CHARACTERIZATIONS AND PROPERTIES OF WORD-REPRESENTABLE GRAPH CLASSES

ESHWAR SRINIVASAN

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI - 781039, INDIA

JANUARY 2026


भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Characterizations and Properties of Word-Representable Graph Classes

by

Eshwar Srinivasan

Roll No. 226123005

Department of Mathematics

submitted in fulfillment of the requirements
of the degree of Doctor of Philosophy

to the

Indian Institute of Technology Guwahati
Guwahati - 781039, India

January 2026

This work is dedicated

to

My Parents: Mrs. C. Revathi & Mr. R. Srinivasan

My Doctorate Guru: Dr. H. Ramesh

My Brother: Mr. Mruthyunjay Srinivasan

All my Gurus

And to that ever-existing, omnipresent bliss!

Certificate

This is to certify that the thesis entitled **"Characterizations and properties of word-representable graph classes"** submitted by **Mr. Eshwar Srinivasan** to the **Indian Institute of Technology Guwahati**, for the award of the Degree of **Doctor of Philosophy**, is a record of the original bona fide research work carried out by him under my guidance and supervision. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.

Date: January 7, 2026
Guwahati, India

Dr. H. Ramesh
Assistant Professor
Department of Mathematics
Indian Institute of Technology Guwahati

Acknowledgements

”The outcome has already been ordained by my decree; you just be the instrument in every deed.’

-Shri Krishna

First and foremost, I would like to express sincere gratitude towards my supervisor, Dr. H. Ramesh. I deeply appreciate his unwavering support in moments of both progress and setbacks throughout my doctoral journey. His enthusiasm, openness to discussions, and constant encouragement paved the way for me to express my ideas and thoughts with confidence. His eye for precision and perfection in mathematics has always inspired me to improve myself. His deep knowledge, with the aptitude for connecting multiple concepts, helped me gain new perspectives on the subject. I owe him a great deal for his patience and graciousness. I cannot thank him enough for making time to diligently review all my work and enhancing its quality with his valuable insights and feedback. He has exemplified a holistic approach towards academics and life, and I feel highly fortunate to have had such a supervisor.

I express my sincere gratitude to the members of my doctoral committee—Prof. K. V. Krishna, Prof. Bikash Bhattacharjya, Prof. Gautam K. Das, and Dr. K. Benny George—for their valuable evaluation of my work and their constructive suggestions. I am deeply thankful for the atmosphere of warmth and ease they fostered throughout the course of my Ph.D., which greatly helped me to stay focused on my research. I am also profoundly indebted to Prof. Natesan Srinivasan, Prof. Partha Sarathi Mandal, and Dr. Anjan K. Chakrabarty, whose encouragement continually inspired me to strive for greater growth as a student of mathematics. Their guidance, coupled with their genuine kindness and welcoming nature, has left a lasting impression on me.

I express my heartfelt gratitude to the Indian Institute of Technology Guwahati for providing an excellent academic environment and the necessary resources that greatly supported my research. I hold deep appreciation for the Serpentine Lake, whose serene presence comforted me in moments of solitude and infused me with positivity throughout my journey. The tranquil and picturesque campus of the institute will always remain close to my heart. I am sincerely thankful to the Ministry of Human Resource Development, Government of India, for the financial support extended during my doctoral studies. I also wish to acknowledge with appreciation the technical and non-teaching staff of the Department of Mathematics, IIT Guwahati, for their constant assistance and goodwill in numerous ways.

I would like to express my heartfelt gratitude to my seniors—Dr. Ajit Singh, Dr. Gurinder Singh, Dr. H. Laxmi, Dr. V. Shanmugam, Dr. Khyodeno Mozhui, Tithi Dwary, Bivakar Bose, Sunil Kundu, Sachin Kumar, Dr. Aayusman Raina, and Gaurav Kumar—and to the post-doctoral fellows Dr. Rohit Verma and Dr. Parveen Bishnoi for their inspiration and invaluable help in various situations during my stay at IIT Guwahati.

My sincere appreciation also goes to my dear friends and research group members—Koushik Bhakta, Biswajit Das, Pratham Singh, Sumit Kumar, Sohini Pal, Akash Kalita, Ritvika Kushwaha, Suchanda Roy, Sipra Maity, Pratik Sarkar, Ansh Agarwal, Kushal Roy, and Pratyoy Mandal—for fostering a vibrant and supportive research environment. They have been a constant source of encouragement, standing by me through every challenge, and each of them has enriched my journey with their unique qualities. I will always cherish and carry this spirit of camaraderie with me.

Most importantly, I am forever indebted to my family, who have shaped me into the person I am today. Words fall short in expressing my love and gratitude to my parents, Mr. R. Srinivasan and Mrs. C. Revathi, and my brother, Mr. Mruthyunjay Srinivasan, for their unwavering care and encouragement. Despite my shift from engineering and two years spent in professional work, they stood by me with faith and warmth, supporting my decision to follow my passion and my love for learning. Finally, I remain ever grateful to the eternal, omnipresent bliss that pervades all existence, for guiding me and allowing me to witness the beauty of life in every moment.

At last, I extend my deepest gratitude and immeasurable appreciation to all those who, in countless ways, contributed to the successful completion of my Ph.D. and made this journey an unforgettable experience.

Date: January 7, 2026

Guwahati, India

Eshwar Srinivasan

Abstract

A graph G with vertex set $V(G)$ and edge set $E(G)$ is said to be *word-representable* if there exists a word w over the alphabet $V(G)$ such that, for any two distinct letters $x, y \in V(G)$, the letters x and y alternate in w if and only if $xy \in E(G)$. Equivalently, a graph is word-representable if and only if it admits a *semi-transitive orientation*—that is, an acyclic orientation in which, for every directed path $v_0 \to v_1 \to \cdots \to v_m$ with $m \geq 2$, either there is no arc between v_0 and v_m, or, for all $1 \leq i < j \leq m$, there exists an arc from v_i to v_j.

This thesis investigates several aspects of word-representable graphs and semi-transitive orientability, addressing multiple open problems and introducing new structural frameworks. A partial classification of all 2-word-representable graphs is obtained based on the length of their minimum word-representants. In this setting, a relationship is established between the number of occurrences of letters in a minimum-length word-representant and the diameter of the corresponding graph. Furthermore, an upper bound is derived on the minimum possible length of a word-representant under certain word-representability-preserving graph operations, including connecting two graphs by an edge and gluing two graphs at a vertex. The minimum length of word-representants for Cartesian and rooted graph products is then investigated using morphism-based and occurrence-based functions, respectively.

A central theme of this work is the study of semi-transitivity. The semi-transitive orientability of circulant graphs is examined, and an upper bound on the representation number is obtained for certain k-regular circulant graphs.

Let $\mathcal{E}_{i,j}$ denote the class of graphs whose vertex set can be partitioned into at most i independent sets and j cliques. It is known that $\mathcal{E}_{i,j}$ is word-representable for $(i, j) \in$

$\{(1,0),(0,1),(2,0),(3,0)\}$. Consequently, for the classes $\mathcal{E}_{1,1}$ (split graphs) and $\mathcal{E}_{0,2}$ (co-bipartite graphs), complete characterizations of word-representability in terms of forbidden induced subgraphs are obtained in this work.

In this context, a new matrix property, termed the *I-circular property*, is introduced. This property is closely related to the well-known D-circular property introduced by Safe. The I-circular property requires that both the rows of a matrix and the pairwise intersections of its rows form circular intervals under some linear ordering of the columns. Using this property, a direct correspondence is established between the structure of semi-transitive split graphs and the matrix representation of their adjacency relationships. A complete forbidden submatrix characterization of the I-circular property is obtained, which, in turn, yields a characterization of semi-transitive split graphs in terms of minimal forbidden induced subgraphs.

Finally, a structural and algorithmic characterization of word-representable *co-bipartite graphs*—graphs whose vertex set can be partitioned into two cliques—is presented. It is first shown that a co-bipartite graph is a circle graph if and only if it is a permutation graph, leading to a minimal forbidden induced subgraph characterization of co-bipartite circle graphs. The central contribution then establishes a connection between semi-transitivity and the circularly compatible ones property of binary matrices. In addition to these structural results, a linear-time recognition algorithm for semi-transitive co-bipartite graphs is developed by adapting Safe's matrix recognition framework.

Contents

List of Symbols

Symbol	**Description**
$\mathbb{N}$	Set of positive integers
$[k]$	The set $\{1, 2, \ldots, k\}$
$\|S\|$	Cardinality of a set S
w	A word over an alphabet
$\|w\|$	Length of the word w
$w\|_X$	Subword induced by letters in X
$O(w, i)$	Set of letters of w that occur exactly i times
$O_w(x)$	Number of occurrences of letter x in w
$O_{\max}(w)$	Maximum number of occurrences of any letter in w
$O_{\min}(w)$	Minimum number of occurrences of any letter in w
$r(w)$	Reverse of the word w
$\pi(w), \sigma(w)$	Initial and final permutation of word w
G	A simple, undirected graph
$V(G)$	Vertex set of graph G
$E(G)$	Edge set of graph G
$\|G\|$	Number of vertices in G
$N_G(v)$	Open neighbourhood of vertex v

$N_G[v]$	Closed neighbourhood of vertex v
$\deg_G(v)$	Degree of vertex v
$\Delta(G)$, $\delta(G)$	Maximum and minimum degree of G
$d(u,v)$	Distance between vertices u and v
$\mathrm{diam}(G)$	Diameter of G
$G[X]$	Induced subgraph on $X \subseteq V(G)$
$\overline{G}$	Complement of graph G
κ_G	Size of a maximal clique of G
$G(w)$	Graph represented by word w
$\mathcal{R}(G)$	Representation number of graph G
$\ell(G)$	Minimum length of a word-representant of G
$G \cup H$	Union of graphs
$G + H$	Join of graphs
$G \square H$	Cartesian product of graphs
$G \circ H$	Rooted product of graphs
M	Binary matrix
r, s	Rows of matrix M
$r \subseteq s$	Row containment
$r \cap s$	Intersection of rows
$M_{\rho,\sigma}$	Submatrix defined by row and column maps
$\langle n_1, n_2, \ldots, n_s \rangle$	Injective function with domain $[s]$ defined by $i \mapsto n_i$, where $n_1, \ldots, n_s$ are pairwise distinct positive integers
$M_{\langle i \rangle}$	$1 \times \ell$ matrix whose only row is the i-th row of M

List of Figures

List of Tables

Introduction

The theory of word-representable graphs has emerged as a vibrant and growing area within discrete mathematics, uniting concepts from structural graph theory, combinatorics on words, and abstract algebra. Interestingly, the motivation for studying word-representable graphs did not originate within graph theory itself, but rather from a problem in abstract algebra related to the Perkins semigroup. The formal connection between the two areas was first established by Kitaev and Seif in their investigation of the word problem for this particular semigroup [37].

To address the word problem for the Perkins semigroup, Kitaev and Seif introduced a novel combinatorial object known as the *alternation word digraph*, denoted by $\mathrm{Alt}(w)$, which is associated with a word w [37, 23, 22]. The vertices of this directed acyclic graph (DAG) correspond to the non-empty proper subsets of the variables (alphabet) appearing in w. An edge exists from a vertex X to a vertex Y (where X and Y are disjoint non-empty proper subsets of variables) if and only if the letters in X and those in Y alternate within the subword of w induced by $X \cup Y$, and the first letter of this subword belongs to X. The central result of their work established that two words u and v are equivalent over the Perkins semigroup B_2^1 if and only if their corresponding alternation word digraphs, $\mathrm{Alt}(u)$ and $\mathrm{Alt}(v)$, are identical. This result provided a complete solution to the word problem through a graph-theoretic characterization of equivalence.

This algebraic investigation laid the conceptual foundation for the theory of word-representable graphs. The idea of *alternation* between letters—originally defining the edges in the digraph $\mathrm{Alt}(w)$—emerged as a property of intrinsic combinatorial interest. The key creative leap that initiated this new line of study was the simplification of the concept itself. Rather than examining alternation between sets of variables, one could focus on the simplest case: alternation between individual variables. This relationship could then define the edges of a simple graph whose vertices correspond to the letters of the word.

This abstraction effectively detached the notion of alternation from the structural complexity of the full alternation word digraph and from the specific algebraic properties of the Perkins semigroup, allowing it to be studied as a purely graph-theoretic concept. This pivotal shift in perspective—from an algebraic construct to a graph-theoretic definition—was formalized by Kitaev and Pyatkin in 2008, marking the beginning of the systematic exploration of word-representable graphs [33].The formal definition of a word-representable graph captures the essence of the alternation property in a direct and elegant manner (see *Section* 1.5).

This foundational study by Kitaev and Pyatkin established several key properties of this class of graphs:

- The class of word-representable graphs is *hereditary*, meaning that every induced subgraph of a word-representable graph is itself word-representable. This follows directly from the definition: if a word w represents a graph G, and G' is the subgraph induced by $V' \subseteq V(G)$, then the word obtained by deleting all letters not in V' from w represents G'.
- Not all graphs are word-representable. The smallest non-word-representable graph (by number of vertices) is the wheel graph on six vertices, W_5 — a 5-cycle with a central vertex adjacent to all cycle vertices. This graph serves as a fundamental forbidden induced subgraph for the class.
- The definition is independent of vertex labeling. Any relabeling of the vertices of a word-representable graph produces another word-representable graph, since the same permutation can be applied to the letters of the representing word.

A central challenge in the early study of word-representable graphs concerned the nature of their representing words. The definition permits words of arbitrary length and allows letters to appear any number of times, which raises the question of whether finding a representant for a given graph is a finite problem. A foundational result by Kitaev and Pyatkin [33] resolved this issue by showing that the search can be restricted to a well-structured class of words.

A word is said to be *k-uniform* if every letter in its alphabet appears exactly k times. A graph is *k-representable* if there exists a k-uniform word-representant for that graph. Kitaev and Pyatkin proved that a graph is word-representable if and only if it is k-representable for some positive integer k. This theorem is of fundamental importance because it ensures that if a word-representant exists at all, then there also exists a *uniform* one, in which each vertex label appears the same number of times. Consequently, the problem of finding a word-representant is transformed from an infinite search over arbitrary words to a sequence of finite search problems, one for each $k = 1, 2, 3, \ldots$. This property further allows the definition of a natural complexity measure for word-representable graphs, known as the *representation number* (see *Section 1.7*).

The representation number provides a quantitative measure of how "complex" a graph's representation is. The existence of a well-defined minimum is guaranteed by the property that k-representability implies $(k+1)$-representability.

The classes of graphs with small representation numbers have been completely characterized and correspond to well-known graph families. A graph is 1-representable if and only if it is a *complete graph* (K_n). A 1-uniform word is simply a permutation of the vertex set, and in any permutation, every pair of distinct letters alternates exactly once, thereby representing a complete graph [33].

The class of graphs with representation number two coincides exactly with the class of *non-complete circle graphs*. Circle graphs are intersection graphs of chords on a circle, and their correspondence with 2-uniform words reveals a deep connection between word-representable graphs and a classical family of intersection graphs.

The study of graphs with representation number three or higher remains an active and evolving area of research. Notably, the Petersen graph and all prism graphs are known to have a representation number of three [29]. The representation number thus serves as a key tool for organizing the class of word-representable graphs and for exploring the intricate relationship between a graph's structural properties and the complexity of its word encoding.

