Extremal Problems in Digraphs

Blair D. Sullivan

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mathematics

June 2008

Abstract

Let G be a finite simple directed graph on n vertices. Say G is *m-free* if it has no directed cycles of length at most m. In 1978, Caccetta and Häggkvist [3] conjectured that if G has minimum out-degree at least r, then G is not $\lceil n/r \rceil$-free. Finding upper bounds on the minimum out-degree in 3-free digraphs has been of particular interest in recent research. In this thesis, we present new results for several related problems in extremal directed graph theory.

Let $\beta(G)$ denote the size of the smallest subset $X \subseteq E(G)$ such that $G \setminus X$ has no directed cycles, and let $\gamma(G)$ be the number of non-edges of G. Fix an integer $m \geq 3$; let G be m-free. For $3 \leq m \leq 5$, we prove that $\beta(G) \leq \frac{\gamma(G)}{m-2}$. One consequence when $m = 3$ was a new bound of $.3530831|V(G)|$ on the minimum out-degree in 3-free digraphs [7, 21].

We conjecture that $\beta(G) \leq \frac{2}{(m+1)(m-2)}\gamma(G)$ (this would be best possible if true), and prove this conjecture when $V(G)$ is the union of two tournaments. Furthermore, when G is a circular interval digraph (the vertices of G can be arranged in a circle such that if distinct u, v, w are in clockwise order and uw is a directed edge, then so are both uv, vw), we show that $\beta(G) \leq \frac{\gamma(G)}{2(m-2)}$.

Another approach to Caccetta-Häggkvist has been studying relationships between being m-free and having a bounded number of directed t-vertex paths. We prove that in 3-free digraphs, the number of four-vertex directed paths is at most $\frac{4}{75}|V(G)|^4$. Furthermore, Thomassé conjectures that if G is 2-free, then G has at most $(n-1)n(n+1)/15$ induced three-vertex directed paths. We prove that in circular interval digraphs, an upper bound of $n^3/16$ holds.

We also consider the tradeoff between minimum and average out-degree in 3-free digraphs. In particular, we prove that for 3-free G with average out-degree γn and minimum out-degree ϵn, we have $\gamma \leq \min\left(\frac{2(1-2\epsilon)}{2-c}, \frac{1-\epsilon}{2-c}\right)$ for any constant $c > 0$ so that all 3-free digraphs on $n > 0$ vertices have minimum out-degree at most cn.

Acknowledgements

It would not have been possible to write this thesis without the help and support of many people. It is impossible to mention them all by name here, but I am eternally grateful for the help that each and every one of them provided.

I would like to thank my advisor, Dr. Paul Seymour, who shared his expertise and research insights with me. Words cannot express my gratitude for his support during my graduate career.

Additionally, I would like to express my gratitude to Henry Cohn for his advice on many things and especially his council in the process of finding a job.

I am also grateful to the Department of Homeland Security for enabling me to take advantage of many opportunities by providing financial support through their Graduate Fellowship and Dissertation Grant programs.

Above all, I would like to thank my family for their unwavering support: To my husband Jonathan, for believing in me and helping me through many of the struggles of graduate school; To my sisters, especially Caitlin for providing late-night entertainment during long weeks of research and writing; Lastly, to my parents, Mom & Dad, for always having been there. Thank you for all the sacrifices you made so that I might be here today.

Contents

Chapter 1

Introduction

1.1 Notation and Definitions

All digraphs in this thesis are finite. For a digraph $G = (V(G), E(G))$, $V(G)$ and $E(G)$ denote its vertex and edge sets, respectively. Unless otherwise stated, we assume $|V(G)| = n$. The members of $E(G)$ are ordered pairs of vertices. We use the notation uv to denote an ordered pair of vertices (u, v) (whether or not u and v are adjacent). We only consider digraphs which have no loop edges uu, and at most one directed edge uv for all vertices $u \neq v$ (are *simple*). A *non-edge* in G is an unordered pair of distinct vertices u, v so that both uv, vu are not in $E(G)$. We say a simple digraph G is a *tournament* if for all vertices $u \neq v$, exactly one of uv, vu is an edge.

A *directed walk* in a digraph is a sequence $v_1, e_1, v_2, \dots, e_{t-1}, v_t$ where $v_1, \dots, v_t$ are vertices, and $e_i = v_i v_{i+1}$ is an edge for $i = 1, \dots, t-1$; its *length* is $t-1$. A *directed path* in a digraph is a directed walk where $v_1, \dots, v_t$ are distinct vertices (its length is $t-1$). We may denote a directed walk (or path) as v_1-v_2-$\cdots$-v_t. We say a directed path is *induced* if every edge e from v_i to v_j satisfies $j = i+1$ for $0 \leq i, j \leq t$. We say a digraph G *is a directed path* if its vertex set can be labeled $v_1, \dots, v_n$ and its edges $e_1, \dots, e_{n-1}$ so that $v_1, e_1, v_2, \dots, e_{n-1}, v_n$ is an induced directed path in G.

Given a vertex $v \in V(G)$, we define its *out-neighbors* to be $N^+(v) = \{u : vu \in E(G)\}$ and analogously $N^-(v) = \{u : uv \in E(G)\}$ to be its *in-neighbors*. Let $\delta^+(v) = |N^+(v)|$ and $\delta^-(v) = |N^-(v)|$ denote the *out-degree* and *in-degree*, respectively. More generally, let $N_i^+(v)$ be the set of vertices w so the shortest directed path starting with v and ending with w has length i (so $N_1^+(v) = N^+(v)$). Similarly, let $N_i^-(v)$ be the set of vertices whose shortest directed path to v has length i. Let δ_G^+ denote the *minimum out-degree* in a digraph G.

A *directed cycle of length* t is a digraph whose vertices and edges can be ordered as $v_1, e_1, v_2, \ldots, e_{t-1}, v_t, e_t$ with $v_1, \ldots, v_t$ distinct vertices, e_i the directed edge v_iv_{i+1} for $i = 1, \ldots, t-1$, and $e_t = v_tv_1$. We may denote such a cycle as v_1-v_2-$\cdots$-v_t-v_1. For an integer $k \geq 0$, let us say a digraph G is *k-free* if there is no directed cycle of G with length at most k. A digraph is *acyclic* if it has no directed cycle. The *girth* of a digraph G is the minimum integer g so that G is not g-free.

We now give notation for several common ways to form subgraphs. Given $X \subseteq V(G)$, define $G|X$ to be the graph with vertex set X and edge set $\{uv \in E(G) : u, v \in X\}$. For $Y \subseteq E(G)$, we write $G \setminus Y$ for the graph with vertex set $V(G)$ and edge set $E(G) \setminus Y$. If $Y = \{e\}$ where $e = uv$, then we may abbreviate as $G \setminus Y = G \setminus e = G \setminus uv$. Finally, for a vertex v, we write $G \setminus v$ for the graph with vertex set $V(G) \setminus \{v\}$ and edge set $\{uw \in E(G) : u, w \neq v\}$.

Lastly, if Z is a set of non-edges of G, we write $G + Z$ for the graph with vertex set $V(G)$ and edge set $E(G) \cup \{uv : uv \in Z\}$. Analogously, if $Z = \{f\}$ with $f = uv$, we may write $G + f = G + uv = G + Z$.

1.2 Motivation: Caccetta-Häggkvist

The motivating conjecture for my work is from 1978:

Conjecture 1.2.1 (Caccetta-Häggkvist [3]). *If G is a simple directed graph on n vertices, then $\delta^+_G \geq r$ implies G is not $\lceil \frac{n}{r} \rceil$-free for all positive r.*

One should note that if true, the Caccetta-Häggkvist conjecture is best possible. The typical tight example consists of $n = kr + 1$ vertices on a circle, each adjacent to the next $r = (n-1)/k$, and has no directed cycles of length at most k. These tight examples are members of a larger class of directed graphs, which we define here for future reference.

Definition 1.2.2 (Circular Interval Digraph). A digraph G is a *circular interval digraph* if its vertices can be arranged in a circle such that for every triple u, v, w of distinct vertices, if u, v, w are in clockwise order and $uw \in E(G)$, then $uv, vw \in E(G)$. This is equivalent to saying that the vertex set of G can be numbered as $v_1, \ldots, v_n$ such that for $1 \leq i \leq n$, the set of out-neighbors of v_i is $\{v_{i+1}, \ldots, v_{i+a}\}$ for some $a \geq 0$, and the set of in-neighbors of v_i is $\{v_{i-b}, \ldots, v_{i-1}\}$ for some $b \geq 0$, reading subscripts modulo n.

Due to connections with other problems in graph theory and additive number theory, Caccetta-Häggkvist and its variants have drawn considerable attention from the research community. One well-known related problem is Seymour's Second Neighborhood Conjecture:

Conjecture 1.2.3 (Seymour)**.** *If G is a non-null 2-free digraph, then some $v \in V(G)$ satisfies $|N_2^+(v)| \geq |N_1^+(v)|$.*

A proof of Conjecture 1.2.3 would imply a weakening of the $r = n/3$ case of Caccetta-Häggkvist. In the other direction, Caccetta-Häggkvist would imply the Behzad-Chartrand-Wall Conjecture:

Conjecture 1.2.4 (Behzad, Chartrand, Wall [2])**.** *If G is an digraph with girth g and $\delta^-(v) = \delta^+(v) = r$ for all $v \in V(G)$, then $|V(G)| \geq r(g-1)+1$.*

There are also connections with additive number theory, probably best exemplified by Hamidoune's proof [8] of Caccetta-Häggkvist for Cayley graphs (see Section 1.6 for details).

In spite of this attention, Caccetta-Häggkvist remains open for all n and r where $5 < r < n/2$ and $n < 2r^2 - 3r + 1$ (see [23] for a more detailed survey of progress, also [3, 8, 10, 18] for specific results). Most recent research has concentrated on the special case when $r = n/3$, and more specifically the following approximation problem:

Problem 1.2.5. *Find the minimum constant c so that every digraph G with $\delta_G^+ \geq cn$ is not 3-free.*

Caccetta-Häggkvist would imply that $c = 1/3$ is the minimum. A number of papers slowly improved on the best known upper bound for c, culminating in Shen's result of $3 - \sqrt{7}$ in 1998 [17]. In 2006, our work [5] on a related problem for 3-free digraphs led to the only improvement of this bound in the last ten years (due to Hamburger, Haxell, and Kostochka [7]).

1.3 Cycles in Dense Digraphs

An elementary exercise in graph theory is to show that every 3-free tournament is also acyclic. One might then ask whether more generally 3-free graphs which are "almost" tournaments are also "almost" acyclic. One measure of being "almost" acyclic is how many edges must be removed before the graph becomes acyclic:

Definition 1.3.1. In a digraph G, $X \subseteq E(G)$ is a *feedback arc set* if $G \setminus X$ is an acyclic digraph.

The question is then whether graphs which are "close" to being tournaments (have few non-edges) always have small feedback arc sets. More specifically, we define the following parameters:

Definition 1.3.2. For a digraph G, define

$$\begin{aligned}\gamma(G) &= |\{\{u,v\} \subseteq V(G) : u \neq v,\, uv, vu \notin E(G)\}|,\\ \beta(G) &= \min\{|X| : X \subseteq E(G) \text{ is a feedback arc set}\}.\end{aligned}$$

The problem is then to bound $\beta(G)$ by some function of $\gamma(G)$. In joint work with Chudnovsky and Seymour, we showed the following bound:

Theorem 1.3.3 (Chudnovsky, Seymour, Sullivan [5]). *If G is a 3-free digraph, then $\beta(G) \leq \gamma(G)$.*

The proof can be found in Section 2.2. This theorem yielded an improvement in the best known upper bound for c in Problem 1.2.5. Hamburger, Haxell, and Kostochka [7] used Theorem 1.3.3 to show there is a vertex of out-degree less than $\sqrt{2\beta(G)}$, which they combined with techniques from [18] to show $c \leq .35312$. Subsequently, Shen [21] was able to improve their argument and prove $c \leq .3530381$.

Theorem 1.3.3 has also been used in an approximation algorithm for the rank aggregation problem by Hegde, Jain, Williamson, and Zuylen [24]. Rank aggregation can be viewed as a weighted feedback arc set problem, and they use the recursive algorithm given by our proof to generate a good approximate ranking for certain tournaments.

However, Theorem 1.3.3 does not appear to be tight, and we conjecture:

Conjecture 1.3.4 (Chudnovsky, Seymour, Sullivan [5]). *If G is a 3-free digraph, then $\beta(G) \leq \frac{\gamma(G)}{2}$.*

If true, this is best possible for infinitely many values of $\gamma(G)$ (a construction is given in Section 2.1). One consequence of this conjecture would be a strengthening of Hamburger, Haxell, and Kostochka's result to give a vertex of out-degree less than $\sqrt{\beta(G)}$ (thus improving the best known bound for Problem 1.2.5 to .352062).

We have shown Conjecture 1.3.4 for two significant classes of 3-free digraphs:

Theorem 1.3.5 (Chudnovsky,Seymour,Sullivan [5]). *$\beta(G) \leq \frac{\gamma(G)}{2}$ whenever G is a 3-free digraph with $V(G)$ the disjoint union of V_1, V_2 such that $G|V_i$ is a tournament for $i = 1, 2$ or G is a 3-free circular interval digraph.*

The proof for circular interval digraphs can be found in Section 2.4, and the result when G is the union of two tournaments is given in Section 2.5.

One natural extension of this problem is to ask whether requiring G to be m-free (for $m \geq 3$) gives a bound of the form $\beta(G) \leq f(m)\gamma(G)$ for some decreasing function of m which tends to zero as m tends to infinity. By generalizing the classes of examples where Conjecture 1.3.4 is tight, one can show that $f(m) \geq \frac{2}{(m+1)(m-2)}$, and I suspect this is best possible:

Conjecture 1.3.6. *If G is an m-free digraph, then $\beta(G) \leq \frac{2\gamma(G)}{(m+1)(m-2)}$.*

We have several partial results in this direction, including the following theorems, proven in Sections 2.4 and 2.2.

Theorem 1.3.7. *Let G be an m-free circular interval digraph. Then*

$$\beta(G) \leq \frac{\gamma(G)}{2(m-2)}.$$

Theorem 1.3.8. *Let G be an m-free digraph. Then*

$$\begin{aligned} \beta(G) &\leq \gamma(G)/2 && \text{if } m \geq 4 \\ \beta(G) &\leq \gamma(G)/3 && \text{if } m \geq 5. \end{aligned}$$

Finally, in Section 2.6, we look at improving Theorem 1.3.3 by an additive constant, and prove the following results:

Theorem 1.3.9. *Let G be a 3-free digraph. Then*

$$\begin{aligned} \beta(G) &\leq \gamma(G) - 1 && \text{if } \gamma(G) \geq 1 \\ \beta(G) &\leq \gamma(G) - 2 && \text{if } \gamma(G) \geq 3 \\ \beta(G) &\leq \gamma(G) - 3 && \text{if } \gamma(G) \geq 5. \end{aligned}$$

1.4 Directed Paths in k-free Digraphs

Another approach to Caccetta-Häggkvist has been to study relationships between being k-free and having a bounded number of directed s-vertex walks (or paths).

Definition 1.4.1. Let $W_s(G)$ be the number of distinct directed s-vertex walks in a digraph G, $P_s(G)$ the number of distinct s-vertex directed paths, and $\tilde{P}_s(G)$ the number of distinct induced s-vertex directed paths.

In work with Seymour, we made the following observation: if $|V(G)| = n$, then $W_s(G) \geq (\delta_G^+)^{s-1}n$. Then a bound of the form $W_s(G) \leq (c_{s,t}n)^s$ for t-free G would prove that $\delta_G^+ \leq (c_{s,t})^{\frac{s}{s-1}}n$. Letting $t = 3$ (and noting $s = t+1$ implies $W_s(G) = P_s(G)$ in t-free G), we proved:

Theorem 1.4.2 (Seymour, Sullivan)**.** *If G is a 3-free digraph on n vertices, then $P_4(G) \leq \frac{4}{75}n^4 \approx .0533n^4$. Additionally, there is an infinite family of 3-free graphs where $P_4 \to \frac{1}{20.48}n^4 \approx .0488n^4$ as $n \to \infty$.*

Theorem 1.4.2 gives an upper bound of $c \leq \sqrt[3]{4/75} \approx .3764$ for Problem 1.2.5, which doesn't break the current record, but similar results for other values of s might well lead to improvements.

Another related problem on counting directed paths arises from the following conjecture:

Conjecture 1.4.3 (Thomassé)**.** *If G is a 2-free digraph on n vertices, then*

$$\tilde{P}_3(G) \leq \frac{(n-1)n(n+1)}{15}.$$

The best known upper bound on $\tilde{P}_3(G)$ at this time is $2n^3/25$, due to Bondy [22]. However, for the special class of circular interval digraphs, we proved the following strengthening:

Theorem 1.4.4 (Seymour, Sullivan)**.** *If D is a 2-free circular interval digraph on n vertices, then $\tilde{P}_3(D) \leq n^3/16$.*

The details can be found in Section 3.2.

1.5 Other Parameters in k-free Digraphs

The Caccetta-Häggkvist conjecture asserts that if G is k-free, then it has minimum out-degree less than n/k. A natural class of variations is to replace minimum out-degree with another graph parameter, and see what non-trivial bounds can be derived from the lack of short directed cycles. Initially, one might hope that average out-degree (d_{avg}^+) would be a good alternative measure. However, acyclic tournaments are k-free for all k, yet have $d_{avg}^+ = (n-1)/2$ (the maximum possible in a simple digraph). On the other hand, $\delta_G^+ = 0$ for these tournaments, implying that a combination of d_{avg}^+ and δ_G^+ might be a better measure than either one. We were able to show that in

3-free digraphs, there is such a tradeoff between minimum and average out-degrees. In particular,

Theorem 1.5.1 (Seymour, Sullivan)**.** *Let G be a 3-free digraph on n vertices with $\delta^+_{avg} = \gamma n$ and $\delta^+_G = \epsilon n$. If all 3-free graphs have minimum out-degree less than cn, then:*

$$\gamma \leq \min\left(\frac{2(1-2\epsilon)}{2-c}, \frac{1-\epsilon}{2-c}\right).$$

The proof for this, and several other results in this direction can be found in Chapter 4.

1.6 Additive Number Theory

Finally, we address some connections with additive number theory. Hamidoune [8] proved the Caccetta-Häggkvist conjecture for the following graphs:

Definition 1.6.1 (Cayley graph)**.** Given a finite group $(\Gamma, \cdot)$, and $A \subseteq \Gamma$, the *Cayley graph* Cayley(Γ, A) is the digraph with vertex set Γ and edges consisting of all ordered pairs $(x, x \cdot a)$ with $x \in \Gamma$ and $a \in A$.

His proof used the following non-trivial theorem from additive number theory:

Theorem 1.6.2 (Kemperman [11])**.** *Let $(\Gamma, \cdot)$ be a group, and (A, B) a pair of finite subsets so that $1 \in A \cap B$ and if $a \in A, b \in B$ with $ab = 1$, then $a = b = 1$. Then $|AB| \geq |A| + |B| - 1$.*

Reformulating other problems in extremal directed graph theory for Cayley graphs also leads to interesting problems in additive number theory, some of which are still open. Specifically, I propose that a strengthening of Conjecture 1.3.4 should hold for Cayley graphs with $\Gamma = \mathbb{Z}\backslash p\mathbb{Z}$ (p prime). This can be reformulated as a number theoretic conjecture (see Section 5.2 for details):

Conjecture 1.6.3. *Let p be prime. Then*

$$\min_{1 \leq k \leq p-1} \sum_{i=1}^{d} (ka_i \ mod\ p) \leq \frac{p(p-2d-1)}{4}$$

for all $\{a_1, \dots, a_d\}$ such that the $a_i \in \mathbb{Z}\backslash p\mathbb{Z}$ are pairwise distinct, and none of the equations $x_1 + \cdots + x_\ell = 0$ has a solution in $\{a_1, \dots, a_d\}$ for $\ell = 1, 2, 3$.

One possible approach is to think of this as a problem about heights of points in finite projective space, and in joint work with Mel Nathanson [15], we have some partial results in this direction. The details can be found in Chapter 5. There has already been further research on the conjectures this gives rise to in projective space (see [16], [14]).

Chapter 2

Cycles in Dense Digraphs

2.1 Introduction

We begin with some notation.

Definition 2.1.1. Let G be a digraph. For disjoint subsets $A, B \subseteq V(G)$, let $E(A, B)$ denote the set of directed edges ab with $a \in A$ and $b \in B$. Similarly, let $\bar{E}(A, B)$ be the non-edges between A and B (so $|\bar{E}(A, B)| = |A| \cdot |B| - |E(A, B)| - |E(B, A)|$ in a 2-free digraph). If $A = \{a\}$, we may write $E(a, B), E(B, a)$, or $\bar{E}(a, B)$ for convenience.

We investigated the relationship between $\beta(G)$ and $\gamma(G)$ in 3-free digraphs in [5], and made the following conjecture:

Conjecture 2.1.2. *If G is a 3-free digraph then $\beta(G) \leq \frac{1}{2}\gamma(G)$.*

In this chapter, we present a generalization of this conjecture to m-free digraphs, along with several partial results. Consider the following class of digraphs: Fix an integer $n \geq 1$. Let $m \geq 3$ be an integer, and define G_{mn} to be the digraph with vertex set $\{v_1, \ldots, v_{(m+1)n}\}$, and with edge set as follows (reading subscripts modulo $(m+1)n$):

- $v_i v_j \in E(G_{mn})$ for all i, j, k with $1 \leq k \leq m+1$ and $(k-1)n < i < j \leq kn$,
- $v_i v_j \in E(G_{mn})$ for all i, j, k with $1 \leq k \leq m+1$ and $(k-1)n < i \leq kn < j \leq (k+1)n$.

It is easy to see that G_{mn} is m-free and satisfies $\beta(G_{mn}) = n^2$ (certainly $\beta(G_{mn}) \geq n^2$ since G_{mn} has n^2 directed cycles that are pairwise edge-disjoint), and $\gamma(G_{mn}) =$

$\frac{(m+1)(m-2)}{2}n^2$. Then $\beta(G_{mn})/\gamma(G_{mn}) = \frac{2}{(m+1)(m-2)}$, which is independent of our choice of n. We also note that G_{mn} is a circular interval digraph.

We conjecture that the G_{mn}'s attain the largest $\beta(G)/\gamma(G)$ ratio among all m-free digraphs, noting that Conjecture 2.1.2 is the special case of this when $m = 3$.

Conjecture 2.1.3. *If G is an m-free digraph, with $m \geq 3$, then*

$$\beta(G) \leq \frac{2}{(m+1)(m-2)}\gamma(G).$$

We have not been able to prove Conjecture 2.1.3 in general, and in this chapter we prove a number of partial results. First, in Section 2.2 we show that $\beta(G) \leq \gamma(G)/(m-2)$ for $m = 3, 4,$ and 5. In Section 2.4, when G is a circular interval digraph, we show $\beta(G) \leq \frac{1}{2(m-2)}\gamma(G)$ for all $m \geq 3$. Conjecture 2.1.2 for circular interval digraphs follows as a corollary. We prove Conjecture 2.1.3 for graphs whose vertex set is the union of two tournaments in Section 2.5. Finally, in Section 2.6, we return to the case of $m = 3$ and prove several lemmas of the form: if $\gamma(G) \geq k$, then $\beta(G) \leq \gamma(G) - c_k$, where c_k is a constant depending only on k.

2.2 Some Results for 3-, 4-, and 5-free Digraphs

We begin with a general lemma about partitions in minimal counterexamples.

