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)].$$

Proof: Let $x = v$ in $H_{n,k}$. Then

$$
\begin{aligned}
\sigma(x) &= \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= 1+2+3+\cdots+(n-k)+(k-1)(n-k+1). \qquad (5.7)
\end{aligned}
$$

Now, x is an end vertex in $H_{n,k}$, and so, x does not lie on any of the $u-y$ paths in $H_{n,k}$, so,

$$d_{H_{n,k}}(u,y) = d_{H_{n,k}-x}(u,y) \qquad (5.8)$$

for any vertex $u \neq x$ and vertex $y \neq x$ in $H_{n,k}$. So, using (5.7) and (5.8), we have,

$$
\begin{aligned}
\sigma(H_{n,k}) &= \sum_{u\in V(H_{n,k})} \sigma(u) \\
&= \sum_{u\in V(H_{n,k})} \sum_{y\in H_{n,k}} d_{H_{n,k}}(u,y) \\
&= \sum_{u\in V(H_{n,k})} \Big[\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(u,y) + d_{H_{n,k}}(u,x)\Big] \\
&= \sum_{u\in V(H_{n,k})} \Big(\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(u,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})-x} d_{H_{n,k}}(u,y) + d_{H_{n,k}}(x,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{y\in V(H_{n,k})-x} \Big(\sum_{u\in V(H_{n,k})-x} d_{H_{n,k}}(u,y)\Big) + \sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(x,y) \\
&\quad + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + \Big(\sum_{y\in V(H_{n,k})-x} d_{H_{n,k}}(x,y) + d_{H_{n,k}}(x,x)\Big) \\
&\quad + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + \Big(\sum_{y\in V(H_{n,k})} d_{H_{n,k}}(x,y)\Big) + \sum_{u\in V(H_{n,k})} d_{H_{n,k}}(u,x) \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + [1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] \\
&\qquad + [1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}}(u,y) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] \\
&= \sum_{u,y\in V(H_{n,k}-x)} d_{H_{n,k}-x}(u,y) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] \\
&= \sigma(H_{n,k}-x) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] \\
&= \sigma(H_{n-1,k}) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)],
\end{aligned}
$$

as desired. □

Theorem 5.2.15. *In the class of connected k-chromatic graphs G of order n, the sum of all distances $\sigma(G)$ satisfies:*

a) $\sigma(G) \geq \sigma(H)$ where $H \in T(n,k)$;

b) $\sigma(G) \leq \sigma(H_{n,k})$.

Proof: a) Let G be a connected k-chromatic graph of order n and n_i the number of vertices having color i for $1 \leq i \leq k$ with respect to a fixed k-coloring of G with colors $1, 2, \ldots, k$. If vertices x and y have the same color it follows that $d_G(x,y) \geq 2$, and if vertices x and y have different colors, then $d_G(x,y) \geq 1$.
Let V_i denote the set of vertices with color i, $1 \leq i \leq k$ so that $n_i = |V_i|$. Then the order of G is $n = n_1 + n_2 + \cdots + n_k$, and so,

