DISTANCES IN AND BETWEEN GRAPHS

by

TIMOTHY JACKSON BEAN

Submitted in partial fulfilment of the
requirements for the degree of
Master of Science
in the
Department of Mathematics and Applied Mathematics,
University of Natal, Durban.

December 1991

Preface

The research on which this thesis was based was carried out in the Department of Mathematics and Applied Mathematics, University of Natal, Durban, from November 1990 to December 1991, under the supervision of Professor Henda C. Swart.

This thesis represents original work by the author and has not been submitted in any form to another university. Where use was made of the work of others it has been duly acknowledged.

Acknowledgements

I would like to express my sincere thanks to Professor Henda C. Swart for her guidance, encouragement and patience during the preparation of this thesis. Apart from her excellent supervision, her infectious enthusiasm for the subject deepened my interest and made the task of researching this thesis enjoyable and rewarding.

I also thank Dr. Wayne Goddard for the time he spent on discussing with me some of the work he had done on the topics covered in Chapters 2 and 3 of this thesis.

I extend my gratitude to the Foundation for Research Development for the award of a scholarship.

Finally, a sincere thank you to Mrs. Dale Haslop for her typing of this thesis.

Abstract

Aspects of the fundamental concept of distance are investigated in this dissertation. Two major topics are discussed; the first considers metrics which give a measure of the extent to which two given graphs are removed from being isomorphic, while the second deals with Steiner distance in graphs which is a generalization of the standard definition of distance in graphs.

Chapter 1 is an introduction to the chapters that follow. In Chapter 2, the edge slide and edge rotation distance metrics are defined. The edge slide distance gives a measure of distance between connected graphs of the same order and size, while the edge rotation distance gives a measure of distance between graphs of the same order and size. The edge slide and edge rotation distance graphs for a set S of graphs are defined and investigated. Chapter 3 deals with metrics which yield distances between graphs or certain classes of graphs which utilise the concept of greatest common subgraphs. Then follows a discussion on the effects of certain graph operations on some of the metrics discussed in Chapters 2 and 3. This chapter also considers bounds and relations between the metrics defined in Chapters 2 and 3 as well as a partial ordering of these metrics.

Chapter 4 deals with Steiner distance in a graph. The Steiner distance in trees is studied separately from the Steiner distance in graphs in general. The concepts of eccentricity, radius, diameter, centre and periphery are gen-

eralised under Steiner distance. This final chapter closes with an algorithm which solves the Steiner problem and a Heuristic which approximates the solution to the Steiner problem.

Contents

1 Introduction **1**

1.1 Graph Theory Nomenclature 1

1.2 Distances between Graphs and Steiner Distance 3

2 Distance Between Graphs **7**

2.1 Introduction 7

2.2 The Edge Slide Distance d_{es} 8

2.3 The Edge Rotation Distance d_{er} 27

3 Metrics Involving Greatest Common Subgraphs **66**

3.1 Introduction 66

3.2 The Edge Move Distance d_{em} 66

3.3 Bounds and Relations 93

3.4 Graph Operations 108

3.5 Ordering of Metrics 114

4 The Steiner Problem in Graphs **132**

4.1 Introduction 132

4.2 The Steiner Problem in Graphs 135

4.3 Steiner Distance in Trees 140

4.4 Steiner Distance in Graphs 183

4.5 An Algorithm and a Heuristic for the Steiner Problem in Graphs 220

Bibliography **243**

Chapter 1

Introduction

1.1 Graph Theory Nomenclature

The basic text for the graph theory terminology and symbols used here is Chartrand and Lesniak's *Graphs and Digraphs* [CL1]. Here we clarify our conventions.

We denote by Γ the space of all graphs, by $\Gamma(p)$ the space of graphs of order p, by $\Gamma(p,q)$ the space of all graphs of order p and size q and by $\Gamma_c(p,q)$ the space of all connected graphs of order p and size q. The space of all trees of order n is denoted by $\tau(n)$. We denote by S_p the space of all isomorphism classes of graphs on p vertices, by $S_{p,q}$ the space of all isomorphism classes of graphs with p vertices and q edges and by $S^c_{p,q}$ the space of all isomorphism classes of connected graphs with p vertices and q edges.

We use $p(G)$, $q(G)$, $V(G)$ and $E(G)$ to denote the order, size, vertex set and edge set respectively of a graph G. If $v \in V(G)$, the degree of v in G is written as $\deg_G v$ and the minimum degree of G is given by

$\delta(G) = \min\{\deg_G v : v \in V(G)\}$ whereas the maximum degree of G is $\triangle(G) = \max\{\deg_G v : v \in V(G)\}$. The set of all vertices adjacent to v in G is denoted by $N_G(v)$. If S is a set of elements (either edges or vertices) then the number of elements in the set S is written as $|S|$. If $G, H \in \Gamma(p)$ and G and H are defined on the same vertex set then we denote by $|V(G) - V(H)|(|E(G) - E(H)|)$, the number of vertices (edges respectively) which appear in G but not in H. A set $S \subseteq V(G)$ of vertices of G is an *independent* set if no two vertices of S are adjacent in G. If $G \in \Gamma(p, q)$, the *cardinality* of G is defined to be $p + q$, denoted by $|G|$.

If $S \subseteq V(G)$ is a subset of the vertex set of a graph G, then we denote by $\langle S \rangle$ the subgraph of G induced by the vertices of S. We denote by $H < G$ that H is an induced subgraph of the graph G. A *block* B of a graph G is a subgraph of G with maximum order such that B contains no cut-vertex. An *end-block* of G is a block of G which contains exactly one cut-vertex of G. A *branch* B of a graph G at the vertex $w \in V(G)$ is a maximum connected induced subgraph of G containing w as a non cut-vertex.

If G is a given graph with vertex set $V(G) = \{v_1, v_2, ..., v_p\}$ and edge set $E(G) = \{e_1, e_2, ..., e_q\}$ then the *line graph* $L(G)$ of G is the graph obtained as follows: $L(G)$ has vertex set $\{e_1, e_2, ..., e_q\}$ and $e_i e_j \in E(L(G))$ if and only if e_i and e_j are incident with a common vertex in G, for $1 \leq i < j \leq q$.

For graphs G_1 and G_2, the *cartesian product* $G_1 \times G_2$ is a graph which has vertex set $V(G_1) \times V(G_2)$ such that two vertices (u_1, u_2) and (v_1, v_2) are adjacent in $G_1 \times G_2$ if and only if either $u_1 = v_1$ and $u_2 v_2 \in E(G_2)$ or

$u_2 = v_2$ and $u_1v_1 \in E(G_1)$. Again if G_1 and G_2 are vertex-disjoint graphs, then the *join* of G_1 and G_2, denoted by $G_1 + G_2$, is that graph consisting of the disjoint union $G_1 \cup G_2$, together with all edges of the type v_1v_2, where $v_1 \in V(G_1)$ and $v_2 \in V(G_2)$. We denote by $G - \{S\}$, where $S \subseteq V(G)$, the graph obtained from G by deleting the vertices in S from G together with all the edges incident with vertices of S. The *contraction* of graph G along an edge $e = xy \in E(G)$ say, is the graph obtained from G by deleting e and identifying the vertices x and y in G in a single vertex which is adjacent to all vertices in $N_G(x) \cup N_G(y)$.

A *unicyclic* graph is a graph containing only one cycle and the *girth* $g(G)$ of a graph G is the length of a shortest cycle in G. The *star* S_n is isomorphic to the graph $K_{1,n}$.

Other definitions will be given as needed throughout the chapters.

1.2 Distances between Graphs and Steiner Distance

If two graphs G_1 and G_2 are not isomorphic, then how far away from isomorphism are they? In Chapters 2 and 3 we define and develop metrics which can be used to answer this question.

The line of discussion of Section 2.2 is as follows:

Let $G, H \in \Gamma_c(p,q)$; then G can be transformed into H by an *edge slide* if G contains distinct vertices u, v and w such that $uv \in E(G), uw \in$

$E(\bar{G}), vw \in E(G)$ and $H \cong G - uv + uw$. The edge slide distance between G and H in this case, denoted $d_{es}(G, H)$ is 1. The minimum number of edge slides needed to transform one graph into another gives a measure of distance between the graphs. In this section we see that for any two graphs $G_1, G_2 \in \Gamma_c(p, q)$ it is always possible to transform G_1 into G_2 by means of a sequence of edge slides. We also consider edge slide distances between specific classes of graphs. The remainder of this section deals with the *edge slide distance graph* $D_s(S)$ of a set $S \subseteq \Gamma_c(p, q)$ of graphs, where $D_s(S)$ has S as its vertex set and two vertices x and y of $D_s(S)$ are adjacent if and only if $d_{es}(x, y) = 1$. It is shown that every graph is an edge slide distance graph.

The edge rotation distance metric is introduced and discussed in Section 2.3. Let $G, H \in \Gamma(p, q)$; then G can be transformed in H by an *edge rotation* if G contains distinct vertices u, v and w such that $uv \in E(G)$, $uw \in E(\bar{G})$ and $H \cong G - uv + uw$. Here we dispense with the restriction that vw must be an edge of G as is demanded by the edge slide operation. The discussion folowed in this section is similar to the line of discussion in Section 2.2.

A number of metrics which give a measure of distance between nonisomorphic graphs are defined in Chapter 3. These metrics have in common that they are closely linked to the idea of a greatest common subgraph of the graphs in question. In Section 3.2 these metrics are defined and discussed.

In Section 3.3 certain relations between the metrics we have studied are established, together with some results which set bounds on the distances between graphs.

The effects of the application of some simple graph operations on the distances between associated graphs are studied in Section 3.4. The operations considered are: The join, the union and the subdivision.

A partial ordering for the metrics defined in Chapters 2 and 3 is developed in Section 3.5.

Chapter 4 deals with the background to the Steiner problem, which is to connect n given points in the plane by a shortest possible network of line segments. We consider the extension of this problem to graphs which is our main topic of discussion here.

The Steiner problem in graphs may be stated as follows: Consider a connected graph G of order p and a proper subset $S \subseteq V(G)$. The problem is to find a subtree T_S of G of minimum size such that $V(T_S) \supseteq S$. The concepts of eccentricity, radius, diameter, centres and peripheries are generalised under the Steiner distance. The generalisations yield many results which are separated into two categories. Secion 4.3 includes results which apply to trees and Section 4.4 considers graphs in general.

Finally in Section 4.5 we present an algorithm which solves the Steiner problem in graphs exactly but in impractical time for large $|S|$. Hence we also present a more economical approximate algorithm for finding a tree T'_S

in our graph G which is close to optimal, and we show just how accurate this heuristic is.

Chapter 2

Distance Between Graphs

2.1 Introduction

In this chapter we shall define two metrics, both giving a measure of the distance between certain given graphs and/or between certain classes of graphs. We will consider some of the properties exhibited by these metrics and determine distances between specific graphs.

The two metrics to be investigated, namely the edge slide distance metric and the edge rotation distance metric, are similar in nature. They both involve the deformation of a graph G, which translates G into a graph $G' = G - e_1 + e_2$ where $e_1 \in E(G)$ and $e_2 \in E(\overline{G})$. The aim is to transform one graph to another with the least possible number of such deformations. The number of deformations gives a measure of distance between the original graph and the transformed graph.

2.2 The Edge Slide Distance d_{es}

The concepts of edge slide and edge slide distance were defined by M. Johnson in [J1] and independently by Benade, Goddard, McKee and Winter in [BGMW1].

2.2.1 Definitions

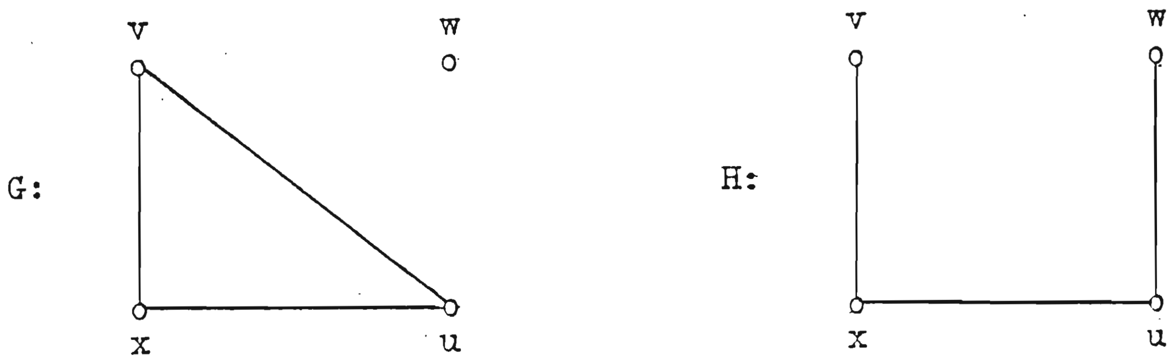
Let G and H be two graphs with the same number, say k, of components, where the components are so labelled that the ith component $(1 \leq i \leq k)$ of each of G and H has the same order and the same size. We say that G can be *transformed into* H by an *edge slide* if G contains distinct vertices u, v and w such that $uv \in E(G), uw \in E(\overline{G}), vw \in E(G)$ and $H \cong G - uv + uw$.

Now let graphs G and H be defined as above; then the *edge slide distance* $d_{es}(G, H)$ between G and H is defined as the smallest nonnegative integer n for which there exists a sequence

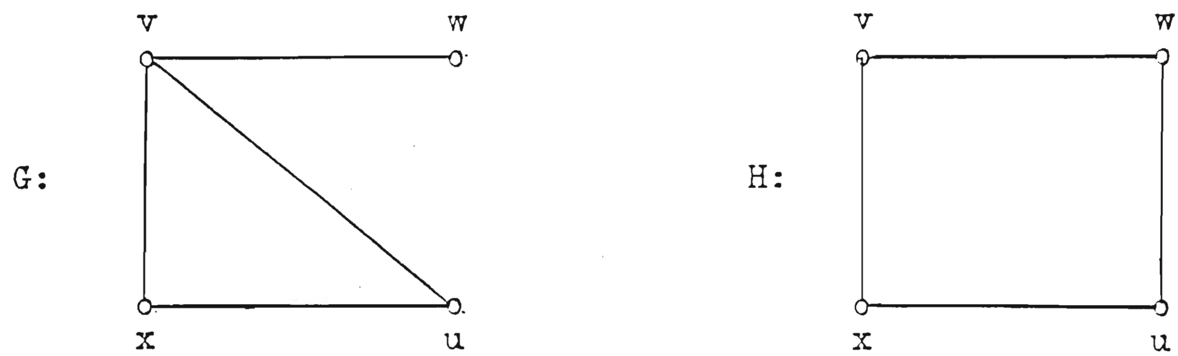
$$G \cong F_0, F_1, F_2, ..., F_n \cong H$$

such that F_i can be transformed into F_{i+1} by an edge slide, for $i = 0, 1, 2, ..., n - 1$.

2.2.2 Example



2.2.3 Figure



2.2.4 Figure

In Figure 2.2.3 the graph G cannot be transformed into the graph H by an edge slide; however in Figure 2.2.4 the graph G can be transformed into the graph H by the edge slide which removes the edge uv from G and then inserts the edge uw in $G - uv$; i.e., $H \cong G - uv + uw$.

To simplify notation we shall denote an edge slide which results in a graph G being transformed into the graph $G - uv + uw$ as in the above example by $t = (u, v, w)$ where $uv \in E(G)$, $uw \in E(\bar{G})$ and $vw \in E(G)$. The graph $G - uv + uw$ will be denoted by tG and to avoid ambiguity t will also be called an edge slide on G.

2.2.5 Remark

Since we may deal with components of G and H with equal size and order separately, when performing the edge slides which transform G into H, it will be sufficient in our discussion of edge slides, to consider only the graphs of $\Gamma_c(p,q)$. We note also, that if there exists no sequence of graphs F_i such that $d_{es}(F_i, F_{i+1}) = 1$ for $i = 0, 1, ..., n-1$, where $F_0 \cong G$ and $F_n \cong H$, then it is usual to say

$$d_{es}(G,H) = \infty.$$

It was shown in [J1] that for any two graphs $A, B \in \Gamma_c(p,q)$ it is possible to transform A into B via a sequence of edge slides. The following result and definition from [J1] will aid us in establishing this.

2.2.6 Lemma

Let $A \in \Gamma_c(p,q)$. Let t be any edge slide on A. Then $tA \in \Gamma_c(p,q)$; i.e., the edge slide preserves order, size and connectivity.

Proof

It is clear that tA has size p and order q since vertices do not undergo any change and edges merely change position. Thus we need only show that tA is connected.

Suppose $t = (u,v,w)$ and let x and y be any distinct vertices of A. Since A is connected there exists a shortest path $P = (x =)x_0x_1...x_n(= y)$ connecting x and y in A. If path P does not pass through the edge uv then the path P connects x and y in tA. Thus we only need consider the case in which uv occurs once in P.

If uv or vu is a subpath of P and vw or wv is not, then construct the walk P' from P by replacing uv or vu by uwv or vwu respectively. If uv or vu is a subpath of P and vw or wv is also a subpath of P, then since P is a shortest path, P contains a subpath of the form uvw or wvu. Form the walk P' by replacing uvw or wvu by uw or wu respectively. In either case P' is a walk connecting x and y in tA. $\square$

2.2.7 Definition

Let $A \in \Gamma_c(p,q)$. We shall say graph G with vertex set $V(G) = \{1, 2, ..., p\}$ and edge set $E(G) = \{e_1, e_2, ..., e_q\}$ is a *standard form of* A if the following conditions are met:

1. $G \cong A$
2. $e_i = ab \Rightarrow a < b$
3. The edges of G are labelled according to the lexicographic ordering. This ordering is obtained by assigning the labels $e_1, e_2, ..., e_{\deg_G 1}$ to the edges incident with 1 (where for $e_{i_1} = 1i$ and $e_{j_1} = 1j$ we have $i_1 < j_1$ if and only if $i < j$), thereafter labelling the edges incident with 2 in a similar fashion, etc. This finally gives that $i < j \Rightarrow e_i < e_j$.

2.2.8 Lemma

For any $A \in \Gamma_c(p,q)$ there exists a standard form G, say, of A.

Proof

Assume there exists no standard form of A. Let G be a graph isomorphic to A with vertex set $V(G) = \{1, 2, ..., p\}$ and edge set $E(G) = \{e_1, e_2, ..., e_q\}$ labelled in such a way that $e_i = ab$ implies that $a < b$ and $e_1 < e_2 < ... < e_k$ for some maximum integer k. Since by assumption G is not a standard form of A, $k < q$ and $e_k > e_{k+1}$. However this implies that

$$e_1 < e_2 < ... < e_{k+1} ... < e_k$$

which, with suitable relabelling contradicts the maximality of k. Therefore $k = q$ and G is a standard form of A. $\square$

The following Theorem is proved in [J1] by means of an adaptation of a method first introduced by Chartrand, Saba and Zou in [CSZ1].

2.2.9 Theorem

For nonisomorphic graphs $A, B \in \Gamma_c(p, q)$, there exists a sequence $t_1, t_2, ..., t_n$ of edge slides such that $t_n ... t_1 A \cong B$.

Proof

For a graph G with vertex set $V(G) = \{1, 2, ..., p\}$ and edge set $E(G) = \{e_1, e_2, ..., e_q\}$, call $S(G) = e_1, e_2, ..., e_q$ the *edge sequence* of G. Let $e_k(G)$ denote the kth edge in $S(G)$. We shall say that G is *1-minimal* if $e_1 = 12$; and for $2 \leq k \leq q$, we shall say that G is *k-minimal* if G is $(k-1)$-minimal with $e_{k-1} = ij$ and $e_k = \begin{cases} i(j+1) & \text{if } j < p \\ (i+1)(i+2) & \text{if } j = p. \end{cases}$

If G and H are both graphs in $\Gamma_c(p, q)$, and if both are q-minimal, then G and H will have the same edge sequence and therefore will be isomorphic. A q-minimal sequence has the following form:

$12, 13, ..., 1p, 23, 24, ..., 2p, ..., (i-1)i, ..., (i-1)p, i(i+1), ..., i(i+m)$ for some integers i and m such that $1 \leq i \leq p-1$ and $1 \leq m \leq p-i$.

We shall show that there exists a standard form G of A, and a sequence $t_1, ..., t_m$ of edge slides such that $t_m...t_2t_1G \cong B$. First we show that there exists a sequence $t_1, t_2, ..., t_n$ of edge slides such that $t_n...t_2t_1G$ is q-minimal.

Assume, to the contrary, that there is a largest k, $k < q$, such that $t'_m...t'_1G$ is k-minimal for some edge slide sequence $t'_1, ..., t'_m$, and some standard form G of A. Let $H = t'_m...t'_1G$. Note that, by Lemma 2.2.6, H is connected. We shall obtain a contradiction by proving the existence of a standard form of H and a sequence $t_1, t_2, ..., t_n$ of edge slides such that $t_n...t_1H$ is $(k+1)$-minimal.

Case 1) Suppose $k < p-1$: Then since H is k-minimal, $e_k(H) = 1(k+1)$, but $e_{k-1}(H) = ij$ where either $i = 1$ and $j > k+2$ or $i > 1$.

There must exist an edge uv where $u \leq k+1 < v$ for otherwise the subgraph induced by the vertex set $\{1, 2, ..., k+1\}$ would form a component of H which would contradict the connectedness property of H. If $v \neq k+2$ form the graph H' by interchanging the labels v and $k+2$; if $v = k+2$ let $H' = H$. Clearly H' is k-minimal. If $u = 1$ then H' is also $(k+1)$-minimal. If $u \neq 1$ then the edge slide $t = (k+2, u, 1)$ exists, since $u(k+2) \in E(H'), (k+2)1 \notin E(H')$ and $u1 \in E(H')$, and tH' is $(k+1)$-minimal.

Case 2) Suppose $p-1 \leq k < q$: Let $e_k(H) = ij$ and $e_{k+1}(H) = uv$ where, as H is not $(k+1)$-minimal,

1. $j < p \Rightarrow e_{k+1}(H) \neq i(j+1)$ and $\sim (1)$
2. $j = p \Rightarrow e_{k+1}(H) \neq (i+1)(i+2)$.

We have therefore that H increases minimally at e_{p-1} and therefore $e_{p-1}(H) = 1p$, and thus the edges $1m$ exist for $m = 2, ..., p$. $\sim$ (2)

Assume $j < p$ and $e_{k+1}(H) = (j+1)v$. Then $t_1 = ((j+1), 1, i)$ exists by (1) and (2) and deletes $1(j+1)$ and creates $i(j+1)$. Also $t_2 = ((j+1), v, 1)$ exists for $t_1 H$, and deletes $(j+1)v$ and recreates $1(j+1)$. Clearly $t_2 t_1 H = H - (j+1)v + i(j+1)$ and it follows that $t_2 t_1 H$ is $(k+1)$- minimal since $e_{k+1}(t_2 t_1 H) = i(j+1)$.

Assume $j < p$ and $e_{k+1}(H) = u(j+1)$. Then $t_1 = ((j+1), 1, i)$ exists on H by (1) and (2). Also $t_2 = ((j+1), u, 1)$ exists on $t_1 H$. Again we clearly have that $t_2 t_1 H = H - u(j+1) + i(j+1)$ and that $t_2 t_1 H$ is $(k+1)$-minimal.

Assume $j < p$ and $e_{k+1}(H) = uv$ where neither u nor v is equal to $j+1$. Then $t_1 = (u, 1, j+1)$ exists on H and $t_2 = (u, v, 1)$ exists on $t_1 H$. Therefore $t_2 t_1 H = H - uv + u(j+1)$ which is k-minimal and has an edge of the form $(j+1)u$ or $u(j+1)$ and thus, as above, $t_2 t_1 H$ and consequently H, can be transformed into a $(k+1)$-minimal graph.

Assume $j = p$. Replacing i by $i+1$ and $j+1$ by $j+2$ in the preceding argument for $j < p$ yields the same result, i.e. that H can be transformed into a $(k+1)$-minimal graph.

We have therefore by contradiction that there exists a sequence $t_1, t_2, ..., t_n$ of edge slides such that $t_n ... t_1 G$ is q-minimal where G is a standard form of

A.

Now for $A, B \in \Gamma_c(p,q)$ we know that there exist edge slide sequences $t_1, t_2, ..., t_n$ and $u_1, u_2, .., u_\ell$ such that, for some standard forms G of A and H of B, we have $t_n...t_1G$ and $u_\ell...u_1H$ are both q-minimal. We note that the *inverse operation of the edge slide* $t = (u,v,w)$ is $t^{-1} = (u,v,w)^{-1} = (u,w,v)$. (If t is defined on G, then $uw \in E(tG)$, $uv \in E(\overline{tG})$ and $wv \in E(tG)$; therefore t^{-1} is defined on tG and $t^{-1}tG = G$.)

We have therefore that $u_1^{-1}u_2^{-1}...u_\ell^{-1}t_n...t_1G \cong H$, which completes the proof. $\square$

The edge slide distance imposes a metric on the set $S^c_{p,q}$ of all isomorphism classes of connected graphs which have p vertices and q edges, as follows: If $\sigma_1, \sigma_2 \in S^c_{p,q}$, then obviously the distance $d_{es}(G_1, G_2)$ is fixed for all $G_1 \in \sigma_1$ and all $G_2 \in \sigma_2$ and is also denoted by $d_{es}(\sigma_1, \sigma_2)$.

2.2.10 Theorem

For any integers $p \geq 1, q \geq 0$, the edge slide distance is a metric on $S^c_{p,q}$.

Proof

Let $\sigma_i \in S^c_{p,q}$ and let $G_i \in \sigma_i$ for $i = 1,2,3$.

i) By definition, $d_{es}(\sigma_1, \sigma_2) \geq 0$ and $d_{es}(\sigma_1, \sigma_2) = 0$ if and only if $d_{es}(G_1, G_2) = 0$, hence if and only if $G_1 \cong G_2$ and $\sigma_1 = \sigma_2$.

ii) If $d_{es}(\sigma_1, \sigma_2) = n$, then $d_{es}(G_1, G_2) = n$ and by definition there exists a sequence $t_1, t_2, ..., t_n$ of edge slides such that $t_n...t_2t_1G_1 \cong G_2$. Consequently $t_1^{-1}t_2^{-1}...t_n^{-1}G_2 \cong G_1$ and so $d_{es}(G_2, G_1) \leq n$; i.e., $d_{es}(\sigma_2, \sigma_1) \leq d_{es}(\sigma_1, \sigma_2)$. A similar argument shows that $d_{es}(\sigma_1, \sigma_2) \leq d_{es}(\sigma_2, \sigma_1)$; hence $d_{es}(\sigma_1, \sigma_2) = d_{es}(\sigma_2, \sigma_1)$.

iii) Let $d_{es}(\sigma_1, \sigma_2) = n$ and $d_{es}(\sigma_2, \sigma_3) = m$; then $d_{es}(G_1, G_2) = n$, $d_{es}(G_2, G_3) = m$ and by definition there exist sequences $t_1, t_2, .., t_n$ and $s_1, s_2, ..., s_m$ of edge slides such that $t_n...t_2t_1G_1 \cong G_2$ and $s_m...s_2s_1G_2 \cong G_3$; hence $s_m...s_2s_1t_n...t_2t_1G_1 \cong G_3$. Therefore $d_{es}(G_1, G_3) \leq n + m$ and so

$$d_{es}(\sigma_1, \sigma_2) \leq d_{es}(\sigma_1, \sigma_2) + d_{es}(\sigma_2, \sigma_3). \qquad \square$$

A useful characterization of the parameters (or properties) of a graph with respect to edge slides was introduced in [BGMW1]. It is useful in that it provides a means of obtaining a lower bound on the edge slide distance between certain graphs and between certain classes of graphs. It is formulated as follows:

2.2.11 Definition

A parameter ψ is said to be *slowly changing* with respect to the edge slide operation if and only if for all graphs $G, H \in \Gamma_c(p, q)$, $d_{es}(G, H) = 1$ implies that $|\psi(G) - \psi(H)| \leq 1$.

For example, suppose $d_{es}(G, H) = 1$ for graphs $G, H \in \Gamma_c(p, q)$. Now since an edge slide can only either increase or decrease the degree of a vertex in G by at most one, only one of the following can hold:

i) $\triangle(H) = \triangle(G)$ ii) $\triangle(H) = \triangle(G) - 1$ iii) $\triangle(H) = \triangle(G) + 1$;

$$\text{i.e., } |\triangle(H) - \triangle(G)| \leq 1.$$

Hence the following proposition has been established.

2.2.12 Proposition

The maximum degree $\triangle(G)$ of a graph G is a slowly changing parameter with respect to the edge slide operation.

The following proposition is immediately obvious and will be of use in determining the edge slide distance between specific pairs of graphs. Our next two results are from [BGMW1].

2.2.13 Proposition

If ψ is a slowly changing parameter with respect to the edge slide operation and if $G, H \in \Gamma_c(p,q)$ are such that

$$|\psi(G) - \psi(H)| = n$$

then $d_{es}(G, H) \geq n$, giving a lower bound on the edge slide distance between G and H.

2.2.14 Theorem

The edge slide distance between the star $S_n = K_{1,n-1}$, on n vertices, and a given tree T on n vertices is

$$\triangle(S_n) - \triangle(T).$$

Proof

Since S_n and T are trees on n vertices we have $S_n, T \in \Gamma_c(n, n-1)$ and thus by Theorem 2.2.9 we can transform S_n into T by means of a sequence of edge slides (or vice versa). From Propositions 2.2.12 and 2.2.13 we have

$$d_{es}(S_n, T) \geq \triangle(S_n) - \triangle(T) \qquad \sim (1)$$

Now let v be a vertex of maximum degree in T. If $\deg_T v = n-1$ then $T \cong S_n$, $d_{es}(S_n, T) = 0$ and $\triangle(S_n) - \triangle(T) = 0$, which satisfies (1). So assume T is not a star and $\deg_T v = d$ $(1 \leq d \leq n-2)$; i.e., $\triangle(T) = d$. Then there exists a vertex $w \in V(T)$ which is not adjacent to v but is adjacent to a vertex x in the neighbourhood of v. Now the edge slide $t_1 = (w, x, v)$ increases the degree of v by 1. Since by Lemma 2.2.6 the edge slide operation preserves connectivity, we have that t_1T is a tree, and $\triangle(t_1T) = \triangle(T) + 1$.

By the same argument if t_1T is not a star then $\triangle(t_1T) \leq n-2$ and there exists an edge slide t_2 such that $\triangle(t_2t_1T) = \triangle(t_1T) + 1 = \triangle(T) + 2$ and t_2t_1T is a tree.

Repeating the procedure $((n-1)-d)$ times gives a tree $t_{(n-1)-d}...t_1T$ where $\triangle(t_{(n-1)-d}...t_1T) = \triangle(T) + (n-1) - d = d + (n-1) - d = n-1$; i.e., $t_{(n-1)-d}...t_1T \cong S_n$, and

$$d_{es}(T, S_n) \leq (n-1) - d = \triangle(S_n) - \triangle(T) \qquad \sim (2)$$

Therefore (1) and (2) imply that $d_{es}(S_n, T) = \triangle(S_n) - \triangle(T)$. $\square$

The following corollary to Theorem 2.2.14 appears in [Z3].

2.2.15 Corollary

The edge slide distance between a path P_n and S_n is $n-3$.

Proof

In this case, $\triangle(S_n)-\triangle(P_n) = n-1-2 = n-3$, hence $d_{es}(S_n, P_n) = n-3$. □

We recall the following terminology. If $G \in \Gamma_c(p,q)$ and diam $G = d$, a pair of vertices $u, v \in V(G)$ is said to be a diametrical pair of vertices of G if $d(u,v) = d$. Every $u-v$ path of length d is called a diametrical $u-v$ path.

The next five results appear in [BGMW1].

2.2.16 Lemma

The diameter of a graph $G \in \Gamma_c(p,q)$ is slowly changing with respect to the edge slide operation.

Proof

Let $G \in \Gamma_c(p,q)$ and let $H = tG$, where $t = (x,y,z)$ is an edge slide. If diam $H >$ diam G, then there exists a diametrical pair of vertices $\{u,v\}$ of G such that each diametrical $u-v$ path contains the edge xy and not the vertex z. Hence a shortest $u-v$ path is obtained in H from a diametrical $u-v$ path in G by replacing xy with xzy and so

$$\text{diam}H = \text{diam}G + 1.$$

If diam $H <$ diam G, then for each diametrical pair of vertices $\{u,v\}$ of G there exists a diametrical $u-v$ path P in G which contains the subpath

xyz. A shortest $u - v$ path is obtain in H by replacing the subpath xyz by the edge xz in P and hence

$$\text{diam}H = \text{diam}G - 1.$$

Thus $d_{es}(G, H) = 1$ implies that $|\text{diam}G - \text{diam}H| \leq 1$ and the lemma is proved. □

2.2.17 Theorem

For the path P_n and a tree T on n vertices, $d_{es}(P_n, T) = \text{diam}P_n - \text{diam}T$.

Proof

Since $P_n, T \in \Gamma_c(n, n-1)$ we have by Theorem 2.2.9, that P_n can be transformed into T via a sequence of edge slides (and vice versa).

From Lemma 2.2.16 the diameter of a graph is a slowly changing parameter with respect to the edge slide operation, therefore from Proposition 2.2.13 we have

$$d_{es}(P_n, T) \geq \text{diam}P_n - \text{diam}T. \qquad \sim (1)$$

The case $T \cong P_n$ is trivial, so assume T is not a path; thus diam $T \leq n - 2$. Let P be a longest path in T; i.e., of length $\ell = \text{diam}T$. Let $P = x_0x_1...x_\ell$. Now since T is connected and not a path there exists a vertex $w \notin V(P)$ and a vertex $x_i (1 \leq i \leq \ell - 1)$ such that $wx_i \in E(T)$. Perform the edge slide $t_1 = (x_{i-1}, x_i, w)$ then the tree t_1T has a longest path $x_0...x_{i-1}wx_i...x_\ell$ of length $\ell + 1$; i.e., diam $t_1T = \ell + 1$. Similarly, if t_1T is not a path, then there exists an edge slide t_2 which will give a tree t_2t_1T with diam $t_2t_1T = \ell + 2$. Repeat this process $(n-1) - \ell$ times to obtain a tree $t_{n-1-\ell}...t_2t_1T$ with diam $t_{n-1-\ell}...t_2t_1T = \ell + (n-1) - \ell = n-1$. Since T has only $n-1$ edges this implies that $t_{(n-1)-\ell}...t_2t_1T \cong P_n$, and

$$d_{es}(P_n, T) \leq (n-1) - \ell = \text{diam}P_n - \text{diam}T \qquad \sim (2)$$

Together (1) and (2) yield

$$d_{es}(P_n, T) = \text{diam}P_n - \text{diam}T. \qquad \square$$

The following lemma will aid us in establishing the edge slide distance between the n-cycle C_n and any unicyclic graph U of order n.

2.2.18 Lemma

The girth $g(G)$ of a connected graph G is slowly changing with respect to the edge slide operation.

Proof

Let the graph G' be obtained from the graph G by an edge slide. By symmetry, G may be obtained from G' by an edge slide. Without loss of generality, let $g(G) \leq g(G')$. Let C be a shortest cycle in G. Let e be the edge that was removed from G to form G'. If e does not appear on the cycle C, then C is a cycle in G', and thus $g(G') \leq g(G)$. If e does appear on C, then $e = xy$, say, was removed from G and a new edge $e' = xm$, say, was added to G to form G', where $x, y, m \in V(G)$, $xm \in E(\overline{G})$ and $ym \in E(G)$. By removing e from C and replacing it with the path xmy one obtains a closed trail in G'. Thus, in either case $g(G') \leq g(G) + 1$ and hence $|g(G') - g(G)| \leq 1$, implying that the girth of a connected graph is a slowly changing parameter. $\square$

2.2.19 Theorem

The edge slide distance between the n-cycle C_n and any unicyclic graph U of order n is given by $g(C_n) - g(U)$; i.e.,

$$d_{es}(U, C_n) = g(C_n) - g(U).$$

Proof

By Lemma 2.2.18 we observe immediately that $d_{es}(U, C_n) \geq g(C_n) - g(U)$. If U has girth n then $U \cong C_n$ and the theorem holds trivially.

Assume then, that U does not have girth n. Let $C = x_1x_2...x_gx_1$ be the unique cycle in U where $g \leq n-1$. Since U is connected there exists a vertex $y \in U$ which is not on C but is adjacent to a vertex x_j say, of C. Form the graph U' by performing the edge slide $t = (x_{j-1}, x_j, y)$ on U; i.e., $U' = tU$. Now U' is unicyclic such that $g(U') = g(U) + 1$. Repeating this procedure $g(C_n) - g(U)$ times will result in a graph which is isomorphic to C_n. Thus $d_{es}(U, C_n) \leq g(C_n) - g(U)$ and hence

$$d_{es}(U, C_n) = g(C_n) - g(U). \qquad \square$$

2.2.20 Theorem

For every nonnegative integer n, there exist graphs $G_1, G_2 \in \Gamma_c(p, q)$ such that $d_{es}(G_1, G_2) = n$.

Proof

If $n = 0$ then for any connected graph $G, d_{es}(G, G) = 0$ so that $G_1 = G_2 = G$.

Assume then that $n \geq 1$ is given. Then we construct G_1 as follows: Take two disjoint paths $P_{2n+2} = x_1 x_2 ... x_{2n+2}$ and $P_{n+2} = y_1 y_2 ... y_{n+2}$. Then let $G_1 = P_{2n+2} \cup P_{n+2} + y_{n+2} x_{n+2}$ and let $G_2 = P_{3n+4}$. Therefore G_1 and G_2 are connected and have order $3n + 4$ and size $3n + 3$.

Now let the edge slide t_i be given by $t_i = (y_{n+2}, x_{n+2+i}, x_{n+3+i})$ for $i = 0, 1, ..., n-1$. Then $t_{n-1} t_{n-2} ... t_1 t_0 G_1 \cong G_2$ and hence

$$d_{es}(G_1, G_2) \leq n. \qquad \sim (1)$$

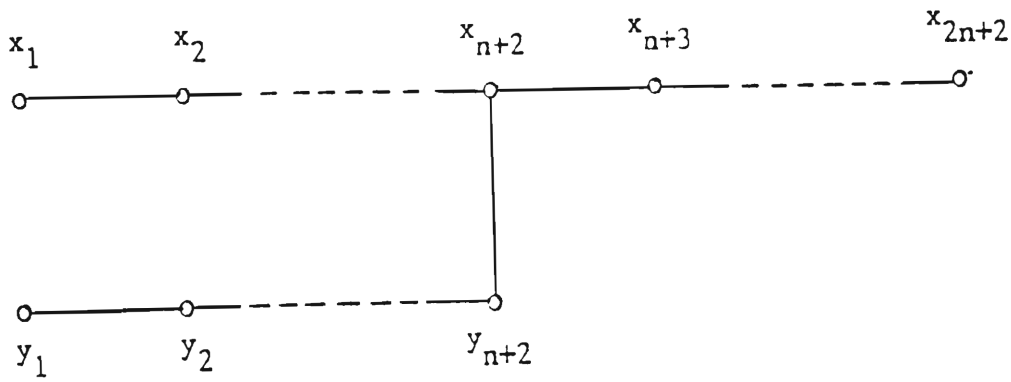
By Lemma 2.2.16, the diameter of a graph is a slowly changing parameter with respect to edge slides. Now diam $G_1 = 2n + 3$ and diam $G_2 = 3n + 3$. Therefore, by Proposition 2.2.13, we have

$$d_{es}(G_1, G_2) \geq |\text{diam} G_2 - \text{diam} G_1| = |(3n + 3) - (2n + 3)| = n. \qquad \sim (2)$$

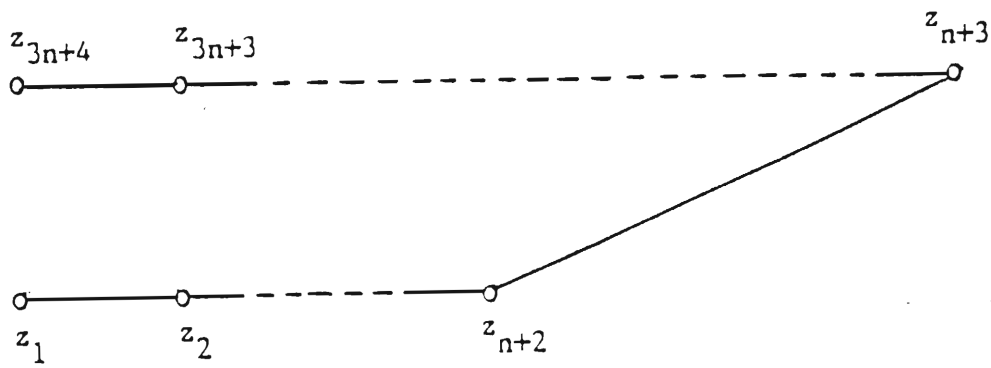
Thus from (1) and (2) we have

$$d_{es}(G_1, G_2) = n. \qquad \square$$

G_1:



G_2:



2.2.21 Figure

The graphs G_1 and G_2 of Theorem 2.2.20

2.2.22 Definition

Let $S = \{s_1, s_2, ..., s_n\}$ be a set of (nonisomorphic) graphs having the same number of components which are labelled in such a way that the ith component $(1 \leq i \leq n)$ of all graphs in S have the same size and order. Then following Chartrand, Goddard, Henning, Lesniak, Swart and Wall in [CGHLSW1], the *edge slide distance graph* $D_s(S)$ of S is defined to be that graph with vertex set S such that two vertices s_i and s_j of $D_s(S)$ are adjacent if and only if $d_{es}(s_i, s_j) = 1$ for $1 \leq i, j \leq n$.

This definition leads naturally to the question: which graphs are edge slide distance graphs? This question was answered in [CGHLSW1].

2.2.23 Theorem

Every graph is an edge slide distance graph.

Proof

Suppose we are given an arbitrary graph G with $V(G) = \{v_1, v_2, ..., v_p\}$. Let H be the graph obtained from G by adding two new vertices each adjacent only to v_1, four new vertices each adjacent only to v_2 and, in general, $2i$ new vertices each adjacent only to v_i for $i = 1, 2, ..., p$. Then for $i = 1, 2, ..., p$, let H_i be the graph obtained from H by adding another new vertex u_i adjacent only to v_i. In H_i we now have that the only vertices that are not end-vertices are those originally in G; i.e., $\{v_1, v_2, ..., v_p\}$. Therefore H_i contains exactly p vertices that are adjacent to end-vertices of H_i. In fact, for H_i, the sequence which displays the number of end-vertices adjacent to the vertices $v_1, v_2, ..., v_p$ is respectively

$$2, 4, ..., 2i - 2, 2i + 1, 2i + 2, ..., 2p.$$

For $j > i$ the analogous sequence for H_j is

$$2, 4, ..., 2i - 2, 2i, 2i + 2, ..., 2j - 2, 2j + 1, 2j + 2, ..., 2p$$

It is now easy to see by looking at these "end-vertex degree" sequences that $d_{es}(H_i, H_j) = 1$ if and only if H_j can be obtained from H_i by the edge slide $t = (u_i, v_i, v_j)$ which exists if and only if $v_i v_j \in E(G)$. Thus $H_i H_j \in E(D_s\{H_1, H_2, ..., H_p\})$ if and only if $v_i v_j \in E(G)$;

$$\text{i.e., } G \cong D_s(\{H_1, H_2, ..., H_p\}) \qquad \square$$

2.3 The Edge Rotation Distance d_{er}

2.3.1 Definition

Let G and H be two graphs having the same order and size. Then, following Chartrand, Saba and Zou in [CSZ1], we say that G can be *transformed into* H by an *edge rotation* if G contains distinct vertices u, v and w such that $uv \in E(G)$, $uw \in E(\bar{G})$ and $H \cong G - uv + uw$.

The edge rotation is similar to the edge slide in that the edge uv is deleted and the edge uv is created, but the restriction $vw \in E(G)$ is dropped in the definition of edge rotation. This is demonstrated by looking again at Figures 2.2.3 and 2.2.4 in Section 2.2. In both Figures 2.2.3 and 2.2.4 it is possible to transform the graph G into the graph H by an edge rotation, whereas only in Figure 2.2.4 is it possible to transform graph G into graph H by an edge slide.

2.3.2 Definition

For graphs G and H of the same order and same size, the *edge rotation distance* $d_{er}(G, H)$ between G and H is the smallest nonnegative integer n for which there exists a sequence

$$G \cong H_0, H_1, H_2, ..., H_n \cong H,$$

such that H_i can be transformed into H_{i+1} $(i = 0, 1, ..., n-1)$ by an edge rotation.

As for edge slides we define $t = (u, v, w)$ to be an edge rotation which deletes edge uv and creates uw when operating on some graph G, where $uv \in E(G)$ and $uw \in E(\overline{G})$. We again denote the graph $G - uv + uw$ by tG. The inverse operation of the edge rotation $t = (u, v, w)$ is denoted by $t^{-1} = (u, w, v)$ which reverses the operation performed by t; i.e., it creates uv and deletes uw; so $t^{-1}tG = G$, as required.

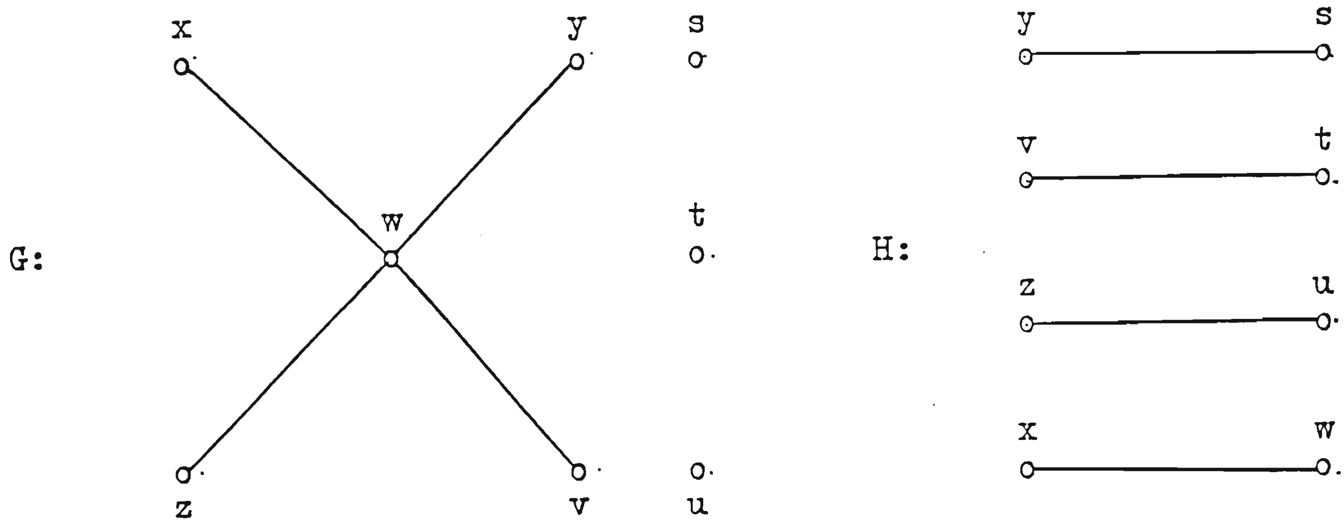
It is immediately obvious that any edge slide is an edge rotation and therefore for $G, H \in \Gamma_c(p, q)$, the following proposition needs no further justification.

2.3.3 Proposition

$$d_{er}(G, H) \leq d_{es}(G, H).$$

We also note that by dropping the restriction $vw \in E(G)$ as described above, we are no longer restricted to studying distances between connected graphs of the same order and size; hence we can dispense with restriction of connectedness. We may therefore consider distances between all pairs of graphs of the same order and the same size.

2.3.4 Example



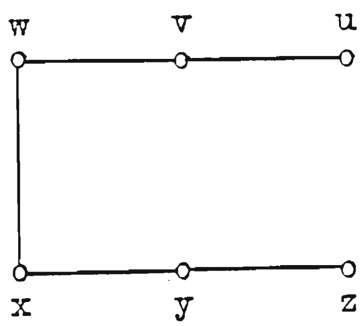
2.3.5 Figure

Let $t_1 = (y, w, s)$, $t_2 = (v, w, t)$ and $t_3 = (z, w, u)$ be edge rotations. Then considering the graphs G and H of Figure 2.3.5, we have $t_3 t_2 t_1 G \cong H$.

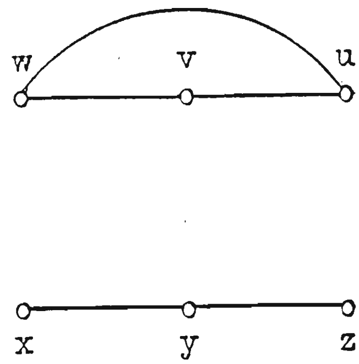
The following example demonstrates that the edge rotation operation does not, in general, preserve connectivity.

2.3.6 Example

G:



H:



2.3.7 Figure

In Figure 2.3.7 we see that $tG \cong H$ where t is the edge rotation (w, x, u). Since G is connected and H is not connected, this illustrates the fact that edge rotations unlike edge slides do not in general preserve connectivity.

It was shown in [CSZ1] that for any two nonisomorphic graphs G and H which have the same order and size it is always possible to transform G into H via a finite sequence of edge rotations. The proof is essentially similar to the proof of Theorem 2.2.9, however we shall present it here in a slightly different way.

2.3.8 Theorem

Let $G_1, G_2 \in \Gamma(p,q)$, where $p \geq 4$ and $q \geq 2$; then there exists a sequence $t_1, t_2, ..., t_n$ of edge rotations such that $t_n ... t_2 t_1 G_1 \cong G_2$.

Proof

The theorem holds trivially for $G_1 \cong G_2$; so suppose that $G_1 \ncong G_2$. Assume, without loss of generality, that G_1 and G_2 are defined on the same vertex set; i.e., $V(G_1) = V(G_2) = \{1, 2, ..., p\}$. Let G_1 have edge set $E(G_1) = \{e_1, e_2, ..., e_q\}$.

Now, by Lemma 2.2.8 there exists a standard form H say, of G_1 and a standard form H' say, of G_2. We now show that there exists a sequence $t_1, t_2, ..., t_n$ of edge rotations such that $t_n ... t_1 H$ is q-minimal.

Assume that there exists no such sequence; then let $k (k < q)$ be the maximum positive integer for which there exists a sequence $t'_1, t'_2, ..., t'_m$ such that $t'_m ... t'_2 t'_1 H$ is k-minimal. Let $F = t'_m ... t'_2 t'_1 H$. Since F is not q-minimal there exists edges ab and cd such that $ab \in E(\bar{F})$ while $cd \in E(F)$ and $ab < cd$.

Case 1) Suppose $a = c$. Then $b < d$. Let t be the edge rotation given by (a, d, b). Then tF is $(k+1)$-minimal which contradicts our assumption and therefore there exists a sequence $t_1, t_2, ..., t_m$ of edge rotations such that $t_m ... t_2 t_1 H$ is q-minimal.

Case 2) Suppose $a < c$. If $b = d$ or $b = c$ then as in Case 1 the edge rotations $t = (b, c, a)$ or $t = (c, d, a)$ respectively, show in both cases that tF is

$(k+1)$-minimal.

Assume then that $b \neq d$ and $b \neq c$ so that a, b, c and d are four distinct vertices. Now if $bd \in E(\bar{F})$ then the edge rotations $t_1 = (d, c, b)$ and $t_2 = (b, d, a)$ operating respectively on the graphs F and t_1F, result in t_2t_1F being $(k+1)$-minimal.

If $bd \in E(F)$ then the edge rotations $t_1 = (b, d, a)$ and $t_2 = (d, c, b)$ again result in t_2t_1F being $k+1$-minimal.

Thus, as in Case 1, we obtain a contradiction and thus there must exist a sequence $t_1, t_2, ..., t_m$ of edge rotations such that $t_m...t_2t_1H$ is q-minimal.

Similarly there exists a sequence $u_1, u_2, ..., u_\ell$ of edge rotations such that $u_\ell...u_2u_1H'$ is q-minimal which, since any two q-minimal graphs are isomorphic yields

$$u_1^{-1}u_2^{-1}...u_\ell^{-1}t_m...t_2t_1H \cong H'. \text{ Since } H \cong G_1 \text{ and } H' \cong G_2,$$

we have that there exists a sequence $t_1, t_2, ..., t_n$ of edge rotations such that

$$t_n...t_1G_1 \cong G_2. \qquad \square$$

The edge rotation distance imposes a metric on the set $S_{p,q}$ of all isomorphism classes of graphs, which have p vertices and q edges, as follows: If $\sigma_1, \sigma_2 \in S_{p,q}$, then obviously the distance $d_{er}(G_1, G_2)$ is fixed for all $G_1 \in \sigma_1$ and all $G_2 \in \sigma_2$ and is also denoted by $d_{er}(\sigma_1, \sigma_2)$.

2.3.9 Theorem

For any integers $p \geq 1, q \geq 0$, the edge rotation distance is a metric on $S_{p,q}$.

Proof

Let $\sigma_i \in S_{p,q}$ and let $G_i \in \sigma_i$ for $i = 1, 2, 3$.

i) By definition, $d_{er}(\sigma_1, \sigma_2) \geq 0$ and $d_{er}(\sigma_1, \sigma_2) = 0$ if and only if $d_{er}(G_1, G_2) = 0$, hence if and only if $G_1 \cong G_2$ and $\sigma_1 = \sigma_2$.

ii) If $d_{er}(\sigma_1, \sigma_2) = n$, then $d_{er}(G_1, G_2) = n$ and by definition there exists a sequence $t_1, t_2, ..., t_n$ of edge rotations such that $t_n...t_2t_1G_1 \cong G_2$. Consequently $t_1^{-1}t_2^{-1}...t_n^{-1}G_2 \cong G_1$ and so $d_{er}(G_2, G_1) \leq n$; i.e., $d_{er}(\sigma_2, \sigma_1) \leq d_{er}(\sigma_1, \sigma_2)$. A similar argument shows that $d_{er}(\sigma_1, \sigma_2) \leq d_{er}(\sigma_2, \sigma_1)$, hence $d_{er}(\sigma_1, \sigma_2) = d_{er}(\sigma_2, \sigma_1)$.

iii) Let $d_{er}(\sigma_1, \sigma_2) = n$ and $d_{er}(\sigma_2, \sigma_3) = m$, then $d_{er}(G_1, G_2) = n$ and $d_{er}(G_2, G_3) = m$ and by definition there exist sequences of edge rotations $t_1, t_2, ..., t_n$ and $s_1, s_2, ..., s_m$ such that $t_n...t_2t_1G_1 \cong G_2$ and $s_m...s_2s_1G_2 \cong G_3$; hence $s_m...s_2s_1t_n...t_2t_1G_1 \cong G_3$. Therefore $d_{er}(G_1, G_3) \leq n + m$ and so

$$d_{er}(\sigma_1, \sigma_3) \leq d_{er}(\sigma_1, \sigma_2) + d_{er}(\sigma_2, \sigma_3). \qquad \square$$

The following interesting observation concerning complements of graphs was made in [CSZ1].

2.3.10 Theorem

For graphs $G_1, G_2 \in \Gamma(p, q)$, $d_{er}(G_1, G_2) = d_{er}(\bar{G}_1, \bar{G}_2)$.