While the definition of word-representable graphs is rooted in combinatorics on words, the most significant theoretical breakthrough in the field came from structural graph theory. The discovery of an equivalent characterization in terms of graph orientations transformed the study of these graphs, providing a powerful analytical tool and connecting the theory to broader concepts in graph algorithms and structure. This tool is the semi-transitive orientation (See *Section* 1.8). A semi-transitive orientation is a specific type of acyclic orientation that generalizes the well-known concept of transitive orientation.

Other orientations have been defined to capture generalizations of transitive orientation. While transitive orientations impose constraints on the orderings of induced P_3, these generalizations impose constraints on the orderings of induced P_4. These include *perfectly orderable graphs* (and their subclasses) and *opposition graphs*. Classes such as chordal graphs are defined in terms of vertex orderings and, therefore, imply acyclic orientations indirectly. However, none of these generalizations characterize word-representable graphs. In [23, 24], Halldórsson et al. introduced another generalization of transitive orientation, known as the *semi-transitive orientation*.

The fundamental connection between the realm of words and the realm of orientations was established by Halldórsson, Kitaev, and Pyatkin through a series of influential papers that form the cornerstone of the modern theory. In their work [24], Halldórsson et al. provided a necessary and sufficient condition stating that a graph is word-representable if and only if it admits a *semi-transitive orientation*. This result forms a powerful bridge between two seemingly distinct domains, enabling researchers to use the structural properties of oriented graphs to derive results about word-representability, and conversely, to apply combinatorial reasoning on words to orientation problems. This equivalence has several profound and immediate implications that have fundamentally shaped the development and direction of the field.

A natural question arises regarding how the class of word-representable graphs relates to other generalizations of circle graphs. One such class is that of *polygon-circle graphs*, defined as the intersection graphs of polygons inscribed in a circle. The work of Enright and Kitaev [15] provides a clear understanding of this relationship. They presented explicit examples showing that the classes of word-representable graphs and polygon-circle graphs are *incomparable*—neither class is a subset of the other.

For instance, the Petersen graph and the family of crown graphs are known to be word-representable but were shown not to be polygon-circle graphs. Conversely, the wheel graph W_5—the canonical example of a small non-word-representable graph—is a polygon-circle graph. This distinction is significant because it establishes word-representable graphs as a separate and independent generalization of circle graphs, possessing properties that differ fundamentally from those of geometric intersection-based generalizations.

Until 2024, there was no known characterization of the classes $\mathcal{R}_k$ for $k \geq 3$. In [17], Fleischmann et al. introduced a generalization of circle graphs that characterizes the classes $\mathcal{R}_k$ for $k \geq 3$. They defined a notion called the *k-circle representation*, as follows: for $k \in \mathbb{N}$, a k-circle representation of a graph G is a circle with an inscribed k-gon corresponding to each vertex of G, such that the sides of any two k-gons intersect $2k$ times if and only if there is an edge between the corresponding vertices of G. Using this notion, they provided a characterization of the classes $\mathcal{R}_k$ for $k \geq 3$; that is, a graph is k-representable if and only if it admits a k-circle representation.

As the theory of word-representable graphs has matured, research has branched out in several directions. These include integrating the theory with mainstream graph theory by examining its behavior under standard graph operations. A central question for any newly defined graph class concerns its behavior under standard graph operations. Understanding which operations preserve membership in the class is essential for constructing larger examples and for developing a coherent structural theory. For word-representable graphs, this question takes the form: which graph operations preserve semi-transitive orientability? This problem was systematically studied by Choi et al. [10].

Their work established both positive and negative results. On the positive side, several fundamental local operations were shown to preserve semi-transitive orientability. In particular, operations such as vertex deletion, edge deletion, edge addition, and edge lifting preserve semi-transitive orientability under certain conditions [10].

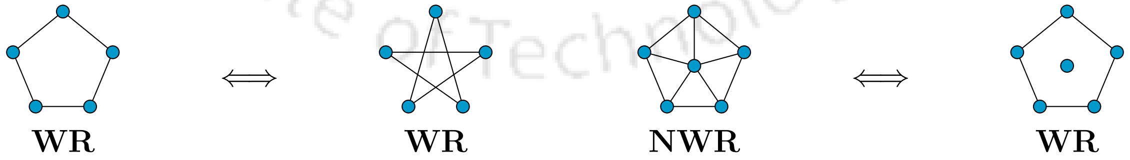


Figure 1: Examples of complement operation

Conversely, they resolved a question posed by Kitaev and Lozin [32] by demonstrating that several standard graph products—such as the tensor, lexicographic, and strong products—as well as operations such as complementation and edge contraction, do not, in general, preserve semi-transitive orientability. Starting with either a word-representable or a

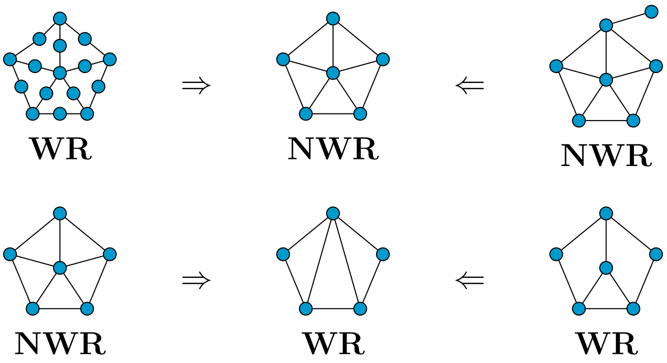


Figure 2: Examples of edge contraction

non-word-representable graph, the graph obtained by taking its complement or by performing an edge contraction may be word-representable or non-word-representable. Figures 1 and 2 illustrate this observation.

These negative results are significant because they delineate the boundaries of compositional methods for proving word-representability. In particular, one cannot establish that a graph is word-representable merely by expressing it as a product of representable factors under these operations.

With the powerful characterization provided by semi-transitive orientations, much of the subsequent research in the field has turned toward a systematic exploration of the graph universe. This line of investigation seeks to determine which well-known graph families are word-representable and which are not, thereby delineating the boundaries of the class and clarifying its position within the broader framework of graph theory. Considerable effort has been devoted to studying the word-representability and representation numbers of specific, highly structured graph families. These investigations not only classify such families but also reveal deeper structural properties and frequently require the development of new constructive techniques.

The *n-prism graph*, denoted by Pr_n, is the Cartesian product of a cycle C_n and an edge K_2. Kitaev showed that these graphs form an interesting family with a representation number of three [29]. This result was among the first to establish a non-trivial infinite family of graphs belonging to $\mathcal{R}_3$, the class of graphs with representation number three.

The *k-dimensional cube* (or hypercube), Q_k, is defined as the Cartesian product of k copies of K_2. Determining its representation number posed a natural and intriguing problem. Broere and Zantema provided a constructive proof that established a tight upper bound [6]. Their result followed as a corollary of a more general theorem concerning Cartesian products: if a graph G is k-representable (for $k > 1$), then the Cartesian product $G \square K_n$ is $(k+n-1)$-representable. Since $Q_k = Q_{k-1} \square K_2$ and $Q_1 = K_2$ is 1-representable, a straightforward induction shows that Q_k is k-representable [6, 5].

More recently, Hefty et al. [25] improved this understanding by proving that the representation number of the k-dimensional cube is $O\left(\frac{\log k}{\log \log k}\right)$, providing the best known asymptotic bound to date.

A *crown graph* $H_{n,n}$ is obtained from the complete bipartite graph $K_{n,n}$ by removing a perfect matching. Since crown graphs are bipartite, they are also comparability graphs, and hence their word-representability is guaranteed. However, their representation number turns out to be surprisingly large, as shown by Glen, Kitaev, and Pyatkin [20]. For $n \geq 5$, the representation number of the crown graph $H_{n,n}$ is $\lceil n/2 \rceil$.

This result is particularly significant for several reasons. First, it provides a complete solution to an open problem posed in the monograph by Kitaev and Lozin [32]. Second, it identifies a family of bipartite graphs with arbitrarily high representation numbers, thereby offering a negative answer to another question from the same monograph—whether all bipartite graphs are 3-word-representable. The existence of crown graphs thus reveals that even within the structurally simple class of bipartite graphs, the complexity of word-representations can grow as the size of the graph increases.

A particularly challenging area in the study of word-representable graphs concerns *triangle-free graphs*. Since all 3-colorable graphs are word-representable, any non-word-representable graph must have a chromatic number of at least four. This naturally raises the question: do there exist triangle-free graphs that are not word-representable?

The existence of such graphs was first established non-constructively by Halldórsson, Kitaev, and Pyatkin [23]. Their proof relied on Erdös's celebrated theorem guaranteeing the existence of graphs with arbitrarily large girth and chromatic number. They showed that graphs with both high girth and high chromatic number must contain induced subgraphs that admit no semi-transitive orientation, thereby proving the existence of triangle-free non-word-representable graphs—albeit without providing explicit examples.

The search for concrete examples led researchers to investigate *Kneser graphs*. The Kneser graph $K(n, k)$ has as its vertices the k-element subsets of an n-element set, with two vertices adjacent if and only if their corresponding subsets are disjoint. For $n < 3k$, the graph $K(n, k)$ is triangle-free. Through computational analysis, Kitaev and Saito [36] provided the first explicit example of a triangle-free non-word-representable graph within this family, namely $K(8, 3)$. Their work also established general conditions for the semi-transitive orientability of Kneser graphs and their complements, showing, for instance, that $K(n, k)$ is not semi-transitively orientable for $n \geq 15k - 24$.

Following this discovery, Kitaev and Pyatkin [34] investigated the smallest possible triangle-free non-word-representable graphs. They proved that the Grötzsch graph, which has 11 vertices, is the smallest such graph. Moreover, they showed that the Chvátal graph, on 12 vertices, is the smallest 4-regular triangle-free non-word-representable graph. This result provides a definitive answer to the question of minimality, grounding the previously abstract existence proofs in concrete and structurally significant examples that play a central role in the study of graph coloring.

Cheon et al. [9] initiated the study of word-representability for *Toeplitz graphs*, which are graphs whose adjacency matrices are Toeplitz matrices (that is, constant along each diagonal). These graphs form a specific subclass of *Riordan graphs*. This work unified the theory of Riordan matrices with the study of word-representable graphs, offering the first systematic exploration of graph representability based on structural patterns in adjacency matrices. A key contribution of the paper was the introduction of the concept of an *infinite word-representable graph*, which extended the framework of word-representability beyond finite graphs for the first time.

De Bruijn graphs are fundamental objects in combinatorics on words and have important applications in areas such as genome assembly. Petyuk [38] introduced the notion of a *simplified de Bruijn graph* $S(n, k)$, defined as the simple graph underlying the standard directed de Bruijn graph $B(n, k)$, and investigated its word-representability. He conjectured that $S(n, k)$ is non-word-representable for all $n \geq 4$ and $k \geq 3$.

To advance the study of complex graphs such as de Bruijn graphs, Huang, Kitaev, and Pyatkin [26] introduced a novel approach based on *graph homomorphisms*. The central idea is that a homomorphism $f : G \to H$ can be used to obtain a semi-transitive orientation from H to induce a semi-transitive orientation on G. When H is a simpler graph—such as a complete graph—with a known semi-transitive orientation, this technique can guide the search for a suitable orientation on a more complex graph G. This marks the first systematic use of graph homomorphisms in the study of word-representable graphs.

As a proof of concept, the authors applied this method to demonstrate the word-representability of a newly introduced graph, the *simplified graph of overlapping permutations*, and to establish the representability of certain subgraphs of the de Bruijn graphs previously examined by Petyuk in [38]. This embedding technique introduces a powerful new tool for addressing word-representability problems in structurally intricate graph families.

A *split graph* is a graph whose vertex set can be partitioned into a clique C and an independent set I. Despite the extensive study of both split graphs and word-representable graphs, the intersection of these two areas remained largely unexplored until a series of recent papers initiated a systematic investigation [31, 8, 28, 27, 35]. The works of Kitaev et al. [31] and Chen, Kitaev, and Saito [8] laid the foundation for this subfield. Their approach characterized word-representable split graphs via *forbidden induced subgraphs*, a classical method in structural graph theory. By leveraging the restrictions imposed by semi-transitive orientations on the clique part, they derived complete theoretical and computational characterizations for split graphs in which the clique has a small, fixed size.

For split graphs with a clique of size four, a complete characterization in terms of forbidden induced subgraphs was obtained in [31]. This result was later extended to split graphs with a clique of size five, which were characterized by a family of nine forbidden induced subgraphs [8].

A novel approach to studying families of split graphs was introduced by Iamthong, drawing on techniques from combinatorics on words and formal language theory [27]. This method employs *morphisms* (substitutions on an alphabet) to iteratively generate infinite families of graphs. A morphism can be applied to the adjacency matrix of a graph to produce a larger graph, and by analyzing the properties of the morphism, one can deduce structural properties of the entire family of graphs generated through its iteration.

Iamthong [27] provided a new characterization of word-representable split graphs based on permutations of the columns of the adjacency matrix corresponding to the independent set. This characterization is particularly amenable to analysis via morphisms. The main contribution of this work was a complete classification of word-representable split graphs generated by morphisms defined by 2×2 matrices.

This powerful technique was later extended in a joint work by Iamthong and Kitaev [28] to the study of *directed* split graphs. They examined the property of semi-transitivity for families of directed split graphs generated by morphisms and achieved a full classification of semi-transitive infinite directed split graphs generated by a broad class of morphisms involving matrices over $\{-1, 0, 1\}$. This line of research highlights the deep interplay between the structural graph theory of semi-transitive orientations and the algebraic and combinatorial framework of morphisms.

The most significant result concerning word-representable split graphs is the resolution of their recognition complexity. While the general problem of recognizing word-representable graphs is NP-complete, the strong structural constraints of split graphs suggested that their recognition might be tractable. This conjecture was confirmed by Kitaev and Pyatkin [35], who proved that the problem of determining whether a given split graph is word-representable can indeed be solved in polynomial time.

This result represents a major breakthrough, establishing split graphs as one of the most extensive and structurally rich classes of graphs for which word-representability can be efficiently decided. The recognition problem is reduced to checking specific structural properties of how vertices in the independent set connect to the transitively oriented clique, which can be modeled as a matrix problem involving the *circular-ones property*, solvable in polynomial time.

The successful research program on split graphs provided a clear methodological template: identify a graph class with a strong vertex partition, leverage the constraints this partition imposes on semi-transitive orientations, and derive structural characterizations or recognition algorithms. A natural next step is to generalize the structure of split graphs. Instead of partitioning the vertices into a clique and an independent set, one can consider a partition into two cliques.

This direction was pursued by Chen, Hameed, and Kitaev in their recent work on the word-representability of $K_m - K_n$ graphs [7]. A $K_m - K_n$ graph is defined as a graph whose vertex set can be partitioned into two cliques, K_m and K_n. Following the established methodology, the authors provided a complete characterization of word-representable $K_m - K_n$ graphs in terms of forbidden induced subgraphs for cases where one of the cliques is small ($m \leq 4$). For the class of $K_4 - K_n$ graphs, they proved that there are exactly seven minimal non-word-representable graphs. The table below summarizes the word-representability of various graph classes.