$$|\{\{x,y\} \mid x,\, y \in V_i\}| = \frac{n_i(n_i-1)}{2}.$$

We have,

$$
\begin{aligned}
\sigma(G) &= \sum_{x,y\in V(G)} d_G(x,y) \\
&= \sum_{x,y\in V_1} d_G(x,y) + \sum_{x,y\in V_2} d_G(x,y) + \cdots + \sum_{x,y\in V_k} d_G(x,y) \\
&\quad + \sum_{x\in V_1, y\in V(G)-V_1} d_G(x,y) + \sum_{x\in V_2, y\in V(G)-V_2} d_G(x,y) + \cdots + \sum_{x\in V_k, y\in V(G)-V_k} d_G(x,y) \\
&= 2\sum_{\{x,y\}\subseteq V_1} d_G(x,y) + 2\sum_{\{x,y\}\subseteq V_2} d_G(x,y) + \cdots + 2\sum_{\{x,y\}\subseteq V_k} d_G(x,y) \\
&\quad + \sum_{x\in V_1, y\in V(G)-V_1} d_G(x,y) + \sum_{x\in V_2, y\in V(G)-V_2} d_G(x,y) + \cdots + \sum_{x\in V_k, y\in V(G)-V_k} d_G(x,y) \\
&= 2\Big[\sum_{\{x,y\}\subseteq V_1} d_G(x,y) + \sum_{\{x,y\}\subseteq V_2} d_G(x,y) + \cdots + \sum_{\{x,y\}\subseteq V_k} d_G(x,y)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{x\in V_1, y\in V(G)-V_1} d_G(x,y) + \sum_{x\in V_2, y\in V(G)-V_2} d_G(x,y) + \cdots + \sum_{x\in V_k, y\in V(G)-V_k} d_G(x,y) \\
&\geq 2\Big[\sum_{\{x,y\}\subseteq V_1} 2 + \sum_{\{x,y\}\subseteq V_2} 2 + \cdots + \sum_{\{x,y\}\subseteq V_k} 2\Big] \\
&+ \sum_{x\in V_1, y\in V(G)-V_1} 1 + \sum_{x\in V_2, y\in V(G)-V_2} 1 + \cdots + \sum_{x\in V_k, y\in V(G)-V_k} 1 \\
&= 2\Big[2\frac{n_1(n_1-1)}{2} + 2\frac{n_2(n_2-1)}{2} + \cdots + 2\frac{n_k(n_k-1)}{2}\Big] + n_1(n-n_1) \\
&+ n_2(n-n_2) + \cdots + n_k(n-n_k) \\
&= 2[n_1(n_1-1) + n_2(n_2-1) + \cdots + n_k(n_k-1)] + n_1(n-n_1) \\
&+ n_2(n-n_2) + \cdots + n_k(n-n_k) \\
&= n_1(n-n_1) + 2n_1(n_1-1) + n_2(n-n_2) + 2n_2(n_2-1) + \cdots + n_k(n-n_k) + 2n_k(n_k-1) \\
&= n_1[(n-n_1) + 2(n_1-1)] + n_2[(n-n_2) + 2(n_2-1)] \\
&+ \cdots + n_k[(n-n_k) + 2(n_k-1)] \\
&= \sum_{i=1}^{k} n_i(n - n_i + 2(n_i-1)) \\
&= \sum_{i=1}^{k} n_i(n + n_i - 2) \\
&= n\sum_{i=1}^{k} n_i + \sum_{i=1}^{k} n_i^2 - 2\sum_{i=1}^{k} n_i \\
&= n^2 - 2n + \sum_{i=1}^{k} n_i^2.
\end{aligned}
$$

Now we show, for any graph $H \in T(n,k)$, that $\sigma(H) = n^2 - 2n + \sum_{i=1}^{k} n_i^2$.

Let $H \in T(n,k)$. Then two vertices in H are connected by an edge if and only if they belong to different classes of vertices as stated in Notation 5.1.4. We have that, vertices that are in the same class are a distance 2 from one another, and vertices in different classes are a distance 1 from one another. Thus, we can color each class of H with one color. By Notation 5.1.4, H has k classes. Let V_i denote the set of vertices of class i, $1 \leq i \leq k$ in H and let $n_i = |V_i|$. So, vertices $x, y \in V_i$ are a distance $d_H(x,y) = 2$ in H, and if $x \in V_i$ and $y \in V(H) - V_i$ then $d_H(x,y) = 1$. Also, the order of H is $n = n_1 + n_2 + \cdots + n_k$, and so,

$$|\{\{x,y\} \mid x,\, y \in V_i\}| = \frac{n_i(n_i-1)}{2}.$$

We have,

$$
\begin{aligned}
\sigma(H) &= \sum_{x,y\in V(H)} d_H(x,y) \\
&= \sum_{x,y\in V_1} d_H(x,y) + \sum_{x,y\in V_2} d_H(x,y) + \cdots + \sum_{x,y\in V_k} d_H(x,y) \\
&+ \sum_{x\in V_1, y\in V(H)-V_1} d_H(x,y) + \sum_{x\in V_2, y\in V(H)-V_2} d_H(x,y) + \cdots + \sum_{x\in V_k, y\in V(H)-V_k} d_H(x,y)
\end{aligned}
$$