Lemma 2.2.1. *Let G be a minimal counterexample to being m-free with $\beta(G) \leq c\gamma(G)$ for a positive constant c. Then every partition of $V(G)$ into non-empty sets V_1, V_2 satisfies $|E(V_1, V_2)| > c|\bar{E}(V_1, V_2)|$.*

Proof. Let G, V_1, V_2 satisfy the hypotheses of Lemma 2.2.1. Choose $X_i \subseteq E(G|V_i)$ so that $(G|V_i) \setminus X_i$ is acyclic and $|X_i| = \beta(G|V_i)$, for $i = 1, 2$. Deleting the edge-set $X_1 \cup X_2 \cup E(V_1, V_2)$ will make G acyclic. Then $|X_1| + |X_2| + |E(V_1, V_2)| \geq \beta(G) > c\gamma(G)$. Also, $\gamma(G) = \gamma(G|V_1) + \gamma(G|V_2) + |\bar{E}(V_1, V_2)|$. Substitution now gives

$$\beta(G|V_1) + \beta(G|V_2) + |E(V_1, V_2)| > c\left(\gamma(G|V_1) + \gamma(G|V_2) + |\bar{E}(V_1, V_2)|\right).$$

By the minimality of G, $G|V_1$ and $G|V_2$ satisfy $\beta(G|V_i) \leq c\gamma(G|V_i)$ for $i = 1, 2$. This proves $|E(V_1, V_2)| > c|\bar{E}(V_1, V_2)|$, and completes the proof of Lemma 2.2.1. □

Theorem 2.2.2. *If G is a 3-free digraph then $\beta(G) \leq \gamma(G)$.*

Proof. Let G be a minimal counterexample. We may assume $V(G) \neq \emptyset$. For each vertex v, let $f(v)$ denote the number of triples of distinct vertices (v, y, z) so that v-y-z is an induced directed path. Let $g(v)$ be the number of triples of distinct vertices (x, v, z) so that x-v-z is an induced directed path. Since $V(G) \neq \emptyset$ and $\sum_{v \in V(G)} f(v) = \sum_{v \in V(G)} g(v)$, there exists $v \in V(G)$ such that $f(v) \leq g(v)$. Choose some such vertex v, and let $C = V(G) \setminus (N^+(v) \cup N^-(v) \cup \{v\})$. Consider the partition of $V(G)$ given by $V_1 = N^+(v) \cup \{v\}$ and $V_2 = N^-(v) \cup C$. Since $g(v)$ is the number of pairs (x, z) with $x \in N^+(v)$ and $z \in N^-(v)$ such that x, z are non-adjacent, it follows from Lemma 2.2.1 (with $c = 1$) that

$$|E(V_1, V_2)| > |\bar{E}(V_1, V_2)| \geq g(v). \tag{2.1}$$

Since there is no edge $ab \in E(G)$ with $a \in N^+(v)$ and $b \in N^-(v)$ (because G is 3-free), it follows that every edge in $E(V_1, V_2)$ is of the form yz with $y \in N^+(v)$, and $z \in C$. Then by definition $|E(V_1, V_2)| = f(v)$. Together with (2.1), this implies $f(v) > g(v)$, a contradiction to our choice of v. This proves Theorem 2.2.2. □

Before proving the remaining results of this section, we establish some notation. For a vertex v of a digraph G, let $P(v)$ be the number of triples of distinct vertices (a, b, c) so that for some $x \in V(G)$, a-x-b-c is an induced directed path, and $a = v$. Similarly, let $Q(v)$ be the number of triples of distinct vertices (a, b, c) so that for some $x \in V(G)$, a-x-b-c is an induced directed path, and $b = v$, and $R(v)$ be the number of such triples with $c = v$. Also, let $P'(v)$ be the number of triples of distinct vertices (a, b, c) so that for some $x \in V(G)$, a-b-x-c is an induced directed path, and $a = v$. Again, let $Q'(v)$ and $R'(v)$ be the number of such triples with $b = v$ and $c = v$, respectively.

We note that from their definitions, one can easily verify

$$\sum_{v \in V(G)} P(v) = \sum_{v \in V(G)} Q(v) = \sum_{v \in V(G)} R(v), \tag{2.2}$$

and

$$\sum_{v \in V(G)} P'(v) = \sum_{v \in V(G)} Q'(v) = \sum_{v \in V(G)} R'(v), \tag{2.3}$$

for every digraph G.

Finally, let $C(v)$ be the vertices whose shortest directed path to or from v has length at least three (that is, $C(v) = V(G) \setminus (\{v\} \cup N_1^+(v) \cup N_2^+(v) \cup N_1^-(v) \cup N_2^-(v))$).

Theorem 2.2.3. *If G is a 4-free digraph, then $\beta(G) \leq \frac{\gamma(G)}{2}$.*

Proof. If G is 4-free, then we have the following bounds on $P(v), Q(v), Q'(v)$, and $R'(v)$ in terms of $C(v)$ and $N_i^-(v), N_i^+(v)$ for $i = 1, 2$:

$$\left.\begin{aligned} P(v) &= |E(N_2^+(v), C(v) \cup N_2^-(v))| \\ Q(v) &\le |\bar{E}(N_2^-(v), N_1^+(v))| \\ Q'(v) &\le |\bar{E}(N_2^+(v), N_1^-(v))| \\ R'(v) &= |E(C(v) \cup N_2^+(v), N_2^-(v))| \end{aligned}\right\} \tag{2.4}$$

Now, let G be a minimal counterexample to Theorem 2.2.3 (G so that $|V(G)|$ is minimal among all 4-free digraphs with $\beta(G) > \frac{\gamma(G)}{2}$). Consider the following two partitions of $V(G)$, given by v:

$$\begin{aligned} P_1 &= (\{v\} \cup N_1^+(v) \cup N_2^+(v), N_1^-(v) \cup N_2^-(v) \cup C(v)). \\ P_2 &= (N_1^+(v) \cup N_2^+(v) \cup C(v), \{v\} \cup N_1^-(v) \cup N_2^-(v)). \end{aligned}$$

If any of these four sets is empty, then G is acyclic, and $\beta(G) = 0 \le \gamma(G)$, so we may assume not. We apply Lemma 2.2.1 with $c = 1/2$ and $m = 4$ to P_1 and P_2, using substitutions from (2.4):

$$P(v) > (Q(v) + Q'(v))/2. \tag{2.5}$$

$$R'(v) > (Q'(v) + Q(v))/2. \tag{2.6}$$

Adding together inequalities (2.5) and (2.6), and summing over all vertices, we have:

$$\sum_{v \in V(G)} P(v) + R'(v) > \sum_{v \in V(G)} 2(Q(v) + Q'(v))/2.$$

Yet equations (2.2) and (2.3) tell us this is the same as

$$\sum_{v \in V(G)} P(v) + R'(v) > \sum_{v \in V(G)} 2(P(v) + R'(v))/2,$$

a contradiction. This proves that Theorem 2.2.3 holds for all 4-free digraphs. □

Theorem 2.2.4. *If G is a 5-free digraph, then $\beta(G) \le \frac{\gamma(G)}{3}$.*

Proof. Since G is 5-free, we have the following bounds on $P(v), Q(v), R(v), P'(v)$, $Q'(v)$, and $R'(v)$ in terms of the sets $C(v)$ and $N_i^-(v), N_i^+(v)$ for $i = 1, 2$:

$$\left.\begin{aligned}P(v) &= |E(N_2^+(v), C(v) \cup N_2^-(v))|\\ Q(v) &\leq |\bar{E}(N_2^-(v), N_1^+(v))|\\ R(v) &\leq |\bar{E}(C(v), N_1^-(v))|\\ P'(v) &\leq |\bar{E}(C(v), N_1^+(v))|\\ Q'(v) &\leq |\bar{E}(N_2^+(v), N_1^-(v))|\\ R'(v) &= |E(C(v) \cup N_2^+(v), N_2^-(v))|\end{aligned}\right\} \tag{2.7}$$

Now, let G be a minimal counterexample to Theorem 2.2.4 (G so that $|V(G)|$ is minimal among all 5-free digraphs with $\beta(G) > \frac{\gamma(G)}{3}$). Consider the following two partitions of $V(G)$, given by v:

$$\begin{aligned}P_1 &= (\{v\} \cup N_1^+(v) \cup N_2^+(v), N_1^-(v) \cup N_2^-(v) \cup C(v)).\\ P_2 &= (N_1^+(v) \cup N_2^+(v) \cup C(v), \{v\} \cup N_1^-(v) \cup N_2^-(v)).\end{aligned}$$

If any of these four sets is empty, then G is acyclic, and $\beta(G) = 0 \leq \gamma(G)$, so we may assume not. We apply Lemma 2.2.1 with $c = 1/3$ and $m = 5$ to P_1 and P_2, using substitutions from (2.7):

$$P(v) > (Q(v) + Q'(v) + P'(v))/3. \tag{2.8}$$

$$R'(v) > (Q'(v) + Q(v) + R(v))/3. \tag{2.9}$$

Adding together inequalities (2.8) and (2.9), and summing over all vertices, we have:

$$\sum_{v \in V(G)} P(v) + R'(v) > \sum_{v \in V(G)} (2(Q(v) + Q'(v)) + R(v) + P'(v))/3.$$

Yet equations (2.2) and (2.3) tell us this is the same as

$$\sum_{v \in V(G)} P(v) + R'(v) > \sum_{v \in V(G)} (2(P(v) + R'(v)) + P(v) + R'(v))/3,$$

which simplifies to

$$\sum_{v \in V(G)} P(v) + R'(v) > \sum_{v \in V(G)} P(v) + R'(v),$$

a contradiction. This proves that Theorem 2.2.4 holds for all 5-free digraphs. □

Note that in the proof of Theorem 2.2.2, we find a partition of the vertex set of

G into two non-empty sets X, Y so that $|E(X,Y)| \leq |\bar{E}(X,Y)|$; and given such a partition, Theorem 2.2.2 follows by applying induction to $G|X$ and $G|Y$. Bruce Reed (private communication) asked whether such a partition existed with $|E(X,Y)| \leq \frac{1}{2}|\bar{E}(X,Y)|$ for all 3-free G, and this question is still open (even for the special classes where Conjecture 2.1.2 is known). We ask whether an analogous strengthening could be true for all m:

Conjecture 2.2.5. *If G is an m-free digraph with $|V(G)| \geq 2$, then there is a partition (X,Y) of $V(G)$ with $X, Y \neq \emptyset$, such that the number of edges with tail in X and head in Y is at most $\frac{2}{(m+1)(m-2)}$ times the number of non-adjacent pairs (x,y) with $x \in X$ and $y \in Y$.*

This is still open, yet it is worth noting that our proofs of Theorems 2.2.3 and 2.2.4 produce partitions with $|E(X,Y)| \leq \frac{1}{2}|\bar{E}(X,Y)|$ when the digraph is 4-free, and $|E(X,Y)| \leq \frac{1}{3}|\bar{E}(X,Y)|$ when the digraph is 5-free.

2.3 A Lemma for Circular Interval Digraphs

We now turn to the case of circular interval digraphs, where we show that for all $m \geq 3$, if G is m-free, then $\beta(G) \leq \frac{1}{2(m-2)}\gamma(G)$. The proof is in the next section; in this section we prove a lemma which is the main step of the proof.

First we need some notation. We use $\mathbf{R}_+$ to denote the set of non-negative real numbers. Let $t \geq 1$ be an integer and let $s = mt + 1$. If n is an integer, let n mod s denote the integer n' with $0 \leq n' < s$ such that $n - n'$ is a multiple of s. If $0 \leq i, j < s$ and i, j are distinct, let $q_{ij} > 0$ be minimum such that $(i + q_{ij})$ mod $s = j$ (so $q_{ij} = j - i$ if $j > i$, and $q_{ij} = j - i + s$ if $j < i$). We define $D_s(ij) = \{(i+p) \bmod s : 0 \leq p < q_{ij}\}$. Let E_s denote the set of all ordered pairs ij with $0 \leq i, j < s$ and $j \neq i$ such that $|D_s(ij)| \leq t$, and let F_s be the set of all unordered pairs $\{i, j\}$ such that $0 \leq i, j < s$ and $j \neq i$ and $ij, ji \notin E_s$. For $0 \leq k < s$, let $C_s(k)$ be the set of all pairs $ij \in E_s$ such that $k \in D_s(ij)$.

Lemma 2.3.1. *Let $t > 0, m \geq 3$ be integers, set $s = mt + 1$, and for $0 \leq i < s$ let $n_i \in \mathbf{R}_+$. Then there exists k with $0 \leq k < s$ such that*

$$\sum_{ij \in C_s(k)} n_i n_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_s} n_i n_j.$$

Proof. Let Q_s be the set of all sequences $(n_0, \ldots, n_{s-1})$ of members of $\mathbf{R}_+$. We say that $(n_0, \ldots, n_{s-1}) \in Q_s$ is *good* if there exists k with $0 \leq k < s$ such that

$$\sum_{ij \in C_s(k)} n_i n_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_s} n_i n_j.$$

Thus we must show that every member of Q_s is good. We prove this by induction on t, with our first step being a base case:

Step 1. *If $t = 1$ then every member of Q_s is good.*

Let $t = 1$ and $(n_0, \ldots, n_m) \in Q_s$; we need to show that there exists k with $0 \leq k \leq m$ such that

$$n_k n_{k+1} \leq \frac{1}{2(m-2)} \left(\sum_{i=1}^{m-2} \sum_{j=i+2}^{m} n_i n_j + \sum_{j=2}^{m-1} n_0 n_j \right).$$

We note that the right hand side is a sum of $\frac{(m-1)(m-2)}{2} + (m-2) = \frac{(m-2)(m+1)}{2}$ terms, so applying the arithmetic-geometric mean inequality gives

$$\begin{aligned}
\frac{2}{(m+1)(m-2)} \left(\sum_{i=1}^{m-2} \sum_{j=i+2}^{m} n_i n_j + \sum_{j=2}^{m-1} n_0 n_j \right) &\geq \left(\prod_{i=0}^{m} n_i^{m-2} \right)^{\frac{2}{(m-2)(m+1)}} \\
&= \left(\prod_{i=0}^{m} n_i \right)^{2/(m+1)} \\
&= \left(\prod_{i=0}^{m} n_i n_{i+1} \right)^{1/(m+1)} \\
&\geq \min_i \{ n_i n_{i+1} \}.
\end{aligned}$$

and Step 1 follows, as $\frac{1}{2(m-2)} \geq \frac{2}{(m+1)(m-2)}$ whenever $m \geq 3$. $\square$

Henceforth we assume that $t > 1$. Our next step allows us to restrict to members of Q_s with strictly positive n_i's.

Step 2. *If $(n_0, \ldots, n_{s-1}) \in Q_s$ and some $n_i = 0$ then $(n_0, \ldots, n_{s-1})$ is good.*

By symmetry, we may assume that $n_0 = 0$. Define m_i for $0 \leq i \leq m(t-1)$ as follows:

$$\begin{aligned}
m_0 &= n_{mt};\\
m_i &= n_i \text{ for } 1 \leq i \leq t-1;\\
m_t &= n_t + n_{t+1};\\
m_i &= n_{i+1} \text{ for } t+1 \leq i \leq 2t-2;\\
m_{2t-1} &= n_{2t} + n_{2t+1};\\
\vdots &= \vdots\\
m_i &= n_{i+(m-2)} \text{ for } (m-2)t-(m-4) \leq i \leq (m-1)t-(m-1);\\
m_{(m-1)t-(m-2)} &= n_{(m-1)t} + n_{(m-1)t+1};\\
m_i &= n_{i+(m-1)} \text{ for } (m-1)t-(m-3) \leq i \leq m(t-1).
\end{aligned}$$

From the inductive hypothesis (and since $t > 1$), the sequence $(m_0, \ldots, m_{m(t-1)}) \in Q_{s-m}$ satisfies Lemma 2.3.1, and so there exists k' with $0 \leq k' < s-m$ such that

$$\sum_{ij \in C_{s-m}(k')} m_i m_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_{s-m}} m_i m_j.$$

Let $k = k' + i$ where $it \leq k' < (i+1)t - i$, and $0 \leq i \leq m-1$. Since $n_0 = 0$, in each case it follows easily (we leave checking this to the reader) that

$$\sum_{ij \in C_s(k)} n_i n_j \leq \sum_{ij \in C_{s-m}(k')} m_i m_j.$$

But

$$\sum_{\{i,j\} \in F_{s-m}} m_i m_j = \sum_{\{i,j\} \in F_s} n_i n_j - \sum_{i=1}^{m-2} n_{it} n_{(i+1)t+1} \leq \sum_{\{i,j\} \in F_s} n_i n_j,$$

as we can check by rewriting the left side in terms of the n_i's, expanding and using that $n_0 = 0$. Consequently,

$$\sum_{ij \in C_s(k)} n_i n_j \leq \sum_{ij \in C_{s-m}(k')} m_i m_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_{s-m}} m_i m_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_s} n_i n_j,$$

and so $(n_0, \ldots, n_{s-1})$ is good. This concludes Step 2. □

We now prove a key inequality for members of Q_s.

Step 3. *Let* $(n_0, \ldots, n_{mt}) \in Q_s$*, such that*

$$\sum_{ij \in C_s(mt)} n_i n_j \le \sum_{ij \in C_s(k)} n_i n_j$$

for all k *with* $0 \le k \le mt$*. Then*

$$\sum_{0 \le i < t} (t-i)(n_{mt-i} + n_i) \le \frac{t}{2} \sum_{0 \le i < s} n_i.$$

Let $0 \le k \le t-1$. For $0 \le i \le k$, define

$$a_i = \sum_{k+1 \le j \le i+t} n_j - \sum_{i+(m-1)t+1 \le j \le mt} n_j.$$

Then

$$\sum_{ij \in C_s(k)} n_i n_j - \sum_{ij \in C_s(mt)} n_i n_j = \sum_{0 \le i \le k} a_i n_i.$$

Since the left side of this is non-negative, and $a_0 \le a_1 \le \cdots \le a_k$, it follows that $a_k \ge 0$, that is,

$$\sum_{k+1 \le j \le k+t} n_j - \sum_{k+(m-1)t+1 \le j \le mt} n_j \ge 0. \tag{2.10}$$

Similarly, for $(m-1)t+1+k \le i \le mt$ let

$$b_i = \sum_{i-t \le j \le (m-1)t+k} n_j - \sum_{0 \le j \le i-(m-1)t-1} n_j;$$

then

$$\sum_{ij \in C_s((m-1)t+k)} n_i n_j - \sum_{ij \in C_s(mt)} n_i n_j = \sum_{(m-1)t+1+k \le i \le mt} b_i n_i.$$

Since $b_{mt} \le b_{mt-1} \le \cdots \le b_{(m-1)t+1+k}$, we deduce similarly that $b_{(m-1)t+1+k} \ge 0$, that is,

$$\sum_{(m-2)t+1+k \le j \le (m-1)t+k} n_j - \sum_{0 \le j \le k} n_j \ge 0. \tag{2.11}$$

Summing (2.10) and (2.11), we have

$$\sum_{k+1 \le j \le k+t} n_j - \sum_{k+(m-1)t+1 \le j \le mt} n_j + \sum_{k+(m-2)t+1 \le j \le k+(m-1)t} n_j - \sum_{0 \le j \le k} n_j \ge 0,$$

that is,

$$\sum_{k+(m-1)t+1\le j\le mt} n_j + \sum_{0\le j\le k} n_j \le \sum_{k+1\le j\le k+t} n_j + \sum_{k+(m-2)t+1\le j\le k+(m-1)t} n_j.$$

But both sides are non-negative, and their sum is less than $\sum_{0\le i<s} n_i$, so the left side is less than $\frac{1}{2}\sum_i n_i$. Summing over all k with $0\le k\le t-1$, we deduce that

$$\sum_{0\le i<t}(t-i)(n_{mt-i}+n_i) < \frac{t}{2}\sum_{0\le i<s} n_i.$$

This finishes Step 3. □

Now to complete the proof, let $(n_0,\dots,n_{s-1})\in Q_s$. Choose h with $0\le h<s$ such that $n_h\le n_i$ for all i with $0\le i<s$. Let $n_h=x$, and for $0\le i<s$, define $p_i=n_i-x$. Thus $(p_0,\dots,p_{s-1})\in Q_s$. We may assume that

$$\sum_{ij\in C_s(mt)} p_ip_j \le \sum_{ij\in C_s(k)} p_ip_j$$

for all k with $0\le k\le mt$, by cyclically permuting $n_0,\dots,n_{mt}$.
By Step 2, $(p_0,\dots,p_{s-1})$ is good, since $p_h=0$. Hence

$$\sum_{ij\in C_s(mt)} p_ip_j \le \frac{1}{2(m-2)}\sum_{\{i,j\}\in F_s} p_ip_j. \tag{2.12}$$

But

$$\begin{aligned}\sum_{ij\in C_s(mt)} n_in_j &= \sum_{ij\in C_s(mt)}(p_i+x)(p_j+x)\\ &= \sum_{ij\in C_s(mt)} p_ip_j + \sum_{0\le k<t} x(t-k)(p_{mt-k}+p_k) + |C_s(mt)|x^2\\ &\le \sum_{ij\in C_s(mt)} p_ip_j + \frac{1}{2}tPx + \frac{1}{2}t(t+1)x^2,\end{aligned}$$

where the last inequality is by Step 3, and $P = \sum_{0 \le i < s} p_i$. Moreover,

$$\begin{aligned}
\sum_{\{i,j\} \in F_s} n_i n_j &= \sum_{\{i,j\} \in F_s} (p_i + x)(p_j + x) \\
&= \sum_{\{i,j\} \in F_s} p_i p_j + \sum_{\{i,j\} \in F_s} x(p_i + p_j) + |F_s| x^2 \\
&= \sum_{\{i,j\} \in F_s} p_i p_j + (m-2)tPx + \frac{1}{2}(m-2)(mt+1)tx^2.
\end{aligned}$$

Using inequality (2.12), it now suffices to show:

$$\frac{1}{2}tPx + \frac{1}{2}t(t+1)x^2 \le \frac{1}{2(m-2)} \left((m-2)tPx + \frac{1}{2}(m-2)(mt+1)tx^2 \right).$$

After multiplying out the right hand side, we see the linear x terms cancel, leaving us with:

$$\frac{t+1}{2}tx^2 \le \frac{mt+1}{4}tx^2.$$

This holds whenever $m \ge 3$ and $t \ge 1$, so it follows that $(n_0, \ldots, n_{mt})$ is good. This completes the proof of Lemma 2.3.1. □

We note that this proof would suffice for a strengthening of Lemma 2.3.1, with $\frac{2}{(m-2)(m+1)}$ instead of $\frac{1}{2(m-2)}$, if Step 3 were replaced by the following:

Conjecture 2.3.2. *Let* $(n_0, \ldots, n_{s-1}) \in Q_s$, *such that*

$$\sum_{ij \in C_s(mt)} n_i n_j \le \sum_{ij \in C_s(k)} n_i n_j$$

for all k *with* $0 \le k \le mt$. *Then*

$$\sum_{0 \le i < t} (t-i)(n_{mt-i} + n_i) \le \frac{2t}{m+1} \sum_{0 \le i < s} n_i.$$

2.4 Circular Interval Digraphs

The main result of this section is:

Theorem 2.4.1. $\beta(G) \le \frac{1}{2(m-2)}\gamma(G)$ *for every* m*-free circular interval digraph.*

We start by defining a special kind of circular interval digraph.

Definition 2.4.2. Let $t \geq 1$ and $n_0, \ldots, n_{mt} \geq 0$ be integers. Let $N_0, \ldots, N_{mt}$ be disjoint sets of cardinalities $n_0, \ldots, n_{mt}$ respectively, and $N = N_0 \cup \cdots \cup N_{mt}$. Define $n = |N| = \sum_i n_i$. Let $N = \{v_1, \ldots, v_n\}$, where

$$N_i = \{v_j \ : \ n_0 + n_1 + \cdots + n_{i-1} < j \leq n_0 + n_1 + \cdots + n_{i-1} + n_i\},$$

for $0 \leq i \leq mt$. We define $G(n_0, \ldots, n_{mt})$ to be the circular interval digraph with vertex set N and adjacency as follows:

- for $0 \leq k \leq mt$, if $i < j$ and $v_i, v_j \in N_k$ then $v_i v_j \in E(G)$,
- for $0 \leq h \leq mt$ and $k \in \{(h+i) \bmod n \, ; 1 \leq i \leq t\}$, every vertex in N_h is adjacent to every vertex in N_k.

Our first lemma shows these special digraphs satisfy Theorem 2.4.1.

Lemma 2.4.3. *For all $m \geq 3$, $t \geq 1$ and all choices of $n_0, \ldots, n_{mt} \geq 0$, if $G = G(n_0, \ldots, n_{mt})$ then $\beta(G) \leq \frac{1}{2(m-2)}\gamma(G)$.*

Proof. By Lemma 2.3.1, given $t, n_0, \ldots, n_{mt}$, there exists k with $0 \leq k \leq mt$ such that

$$\sum_{ij \in C_s(k)} n_i n_j \leq \frac{1}{2(m-2)} \sum_{\{i,j\} \in F_s} n_i n_j,$$

with notation as in Section 2.3. But the left side of this is at least $\beta(G)$, since every directed cycle of G contains an edge uv with $u \in N_i$ and $v \in N_j$ for some $ij \in C_s(k)$; and the right side equals $\frac{1}{2(m-2)}\gamma(G)$. This proves Lemma 2.4.3. ☐

We need a few additional definitions. Let us say an m-free circular interval digraph G is *maximal* if G is not a proper subgraph of any m-free circular interval digraph with the same vertex set.