Proof

If $d_{er}(G_1, G_2) = 0$, then $G_1 \cong G_2$, and thus $\bar{G}_1 \cong \bar{G}_2$ which implies $d_{er}(\bar{G}_1, \bar{G}_2) = 0$, which satisfies the statement of the theorem. Assume then that $d_{er}(G_1, G_2) = n \geq 1$. This implies that there exists a sequence of graphs

$$G_1 \cong H_0, H_1, ..., H_n \cong G_2$$

where H_i can be transformed into H_{i+1} by an edge rotation for $i = 0, 1, ..., n-1$. Let $H_{i+1} = H_i - u_i v_i + u_i w_i$. Then note that $\bar{H}_{i+1} = \bar{H}_i - u_i w_i + u_i v_i$; i.e., $\bar{H}_i$ can be transformed into $\bar{H}_{i+1}$ by an edge rotation. Thus the sequence of graphs

$$\bar{G}_1 \cong \bar{H}_0, \bar{H}_1, ..., \bar{H}_n \cong \bar{G}_2$$

has the property that $d_{er}(\bar{H}_i, \bar{H}_{i+1}) = 1$ for $i = 0, 1, ..., n-1$. This implies that

$$d_{er}(\bar{G}_1, \bar{G}_2) \leq d_{er}(G_1, G_2) = n. \qquad \sim (1)$$

However, by (1)

$$d_{er}(\bar{\bar{G}}_1, \bar{\bar{G}}_2) \leq d_{er}(\bar{G}_1, \bar{G}_2)$$

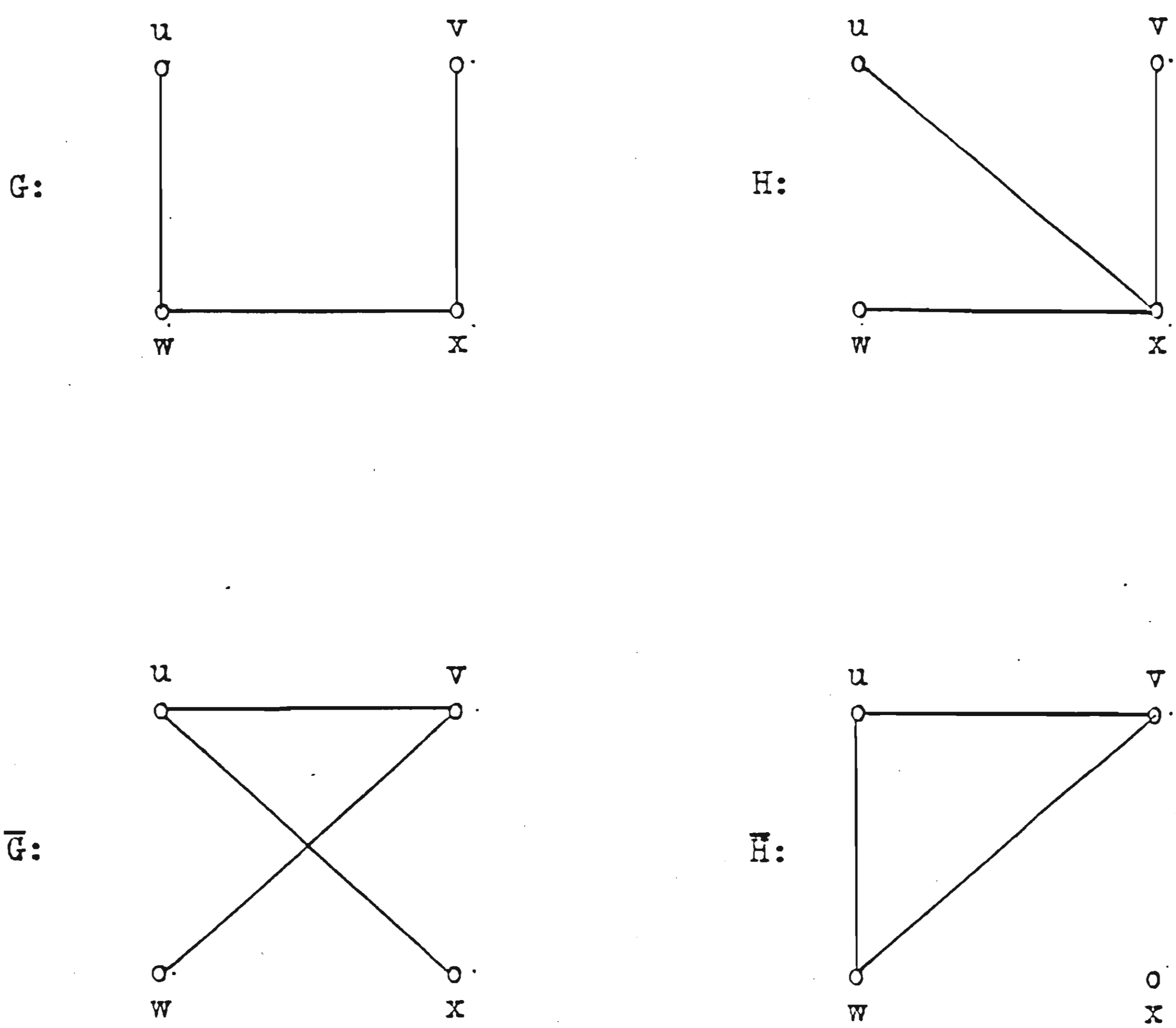
which implies

$$d_{er}(G_1, G_2) \leq d_{er}(\bar{G}_1, \bar{G}_2). \qquad \sim (2)$$

Thus (1) and (2) together imply $d_{er}(G_1, G_2) = d_{er}(\bar{G}_1, \bar{G}_2)$. $\square$

The following example demonstrates that there is no similar result to Theorem 2.3.10 for edge slides.

2.3.11 Example



2.3.12 Figure

Referring to the graphs $G, \bar{G}, H$ and $\bar{H}$ of Figure 2.3.12, the edge slide $t = (u, w, x)$ yields

$$tG \cong H.$$

However $\bar{G}$ has one component whereas $\bar{H}$ has two components. Since, by Lemma 2.2.6, the edge slide operation preserves connectivity, it is not even possible to transform $\bar{G}$ into $\bar{H}$; i.e., $d_{es}(\bar{G}, \bar{H}) = \infty$. Thus, in general, $d_{es}(G, H) \neq d_{es}(\bar{G}, \bar{H})$.

Both the edge slide and edge rotation operations may be considered to be deformations which translate a graph G into a graph $G' = G - e_1 + e_2$ where $e_1 \in E(G)$ and $e_2 \in E(\bar{G})$. The edge move distance which will be defined in Chapter 3 also falls into this category. Definition 2.2.11 is generalized in [BGMW1] to include all such operations.

2.3.13 Definition

We say that a parameter ψ is *slowly changing* with respect to a particular deformation if and only if for all graphs G and deformations G' of G it holds that $|\psi(G') - \psi(G)| \leq 1$.

We now formalize the technique used in the proofs of Theorems 2.2.14, 2.2.17 and 2.2.19, in the form of a lemma. This will simplify the work in determining some specific formulae for distances. The following two results are from [BGMW1].

2.3.14 Lemma

Let $\mathcal{G}$ be a collection of graphs and let $F \in \mathcal{G}$ be a designated element. Further, let μ be an integer valued graphical parameter and consider a particular deformation. Then for that deformation, with distance between graphs G and H denoted by $\delta(G, H)$, it holds that:

$$\delta(F, G) = |\mu(G) - \mu(F)|, \text{ for all } G \in \mathcal{G},$$

if the following three properties are satisfied:

P1 The parameter μ is slowly changing with respect to that particular deformation;

P2 F is the only element of $\mathcal{G}$ with that value of μ; and

P3 given any $G \in \mathcal{G}$ with $\mu(G) \neq \mu(F)$ there exists a deformation (of the required type) yielding $G' \in \mathcal{G}$ such that $|\mu(G') - \mu(F)| < |\mu(G) - \mu(F)|$.

Proof

Property P1 establishes that $|\mu(G) - \mu(F)|$ is a lower bound, while properties P2 and P3 together show that the value $|\mu(G) - \mu(F)|$ is an upper bound for the distance. □

2.3.15 Lemma

The maximum degree $\Delta(G)$ of a graph G is slowly changing with respect to the edge rotation operation.

Proof

Any edge rotation $t = (u, v, w)$ lowers the degree of the vertex $v \in V(G)$ by 1, and increases the degree of the vertex $w \in V(G)$ by one, when operating on some graph G with $uv \in E(G)$ and $uw \in E(\bar{G})$. Hence $| \Delta\ (tG) - \Delta(G)| \leq 1$, and the maximum degree $\Delta(G)$ of a graph G is slowly changing with respect to edge rotations. $\square$

The following result appears in [CSZ1].

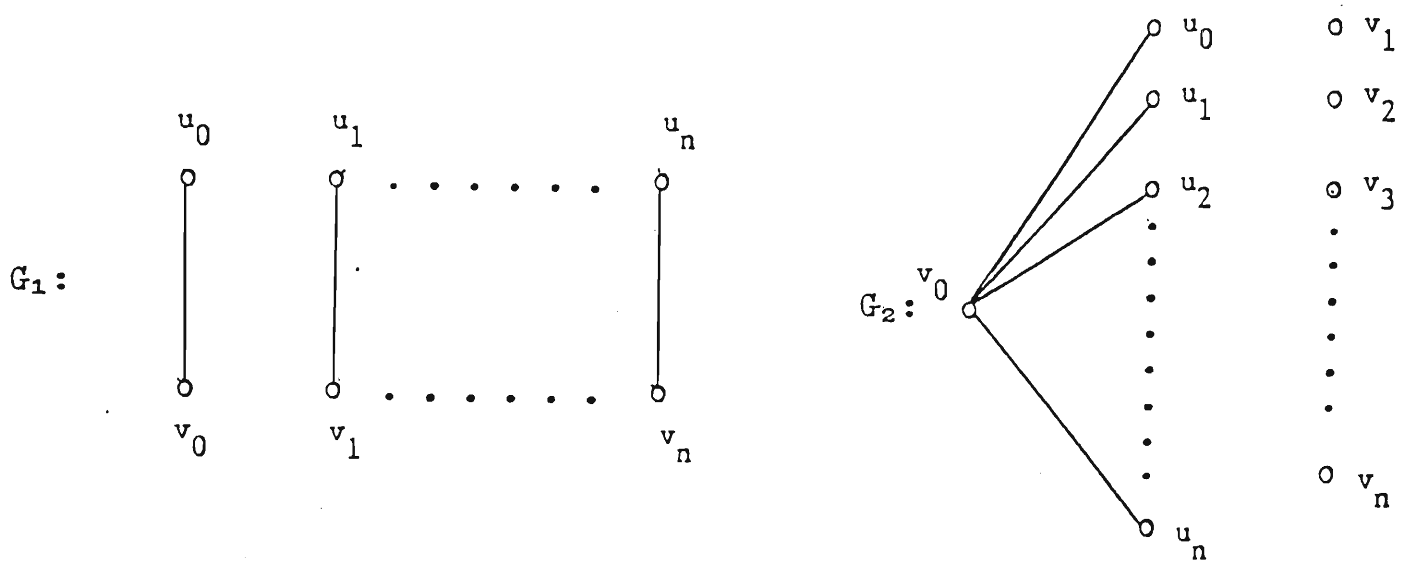
2.3.16 Theorem

For every nonnegative integer n, there exist graphs G_1 and G_2 such that $d_{er}(G_1, G_2) = n$.

Proof

If $n = 0$ then for every graph $G, d_{er}(G, G) = 0$; so let $G_1 = G_2 = G$ in this case.

If n is a given positive integer let $G_1 = (n+1)K_2$ and $G_2 = K_{1,n+1} \cup nK_1$, so that G_1 and G_2 are graphs of order $2n + 2$ and size $n + 1$. Let the edge set of G_1 be given by $E(G_1) = \{u_0v_0, u_1v_1, ..., u_nv_n\}$.



2.3.17 Figure

The graphs G_1 and G_2 of Theorem 2.3.16.

For $i = 1, 2, ..., n$ define the edge rotation $t_i = (u_i, v_i, v_0)$; then $t_n ... t_1 G_1 \cong G_2$ which implies that

$$d_{er}(G_1, G_2) \leq n. \qquad \sim (1)$$

By Lemma 2.3.15, the maximum degree $\triangle(G)$ of a graph G is slowly changing with respect to edge rotations. Now $\triangle(G_2) = n+1$ while $\triangle(G_1) = 1$, and hence a simple generalization of Proposition 2.2.13 to include edge rotations yields

$$d_{er}(G_1, G_2) \geq |\triangle(G_2) - \triangle(G_1)| = n. \qquad \sim (2)$$

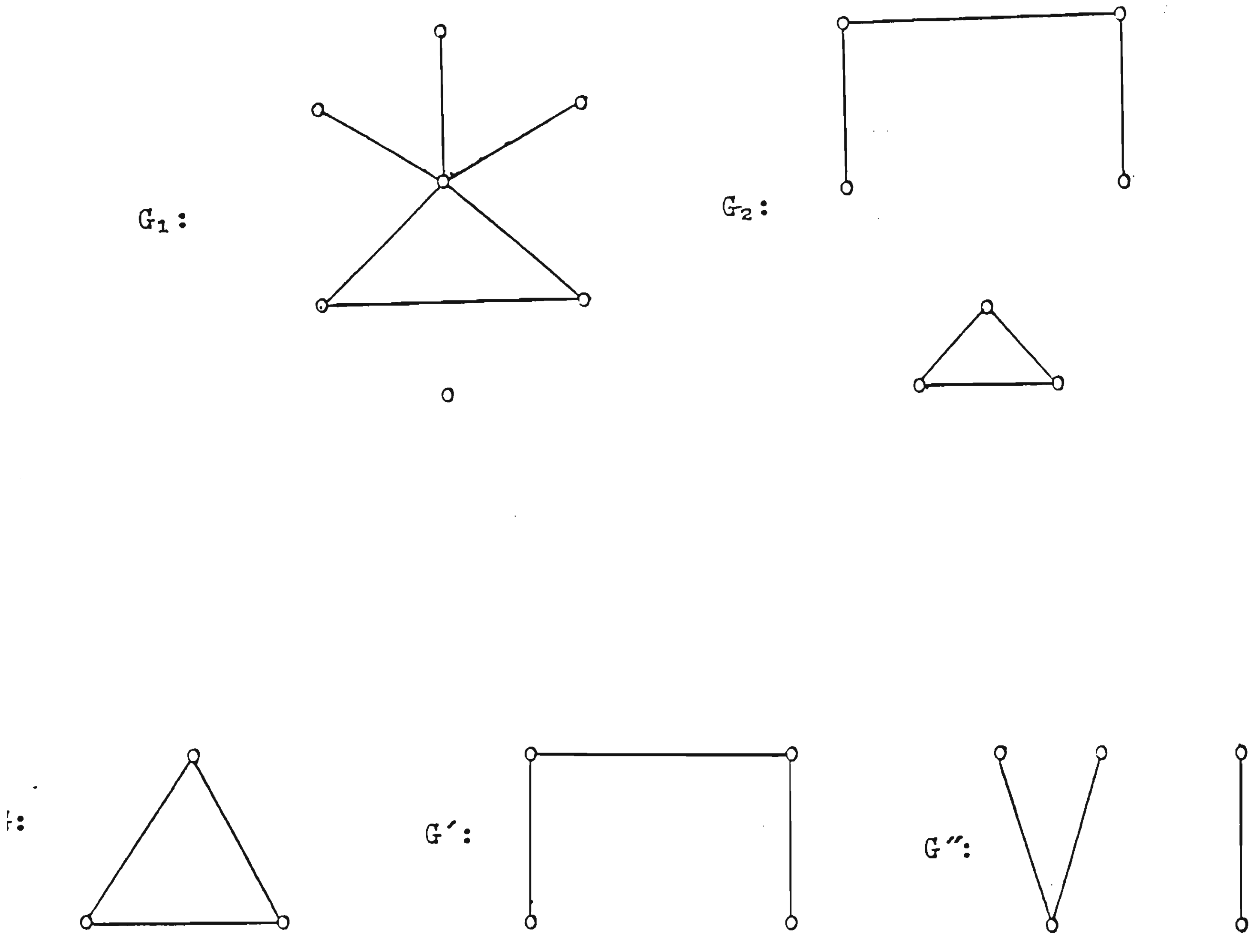
Therefore (1) and (2) together imply that $d_{er}(G_1, G_2) = n$. $\square$

In order to present upper and lower bounds on the edge rotation distance between graphs (having the same order and size), we introduce the concept of a *greatest common subgraph* which first appeared in [CSZ1].

2.3.18 Definition

For nonempty graphs G_1 and G_2 a *greatest common subgraph* of G_1 and G_2 is defined as any graph G of maximum size without isolated vertices, that is a subgraph of both G_1 and G_2.

While every pair of graphs G_1 and G_2 of nonempty graphs has a greatest common subgraph G say, this graph G need not be unique. For example the graphs G_1 and G_2 shown in Figure 2.3.19 below, have three greatest common subgraphs G, G' and G''. These graphs are all pairwise nonisomorphic but all have the same maximum size, 3.



2.3.19 Figure

Utilising this concept we now prove the following result from [CSZ1] which sets upper and lower bounds for the edge rotation distance.

2.3.20 Theorem

Let $G_1, G_2 \in \Gamma(p,q), q \geq 1$ and let G be a greatest common subgraph of G_1 and G_2, where G has size s say. Then

$$q - s \leq d_{er}(G_1, G_2) \leq 2(q - s).$$

Proof

First we prove $d_{er}(G_1, G_2) \geq q - s$. Let G_1 and G_2 be defined on the same vertex set, so that the subgraphs G' and G'' of G_1 and G_2 respectively, which are isomorphic to G, are identically labelled; i.e., $V(G_1) = V(G_2) = \{v_1, v_2, ..., v_p\}$ and $V(G') = V(G'') = \{v_{i_1}, ..., v_{i_s}\}$ with $v_{i_j}v_{i_k} \in E(G')$ if and only if $v_{i_j}v_{i_k} \in E(G'')$.

Now $E(G_1) - E(G_2)$ contains $q - s$ edges; similarly $E(G_2) - E(G_1)$ contains $q - s$ edges. Therefore in the transformation of G_1 into G_2 via edge rotations at least one edge rotation will be needed for each of the $q-s$ steps in replacing an edge of $E(G_1) - E(G_2)$ by an edge of $E(G_2) - E(G_1)$. Therefore

$$d_{er}(G_1, G_2) \geq q - s. \qquad \sim (1)$$

For the upper bound $d_{er}(G_1, G_2) \leq 2(q - s)$ we note that if $s = q$ then $G_1 \cong G_2$ and $d_{er}(G_1, G_2) = 0$. Thus we assume that $1 \leq s < q$. Let G_1 and G_2 be labelled as before. Now, since $G_1 \ncong G_2$, the graph G_1 contains an edge $v_iv_j \notin E(G_2)$ and G_2 contains an edge $v_kv_\ell \notin E(G_1)$. We now show that the step of transforming G_1 into $G_1 - v_iv_j + v_kv_\ell = H_1$ requires at most two edge rotations, and since there are $q - s$ such transformations necessary, the result will follow immediately.

Case 1) Suppose that $\{v_i, v_j\} \cap \{v_k, v_\ell\} \neq \emptyset$; say $v_j = v_k$. Then G_1 can be transformed into $H_1 = G_1 - v_iv_j + v_jv_\ell$ by a single edge rotation $t = (v_j, v_i, v_\ell)$ and $d(G_1, H_1) = 1$. Hence we may assume that $\{v_k, v_j\} \cap \{v_k, v_\ell\} = \emptyset$.

Case 2) Suppose that at least one of v_i and v_j is nonadjacent in G_1 to at least one of v_k and v_ℓ, say $v_iv_k \notin E(G_1)$. Then G_1 can be transformed into H_1 by the edge rotations $t_1 = (v_i, v_j, v_k)$ and $t_2 = (v_k, v_i, v_\ell)$ where $t_2t_1G_1 \cong H_1$. Thus $d_{er}(G_1, H_1) \leq 2$.

Assume then that each of v_i and v_j is adjacent to both v_k and v_ℓ. Then G_1 can be transformed into H_1 by the edge rotation $t'_1 = (v_k, v_i, v_\ell)$ and $t'_2 = (v_i, v_j, v_k)$, where

$$t'_2t'_1G_1 \cong H_1.$$

Therefore $d_{er}(G_1, H_1) \leq 2$.

Thus in both cases G_1 can be transformed into H_1 and $d_{er}(G_1, H_1) \leq 2$. Now H_1 and G_2 have $s + 1$ edges in common. Proceeding as above, we construct graph H_2 such that $d_{er}(H_1, H_2) \leq 2$; hence $d_{er}(G_1, H_2) \leq 4$ where H_2 and G_2 have $s+2$ edges in common. Continuing in this way, we construct a graph $H_{q-s} \cong G_2$ where $d_{er}(G_1, H_{q-s}) \leq 2(q - s)$. Hence

$$d_{er}(G_1, G_2) \leq 2(q - s), \qquad \sim (2)$$

and (1) and (2) together yield

$$q - s \leq d_{er}(G_1, G_2) \leq 2(q - s). \qquad \square$$

To show that both the upper and lower bounds presented by Theorem 2.3.20 cannot, in general, be improved, we consider two examples from [CSZ1],in which equality is attained for both the upper and lower bounds respectively.

2.3.21 Example

For $n \geq 1$ define $G_1 = K_{2n} \cup \bar{K}_{4n^2-4n}$ and $G_2 = (2n^2 - n)K_2$.
Both G_1 and G_2 have order $4n^2 - 2n$ and size $q = 2n^2 - n$. Now G_1 and G_2 have a unique greatest common subgraph namely $G = nK_n$ which has size $s = n$. Therefore

$$2(q - s) = 2[(2n^2 - n) - n] = 4n^2 - 4n.$$

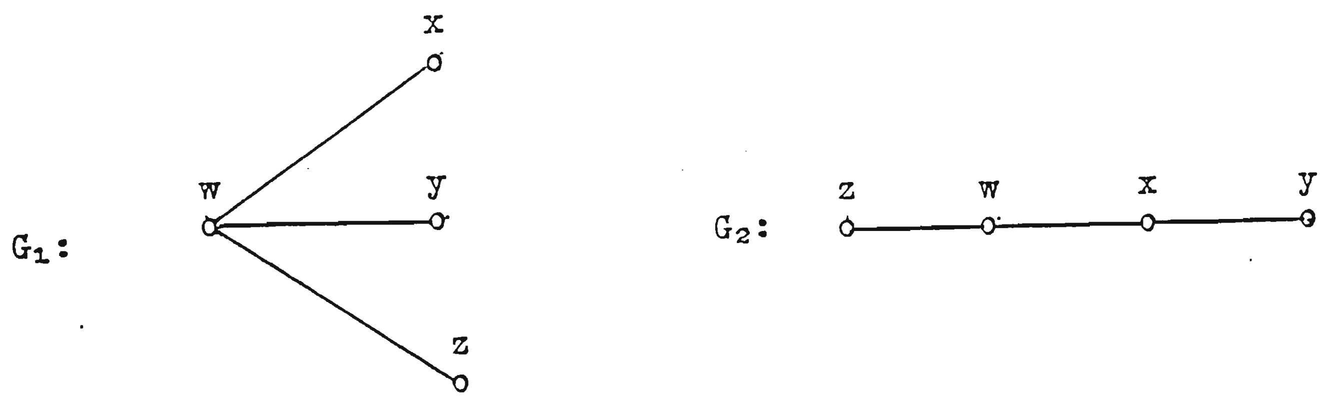
Now G_2 is 1-regular, while G_1 contains $4n^2 - 4n$ isolated vertices. Therefore $d_{er}(G_1, G_2) \geq 4n^2 - 4n$. By Theorem 2.3.20, we also have

$$d_{er}(G_1, G_2) \leq 2(q - s) = 4n^2 - 4n,$$

which implies that $d_{er}(G_1, G_2) = 4n^2 - 4n = 2(q - s)$.

2.3.22 Example

Let $G_1 \cong S_4$ and $G_2 \cong P_4$, then G_1 and G_2 have order 4 and size $q = 3$.



2.3.23 Figure

The graphs G_1 and G_2 of Example 2.3.22

Now a greatest common subgraph of G_1 and G_2 is $G \cong P_3$ which has size 2. In this case $q - s = 3 - 2 = 1$. Now the edge rotation $t = (y, w, x)$ transforms G_1 into G_2; i.e., $tG_1 \cong G_2$; hence $d_{er}(G_1 G_2) \leq 1 = q - s$.

However from Theorem 2.3.20 we have $d_{er}(G_1, G_2) \geq q - s$ and so in this case

$$d_{er}(G_1, G_2) = q - s = 1.$$

Examples 2.3.21 and 2.3.22 show that the bounds presented in Theorem 2.3.20 are sharp.

We now return to the concept of slowly changing parameters to establish the edge rotation distance between certain graphs, following [BGMW1].

2.3.24 Theorem

The edge rotation distance between the star S_n and any tree T on n vertices is equal to the difference of their maximum degrees; i.e.,

$$d_{er}(S_n, T) = \triangle(S_n) - \triangle(T).$$

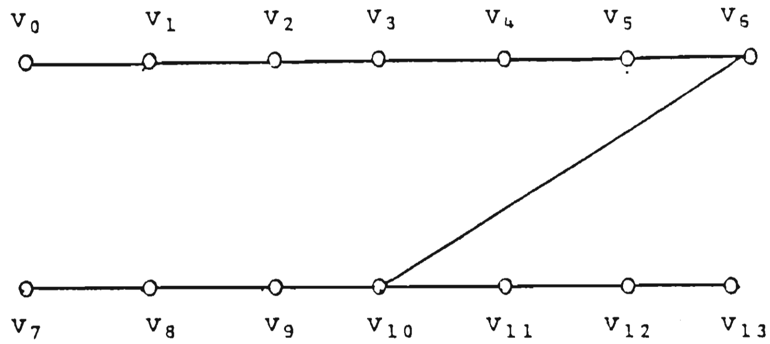
Proof

As every edge slide is an edge rotation, it follows from Theorem 2.2.14 that $d_{er}(S_n, T) \leq d_{es}(S_n, T) = \triangle(S_n) - \triangle(T)$. Since the maximum degree of a graph is a slowly changing parameter with respect to the edge rotation operation, $d_{er}(S_n) \geq \triangle(S_n) - \triangle(T)$; hence $d_{er}(S_n, T) = \triangle(S_n) - \triangle(T)$. □

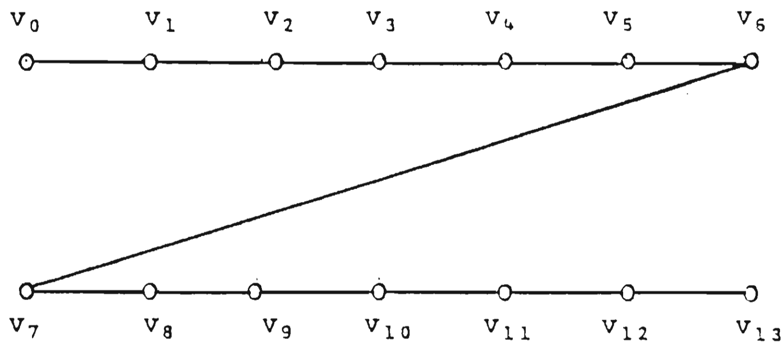
In Theorem 2.2.17 we saw that $d_{es}(P_n, T) = \mathrm{diam}P_n - \mathrm{diam}T$. However it is not true in general for a path P_n and a tree T on n vertices that $d_{er}(P_n, T) = \mathrm{diam}P_n - \mathrm{diam}T$. This is demonstrated in the following example which also shows that the diameter of a graph is not slowly changing with respect to the edge rotation operation.

2.3.25 Example

G:



P_{14} :



2.3.26 Figure

Now diam$G$ = 10 while diam P_{14} = 13, hence $|\text{diam}G - \text{diam}P_{14}| = 3$. Define the edge rotation $t = (v_6, v_{10}, v_7)$, then $tG \cong P_{14}$. That is,

$$d_{er}(G, P_{14}) = 1 \neq 3 = \text{diam}P_{14} - \text{diam}G.$$

Thus in order to find a general formula for $d_{er}(P_n, T)$ we need to introduce a parameter which is slowly changing with respect to the edge rotation operation.

2.3.27 Definition

Let G be any graph then we define $end(G)$ to be the cardinality of the set $\{v \in V(G) : \deg_G v = 1\}$ and $end'(G)$ to be the cardinality of the set $\{v \in V(G) : \deg_G v \leq 1\}$;

$$\text{i.e.,} \quad \begin{aligned} end(G) &= |\{v \in V(G) : \deg_G v = 1\}|, \qquad \text{and} \\ end'(G) &= |\{v \in V(G) : \deg_G v \leq 1\}|. \end{aligned}$$

We now show that the parameter $end'(G)$ is slowly changing with respect to edge rotations.

2.3.28 Lemma

For any graph G the parameter $end'(G)$ is slowly changing with respect to the edge rotation operation.

Proof

Consider any edge rotation $t = (u, v, w)$ on G; then $uv \in E(G)$ and $uw \in E(\bar{G})$. Thus $\deg_{tG} v = \deg_G v - 1$, $\deg_{tG} w = \deg_G w + 1$, and $\deg_{tG} z = \deg_G z$ for all $z \in V(G) - \{v, w\}$. It follows that $|end'(G) - end'(tG)| \leq 2$ with equality if and only if either $v, w \in end'(G)$ and $v, w \notin end'(tG)$, or $v, w \notin end'(G)$ and $v, w \in end'(tG)$. However if $v \in end'(G)$, then $v \in end'(tG)$ while if $w \notin end'(tG)$, then $w \notin end'(tG)$. We conclude that $|end'(G) - end'(tG)| \leq 1$. That is to say, $end'(G)$ is a slowly changing parameter with respect to edge rotations. $\square$

With the aid of Lemma 2.3.28 we prove the following result.

2.3.29 Theorem

For all trees T of order n, $d_{er}(P_n, T) = end(T) - end(P_n)$.

Proof

If $n = 1$ the result is clearly true; therefore we assume that $n \geq 2$. We note first that for nontrivial trees T the parameters $end'(T)$ and $end(T)$ coincide. Thus we refer to $end(T)$ where, in fact $end'(T)$ is the slowly changing parameter (from Lemma 2.3.28). Further, paths are the only trees with exactly two end-vertices; thus properties P1 and P2 of Lemma 2.3.14 have been verified. Property P3 of Lemma 2.3.14 will follow when we show that for any tree T with more than two end-vertices there exists a tree T' formed by a single edge rotation which has one less end-vertex than T.

Let x be an end-vertex of T, and let y be the vertex of T, of degree at least three, nearest to x. Let z be any neighbour of y not on the $x - y$ path in T. Then define the edge rotation $t = (z, y, x)$. Let $T' = tT$. Since T' is connected it is a tree, and as x is no longer an end-vertex, $end(T') = end(T) - 1$. Thus property P3 of Lemma 2.3.14 is verified and hence

$$d_{er}(P_n, T) = end(T) - end(P_n). \qquad \square$$

2.3.30 Definition

Let S be a set of (nonisomorphic) graphs of the same order and size. Then following [CGHLSW1], the *edge rotation distance graph* $D_{er}(S)$ of S is defined to be that graph with vertex set S such that two vertices G and H of $D_{er}(S)$ are adjacent if and only if $d_{er}(G, H) = 1$.

The question of which graphs are edge rotation distance graphs arises naturally from this definition. This question was discussed in [CGHLSW1] where it was conjectured that all graphs are edge rotation distance graphs; however this problem remains unsolved. Partial results were however obtained and we now discuss these. Apart from Example 2.3.37, all the results in the remainder of this section first appeared in [CGHLSW1].

2.3.31 Lemma

K_p is an edge rotation distance graph.

Proof

Let G be any graph of order p with vertex set $V(G) = \{v_1, v_2, ..., v_p\}$ and let $\{H_1, H_2, ..., H_p\}$ be the set of graphs described in the proof of Theorem 2.2.23. Then for $1 \leq i \neq j \leq p$ the edge rotation $t = (u_i, v_i, v_j)$ transforms the graph H_i into H_j; i.e.,

$$tH_i = H_j.$$

Therefore every pair of vertices in $D_{er}(\{H_1, H_2, ..., H_p\})$ are adjacent; i.e., $D_{er}(\{H_1 H_2 ... H_p\}) \cong K_p$. $\square$

2.3.32 Lemma

For $n \geq 3$, C_n is an edge rotation distance graph.

Proof

Let $C = x_1x_2...x_{2n+2}x$ be a $(2n+2)$-cycle and for $i = 1, 2, ..., n$ let $F_i = C + x_1x_{i+2}$. For $i = 1, 2, ..., n-1$, define $H_i = F_i \cup F_{i+1}$ and define $H_n = F_n \cup F_1$. Then for $i = 1, 2, ..., n-1$ the edge rotation $t = (x_1, x_{i+2}, x_{i+4})$ transforms the graph H_i into H_{i+1}; i.e.,

$$tH_i \cong H_{i+1}.$$

Therefore we have that

$$d_{er}(H_i, H_{i+1}) = 1 \text{ for all } i = 1, 2, ..., n-1. \quad \sim (1)$$

Now the edge rotation $t' = (x_1, x_3, x_{n+2})$ transforms H_1 into H_n; i.e., $t'_1H_1 \cong H_n$. Therefore

$$d_{er}(H_1, H_n) = 1. \quad \sim (2)$$

Now for any other pair H_j, H_k where $1 \leq j < k \leq n$, $(j,k) \neq (1,n)$ and $k - j \neq 1$ we have that

$$d_{er}(H_j, H_k) \geq 2. \quad \sim (3)$$

Therefore (1), (2) and (3) together imply that

$$C_n \cong D_{er}(\{H_1, H_2, ..., H_n\}). \quad \square$$

2.3.33 Lemma

For $n \geq 1$, P_n is an edge rotation distance graph.

Proof

Let $C = x_1x_2...x_{2n+6}x_1$ be a $(2n+4)$-cycle and for $i = 1, 2, ..., n+1$ define $F_i = C + x_1x_{i+2}$. For $i = 1, 2, ..., n$, define $H_i = F_i \cup F_{i+1}$. Then, as in Lemma 2.3.32, it is easily shown that

$$P_n \cong D_{er}(\{H_1, H_2, ..., H_n\}). \qquad \square$$

The following two lemmas help to establish that a number of large classes of graphs are edge rotation distance graphs.

2.3.34 Lemma

Let $G, H \in \Gamma(p,q)$. Then $d_{er}(G,H) = 1$ if and only if $d_{er}(G+K_1, H+K_1) = 1$.

Proof

If $d_{er}(G,H) = 1$ then there exists an edge rotation $t = (u,v,w)$ such that $tG \cong H$ (where $u,v,w \in V(G), uv \in E(G)$ and $uw \in E(\bar{G})$); then $G + K_1 \not\cong H + K_1$ and $t(G+K_1) \cong H + K_1$; hence $d_{er}(G,H) = 1$ implies that $d_{er}(G+K_1, H+K_1) = 1$.

Now suppose that $d_{er}(G+K_1, H+K_1) = 1$.

Case 1) Assume that there exist vertices $u, v, w, x \in V(G+K_1)$ such that

i) $\deg_{G+K_1} u = p$;

ii) $v, w, x \in V(G+K_1) - \{u\}$, $vw \in E(G)$ and $vx \notin E(G)$; and

iii) $G + K_1 + vx - vw \cong H + K_1$.

Then necessarily, $G + vx - vw \cong H$ and $d_{er}(G, H) = 1$.

Case 2) Assume Case 1 does not occur. Now we know that there exists a vertex $u \in V(G + K_1)$ such that $\deg_{G+K_1} u = p$ and $u \notin V(G)$. Then by assumption there exist nonadjacent vertices v and w of G such that $G + K_1 - uv + vw \cong H + K_1$. Now assume there exists another vertex z say such that $\deg_{G+K_1} z = p$. Then $z \in V(G)$ and $G + K_1 - zv + vw \cong H + K_1$, as in Case 1, a contradiction. Therefore $G + K_1$ has only one vertex, namely u, of degree p.

Since u is the only vertex of degree p in $G + K_1$ it follows that w is the only vertex of degree p in $G + K_1 - uv + vw \cong H + K_1$. This implies that $(G + K_1 - uv + vw) - w \cong H$. Now in $G + K_1$, $\deg_{G+K_1} u = p$ and $\deg_{G+K_1} w = p - 1$, where $vw \in E(\bar{G})$. On the other hand in $G + K_1 - uv + vw$ we have $\deg_{G+K_1-uv+vw} w = p$ and $\deg_{G+K_1-uv+vw} u = p - 1$ where $uv \notin E(G + K_1 - uv + vw)$. It follows that

$$(G + K_1 - uv + vw) - w \cong G.$$

This however implies that $G \cong H$ which contradicts the fact that $d_{er}(G + K_1, H + K_1) = 1$. Therefore Case 2 cannot occur which completes the proof. □

2.3.35 Lemma

For any edge rotation distance graph G and any positive integer n there exists a set S_n of n-connected graphs such that $G \cong D_{er}(S_n)$.

Proof

Let τ be a set of graphs such that $G \cong D_{er}(\tau)$. Now let $S_1 = \{H + K_1 : H \in \tau\}$. Since for $A, B \in \tau$, $d_{er}(A + K_1, B + K_1) = 1$ if and only if $d_{er}(A, B) = 1$ by Lemma 2.3.34, we have that

$$G \cong D_{er}(S_1).$$

Similarly, letting $S_i = \{H + K_i : H \in \tau\}$ and noting that $H + K_i = (H + K_{i-1}) + K_1$, we find by repeated application of Lemma 2.3.34 that $d_{er}(A + K_i, B + K_i) = 1$ if and only if $d_{er}(A + K_{i-1}, B + K_{i-1}) = 1$ ($i = 2, 3, ..., n$) for $A, B \in \tau$. Therefore $G \cong D_{er}(S_i)$ for $1 \leq i \leq n$.

Taking $i = n$ we now have that $G \cong D_{er}(S_n)$, where $S_n = \{H + K_n : H \in \tau\}$. Since for any graph G, $G + K_n$ is n-connected, it follows that S_n is a set of n-connected graphs. $\square$

2.3.36 Lemma

Let G_0, G_1, H_0, H_1 be 2-connected graphs of the same order and the same size. Then $d_{er}(G_0 \cup G_1, H_0 \cup H_1) = 1$ if and only if $G_i \cong H_j$ for some $i, j \in \{0, 1\}$ and $d_{er}(G_{1-i}, H_{1-j}) = 1$.

Proof

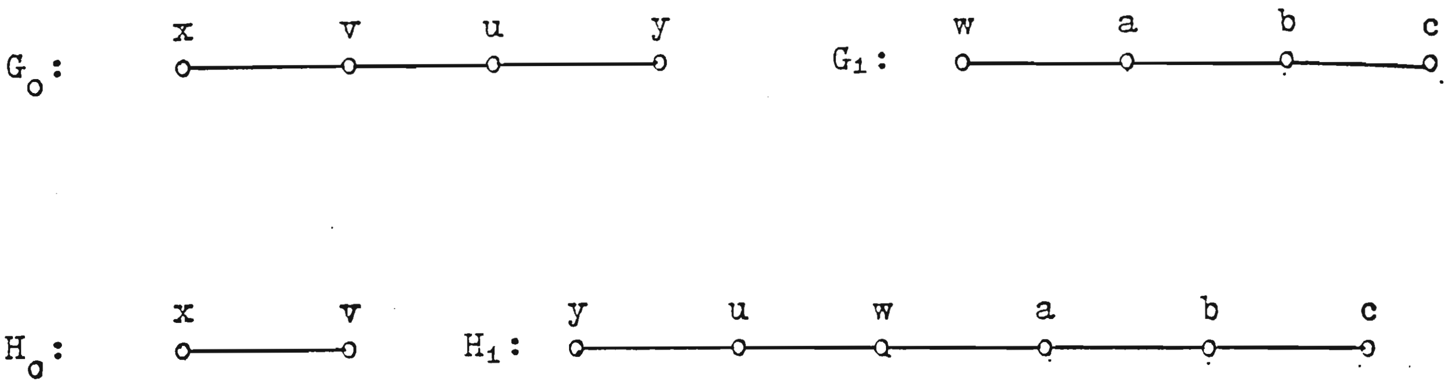
Suppose that, for some $i, j \in \{0, 1\}$, $G_i \cong H_j$ and that $d_{er}(G_{1-i}, H_{1-j}) = 1$, where $t = (u, v, w)$ is an edge rotation such that $tG_{1-i} \cong H_{1-j}$; then $t(G_0 \cup G_1) \cong H_0 \cup H_1$.

Conversely, suppose that $d_{er}(G_0 \cup G_1, H_0 \cup H_1) = 1$ and let $t = (u, v, w)$ be an edge rotation such that $t(G_0 \cup G_1) \cong H_0 \cup H_1$, where without loss of generality, we assume $uv \in E(G_0)$ and $uw \notin E(G_0 \cup G_1)$. Since G_0 is 2-connected, $G_0 - uv$ is connected; so, if $w \in V(G_1)$ then $t(G_0 \cup G_1)$ is connected. However, $t(G_0 \cup G_1)$ is isomorphic to $H_0 \cup H_1$, which is disconnected. So $w \in V(G_0)$ and $t(G_0 \cup G_1) = tG_0 \cup G_1 \cong H_0 \cup H_1$, from which we obtain $tG_0 \cong H_j$ for some $j \in \{0, 1\}$ and $G_1 \cong H_{1-j}$. $\square$

The following example shows that in the above lemma, we may not dispense with the condition that G_0, G_1, H_0, H_1 be 2-connected.

2.3.37 Example

Let $G_0 \cong G_1 \cong P_4$, $H_0 \cong P_2$ and $H_1 \cong P_6$, as in Figure 2.3.38, then $d_{er}(G_0 \cup G_1, H_0 \cup H_1) = 1$ where $t(G_0 \cup G_1) \cong H_0 \cup H_1$ and $t = (u, v, w)$. Now $G_i \ncong H_j$ for any $i, j \in \{0, 1\}$ and hence the condition that G_0, G_1, H_0, H_1 be 2-connected, is necessary for Lemma 2.3.36 to hold.



2.3.38 Figure

The graphs G_0, G_1, H_0 and H_1 of Example 2.3.37.

2.3.39 Theorem

Every induced subgraph of an edge rotation distance graph is an edge rotation distance graph.

Proof

Let G be an edge rotation distance graph of order n say, with vertex set $V(G) = \{v_1, v_2, ..., v_n\}$. Then by definition there exists a set $S = \{S_1, S_2, ..., S_n\}$ of graphs with the same order and size such that

$$G \cong D_{er}(\{S_1, S_2, ..., S_n\}).$$

We may assume without loss of generality that the graphs S_i, for $i = 1, 2, ..., n$, are labelled in such a way that $v_j v_k \in E(G)$ if and only if $d_{er}(S_j, S_k) = 1$; i.e., the vertex S_i in $D_{er}(\{S_1, S_2, ..., S_n\})$ corresponds to the vertex v_i in G.

Let $H = (v_{i_1}, v_{i_2}, ..., v_{i_\ell})$, where $1 \leq \ell < n$, be a proper subset of $V(G)$. Then for $v_{i_j}, v_{i_k} \in V(\langle H \rangle), v_{i_j} v_{j_k} \in E(\langle H \rangle)$ if and only if $v_{i_j} v_{j_k} \in E(G)$ if and only if $d_{er}(S_{i_j}, S_{i_k}) = 1$.

Therefore we have that $\langle H \rangle \cong D_{er}(S_{i_1}, S_{i_2}, ..., S_{i_\ell})$ where $1 \leq \ell < n$. $\square$

Using Lemmas 2.3.35 and 2.3.36 we are able to prove the next result which concerns the union and cartesian product of two edge rotation distance graphs.

2.3.40 Theorem

Let G and H be edge rotation distance graphs. Then

a) $G \cup H$ is an edge rotation distance graph;

b) $G \times H$ is an edge rotation distance graph; and

c) for every pair $\{v, u\}$ where $v \in V(G)$ and $w \in V(H)$, the graph obtained from G and H by identifying v and w is an edge rotation distance graph.

Proof

a) By Lemma 2.3.35 there exist sets $\mathcal{S}$ and $\mathcal{T}$ of 2-connected graphs such that $D_{er}(\mathcal{S}) \cong G$ and $D_{er}(\mathcal{T}) \cong H$. We ensure that the order of the graphs in $\mathcal{S}$ is different from the order of the graphs in $\mathcal{T}$ to ensure that if A and B are graphs in $\mathcal{S} \cup \mathcal{T}$ with $d_{er}(A, B) = 1$, then A and B are either both in $\mathcal{S}$ or both in $\mathcal{T}$. This is done by choosing a suitable n when applying Lemma 2.3.35 to establish the 2-connected sets of graphs $\mathcal{S}$ and $\mathcal{T}$. It then follows that

$$G \cup H \cong D_{er}(\mathcal{S} \cup \mathcal{T})$$

b) The graph described in c) in the statement of the theorem is an induced subgraph of $G \times H$, thus by Theorem 2.3.39, to complete the proof it is sufficient to establish b). By Lemma 2.3.35, we may assume as in c), that there exist disjoint sets $\mathcal{S}$ and $\mathcal{T}$ of 2-connected graphs for which $D_{er}(\mathcal{S}) \cong G$ and $D_{er}(\mathcal{T}) \cong H$. Let $\mathcal{S} = \{G_u : u \in V(G)\}$ with $d_{er}(G_u, G_w) = 1$ if and only if $uw \in E(G)$. Similarly let $\mathcal{T} = \{H_v : v \in V(H)\}$ with $d_{er}(H_v, H_w) = 1$ if and only if $vw \in E(H)$.

By Lemma 2.3.36 for $u, u' \in V(G)$ and $v, v' \in V(H)$, we have that $d_{er}(G_u \cup H_v, G_{u'} \cup H_{v'}) = 1$ if and only if either $G_u \cong G_{u'}$ and $d_{er}(H_v, H_{v'}) = 1$ or $H_v \cong H_{v'}$ and $d_{er}(G_u, G_{u'}) = 1$. Therefore $d_{er}(G_u \cup H_v, G_{u'} \cup H_{v'}) = 1$ if and only if either $u = u'$ and $vv' \in E(H)$ or $v = v'$ and $uu' \in E(G)$. Now $(u, v)(u', v') \in E(G \times H)$ if and only if either $u = u'$ and $vv' \in E(H)$ or $v = v'$ and $uu' \in E(G)$. Therefore it follows that $d_{er}(G_u \cup H_v, G_{u'} \cup H_{v'}) = 1$ if and only if $\{(u, v), (u', v')\} \in E(G \times H)$. Hence $G \times H \cong D_{er}(\{G_u \cup H_v : u \in V(G), v \in V(H)\})$.$\square$

There are two immediate consequences of Theorem 2.3.40.

2.3.41 Corollary

If the blocks of a connected graph G are all edge rotation distance graphs, then G is an edge rotation distance graph.

Proof

We shall proceed by induction on the number of blocks b, of G. If $b = 1$ the statement is obviously valid. Suppose that the statement is true for all graphs with fewer than b blocks and let G be a connected graph with b blocks, all of which are edge rotation distance graphs. Let B be an end block of G containing the cut vertex v of G and let $H = G - (V(B) - \{v\})$. Then H has $b - 1$ blocks, all of which are edge rotation distance graphs. Now since B is an edge rotation graphs, the fact that G is an edge rotation distance graph now follows from part c) of Theorem 2.3.40. □

2.3.42 Corollary

Every tree is an edge rotation distance graph.

Proof

Every block of a tree T is isomorphic to K_2 which by Lemma 2.3.31 is an edge rotation distance graph. That T is an edge rotation distance graph now follows directly from Corollary 2.3.41. □

2.3.43 Theorem

Every line graph is an edge rotation distance graph.

Proof

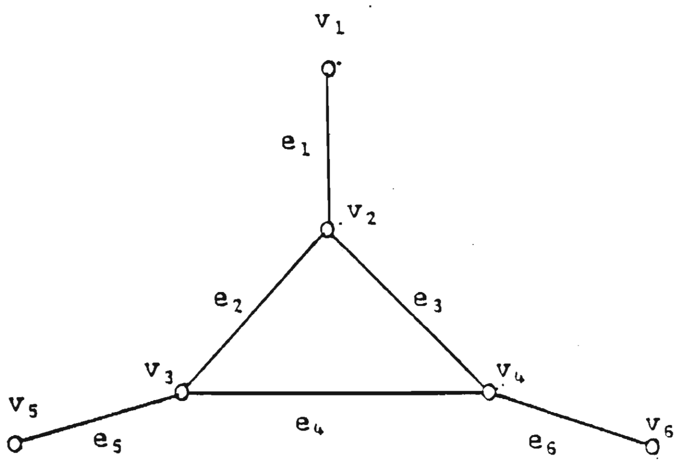
Let G be any given line graph. Therefore by definition there exists a graph H with $V(H) = \{v_1, v_2, ..., v_p\}$ and $V(G) = E(H) = \{e_1, e_2, ..., e_q\}$ where $G \cong L(H)$. Let $\{F_1, F_2, ..., F_p\}$ be a set of 2-connected graphs with the property that $d_{er}(F_i, F_j) = 1$ for $1 \leq i < j \leq p$. That such a set exists, follows from Lemmas 2.3.31 and 2.3.35.

For $k = 1, 2, ..., q$ define the graph G_k to be $F_i \cup F_j$, where $e_k = v_i v_j$. Let $S = \{G_1, G_2, ..., G_q\}$. Observe that by Lemma 2.3.36, $d_{er}(G_i, G_j) = 1$ if and only if G_i and G_j have exactly one common component. Thus, $d_{er}(G_i, G_j) = 1$ if and only if there exist three distinct integers s, t and u such that $G_i = F_s \cup F_t$ and $G_j = F_t \cup F_u$ say; i.e., $e_i = v_s v_t$ and $e_j = v_t v_u$. It follows that $d_{er}(G_i, G_j)$=1 if and only if $e_i e_j \in E(G)$ and therefore

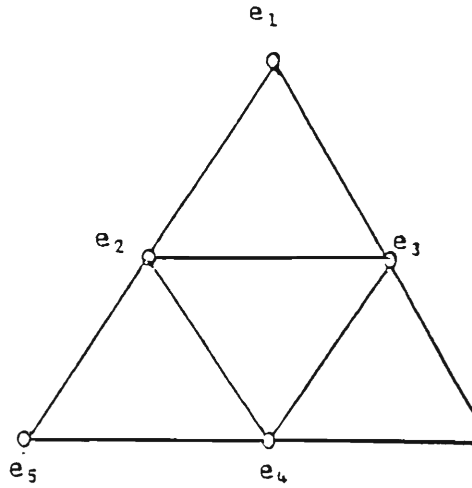
$$G \cong D_{er}(S). \qquad \square$$

2.3.44 Example

H:



G=L(H):



2.3.45 Figure

Let G and H be as in Figure 2.3.45. Let G' be any graph of order 6 with $V(G') = \{u_1, u_2, ..., u_6\}$. Now add two new vertices adjacent to u_1, four new vertices adjacent to u_2 and, in general, $2i$ new vertices adjacent to v_i, for $i = 1, 2, ..., 6$; call this graph F. Now for $i = 1, 2, ..., 6$ let F_i denote the graph obtained from F by adding another new vertex adjacent only to u_i. Then, as this is a special case of the construction in the proofs of Theorem 2.3.23 and Lemma 2.3.31, we have that

$$d_{er}(F_i, F_j) = 1 \text{ for } 1 \leq i < j \leq 6.$$

For $k = 1, 2, ..., 6$ define $G_k = F_i \cup F_j$ if $e_k = v_i v_j$. Therefore we have $G_1 = F_1 \cup F_2$, $G_2 = F_2 \cup F_3$, $G_3 = F_2 \cup F_4$, $G_4 = F_3 \cup F_4$, $G_5 = F_3 \cup F_5$

and $G_6 = F_4 \cup F_6$. Let $\mathcal{S} = \{G_1, G_2, ..., G_6\}$. Then

$$G \cong D_{er}(\mathcal{S}).$$

We note that the construction in the proof of Lemma 2.3.32 is a special case of the construction of $\mathcal{S}$ in Theorem 2.3.43.

In the proofs of Lemmas 2.3.32, 2.3.33 and Theorems 2.3.40 and 2.3.43, we used the fact that if G_0, G_1, H_0 and H_1 are 2-connected graphs of the same order and the same size, then $d_{er}(G_0 \cup G_1, H_0 \cup H_1) = 1$ if and only if for some $i, j \in \{0, 1\}$, $G_{1-i} \cong H_{1-j}$ and $d_{er}(G_i, H_j) = 1$ as stated in Lemma 2.3.36. We now extend this concept to include n components.

2.3.46 Remark

Let $G_1, G_2, ..., G_n$, $H_1, H_2, ..., H_n$ be 2-connected graph of the same order and the same size, then $d_{er}(G_1 \cup G_2 \cup ... \cup G_n,\ H_1 \cup H_2 \cup ... \cup H_n) = 1$ if and only if for some $i, j \in \{1, 2, ..., n\}$ we have

$$G_1 \cup G_2 \cup ...G_{i-1} \cup G_{i+1} \cup ... \cup G_n \cong H_1 \cup H_2 \cup ... \cup G_{j-1} \cup H_{j+1} \cup ... \cup H_n.,$$

and $d_{er}(G_i, H_j) = 1$.

2.3.47 Definition

For k and m fixed positive integers, let $\mathcal{F}$ be a set of m 2-connected graphs which are pairwise at an edge rotation distance of one from each other. Define $\mathcal{G}_m^k(\mathcal{F})$ to be the set of all graphs with k components, each of which is a (not necessarily distinct) element of $\mathcal{F}$. It is obvious that if $\mathcal{F}'$ is another distinct set of m 2-connected graphs, pairwise at an edge rotation distance of one, then

$$|\mathcal{G}_m^k(\mathcal{F})| = |\mathcal{G}_m^k(\mathcal{F}')|, \qquad \text{and}$$
$$D_{er}(\mathcal{G}_m^k(\mathcal{F})) \cong D_{er}(\mathcal{G}_m^k(\mathcal{F}')).$$

Thus for convenience we shall write $\mathcal{G}_m^k$ and $D_{er}(\mathcal{G}_m^k)$ without reference to the set $\mathcal{F}$.

2.3.48 Note

i) From Lemma 2.3.31, we have that $K_n \cong D_{er}(\mathcal{G}_n^1)$.

ii) From Lemmas 2.3.32 and 2.3.33, $P_{n-1} < C_n < D_{er}(\mathcal{G}_n^2)$.

iii) From the proof of Theorem 2.3.43, we have that if H is a graph of order p, then $L(H) < D_{er}(\mathcal{G}_p^2)$.

Note 2.3.48 suggests that by considering $D_{er}(\mathcal{G}_m^k)$ for $k > 2$ it may be possible to establish many more graphs as edge rotation distance graphs, that is, show that $G < D_{er}(\mathcal{G}_m^k)$ for some k and m for many graphs G.

This approach, however, was found by the authors of [CGHLSW1] to be limited, as the following result shows.

2.3.49 Theorem

Let $G \cong D_{er}(\mathcal{G}_m^k)$ and let $x, y \in V(G)$. If $d_G(x, y) = 2$, then $\langle N_G(x) \cap N_G(y)\rangle$ is isomorphic to one of K_1, K_2 or C_4.

Proof

Let $\mathcal{F}$ be a set of m 2-connected graphs pairwise at a distance 1; while $\mathcal{F} = \{F_1, F_2, ..., F_m\}$ say. Then any graph in $\mathcal{G}_m^k(\mathcal{F})$ has the form $n_1F_1 \cup n_2F_2 \cup ... \cup n_mF_m$, where $n_1, n_2, ..., n_m$ are nonnegative integers, and $\sum_{i=1}^{m} n_i = k$. Thus there is a one-to-one correspondence between the vertices of G and the set of m-tuples $(n_1, n_2, ..., n_m)$ of nonnegative integers with $\sum_{i=1}^{m} n_i = k$.

Suppose that $x, y \in V(G)$ where x corresponds to $(n_1, n_2, ..., n_m)$, y corresponds to $(t_1, t_2, ..., t_m)$ and $d_G(x, y) = 2$; then these two m-tuples differ in either two, three or four entries. We may assume, without loss of generality, that one of the following situations occurs:

Case 1) $x : (n_1, n_2, n_3, n_4, ..., n_m)$ and $y : (n_1 - 2, n_2 + 2, n_3, n_4, ..., n_m)$,

Case 2) $x : (n_1, n_2, n_3, n_4, ..., n_m)$ and $y : (n_1 - 2, n_2 + 1, n_3 + 1, n_4, ..., n_m)$

Case 3) $x : (n_1, n_2, n_3, n_4, ..., n_m)$ and $y : (n_1 - 1, n_2 - 1, n_3 + 1, n_4 + 1, n_5, ..., n_m)$

Case 4) $x : (n_1, n_2, n_3, n_4, ..., n_m)$ and $y : (n_1 - 1, n_2 - 1, n_3 + 2, n_4, ..., n_m)$.

In all cases x and y are nonadjacent. In Case 1), x and y would both be adjacent in G to the unique vertex z corresponding to the m-tuple $(n_1 - 1, n_2 + 1, n_3, n_4, ..., n_m)$. In Case 2), x and y are adjacent in G to exactly the two vertices corresponding to $(n_1 - 1, n_2 + 1, n_3, n_4, ..., n_m)$ and $(n_1 - 1, n_2, n_3 + 1, n_4, ..., n_m)$. In Case 3) x and y are adjacent in G to the vertices corresponding to $(n_1 - 1, n_2, n_3, n_4 + 1, n_5, ..., n_m)$, $(n_1, n_2 - 1, n_3 + 1, n_4, ..., n_m)$, $(n_1 - 1, n_2 - 1, n_3, n_4, ..., n_m)$ and $(n_1, n_2, n_3 + 1, n_4 + 1, n_5, ..., n_m)$. In Case 4) x and y are adjacent to the vertices in G corresponding to $(n_1, n_2 - 1, n_3 + 1, n_4, ..., n_m)$ and $(n_1 - 1, n_2, n_3 + 1, n_4, ..., n_m)$.

Therefore in Case 1), $\langle N_G(x) \cap N_G(y) \rangle \cong K_1$. In Cases 2) and 4), $\langle N_G(x) \cap N_G(y) \rangle \cong K_2$ and in Case 3), $\langle N_G(x) \cap N_G(y) \rangle \cong C_4$. $\square$

2.3.50 Corollary

If $G < D_{er}(\mathcal{G}_m^k)$ and $x, y \in V(G)$ where $d(x, y) = 2$, then

$$\langle N_G(x) \cap N_G(y) \rangle < C_4.$$

We close this chapter with the following:

2.3.51 Conjecture

All graphs are edge rotation distance graphs.