$$
\begin{aligned}
&= 2\sum_{\{x,y\}\subseteq V_1} d_H(x,y) + 2\sum_{\{x,y\}\subseteq V_2} d_H(x,y) + \cdots + 2\sum_{\{x,y\}\subseteq V_k} d_H(x,y) \\
&\quad + \sum_{x\in V_1, y\in V(H)-V_1} d_H(x,y) + \sum_{x\in V_2, y\in V(H)-V_2} d_H(x,y) + \cdots + \sum_{x\in V_k, y\in V(H)-V_k} d_H(x,y) \\
&= 2\Big[\sum_{\{x,y\}\subseteq V_1} d_H(x,y) + \sum_{\{x,y\}\subseteq V_2} d_H(x,y) + \cdots + \sum_{\{x,y\}\subseteq V_k} d_H(x,y)\Big] \\
&\quad + \sum_{x\in V_1, y\in V(H)-V_1} d_H(x,y) + \sum_{x\in V_2, y\in V(H)-V_2} d_H(x,y) + \cdots + \sum_{x\in V_k, y\in V(H)-V_k} d_H(x,y) \\
&= 2\Big[\sum_{\{x,y\}\subseteq V_1} 2 + \sum_{\{x,y\}\subseteq V_2} 2 + \cdots + \sum_{\{x,y\}\subseteq V_k} 2\Big] \\
&\quad + \sum_{x\in V_1, y\in V(H)-V_1} 1 + \sum_{x\in V_2, y\in V(H)-V_2} 1 + \cdots + \sum_{x\in V_k, y\in V(H)-V_k} 1 \\
&= 2\Big[2\frac{n_1(n_1-1)}{2} + 2\frac{n_2(n_2-1)}{2} + \cdots + 2\frac{n_k(n_k-1)}{2}\Big] \\
&\quad + n_1(n-n_1) + n_2(n-n_2) + \cdots + n_k(n-n_k) \\
&= 2[n_1(n_1-1) + n_2(n_2-1) + \cdots + n_k(n_k-1)] \\
&\quad + n_1(n-n_1) + n_2(n-n_2) + \cdots + n_k(n-n_k) \\
&= n_1(n-n_1) + 2n_1(n_1-1) + n_2(n-n_2) \\
&\quad + 2n_2(n_2-1) + \cdots + n_k(n-n_k) + 2n_k(n_k-1) \\
&= n_1[(n-n_1) + 2(n_1-1)] + n_2[(n-n_2) + 2(n_2-1)] + \cdots + n_k[(n-n_k) + 2(n_k-1)] \\
&= \sum_{i=1}^{k} n_i(n - n_i + 2(n_i - 1)) \\
&= \sum_{i=1}^{k} n_i(n + n_i - 2) \\
&= n\sum_{i=1}^{k} n_i + \sum_{i=1}^{k} n_i^2 - 2\sum_{i=1}^{k} n_i \\
&= n^2 - 2n + \sum_{i=1}^{k} n_i^2.
\end{aligned}
$$

So, we have that $\sigma(G) \geq n^2 - 2n + \sum_{i=1}^{k} n_i^2 = \sigma(H)$, where $H \in T(n,k)$, as desired.

b) We proceed by contradiction. Suppose G is a k-chromatic graph of order n which is a counter example to the statement of b). Then, we have that $\sigma(G) > \sigma(H_{n,k})$.

Amongst all such counter examples, choose G to be a counter example where $n+k$ is minimum. Therefore, if H is k'-chromatic and of order n' where $n'+k' < n+k$, then H is not a counter example, i.e., H satisfies b).

By Theorem 5.2.12, there is a vertex x such that $G-x$ is connected. Suppose that $G-x$ is connected. Then $\chi(G-x)$ is either equal to $k-1$ or k, and $G-x$ has order $n-1$.

Note 5.2.16. If $\chi(G-x) = k-1$, we have that $k-1+n-1 = k+n-2 < n+k$, so, $G-x$ is not a counter example of b). Thus, if $\chi(G-x) = k-1$, then $\sigma(G-x) \leq \sigma(H_{n-1,k-1})$.

Note 5.2.17. If $\chi(G-x) = k$, we have that $k+n-1 < k+n$, and so, $G-x$ is not a counter example of b). Thus, if $\chi(G-x) = k$, then $\sigma(G-x) \leq \sigma(H_{n-1,k})$.