Definition 2.4.4. Given a circular interval digraph G with vertices $v_1, \ldots, v_n$ (in circular order), let us say that $X \subseteq V(G)$ is a *cluster* if it satifies all of the following conditions:

i. X is non-empty.

ii. Every two vertices in X are adjacent.

iii. X can be written in the form $\{v_a, v_{a+1}, \ldots, v_b\}$ for some a, b.

iv. For every vertex $v \notin X$, either $X \subseteq N^+(v)$, or $X \subseteq N^-(v)$, or $X \cap (N^+(v) \cup N^-(v)) = \emptyset$.

Lemma 2.4.5. *Let G be a maximal m-free circular interval graph. Then either G is an acyclic tournament, or G is isomorphic to $G(n_0, \ldots, n_{mt})$ for some choice of $t, n_0, \ldots, n_{mt}$.*

Proof. Let the vertices of G be $v_1, \ldots, v_n$, numbered as in Definition 1.2.2, and throughout we read these subscripts modulo n.

Step 1. *If $N^-(v) = \emptyset$ or $N^+(v) = \emptyset$ for some vertex v, then G is an acyclic tournament.*

Suppose that $N^-(v) = \emptyset$ for some vertex v, say v_1. If $v_k v_j \in E(G)$ for some j, k with $1 \leq j < k \leq n$, then $j > 1$ and v_k, v_1, v_j are in clockwise order, so $v_k v_1 \in E(G)$, a contradiction. Thus G is acyclic, and it follows that G is a subgraph of some acyclic tournament H. An acyclic tournament is an m-free circular interval digraph. But then the maximality of G implies $G = H$. Similarly, if $N^+(v) = \emptyset$ for some vertex v, then G is an acyclic tournament. This concludes Step 1. □

We may therefore assume that $v_i v_{i+1} \in E(G)$ for $1 \leq i \leq n$.

Step 2. *For $1 \leq i \leq n$, if $\{v_i, v_{i+1}\}$ is not a cluster, then $N^+(v_{i+1}) \not\subseteq N^+(v_i)$ and $N^-(v_i) \not\subseteq N^-(v_{i+1})$.*

Let $N^+(v_i) = \{v_{i+1}, \ldots, v_{i+a}\}$, where $a \geq 1$. Suppose that $N^+(v_{i+1}) \subseteq N^+(v_i)$. We will prove $\{v_i, v_{i+1}\}$ must be a cluster. We must have $N^+(v_{i+1}) = \{v_{i+2}, \ldots, v_{i+a}\}$. Let $N^-(v_i) = \{v_{i-b}, \ldots, v_{i-1}\}$, where $b \geq 1$, and $N^-(v_{i+1}) = \{v_{i-c}, \ldots, v_i\}$. We know $c \leq b$; suppose that $c < b$. Then $v_{i-c-1} v_{i+1} \notin E(G)$, and also $v_{i+1} v_{i-c-1} \notin E(G)$ since G is m-free and $v_{i-c-1} v_i, v_i v_{i+1} \in E(G)$. Since $v_{i-c-1} v_i, v_{i-c} v_{i+1} \in E(G)$, it follows that $v_{i-c-1} v_h, v_h v_{i+1} \in E(G)$ for all $h \in \{(i-k) \bmod n : 0 \leq k \leq c\}$. Consequently, the digraph $G' = G + v_{i-c-1} v_{i+1}$ is a circular interval digraph. From the maximality of G, G' is not m-free, and so there exists a path of length at most $m-1$ from v_{i+1} to v_{i-c-1}, but since $N^+(v_{i+1}) \subseteq N^+(v_i)$, there is also a path of length at most $m-1$ from v_i to v_{i-c-1}, a contradiction to G being m-free. This proves that $c = b$, and so $\{v_i, v_{i+1}\}$ is a cluster. Similarly if $N^-(v_i) \subseteq N^-(v_{i+1})$ then $\{v_i, v_{i+1}\}$ is a cluster. This finishes Step 2. □

If X, Y are clusters with $X \cap Y \neq \emptyset$, it follows easily that $X \cup Y$ is a cluster. Consequently every two maximal clusters are disjoint. Since $\{v\}$ is a cluster for every vertex v, it follows that the maximal clusters form a partition of $V(G)$. Let the maximal clusters be $N_0, \dots, N_{s-1}$, numbered in their natural circular order, and let $|N_i| = n_i$ for $0 \le i < s$. From the definition of a cluster, if X, Y are disjoint clusters and there exists $xy \in E(G)$ with $x \in X$ and $y \in Y$, then $xy \in E(G)$ for all $x \in X$ and $y \in Y$; we denote this by $X \to Y$. For $0 \le h < s$, let T_h be the set of all $k \in \{0, \dots, s-1\} \setminus \{h\}$ such that $N_h \to N_k$; then $T_h = \{(h+i) \bmod s \ : 1 \le i \le t_h\}$ say, for some $t_h \ge 0$. Choose h with $0 \le h < s$, and choose i such that $v_i \in N_h$ and $v_{i+1} \in N_{h+1}$. Since $\{v_i, v_{i+1}\}$ is not a cluster (because maximal clusters are disjoint), it follows from Step 2 that $N^+(v_{i+1}) \not\subseteq N^+(v_i)$, and so $t_{h+1} \ge t_h$. Since this holds for all choices of h, and $t_0 \ge t_{s-1}$, we deduce that $t_0 = t_1 = \cdots = t_{s-1}$. Say $t_0 = t$. We claim that $s = mt + 1$. Since G is m-free, it follows that $s \ge mt + 1$. We now prove the reverse inequality. Let $i = n_0$ and $j = n_0 + \cdots + n_t + 1$; thus $v_i \in N_0, v_{i+1} \in N_1$, $v_{j-1} \in N_t$ and $v_j \in N_{t+1}$. Since G is maximal, adding the edge $v_i v_j$ does not result in an m-free circular interval digraph, and it follows that there is a path of length at most $m-1$ from v_j to v_i. Since edges leaving N_i can reach only sets N_j with $j \le i+t$, and there are $s-t-1$ sets between N_{t+1} and N_0, $s-t-1 \le (m-1)t$, or $s \le mt+1$. This proves that $s = mt+1$, and so G is isomorphic to $G(n_0, \dots, n_{mt})$. This proves Lemma 2.4.5. □

We can now prove the main result of this section.

Proof of Theorem 2.4.1. We proceed by induction on $\gamma(G)$. Suppose that G is not a maximal m-free circular interval graph. Then we can add an edge to G forming a m-free circular interval graph G' with $\gamma(G') = \gamma(G) - 1$. Thus $\beta(G') \le \frac{1}{2(m-2)}\gamma(G')$ from the inductive hypothesis. Then

$$\beta(G) \le \beta(G') \le \frac{1}{2(m-2)}\gamma(G') \le \frac{1}{2(m-2)}\gamma(G),$$

as required. Thus we may assume that G is maximal, and that G is not an acyclic tournament. Applying Lemmas 2.4.3 and 2.4.5, this proves Theorem 2.4.1. □

Although this is not tight in general, it is for $m = 3$ (see the definition of G_{3n} with $\beta(G_{3n}) = \gamma(G_{3n})/2$ in Section 2.1):

Corollary 2.4.6. *If G is a 3-free circular interval digraph, then $\beta(G) \le \gamma(G)/2$.*

2.5 The Two Tournaments Result

We first need a technical theorem, and begin with some necessary definitions for this result.

Definition 2.5.1. For integers $m, n \geq 1$, define P_{mn} to be the set of all pairs (i, j) of integers with $1 \leq i \leq m$ and $1 \leq j \leq n$. If $f : P_{mn} \to \mathbf{R}_+$, and $X \subseteq P_{mn}$, we define $f(X)$ to mean $\sum_{x \in X} f(x)$.

Definition 2.5.2. For $(i, j), (i', j') \in P_{mn}$, we say that (i', j') *dominates* (i, j) if $i < i'$ and $j < j'$. Furthermore, if $a, b : P_{mn} \to \mathbf{R}_+$ are functions, we say that b *dominates* a if

i. $a(P_{mn}) = b(P_{mn})$.

ii. For all $X, Y \subseteq P_{mn}$, if $a(X) + b(Y) > a(P_{mn})$ then there exist $x \in X$ and $y \in Y$ such that y dominates x.

The following theorem's proof is given in Appendix B.

Theorem 2.5.3. *Let $m, n \geq 1$ be integers, let $P = P_{mn}$, and let a, b, c, d be functions from P to $\mathbf{R}_+$, satisfying the following:*

i. $a(i, j)b(i', j') \leq c(i', j)d(i, j')$ for $1 \leq i < i' \leq m$ and $1 \leq j < j' \leq n$.

ii. b dominates a.

Then $a(P)b(P) \leq c(P)d(P)$.

We first prove a lemma about digraphs which are the disjoint union of k tournaments.

Lemma 2.5.4. *Let $k \geq 1$ and G be a digraph with $V_1, \dots, V_k$ a partition of $V(G)$ such that $G|V_i$ is a tournament for $1 \leq i \leq k$. Then either G is acyclic, or G has a directed cycle of length at most $2k$; that is, G is not $2k$-free.*

Proof. We may assume G has a directed cycle. Then it has a shortest directed cycle, which must be induced. Yet an induced directed cycle in G contains at most two vertices from each V_i (since they are tournaments), so the cycle has length at most $2k$. This proves Lemma 2.5.4. $\square$

We can now prove the main result of this section.

Theorem 2.5.5. *Let G be a 3-free digraph and let M, N be a partition of $V(G)$ such that $G|M, G|N$ are both tournaments. Then there is a set $X \subseteq E(M, N) \cup E(N, M)$ such that $|X| \leq \gamma(G)/2$, and $G \setminus X$ is acyclic. In particular, $\beta(G) \leq \gamma(G)/2$. If G is also m-free for $m > 3$, then $\beta(G) = 0$.*

Proof. If G is m-free for $m > 3$, then Lemma 2.5.4 proves G is acyclic, so $\beta(G) = 0$, and we may restrict our attention to 3-free G. The second assertion in Theorem 2.5.5 follows immediately from the first, so we just prove the first. Let G be a 3-free digraph and let M, N be a partition of $V(G)$ such that $G|M, G|N$ are both tournaments. Since the restriction of G to M is a 3-free tournament, we can number $M = \{u_1, \dots, u_m\}$ such that $u_i u_{i'} \in E(G)$ for $1 \le i < i' \le m$. The same holds for N, but it is convenient to number its members in reverse order; thus we assume that $N = \{v_1, \dots, v_n\}$, where $v_{j'} v_j \in E(G)$ for $1 \le j < j' \le n$. Let $P = P_{mn}$. For $a = (i, j) \in P$ and $b = (i', j') \in P$, let us say that (a, b) is a *cross* if $v_j u_i, u_{i'} v_{j'} \in E(G)$ and $1 \le i < i' \le m$ and $1 \le j < j' \le n$. Let $A_0 = E(N, M)$ and $B_0 = E(M, N)$. Let k be the minimum cardinality of a subset $X \subseteq A_0 \cup B_0$ such that $G \setminus X$ is acyclic. (Such a number exists since $G \setminus (A_0 \cup B_0)$ is acyclic.)

Step 1. *There are k crosses $(a_1, b_1), \dots, (a_k, b_k)$ such that $a_1, \dots, a_k, b_1, \dots, b_k$ are all distinct.*

Suppose not. Let H be the bipartite graph with vertex set $A_0 \cup B_0$, in which $v_j u_i \in A_0$ and $u_{i'} v_{j'} \in B_0$ are adjacent if $((i, j), (i', j'))$ is a cross. Then H has no k-edge matching, and so by König's theorem, there exists $X \subseteq A_0 \cup B_0$ with $|X| < k$ meeting every edge of H; that is, such that for every cross $((i, j), (i', j'))$, X contains at least one of the edges $v_j u_i, u_{i'} v_{j'}$. We claim that $G \setminus X$ is acyclic. Assume not, and let θ be a directed cycle of $G \setminus X$ with vertices $c_1, \dots, c_t$ in order. We will show that some two edges of θ correspond to a cross, contradicting the choice of X. We may assume that $c_t = v_j$, and none of $v_1, \dots, v_{j-1}$ are vertices of θ. Thus $c_1 \in M$, say $c_1 = u_i$. If $c_2 \in N$, say $c_2 = v_{j'}$, then $j' > j$ and so $c_2 c_t \in E(G)$; but then the vertices c_t, c_1, c_2 are the vertices of a directed cycle of G, contradicting that G is 3-free. Thus $c_2 \in M$. Since $c_t \notin M$, we may choose s with $3 \le s \le t$, minimum such that $c_s \in N$. Let $c_s = v_{j'}$, and $c_{s-1} = u_{i'}$ say. Since $c_2, \dots, c_{s-1} \in M$ and form a directed path in this order, and $G|M$ is acyclic, it follows that $i' > i$. Also, since none of $v_1, \dots, v_{j-1}$ are vertices of θ, it follows that $j' \ge j$. If $j' = j$ then $s = t$ and c_{t-1}, c_t, c_1 are the vertices of a directed cycle, a contradiction; so $j' > j$. Hence $((i, j), (i', j'))$ is a cross, and X contains neither of the edges $v_j u_i, u_{i'} v_{j'}$, a contradiction. Thus $G \setminus X$ is acyclic, and this proves Step 1. □

Let $(a_1, b_1), \dots, (a_k, b_k)$ be crosses as in Step 1. Let $A = \{a_1, \dots, a_k\}$, and $B = \{b_1, \dots, b_k\}$. Let C be the set of all $(i', j) \in P$ such that there exist i, j' with $1 \le i < i' \le m$ and $1 \le j < j' \le n$ satisfying $(i, j) \in A$ and $(i', j') \in B$. Let D be the set of all $(i, j') \in P$ such that there exist i', j with $1 \le i < i' \le m$ and $1 \le j < j' \le n$

satisfying $(i, j) \in A$ and $(i', j') \in B$.

Step 2. $C \cap D = \emptyset$, and $|C| + |D| \leq \gamma(G)$.

First suppose that $(i, j) \in C \cap D$. Since $(i, j) \in C$, there exists $j' > j$ such that $(i, j') \in B$; and since $(i, j) \in D$, there exists $j'' < j$ such that $(i, j'') \in A$. Then $v_{j'}v_{j''} \in E(G)$ since $j'' < j < j'$. Yet $v_{j''}u_i \in E(G)$ since $(i, j'') \in A$; and $u_iv_{j'} \in E(G)$ since $(i, j') \in B$, contradicting that G is 3-free. This proves that $C \cap D = \emptyset$. Moreover, if $(i', j) \in C$, we claim that $u_{i'}, v_j$ are non-adjacent in G. Choose i, j' with $1 \leq i < i' \leq m$ and $1 \leq j < j' \leq n$ such that $(i, j) \in A$ and $(i', j') \in B$. Since $\{v_j, u_i, u_{i'}\}$ is not the vertex set of a directed cycle, it follows that $u_{i'}v_j \notin E(G)$. Similarly, since $\{u_{i'}, v_{j'}, v_j\}$ is not the vertex set of a directed cycle, $v_ju_{i'} \notin E(G)$. This proves that $u_{i'}, v_j$ are non-adjacent. Similarly $u_i, v_{j'}$ are non-adjacent for all $(i, j') \in D$. Since $C \cap D = \emptyset$, it follows that $|C| + |D| \leq \gamma(G)$. This proves Step 2. $\square$

Let $a : P \to \mathbf{R}_+$ be defined by $a(x) = 1$ if $x \in A$, and $a(x) = 0$ if $x \in P \setminus A$; thus, a is the characteristic function of A. Similarly let b, c, d be the characteristic functions of B, C, D respectively.

Step 3. *The functions a, b, c, d satisfy the hypotheses of Theorem 2.5.3.*

First, we check condition *(i)*. If $1 \leq i < i' \leq m$, $1 \leq j < j' \leq n$, and $a(i, j)b(i', j') > 0$, then $(i, j) \in A$ and $(i', j') \in B$; hence $v_ju_i, u_{i'}v_{j'} \in E(G)$, and so $(i', j) \in C$ and $(i, j') \in D$ from the definitions of C, D. Then $c(i', j)d(i, j') = 1 = a(i, j)b(i', j')$, as required.

For condition *(ii)*, note first that $a(P) = k = b(P)$. Let $X, Y \subseteq P$ with $a(X) + b(Y) > a(P) = k$. We recall that $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$ where (a_i, b_i) is a cross for $1 \leq i \leq k$. Thus, $a(X) = |A \cap X|$ is the number of values of $h \in \{1, \ldots, k\}$ such that $a_h \in X$, and similarly $b(Y)$ is the number of h with $b_h \in Y$. Since $a(X) + b(Y) > k$, there exists h such that $a_h \in X$ and $b_h \in Y$, and so b_h dominates a_h. This proves that b dominates a, and completes the proof of Step 3. $\square$

Now, applying Theorem 2.5.3 to a, b, c, d, it follows that $a(P)b(P) \leq c(P)d(P)$, and so $|A||B| \leq |C||D|$. But $|A| = |B| = k$, and so $|C||D| \geq k^2$. Consequently $|C| + |D| \geq 2k$, and hence by Step 2, $k \leq \frac{1}{2}\gamma(G)$. This proves Theorem 2.5.5. $\square$

2.6 Additive Constants

We begin with a few definitions. Let G be a 3-free digraph. For $v \in V(G)$, let $A(v)$ denote the set of out-neighbors of v, $B(v)$ the set of in-neighbors, and $C(v)$ the set of non-neighbors (not including v). Also, let $\alpha(v)$ be the number of induced directed paths of the form v-x-y, and $\delta(v)$ the number of form x-v-y, with $x, y \in V(G)$. We note that $\alpha(v) = |E(A(v), C(v))|$, and $\delta(v) = |\bar{E}(A(v), B(v))|$. Finally, given a digraph G and a vertex v, let $G_1(v) = G|(\{v\} \cup A(v))$ and $G_2(v) = G|(B(v) \cup C(v))$.

Since

$$\sum_{v \in V(G)} \alpha(v) = \sum_{v \in V(G)} \delta(v), \tag{2.13}$$

the set $\Delta = \{v : \alpha(v) \leq \delta(v)\}$ is non-empty for every non-null G.

Theorem 2.6.1. *If G is a 3-free digraph with $\gamma(G) \geq 1$, then $\beta(G) \leq \gamma(G) - 1$.*

Proof. Assume not. Let G be such that $|V(G)|$ is minimal among all 3-free digraphs with $\gamma(G) \geq 1$ and $\beta(G) > \gamma(G) - 1$. By minimality, every vertex must be in a directed cycle, and since G is 3-free this implies $|C(v)| > 0$ for all v. If $\gamma(G) = 1$, then G must be acyclic (every cycle has length at least 4; the smallest cycle is induced and therefore has at least 2 non-edges), so $\beta(G) = 0 = \gamma(G) - 1$. So we may assume $\gamma(G) \geq 2$. Consider a vertex $v^* \in \Delta$. Since $|C(v^*)| \geq 1$, by applying Theorem 2.2.2 to $G_1(v^*)$ and $G_2(v^*)$, we can make G acyclic by deleting a set of size $\beta(G_1(v^*)) + \beta(G_2(v^*)) + |E(A(v^*), C(v^*))|$, which is at most $x = \gamma(G_1(v^*)) + \gamma(G_2(v^*)) + \delta(v^*)$. But $\gamma(G) \geq \gamma(G_1(v^*)) + \gamma(G_2(v^*)) + \delta(v^*) + |C(v^*)|$, so $x \leq \gamma(G) - 1$. This proves Theorem 2.6.1. $\square$

Theorem 2.6.2. *If G is a 3-free digraph with $\gamma(G) \geq 3$, then $\beta(G) \leq \gamma(G) - 2$.*

Proof. Assume not. Let G be such that $|V(G)|$ is minimal among all 3-free digraphs with $\gamma(G) \geq 3$ and $\beta(G) > \gamma(G) - 2$. As in the proof of Theorem 2.6.1, it follows that $C(v)$ is non-empty for all vertices v. We again consider a vertex $v^* \in \Delta$. Applying Theorem 2.2.2 to $G_1 = G_1(v^*)$ and $G_2 = G_2(v^*)$, we can make G acyclic by deleting a set of size $\beta(G_1) + \beta(G_2) + |E(A(v^*), C(v^*))|$, or $\beta(G_1) + \beta(G_2) + \alpha(v^*)$. This is at most $\gamma(G_1) + \gamma(G_2) + \delta(v^*)$, which is bounded above by $\gamma(G) - |C(v^*)|$. So we must have $|C(v^*)| \leq 1$. Furthermore, if $\alpha(v^*) < \delta(v^*)$, then $|C(v^*)| = 0$, a contradiction. Thus $|C(v^*)| = 1$, and $\alpha(v^*) = \delta(v^*)$ for all $v^* \in \Delta$. It now follows from equation (2.13) that $\alpha(v) = \delta(v)$ for all $v \in V(G)$. If $\gamma(G_1(v))$ or $\gamma(G_2(v))$ is at least one for some vertex v, applying Theorem 2.6.1 to it (and Theorem 2.2.2 to the other G_i)

proves

$$\beta(G_1(v)) + \beta(G_2(v)) + |E(A(v), B(v))| \leq \gamma(G_1(v)) + \gamma(G_2(v)) - 1 + \delta(v).$$

Yet the left hand side is at least $\beta(G)$, and the right hand side is at most $\gamma(G) - 1 - |C(v)| = \gamma(G) - 2$, giving a contradiction. To summarize thus far, all vertices v of G satisfy $\alpha(v) = \delta(v)$, $|C(v)| = 1$, and $\gamma(G_1(v)) = \gamma(G_2(v)) = 0$. But now vertices in $A(v)$ each have exactly one non-neighbor, which must be in $B(v)$, so $\delta(v) = |A(v)|$. Analogously, $\delta(v) = |B(v)|$, so G is regular with in-degree and out-degree $(n-2)/2$. However, $\gamma(G) \geq 3$, and every vertex is in at most one non-edge, so $n \geq 6$. Then $|A(v)| = |B(v)| = (n-2)/2 \geq 2$. Since G is 3-free, there are no edges from $A(v)$ to $B(v)$, so every vertex in $B(v)$ must have one non-neighbor and $(n-2)/2 - 1$ out-neighbors in $A(v)$. Since $|B(v)| \geq 2$, let $u, w \in B(v)$. Since $\gamma(G_2(v)) = 0$, either $uw \in E(G)$ or $wu \in E(G)$, yet this is a contradiction since either u or w will then have out-degree $(n-2)/2 + 1$. This proves Theorem 2.6.2. □

Theorem 2.6.3. *If G is a 3-free digraph with $\gamma(G) \geq 5$, then $\beta(G) \leq \gamma(G) - 3$.*

Proof. Assume not, and let G be a counterexample. If there are non-adjacent vertices u, v such that $G + uv$ is 3-free, then we can apply Theorem 2.6.2 to $G + uv$ and prove the desired result. So we may assume that for all non-adjacent u, v, there is a 2-edge directed path from u to v, and one from v to u. If some vertex v has no non-neighbors, then v is not in any directed cycles, and the result follows by deleting v. Now, if some vertex u has out-degree zero, then $C(u) = \emptyset$ (since no 2-edge-path starts at u). Thus we may assume every vertex has positive out-degree and in-degree. We have now shown for every vertex v, the sets $A(v), B(v)$, and $C(v)$ are all non-empty.