Chapter 3

Metrics Involving Greatest Common Subgraphs

3.1 Introduction

In this chapter we shall deal with distances between graphs, where the measure of distance, in each case, may be determined by a method involving a greatest common subgraph of some type.

3.2 The Edge Move Distance d_{em}

3.2.1 Definitions

Let G and H be graphs of the same order and size. Then following [BGMW1] we say that G can be transformed into H by an *edge move* if G contains four vertices u, v, w and x, at least three of which are distinct, such that $uv \in E(G)$, $wx \in E(\bar{G})$ and $H \cong G - uv + wx$.

For graphs G and H of the same order and size, the *edge move distance* $d_{em}(G, H)$ between G and H is the smallest nonnegative integer n for which there exists a sequence

$$G \cong H_0, H_1, H_2, ..., H_n \cong H$$

such that H_i can be transformed into $H_{i+1}(i = 0, 1, ..., n-1)$ by an edge move.

It is immediately clear that the edge slide and edge rotation operations are merely special cases of the edge move operation, that is, every edge slide and every edge rotation is also an edge move. It also follows immediately that for graphs G and H of the same order and size

$$d_{em}(G, H) \leq d_{er}(G, H) \leq d_{es}(G, H).$$

We denote an edge move on a graph G by $t = (u, v, w, x)$ where $tG = G - uv + wx$. We denote the inverse of the edge move t by $t^{-1} = (w, x, u, v)$. That t^{-1} is also an edge move is obvious.

The edge move distance imposes a metric on the set $S_{p,q}$ of all isomorphism classes of graphs which have p vertices and q edges, as follows: If $\sigma_1, \sigma_2 \in S_{p,q}$, then obviously the distance $d_{em}(G_1, G_2)$ is fixed for all $G_1 \in \sigma_1$ and all $G_2 \in \sigma_2$ and is also denoted by $d_{em}(\sigma_1, \sigma_2)$.

3.2.2 Theorem

For any integers $p \geq 1, q \geq 0$, the edge move distance is a metric on $S_{p,q}$.

Proof

Let $\sigma_i \in S_{p,q}$ and let $G_i \in \sigma_i$ for $i = 1, 2, 3$.

i) By definition, $d_{em}(\sigma_1, \sigma_2) \geq 0$ and $d_{em}(\sigma_1, \sigma_2) = 0$ if and only if $d_{em}(G_1, G_2) = 0$, hence if and only if $G_1 \cong G_2$ and $\sigma_1 = \sigma_2$.

ii) If $d_{em}(\sigma_1, \sigma_2) = n$, then, $d_{em}(G_1, G_2) = n$ and by definition there exists a sequence $t_1, t_2, ..., t_n$ of edge moves such that $t_n...t_1G_1 \cong G_2$. Consequently $t_1^{-1}...t_n^{-1}G_2 \cong G_1$ and so $d_{em}(G_2, G_1) \leq n$; i.e., $d_{em}(\sigma_2, \sigma_1) \leq d_{em}(\sigma_1, \sigma_2)$. A similar argument shows that $d_{em}(\sigma_1, \sigma_2) \leq d_{em}(\sigma_2, \sigma_1)$; hence $d_{em}(\sigma_1, \sigma_2) = d_{em}(\sigma_2, \sigma_1)$.

iii) Let $d_{em}(\sigma_1, \sigma_2) = n$ and $d_{em}(\sigma_2, \sigma_3) = m$; then $d_{em}(G_1, G_2) = n$, $d_{em}(G_2, G_3) = m$ and by definition there exist edge move sequences $t_1, t_2, ..., t_n$ and $s_1, s_2, ..., s_m$ such that $t_n...t_1G_1 \cong G_2$ and $s_m...s_1G_2 \cong G_3$; hence $s_m...s_1t_n...t_1G_1 \cong G_3$. Therefore $d_{em}(G_1, G_3) \leq n + m$ and so $d_{em}(\sigma_1, \sigma_3) \leq d_{em}(\sigma_1, \sigma_2) + d_{em}(\sigma_2, \sigma_3)$. □

Since the edge slide and edge rotation operations are special cases of the edge move operation the following result needs no further proof (see Theorem 2.3.8).

3.2.3 Theorem

Let $G_1, G_2 \in \Gamma(p, q)$, then there exists an edge move sequence $t_1, t_2, ..., t_n$ such that $t_n...t_2t_1G_1 \cong G_2$.

We now show that the edge move distance is preserved by complementation.

3.2.4 Theorem

For graphs $G_1, G_2 \in \Gamma(p, q)$, $d_{em}(G_1, G_2) = d_{em}(\bar{G}_1, \bar{G}_2)$.

Proof

If $d_{em}(G_1, G_2) = 0$ then $G_1 \cong G_2$ and hence $\bar{G}_1 \cong \bar{G}_2$ which implies that $d_{em}(\bar{G}_1, \bar{G}_2) = 0$, which satisfies the statement of the theorem. Assume then that $d_{em}(G_1, G_2) = n \geq 1$. By definition there exists a sequence of graphs

$$G_1 \cong H_0, H_1, ..., H_n \cong G_2$$

where H_i can be transformed into H_{i+1} by an edge move for $i = 0, 1, ..., n-1$. Let $H_{i+1} = H_i - u_i v_i + w_i x_i$. Then note that $\bar{H}_{i+1} = \bar{H}_i - w_i x_i + u_i v_i$; i.e., $\bar{H}_i$ can be transformed into $\bar{H}_{i+1}$ by an edge move. Thus the sequence of graphs

$$\bar{G}_1 \cong \bar{H}_0, \bar{H}_1, ..., \bar{H}_n \cong \bar{G}_2$$

has the property that $d_{em}(\bar{H}_i, \bar{H}_{i+1}) = 1$ for $i = 0, 1, ..., n-1$. This implies that

$$d_{em}(\bar{G}_1, \bar{G}_2) \leq d_{em}(G_1, G_2) = n. \qquad \sim (1)$$

However, by (1)

$$d_{em}(\bar{\bar{G}}_1, \bar{\bar{G}}_2) \leq d_{em}(\bar{G}_1, \bar{G}_2)$$

which implies that

$$d_{em}(G_1, G_2) \leq d_{em}(\bar{G}_1, \bar{G}_2). \qquad \sim (2)$$

Together (1) and (2) imply $d_{em}(G_1, G_2) = d_{em}(\bar{G}_1, \bar{G}_2)$. □

Once again we consider slowly changing parameters; this time with respect to edge moves (see Definition 2.3.13).

Since a single edge move can only increase or decrease the degree of a vertex at most by one, the maximum degree $\triangle(G)$ of a graph G is a slowly changing parameter with respect to edge moves. Therefore the proof of the following theorem may be obtained immediately from that of Theorem 2.3.16.

3.2.5 Theorem

For any nonnegative integer n there exist graphs $G_1, G_2 \in \Gamma(p,q)$ such that $d_{em}(G_1, G_2) = n$.

The following theorem appears in [BGMW1].

3.2.6 Theorem

The edge move distance between the star S_n and a tree T on n vertices is $\triangle(S_n) - \triangle(T)$.

Proof

Since the maximum degree of a graph is a slowly changing parameter with respect to the edge move operation

$$d_{em}(S_n, T) \geq \triangle(S_n) - \triangle(T). \qquad \sim (1)$$

As every edge rotation is an edge move, it follows from Theorem 2.3.24 that

$$d_{em}(S_n, T) \leq d_{er}(S_n, T) = \triangle(S_n) - \triangle(T). \qquad \sim (2)$$

Therefore (1) and (2) together yield $d_{em}(S_n, T) = \triangle(S_n) - \triangle(T)$. $\square$

Using the concept of the greatest common subgraph of two given graphs (see Definition 2.3.18), we obtain an equivalent formulation of the edge move distance between two graphs.

3.2.7 Theorem

Let $G_1, G_2 \in \Gamma(p, q)$. Let G be a greatest common subgraph of G_1 and G_2 of size $q(G) = s$. Then $d_{em}(G_1, G_2) = q - s$.

Proof

If $s = q$ then $G_1 \cong G_2$ and $d_{em}(G_1, G_2) = 0$.

Assume then that $1 \leq s < q$. Let the vertices of G_1 and G_2 be labelled $v_1, v_2, ..., v_p$ so that the vertices of the subgraphs of G_1 and G_2 isomorphic to G are identically labelled. Since $q > s$ we have that G_1 has $q - s$ edges $v_i v_j$ not contained in G_2 and G_2 has $q - s$ edges $v_k v_\ell$ not contained in G_1. Therefore

$$d_{em}(G_1, G_2) \geq q - s. \qquad \sim (1)$$

Let $u, v, w, x \in V(G_1) = V(G_2)$ such that $uv \in E(G_1)$, $uv \in E(\bar{G}_2)$, $wx \in E(\bar{G}_1)$ and $wx \in E(G_2)$. The edge move $t_1 = (u, v, w, x)$ on G_1 results in $t_1 G_1$ and G_2 having a greatest common subgraph with $s + 1$ edges.

Repeating this process for each of the $q - s$ pairs of edges $v_i v_j, v_k v_\ell$ where $v_i v_j \in E(G_1), v_i v_j \in E(\bar{G}_2), v_k v_\ell \in E(\bar{G}_1)$ and $v_k v_\ell \in E(G_2)$, gives a sequence $t_1, t_2, ..., t_{q-s}$ of edge moves such that $t_{q-s} ... t_2 t_1 G_1 \cong G_2$; i.e.,

$$d_{em}(G_1, G_2) \leq q - s. \qquad \sim (2)$$

Together (1) and (2) imply that $d_{em}(G_1, G_2) = q - s$. $\square$

Theorem 3.2.6 suggests an alternative method of finding the edge move distance between two graphs $G_1, G_2 \in \Gamma(p, q)$. This method involves finding a greatest common subgraph G of G_1 and G_2 and determining its size, s say. Once we have done this we perform the subtraction $q - s$ to obtain $d_{em}(G_1, G_2)$. The difficulty of this method lies in finding a greatest common subgraph of two graphs. Unfortunately no efficient algorithm exists which does this. This is seen as follows : suppose an efficient algorithm for finding a greatest common subgraph of two graphs did exist. Then to see whether a graph G is Hamiltonian, just use this algorithm to see if C_p is a greatest common subgraph of G and $H \cong C_p$. This would solve the travelling salesman problem efficiently and we know that to be NP-complete.

3.2.8 Definition

The simplest metric possible when considering distances between graphs was defined by Johnson in [J1]. The *discrete metric* $d_d : \Gamma \times \Gamma \to \{0, 1\}$ is defined by $d_d(G, H) = 0$ if $G \cong H$ and $d_d(G, H) = 1$ otherwise. The metric serves merely to distinguish between isomorphic and nonisomorphic graphs. This metric is not very interesting and perhaps deserves the title of The Trivial Metric.

3.2.9 Definition

Define the cardinality $|G|$ of a graph G to be $|V(G)| + |E(G)|$. Johnson [J1] defined the *subgraph metric* $d_s : \Gamma \times \Gamma \to \mathbf{Z}^+$ such that $d_s(G, H)$ is the minimum of $|G| + |H| - 2|C|$ taken over all graphs C which are isomorphic to subgraphs of both G and H.

3.2.10 Definition

Zelinka [Z1] introduced a distance on the space S_n of all isomorphism classes of graphs with n vertices. The *induced subgraph distance* d_i is defined so that if $\sigma_1, \sigma_2 \in S_n$ and $n+k$ is the least possible number of vertices of a graph containing an induced subgraph from each of the classes σ_1 and σ_2, then $d_i(\sigma_1, \sigma_2) = k$.

If G_1 and G_2 are two graphs of order n, the induced subgraph distance between G_1 and G_2, denoted by $d_i(G_1, G_2)$, is defined to be the induced distance between the isomorphism classes of graphs containing G_1 and G_2, respectively. Hence $d_i(G_1, G_2)$ is the smallest number k for which there exists a graph G of order $k+n$ which contains induced subgraphs G_1' and G_2', isomorphic to G_1 and G_2, respectively.

Since the graph H obtained from the disjoint union of G_1 and G_2 by identifying a vertex of G_1 with a vertex of G_2 clearly contains both G_1 and G_2 as induced subgraphs, it follows immediately that $d_i(G_1, G_2)$ exists and that $d_i(G_1, G_2) \leq n-1$.

The following three results are from [Z1].

3.2.11 Theorem

Let n be a positive integer and k a nonnegative integer. Let G_1 and G_2 be graphs of order n. Then the following two assertions are equivalent:

1. There exists a graph G of order at most $n+k$ having induced subgraphs G_1' and G_2' such that $G_1' \cong G_1$ and $G_2' \cong G_2$.

2. There exist isomorphic graphs G_1'' and G_2'', each of order at least $n-k$, such that G_1'' is an induced subgraph of G_1 and G_2'' is an induced subgraph of G_2.

Proof

$1 \Rightarrow 2$: Since the graph $G_1 \cup G_2$ of order $2n$ vertices contains G_1 and G_2 as induced subgraphs, we have $k \leq n$. If two sets, each containing n elements, are subsets of a set with at most $n+k$ elements, where $k \leq n$, then their intersection contains at least $n-k$ elements. Thus the intersection of the vertex sets $V(G_1')$ and $V(G_2')$ contains at least $n-k$ elements; this set of vertices induces a subgraph G'' of G and of both G_1' and G_2', which is isomorphic to an induced subgraph G_1'' of G_1 and to an induced subgraph G_2'' of G_2.

$2 \Rightarrow 1$: Assume, without loss of generality, that G_1 and G_2 are vertex disjoint. Let φ be an isomorphism from G_1'' into G_2''. Let G be the graph obtained from $G_1 \cup G_2$ by identifying each vertex v of G_1'' with its image $\varphi(v)$ in G_2''. Evidently G has at most $n+k$ vertices. If we let $G_1' = G_1$ and $G_2' = G_2$, then statement 1 is satisfied and the proof is complete. □

We show now that the induced subgraph distance d_i is a metric on S_n.

3.2.12 Theorem

Let S_n be the space of all isomorphism classes of graphs with n vertices. Then the induced subgraph distance d_i together with S_n is a metric space.

Proof

Let $\sigma_1, \sigma_2, \sigma_3 \in S_n$.

i) Suppose $\sigma_1 = \sigma_2$, then the least possible number of vertices of a graph containing subgraphs from σ_1 and σ_2 is n, because any graph from the class $\sigma_1 = \sigma_2$ will be such a graph. Thus $d_i(\sigma_1, \sigma_2) = 0$. If $d_i(\sigma_1, \sigma_2) = 0$, then there exists a graph with n vertices containing induced subgraphs from σ_1 and σ_2. Each graph from a class of S_n has n vertices, and a graph with n vertices contains exactly one induced subgraph with n vertices, namely itself. Therefore the graph belongs to both of the isomorphism classes σ_1 and σ_2, thus $\sigma_1 = \sigma_2$. Therefore $d_i(\sigma_1, \sigma_2) = 0$ if and only if $\sigma_1 = \sigma_2$.

ii) That $d_i(\sigma_1, \sigma_2) = d_i(\sigma_2, \sigma_1)$ follows immediately from the definition of d_i.

iii) Let $d_i(\sigma_1, \sigma_2) = k_{12}$ and let $d_i(\sigma_2, \sigma_3) = k_{23}$. Then there exists a graph G_{12} with $n + k_{12}$ vertices which contains an induced subgraph $G_1 \in \sigma_1$ and an induced subgraph $G_2 \in \sigma_2$; and there exists a graph G_{23} with $n + k_{23}$ vertices which contains an induced subgraph $H_2 \in \sigma_2$ and an induced subgraph $H_3 \in \sigma_3$. Since both H_2 and G_2 belong to σ_2 we have that $G_2 \cong H_2$ and there exists an isomorphism ψ of G_2 onto H_2. Let G be the graph obtained from G_{12} and G_{23} by identifying each vertex v in G_2 with its image $\psi(v)$ in H_2. This graph has $n + k_{12} + n + k_{23} - n = n + k_{12} + k_{23}$ vertices and contains $G_1 \in \sigma_1$ and $H_3 \in \sigma_3$ as induced subgraphs; hence

$$d_i(\sigma_1, \sigma_3) \leq k_{12} + k_{23} = d_i(\sigma_1, \sigma_2) + d_i(\sigma_2, \sigma_3).$$

Thus the triangle inequality holds for d_i and the proof is complete.

□

The following theorem proves that the induced subgraph distance between two graphs $G_1, G_2 \in \Gamma(p)$ from classes $\sigma_1, \sigma_2 \in S_p$, respectively, is the same as the induced subgraph distance between the isomorphism classes containing the complements to the graphs of G_1 and G_2.

3.2.13 Theorem

Let $G_1, G_2 \in \Gamma(p)$, then $d_i(G_1, G_2) = d_i(\bar{G}_1, \bar{G}_2)$.

Proof

There exists a graph G with $p + d_i(G_1, G_2)$ vertices containing G_1 and G_2 as induced subgraphs. Now the complement $\bar{G}$ of G contains $\bar{G}_1$ and $\bar{G}_2$ as induced subgraphs and $\bar{G}$ has $p + d_i(G_1, G_2)$ vertices, therefore $d_i(\bar{G}_1, \bar{G}_2) \leq d_i(G_1, G_2)$. However, interchanging G_1 with $\bar{G}_1$ and G_2 with $\bar{G}_2$ in our argument, we obtain $d_i(G_1, G_2) \leq d_i(\bar{G}_1, \bar{G}_2)$ and therefore

$$d_i(G_1, G_2) = d_i(\bar{G}_1, \bar{G}_2). \qquad \square$$

Zelinka [Z2] introduced a metric analogous to the induced subgraph metric to study a distance between isomorphism classes of trees. Apart from Theorem 3.2.32 all results in the remainder of this section are from [Z2].

3.2.14 Definitions

Consider the set $\mathcal{T}_n$ of all isomorphism classes of trees with n vertices, $n \geq 3$. Let $\tau_1, \tau_2 \in \mathcal{T}_n$, then define the *tree metric* $d_T : \mathcal{T}_n \times \mathcal{T}_n \to \mathbf{Z}^+ \cup \{0\}$ such that $d_T(\tau_1, \tau_2)$ is the least integer with the property that there exists a tree with $n + d_T(\tau_1, \tau_2)$ vertices which contains a subtree, $T_1 \in \tau_1$ and a subtree $T_2 \in \tau_2$.

If T_1 and T_2 are two trees of order n, the *tree distance* between T_1 and T_2 denoted by $d_T(T_1, T_2)$, is defined to be the smallest integer k for which there exists a tree T of order $k + n$ which contains subtrees T_1' and T_2', isomorphic to T_1 and T_2 respectively.

Since the tree H obtained from the disjoint union of T_1 and T_2 by identifying a vertex of T_1 with a vertex of T_2 clearly contains both T_1 and T_2 as subtrees, it follows immediately that $d_T(T_1, T_2)$ exists and that

$$d_T(T_1, T_2) \leq n - 1.$$

3.2.15 Theorem

The functional d_T is a metric on the set $\mathcal{T}_n$.

Proof

Let $\tau_1, \tau_2, \tau_3 \in \mathcal{T}_n$.

i) By definition $d_T(\tau_1, \tau_2) \geq 0$ and $d_T(\tau_1, \tau_2) = 0$ if and only if there exists a tree T with n vertices such that $T \in \tau_1$ and $T \in \tau_2$; i.e., if and only if $\tau_1 = \tau_2$.

ii) Let $d_T(\tau_1, \tau_2) = m$. Then there exists a tree T with $n + m$ vertices which contains a subtree $T_2 \in \tau_2$ and a subtree $T_1 \in \tau_1$. Therefore $d_T(\tau_2, \tau_1) \leq m = d_T(\tau_1, \tau_2)$. Similarly $d_T(\tau_1, \tau_2) \leq d_T(\tau_2, \tau_1)$ and therefore $d_T(\tau_1, \tau_2) = d_T(\tau_2, \tau_1)$.

iii) There exists a tree T_{12} with $n + d_T(\tau_1, \tau_2)$ vertices which contains a subtree $T_1 \in \tau_1$ and a subtree $T_2 \in \tau_2$, and there exists a tree T_{23} with $n + d_T(\tau_2, \tau_3)$ vertices which contains a subtree $T_2' \in \tau_2$ and a subtree $T_3 \in \tau_3$. The trees T_2 and T_2' are isomorphic. From T_{12} and T_{23} we obtain the graph T by taking an isomorphic mapping of T_2 onto T_2' and identifying each vertex of T_2 with its image in this mapping. Now T is connected, has $n + d_T(\tau_1, \tau_2) + d_T(\tau_2, \tau_3)$ vertices and has $(n + d_T(\tau_1, \tau_2) - 1) + (n + d_T(\tau_2, \tau_3) - 1) - (n - 1) = n + d_T(\tau_1, \tau_2) + d_T(\tau_2, \tau_3) - 1$ edges.

Therefore $q(T) = p(T) - 1$ and therefore T is a tree. Now T contains a subgraph $T_1 \in \tau_1$ and a subgraph $T_3 \in \tau_3$, i.e. $d_T(\tau_1, \tau_3) \leq d_T(\tau_1, \tau_2) + d_T(\tau_2, \tau_3)$ and the triangle inequality holds. □

3.2.16 Definition

Denote by $\tau(n)$ the set of all trees of order n.

3.2.17 Theorem

Let $T_1, T_2 \in \tau(n)$ and let k be a nonnegative integer, $k < n$. Then the following two statements are equivalent:

1. There exists a tree T with $n + k$ vertices which contains a subtree isomorphic to T_1 and a subtree isomorphic to T_2.

2. There exists a tree T_0 with $n - k$ vertices such that both T_1 and T_2 contain subtrees isomorphic to T_0.

Proof

$1 \Rightarrow 2$: Suppose statement 1 is true. Let T_1' and T_2' be subtrees of T isomorphic to T_1 and T_2 respectively. Since $k < n$, T_1' and T_2' have a nonempty intersection and this intersection is a subtree T_0' of T. Now there are $n + k - n = k$ vertices not in T_1', therefore T_0' must have at least $n - k$ vertices. Choose a subtree T_0 of T_0' which has exactly $n - k$ vertices. Taking isomorphic mappings of T_1' onto T_1 and of T_2' onto T_2, the images of T_0 in these mappings must be subtrees of T_1 and T_2 and are isomorphic to one another and of course to T_0.

$2 \Rightarrow 1$ Suppose statement 2 is true. We may assume without loss of generality that T_1 and T_2 are vertex disjoint. Let T_0' and T_0'' be subtrees of T_1 and T_2 respectively, which are both isomorphic to T_0. Let T be the graph obtained from T_1 and T_2 by taking an isomorphic mapping of T_0' onto T_0'' and identifying each vertex of T_0' with its image in this mapping. The graph T constructed as in iii) in the proof of Theorem 3.2.15 is a tree. Now T has $n + n - (n - k) = n + k$ vertices and it contains T_1 and T_2 as subtrees. □

The tree metric distance graph $D_T(\mathcal{J}_n)$ is defined to be the graph whose vertex set is $\mathcal{J}_n$ and in which $\tau_1\tau_2 \in E(D_T(\mathcal{J}_n))$ if and only if $d_T(\tau_1, \tau_2) = 1$.

3.2.18 Theorem

The distance between any two vertices τ_1, τ_2 of $D_T(\mathcal{J}_n)$ is equal to $d_T(\tau_1, \tau_2)$.

Proof

Let $\tau_1, \tau_2 \in V(D_T(\mathcal{F}_n))$ and let $d_T(\tau_1, \tau_2) = k$. Then there exists a tree T with $n+k$ vertices which contains a subtree $T_1 \in \tau_1$ and a subtree $T_2 \in \tau_2$. Now since $n \geq 3$, P_3 is a subtree of every graph in $\mathcal{F}_n$ and we have from Theorem 3.2.17 that $n-k \geq 3$. Therefore $k \leq n-3$ and therefore T_1 and T_2 have a nonempty intersection containing, by Theorem 3.2.17, exactly $n-k$ vertices of T. Thus, there are k vertices of T_1 not belonging to T_2 and k vertices of T_2 not belonging to T_1. Let $\{u_1, ..., u_k\}$ be the set of vertices of T_1 not belonging to T_2 where each u_i is adjacent to either a common vertex of T_1 and T_2 or to a vertex u_j with $j < i$. Let $\{v_1, ..., v_k\}$ be the set of vertices of T_2 not belonging to T_1 such that each v_i is adjacent to either a common vertex of T_1 and T_2 or to a vertex v_j with $j > i$. For each $j = 1, 2, ..., k$, let S_j be the graph obtained from T_2 by deleting the vertices v_i for $i \leq j$ and adding the vertices u_i for $i \leq j$ together with the edges which join them and the edges which join them to the common vertices of T_1 and T_2 in T. Each graph S_i is a tree since S_i is a connected subgraph of T. It is evident that $S_k = T_1, d_T(T_2, S_1) = 1$ and $d_T(S_i, S_{i+1}) = 1$ for $i = 1, 2, ..., k-1$. The vertices $T_2, S_1, ..., S_k = T_1$ of $D_T(\mathcal{F}_n)$ (where the trees $T_2, S_1, ..., S_k$ represent the isomorphism classes containing them) form a path of length k in $D_T(\mathcal{F}_n)$. Therefore in $D_T(\mathcal{F}_n)$

$$d(\tau_1, \tau_2) \leq d_T(\tau_1, \tau_2). \qquad \sim (1)$$

Now suppose that $d(\tau_1, \tau_2) = \ell$ in $D_T(\mathcal{F}_n)$. Then there exists a path of length ℓ in $D_T(\mathcal{F}_n)$ consisting of the vertices $T_1 = S'_0, S'_1, ..., S'_\ell = T_2$. Thus $d_T(S'_i, S'_{i+1}) = 1$ for $i = 0, 1, ..., \ell - 1$. Let S''_i be a tree with $n+1$ vertices which contains a subtree isomorphic to S'_i and a tree isomorphic to S'_{i+1}.

For each $i = 0, ..., \ell - 2$ we choose an isomorphism of the subtree of S_i'' isomorphic to S_{i+1}' onto the subtree of S_{i+1}'' isomorphic to S_{i+1}' and identify each vertex of the domain of this mapping with its image. Then we obtain a tree with $n + \ell$ vertices which contains a subtree from τ_1 and a subtree from τ_2. Thus $d_T(\tau_1, \tau_2) \leq \ell$ and therefore

$$d_T(\tau_1, \tau_2) \leq d(\tau_1, \tau_2). \qquad \sim (2)$$

Together (1) and (2) imply that $d_T(\tau_1, \tau_2) = d(\tau_1, \tau_2)$. □

Thus according to Theorem 3.2.18 in order to determine the diameter of $D_T(\mathcal{F}_n)$, we may look for isomorphism classes in $\mathcal{F}_n$ which are furthest apart with respect to the tree distance. That is, if say $\tau_1, \tau_2 \in \mathcal{F}_n$ such that $d_T(\tau_1, \tau_2)$ is a maximum then

$$\text{diam} D_T(\mathcal{F}_n) = d_T(\tau_1, \tau_2).$$

This problem is resolved in the following theorem.

3.2.19 Theorem

The diameter of $D_T(\mathcal{F}_n)$ is $n - 3$. There is exactly one pair of vertices in $D_T(\mathcal{F}_n)$ between which the distance is $n - 3$.

Proof

We have already seen in the proof of Theorem 3.2.18 that every tree in $\mathcal{F}_n$ contains P_3 as a subtree. Let $\tau_1, \tau_2 \in \mathcal{F}_n$ and let $T_1 \in \tau_1$ and $T_2 \in \tau_2$; then by Theorem 3.2.17 there exists a tree with $n + (n - 3) = 2n - 3$ vertices which contains a subtree isomorphic to T_1 and a subtree isomorphic to T_2. Thus $d_T(\tau_1, \tau_2) \leq n-3$ for all $\tau_1, \tau_2 \in \mathcal{T}_n$ and thus by Theorem 3.2.18

$$d_{D_T(\mathcal{T}_n)}(\tau_1, \tau_2) \leq n-3, \qquad \sim (1)$$

for any pair of vertices $\tau_1, \tau_2 \in V(D_T(\mathcal{T}_n))$.

Now the path P_n and the star S_n are trees of order n and are therefore elements of $\mathcal{T}_n$, and thus vertices of $D_T(\mathcal{T}_n)$. Any subtree of P_n (S_n) with more than three vertices is a path (a star, respectively) with more than two edges. Therefore for $T_1 \cong P_n$ and $T_2 \cong S_n$ statement 2 of Theorem 3.2.17 holds only for $n-k \leq 3$; i.e., for $k \geq n-3$. Thus statement 2 of Theorem 3.2.17 does not hold for $k < n-3$ and thus statement 1 does not hold either for $k < n-3$. Thus

$$d_T(P_n, S_n) = n-3 \qquad \sim (2)$$

and so the isomorphism classes containing P_n and S_n have tree distance $n-3$ between them.

Together (1) and (2) imply that the diameter of $D_T(\mathcal{T}_n)$ is $n-3$.

Finally we show that P_n and S_n are unique in that the isomorphism classes containing them are the only ones to have a tree distance of $n-3$ between them. Any tree $T_1 \in \tau_1$ with $n \geq 4$ vertices which is neither a path nor a star contains P_4 and S_4 as subtrees. Let $\tau_2 \in \mathcal{T}_n$ such that $\tau_1 \neq \tau_2$. Then P_4 and S_4 are subgraphs of T_1 and T_2 for any $T_1 \in \tau_1$ and $T_2 \in \tau_2$. Thus statement 2 of Theorem 3.2.17 holds for $n-k=4$ and therefore there exists a tree T with $n+(n-4)$ vertices which contains a subtree isomorphic to T_1 and a subtree isomorphic to T_2. Therefore $d_T(\tau_1, \tau_2) \leq n-4$ and by Theorem 3.2.18 the distance of τ_1 from any other vertex in $D_T(\mathcal{T}_n)$ is at most $n-4$. □

3.2.20 Corollary

The tree distance between any two isomorphism classes $\tau_1, \tau_2 \in \mathcal{F}_n$ is at most $n-3$. The isomorphism classes τ_1 and τ_2 which contain P_n and S_n, respectively, are unique in the sense that $d_T(\tau_1, \tau_2) = n-3$.

Proof

Immediately from Theorems 3.2.18 and 3.2.19. □

3.2.21 Definition

We define the tree $T(k)$ for all positive integers $k \geq 3$ as follows: First we define the graph $T_0(k)$. The vertex set of $T_0(k)$ consists of all vectors with dimensions $0, 1, 2, ..., \lceil \frac{k}{2} \rceil - 1$, whose coordinates are numbers from the set $\{1, 2, ..., k-1\}$. Thus $T_0(k)$ has $1 + \sum_{i=1}^{\lceil \frac{k}{2} \rceil - 1} (k-1)^i$ vertices. Two vectors u, v are adjacent in $T_0(k)$ if and only if one of them can be obtained from the other by adding one coordinate. If k is odd, take two disjoint copies of $T_0(k)$ and add an edge between the vertices which correspond to the zero vector in both of them. If k is even, we take a new vertex c and k pairwise disjoint copies of $T_0(k)$ and insert edges between c and the vertices corresponding to the zero vector in all of them. The tree obtained in this way will be denoted by $T(k)$. For k odd, $T(k)$ has $2 + 2 \sum_{i=1}^{\lceil \frac{k}{2} \rceil - 1} (k-1)^i$ vertices and for k even $T(k)$ has $1 + k + k \sum_{i=1}^{\frac{k}{2}-1} (k-1)^i = 1 + k \sum_{i=1}^{\frac{k}{2}} (k-1)^{i-1}$.

3.2.22 Lemma

The tree $T(k)$ has the maximum number of vertices among all trees with diameter at most k and with maximum degree at most k.

Proof

Let T be a tree with diameter k and maximum degree k. If k is even, then T has one central vertex c, and for all $v \in V(T)$, $d(c,v) \leq \frac{k}{2}$. Since $\triangle(T) = k$, for each $i = 1, 2, ..., \frac{k}{2}$ there are at most $k(k-1)^{i-1}$ vertices of T whose distance from c is i. Thus T has at most $1 + k \sum_{i=1}^{\frac{k}{2}} (k-1)^{i-1}$ vertices and this is the number of vertices in $T(k)$.

If k is odd, then T has two centres c_1 and c_2 which are adjacent. For each $i = 1, 2, ..., \lceil \frac{k}{2} \rceil$ -1 there are at most $2(k-1)^i$ vertices v of T such that $\min_{j=1,2}\{d(c_j, v)\} = i$. Thus T has at most $2 + 2 \sum_{i=1}^{\lceil \frac{k}{2} \rceil - 1} (k-1)^i$ vertices and this is the number of vertices in $T(k)$. $\square$

Let $\alpha(k)$ denote the number of vertices in $T(k)$ for any integer $k \geq 3$. By Definition 3.2.21

$$\alpha(k) = 1 + k \sum_{i=1}^{\frac{k}{2}} (k-1)^{i-1} \text{ for } k \text{ even and,}$$

$$\alpha(k) = 2 \sum_{i=1}^{\lceil \frac{k}{2} \rceil} (k-1)^{i-1} \text{ for } k \text{ odd.}$$

Further, for $n \geq 6$ we denote

$$\sigma(n) = \max\{k \in \mathbf{Z}^+ : \alpha(k) \leq n\}.$$

3.2.23 Theorem

Let the radius of $D_T(\mathcal{J}_n)$ be ρ. Then $\rho \leq n - \sigma(n) - 1$.

Proof

Let $k = \sigma(n)$. We shall construct a tree C: If $\alpha(k) = n$, then let $C \cong T(k)$. If $\alpha(k) < n$, then let C be an arbitrary tree with n vertices which contains $T(k)$ as a subtree. Let T be any tree with n vertices.

Case 1) Suppose diam $T \geq k$. The tree C contains $T(k)$ as a subtree, and we know from Lemma 3.2.22 that the diameter of $T(k)$ is k. Therefore both C and T contain P_{k+1} as a subtree. Hence by Theorem 3.2.17 there exists a tree with $n + (n - k - 1)$ vertices containing C and T as subtrees. Therefore if z and τ are the isomorphism classes containing C and T respectively, it is evident that

$$d_T(z, \tau) \leq n - k - 1.$$

Case 2) Suppose diam $T < k$, then since T has $n \geq \alpha(k)$ vertices, by Lemma 3.2.22, its maximum degree must be greater than k. But since C contains $T(k)$ as a subgraph, and $\triangle(T(k)) = k$, we must have $\triangle(C) \geq k$. Therefore both C and T contain S_{k+1} as a subtree and again

$$d_T(z, \tau) \leq n - k - 1.$$

The tree distance of z from the isomorphism class containing P_n and from the isomorphism class containing S_n is exactly $n - k - 1$ (by Theorem 3.2.17). Therefore the radius of $D_T(\mathcal{J}_n)$ is at most $n - k - 1 = n - \sigma(n) - 1$.□

3.2.24 Conjecture

The radius of $D_T(\mathcal{F}_n)$ is equal to $n - \sigma(n) - 1$.

We now study the class of trees called caterpillars. Recall that a caterpillar is a tree with the property that after deleting all of its end-vertices we are left with a path, called the body of the caterpillar. (A graph consisting of one vertex is considered a path.)

3.2.25 Theorem

Let T_1 and T_2 be caterpillars of order n and let $d_T(T_1, T_2) = k$; then there exists a caterpillar T with $n+k$ vertices which contains a subtree isomorphic to T_1 and a subtree isomorphic to T_2.

Proof

As $d_T(T_1, T_2) = k$, we have, by Theorem 3.2.17, that there exists a tree T_0 with $n - k$ vertices such that both T_1 and T_2 contain subtrees isomorphic to T_0. Since P_3 is a subtree of all trees in $\mathcal{F}_n$ we have $n - k \geq 3$, and T_0 has at least two edges. T_0 is a subtree of a caterpillar, and thus T_0 itself is a caterpillar. Let $B(T_1), B(T_2)$ and $B(T_0)$ be the bodies of the caterpillars T_1, T_2 and T_0 respectively. Let T_0' and T_0'' be subtrees of T_1 and T_2 respectively, which are both isomorphic to T_0. Take an isomorphism of T_0' onto T_0'' and let T be the tree obtained from T_1 and T_2 by identifying each vertex of T_0' with its image in this isomorphism. If T is not a caterpillar, then there exists an edge e_1 of $B(T_1)$ not belonging to $B(T_2)$ and an edge e_2 of $B(T_2)$ not belonging to $B(T_1)$ such that they are both incident to a vertex v_0 of $B(T_0)$. Let v_1 (v_2) be the vertex incident with e_1 (e_2, respectively) distinct from v_0. Now by identifying the vertices v_1 and v_2 in T, we obtain a tree with $n+k-1$ vertices which contains T_1 and T_2 as subtrees; this contradicts the fact that $d_T(T_1, T_2) = k$. Thus T is a caterpillar and the theorem is proved. $\square$

3.2.26 Corollary

The set of all isomorphism classes of caterpillars with n vertices induces a subgraph $\tilde{D}_T(\mathcal{T}_n)$ of $D_T(\mathcal{T}_n)$ with the property that the distance between two vertices in $\tilde{D}_T(\mathcal{T}_n)$ is the same as in $D_T(\mathcal{T}_n)$. The diameter of $\tilde{D}_T(\mathcal{T}_n)$ is $n-3$.

Proof

Let $\tau_1, \tau_2 \in \mathcal{T}_n$ such that $T_1 \in \tau_1$, $T_2 \in \tau_2$ and T_1 and T_2 are caterpillars. Let $d_T(\tau_1, \tau_2) = k$. Then, by Theorem 3.2.18, $d(\tau_1, \tau_2) = k$ in $D_T(\mathcal{T}_n)$. Now by Theorem 3.2.25 we have that there exists a caterpillar T with $n+k$ vertices which contains a subtree isomorphic to T_1, and a subtree isomorphic to T_2, therefore the following analogue to Theorem 3.2.18 holds:

The distance between any two vertices τ_1, τ_2 of $\tilde{D}_T(\mathcal{T}_n)$ is equal to $d_T(\tau_1, \tau_2)$. The proof is exactly as in Theorem 3.2.18 with the word tree replaced by caterpillar. Therefore

$$d_{D_T(\mathcal{T}_n)}(\tau_1, \tau_2) = d_{\tilde{D}_T(\mathcal{T}_n)}(\tau_1, \tau_2).$$

Since the star S_n and the path P_n are both caterpillars we have by Theorem 3.2.19 that the diameter of $\tilde{D}_T(\mathcal{T}_n)$ is $n-3$. $\square$

Now for every positive integer k we construct the caterpillar $\tilde{T}(k)$. Let the body of $\tilde{T}(k)$ consist of a path of length $k-2$. Let the degree of every vertex of the body in $\tilde{T}(k)$ be k. The number of vertices in $\tilde{T}(k)$ is $(k-1)+(k-3)(k-2)+2(k-1)=k^2-2k+3$.

3.2.27 Lemma

The caterpillar $\tilde{T}(k)$ has maximum order among all caterpillars with diameter at most k and maximum degree at most k.

Proof

The diameter of a caterpillar is always the length of its body plus two. Thus $\tilde{T}(k)$ has diameter k. In a caterpillar the only vertices which can have degree greater than one are the vertices which belong to the body of the caterpillar. Since every vertex contained in the body of $\tilde{T}(k)$ has degree k, the lemma is proved. $\square$

3.2.28 Theorem

Let $\tilde{\rho}$ be the radius of $\tilde{D}_T(\mathcal{F}_n)$. Then

$$\tilde{\rho} \leq n - \tilde{\sigma}(n) - 1 \text{ where}$$

$$\tilde{\sigma}(n) = \max\{k \in \mathrm{Z}^+ : \; k^2 - 2k + 3 \leq n\}.$$

Proof

Let $k = \tilde{\sigma}(n)$. We construct the caterpillar $\tilde{C}$. If $k^2-2k+3=n$ then let $\tilde{C} \cong \tilde{T}(k)$. If $k^2-2k+3<n$ then let $\tilde{C}$ be an arbitrary caterpillar with n vertices which contains $\tilde{T}(k)$ as a subtree. Let $\tilde{T}$ be any caterpillar with n vertices.

Case 1) Suppose diam $\tilde{T} \geq k$. The caterpillar $\tilde{C}$ contains $\tilde{T}(k)$ as a subtree, and we know from Lemma 3.2.27 that the diameter of $\tilde{T}(k)$ is k. Therefore both $\tilde{C}$ and $\tilde{T}$ contain the caterpillar P_{k+1} as a subtree. Hence by Theorem 3.2.17 there exists a tree $\tilde{T}'$ with $n + (n - k - 1)$ vertices containing $\tilde{C}$ and $\tilde{T}$ as subtrees. That $\tilde{T}'$ is a caterpillar follows from Theorem 3.2.25. Therefore if $\tilde{z}$ and $\tilde{\tau}$ are the isomorphism classes containing $\tilde{C}$ and $\tilde{T}$ respectively then

$$d_T(\tilde{z}, \tilde{\tau}) \leq n - k - 1.$$

Case 2) Suppose diam $\tilde{T} < k$, then since $\tilde{T}$ has $n \geq k^2 - 2k + 3$ vertices by Lemma 3.2.27, $\triangle(\tilde{T}) > k$. But since $\tilde{C}$ contains $\tilde{T}(k)$ as a subgraph, and $\triangle(\tilde{T}(k)) = k$, we must have $\triangle(\tilde{C}) \geq k$. Therefore both $\tilde{C}$ and $\tilde{T}$ contain the caterpillar S_{k+1} as a subtree and again

$$d_T(\tilde{z}, \tilde{\tau}) \leq n - k - 1.$$

The tree distance of $\tilde{z}$ from the isomorphism class containing P_n and from the isomorphism class containing S_n is exactly $n - k - 1$. Thus the radius of $\tilde{D}_T(\mathcal{J}_n)$ is at most

$$n - k - 1 = n - \tilde{\sigma}(n) - 1;$$

i.e.,

$$\tilde{\rho} \leq n - \tilde{\sigma}(n) - 1. \qquad \square$$

3.2.29 Conjecture

The radius of $\tilde{D}_T(\mathcal{J}_n)$ is equal to $n - \tilde{\sigma}(n) - 1$.

We now compare the tree distance d_T with the induced subgraph metric d_i (see Definition 3.2.10).

3.2.30 Theorem

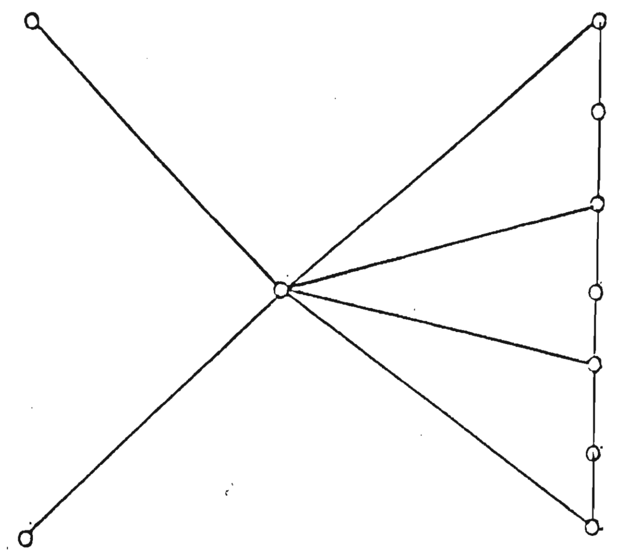
For elements $T_1, T_2 \in \tau(n)$ for $n \geq 7$ the distances $d_T(T_1, T_2)$ and $d_i(T_1, T_2)$ are different in general.

Proof

Let $T_1 \cong P_n$ and T_2 be the star on n-vertices. Then by Theorem 3.2.19, $d_T(T_1, T_2) = n - 3$. In T_1 take a maximal independent set of vertices. This set will obviously contain $\lceil \frac{n}{2} \rceil$ vertices. Identify each vertex of this set with one end-vertex of T_2. We obtain a graph G_n with $\lfloor \frac{3n}{2} \rfloor$ vertices which contains T_1 and T_2 as induced subgraphs. Thus

$$d_i(T_1, T_2) \leq \lfloor \frac{3n}{2} \rfloor - n = \lfloor \frac{n}{2} \rfloor < n - 3 = d_T(T_1, T_2). \qquad \square$$

G_7:



3.2.31 Figure

The graph G_n described in Theorem 3.2.30 for $n = 7$.

To end this section we present a new result analogous to Theorem 3.2.18 for the induced subgraph distance (see Definition 3.2.10).

We define the *induced subgraph metric distance graph* $D_{d_i}(S_n)$ to be the graph with vertex set S_n and in which $\sigma_1\sigma_2 \in E(D_{d_i}(S_n))$ if and only if $d_i(\sigma_1, \sigma_2) = 1$.

3.2.32 Theorem

The distance between any two vertices σ_1, σ_2 in $D_{d_i}(S_n)$ is equal to $d_i(\sigma_1, \sigma_2)$.

Proof

Let $\sigma_1, \sigma_2 \in S_n$ such that $d_i(\sigma_1, \sigma_2) = k$, then by definition there exists a graph G with $n+k$ vertices which contains induced subgraphs G_1 and G_2 such that $G_1 \in \sigma_1$ and $G_2 \in \sigma_2$. By Theorem 3.2.11, there exist isomorphic graphs G_1' and G_2' with $n-k$ vertices such that G_1' is an induced subgraph of G_1 and G_2' is an induced subgraph of G_2.

In G there are $n-k$ vertices common to G_1 and G_2. Therefore there are k vertices in G_1 which are not in G_2 and k vertices in G_2 not in G_1. Label arbitrarily the vertices in $V(G_1) - V(G_2)$ as $u_1, u_2, .., u_k$ and those in $V(G_2) - V(G_1)$ as $v_1, v_2, ..., v_k$. Now for each $j = 1, 2, ..., k$ let F_j be the subgraph of G induced by the vertex set $V(F_j) = (V(G_1) - \{u_1, u_2, ..., u_j\}) \cup \{v_1, v_2, ..., v_j\}$. Let β_j be the isomorphism class containing F_j. Then it is evident that $\beta_j \in S_n$ and $F_j \neq F_i$ for $j \neq i$. It is also evident that $F_k \cong G_2$, $d_i(F_1, G_1) = 1$ and $d_i(F_j, F_{j+1}) = 1$ for $j = 1, 2, ..., k-1$. Therefore we have constructed a path $G_1 F_1 F_2 ... F_k \cong G_2$ of length k in $D_{d_i}(S_n)$ and therefore

$$d_i(G_1, G_2) \geq d_{D_{d_i}(S_n)}(G_1, G_2). \qquad \sim (1)$$

Now suppose that the distance between G_1 and G_2 in $D_{d_i}(S_n)$ is ℓ. Then there exists a path of length ℓ in $D_{d_i}(S_n)$ of the form

$$G_1 \cong F_0' F_1' ... F_\ell' \cong G_2.$$

We have that $d_i(F_j', F_{j+1}') = 1$ for $j = 0, 1, ..., \ell - 1$. Let F_j'' be the graph with $n+1$ vertices which contains an induced subgraph isomorphic to F_j'

and an induced subgraph ismorphic to F'_{j+1} for $j = 0, 1, \ldots, \ell - 1$. For each $j = 0, 1, \ldots, \ell - 2$ let F'''_j be the graph obtained from F''_j and F''_{j+1} by choosing an isomorphism of the induced subgraph of F''_j which is isomorphic to F'_{j+1} and identifying each vertex of the domain of this mapping with its image. Then $F'''_{\ell-2}$ is a graph with $n + \ell$ vertices containing G_1 and G_2 as induced subgraphs. Therefore

$$d_i(G_1, G_2) \leq \ell = d_{D_{d_i}(S_n)}(G_1, G_2). \tag{2}$$

Together (1) and (2) imply that

$$d_i(G_1, G_2) = d_{D_{d_i}(S_n)}(G_1, G_2). \qquad \square$$

3.3 Bounds and Relations

We now present some results which give bounds and relationships between some of the metrics we have studied.

As we have already noted for all graphs $G, H \in \Gamma_c(p, q)$

$$d_{em}(G, H) \leq d_{er}(G, H) \leq d_{es}(G, H).$$

Our first result from [BGMW1] stems from the fact that any edge move can be simulated by a maximum of two edge rotations as was demonstrated in the proof of Theorem 2.3.20.

3.3.1 Lemma

For graphs $G_1, G_2 \in \Gamma(p, q)$

$$d_{er}(G_1, G_2) \leq 2d_{em}(G_1, G_2).$$

Proof

Let G be a greatest common subgraph of G_1 and G_2 with size s say. According to Theorem 3.2.7, $d_{em}(G_1, G_2) = q - s$. From Theorem 2.3.20, we have $d_{er}(G_1, G_2) \leq 2(q - s)$, and hence

$$d_{er}(G_1, G_2) \leq 2d_{em}(G_1, G_2) \qquad \square$$

The following lemma demonstrates that the ratio $\frac{d_{es}(G_1,G_2)}{d_{er}(G_1,G_2)}$ can be made arbitrarily large for graphs $G_1, G_2 \in \Gamma_c(p, q)$ and sufficiently large values of p.

3.3.2 Lemma

There exist graphs $G_1, G_2 \in \Gamma_c(p, q)$ such that for any integer $a > 0$ $\frac{d_{es}(G_1,G_2)}{d_{er}(G_1,G_2)} = a$.

Proof

Let graph G_1 consist of two disjoint paths $P_1 = v_1 v_2 ... v_{2a+1}$ and $P_2 = u_1 u_2 ... u_{2a+1}$ of length $2a$, together with the edge $u_{a+1}v_{a+1}$ and let graph G_2 be a path of length $4a + 1$. Now $d_{em}(G_1, G_2) = 1$. The edge rotations $t_1 = (v_{a+1}, u_{a+1}, u_{2a+1})$ on G_1 and $t_2 = (u_{2a+1}, v_{a+1}, v_{2a+1})$ on t_1G_1 transform G_1 into G_2; i.e., $t_2t_1G_1 \cong G_2$. Hence $d_{er}(G_1, G_2) \leq 2$. That $d_{er}(G_1, G_2) \geq 2$ is shown as follows:

The parameter $end'(G)$ of a graph G is slowly changing with respect to the edge rotation operation (see Lemma 2.3.28) and

$$|end'(G_1) - end'(G_2)| = |4 - 2| = 2,$$

hence $d_{er}(G_1, G_2) \geq 2$. It now follows that $d_{er}(G_1, G_2) = 2$.

Define the edge slides $t_i = (u_{a+1}, v_{a+i}, v_{a+i+1})$ for $i = 1, 2, ..., a$ and $s_j = (v_{2a+1}, u_{a+j}, u_{a+j+1})$ for $j = 1, 2, ..., a$. Then $s_a ... s_2 s_1 t_a ... t_2 t_1 G_1 \cong G_2$, and hence

$$d_{es}(G_1, G_2) \leq 2a.$$

However since the diameter of a graph is slowly changing with respect to edge slides (see Lemma 2.2.16), we have that $d_{es}(G_1, G_2) \geq |\text{diam}G_1 - \text{diam}G_2)| = |(2a+1) - (4a+1)| = 2a$. Hence $d_{es}(G_1, G_2) = 2a$.

Therefore $\frac{d_{es}(G_1,G_2)}{d_{er}(G_1,G_2)} = \frac{2a}{2} = a$ for arbitrary a. $\square$

Note that if we do not restrict ourselves to connected graphs; i.e., consider graphs $G_1, G_2 \in \Gamma(p,q)$, then the ratio $\frac{d_{es}(G_1,G_2)}{d_{er}(G_1,G_2)}$ can be made infinite by taking any disconnected graph $G_1 \in \Gamma(p,q)$ and any connected graph $G_2 \in \Gamma_c(p,q)$.

Our next four results are from [GS1]. We aim to provide a relationship between the edge slide and edge rotation distances between two graphs $G, H \in \Gamma(p,q)$.

3.3.3 Theorem

Let $G_1, G_2 \in \Gamma(p,q)$ then

$$d_{er}(G_1 \cup K_1, G_2 \cup K_1) = d_{er}(G_1 + K_1, G_2 + K_1) = d_{er}(G_1, G_2).$$

Proof

Since complementation preserves the edge rotation distance

$$\begin{aligned} d_{er}(G_1 + K_1, G_2 + K_1) &= d_{er}(\overline{G_1 + K_1}, \overline{G_2 + K_1}) = d_{er}(\bar{G}_1 \cup K_1, \bar{G}_2 \cup K_1) \\ &= d_{er}(\bar{G}_1, \bar{G}_2) = d_{er}(G_1, G_2). \end{aligned}$$

Hence we need only prove that $d_{er}(G_1 \cup K_1, G_2 \cup K_1) = d_{er}(G_1, G_2.)$ Trivially

$$d_{er}(G_1 \cup K_1, G_2 \cup K_1) \leq d_{er}(G_1, G_2) \qquad \sim (1)$$

Consider a sequence of edge rotations $t_1, t_2, ..., t_n$ such that $t_n...t_2t_1(G_1 \cup K_1) \cong G_2 \cup K_1$. Let x be the designated isolated vertex at the start of the transformation and let y be the designated isolated vertex at the end of the transformation. Consider all edge rotations of the form $t_i = (u_i, y, w_i)$ where $u_i, w_i \in V(G_1)$ (if an edge incident with y is to be rotated twice, arrange it so that it releases from y first). Thus all such operations reduce the neighbourhood of y. Of these, perform all the edge rotations which do not involve x. Now interchange the labels y and x, this will affect all the rotations including y and x. Now continue with the edge rotation sequence. Note that there is a designated isolated vertex throughout the transformation and therefore

$$d_{er}(G_1 \cup K_1, G_2 \cup K_1) \geq d_{er}(G_1, G_2) \qquad \sim (2)$$

Together (1) and (2) imply that $d_{er}(G_1 \cup K_1, G_2 \cup K_1) = d_{er}(G_1, G_2)$. □

3.3.4 Theorem

Let $G, H \in \Gamma(p, q)$, then $d_{es}(G + K_1, H + K_1) \leq 2d_{er}(G, H)$.

Proof

Both $G + K_1$ and $H + K_1$ are connected graphs and thus by Theorem 2.2.9, we know that $d_{es}(G + K_1, H + K_1)$ is finite. It is thus only necessary to show that for $x, y, z \in V(G), xy \in E(G)$ and $xz \in E(\bar{G})$, any edge rotation $t = (x, y, z)$ on G can be simulated by two edge slides t_1 and

t_2 on $G + K_1$ and $t_1(G + K_1)$, respectively. Let v be the joined vertex in $G + K_1$. Define the edge slides $t_1 = (x, v, z)$ and $t_2 = (x, y, v)$. Then $xy \in E(t_2t_1(G + K_1))$, $xz \in E(t_2t_1(G+K_1))$, while vertex v still has degree p, in $t_2t_1(G + K_1)$. Therefore $d_{es}(G + K_1, H + K_1) \leq 2d_{er}(G, H)$. □

3.3.5 Theorem

Let $G, H \in \Gamma_c(p, q)$ where $p \geq 3$ and $q \geq 2$, then

$$d_{es}(G, H) \leq 2d_{er}(G, H) - (\triangle(G) + \triangle(H)) + 6p - 6.$$

Proof

Let $v(w)$ be a vertex of maximum degree in G (H, respectively). Let $T_v(T_w)$ be a spanning tree of $G(H$, respectively) containing all the edges incident with $v(w$, respectively) (e.g. a spanning tree of $G(H)$ that is distance preserving from $v(w$, respectively)). Let $G' = G - E(T_v)(H' = H - E(T_w))$ and let $G'' = G' + \{vu : u \in V(G) - \{v\}\}$ $(H'' = H' + \{wu : u \in V(G) - \{w\}\})$.