We have that G and $G-x$ are connected graphs. Now, assume that $\chi(G-x)=k-1$. By Property 5.2.10, we have that $\sigma(x) \leq k-1+2+3+\cdots+(n-k+1)$ and equality holds if G is isomorphic to the graph obtained from $H_{n,k}$ by adding the edges from a set $E_n \subset \{uz : z \in V(K_k), z \neq x\}$ with $x \in V(K_k)$, $x \neq w$.

Now, if $E_n \neq \emptyset$, then the resulting graph is obtained by adding at least one edge from the set $E_n \subset \{uz : z \in V(K_k), z \neq x\}$ with $x \in V(K_k)$, $x \neq w$ to $H_{n,k}$. Denote the resulting graph by $H^*_{n,k}$. It follows that there is some vertex $z \neq x$ in $H^*_{n,k}$ that belongs to the k-clique of $H^*_{n,k}$ such that the distance between z and v is $n-k$ in $H^*_{n,k}$. But in $H_{n,k}$, we have that $d_{H_{n,k}}(z,v) = n-k+1$ for every vertex z in the k-clique of $H_{n,k}$, and this makes $\sigma(H_{n,k}) > \sigma(H^*_{n,k})$. Thus, showing that $\sigma(G) \leq \sigma(H_{n,k})$ will be sufficient. This is because both $H_{n,k}$ and $H^*_{n,k}$ have order n and chromatic number k with the property that $\chi(H^*_{n,k}-x) = \chi(H_{n,k}-x) = k-1$.
Since we aim to show the $H_{n,k}$ is the extremal graph, we assume that $E_n = \emptyset$. Let $E_n = \emptyset$. By Proposition 5.2.11, we have, $\sigma(G) \leq \sigma(G-x) + 2\sigma(x)$. So, we have that,

$$\begin{aligned}
\sigma(G) &\leq \sigma(G-x) + 2\sigma(x) \\
&\leq \sigma(H_{n-1,k-1}) + 2\sigma(x) && \text{(by Note 5.2.16)} \\
&\leq \sigma(H_{n-1,k-1}) + 2[k-1+2+3+\cdots+(n-k+1)] && \text{(by Property 5.2.10)} \\
&= \sigma(H_{n,k}). && \text{(by Lemma 5.2.13)}
\end{aligned}$$

We have that G and $G-x$ are connected graphs. Now, assume that $\chi(G-x)=k$. By Property 5.2.5, we have that $\sigma(x) \leq 1+2+3+\cdots+(n-k)+(k-1)(n-k+1)$ and equality holds if G is isomorphic to the graph $H_{n,k}$ with $x=v$. By Proposition 5.2.11, we have, $\sigma(G) \leq \sigma(G-x) + 2\sigma(x)$. So, we have that,

$$\begin{aligned}
\sigma(G) &\leq \sigma(G-x) + 2\sigma(x) \\
&\leq \sigma(H_{n-1,k}) + 2\sigma(x) && \text{(by Note 5.2.17)} \\
&\leq \sigma(H_{n-1,k}) + 2[1+2+3+\cdots+(n-k)+(k-1)(n-k+1)] && \text{(by Property 5.2.5)} \\
&= \sigma(H_{n,k}). && \text{(by Lemma 5.2.14)}
\end{aligned}$$

Since $\sigma(G) \leq \sigma(H_{n,k})$ whenever $\chi(G-x)=k-1$ or $\chi(G-x)=k$ and $G-x$ is connected, we have that G is not a counter example of b), contradicting our choice of G. □

5.3 Conclusion

In this chapter, we presented proofs to a two part theorem. The first part of the theorem gave a lower bound on the average distance in terms of order and chromatic number of the graph. The second part of the theorem gave an upper bound on average distance in terms of order and chromatic number of the graph. The theorem we proved is given by Theorem 5.2.15, and is stated as follows: "In the class of connected k-chromatic graphs G of order n the sum of all distances $\sigma(G)$ satisfies:

a) $\sigma(G) \geq \sigma(H)$ where $H \in T(n,k)$;

b) $\sigma(G) \leq \sigma(H_{n,k})$."