For a vertex v, we may choose $X_i \subseteq E(G_i(v))$ so that $|X_i| = \beta(G_i(v))$ and $G_i(v) \setminus X_i$ is acyclic for $i = 1, 2$. Then $G \setminus X(v)$ is acyclic when $X(v) = X_1 \cup X_2 \cup E(A(v), C(v))$. This proves

$$\beta(G) \leq |X(v)| = \beta(G_1(v)) + \beta(G_2(v)) + \alpha(v). \tag{2.14}$$

Also, we see that

$$\gamma(G) = \gamma(G_1(v)) + \gamma(G_2(v)) + \delta(v) + |C(v)| + |\bar{E}(C(v), A(v))|. \tag{2.15}$$

Since G is a counterexample, we can deduce from (2.14) and (2.15) that

$$\begin{aligned} 2 \geq \gamma(G) - \beta(G) \quad &\geq \quad \gamma(G_1(v)) - \beta(G_1(v)) + \gamma(G_2(v)) - \beta(G_2(v)) \\ &\quad + \delta(v) - \alpha(v) + |C(v)| + |\bar{E}(C(v), A(v))|. \end{aligned} \tag{2.16}$$

First, suppose that there is a vertex v^* with $|C(v^*)| = 1$, say $C(v^*) = \{c\}$. For ease of notation, we let $A = A(v^*)$, and $B = B(v^*)$. We also set $\alpha^* = \alpha(v^*)$ and $\delta^* = \delta(v^*)$. Suppose $|E(A, c)| > \delta^*$. Then some vertex u in A must have $uc \in E(G)$ and no non-neighbors in B (thus being adjacent from every vertex in B). Consequently c has no out-neighbor in B, contradicting the existence of a 2-edge path from c to v^*. This proves that $|E(A, c)| \leq \delta^*$. Similarly $|E(c, B)| \leq \delta^*$.

Now let $G_1 = G_1(v^*)$ and $G_2 = G_2(v^*)$. We can rewrite inequality (2.16) as

$$1 \geq \gamma(G_1) - \beta(G_1) + \gamma(G_2) - \beta(G_2) + \delta^* - \alpha^* + |\bar{E}(c, A)|. \tag{2.17}$$

From inequality (2.17) and Theorem 2.2.2, it follows that $\delta^* - \alpha^* \leq 1$, or $|E(A, c)| \geq \delta^* - 1$. By reversing all edges, this also proves $|E(c, B)| \geq \delta^* - 1$. Since $\delta^* \geq \alpha^* |E(c, B)|$, this proves $\delta^* \leq 2$. Using equation (2.15), we have

$$\gamma(G_1) + \gamma(G_2) + |\bar{E}(c, A)| = \gamma(G) - 1 - \delta^* \geq 2, \tag{2.18}$$

since $\gamma(G) \geq 5$. If $|\bar{E}(c, A)| > 0$, then inequality (2.17) implies that $\gamma(G_i) - \beta(G_i) = 0$ for $i = 1, 2$ (and hence $\gamma(G_i) = 0$ by Theorem 2.6.1), and $|\bar{E}(c, A)| \leq 1$, a contradiction to equation (2.18). So $|\bar{E}(c, A)| = 0$. Since $\gamma(G) \geq 5$, equation (2.15) implies $\gamma(G_i) > 0$ for some $i \in \{1, 2\}$, and hence $\gamma(G_i) - \beta(G_i) \geq 1$ (by Theorem 2.6.1). Then $\delta^* - \alpha^* = 0$ by inequality (2.17). Since $|E(A, c)| = \delta^*$, by reversing all edges, we also have $|E(c, B)| \geq \delta^*$, and since $\delta^* \geq \alpha^* |E(c, B)|$ it follows that $\delta^* \leq 1$. This proves $\gamma(G_1) + \gamma(G_2) \geq 3$. But from inequality (2.17), $\gamma(G_1) - \beta(G_1) + \gamma(G_2) - \beta(G_2) \leq 1$, and it follows from Theorems 2.6.1 and 2.6.2 that $\gamma(G_1) + \gamma(G_2) \leq 2$, a contradiction.

Hence we have shown $|C(v)| \geq 2$ for every vertex v. Choose a vertex v so that $\alpha(v) \leq \delta(v)$. By (2.16), we have

$$\begin{aligned} \alpha(v) &= \delta(v), \\ |\bar{E}(C(v), A(v))| &= 0, \\ \beta(G_1(v)) &= \gamma(G_1(v)), \\ \beta(G_2(v)) &= \gamma(G_2(v)). \end{aligned}$$

By Theorem 2.6.1, the last two equalities imply $\gamma(G_1(v)) = \gamma(G_2(v)) = 0$. Let $c \in C(v)$. Then since $|\bar{E}(C(v), A(v))| = 0$, and $G_2(v) = G|(B(v) \cup C(v))$ is a tournament, c has only one non-neighbor, a contradiction. This proves Theorem 2.6.3. $\square$

Chapter 3

Counting Paths

3.1 Introduction

We first study a conjecture of Thomassé on inducted directed paths in simple digraphs.

Conjecture 3.1.1 (Thomassé)**.** *If G is a 2-free digraph on n vertices, then*

$$\tilde{P}_3(G) \leq \frac{(n-1)n(n+1)}{15}.$$

This is tight on the following infinite family of digraphs: Let G_0 be the digraph consisting of a single vertex and no edges. Define G_i for $i \geq 1$ to be the digraph obtained by taking four disjoint copies of G_{i-1} (call them D_1, D_2, D_3, D_4) and forming the digraph with vertex set $V(G_i) = \bigcup_{j=1}^{4} V(D_j)$ and edge set

$$E(G_i) = \left(\bigcup_{j=1}^{4} E(D_j)\right) \cup \{uv : u \in D_j, v \in D_{j+1}, j = 1, 2, 3, 4\},$$

where D_5 means D_1. In other words, arrange four copies of G_{i-1} in a square and put in all edges between consecutive copies in a clockwise direction. It is easy to check inductively that $\tilde{P}_3(G_i) = (n-1)n(n+1)/15$, where $n = 4^i = |V(G_i)|$.

The best known approximate result is due to Bondy; his proof is presented in Appendix A.

Theorem 3.1.2 (Bondy [22])**.** *If G is a 2-free digraph on n vertices, then $\tilde{P}_3(G) \leq \frac{2}{25}n^3$.*

For the class of circular interval digraphs, we prove an upper bound of $n^3/16$ in Theorem 3.2.1.

In the second half of this chapter, we study a problem of (non-induced) directed paths in t-free digraphs:

Problem 3.1.3. *Given integers s, t, find the smallest $c_{s,t}$ so that $\sqrt[s]{W_s(G)} \leq c_{s,t}n$ for every t-free digraph G on n vertices.*

The connections to Caccetta-Häggkvist are explained in Section 3.3, and we prove in Theorem 3.3.2 that $c_{4,3} \leq \sqrt[4]{\frac{4}{75}}$. We note that a lower bound on $c_{4,3}$ is given by an infinite family of 3-free graphs where $P_4(G) \to \frac{1}{20.48}n^4 \approx .0488n^4$ as $n \to \infty$ $\left(\text{so } c_{4,3} \geq \sqrt[4]{\frac{1}{20.48}}\right)$. These graphs are given by taking four acyclic tournaments $S_1, \ldots, S_4$, each on $n/4$ vertices and adding the edges uv where $u \in S_i$ and $v \in S_{i+1}$ for $i = 1, 2, 3$, as well as those from S_4 to S_1.

3.2 Induced 3-vertex Paths in Circular Interval Digraphs

The main result of this section is:

Theorem 3.2.1. *If G is a 2-free circular interval digraph on n vertices, then $\tilde{P}_3(G) \leq n^3/16$.*

We first show this is best possible. Let G be a circular interval digraph.

Definition 3.2.2. For $u, v \in V(G)$ let

$$d(u,v) = \begin{cases} 1 + |\{w : u, w, v \text{ are distinct vertices in clockwise order in } G\}| & \text{if } u \neq v \\ 0 & \text{if } u = v. \end{cases}$$

For any edge or non-edge uv, we say it has *length* $d(u, v)$.

Definition 3.2.3. For integer β, let G_β be the circular interval digraph on n vertices with $E(G_\beta) = \{uv : 0 < d(u,v) \leq \beta\}$.

Lemma 3.2.4. *There are circular interval digraphs on n vertices with $n^3/16$ induced 3-vertex paths.*

Proof. Let n be chosen so that $\beta = (3n-4)/8$ is an integer. A straightforward computation shows the number of induced 3-vertex paths in G_β is $n(n-2\beta-1)(2\beta-n/2+1)$. Then $G_{(3n-4)/8}$ has $(n-(3n-4)/4-1)((3n-4)/4-n/2+1) = n^3/16$ induced 3-vertex paths. □

To prove Theorem 3.2.1, we first need a few definitions and lemmas.

Definition 3.2.5. Let G be a 2-free circular interval digraph. If G has a non-edge, we define α_G to be the length of a shortest non-edge in G, and otherwise set $\alpha_G = \infty$. If G has an edge, we define β_G to be the length of a longest edge in G, and otherwise set $\beta_G = 0$.

Observation 3.2.6. *Let $G = (V, E)$ be a 2-free circular interval digraph. Let X be a set of longest edges in G and Y a set of shortest non-edges in G so that uv and vu are not both in Y for any $u, v \in V(G)$. Then $G \setminus X$ and $G + Y$ are 2-free circular interval digraphs. Additionally, if $\alpha_G \leq \beta_G$, then the digraph $(G \setminus X) + Y$ is also a 2-free circular interval digraph.*

Definition 3.2.7. Let $\xi(G)$ denote the number of pairs (uv, wx) where uv is an edge of G, wx is a non-edge and $d(u, v) > d(w, x)$ (u, v are not necessarily distinct from w, x).

Definition 3.2.8 (Optimal Digraph)**.** For a fixed $n \geq 4$, say a digraph G is *optimal* if among all 2-free circular interval digraphs on n vertices, it has the maximum number of 3-vertex induced directed paths and subject to this, $\xi(G)$ is minimum.

We now show optimal digraphs do not have edges with length at least $n/2$.

Lemma 3.2.9. *If G is an optimal digraph on n vertices, then $\beta_G < n/2$.*

Proof. If G has no edges, then $\beta_G = 0 < n/2$. We may assume $E(G) \neq \emptyset$. Suppose $\beta = \beta_G \geq n/2$ and let $e = uv$ be an edge of length β. Let $G' = G \setminus e$, which is also a 2-free circular interval digraph by Observation 3.2.6. Define $c = |N^+(v) \cap N^-(u)|$. Then

$$\tilde{P}_3(G') = \tilde{P}_3(G) - (\delta^-(u) + \delta^+(v) - 2c) + (\beta - 1 + c). \tag{3.1}$$

Since $N^+(v), N^-(u) \setminus N^+(v), \{u, v\}$, and $\{w : u, w, v \text{ are in clockwise order}, w \neq u, v\}$ are disjoint sets in $V(G)$, we have

$$\delta^+(v) + (\delta^-(u) - c) + 2 + (\beta - 1) \leq n. \tag{3.2}$$

Rearranging (3.2) gives $\delta^-(u) + \delta^+(v) \leq n - 1 - \beta + c$, and substituting for $\delta^-(u) + \delta^+(v)$ in (3.1) gives $\tilde{P}_3(G') \geq \tilde{P}_3(G) - (n - 1 - \beta + c - 2c) + (\beta - 1 + c)$ or

$$\tilde{P}_3(G') \geq \tilde{P}_3(G) - (n - 2\beta - 2c).$$

Since $\tilde{P}_3(G') \leq \tilde{P}_3(G)$ because G is optimal, we have $n - 2\beta - 2c \geq 0$. Since $\beta \geq n/2$, it follows that $\beta = n/2$, $c = 0$, and $\tilde{P}_3(G') = \tilde{P}_3(G)$. Hence $d(v, u) = n - d(u, v) = n/2 \geq 2$ ($n \geq 4$ since G is optimal). Then there is at least one vertex w so that u, v, w appear in clockwise order. Since $c = 0$, one of vw, wu must be a non-edge, and thus $\alpha_G < n/2$. But then $\xi(G') < \xi(G)$, contradicting the optimality of G. This proves Lemma 3.2.9. □

We now prove a straightforward lemma giving upper and lower bounds on the vertex degrees in an optimal digraph.

Lemma 3.2.10. *For every vertex v in an optimal digraph G,*

$$\alpha_G - 1 \leq \delta^+(v), \delta^-(v) \leq \beta_G.$$

Proof. Let v be a vertex in an optimal digraph G on n vertices. Certainly α_G is finite, as otherwise G has no induced directed paths of length greater than one. Suppose $\delta^+(v) < \alpha_G - 1$. We first show that v has a non-neighbor. We have $\alpha_G \leq \beta_G + 1$, since if $v_i v_j$ is a shortest non-edge then $v_i v_{j-1}$ is an edge, and therefore has length at most β_G. Then since $\beta_G < n/2$ by Lemma 3.2.9, it follows that $\alpha_G < n/2 + 1$. Now $\delta^+(v) \leq \alpha_G - 2$ implies $\delta^+(v) < n/2 - 1$. Since v has in-degree at most $(n-1)/2$ by Lemma 3.2.9, it has less than $(n/2 - 1) + (n-1)/2 = n - 3/2$ neighbors. Consequently v has a non-neighbor, and we let u be the first vertex following v in the clockwise order for which vu is a non-edge. Then vu has length $\delta^+(v) + 1 < \alpha_G$, a contradiction to the definition of α_G. Analogously, $\delta^-(v) \geq \alpha_G - 1$. Now, suppose $\delta^+(v) > \beta_G$. Then the edge from v to its last clockwise out-neighbor has length $1 + (\delta^+(v) - 1) > \beta_G$. This contradicts the definition of β_G. Again, $\delta^-(v) \leq \beta_G$ by an analogous argument. □

This allows us to give a lower bound on α_G in optimal digraphs G.

Lemma 3.2.11. *If G is an optimal digraph on n vertices, then $\alpha_G > n/4$.*

Proof. If G has no non-edges, then $\alpha_G = \infty > n/4$. We may assume G has a non-edge. Suppose $\alpha = \alpha_G \leq n/4$ and let $e = uv$ be a non-edge of length α. Let $G' = G + e$, which is also a 2-free circular interval digraph by Observation 3.2.6. Define $c = |N^+(v) \cap N^-(u)|$. Then

$$\tilde{P}_3(G') = \tilde{P}_3(G) - (\alpha - 1 + c) + (\delta^-(u) + \delta^+(v) - 2c).$$

Since $\tilde{P}_3(G') \leq \tilde{P}_3(G)$ because G is optimal, we have

$$\alpha - 1 + 3c \geq \delta^-(u) + \delta^+(v). \tag{3.3}$$

Suppose $c = 0$. Then since $\delta^+(v), \delta^-(u) \geq \alpha - 1$ by Lemma 3.2.10, we have $\alpha = 1$ and $\delta^-(u) = \delta^+(v) = 0$. Letting w be the vertex immediately following v in the circular order, we see that $G + \{uv, vw\}$ has more induced 2-edge paths than G, contradicting its optimality. Thus $c > 0$. Now $N^+(v), N^-(u) \setminus N^+(v)$, $\{w : u, w, v$ are in clockwise order, $w \neq u, v\}$, and $\{u, v\}$ form a partition of $V(G)$, so

$$\delta^+(v) + \delta^-(u) - c + (\alpha - 1) + 2 = n. \tag{3.4}$$

We observe that Lemmas 3.2.9 and 3.2.10 imply

$$\delta^+(v) + \delta^-(u) \leq 2\beta \leq n - 1. \tag{3.5}$$

Taking the combination (3.3)+(3.4)−2·(3.5) and simplifying gives $4\alpha \geq n$, so $\alpha = n/4$ and we have equality in both (3.3) and (3.5). The equality in (3.5) implies $\beta_G \geq (n-1)/2$, so $\beta_G > \alpha = n/4$. It follows that $\xi(G') < \xi(G)$. Yet the equality in (3.3) tells us $\tilde{P}_3(G) = \tilde{P}_3(G')$, contradicting the optimality of G. This proves Lemma 3.2.11. □

Lemma 3.2.12. *If G is an optimal digraph and uv is a shortest non-edge in G, then $N^+(v) \cap N^-(u) \neq \emptyset$.*

Proof. Suppose not, and let uv be a non-edge of length α_G in an optimal digraph G on n vertices. Then by Lemma 3.2.11 and the fact that $n \geq 4$, $\alpha_G > n/4 \geq 1$. Let $G' = G + uv$ and $\alpha = \alpha_G$. Then

$$\tilde{P}_3(G') = \tilde{P}_3(G) + \delta^+(v) + \delta^-(u) - (\alpha - 1).$$

Since $\tilde{P}_3(G') \leq \tilde{P}_3(G)$ by optimality, $\delta^+(v) + \delta^-(u) \leq \alpha - 1$. But by Lemma 3.2.10, $\delta^+(v), \delta^-(u) \geq \alpha - 1$. Then $2\alpha - 2 \leq \alpha - 1$, or $\alpha \leq 1$, a contradiction. This proves Lemma 3.2.12. □

We can now prove that in an optimal digraph G, $\alpha_G + \beta_G$ is approximately $3|V(G)|/4$.

Definition 3.2.13. For an circular interval digraph G, we let

$$\epsilon_G = \begin{cases} 0 & \beta_G > \alpha_G \\ 1 & \beta_G \leq \alpha_G. \end{cases}$$

Lemma 3.2.14. *If G is an optimal digraph on n vertices, then*

$$\frac{3n}{4} - \frac{1}{2} - \frac{\epsilon_G}{4} < \alpha_G + \beta_G < \frac{3n}{4} + \frac{1}{2} + \frac{\epsilon_G}{4}.$$

Additionally, if some vertex is incident with a longest edge, but with no shortest non-edge, then $\alpha_G + \beta_G < 3n/4 + \epsilon_G/4$, and if some vertex is incident with a shortest non-edge but with no longest edge, then $\alpha_G + \beta_G > 3n/4 - \epsilon_G/4$.

Proof. Let G be an optimal digraph on n vertices with $\alpha = \alpha_G$, $\beta = \beta_G$, and $\epsilon = \epsilon_G$.

Step 1. $\alpha + \beta < 3n/4 + 1/2 + \epsilon/4$, *and if some vertex is incident with a longest edge but no shortest non-edge, then* $\alpha + \beta < 3n/4 + \epsilon/4$.

If G has no edges, then $\beta = 0$, and since $\alpha \leq \beta + 1$ by Lemma 3.2.10, we have $\alpha + \beta \leq 1 \leq 3n/4$ (since $n \geq 4$), as required. Thus $E(G) \neq \emptyset$. Let uv be a longest edge in G and $G' = G \setminus uv$. For notational convenience, set $\delta^+(v) = a$, $\delta^-(u) = b$, and $|N^+(v) \cap N^-(u)| = c$. The number of induced 3-vertex paths in G' is

$$\tilde{P}_3(G') = \tilde{P}_3(G) + (b - c) + (a - c) + (\beta - 1 + c).$$

Since G is optimal, $\tilde{P}_3(G') \leq \tilde{P}_3(G)$, and strict inequality holds if $\beta > \alpha$ (because then $\xi(G') < \xi(G)$), it follows that:

$$3c < a + b - \beta + 1 + \epsilon. \tag{3.6}$$

Suppose that no vertex is non-adjacent to both u and v. Then counting the vertices, we have $(c + \beta - 1) + (a - c) + (b - c) + 2 = n$, or $c = a + b + \beta + 1 - n$. Substituting for c in (3.6) gives $2(a + b + 1) + 3(\beta - n) < \epsilon - \beta$, or

$$4\beta + 2 + 2(a + b) < 3n + \epsilon.$$

Since $a, b \geq \alpha - 1$ by Lemma 3.2.10, it follows that $4\beta + 2 + (4\alpha - 4) < 3n + \epsilon$, or

$$\alpha + \beta < 3n/4 + 1/2 + \epsilon/4.$$

Note that if $a \geq \alpha$ or $b \geq \alpha$ (one of the endpoints of uv is not incident with a shortest non-edge), we have $4\beta + 2 + (4\alpha - 2) < 3n + \epsilon$, or

$$\alpha + \beta < 3n/4 + \epsilon/4.$$

Thus we may assume there is a vertex y non-adjacent to u and v. In this case, $a + b + (\beta - 1) + 3 \leq n$, so $2\alpha + \beta \leq n$. We know $\alpha \geq (n+1)/4$ by Lemma 3.2.11 (since $\alpha \in \mathbb{Z}$), proving

$$\alpha + \beta < 3n/4 - 1/4 < 3n/4 + \epsilon/4.$$

This proves Step 1. $\square$

Step 2. $\alpha + \beta > 3n/4 - 1/2 - \epsilon/4$, *and if some vertex is incident with a shortest non-edge but no longest edge, then* $\alpha + \beta > 3n/4 - \epsilon/4$.

If G has no non-edges, then $\alpha = \infty$, yet $\alpha \leq \beta + 1 < n/2 + 1$ by Lemmas 3.2.10 and 3.2.9, a contradiction. Then let uv be a shortest non-edge in G, and $G' = G + uv$. For notational convenience, set $\delta^+(v) = a$, $\delta^-(u) = b$, and $|N^+(v) \cap N^-(u)| = c$. The number of induced 3-vertex paths in G' is

$$\tilde{P}_3(G') = \tilde{P}_3(G) + (b - c) + (a - c) - (c + \alpha - 1).$$

Since G is optimal, $\tilde{P}_3(G') \leq \tilde{P}_3(G)$, and strict inequality holds if $\beta > \alpha$ (because then $\xi(G') < \xi(G)$), it follows that:

$$a + b - \alpha + 1 < 3c + \epsilon. \tag{3.7}$$

We know $\alpha + 1 + a + b - c = n$ by Lemma 3.2.12. We solve for $c = \alpha + a + b + 1 - n$, and substitute into equation (3.7). This gives $a + b - \alpha + 1 - \epsilon < 3(\alpha + a + b + 1 - n)$, or

$$3n < 4\alpha + 2(a + b) + 2 + \epsilon.$$

Since $a, b \leq \beta$ by Lemma 3.2.10, $3n < 4(\alpha + \beta) + 2 + \epsilon$, or

$$\alpha + \beta > 3n/4 - 1/2 - \epsilon/4.$$

If $a < \beta$ or $b < \beta$ (one of the endpoints of uv is not incident with a longest edge), we instead have $3n < 4\alpha + 2\beta + 2(\beta - 1) + \epsilon$, or

$$\alpha + \beta > 3n/4 - \epsilon/4,$$

as desired. This proves Step 2, and completes the proof of Lemma 3.2.14. $\square$

Definition 3.2.15. Let G be a circular interval digraph on n vertices. We define $\gamma_G = 4(\alpha_G + \beta_G) - 3n$.

Lemma 3.2.16. *Let G be an optimal digraph on n vertices with $\beta_G > \alpha_G$. Then $-1 \leq \gamma_G \leq 1$. Furthermore, $\gamma_G = -1$ if some vertex is incident with a longest edge and with no shortest non-edge, and $\gamma_G = 1$ if some vertex is incident with a shortest non-edge and with no longest edge.*

Proof. Since $\gamma_G = 4(\alpha_G + \beta_G) - 3n$ and $\epsilon_G = 0$, Lemma 3.2.14 implies

$$-2 = 4(3n/4 - 1/2) - 3n < \gamma_G < 4(3n/4 + 1/2) - 3n = 2.$$

Since γ_G is an integer, this is equivalent to $-1 \leq \gamma_G \leq 1$. Furthermore, if some vertex is incident with a longest edge and no shortest non-edge, Lemma 3.2.14 proves $\gamma_G < 0$. Since $\gamma_G \geq -1$, this implies $\gamma_G = -1$. Similarly, if some vertex is incident with a shortest non-edge and no longest edge, we have $\gamma_G > 0$, which combined with $\gamma_G \leq 1$ implies $\gamma_G = 1$. This proves Lemma 3.2.16 $\square$

We now prove several more facts about optimal digraphs.

Definition 3.2.17. Given a digraph G, say a set of vertices $X \subseteq V(G)$ is *stable* if $uv \notin E(G)$ for all $u, v \in X$; that is, $G|X$ has no edges.

Lemma 3.2.18. *If G is an optimal digraph, it has no stable set of size at least* 3.

Proof. Suppose not and let $|V(G)| = n$. Take a stable set $\{u, v, w\}$ so that $d(u, v)$ is minimum, and let $k = d(u, v)$. Then by Lemma 3.2.11, $\alpha_G > n/4 \geq 1$, so every vertex has at least one out-neighbor. It follows that $k \geq 2$. Let $G' = G + uv$. Then G' is a circular interval digraph, since $d(u, v)$'s minimality implies uw and wv are edges for all w between u and v in the circular order. Since $|N^+(v) \cap N^-(u)| = 0$, it follows that

$$\tilde{P}_3(G') = \tilde{P}_3(G) + \delta^+(v) + \delta^-(u) - (k - 1).$$

Since $\tilde{P}_3(G') \leq \tilde{P}_3(G)$ by optimality, $\delta^+(v) + \delta^-(u) \leq k - 1$.