Graph Classes	**WR or NWR**	**Characterization**	**References**
Toeplitz graphs	Contains both	No	[9]
Kneser graphs	Contains both	No	[36]
De Bruijn graphs	Contains both	No	[38, 26]
Split graphs	Contains both	Yes	[31, 28, 27, 8, 35]
Triangle-free graphs	Contains both	No	[34]
Circulant graphs	Contains both	No	[34]
Co-bipartite graphs	Contains both	Yes	[14]

Table 2: Word-representability of certain graph classes

While much of the research in this field has employed graph-theoretic tools to study the properties of representable graphs, an emerging trend shifts the focus back to the combinatorial structure of the words themselves. This *word-centric perspective* asks not which graphs are representable, but rather what intrinsic properties the representing words possess.

This direction was pursued by Gaetz and Ji, who investigated the properties of minimal-length word-representants [18]. Instead of restricting attention to k-uniform words, they examined words of the absolute minimal length required to represent a given graph. For trees and cycles on n vertices, they derived explicit formulas for both the length of a minimal word-representant and the number of distinct minimal words representing each graph.

Overall, the study of word-representable graphs has grown into an active area of research that brings together ideas from combinatorics on words and graph theory. Although many important results have been obtained on representation numbers, semi-transitive orientations, and specific graph families, several fundamental questions remain open. In particular, there is a need for clearer structural characterizations, efficient recognition algorithms, and a deeper understanding of the relationship between words and graphs. Motivated by these challenges, this thesis focuses on three main directions: the study of minimal-length word-representants, the semi-transitive orientability of structured graph classes, and forbidden induced subgraph characterizations using matrix-based methods.

Organization of the Thesis

We present the entire work of this thesis in six chapters as described below.

- Chapter 1: Preliminaries
- Chapter 2: Minimum length word-representants of word-representable graphs and graph products
- Chapter 3: On semi-transitive orientability of circulant graphs
- Chapter 4: Forbidden induced subgraph characterization of word-representable split graphs
- Chapter 5: Forbidden induced subgraph characterization of word-representable co-bipartite graphs

Results Obtained in the Thesis

In Chapter 1, we discuss the necessary background and foundational concepts required to understand the work presented in the subsequent chapters. This includes essential definitions and results concerning word-representability, graph-theoretic notions, and relevant matrix properties that will be utilized throughout the rest of the thesis.

In Chapter 2, we address an open problem posed by Gaetz and Ji [18], solving two of its three cases and providing a partial solution to the remaining one. The problem concerns the classification of all 2-word-representable graphs based on the length of their minimal word-representants. In this context, we also establish a relationship between the maximum and minimum number of occurrences of letters in minimum length word-representants and the diameter of the corresponding graph.

Furthermore, we derive an upper bound on the minimum possible length of a word-representant under certain word-representability preserving graph operations, such as connecting two graphs by an edge and gluing two graphs at a vertex. We then investigate the minimum length of word-representants for Cartesian and rooted graph products using morphism-based and occurrence-based functions, respectively.

Finally, we resolve an open problem posed by Broere in his master's thesis [5], which asks for the construction of a word representing the Cartesian product of two arbitrary word-representable graphs.

In [34], Kitaev and Pyatkin showed that every 4-regular circulant graph is semi-transitive. They also posed an open problem concerning the semi-transitive orientability of circulant graphs in which the elements of the set $\{a_1, a_2, \ldots, a_k\}$ are consecutive positive integers. In Chapter 3, we investigate the semi-transitive orientability of circulant graphs and demonstrate that certain circulant graphs are semi-transitive, while others are not, under specific conditions. Additionally, we provide an upper bound on the representation number of particular k-regular circulant graphs.

In Chapter 4, we introduce a new matrix property, termed the I-circular property, which is closely related to the well-known D-circular property introduced by Safe in [40]. The I-circular property requires that both the rows of a matrix and the pairwise intersections of its rows form circular intervals under some linear ordering of the columns. Leveraging this property, we establish a direct correspondence between the structure of semi-transitive split graphs and the matrix representation of their adjacency relationships. Our main result is a complete forbidden submatrix characterization of the I-circular property, which, in turn, yields a characterization of semi-transitive split graphs in terms of minimal forbidden induced subgraphs.

In Chapter 5, we present a comprehensive structural and algorithmic characterization of word-representable *co-bipartite graphs*—graphs whose vertex set can be partitioned into two cliques. Building upon Bouchet's theory of circle graphs [4] and Tucker's framework of circular-ones matrices [44, 43], this work unifies graph-theoretic and matrix-theoretic perspectives. We first establish that a co-bipartite graph is a circle graph if and only if it is a permutation graph, thereby obtaining a minimal forbidden induced subgraph characterization for co-bipartite circle graphs. The central contribution then links semi-transitivity with the circularly compatible ones property of binary matrices. Beyond the structural results, we also develop a linear-time recognition algorithm for semi-transitive co-bipartite graphs by adapting Safe's matrix recognition framework.

1

Preliminaries

For each positive integer k, we denote the set $\{1, 2, \ldots, k\}$ by $[k]$. If $k = 0$, we define $[k]$ to be the empty set. We denote the identity function by id_k. A set S is said to be properly contained in T if $S \subseteq T$ and $S \neq T$. Whenever we say that i is modulo k, or that addition and subtraction involving i are taken modulo k, we mean that i, or any addition or subtraction involving i, is taken modulo k and yields the unique element $j \in [k]$ with the same remainder as j upon division by k. Let $gcd(a_1, a_2, \ldots, a_n)$ denote the greatest common divisor of the integers $a_1, a_2, \ldots, a_n$.

1.1 Sequences

Let $a = a_1a_2 \ldots a_k$ be a sequence of length k. The shift of a is the sequence $a_2a_3 \ldots a_ka_1$. The length of any sequence a is denoted by $|a|$. A *binary bracelet* [41] is a lexicographically smallest element in an equivalence class of binary sequences under shifts and reversals. For each $k \geq 4$, let A_k be the set of binary bracelets of length k. Let $A_3 = \{000, 111\}$.

The elements of A_3 are binary bracelets but the two other binary bracelets of length 3 (001 and 011) do not belong to A_3. If b is a sequence and $i \in [k]$, then b is said to occur circularly in a at position i if $|b| \leq k$ and $a_i a_{i+1} \dots a_{i+|b|-1} = b$, where the subscripts are taken modulo k. If a is binary (that is, each $a_i \in \{0, 1\}$), then the complement of a is defined as the sequence $\overline{a}$ obtained from a by interchanging 0's and 1's. An index map of a is an injective function $\rho \colon [k'] \to [k]$. If ρ is an index map of a, then the sequence $a_{\rho(1)} a_{\rho(2)} \dots a_{\rho(k')}$ is denoted by a_ρ. A sequence is quaternary if each of its elements lies in the set $\{0, 1, 2, 3\}$.

1.2 Graphs

Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. For a vertex $v \in V(G)$, the *neighbourhood* of v in G, denoted by $N_G(v)$, is the set of vertices adjacent to v in G. The *closed neighbourhood* of v, denoted by $N_G[v]$, is defined as $N_G(v) \cup \{v\}$. The *degree* of a vertex v in G, denoted by $\deg_G(v)$, is the cardinality of the set $N_G(v)$. We denote by $\Delta(G)$ (respectively, $\delta(G)$) the maximum (respectively, minimum) degree of the graph G. For two vertices x and y, $d(x, y)$ denotes the distance between them, i.e., the length of a shortest path connecting x and y. The *diameter* of a graph G, denoted by $\operatorname{diam}(G)$, is the maximum possible distance between any two vertices of G. For any subset $X \subseteq V(G)$, the *subgraph of G induced by X* is the graph with vertex set X and edge set consisting of all edges of G whose endpoints both lie in X. A graph G is said to *contain* a graph H as an *induced subgraph* if H is isomorphic to some induced subgraph of G. For a graph G, $|G|$ denotes the *order* of the graph.

An *independent set* (respectively, *clique*) of a graph G is a set of vertices that are pairwise nonadjacent (respectively, pairwise adjacent). A graph class is said to be *hereditary* if it is closed under taking induced subgraphs. The *complement* of a graph G, denoted by $\overline{G}$, is the graph with vertex set $V(G)$ and edge set $E(\overline{G}) = \{xy \colon xy \notin E(G)\}$. The edges of $\overline{G}$ are called the *non-edges* of G.

The *union* of two graphs G_1 and G_2, denoted by $G_1 \cup G_2$, is the graph with vertex set $V(G_1) \cup V(G_2)$ and edge set $E(G_1) \cup E(G_2)$. The *join* of G_1 and G_2, denoted by $G_1 + G_2$, is obtained from $G_1 \cup G_2$ by adding all edges between every vertex of G_1 and every vertex of

G_2. A vertex v of a graph G is called *universal* if it is adjacent to every other vertex of G, and *isolated* if it is adjacent to none. For an edge $xy \in E(G)$, let $G - xy$ denote the graph obtained by deleting xy. For a non-edge $xy \in E(\overline{G})$, let $G + xy$ denote the graph obtained by adding xy. More generally, for $S \subseteq E(G)$ (respectively, $S \subseteq E(\overline{G})$), we write $G - S$ (respectively, $G + S$) for the graph obtained from G by deleting (respectively, adding) all edges in S.

A graph is called an *intersection graph* if there exists a set of objects such that each vertex corresponds to an object, and two vertices are adjacent if and only if the corresponding objects have a non-empty intersection. A graph G is said to be *bipartite* if its vertex set can be partitioned into two subsets such that each subset induces an independent set. A graph is *co-bipartite* if its complement is bipartite. A graph is called a *circle graph* if its vertices can be associated with the chords of a circle such that two vertices are adjacent if and only if the corresponding chords intersect.

Let π be a permutation over $[n]$. The *permutation graph* of π is the graph with vertex set $[n]$, where two vertices i and j with $i < j$ are adjacent if and only if j occurs before i in π. The class of permutation graphs forms a subclass of the class of circle graphs.

A *local complementation* of a graph G at a vertex v is the graph, denoted by $G * v$, obtained by replacing the subgraph induced by the neighbours of v with its complement. A sequence of local complementations at vertices $v_1, v_2, \ldots, v_k$ is denoted by $G * v_1 v_2 \ldots v_k$. Two graphs G and G' are said to be *locally equivalent* if there exists a sequence m of vertices such that $G' = G * m$. An *ℓ-reduction* of G is an induced subgraph of a graph locally equivalent to G.

All graphs in this study are assumed to be simple, that is, finite, undirected, and having no loops or multiple edges. For additional graph-theoretic terminology and notation, we refer the reader to [48].

1.3 Matrices

All the matrices considered in this thesis are binary, that is, every entry of the matrices is either 0 or 1. Each row r of a matrix M is defined as the set of indices of all columns of M having a 1 in row r. A row r is said to be contained in a row s, denoted by $r \subseteq s$, if s

has a 1 in every column where r has a 1. Similarly, if r and s are rows of M, then $r \cap s$ denotes the set of columns having a 1 in both r and s. A row is empty if all its entries are 0. A row is trivial if all its entries are equal to the same value. Analogous conventions apply to the columns of M.

Let M be a $k \times \ell$ binary matrix with rows and columns indexed by $[k]$ and $[\ell]$, respectively. The *complement* of a row i is obtained by interchanging the 0's and 1's in that row. The *complement* of M, denoted by $\overline{M}$, is the matrix obtained by complementing every row of M.

For a binary sequence a of length k, we denote by $a \oplus M$ the matrix obtained from M by complementing exactly those rows $i \in [k]$ for which $a_i = 1$.

A *row map* of a matrix is an injective function $\rho \colon [k'] \to [k]$ for some positive integer k', and a *column map* is an injective function $\sigma \colon [\ell'] \to [\ell]$ for some positive integer ℓ'. Given a row map ρ and a column map σ, we denote by $M_{\rho,\sigma}$ the $k' \times \ell'$ matrix whose (i, j)-entry is equal to the $(\rho(i), \sigma(j))$-entry of M.

Let M and M' be matrices. We say that M *contains* M' as a configuration if some submatrix of M is equal to M' up to permutations of rows and columns. Equivalently, M contains M' as a configuration if and only if there exist a row map ρ and a column map σ such that $M_{\rho,\sigma} = M'$. Two matrices M and M' are said to *represent the same configuration* if one can be obtained from the other by permuting rows and columns.

Let a be a binary sequence whose length equals the number of rows of a matrix M. Then

$$(a \oplus M)_{\rho,\sigma} = a_\rho \oplus M_{\rho,\sigma}.$$

Moreover, if ρ and σ are a row map and a column map of M, and if ρ' and σ' are a row map and a column map of $M_{\rho,\sigma}$, then the compositions $\rho' \circ \rho$ and $\sigma' \circ \sigma$ are a row map and a column map of M, respectively, and

$$M_{\rho' \circ \rho, \sigma' \circ \sigma} = \left(M_{\rho,\sigma}\right)_{\rho',\sigma'}.$$

For a positive integer s and pairwise distinct positive integers $n_1, n_2, \ldots, n_s$, we denote by $\langle n_1, n_2, \ldots, n_s \rangle$ the injective function from $[s]$ to the positive integers that maps i to n_i for each $i \in [s]$. In particular, for $i \in [k]$, the notation $M_{\langle i \rangle}$ denotes the $1 \times \ell$ matrix whose single row is the i-th row of M.

Finally, for any matrix M, we denote by M^* the matrix obtained from M by appending an empty column.