We conclude the dissertation in the next chapter.

Chapter 6

Conclusion

In this dissertation, we studied upper bounds on the average eccentricity and upper bounds on average distance of a graph G in terms of order and independence number of the graph, and also, we studied upper bounds on average eccentricity and upper bounds on the average distance of a graph G in terms of order and chromatic number of the graph. We gave a lower bound on the average distance in terms of order and chromatic number of the graph as well. We began with an introductory chapter, Chapter 1, where we gave neccessary definitions that were used throughout the dissertation, and where we introduced the reader to the concepts that we used in the results of this dissertation. To give an idea to why we would study the results of this dissertation, we gave a detailed literature review in Chapter 2, where we discussed motivation and background work done on the bounds of average distance and average eccentricity in terms of order, independence number and chromatic number of the graph. We also gave a survey of important results in Chapter 2.

In Chapter 3, we presented upper bounds on the average eccentricity in terms of order and independence number of the graph. What makes the upper bounds difficult to establish is the fact that there are two extremal graphs depending on the size of the independence number. In particular, in an in-depth analysis, we showed that if G is a connected graph of order n and independence number α, where $\alpha \leq \frac{n}{2}$, then $\zeta(G) \leq (2\alpha - 1)n - \alpha^2 + \alpha$, whilst, if G is a connected graph of order n and independence number α, where $\alpha > \frac{n}{2}$, then $\zeta(G) \leq n^2 - \alpha^2$. We showed that if α is at most $\frac{n}{2}$ then an extremal graph is made up of a path of order $2\alpha - 2$ whose one end is joined to a complete graph of order a and the other end is joined to a complete graph of order b where $a + b = n - 2\alpha + 2$. For values of α greater than $\frac{n}{2}$, an extremal graph is a tree T, obtained from a path of order $2n - 2\alpha$ and attaching $2\alpha - n$ end vertices to one end of the path. We proceeded to present upper bounds on average eccentricity in terms of order and chromatic number of the graph in Chapter 3. In particular, we showed that if G is a connected graph of order n and chromatic number χ, then

$$avec(G) \leq \frac{1}{n}\left(\left\lfloor \frac{3}{4}n^2 - \frac{1}{4}\chi^2 - \frac{1}{2}\chi n + \frac{1}{2}n + \frac{1}{2}\chi \right\rfloor\right),$$

with an extremal graph being the graph obtained from combining the complete graph K_χ and a path $P_{n-\chi}$ with an edge that joins an end vertex of $P_{n-\chi}$ to one of the vertices of K_χ.

In Chapter 4, we presented upper bounds on the average distance in terms of order and the independence number of the graph. We showed that if G is a connected graph of order n and independence number α, where $\alpha < \frac{n+1}{2}$, then $\mu(G) \leq \alpha + 2$, whilst if $\alpha \geq \frac{n+1}{2}$, then $\mu(G) \leq \frac{2}{3}\alpha$. We showed, by construction, that these bounds are best possible. For values of α less than $\frac{n+1}{2}$, an extremal graph is made up of a path of order $2\alpha - 2$ whose ends are joined to two complete graphs with order that differ by at most one. For values of α bigger than $\frac{n+1}{2}$, an extremal graph is made up of a path of order $2n - 2\alpha$ whose ends are joined to two complete graphs with order that differ by at most one.

In Chapter 5, we presented lower and upper bounds on the average distance in terms of order and chromatic number. We started by showing that if G is a connected k-chromatic graph of order n then $\mu(G) \geq \mu(H)$, where H is a graph in the set $T(n, k)$. Using contradiction, we showed that if G is a connected k-chromatic graph of order n, then $\mu(G) \leq \mu(H_{n,k})$.

To further improve the bounds on average distance and average eccentricity, an area for future research is bounds on average distance and average eccentricity in terms of order, independence number and minimum degree of the graph, and also, in terms of order, chromatic number and minimum degree of the graph. It would also be interesting to apply our findings on triangle free graphs, and find extremal graphs of the improved bounds for triangle free graphs.