Let y be the furthest out-neighbor of v in G. Then v is non-adjacent to the next vertex in the circular order (call it a), and the non-edge va is part of a stable set of size three, namely $\{v, a, u\}$ (if u were adjacent to a then there could not be a vertex w to which both u and v were non-adjacent). This means the length of va is at least k by choice of uv, so $\delta^+(v) \geq k - 1$. An analogous argument shows $d^-(u) \geq k - 1$. Since $\delta^+(v) + \delta^-(u) \leq k - 1$, it follows that $k = 1$, a contradiction. This proves Lemma 3.2.18. □

Definition 3.2.19. Let G be an optimal digraph. Say a pair of vertices uv is *extreme* if uv is a longest edge or a shortest non-edge in G.

Lemma 3.2.20. *Let G be an optimal digraph with $\beta_G > \alpha_G$, and u, v, w vertices appearing in clockwise order. Then not all of uv, vw, wu are extreme pairs. Additionally, if some two of them are extreme, then either all three pairs are edges, or two are edges and the third is a shortest non-edge.*

Proof. Let G be an optimal digraph on n vertices with $\beta_G > \alpha_G$. We begin by proving that no vertex is in two shortest non-edges. Assume not, and let v be a vertex with uv and vw non-edges of length α_G. Hence u, v, w appear in clockwise order. Then by Lemma 3.2.18, u and w must be adjacent. If $uw \in E(G)$, G is not a circular interval graph, a contradiction. Thus $wu \in E(G)$, and say it has length L. Let $G' = G + vw$. This is a circular interval digraph by Observation 3.2.6. Then

$$\tilde{P}_3(G') = \tilde{P}_3(G) - (\alpha_G - 1 + \delta^+(w) - L) + L + (\alpha_G - 1) - (\delta^+(w) - L).$$

Since $\tilde{P}_3(G') \leq \tilde{P}_3(G)$ by optimality, $3L - 2\delta^+(w) \leq 0$, or

$$L \leq \frac{2\delta^+(w)}{3} \leq \frac{2\beta_G}{3}.$$

Now since $2\alpha_G + L = n$, we have

$$n \leq 2\alpha_G + \frac{2\beta_G}{3}. \tag{3.8}$$

We note that since $\beta_G > \alpha_G$ and v is incident with two shortest non-edges, v cannot be incident with a longest edge. Then Lemma 3.2.16 gives $\gamma_G = 1$, or

$$\alpha_G + \beta_G = \frac{3n + 1}{4}. \tag{3.9}$$

Combining equations (3.8) and (3.9), we have

$$n \leq \frac{2(\alpha_G + \beta_G)}{3} + \frac{4\alpha_G}{3} = \frac{3n+1}{6} + \frac{4\alpha_G}{3}.$$

This implies

$$\alpha_G \geq \frac{3n-1}{8}.$$

Then since $\alpha_G + \beta_G = (3n+1)/4$, $\beta_G \leq (3n+3)/8$. Yet $\beta_G > \alpha_G$, and both are integers. This implies there are two integers in the range $[(3n-1)/8, (3n+3)/8]$, a contradiction. This proves no vertex is incident with two shortest non-edges.

We now show no triple of vertices forms three longest edges. Let u, v, w be in clockwise order, and assume uv, vw, wu are all edges of length β_G. Then $3\beta_G = n$. Since $\beta_G > \alpha_G$, this implies $n > 2\beta_G + \alpha_G$. Now, Lemma 3.2.14 gives $\alpha_G + \beta_G > 3n/4 - 1/2$ (since $\epsilon_G = 0$), so $n > \beta_G + 3n/4 - 1/2$, or $\beta_G < n/4 + 1/2$. Then $\alpha_G < \beta_G < (n+2)/4$ implies $\alpha_G < (n+1)/4$, a contradiction to Lemma 3.2.11.

This proves that for u, v, w in clockwise order, not all of uv, vw, wu are extreme pairs. Additionally, it proves that if two are extreme pairs, at least one must be a longest edge. To complete the proof of the theorem, we need to show that if two of uv, vw, wv are longest edges, the third pair must be an edge and that if there is a shortest non-edge among uv, vw, wv, the other two pairs must be edges.

We first prove there do not exist $u, v, w \in V(G)$ so that uv, vw are edges of length β_G and wu is a non-edge. Suppose such u, v, w exist. It follows that u, v, w are in clockwise order. Since all three pairs are not extreme, wu has length $L > \alpha_G$. Then $2\beta_G + L = n$, or $2\beta_G + \alpha_G < n$. Since v is incident with two longest edges, it cannot be incident with a shortest non-edge, and Lemma 3.2.16 implies $\gamma_G = -1$, or $\alpha_G + \beta_G = (3n-1)/4$. We then have $(3n-1)/4 + \beta_G < n$, or $\beta_G < (n+1)/4$. Since $\alpha_G < \beta_G$, this contradicts Lemma 3.2.11.

Finally, suppose there are $u, v, w \in V(G)$ so that uv, vw, and wu consist of a shortest non-edge, a longest edge, and a non-edge of length $L > \alpha_G$. Then $\alpha_G + \beta_G + L = n$ implies $2\alpha_G + \beta_G < n$. Since $\alpha_G + \beta_G > (3n-2)/4$ by Lemma 3.2.14, we see that $\alpha_G + (3n-1)/4 < n$, or $\alpha_G < (n+1)/4$. Again, this contradicts Lemma 3.2.11.

This proves Lemma 3.2.20. □

Definition 3.2.21 (Alternating Sequence)**.** Given an optimal digraph G, let $S = v_1, f_1, v_2, f_2, v_3, f_3, \ldots, f_k, v_{k+1}$ be a sequence where the v_i are vertices of G, and the $f_i = v_i v_{i+1}$ are extreme pairs of G. We say S is an *alternating sequence* if it satisfies the conditions:

i. $f_i \neq f_j$ for $i \neq j$.

ii. For $1 \leq i \leq k-1$, if f_i is an edge, then f_{i+1} is a non-edge.

iii. For $1 \leq j \leq k-1$, if f_j is a non-edge, then f_{j+1} is an edge.

In other words, S is an alternating sequence of longest edges and shortest non-edges. Let k be the *length* of S, and define X_S to be the set of longest edges in S and Y_S to be the set of shortest non-edges.

Definition 3.2.22 (Augmenting Sequence)**.** Say a sequence S is an *augmenting sequence* if it is a maximal alternating sequence.

Lemma 3.2.23. *Let G be an optimal digraph on n vertices with $\beta_G > \alpha_G$. If $\alpha_G + \beta_G \leq 3n/4$, then every shortest non-edge has a longest edge incident with each of its endpoints. If $\alpha_G + \beta_G \geq 3n/4$, every longest edge has a shortest non-edge incident with each of its endpoints. Consequently, every augmenting sequence in G has length at least* 3.

Proof. Let G be an optimal digraph on n vertices with $\beta_G > \alpha_G$. If some longest edge uv is not incident with a shortest non-edge at both u and v, since $\beta_G > \alpha_G$, Lemma 3.2.14 gives $\alpha_G + \beta_G > 3n/4$. Similarly, if some shortest non-edge uv is not incident with a longest edge at both u and v, Lemma 3.2.14 gives $\alpha_G + \beta_G < 3n/4$. Clearly, these cannot hold simultaneously. This proves Lemma 3.2.23. □

Lemma 3.2.24. *Let G be an optimal digraph with $\beta_G > \alpha_G$. For every augmenting sequence $S = v_1, f_1, v_2, f_2, v_3, f_3, \ldots, f_k, v_{k+1}$ in G, $v_i \neq v_j$ for $i \neq j$, except possibly $v_{k+1} = v_1$.*

Proof. Suppose not, and let $v_h = v_j$ with $1 \leq h, j \leq k+1$, and h different from j. Suppose $h, j > 1$. Then one of f_{h-1}, f_{j-1} is a longest edge and one is a shortest non-edge (since there cannot be two longest edges ending at a given vertex). However, this contradicts G being a circular interval digraph with $\beta_G > \alpha_G$. Analogously, if $h, j < k+1$, one of f_{h+1}, f_{j+1} is a longest edge, and the other is a shortest non-edge. The same contradiction is reached. This proves that if $v_h = v_j$, then $\{h, j\} = \{1, k+1\}$. □

Theorem 3.2.25. *If G is an optimal digraph, then $\beta_G \leq \alpha_G$. Furthermore, either $\alpha_G = \beta_G$ or $\alpha_G = \beta_G + 1$.*

Proof. The second statement in Theorem 3.2.25 follows from the first since $\alpha_G \leq \beta_G + 1$ by Lemma 3.2.10, so we just prove the first. Suppose G is an optimal digraph on n vertices with $\beta = \beta_G > \alpha_G = \alpha$. Also let $\gamma = \gamma_G$. Let S be an augmenting sequence in G, and set $X = X_S$ and $Y = Y_S$. Let $S = v_1, f_1, v_2, f_2, \ldots, f_k, v_{k+1}$. Define $G' = (G + X) \setminus Y$, and note this is also a 2-free circular interval digraph by Observation 3.2.6.

Fix an extreme pair $uv \in X \cup Y$. Define $G_{uv} = G + uv$ if $uv \in Y$ and $G_{uv} = G \setminus uv$ if $uv \in X$. For vertices w different from u, v, define $p(w) = 1$ if $G|\{u, v, w\}$ is a directed path (and $p(w) = 0$ otherwise), and $q(w) = 1$ if $G_{uv}|\{u, v, w\}$ is a directed path (and $q(w) = 0$ otherwise). Finally, let

$$R(uv) = \sum_{w \neq u,v} q(w) - p(w).$$

We now define

$$R = \sum_{uv \in X \cup Y} R(uv).$$

By Lemma 3.2.20, no triple of vertices in G contains three extreme pairs. We then let T_1 be the number of triples of vertices $\{u, v, w\}$ such that u, v, w are in clockwise order in G and two of uv, vw, wu are in X. Let T_2 be the number of $\{u, v, w\}$ such that u, v, w are in clockwise order in G, one of uv, vw, wu is in X, and one is in Y.

Finally, for a vertex v, define $s^+(v) = |N^+(v)| - (\alpha - 1)$ and $s^-(v) = |N^-(v)| - (\alpha - 1)$. Also define $t^+(v) = \beta - |N^+(v)|$ and $t^-(v) = \beta - |N^-(v)|$. Then $s^+(v), s^-(v), t^+(v)$, and $t^-(v)$ are non-negative by Lemma 3.2.10.

Step 1. $\tilde{P}_3(G') - \tilde{P}_3(G) = R - 2T_1 + 2T_2$.

This follows from the definitions, using Lemma 3.2.20 to characterize those pairs with two extreme pairs in $X \cup Y$. □

Step 2. *For* $uv \in X$, $R(uv) = \gamma + 2s^+(v) + 2s^-(u) - 2$.

There are $\beta - 1 + |N^+(v) \cap N^-(u)|$ vertices w where $q(w) = 1$ and $p(w) = 0$. Since $\beta > \alpha$, Lemma 3.2.14 implies $\alpha + \beta > 3n/4 - 1/2$, and combined with Lemma 3.2.11, this gives $2\alpha + \beta \geq n$. By Lemma 3.2.20, this implies no vertex is non-adjacent to both u and v. We can now count that there are $n - (\beta - 1) - 2 - |N^+(v) \cap N^-(u)|$ vertices w where $q(w) = 0$ but $p(w) = 1$. Recalling that $R(uv) = \sum_{w \neq u,v} q(w) - p(w)$, we have

$$R(uv) = \beta - 1 + |N^+(v) \cap N^-(u)| - (n - \beta - 1 - |N^+(v) \cap N^-(u)|),$$

or

$$R(uv) = 2\beta + 2|N^+(v) \cap N^-(u)| - n.$$

We know $|N^+(v) \cap N^-(u)| = |N^+(v)| + |N^-(u)| + (\beta - 1) + 2 - n$, which we can rewrite as $2\alpha + \beta - n - 1 + (|N^+(v)| - (\alpha - 1)) + (|N^-(u)| - (\alpha - 1))$. Then using the definitions of $s^+(v)$ and $s^-(u)$, we have

$$R(uv) = 2\beta + 2(2\alpha + \beta - n - 1) + 2s^+(v) + 2s^-(u) - n,$$

which simplifies to

$$R(uv) = 4(\alpha + \beta) - 3n - 2 + 2s^+(v) + 2s^-(u).$$

From the definition of $\gamma = 4(\alpha + \beta) - 3n$, this proves Step 2. ☐

Step 3. *For $uv \in Y$, $R(uv) = 2t^+(v) + 2t^-(u) - \gamma - 2$.*

There are $n - (\alpha + 1) - |N^+(v) \cap N^-(u)|$ vertices w where $q(w) = 1$ but $p(w) = 0$. We can also count there are $\alpha - 1 + |N^+(v) \cap N^-(u)|$ vertices w where $q(w) = 0$ and $p(w) = 1$. Recalling that $R(uv) = \sum_{w \neq u,v} q(w) - p(w)$, we have

$$R(uv) = n - \alpha - 1 - |N^+(v) \cap N^-(u)| - (\alpha - 1 + |N^+(v) \cap N^-(u)|),$$

or

$$R(uv) = n - 2\alpha - 2|N^+(v) \cap N^-(u)|.$$

We know $|N^+(v) \cap N^-(u)| = |N^+(v)| + |N^-(u)| + (\alpha - 1) + 2 - n$, which we can rewrite as $\alpha + 2\beta - n + 1 - (\beta - |N^+(v)|) - (\beta - |N^-(u)|)$. Then using the definitions of $t^+(v), t^-(u)$, we have

$$R(uv) = n - 2\alpha - 2(\alpha + 2\beta - n + 1) + 2t^+(v) + 2t^-(u),$$

which simplifies to

$$R(uv) = 3n - 4(\alpha + \beta) + 2t^+(v) + 2t^-(u) - 2.$$

From the definition of $\gamma = 4(\alpha + \beta) - 3n$, this proves Step 3. ☐

Step 4. *$\tilde{P}_3(G') - \tilde{P}_3(G) \geq 0$.*

For $1 \leq i \leq k$, if $v_i v_{i+1} \in X$, then v_i has out-degree β and $t^+(v_i) = 0$; similarly, if $v_i v_{i+1} \in Y$, then $s^+(v_i) = 0$. For $2 \leq i \leq k+1$, if $v_{i-1}v_i \in X$, then v_i has in-degree β and $t^-(v_i) = 0$; analogously, if $v_{i-1}v_i \in Y$, then $s^-(v_i) = 0$. For $2 \leq i \leq k$, we note that the definition of an augmenting sequence implies that one of $v_{i-1}v_i, v_i v_{i+1}$ is in X and the other in Y. Then by Steps 2 and 3, for $3 \leq i \leq k$

$$R(v_{i-1}v_i) = \begin{cases} \gamma - 2 & \text{if } v_{i-1}v_i \in X \\ -\gamma - 2 & \text{if } v_{i-1}v_i \in Y. \end{cases} \tag{3.10}$$

We see $T_1 = 0$ unless $v_1 = v_{k+1}$, k is odd, and v_1v_2, v_kv_{k+1} are both in X, and in that case $T_1 = 1$. Also $T_2 = k - 1$ unless $v_1 = v_{k+1}$ and k is even, and in that case $T_2 = k$.

First, suppose $v_1 = v_{k+1}$ and k is odd. Then v_1v_2 and v_kv_1 must be in X by Lemma 3.2.20. By Step 1, $\tilde{P}_3(G') - \tilde{P}_3(G) = R - 2T_1 + 2T_2$, which in conjunction with equation (3.10) and the earlier argument giving $s^+(v_2) = s^-(v_k) = 0$ implies

$$\tilde{P}_3(G') - \tilde{P}_3(G) = (\gamma + 2s^-(v_1) - 2) + (\gamma + 2s^+(v_{k+1}) - 2) - 2(k-2) - \gamma - 2T_1 + 2T_2.$$

Using that $v_1 = v_{k+1}$ and substituting $T_1 = 1$ and $T_2 = k - 1$, we have

$$\tilde{P}_3(G') - \tilde{P}_3(G) = \gamma + 2s^-(v_1) + 2s^+(v_1) - 2k - 2 + 2(k-1) = \gamma + 2s^-(v_1) + 2s^+(v_1) - 4.$$

We know $|N^-(v_1)| = |N^+(v_1)| = \beta$, so $s^-(v_1) = s^+(v_1) = \beta - \alpha + 1$. Since $\beta > \alpha$, $s^-(v_1), s^+(v_1) \geq 2$. Substituting in the above inequality, we have

$$\tilde{P}_3(G') - \tilde{P}_3(G) \geq \gamma + 4.$$

Since $\gamma \geq -1$ by Lemma 3.2.16, $\tilde{P}_3(G') > \tilde{P}_3(G)$, as required.

Now suppose that $v_1 = v_{k+1}$ and k is even. Since $v_1 = v_{k+1}$, equation (3.10) holds for $2 \leq i \leq k+1$. Also, by Step 1,

$$\tilde{P}_3(G') - \tilde{P}_3(G) = R - 2T_1 + 2T_2 = -2k - 2T_1 + 2T_2. \tag{3.11}$$

Substitution of $T_1 = 0$, $T_2 = k$ in equation (3.11) yields $\tilde{P}_3(G') - \tilde{P}_3(G) = 0$, as required.

Thus we may assume $v_1 \neq v_{k+1}$. We note Lemma 3.2.23 implies that if $v_1 \neq v_{k+1}$, then v_1v_2, v_kv_{k+1} are either both in X or both in Y, and k is odd. Let $\mu = -1$ if $v_1v_2 \in X$ and $\mu = 1$ otherwise.

By Step 1, $\tilde{P}_3(G') - \tilde{P}_3(G) = R - 2T_1 + 2T_2$, which in conjunction with equation (3.10) implies

$$\tilde{P}_3(G') - \tilde{P}_3(G) = R(v_1v_2) + R(v_kv_{k+1}) - 2(k-2) + \mu\gamma - 2T_1 + 2T_2, \tag{3.12}$$

where

$$R(v_1v_2) = \begin{cases} \gamma + 2s^-(v_1) - 2 & \text{if } v_1v_2 \in X \\ 2t^-(v_1) - \gamma - 2 & \text{if } v_1v_2 \in Y, \end{cases}$$

and

$$R(v_kv_{k+1}) = \begin{cases} \gamma + 2s^+(v_{k+1}) - 2 & \text{if } v_kv_{k+1} \in X \\ 2t^+(v_{k+1}) - \gamma - 2 & \text{if } v_kv_{k+1} \in Y. \end{cases}$$

Then equation (3.12) can be simplified to

$$\tilde{P}_3(G') - \tilde{P}_3(G) = \begin{cases} \gamma + 2s^-(v_1) + 2s^+(v_{k+1}) - 2k - 2T_1 + 2T_2 & \text{if } v_1v_2 \in X \\ -\gamma + 2t^-(v_1) + 2t^+(v_{k+1}) - 2k - 2T_1 + 2T_2 & \text{if } v_1v_2 \in Y. \end{cases}$$

Recalling that $T_1 = 0$ and $T_2 = k - 1$, we have

$$\tilde{P}_3(G') - \tilde{P}_3(G) = \begin{cases} \gamma + 2s^-(v_1) + 2s^+(v_{k+1}) - 2 & \text{if } v_1v_2 \in X \\ -\gamma + 2t^-(v_1) + 2t^+(v_{k+1}) - 2 & \text{if } v_1v_2 \in Y. \end{cases}$$

First, suppose $v_1v_2 \in X$. Then $s^-(v_1), s^+(v_{k+1}) \geq 1$, since otherwise S is not maximal. Since $\gamma \geq -1$ by Lemma 3.2.16, this proves $\tilde{P}_3(G') - \tilde{P}_3(G) > 0$. On the other hand, suppose $v_1v_2 \in Y$. Then $t^-(v_1), t^+(v_{k+1}) \geq 1$ by the maximality of S. Now $\gamma \leq 1$ by Lemma 3.2.16, and again $\tilde{P}_3(G') - \tilde{P}_3(G) > 0$. This completes the proof of Step 4. □

We observe that $\xi(G') < \xi(G)$ follows immediately from the definition of $\xi(G)$ and the fact $\beta < \alpha$. Yet we have now contradicted the optimality of G. This proves Theorem 3.2.25. □

Finally, we prove a lemma relating the number of induced 3-vertex paths in a general circular interval digraph with longest edge of length β to the number in G_β (from Definition 3.2.3). We need two further definitions.

Definition 3.2.26. Let H_β be the subgraph of G_β with the same vertex set, and $E(H_\beta) = \{uv : d(u,v) = \beta\}$. Finally, for $X \subseteq E(H_\beta)$, let $t(X)$ be the number of vertices of H_β which are incident with exactly one edge in X.

Lemma 3.2.27. *Let* $n \geq 4$*, and let* β *be an integer satisfying* $-2 \leq 8\beta - 3n \leq 2$*. Then for all* $X \subseteq E(H_\beta)$*,*

$$|X|(8\beta - 3n) + t(X) + n(n - 2\beta - 1)(2\beta - n/2 + 1) \leq n^3/16.$$

Proof. Let $\delta = 8\beta - 3n$. Then $-2 \leq \delta \leq 2$, and (eliminating β) we must show that

$$|X|\delta + t(X) + n(n - \delta - 4)(n + \delta + 4)/16 \leq n^3/16,$$

that is,

$$|X|\delta + t(X) \leq n(\delta + 4)^2/16 \tag{3.13}$$

for all $X \subseteq E(H_\beta)$.

Let $t = t(X)$, and $Y = E(H_\beta) \setminus X$. In G_β, every vertex is incident with two edges of length β. Since $X \cup Y = E(H_\beta)$, and t counts vertices which are incident with exactly one edge in X, we have that $2|Y| \geq t$, $2|X| \geq t$, and $|X| + |Y| = n$.

Case 1. $\delta = 0$.

Since $\delta = 0$, equation (3.13) becomes $t \leq n$, which is clear since G has n vertices.

Case 2. $\delta = 1$.

Substituting into inequality (3.13), we must show that $|X| + t \leq 25n/16$. Since $2|Y| \geq t$ and $2|X| \geq t$, it follows that $6|Y| + 2|X| \geq 4t$. Using $|X| + |Y| = n$ to eliminate $|Y|$ gives $6(n - |X|) + 2|X| \geq 4t$, that is, $|X| + t \leq 3n/2 < 25n/16$, as required.

Case 3. $\delta = 2$.

In this case, equation (3.13) becomes $2|X| + t \leq 9n/4$. But since $2|Y| \geq t$ and $|Y| = n - |X|$, we have $2(n - |X|) \geq t$, or $2|X| + t \leq 2t \leq 2n < 9n/4$, as required.

Case 4. $\delta = -1$.

When $\delta = -1$, we need to show $t - |X| \leq 9n/16$ to prove the inequality in (3.13). If $|X| \leq n/2$ then $t \leq 2|X| \leq |X| + n/2$. If $|X| > n/2$, then $t \leq 2|Y| = 2(n - |X|) \leq n/2 + |X|$. In both cases, $t \leq |X| + n/2 < |X| + 9n/16$, as required.

Case 5. $\delta = -2$.

Finally, when $\delta = -2$, proving (3.13) requires $t - 2|X| \leq n/4$. But $2|X| \geq t$, so this is trivial. This proves Lemma 3.2.27. □

Lemma 3.2.28. *Let* $G = G_\beta \setminus X$*, where* $X \subseteq E(H_\beta)$ *and* $|V(G)| = n$*. Then* $\tilde{P}_3(G) = \tilde{P}_3(G_\beta) + |X|(8\beta - 3n) + t(X)$*.*

Proof. First, consider all triples of vertices u, v, w where exactly one pair of vertices is at distance β (say uv). Then the gain in induced 3-vertex paths among these triples from deleting the edge uv is $\beta-1+(3\beta-n-1)$ and the loss is $2(n-2\beta-1)$. This gives a net gain of $8\beta-3n$ for each edge in X. Now consider the triples u, v, w where there are two pairs at distance β. Among these, you gain an induced 3 vertex path for each triple where exactly one of the two edges of length β is deleted. This is precisely $t(X)$, by definition. Finally, we note that no other triple can change how many induced 3-vertex paths it contributes when X is deleted, giving $\tilde{P}_3(G) = \tilde{P}_3(G_\beta)+|X|(8\beta-3n)+t(X)$, as desired. □

Proof of Theorem 3.2.1. Let G be a digraph on n vertices. If $n \leq 2$, then $\tilde{P}_3(G) = 0 \leq n^3/16$. If $n = 3$, then $\tilde{P}_3(G) \leq 1 \leq 27/16$. So we may assume $n \geq 4$, and that G is optimal. It follows from Theorem 3.2.25 that every optimal digraph G with maximum edge length β can be written as $G_\beta \setminus X$ for some set X of edges with length β. Let $\alpha = \alpha_G$ and $\beta = \beta_G$. We now show that every choice of X gives $\tilde{P}_3(G) \leq n^3/16$. By Lemma 3.2.25, either $\alpha = \beta$, or $\alpha = \beta + 1$.