Note, by Theorem 2.2.14, that

$$\left.\begin{aligned} d_{es}(G, G'') &= d_{es}(T_v, K_{1,p-1}) = \triangle(S_p) - \triangle(T_v) = p - 1 - \triangle(G); \\ d_{es}(H, H'') &= d_{es}(T_w, K_{1,p-1}) = \triangle(S_p) - \triangle(T_w) = p - 1 - \triangle(H). \end{aligned}\right\} \sim (1)$$

Let $t_1, t_2, ..., t_n$ be a sequence of $n = d_{er}(G, H)$ edge rotations such that $t_n...t_1G \cong H$. Let H^* be the graph obtained by restricting the edge rotation sequence $t_1, t_2, ..., t_n$ to G'; then since H^* and H' must have a greatest common subgraph of size at least $q - (p - 1)$

$$d_{em}(H^*, H') \leq p - 1,$$

hence by Lemma 3.3.1,

$$d_{er}(H^*, H') \leq 2(p-1) \text{ and so}$$

$$d_{er}(G', H') \leq d_{er}(G, H) + 2(p-1).$$

We note that v and w are isolated vertices of G' and H' respectively, so from Theorem 3.3.3,

$$d_{er}(G' - v, H' - w) \leq d_{er}(G, H) + 2(p-1). \qquad \sim (2)$$

Now

$$\begin{aligned} d_{es}(G'', H'') &= d_{es}((G' - v) + K_1, (H' - w) + K_1) \\ &= 2d_{er}(G' - v, H' - w) \text{ (by Theorem 3.3.4)} \\ &\leq 2[d_{er}(G, H) + 2(p-1)] \text{ (from (2))} \\ &= 2d_{er}(G, H) + 4(p-1). \end{aligned}$$

$$\sim (3)$$

It is evident that

$$d_{es}(G, H) \leq d_{es}(G, G'') + d_{es}(G'', H'') + d_{es}(H'', H);$$

therefore from (1) and (3)

$$\begin{aligned} d_{es}(G, H) &\leq p - 1 - \triangle(G) + 2d_{er}(G, H) + 4(p-1) + p - 1 - \triangle(H) \\ &= 2d_{er}(G, H) - (\triangle(G) + \triangle(H)) + 6p - 6. \end{aligned}$$

$\square$

3.3.6 Corollary

Let $G, H \in \Gamma_c(p, q)$ where $p \geq 3$ and $q \geq 2$; then

$$d_{es}(G, H) \leq 2d_{er}(G, H) + 6p - 10.$$

Proof

Since both G and H are connected and have $q \geq 2$ edges, $\triangle(G) \geq 2$ and $\triangle(H) \geq 2$. Hence $\triangle(G) + \triangle(H) \geq 4$ and the result follows directly from Theorem 3.3.5. $\square$

The following theorem by Zelinka [Z3] shows that the induced subgraph distance between two graphs is bounded above by the edge rotation distance of those two graphs (cf. [Z3] for results 3.3.7 to 3.3.11).

3.3.7 Theorem

Let $\sigma_1, \sigma_2 \in S_{p,q}, G_1 \in \sigma_1$ and $G_2 \in \sigma_2$. Then

$$d_i(G_1, G_2) \leq d_{er}(G_1, G_2)$$

where equality may occur.

Proof

If $d_{er}(G_1, G_2) = 1$, then G_2 may be obtained from G_1 by a single edge rotation. Hence there exists a graph G with p vertices and $q - 1$ edges which is isomorphic to a subgraph of G_1 and to a subgraph of G_2. Label the vertices of G_1 and G_2 so that the subgraphs isomorphic to G in G_1 and G_2 respectively, are identical. Suppose then that $V = V(G_1) = V(G_2) = \{v_0, v_1, ..., v_{p-1}\}$. Then there exist vertices, say, $v_0, v_1, v_2 \in V$ such that $v_0v_1 \in E(G_1)$ and $v_0v_2 \in E(\bar{G}_1)$, while $v_0v_1 \in E(\bar{G}_2)$ and $v_0v_2 \in E(G_2)$. For any other pair of vertices v_i, v_j, either $v_iv_j \in E(G_1)$ and $v_iv_j \in E(G_2)$ or $v_iv_j \in E(\bar{G}_1)$ and $v_iv_j \in E(\bar{G}_2)$. The set $V - \{v_0\}$ induces the same subgraph in both G_1 and G_2 and thus $d_i(G_1, G_2) = d_{er}(G_1, G_2) = 1$.

Now let $k \geq 2$ be an integer, and let $d_{er}(G_1, G_2) = k$. Then there exist graphs $H_0, H_1, ..., H_k$ such that $H_0 \cong G_1$ and $H_k \cong G_2$ and the graph H_i may be obtained from H_{i-1} by a single edge rotation for $i = 1, 2, ..., k$. We have $d_{er}(H_{i-1}, H_i) = 1$ and hence by the above $d_i(H_{i-1}, H_i) = 1$ for $i = 1, 2, ..., k$. Inductively from the triangle inequality we obtain

$$d_i(H_0, H_k) = d_i(G_1, G_2) \leq k = d_{er}(G_1, G_2). \qquad \square$$

The following result demonstrates that we can construct two graphs for which the difference between their edge rotation distance and induced subgraph distance may be chosen to be arbitrarily large.

3.3.8 Theorem

Let N be a positive integer. Let $\sigma_1, \sigma_2 \in S_{p,q}$, then there exist graphs $G_1 \in \sigma_1$ and $G_2 \in \sigma_2$ such that

$$d_{er}(G_1, G_2) - d_i(G_1, G_2) = N.$$

Proof

We construct graphs G_1 and G_2 with a common vertex set $V = \{u_1, u_2, ..., u_{N+1}, v_1, v_2, ..., v_{N+1}, w\}$. In G_1 the set $\{u_1, u_2, ..., u_{N+1}, w\}$ induces a clique and the vertices $v_1, v_2, ..., v_{N+1}$ are isolated. In G_2 two vertices are adjacent if and only if either they both belong to the set $\{u_1, u_2, ..., u_{N+1}\}$, or one of them is w and the other belongs to the set $\{v_1, v_2, ..., v_{N+1}\}$. Each of the graphs G_1 and G_2 has $\frac{1}{2}(N+1).(N+2)$ edges. The set $V - \{w\}$ induces the same subgraph in both G_1 and G_2, hence $d_i(G_1, G_2) = 1$. The graph G_2 can be obtained from G_1 by $N+1$ edge rotations; each rotation is of the form $t_i = (w, u_i, v_i)$ for $i = 1, 2, ..., N+1$;

i.e., $t_{N+1}, t_N, ..., t_2, t_1 G_1 \cong G_2$. Hence $d_{er}(G_1, G_2) \leq N + 1$. Now performing fewer than $N + 1$ edge rotations on G_1 will result in a graph with at least one isolated vertex. Since G_2 has no isolated vertices we have that $d_{er}(G_1, G_2) \geq N + 1$; and hence

$$d_{er}(G_1, G_2) = N + 1 \text{ and the result follows.} \qquad \square$$

The following lemma will aid us in proving that the edge rotation distance between two trees is bounded above by their tree distance.

3.3.9 Lemma

Let T be a tree with p vertices and edge set $E(T)$; let T_0 be a proper subtree of T with edge set $E(T_0)$. Then there exists a bijection f of the set $E(T) - E(T_0)$ onto the number set $\{1, 2, ..., |E(T) - E(T_0)|\}$ with the property that the set $E_i = E(T_0) \cup \{e \in E(T) - E(T_0) : f(e) \leq i\}$ is the edge set of a subtree of T for each $i \in \{1, 2, ...|E(T) - E(T_0)|\}$.

Proof

We proceed by induction on the cardinality of $E(T) - E(T_0)$. If $|E(T) - E(T_0)| = 1$, then $E_1 = E(T)$ and the assertion holds trivially. Let $k \geq 2$ be an integer and suppose that the assertion is true for $|E(T) - E(T_0)| \leq k - 1$.

Suppose $|E(T) - E(T_0)| = k$. There exists at least one edge $e_1 \in E(T) - E(T_0)$ which is incident with a vertex which is in T_0. Evidently $E(T_0') = E(T_0) \cup \{e_1\}$ is the edge set of a subtree T_0' of T. We have $|E(T) - E(T_0')| = k - 1$ and by the induction hypothesis there exists a bijection f' of $E(T) - E(T_0')$ onto $\{1, 2, ..., k-1\}$ such that the set $E_i' = E(T_0) \cup \{e \in E(T) - E(T_0') : f'(e) \leq i\}$ is the edge set of a subtree of T for each $i = 1, 2, ..., k - 1$. We define a bijection f of $E(T) - E(T_0)$ onto $\{1, 2, ..., k\}$ in such a way that $f(e_1) = 1$ and $f(e) = f'(e) + 1$ for each $e \in E(T) - E(T_0')$. Then evidently $E_1 = e_1$ and $E_i = E'_{i-1}$ for $i = 2, 3, ..., k$ and the assertion holds. □

3.3.10 Theorem

Let $n \in \mathrm{N}$, $\tau_1, \tau_2 \in \tau(n), T_1 \in \tau_1$ and $T_2 \in \tau_2$. Then

$$d_{er}(T_1, T_2) \leq d_T(T_1, T_2).$$

Proof

Let $d_T(T_1, T_2) = k$. This means, by Theorem 3.2.17, that the maximum number of vertices of a tree which is isomorphic to a subtree of T_1 and simultaneously isomorphic to a subtree of T_2 is equal to $n - k$. Suppose T_0 is a tree with $n - k$ vertices that is a subtree of both T_1 and T_2. Let $f_1(f_2)$ be a mapping corresponding to the mapping f from Lemma 3.3.9 where we consider $T_1(T_2$, respectively) instead of T. Both f_1 and f_2 are bijections onto the set $\{1, 2, ..., k\}$. For each $i = 1, 2, ..., k$ let $e_1(i)$ $(e_2(i))$ be the edge which is mapped by f_1 (f_2, respectively) onto the number i. The vertices incident with the edge $e_1(i)(e_2(i))$ will be denoted by $v_1(i)$ and $w_1(i)$ ($v_2(i)$ and $w_2(i)$, respectively) in such a way that the distance of $w_1(i)(w_2(i))$ from a vertex of T_0 is greater than the distance of $v_1(i)(v_2(i)$, respectively) from the same vertex. Now we identify $w_2(i)$ with $w_1(k+1-i)$ for $i = 1, 2, ..., k$. After this identification the trees T_1 and T_2 have the same vertex set. For $i = 1, 2, ..., k$ define the edge rotation t_i which deletes the edge $e_1(i) = v_1(i)w_1(i)$ and adds the edge $e_2(k + 1 - i) = v_2(k + 1 - i)w_2(k + 1 - i) = v_2(k + 1 - i)w_1(i)$, i.e. $t_i = (w_1(i), v_1(i), v_2(k+1-i))$. Then $t_k, t_{k-1}, ..., t_2, t_1 T_1 \cong T_2$. and hence

$$d_{er}(T_1, T_2) \leq k = d_T(T_1, T_2). \qquad \square$$

We now present a result which is similar to that of Theorem 3.3.8.

3.3.11 Theorem

Let N be a positive integer. Then there exist trees T_1 and T_2 of order n such that

$$d_T(T_1, T_2) - d_{er}(T_1, T_2) = N.$$

Proof

We construct trees T_1 and T_2 with a common vertex set
$V = \{u_1, u_2, u_3, u_4, u_5, u_6, v_1, v_2, ..., v_{2N+4}, w_1, w_2 ..., w_{N+2}\}$. Both T_1 and T_2 contain the edges $u_1u_2, u_2u_3, u_3u_4, u_4u_5, u_5u_6$ and u_2v_i for $i = 1, 2, ..., 2N+4$. Further, T_1 contains the edges u_4w_i and T_2 contains the edges u_5w_i for $i = 1, 2, ..., N+2$.

The subtree T_0 induced in both T_1 and T_2 by the set $\{u_1, u_2, u_3, u_4, u_5, u_6, v_1, ..., v_{2N+4}\}$ has $2N + 10$ vertices; evidently no tree with more vertices than T_0 can be isomorphic simultaneously to a subtree of T_1 and to a subtree of T_2. The set V contains $3N + 12$ vertices, hence

$$d_T(T_1, T_2) = (3N + 12) - (2N + 10) = N + 2.$$

Let $t_1 = (u_2, u_3, u_4)$ and $t_2 = (u_3, u_4, u_6)$ be edge rotations and let $T_1' = t_2 t_1 T_2$. Define the bijection $f : V \to V$ such that $f(u_3) = u_6, f(u_4) = u_3, f(u_5) = u_4, f(u_6) = u_5$ and $f(x) = x$ for each $x \in V - \{u_3, u_4, u_5, u_6\}$. The mapping f is an isomorphism of T_1' onto T_1. Hence $T_1' \cong T_1$ and T_1' was obtained from T_2 by two edge rotations. Evidently no single edge rotation will transform T_1 into T_2 (or vice versa) and hence $d_{er}(T_1, T_2) = 2$. The result follows. □

The following theorem gives an upper bound for the edge move distance between two graphs $G, H \in \Gamma(p, q)$. The next four results first appeared in [GS1].

3.3.12 Theorem

Let G and H have order p and size $\pi\binom{p}{2}$ where $0 \leq \pi \leq 1$. Then $d_{em}(G, H) \leq \binom{p}{2}\pi(1-\pi)$.

Proof

Consider a random bijection ϕ from G onto H. We want to determine the size of the greatest common subgraph of $\phi(G)$ and H that is induced by ϕ. For any edge e in G, the probability that ϕ maps e to an edge in H is π. Thus it is expected that $\pi\pi\binom{p}{2}$ edges of G will be mapped to edges in H and thus the expected size of the greatest common subgraph of $\phi(G)$ and H is $\pi^2\binom{p}{2}$. Thus by the probablistic method there exists a bijection ϕ from G onto H such that a greatest common subgraph of $\phi(G)$ and H has size at least $\pi^2\binom{p}{2}$. Hence G and H have a greatest common subgraph with size at least $\pi^2\binom{p}{2}$; thus

$$d_{em}(G, H) \leq \pi\binom{p}{2} - \pi^2\binom{p}{2} = \pi\binom{p}{2}(1-\pi) \text{ as required.} \qquad \square$$

Since the expression $\binom{p}{2}\pi(1-\pi)$ is maximised for $\pi = \frac{1}{2}$ the following result needs no further proof.

3.3.13 Corollary

The maximum distance between two graphs in $\Gamma(p,q)$ under the edge move distance is at most $\frac{p^2-p}{8}$.

From Corollary 3.3.13 above and Lemma 3.3.1 we also obtain the following result.

3.3.14 Corollary

The maximum distance between two graphs in $\Gamma(p,q)$ under the edge rotation distance is at most $\frac{p^2-p}{4}$.

The following result determines an upper bound for the edge slide distance between two graphs.

3.3.15 Corollary

The maximum distance between two graphs $G_1, G_2 \in \Gamma_c(p,q)$ under the edge slide distance is at most $\frac{p^2+11p}{2} - 10$.

Proof

From Corollary 3.3.6,

$$\begin{aligned} d_{es}(G_1,G_2) &\leq 2d_{er}(G,H) + 6p - 10 \\ &\leq 2\tfrac{p^2-p}{4} + 6p - 10 \text{ (by Collary 3.3.14)} \\ &\leq \tfrac{p^2-p}{2} + 6p - 10 = \tfrac{p^2+11p}{2} - 10. \end{aligned}$$

□

The following lemma will aid us in determining a lower bound for the edge rotation distance.

3.3.16 Lemma

If $\{a_i\}_{i=1,2,\ldots,n}$ and $\{b_i\}_{i=1,2,\ldots,n}$ are two sequences of n nonnegative integers with $a_1 \geq a_2 \geq \ldots \geq a_n$ and $D = \sum_{i=1}^{n} |a_i - b_i|$, then D is minimised when the sequence $\{b_i\}$ is arranged in nonascending order.

Proof

Suppose $\{b_i\}$ is given in some, not necessarily descending order. It can then be rearranged in nonascending order by means of a finite number of term-interchanges each of which involves a pair of terms in ascending order; i.e., b_i and b_j are interchanged if $b_i < b_j$ and $i < j$. Specifically, interchange the smallest b_i with b_n and the second smallest b_i with b_{n-1}, repeating this procedure until we have a nonascending sequence.

It remains to show that a single interchange, as described above, decreases the value of $\sum_{i=1}^{n} |a_i - b_i|$. Suppose that $b_p < b_q$ where $p < q$, then obviously it suffices to look at the sign of
$d = (|a_p - b_q| + |a_q - b_p|) - (|a_p - b_p| + |a_q - b_q|)$. If $d \leq 0$ then the lemma is proved. There are six cases to consider.

Case 1) Suppose $a_q \geq b_q$, then $d = 0$.

Case 2) Suppose $a_p \geq b_q > a_q \geq b_p$, then $d = -2|b_q - a_q|$.

Case 3) Suppose $a_p \geq b_q$ and $b_p \geq a_q$, then $d = -2|b_q - b_p|$.

Case 4) Suppose $b_q \geq a_p$ and $a_q \geq b_p$, then $d = -2|a_p - a_q|$.

Case 5) Suppose $b_q \geq a_p > b_p \geq a_q$, then $d = -2|a_p - b_p|$

Case 6) Suppose $b_p \geq a_p$, then $d = 0$. □

The following theorem is from [GS1].

3.3.17 Theorem

Let $G_1, G_2 \in \Gamma(p, q)$. Let the graph G_1 have degree sequence $d_1 \geq d_2 \geq \ldots. \geq d_p$ and let the graph G_2 have the degree sequence $e_1 \geq e_2 \geq \ldots \geq e_p$. Then

$$d_{er}(G_1, G_2) \geq \frac{1}{2}\sum_{i=1}^{p} |d_i - e_i|.$$

Proof

Let $V(G_1) = V(G_2) = \{v_1, v_2, \ldots, v_p\}$, where the vertices of G_1 and G_2 are labelled in such a way that $\deg_{G_1} v_i = d_i$ and $\deg_{G_2} v_i = e_i$ for $i = 1, 2, \ldots, p$.

Since each edge rotation increases the degree of exactly one vertex by 1 and decreases the degree of exactly one vertex by 1, it follows that $d_{er}(G_1, G_2) \geq \frac{1}{2}\sum_{i=1}^{p} |\deg_{G_1} v_i - \deg_{G_2} v_i|$. It is then an immediate consequence of Lemma 3.3.16 that

$$d_{er}(G_1, G_2) \geq \frac{1}{2}\sum_{i=1}^{p} |d_i - e_i|.$$ □

3.4 Graph Operations

Following [GS1] in this section, we now determine what effect some simple graph operations on graphs G_1 and G_2 have on the distance between them. The simplest operations are joining a vertex to a graph and adding an isolated vertex to a graph. We denote the joining of a vertex to G by $G + K_1$ and the adding of an isolated vertex to G by $G \cup K_1$.

It is clear that edge slide distance between graphs $G_1, G_2 \in \Gamma_c(p,q)$ is preserved when the same number of isolated vertices are added to both G_1 and G_2; i.e.,

$$d_{es}(G_1 \cup K_1, G_2 \cup K_1) = d_{es}(G_1, G_2).$$

On the other hand Theorem 3.3.4 showed that joining a vertex to both $G_1, G_2 \in \Gamma(p,q)$ can considerably reduce the edge slide distance between G_1 and G_2. In some cases (e.g. for $G_1 \in \Gamma_c(p,q)$ and $G_2 \in \Gamma(p,q)$ disconnected) it is possible to reduce the edge slide distance between two graphs from being infinite between G_1 and G_2, to being finite between $G_1 + K_1$ and $G_2 + K_1$.

By the greatest common subgraph formulation, we see immediately that both of the operations above preserve the edge move distance. These two operations are in fact complementary, and as we saw in Theorem 3.2.4 the edge move distance is preserved by complementation. While edge slide distance is not preserved by complementation, edge rotation distance is.

3.4.1 Lemma

Let $G_1, G_2 \in \Gamma(p,q)$, then allowing multigraphs in the intermediate steps in transforming G_1 into G_2 via edge rotations does not affect the edge rotation distance.

Proof

Assume we have a minimum edge rotation sequence $t_1, t_2, ..., t_n$ where $t_n....t_2t_1G_1 \cong G_2$ and $t_it_{i-1}...t_2t_1G_1$ is a multigraph for some $i(1 \leq i \leq n-1)$ and where i is a maximum; i.e., we assume that a multigraph is formed as late as possible in the sequence. Assume that the edge rotation $t_i = (x, v, y)$ results in there being two edges between vertices x and y where $x, v, y \in V(t_{i-1}...t_2t_1G_1)$, $vx \in E(t_{i-1}...t_2t_1G_1)$ and $xy \in E(\overline{t_{i-1}...t_2t_1G_1})$. Now since G_2 is not a multigraph one of these two edges xy must be rotated to a new position by the edge rotation $t_j = (y, x, w)$ say, where $n \geq j > i$. Now $yw \in E(\overline{t_i...t_2t_1G_1})$ otherwise we would eventually rotate the same edge three times contradicting the minimality of the sequence $t_1t_2...t_n$. But consider the edge rotation sequence $t_1, t_2, ..., t_{i-1}, t_j, t_i, t_{j+1}, ..., t_n$. This edge rotation sequence transforms G_1 into G_2, however the formation of a multigraph is delayed, contradicting the maximality of i. It is clear that by repeating the process above we can obtain a sequence of n edge rotations which do not involve multigraphs and the result is true. □

Consider $G_1, G_2 \in \Gamma_c(p, q)$, then it is immediately obvious that $d_{es}(2G_1, 2G_2) = 2d_{es}(G_1, G_2)$. However, quite surprisingly, the analogous result for the edge move and edge rotation distances does not hold.

3.4.2 Theorem

There exist graphs $G_1, G_2 \in \Gamma(p, q)$ such that

i) $d_{em}(2G_1, 2G_2) < 2d_{em}(G_1, G_2)$

ii) $d_{er}(2G_1, 2G_2) < 2d_{er}(G_1, G_2)$.

Proof

i) Let F be a 4-connected graph with three distinguishable vertices x, y and z say. Let G_1 be the disjoint union of four copies of F, namely F_1, F_2, F_3, F_4 say, together with the four edges connecting $F_1(x)$ to $F_2(x)$, $F_3(x)$ to $F_4(x)$, $F_1(y)$ to $F_3(y)$ and $F_2(y)$ to $F_4(y)$ (where for example, $F_1(x)$ denotes the distinguishable vertex $x \in V(G_1)$ in F_1, see Figure 3.4.3). Let G_2 be the disjoint union of four copies of F, namely F'_1, F'_2, F'_3 and F'_4 say, together with the four edges connecting $F'_1(x)$ to $F'_2(x)$, $F'_3(y)$ to $F'_4(y)$, $F'_1(z)$ to $F'_3(z)$ and $F'_2(z)$ to $F'_4(z)$.

It is obvious that no two edge moves will transform G_1 into G_2, hence $d_{em}(G_1, G_2) \geq 3$. Define the edge moves $t_1 = (F_1(y), F_3(y), F_1(z), F_3(z))$, $t_2 = (F_2(y), F_4(y), F_2(z), F_4(z))$ and $t_3 = (F_3(x), F_4(x), F_3(y), F_4(y))$. Then

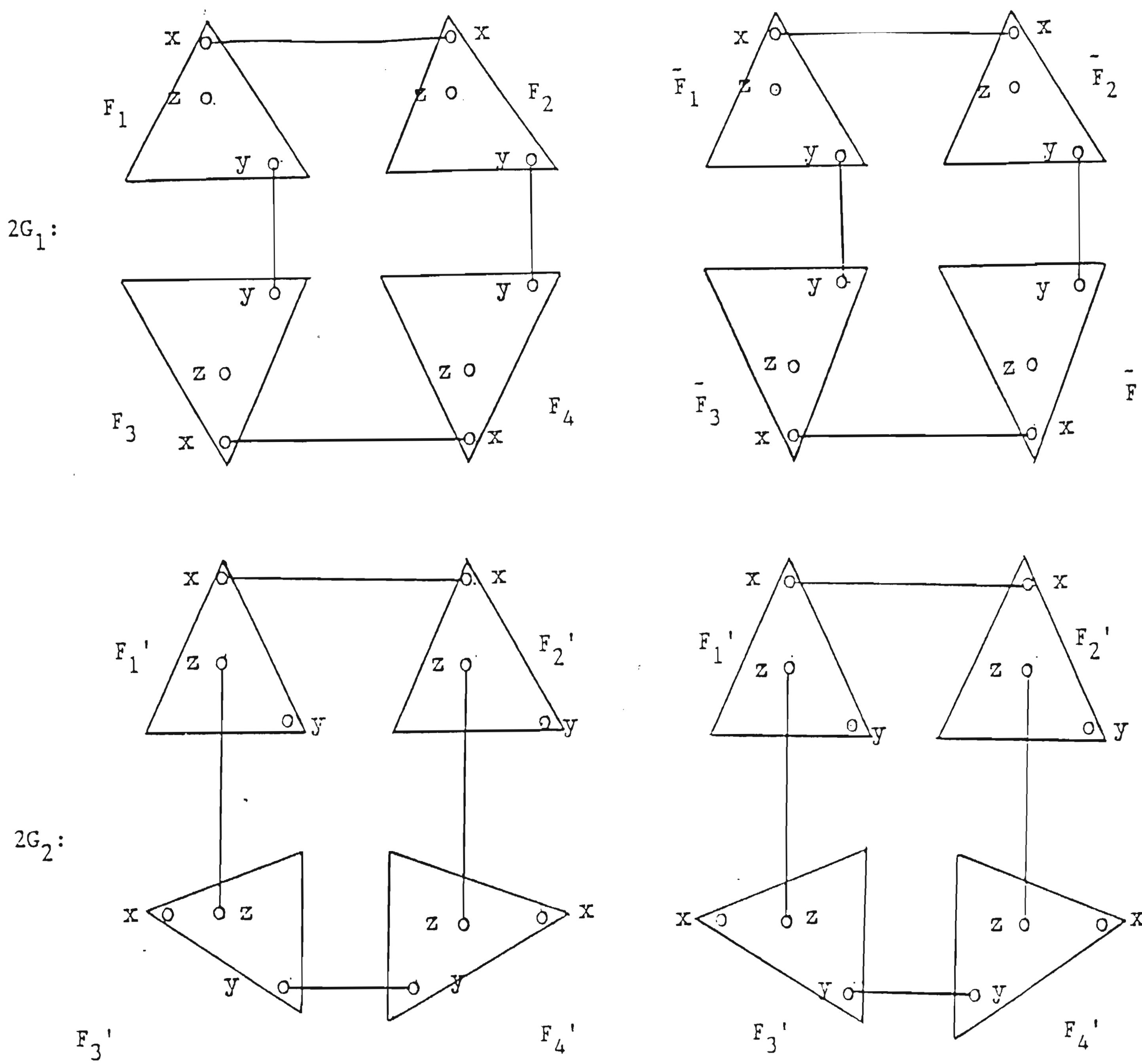
$$t_3t_2t_1G_1 \cong G_2 \text{ and hence } d_{em}(G_1, G_2) = 3.$$

Label the graph $2G_1$ as shown in Figure 3.4.3. Define the edge moves $t'_1 = (F_1(y), F_3(y), F_1(z), \tilde{F}_3(z))$, $t'_2 = (\tilde{F}_1(x), \tilde{F}_2(x), F_2(z), \tilde{F}_1(z))$, $t'_3 = (\tilde{F}_3(x), \tilde{F}_4(x), \tilde{F}_2(z), F_3(z)$ and $t'_4 = (F_2(y), F_4(y), F_4(z), \tilde{F}_4(z))$, then $t'_4t'_3t'_2t'_1G_1 \cong 2G_2$ and hence

$$d_{em}(2G_1, 2G_2) \leq 4 < 2d_{em}(G_1, G_2) = 6.$$

ii) Define the edge rotation $t_1 = (F_4(x), F_3(x), F_2(z))$, $t_2 = (F_2(z), F_4(x), F_4(z))$, $t_3 = (F_4(y), F_2(y), F_3(y))$, $t_4 = (F_1(y), F_3(y), F_3(z))$ and $t_5 = (F_3(z), F_1(y), F_1(z))$, then $t_5t_4t_3t_2t_1G_1 \cong G_2$, hence $d_{er}(G_1, G_2) \leq 5$. That $d_{er}(G_1, G_2) \geq 5$ follows from the fact that there exist no four edge rotations which transform G_1 into G_2. Hence $d_{er}(G_1, G_2) = 5$. Now by Lemma 3.3.1,

$$d_{er}(2G_1, 2G_2) \leq 2d_{em}(2G_1, 2G_2) \leq 8 < 10 = 2d_{er}(G_1, G_2). \qquad \square$$



3.4.3 Figure

3.4.4 Definition

For any graph G, the subdivision graph of G denoted by $S(G)$, is the graph obtained from G by replacing each subpath uv of length one in G by a path of length two, having u and v as end-vertices. Hence the order of $S(G)$ exceeds that of G by $q(G)$

. We now look at the effect of the subdivision operation on the edge move and edge rotation distances.

3.4.5 Theorem

For all graphs $G_1, G_2 \in \Gamma(p,q)$, $d_{er}(S(G_1), S(G_2)) \leq d_{er}(G_1, G_2)$.

Proof

For $x, y, z \in V(G_1)$, let $t = (x, y, z)$ be any edge rotation in an edge rotation sequence which transforms G_1 into G_2. It is sufficient to prove that for each such edge rotation t there is a corresponding edge rotation t' which deletes the subdivided edge xy and creates a subdivided edge xz in the transformation of $S(G_1)$ into $S(G_2)$. Suppose e is the vertex subdividing the edge xy then the edge rotation $t' = (e, y, z)$ does what is required. $\square$

3.4.6 Corollary

For all graphs $G_1, G_2 \in \Gamma(p,q)$

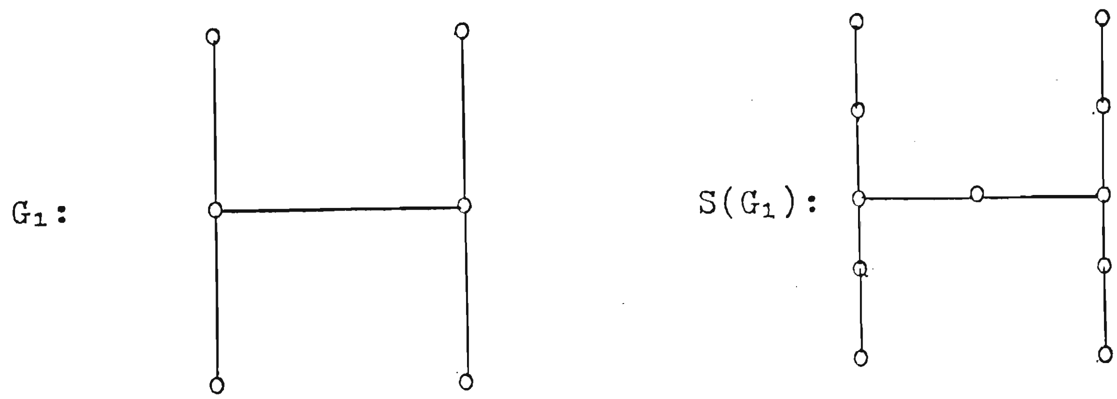
$$d_{em}(S(G_1), S(G_2)) \leq 2d_{em}(G_1, G_2).$$

Proof

Evidently $d_{em}(S(G_1), S(G_2)) \leq d_{er}(S(G_1), S(G_2))$, while from Theorem 3.4.5, $d_{er}(S(G_1), S(G_2)) \leq d_{er}(G_1, G_2)$ and from Lemma 3.3.1, $d_{er}(G_1, G_2) \leq 2d_{er}(G_1, G_2)$. Therefore $d_{em}(S(G_1), S(G_2)) \leq 2d_{em}(G_1, G_2)$. $\square$

3.4.7 Example

Let G_1 be the graph shown in Figure 3.4.8 and let $G_2 = P_6$.



3.4.8 Figure

Obviously $d_{em}(G_1, G_2) = 1$ while $d_{em}(S(G_1), S(G_2)) = 2$. This example shows that the upper bound in Corollary 3.4.6 is sharp since $d_{em}(S(G_1), S(G_2)) = 2d_{em}(G_1, G_2)$ in this case.

3.5 Ordering of Metrics

In [J1] Johnson presents a means of partially ordering some of the metrics that we have studied so far; namely the induced subgraph metric d_i, the edge rotation distance metric d_{er}, the edge slide distance metric d_{es}, the subgraph metric d_s, and the discrete metric d_d. Throughout the remainder of this section the results obtained are essentially from [J1] with the following exceptions: Lemma 3.5.20 and Theorem 3.5.21 are new, while Theorems 3.5.15 and 3.5.22 have been modified to include the edge move distance in the ordering.

3.5.1 Definition

A metric $d : W \times W \to \mathrm{N} \cup \{0\}$ will be called an *integer metric* with *unit* λ if $\lambda = \min\{d(w, w') : w, w' \in W \text{ and } w \neq w'\}$. If an integer metric is defined on a singleton set then we say that it has unit λ for any $\lambda \in \mathrm{N}$. Now for any integer metric d defined on W we associate with it the graph $M(d)$ which has W as its vertex set, and for $w, w' \in W$, $ww' \in E(M(d))$ if and only if $d(w, w') = \lambda$.

3.5.2 Remark

For the edge rotation and edge slide distance metrics with unit $\lambda = 1$ we note that the graphs $M(d_{er})$ and $M(d_{es})$ are the edge rotation distance and edge slide distance graphs defined in Sections 2.3 and 2.2 respectively.

3.5.3 Definition

Let d and d' be distinct integer metrics defined on W; then if $M(d)$ is a subgraph of $M(d')$, we say that d *expands* d', denoted by $d \geq d'$. Since the subgraph relation is a partial order, this expansion relation is also a partial order. We shall say that d *strictly expands* d', denoted by $d > d'$, if d expands d', but not vice versa.

Thus to obtain a partial ordering of the set $D = \{d_{es}, d_{er}, d_{em}, d_s, d_i, d_d\}$ it will be necessary to restrict ourselves to $\Gamma_c(p, q)$, or if we wish to obtain a partial ordering of a proper subset of D, we will accordingly work with the most restricted domain of this subset.

We recall that the discrete metric $d_d : \Gamma \times \Gamma \to \{0,1\}$ is defined by $d_d(G_1, G_2) = 0$ if $G_1 \cong G_2$ and $d_d(G_1, G_2) = 1$, otherwise.

3.5.4 Lemma

Let d and d_d be integer metrics defined on W such that d has unit λ and d_d is the discrete metric. Then $d \geq d_d$ and if $d(w, w') > \lambda$ for any $w, w' \in W$, then $d > d_d$.

Proof

The graph $M(d_d)$ associated with d_d is the complete graph, since by definition $w, w' \in W$ and $w \neq w'$ implies that $d_d(w, w') = 1$. Since all graphs with vertex set W are subgraphs of the complete graph with vertex set W, it follows that $M(d)$ is a subgraph of $M(d_d)$ and $d \geq d_d$. If $d(w, w') > \lambda$ for any $w, w' \in W$, then $ww' \in \overline{E(M(d))}$ and thus $M(d)$ is not complete; so $M(d_d)$ is not a subgraph of $M(d)$ and $d > d_d$. □

3.5.5 Lemma

Let d and d' be integer metrics defined on W with units λ and λ' respectively. If for $w, w' \in W, d(w, w') = \lambda$ implies that $d'(w, w') \leq \lambda^*$ where λ^* is an integer and $\lambda' \geq \lambda^*$, then $\lambda' = \lambda^*$ and $d \geq d'$.

Proof

If W is a singleton set, the proposition is true by setting $\lambda^* = \lambda'$. Suppose then that W is not a singleton set. Then there exist $w, w' \in W$ such that $d(w, w') = \lambda$. This implies $d'(w, w') \leq \lambda^*$, and by definiton of λ', that $d'(w, w') \geq \lambda'$, hence $\lambda' \leq \lambda^*$. However, $\lambda' \geq \lambda^*$; therefore $\lambda' = \lambda^*$. Hence $d(w, w') = \lambda$ implies $d'(w, w') = \lambda'$; thus $M(d)$ is a subgraph of $M(d')$ and $d \geq d'$. □

We note that if d and d' are integer metrics defined on W with units λ and λ' respectively, then

i) If W is a singleton set then $d \geq d'$ and $d' \geq d$.

ii) If W contains two distinct elements then $d \geq d'$ if and only if $\lambda \leq \lambda'$.

3.5.6 Lemma

Let d and d' be integer metrics defined on W with units λ and λ' respectively. Let $W' \subset W$. If $d \geq d'$ and if $d|W'$ (the restriction of d to W') has unit λ then $d'|W'$ has unit λ' and $d|W' \geq d'|W'$.

Proof

Let $w, w' \in W$. Then since $d \geq d'$, $ww' \in E(M(d))$ implies $ww' \in E(M(d'))$. Let $v, v' \in W'$ such that $d(v, v') = \lambda$; however this implies that $d'(v, v') = \lambda'$. Hence $vv' \in E(M(d|W'))$ implies $vv' \in E(M(d'|W')$, and thus $M(d|W')$ is a subgraph of $M(d'|W')$ and $d|W' \geq d'|W'$. □

3.5.7 Definition

For any metric d defined on W, we say that d is *connected* if $M(d)$ is a connected graph. If $M(d)$ is connected then for any $w, w' \in W$ there exists a shortest path connecting w and w' in $M(d)$, the length of which we denote by $\delta(w, w')$. The function $\delta : W \times W \to \mathbf{N} \cup \{0\}$ associated with d is a metric defined on W which we call the *path metric associated* with d.

3.5.8 Example

The metric d defined on $\{1, 2, 3\}$ by $d(1, 2) = 1$, $d(1, 3) = 2$ and $d(2, 3) = 2$ has unit $\lambda = 1$ and $M(d) \cong K_2 \cup K_1$, which is not connected.

3.5.9 Lemma

Let d be a connected integer metric with unit λ defined on W, and let δ be the path metric associated with d. Then for every $w, w' \in W$

$$d(w, w') \leq \lambda\delta(w, w').$$

Proof

We proceed by induction on $\delta(w, w')$. The statement obviously holds if $\delta(w, w') = 1$.

Suppose $w = w_0, w_1, w_2, ..., w_n = w'$ is a path of length $\delta(w, w') = n$ in $M(d)$. Then $d(w_0, w_2) \leq d(w_0, w_1) + d(w_1, w_2) = 2\lambda$ by the triangle inequality. Similarly $d(w_0, w_3) \leq d(w_0, w_2) + d(w_2, w_3) = 2\lambda + \lambda = 3\lambda$. Assume $d(w_0, w_k) \leq k\lambda$ for $3 \leq k \leq n - 1$ then

$$d(w_0, w_{k+1}) \leq d(w_0, w_k) + d(w_k, w_{k+1}) = k\lambda + \lambda = (k + 1)\lambda.$$

Therefore

$$d(w_0, w_n) = d(w, w') = \lambda n = \lambda\delta(w, w') \text{ as required.} \qquad \square$$

The following lemma will be useful in proving that the metrics we have studied, subject to their various restrictions, are connected.

3.5.10 Lemma

Let d and d' be any two integer metrics defined on W. If $d \geq d'$ and if d is connected, then d' is connected.

Proof

Since $d \geq d'$, $M(d)$ is a connected subgraph of $M(d')$ and since $M(d)$ and $M(d')$ have the same vertex set, the result follows. $\square$

3.5.11 Definition

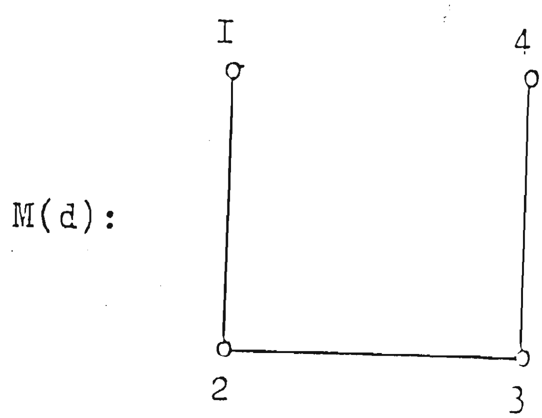
Let d be any connected integer metric on W with associated path metric δ. If $d(w, w') = \lambda\delta(w, w')$ for all $w, w' \in W$ then d will be said to be *graphable*. Note that δ is always graphable with unit 1.

Note that if we define $d' : W \times W \to \mathrm{N} \cup \{0\}$ by $d'(w, w') = \frac{d(w,w')}{\lambda}$, we see that any graphable metric with unit λ is equivalent to a graphable metric with unit 1. However, if d is not graphable, d' may not be an integer metric. From now on we assume, unless stated otherwise, that all metrics have unit 1.

The following example shows that a metric can be connected, but not graphable.

3.5.12 Example

Define the metric d on $\{1,2,3,4\}$ by $d(i,i+1) = 1$ for $i = 1,2,3$, $d(i,i+2) = 2$ for $i = 1,2$, and $d(1,4) = 2$. See Figure 3.5.13.



3.5.13 Figure

To see that $M(d)$ misrepresents d in the sense that d is not graphable, note that $\delta(1,4) = 3$ and $d(1,4) = 2 \neq 3 = \delta(1,4)$. The metric d can be changed into a graphable metric by redefining $d(1,4) = 3$.

3.5.14 Lemma

Let d and d' be integer metrics defined on W with units λ and λ' respectively. Let d be graphable. If $d \geq d'$, then $d(w,w') \geq \left(\frac{\lambda}{\lambda'}\right) d'(w,w')$ for all $w, w' \in W$.

Proof

By assumption, d is graphable and hence

$$\frac{d(w,w')}{\lambda} = \delta(w,w'). \qquad \sim (1)$$

Lemma 3.5.10 and $d \geq d'$ together imply that d' is connected. Thus δ' is a well defined metric. $M(d)$ is a subgraph of $M(d')$ therefore $\delta(w,w') \geq \delta'(w,w')$, and hence from (1)

$$\frac{d(w,w')}{\lambda} \geq \delta(w,w'). \qquad \sim (2)$$

From Lemma 3.5.9

$$\delta'(w,w') \geq \frac{d'(w,w')}{\lambda'}. \qquad \sim (3)$$

The equations (2) and (3) together imply that $\frac{d(w,w')}{\lambda} \geq \frac{d(w,w')}{\lambda'}$,

i.e. $d(w,w') \geq \left(\frac{\lambda}{\lambda'}\right) d'(w,w')$ as required. □

It is now possible to start developing the partial ordering of the metrics d_i, d_s, d_{em}, d_{er} and d_{es}.

3.5.15 Theorem

The integer metrics, d_i, d_s, d_{em}, d_{er} and d_{es} all strictly expand d_d on their respective domains.

Proof

These are just special cases of lemma 3.5.4. □

3.5.16 Theorem

The restriction of d_s to $\Gamma(p)$ expands the restriction to $\Gamma(p)$ of d_i for any $p \in \mathbf{Z}^+$.

Proof

Let e be any edge of K_p. Then $d_s(K_p, K_p - e) = 1$, Thus $d_s|\Gamma(p)$ has unit 1. Let $M(d_s|\Gamma(p))$ be the graph associated with the metric d_s restricted to $\Gamma(p)$.

Let GH be any edge in $M(d_s|\Gamma(p))$; then $d_s(G, H) = 1$. Therefore either G is a proper subgraph of H or vice versa. Without loss of generality assume G is a proper subgraph of H, then either $V(G) = V(H)$ and $|E(H) - E(G)| = 1$ or $E(G) = E(H)$ and $|V(H) - V(G)| = 1$.

Case 1) Assume $V(G) = V(H)$ and $|E(H) - E(G)| = 1$. Let $uv \in E(H) - E(G)$. Then $G - u$ is an induced subgraph of both G and H, therefore $d_i(G, H) = 1$.

Case 2) Assume $E(G) = E(H)$ and $|V(H) - V(G)| = 1$. Then G is an induced subgraph of both G and H and therefore $d_i(G, H) = 1$.

In both cases we have $d_i(G, H) = 1$; hence $d_s(G, H) = 1$ implies $d_i(G, H) = 1$ and therefore $GH \in E(M(d_i))$. Hence $M(d_s(\Gamma(p))$ is a subgraph of $M(d_i)$ and consequently $d_s|\Gamma(p)$ expands $d_i|\Gamma(p)$; i.e., $d_s|\Gamma(p) \geq d_i|\Gamma(p)$.

To show $d_s|\Gamma(p) > d_i|\Gamma(p)$ for $p \geq 3$ we note that for $p \geq 3$, there exist $u, v, w \in V(K_p)$. Let $G = K_p$ and $H = K_p - uv - uw$. Then $d_s(G, H) = 2 > 1 = d_i(G, H)$ and therefore d_i does not expand d_s since $GH \in E(M(d_i))$ but $GH \in E(\overline{M(d_s)\Gamma(p)})$ and hence $d_s > d_i$ for $p \geq 3$. □

3.5.17 Lemma

The metric d_s restricted to $\Gamma(p, q)$ has unit $\lambda \geq 2$.

Proof

Suppose $G, H \in \Gamma(p, q)$ such that $d_s(G, H) = 1$. Then G is a proper subgraph of H or vice versa, and hence either $G \notin \Gamma(p, q)$ or $H \notin \Gamma(p, q)$. The contradiction establishes the lemma. □

3.5.18 Theorem

The edge rotation metric d_{er} on $\Gamma(p, q)$ expands the restriction of d_s to $\Gamma(p, q)$, and there exist integers p and q such that $d_{er}|\Gamma(p, q) > d_s|\Gamma(p, q)$.

Proof

We establish the conditions of Lemma 3.5.5. Let $d_s|\Gamma(p, q)$ have unit λ'; then, by Lemma 3.5.17, $\lambda' \geq 2$.

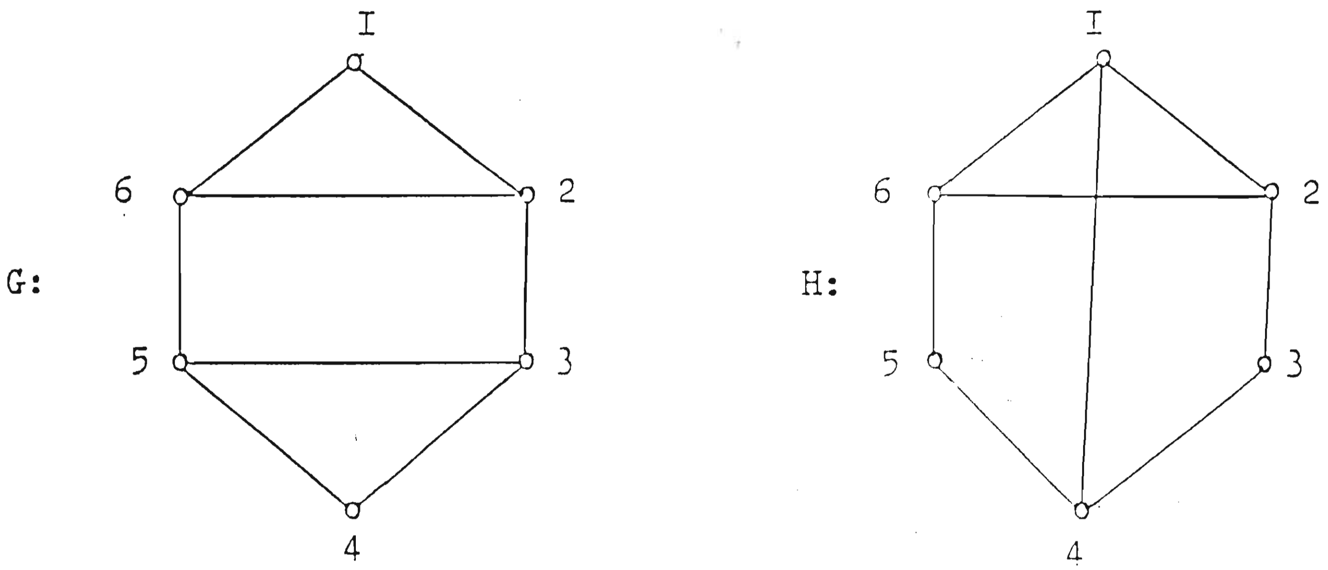
Let $G, H \in \Gamma(p, q)$ such that GH is an edge of $M(d_{er})$; i.e., $d_{er}(G, H) = 1$. By the definition of the edge rotation distance we may assume without loss of generality that $H \cong G - uv + vw$ where $u, v, w \in V(G), uv \in E(G)$ and $vw \in E(\bar{G})$.

Since $H \cong G - uv + vw$ it follows that $uv \in E(\bar{H})$ while $vw \in E(H)$. Hence $G - uv$ is a subgraph of both G and H, and, therefore $d_s(G, H) \leq 2$.

Thus we have that $d_{er}(G,H) = 1$ implies that $d_s(G,H) \leq 2$ where the unit of $d_s|\Gamma(p,q)$ is $\lambda' \geq 2$. Applying Lemma 3.3.5 with $\lambda^* = 2$, we obtain

$$d_{er} \geq d_s|\Gamma(p,q).$$

To show that there exists p and q such that $d_{er} > d_s|\Gamma(p,q)$ consider the graphs G and H shown in Figure 3.5.19.



3.5.19 Figure

Any single edge rotation of any edge on the 6-cycle $C_6 = 1234561$ in G, will not transform G into H since it will produce a graph with no 6-cycle. A single edge rotation t involving either the 26 or 35 edge in G eliminates the existence of a 4-cycle in tG unless a vertex of degree four is formed, and H contains no vertex of degree four. Thus no single edge rotation will transform G into H, hence $d_{er}(G, H) \geq 2$, and therefore

$$GH \notin E(M(d_{er})). \qquad \sim (1)$$

Since $G - 35$ is a subgraph of G and H, we have

$$d_s(G, H) \leq p + q + p + q - 2(p + q - 1) = 2$$

and since $G \ncong H$ and since $\lambda' \geq 2$ we have that $d_s(G, H) = \lambda' = 2$. Hence

$$GH \in E(M(d_s|\Gamma(6, 8)). \qquad \sim (2)$$

Together (1) and (2) imply that $M(d_{er}|\Gamma(6, 8))$ is not a subgraph of $M(d_s)|\Gamma(6, 8)$ and hence $d_s < d_{er}$, in this case. □

We will now show that the subgraph metric d_s restricted to $\Gamma(p, q)$ is equivalent to the edge move distance d_{em} (which is defined on $\Gamma(p, q)$) in the sense that $M(d_s|\Gamma(p, q)) \cong M(d_{em})$ for all p and q; i.e., $d_s|\Gamma(p, q) \geq d_{em}$ and $d_{em} \geq d_s|\Gamma(p, q)$, we denote this equivalence by $d_{em} \sim d_s$. Consequently the fact that $d_{em} \geq d_i|\Gamma(p, q)$ and $d_{er} \geq d_{em}$ will need no further justification.

The following lemma will aid us in establishing this equivalence.

3.5.20 Lemma

Let $G, H \in \Gamma(p, q)$ then $d_s(G, H) = 2d_{em}(G, H)$.

Proof

Recall that $d_s(G,H) = \min\{|G| + |H| - 2|F|\}$ taken over all graphs F which are isomorphic to subgraphs of both G and H. Suppose F^* is a graph which minimises the expression above. Obviously $|V(F^*)| = p$ and suppose $|E(F^*)| = s$, then

$$d_s(G,H) = 2p + 2q - 2(p+s) = 2(q-s).$$

Now obviously deleting any isolated vertices from F^* yields a greatest common subgraph of G and H which contains no isolated vertex and hence by Theorem 3.2.7, $d_{em}(G,H) = q - s$ and the result follows. □

3.5.21 Theorem

The graph $M(d_s|\Gamma(p,q)$ is isomorphic to the graph $M(d_{em})$; i.e., $d_{em} \sim d_s|\Gamma(p,q)$.

Proof

Let $G, H \in \Gamma(p,q)$ such that $d_{em}(G,H) = 1$; i.e., $GH \in E(M(d_{em}))$. Lemmas 3.5.17 and 3.5.20 together imply that $d_s(G,H) = 2$ and $d_s|\Gamma(p,q)$ has unit $\lambda = 2$. Hence $GH \in E(M(d_s|\Gamma(p,q)))$ and $M(d_{em})$ is a subgraph of $M(d_s|\Gamma(p,q))$.

Conversely suppose $d_s|\Gamma(p,q)$ has unit $\lambda \geq 2$ and that $G, H \in \Gamma(p,q)$ such that $d_s(G,H) = \lambda$, i.e. $GH \in E(M(d_s|\Gamma(p,q)))$. By Lemma 3.5.20, d_{em} has unit $\frac{\lambda}{2}$, and $d_{em}(G,H) = \frac{\lambda}{2}$, hence $GH \in E(M(d_{em}))$. Thus $M(d_s|\Gamma(p,q))$ is a subgraph of $M(d_{er})$.

It now follows that $M(d_s|\Gamma(p,q)) \cong M(d_{em})$. □

3.5.22 Theorem

If we restrict the metrics $d_d, d_i, d_s, d_{em}, d_{er}$ and d_{es} to $\Gamma_c(p,q)$ then $d_d \leq d_i \leq d_s \sim d_{em} \leq d_{er} \leq d_{es}$. Moreover, there exist integers p and q such that $d_d < d_i < d_s \sim d_{em} < d_{er} < d_{es}$.

Proof

First we show that the expansion relation is transitive. Suppose d, d' and d'' are distinct integer metrics such that $d \geq d'$ and $d' \geq d''$. Then $M(d)$ is a subgraph of $M(d')$ and $M(d')$ is a subgraph of $M(d'')$. It is thus clear that $M(d)$ must be a subgraph of $M(d'')$ and therefore $d \geq d''$. Thus we have $d \geq d' \geq d''$.

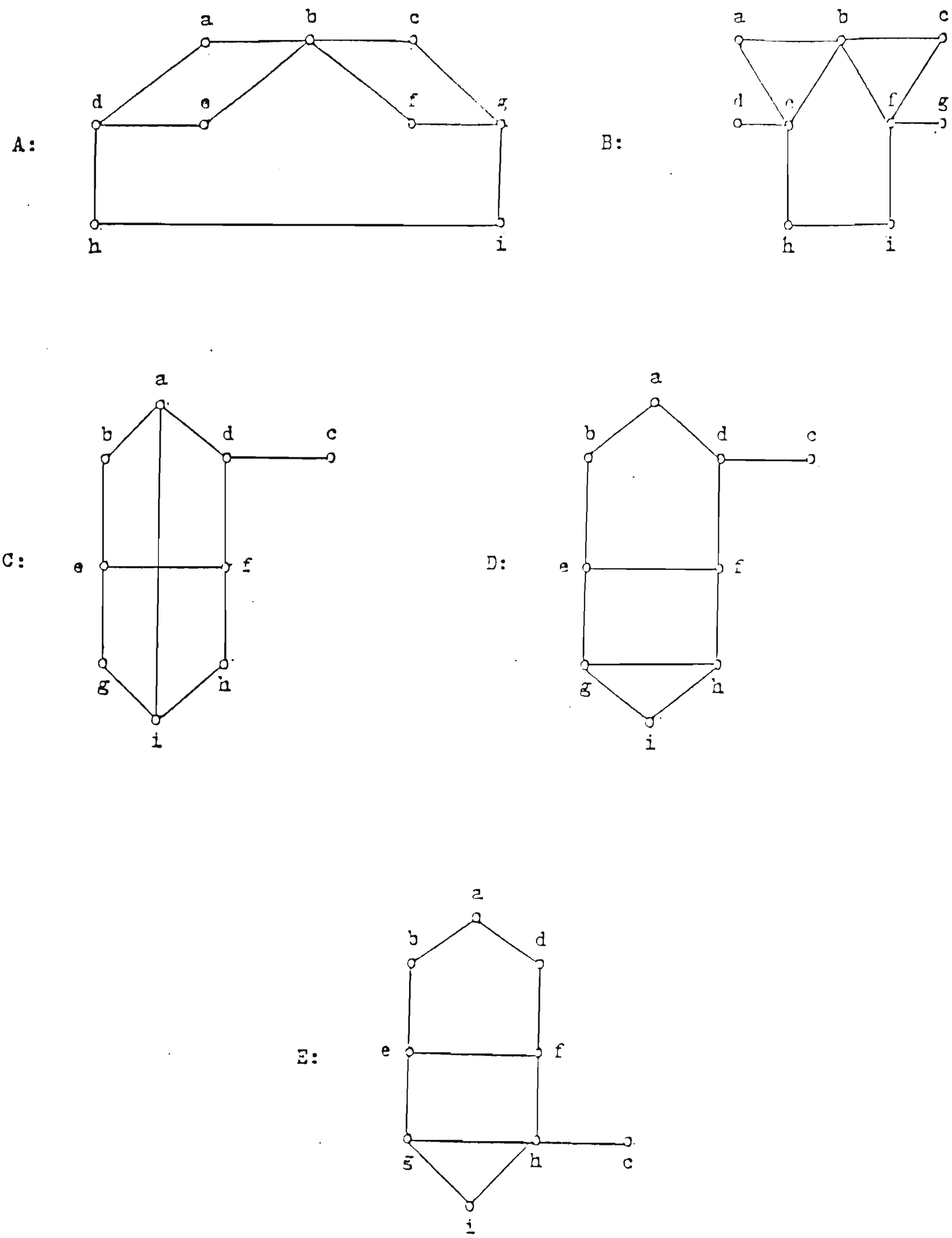
From Lemma 3.5.5 and Theorems 3.5.15, 3.5.16, 3.5.18 and 3.5.21, together with the transitivity of the expansion relation, we have that

$$d_d \leq d_i \leq d_s \sim d_{em} \leq d_{er}$$

where we will assume that all metrics are restricted to $\Gamma_c(p,q)$. Therefore to establish the first statement of the theorem we need only show that $d_{es} \geq d_{er}$. However this is trivial since the edge slide is just a special case of the edge rotation, and therefore $d_{es}(G,H) = 1$ implies $d_{er}(G,H) = 1$. Hence $M(d_{es}|\Gamma_c(p,q))$ is a subgraph of $M(d_{er}|\Gamma_c(p,q))$, and we have

$$d_d \leq d_i \leq d_s \sim d_{em} \leq d_{er} \leq d_{es}.$$

To establish the strict expansion relation for some (p,q), we consider the graphs in Figure 3.5.23.