Bibliography

[1] Aouchiche, M., Bonnefoy, J.-M., Fidahoussen, A., Caporossi, G., Hansen, P., Hiesse, L., Lacheré, J., and Monhait, A. Variable neighborhood search for extremal graphs 14: the autographix 2 system. *Global Optimization* (2006), 281–310.

[2] Chartrand, G., and Zhang, P. *Chromatic Graph Theory.* Discrete Mathematics and Its Applications. CRC Press, 2008.

[3] Chung, F. R. The average distance and the independence number. *Journal of Graph theory 12*, 2 (1988), 229–235.

[4] Cvetković, D., Kraus, L., and Simić, S. Discussing graph theory with a computer i. implementation of graph theoretic algorithms/razgovor sa kompjuterom o teoriji grafova i. implementacija algoritama teorije grafova. *Publikacije Elektrotehničkog fakulteta. Serija Matematika i fizika* (1981), 100–104.

[5] Cvetković, D., and Simić, S. Graph theoretical results obtained by the support of the expert system" graph". *Bulletin (Académie serbe des sciences et des arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)* (1994), 19–41.

[6] Dankelmann, P. Average distance and independence number. *Discrete Applied Mathematics 51*, 1-2 (1994), 75–83.

[7] Dankelmann, P. Average distance and domination number. *Discrete Applied Mathematics 80*, 1 (1997), 21–35.

[8] Dankelmann, P., Goddard, W., and Swart, C. S. The average eccentricity of a graph and its subgraphs. *Utilitas Mathematica 65* (2004), 41–52.

[9] Dankelmann, P., and Mukwembi, S. Upper bounds on the average eccentricity. *Discrete Applied Mathematics 167* (2014), 72–79.

[10] DeLaVina, E. Graffiti. pc. *Graph Theory Notes of New York 42*, 3 (2002), 26–30.

[11] Dobrynin, A. A., Entringer, R., and Gutman, I. Wiener index of trees: theory and applications. *Acta Applicandae Mathematica 66*, 3 (2001), 211–249.

[12] Doyle, J., and Graver, J. Mean distance in a graph. *Discrete Mathematics 17*, 2 (1977), 147–154.

[13] Du, Z., and Ilic, A. On autographix conjectures regarding average eccentricity. *Match-Communications in Mathematical and Computer Chemistry 69*, 3 (2013), 597–609.

[14] Entringer, R. C., Jackson, D. E., and Snyder, D. Distance in graphs. *Czechoslovak Mathematical Journal 26*, 2 (1976), 283–296.

[15] Fajtlowicz, S. On conjectures of graffiti. *Annals of Discrete Mathematics 38* (1988), 113–118.

[16] Ilić, A. On the extremal properties of the average eccentricity. *Computers & Mathematics with Applications 64*, 9 (2012), 2877–2885.

[17] Kouider, M., and Winkler, P. Mean distance and minimum degree. *Journal of Graph Theory 25*, 1 (1997), 95–99.

[18] Mukwembi, S. Average distance, independence number, and spanning trees. *Journal of Graph Theory 76*, 3 (2014), 194–199.

[19] Plesník, J. On the sum of all distances in a graph or digraph. *Journal of Graph Theory 8*, 1 (1984), 1–21.

[20] Quadras, J., Christy, K. A., Nelson, A., and Surya, S. S. Average distance of certain graphs. *Recent Trends in Graph Theory and Combinatorics* (2017), 87–94.

[21] Sampathkumar, E., and Walikar, H. The connected domination number of a graph. *Journal of Mathematical and Physical Sciences 13*, 6 (1979), 607–613.

[22] Tang, Y., and Qi, X. Ordering graphs with large eccentricity-based topological indices. *Journal of Inequalities and Applications* (2021), 1–11.

[23] Tang, Y., and Zhou, B. On average eccentricity. *Match-Communications in Mathematical and Computer Chemistry 67*, 2 (2012), 405.

[24] TOMESCU, I., and MELTER, R. A. On distances in chromatic graphs. *The Quarterly Journal of Mathematics 40*, 4 (1989), 475–480.

[25] Xu, K., Liu, M., Das, K. C., Gutman, I., and Furtula, B. A survey on graphs extremal with respect to distance-based topological indices. *Match-Communications in Mathematical and Computer Chemistry 71*, 3 (2014), 461–508.