1.4 Matrix Properties

Definition 1.4.1. *Let $\preccurlyeq_x$ be a linear order on a set X. For elements $a, b \in X$ with $a \preccurlyeq_x b$, the* linear interval of $\preccurlyeq_x$ with endpoints a and b, *denoted by $[a, b]_{\preccurlyeq_x}$, is defined as the set $\{x \in X : a \preccurlyeq_x x \preccurlyeq_x b\}$. A* linear interval of $\preccurlyeq_x$ *is either the empty set or an interval of the form $[a, b]_{\leq_X}$ for some $a, b \in X$ satisfying $a \preccurlyeq_x b$. A sequence $a_1a_2 \ldots a_k$ is said to be* monotone *on X if $a_1, a_2, \ldots, a_k \in X$ and $a_1 \preccurlyeq_x a_2 \preccurlyeq_x \cdots \preccurlyeq_x a_k$.*

Definition 1.4.2. *A matrix M has the consecutive-ones property for rows if there is a linear order $\preccurlyeq_C$ of the columns of M such that each row of M is a linear interval of $\preccurlyeq_C$, and $\preccurlyeq_C$ is called a consecutive-ones order. Analogous definition apply to the columns of M. If no mention is made for rows or columns, we mean the corresponding property for the rows.*

Definition 1.4.3. *Let $\preccurlyeq_x$ be a linear order on a set X, and let $a, b \in X$. The* circular interval of $\preccurlyeq_x$ with endpoints a and b, *denoted by $[a, b]_{\preccurlyeq_x}$, is defined as follows: if $a \preccurlyeq_x b$, then*

$$[a, b]_{\preccurlyeq_x} = \{x \in X \mid a \preccurlyeq_x x \preccurlyeq_x b\},$$

while if $b \prec_X a$, then

$$[a, b]_{\preccurlyeq_x} = \{x \in X \mid x \preccurlyeq_x b \text{ or } a \preccurlyeq_x x\}.$$

A circular interval of $\preccurlyeq_x$ *is either the empty set or a set of the form $[a, b]_{\preccurlyeq_x}$ for some $a, b \in X$. A sequence $a_1a_2 \ldots a_k$ is said to be* circularly monotone with respect to $\preccurlyeq_x$ *if $a_1, a_2, \ldots, a_k \in X$ and the relation $a_i \preccurlyeq_x a_{i+1}$ holds for all but at most one index $i \in \{1, 2, \ldots, k\}$, where a_{k+1} is identified with a_1.*

Definition 1.4.4. *A matrix M has the circular-ones property for rows if there is a linear order $\preccurlyeq_C$ of the columns of M such that each row of M is a circular interval of $\preccurlyeq_C$, and $\preccurlyeq_C$ is called a circular-ones order. Analogous definition apply to the columns of M. If no mention is made for rows or columns, we mean the corresponding property for the rows.*

Theorem 1.4.5 ([44]). *Let M be a matrix, and let M' denote any matrix obtained from M by complementing some of its rows so that M' contains at least one column consisting entirely of zeros. Then, M has the circular-ones property if and only if M' has the consecutive-ones property.*

$$M_I(k) = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

(a) $M_I(k)$ for each $k \geq 3$

$$M_{II}(k) = \begin{pmatrix} 1 & 1 & & & & 0 \\ & 1 & 1 & & & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & 1 & 1 & 0 \\ 1 & 1 & \cdots & 1 & 0 & 1 \\ 0 & 1 & \cdots & 1 & 1 & 1 \end{pmatrix}$$

(b) $M_{II}(k)$ for each $k \geq 4$

$$M_{III}(k) = \begin{pmatrix} 1 & 1 & & & & 0 \\ & 1 & 1 & & & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & 1 & 1 & 0 \\ 0 & 1 & \cdots & 1 & 0 & 1 \end{pmatrix}$$

(c) $M_{III}(k)$ for each $k \geq 3$

$$M_{IV} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

(d) M_{IV}

$$M_V = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

(e) M_V

Figure 1.1: Tucker matrices, where k denotes the number of rows and omitted entries are 0's

Theorem 1.4.6 ([39]). *A matrix M has the circular-ones property if and only if M contains no matrix in the set $\mathcal{F}_{\text{circR}}$ as a configuration. The corresponding set of forbidden submatrices is*

$$\mathcal{F}_{\text{circR}} = \{\, a \oplus M_I^*(k) \colon k \geq 3 \text{ and } a \in A_k \,\} \cup \{\, M_{IV},\, \overline{M_{IV}},\, M_V^*,\, \overline{M_V^*} \,\},$$

where $M_I^*(k)$ and M_V^* denote $(M_I(k))^*$ and $(M_V)^*$, respectively, $A_3 = \{000, 111\}$, and for each $k \geq 4$, A_k is the set of all binary bracelets of length k. Notice that 001 and 011 *are binary bracelets of length* 3 *but do not belong to* A_3*. Moreover, there exists a linear-time algorithm that, given any matrix* M *not having the circular-ones property, outputs a matrix in* $\mathcal{F}_{\mathrm{circR}}$ *contained in* M *as a configuration. In addition, every matrix in* $\mathcal{F}_{\mathrm{circR}}$ *is a minimal forbidden submatrix for the circular-ones property. Hence, for each* $M \in \mathcal{F}_{\mathrm{circR}}$ *and each binary sequence* a *whose length equals the number of rows of* M*, the matrix* $a \oplus M$ *represents the same configuration as some matrix in* $\mathcal{F}_{\mathrm{circR}}$*.*

Definition 1.4.7. *A matrix* M *is said to have the* D-circular property *if there exists a linear order* $\preccurlyeq_c$ *of its columns such that every row of* M *is a circular interval with respect to* $\preccurlyeq_c$*, and for any two rows* r *and* s *of* M*, the set difference* $s - r$ *also forms a circular interval of* $\preccurlyeq_c$*. In this case,* $\preccurlyeq_c$ *is called a* D-circular order *of* M*.*

Definition 1.4.8. *A matrix* M *is said to have the* circularly compatible ones property *if there exists a biorder* $(\preccurlyeq_r, \preccurlyeq_c)$ *such that:*

(i) *each row of* M *forms a circular interval with respect to* $\preccurlyeq_c$*;*

(ii) *each column of* M *forms a circular interval with respect to* $\preccurlyeq_r$*; and*

(iii) *if* $r_1, r_2, \ldots, r_p$ *denote all nontrivial rows of* M *listed in ascending order of* $\preccurlyeq_r$*, and if* $r_i = [d_i, e_i]_{\preccurlyeq_c}$ *for each* $i \in \{1, 2, \ldots, p\}$*, then both sequences* $d_1 d_2 \ldots d_p$ *and* $e_1 e_2 \ldots e_p$ *are circularly monotone with respect to* $\preccurlyeq_c$*.*

In this case, the pair $(\preccurlyeq_r, \preccurlyeq_c)$ *is called a* circularly compatible ones biorder.

In [1], Basu et al. proved that, for a matrix M without trivial rows, the D-circular property coincides with the monotone circular property. We now state the definitions and results related to the monotone circular property.

Definition 1.4.9. *Let* $X = \{x_1, x_2, \ldots, x_k\}$ *with* $x_1 \preccurlyeq_x x_2 \preccurlyeq_x \cdots \preccurlyeq_x x_k$*. Define the set* $X^+ = \{x_1, x_2, \ldots, x_k, x_1^+, x_2^+, \ldots, x_k^+\}$*, and let* $\preccurlyeq_x^+$ *be the linear order on* X^+ *given by*

$$x_1 \preccurlyeq_x^+ x_2 \preccurlyeq_x^+ \cdots \preccurlyeq_x^+ x_k \preccurlyeq_x^+ x_1^+ \preccurlyeq_x^+ x_2^+ \preccurlyeq_x^+ \cdots \preccurlyeq_x^+ x_k^+.$$

For each nontrivial circular interval $[a, b]_{\preccurlyeq_x}$*, we define its* unwrapped interval relative to $\preccurlyeq_x$*, denoted by* $[a, b]^+_{\preccurlyeq_x}$*, as the linear interval* $[a, c]_{\preccurlyeq_x^+}$*, where* $c = b$ *if* $a \preccurlyeq_x b$*, and* $c = b^+$ *if* $b \prec_x a$*.*

Definition 1.4.10. *Let M be a matrix and let $(\preccurlyeq_r, \preccurlyeq_c)$ be a biorder of M. Let $r_1, r_2, \ldots, r_p$ denote all nontrivial rows of M listed in ascending order with respect to $\preccurlyeq_r$. If $\preccurlyeq_c$ is a circular-ones order of M, $r_i = [d_i, e_i]_{\preccurlyeq_c}$ for each $i \in [p]$, and $[d_i, f_i]^+_{\preccurlyeq_c}$ denotes the unwrapped interval of $[d_i, e_i]_{\preccurlyeq_c}$ with respect to $\preccurlyeq_c$, we say that M has* monotone left endpoints *with respect to $(\preccurlyeq_r, \preccurlyeq_c)$ if the sequence $d_1 d_2 \ldots d_p$ is monotone with respect to $\preccurlyeq_c$, and we say that M has* monotone unwrapped right endpoints *with respect to $(\preccurlyeq_r, \preccurlyeq_c)$ if the sequence $f_1 f_2 \ldots f_p$ is monotone with respect to $\preccurlyeq_c^+$.*

If M has no trivial rows, a monotone circular biorder *of M is a biorder $(\preccurlyeq_r, \preccurlyeq_c)$ satisfying all of the following conditions:*

- *(i) M has monotone left endpoints $d_1 d_2 \ldots d_p$ with respect to $(\preccurlyeq_r, \preccurlyeq_c)$;*
- *(ii) M has circularly monotone unwrapped right endpoints $f_1 f_2 \ldots f_p$ with respect to $(\preccurlyeq_r, \preccurlyeq_c)$;*
- *(iii) either $f_1 = e_1^+$, or both $f_1 = e_1$ and $f_p \preccurlyeq_c^+ e_1^+$.*

A matrix M having no trivial rows has the monotone circular property *if it admits a monotone circular biorder.*

The following result from [1] relates the D-circular property to the monotone circular property.

Theorem 1.4.11 ([1]). *If M is a matrix with no trivial rows, the following statements are equivalent:*

- *(i) M has the D-circular property;*
- *(ii) M has the monotone circular property.*

Definition 1.4.12. *A matrix M is said to have the* doubly D-circular property *if both M and its transpose M^t possess the D-circular property.*

Definition 1.4.13. *We denote*

$$\mathcal{F}^{\infty}_{\text{CCO}} = \mathcal{F}_{\text{CCO}} \cup \bigcup_{k=3}^{\infty} \{M_I^*(k), \overline{M_I^*(k)}, M_I^*(k)^t, \overline{M_I^*(k)}^t\},$$

where

$$\mathcal{F}_{\mathrm{CCO}} = \{Z_2^*, Z_3^*, Z_4^*, Z_5, \overline{Z_2^*}, \overline{Z_4^*}, (Z_2^*)^t, (Z_3^*)^t, (Z_4^*)^t, Z_5^t, (\overline{Z_2^*})^t, (\overline{Z_4^*})^t\}.$$

(See Figure 1.2)

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(a) Z_2^*

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

(b) Z_3^*

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

(c) Z_4^*

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(d) Z_5

$$\begin{pmatrix} 1 & 1 & & & & 0 \\ & 1 & 1 & & & 0 \\ & & \ddots & \ddots & & \vdots \\ & & & 1 & 1 & 0 \\ 1 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

(e) $M_I^*(k)$ for each $k \geq 3$, where omitted entries are 0's

Figure 1.2: Some matrices in $\mathcal{F}_{\mathrm{CCO}}^{\infty}$

The following result from [40] provides a minimal forbidden submatrix characterization of matrices possessing the circularly compatible ones property.

Theorem 1.4.14 ([40]). *For any matrix M, the following statements are equivalent:*

(i) M has the circularly compatible ones property;

(ii) M contains no matrix from $\mathcal{F}_{\mathrm{CCO}}^{\infty}$ as a configuration;

(iii) M has the circular-ones property for both rows and columns, and contains no matrix from $\mathcal{F}_{\mathrm{CCO}}$ as a configuration;

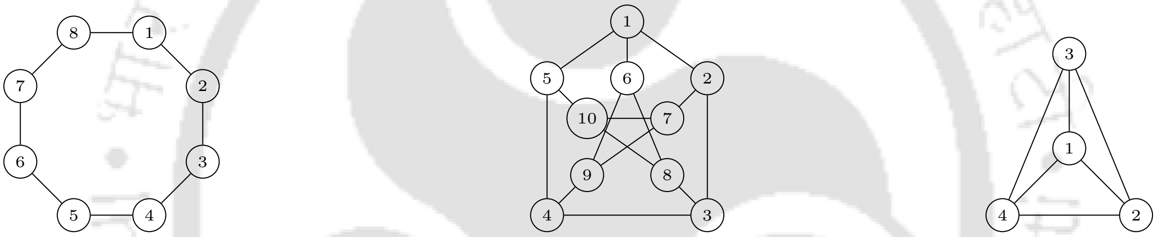
(iv) M has the doubly D-circular property.

Theorem 1.4.15 ([40]). *There exists a linear-time algorithm that, for any given matrix M, either produces a circularly compatible ones biorder of M or identifies a matrix in $\mathcal{F}_{\mathrm{CCO}}^{\infty}$ that occurs in M as a configuration.*

1.5 Word-Representable Graphs

Suppose that w is a word over some alphabet, and x and y are two distinct letters in w. We say that x and y *alternate* in w if, after deleting all letters except the copies of x and y in w, we either obtain a word $xyxy\dots$ (of odd or even length) or a word $yxyx\dots$ (of odd or even length). Hence, if w has a single occurrence of x and a single occurrence of y, then x and y alternate in w.

Definition 1.5.1. *A graph G is said to be word-representable if a word w can be formed using the letters of the alphabet $V(G)$ such that for every pair of vertices x and y, $xy \in E(G)$ if and only if x and y alternate in w. We say that w represents G, and w is called a word-representant of G. Also, w must contain each letter of $V(G)$ at least once.*



$w = 1821324354657687$ $w = 1387296(10)7493541283(10)7685(10)194562$ $w = 1234$

Figure 1.3: Three word-representable graphs Cycle C_8 (left), the Peterson graph (middle), and the complete graph K_4 (right)

Figure 1.3 illustrates examples of word-representable graphs along with their word-representants. For instance, each *complete graph* K_n can be represented by any permutation π of $\{1, 2, \dots, n\}$, or by π concatenated any number of times. Similarly, the *empty graph* E_n (also known as the *edgeless graph* or *null graph*) on the vertex set $\{1, 2, \dots, n\}$ can be represented by $1\ 2 \cdots (n-1)\ n\ n\ (n-1) \cdots 2\ 1$, or by any permutation concatenated with the same permutation written in reverse order.

Definition 1.5.2. *The reverse of the word $w = w_1 w_2 \dots w_n$ is the word $r(w) = w_n \dots w_2 w_1$.*

Proposition 1 ([32])**.** If w is a word-representant of a graph G, then $r(w)$ also represents the graph G.

For a word w, let $\pi(w)$ denote the permutation obtained from w by removing all but the leftmost occurrence of each letter x. We call $\pi(w)$ the *initial permutation* of w. Similarly, let

$\sigma(w)$ denote the permutation obtained from w by removing all but the rightmost occurrence of each letter x. We call $\sigma(w)$ the *final permutation* of w. Furthermore, the restriction of a word w to certain letters $x_1, \ldots, x_m$ is denoted by $w|_{\{x_1,\ldots,x_m\}}$. Let $A(w)$ denote the set of letters occurring in the word w. For example, if $w = 35423214$, then $\pi(w) = 35421$, $\sigma(w) = 53214$, $w|_{\{1,2\}} = 221$, and $A(w) = \{1, 2, 3, 4, 5\}$.

Definition 1.5.3. *Let w be a word. $O(w, i)$ is defined as the set of letters of w which occur exactly i times in it.*

Definition 1.5.4. *Let w be a word-representant for the graph $G = (V, E)$. Let $x \in V$. Define $O_w(x)$ as the number of occurrences of the letter x in the word w.*

Example 1.5.5. *In the word $w = 322414$, $O_w(3) = 1$ and $O_w(2) = 2$. $O(w, 2) = \{2, 4\}$ and $O(w, 1) = \{1, 3\}$.*

Proposition 2 ([29]). Let $w = w_1xw_2xw_3$ be a word-representant of a graph G such that w_1, w_2 and w_3 are possibly empty words, and w_2 contains no x. Then the possible neighbours of x in G are the letters in $O(w_2, 1)$.

Remark 1 ([30]). The class of word-representable graphs is *hereditary*. That is, removing a vertex v in a word-representable graph G results in a word-representable graph G'. Indeed, if w represents G then w with v removed represents G'.

1.5.1 k-Representability

A word w is *k-uniform* if each letter in w occurs k times. For example, 243321442311 is a 3-uniform word, while 23154 is a 1-uniform word (a permutation).