3.5.23 Figure

First note that $C - v$ contains a 5-cycle for all $v \in V(C)$. Since A does not contain a 5-cycle, we have that a maximum induced subgraph (with respect to order) of both A and C has order less than 8. But $d_i(A, C) = n$ where by definition n is such that $p(A) - n$ is the maximum order of an induced subgraph of both A and C. Therefore $p(A) - n = 9 - n < 8$ and hence $n > 1$. Thus $d_i(A, C) > 1$. Thus it follows that $M(d_i|\Gamma_c(9, 11)$ is not complete and therefore a subgraph of $M(d_d|\Gamma_c(9, 11))$,hence

$$d_i|\Gamma_c(9, 11) > d_d\Gamma_c(9, 11). \qquad \sim (1)$$

Now it is clear that $A - b \cong B - b$. Thus $d_i(A, B) = 1$. However $B - e'$ contains a 3-cycle for all $e' \in E(B)$, while A contains no 3-cycles. Therefore $d_s(A, B) > (11 + 9) + (11 + 9) - 2(9 + 10) = 2$. Since the unit of $d_s|\Gamma_c(9, 11)$ is $\lambda = 2$ we have that $AB \in E(M(d_i|\Gamma_c(9, 11)))$ while $AB \notin E(M(d_s|\Gamma_c(9, 11)))$ and thus $M(d_i|\Gamma_c(9, 11))$ is not a subgraph of $M(d_s|\Gamma_c(9, 11))$, this together with Theorem 3.5.21 yields

$$d_{em}|\Gamma_c(9, 11) \sim d_s|\Gamma_c(9, 11) > d_i|\Gamma_c(9, 11). \qquad \sim (2)$$

Since $C - ai \cong D - gh$ we have that $d_s(C, D) = \lambda = 2$. However there exists no single edge rotation that will transform C into D. This is because a rotation of any edge on the 8-cycle in C, creates a graph in which there is no 8-cycle, and D has an 8-cycle . Also rotating any one of the edges dc, ef or ai so as to form a 3-cycle, as in D, will either create a graph with a vertex of degree 4 or a graph with no end-vertex. Therefore $d_{er}(C, D) > 1$. It follows that

$$d_{er}|\Gamma_c(9, 11) > d_s|\Gamma_c(9, 11) \sim d_{em}|\Gamma_c(9, 11) \qquad \sim (3)$$

Finally let $t = (c, d, h)$ be an edge rotation. Then $tD \cong E$. Thus $d_{er}(D, E) = 1$. To show $d_{es}(D, E) > 1$, note that there exist no edge slides of the form $t_1 = (g, i, v)$ or $t_2 = (h, i, v)$ where $v \in V(D)$, since these create multigraphs. Any other edge slide of any edge lying on the 8-cycle of D will create either a graph in which there exists no 8-cycle or a graph with no end-vertex, and E has both an 8-cycle and an end-vertex. The edges gh, ef and dc of D cannot be slid to form a 3-cycle with one of the vertices on the cycle adjacent to an end-vertex, and D has such a vertex. It follows that $d_{es}(D, E) > 1$, and that

$$d_{es}|\Gamma_c(9, 11) > d_{er}|\Gamma_c(9, 11). \qquad \sim (4)$$

Together (1), (2), (3) and (4) imply that there exist integers p and q such that

$$d_{es}|\Gamma_c(p, q) > d_{er}|\Gamma_c(p, q) > d_s|\Gamma_c(p, q) \sim d_{em}|\Gamma_c(p, q) > d_i|\Gamma_c(p, q) > d_d|\Gamma_c(p, q).$$

□

Chapter 4

The Steiner Problem in Graphs

4.1 Introduction

The original Steiner problem is easy to state: In a Euclidean space (usually a Euclidean plane) draw the shortest possible network of line segments interconnecting, say, 100 given points. However this problem is unsolvable in many cases. The practical importance to designers of telephone networks, for example, is obvious, and has led to the development of algorithms that yield rough solutions quickly.

The Steiner problem generally cannot be solved by simply drawing lines between the given points, but it can be solved by adding new points, called Steiner points, that serve as junctions in a shortest network. To determine the location and number of Steiner points, mathematicians and computer scientists have developed algorithms or precise procedures. However even the best of these algorithms running on the fastest computers cannot provide a solution for a large set of given points because the time it would take to solve such a problem is impractically long. Furthermore, the

Steiner problem belongs to the class of problems known as NP-Hard problems, for which many computer scientists now believe an efficient algorithm may never exist. Thus the importance of approximate solutions becomes apparent. Such approximate solutions are used routinely in designing integrated circuits, determining the evolutionary trees of groups of organisms and minimizing materials used for networks of telephone lines, pipelines and roadways.

The Steiner problem, in its general form, first appeared in a paper by Miloš, Kössler and Vojtěch Jarník in 1934, but the problem did not become popular until 1941 when Richard Courant and Herbert E. Robbins [CR1] included it in their book *What Is Mathematics?* Courant and Robbins link this problem to the work of Jacob Steiner, the famous geometer at the University of Berlin in the early nineteenth century. Steiner worked on the following problem. Three villages A, B, C are to be joined by a system of roads of minimum total length. Mathematically, three points A, B, C are given in a plane, and a fourth point P in the plane is sought so that the sum $a + b + c$ is a minimum, where a, b and c denote the three distances from P to A, B and C respectively. Evangelista Torricelli (1608- 1647) and Francesco Cavalieri (1598 - 1647) solved the problem independently. They deduced that if in a triangle ABC all angles are less than $120°$ then P is the point from which each of the three sides, AB, BC and CA subtends an angle of $120°$. If, however, an angle of triangle ABC, say the angle at C, is greater than or equal to $120°$, then the point P coincides with the vertex C. Torricelli and Cavalieri also developed geometric constructions for finding P. (See pages 356- 358 of [CR1].)

It is natural to generalize the problem to the case of n given points,

$A_1, A_2, \ldots, A_n$; we need to find the point P in the plane for which the sum of the distance $a_1 + a_2 + \ldots + a_n$ is a minimum, where a_i is the distance from point P to point A_i. This problem, which was also treated by Steiner, does not lead to interesting results. To find a significant extension of Steiner's problem we abandon the search for a single point P. The extension we are looking for is expressed by Courant and Robbins as follows: Given n points $A_1, A_2, \ldots, A_n$ we seek a connected system of straight line segments of shortest total length such that any two of the given points can be joined by a polygon consisting of segments of the system. This problem is called The Steiner problem.

A similar problem to the Steiner problem, which was proposed by Z.A. Melzak in [M1], will lead us to the extension of the Steiner problem which we shall study in some detail. Melzak proposed the problem of connecting n given points in the plane by line segments between these n points, so that the sum of these distances is a minimum. Extending this problem a bit further, to include graphs, we will arrive at the problem which we will call the Steiner problem in graphs. This is the problem which we shall study in detail in this chapter.

We shall define the Steiner distance of a set of vertices in a connected graph G (which is a generalization of the well-known concept of distance) and then investigate properties of the Steiner distance and its related structures. Since simplifications occur if the graph G is a tree and the related results differ significantly from those pertaining to graphs which contain cycles, we shall deal with trees in Section 4.3 and consider more general graphs in Section 4.4. Representative techniques of calculating Steiner distances (precisely or approximately) are provided in Section 4.5.

4.2 The Steiner Problem in Graphs

The distance between two vertices x and y in a given graph G may be defined as the minimum size among all connected subgraphs of G whose vertex sets contain x and y. It is clear that every such subgraph of G would be a shortest path between x and y, as demanded by the standard definition of distance. This leads naturally to a generalization of distance, where we may consider a distance among a set of two, three or more vertices (see [COTZ1]).

4.2.1 Definition

Let G be a connected graph of order at least two and let S be a nonempty set of vertices of G. Then the *Steiner distance $d(S)$* among the vertices of S (or simply the distance of S) is the minimum size among all connected subgraphs of G whose vertex sets contain S.

4.2.2 Note

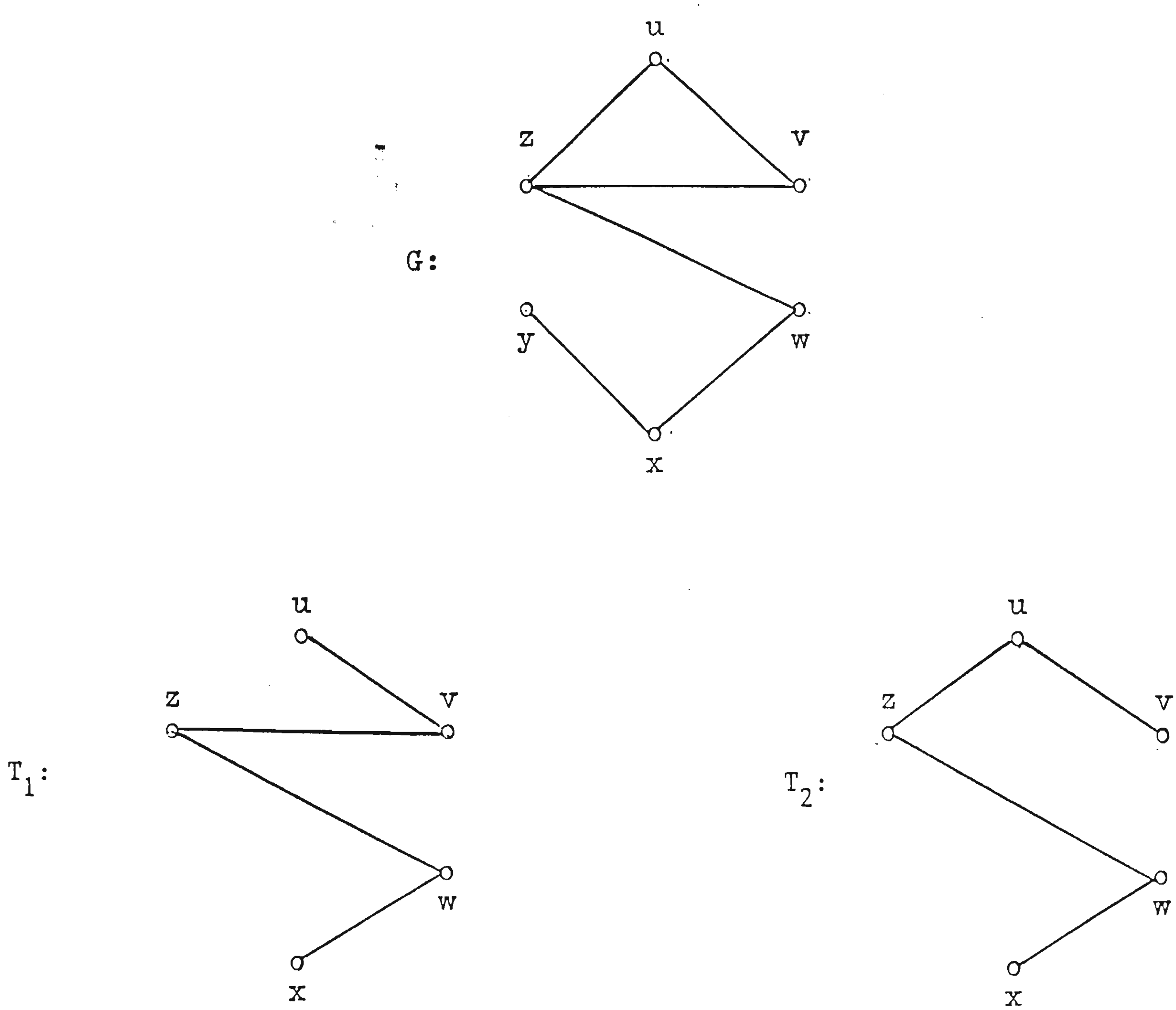
If H is a connected subgraph of G such that $S \subseteq V(H)$ and $|E(H)| = d(S)$, then H must be a tree, since if H contained cycles then removing an edge from one of these cycles would yield a connected graph H' with fewer edges than H, with $V(H') = V(H) \supseteq S$, contradicting the minimality of H. Such a tree is referred to as a Steiner tree for the set S. Further, if $S = \{u, v\}$ where $u, v \in V(G)$, then $d(S) = d(u, v)$, while if $|S| = n$ then $d(S) \geq n - 1$.

4.2.3 Remark

The Steiner problem in graphs is thus to connect a subset S of vertices of a graph G with a tree of minimum size which is a subgraph of G. The difference between this problem and the original Steiner problem is that no new points (vertices) are added here and the network of line segments (edges) are already present in the graph G. The problem is to find which edges of the graph G are to be used in the Steiner tree.

4.2.4 Example

If G is the graph of Figure 4.2.5 and $S = \{u, v, x\}$, then $d(S) = 4$; there are two trees T_1 and T_2 of size 4 containing S, both of which are shown in Figure 4.2.5.



4.2.5 Figure

4.2.6 Example

Let $G_1 \cong K_{n-1,n-1}$, with partite sets $V_1 = \{u_1, u_2, ..., u_{n-1}\}$ and $V_2 = \{v_1, v_2, ..., v_{n-1}\}$ and let S be any set of n vertices of G. Then $\langle S \rangle$ is connected and hence $d(S) = n - 1$.

4.2.7 Remark

The usual distance between pairs of vertices defined in a connected graph G is a metric on its vertex set. Thus for vertices $u, v, w \in V(G)$ the properties (1) $d(u,v) \geq 0$ and $d(u,v) = 0$ if and only if $u = v$, (2) $d(u,v) = d(v,u)$ and (3) $d(u,w) \leq d(u,v) + d(v,w)$ hold.

Chartrand, Oellermann, Tian and Zou [COTZ1] extended properties (1) and (3) to include the Steiner distance. Let G be a connected graph and let $S \subseteq V(G)$, where $S \neq \emptyset$. Then $d(S) \geq 0$, while $d(S) = 0$ if and only if $|S| = 1$, which extends property (1) above. To obtain an extension of (3), let S, S_1 and S_2 be subsets of $V(G)$ such that $\emptyset \neq S \subseteq S_1 \cup S_2$ and $S_1 \cap S_2 \neq \emptyset$. Then $d(S) \leq d(S_1) + d(S_2)$. To see this, let T_i be a tree of size $d(S_i)$ such that $S_i \subseteq V(T_i)$ for $i = 1, 2$. Let H be the graph with vertex set $V(T_1) \cup V(T_2)$ and edge set $E(T_1) \cup E(T_2)$. Now T_1 and T_2 are connected and $V(T_1) \cap V(T_2) \neq \emptyset$, hence H is connected. Since $S \subseteq V(H)$ and since H is connected, $d(S) \leq q(H) \leq d(S_1) + d(S_2)$. (The extension of (2) to Steiner distance is obviously tautologous: $d(S) = d(S)$.)

In [COTZ1] the concepts of eccentricity, radius and diameter were generalized to accommodate the Steiner distance.

4.2.8 Definition

Let G be a connected graph of order $p \geq 2$ and let n be an integer with $2 \leq n \leq p$. The *n-eccentricity* $e_n(v)$ of a vertex $v \in V(G)$ is defined by

$$e_n(v) = \max\{d(S) : S \subseteq V(G), |S| = n, \text{ and } v \in S\}.$$

The n-radius of G is defined by

$$\text{rad}_n G = \min\{e_n(v) : v \in V(G)\}$$

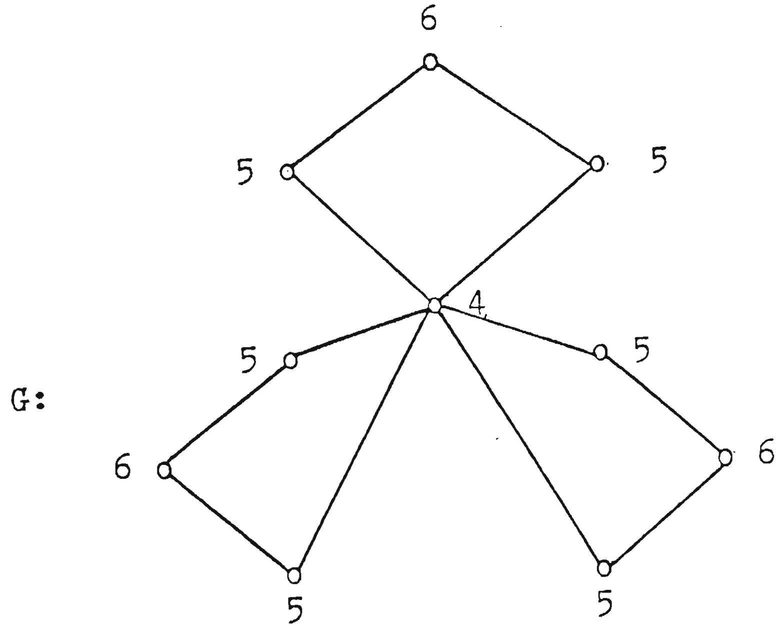
and the n-diameter of G is defined by

$$\text{diam}_n G = \max\{e_n(v) : v \in V(G)\}.$$

Note that for $n = 2$ we have $e_2(v) = e(v)$ for all $v \in V(G)$ while $\text{rad}_2 G = \text{rad} G$ and $\text{diam}_2 G = \text{diam} G$.

4.2.9 Example

In Figure 4.2.10 each vertex of the graph G is labelled with its 3-eccentricity, so that $\text{rad}_3 G = 4$ and $\text{diam}_3 G = 6$.



4.2.10 Figure

4.3 Steiner distance in Trees

We now focus our discussion of Steiner distance on trees. We study trees separately mainly due to the fact that there is a unique path between every pair of vertices in a given tree T. This simplifies the search for a connected subgraph of T of minimum size containing a given set $S \subseteq V(G)$. Thus it is possible to obtain various properties and results related to Steiner distance, which hold for trees but not for graphs in general. We follow [COTZ1].

4.3.1 Lemma

Let T be a nontrivial tree and let $S \subseteq V(T)$ where $|S| \geq 2$; then there is a unique subtree T_S of T of size $d(S)$ containing the vertices of S.

Proof

Suppose to the contrary that there exist two nonisomorphic trees T_S and $T_{S'}$ both of size $d(S)$ which contain the vertices of S. Since $T_S \ncong T_{S'}$ there exists an edge $e \in E(T)$ say, such that $e \in E(T_S)$ and $e \notin E(T_{S'})$. Now since T_S is a tree of minimum size that contains S, there exists a pair of vertices u, v say, of S such that the $u - v$ path in T_S contains the edge e. However $T_{S'}$ contains a $u - v$ path which does not contain the edge e and hence there are at least two distinct $u - v$ paths in T, which is impossible.

Hence $T_S \cong T_{S'}$. □

4.3.2 Definition

Let T be a nontrivial tree and let $S \subseteq V(T)$ where $|S| \geq 2$, then the unique subtree T_S of T of size $d(S)$ containing the vertices of S is defined to be the *tree generated by* S denoted by T_S.

4.3.3 Note

If S and S' are sets of vertices of a tree T such that $S \subset S'$, then $T_S \subset T_{S'}$; otherwise T_S contains an edge e say, that does not belong to $T_{S'}$ and a similar discussion to that followed in the proof of Lemma 4.3.1 provides a contradiction. Hence if S is a subset of the vertex set of a tree T and $v \in V(T) - S$ then $T_{S \cup \{v\}} \supset T_S$. Let w be the (unique) vertex of T_S whose distance from v is a minimum. Then $T_{S\cup\{v\}}$ contains the unique $v-w$ path in T and

$$d(S \cup \{v\}) = d(S) + d(v,w)$$

or equivalently

$$d(S \cup \{v\}) = d(S) + d(v,T_S)$$

where $d(v,T_S)$ denotes the minimum distance from v to a vertex of T_S in T.

We denote by $V_1(T)$ the set of end-vertices of a tree T and the number of end-vertices in T is denoted by $p_1(T)$.

4.3.4 Lemma

Let T be a tree and let S be the set of end-vertices of T; i.e., $S = V_1(T)$, then $T_S = T$.

Proof

Suppose to the contrary, that there exists a vertex $v \in V(T)$ such that $v \notin V(T_S)$. Since $v \notin V(T_S)$, $v \notin S$ and hence $\deg_T v \geq 2$. Let x and y be distinct vertices of T which are adjacent to v. Let $P_1 = vxx_1...x_n$ be a longest path in T which begins with the edge $e_1 = vx$. Then x_n must be an end-vertex of T; i.e., $x_n \in S$. Similarly, let $P_2 = vyy_1y_2...y_m$ be a longest path in T which begins with the edge $e_2 = vy$. Then $y_m \in V_1(T) = S$. Hence the unique $x_n - y_m$ path in T contains the vertex v. However since $v \notin T_S$ there is no $x_n - y_m$ path in T_S which implies that T_S is not connected, which contradicts the definition of T_S. Therefore $T_S = T$. □

The following corollaries follow directly from Lemma 4.3.4 and require no further proof.

4.3.5 Corollary

Let T be a tree, and let $S = V_1(T)$. Then $d(S) = q(T)$ and $d(S \cup \{v\}) = d(S) = q(T)$ for all $v \in V(T)$.

4.3.6 Corollary

Let T be a tree and $n \geq 2$ an integer with $p_1(T) < n$, then $e_n(v) = q(T)$ for all $v \in V(T)$.

The following result considers n-eccentricities of vertices in trees with at least n end- vertices.

4.3.7 Theorem

Let $n \geq 2$ be an integer and suppose that T is a tree of order p with $p_1(T) \geq n$. Let $v \in V(T)$. If $S \subseteq V(T)$, such that $v \notin S$, $|S| = n - 1$ and $d(S \cup \{v\}) = e_n(v)$, then $S \subseteq V_1(T)$.

Proof

Suppose, to the contrary, that there exists a set S of T which satisfies the hypothesis of the theorem such that $S \not\subseteq V_1(T)$. Then there exists a vertex $w \in S$ such that $\deg_T w \geq 2$. Let T_0 denote the subtree of T generated by $S_0 = S \cup \{v\}$, and let T_0' be the branch of T at w that contains v. Suppose there exists an end-vertex x of T in a branch of T at w which is different from T_0', such that $x \notin S$. Then

$$d(S_0 \cup \{x\}) - \{w\}) > d(S_0) = e_n(v),$$

since the only path from x to v in T includes the vertex w and hence $w \in V(T_{S_0 \cup \{x\} - \{w\}})$, which produces a contradiction. Hence there is no such end-vertex x.

Thus every end-vertex y of T in a branch of T at w different from T_0' is in S. Now $\deg_T w \geq 2$; hence there are at least two branches T_1 and T_2 of T at w, and there exist vertices $z_1, z_2 \in S - \{w\}$ such that $z_1 \in V(T_1)$ and $z_2 \in V(T_2)$. Now the unique $z_1 - z_2$ path in T contains the vertex w. But then T_0 is also the tree generated by $S_1 = S_0 - \{w\}$. Let $u \in V_1(T)$ such that $y \notin S$. Then

$$d(S_1 \cup \{y\}) > d(S_0) = e_n(v), \text{ again a contradiction.} \qquad \square$$

4.3.8 Corollary

Let $n \geq 2$ be an integer and T a tree with $p_1(T) \geq n$. Then $\text{diam}_n T = d(S)$, for some set S of n end-vertices of T.

Proof

If $n = 2$, then $\text{diam}_2 T = \text{diam} T$, and S consists of a pair of end-vertices of T between which there is a path of maximum length in T. Assume then that $n \geq 3$. Suppose that $v \in V(T)$ with $e_n(v) = \text{diam}_n T$. Let S' be a set of $n-1$ vertices of T such that $d(S' \cup \{v\}) = e_n(v)$. By Theorem 4.3.7, $S' \subseteq V_1(T)$. If $u \in S'$, then $e_n(u) \geq d(S' \cup \{v\}) = e_n(v)$, which implies that $e_n(u) = \text{diam}_n T$. However, then $S'' = S' \cup \{v\} - \{u\}$ is a set of $n-1$ vertices of T such that $d(S'' \cup \{u\}) = e_n(u)$ and again by Theorem 4.3.7 we have that $S'' \subseteq V_1(T)$ and hence $v \in V_1(T)$. Therefore $S = S' \cup \{v\}$ is a set of n end-vertices of T with $d(S) = \text{diam}_n T$. □

4.3.9 Lemma

Let S be a set of $n \geq 3$ end-vertices of a tree T and suppose that $v \in S$. Then $T_{S-\{v\}}$ can be obtained from T_S by deleting v and every vertex of degree 2 on a shortest path from v to a vertex of degree at least 3 in T_S.

Proof

Let $S' = S - \{v\}$. Then from Note 4.3.3 we have that

$$d(S' \cup \{v\}) = d(S') + d(v, T_{S'}); \text{ hence}$$

$$d(S - \{v\}) = d(S) - d(v, T_{S'}). \qquad \sim (1)$$

Now $d(v, T_{S'})$ is the length of a shortest path $P = vv_1v_2...v_n$ in T such that $v_n \in V(T_{S'})$ and $v_i \in T_S$ for $1 \leq i \leq n-1$. Now $\deg_{T_{S'}} v_n \geq 2$, otherwise v_n is an end-vertex of $T_{S'}$ and hence of T. But then $\deg_{T_S} v_n \geq 2$ and hence $\deg_T v_n \geq 2$ which is a contradiction. Therefore $\deg_{T_S} v_n \geq 3$.

Now $T_S - \{v, v_1, v_2, ..., v_{n-1}\} = T_{S-\{v\}}$ is a tree with $d(S) - d(v, T_{S'}) = d(S - \{v\})$ edges, and since by Lemma 4.3.1 this tree is the unique subtree of T of size $d(S - \{v\})$, the result follows. □

The following result proves to be a useful tool in establishing some important properties in the remainder of this section. We present a slightly different proof to that which appears in [COTZ1].

4.3.10 Theorem

Let $n \geq 3$ be an integer and suppose that T is a tree with $p_1 \geq n$ end-vertices. If v is a vertex of T such that $e_n(v) = \text{rad}_n T$, then there exists a set S of $n-1$ end-vertices of T such that $d(S \cup \{v\}) = e_n(v)$ and $v \in V(T_S)$.

Proof

Assume that the proposition is false. Then there exists a tree T that is a counterexample to the proposition and a vertex v in T for which the conclusion fails.

By Theorem 4.3.7, there exists a set S of $n-1$ end-vertices of T such that $e_n(v) = d(S \cup \{v\})$. From our assumption it follows that $v \notin V(T_S)$; hence S is contained in a single component, T_1 say, of $T - v$. Let u be the unique vertex of T_1 that is adjacent to v in T and let T_2 be the component of $T - u$ that contains v. Then T_1 and T_2 are the two components of $T - uv$ and T is decomposed into T_1, T_2 and the complete graph of order 2 with u and v as vertices. Note that $d(S \cup \{u\}) = d(S \cup \{v\}) - 1 = e_n(v) - 1$.

Now since $e_n(v) = \text{rad}_n T$ we have $e_n(u) \geq e_n(v)$; let R be a set of $n-1$ end-vertices of T such that $d(R \cup \{u\}) = e_n(u)$. Then $R \not\subseteq V(T_1)$, since otherwise, if $R \subseteq V(T_1)$, then $d(R \cup \{v\}) = d(R \cup \{u\}) + 1 > e_n(u)$, which implies that $e_n(v) > e_n(u)$. Furthermore, if R contains at least one vertex from each of $V(T_1)$ and $V(T_2)$ then T_R contains the vertices u and v and so

$$d(R \cup \{v\}) = d(R) = d(R \cup \{u\}) = e_n(u) \geq e_n(v) \geq d(R \cup \{v\}),$$

so that $d(R \cup \{v\}) = e_n(v)$, which contradicts our assumption about T and v. Hence $R \subseteq V(T_2)$. Furthermore

$$d(R \cup \{v\}) = d(R \cup \{u\}) - 1 = e_n(u) - 1 \geq e_n(v) - 1.$$

Now let $x \in S, y \in R$ and let $S_1 = (S - \{x\}) \cup \{y\}$ and $S_2 = (R - \{y\}) \cup \{x\}$. We note that $S \cup R = S_1 \cup S_2$, so $T_{S \cup R} = T_{S_1 \cup S_2}$ and $uv \in E(T_{S_1}) \cap E(T_{S_2})$. Now

$$\begin{aligned} |E(T_{S_1 \cup S_2})| = |E(T_{S \cup R})| &= |E(T_{S \cup \{u\}})| + |E(T_{R \cup \{v\}})| + 1 \\ &= e_n(v) - 1 + e_n(u) - 1 + 1 = e_n(v) + e_n(u) - 1. \end{aligned}$$

But

$$\begin{aligned} |E(T_{S_1 \cup S_2})| &= |E(T_{S_1})| + |E(T_{S_2})| - |E(T_{S_1}) \cap E(T_{S_2})| \\ &\leq |E(T_{S_1})| + |E(T_{S_2})| - 1; \end{aligned}$$

hence $e_n(v) + e_n(u) - 1 \leq |E(T_{S_1})| + |E(T_{S_2})| - 1$. So $|E(T_{S_1})| + |E(T_{S_2})| \geq e_n(v) + e_n(u) \geq 2e_n(v)$.

It follows that $|E(T_{S_1})| \geq e_n(v)$ or $|E(T_{S_2})| \geq e_n(v)$; assume without loss of generality that $|E(T_{S_1})| \geq e_n(v)$. However, S_1 is a set of $n-1$ vertices, so

$$e_n(v) \geq |E(T_{S_1 \cup \{v\}}| \geq |E(T_{S_1})| \geq e_n(v);$$

hence $e_n(v) = |E(T_{S_1 \cup \{v\}})|$ and $v \in T_{S_1}$, contrary to our assumption about v and T. Hence the validity of the theorem follows. □

4.3.11 Corollary

Let $n \geq 3$ be an integer and suppose that T is a tree such that $p_1(T) \geq n$. If v is a vertex of T with $e_n(v) = \text{rad}_n T$, then v is not an end-vertex of T.

Proof

Let S be a set of $n-1$ end-vertices of T such that $e_n(v) = d(S \cup \{v\}) = \text{rad}_n T$. From Theorem 4.3.10 we have that $d(S) = d(S \cup \{v\})$. Assume to the contrary that v is an end-vertex of T. Then $v \notin S$ and $d(S \cup \{v\}) \geq d(S) + 1 > d(S)$ which is a contradiction. Hence $v \notin V_1(T)$.□

Before presenting the next result we introduce some additional terminology, which was used by Oellermann and Tian in [OT1].

4.3.12 Definition

For any tree T having at least three end-vertices, a shortest path from an end-vertex v of T to a vertex of degree at least 3 in T is called a *stem* of T.

A relationship between the n-diameter and the $(n-1)$-diameter of a tree, where $n \geq 3$ is an integer, was established in [COTZ1] and we now present this result.

4.3.13 Theorem

Let $n \geq 3$ be an integer and T a tree of order $p \geq n$. then

$$\mathrm{diam}_{n-1}T \leq \mathrm{diam}_n T \leq \left(\frac{n}{n-1}\right) \mathrm{diam}_{n-1}T.$$

Proof

Suppose S is a set of $n-1$ vertices of T such that $d(S) = \mathrm{diam}_{n-1}T$. Then for every set S' of n vertices of T, where $S \subseteq S'$, we have

$$\mathrm{diam}_{n-1}T = d(S) \leq d(S') \leq \mathrm{diam}_n T.$$

Hence the left inequality in the statement of the Theorem follows.

To verify that $\text{diam}_n T \leq \left(\frac{n}{n-1}\right) \text{diam}_{n-1}T$, we note firstly that if T has at most $n-1$ end- vertices then $\text{diam}_n T = \text{diam}_{n-1}T = p-1$ and hence $\text{diam}_n T < \frac{n}{n-1}\text{diam}_{n-1}T$ in this case.

Assume now that T has $p_1(T) \geq n$. By Corollary 4.3.8, there exists a set S of n end-vertices of T such that $\text{diam}_n T = d(S)$. Let $S = \{v_1, v_2, ..., v_n\}$ and let $\ell_i (1 \leq i \leq n)$ denote the length of the stem in T_S which contains v_i.

We now show that there exists at least one $i(1 \leq i \leq n)$ such that $\ell_i \leq \left(\frac{1}{n-1}\right) \text{diam}_{n-1}T$. Suppose that $\ell_i > \left(\frac{1}{n-1}\right) \text{diam}_{n-1}T$ for all $i(1 \leq i \leq n)$. Since by Lemma 4.3.9, $T_{S-\{v_n\}}$ can be obtained from T_S by deleting v_n and every vertex of degree 2 on the stem of T_S containing v_n, it follows that

$$q(T_{S-\{v_n\}}) \geq \sum_{i=1}^{n-1} \ell_i > (n-1)\left(\frac{1}{n-1}\right) \text{diam}_{n-1}T = \text{diam}_{n-1}T,$$

which is not possible since

$$\text{diam}_{n-1}T \geq d(S - \{v_n\}) = q(T_{S-\{v_n\}}).$$

Hence we may assume without loss of generality that $\ell_n \leq \left(\frac{1}{n-1}\right) \text{diam}_{n-1}T$. Then from Note 4.3.3 we have that

$$\begin{aligned} \text{diam}_n T = d(S) &= d(S - \{v_n\}) + d(v_n, T_{S-\{v_n\}}) \\ &\leq \text{diam}_{n-1}T + \tfrac{1}{n-1}\text{diam}_{n-1}T = \tfrac{n}{n-1}\text{diam}_{n-1}T. \end{aligned}$$

□

The following result from [COTZ1] provides a relationship between the n-diameter and n-radius of a tree (cf. [COTZ1]).

4.3.14 Theorem

Let $n \geq 3$ be an integer and T a tree of order $p \geq n$. Then $\text{diam}_{n-1}T = \text{rad}_n T$.

Proof

If $p_1(T) \leq n-1$, then $\text{rad}_n T = \text{diam}_{n-1} T = p - 1$. Assume then that $p_1(T) \geq n$. We show first that $\text{rad}_n T \geq \text{diam}_{n-1} T$. Let v be any vertex of T and let S be a set of $n-1$ end-vertices of T such that $d(S) = \text{diam}_{n-1} T$. Then $e_n(v) \geq d(S \cup \{v\}) \geq d(S) = \text{diam}_{n-1} T$. Hence

$$\text{rad}_n T = \min_{v \in V(T)} e_n(v) \geq \text{diam}_{n-1} T.$$

We now verify that $\text{diam}_{n-1} T \geq \text{rad}_n T$. Let v be a vertex of T such that $e_n(v) = \text{rad}_n T$. By Theorem 4.3.10 there exists a set S of $n-1$ end-vertices of T such that $d(S \cup \{v\}) = d(S) = \text{rad}_n(T)$ and $v \in V(T_S)$. Therefore,

$$\text{diam}_{n-1} T = \max\{d(S') : |S'| = n-1,\ S' \subseteq V_1(T)\} \geq d(S) = \text{rad}_n T.$$

Hence $\text{diam}_{n-1} T = \text{rad}_n T$. □

4.3.15 Corollary

If $n \geq 2$ is an integer and T a tree of order $p \geq n$, then

$$\text{rad}_n T \leq \text{diam}_n T \leq \frac{n}{n-1} \text{rad}_n T.$$

Proof

The result for $n = 2$ is well-known. If $n \geq 3$ then the result follows directly from Theorems 4.3.13 and 4.3.14. □

For a connected graph G of order $p \geq 2$ the relationship $\text{rad}_n G \leq \text{diam}_n G \leq \frac{n}{n-1} \text{rad}_n G$, surprisingly does not hold and we will see this in Section 4.4.

4.3.16 Definitions

For a connected graph G of order $p \geq 2$ the *diameter sequence* of G is defined to be the sequence

$$\mathrm{diam}_2 G, \mathrm{diam}_3 G, ..., \mathrm{diam}_p G,$$

while the *radius sequence* is the sequence

$$\mathrm{rad}_2 G, \mathrm{rad}_3 G, ..., \mathrm{rad}_p G.$$

Let G be a connected graph of order p. Let $n (2 \leq n \leq p)$ be an integer. A set S consisting of n vertices of G is called an *n-diameter set* of G if $d(S) = \mathrm{diam}_n(G)$.

4.3.17 Note

If T is a tree with $p_1(T)$ end-vertices, then for every integer n $(2 \leq n \leq p_1(T))$ there exists, by Corollary 4.3.8, a set S of n end-vertices of T such that $\mathrm{diam}_n T = d(S)$; i.e., there exist n-diameter sets consisting of only end-vertices of T for all $2 \leq n \leq p_1(T)$.

We now present a result which appears in [OT1] that will aid us in characterizing the diameter sequences of trees.

4.3.18 Theorem

Let T be a nontrivial tree. Then there exists for every integer n with $2 \leq n \leq p_1(T)$, an n- diameter set consisting of only end-vertices of T such that

$$S_2 \subset S_3 \subset ... \subset S_{p_1(T)}.$$

Proof

Since T is a nontrivial tree, $p_1(T) \geq 2$; hence T contains a pair u, v of end-vertices such that diam $T = d(u,v)$. Hence if $S_2 = \{u, v\}$, then $d(S_2) =$ diam$T$ = diam$_2T$ and S_2 is thus a 2-diameter set. If $p_1(T) = 2$ the proof is complete. Assume then that $p_1(T) > 2$. We proceed inductively to complete the proof. Suppose for some integer n and every integer k with $2 \leq k \leq n < p_1(T)$ that there exists a k-diameter set S_k where $S_k \subset V_1(T)$, such that $S_2 \subset S_3 \subset ... \subset S_k \subset ... \subset S_n$. We show now that an $(n+1)$-diameter set $S_{n+1} \subseteq V_1(T)$ containing $S_n \subset V_1(T)$ can be obtained from S_n.

For every vertex $v \in V(T) - V(T_{S_n})$, let $\ell_v = d(v, T_{S_n})$.
Let $w \in V(T) - V(T_{S_n})$ be such that

$$\ell_w = \max\{\ell_v : v \in V(T) - V(T_{S_n}\}.$$

Define $S_{n+1} = S_n \cup \{w\}$. Then $|S_{n+1}| = n + 1$ and w must be an end-vertex of T; hence $S_n \subset S_{n+1} \subseteq V_1(T)$. It remains to be shown that $d(S_{n+1}) = \text{diam}_{n+1}T$.

Let S' be an (n+1)-diameter set of T such that $|S_n \cap S'|$ is as large as possible. Since $|S'| = n + 1$ and $|S_n| = n$, the set $S' - S_n$ is nonempty. Let $v_0 \in S' - S_n$ and let $P = v_0v_1...v_k$ be the stem of $T_{S'}$ containing v_0. Since v_0 is an end-vertex of T, $k \geq 1$.

We now show $k \leq \ell_w$. Assume, to the contrary, that $k > \ell_w$. Since $v_0 \in V(T) - V(T_{S_n})$, it follows from our choice of w that $\ell_{v_0} \leq \ell_w$, which implies that $\ell_{v_0} < k$. Let T_1 be the component of $T - v_{k-1}v_k$ that contains

v_0 and let T_2 be the other component of $T - v_{k-1}v_k$. There exists at least one vertex in $V(T_1) \cap V(T_{S_n})$; otherwise, $T_{S_n} \subset T_2$ and hence the length of the shortest path from v_0 to a vertex of T_{S_n} is at least k, which implies that $\ell_{v_0} \geq k > \ell_w$ which contradicts our choice of w. Therefore there exists a vertex $u \in V(T_1) \cap V(T_{S_n})$ such that $d_T(u, v_0) = \ell_{v_0}$. (Note that u cannot be an end-vertex of T_{S_n}, and hence of T, for otherwise the unique vertex adjacent to u must be a vertex of T_{S_n} and hence $\ell_{v_0} \leq d(u, v_0) - 1$, which is impossible.) Since $V(T_1) \cap V(T_{S_n}) \neq \emptyset$, we have that $S_n \cap V(T_1) \neq \emptyset$, for otherwise $S_n \subset V(T_2)$ and hence $T_{S_n} \subset T_2$, which, as we saw earlier, is not possible. Let $v \in S_n \cap V(T_1)$, then since by assumption $S_n \subset V_1(T_1)$ we have that v is an end-vertex of T. By definition of the path P, $T_{S'-\{v_0\}} \subset T_2$ hence v_0 is the only vertex of S' in T_1, therefore $v \notin S'$. Since by Corollary 4.3.7 S' consists of end-vertices of T we have that both v and v_0 are end-vertices of T. Also since u is not an end-vertex of T we have that $u \neq v$. To complete the proof, we consider two cases.

Case 1) Suppose that $d_T(u, v) < \ell_{v_0}$. Then by Note 4.3.3

$$\begin{aligned} d((S_n - \{v\}) \cup \{v_0\}) &= \ell_{v_0} + d(S_n - \{v\}) \\ &> d(u, v) + d(S_n - \{v\}) \\ &\geq d(S_n) \qquad \text{(since } u \in V(T_{S_n})\text{)}. \end{aligned}$$

which is impossible since $|(S_n - \{v\}) \cup \{v_0\}| = n$. Therefore Case 1 cannot occur.

Case 2) Suppose then that $d_T(u, v) \geq \ell_{v_0}$. If $d_T(u, v) > \ell_{v_0}$ then

$$\begin{aligned} d((S' - \{v_0\}) \cup \{v\}) &= d(S') - d_T(u, v_0) + d_T(u, v) \\ &= d(S') - \ell_{v_0} + d_T(u, v) > d(S') \end{aligned} \quad \sim (1)$$

which is not possible. Hence $d_T(u,v) = \ell_{v_0}$. Further, $u \in V(P)$, otherwise $d_T(v,v_k) > d_T(v_0,v_k)$, which implies that

$$\begin{aligned} d((S'-\{v_0\}) \cup \{v\}) &= d(S'-\{v_0\}) + d_T(v,v_k) \\ &> d(S'-\{v_0\}) + d_T(v_0,v_k) \\ &= d(S'), \end{aligned}$$

which is impossible. However, by (1), $d((S'-\{v_0\}) \cup \{v\}) = d(S')$, which contradicts our choice of S' since

$$|((S'-\{v_0\}) \cup \{v\}) \cap S_n| > |S' \cap S_n|.$$

Hence $k \leq \ell_w$, which implies that

$$\begin{aligned} \mathrm{diam}_{n+1}T &= d(S') \\ &= q(T_{S'-\{v_0\}}) + d(v_0, T_{S'-\{v_0\}}) \\ &= q(T_{S'-\{v_0\}}) + k \leq q(T_{S_n}) + \ell_w = d(S_{n+1}) \leq \mathrm{diam}_{n+1}T. \end{aligned}$$

Hence $d(S_{n+1}) = \mathrm{diam}_{n+1}T$. □

We now present the characterization of diameter sequences of trees, following [COTZ1].

4.3.19 Theorem

A sequence $a_2, a_3, \ldots, a_p$ of positive integers is the diameter sequence of a tree of order p with $p_1(T)$ end-vertices if and only if

(1) $a_{n-1} < a_n \leq \left(\frac{n}{n-1}\right) a_{n-1}$ for $3 \leq n \leq p_1(T)$.

(2) $a_n = p-1$ for $p_1(T) \leq n \leq p$, and

(3) $a_{n+1} - a_n \leq a_n - a_{n-1}$ for $3 \leq n \leq p-1$.

Proof

Let T be a tree of order p with $p_1(T) \geq 2$ end-vertices and diameter sequence $a_2, a_3, ..., a_p$. Assume that $3 \leq n \leq p_1(T)$. By Theorem 4.3.13, $a_{n-1} \leq a_n \leq \left(\frac{n}{n-1}\right) a_{n-1}$. Now by Theorem 4.3.18, there exists an n-diameter set S_n and an $(n-1)$-diameter set S_{n-1}, each consisting of only end-vertices of T, such that $S_{n-1} \subset S_n$; hence $S_n = S_{n-1} \cup \{v\}$ for some end-vertex $v \in V(T) - S_{n-1}$. Thus,

$$a_n = \text{diam}_n T = d(S_n \cup \{v\}) \geq d(S_{n-1}) + 1 > d(S_{n-1}) = \text{diam}_{n-1} T = a_{n-1},$$

which verifies (1).

If $n \geq p_1(T)$, then $\text{diam}_n T = p - 1$, so that $a_{p_1(T)} = a_{p_1(T)+1} = ... = a_p = p - 1$ and hence (2) is established.

To verify (3), we again employ Theorem 4.3.18. Let $a_{n-1} = d(S_{n-1}), a_n = d(S_n)$ and $a_{n+1} = d(S_{n+1})$, where

$$S_n = S_{n-1} \cup \{v\}, S_{n+1} = S_n \cup \{u\} \text{ and } 3 \leq n \leq p - 1.$$

By Note 4.3.3 we have

$$d(S_n) = d(S_{n-1}) + d(v, T_{S_{n-1}}), \text{ so that}$$

$a_n = d(S_n) = d(S_{n-1} \cup \{v\}) = d(S_{n-1}) + d(v, T_{S_{n-1}}) = a_{n-1} + d(v, T_{S_{n-1}})$.
Therefore $a_n - a_{n-1} = d(v, T_{S_{n-1}})$.

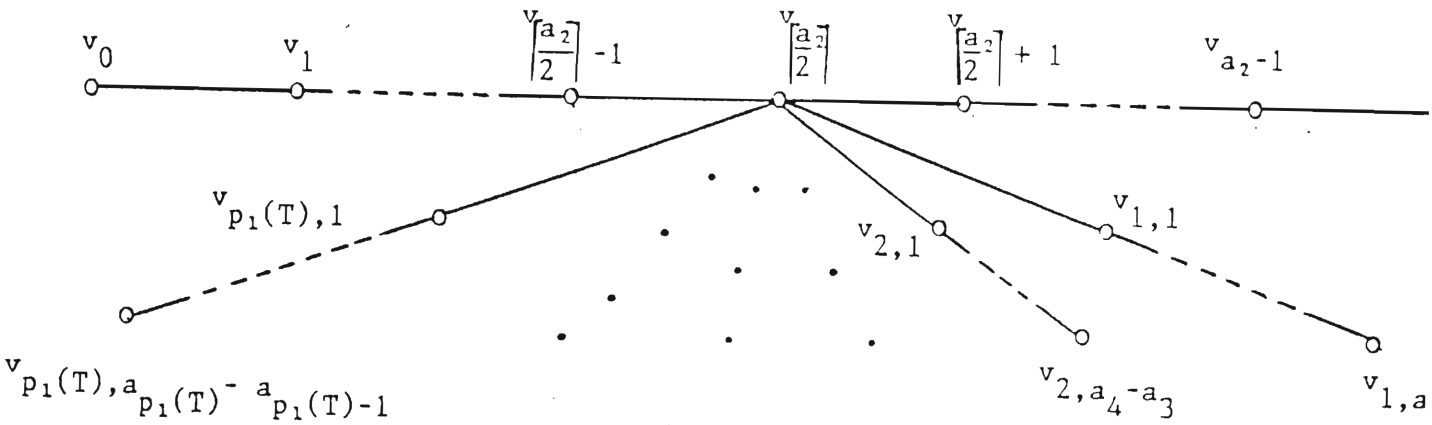
Similarly, $a_{n+1} = a_n + d(u, T_{S_n})$. Therefore,

$$a_{n+1} - a_n = d(u, T_{S_n}).$$

Now $d(u, T_{S_n}) \leq d(u, T_{S_{n-1}})$ since $T_{S_{n-1}} \subset T_{S_n}$ and $d(u, T_{S_{n-1}}) \leq d(v, T_{S_{n-1}})$ otherwise $d(S_{n-1} \cup \{u\}) > d(S_{n-1} \cup \{v\})$.

Therefore $a_{n+1} - a_n = d(u, T_{S_n}) \leq d(v, T_{S_{n-1}}) = a_n - a_{n-1}$ which verifies (3).

For the converse, suppose that $a_2, a_3, ..., a_p$ is a sequence of positive integers satisfying properties (1) - (3). Let H_2 be a path of length a_2 and suppose $H_2 = v_0 v_1 ... v_{a_2}$. For $3 \leq i \leq p_1(T)$, let $H_i = v_{i,0} v_{i,1} ... v_{i,a_i - a_{i-1}}$ be a path of length $a_i - a_{i-1}$. Define T to be the tree obtained by identifying $v_{i,0} (3 \leq i \leq p_1(T))$ with v_r where $r = \lceil \frac{a_2}{2} \rceil$. Then T has size $a_2 + (a_3 - a_2) + (a_4 - a_3) + ... + (a_{p_1(T)} - a_{p_1(T)-1}) = a_{p_1(T)} = p - 1$, and therefore has order p. Further, T has diameter sequence $a_2, a_3, ..., a_p$. □



4.3.20 Figure

The tree T constructed in Theorem 4.3.19.

This leads to a similar characterization of the radius sequences of trees, as stated in [COTZ1].

4.3.21 Corollary

A sequence $a_2, a_3, ..., a_p$ of positive integers is the radius sequence of a tree of order $p \geq 2$ with $p_1(T)$ end-vertices if and only if

(1) $a_3 = 2a_2$ or $a_3 = 2a_2 - 1$

(2) $a_n < a_{n+1} \leq \frac{n}{n-1} a_n$ for $3 \leq n + 1 \leq p_1(T)$

(3) $a_n = p - 1$ for $p_1(T) + 1 \leq n \leq p$ and

(4) $a_{n+1} - a_n \leq a_n - a_{n-1}$ for $4 \leq n \leq p$.

Proof

Let T be a tree of order p with $p_1(T) \geq 2$ end-vertices which has radius sequence $R = a_2, a_3, ..., a_p$. Now by Theorem 4.3.14, $\mathrm{rad}_3 T = \mathrm{diam}_2 T =$ $\mathrm{diam} T$. However if T is a central tree then $\mathrm{diam}\ T = 2\mathrm{rad}_2 T = 2a_2$, while if T is bicentral then $\mathrm{diam}\ T = 2\mathrm{rad}_2 T - 1 = 2a_2 - 1$. Therefore $a_3 = 2a_2$ or $a_3 = 2a_2 - 1$ which verifies (1).

By Theorem 4.3.14 and Corollary 4.3.15, $a_{n+1} \leq \left(\frac{n}{n-1}\right) a_n$ for $3 \leq n+1 \leq$ $p_1(T)$. Let S be any set of n vertices of T. Then for any end-vertex $v \in V(T) - S$

$$d(S \cup \{v\}) \geq d(S) + 1 > d(S).$$

Hence $\mathrm{rad}_{n+1} T = \min\{d(S \cup \{v\}) : v \in V(T) - S\} > \min\{d(S)\} = \mathrm{rad}_n T$ thus $a_{n+1} > a_n$ which verifies (2).

If $n \geq p_1(T) + 1$, then $\mathrm{rad}_n T = p - 1$, so that $a_{p_1(T)+1} = a_{p_1(T)+2} = ... = a_p = p - 1$ and (3) is established.

To verify (4), we note from Theorem 4.3.14 that for $3 \leq n \leq p$, $a_n = \mathrm{diam}_{n-1} T$. Therefore the subsequence $a_3, a_4, ..., a_p$ of R is a subsequence of the diameter sequence of T, and hence by Theorem 4.3.19 $a_{n+1} - a_n \leq a_n - a_{n-1}$ for $4 \leq n \leq p$ which establishes (3).

For the converse suppose that $a_2, a_3, ..., a_p$ is a sequence of positive integers satisfying properties (1) - (4). If $a_3 = 2a_2$ then let H_3 be a path of length $2a_2$ and suppose $H_3 = u_{a_2} u_{a_2-1} ... u_1 c v_1 v_2 ... v_{a_2-1} v_{a_2}$. If $a_3 = 2a_2 - 1$ then let H_3 be a path of length $2a_2 - 1$ and suppose $H_3 = v_{i,0} v_{i,1} ... v_{i,a_i - a_{i-1}}$ be a path of length $a_i - a_{i-1}$. Now define T to be the tree obtained by identifying the vertices $v_{i,0}$ $(4 \leq i \leq p_1(T) + 1)$ in H_i with c in H_3. In either case T has size

$$a_3 + (a_4 - a_3) + (a_5 - a_4) + ... + (a_{p_1(T)+1} - a_{p_1(T)}) = a_{p_1(T)+1} = p - 1$$

and therefore T has order p. Further T has radius sequence R. $\square$

The concept of the centre of a connected graph was generalized in [OT1] as follows:

4.3.22 Definition

The Steiner n-centre $C_n(G), n \geq 2$ of a connected graph G is the subgraph of G induced by the vertices v of G with $e_n(v) = \mathrm{rad}_n G$.

Hence the Steiner 2-centre of a graph is simply its centre.

4.3.23 Note

We now employ a slight variation in notation for the sake of clarity. Since we will often need to look at the n-eccentricity of a vertex in a given graph G, as well as its n-eccentricity in some induced subgraphs of G, we denote by $e_n(v, G)$ the n-eccentricity of the vertex v in the graph G.

The next eleven results first appeared in [OT1].

4.3.24 Lemma

Let T be a tree of order $p \geq 3$ and n an integer with $3 \leq n \leq p$. Let T' be the tree obtained by deleting the end-vertices from T. If T has at least n end-vertices, then

$$C_n(T) \subset C_n(T').$$

Proof

If $v \in V(T')$, then

$$e_n(v, T') \geq e_n(v, T) - (n-1). \qquad \sim (1)$$

Let u be a vertex which is contained in the n-centre of T; i.e., $e_n(u, T) = \mathrm{rad}_n T$. (Note by Corollary 4.3.11 that u is not an end- vertex of T.) Then, by Theorem 4.3.10, there exists a set S of $n-1$ end-vertices of T such that $d(S \cup \{u\}) = e_n(u, T)$ and $u \in V(T_S)$, that is,

$$d(S \cup \{u\}) = d(S) = e_n(u, T).$$

For every $v \in S$, let ℓ_v be the length of the stem in $T_{S \cup \{u\}}$ containing v. (Note that $T_{S \cup \{u\}} = T_S$). Let $\ell = \min_{v \in S} \ell_v$. There are two cases to consider.

Case 1) Suppose $\ell = 1$ and let $v \in S$ such that $\ell_v = 1$. Then every end-vertex of T that does not belong to S is adjacent to a vertex of T_S; otherwise, suppose $w \in V_1(T)$ such that $d(w, T_S) \geq 2$, then $e_n(u,T) \geq d(S \cup \{u\} \cup \{w\} - \{v\}) > d(S \cup \{u\}) = e_n(u,T)$, which is impossible. Hence the end-vertices of T' are exactly the end-vertices of $T_S - S$. Since T_S has $n-1$ end-vertices, T' has at most $n-1$ end-vertices, implying, by Corollary 4.3.6, that for all $x \in V(T')$, $e_n(x,T') = q(T')$. Hence $T' = C_n(T')$. Because T has at least n end-vertices we have by Corollary 4.3.11 that $C_n(T) \subset T'$ so that $C_n(T) \subset C_n(T')$.

Case 2) Assume now that $\ell \geq 2$. Let S' be the set of end- vertices of $T_S - S$. Since $\ell \geq 2$, it follows that for $y \in S$, $xy \in E(T_S)$ if and only if $x \in V_1(T_S - S)$ and hence $|S'| = |S| = n-1$. Further,

$$\begin{aligned} d_{T'}(S' \cup \{u\}) &= d_T(S \cup \{u\}) - (n-1) \\ &= e_n(u,T) - (n-1). \end{aligned}$$

By Corollary 4.3.6, $d_{T'}(S' \cup \{u\}) = e_n(u,T')$ and hence $e_n(u,T') = e_n(u,T) - (n-1)$. Therefore, by (1), $e_n(u,T') = \mathrm{rad}_n T'$ and u belongs to $C_n(T')$ and $C_n(T) \subset C_n(T')$. □

4.3.25 Lemma

Let $n \geq 2$ be an integer and T a tree of order $p \geq n$. Then $C_n(T)$ is a tree.

Proof

If $n = 2$, then the n-centre is simply the centre of T. Since the centre of a tree is isomorphic to either K_1 or K_2 (see [K3]), it follows that the 2-centre of a tree is a tree.

Assume now that $n \geq 3$. If T has at most $n-1$ end-vertices, then, by Corollary 4.3.6, $C_n(T) = T$ so the lemma follows in this case. Suppose thus that T has at least n end-vertices.