Suppose $\alpha = \beta + 1$. Then $X = \emptyset$, and $G = G_\beta$. A straightforward calculation gives that

$$\tilde{P}_3(G_\beta) = n(n - 2\beta - 1)(2\beta - n/2 + 1). \tag{3.14}$$

Let $x = 2\beta + 1$. Then we need to show $n(n - x)(x - n/2) \leq n^3/16$, or $x(3n/2 - x) \leq 9n^2/16$. Now, Lemma 3.2.14 implies that $3n/4 - 1/2 \leq x \leq 3n/4 + 1/2$. We see that $x(3n/2 - x)$ is maximized when $x = 3n/4$, where it is equal to $9n^2/16$. This proves that when $\alpha = \beta + 1$, $\tilde{P}_3(G) \leq n^3/16$.

Thus we may assume $\alpha = \beta$. Lemma 3.2.14 now gives $3n/8 - 1/4 \leq \beta \leq 3n/8 + 1/4$, or $3n - 2 \leq 8\beta \leq 3n + 2$. Theorem 3.2.1 then follows directly from equation 3.14, together with Lemmas 3.2.27 and 3.2.28. □

3.3 Four-Vertex Paths in 3-free Digraphs

For integer t, let α_t be the minimum constant so that all n-vertex digraphs with minimum out-degree at least $\alpha_t n$ have a directed cycle of length at most t. The Caccetta-Häggkvist conjecture is that $\alpha_t = 1/t$. A number of papers have focused on the special case of getting an upper bound on α_3 that is as close to $1/3$ as possible.

The most recent result by Shen [21] slightly tightens an argument of Hamburger, Haxell, and Kostochka [7] and proves $\alpha_3 \leq .3530381$.

One possible approach for finding such constants is to find bounds on the number of directed walks in t-free digraphs. It is easy to see that if G is a digraph on n vertices with minimum out-degree d, then $W_s(G) \geq d^{s-1}n$. Hence a bound of the form $W_s(G) \leq (c_{s,t}n)^s$ for t-free digraphs G would prove there is a vertex of out-degree at most $(c_{s,t})^{s/(s-1)}$. In this light, the following problem is of interest:

Problem 3.3.1. *Given integers s, t, find the smallest $c_{s,t}$ so that $\sqrt[s]{W_s(G)} \leq c_{s,t}n$ for every t-free digraph G on n vertices.*

We also note that if G is t-free and $s = t+1$, then $W_s(G) = P_s(G)$. We were able to show the following when $t = 3$:

Theorem 3.3.2. *If G is a 3-free digraph on n vertices, then $P_4(G) \leq \frac{4}{75}n^4$.*

Proof. Let G be a 3-free digraph on n vertices. We will use the term *square* to refer to a subgraph of G which is a directed cycle of length four. If $X \subseteq V(G)$ with $|X| = 4$, let $t(X)$ be the number of 4-vertex directed paths with vertex set X. We observe that since G is 3-free, $t(X) \in \{0, 1, 4\}$ for every such X. This motivates the following definitions. Let R be the number of four-tuples of distinct vertices (a, b, c, d) such that $t(\{a, b, c, d\}) = 1$. Let S be the number of four-tuples of distinct vertices (a, b, c, d) such that $G|\{a, b, c, d\}$ is a square (equivalently, $t(\{a, b, c, d\}) = 4$). Then S is 24 times the number of squares. Define N to be the set of four-tuples of vertices not counted by either R or S, so $|N| = n^4 - R - S$. For distinct vertices u, v, let $M(u, v)$ be the set of all vertices x so that (u, x, v) is an induced 3-vertex path. Set $m(u, v) = |M(u, v)|$, the number of induced directed 3-vertex paths starting at u and ending at v. Finally, define $T = \tilde{P}_3(G)$ to be the total number of induced directed 3-vertex paths in G.

Note that $n^4 = R + S + |N|$ by definition. We can also express the number of 4-vertex paths $P_4(G)$ in terms of these parameters, as $24P_4(G) = 4S + R$. Combining these equalities, we write

$$24P_4(G) = n^4 + 3S - |N|. \tag{3.15}$$

To prove an upper bound for $P_4(G)$, it then suffices to bound S from above and $|N|$

from below. In fact, we will prove

$$S \leq \frac{3n}{2}T, \tag{3.16}$$

and

$$|N| \geq \frac{2}{3}S. \tag{3.17}$$

Combining (3.16) and (3.17) with (3.15), we see that:

$$24P_4(G) \leq n^4 + \frac{7}{3}S \leq n^4 + \frac{7}{2}nT.$$

But $T \leq \frac{2}{25}n^3$ by Theorem 3.1.2, and so $24P_4(G) \leq (1 + 7/25)n^4$, or $P_4(G) \leq \frac{4}{75}n^4$, as desired. The proofs of inequalities (3.16) and (3.17) follow.

Proof of inequality (3.16): We will write $P \sqsubset G$ to mean P is a (directed) path of G, and $\sum_P, \sum_\Gamma$ to mean the sum over all induced 3-vertex paths in G and the sum over all squares in G, respectively. For each square $\Gamma = a\text{-}b\text{-}c\text{-}d\text{-}a$ in G, define

$$\omega(\Gamma) = \frac{1}{m(c,a)} + \frac{1}{m(d,b)} + \frac{1}{m(a,c)} + \frac{1}{m(b,d)}.$$

Now, since $m(a,c) + m(c,a) + m(b,d) + m(d,b) \leq n$ (each path has a middle vertex, and no vertex can serve as the middle of two of the paths counted or G would have a directed cycle of length at most three), $\omega(\Gamma) \geq 16/n$ for all Γ. Since there are $S/24$ squares, it follows that:

$$\sum_\Gamma \omega(\Gamma) \geq \frac{16}{n}\left(\frac{S}{24}\right) = \frac{2}{3n}S.$$

For an induced 3-vertex path $P = u\text{-}w\text{-}v$ in G, let

$$\omega(P) = \frac{1}{m(v,u)}\left|\{\text{squares } \Gamma : P \sqsubset \Gamma\}\right|.$$

We claim that $\omega(P) = 1$ for all P. The squares containing P are of the form $u\text{-}w\text{-}v\text{-}x\text{-}u$ where (v, x, u) is also an induced 3-vertex path. Since G is 3-free, every 4-cycle is induced, so every choice of $x \in M(v,u)$ gives a square, proving $\omega(P) = m(v,u) \cdot \frac{1}{m(v,u)} = 1$. Then $\sum_P \omega(P) = \sum_P 1 = T$ by definition.

Finally, we show $\sum_P \omega(P) = \sum_\Gamma \omega(\Gamma)$. Below, let P be u-w-v. Then

$$\begin{aligned}\sum_P \omega(P) &= \sum_P \frac{1}{m(v,u)} \left|\{\text{squares } \Gamma : P \sqsubset \Gamma\}\right| \\ &= \sum_P \sum_{\Gamma \sqsupset P} \frac{1}{m(u,v)} \\ &= \sum_\Gamma \sum_{P \sqsubset \Gamma} \frac{1}{m(u,v)} \\ &= \sum_\Gamma \omega(\Gamma).\end{aligned}$$

We now have $T = \sum_P \omega(P) = \sum_\Gamma \omega(\Gamma) \geq \frac{2}{3n} S$, or

$$S \leq \frac{3n}{2} T.$$

This proves inequality (3.16). □

Proof of inequality (3.17). Let $\Gamma = a$-b-c-d-a be a square in G. Define

$$\omega(\Gamma) = 2\left(\frac{(m(b,d)+m(d,b))^2}{m(a,c)m(c,a)} + \frac{(m(a,c)+m(c,a))^2}{m(b,d)m(d,b)}\right).$$

Again, $m(a,c)+m(c,a)+m(d,b)+m(b,d) \leq n$, and by Cauchy-Schwarz, $\omega(\Gamma) \geq 16$ (since we know that $m(u,v) > 0$ for each relevant u, v). Since there are $S/24$ squares, we have

$$\frac{2}{3} S \leq \sum_\Gamma \omega(\Gamma). \tag{3.18}$$

Given a four-tuple of vertices $\pi = (p,q,r,s)$ and a square Γ, we say they are *associated* if there exist vertices u, v such that $\Gamma = p$-u-q-v-p and $r, s \in M(u,v) \cup M(v,u)$. We denote association as $\pi \sim \Gamma$. Note that for a square $\Gamma = a$-b-c-d-a, the four-tuples associated with it are precisely those of the forms (a,c,x,y) or (c,a,x,y) where $x, y \in M(d,b) \cup M(b,d)$, and (b,d,x,y) or (d,b,x,y) with $x,y \in M(a,c) \cup M(c,a)$.

Now, for a four-tuple of vertices $\pi = (p,q,r,s)$, define $\omega(\pi)$ as follows:

$$\omega(\pi) = \frac{|\{\Gamma : \Gamma \sim \pi\}|}{m(p,q)m(q,p)}.$$

Note that $\omega(\pi) \leq 1$, since the number of squares associated with π is at most

$m(p,q)m(q,p)$ by definition. Then

$$\sum_{\pi \in N} \omega(\pi) \leq \sum_{\pi \in N} 1 \leq |N|. \tag{3.19}$$

Next, if $\Gamma = a\text{-}b\text{-}c\text{-}d\text{-}a$ is a square in G, we show that $\pi \sim \Gamma$ implies $\pi \in N$. Without loss of generality, we may let $\pi = (a, c, x, y)$. We need to show that there is no 4-vertex path with vertex set $\{a, c, x, y\}$. This is clear if a, c, x, y are not all distinct, so we assume they are distinct. Since b is adjacent to every vertex in $M(b,d)$ and from every vertex in $M(d,b)$, there is no edge from $M(b,d)$ to $M(d,b)$, since otherwise there would be a directed triangle. Similarly, there is no edge from a vertex in $M(d,b)$ to a vertex in $M(b,d)$. Consequently, if X is a set of four vertices so that $G|X$ has a 4-vertex path as a subgraph and $X \subseteq M(b,d) \cup M(b,d)$, then $X \subseteq M(b,d)$ or $X \subseteq M(b,d)$. So not both of a, c are in X. This proves that every $\pi \sim \Gamma$ belongs to N.

This observation allows us to relate $\sum_\Gamma \omega(\Gamma)$ to $\sum_{\pi \in N} \omega(\pi)$. Assuming $\pi = (p, q, r, s)$ for the purposes of writing $\omega(\pi)$,

$$\sum_{\Gamma} \omega(\Gamma) = \sum_{\Gamma} \sum_{\pi \sim \Gamma} \frac{1}{m(p,q)m(q,p)} = \sum_{\pi \in N} \sum_{\Gamma \sim \pi} \frac{1}{m(p,q)m(q,p)} = \sum_{\pi \in N} \omega(\pi).$$

Combining this with (3.18) and (3.19), we have

$$\frac{2S}{3} \leq \sum_{\Gamma} \omega(\Gamma) = \sum_{\pi \in N} \omega(\pi) \leq |N|.$$

This proves inequality (3.17), and thus Theorem 3.3.2. □

Corollary 3.3.3. *The constant* α_3 *is bounded above by* $\sqrt[3]{4/75} \approx .3764$.

Observation 3.3.4. *If Conjecture 3.1.1 holds, we could replace Bondy's bound on* P_3 *by* $n^3/15$, *and the proof of 3.3.2 would then give* $P_4(G) < \frac{1}{19.45} n^4 \approx .0514 n^4$, *and* $\alpha_3 \leq \sqrt[3]{\frac{1}{19.45}} \approx .37184$.

Chapter 4

Out-degree Relationships in 3-free Digraphs

4.1 Introduction

Recall that the Caccetta-Häggkvist conjecture asserts that if G is 3-free, then it has minimum out-degree less than $n/3$. This chapter focuses on showing that this is not independent of the average out-degree. That is, in a 3-free digraph on n vertices, the minimum out-degree is bounded above by functions of the form $c_1 n - c_2 \delta^+_{avg}$ for positive constants $c_1 \leq 1$ and c_2. We also show that if G is Eulerian, we can prove better constants.

4.2 Average and Minimum Out-Degree

Theorem 4.2.1. *Let $\delta < 1$ be any constant so that if a digraph H has $\delta^+_H \geq \delta|V(H)| > 0$,then H is not 3-free. Let G be a non-null 3-free digraph on n vertices with average out-degree γn and minimum out-degree at least ϵn. Then*

$$\gamma \leq \frac{1-\epsilon}{2-\delta}.$$

Proof. In a 2-free digraph, the maximum number of edges is $n(n-1)/2$, so

$$\gamma \leq \frac{n-1}{2n} \leq \frac{1}{2}.$$

Then we may assume

$$\frac{1-\epsilon}{2-\delta} < \frac{1}{2},$$

as otherwise the theorem holds. Since $\delta \geq 0$, this implies $\epsilon > 0$. Thus G has minimum out-degree at least one. Let $\lambda = \frac{1}{1-\delta}$. Choose $v \in V(G)$ such that

$$a + \lambda b \geq (1+\lambda)\gamma n, \tag{4.1}$$

where $a = \delta^+(v)$, $b = \delta^-(v)$, and $c = n - 1 - a - b$ (the number of non-neighbors of v). Such a vertex exists since

$$\sum_{v \in V(G)} \delta^+(v) + \lambda \delta^-(v) = (1+\lambda)|E(G)| = (1+\lambda) \sum_{v \in V(G)} \gamma n.$$

Now we prove the following inequalities also hold:

$$\begin{aligned} a+b+c &< n, & (4.2) \\ c &> \epsilon n - \delta a. & (4.3) \end{aligned}$$

Equation (4.2) follows from the definition of c. To prove inequality (4.3), consider $A = G|N^+(v)$. It is non-null (no vertex has out-degree zero) and 3-free, so there must be a vertex $w \in A$ with less than δa out-neighbors in A from the definition of δ. But $\delta^+(w) \geq \epsilon n$, so w must have more than $\epsilon n - \delta a$ out-neighbors not in A. Since G is 3-free, these must all be non-neighbors of v, and so $c > \epsilon n - \delta a$, as desired.

We now consider $(1-\delta)$(4.1)−(4.2)+(4.3):

$$(1-\delta)a + b - (a+b+c) + c \geq (1-\delta+1)\gamma n - n + \epsilon n - \delta a.$$

This simplifies to:

$$((2-\delta)\gamma + \epsilon - 1)n \leq 0.$$

We can cancel $n > 0$, and since $\delta < 2$, this is the same as

$$\frac{1-\epsilon}{2-\delta} \geq \gamma.$$

This proves Theorem 4.2.1. □

We can also prove a slightly different relationship between γ, ϵ, and δ, which gives a better upper bound when $\epsilon > 1/3$.

First, we present the proof of an inequality satisfied by every edge of a 3-free digraph with large enough minimum out-degree, which is a slight extension of a result shown by Shen [17].

Definition 4.2.2. For $uv \in E(G)$ define $t(u,v) = |N^+(u) \cap N^+(v)|$, and $q(u,v) = |N^-(u) \setminus N^-(v)|$.

Lemma 4.2.3. *Let δ be any constant so that if a digraph H has $\delta^+_H \geq \delta|V(H)| > 0$, then H is not 3-free. Let G be a 3-free digraph on n vertices with minimum out-degree at least ϵn. For every $uv \in E(G)$,*

$$n > \epsilon n + \delta^-(v) + q(u,v) + (1-\delta)t(u,v).$$

Proof. If $t(u,v) = 0$, then the inequality holds because $N^+(v)$, $N^-(v)$, and $N^-(u) \setminus N^-(v)$ are pairwise-disjoint sets of cardinalities at least ϵn, $\delta^-(v)$, and $q(u,v)$, respectively, so we assume $t(u,v) > 0$. Then some vertex $w \in N^+(u) \cap N^+(v)$ has out-degree less than $\delta t(u,v)$ in $G|(N^+(u) \cap N^+(v))$ (otherwise this subgraph would contain a directed 3-cycle). Thus w is joined to at least $\delta^+(w) - \delta t(u,v) \geq \epsilon n - \delta t(u,v)$ vertices outside $N^+(u) \cap N^+(v)$. Since G is 3-free, these vertices are neither in $N^-(v)$ nor in $N^-(u) \setminus N^-(v)$. Then $|N^+(w) \cup N^+(v)| \geq \epsilon n - \delta t(u,v) + t(u,v)$. Thus

$$n \geq |N^+(v) \cup N^+(w)| + |N^-(u) \setminus N^-(v)| + |N^-(v)| + |\{v\}|,$$

since all the terms on the right hand side count pairwise disjoint sets of vertices. We then have

$$n \geq \epsilon n + (1-\delta)t(u,v) + q(u,v) + \delta^-(v) + 1.$$

This proves Lemma 4.2.3. □

Theorem 4.2.4. *Let $\delta < 1$ be any constant so that if a digraph H has $\delta^+_H \geq \delta|V(H)| > 0$,then H is not 3-free. Let G be a 3-free digraph on n vertices with average out-degree γn and minimum out-degree at least ϵn. Then*

$$\gamma \leq \frac{2(1-2\epsilon)}{2-\delta}.$$

Proof. By Lemma 4.2.3, for every $uv \in E(G)$, we have

$$n > \epsilon n + q(u,v) + (1-\delta)t(u,v) + \delta^-(v). \tag{4.4}$$

Summing this over all edges, we have

$$(n - \epsilon n)|E(G)| > \tilde{P}_3(G) + (1-\delta)T + \sum_{uv \in E(G)} \delta^-(v) \tag{4.5}$$

where T is the number of transitive triangles ($\{u, v, w\} \subseteq V(G)$ so that $uv, vw, uw \in E(G)$). We observe that

$$\tilde{P}_3(G) + T = \sum_{v \in V(G)} \delta^-(v)\delta^+(v), \tag{4.6}$$

and

$$\sum_{uv \in E(G)} \delta^-(v) = \sum_{v \in V(G)} \delta^-(v)^2. \tag{4.7}$$

Using equalities (4.6) and (4.7), we can rewrite equation (4.5) as

$$(1-\epsilon)n|E(G)| > \sum_{v \in V(G)} \delta^-(v)^2 + \sum_{v \in V(G)} \delta^-(v)\delta^+(v) - \delta T. \tag{4.8}$$

The following inequalities are easily verified using definitions:

$$\begin{aligned} T &\leq \frac{1}{2} \sum_{v \in V(G)} \delta^-(v)^2, \\ \sum_{v \in V(G)} \delta^-(v)\delta^+(v) &\geq \epsilon n |E(G)|, \\ \sum_{v \in V(G)} \delta^-(v)^2 &\geq \frac{|E(G)|^2}{n}. \end{aligned}$$

Substituting for T in inequality (4.8), we have

$$(1-\epsilon)n|E(G)| > \left(1 - \frac{\delta}{2}\right) \sum_{v \in V(G)} \delta^-(v)^2 + \sum_{v \in V(G)} \delta^-(v)\delta^+(v).$$

Using the other two inequalities, this implies

$$(1-\epsilon)n|E(G)| > \left(1 - \frac{\delta}{2}\right) \frac{|E(G)|^2}{n} + \epsilon n |E(G)|.$$

This simplifies to

$$(1-2\epsilon)n^2 > \left(1 - \frac{\delta}{2}\right) |E(G)|.$$

Since $|E(G)| = \gamma n^2$, we can substitute, giving $n^2(1-2\epsilon) > \left(1-\frac{\delta}{2}\right)\gamma n^2$, or

$$\gamma \leq \frac{1-2\epsilon}{1-\frac{\delta}{2}} = \frac{2(1-2\epsilon)}{2-\delta}.$$

This proves Theorem 4.2.4. $\square$

We note that if G is Eulerian (that is, $\delta^+(v) = \delta^-(v)$ for every vertex v), we can improve Theorem 4.2.4 as follows:

Corollary 4.2.5. *Let $\delta < 1$ be any constant so that if a digraph H has $\delta_H^+ \geq \delta|V(H)| > 0$, then H is not 3-free. Let G be an Eulerian 3-free digraph on n vertices with average out-degree γn and minimum out-degree at least ϵn. Then*

$$\gamma \leq \frac{2(1-\epsilon)}{4-\delta}.$$

Proof. Since G is Eulerian, then $\delta^+(v) = \delta^-(v)$ for all v, and in particular,

$$\sum_{v\in V(G)} \delta^-(v)\delta^+(v) = \sum_{v\in V(G)} \delta^-(v)^2.$$

Then we proceed as in the proof of Theorem 4.2.4, until we get to (4.8), where we instead have

$$(1-\epsilon)n|E(G)| > \delta\left(\frac{1}{2}\sum_{v\in V(G)} \delta^-(v)^2 - T\right) + (2-\frac{\delta}{2})\sum_{v\in V(G)} \delta^-(v)^2. \tag{4.9}$$

We now apply the following inequalities:

$$\begin{aligned} T &\leq \frac{1}{2}\sum_{v\in V(G)} \delta^-(v)^2, \\ \sum_{v\in V(G)} \delta^-(v)^2 &\geq \frac{|E(G)|^2}{n}. \end{aligned}$$

Together with (4.9), these give

$$(1-\epsilon)n|E(G)| > \left(2-\frac{\delta}{2}\right)\frac{|E(G)|^2}{n},$$

which simplifies to

$$(1-\epsilon)n^2 > \left(2-\frac{\delta}{2}\right)|E(G)|.$$

Since $|E(G)| = \gamma n^2$ by definition, we can substitute, giving $n^2(1-\epsilon) > \left(2 - \frac{\delta}{2}\right)\gamma n^2$, or

$$\gamma \leq \frac{1-\epsilon}{2-\frac{\delta}{2}} = \frac{2(1-\epsilon)}{4-\delta}.$$

This proves Corollary 4.2.5. □

4.3 Numerical Implications

The best known lower bound on δ at this time was given by Shen in 2007.

Lemma 4.3.1. *(Shen [21]) If G is a digraph on n vertices with minimum out-degree at least $.3530381n$, then G is not 3-free.*

This gives the following corollaries of Theorems 4.2.1 and 4.2.4:

Corollary 4.3.2. *Let G be a 3-free digraph on n vertices with average out-degree γn and minimum out-degree at least ϵn. Then $\gamma \leq .6071679(1-\epsilon)$.*

Proof. This follows from Lemma 4.3.1 and setting $\delta = .3530381$ in Theorem 4.2.1. □

Corollary 4.3.3. *Let G be a 3-free digraph on n vertices with average out-degree γn and minimum out-degree at least ϵn. Then $\gamma \leq 1.21240117(1-2\epsilon)$.*

Proof. This follows from Lemma 4.3.1 and setting $\delta = .3530381$ in Theorem 4.2.4. □

Chapter 5

Additive Number Theory

5.1 Introduction

Recall that Conjecture 2.1.2 claims that in every 3-free digraph G, there is a feedback arc set of size at most $\gamma(G)/2$. One natural class of digraphs for which Conjecture 2.1.2 is still open is the class of Cayley graphs. In this chapter, we present a reformulation for some special Cayley graphs which gives an interesting problem in additive number theory. In Section 5.2 we prove some elementary results showing that feedback arc sets naturally arise from permutations on the vertex set of a digraph. In Section 5.3 we look at a related problem from the perspective of finite projective space, with Section 5.4 focusing on the case of dimension one. Finally, in Section 5.5 we apply the results from projective space to obtain several partial results for Cayley graphs.

5.2 Permutations and Feedback Arc Sets

For notational convenience, we make the following definitions for the remainder of this chapter:

Definition 5.2.1. For $n \in \mathbb{Z}$, let $I_n = \{0, 1, \ldots, n-1\}$, and $V_n = \{v_0, \ldots, v_{n-1}\}$.

We will also make use of the following elementary fact from graph theory:

Observation 5.2.2. *If G is a finite acyclic digraph, then some vertex has out-degree zero.*

First, we show that in every acyclic digraph, there is a linear ordering of the vertices so that all edges go from left to right.