Definition 1.5.6. *A graph G is* k-word-representable, *or* k-representable *for brevity, if there exists a k-uniform word w representing it. We say that w* k-represents *G.*

Let us discuss some properties of k-representable graphs.

Proposition 3 ([32]). Let $w = uv$ be a k-uniform word representing a graph G, where u and v are two, possibly empty, words. Then, the word $w' = vu$ also represents G.

Remark 2. This is not true for non-uniform words. The word 121 represents K_2, but the word 112 represents an empty graph on 2 vertices.

The following result establishes the equivalence of Definitions 1.5.1 and 1.5.6.

Theorem 1.5.7 ([33]). *A graph is word-representable if and only if it is k-representable for some k.*

Observation 1.5.8 ([33]). *Let w be a word-representant of G. Then $\pi(w)w$ also represents G. In particular, for every $l > k$, a k-representable graph is also l-representable.*

Thus, from the above observation, it is clear that a graph can have infinitely many words representing it.

1.6 Permutationally Representable Graphs and Non-Word-Representable Graphs

Are there any non-representable graphs? In this section, we provide a positive answer to this question. An orientation of a graph is *transitive* if the presence of edges $u \to v$ and $v \to z$ implies the presence of the edge $u \to z$. An undirected graph is called a *comparability graph* if it admits a transitive orientation. It is well known [32], and easy to verify, that the smallest non-comparability graph is the cycle graph C_5.

Definition 1.6.1. *A graph $G = (V, E)$ is* permutationally representable *if it can be represented by a word of the form $p_1 \cdots p_k$ where p_i is a permutation. We say that G is* permutationally k-representable.

The following theorem is an easy corollary of the fact that any partially ordered set can be represented as an intersection of linear orders.

Theorem 1.6.2 ([37]). *A graph is permutationally representable if and only if it is a comparability graph.*

Theorem 1.6.3 ([33]). *A graph G obtained from a graph H by adding an apex is word-representable if and only if H is permutationally representable.*

A *wheel graph* W_n is the graph obtained from a cycle graph C_n by adding an apex vertex. It is easy to see that none of the cycle graphs C_{2n+1}, for $n \geq 2$, is a comparability graph, and thus none of the wheel graphs W_{2n+1}, for $n \geq 2$, is word-representable. In fact,

W_5 is the smallest example of a non-word-representable graph (and the only one on six vertices).

As a direct corollary to Theorem 1.6.3, we have the following important result revealing the structure of neighbourhoods of vertices in a word-representable graph.

Theorem 1.6.4 ([33]). *If a graph G is word-representable then the neighbourhood of each vertex in G is permutationally representable (is a comparability graph by Theorem 1.6.2).*

The converse to Theorem 1.6.4 is *not* true as demonstrated by the counterexamples in Figure 1.4 taken from [11] and [22], respectively.

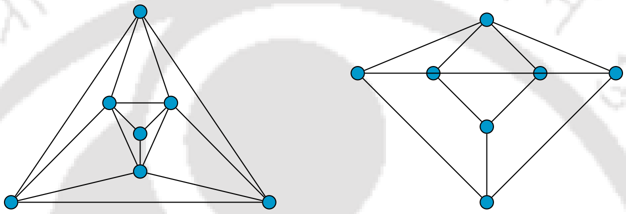


Figure 1.4: Non-word-representable graphs in which each neighbourhood is permutationally representable

1.7 Representation Number

From *Observation* 1.5.8, it is evident that if a graph is k-representable, then it is also l-representable for all $l > k$. Therefore, it remains to determine the minimum k for which a word-representable graph admits a k-uniform word-representant.

Definition 1.7.1. *The representation number of a graph is the least k such that the graph is k-representable. For graphs that are not word-representable (whose existence will be discussed below), we define $k = \infty$. We denote the representation number of a graph G by $\mathcal{R}(G)$, and let $\mathcal{R}_k = \{ G : \mathcal{R}(G) = k \}$.*

The classes of graphs with small representation numbers have been fully characterized and correspond to well-known graph families.

- $\mathcal{R}_k = 1$: A graph is 1-representable if and only if it is a *complete graph* (K_n). A 1-uniform word is a permutation of the vertices. In any permutation, any pair of distinct letters appears exactly once, so they trivially alternate. Thus, any permutation represents a complete graph.

- $\mathcal{R}_k = 2$: The class of graphs with representation number 2 is precisely the class of *non-complete circle graphs*. Circle graphs are the intersection graphs of chords in a circle, and their connection to 2-uniform words establishes an important link between word-representable graphs and a classical family of intersection graphs.

The following theorem gives a characterization for the class $\mathcal{R}_2$.

Theorem 1.7.2 ([30]). $\mathcal{R}_2 = \{G : G$ *is a non-complete circle graph*$\}$.

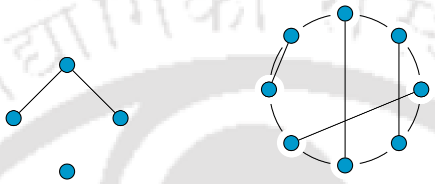


Figure 1.5: A circle graph on four vertices and its associated chords

1.7.1 Graphs with representation number 3

The study of graphs with representation number 3 and higher is an active area of research. For instance, it is known that the Petersen graph and all prism graphs have a representation number of 3 [29].

The following lemma and proposition give a useful tool to construct 3-word-representable graphs, that is, graphs with a representation number of at most 3.

Lemma 1.7.3 ([33]). *Let* $G = (V, E)$ *be a 3-word-representable graph and* $x, y \in V$. *Denote by* H *the graph obtained from* G *by adding a path of length at least* 3 *connecting* x *and* y. *Then* H *is also* 3*-word-representable.*

Proposition 4 ([29]). Let $G \in \mathcal{R}_k$, where $k \geq 2$, and $x \in V(G)$. Also, let G' be the graph obtained from G by adding an edge xy, where $y \notin V(G)$. Then $G' \in \mathcal{R}_k$.

1.8 Semi-Transitive Orientation

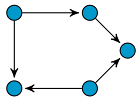
In Section 1.6, we saw a characterization in terms of orientability, which implies that alternation corresponds to a property of a digraph obtained by directing the edges in a certain way. It is known that a graph is a permutationally-representable graph if and only if it has a transitive orientation (i.e., it is a comparability graph). In [23, 24], Halldórsson et al. gave another generalization for transitive orientation known as *semi-transitive orientation.*

We now turn to the characterization and begin with definitions of certain directed graphs. A *shortcut* is an *acyclic, non-transitively oriented* graph obtained from a directed cycle on at least four vertices by reversing the orientation of one of its edges, and possibly by adding additional directed edges connecting some of the vertices (while keeping the graph acyclic and non-transitive). Thus, any shortcut

- is *acyclic* (that is, it contains *no directed cycles*);
- has *at least four* vertices;
- has *exactly one* source (a vertex with no incoming edges) and *exactly one* sink (a vertex with no outgoing edges), with a *directed path* from the source to the sink that passes through *every* vertex of the graph;
- contains an edge connecting the source to the sink, referred to as the *shortcutting edge*;
- is *not transitive* (that is, there exist vertices u, v, and z such that $u \to v$ and $v \to z$ are edges, but there is *no* edge $u \to z$).

Definition 1.8.1. *An orientation of a graph is* semi-transitive *if it is* acyclic *and* shortcut-free.

It is easy to see from definitions that *any* transitive orientation is necessarily semi-transitive. The converse is *not* true, e.g. the following semi-transitively oriented graph is *not* transitively oriented:



Thus, semi-transitive orientations generalize transitive orientations.

A way to check if a given oriented graph G is semi-transitively oriented is as follows. First, check that if G is acyclic; if not, the orientation is not semi-transitive. Next, for a directed edge from a vertex x to a vertex y, consider *each* directed path P having at least three edges without repeated vertices from x to y, and check that the subgraph of G induced by P is transitive. If such a non-transitive subgraph is found, the orientation is not semi-transitive. This procedure needs to be applied to each edge in G, and if no non-transitivity is discovered, G's orientation is semi-transitive.

As we will see in Theorem 1.8.2, finding a semi-transitive orientation is equivalent to recognising whether a given graph is word-representable, and this is an NP-hard problem. Thus, there is no efficient way to construct a semi-transitive orientation in general, and such a construction would rely on an exhaustive search orienting edges one by one, and thus branching the process. Having said that, there are several situations in which branching is not required.

The main characterization theorem to date for word-representable graphs is the following result.

Theorem 1.8.2 ([24]). *A graph G is word-representable if and only if it is semi-transitively orientable. Moreover, every non-complete word-representable graph is $2(n-\kappa)$-word-representable, where κ denotes the size of a maximum clique in G.*

Since each complete graph is 1-word-representable and each edgeless graph is 2-word-representable, we have the following statement.

Corollary 1 ([24]). Each word-representable graph G on $n \geq 3$ vertices is $2(n-2)$-word-representable.

1.9 Graph Operations & Word-Representability

In this section, we consider some of the most basic operations on graphs, namely, connecting two graphs by an edge and gluing two graphs in a clique, taking a line graph, taking the complement, edge contraction, edge addition, edge deletion, and various graph products.

1.9.1 Connecting two graphs by an edge and gluing two graphs in a clique

The operations of connecting two graphs, G_1 and G_2, by an edge and gluing two graphs at a vertex are presented schematically in Figure 1.6. It follows directly from Theorem 1.8.2 that if both G_1 and G_2 are word-representable, then the resulting graphs will be word-representable too, while if at least one of G_1 or G_2 is non-word- representable then the resulting graphs will be non-word-representable. Indeed, if G_1 and G_2 are oriented semi-transitively, then orienting the edge xy in either direction will give no chance for the resulting graph to have a shortcut, thus resulting in a semi-transitively oriented graph; similarly, no shortcut is possible when semi-transitively oriented G_1 and G_2 are glued at a vertex z.

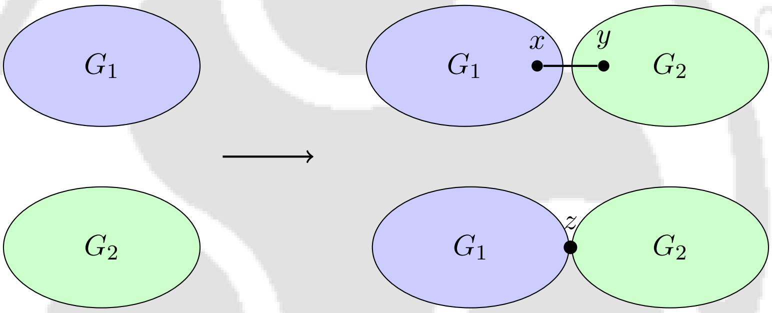


Figure 1.6: Connecting graphs by an edge and gluing graphs at a vertex

Moreover, the following theorem provides an answer to the question: if G_1 is k_1-word-representable, G_2 is k_2-word-representable, and G is k-word-representable then what can be said about k?

Theorem 1.9.1 ([32]). *For $k \geq 2$, let w_1 and w_2 be k-uniform words representing graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, respectively. where V_1 and V_2 are disjoint. Suppose that $x \in V_1$ and $y \in V_2$. Let H_1 be the graph formed by connecting graphs G_1 and G_2 with an edge. Let H_2 be the graph formed from graphs G_1 and G_2 by identifying vertex x and vertex y into a new vertex z. Then H_1 and H_2 are k-word-representable.*

Theorem 1.9.2 ([32]). *Suppose that for graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, $\mathcal{R}(G_1) = k_1$ and $\mathcal{R}(G_2) = k_2$, $x \in V_1$, $y \in V_2$ and $k = \max(k_1, k_2)$. Also, let the graph G' be obtained by connecting G_1 and G_2 by the edge xy, and the graph G'' be obtained from G_1 and G_2 by identifying the vertices x and y into a single vertex z. The following holds.*

1. *If* $|V_1| = |V_2| = 1$ *then both* G' *and* G'' *are cliques and thus* $k_1 = k_2 = 1$*. In this case,* $\mathcal{R}(G') = \mathcal{R}(G'') = 1$.
2. *If* $\min(|V_1|, |V_2|) = 1$ *but* $\max(|V_1|, |V_2|) > 1$ *then* $\mathcal{R}(G') = k$ *and* $\mathcal{R}(G'') = \max(k, 2)$.
3. *If* $\min(|V_1|, |V_2|) > 1$ *then* $\mathcal{R}(G') = \mathcal{R}(G'') = \max(k, 2)$.

1.9.2 Various graph products

There are graph products that preserve word-representability, and there are graph products that do not preserve word-representability. Table 1.1 provides the definition of various graph products, and Table 1.2 provides the details about the products and the word-representability.

Graph product	**Denoted by**	**Edge condition for** $(u_1, u_2) \sim (v_1, v_2)$
Cartesian Product	$G \,\square\, H$	$(u_1 = v_1$ and $u_2 \sim v_2)$ or $(u_1 \sim v_1$ and $u_2 = v_2)$
Rooted Product	$G \circ H$	$(u_1 \sim v_1$ and $u_2 = v_2 = r)$ or $(u_1 = v_1$ and $u_2 \sim v_2)$ where r is the root vertex in H
Tensor Product	$G \times H$	$u_1 \sim v_1$ and $u_2 \sim v_2$
Lexicographical Product	$G \bullet H$	$u_1 \sim v_1$ or $(u_1 = v_1$ and $u_2 \sim v_2)$
Strong Product	$G \boxtimes H$	$(u_1 = v_1$ and $u_2 \sim v_2)$ or $(u_1 \sim v_1$ and $u_2 = v_2)$ or $(u_1 \sim v_1$ and $u_2 \sim v_2)$

Table 1.1: Definitions of various products of two graphs G and H.

Graph product	**Preserves Word-representability**	**Reference**
Cartesian Product	Yes	[37]
Rooted Product	Yes	[37]
Tensor Product	Not necessarily	[10]
Lexicographical Product	Not necessarily	[10]
Strong Product	Not necessarily	[10]

Table 1.2: Word-representability & various products of two word-representable graphs G and H

In [5], Broere has given the k-representability of the Cartesian product of two word-representable graphs. The following results give the information on the upper bound of the representation number of Cartesian products of a word-representable graph with K_2, K_n and a permutationally-representable graph, respectively.

Theorem 1.9.3 ([5]). *Let G be a k-representable graph for $k > 1$. Then the graph $G\Box K_n$ is $(k+(n-1))$-representable*

Theorem 1.9.4 ([5]). *Let $G = (V_G, E_G)$ be a k-representable graph for $k > 1$. Also, let $H = (V_H, E_H)$ be l-permutationally representable graph. Then the graph $G\Box H$ is $l(k+(n-1))$-representable.*

1.10 Minimum Length Word-Representants

In [24], Halldórsson, Kitaev, and Pyatkin study the minimal k for which certain graphs are k-word-representable. In [18], Gaetz and Ji have investigated the absolute minimal length word-representants of trees and cycles, where we do not require our word-representants to be k-uniform for any k. In this section, we present the results regarding the lengths of these absolute minimum length word-representants.