Since any induced connected subgraph of a tree is itself a tree it suffices to prove that $C_n(T)$ is connected. Assume, to the contrary, that $C_n(T)$ is disconnected. Let $P = v_0 v_1 ... v_k$ be the shortest path in T between vertices of two components of $C_n(T)$. Then $k \geq 2, v_i \notin V(C_n(T))$ for $1 \leq i \leq k-1$ and $v_0, v_k \in V(C_n(T))$. Let T_1 be the component of $T - v_{k-1}$ containing v_k. By Theorem 4.3.7, and since $v_{k-1} \notin V(C_n(T))$, there exists a set S' of $n-1$ end-vertices of T such that

$$d(S' \cup \{v_{k-1}\}) = e_n(v_{k-1}, T) \geq e_n(v_k, T) + 1 = e_n(v_0, T) + 1. \qquad \sim (1)$$

Observe that $S' \subseteq V(T_1)$, otherwise

$$\begin{aligned} e_n(v_k, T) &\geq d(S' \cup \{v_k\}) \\ &\geq d(S' \cup \{v_{k-1}\}) \\ &= e_n(v_{k-1}, T) \\ &\geq e_n(v_k, T) + 1, \text{ which is not possible.} \end{aligned}$$

Let $S'' = S' \cup \{v_0\}$. Then $T_{S''}$ contains v_{k-1}, implying that $T_{S' \cup \{v_{k-1}\}} \subset T_{S''}$. Hence

$$e_n(v_0, T) \geq d(S'') \geq d(S' \cup \{v_{k-1}\}) = e_n(v_{k-1}, T) \geq e_n(v_0, T) + 1,$$

which is impossible. Hence $C_n(T)$ is connected. □

It is well-known (see [CL1]), that a tree T is the 2-centre of a tree if and only if $T \cong K_1$ or K_2. The following theorem characterizes those trees that are n-centres of trees for $n \geq 3$.

4.3.26 Theorem

Let $n \geq 3$ be an integer and T a tree. Then T is the n-centre of some tree if and only if $p_1(T) \leq n - 1$.

Proof

Suppose that T is the n-centre of some tree H. Let u be a vertex of $T = C_n(H)$. By Theorem 4.3.10, there exists a set S of $n-1$ end- vertices of H such that

$$d_H(S \cup \{u\}) = e_n(u, H) \text{ and } u \in V(H_S)$$

where H_S is the subtree of H generated by S; hence $d(S \cup \{u\}) = d(S)$.

We show first that $V(T) = V(C_n(H)) \subseteq V(H_S)$. Let $v \in V(H) - V(H_S)$ and let $S' = S \cup \{v\}$. Since $v \notin V(H_S)$, it follows that

$$d(S') \geq d(S) + 1 = d(S \cup \{u\}) + 1 = e_n(u, H) + 1.$$

Therefore $e_n(v, H) \geq e_n(u, H) + 1$, which implies that $v \notin V(C_n(H))$. Hence $V(T) = V(C_n(H)) \subseteq V(H_S)$. Since $n \geq 3$, the tree H_S has $n-1$ end-vertices. Therefore T has at most $n-1$ end-vertices; i.e., $p_1(T) \leq n-1$.

For the converse, let T be a tree with $p_1(T) \leq n-1$. If $2p_1(T) \geq n$, then let H be the tree obtained from T by joining two new vertices to each end-vertex of T. If $2p_1(T) < n$ and $T \ncong K_1$, then let H be the tree obtained by joining two new vertices to each of $p_1(T) - 1$ end-vertices of T and then joining $n - 2(p_1(T) - 1)$ new vertices to the remaining end-vertex of T. If $T \cong K_1$, then let H be obtained by joining n new vertices to the vertex of T. In all of the above three cases, $p_1(H) \geq n$. Now let S_1 be the set of end-vertices of H, in any one of the above three cases. By Lemma 4.3.24, $C_n(H) \subset C_n(H - S_1)$. Now $H - S_1 = T$ and since $p_1(T) \leq n-1$, $C_n(T) = T$. Hence $C_n(T) \subset H$. For every vertex v of T we have, by Corollary 4.3.6, that $e_n(v, T) = q(T)$ so that

$$e_n(v, H) = e_n(v, T) + n - 1 = q(T) + n - 1.$$

Since all the vertices of T have the same n-eccentricity in H and $C_n(H) \subset T$, the n-centre of H is precisely T. □

The following corollary follows straight from Theorem 4.3.26 and the fact that every branch at a vertex v of T must contain an end-vertex.

4.3.27 Corollary

If T is a tree that is the n-centre ($n \geq 3$) of some tree, then the maximum degree $\triangle(T)$ of T is at most $n - 1$.

The following Corollary follows as a direct consequence of Theorems 4.3.18 and 4.3.14 together with Lemma 4.3.9.

4.3.28 Corollary

Let T be a tree such that $p_1(T) \geq 3$, and suppose that n is an integer with $3 \leq n \leq p_1(T)$. Let S_{n-1} be an $(n-1)$-diameter set and S_n an n-diameter set, so that $S_{n-1} \subset S_n$. Then

(1) a vertex v of T_{S_n} is an end-vertex of T_{S_n} if and only if $v \in S_n$;

(2) $e_n(v, T_{S_n}) = \mathrm{diam}_n(T_{S_n})$ if and only if v is an end -vertex of T_{S_n};

(3) $\operatorname{diam}_n T_{S_n} = q(T_{S_n}) = \operatorname{diam}_n T$;

(4) every vertex v of T_{S_n} is such that $e_n(v, T_{S_n}) \leq e_n(v, T)$;

(5) if ℓ is the length of a shortest stem in T_{S_n}, then

$$\operatorname{diam}_{n-1} T_{S_n} = \operatorname{diam}_{n-1} T = \operatorname{rad}_n T_{S_n} = \operatorname{rad}_n T = q(T_{S_n}) - \ell.$$

4.3.29 Lemma

Let T be a tree such that $p_1(T) \geq 3$ and suppose that n is an integer with $3 \leq n \leq p_1(T)$. Let S_{n-1} be an $(n-1)$-diameter set and S_n an n-diameter set of T with $S_{n-1} \subset S_n$. Suppose that ℓ is the length of a shortest stem of T_{S_n}. If

$$\mathcal{U} = \{u \in V(T_{S_n}) : \text{ there exists } v \in S_n \text{ with } d(u, v) \leq \ell - 1\},$$

then $C_n(T_{S_n}) = T_{S_n} - \mathcal{U}$.

Proof

Suppose $w \in S_n$ such that the length of the stem in T_{S_n} containing w is ℓ. Then let $S = S_n - w$. If $v \in \mathcal{U}$, then

$$e_n(v, T_{S_n}) \geq d(S \cup \{v\}) \geq q(T_{S_n}) - \ell + 1 = \operatorname{rad}_n T_{S_n} + 1.$$

If $v \in V(T_{S_n}) - \mathcal{U}$, then $\operatorname{rad}_n T_{S_n} \leq e_n(v, T_{S_n}) \leq q(T_{S_n}) - \ell = \operatorname{rad}_n T_{S_n}$. Hence $e_n(v, T_{S_n}) = \operatorname{rad}_n T_{S_n}$ for all $v \in V(T_{S_n}) - \mathcal{U}$. Therefore

$$C_n(T_{S_n}) = T_{S_n} - U. \qquad \square$$

4.3.30 Theorem

Let T be a tree such that $p_1(T) \geq 3$, and suppose that n is an integer with $3 \leq n \leq p_1(T)$. Let S_{n-1} be an $(n-1)$-diameter set and S_n an n-diameter set of T such that $S_{n-1} \subset S_n$. Then for every vertex $v \in V(T_{S_n})$,

$$e_n(v, T_{S_n}) = e_n(v, T).$$

Proof

By Corollary 4.3.28(4), $e_n(v, T_{S_n}) \leq e_n(v, T)$, for all $v \in V(T_{S_n})$. Therefore we have only to show that $e_n(v, T_{S_n}) \geq e_n(v, T)$ for all $v \in V(T_{S_n})$.

Assume, to the contrary, that there exists a vertex v of T_{S_n} such that $e_n(v, T_{S_n}) < e_n(v, T)$. By Corollary 4.3.28(2) and 4.3.28(3), such a vertex v is not an end-vertex of T_{S_n}, that is, $v \notin S_n$. By Theorem 4.3.7, there exists a set S of $n-1$ end-vertices of T such that $d(S \cup \{v\}) = e_n(v, T)$, and $|S \cap S_n|$ is as large as possible. Since $e_n(v, T_{S_n}) < e_n(v, T)$, it follows that $S \not\subset S_n$; otherwise $T_{S \cup \{v\}} \subset T_{S_n}$ and hence $e_n(v, T_{S_n}) \geq e_n(v, T)$, which is not possible. Therefore $S - S_n \neq \emptyset$ and further, since $|S_n| - |S| \geq 1$, we have $S_n - S \neq \emptyset$. Let $u \in S_n - S$ and $w \in S - S_n$. Let ℓ_u and ℓ_w be the lengths of the stems of $T' = T_{S \cup \{v,u\}}$ containing u and w respectively.

We show that $\ell_w > \ell_u$. Assume that $\ell_w \leq \ell_u$. If $\ell_w < \ell_u$, then $(S - \{w\}) \cup \{u, v\}$ is a set of n vertices of T containing v with $d((S - \{w\}) \cup \{u, v\}) > d(S \cup \{v\}) = e_n(v, T)$, which is not possible. If $\ell_u = \ell_w$, then $(S - \{w\}) \cup \{u, v\}$ is a set of n vertices of T containing v such that $d((S - \{w\}) \cup \{u, v\}) = d(S \cup \{v\})$ and $|((S - \{w\}) \cup \{u\}) \cap S_n| > |S \cap S_n|$ which contradicts our choice of S. Hence $\ell_w > \ell_u$.

Let $T'' = T_{S_n \cup \{w\}}$ and let ℓ'_u and ℓ'_w be the lengths of the stems containing

u and w, respectively, in T''.

We show that $\ell'_w \leq \ell'_u$. Suppose that $\ell'_w > \ell'_u$; then $q(T_{(S_n-\{u\})\cup\{w\}}) = d((S_n-\{u\})\cup\{w\}) > q(T_{S_n}) = \text{diam}_n T$, which is impossible. Hence $\ell'_w \leq \ell'_u$.

Since $v \in V(T') \cap V(T_{S_n})$, the tree T' contains a path from every vertex of $S - S_n$ to a vertex of T_{S_n}. In fact, since we are dealing with trees, there is a unique path in T between a vertex of T which is not in T_{S_n} and a vertex of T_{S_n}. Hence T' contains the unique shortest path from any vertex of $S - S_n$ to T_{S_n}.

We show now that if $w' \in S - S_n$, then a shortest path from w' to a vertex of T_{S_n} does not contain a vertex from a stem of T_{S_n} that contains a vertex of $S_n - S$. Assume, to the contrary, that there exists a vertex $w' \in S - S_n$ such that a shortest path from w' to a vertex of T_{S_n} contains a vertex x, say, of a stem of T_{S_n} that contains a vertex $u' \in S_n - S$. Choose such a vertex w' so that $d(w', x)$ is as small as possible. Then $d(u', x) \geq d(w', x)$; otherwise,

$$d((S_n - \{u'\}) \cup \{w'\}) > d(S_n), \text{ which is not possible.}$$

Let k be the shortest distance from w' to a vertex of degree at least 3 in $T_{S\cup\{v,u'\}}$. We establish that $k \leq d(w', x)$. $T_{S\cup\{v,u'\}}$ contains the unique $w' - u'$ path in T which contains x as well as the unique $u' - v$ path which also contains x as an internal vertex. Since x is the only vertex common to the $w' - u'$ path and the $u' - v$ path, it follows that x has degree at least 3 in $T_{S\cup\{v,u'\}}$.

Hence

$$k \leq d(w', x).$$

Now $d(u', x) < k$; otherwise if $d(u', x) > k$, then $d((S \cup \{v, u'\}) - \{w'\}) > d(S \cup \{v\}) = e_n(v, T)$ which is impossible, and if $d(u', x) = k$, then $|(S \cup \{u'\} - \{w\}) \cap S_n| > |S \cap S_n|$ and $d((S \cup \{v, u'\}) - \{w'\}) = d(S \cup \{v\})$ which contradicts our choice of S. This implies that

$$d(w', x) \leq d(u', x) < k \leq d(w', x)$$

which is impossible. Hence if P is a $w - x$ path that is a shortest path from w to a vertex of T_{S_n}, then x does not belong to a stem of S_n that contains u. Hence the distance from u to a vertex having degree at least 3 in T_{S_n} is simply ℓ'_u, and the distance from u to a vertex having degree at least 3 in T' is at least ℓ'_u, that is, $\ell_u \geq \ell'_u$. Hence

$$\ell_w > \ell_u \geq \ell'_u \geq \ell'_w \text{ so that } \ell_w > \ell'_w.$$

Thus, if P' is a shortest path from w to a vertex z', say, having degree at least 3 in T', then P' contains $x \in V(T_{S_n})$. Now x has degree 2 in T', since z' is the first vertex of degree at least 3 on P', and $x \in V(T_{S_n})$. Therefore x must lie on a path between v and some vertex $y \in S_n$. Suppose $y \in S_n \cap S$, then the $y - v$ path (and hence also x) is contained in $T' = T_{S \cup \{u,v\}}$. However x is on the $v - w$ path in T' and $w \neq y$, therefore (since $\ell_w \geq d(w, x)$) $\deg_{T'} x \geq 3$ which is a contradiction. Therefore $y \in S_n - S$. Because x has degree 2 in T',the path from y to x does not contain vertices of T' other than x. Hence the distance from y to a vertex of T' is $d(y, x)$. Now $d(y, x) < d(w, x)$; otherwise if $d(y, x) > d(w, x)$ then $d((S \cup \{v, y\}) - \{w\}) > d(S \cup \{v\}) = e_n(v, T)$ which is impossible, and if $d(y, x) = d(w, x)$ then

$$|((S \cup \{y\}) - \{w\} \cap S_n)| > |S \cap S_n| \text{ and } d(S \cup \{v, y\} - \{w\}) = d(S \cup \{v\})$$

which contradicts our choice of S.

Since $y \in S_n - S$, the vertex x does not belong to a stem of T_{S_n} that contains y. Hence the distance from y to a vertex having degree at least 3 in $T_{S_n \cup \{w\}}$ is less than $d(y, x) < d(w, x)$. However, then $d((S_n - \{y\}) \cup \{w\}) > d(S_n)$, which is impossible. Therefore $e_n(v, T_{S_n}) \geq e_n(v, T)$ for every vertex $v \in V(T_{S_n})$ and hence $e_n(v, T_{S_n}) = e_n(v, T)$ for every vertex $v \in V(T_{S_n})$. □

With the aid of Theorem 4.3.30, we now obtain the following result.

4.3.31 Theorem

Let T be a tree such that $p_1(T) \geq 3$ and suppose that n is an integer with $3 \leq n \leq p_1(T)$. Let S_n be an n-diameter set of T which contains an $(n-1)$-diameter set of T. Then $C_n(T) = C_n(T_{S_n})$.

Proof

Suppose there exists a vertex $u \in V(C_n(T))$ such that $u \notin V(C_n(T_{S_n}))$. Since T is acyclic every path between u and a vertex of T_{S_n} must contain the vertex $x \in V(T_{S_n})$, say, where the $u - x$ path has length $d(u, T_{S_n})$.

By Theorem 4.3.7 there exists a set S' of $n-1$ end-vertices of T_{S_n} such that $d(S' \cup \{x\}) = e_n(x, T_{S_n})$. But $e_n(s, T_{S_n}) = e_n(x, T)$, Theorem 4.3.30, and hence $d(S' \cup \{x\}) = e_n(x, T)$. Note that $T_{S' \cup \{u\}}$ must contain x and therefore $T_{S' \cup \{u\}} \supset T_{S' \cup \{x\}}$. Hence

$$d(S' \cup \{u\}) \geq 1 + d(S' \cup \{x\}) = 1 + e_n(x, T) > \mathrm{rad}_n T$$

which contradicts the fact that $u \in V(C_n(T))$. Therefore

$$C_n(T_{S_n}) \supseteq C_n(T). \qquad \sim (1)$$

Let $v \in V(C_n(T_{S_n}))$. Then $e_n(v, T_{S_n}) = \text{rad}_n T_{S_n}$. However, $e_n(v, T_{S_n}) = e_n(v, T)$ by Theorem 4.3.30, and $\text{rad}_n T_{S_n} = \text{rad}_n T$ by Corollary 4.3.28(5). Hence $e_n(v, T) = \text{rad}_n T$, and thus $v \in V(C_n(T))$. Therefore

$$C_n(T_{S_n}) \subseteq C_n(T) \tag{2}$$

and the result follows from (1) and (2). □

We are now in a position to present a relationship betwee the n-centre and $(n-1)$-centre of a tree for $n \geq 3$ an integer.

4.3.32 Theorem

Let $n \geq 3$ be an integer and T a tree of order $p \geq n$. Then $C_{n-1}(T) \subseteq C_n(T)$.

Proof

If T has at most $n-1$ end- vertices, then $C_n(T) = T$, so trivially $C_{n-1}(T) \subseteq C_n(T)$. Suppose now that T has at least n end-vertices. Let $S_{n-1} \subset V_1(T)$ be an $(n-1)$-diameter set and $S_n \subseteq V_1(T)$ an n-diameter set of T such that $S_{n-1} \subseteq S_n$. Assume first that $n = 3$. Then $S_{n-1} = S_2 = \{u, v\}$, say, and $S_3 = \{u, v, w\}$. Let $T_{S_2} = (u =)u_0u_1...u_k(= v)$. It is known (see [K3], pg 65) that the 2- centre of T_{S_2} is the 2-centre of every subtree H of T that contains T_{S_2}. Hence $C_2(T_{S_2}) = C_2(T) = C_2(T_{S_3})$. The 2-centre of T is therefore $\{u_{\frac{k}{2}}\}$ if k is even and $\langle\{u_{\frac{k-1}{2}}, u_{\frac{k+1}{2}}\}\rangle$ if k is odd. Let x be the vertex of degree 3 in T_{S_3}. Then $d(w, x) \leq \min\{d(u, x), d(v, x)\}$. Therefore $d(w, x) \leq \frac{k}{2}$ if k is even and $d(w, x) \leq \frac{k-1}{2}$ if k is odd. By Theorem 4.3.29, the 3-centre of T_{S_3} can be obtained from T_{S_3} by deleting the vertices of

$$\mathcal{U} = \{z \in V(T_{S_3}) : d(y,z) \leq d(w,x) - 1 \text{ for } y \in S_3\}.$$

Hence if k is even then $u_{\frac{k}{2}} \in V(T_{S_3} - \mathcal{U}) = V(C_3(T_{S_3}))$ and if k is odd then $\{u_{\frac{k-1}{2}}, u_{\frac{k+1}{2}}\} \in V(T_{S_3} - \mathcal{U}) = V(C_3(T_{S_3}))$. That is $C_2(T) \subset C_3(T_{S_3})$. By Theorem 4.3.31, $C_3(T_{S_3}) = C_3(T)$ so that $C_2(T) \subset C_3(T)$.

Suppose now that $n \geq 4$. Let ℓ' and ℓ'' be the lengths of the shortest stems of $T_{S_{n-1}}$ and T_{S_n}, respectively. Let

$$U' = \{u \in V(T_{S_{n-1}}) \ : \ \text{there exists } v \in S_{n-1} \text{ with } d(u,v) \leq \ell' - 1\}, \text{and}$$
$$U'' = \{u \subset V(T_{S_n}) \ : \ \text{there exists } v \in S_n \text{ with } d(u,v) \leq \ell'' - 1\}.$$

By Lemma 4.3.29,

$$C_{n-1}(T_{S_{n-1}}) = T_{S_{n-1}} - U' \text{ and } C_n(T_{S_n}) = T_{S_n} - U''.$$

Since $T_{S_{n-1}} \subset T_{S_n}$ we have $\ell'' \leq \ell'$ and therefore $C_{n-1}(T_{S_{n-1}}) \subset C_n(T_{S_n})$. Therefore, by Theorem 4.3.31,

$$C_{n-1}(T) = C_{n-1}(T_{S_{n-1}}) \subset C_n(T_{S_n}) = C_n(T);$$

that is

$$C_{n-1}(T) \subset C_n(T).$$

□

4.3.33 Definition

Let $n \geq 2$ be an integer and T a tree of order $p \geq \max\{3, n\}$. If $p_1(T) \geq n$, then define the derivative of T, denoted by T', as the tree obtained by deleting the end-vertices of T. Suppose the k^{th}- derivative $T^{(k)}$ of T has been defined. If $T^{(k)} \not\cong K_2$ and has $p_1(T^{(k)}) \geq n$, then the $(k+1)^{\text{st}}$-derivative

$T^{(k+1)}$ is defined as the derivative of $T^{(k)}$.

It is well-known that, by successively deleting the end-vertices of the trees produced (beginning with T) until a tree isomorphic to K_1 or K_2 results, we obtain the centre of a tree T. Hence $C(T) = T^{(k)}$ for some $k \geq 1$.

4.3.34 Theorem

Let $n \geq 3$ be an integer and let T be a tree of order $p \geq n$. Then there exists an integer ℓ such that $C_n(T) = T^{(\ell)}$ where $T^{(\ell)}$ has at most $n-1$ end-vertices.

Proof

By Theorem 4.3.18 there exists an n-diameter set $S_n \subseteq V_1(T)$ which contains an $(n-1)$-diameter set $S_{n-1} \subset V_1(T)$.Let $S_n \subseteq V_1(T)$ be an n-diameter set of T. Let ℓ be the length of a shortest stem in T_{S_n}. By Lemma 4.3.29 $C_n(T_{S_n}) \cong T_{S_n}^{(\ell)}$. Since the end-vertices of T_{S_n} are also end-vertices of T we have that $T_{S_n}^{(\ell)} \subset T^{(\ell)}$ and hence $C_n(T_{S_n}) = C_n(T) \subset T^{(\ell)}$. Note that $T^{(\ell)}$ has at most $n-1$ end- vertices. By Corollary 4.3.6, for every vertex $v \in V(T^{(\ell)})$, $e_n(v, T^{(\ell)}) = q(T^{(\ell)})$. Since $C_n(T_{S_n}) \subset T^{(\ell)} \subset T$, it is clear that for every vertex $u \in V(C_n(T_{S_n})) \cap V(T^{(\ell)})$, $e_n(u, C_n(T_{S_n})) = e_n(u, T_{S_n}) \leq e_n(u, T^{(\ell)}) \leq e_n(u, T)$. However by Theorem 4.3.30, $e_n(u, T_{S_n}) = e_n(u, T)$, hence $q(T^{(\ell)}) = e_n(u, T^{(\ell)}) = e_n(u, T_{S_n}) = e_n(u, T) = \text{rad}_n T$. Thus every vertex $v \in V(T^{(\ell)})$ has $e_n(v, T) = \text{rad}_n T$ and hence $T^{(\ell)} \subset C_n(T)$. Therefore $C_n(T_{S_n}) \cong T^{(\ell)}$. $\square$

The following definition appears in [OT2]

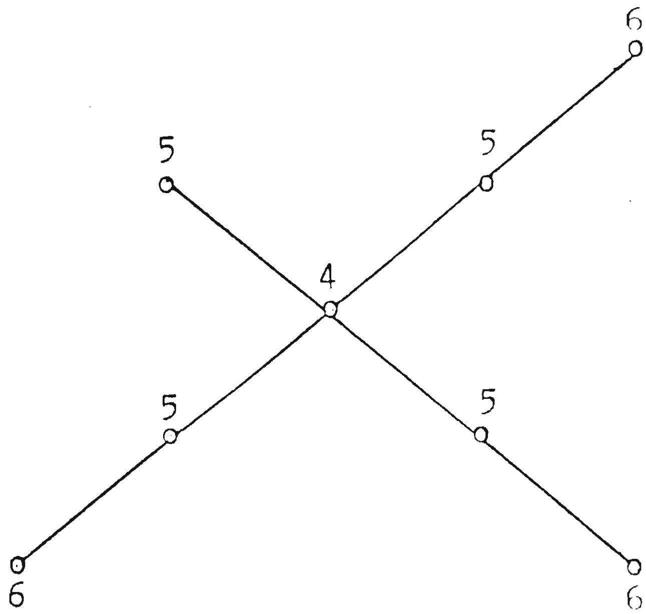
4.3.35 Definition

A non-decreasing sequence $S : a_1, a_2, ..., a_p$ of nonnegative integers is called a *Steiner n-eccentric sequence* or simply an *n-eccentric sequence*, $n \geq 3$, if there exists a connected graph G whose vertices can be labelled $v_1, v_2, ..., v_p$ such that $e_n(v_i) = a_i$ for $1 \leq i \leq p$. In this case, we call S the *Steiner n-eccentricity sequence* of G.

The 2-eccentricity sequence of a connected graph is therefore its eccentricity sequence.

4.3.36 Example

The graph of Figure 4.3.37 has 3-radius 4, 3-diameter 6 and 3-eccentricity sequence $S : 4, 5, 5, 5, 5, 6, 6, 6$.



4.3.37 Figure

The following lemma from [OT1] which holds not only for trees but for graphs in general will prove to be useful.

4.3.38 Lemma

Let G be a graph of order p and n an integer satisfying $2 \leq n \leq p$. If $uv \in E(G)$, then $|e_n(u, G) - e_n(v, G)| \leq 1$.

Proof

We may assume without loss of generality that $e_n(u,G) \leq e_n(v,G)$. Let S be any set of n vertices containing v. If $u \in S$, then $d(S) \leq e_n(u,G)$. If $u \notin S$, then let $S' = (S - \{v\}) \cup \{u\}$. Since $uv \in E(G)$, it follows that

$$d(S) \leq d(S') + 1 \leq e_n(u,G) + 1.$$

Hence $e_n(v) = \max\{d(S) : S \subseteq V(G), |S| = n \text{ and } v \in S\} \leq e_n(u,G) + 1$; that is, $e_n(v,G) - e_n(u,G) \leq 1$ which implies that

$$|e_n(u,G) - e_n(v,G)| \leq 1. \qquad \square$$

The following results concerning n- eccentricities of trees were presented by Oellermann and Tian in [OT2] and will culminate in a characterization of the n-eccentricity sequences of trees (cf. [OT2]).

4.3.39 Lemma

Let T be a tree with $p_1(T) \geq 3$ and suppose that n is an integer with $3 \leq n \leq p_1(T)$. Let $S_{n-1} \subset V_1(T)$ be an $(n-1)$- diameter set and $S_n \subseteq V_1(T)$ an n-diameter set of T such that $S_{n-1} \subset S_n$. If $a_1, a_2, ..., a_p$ is the n-eccentricity sequence of T_{S_n}, then for every integer k with $a_1 < k \leq a_p$ there exists some i where $p - n(a_p - a_1) \leq i \leq p - n + 1$, such that $a_i = a_{i+1} = ... = a_{i+n-1} = k$.

Proof

Let ℓ be the length of a shortest stem of T_{S_n}. Then for every vertex v in S_n there exists exactly one vertex v' of T_{S_n} such that $d(v,v') = \ell$. Let P_v be the $v - v'$ path in T_{S_n}. By Lemma 4.3.29, $C_n(T_{S_n})$ can be obtained by deleting, for every $v \in S_n$, all the vertices of P_v, except v', from T_{S_n}. Now $e_n(v, T_{S_n}) = \text{diam}_n T_{S_n} = a_p$ and since $v' \in V(C_n(T_{S_n}))$, $e_n(v', T_{S_n}) = \text{rad}_n T_{S_n} = a_1$ for all $v \in S_n$; hence, by Lemma 4.3.38, there is at least one vertex of $P_v - v'$ that has n-eccentricity k for $a_1 < k \leq a_p$. By our choice of S_n we have

$$\text{diam}_{n-1} T_{S_n} = q(T_{S_n}) - \ell = \text{diam}_n T_{S_n} - \ell = a_p - \ell = \text{rad}_n T_{S_n} = a_1.$$

Hence $\ell = a_p - a_1$, which implies that for each k with $a_1 < k \leq a_p$ there is exactly one vertex on $P_v - v'$ whose n- eccentricity is k. Since $|S_n| = n$ we have n end-vertices in T_{S_n} and thus n stems of T_{S_n} and hence n paths P_v. Thus on each of the n paths P_v there is a vertex with n-eccentricity k where $a_1 < k \leq a_p$, and hence the lemma follows. $\square$

4.3.40 Lemma

Let T be a tree of order $p \geq 3$ and n an integer with $3 \leq n \leq p$. Suppose $S : a_1, a_2, ..., a_p$ is the n-eccentricity sequence of T. Then

(1) $a_1 \geq n - 1$, and

(2) for every integer k with $a_1 < k \leq a_p$ there exist at least n consecutive elements of S equal to k.

Proof

For every vertex $v \in V(T), e_n(v, T) \geq n - 1$, since a tree with n vertices has size $n - 1$; hence it follows that $a_1 \geq n - 1$, which establishes (1). If T has at most $n - 1$ end-vertices, then $a_1 = a_p$ and (2) holds vacuously. Suppose therefore that T has at least n end-vertices. Then $a_1 < a_p$. Let

$S_{n-1} \subset V_1(T)$ and $S_n \subseteq V_1(T)$ be $(n-1)$-diameter and n-diameter sets of T, respectively, such that $S_{n-1} \subset S_n$. Then, by Theorem 4.3.30, $e_n(v,T) = e_n(v,T_{S_n})$ for every vertex $v \in V(T_{S_n})$. Since $\text{diam}_n T_{S_n} = \text{diam}_n T = a_p$ and $\text{rad}_n T_{S_n} = \text{rad}_n T = a_1$, Lemma 4.3.39 implies that T_{S_n} and thus T contain at least n vertices whose n-eccentricity is k for $a_1 < k \leq a_p$. Hence (2) is established. □

4.3.41 Lemma

Let $n \geq 3$ be an integer and suppose that $S : a_1, a_2, ..., a_p$ is the n-eccentricity sequence of a tree T of order $p \geq 3$. Then if $a_1 \neq a_p$, $\left(\frac{n}{n-1}\right) a_1 = a_p + \frac{m_0-1}{n-1}$ where m_0 is the largest integer such that $a_1 = a_{m_0}$.

Proof

Let $S_n \subseteq V_1(T)$ be an n-diameter set of T which contains an $(n-1)$-diameter set. Then by Theorem 4.3.31, the n-centre of T is isomorphic to the n-centre of T_{S_n} and, by Lemma 4.3.29, can be obtained from T_{S_n} by deleting end-vertices until a tree with at most $n-1$ end-vertices remains. Let ℓ be the length of a shortest stem in T_{S_n}. Then the ℓ^{th} derivative of T_{S_n} is $C_n(T)$. As we saw in the proof of Lemma 4.3.39, $\ell = a_p - a_1$ and obviously m_0 is the order of $C_n(T)$. Thus the number of edges in T_{S_n} is equal to the number of edges in $C_n(T_{S_n})$ (which is $m_0 - 1$), plus ℓn;

$$\begin{aligned} \text{i.e., } a_p &= \ell n + m_0 - 1 = (a_p - a_1)n + m_0 - 1 \\ \text{so } n a_1 &= (n-1)a_p + m_0 - 1. \end{aligned}$$

$$\text{Hence } \left(\frac{n}{n-1}\right) a_1 = a_p + \frac{m_0 - 1}{n-1}. \qquad \square$$

Lesniak [L1] characterized the eccentricity (or 2-eccentricity) sequences of trees. A necessary and sufficient condition for a nondecreasing sequence of integers to be the n-eccentricity sequence of a tree for $n \geq 3$, is now presented, following [OT2].

4.3.42 Theorem

Let $n \geq 3$ be an integer. A nondecreasing sequence $S : a_1, a_2, ..., a_p$ of $p \geq n$ positive integers is the n- eccentricity sequence of a tree if and only if

(1) $a_1 \geq n - 1$,

(2) $\left(\frac{n}{n-1}\right) a_1 = a_p + \frac{m_0-1}{n-1}$ if $a_1 \neq a_p$ where m_0 is the largest integer such that $a_1 = a_{m_0}$, and

(3) if $a_1 < a_p$ and k is an integer with $a_1 < k \leq a_p$, then there exists an integer $i(2 \leq i \leq p - n + 1)$ such that $a_i = a_{i+1} = ... = a_{i+n-1} = k$; otherwise, if $a_1 = a_p$, then $p = a_1 + 1$.

Proof

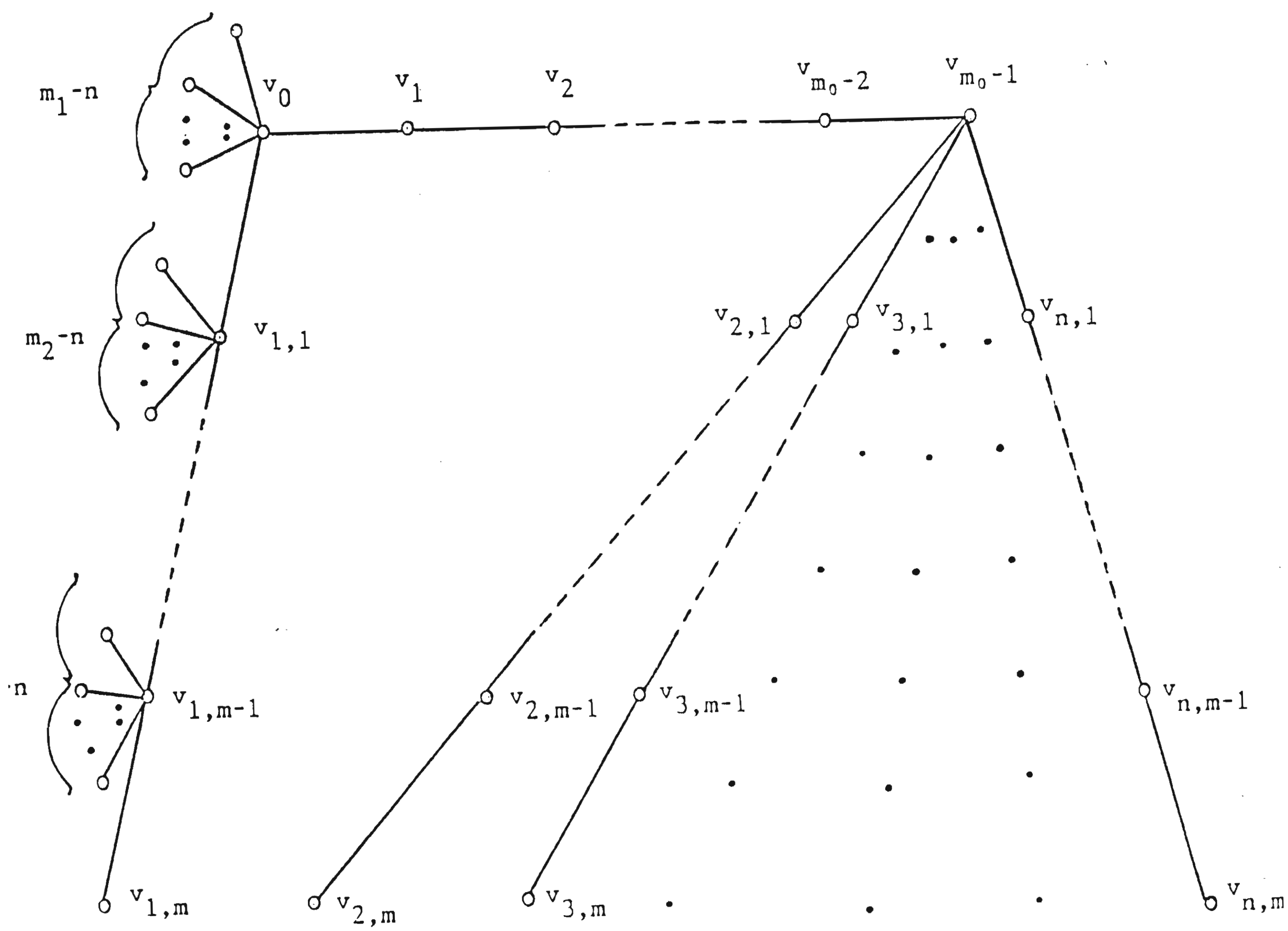
Suppose that $S : a_1, a_2, ..., a_p$ is the n-eccentricity sequence of some tree of order $p \geq n$. Note that, by Corollary 4.3.6, if $a_1 = a_p$, then $a_1 = q(T) = p - 1$; therefore $p = a_1 + 1$. Hence together with Lemmas 4.3.39, 4.3.40 and 4.3.41, this implies that conditions (1), (2) and (3) of the theorem hold.

For the converse suppose that $S : a_1, a_2, ..., a_p$ is a nondecreasing sequence of positive integers satisfying conditions (1), (2) and (3) of the theorem. If $a_1 = a_p$, then $p = a_1 + 1$. Let T be a path of length a_1. Then each vertex of T has n-eccentricity $p - 1 = a_1$ (since $n \geq 3$), that is, S is the n-eccentricity sequence of T.

Assume now that $a_1 < a_p$. Let $A = \{a_{i_0}, a_{i_1}, a_{i_2}, ..., a_{i_m}\}$ with $a_{i_0} < a_{i_1} < ... < a_{i_m}$ be the distinct elements of the sequence S. Then $a_{i_0} = a_1$ and by Lemma 4.3.38, $a_{i_m} = a_p = a_1 + m$. Hence $a_1 = nm + m_0 - m - 1, a_p = nm + m_0 - 1$ and S contains all the integers between a_1 and a_p. Let m_j be the number of occurrences of a_{i_j} in S. Let $H_0 = v_0 v_1 v_2 ... v_{m_0 - 1}$ be a path of order m_0 and let $T_1 \cong T_2 \cong ... \cong T_n \cong P_{m+1}$, where $T_i = v_{i,0} v_{i,1} ... v_{i,m}$ for $i = 1, 2, ..., n$. Denote by H the tree obtained from H_0 and $T_1, T_2, ..., T_n$ by identifying $v_{1,0}$ with one end-vertex of H_0 and then identifying $v_{i,0}$ with the other end-vertex of H_0 for $2 \leq i \leq n$ if $m_0 \geq 2$, otherwise, identify $v_{i,0}$ with the only vertex of H_0 for $1 \leq i \leq n$. Finally, join $m_j - n$ new vertices to $v_{1,j-1}$ for $1 \leq j \leq m$ and let T be the resulting tree. Then T has order

$$
\begin{aligned}
& m_0 + mn + (m_1 - n) + (m_2 - n) + ... + (m_m - n) \\
= \quad & m_0 + mn + (m_1 + m_2 + ... + m_m) - mn = m_0 + m_1 + m_2 + ... + m_m = p,
\end{aligned}
$$

and we verify now that T has n-eccentricity sequence S. Referring to Figure 4.3.43 each vertex v_i ($0 \leq i \leq m_0 - 1$) has $\{v_i, v_{j,m}\}_{j=1,2,...,n-1}$ as an n-eccentric set and hence $e_n(v_i, T) = nm + m_0 - m - 1 = a_1$. Hence we have m_0 vertices with n-eccentricity equal to a_1 as required. The vertices $v_{i,1}$ $(1 \leq i \leq n)$ and the $m_1 - n$ end- vertices of T adjacent to v_0 have n-eccentricity equal to $nm + m_0 - m = a_1 + 1$. Hence we have m_1 vertices with n-eccentricity $a_1 + 1$. Continuing in this manner we see that there are m_j vertices with n-eccentricity $a_1 + j$ $(0 \leq j \leq m)$ and it follows that T has n-eccentricity sequence S. $\square$



4.3.43 Figure

The graph constructed in Theorem 4.3.42 for $m_0 \geq 2$.

The following definitions were introduced by Oellermann in [O1].

4.3.44 Definition

The *Steiner number* $S(G)$ of a connected graph G of order p is the least positive integer m for which there exists a set S of m vertices of G such that $d(S) = p - 1$.

Thus the Steiner number of a connected graph G is the smallest cardinality of a set S of vertices of G such that every connected subgraph of minimum size that contains S is a spanning tree of G.

4.3.45 Definition

The k^{th} *Steiner number* $S_k(G)$ of a graph G is the smallest positive integer m for which there exists a set S of m vertices of G such that $d(S) = k$.

For every connected graph G of order p, the sequence $S_0(G), S_1(G), ..., S_{p-1}(G)$ is called the *Steiner sequence* of G. Note that $S_{p-1}(G) = S(G)$. For example the Steiner sequence of the path P_n is 1,2,2,...,2.

The following Theorem which appears in [O1] gives necessary and sufficient conditions for a sequence of positive integers to be the Steiner sequence of a tree.

4.3.46 Theorem

Let $S : s_0, s_1, ..., s_{p-1}$ be a sequence of $p \geq 3$ positive integers. Then S is a Steiner sequence of a tree if and only if the following conditions are satisfied:

(1) $s_0 = 1, s_1 = 2, s_2 = 2;$

(2) S is a nondecreasing sequence such that $0 \leq s_{k+1} - s_k \leq 1$ for $2 \leq k < p-1$; and

(3) if $3 \leq n \leq s_{p-1}$, ℓ is the largest positive integer such that $S_\ell(T) = n$, and k is the largest positive integer such that $S_k(T) = n-1$, then $\ell \leq \frac{n}{n-1}k$.

Proof

Suppose T is a tree of order $p \geq 3$, with Steiner sequence S. Then since K_1, K_2 and P_3 are subtrees of T it follows that $s_0 = 1, s_1 = 2$ and $s_2 = 2$. Hence (1) is established.

Let S' be a set of s_{r-1} vertices of T such that $d(S') = r-1$, and let S'' be a set of s_r vertices of T such that $d(S'') = r$. Take any $v \in S''$. Then v is an end-vertex of $T_{S''}$. Let u be the unique vertex adjacent to v in $T_{S''}$. Let $S''' = S'' - \{v\} \cup \{u\}$. Then $d(S''') = d(S'') - 1 = r - 1$. Therefore

$$s_{r-1} \leq |S'''| = s_r.$$

Hence S is nondecreasing. We show now that $S_{k+1}(T) = S_k(T)$ or $S_{k+1}(T) = S_k(T) + 1$ for all k, $1 \leq k < p-1$, which will establish (2). Suppose that $1 \leq k < p-1$ and $S_k(T) = m$. Then there exists a set S' of m vertices of T such that $T_{S'}$ has k edges. Further, $v \in S'$ if and only if $v \in V_1(T_{S'})$. If T contains a vertex u that is adjacent to a vertex w of S' but $u \notin V(T_{S'})$, then

$$d((S - \{w\}) \cup \{u\}) = k + 1$$

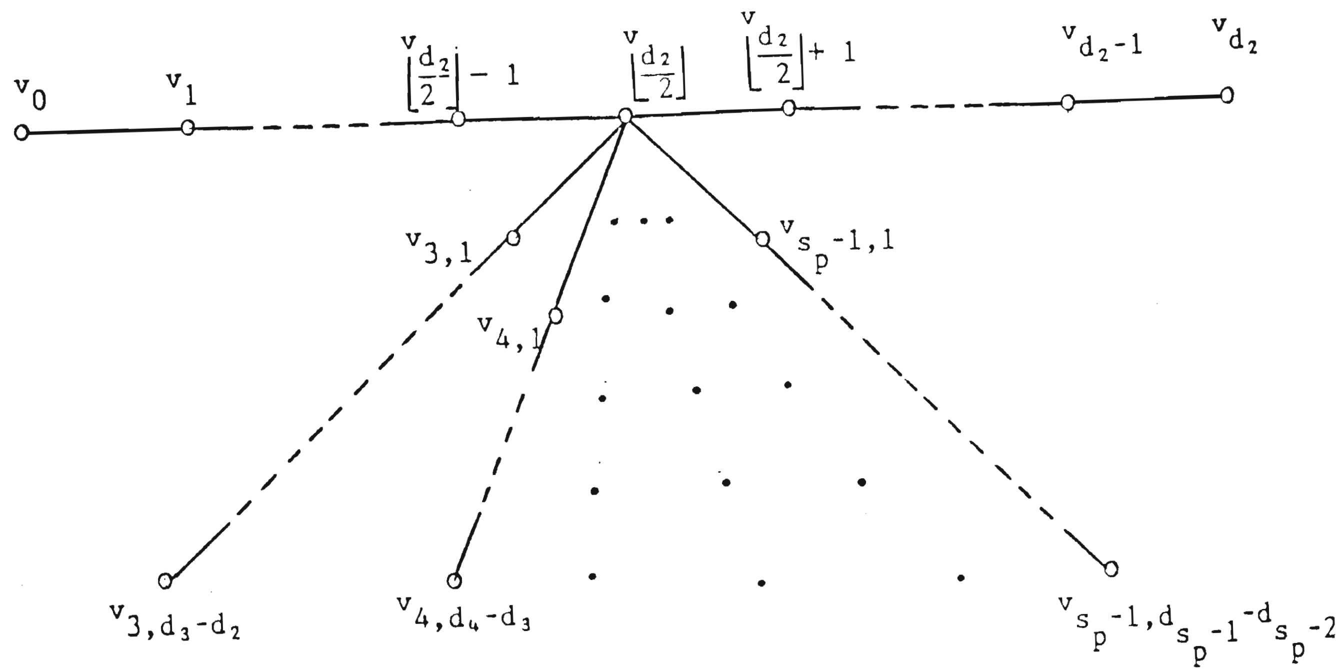
so that $S_{k+1}(T) = m$. However, if the end-vertices of $T_{S'}$ are also end-vertices of T, then it follows, since $k < p-1$, that there is a vertex x in

$V(T)$ such that $x \notin V(T_{S'})$ and x is adjacent with a vertex of $T_{S'}$. Hence $d(S' \cup \{x\}) = k+1$, so $S_{k+1}(T) \leq m+1$. We may thus conclude that if $1 \leq k < p-1$, then $S_{k+1}(T) = S_k(T)$ or $S_{k+1}(T) = S_k(T)+1$. Hence $0 \leq s_{k+1} - s_k \leq 1$ for $2 \leq k < p-1$.

To establish (3) we note that if $2 \leq n \leq S(G)$ and k is the largest positive integer such that $S_k(T) = n$ then $k = \text{diam}_n T$. Hence, if $n \geq 3$ and ℓ is the largest positive integer such that $S_\ell(T) = n$ while k is the largest positive integer such that $S_k(T) = n-1$, then $\ell = \text{diam}_n T$ and $k = \text{diam}_{n-1} T$. Hence by Theorem 4.3.13 we have that $\ell \leq \frac{n}{n-1} k$.

Conversely suppose that $S : s_0, s_1, ..., s_{p-1}$ is a sequence of positive integers that satisfies conditions (1), (2) and (3) of the theorem. For $i = 1, 2, ..., s_{p-1}$ let d_i denote the largest integer such that $S_{d_i}(T) = i$. Note since S is nondecreasing and since consecutive terms of S differ by at most 1, that d_i is defined for all $i = 1, 2, ..., s_{p-1}$. By condition (3) of the theorem, $d_n \leq \frac{n}{n-1} d_{n-1}$ for $2 \leq n \leq s_{p-1}$. Let $m = \lfloor \frac{d_2}{2} \rfloor$ and let $P = v_0 v_1 ... v_{d_2}$ be a path of length d_2. For $i = 3, 4, ..., s_{p-1}$, let $H_i = u_{i,0} u_{i,1} u_{i,2} ... u_{i, d_i - d_{i-1}}$ be a path of length $d_i - d_{i-1}$. Let T be the tree obtained from $P \cup H_3 \cup H_4 \cup ... \cup H_{s_{p-1}}$ by identifying v_m and the vertices $u_{i,0}$ for $i = 3, 4, ..., s_{p-1}$. Then T has Steiner sequence S. □

T:



4.3.47 Figure

The graph T constructed in Theorem 4.3.46.

4.4 Steiner distance in Graphs

Our discussion of Steiner distance is now broadened so as to include graphs in general. Obviously, any results obtained for graphs in general will also hold for trees.

Given a graph G of order p, and any subset $S \subseteq V(G)$, the minimum possible value for $d(S)$ is $|S|-1$, this being the size of a tree with $|S|$ vertices.

Now $d(S) = |S| - 1$ for every subset S of G if and only if G is complete; for otherwise, if $S^* = \{u, v\}$ where $uv \notin E(G)$, then $d(S^*) \geq 2 = |S^*|$. The related problem of determining the minimum size of a graph G of order p having the property that $d(S) = |S| - 1$ for all subsets $S \subseteq V(G)$ with $|S| = n$ for a fixed n where $2 \leq n \leq p$, was discussed in [COTZ1].

4.4.1 Definition

Let n and p be integers with $2 \leq n \leq p$. A graph G of order p is called *(n;p)-complete* if it is of minimum size with the property that $d(S) = n - 1$ for all such $S \subseteq V(G)$ with $|S| = n$.

The goal is thus to determine the size of an $(n; p)$-complete graph for each pair n, p of integers with $2 \leq n \leq p$. The following results appear within a proof of a theorem by Harary [H2] and will prove to be useful.

4.4.2 Theorem

(i) If $2 \leq 2k = n < p$, then C_p^k is n-connected.

(ii) Let p be an even integer satisfying $p > n = 2k + 1 \geq 3$. If G is the graph obtained by joining diametrically opposite vertices of C_p in C_p^k, then G is n-connected.

(iii) Let p be an odd integer such that $p > n = 2k + 1 \geq 3$, and let C_p be the cycle $v_0, v_1, v_3, ..., v_{p-1}, v_0$. If G is the graph obtained by adding $\frac{(p+1)}{2}$ edges to C_p^k, namely those edges joining v_i and v_j, where $j - i = \frac{(p-1)}{2}$, then G is n-connected.

The following lemma will aid us in the determination of the size of an $(n;p)$-complete graph, following [COTZ1].

4.4.3 Lemma

Let n and p be integers with $2 \leq n \leq p$. Every $(n;p)$-complete graph is $(p-n+1)$-connected.

Proof

Suppose, to the contrary, that there exists an $(n;p)$-complete graph G which is not $(p-n+1)$-connected. Then there exists a vertex cutset X of cardinality $p-n$ such that $G-X$ has two or more components. Let $S = V(G) - X$. Since $|S| = n$ and $\langle S \rangle$ is disconnected, G is not $(n;p)$–complete, producing a contradiction. □

4.4.4 Note

If G is $(n;p)$- complete, where $2 \leq n \leq p$, then $\delta(G) \geq p - n + 1$; since otherwise if $\delta(G) < p - n + 1$, then removing all the vertices adjacent to a vertex of degree $\delta(G)$ would result in a disconnected graph or K_1 implying that G is not $(p-n+1)$-connected.

4.4.5 Theorem

Let n and p be integers with $2 \leq n \leq p$. The size of an $(n;p)$- complete graph G is $n-1$ if $p = n$ and $\left\lceil \frac{(p-n+1)p}{2} \right\rceil$ if $p > n$.

Proof

Assume $p = n$. Then a graph of order n which has minimum size having the property that its vertex set induces a connected graph of size $n - 1$, is a tree. Conversely any tree of order n is an $(n; n)$-complete graph. Hence a graph is $(n; n)$-complete if and only if it is a tree of order n. Therefore for $p = n$ the size of an $(n; p)$-complete graph is $n - 1$.

Assume, then, that $p > n$. By Note 4.4.4, if G is $(n; p)$- complete, then $\delta(G) \geq p - n + 1$. Therefore, if for given integers n and p, with $2 \leq n \leq p$, we can exhibit either a $(p-n+1)$-regular $(n; p)$-complete graph or an $(n; p)$-complete graph all of whose vertices have degree $p - n + 1$ except at most one, which has degree $p - n + 2$, then the result will follow.

Suppose first that there exists an integer $k \geq 2$ such that $p = (n-1)k$. Consider the graph $\overline{kK_{n-1}}$. Since any set S of n vertices of $\overline{kK_{n-1}}$ induces a connected graph, we have that $d(S) = n - 1$. Since $\overline{kK_{n-1}}$ has $k(n-1) = p$ vertices and is $(p - n + 1)$-regular, it is an appropriate $(n; p)$-complete graph. Hence assume that $n - 1$ does not divide p. Thus $p = (n-1)q + r$, where $2 \leq r \leq n$, $r \neq n - 1$, $q \geq 1$ and q and r are integers. For each such integer r, we describe the appropriate Harary Graph H_r which is an $(n; n - 1 + r)$- complete graph with the desired properties. It will then folow that $H_r + \overline{(q-1)K_{n-1}}$ is an $(n; p)$-complete graph with the required properties, which will complete the proof.

Case 1) Assume r is even, so that that $r = 2k \geq 2$. By Theorem 4.4.2 (i), the graph $H_r \cong C_{n-1+r}^{k}$ is r-connected. Let S be a set of n vertices of H_r. Since $|V(H_r) - S| = r - 1$, removing the $r - 1$ vertices of $V(H_r) - S$ from H_r will result in a connected graph with vertex set S; i.e., $\langle S \rangle$ is connected. Therefore H_r is a $2k = r$-regular, $(n; n-1+r)$-complete graph. Hence $H_r + \overline{(q-1)K_{n-1}}$ has order p, is $r + (q-1)(n-1) = r + q(n-1) - n + 1 = p - n + 1$-regular, and is thus an $(n;p)$-complete graph with size $\frac{p(p-n+1)}{2}$.

Case 2) Assume r is odd, so that $r = 2k+1 \geq 3$. We consider two subcases.

Subcase 2.1) Assume n is even. Let H_r be the graph obtained by joining diametrically opposite vertices of C_{n-1+r} in C^k_{n-1+r}. By Theorem 4.4.2 (ii), H_r is r-connected. The proof follows as in Case 1.

Subcase 2.2) Assume n is odd. Let H_r be the graph obtained as follows. First draw C^k_{n-1+r} and label its vertices $v_0, v_1, ..., v_{n-2+r}, v_0$. Then, to C^k_{n-1+r} add the edges joining vertex v_0 to vertices $v_{\frac{n-2-r}{2}}$ and $v_{\frac{n+r}{2}}$, together with the edges joining the vertex v_i to the vertices $v_{i+\frac{n+r}{2}}$ (where all additions are taken modulo n), for $1 \leq i < \frac{n-2+r}{2}$. By Theorem 4.4.2 (iii), H_r is r-connected and, again the proof follows as in Case 1. □

In Theorem 4.3.42 we established a necessary and sufficient condition for a nondecreasing sequence of positive integers to be an n-eccentricity sequence of a tree. Although far less descriptive, a necessary and sufficient condition for a nondecreasing sequence of positive integers with m distinct values to be the n-eccentricity sequence of a graph, was established in [OT2].

4.4.6 Theorem

A nondecreasing sequence $S : a_1, a_2, ..., a_p$ with m distinct values is the n-eccentricity sequence , $n \geq 2$, of a graph if and only if some subsequence of S with m distinct values is the n- eccentricity sequence of some graph.

Proof

Suppose S is a sequence with m distinct values, which is the n- eccentricity sequence of some graph G; then since S is a subsequence of itself, we have S is the n-eccentricity sequence of a graph and S has m distinct values.

For the converse, suppose that S' is a subsequence of S that has the same m distinct values as S and suppose that S' is the n-eccentricity sequence of some graph G with $V(G) = \{v_1, v_2, ..., v_p\}$. Let the distinct values of S' be given by $t_1, t_2, ..., t_m$. Let s_i $(1 \leq i \leq m)$ be the number of occurrences of t_i in S, and let s'_i $(1 \leq i \leq m)$ be the number of occurrences of t_i in S'. Now for each t_i $(1 \leq i \leq m)$ select a vertex v_i of G whose n-eccentricity in G is t_i. Let $n_i = s_i - s'_i + 1$. In G replace v_1 with a copy of K_{n_1} and join each vertex of K_{n_1} to all the vertices adjacent to v_1 in G. Call this graph G_1. Note that each vertex v of the copy of K_{n_1} has $e_n(v, G_1) = e_n(v_1, G) = t_1$ while for $2 \leq i \leq m$, $e_n(v_i, G_1) = t_i$. In G_1, replace v_2 with a copy of K_{n_2} and join each vertex of K_{n_2} to all the vertices adjacent to v_2 in G_1. Call this graph G_2. Again each vertex v of the copy of K_{n_1} has $e_n(v; G_2) = t_1$, while each vertex w of the copy of K_{n_2} has $e_n(w, G_2) = t_2$ and, for $3 \leq i \leq m$, $e_n(v_i, G_2) = t_i$. Continue in this fashion to obtain the graph G_m. Then G_m has order p and has S as its n-eccentricity sequence. $\square$

The Steiner number $S(G)$ of a connected graph G was introduced in Definition 4.3.44, as the smallest positive integer m for which there exists a set S of m vertices of G such that $d(S) = p - 1$. We now take a closer look at the characteristics of such sets S for which $d(S) = p - 1$, following [O1].

4.4.7 Definition

If G is a connected graph of order p and S is a set of $S(G)$ vertices such that $d(S) = p - 1$, then S is called a *Steiner spanning set* of G.

The following theorem shows that every connected graph has a unique Steiner spanning set.

4.4.8 Theorem

Let G be a connected graph of order $p \geq 2$. A vertex v of G belongs to a Steiner spanning set of G if and only if v is not a cut-vertex of G.

Proof

Suppose that v is not a cut-vertex of G. Since $G - v$ is connected, we have for all nonempty subsets S of $V(G - v)$ that $d_G(S) \leq p - 2$. Hence v is contained in every Steiner spanning set of G.