Theorem 5.2.3. *Let G be a digraph with $V(G) = V_m$. Then G is acyclic if and only if there is a permutation σ of I_m such that if $v_{\sigma(i)}v_{\sigma(j)}$ is an edge of G, then $i < j$.*

Proof. Let σ be a permutation of I_m such that if $v_{\sigma(i)}v_{\sigma(j)} \in E(G)$, then $i < j$. If $v_{\sigma(i_1)}, \ldots, v_{\sigma(i_n)}$ is a directed walk in G, then $i_1 < \cdots < i_n$ and so $i_n \neq i_1$, that is, $v_\sigma(i_n) \neq v_\sigma(i_1)$, and so there is no closed walk in G of length greater than zero.

To prove the converse, we use induction on m. Clearly, Theorem 5.2.3 holds for $m = 1, 2$; let $m \geq 2$ and assume the theorem holds for every acyclic digraph on m vertices. Let G be an acyclic digraph with $m + 1$ vertices, say $V(G) = V_{m+1}$. By Observation 5.2.2 there exists a vertex with out-degree zero, and without loss of generality, we may assume it is v_m. Consider $G' = G \setminus v_m$. By induction, there is a permutation σ' of I_m such that if $v_{\sigma'(i)}v_{\sigma'(j)}$ is an edge of G', then $\sigma'(i) < \sigma'(j)$. Extend this map to a permutation σ of I_{m+1} by defining $\sigma(i) = \sigma'(i)$ if $i \in I_m$, and $\sigma(m) = m$. It now suffices to show there is no edge of the form $v_{\sigma(m)}v_{\sigma(j)}$ for $j \leq m$. Since $v_m = v_{\sigma(m)}$ has out-degree zero, no such edge exists. This proves Theorem 5.2.3. □

We can now make the following definitions:

Definition 5.2.4. Let G be a digraph with vertex set V_m and σ a permutation of I_m. We define the set of edges

$$B_\sigma(G) = \left\{ v_{\sigma(i)}v_{\sigma(j)} \in E(G) \,:\, i, j \in I_m,\, i > j \right\}.$$

Also, for a set Σ of permutations of I_m, we let

$$B_\Sigma(G) = \min \left\{ |B_\sigma(G)| : \sigma \in \Sigma \right\}.$$

A straightforward corollary of Theorem 5.2.3 is that every permutation σ of the vertex set naturally gives rise to a feedback arc set (namely those edges which go from right to left in the linear order given by σ).

Corollary 5.2.5. *Let G be a digraph with vertex set V_m, and let σ be a permutation of I_m. The digraph $G \setminus B_\sigma(G)$ is acyclic.*

It then follows that $\beta(G)$ is bounded above by the minimum size of the feedback arc sets "generated" by a set of permutations.

Corollary 5.2.6. *Let G be a digraph with vertex set V_m and let Σ be a set of permutations of I_m. Then $\beta(G) \leq B_\Sigma(G)$.*

Let p be a prime. We wish to consider the special case of Conjecture 2.1.2 in which the 3-free digraph is $G = \text{Cayley}(\mathbf{F}_p, A)$ where $\mathbf{F}_p = \mathbb{Z}/p\mathbb{Z}$ and $A = \{a_1, a_2, \ldots, a_d\} \subseteq \mathbf{F}_p^* = \mathbf{F}_p \setminus \{0\}$. We will use the notation $i + p\mathbb{Z}$ to denote the equivalence class of i in $\mathbf{F}_p$. Using Corollary 5.2.6, we want to find a set Σ of permutations of I_p where we know how to bound $B_\Sigma(G)$. We will consider the following set:

Definition 5.2.7. Let p be prime. Define Σ_p to be the set of permutations of I_p given by

$$\Sigma_p = \{\sigma_k : \sigma_k(i) = ki \mod p, 1 \le k \le p-1\}.$$

Also, for a digraph G on vertex set V_p, let

$$\ell_p(G) = B_{\Sigma_p}(G) = \min\{|B_{\sigma_k}(G)| : \sigma_k \in \Sigma_p\}.$$

First, we prove a formula for $\ell_p(\text{Cayley}(\mathbf{F}_p, A))$.

Theorem 5.2.8. *Let p be prime, $A = \{a_1, a_2, \ldots, a_d\} \subseteq \mathbf{F}_p^*$, and $G = \mathit{Cayley}(\mathbf{F}_p, A)$. Then*

$$\ell_p(G) = \min\left\{\sum_{j=1}^{d}(ka_j \mod p) : k = 1, \ldots, p-1\right\}.$$

Proof. Let $V(G) = V_p$ with v_i corresponding to $i + p\mathbb{Z}$ in $\mathbf{F}_p$. Then G has edges $v_i v_{i+a_j}$ with vertex subscripts read modulo p (and $i \in I_p$, $1 \le j \le d$). Fix $k \in \{1, 2, \ldots, p-1\}$. Then the edges of G are $v_{\sigma_k(i)} v_{\sigma_k(i)+a_j}$, and we define $t_{ij} \in I_p$ so that

$$\sigma_k(i) + a_j \equiv \sigma_k(t_{ij}) \bmod p, \tag{5.1}$$

which lets us write

$$E(G) = \{v_{\sigma_k(i)} v_{\sigma_k(t_{ij})} : i \in I_p, 1 \le j \le d\}.$$

If we let u_k denote the least non-negative integer such that $ku_k \equiv 1 \bmod p$, equation (5.1) can be rewritten as:

$$i + u_k a_j \equiv t_{ij} \bmod p.$$

To find $|B_{\sigma_k}|$, we need to count edges $v_{\sigma_k(i)} v_{\sigma_k(t_{ij})}$ where $t_{ij} \bmod p < i \bmod p$. That is, $(i + u_k a_j) \bmod p < i \bmod p$, which happens precisely when $i + u_k a_j \ge p$, or $i \ge p - u_k a_j$. So there are $u_k a_j$ choices of i for which $v_{\sigma_k(i)} v_{\sigma_k(t_{ij})}$ is an edge (since

$i \le p-1$ by definition), giving

$$|B_{\sigma_k}| = \sum_{j=1}^{d} (u_k a_j \bmod p).$$

Now, since p is prime, we have $\{u_1, \dots, u_{p-1}\} = \{1, \dots, p-1\}$, so

$$\begin{aligned} \ell_p(G) &= \min\{|B_{\sigma_k}| : k = 1, \dots, p-1\} \\ &= \min\left\{ \sum_{j=1}^{d} (u_k a_j \bmod p) : k = 1, \dots, p-1 \right\} \\ &= \min\left\{ \sum_{j=1}^{d} (k a_j \bmod p) : k = 1, \dots, p-1 \right\}. \end{aligned}$$

This proves Theorem 5.2.8. □

We now study the problem of bounding $\ell_p(G)$ by translating to finite projective spaces.

5.3 Heights in Finite Projective Space

Let F be a field and let $F^* = F \setminus \{0\}$. For $d \ge 2$, we define an equivalence relation on the set of non-zero d-tuples $F^d \setminus \{(0, \dots, 0)\}$ as follows: $(a_1, \dots, a_d) \sim (b_1, \dots, b_d)$ if there exists $k \in F^*$ such that $(b_1, \dots, b_d) = (ka_1, \dots, ka_d)$. We denote the equivalence class of $(a_1, \dots, a_d)$ by $\langle a_1, \dots, a_d \rangle$. The set of equivalence classes is called the $(d-1)$-dimensional projective space over the field F, and denoted $\mathbf{P}^{d-1}(F)$. We consider projective space over the finite field $\mathbf{F}_p$ for prime p.

Definition 5.3.1. The *height* of a point $\mathbf{a} = \langle a_1, \dots, a_d \rangle \in \mathbf{P}^{d-1}(\mathbf{F}_p)$ is

$$h_p(\mathbf{a}) = \min\left\{ \sum_{i=1}^{d} (k a_i \bmod p) : k = 1, \dots, p-1 \right\}.$$

For a non-empty set $\mathcal{A} \subseteq \mathbf{P}^{d-1}(\mathbf{F}_p)$, we let

$$H_p(\mathcal{A}) = \max\left(\{h_p(\mathbf{a}) : \mathbf{a} \in \mathcal{A}\}\right).$$

These definitions are motivated by the following observation:

Observation 5.3.2. *If* $G = \mathit{Cayley}(\mathbf{F}_p, A)$ *with* $A = \{a_1, \dots, a_d\}$*, then* $\ell_p(G) = h_p(\langle a_1, \dots, a_d \rangle)$*.*

For $\mathbf{a} = \langle a_1, \dots, a_d \rangle \in \mathbf{P}^{d-1}(\mathbf{F}_p)$, let $d^*(\mathbf{a})$ denote the number of non-zero components of $\mathbf{a}$, that is, the number of i such that $a_i \neq 0$. The function $d^*(\mathbf{a})$ is well-defined, that is, independent of the representative of the equivalence class of $\mathbf{a}$. Then $h_p(\mathbf{a}) \leq d^*(\mathbf{a})(p-1)$ for all $\mathbf{a} \in \mathbf{P}^{d-1}(\mathbf{F}_p)$. We can reduce this upper bound by an averaging argument.

Lemma 5.3.3. *For every point* $\mathbf{a} \in \mathbf{P}^{d-1}(\mathbf{F}_p)$*,*

$$h_p(\mathbf{a}) \leq \left\lfloor \frac{d^*(\mathbf{a})p}{2} \right\rfloor.$$

Proof. If $a \in \mathbf{F}_p^*$, then $\{ka \bmod p : k = 1, \dots, p-1\} = \{1, \dots, p-1\}$ and so

$$\sum_{k=1}^{p-1} (ka \bmod p) = \sum_{k=1}^{p-1} k = \frac{p(p-1)}{2}.$$

It follows that for every $\mathbf{a} = \langle a_1, \dots, a_d \rangle \in \mathbf{P}^{d-1}(\mathbf{F}_p)$, we have

$$\sum_{k=1}^{p-1} \sum_{i=1}^{d} (ka_i \bmod p) = \sum_{i=1}^{d} \sum_{k=1}^{p-1} (ka_i \bmod p) = \frac{d^*(\mathbf{a})p(p-1)}{2}.$$

Since the minimum of a set of numbers does not exceed the average of the set, we have

$$h_p(\mathbf{a}) \leq \frac{1}{p-1} \sum_{k=1}^{p-1} \sum_{i=1}^{d} (ka_i \bmod p) = \frac{d^*(\mathbf{a})p}{2}.$$

The lemma follows from the fact that heights are positive integers. □

Lemma 5.3.4. *For every odd prime* p *and* $d \geq 2$*,*

$$H_p(\mathbf{P}^{d-1}(\mathbf{F}_p)) = \frac{dp}{2} \qquad \text{if } d \text{ is even,}$$

$$\frac{(d-1)p}{2} + 1 \leq H_p(\mathbf{P}^{d-1}(\mathbf{F}_p)) \leq \frac{dp-1}{2} \qquad \text{if } d \text{ is odd.}$$

Proof. Let $r = \left\lfloor \frac{d}{2} \right\rfloor$, and choose $b_1, \dots, b_r, b_{2r+1}, \dots, b_d$ non-zero elements of the field

$\mathbf{F}_p$ with $b_{2r+1},\ldots,b_d$ distinct elements. Define $\mathbf{a} = \langle a_1,\ldots,a_d\rangle$ with

$$a_i = \begin{cases} b_i & 1 \le i \le r \text{ or } 2r+1 \le i \le d \\ -b_{i-r} & r+1 \le i \le 2r. \end{cases}$$

Then $d^*(\mathbf{a}) = d$ and

$$\sum_{i=1}^{d}(ka_i \bmod p) = \sum_{i=1}^{r}((ka_i \bmod p) + (-ka_i \bmod p)) + \sum_{i=2r+1}^{d}(ka_i \bmod p),$$

for all $k = 1,\ldots,p-1$. Since $(x \bmod p) + (-x \bmod p) = p$ for all $x \not\equiv 0 \bmod p$, and the a_i are distinct for $i > 2r$, we have

$$\sum_{i=1}^{d}(ka_i \bmod p) \ge rp + \frac{(d-2r)(d-2r+1)}{2},$$

that is

$$\sum_{i=1}^{d}(ka_i \bmod p) \ge \left\lfloor \frac{d}{2} \right\rfloor p + \frac{(d-2\left\lfloor \frac{d}{2}\right\rfloor)(d-2\left\lfloor \frac{d}{2}\right\rfloor+1)}{2}. \tag{5.2}$$

Applying Lemma 5.3.3 to inequality (5.2) we obtain $h_p(\mathbf{a}) = dp/2$ if d is even and

$$\frac{(d-1)p}{2} + 1 \le h_p(\mathbf{a}) \le \frac{dp-1}{2}$$

if d is odd. This completes the proof of Lemma 5.3.4. □

5.4 The Finite Projective Line

Finite projective space $\mathbf{P}^{d-1}(\mathbf{F}_p)$ when $d = 2$ is called the finite projective line, and consists of all equivalence classes of pairs (a_1, a_2), where $a_1, a_2 \in \mathbf{F}_p$ and a_1 and a_2 are not both 0. In this setting, we can give several more upper bounds on heights.

First, if $a_1 = 0$, then $\langle 0, a_2\rangle = \langle 0, 1\rangle$ and $h_p(\langle 0,1\rangle) = 1$. Similarly, if $a_2 = 0$, then $\langle a_1, 0\rangle = \langle 1, 0\rangle$ and $h_p(\langle 1,0\rangle) = 1$. If $a_1 \ne 0$ and $a_2 \ne 0$, then $\langle a_1, a_2\rangle = \langle 1, a_1^{-1}a_2\rangle$. Thus, for all $\mathbf{a} \in \mathbf{P}^1(\mathbf{F}_p)$, either $\mathbf{a} \in \{\langle 1,0\rangle, \langle 0,1\rangle\}$, or $\mathbf{a} = \langle 1, a\rangle$ for some $a \in \mathbf{F}_p^*$, and $h_p(\langle 1,a\rangle) \ge 2$.

We now give more specific bounds on $h_p(\langle 1,a\rangle)$ for $a \ne 0$:

Lemma 5.4.1. *Let p be an odd prime and $1 \leq a \leq p-1$. Then*

$$
\begin{array}{rll}
(i) & h_p(\langle 1,a\rangle) \leq 1+a & \textit{for all } a, \\
(ii) & h_p(\langle 1,a\rangle) = 1+a & \textit{if } a < \sqrt{p}, \\
(iii) & h_p(\langle 1,a\rangle) = 2 & \textit{if and only if } a = 1, \\
(iv) & h_p(\langle 1,a\rangle) = 3 & \textit{if and only if } a = 2 \textit{ or } a = (p+1)/2, \\
(v) & h_p(\langle 1,a\rangle) = p & \textit{if and only if } a = -1, \\
(vi) & h_p(\langle 1,a\rangle) \leq \dfrac{p+(b-1)^2}{b} & \textit{if } a = p-b+p\mathbb{Z},\ 1 \leq b \leq p-1.
\end{array}
$$

Proof. For all $1 \leq a \leq p-1$ and $k \in \{1, \ldots, p-1\}$, we have

$$h_p(\langle 1,a\rangle) = \min\{k + (ka \bmod p) : k = 1, \ldots, p-1\} \leq 1+a.$$

If $k \geq a$, then $k + (ka \bmod p) \geq a+1$. We note that $ka \bmod p \leq ka$ for all $k \geq 1$.

If $1 \leq k \leq a-1$ and $a < \sqrt{p}$, then

$$ka \leq (a-1)a \leq a^2 < p.$$

It follows that $ka \bmod p = ka$ and

$$k + (ka \bmod p) = k + ka \geq 1+a.$$

and so $h_p(\langle 1,a\rangle) = 1+a$. This proves *(i)* and *(ii)*.

We have $k+(ka \bmod p) = 2$ if and only if $k=1$ and $ka \bmod p = a = 1$. Similarly, $k + (ka \bmod p) = 3$ if and only if either $k = 1$ and $ka \bmod p = a = 2$, or $k = 2$ and $ka \bmod p = 2a \bmod p = 1$, that is $a = (p+1)/2$. This proves *(iii)* and *(iv)*.

If $a = -1$, then $k + (ka \bmod p) = k + (p-k) = p$ for all $k = 1, \ldots, p-1$ and so $h_p(1,a) = p$. Conversely, if $h_p(1,a) = p$, then $k + (ka \bmod p) = p$ for all k, and so $ka \bmod p = -k \bmod p$ and $a = -1$. This proves *(v)*.

Finally, to prove *(vi)*, we let $p = qb + r$, where $q = \left\lfloor \frac{p}{b} \right\rfloor$ and $1 \leq r \leq b-1$. Then

$$qa = \left\lfloor \frac{p}{b} \right\rfloor (p-b) + p\mathbb{Z} = p - \left\lfloor \frac{p}{b} \right\rfloor b + p\mathbb{Z} = r + p\mathbb{Z},$$

and so $qa \bmod p = r$. Therefore,

$$h_p(\langle 1,a\rangle\rangle) \leq q + r = \frac{p + r(b-1)}{b} \leq \frac{p+(b-1)^2}{b}.$$

This completes the proof. ☐

We can now improve Lemma 5.3.3 for $\mathbf{P}^1(\mathbf{F}_p)$:

Theorem 5.4.2. *Let p be an odd prime and $a \in \mathbf{F}_p^*$. Then*

$$h_p(\langle 1, a\rangle) = \begin{cases} (p+1)/2 & \text{if } a = (p-1)/2 + p\,\mathbb{Z} \ \text{ or } a = p-2+p\mathbb{Z} \\ p & \text{if } a = p-1+p\mathbb{Z}, \end{cases}$$

and

$$h_p(\langle 1, a\rangle) \leq (p-1)/2 \ \text{ otherwise.}$$

Proof. The theorem is true for $p = 3, 5$, and 7, so we can assume that $p \geq 11$.

Step 1. *If $a = p-2+p\mathbb{Z}$, then $h_p(\langle 1, a\rangle) = (p+1)/2$.*

If $1 \leq k \leq (p-1)/2$, then

$$k + (ka \bmod p) = k + (p-2k) = p - k \geq \frac{p+1}{2}$$

and $k + (ka \bmod p) = (p+1)/2$ when $k = (p-1)/2$. If $k \geq (p+1)/2$, then $k + (ka \bmod p) \geq (p+3)/2$. Therefore, $h_p(\langle 1, a\rangle)) = (p+1)/2$, which proves Step 1. ☐

Step 2. *If $a = (p-1)/2 + p\mathbb{Z}$, then $h_p(\langle 1, a\rangle) = (p+1)/2$.*

Let $j \in \{1, \ldots, (p-1)/2\}$. If $k = 2j$, then

$$k + (ka \bmod p) = 2j + (j(p-1) \bmod p) = 2j + (p-j) = p + j \geq p+1.$$

If $k = 2j-1$, then

$$\begin{aligned} k + (ka \bmod p) &= (2j-1) + \left(\frac{(2j-1)(p-1)}{2} \bmod p\right) \\ &= (2j-1) + \left(\frac{p+1}{2} - j\right) \\ &= \frac{p+2j-1}{2} \geq \frac{p+1}{2}. \end{aligned}$$

Since $1 + (a \bmod p) = (p+1)/2$, we have $h_p(\langle 1, a\rangle) = (p+1)/2$, proving Step 2. ☐

Step 3. *If $a = p-1+p\mathbb{Z}$, then $h_p(\langle 1, a\rangle) = p$.*

This follows immediately from part (v) of Lemma 5.4.1. ☐

Step 4. *For all $a \in \mathbf{F}_p^*$ not considered in Steps 1-3, $h_p(\langle 1, a\rangle) \le (p-1)/2$.*

First, if $a \bmod p \le (p-3)/2$, then by Lemma 5.4.1 *(i)*,

$$h_p(\langle 1, a\rangle) \le 1 + (a \bmod p) \le \frac{p-1}{2}.$$

If $a \bmod p = (p+1)/2$, then $h_p(\langle 1, a\rangle) = 3 < (p-1)/2$ by Lemma 5.4.1 *(iv)*. Finally, if $(p+3)/2 \le a \bmod p \le p-3$, there is an integer b such that

$$3 \le b \le \frac{p-3}{2} \qquad \text{and} \qquad a = p - b + p\,\mathbb{Z}\,.$$

By Lemma 5.4.1 *(vi)* we have $h_p(\langle 1, a\rangle) \le \left(p + (b-1)^2\right)/b$. Then $h_p(\langle 1, a\rangle) \le (p-1)/2$ if $2b + 1 + \frac{4}{b-2} \le p$. If $4 \le b \le (p-3)/2$, then $2b + 1 + \frac{4}{b-2} \le 2b + 3 \le p$. Otherwise $b = 3$, and we have $2b + 1 + \frac{4}{b-2} = 11 \le p$ by assumption. This handles all values of $a \bmod p$, and thus completes the proof of Step 4 and Theorem 5.4.2. □

5.5 Cayley Graphs and Heights

Let $G = \text{Cayley}(\mathbf{F}_p, A)$ be a 3-free digraph with p prime, $|A| = d$. Since G is 2-free , we have $|E(G)| = dp$, and

$$\gamma(G) = \binom{p}{2} - dp = \frac{p(p-1-2d)}{2}.$$

In this case Conjecture 2.1.2 asserts that

$$\beta(G) \le \frac{p(p-1-2d)}{4}.$$

We now apply the results from Sections 5.3 and 5.4 to 3-free Cayley graphs.

Lemma 5.5.1. *Let $G = \text{Cayley}(\mathbf{F}_p, A)$ with p prime and $A = \{a_1, \ldots, a_d\}$. Then $\beta(G) \le h_p(\langle a_1, \ldots, a_d\rangle)$.*

Proof. This follows from Observation 5.3.2 and Corollary 5.2.6 with $\Sigma = \Sigma_p$. □

Using Lemma 5.3.4, we immediately have the following upper bound:

Corollary 5.5.2. *Let p be prime, $G = \text{Cayley}(\mathbf{F}_p, A)$, and $|A| = d$. Then $\beta(G) \le \left\lfloor \frac{dp}{2} \right\rfloor$.*

We can now restrict the range of d for which Conjecture 2.1.2 is still open for Cayley graphs:

Lemma 5.5.3. *Let $G =$ Cayley$(\mathbf{F}_p, A)$ with p prime, $|A| = d$, and G a 3-free digraph. Then if $d \leq (p-1)/4$ or $d \geq p/3$, G satisfies $\beta(G) \leq \gamma(G)/2$.*

Proof. First, if $d \leq (p-1)/4$, then $\lfloor dp/2 \rfloor \leq p(p-1-2d)/4$, so by Corollary 5.5.2, $\beta(G) \leq p(p-1-2d)/4 = \gamma(G)/2$.

Alternatively, when $d \geq p/3$, G cannot be 3-free by a result of Hamidoune [8] (he showed that when $A \subseteq \mathbf{F}_p^*$ and $d = |A| \geq p/r$, the Cayley graph $G =$ Cayley$(\mathbf{F}_p, A)$ contains a cycle of length at most r). $\square$

Finally, we use our results on the projective line to prove Conjecture 2.1.2 when G is a Cayley graph with out-degree two.

Theorem 5.5.4. *Let p be a prime number, $p \geq 7$, and let $A = \{a_1, a_2\} \subseteq \mathbf{F}_p^*$ with $a_1 \neq a_2$. If $G =$ Cayley$(\mathbf{F}_p, A)$ is a 3-free digraph, then*

$$\beta(G) \leq \frac{p-1}{2} \leq \frac{\gamma(G)}{2}.$$

Proof. Since $\langle a_1, a_2 \rangle = \langle 1, a \rangle$ in $\mathbf{P}^1(\mathbf{F}_p)$ with $a = a_1^{-1}a_2 \neq 1$, and since $\beta(G) \leq h_p(\langle a_1, a_2 \rangle) = h_p(\langle 1, a \rangle)$, it suffices to consider the case $A = \{1, a\}$. The Cayley graph G is 3-free if and only if none of the equations

$$\begin{aligned} x &= 0 \\ x + y &= 0 \\ x + y + z &= 0 \end{aligned}$$

has a solution with $x, y, z \in \{1, a\}$. The first equation implies that $a \neq 0$, the second that $a \neq p - 1 + p\mathbb{Z}$, and that third that $2a + 1 \neq 0$ and $a + 2 \neq 0$, or, equivalently, that $a \neq (p-1)/2 + p\mathbb{Z}$ or $p - 2$. It follows from Theorem 5.4.2 that

$$\beta(G) \leq h_p(\langle 1, a \rangle) \leq \frac{p-1}{2} \leq \frac{p(p-5)}{4} = \frac{\gamma(G)}{2}$$

if $p \geq 7$. This completes the proof. $\square$

Appendix A

Bondy's Result on 3-Vertex Paths

Theorem A.1 (Bondy [22])**.** *If D is a 2-free digraph on n vertices, then* $\tilde{P}_3(D) \leq \frac{2}{25}n^3$.