Definition 1.10.1. *Let G be a word-representable graph, where w is a word-representant of G. Then, $l(G)$ is defined as the minimum possible length of the word w. Let the word w be called as a minimum length word-representant of graph G.*

Theorem 1.10.2 ([18]). *Let $G = (V, E)$ be a word-representable graph with connected components $\{G_i = (V_i, E_i)\}_{i=1}^{k}$. Then*

$$l(G) \leq \sum_{i=1}^{k}(l(G_i) + |V_i|) - \max_{1\leq j\leq k} |V_j|.$$

In the case of trees and cycles, a precise value for $\ell(G)$ is given. First, we present a simple bound for all triangle-free graphs.

Lemma 1.10.3 ([18]). *Let G be a triangle-free graph with n vertices. Then $l(G) \geq 2n-2$.*

Theorem 1.10.4 ([18]). *Let $T = (V, E)$ be a tree with $n \geq 2$ vertices. Then $l(T) = 2n-2$.*

Theorem 1.10.5 ([18]). *Let $C_n = (V, E)$ be the cycle graph on $n \geq 4$ vertices. Then, $l(C_n) = 2n-2$.*

1.11 Characterization of Word-Representable Split Graphs

It is clear from Table 2 that, so far, only split graphs have been characterized in terms of word-representability. In this subsection, we discuss the characterization of split graphs that are word-representable. Moreover, the recognition problem for word-representable split graphs can be solved in polynomial time.

A graph G is called a *split graph* if its vertex set $V(G)$ can be partitioned into two parts $I \cup C$ such that C induces a clique, and I induces an independent set.

Theorem 1.11.1 ([35]). *A split graph G is semi-transitive if and only if the vertices of C can be labeled from 1 to $k = |C|$ in such a way that for each $v, u \in I$:*

- *either $N_G(v) = [a, b]$ for $a \leq b$ or $N_G(v) = [1, a] \cup [b, k]$ for $a < b$.*
- *If $N_G(u) = [a_1, b_1]$ and $N_G(v) = [1, a_2] \cup [b_2, k]$, for $a_1 \leq b_1$, $a_2 < b_2$, then $a_1 > a_2$ or $b_1 < b_2$.*
- *If $N_G(u) = [1, a_1] \cup [b_1, k]$ and $N_G(v) = [1, a_2] \cup [b_2, k]$, for $a_1 < b_1$ and $a_2 < b_2$, then $a_2 < b_1$ and $a_1 < b_2$.*

Determining the semi-transitivity of split graphs is closely related to the well-known circular ones property of $(0, 1)$-matrices. A $(0, 1)$-matrix has the *consecutive ones* property (for columns) if, after some permutation of its rows in all columns, the ones are consecutive. A $(0, 1)$-matrix has the *circular ones* property (for columns) if, after some permutation of its rows in all columns, either ones or zeroes are consecutive. Note that if zeroes are consecutive, then ones are "almost consecutive" in the sense that they are allowed to wrap around from the bottom of a column to its top.

Given a split graph G with $C = \{u_1, \dots, u_k\}$ and $I = \{v_1, \dots, v_t\}$, consider a $(0, 1)$-matrix $M(G)$ with k rows and t columns where $m_{ij} = 1$ if and only if u_i is adjacent to v_j. Then, clearly, any labelling of the vertices in C defines a permutation of the rows of the matrix $M(G)$. Moreover, such a labelling satisfies condition (1) of Theorem 1.11.1 if and only if the corresponding permutation provides a circular ones property.

Theorem 1.11.2 ([35]). *A split graph G is semi-transitive if and only if the rows of matrix $M(G)$ can be permuted in such a way that:*

(i) *The ordering has a circular ones property for all columns.*

(ii) *If a column has the form* $1^a0^b1^c$ *where* $a+b+c=k$ *and* $a,b,c \geq 1$*, then no other column may contain ones in all positions from* a *to* $a+b+1$*.*

It is easy to verify that property (ii) of Theorem 1.11.2 is equivalent to conditions (2)–(3) of Theorem 1.11.1.

Using the above-mentioned results, the following result can be stated.

Theorem 1.11.3 ([35]). *The recognition problem for word-representable split graphs can be solved in time* $\mathcal{O}(t^2k)$*.*

1.12 Characterization of Word-Representable Co-Bipartite Graphs

Similar to split graphs, the class of word-representable co-bipartite graphs has also been characterized. This characterization was obtained by Das and Ramesh [14]. Moreover, they proved that the representation number of any word-representable co-bipartite graph is at most three [13].

A graph G is called a *co-bipartite graph* if its vertex set $V(G)$ can be partitioned into two cliques, say $V(K_m)$ and $V(K_n)$.

Theorem 1.12.1 ([13]). *A co-bipartite graph* G *is semi-transitive if and only if the vertices of* $V(K_n)$ *can be labeled by the integers* 1 *to* n *in such a way that, for every* $u, v \in V(K_m)$*, the following conditions hold, where* $N(u) = N_G(u) \cap V(K_n)$*.*

- $N(v) = [a,b]$ *for some* $a \leq b$*, or* $N(v) = [1,a] \cup [b,n]$ *for some* $a < b$*, or* $N(v) = \varnothing$*.*
- *If* $N(u) = [a_1,b_1]$ *and* $N(v) = [1,a_2] \cup [b_2,n]$*, where* $a_1 \leq b_1$ *and* $a_2 < b_2$*, then* $u < v$*.*
- *If* $N(u) = [a_1,b_1]$ *and* $N(v) = [a_2,b_2]$*, where* $u < v$*,* $a_1 \leq b_1$*, and* $a_2 \leq b_2$*, then* $a_1 \leq a_2$ *and* $b_1 \leq b_2$*.*
- *If* $N(u) = [a_1,b_1]$ *and* $N(v) = [1,a_2] \cup [b_2,n]$*, where* $a_1 \leq b_1$ *and* $a_2 < b_2$*, then* $a_1 \leq b_2$ *and* $a_2 \leq b_1$*.*

- *If* $N(u) = [1, a_1] \cup [b_1, n]$ *and* $N(v) = [1, a_2] \cup [b_2, n]$*, where* $u < v$*,* $a_1 < b_1$*, and* $a_2 < b_2$*, then* $a_1 \leq a_2$ *and* $b_1 \leq b_2$*.*
- *Suppose* $N(u) = \varnothing$*. Then, for all* $v_1 < u < v_2$*, we have* $N(v_1) = [a_1, b_1]$ *with* $a_1 \leq b_1$ *and* $N(v_2) = [1, a_2] \cup [b_2, n]$ *with* $a_2 < b_2$*. Moreover,* $b_1 < b_2$*.*

2

Minimum length word-representants of word-representable graphs and graph products

2.1 Introduction

In this chapter, we establish a relationship between the maximum and minimum number of occurrences of letters in minimum length word-representants and the diameter of the corresponding graph. Furthermore, we derive an upper bound on the minimum possible length of a word-representant under certain word-representability-preserving graph operations, such as connecting two graphs by an edge and gluing two graphs at a vertex. We also determine the minimum length of word-representants for the Cartesian and rooted products of word-representable graphs, employing morphism-based and occurrence-based functions, respectively. Finally, we address an open problem posed by Broere in [5], which concerns the construction of a word representing the Cartesian product of two arbitrary word-representable graphs.

In Section 2.2, we establish a lower bound on the minimum possible length of word-representants for certain classes of graphs, such as bipartite and ladder graphs. We also provide a partial characterization of 2-word-representable graphs based on the minimum length of their word-representants and demonstrate a relationship between the maximum and minimum number of occurrences of letters in a minimum-length word-representant and the diameter of the graph.

In Section 2.3, we derive an upper bound on the minimum possible length of a word-representant under specific graph operations that preserve word-representability. Furthermore, in Section 2.3.2, we establish an upper bound for the minimum length of words representing the Cartesian product of word-representable graphs in terms of their individual minimum lengths. Similarly, in Section 2.3.3, we present an upper bound for the minimum length of words representing the rooted product of word-representable graphs in terms of their respective minimum lengths.

2.2 Minimum Length Words

In this section, we investigate the relationship between the minimum possible length of a word-representant of a graph and its structural properties. We establish several lower and upper bounds on this length for various graph classes, including complete bipartite, ladder, multipartite, and K_p-free graphs. We also explore the role of the representation number in determining the minimum length and discuss the implications for open problems posed by Gaetz and Ji [18].

Lemma 2.2.1. *Let w be a minimum length word-representant of a word-representable graph G. Then*

$$|O(w,1)| \leq \kappa_G$$

where κ_G is the size of a maximal clique of G.

Proof. Suppose there are more than κ_G letters in w that appear exactly once. This implies the existence of a clique of size $|O(w,1)|$, which contradicts the fact that the size of a maximal clique in G is κ_G. □

Theorem 2.2.2. *Let $G = (X \cup Y, E)$ be a complete bipartite graph on n vertices with $|X| = n_1$ and $|Y| = n_2$. Then $\ell(G) = 2n - 2$.*

Proof. Since G is bipartite, we can say that $\ell(G) \geq 2n-2$ by *Lemma* 1.10.3. Therefore, our objective is to show that $\ell(G) \leq 2n-2$. It is enough to show that a word of length $2n-2$ can represent G. Let X be the set $\{1, 2, \ldots, n_1\}$ and Y be the set $\{n_1+1, n_1+2, \ldots, n = n_1+n_2\}$. Our claim is that the word $w = 234 \ldots n_1(n_1+1)(n_1+2) \ldots nn_1(n_1-1) \ldots 1n(n-1) \ldots (n_1+2)$ represents G.

If $2 \leq x, y \leq n_1$ or $n_1 + 2 \leq x, y \leq n$, then $w|_{\{x,y\}} = xyyx$ or $yxxy$. Therefore, there are no edges between any two vertices of X or vertices of Y. If $2 \leq x \leq n_1$ and $n_1 + 2 \leq y \leq n$, then $w|_{\{x,y\}} = xyxy$. Therefore, $xy \in E$ for all $x \in X$ and $y \in Y$. If $x = 1$ and $n_1 + 2 \leq y \leq n$, then $w|_{\{1,y\}} = y1y$. Hence, $1y \in E$ for all $y \in Y$. Similar arguments show that $x(n_1+1) \in E$ for all $x \in X$. If $x = 1$ and $y = n_1+1$, then $w|_{\{1,n_1+1\}} = (n_1+1)1$. Hence, $1(n_1 + 1) \in E$. As a result, w represents G. □

Theorem 2.2.3 ([29]). *Let L_n be the ladder graph on $2n \geq 4$ vertices. Then for all $n \geq 2$, $\mathcal{R}(L_n) = 2$.*

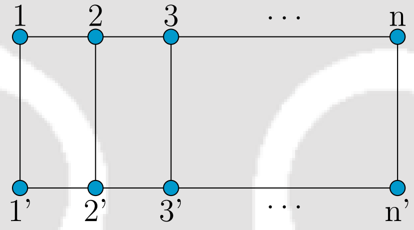


Figure 2.1: Ladder graph L_n

Theorem 2.2.4. *Let L_n be the ladder graph on $2n \geq 4$ vertices. Then $\ell(L_n) = 4n - 2$.*

Proof. Let L_n be the ladder graph shown in *Figure* 2.1 with the vertex set $V(L_n) = \{1, 2, \ldots, n, 1', 2', \ldots, n'\}$. Let $u = 123 \ldots (n-1)n$ be a word over the alphabet $A = \{1, 2, \ldots, n-1, n\}$. Let us define the morphisms $h : A^* \to V^*$ and $g : A^* \to V^*$ as follows.

$$h(i) = \begin{cases} (i+1)'i' & \text{if } i < n \text{ and } i \text{ is odd,} \\ (i+1)i & \text{if } i < n \text{ and } i \text{ is even,} \\ i' & \text{if } i = n \text{ and } i \text{ is odd,} \\ i & \text{if } i = n \text{ and } i \text{ is even.} \end{cases}$$

and

$$g(i) = \begin{cases} i(i+1) & \text{if } i < n \text{ and } i \text{ is odd,} \\ i'(i+1)' & \text{if } i < n \text{ and } i \text{ is even,} \\ i & \text{if } i = n \text{ and } i \text{ is odd,} \\ i' & \text{if } i = n \text{ and } i \text{ is even.} \end{cases}$$

We claim that the word $w = h(u)g(r(u))$ represents the ladder graph L_n. Recall that $r(u)$ denotes the reversal of the word u. According to the definition of a ladder graph, vertex i is adjacent to the vertices $\{i+1, i-1, i'\}$ for all $i = \{2, 3, \dots n-1\}$. Vertex 1 is adjacent to $\{2, 1'\}$, and vertex n is adjacent to $\{n-1, n'\}$. This is also the case for vertices $\{1', 2', \dots, n'\}$. From *Theorem* 2.2.3 and *Proposition* 2, only the adjacent vertices of vertex i occur exactly once between two occurrences of i. Therefore, it is sufficient to show that only the vertices $\{i+1, i-1, i'\}$ $(\{(i+1)', (i-1)', i\})$ occur once between two instances of the vertex $i(i')$.

For each vertex i (i'), all vertices $\{j, (j-1)'\}$ $(\{j', (j-1)\})$, where $j > i+1$, appear to the right of i (i') in $h(u)$ and to the left of i (i') in $g(r(u))$. Similarly, for each vertex i (i'), all vertices $\{j, (j+1)'\}$ $(\{j', (j+1)\})$, where $j < i-1$, appear to the left of i (i') in $h(u)$ and appear to the right of i (i') in $g(r(u))$. Therefore, all vertices apart from $\{i+1, i-1, i'\}$ $(\{(i+1)', (i-1)', i\})$ either appear twice or do not appear between the two instances of i (i'). The vertices $\{i+1, i-1, i'\}$ $(\{(i+1)', (i-1)', i\})$ appear to the right of i (i') in both $h(u)$ and $g(r(u))$. Therefore, they occur exactly once between the two instances of i (i'). Therefore, w represents the graph L_n.

Since $|h(u)| = 2n-1$ and $|g(r(u))| = 2n-1$, we have $\ell(L_n) \leq 4n-2$. As L_n is triangle-free, we can say that $\ell(L_n) = 4n-2$ by *Lemma* 1.10.3. □

Lemma 1.10.3 provides a result for K_3-free graphs. This raises the question of whether the result can be extended to all K_p-free graphs. The following lemma provides a positive answer to this question.

Theorem 2.2.5. *Let $G = (V, E)$ be a K_p-free graph on n vertices, where $p \leq n$. Then $\ell(G) \geq 2n - p + 1$.*

Proof. Let w be a minimum length word-representant of G. By *Definition* 1.5.1, all elements of V occur at least once in w. We claim that there are at most $p-1$ letters from V

that occur only once in w. Suppose that the vertices $x_1, x_2, \ldots, x_p \in V$ appear only once in w. Then the vertices $\{x_1, x_2, \ldots, x_p\}$ induce a K_p in G, which is a contradiction. □

Lemma 2.2.6. *Every multi-partite graph, $H_{n,p}$, on n vertices and p partitions is K_{p+1}-free.*

Proof. Assume that $H_{n,p}$ contains a K_{p+1} subgraph, which implies that there are $p+1$ vertices that are adjacent to every other vertex. These $p+1$ vertices must be chosen from p partitions. By applying the pigeonhole principle, we can deduce that there must be at least two vertices that are selected from the same partition. However, they cannot be adjacent as they are from the same partition, which is a contradiction. □

Theorem 2.2.7. *Let $H_{n,p}$ be a multi-partite graph on n vertices and p partitions. Then $\ell(H_{n,p}) \geq 2n - p$.*

Proof. Proof follows directly from *Theorem* 2.2.5 and *Lemma* 2.2.6. □

Theorem 2.2.8. *Let $G = (V, E)$ be a k-word representable graph with $k \geq 2$. Then $\ell(G) \leq kn - 1$.*

Proof. Since G is k-word representable, we can say that $\ell(G) \leq kn$. Let w be a word-representant of G. Then, by *Proposition* 3, for a vertex $x \in V(G)$, the word w can be considered as $w = x^1 A_1 x^2 A_2 \ldots x^k A_k$, where x^i denotes the i^{th} occurrence of the letter x. We claim that $w' = A_1 x^2 A_2 \ldots x^k A_k$ also represents G.