Let S be a Steiner spanning set of G. Then $d(S) = p - 1$. Assume, to the contrary, that S contains a cut-vertex u of G. It follows that S does not contain vertices from distinct components of $G - u$, since otherwise if v_1 and v_2 are vertices of S belonging to distinct components of $G - u$, then every connected subgraph of G that contains v_1 and v_2 must also contain the vertex u, which implies that $d(S) = d(S - \{u\}) = p - 1$. This implies however, that $S(G) \leq |S - \{u\}| = |S| - 1$, which contradicts the fact that $S(G) = |S|$. Let w be a vertex of a component of $G - u$ which contains no vertices of S. Then any connected subgraph of minimum size that contains S does not contain w, which implies that $d(S) \leq p-2$, producing a contradiction. Hence S contains no cut-vertices of G. □

4.4.9 Remark

Since the only vertices of a tree T, which are not cut-vertices are the end-vertices, Theorem 4.4.8 implies that the set $S = V_1(T)$ is the unique Steiner spanning set of T. We also note that since there exists an efficient algorithm for determining the cut-vertices of a graph (see [E1]) it follows that there exists an efficient algorithm for determining the Steiner number of a connected graph.

Referring to Definition 4.3.45 we now determine the Steiner sequences of the graphs K_p and C_p as stated in [O1].

4.4.10 Theorem

The complete graph K_p on p vertices has Steiner sequence 1,2,3,...,p, while the cycle C_p has $S_k(C_p) = \lceil \frac{p}{p-k} \rceil$ for $0 \leq k \leq p-1$.

Proof

Since every pair of vertices are adjacent in K_p it follows trivially that $S_k(K_p) = k+1$ for $0 \leq k \leq p-1$.

Let m, p be integers with $2 \leq m \leq p-1$, and let C_p be the cycle $v_1, v_2, ..., v_p, v_1$. Let $S = \{v_{i_1}, v_{i_2}, ..., v_{i_m}\} \subseteq V(C_p)$, where $1 \leq i_1 \leq i_2 \leq ... \leq i_m \leq p$. Let $\ell = \max\{i_2 - i_1, i_3 - i_2, ..., i_m - i_{m-1}, i_1 + p - i_m\}$. Then $d(S) = p - \ell$. Note that $\max\{d(S)\}$ is obtained when ℓ is as small as possible, and this occurs when $\ell = \lceil \frac{p}{m} \rceil$. Hence

$$\max\{d(S)\} = p - \lceil \frac{p}{m} \rceil.$$

Now given the integer k with $1 \leq k \leq p-1$, we seek the smallest value for m (i.e., $S_k(C_p)$) for which $k = p - \lceil \frac{p}{m} \rceil$ and $p - \lceil \frac{p}{m-1} \rceil < k$. Hence we seek the smallest m such that $\lceil \frac{p}{m} \rceil = p - k$; i.e.,

$$p - k \geq \frac{p}{m} \qquad \sim (1)$$

and $\frac{p}{m-1} > p - k$ which implies that

$$m - 1 < \frac{p}{p-k}. \qquad \sim (2)$$

Thus from (1) $m \geq \frac{p}{p-k}$, while from (2) $m < \frac{p}{p-k} + 1$ and, since m is an integer,

$$m = \lceil \frac{p}{p-k} \rceil. \qquad \square$$

In [CJO1] the periphery $P(G)$ of a graph G was defined as the subgraph of G induced by those vertices whose eccentricity in G equals diam G. The following characterization of peripheries, established in [CJO1], we state without proof.

4.4.11 Theorem

A graph G of order p is the periphery of a graph if and only if $G \cong K_p$ or the maximum degree $\Delta(G)$ of G is less than $p-1$.

Note that if G is a graph of order p, then $G \cong K_p$ if and only if diam $G = 1$. Further, $\Delta(G) < p - 1$ if and only if rad $G \geq 2$. Thus, by Theorem 4.4.11, a graph of order p is the periphery of a graph if and only if diam $G = 1$ or rad $G \geq 2$.

The following generalization of the periphery of a graph was introduced by Henning, Oellermann and Swart in [HOS2]. We follow [HOS2] up until and including Theorem 4.4.22.

4.4.12 Definition

Let $n \geq 2$ be an integer. Then the *n-periphery* $P_n(G)$ of a graph G is the subgraph of G induced by those vertices v of G with $e_n(v, G) = \text{diam}_n G$.

Hence $P(G) = P_2(G)$, that is, the 2-periphery of a graph is its periphery. We now consider a generalization of Theorem 4.4.11.

4.4.13 Theorem

Let $n \geq 2$ be an integer. A graph G of order $p \geq n$ is the n-periphery of a graph if and only if $\text{diam}_n G = n - 1$ or $\text{rad}_n G \geq n$.

Proof

Suppose G is the n-periphery of some graph H and that $\text{rad}_n G < n$. Then there exists a vertex $v \in V(G)$ such that $e_n(v, G) = \text{rad}_n G < n$. This implies that $e_n(v, G) = n-1$, since for a set S' of n vertices containing v such that $d(S') = e_n(v, G)$, we must have that $T_{S'}$ is connected and hence $d(S') = q(T_{S'}) = n - 1$. Thus, if S is any set of n vertices of G that contains v, then the subgraph $\langle S \rangle$ induced by S is connected. Let D be a set of n-vertices of

H containing v such that $d_H(D) = e_n(v, H) = \text{diam}_n H$. Then every vertex u of D has $e_n(u, H) = \text{diam}_n H$. So $D \subseteq V(P_n(H)) = V(G)$. Thus since $D \subseteq V(G)$ contains v we have that $\langle D \rangle$ is a connected subgraph of G and thus of H. However, then $e_n(v, H) = d_H(D) = n - 1 = \text{diam}_n H$. Therefore the subgraph induced by every set of n vertices in H, and therefore in G, is connected. So $\text{diam}_n G = n - 1$.

For the converse, suppose first that $\text{diam}_n G = n - 1$. Then every set of n vertices of G induces a connected subgraph and every vertex $v \in V(G)$ has $e_n(v, G) = \text{diam}_n G$. Let $H = G$. Then necessarily $P_n(H) = G$. Suppose now that $\text{rad}_n G \geq n$. Let $H = G + K_1$ and suppose that v is the vertex of degree p in H. Then $e_n(v, H) = n - 1$. Let $u \in V(G)$. Since $\text{rad}_n G \geq n$, there exists a set of n vertices of G, containing u, that induces a disconnected subgraph. However, if S is any set of n vertices of G, containing u, then $\langle S \cup \{v\} \rangle$ is connected, and has size n. Therefore $e_n(u, H) = n$. Thus $G = P_n(H)$. □

4.4.14 Definition

A graph G of order $p \geq n \geq 2$ is self n-centred if $\text{rad}_n G = \text{diam}_n G$.

4.4.15 Remark

Self 2-centred graphs are also called self-centred graphs. Jordan [J3] showed that the only self-centred trees are K_1 and K_2. By Theorem 4.3.26 we have that a tree T is self n-centred, $n \geq 3$, if and only if T is a tree of order $p \geq n \geq 3$ with at most $n - 1$ end-vertices. Thus if T is a tree of order $p \geq n \geq 3$ with at most $n - 1$ end-vertices, then T is its own n-periphery.

We now characterize those graphs which are n-peripheries of trees for $n \geq 2$. Note that the n-periphery of a tree is acyclic and is thus a forest.

4.4.16 Theorem

Let F be a forest and $n \geq 2$ an integer. Then F is the n-periphery of a tree T if and only if

1) $n = 2$ and $F \cong K_1$ or K_2; or

2) $n \geq 3$ and F is a tree with at most $n - 1$ end-vertices; or

3) $F \cong \overline{K_m}$ for some $m \geq n$.

Proof

Suppose F is the n-periphery of a tree T. Suppose that neither 1) nor 2) holds. We show $F \cong \overline{K_m}$. Since the n-periphery of a tree of order $p \geq n$ contains at least n vertices (by Corollary 4.3.8, and since each vertex of an n-diametral set of T has eccentricity equal to $\text{diam}_n T$), it follows that F has order at least n. It remains to be shown that F contains no edges. Let S be any set of n vertices of T such that $d(S) = \text{diam}_n T$. Then $S \subseteq V(F)$. Since 2) does not hold, we have, by Theorem 4.3.7, for every $v \in S$, that $S - \{v\}$ consists of end-vertices of T. Similarly for any vertex $u \in S$ where $u \neq v$ we have that all vertices of $S - \{u\}$ are end-vertices of T and hence S consists of only end-vertices of T. Since every vertex of F belongs to some set S of n vertices of T for which $d(S) = \text{diam}_n T$, every vertex of F must be an end-vertex of T. Since 1) does not hold, $T \not\cong K_2$. Therefore no two end-vertices of T are adjacent. Hence F contains no edges and so $F \cong \overline{K_m}$ for some $m \geq n$.

For the converse, suppose first that 1) or 2) holds. In either case let $T = F$. From Remark 4.4.15 it follows that $P_n(T) = F$, so that F is the n-periphery of a tree in this case. Suppose now that F satisfies 3). Then let $T \cong K_{1,m}$, label the end-vertices of T, $v_1, v_2, ..., v_m$ and label the vertex of degree m in T, u. Since $m \geq n$ for each v_i $(1 \leq i \leq m)$ there exists a set S_i consisting of v_i and $n-1$ other end-vertices of T such that $d(S_i) = n = \text{diam}_n T$. Therefore for $1 \leq i \leq m$, $e_n(v_i, T) = \text{diam}_n T$ while $e_n(u, T) = n - 1$. Thus $P_n(T) \cong \overline{K_m} = F$ and F is the n-periphery of the tree $K_{1,m}$. □

We saw in Theorem 4.3.32 that for a tree T, $C_{n-1}(T) \subseteq C_n(T)$. We now consider for an integer $n \geq 3$, relationships between the $(n-1)$-centre and n-centre of a graph as well as relationships between the $(n-1)$-periphery and n- periphery of a graph. The following result demonstrates that the $(n-1)$-centre of a graph is not in general contained in the n- centre of that graph.

4.4.17 Theorem

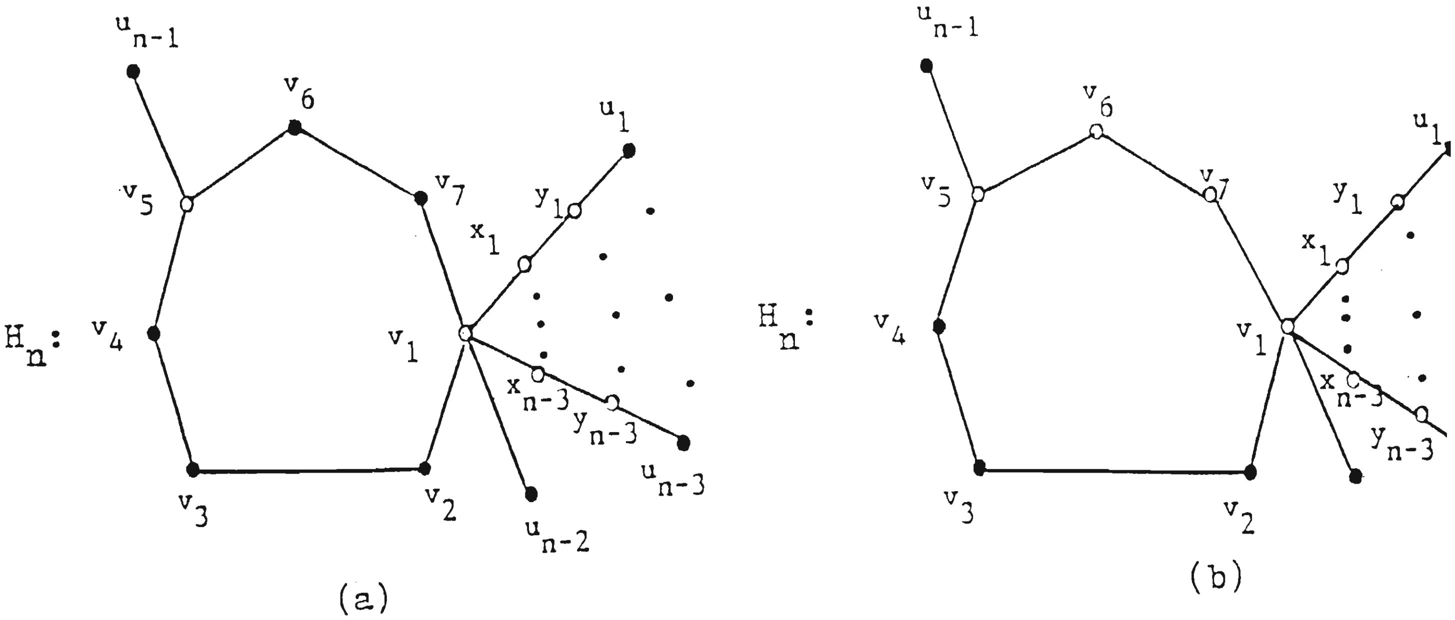
For every integer $n \geq 4$, there exists a graph H_n such that

$$C_{n-1}(H_n) \not\subseteq C_n(H_n).$$

Proof

Consider the 7-cycle $C = v_1v_2v_3v_4v_5v_6v_7v_1$. Add $n-1$ new vertices $u_1, u_2, ..., u_{n-1}$ to C and join $u_1, u_2, ..., u_{n-2}$ to v_1 and u_{n-1} to v_5. For $i = 1, 2, ..., n-3$ subdivide the edge v_1u_i twice and let v_1, x_i, y_i, u_i be the path thus produced. Let H_n be the resulting graph.

Observe that $\text{diam}_n H_n = 3(n-3)+2+2+2 = 3n-3$ and $\text{rad}_n H_n = 3(n-3)+4+1 = 3n-4$. The vertices with n-eccentricity equal to $\text{diam}_n H_n$ are $u_1, u_2, ..., u_{n-1}, v_2, v_3, v_4, v_6, v_7$. All the remaining vertices of H_n have n-eccentricity $3n-4$. Figure 4.4.18 a) shows H_n where vertices of maximum n-eccentricity are darkened. Note also that $\text{diam}_{n-1} H_n = 3(n-3)+4+1 = 3n-4$ and $\text{rad}_{n-1} H_n = 3(n-3)+4 = 3n-5$. The vertices of H_n with $(n-1)$-eccentricity equal to $\text{diam}_{n-1} H_n$ are $u_1, u_2, ..., u_{n-1}, v_2, v_3, v_4$. All the remaining vertices have $(n-1)$-eccentricity $3n-5$. Figure 4.4.18 b) shows H_n and vertices of maximum $(n-1)$-eccentricity have been darkened. Hence $C_{n-1}(H_n) \not\subset C_n(H_n)$. $\square$



4.4.18 Figure

The graph H_n of Theorem 4.4.17

The next result shows that in general the $(n-1)$-periphery of a graph is not contained in its n-periphery where $n \geq 3$.

4.4.19 Theorem

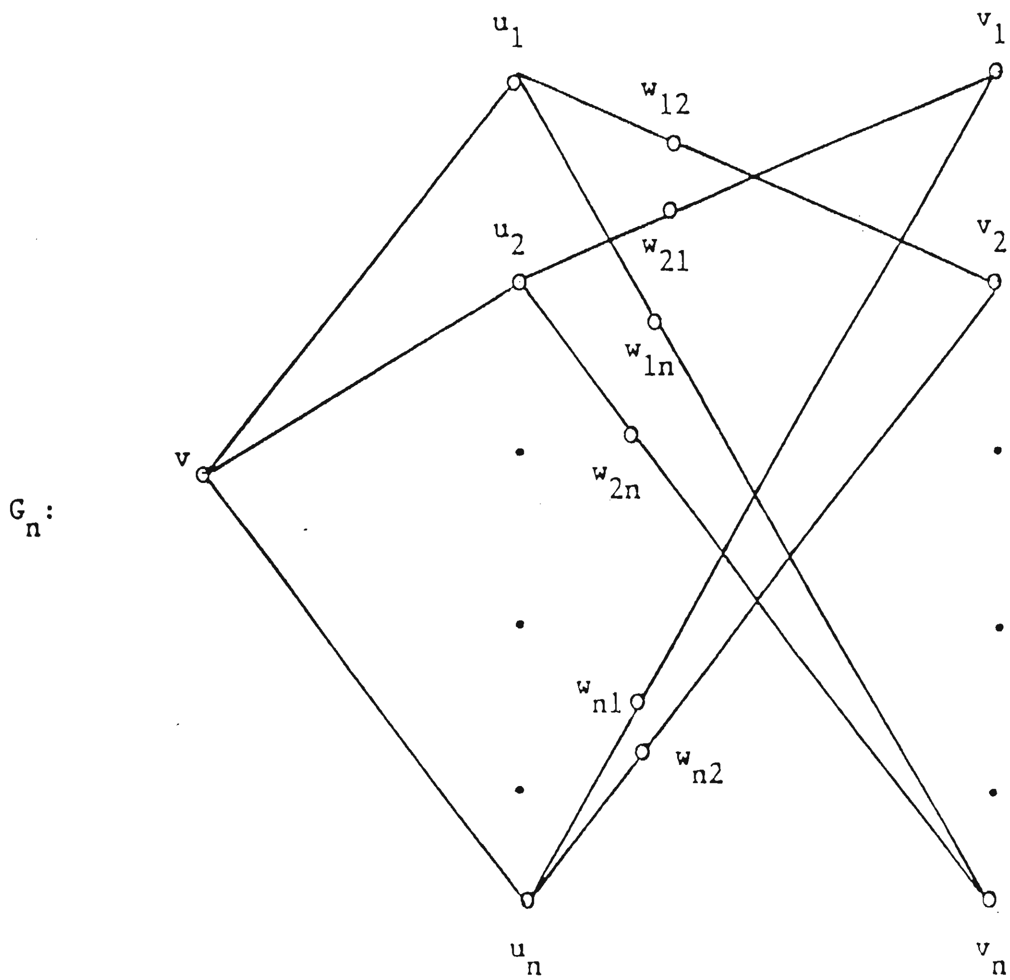
For every integer $n \geq 3$ there exists a graph G_n such that $P_{n-1}(G_n) \not\subset P_n(G_n)$.

Proof

Consider the complete bipartite graph $K_{n,n}$ with partite sets $U = \{u_0, u_1, ..., u_{n-1}\}$ and $V = \{v_0, v_1, ..., v_{n-1}\}$. Join a new vertex v to every vertex in U and delete the edges of the type $u_i v_i$ for $0 \leq i \leq n-1$. Finally subdivide each edge of the type $u_i v_j$ for $0 \leq i, j \leq n-1$ with $i \neq j$ exactly once and let w_{ij} be the vertex of degree 2 that is adjacent with u_i and v_j. Let G_n be the resulting graph. Then $e_n(v, G_n) = d(\{v\} \cup \{v_0, v_1, ..., v_{n-2}\}) = 1 + 2(n-1) = 2n-1$, $e_n(v_i, G_n) = d(V) = 2(n-1)+4 = 2n+2$, $e_n(u_i, G_n) = d(\{u_i\} \cup \{v_i, v_{i+1}, ..., v_{i+n-2}\}) = 4 + 2(n-2) = 2n$ for $0 \leq i \leq n-1$ where addition of indices is taken modulo n and $e_n(w_{ij}, G_n) = d(\{w_{ij}\} \cup \{v_i, v_{i+1}, ..., v_{i+n-2}\}) = 5 + 2(n-2) = 2n+1$ for $0 \leq i, j \leq n-1$ and $i \neq j$. Further, $e_{n-1}(v, G_n) = d(\{v\} \cup \{v_0, v_1, ..., v_{n-3}\}) = 1 + 2(n-2) = 2n-3$, $e_{n-1}(v_i, G_n) = d(\{v_i\} \cup \{w_{ij}\} \cup \{v_{i+1}, ..., v_{i+n-3}\}) = 5 + 2(n-2) = 2n-1$ for $j \neq i$ where again addition of indices is taken modulo n, $e_{n-1}(u_i, G_n) = d(\{u_i\} \cup \{v_i\} \cup \{v_{i+1}, ..., v_{i+n-3}\}) = 4 + 2(n-3) = 2n-2$ for $0 \leq i \leq n-1$ and $e_{n-1}(w_{ij}, G_n) = d(\{w_{ij}\} \cup \{v_i\} \cup \{v_{i+1}, ..., v_{i+n-3}\}) = 5 + 2(n-3) = 2n-1$ for $0 \leq i, j \leq n-1$ where $i \neq j$. Therefore $P_{n-1}(G_n) = \langle \{V \cup \{w_{ij}\}\} \rangle$ for $0 \leq i, j \leq n-1$ where $i \neq j$ while $P_n(G_n) = \langle \{V\} \rangle$, and hence it follows that

$$P_{n-1}(G_n) \not\subset P_n(G_n).$$

□



4.4.20 Figure

The graph G_n of Theorem 4.4.19.

However, if T is a tree, then the next result shows that there is a relationship between the $(n-1)$-periphery and n-periphery of T.

4.4.21 Theorem

Let T be a tree and $n \geq 3$ an integer. Then $P_{n-1}(T) \subset P_n(T)$.

Proof

If T has at most $n-1$ end-vertices, then by Corollary 4.3.6, $P_n(T) = T$, so the result follows in this case. Suppose now that T has at least n end-vertices. Let S_{n-1} be an $(n-1)$-diameter set. By Theorem 4.3.18, an n-diameter set S_n containing S_{n-1} can be obtained from S_{n-1}. Since the union of all $(n-1)$-diameter sets (or n-diameter sets) is the vertex set of the $(n-1)$-periphery (or n-periphery) of T it follows that

$$P_{n-1}(T) \subset P_n(T). \qquad \square$$

We showed in Theorem 4.3.13 that if $n \geq 3$ is an integer and T is a tree, then $\text{diam}_n T \leq \frac{n}{n-1}\text{diam}_{n-1} T$. However, this inequality does not hold for graphs in general. For example, for the graph G_n described in Theorem 4.4.19, $\text{diam}_n G_n = 2n+2$ while $\text{diam}_{n-1} G_n = 2n-1$. So, in this case, $\text{diam}_n G_n = \frac{2n+2}{2n-1}\text{diam}_{n-1} G_n > \frac{n}{n-1}\text{diam}_n G_n$. However a bound for the n-diameter of a graph in terms of its $(n-1)$-diameter was established in [HOS2].

4.4.22 Theorem

Let G be a connected graph and $n \geq 3$ an integer. Then

$$\text{diam}_n G \leq \frac{n+1}{n-1}\text{diam}_{n-1} G.$$

Proof

Let $S = \{v_1, v_2, ..., v_n\}$ be an n-diameter set of G. For $1 \leq i \leq n$, let $S_i = S - \{v_i\}$ and let T_i be a tree of minimum size containing S_i. Construct an Eulerian multigraph H by duplicating every edge of T_n. Observe that $q(T_i) \leq \mathrm{diam}_{n-1}G$ for $1 \leq i \leq n$ and $q(H) \leq 2\mathrm{diam}_{n-1}G$.

We now construct $n-1$ connected subgraphs $G_1, G_2, ..., G_{n-1}$ (from $T_1, T_2, ..., T_{n-1}$ and H) each of which contains the vertices of S. Let C be an Eulerian $v-v$ circuit of H. Let $v_{i_1}, v_{i_2}, ..., v_{i_{n-1}}$ be the vertices of S_n in the order in which they appear on C for the first time. For $j = 1, 2, ..., n-2$, let R_j be a $v_{i_j} - v_{i_{j+1}}$ trail of C between the first appearance of v_{i_j} and the first appearance of $v_{i_{j+1}}$. Let R_{n-1} be the $v_{i_{n-1}} - v_{i_1}$ trail of C between the first appearance of $v_{i_{n-1}}$ and the first appearance of v_{i_1}. Since T_{i_j} contains $v_{i_{j+1}}$, the edges of T_{i_j} and R_j induce a connected graph G_j for $1 \leq j \leq n-2$. Further, since v_{i_1} is a vertex of $T_{i_{n-1}}$, the edges of $T_{i_{n-1}}$ and R_{n-1} also induce a connected graph G_{n-1}.

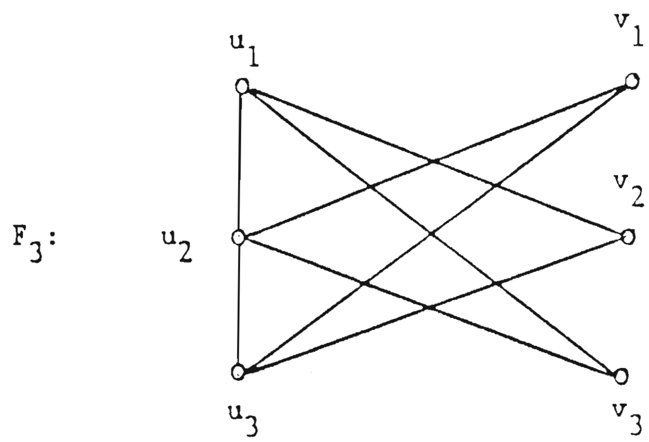
Note that each one of the connected graphs $G_j (1 \leq j \leq n-1)$ contains the vertices of S. Hence $\mathrm{diam}_n G \leq q(G_j)$ for $1 \leq j \leq n-1$. Therefore

$$\begin{aligned}
(n-1)\mathrm{diam}_n G &\leq \sum_{j=1}^{n-1} q(G_j) \\
&\leq \sum_{i=1}^{n-1} q(T_i) + 2q(T_n) = \sum_{i=1}^{n-1} q(T_i) + q(H) \\
&\leq (n-1)\mathrm{diam}_{n-1}G + 2\mathrm{diam}_{n-1}G = (n+1)\mathrm{diam}_{n-1}G,
\end{aligned}$$

so that $\mathrm{diam}_n G \leq \frac{n+1}{n-1}\mathrm{diam}_{n-1}G$. □

The authors of [HOS2] showed that the bound presented in Theorem

4.4.22 is sharp. Consider for $n \geq 4$ the complete bipartite graph $K_{n,n}$ with partite sets $U = \{u_1, u_2, ..., u_n\}$ and $V = \{v_1, v_2, ..., v_n\}$. Let F_n be the graph obtained from $K_{n,n}$ by deleting edges of the type $u_i v_i$ for $1 \leq i \leq n$. Observe that $\text{diam}_n F_n = d(V) = n - 1 + 2 = n + 1$ and that $\text{diam}_{n-1} F_n = d(V - v_1) = n - 1$. Hence $\text{diam}_n F_n = \frac{n+1}{n-1} \text{diam}_{n-1} F_n$. For $n = 3$, the graph F_3 of Figure 4.4.23 has $\text{diam}_3 F_3 = d(\{v_1, v_2, u_1\}) = 4$ and $\text{diam}_2 F_3 = 2$, hence $\text{diam}_3 F_3 = 2\text{diam}_2 F_3$.



4.4.23 Figure

It was conjectured in [COTZ1] that Corollary 4.3.15 can be extended to all connected graphs; i.e., if $n \geq 3$ is an integer and G is a connected graph, then $\text{diam}_n G \leq \frac{n}{n-1} \text{rad}_n G$. However Henning, Oellermann and Swart disproved this conjecture in [HOS1] with the following result.

4.4.24 Theorem

Let $n \geq 3$ be an integer. Then there exists a graph G_n such that

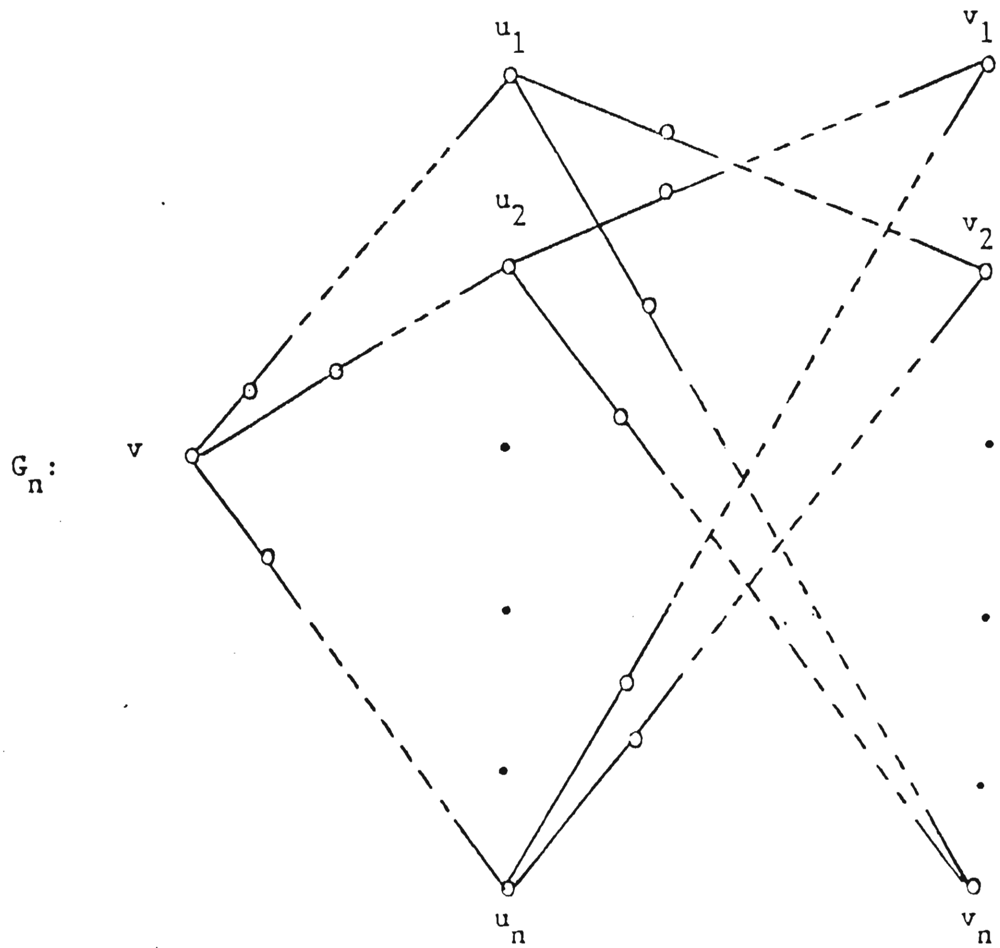
$$\text{diam}_n G_n = \frac{2(n+1)}{2n-1} \text{rad}_n G_n.$$

Proof

Let H be the complete bipartite graph $K_{n,n}$ with partite sets $U = \{u_1, u_2, ..., u_n\}$ and $V = \{v_1, v_2, ..., v_n\}$. Let H_n be obtained from $H - \{u_i v_i : 1 \leq i \leq n\}$ by joining a new vertex v to every vertex of U. Let k be a positive integer. Let G_n be obtained from H_n by subdividing $k-1$ times, every edge of the type vu_i for $1 \leq i \leq n$ and subdividing $2k-1$ times, every edge of the type $u_i v_j$ for $1 \leq i, j \leq n$ and $i \neq j$.

Then $\text{diam}_n G_n = d(V) = 2k(n-1)+4k = 2k(n+1)$. Further, $\text{rad}_n G_n = d(\{v\} \cup (V - \{v_1\}) = 2k(n-1) + k = k(2n-1)$. Hence

$$\text{diam}_n G_n = \frac{2(n+1)}{2n-1} \text{rad}_n G_n. \qquad \square$$



4.4.25 Figure

The graph G_n of Theorem 4.4.24.

It is immediately evident from Theorem 4.4.24 that there exists a graph G such that $\mathrm{diam}_3 G = \frac{8}{5}\mathrm{rad}_3 G$. The following result from [HOS1]shows that the 3-diameter of a graph never exceeds $\frac{8}{5}$ths its 3-radius.

4.4.26 Theorem

If G is a connected graph of order $p \geq 3$, then

$$\mathrm{diam}_3 G \leq \frac{8}{5}\mathrm{rad}_3 G.$$

Proof

Assume, to the contrary, that there exists a connected graph G such that $\mathrm{diam}_3 G > \frac{8}{5}\mathrm{rad}_3 G$. Let v_1, v_2 and v_3 be three vertices of G such that $d(\{v_1, v_2, v_3\}) = \mathrm{diam}_3 G$, and let $v_0 \in C_3(G)$.

Then $d(\{v_i, v_j, v_0\}) \leq \mathrm{rad}_3 G$ for $1 \leq i < j \leq 3$. Let T_i be a Steiner tree for $\{v_0, v_1, v_2, v_3\} - \{v_i\}$ for $i = 1, 2, 3$. Since $q(T_i) \leq \mathrm{rad}_3 G$, it follows that $q(T_i) < \frac{5}{8}\mathrm{diam}_3 G$. The tree T_i together with a shortest path from vertex v_i to a vertex of T_i contains the vertices v_1, v_2 and v_3, therefore $\mathrm{diam}_3 G \leq q(T_i) + d(v_i, T_i) < \frac{5}{8}\mathrm{diam}_3 G + d(v_i, T_i)$. Thus the shortest distance from v_i to every vertex of T_i must be greater than $\frac{3}{8}\mathrm{diam}_3 G$ for $i = 1, 2, 3$. In particular $d(v_k, v_j) > \frac{3}{8}\mathrm{diam}_3 G$ for $0 \leq k < j \leq 3$.

Note that T_i cannot be a path; otherwise $q(T_i) > \frac{3}{8}\mathrm{diam}_3 G + \frac{3}{8}\mathrm{diam}_3 G = \frac{6}{8}\mathrm{diam}_3 G$ which contradicts our assumption. Hence T_i has exactly three end-vertices. Let P_1 be the $v_2 - v_3$ path in T_1, P_2 the $v_1 - v_3$ path in T_2 and P_3 the $v_1 - v_2$ path in T_3. Then at least two of the paths P_1, P_2 and P_3 have size at least $\frac{1}{2}\mathrm{diam}_3 G$, otherwise if say P_2 and P_3 were both of size less than $\frac{1}{2}\mathrm{diam}_3 G$ then we could find a Steiner tree T' for $\{v_1, v_2, v_3\}$ with size less than $\frac{1}{2}\mathrm{diam}_3 G + \frac{1}{2}\mathrm{diam}_3 G = \mathrm{diam}_3 G$, which is a contradiction. Suppose P_2 and P_3 are such paths. Let $\ell_2 = d_{T_2}(v_0, P_2)$ and $\ell_3 = d_{T_3}(v_0, P_3)$. Then

$$\ell_2 = d_{T_2}(v_0, P_2) \quad = \quad q(T_2) - q(P_2)$$

$$< \frac{5}{8}\text{diam}_3G - \frac{1}{2}\text{diam}_3G = \frac{1}{8}\text{diam}_3G.$$

Similarly $\ell_3 < \frac{1}{8}\text{diam}_3G$.

Let Q_i be the $v_0 - v_i$ path in T_2 for $i = 1, 3$. Then $q(T_i) + q(Q_i) \geq \text{diam}_3G$ for $i = 1, 3$. Hence $\sum_{i=1}^{3} q(T_i) + \ell_2 = q(T_1) + q(T_2) + q(T_3) + \ell_2 = q(T_1) + q(T_3) + q(Q_1) + q(Q_3) \geq 2\text{diam}_3G$. Thus

$$3\text{rad}_3G + \ell_2 \geq 2\text{diam}_3G, \text{ so that}$$

$$\ell_2 \geq 2\text{diam}_3G - 3\text{rad}_3G > 2\text{diam}_3G - \frac{15}{8}\text{diam}_3G = \frac{1}{8}\text{diam}_3G.$$

This contradiction establishes the theorem. □

From Theorem 4.4.24 we also know that there exists a graph such that $\text{diam}_4G = \frac{10}{7}\text{rad}_4G$. Following [HOS1] we show next that the 4-diameter of every connected graph is bounded above by $\frac{10}{7}$ths its 4-radius.

4.4.27 Theorem

If G is a connected graph of order $p \geq 4$, then

$$\text{diam}_4G \leq \frac{10}{7}\text{rad}_4G.$$

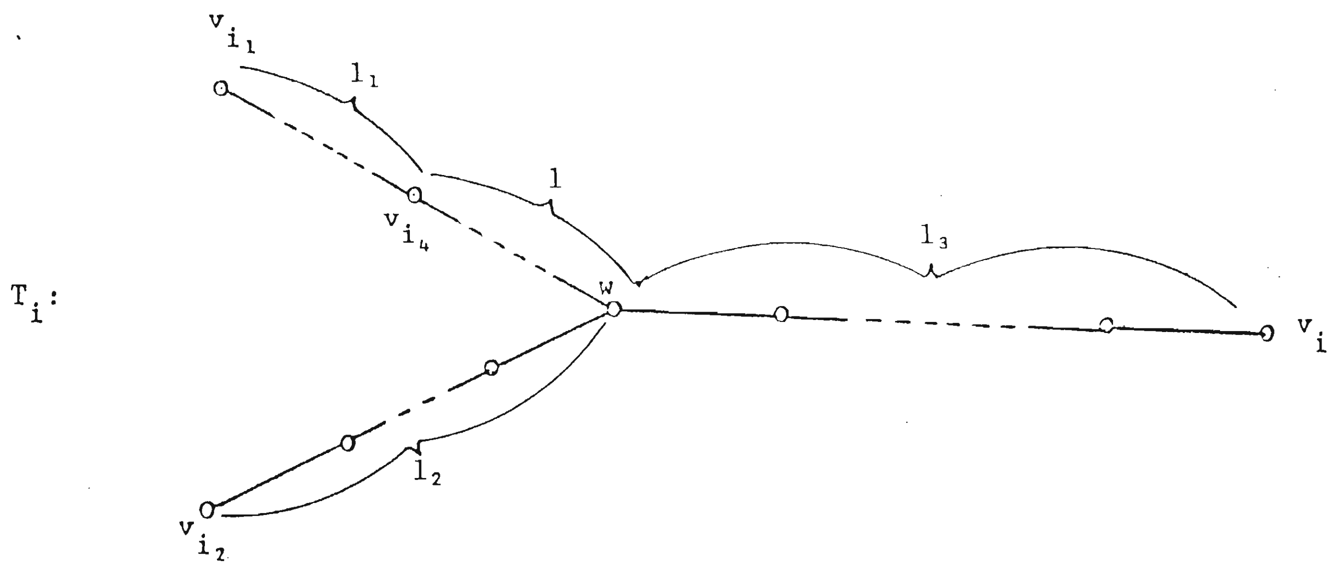
Proof

Assume, to the contrary, that there exists a graph G of order at least 4, for which $\text{diam}_4G > \frac{10}{7}\text{rad}_4G$. Let $D = \{v_1, v_2, v_3, v_4\}$ be a 4-diameter set; i.e., $d(D) = \text{diam}_4G$. Let $v_0 \in C_4(G)$. For each i with $1 \leq i \leq 4$, let T_i be a Steiner tree for $D_i = (D - \{v_i\}) \cup \{v_0\}$. Then since T_i contains $v_0 \in C_4(G)$ it follows that $q(T_i) \le \text{rad}_4 G < \frac{7}{10}\text{diam}_4 G$. Taking T_i, together with a shortest path from v_i to T_i, we obtain a tree containing D and hence of size at least $\text{diam}_4 G$. Therefore $\text{diam}_4 G \le q(T_i) + d(v_i, T_i) < \frac{7}{10}\text{diam}_4 G + d(v_i, T_i)$, hence $d(v_i, T_i) > \text{diam}_4 G - \frac{7}{10}\text{diam}_4 G = \frac{3}{10}\text{diam}_4 G$. In particular this implies that

$$d(v_i, v_j) > \frac{3}{10}\text{diam}_4 G \qquad \sim (1)$$

for $0 \le i < j \le 4$.

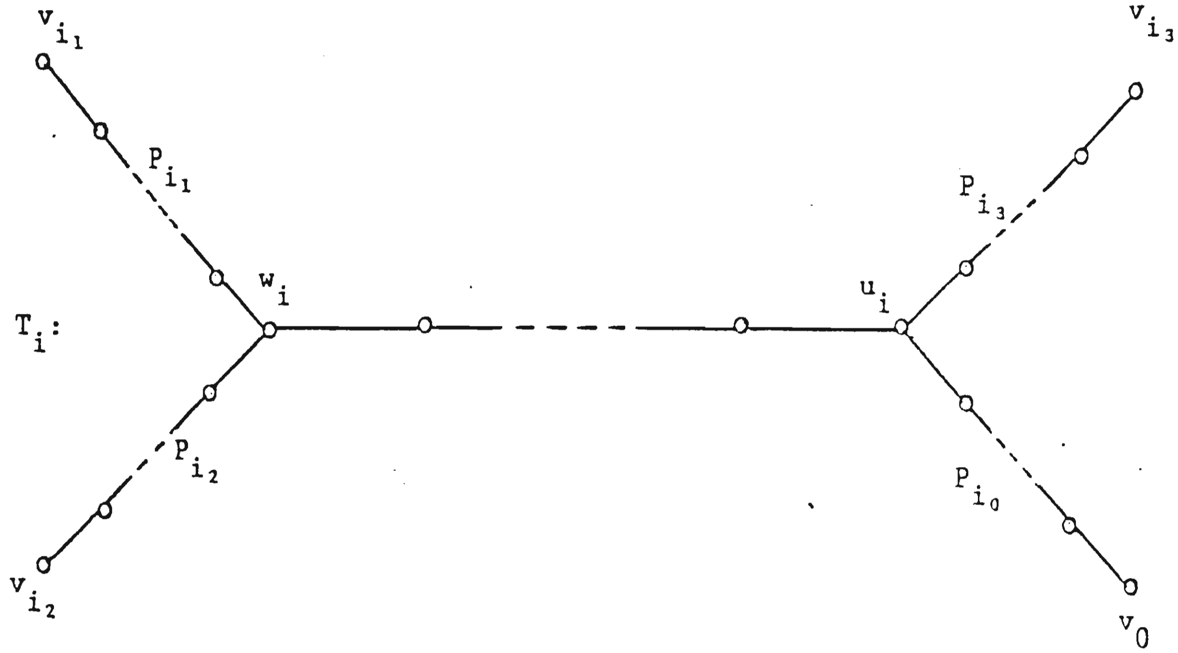
We show next that every $T_i (1 \le i \le 4)$ has exactly four end-vertices, namely, the vertices in D_i. Observe first that T_i is not a path, otherwise by (1), $q(T_i) > 3(\frac{3}{10}\text{diam}_4 G) = \frac{9}{10}\text{diam}_4 G$, which contradicts our assumption. Suppose now that some T_i has exactly three end-vertices v_{i_1}, v_{i_2} and v_{i_3}, each of which is necessarily in D_i. Let v_{i_4} be the remaining vertex of D_i in T_i and let w be the vertex of degree 3 in T_i. We may assume that v_{i_4} lies on the $v_{i_1} - w$ path of T_i. Let ℓ_1 be the length of the $v_{i_1} - v_{i_4}$ path of T_i, and ℓ the length of the $v_{i_4} - w$ path. Further let ℓ_2 and ℓ_3 be the lengths of the $v_{i_2} - w$ and $v_{i_3} - w$ paths, respectively (see Figure 4.4.28).



4.4.28 Figure

Since by (1), $d(v_i, v_j) > \frac{3}{10}\text{diam}_4 G$ for $0 \leq i < j \leq 4$, it follows that each of $\ell_2 + \ell_3, \ell + \ell_2$ and $\ell + \ell_3$ is greater than $\frac{3}{10}\text{diam}_4 G$. Hence $\ell + \ell_2 + \ell + \ell_3 + \ell_2 + \ell_3 = 2(\ell + \ell_2 + \ell_3) > \frac{9}{10}\text{diam}_4 G$, or equivalently $\ell + \ell_2 + \ell_3 > \frac{9}{20}\text{diam}_4 G$. Since by (1), $\ell_1 > \frac{3}{10}\text{diam}_4 G$ we now have $q(T_i) = \ell + \ell_1 + \ell_2 + \ell_3 > \frac{3}{10}\text{diam}_4 G + \frac{9}{20}\text{diam}_4 G = \frac{15}{20}\text{diam}_4 G > \frac{7}{10}\text{diam}_4 G$, which produces a contradiction. Hence T_i has exactly four end-vertices, namely, the vertices in $D_i (1 \leq i \leq 4)$.

Suppose v_0, v_{i_1}, v_{i_2} and v_{i_3} are the end-vertices of T_i. Let P_{i_j} be the shortest path from v_{i_j} to a vertex of degree at least 3 in T_i for $1 \leq j \leq 3$. Further, let P_{i_0} be a shortest path from v_0 to a vertex of degree at least 3 in T_i. We may assume that $P_{i_1}(P_{i_2})$ is a $v_{i_1} - w_i$ path ($v_{i_2} - w_i$ path, respectively) and that $P_{i_3}(P_{i_0})$ is a $v_{i_3} - u_i$ path ($v_0 - u_i$ path, respectively). It is possible that $v_i = w_i$. Let P_i be the $u_i - w_i$ path in T_i. For $i = 1, 2, 3, 4$ and $j = 0, 1, 2, 3$ let $q(P_{i_j}) = \ell_{i_j}$. Further, let $q(P_i) = \ell_i$, and observe that ℓ_i could be zero. T_i is illustrated in Figure 4.4.29.



4.4.29 Figure

From (1) it follows that each of $\ell_{i_1}+\ell_{i_2}, \ell_{i_1}+\ell_i+\ell_{i_3}$ and $\ell_{i_2}+\ell_i+\ell_{i_3}$ is greater than $\frac{3}{10}\mathrm{diam}_4 G$. Thus $2(\ell_{i_1}+\ell_{i_2}+\ell_{i_3}+\ell_i) > \frac{9}{10}\mathrm{diam}_4 G$, or equivalently, $\ell_{i_1}+\ell_{i_2}+\ell_{i_3}+\ell_i > \frac{9}{20}\mathrm{diam}_4 G$. Since $q(T_i) < \frac{7}{10}\mathrm{diam}_4 G$, we conclude that $\ell_{i_0} = q(T_i) - (\ell_{i_1}+\ell_{i_2}+\ell_{i_3}+\ell_i) < (\frac{7}{10}-\frac{9}{20})\mathrm{diam}_4 G = \frac{\mathrm{diam}_4 G}{4}$. Further by (1), $\ell_{i_3}+\ell_{i_0} > \frac{3}{10}\mathrm{diam}_4 G$ and so $\ell_{i_3} > (\frac{3}{10}-\frac{1}{4})\mathrm{diam}_4 G = \frac{\mathrm{diam}_4 G}{20}$. Now, interchanging the roles of ℓ_{i_0} and ℓ_{i_3} in the above argument, we obtain $\ell_{i_3} < \frac{\mathrm{diam}_4 G}{4}$ and $\ell_{i_0} > \frac{\mathrm{diam}_4 G}{20}$. Hence for $i = 1,2,3,4$ we have

$$\begin{aligned} \frac{\mathrm{diam}_4 G}{20} < \ell_{i_0} < \frac{\mathrm{diam}_4 G}{4}, \qquad & \text{and} \\ \frac{\mathrm{diam}_4 G}{20} < \ell_{i_3} < \frac{\mathrm{diam}_4 G}{4}. & \end{aligned} \qquad \sim (2)$$

For $i = 1,2,3,4$ let T_i' be the tree obtained from T_i by deleting all the vertices of P_{i_0} except u_i and let T_i'' be obtained from T_i by deleting all the vertices of P_{i_3} except u_i. Observe from (2) that

$$q(T_i'') = q(T_i) - \ell_{i_3} < \left(\frac{7}{10} - \frac{1}{20}\right) = \frac{13}{20}\mathrm{diam}_4 G. \qquad \sim (3)$$

We now consider two cases.

Case 1) Suppose that for every $i = 1, 2, 3, 4$ we have $\ell_{i_0} \geq \frac{\mathrm{diam}_4 G}{10}$. Then

$$q(T_i') = q(T_i) - \ell_{i_0} < \left(\frac{7}{10} - \frac{1}{10}\right) \mathrm{diam}_4 G = \frac{6}{10}\mathrm{diam}_4 G. \qquad \sim (4)$$

Let i be some fixed element of $\{1, 2, 3, 4\}$. Observe that T_{i_1}' together with the $v_{i_1} - v_{i_2}$ paths of T_i produces a connected graph containing the vertices of D. Hence

$$q(T_{i_1}') + \ell_{i_1} + \ell_{i_2} \geq \mathrm{diam}_4 G. \qquad \sim (5)$$

Similarly it follows that

$$q(T_{i_1}') + \ell_{i_1} + \ell_i + \ell_{i_3} \geq \mathrm{diam}_4 G \qquad \sim (6)$$

and

$$q(T_{i_2}') + \ell_{i_2} + \ell_i + \ell_{i_3} \geq \mathrm{diam}_4 G. \qquad \sim (7)$$

From (5), (6) and (7) we obtain

$$2q(T_{i_1}') + q(T_{i_2}') + 2(\ell_i + \ell_{i_1} + \ell_{i_2} + \ell_{i_3}) = 2q(T_{i_1}') + q(T_{i_2}') + 2q(T_i') \geq 3\mathrm{diam}_4 G$$

while from (4) we obtain

$$5(\frac{6}{10}\mathrm{diam}_4 G) = 3\mathrm{diam}_4 G > 2q(T_{i_1}') + q(T_{i_2}') + 2q(T_i') \geq 3\mathrm{diam}_4 G,$$

which produces a contradiction.

Case 2) Suppose that for some $i \in \{1,2,3,4\}$ we have $\ell_{i_0} < \frac{\mathrm{diam}_4 G}{10}$. Since $\ell_{i_0} + \ell_{i_3} > \frac{3}{10}\mathrm{diam}_4 G$, $\ell_{i_3} > \left(\frac{3}{10} - \frac{1}{10}\right)\mathrm{diam}_4 G = \frac{\mathrm{diam}_4 G}{5}$. Hence

$$q(T_i'') = q(T_i) - \ell_{i_3} < \left(\frac{7}{10} - \frac{2}{10}\right)\mathrm{diam}_4 G = \frac{\mathrm{diam}_4 G}{2}. \qquad \sim (8)$$

As in Case 1

$$\begin{aligned} q(T_{i_1}'') + \ell_{i_1} + \ell_{i_2} &\geq \mathrm{diam}_4 G, && \sim (9)\\ q(T_{i_1}'') + \ell_{i_1} + \ell_i + \ell_{i_0} &\geq \mathrm{diam}_4 G, && \sim (10)\\ \text{and} \quad q(T_{i_2}'') + \ell_{i_2} + \ell_i + \ell_{i_0} &\geq \mathrm{diam}_4 G. && \sim (11) \end{aligned}$$

Thus by (9), (10) and (11) we have

$$2q(T_{i_1}'') + q(T_{i_2}'') + 2(\ell_i + \ell_{i_1} + \ell_{i_2} + \ell_{i_0}) = 2q(T_i'') + q(T_{i_2}'') + 2q(T_i'') \geq 3\mathrm{diam}_4 G,$$

while from (3) and (8) we obtain

$$3(\frac{13}{20}\mathrm{diam}_4 G) + 2\frac{\mathrm{diam}_4 G}{2} = \frac{59}{20}\mathrm{diam}_4 G > 2q(T_i'') + q(T_{i_2}'') + 2q(T_i'') \geq 3\mathrm{diam}_4 G$$

which produces a contradiction.

Therefore $\mathrm{diam}_4 G \leq \frac{10}{7}\mathrm{rad}_4 G$ for all connected graphs G of order $p \geq 4$. □

In view of Theorems 4.4.24, 4.4.26 and 4.4.27, the following conjecture appears in [HOS1].

4.4.30 Conjecture

For all integers $n \geq 2$ and every connected graph G of order $p \geq n$

$$\mathrm{diam}_n G \leq \frac{2(n+1)}{2n-1}\mathrm{rad}_n G$$

4.4.31 Definition

A vertex v of a connected graph G is *n-eccentric* if there exists a vertex u in $C_n(G)$ and a set S of n vertices of G that contains both u and v such that $d(S) = e_n(u, G) = \text{rad}_n G$. The subgraph induced by the n-eccentric vertices is called the *n-eccentricity* of G and is denoted by $EC_n(G)$.

We now study relationships between the n-periphery and n- eccentricity of connected graphs. Since trees are the simplest connected graphs we begin by comparing their n-peripheries and n-eccentricities. The following result, established by Buckley and Lewinter in [BL1], is stated here without proof.

4.4.32 Theorem

If T is a tree, then $EC_2(T) = P_2(T)$ (i.e., the eccentricity of a tree is equal to its periphery).

The following extension of Theorem 4.4.32 was established by Oellermann and Swart in [OS1]

4.4.33 Theorem

Let $n \geq 3$ be an integer and T a tree of order at least n. Then $P_n(T) = EC_n(T)$.

Proof

If T has at most $n-1$ end-vertices, then, by Corollary 4.3.6, $e_n(v, T) = q(T)$ for all $v \in V(T)$ and hence $C_n(T) = P_n(T) = T$. Let w be any vertex of T.

If S is any set of n vertices that contains w and all the end-vertices of T, then S must contain a vertex $u \neq w$. Now $d(S) = e_n(u,T) = \text{rad}_n T$, and u belongs to $C_n(T)$. Hence $w \in EC_n(T)$ which implies that $T = EC_n(T)$; i.e., $P_n(T) = EC_n(T)$.

Suppose now that T has at least n end-vertices. We show first that $P_n(T) \subseteq EC_n(T)$. Let v be a vertex of $P_n(T)$ and S_n an n-diameter set of T containing v. Then $S_n \subseteq V_1(T)$. So every vertex of S_n is an end-vertex of T_{S_n}. For each vertex x in S_n let P_x be the stem of T_{S_n} which contains the vertex x, and let ℓ_x denote the length of P_x. Suppose u is a vertex of S_n such that $\ell_u = \min\{\ell_x : x \in S_n\}$. Referring to Definition 4.3.33, we have by Lemma 4.3.29 that $C_n(T_{S_n}) \cong (T_{S_n})^{\ell_u}$. Hence by Theorem 4.3.31, $(T_{S_n})^{\ell_u} \cong C_n(T)$. Therefore

$$\text{diam}_n T = q(C_n(T)) + n\ell_u. \qquad \sim (1)$$

We show next that $\text{rad}_n T = q(C_n(T)) + (n-1)\ell_u$. Let S'_{n-1} be any $(n-1)$-diameter set of T. Then, by Theorem 4.3.18, there exists an n-diameter set S'_n of T such that $S'_n \supset S'_{n-1}$. Let a be the vertex of $S'_n - S'_{n-1}$. For each vertex $z \in S'_n$, let ℓ'_z be the length of the stem of $T_{S'_n}$ containing z. Then necessarily $\ell'_a = \min\{\ell'_z : z \in S'_n\}$, otherwise, if say $y \in S'_n$ has $\ell'_y < \ell'_a$ then $y \neq a$ and $d(S'_{n-1} - \{y\} \cup \{a\}) > d(S'_{n-1})$, which is not possible. As before $\text{diam}_n T = q(T_{S'_n}) = q(C_n(T)) + n\ell'_a$. Thus from (1) $\ell'_a = \ell_u$, and $\text{diam}_{n-1} T = q(T'_{S_n}) - \ell'_a = \text{rad}_n T$. So $\text{rad}_n T = \text{diam}_{n-1} T = \text{diam}_n T - \ell_u = q(C_n(T)) + n\ell_u - \ell_u = q(C_n(T) + (n-1)\ell_u$.

Let $y \neq v$ be a vertex of S_n, and suppose w is a vertex on P_y such that $d_{T_{S_n}}(y,w) = \ell_u$. Observe that w is a vertex of $(T_{S_n})^{\ell_u}$. So $w \in V(C_n(T))$. Further, $S_n - \{y\} \cup \{w\}$ is a set of n vertices of T that contains both v and w such that $d(S_n - \{y\} \cup \{w\}) = q(T_{S_n}) - \ell_u = \text{diam}_n T - \ell_u =$

$q(C_n(T)) + (n-1)\ell_u = \text{rad}_n T$. Therefore $v \in V(EC_n(T))$. Hence

$$P_n(T) \subseteq EC_n(T). \qquad \sim (2)$$

We now show that $EC_n(T) \subseteq P_n(T)$. Assume, to the contrary, that there is a vertex $v \in V(EC_n(T))$ such that $v \notin V(P_n(T))$. Thus there exists a vertex $u \in V(C_n(T))$ and a set S of n vertices containing u and v such that $d(S) = e_n(u,T) = \text{rad}_n T$. By Theorem 4.3.7, v is an end-vertex of T, since T has at least n end-vertices. Observe that u is an end-vertex of T_S; otherwise u belongs to $T_{S-\{u\}}(= T_S)$ and, by Theorem 4.3.14, $S - \{u\}$ is an $(n-1)$-diameter set of T which contains v. By Theorem 4.3.18, there exists some n-diameter set of T which contains $S - \{u\}$ and hence v. However, this contradicts our assumption that $v \notin V(P_n(T))$.

Let S' be an n-diameter set of T such that $|S' \cap S|$ is as large as possible. Then $S' \subseteq V_1(T)$. Since $v \in S$ is not contained in any n-diameter set of T, it follows that $S' - S \neq \emptyset$. In fact, since $u \notin S', |S' - S| \geq 2$. Let ℓ_v be the length of the stem P_v of T_s which contains v. Suppose $P_v = (v =)v_0v_1...v_m$ $(= w)$. Observe that the shortest path from every vertex $y \in S' - S$ to a vertex of T_S must have length at most ℓ_v; otherwise if $\ell_y > \ell_v$ then $d(S - \{v\} \cup \{y\}) > d(S) = e_n(u,T) = \text{rad}_n T$, which is not possible. Let $P = (u =)u_0u_1...u_k$ be the stem of T_S which contains u. Let T_1 and T_2 be the two components of $T - u_{k-1}u_k$, and assume that $u \in V(T_2)$. By our choice of P, u is the only vertex of S in T_2. Since $u \in V(C_n(T))$ we have by Theorem 4.3.31 that $u \in V(T_{S'})$, hence it follows that T_2 contains a vertex, z say, of $S' - S$ such that the $z - v$ path in T contains P.