Proof. There are seven digraphs on three vertices, which we call types $1, \ldots, 7$ as shown in Figure A.1. Given a digraph D with vertex set $\{v_i : 1 \leq i \leq n\}$, let d_i^- and d_i^+ denote the indegree and outdegree of v_i $(1 \leq i \leq n)$ and s_j the number of induced subgraphs of type j in D $(1 \leq j \leq 7)$.

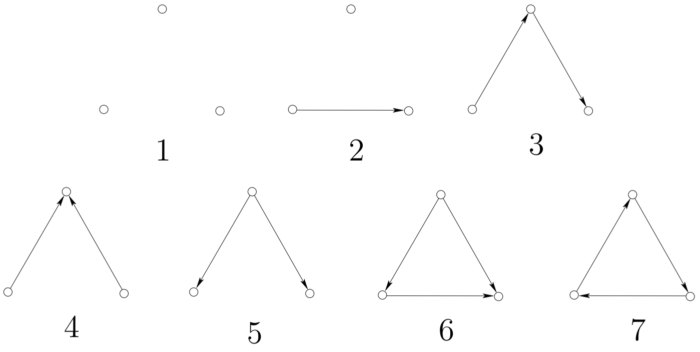

Figure A.1: The 3-vertex digraphs.

The following system of five equations holds:

$$\begin{aligned}
s_1 + s_2 + s_3 + s_4 + s_5 + s_6 + s_7 &= \binom{n}{3} \\
s_2 + 2s_3 + 2s_4 + 2s_5 + 3s_6 + 3s_7 &= \frac{1}{2}(n-2)\sum_i (d_i^- + d_i^+) \\
s_3 + s_6 &= \sum_i \binom{d_i^-}{2} \\
s_4 + s_6 + 3s_7 &= \sum_i d_i^- d_i^+ \\
s_5 + s_6 &= \sum_i \binom{d_i^+}{2}.
\end{aligned}$$

We prove an upper bound on $s_4 = \tilde{P}_3(D)$ as follows:

$$\begin{aligned}
s_4 &\leq \frac{2}{5}s_2 + \frac{1}{10}s_3 + s_4 + \frac{1}{10}s_5 + \frac{9}{5}s_7 \\
&= \frac{2}{5}(s_2 + 2(s_3 + s_4 + s_5) + 3(s_6 + s_7)) - \frac{7}{10}(s_3 + s_5 + 2s_6) + \frac{1}{5}(s_4 + s_6 + 3s_7) \\
&= \frac{n-2}{5}\sum_i (d_i^- + d_i^+) - \frac{7}{20}\sum_i (((d_i^-)^2 - d_i^-) + (d_i^+)^2 - d_i^+)) + \frac{1}{5}\sum_i d_i^- d_i^+ \\
&= \frac{n}{5}\sum_i (d_i^- + d_i^+) - \frac{7}{20}\sum_i ((d_i^-)^2 + (d_i^+)^2) + \frac{1}{5}\sum_i d_i^- d_i^+ - \frac{1}{20}\sum_i (d_i^+ + d_i^-) \\
&= \frac{2n^3}{25} - \frac{1}{10}\sum_i (d_i^- - d_i^+)^2 - \frac{1}{4}\sum_i \left(\frac{2n}{5} - d_i^-\right)^2 - \frac{1}{4}\sum_i \left(\frac{2n}{5} - d_i^+\right)^2 \\
&\leq \frac{2n^3}{25}.
\end{aligned}$$

This proves 3.1.2. □

Appendix B

A Four Functions Theorem

In this appendix we prove a result used in Section 2.5. We begin with an elementary lemma, with the notation that $\mathbf{R}_+$ denotes the set of non-negative real numbers.

Lemma B.1. *If* $a_1, a_2, c_1, c_2, d_1, d_2 \in \mathbf{R}_+$ *and* $a_k^2 \leq c_k d_k$ *for* $k = 1, 2$*, then* $(a_1+a_2)^2 \leq (c_1 + d_1)(c_2 + d_2)$.

Proof. If $c_1 = 0$, then since $a_1^2 \leq c_1 d_1$, it follows that $a_1 = 0$, and so

$$(a_1 + a_2)^2 = a_2^2 \leq c_2 d_2 \leq (c_1 + c_2)(d_1 + d_2),$$

as required. The same argument applies for c_2, so we may assume that c_1, c_2 are nonzero. Now

$$\begin{aligned} (c_1 + c_2)(d_1 + d_2) &= c_1 d_1 + c_1 d_2 + c_2 d_1 + c_2 d_2 \\ &\geq a_1^2 + c_1(a_2^2/c_2) + c_2(a_1^2/c_1) + a_2^2 \\ &= (a_1 + a_2)^2 + c_1 c_2 (a_2/c_2 - a_1/c_1)^2 \\ &\geq (a_1 + a_2)^2. \end{aligned}$$

This proves Lemma B.1. □

Before the main result of this section we must set up some notation.

Definition B.2. For any integers $m, n \geq 1$, define P_{mn} to be the set of all pairs (i, j) of integers with $1 \leq i \leq m$ and $1 \leq j \leq n$.

If $f : P_{mn} \to \mathbf{R}_+$, and $X \subseteq P_{mn}$, we define $f(X)$ to mean $\sum_{x \in X} f(x)$. For $(i, j), (i', j') \in P_{mn}$, we say that (i', j') *dominates* (i, j) if $i < i'$ and $j < j'$.

Definition B.3. Let $a, b : P_{mn} \to \mathbf{R}_+$ be functions. We say that b *dominates* a if

- $a(P_{mn}) = b(P_{mn})$.
- For all $X, Y \subseteq P_{mn}$, if $a(X) + b(Y) > a(P_{mn})$ then there exist $x \in X$ and $y \in Y$ such that y dominates x.

The main result of this appendix is the following. (It is reminiscent of the "four functions" theorem of Ahlswede and Daykin [1], but we were not able to derive it from that theorem.)

Theorem B.4. *Let $m, n \geq 1$ be integers, let $P = P_{mn}$, and let a, b, c, d be functions from P_{mn} to $\mathbf{R}_+$, satisfying the following:*

i. $a(i,j)b(i',j') \leq c(i',j)d(i,j')$ for $1 \leq i < i' \leq m$ and $1 \leq j < j' \leq n$.

ii. b dominates a.

Then $a(P)b(P) \leq c(P)d(P)$.

Proof. We proceed by induction on $m+n$. Let Q be the set of all quadruples (a, b, c, d) of functions from P to $\mathbf{R}_+$ that satisfy conditions *(i)* and *(ii)* above. We say that $(a, b, c, d) \in Q$ is *good* if

$$a(P)b(P) \leq c(P)d(P).$$

Thus, we need to show that every member of Q is good. Certainly if $m = 1$ or $n = 1$ then condition *(ii)* implies that $a(P) = b(P) = 0$, and therefore (a, b, c, d) is good; so we may assume that $m, n \geq 2$.

Step 1. *If $(a, b, c, d) \in Q$ then $b(i, 1) = 0$ for $1 \leq i \leq m$, and $a(m, j) = 0$ for $1 \leq j \leq n$.*

For let $X = P$, and let Y be the set of all pairs $(i, 1)$ with $1 \leq i \leq m$. There do not exist $x \in X$ and $y \in Y$ such that y dominates x, and since b dominates a it follows that $a(X) + b(Y) \leq a(P)$. Since $a(X) = a(P)$ we deduce that $b(Y) = 0$. This proves the first statement, and the second follows similarly. This proves Step 1. □

Step 2. *If $(a, b, c, d) \in Q$ and $a(i, 1) = 0$ for all $i \in \{1, \ldots, m\}$ then (a, b, c, d) is good.*

This follows from Step 1 and the inductive hypothesis applied to the restriction of a, b, c, d to the set of all $(i, j) \in P$ with $j > 1$ (relabeling appropriately). □

For $(a, b, c, d) \in Q$, let us define its *margin* to be the number of pairs (i, j) such that either $j = 1$ and $a(i, j) > 0$, or $i = m$ and $b(i, j) > 0$. For fixed m, n we proceed by induction on the margin. Thus, we assume that $t \geq 0$ is an integer, and every $(a, b, c, d) \in Q$ with margin smaller than t is good. We must show that every $(a, b, c, d) \in Q$ with margin t is good.

Step 3. *Let $(a, b, c, d) \in Q$ with margin t, and suppose that there exist $X, Y \subseteq P$ such that*

- *$a(X) + b(Y) = a(P)$,*
- *there do not exist $x \in X$ and $y \in Y$ such that y dominates x,*
- *there exists $i \in \{1, \ldots, m\}$ such that $(i, 1) \notin X$ and $a(i, 1) > 0$, and there exists $j \in \{1, \ldots, n\}$ such that $(m, j) \notin Y$ and $b(m, j) > 0$.*

Then (a, b, c, d) is good.

Let $A_1 = X$ and $A_2 = P \setminus X$. Let $B_1 = P \setminus Y$ and $B_2 = Y$. For $k = 1, 2$, let C_k be the set of all pairs $(i', j) \in P$ such that there exist i, j' with $i < i'$ and $j < j'$ and $(i, j) \in A_k$ and $(i', j') \in B_k$; and let D_k be the set of all pairs (i, j') such that there exist i', j with $i < i'$ and $j < j'$ and $(i, j) \in A_k$ and $(i', j') \in B_k$. We observe first that $C_1 \cap C_2 = \emptyset$; for suppose that $(i', j) \in C_1 \cap C_2$. Since $(i', j) \in C_1$, there exists $i < i'$ such that $(i, j) \in X$; and since $(i', j) \in C_2$, there exists $j' > j$ such that $(i', j') \in Y$. But then $(i', j') \in Y$ dominates $(i, j) \in X$, contradicting the second hypothesis about X, Y. This proves that $C_1 \cap C_2 = \emptyset$, and similarly $D_1 \cap D_2 = \emptyset$. For $k = 1, 2$, and $x \in P$, define $a_k(x) = a(x)$ if $x \in A_k$, and $a_k(x) = 0$ otherwise. Define $b_k(x), c_k(x), d_k(x)$ similarly. Since $a_1(P) + a_2(P), b_1(P) + b_2(P)$ and $a_1(P) + b_2(P)$ all equal $a(P)$, it follows that $a_1(P) = b_1(P)$ and $a_2(P) = b_2(P)$. We claim that $(a_k, b_k, c_k, d_k) \in Q$ for $k = 1, 2$. To see this, let $i < i'$ and $j < j'$; we must show first that $a_k(i, j)b_k(i', j') \leq c_k(i', j)d_k(i, j')$. Hence we may assume that $a_k(i, j)$ and $b_k(i', j') \neq 0$, and therefore $(i, j) \in A_k$ and $(i', j') \in B_k$. From the definition of C_k, D_k it follows that $(i', j) \in C_k$ and $(i, j') \in D_k$. Hence $a_k(i, j) = a(i, j)$, and $b_k(i', j') = b(i', j')$, and $c_k(i', j) = c(i', j)$, and $d_k(i, j') = d(i, j')$; and since $a(i, j)b(i', j') \leq c(i', j)d(i, j')$, this proves the claim. Second, we must show that b_k dominates a_k. We have already seen that $a_k(P) = b_k(P)$. Let $X', Y' \subseteq P$ with $a_k(X') + b_k(Y') > a_k(P)$; we must show that there exist $x \in X'$ and $y \in Y'$ such that y dominates x. From the symmetry we may assume that $k = 1$. Now

$a(X \cap X') = a_k(X')$, and $b(Y \cup Y') = b(Y) + b_k(Y')$, and so

$$a(X \cap X') + b(Y \cup Y') = a_k(X') + b(Y) + b_k(Y') > a_k(P) + b(Y) = a(X) + b(Y) = a(P).$$

Since b dominates a, there exist $x \in X \cap X'$ and $y \in Y \cup Y'$ such that y dominates x. No vertex in Y dominates a vertex in X, from the choice of X, Y, and it follows that $y \in Y'$, as required. This proves that b_k dominates a_k, and consequently $(a_k, b_k, c_k, d_k) \in Q$, for $k = 1, 2$.

We claim that for $k = 1, 2$, the margin of (a_k, b_k, c_k, d_k) is less than t. For from the third hypothesis about X, Y, there exists $i \in \{1, \ldots, m\}$ such that $a(i, 1) > 0$ and $(i, 1) \notin X$ (and hence $a_1(i, 1) = 0$); this shows that the margin of (a_1, b_1, c_1, d_1) is less than that of (a, b, c, d), and so less than t. Also, there exists $j \in \{1, \ldots, n\}$ such that $b(m, j) > 0$ and $(m, j) \notin Y$; and so similarly the margin of (a_2, b_2, c_2, d_2) is less than t. Hence from the second inductive hypothesis, we deduce that $a_k(P) b_k(P) \leq c_k(P) d_k(P)$ for $k = 1, 2$. But $a_k(P) = b_k(P)$ for $k = 1, 2$; thus $a_k(P)^2 \leq c_k(P) d_k(P)$ for $k = 1, 2$. Since $a_1(P) + a_2(P) = a(P) = b(P)$ and since $c(P) \geq c_1(P) + c_2(P)$ (because $C_1 \cap C_2 = \emptyset$), and similarly $d(P) \geq d_1(P) + d_2(P)$, it suffices to show that

$$(a_1(P) + a_2(P))^2 \leq (c_1(P) + c_2(P))(d_1(P) + d_2(P)),$$

and this follows from Lemma B.1. This proves Step 3. $\square$

Step 4. *If $(a, b, c, d) \in Q$ with margin t, and there exists $j \geq 3$ such that $b(m, j) > 0$, then (a, b, c, d) is good.*

For let ϵ satisfy $0 \leq \epsilon \leq 1$. For $1 \leq i \leq m$, define

$$\begin{aligned} a_1(i,1) &= (1-\epsilon)a(i,1), \\ a_1(i,2) &= \epsilon a(i,1) + a(i,2), \\ a_1(i,j) &= a(i,j) \text{ for } 3 \leq j \leq n, \\ c_1(i,1) &= (1-\epsilon)c(i,1), \\ c_1(i,2) &= \epsilon c(i,1) + c(i,2), \\ c_1(i,j) &= c(i,j) \text{ for } 3 \leq j \leq n. \end{aligned}$$

Since b dominates a, by compactness we may choose $\epsilon \leq 1$ maximum such that b dominates a_1. We claim that $(a_1, b, c_1, d) \in Q$; for let $i < i'$ and $j < j'$. We must check that $a_1(i,j)b(i',j') \le c_1(i',j)d(i,j')$. If $j=1$, then

$$a_1(i,j)b(i',j') = (1-\epsilon)a(i,1)b(i',j')$$

and

$$c_1(i',j)d(i,j') = (1-\epsilon)c(i,1)d(i,j'),$$

and since $a(i,1)b(i',j') \le c(i,1)d(i,j')$ it follows that $a_1(i,j)b(i',j') \le c_1(i',j)d(i,j')$ as required. If $j=2$, then

$$a_1(i,j)b(i',j') = (\epsilon a(i,1) + a(i,2))b(i',j')$$

and

$$c_1(i',j)d(i,j') = (\epsilon c(i,1) + c(i,2))d(i,j'),$$

and since $a(i,1)b(i',j') \le c(i,1)d(i',j')$ and $a(i,2)b(i',j') \le c(i,2)d(i',j')$, it follows that $a_1(i,j)b(i',j') \le c_1(i',j)d(i,j')$ as required. Finally, if $j>2$ the claim is clear, since $a_1(i,j) = a(i,j)$ and $c_1(i',j) = c(i',j)$. This proves that $(a_1,b,c_1,d) \in Q$.

We claim that (a_1,b,c_1,d) is good. If $\epsilon = 1$, then $a_1(i,1) = 0$ for $1 \le j \le m$, and therefore (a_1,b,c_1,d) is good by Step 2. We may therefore assume that $\epsilon < 1$. From the maximality of ϵ, there exist $X, Y \subseteq P$ such that

- there does not exist $x \in X$ and $y \in Y$ such that y dominates x,
- $a_1(X) + b(Y) = a_1(P)$,
- for some i with $1 \le i \le m$, $(i,1) \notin X$ and and $(i,2) \in X$ and $a(i,1) > 0$.

(The third statement follows from the fact that increasing ϵ will cause $a_1(X)$ strictly to increase.) Now we recall that there exists $j \ge 3$ such that $b(m,j) > 0$. Since $(i,2) \in X$ is dominated by (m,j) (for $i < m$ by Step 1, since $a(i,2) > 0$), it follows that $(m,j) \notin Y$. But then (a_1,b,c_1,d) satisfies the hypotheses of Step 3, and therefore (a_1,b,c_1,d) is good. This proves the claim. Since $a_1(P) = a(P)$ and $c_1(P) = c(P)$, we deduce that (a,b,c,d) is good. This proves Step 4. □

Now let $(a,b,c,d) \in Q$ with margin t; we shall prove that it is good. By Step 4 we may assume that $b(m,j) = 0$ for $3 \le j \le m$, and similarly that $a(i,1) = 0$ for $1 \le i \le m-2$. Since $a(m,1) = b(m,1) = 0$ by Step 1, it follows that $a(i,1) \ne 0$ only if $i = m-1$, and $b(m,j) \ne 0$ only if $j = 2$. Let $X = \{(m-1,1)\}$ and let Y be the set of all $(i,j) \in P$ with $i < m$; then there do not exist $x \in X$ and $y \in Y$ such that y dominates x. Consequently $a(X) + b(Y) \le a(P)$. But $a(X) = a(m-1,1)$ and

$b(Y) \geq a(P) - b(m, 2)$, and so $a(m-1, 1) \leq b(m, 2)$. Similarly the reverse inequality holds, and so $a(m-1, 1) = b(m, 2)$. For $(i, j) \in P$, if either $i = m$ or $j = 1$, define

$$a_1(i, j) = b_1(i, j) = c_1(i, j) = d_1(i, j) = 0.$$

If $i < m$ and $j > 1$ let $a_1(i, j) = a(i, j), b_1(i, j) = b(i, j)$, and $c_1(i, j) = c(i, j)$; and let $d_1(i, j) = d(i, j)$ except that $d_1(m-1, 2) = 0$. We claim that $(a_1, b_1, c_1, d_1) \in Q$. For let $i < i'$ and $j < j'$. We must check that $a_1(i, j)b_1(i', j') \leq c_1(i', j)d_1(i, j')$. If $i' < m$ and $j > 1$ then $a_1(i, j) = a(i, j)$ and so on, and the claim is clear. If $i' = m$ or $j = 1$ then $a_1(i, j)b_1(i', j') = 0$ and again the claim is clear. Thus $a_1(i, j)b_1(i', j') \leq c_1(i', j)d_1(i, j')$. Next we must check that b_1 dominates a_1. Certainly

$$a_1(P) = a(P) - a(m-1, 1) = b(P) - b(m, 2) = b_1(P).$$

Let $X, Y \subseteq P$ such that $a_1(X) + b_1(Y) > a_1(P)$. We must show that there exist $x \in X$ and $y \in Y$ such that y dominates x. We may therefore assume that $a_1(x) > 0$ for all $x \in X$, and $b_1(y) > 0$ for all $y \in Y$. In particular, since $a_1(m-1, 1) = b_1(m, 2) = 0$, it follows that $(m-1, 1) \notin X$ and $(m, 2) \notin Y$. Let $X' = X \cup \{(m-1, 1)\}$. Then $a(X') = a_1(X) + a(m-1, 1)$, and so

$$a(X') + b(Y) = a_1(X) + a(m-1, 1) + b(Y) > a_1(P) + a(m-1, 1) = a(P).$$

Hence there exist $x \in X'$ and $y \in Y$ such that y dominates x. If $x = (m-1, 1)$, then $y = (m, j)$ for some $j > 1$, and therefore $b_1(y) = 0$, a contradiction, since $b_1(y) > 0$ for all $y \in Y$. Thus $x \neq (m-1, 1)$, and so $x \in X$, as required. This proves that b_1 dominates a_1.

By Step 2, (a_1, b_1, c_1, d_1) is good, and so $a_1(P)b_1(P) \leq c_1(P)d_1(P)$. Moreover,

$$a(m-1, 1)b(m, 2) \leq c(m, 1)d(m-1, 2)$$

and $b(m, 2) = a(m-1, 1)$, and so $a(m-1, 1)^2 \leq c(m, 1)d(m-1, 2)$. Hence Lemma B.1 implies that

$$(a_1(P) + a(m-1, 1))^2 \leq (c_1(P) + c(m, 1))(d_1(P) + d(m-1, 2)).$$

But $a(P) = a_1(P) + a(m-1, 1) = b(P)$, and $c(P) \geq c_1(P) + c(m, 1)$, and $d(P) \geq d_1(P) + d(m-1, 2)$; and it follows that (a, b, c, d) is good. This completes the inductive proof that every member of Q is good, and so proves Theorem B.4. □

Bibliography

[1] R. Ahlswede and D.E. Daykin. "An inequality for the weights of two families, their unions and intersections," *Z. Wahrscheinlichkeitsth. verw. Gebiete*, 43:183-185, 1978.

[2] M. Behzad, G. Chartrand, and C. Wall. "On minimal regular digraphs with given girth," Fundamenta Mathematicae, 69:227-231, 1970.

[3] L. Caccetta and R. Häggkvist. "On minimal digraphs with given girth," *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory and Computing* (Florida Atlantic University, Boca Raton, Florida, 1978), Congressus Numerantium XXI, Utilitas Math, 181-187, 1978.

[4] G. Chen, J. Shen, and R. Yuster. "Second neighborhood via first neighborhood in digraphs," Annals of Combinatorics 7(1):15-20, 2003.

[5] M. Chudnovsky, P. Seymour, and B. Sullivan. "Cycles in Dense Digraphs," to appear, Combinatorica, 2006.

[6] V. Chvatal and E. Szemeredi. "Short cycles in directed graphs," Journal of Combinatorial Theory, 35(3):323-327, 1983.

[7] P. Hamburger and P. Haxell and A. Kostochka. "On Directed Triangles in Digraphs," Electronic Journal of Combinatorics, 14(1), 2007.

[8] Y. O. Hamidoune. "An application of connectivity theory in graphs to factorizations of elements in groups," European Journal of Combinatorics, 2(4):349-355, 1981.

[9] Y. O. Hamidoune. "A note on minimal directed graphs with given girth," Journal of Combinatorial Theory, Series B, 43(3):343-348, 1987.

[10] C. Hoáng and B. Reed. "A note on short cycles in digraphs," Discrete Math, 66(1-2):103-107, 1987.

[11] J. Kemperman. “On complexes in a semigroup,” Indagationes Mathematicae, 18:247-254, 1956.

[12] A. Kostochka and M. Stiebitz. In a lecture by Kostochka at the American Institute of Mathematics in January 2006.

[13] M. B. Nathanson. “The Caccetta-Häggkvist conjecture and additive number theory,” arXiv:math.CO/0603469, 2006.

[14] M. B. Nathanson. “Heights on the finite projective line,” arXiv:math/0703646, 2007.

[15] M. B. Nathanson and B. D. Sullivan. “Heights in finite projective space and a problem on directed graphs,” Integers: Electronic Journal of Combinatorial Number Theory, 8:A13, 2008.

[16] K. O’Bryant. “Gaps in the spectrum of heights of projective points,” arXiv:math/0703625, 2007.

[17] J. Shen. “Directed triangles in digraphs,” Journal of Combinatorial Theory, Series B, 74(2):405-407, 1998.

[18] J. Shen. “On the girth of digraphs,” *Discrete Math*, 211(1-3):167-181, 2000.

[19] J. Shen. “On the Caccetta-Häggkvist conjecture,” Graphs and Combinatorics 18(3):645-654, 2002.

[20] J. Shen. “Short cycles in digraphs with local average outdegree at least two,” Electronic Journal of Combinatorics, 10, 2003.

[21] J. Shen. “Paths in 3-free Digraphs,” personal communication, 2006.

[22] P. Seymour, B. D. Sullivan. “Counting Paths in Digraphs,” in preparation, 2008.

[23] B. D. Sullivan. “A summary of problems and results related to the Caccetta-Häggkvist conjecture,” AIM Preprint, 13, 2006.

[24] A. van Zuylen, R. Hegde, K. Jain, and D. P. Williamson. “Deterministic pivoting algorithms for constrained ranking and clustering problems,” SIAM SODA Proceedings (N. Bansal, K. Pruhs, and C. Stein, editors), 405-414, 2007.