If $xy \in E$, then x and y alternate in w. This means that $y \in A_i$ for all $i = 1, 2, \ldots k-1$. Therefore, it is clear that x and y alternate in w'. Similarly if $xy \notin E$, then there is an A_i in w where $y \notin A_i$. If $i \neq 1$, then it is clear that x and y do not alternate in w'. If $y \notin A_1$, then according to the pigeonhole principle, since G is k-word representable, there must be some A_j that has at least two y. Therefore, x and y do not alternate in w'. This shows that w' represents G. □

Theorem 2.2.9. *Let $G = (V, E)$ be a 2-word-representable graph on n vertices that contains at least one edge. Then $\ell(G) \leq 2n - 2$.*

Proof. Since $\mathcal{R}(G) = 2$, we have $\ell(G) \leq 2n - 1$ by *Theorem* 2.2.8. Therefore, it suffices to show that there exists a word-representant of G of length $2n - 2$. Let w' be a 2-uniform word-representant of the graph G, and let x be a letter such that no other letter occurs twice between x^1 and x^2, but at least one letter occurs between them. Such a word always exists; otherwise, the graph would be empty, contradicting our assumption that G contains at least one edge and is 2-word-representable.

Hence, all the letters between the two occurrences of x are its neighbours. Let y be a letter adjacent to x. Then, by *Proposition* 3, the word w' can be written as $w' = BAxy$. We claim that the word $w = BA$ also represents the graph G. To prove this, we will show that two vertices i and j alternate in w if and only if they alternate in w'.

Suppose $i, j \in V \setminus \{x, y\}$. Then it is clear that, $w'|_{\{i,j\}} = (BA)|_{\{i,j\}} = w|_{\{i,j\}}$. Therefore, in this case, i and j alternate in w if and only if they alternate in w'.

Without loss of generality, let us assume that $i = x$ and $j \in V \setminus \{x, y\}$. Then, we have $w|_{\{x,j\}} = (BA)|_{\{x,j\}}$. If x and j do not alternate in w, then they clearly do not alternate in w'. Suppose x and j do not alternate in w'. Then the factor BA contains xjj or jjx as a subword. Therefore, x and j do not alternate in w either. As a result, x and j alternate in w if and only if they alternate in w'. A similar argument shows that y and $j \in V \setminus \{x, y\}$ alternate in w if and only if they alternate in w'.

Suppose $i = x$ and $j = y$. They alternate in w as they alternate in w'. Thus w represents the graph G. Since $|w| = 2n - 2$, $\ell(G) \leq |w| = 2n - 2$. ☐

Theorem 2.2.10. *If G is a K_3-free 2-word-representable graph on n vertices, then* $\ell(G) = 2n - 2$.

Proof. Assume that G is a triangle-free graph on n vertices. Since $\mathcal{R}(G) = 2$, we have $\ell(G) = 2n - 2$ by *Theorem* 2.2.5 and *Theorem* 2.2.9. ☐

From *Theorem* 2.2.10, we can ask whether $\ell(G) = 2n - p + 1$ if G is a K_p-free graph on n vertices. However, the following example disproves this statement.

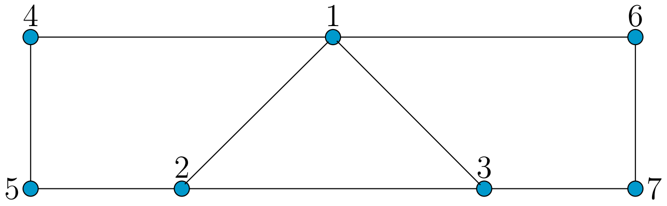


Figure 2.2: Minimum length word-representant of the graph, $w = 415425736716$

Example 2.2.11. *Consider the graph G shown in Figure 2.2, which is K_4-free. We expect that $\ell(G) = 2n-3 = 11$. However, consider the graph $G \setminus 6$, and let u be a minimum-length word-representant of $G \setminus 6$ with $\ell(G \setminus 6) = 9$. It can be verified that the two occurrences of the vertex 6 cannot be inserted into u while preserving the existing alternations and non-alternations. Hence, we require one additional occurrence of at least one of the vertices in $\{1, 2, 3\}$. Therefore, the word $w = 415425736716$ is a minimum-length word-representant of G, with $\ell(G) = 12 = 2n - 2$.*

Theorem 2.2.12. *Let G be a 2-word-representable graph on n vertices. Then $\ell(G) = 2n-1$ if and only if G is a K_2-free graph, i.e., an empty graph.*

Proof. Suppose that G is K_2-free. Then by *Theorem* 2.2.5 and *Theorem* 2.2.8, $\ell(G) = 2n - 1$.

Now assume that G is a graph with $\ell(G) = 2n - 1$. For the sake of contradiction, assume that G is not an empty graph. Then by *Theorem* 2.2.9, $\ell(G) \leq 2n - 2$, which contradicts the assumption. Therefore, G is a K_2-free graph. ☐

In [18], Gaetz and Ji have posed an open problem (*Problem 2*) that asks to classify the 2-word-representable graphs based on the minimum possible length of their word-representants. The following result completely answers two of the three cases and partially answers the remaining case of the problem mentioned above.

Theorem 2.2.13. *Let G be a 2-word-representable graph on n vertices. Then*

1. $\{G : \ell(G) = 2n\} = \emptyset$.
2. $\{G : \ell(G) = 2n - 1\} = \{$*empty graph on n vertices*$\}$.
3. $\{G : \ell(G) = 2n - 2\} \supset \{$*All K_3-free 2-representable graphs on n vertices*$\}$.

Proof. By *Theorem* 2.2.8, it is clear that the set $\{G\colon \ell(G) = 2n\}$ is empty. The second and third one directly follow from *Theorems* 2.2.12 and 2.2.10, respectively. ☐

The following result positively answers an open problem (*Problem 1*) posed by Gaetz and Ji in [18].

Theorem 2.2.14. *Let* $G = (V, E)$ *be a word-representable graph with connected components* $\{G_i = (V_i, E_i)\}_{i=1}^k$*. Let* w_i *be a minimum length word-representant of the component* G_i*. Assume that the components are ordered in such a way that* $|O(w_i, 1)| \geq |O(w_j, 1)|$ *for all* $i > j$*. Then,*

$$\ell(G) \leq \sum_{i=1}^{k} \ell(G_i) + \sum_{i=1}^{k-1} \kappa_i$$

where κ_i *is the size of a maximal clique of the graph* G_i

Proof. Here we will employ the proof technique used in *Theorem* 1.10.2. Let w_i be a minimum length word-representant of the component G_i , where $i \in \{1, 2, \ldots, k\}$. Without loss of generality, we assume that G_k is the graph component having the largest size of a maximal clique among all components. We claim that the word

$$w = w_1\ \sigma(w_1|_{O(w_1,1)})\ w_2\ \sigma(w_2|_{O(w_2,1)}) \ldots w_{k-1}\ \sigma(w_{k-1}|_{O(w_{k-1},1)})\ w_k$$

represents G.

Note that the argument below shows that $w_i\sigma(w_i|_{O(w_i,1)})$ represents G_i. If $x, y \notin O(w_i, 1)$, then x and y alternate in $w_i\sigma(w_i|_{O(w_i,1)})$ if and only if $xy \in E_i$. If $x, y \in O(w_i, 1)$, then we can make x and y alternate in $w_i\sigma(w_i|_{O(w_i,1)})$ by concatenating $\sigma(w_i|_{O(w_i,1)})$ with w_i. Without loss of generality, let us consider $x \in O(w_i, 1)$ and $y \notin O(w_i, 1)$. If $xy \in E_i$, then $(w_i\sigma(w_i|_{O(w_i,1)}))|_{\{x,y\}} = yxyx$ since $w_i|_{\{x,y\}} = yxy$. If $xy \notin E_i$, then $(w_i\sigma(w_i|_{O(w_i,1)}))|_{\{x,y\}} = yyxx$ or $xyyx$ since $w_i|_{\{x,y\}} = yyx$ or xyy. Hence, x and y alternate in $w_i\sigma(w_i|_{O(w_i,1)})$ if and only if they alternate in w_i.

Furthermore, for all pairs (i, j), where $1 \leq i < j \leq k$, each vertex of V_i occurs at least twice in w before any vertex of V_j appears. Therefore, there is no edge between two vertices of two distinct components if and only if the two vertices do not alternate in w. As a result, w represents the graph G. Therefore, $\ell(G) \leq |w| = \sum_{i=1}^k \ell(G_i) + \sum_{i=1}^{k-1} |O(w_i, 1)|$. Hence, we get $\ell(G) \leq \sum_{i=1}^k \ell(G_i) + \sum_{i=1}^{k-1} \kappa_i$ by *Lemma* 2.2.1. □

Remark 3. The theorem mentioned above gives a tighter upper bound on $\ell(G)$ than the one provided by *Theorem* 1.10.2, as the size of a maximal clique cannot exceed the number of vertices of the graph.

Definition 2.2.15. *Let* w *be a word. The minimum and maximum number of occurrences of a letter in* w *are denoted by* $O_{min}(w)$ *and* $O_{max}(w)$*, respectively.*

Lemma 2.2.16. *Let w be a minimum length word-representant of a graph G. Let $x \in V(G)$ be a vertex in w such that $O_w(x) = O_{max}(w)$. Define L_i as the set of vertices which are at distance i from vertex x. Then, for every vertex $y \in L_i$,*

$$O_w(y) \geq O_{max}(w) - i.$$

Proof. We prove this result using induction on i. It is evident that the occurrence of the only vertex in L_0, meaning $O_w(x) = O_{max}(w) \geq O_{max}(w) - 0$, which proves the base case. Now let us assume that the occurrence of each vertex in L_{k-1} is at least $O_{max}(w) - (k-1)$. We can now consider a vertex $z \in L_k$. Suppose that $O_w(z) < O_{max}(w) - k$. By the definition of the above set, there is at least one vertex $a \in L_{k-1}$ such that $d(a, z) = 1$, i.e. a and z are adjacent. Since a and z are adjacent, they should alternate in w. By induction hypothesis, $O_w(a) \geq O_{max}(w) - k + 1$. Since $O_w(z) \leq O_{max}(w) - k - 1$, there can be a maximum of $O_{max}(w) - k$ gaps between z where a can be placed to make them alternate. However, considering there are at most $O_{max}(w) - k$ gaps and at least $O_{max}(w) - k + 1$ a's to be placed, by applying the pigeonhole principle, there exists a factor in w where at least two a's occur between two z's, which contradicts the fact that a and z alternate in w. Therefore, $O_w(z) \geq O_{max}(w) - k$. As a result, $O_w(y) \geq O_{max}(w) - i$ for all $y \in L_i$. □

Theorem 2.2.17. *Let w be a minimum length word-representant of a word-representable connected graph G. Then*

$$O_{max}(w) - O_{min}(w) \leq diam(G)$$

where $diam(G)$ is the diameter of the graph G.

Proof. Let $x \in V(G)$ be a vertex with maximum number, $O_{max}(w)$, of occurrences in w. Let L_i denote the set of vertices which are at distance i from the vertex x. Let $x' \in V(G)$ be a vertex with minimum number, $O_{min}(w)$, of occurrences in w. Since the graph is connected, the vertex x' will be present in the set L_i for some i. Hence, $O_{min}(w) \geq O_{max}(w) - i$, by *Lemma* 2.2.16. Since $i \leq diam(G)$, we have

$$O_{max}(w) - O_{min}(w) \leq diam(G).$$

□

The below example shows that the above-established bound is tight.

Example 2.2.18. *Consider Pr_3, the 3-Prism shown in Figure 2.4. $w = 52634561425$ is a minimum length word-representant of the graph Pr_3. Therefore, $O_{max}(w) = 3$ and $O_{min}(w) = 1$. It follows that $O_{max}(w) - O_{min}(w) = 2$. We know that $diam(Pr_3) = 2$.*

A crown graph, $H_{n,n}$, is formed by removing a perfect matching from the complete bipartite graph, $K_{n,n}$. It is easy to see that $diam(H_{n,n}) = 3$.

Theorem 2.2.19 ([20]). *If $n \geq 5$, then the representation number of $H_{n,n}$ is $\lceil \frac{n}{2} \rceil$.*

Theorem 2.2.20. *The minimum possible length of a word-representant of the crown graph, $H_{n,n}$, is bounded below by*

$$\ell(H_{n,n}) \geq 2n\lceil \frac{n}{2} \rceil - 3n.$$

Proof. Let w be a minimum length word-representant of the graph $H_{n,n}$. According to *Theorem* 2.2.19, $O_{max}(w) = \lceil \frac{n}{2} \rceil$ for all $n \geq 5$. Let X and Y be the bipartition of the vertex set. Without loss of generality, assume that the vertex $x \in X$ occurs maximum number of times in w. Let L_i denote the set of vertices that are at a distance i from the vertex x. The set L_1 consists of all the adjacent vertices of x. Therefore, $|L_1| = n - 1$. Similarly, set L_2 contains all vertices in the partition X except for vertex x. Hence, $|L_2| = n - 1$. The set L_3 contains the remaining vertex in Y that is not adjacent to vertex x. Hence, $|L_3| = 1$. According to *Lemma* 2.2.16,

$$\begin{aligned}
\ell(H_{n,n}) &= \sum_{x \in V} O_w(x) \\
&= \sum_{i=0}^{3} O_w(x)|L_i|, \quad x \in L_i \\
&\geq \lceil \frac{n}{2} \rceil + (n-1)(\lceil \frac{n}{2} \rceil - 1) + (n-1)(\lceil \frac{n}{2} \rceil - 2) + (\lceil \frac{n}{2} \rceil - 3) \\
&= 2n\lceil \frac{n}{2} \rceil - 3n.
\end{aligned}$$

It is easy to verify that the provided lower bound holds for $n \leq 4$ as well. □

The graph G_n with $2n + 1$ vertices is formed by adding an all-adjacent vertex to the crown graph $H_{n,n}$.