We show now that no vertex of P_v except possibly w belongs to $T_{S'}$. Clearly $v_0 = v$ does not belong to $T_{S'}$ since v does not belong to any n-diameter set of T. Assume to the contrary, that there exists a vertex

$a \in S' - S$ such that v_i belongs to the $a - u$ path in $T_{S'}$. We may assume that $i < m$ is the smallest integer such that v_i belongs to the $a - u$ path of $T_{S'}$, where $a \in S' - S$; i.e., if $j < i$ then v_j does not belong to $T_{S'}$. The $v_0 - v_i$ path must have length less than the $a - v_i$ path in $T_{S'}$; otherwise $d(S' - \{a\} \cup \{v_0\}) \geq d(S') = \text{diam}_n T$ and therefore $S' - \{a\} \cup \{v_0\}$ must be an n-diameter set of T that has more vertices in common with S than S'. Observe that the $a - v_i$ path has no vertex of T_S as an internal vertex; otherwise if b is such a vertex, then since $a \notin S$, there must exist a vertex $y \in S$ such that the $y - v_i$ path in T_S contains b; however then $v_0 v_1 ... v_i$ is a path from v to a vertex of degree at least 3 in T_S which is impossible. However, then $d(S \cup \{a\} - \{v\}) > d(S) = e_n(u, T)$ which is not possible. Therefore no internal vertex of P_v belongs to $T_{S'}$.

Suppose now that T_1 contains a vertex $a \in S' - S$. We show that the stem Q of $T_{S'}$ containing the vertex a, does not contain a vertex of T_S as internal vertex. Suppose $Q = (a =) a_0 a_1 ... a_r$ and that some a_j belongs to T_S. Choose j to be as small as possible. Then there exists an end-vertex $x \in S - S'$ of T such that a_j belongs to the $x - u$ path in T_S. As in the case of v we can show that a_j is not an internal vertex of the stem of T_S containing x, and no internal vertex of the $x - a_j$ path belongs to $T_{S'}$. However, then $d(x, a_j) \geq d(a, a_j)$ which implies that $d(S' \cup \{x\} - \{a\}) \geq d(S') = \text{diam}_n T$. So $S' \cup \{x\} - \{a\}$ must be an n-diameter set of T that has more vertices in common with S than S', contrary to assumption. Therefore T_1 contains no vertex of $S' - S$. Hence T_2 must contain at least two vertices of $S' - S$. Let z be a vertex of $S' - S$. Since the shortest path from z to a vertex of T_S has length at most ℓ_v and since u belongs to $T_{S'}$ and to S, the length of the stem of $T_{S'}$ containing z is at most ℓ_v. However, then $d(S' - \{z\} \cup \{v\}) \geq$

$d(S') = \text{diam}_n T$. So $(S' - \{z\}) \cup \{v\}$ is an n- diameter set of T which has more vertices in common with S than S', again a contradiction. Hence

$$EC_n(T) \subseteq P_n(T). \qquad \sim (3)$$

Thus (2) and (3) together imply that $EC_n(T) = P_n(T)$. $\square$

It was shown, by Reid and Weizhen Gu in [RW1], that there exist graphs for which the periphery is properly contained in its eccentricity and vice versa. So Theorem 4.4.32 cannot be extended to graphs in general. The following results show that Theorem 4.4.33 cannot be extended to include graphs in general (cf. [OS1]).

4.4.34 Theorem

For every positive integer $n \geq 3$ there exists a graph G_n such that $P_n(G_n) \subsetneq EC_n(G_n)$.

Proof

Consider the complete bipartite graph $K_{n,n}$ with $U = \{u_0, u_1, ..., u_{n-1}\}$ and $V = \{v_0, v_1, ..., v_{n-1}\}$ as partite sets. Join a new vertex v to every vertex in U and delete the edges of the type $u_i v_i$ for $0 \leq i \leq n-1$. Now, subdivide each edge of the type $u_i v_j$ for $0 \leq i, j \leq n-1$ and $i \neq j$ exactly once. Let $w_{i,j}$ be the vertex of degree 2 that is adjacent to u_i and v_j. Let G_n be the graph thus obtained. Then, $e_n(v, G_n) = d(\{v, v_0, v_2, ..., v_{n-1}\}) = \text{rad}_n G_n = 2n-1$, while $e_n(v_i, G_n) = d(V) = \text{diam}_n G_n = 2n+2$, for $0 \leq i \leq n-1$. It is easily verified that $C_n(G_n) = \langle\{v\}\rangle, P_n(G_n) = \langle V \rangle$. Take any $v_i \in V$, then $d(\{v_i, v, v_{i+2}, ..., v_{i+n-1}\}) = 2n-1 = \text{rad}_n G_n$ for $0 \leq i \leq n-1$, where addition of indices is taken modulo n. Hence $v_i \in V(EC_n(G_n))$

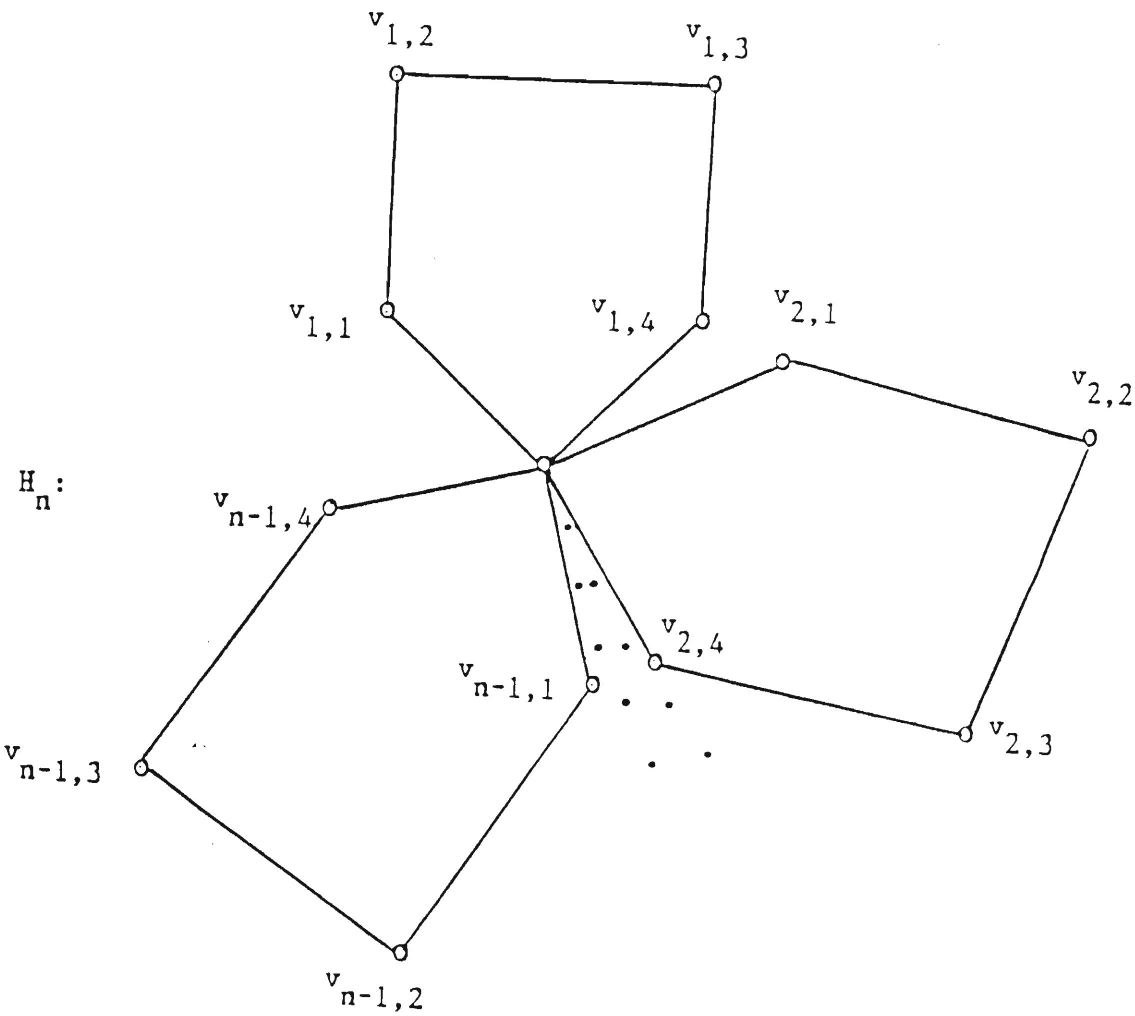
for $0 \leq i \leq n-1$. That $w_{1,2} \in E(C_n(G_n))$ may be seen by noting that $d(\{w_{1,2}, v, v_0, v_2, ..., v_{n-2}\}) = 2n-1$ and it follows by symmetry that $w_{i,j} \in V(EC_n(G_n))$ for $0 \leq i, j \leq n-1$ and $i \neq j$. Hence $EC_n(G_n) = \langle V \cup \{w_{i,j} : 0 \leq i, j \leq n-1 \text{ and } i \neq j\}\rangle$. Therefore it is clear that $P_n(G_n) \subsetneq EC_n(G_n)$.□

4.4.35 Theorem

For every positive integer $n \geq 3$ there exists a graph H_n such that $EC_n(H_n) \subsetneq P_n(H_n)$.

Proof

Let $Q_1, Q_2, ..., Q_{n-1}$ be $n-1$ cycles of length 5 where $Q_i = v_{i,0}v_{i,1}...v_{i,4}v_{i,0}$ $(1 \leq i \leq n)$. Let H_n be obtained by identifying the $n-1$ vertices $v_{1,0}, v_{2,0}, ..., v_{n-1,0}$ is a single vertex v_0. Then $e_n(v_0, H_n) = d(\{v_0, v_{12}, v_{22}, ..., v_{n-1,2}\}) = 2n-2 = \text{rad}_n H_n$, while the n-eccentricity of any vertex of $H_n - \{v_0\}$ is $2n-1$, hence $\text{diam}_n H_n = 2n-1$. Thus $C_n(H_n) = \langle\{v_0\}\rangle$ while $P_n(H_n) = H_n - v_0$. Further, since for all vertices v of the type $v_{i,2}$ and $v_{i,3}$ for $1 \leq i \leq n-1$ there exists a set S of n vertices of H_n including v_0 and v such that $d(S) = \text{rad}_n H_n$, it follows that $EC_n(H_n) = \langle\{v_{i,2}, v_{i,3} : 1 \leq i \leq n-1\}\rangle$. Hence $EC_n(H_n) \subsetneq P_n(H_n)$. □



4.4.36 Figure

The graph H_n of Theorem 4.4.35

4.4.37 Theorem

For every positive integer $n \geq 3$ there exists a graph F_n such that $P_n(F_n) \not\subset EC_n(F_n)$ and $EC_n(F_n) \not\subset P_n(F_n)$.

Proof

Let $R_1, R_2, ..., R_{n-1}$ be $n-1$ cycles of length 7 where $R_i = v_{i,0}v_{i,1}...v_{i,6}v_{i,0}$ for $1 \leq i \leq n-1$. Let F_n be obtained by identifying the $n-1$ vertices $v_{1,0}, v_{2,0}, ..., v_{n-1,0}$ in a single vertex v_0 and then joining a new vertex u to $v_{1,1}$. Then again it can be shown that $C_n(F_n) = \langle\{v_0\}\rangle, EC_n(F_n) = \langle\{v_{i,3}, v_{i,4} : 1 \leq i \leq n-1\}\rangle$ and $P_n(F_n) = \langle\{v_{i,3}, v_{i,4} : 2 \leq i \leq n-1\} \cup \{u, v_{1,4}\}\rangle$. Since u belongs to $P_n(F_n)$ but not to $EC_n(F_n)$ and $v_{1,3}$ belongs to $EC_n(F_n)$ but not to $P_n(F_n)$ the theorem now follows. □

4.5 An Algorithm and a Heuristic for the Steiner Problem in Graphs

Given a graph G and a nonempty set S of vertices of G, we now consider the practical problem of determining the Steiner distance $d(S)$ of S, as well as the problem of locating a Steiner tree with size $d(S)$, which is a subgraph of G. Our discussion here will be extended to include weighted graphs, hence the theory discussed thus far in Chapter 4 reduces to the special case where every edge has weight 1. We call this problem the Steiner Problem in Graphs (abbreviated SPG). The SPG was formally formulated by Winter [W1] as follows:

GIVEN: A weighted graph $G = (V(G), E(G), c)$ with p vertices, m edges, the weight function or cost function $c : E(G) \rightarrow \mathbb{R}$, and a subset $S = \{v_1, v_2, ..., v_n\} \subseteq V(G)$ of n vertices.

FIND: A weighted graph G_S which is a subgraph of G such that there is a path in G_S between every pair of vertices of S, and the sum of the costs of the edges of G_S is a minimum. (We call this sum the cost of the graph G_S.)

A number of exact algorithms for the SPG exist. We shall mention a few here and consider an algorithm described by Dreyfus and Wagner [DW1] in some detail. No polynomial time algorithms for solving the SPG are likely to exist, since Karp [K1] showed that this problem is NP-complete. Thus all known algorithms are only useful for small values of n. Hence it is of practical importance to obtain approximation methods which find trees whose costs are close to optimal. There are a number of such heuristics which are known, and we shall discuss one which was presented by Takahashi and Matsuyama in [MT1]. For a detailed survey of known algorithms and heuristics for the SPG see [W1].

We examine first some special cases of the SPG, see [W1].

4.5.1 Special Cases

Let $G = (V(G), E(G), c)$ be a connected graph with p vertices and m edges and cost function $c : E(G) \to \mathbf{R}$. Let $S \subseteq V(G)$ be a nonempty set of n vertices of G.

a) Suppose G contains edges with nonpositive weight. Let $F = \{e \in E(G) : c(e) \leq 0\}$. Consider the network $\tilde{G}$ obtained by the contraction of G along the edges in F. Given the solution $\tilde{G}_S$ in $\tilde{G}$, we obtain the solution G_S in G by adding to $\tilde{G}_S$ the edges of F. When the edge costs are all positive, every solution is a tree spanning S. Hence for the rest of this discussion we may assume, without loss of generality, that $c(e) > 0$ for all $e \in E(G)$.

b) If $|S| = 1$ then G_S consists of a single vertex.

c) If $|S| = p$ then the SPG reduces to the well-known minimal spanning tree (abbreviated MST) problem. Polynomial time algorithms for this problem are known (e.g. Kruskal [K2] or Prim [P1]).

d) If $|S| = 2$, then the SPG reduces to the well-known shortest path problem. Polynomial time algorithms for this problem are known (e.g. Bellman [B3] or Dijkstra [D1]).

For the rest of Section 4.5, G is assumed to be a weighted connected graph with cost function $c : E(G) \to \mathrm{R}$, $2 \leq n \leq p$, and for the reasons given in a) above $c(e) > 0$ for all $e \in E(G)$. Also $S \subseteq V(G)$ is assumed to be a nonempty set of n vertices of G.

4.5.2 Definition

Let G be a graph with order p and size m, and let $S \subseteq V(G)$ be any proper subset of vertices of G. Then a Steiner tree G_S for the set S, in G, is a connected subgraph of G which has minimum cost among all such subgraphs whose vertex set contains S. (That G_S is a tree is obvious.)

4.5.3 Some Exact Algorithms for the SPG

Hakimi [H1] provided a straightforward algorithm in which the Steiner tree G_S can be found by finding the MST's of subgraphs of G induced by subsets W of $V(G)$ such that $S \subseteq W \subseteq V(G)$. The time complexity of this algorithm is $0(n^2 2^{p-n} + p^3)$. Winter [W1] calls this algorithm the Spanning

Tree Enumeration Algorithm.

Another algorithm is presented by Hakimi in [H1] which Winter [W1] calls the Topology Enumeration Algorithm. Other exact algorithms can be found in [A1], [FGS1] and [B1, B2] to mention but a few.

We shall present the algorithm by Dreyfus and Wagner [DW1] which solves the SPG exactly, in time proportional to

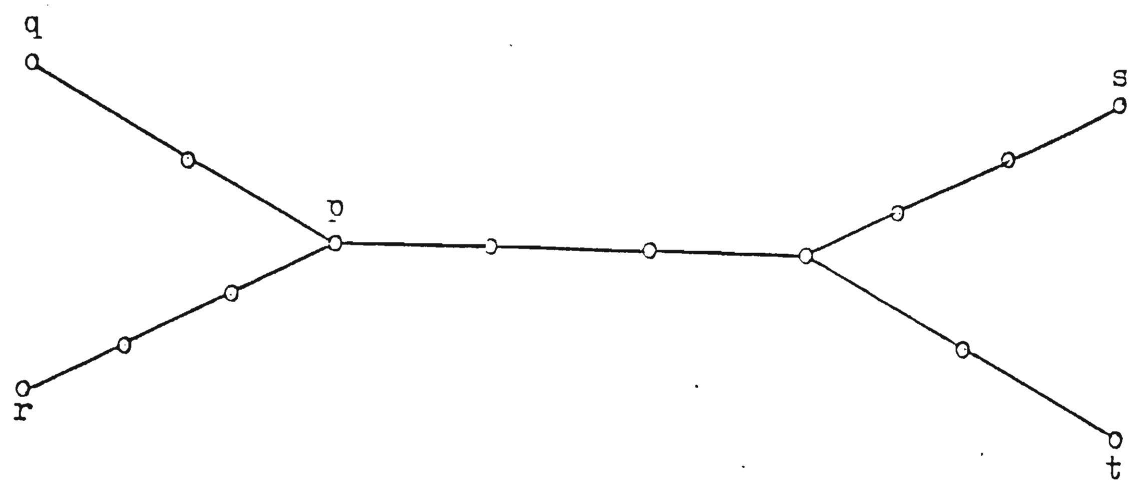
$$\frac{p^3}{2} + p^2(2^{n-1} - n - 1) + p\frac{(3^{n-1} - 2^n + 3)}{2}.$$

The time requirement above includes the term $\frac{p^3}{2}$, which can be eliminated if the set of shortest paths connecting each pair of vertices in the graph is available.

4.5.4 Example

Consider Figure 4.5.5, showing a typical solution G_S to a Steiner problem on $S \subseteq V(G)$ in G.

G_S:



4.5.5 Figure

Here $S = \{q, r, s, t\}$. Note that any vertex belonging to S, say q, belongs to a branch of the solution tree G_S at the vertex p (which we note has degree 3 and does not belong to S).

Clearly the path connecting q and p is the shortest path connecting these two vertices in G, otherwise G_S would not be a subtree of G of minimum cost containing S. Also note that each of the other branches of G_S at the vertex p represent the solution of a Steiner problem connecting fewer vertices than the number in the set S. In other words the subgraph of G_S induced by $\{p, s, t\}$ has the smallest cost among all connected subgraphs of G which contain the vertices $\{p, s, t\}$, similarly $\{p, r\}$ cannot be connected by a shorter path in G than the one which appears in G_S : If they could, again G_S would not be a solution to the Steiner problem for S in G.

The "division" of the Steiner problem by p into three smaller parts as demonstrated in Example 4.5.4 was called the optimal decomposition property by Dreyfus and Wagner in [DW1]. It can be formally stated as follows:

4.5.6 Optimal Decomposition Property

Let G be any connected graph of order p and size q. Suppose G_S is a Steiner tree for the set $S \subseteq V(G)$, in G, and let z be any vertex of S. If S contains at least three vertices then there exists a vertex $x \in V(G)$ and a subset D of S such that

1) D is a proper subset of $S - \{z\}$, and D is nonempty.

2) G_S consists of three edge disjoint subgraphs; S_1, S_2 and S_3.

3) S_1 contains $\{x, z\}$, S_2 contains $\{x\} \cup D$, while S_3 contains $\{x\} \cup (S - D - \{z\})$.

4) S_1, S_2 and S_3 are all Steiner trees for their respective sets in G.

A general proof of the existence of an optimal decomposition of the type described above, covering all degenerate cases, appears in Appendix A page 205 of [DW1].

The solution algorithm from [DW1] which we now describe is based on the dynamic programming methodology and in his survey Winter [W1] gives this algorithm the name: Dynamic Programming Algorithm.

4.5.7 Dynamic Programming Algorithm

The algorithm exploits Property 4.5.6. A straightforward application of Property 4.5.6 would entail choosing a vertex $z \in S$ (any z will do), then searching for the optimal choice of x. In turn, an optimal choice of x requires that an optimal choice of the subset $D \subset S$ be made, and that the Steiner trees S_2 and S_3 for the sets $D \cup \{x\}$ and $(S - D - \{z\}) \cup \{x\}$ in G, respectively, be known. Thus, the original problem could be solved recursively. However, we could also build up the desired solution by means of the following $|S| - 1$ steps. (Note that we assume the lengths of the shortest paths between every pair of vertices of G have been calculated; see [F1].)

Step 1: Remove one vertex, z, from S. Let $C = S - \{z\}$.

Step 2: Solve the Steiner problem for each set of two vertices of C and one vertex $y \in V(G)$. (y can be an element of C, or it can even be the vertex z.)

Step 3: Use this result to solve Steiner problems for each set of three vertices of C and one vertex $y \in V(G)$.

.

.

.

Step |S| -2: Solve Steiner problems for each set of $|S| - 2$ vertices of C and one vertex $y \in V(G)$.

Step |S|-1: Solve the Steiner problem for z and the set C.

Given a subset D of C, and $y \in V(G)$, each step in the solution above involves two searches: Search 1 locates the intermediate vertex, $x \in V(G)$; Search 2 finds the optimal proper subset E of D so that the cost of the Steiner trees containing $\{x\} \cup E$ and $\{P\} \cup (D - E)$, plus the distance from y to x is a minimum.

The efficiency of this procedure stems from the fact that only optimal solutions for the relevant subsets are ever considered. Nonoptimal solutions to smaller subproblems are disposed of at the time that subproblem is solved. The optimal solution is retained for use in solving later subproblems, and the smaller subproblem is never solved again. Straightforward enumeration of all possible solutions to the entire problem would unnecessarily consider nonoptimal solutions many times. This building up of larger optimal solutions from optimal solutions of all possible smaller problems is the fundamental technique in the general methodology of dynamic programming.

Let us discuss in some detail the procedure whereby the Steiner solution for a given subset D consisting of a certain $j (\geq 2)$ vertices of C and one vertex, $w \in V(G)$, is found. Here again we avoid some unnecessary calculation by first solving all possible smaller problems of a certain form. First we associate with each vertex $k \in V(G)$ an integer $S_k(D)$ which is found using the following method:

(1) Divide D into two proper subsets E and F, and add the Steiner distance for the set consisting of the members of E and vertex k to the Steiner distance for the set consisting of the vertices in F and vertex k, and

(2) Minimize this sum over all distinct choices of sets E and F.

(Note that the number $S_k(D)$ is not necessarily the Steiner distance of the set composed of the elements of D and vertex k, since no saving due to coalescing the subsolutions at a vertex other than k is considered.) Having done this for a given D and all k, to solve the Steiner problem for w and D,we let d_{wk} denote the length of a shortest path from w to k in G and we minimize $d_{wk} + S_k(D)$ over all vertices $k \in V(G)$. Let $S(w, D)$ denote the cost of the Steiner tree for the vertices $\{w\} \cup D$. Since $S_k(D)$ does not depend on the choice of the vertex w, knowledge of $S_k(D)$ for all $k \in V(G)$ allows easy computation of Steiner solutions for any $w \in V(G)$, all, of course, for a given D. The computation is repeated, then, for all choices of the set D.

So far we have described a procedure for generating the cost of the Steiner tree, but not the actual tree. To determine the tree, there are two "pure" strategies available. Method 1: For each choice of w and D one can record the value of k that minimized $d_{wk} + S_k(D)$ and the sets E and F that generated $S_k(D)$. Then k is the vertex which produces the optimal decomposition, with w connected to k by a shortest path, while E and F, respectively, are joined to k by paths in the Steiner trees for the sets $\{E \cup \{k\}\}$ and $\{F \cup \{k\}\}$ in G, respectively. Method 2: On the other hand, the values of $S_k(D)$ can be stored and the minimizing value of k and associated sets E and F can be recomputed as needed in the reconstruction of the Steiner tree. In either case, as is typical in dynamic programming procedures, the Steiner tree is constructed (after the optimal cost has been determined) by processing sets in the reverse order of that of the cost-determination algorithm. The first method of tree-construction involves less computation while the second uses less computer strorage. Since tree-construction by Method 2 requires at most $\frac{1}{p}$th the computation time of the cost- generation, this is the recommended and most practical method.

We now present a numerical illustration from [DW1] of the procedure followed in Algorithm 4.5.7.

4.5.8 Example

Let $V(G) = \{1,2,3,4,5,6,7\}$, $S = \{1,2,3,4\}$ and the matrix A of distances ($a_{ij} = a_{ji} =$ the weight of the edge (i,j) between vertices i and j) be

$$A = \begin{array}{c|ccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 1 & X & 2 & 2 & 2 & 1 & 1 & 2 \\ 2 & 2 & X & 2 & 2 & 2 & 1 & 2 \\ 3 & 2 & 2 & X & 2 & 2 & 2 & 1 \\ 4 & 2 & 2 & 2 & X & 1 & 2 & 1 \\ 5 & 1 & 2 & 2 & 1 & X & 2 & 1 \\ 6 & 1 & 1 & 2 & 2 & 2 & X & 1 \\ 7 & 2 & 2 & 1 & 1 & 1 & 1 & X \end{array}$$

First we compute the matrix D of shortest lengths ($d_{ij} = d_{ji} =$ the length of the shortest path between vertices i and j) by the method described in [F1]. Clearly, by our choice of data, matrix D is identical to matrix A.

Step 1: Remove one vertex, say vertex 1, from S. Let $C = \{2,3,4\}$.

Step 2: Let $D = \{2,3\}$. Then $S_1(D) = d_{12} + d_{13} = 2 + 2 = 4$, $S_2(D) = 2$, $S_3(D) = 2$, $S_4(D) = 4$, $S_5(D) = 4$, $S_6(D) = 3$, $S_7(D) = 3$.

Letting $S(w, D)$ denote the cost of the Steiner tree for the set $\{w \cup D\}$, we have $S(1, D) = \min_k (d_{1k} + S_k(D)) = 4$ (with several different trees yielding the result e.g. $k = 2$).
$S(2, D) = 2$, $S(3, D) = 2$, $S(4, D) = 4$, $S(5, D) = 4$, $S(6, D) = 3$, $S(7, D) = 3$.

Now let $D = \{3,4\}$. Then $S_1(D) = 4$, $S_2(D) = 4$, $S_3(D) = 2$, $S_4(D) = 2$, $S_5(D) = 3$, $S_6(D) = 4$, $S_7(D) = 2$. Hence $S(1, D) = 4$, $S(2, D) = 4$, $S(3, D) = 2$, $S(4, D) = 2$, $S(5, D) = 3$, $S(6, D) = 3$, $S(7, D) = 2$.

Finally let $D = \{2,4\}$. Then $S_1(D) = 4$, $S_2(D) = 2$, $S_3(D) = 4$, $S_4(D) = 2$, $S_5(D) = 3$, $S_6(D) = 3$, $S_7(D) = 3$. Hence $S(1, D) = 4$, $S(2, D) = 2$, $S(3, D) = 4$, $S(4, D) = 2$, $S(5, D) = 3$, $S(6, D) = 3$, $S(7, D) = 3$.

We are now ready for Step $|S| - 1 = 4 - 1 = 3$, in this case.

Step 3: Let $D = \{2, 3, 4\}$.
Let $E = \{2\}$ and $F = \{3,4\}$. Then $S_1(D|E, F) = S(1, E) + S(1, F) = 2+4 = 6$. Now let $E = \{3\}, F = \{2,3\}$. Then $S_1(D|E, F) = 2+4 = 6$. Finally let $E = \{4\}, F = \{2, 3\}$. Then $S_1(D|E, F) = 2 + 4 = 6$. Hence $S_1(D) =$ minimum over all choices of E and F of $S_1(D|E, F) = 6$.

Letting $E = \{2\}$ and $F = \{3,4\}$, $S_2(D|E, F) = 4$. Letting $E = \{3\}$ and $F = \{2,4\}, S_2(D|E, F) = 4$. Letting $E = \{4\}$ and $F = \{2,3\}$,

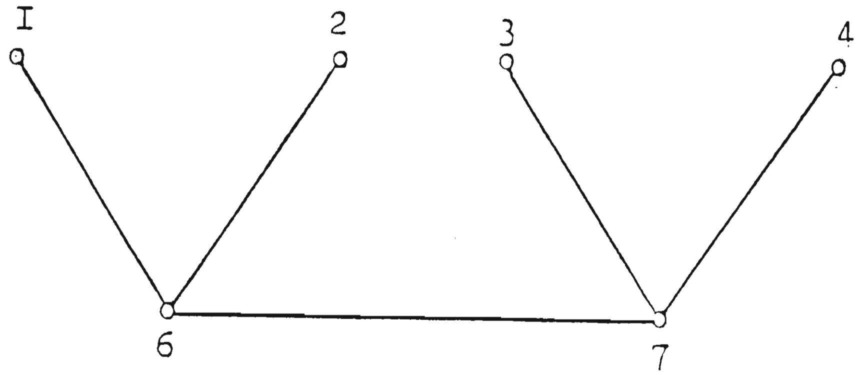
$S_2(D|E,F) = 4$. Hence $S_2(D) = 4$.

Similarly $S_3(D) = \min(4,4,4) = 4$. $S_4(D) = \min(4,5,5) = 4$. $S_5(D) = \min(5,5,5) = 5$. $S_6(D) = \min(4,5,5) = 4$. $S_7(D) = \min(4,4,4) = 4$.

Hence $S(1,D) = \min_k (d_{1k} + S_k(D)) = d_{16} + S_6(D) = 1 + 4 = 5$, which is the size of a Steiner tree for $\{1,2,3,4\}$ in G.

To construct the Steiner tree for $\{1,2,3,4\}$ in G, we note that since $k = 6$ yielded the minimum in the above minimization, vertex 1 is to be connected to vertex 6 by a shortest path in G, which in this case is the edge connecting vertices 1 and 6. Now $S_6(D)$ resulted when $E = \{2\}$ and $F = \{3,4\}$. Hence the shortest path from 6 to 2 is part of the solution Steiner tree for {1,2,3,4}. This is the edge between the vertices 2 and 6. Finally, the Steiner tree for the set {6,3,4} in G must be a subtree of the Steiner tree for {1,2,3,4} in G. To find this subtree we refer to $S(6,D)$ above for $D = \{3,4\}$ we see $S(6,D) = 3$ and the value 3 was obtained when $k = 7$ yielding $\min_k (d_{6k} + S_k(D)) = d_{67} + S_7(D) = 1 + 2 = 3$. Hence vertex 6 must be connected to vertex 7 by a shortest path in G (the edge connecting 6 and 7 in this case) and vertex 7 must be connected to vertices 3 and 4 by shortest paths in G (the edges between 7 and 3, and 7 and 4, respectively, in this case). Hence the Steiner tree for {1,2,3,4} in G, consists of the edges (1,6), (2,6), (6,7), (7,3) and (7,4), see Figure 4.5.9.

G_S :



4.5.9 Figure

The Steiner tree G_S for {1,2,3,4} in G.

The sum of the weights of the edges of G_S is indeed 5, agreeing with our computed value of $S(1, D)$ for $D = \{2, 3, 4\}$.

With reference to Property 4.5.6, if vertex 1 is identified with vertex q in the statement of the property, then vertex 6 is identified with vertex p and vertex 2 constitutes the set D in the statement of the property. Then set S_1 consists of the edge (1,6), set S_2 consists of (2,6) and S_3 consists of {(6,7), (7,3), (7,4)}.

For verification of the time requirement for Algorithm 4.5.7 see [DW1].

Since the time needed to solve the Steiner problem increases exponentially with an increase in the size of our set S, we deduce that Algorithm 4.5.7 is useful only for small $|S|$. Hence we now turn our attention to approximation methods which find trees which have costs close to that of a Steiner tree.

There are many such heuristics available, see [P2], [R1] and [A1] for examples. We shall consider the approximate solution for the Steiner problem in graphs developed by Takahashi and Matsuyama [MT1], which requires at most $0(np^2)$ time and we shall determine the accuracy of the approximation. For the remainder of this section we follow [MT1].

4.5.10 Definition

Let W be a proper subset of $V(G)$, then define Path (W, v) to be a path whose cost is minimum among all paths from vertices in W to vertex $v \notin W$. Denote by $\hat{c}(W, v)$ the cost of Path (W, v).

4.5.11 Approximation Algorithm

Step 1: Start with subgraph $T_1 = (V_1, E_1)$ of G, with vertex set V_1 and edge set E_1 consisting of a single vertex, say v_1, where $v_1 \in S$; i.e., $V_1 = \{v_1\}$ and $E_1 = \emptyset$. Let $i = 2, 3, ..., n$.

Step i: Find a vertex in $S - V_{i-1}$, say v_i, such that $\hat{c}(V_{i-1}, v_i) = \min\{\hat{c}(V_{i-1}, v_j) : v_j \in S - V_{i-1}\}$. Construct tree $T_i = (V_i, E_i)$ with vertex set V_i and edge set E_i, by adding Path (V_{i-1}, v_i) to T_{i-1}; i.e., set $V_i = V_{i-1} \cup V$ (Path (V_{i-1}, v_i)) and $E_i = E_{i-1} \cup E$(Path (V_{i-1}, v_i)).

We assume that when there are ties in step i, they may be broken arbitrarily. At each step in this algorithm, a tree containing a subset of S has been built up, and a new vertex of S is inserted together with a path of minimum cost connecting the tree and the vertex. Hence we end up with a tree T_n which is our approximate solution to the Steiner problem.

We note by Dijkstra's algorithm [D1] that Path (V_{i-1}, v_i) can be computed in time complexity $O(p^2)$; hence this algorithm requires at most $O(np^2)$ time.

4.5.12 Definition

Let OPTIMAL represent the cost of the Steiner tree for the set S in G, and let $d_S(u,v)$ denote the cost of the path between vertices u and v in a Steiner tree T_S.

The following lemma will aid us in determining the accuracy of Approximation Algorithm 4.5.11.

4.5.13 Lemma

Let G_S be a Steiner tree for the set S in a graph G, and let $V(G_S) = \{v_1, v_2, ..., v_n\}$. There exists a permutation $t_1, t_2, ..., t_n$ of 1,2,...,n such that $d_S(v_{t_1}, v_{t_2}) + d_S(v_{t_2}, v_{t_3}) + ... + d_S(v_{t_{n-1}}, v_{t_n}) + d_S(v_{t_n}, v_{t_1}) = 2.$ OPTIMAL and

$$d_S(v_{t_n}, v_{t_1}) \geq (\frac{2}{n}).\text{OPTIMAL}.$$

Proof

Suppose that $v_{i_j} \in S$ is visited after $v_{i_{j-1}} \in S$ for each $2 \leq j \leq n$ in a depth first search of a Steiner tree G_S for S in G, starting from an arbitrary vertex of G_S. Then

$$d_S(v_{i_1}, v_{i_2}) + \ldots + d_S(v_{i_{n-1}}, v_{i_n}) + d_S(v_{i_n}, v_{i_1}) = 2.\text{OPTIMAL}.$$

Assume $d_S(v_{i_{r-1}}, v_{i_r}) = \max\{d_S(v_{i_1}, v_{i_2}), \ldots, d_S(v_{i_{n-1}}, v_{i_n}), d_S(v_{i_n}, v_{i_1})$ for some r, $2 \leq r \leq k$. Then setting $t_1 = i_r, \ldots, t_{n-r+1}, t_{n-r+2} = i_1, \ldots, t_n = i_{r-1}$, we have $d_S(v_{t_{j-1}}, v_{t_j} \leq d_S(v_{t_n}, v_{t_1})$ for all $2 \leq j \leq n$. Hence

$$2.\text{OPTIMAL} = d_S(v_{t_1}, v_{t_2}) + \ldots + d_S(v_{t_{n-1}}, v_{t_n}) + d_S(v_{t_n}, v_{t_1}) \leq n d_S(t_n, v_{t_1})$$

and hence

$$\left(\frac{2}{n}\right).\text{OPTIMAL} \leq d_S(v_{t_n}, v_{t_1}). \qquad \square$$

4.5.14 Definition

Let APPROXIMATE be the cost of the tree T_n obtained by Approximate Algorithm 4.5.11. Then APPROXIMATE is equal to $\sum\limits_{i=1}^{n} \hat{c}(V_{i-1}, v_i)$.

We now show that the tree T_n obtained by Approximate algorithm 4.5.11 has a worst case cost ratio to the Steiner tree for S in G which is less than or equal to $2(1 - \frac{1}{n})$.

4.5.15 Theorem

For all p and n $(2 \leq n \leq p-1)$

$$\frac{\text{APPROXIMATE}}{\text{OPTIMAL}} \leq 2(1 - \frac{1}{n}).$$

Moreover if $n = p$, then APPROXIMATE = OPTIMAL.

Proof

If $n = p$, then the problem reduces to the well-known minimal spanning tree problem, hence the latter half of the theorem is Prim's algorithm [P1].

Since the cost of Path (V_{i-1}, v_i) is minimum among all paths between vertices in V_{i-1} and vertices in $S - V_{i-1}$, we have

$$\hat{c}(V_{i-1}, v_i) \leq d_S(v_p, v_q) \text{ for all } 2 \leq i \leq n, \qquad \sim (1)$$

where $1 \leq \min\{p, q\} \leq i - 1$ and $i \leq \max\{p, q\} \leq n$. By Lemma 4.5.13 there is a permutation $t_1, t_2, ..., t_n$ of $1, 2, ..., n$ such that

$$d_S(v_{t_1}, v_{t_2}) + d_S(v_{t_2}, v_{t_3}) + ... + d_S(v_{t_{n-1}}, v_{t_n}) + d_S(v_{t_n}, v_{t_1}) = 2.\text{OPTIMAL} \qquad \sim (2)$$

and

$$d_S(v_{t_n}, v_{t_1}) \geq \left(\frac{2}{n}\right) \text{OPTIMAL}. \qquad \sim (3)$$

We can construct a one-to- one correspondence between the numbers i, $i = 2, 3, ..., n$ and pairs $(t_{j-1}, t_j), j = 2, 3, ..., n$, such that

$$\hat{c}(V_{i-1}, v_i) \leq d_S(v_{t_{j-1}}, v_{t_j}).$$

Such a correspondence can be established by the method which Rosenkrantz, Stearns and Lewis II used in a more general case in [LRS1, Proof of Lemma 3]. For each i with $2 \leq i \leq n$, consider the longest subsequence $t_{p(i)}, t_{p(i)+1}, ..., i, ..., t_{q(i)-1}, t_{q(i)}$ including i of the sequence $t_1, t_2, ..., t_n$ such that $t_{p(i)} \leq i$, $t_{q(i)} \leq i$ and $t_j \geq i$ for each j, $j = p(i) + 1, ..., q(i) - 1$. In other words, subsequence $t_{p(i)}, ..., t_{q(i)}$ includes i, and all the intermediate numbers except for i in that subsequences are larger than i. We define the critical number i^* for i by

$$i^* = \begin{cases} t_{p(i)} & \text{if } t_{q(i)} = i, \\ t_{q(i)} & \text{if } t_{p(i)} = i, \\ \max\{t_{p(i)}, t_{q(i)}\} & \text{otherwise.} \end{cases}$$

The critical pair for i is defined to be

$$\begin{array}{ll} (t_{p(i)}, t_{p(i)+1}) & \text{if} \quad i^* = t_{p(i)}, \\ (t_{q(i)-1}, t_{q(i)}) & \text{if} \quad i^* = t_{q(i)}. \end{array}$$

Next we show that no two numbers i and j from the sequence $t_1, t_2, ..., t_n$ can have the same critical pair. Assume, to the contrary, that i and $j(i < j)$ have the same critical pair (t_{r-1}, t_r). Assume that $t_r < t_{r-1}$. Then t_r is critical for i and j, and $r = q(i) = q(j)$. Since all the intermediate numbers in the subsequence from j to t_r of subsequence $t_1, t_2, ..., t_n$ are larger than j, number i cannot be in that subsequence. This implies since $i < j$ that number j is in the sequence from i to t_r. Since then by definition $t_r < i$, all the numbers in the sequence from i to j are larger than t_r. Thus $t_{p(j)} > t_r = t_{q(j)}$, contradicting the assumption that t_r is critical for j. The same contradiction is concluded if we assume $t_r > t_{r-1}$.

Let $(t_{r(i)-1}, t_{r(i)})$ be the critical pair for i, then from (1) we have, since $\min\{t_{r(i)-1}, t_{r(i)}\} < i \leq \max\{t_{r(i)-1}, t_{r(i)}\}$ holds,

$$\hat{c}(V_{i-1}, v_i) \leq d_S(v_{t_{r(i)-1}}, v_{t_{r(i)}}). \qquad \sim (4)$$

From (2), (3) and (4), we have

$$\text{APPROXIMATE} \quad = \quad \sum_{i=2}^{n} \hat{c}(V_{i-1}, v_i)$$

$$\leq \sum_{i=2}^{n} d_S(v_{t_{r(i)-1}}, v_{t_{r(i)}})$$

and since no two numbers have the same critical pair

$$\sum_{i=2}^{n} d_S(v_{t_{r(i)-1}}, v_{t_{r(i)}}) = \sum_{p=2}^{n} d_S(v_{t_{p-1}}, v_{t_p})$$

hence

$$\begin{aligned} \text{APPROXIMATE} &\leq \textstyle\sum_{p=2}^{n} d_S(v_{t_{p-1}}, v_{t_p}) = 2.\text{OPTIMAL} - d(v_{t_n}, v_{t_1}) \\ &\leq 2.\text{OPTIMAL} - \left(\tfrac{2}{n}\right).\text{OPTIMAL} = 2.(1 - \tfrac{1}{n}).\text{OPTIMAL} \end{aligned}$$

□

Finally we show that for $n \leq p - 1$, we can construct graphs for which the APPROXIMATE to OPTIMAL ratio is equal to $2(1 - \frac{1}{n})$.

4.5.16 Theorem

For all p and n $(2 \leq n \leq p)$, there exists a graph for which

$$\frac{\text{APPROXIMATE}}{\text{OPTIMAL}} = 2(1 - \frac{1}{n}).$$

Proof

Let $V = \{1, 2, ..., p\}$, $E = \{(i,j) : i = 1, 2, ..., p,\ j = 1, 2, ..., p\}$, and $S = \{1, 2, ..., n\}$. Suppose that

$$c(i,j) = \begin{cases} 1 & i = 1, 2, ..., n,\ j = n + 1, \\ 2 & i = 1, ..., n - 1,\ j = i + 1, \\ 10 & \text{otherwise} \end{cases}$$

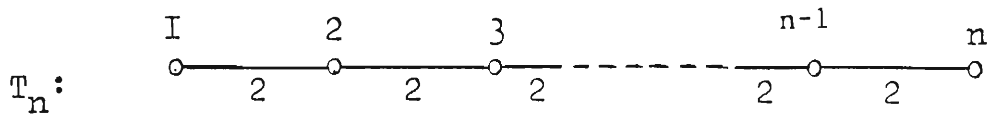
Then let G be the graph with vertex set V, edge set E and cost function c.

It is evident that the tree T_n with vertex set S and edge set $E(T_n) = \{(j,k) : j = 1,2,...,n-1,\ k = j+1\}$ with every edge of T_n having weight 2 (see Figure 4.5.17), is obtainable by Approximation algorithm 4.5.11. The cost of T_n is thus $2(n-1)$. It is also evident that a Steiner tree G_S for S in G has vertex set $V(G_S) = \{S \cup \{n+1\}\}$, edge set $E(G_S) = \{(i, n+1) : i = 1,2,...,n\}$ with every edge of G_S having weight 1 (see Figure 4.5.18). The cost of G_S is then n. Hence using our previous terminology OPTIMAL $= n$ while APPROXIMATE $= 2(n-1)$ and

$$\frac{\text{APPROXIMATE}}{\text{OPTIMAL}} = \frac{2(n-1)}{n} = 2(1 - \frac{1}{n})$$

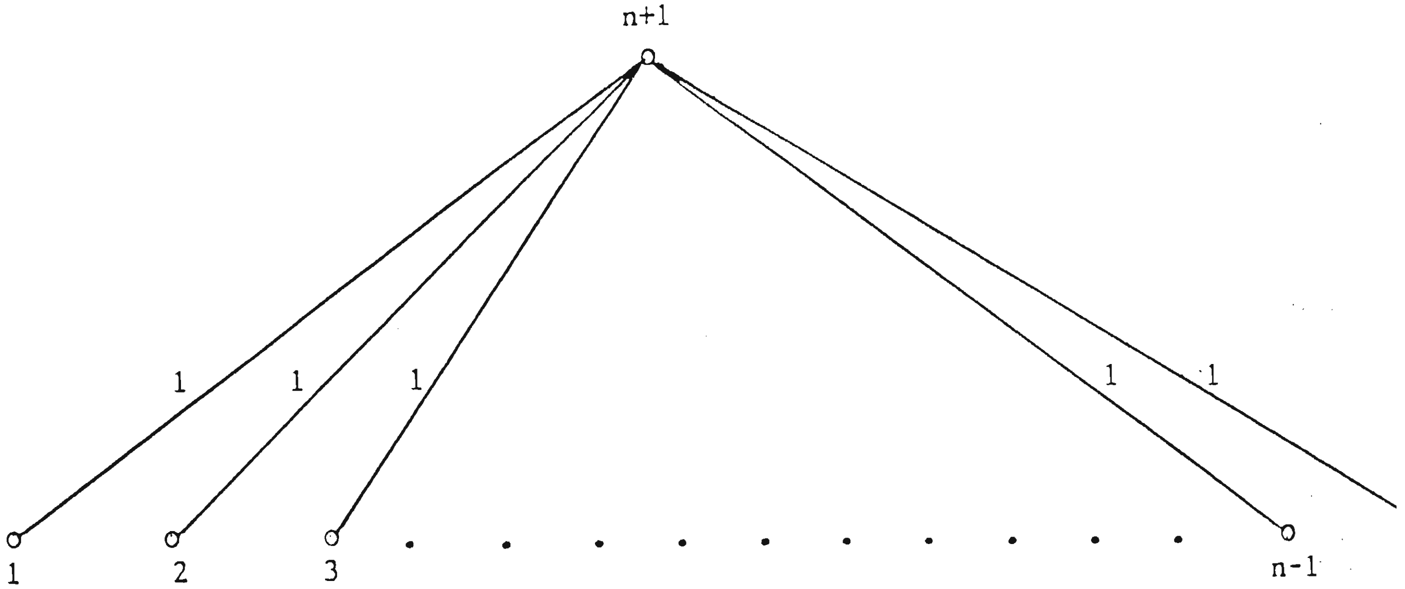
as required. □

By Theorems 4.5.15 and 4.5.16, the worst case ratio of APPROXIMATE to OPTIMAL is $2(1 - \frac{1}{n})$.



4.5.17 Figure

The tree T_n of Theorem 4.5.16.



4.5.18 Figure

The Steiner tree G_S of Theorem 4.5.16.

In view of the complexity of the problem of determining the Steiner distance of a given set of vertices in a graph, efforts have been made to consider graphs with special properties in which Steiner distances of given sets may be found in polynomial time. For instance, for any integer $k \geq 2$, Day, Oellermann and Swart defined a graph G to be k-Steiner distance hereditary if, for every $S \subseteq V(G)$ such that $|S| = k$ and every connected induced subgraph H of G containing S, $d_H(S) = d_G(S)$. In [DOS1] they showed that if G is 2-Steiner distance hereditary, then G is k-Steiner distance hereditary for every integer $k \geq 2$. They then gave efficient algorithms for testing whether a graph is 2-Steiner distance hereditary and for determining the Steiner distance of a set of k vertices in a k-Steiner distance hereditary graph, thereby providing an efficient algorithm for obtaining the Steiner distances of sets of k vertices in 2-Steiner distance hereditary graphs.

The search for further large classes of graphs in which Steiner distances may be determined by means of efficient algorithms presents a challenging new field of research as does the investigation of graphs with specified properties such as the uniquely Steiner n-eccentric graphs investigated by Henning, Oellermann and Swart in [HOS3]. Furthermore, in view of the complexity of the problem of evaluating the n-Steiner radius and n-Steiner diameter of a graph, an investigation of graphs that have maximal or minimal order or size and given n-radius or n-diameter may yield results that have useful applications as would the characterization of the associated classes of extremal graphs.

BIBLIOGRAPHY

The starred entries in the bibliography denote books.

Aneja, Y.P.

[A1] An integer linear programming approach to the Steiner problem in graphs. *Networks* **10** (1980), 167- 178.

Beasely, J.E.

[B1] An algorithm for the Steiner problem in graphs. *Networks* **14** (1984), 147-159.

[B2] An SST-based algorithm for the Steiner problem in graphs. Tech. Rep., Department of Man. Science, Imperial College, London (1985).

Bellman, R.E.

[B3] On a routing problem. *Q. Appl. Math.* **16** (1958), 87 -90.

Benadé, G.

[BGMW1] (with W. Goddard, T.A. McKee and P.A. Winter). On distances between isomorphism classes of graphs. Math. Bohemica **116** (1991), 160 -169.

Buckley, F

[BL1] (with M. Lewinter). Minimal graph embeddings, eccentric vertices and the peripheries. *Proceedings of the Fifth Caribbean Conference on Combinatorics and Computing, University of the West Indies* (1988), 72 -84.

Chartrand, G.

[CGHLSW1] (with W. Goddard, M.A. Henning, L. Lesniak, H.C. Swart and C.E. Wall). Which graphs are distance graphs? *Ars Combin.* **29A** (1990), 225 - 232.

[CJO1] (with G. Johns and O.R. Oellermann). On peripheral vertices in graphs. Topics in Combinatorics and Graph Theory, R. Bodendiek and R. Henn (Eds.). Physica-Verlag, Heidelberg (1990), 193-199.

*[CL1] (with L. Lesniak) Graphs and Digraphs (Second Edition), Wadsworth, Monterey, (1986).

[COTZ1] (with O.R. Oellermann, S. Tian and H.B. Zou) Steiner distance in graphs. *Čas. pěst. mat.* **114** (1989), 399 - 410.

[CSZ1] (with F. Saba and H.B. Zou) Edge rotations and distances between graphs. *Čas. pěst. mat.* **110** (1985), 87 - 91.

Courant, R.

*[CR1] (with H.E. Robbins) What is Mathematics? Oxford University Press (1941) Sixth Printing 1953.

Day, D.P.

[DOS1] (with O.R. Oellermann and H.C. Swart) Steiner Distance Hereditary graphs. (Submitted for publication.)

Dijkstra, E.W.

[D1] A note on two problems in connection with graphs. *Numer. Math.* **1** (1959), 269 - 271.

Dreyfus, S.E.

[DW1] (with R.A. Wagner). The Steiner problem in graphs. *Networks* **1**(1972), 195 - 207.

Even, S.
*[E1] Graph algorithms. Computer Science Press, Rockville, Maryland (1985).

Floyd, R.W.
[F1] Algorithm 97, Shortest path. *Comm. ACM* 5(1962), 345.

Foulds, L.R.
[FGS1] (with P.B. Gibbons and M.L. Shore) An algorithm for the Steiner problem in graphs. *Networks* **12**(1982), 323 - 333.

Gibbons, P.B.
See [FGS1].

Goddard, W
[GS1] (with H.C. Swart) Distances between graphs under edge operations. To appear.

See also [BGMW1] and [CGHLSW1].

Hakimi, S.L.
[H1] Steiner's problem in graphs and its implications. *Networks* **1** (1971), 113 - 133.

Harary, F.
[H2] The maximum connectivity of a graph. *Proc. Acad. Sci. USA* **48** (1962), 1142 - 1146.

Henning, M.A.
[HOS1] (with O.R. Oellermann and H.C. Swart) On the Steiner radius and Steiner diameter of a graph. *Ars Combin.* **29C**(1990), 13 - 19.

[HOS2] (with O.R. Oellermann and H.C. Swart) On vertices with Maximum Steiner eccentricity in graphs. *Graph Theory, Combinatorics, Algorithms and Applications.* Y. Alavi, F.R.K. Chung, R.L. Graham and D.F. Hsu. Siam Publications (1991), 393- 403.

[HOS3] (with O.R. Oellermann and H.C. Swart) Uniquely Steiner-Eccentric graphs. Submitted for publication.

Johnson, M.

[J1] An ordering of some metrics defined on the space of graphs. *Czech. Mat. J.* **37** (112) 1987, 75 - 85.

Johns, G.

[J2] Distance in graphs. Ph.D. Thesis, Western Michigan University, 1988.

See also [CJO1].

Jordan, C.

[J3] Sur les assemblages de ligners. *J. Reine Angev. Math* **70**(1869), 185 - 190.

Karp, R.M.

*[K1] Reducibility among combinatorial problems, Complexity of computer computations, R.E. Miller and J.W. Thatcher, eds., Plenum Press, New York (1972).

Kruskal, J.B.

[K2] On the shortest spanning subtree of a graph and the travelling salesman problem. *Proc. Amer. Math. Soc.* **7** (1956) 48 - 50.

König, D.
*[K3] Theorie der endlichen und unendlichen graphen. Liepzig (1936). (Reprinted Chelsea, New York (1950).

Lesniak, L
[L1] Eccentricity sequences in graphs. *Period. Math. Hungar.* **6**(1975), 287 - 293.

See also [CGHLSW1] and [CL1].

Lewinter, M
See [BL1].

Lewis II, P.M.
[LRS1] (with D.J. Rosenkrantz and R.E. Stearns) An analysis of several heuristics for the travelling salesman problem. To appear in SIAM J. Comput.

Matsuyama, A
[MT1] (with H. Takahashi) An approximate solution for the Steiner problem in graphs. *Math. Japonica* **24**(1980), 573 - 577.

Melzak, Z.A.
[M1] On the problem of Steiner. *Canad. Math. Bull.*, **4**(1961), 143 - 148.

McKee, T.A.
See [BGMW1].

Oellermann, O.R.
[O1] Steiner numbers in graphs. *Quaestiones Mathematicae* **13**(1990), 159 - 164.

[OS1] (with H.C. Swart) On the Steiner periphery and Steiner eccentricity of a graph. Topics in Combinatorics and Graph Theory, R. Bodendick, R. Henn (Eds.). Physica-Verlag Heidelburg 1990, 541 - 547.

[OT1] (with S. Tian) Steiner centers in graphs. *Journal of Graph Theory* **14** (1990), 585 - 597.

[OT2] (with S. Tian) Steiner n-eccentricity sequences of graphs. Recent Studies in Graph Theory (ed. V.R. Kulli). Vishwa International Publications (1989), 206-211.

See also [COS1], [COTZ1], [DOS1], [HOS1], [HOS2], [HOS3] and [CJO1].

Prim, R.C.

[P1] Shortest connection networks and some generalization. *Bell System Tech. J.* **36** (1957), 1389 - 1401.

Plesnik, J.

[P2] A bound for the Steiner tree problem in graphs. *Mat. Slovaca* **31** (1981), 155 - 163.

Rayward-Smith, V.J.

[R1] The computation of nearly minimal Steiner trees in graphs. *Int. J. Math. Educ. Sci.Technol.* **14** (1983), 15 - 23.

Reid, K.B.

[RW1] (with G. Weizhen) Peripheral and eccentric vertices graphs. Submitted for publication.

Robbins, H.

See [CR1].

Rosenkrantz, D.J.

See [LRS1].

Saba, F.

See [CSZ1].

Shore, M.L.

See [FGS1].

Stearns, R.E.

See [LRS1].

Swart, H.C.

See [CGHLSW1], [DOS1], [GS1], [HOS1], [HOS2], [HOS3] and [OS1].

Takahashi, H.

See [MT1].

Tian, S.

See [COTZ1], [OT1] and [OT2].

Wagner, R.A.

See [DW1].

Wall, C.E.

See [CGHLSW1].

Weizhen, G.

See [RW1].

Winter, P.A.

See [BGMW1].

Winter, P.

[W1] Steiner problem in networks: A survey. *Networks* **17** (1987), 129 - 167.

Zelinka, B

[Z1] On a certain distance between isomorphism classes of graphs. *Čas. pěst. mat.* **100** (1975), 371 - 373.

[Z2] A distance between isomorphism classes of trees. *Czech. Mat. J.* **33** (1983), 126 - 130.

[Z3] Comparison of various distances between isomorphism classes of graphs. *Čas. pěst. mat.* **110**(1985), 289 - 293.

Zou, H.B.

see [CSZ1] and [COTZ1].