Theorem 2.2.21 ([23]). *The representation number of the graph G_n is n.*

Theorem 2.2.22. *The minimum possible length of a word-representant of* G_n *is bounded below by*

$$\ell(G_n) \geq 2n^2 - 2n.$$

Proof. Let w be a minimum length word-representant of the graph G_n. According to *Theorem* 2.2.21, it is clear that $O_{max}(w) = n$. Let X and Y be the bipartition of the crown graph $H_{n,n}$, with z as the all-adjacent vertex. There are two cases: either a vertex from either partition X or Y has the maximum number of occurrences in w, or the all-adjacent vertex z has the maximum number of occurrences in w. Without loss of generality, assume a vertex $x \in X$ occurs the maximum number of times in w. Let L_i be the set of vertices that are at a distance i from vertex x. The set L_1 contains all the adjacent vertices of x, including the all-adjacent vertex z. Hence, $|L_1| = n$. As $z \in L_1$, the set L_2 contains all the other vertices that are not adjacent to x. Hence, $|L_2| = n$. According to *Lemma* 2.2.16,

$$\begin{aligned}
\ell(G_n) &= \sum_{x\in V} O_w(x) \\
&= \sum_{i=0}^{2} O_w(x)|L_i|, \quad x \in L_i \\
&\geq n + n(n-1) + n(n-2) \\
&= 2n^2 - 2n.
\end{aligned}$$

Now assume that the all-adjacent vertex z occurs the maximum number of times in w. Let L'_i be the set of vertices that are at a distance i from z. The set L'_1 contains all vertices except z. Hence, $|L'_1| = 2n$. According to *Lemma* 2.2.16,

$$\begin{aligned}
\ell(G_n) &= \sum_{x\in V} O_w(x) \\
&= \sum_{i=0}^{1} O_w(x)|L'_i|, \quad x \in L'_i \\
&\geq n + 2n(n-1) \\
&= 2n^2 - n \\
&\geq 2n^2 - 2n.
\end{aligned}$$

□

2.3 Minimum Length Word-Representants of Graphs Obtained Through Graph Operations

In this section, we establish bounds on the minimum length of word-representants for certain graph operations that preserve word-representability, such as connecting two graphs by an edge and gluing two graphs at a vertex. In [32], Kitaev and Lozin provided insights into graph operations that preserve word-representability. Furthermore, we determine the minimum lengths of word-representants for the Cartesian and rooted products of word-representable graphs, employing morphism-based and occurrence-based functions, respectively. Finally, we address an open problem posed by Broere in [5], concerning the construction of a word representing the Cartesian product of two arbitrary word-representable graphs.

2.3.1 Connecting Two Graphs by an Edge and Gluing Two Graphs in a Vertex

In this subsection, we determine an upper bound on the minimum possible length of a word-representant of a graph obtained by connecting two graphs by an edge and gluing two graphs in a vertex. Let $sign(n)$ be a real function which is defined as follows: $sign(n) = 1$ if $n \geq 0$ and $sign(n) = -1$ if $n < 0$. In this section, we denote the concatenation of words as follows: $w_1 w_2 \dots w_n = \prod_{i=1}^{n} w_i$, where w_i are possibly empty words.

Theorem 2.3.1. *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two word-representable graphs. Let w_1 and w_2 be minimum length word-representants of G_1 and G_2, respectively. Let G' be the graph obtained by connecting G_1 and G_2 with an edge xy, where $x \in V_1$ and $y \in V_2$. Then*

$$\ell(G') \leq \ell(G_1) + \ell(G_2) + 2(|V_1| + |V_2|) + (|O_{w_2}(y) - O_{w_1}(x)|)\alpha.$$

where, $\alpha = \frac{1}{2}(|V_1| + |V_2| + sign(O_{w_2}(y) - O_{w_1}(x))(|V_1| - |V_2|))$

Proof. Let us assume, without loss of generality, that $O_{w_1}(x) = j \geq O_{w_2}(y) = i$. Let $w_{1a} = w_1$ and $w_{2a} = (\prod_{1}^{j-i} \pi(w_2))w_2$ be word-representants of graphs G_1 and G_2, respectively, such that $O_{w_{1a}}(x) = O_{w_{2a}}(y) = j$.

Let the words w_{1a} and w_{2a} be denoted as follows, where x^i denotes the i^{th} occurrence of the letter x.

$$w_{1a} = A_0 x^1 A_1 x^2 \dots A_{j-1} x^j A_j$$

$$w_{2a} = B_0 y^1 B_1 y^2 \dots B_{j-1} y^j B_j$$

Let $\pi(w_1) = P_0 x P_1$ and $\pi(w_2) = P'_0 y P'_1$. Note that P_0 and P'_0 are subwords of A_0 and B_0, respectively. Consider the word

$$w = P_0 x^1 P'_0 y^1 P'_1 P'_0 P_1 P_0 x^2 P_1 A_0 y^2 x^3 P'_1 B_0 y^3 A_1 B_1 x^4 y^4 A_2 B_2 x^5 y^5 \dots A_{j-1} B_{j-1} x^{j+2} y^{j+2} A_j B_j.$$

Claim: w represents G'.

Proof of Claim:

- Since $w|_{V_1} = \pi(w_1)\pi(w_1)w_{1a}$, w represents G_1 as per *Observation* 1.5.8.
- Since $w|_{V_2} = \pi(w_2)\pi(w_2)w_{2a}$, w represents G_2 as per *Observation* 1.5.8.
- Since $w|_{\{x,y\}} = x^1 y^1 \dots x^{j+2} y^{j+2}$, x and y alternate in w.
- Let $u \in V_1$ and $v \in V_2$ such that $\{u, v\} \neq \{x, y\}$. It is enough to prove that u and v do not alternate in w.

Suppose $u \neq x$ and $v \neq y$. Then, u appears in either P_0 or P_1, and v appears in either P'_0 or P'_1. Let us consider these cases separately.

Case 1: u appears in P_0, v appears in P'_0

Consider the factor $P_0 x^1 P'_0 y^1 P'_1 P'_0 P_1 P_0$ of w. Since P'_0 appears twice between two P_0, u and v do not alternate in w.

Case 2: u appears in P_0, v appears in P'_1

Consider the factor $P'_1 P'_0 P_1 P_0 x^2 P_1 A_0 y^2 x^3 P'_1$ of w. Since P_0 is a subword of A_0, every letter that appears in P_0 will also appear in A_0. As both P_0 and A_0 appear between two P'_1, u and v do not alternate in w.

Case 3: u appears in P_1, v appears in P'_0

Consider the factor $P'_0 y^1 P'_1 P'_0 P_1$ of w. Since P'_0 appears twice before the first occurrence of P_1, u and v do not alternate in w.

Case 4: u appears in P_1, v appears in P_1'

Consider the factor $P_1'P_0'P_1P_0x^2P_1A_0y^2x^3P_1'$ of w. Since P_1 appears twice in between two P_1', u and v do not alternate in w.

Suppose $u = x$ and $v \neq y$. Then, v appears in either P_0' or P_1'. Consider the factor $x^2P_1A_0y^2x^3$ of w. Since neither P_0' nor P_1' appear between x^2 and x^3, x and v do not alternate in w.

Finally, suppose $u \neq x$ and $v = y$. Then, u appears either in P_0 or in P_1. Consider the factor $y^2x^3P_1'B_0y^3$ of w. Since neither P_0 nor P_1 appear between y^2 and y^3, u and y do not alternate in w.

As a result, w represents the graph G'. Therefore, the length of a minimum length word-representant of the graph G'

$$\ell(G') \leq |w| = (j-i)|V_2| + \ell(G_1) + \ell(G_2) + 2|V_1| + 2|V_2|.$$

□

Theorem 2.3.2. *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two word representable graphs. Let w_1 and w_2 be minimum length word-representants of G_1 and G_2, respectively. Let G'' be the graph obtained from G_1 and G_2 by identifying vertices $x \in V_1$ and $y \in V_2$ into a single vertex z. Then*

$$\ell(G'') \leq \ell(G_1) + \ell(G_2) + 2(|V_1| + |V_2|) - max\{O_{w_1}(x), O_{w_2}(y)\} + (|O_{w_2}(y) - O_{w_1}(x)|)\alpha.$$

where, $\alpha = \frac{1}{2}(|V_1| + |V_2| + sign(O_{w_2}(y) - O_{w_1}(x))(|V_1| - |V_2|))$

Proof. Let w_{1a} and w_{2a} be word-representants of graphs G_1 and G_2, respectively, such that $O_{w_{1a}}(x) = O_{w_{2a}}(y)$. The word w_{ia} $(i = 1, 2)$ is obtained by concatenating the initial permutation $\pi(w_i)$ with w_i in order to ensure the same number of occurrences of x and y in both words w_{1a} and w_{2a}, respectively. Let us assume, without loss of generality, that $O_{w_1}(x) = j \geq O_{w_2}(y) = i$. Then, by considering w_{2a} as $(\prod_{1}^{j-i} \pi(w_2))w_2$, we can get $O_{w_{1a}}(x) = O_{w_{2a}}(y) = j$.

Let the words w_{1a} and w_{2a} be denoted as follows, where x^i denotes the i^{th} occurrence of the letter x.

$$w_{1a} = A_0 x^1 A_1 x^2 \dots A_{j-1} x^j A_j$$

$$w_{2a} = B_0 y^1 B_1 y^2 \dots B_{j-1} y^j B_j$$

Let $\pi(w_1) = P_0 x P_1$ and $\pi(w_2) = P'_0 y P'_1$. Note that P_0, and P'_0 are subwords of A_0 and B_0, respectively. Consider the word

$$w = P_0\, P'_0\, z^1\, P'_1\, P'_0\, P_1\, P_0\, z^2\, P_1\, A_0\, P'_1\, B_0\, z^3\, A_1\, B_1\, z^4\, A_2\, B_2\, z^5\ \dots\ A_{j-1}\, B_{j-1}\, z^{j+2}\, A_j\, B_j\,.$$

Claim: w represents G''.

Proof of Claim:

- Since $w|_{V_1} = \pi(w_1)\pi(w_1)w_{1a}$ (with vertex x replaced by z), w represents G_1 as per *Observation* 1.5.8.
- Since $w|_{V_2} = \pi(w_2)\pi(w_2)w_{2a}$ (with vertex y replaced by z), w represents G_2 as per *Observation* 1.5.8.
- Let $u \in V_1$ and $v \in V_2$ such that $u \neq v \neq z$. It is enough to prove that u and v do not alternate in w.

Thus, u appears in either P_0 or P_1, and v appears in either P'_0 or P'_1. Let us consider these cases separately.

Case 1: u appears in P_0, v appears in P'_0

Consider the factor $P_0 P'_0 z^1 P'_1 P'_0 P_1 P_0$ of w. Since P'_0 appears twice in between two P_0, u and v do not alternate in w.

Case 2: u appears in P_0, v appears in P'_1

Consider the factor $P'_1 P'_0 P_1 P_0 z^2 P_1 A_0 P'_1$ of w. Since P_0 is a subword of A_0, every letter that appears in P_0 will also appear in A_0. As both P_0 and A_0 appear between two P'_1, u and v do not alternate in w.

Case 3: u appears in P_1, v appears in P'_0

Consider the factor $P'_0 z^1 P'_1 P'_0 P_1$ of w. Since P'_0 appears twice before the first occurrence of P_1, u and v do not alternate in w.

Case 4: u appears in P_1, v appears in P'_1

Consider the factor $P_1'P_0'P_1P_0z^2P_1A_0P_1'$ of w. Since P_1 appears twice in between two P_1', u and v do not alternate in w.

As a result, w represents the graph G''. Therefore, the length of a minimum length word-representant of the graph G''

$$\ell(G'') \leq |w| = (j-i)|V_2| + \ell(G_1) + \ell(G_2) + 2|V_1| + 2|V_2| - j.$$

□

2.3.2 Cartesian Products

Cartesian product is a graph operation that preserves word-representability. In [6], Broere and Zantema have constructed a uniform word representing the Cartesian product of a word-representable graph with a complete graph using an occurrence-based function. In this section, we give an upper bound for the Cartesian product of word-representable graphs G with K_2 and K_n in terms of $\ell(G)$. In addition, we solve an open problem *(Question 6.10)* posed in [5] by finding a word representing the Cartesian product of two arbitrary word-representable graphs. Moreover, we give an upper bound for the minimum length of the words representing the Cartesian product of two arbitrary word-representable graphs. In this section, we will be using u^v to denote the ordered pair of vertices, (u, v), of $G \,\square\, H$, where $u \in V(G)$ and $v \in V(H)$.

Definition 2.3.3. *The Cartesian product of two graphs* $G = (V(G), E(G))$ *and* $H = (V(H), E(H))$ *is a graph* $G \,\square\, H = (V(G \,\square\, H), E(G \,\square\, H))$, *where* $V(G \,\square\, H) = V(G) \times V(H)$ *and*
$E(G \,\square\, H) = \{((u,v),(u',v')) \mid u = u' \text{ and } (v,v') \in E(H) \text{ or } v = v' \text{ and } (u,u') \in E(G)\}$.

Figure 2.3 depicts an example of the Cartesian product of two graphs.

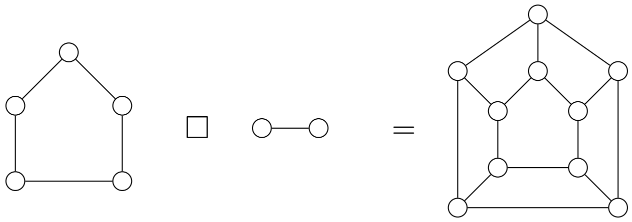


Figure 2.3: Cartesian product of two graphs

Theorem 2.3.4 ([32]). *Let G and H be two word-representable graphs. Then, the Cartesian product $G \,\square\, H$ is also word-representable.*

In [6], Broere and Zantema have provided an upper bound for the representation number of the graph $G \,\square\, K_n$.

Theorem 2.3.5 ([6]). *Let G be a k-representable graph for $k > 1$. Then, the graph $G \,\square\, K_n$ is $(k+n-1)$-representable.*

The following lemmas will help us in proving the main results.

Lemma 2.3.6. *Let $w = xUxZ$ be a word-representant of a graph G, where $x \in V(G)$, such that U is a word containing all elements of the set $V(G) \setminus \{x\}$. Then, $w' = UxZ$ also represents the graph G.*

Proof. Let $y \in V(G) \setminus \{x\}$. Suppose that x is adjacent to y. This implies y occurs exactly once in U, and x, y alternate in xZ. Therefore, $w'|_{\{x,y\}} = yxZ|_{\{x,y\}}$. Hence, x and y alternate in w' as well.

Suppose that x is not adjacent to y. Then either y occurs once or more than once in U, as U contains all elements of $V(G) \setminus \{x\}$. If y occurs only once in U, then x and y do not alternate in xZ as they do not in w. Hence, x and y do not alternate in $w' = UxZ$. If y occurs more than once in U, then x and y do not alternate in $w' = UxZ$. So, $w' = UxZ$ also represents G. ☐

Lemma 2.3.7. *Let $w = UxZx$ be a word-representant of a graph G, where $x \in V(G)$, such that Z is a word containing all elements of the set $V(G) \setminus \{x\}$. Then, $w' = UxZ$ also represents the graph G.*

Proof. By *Proposition* 1, $r(w) = xr(Z)xr(U)$ also represents the graph G. Hence, by *Lemma* 2.3.6, $r(w') = r(Z)xr(U)$ also represents the graph G. Again by *Proposition* 1, $w' = UxZ$ also represents the graph G. ☐

Before establishing the result for Cartesian products, let us introduce certain functions that will be used in the proof. Throughout the chapter, parentheses within a word are used as delimiters of the word parts, and they do not denote m-tuples.

Definition 2.3.8. *Let* $w_G = u_1u_2\dots u_n$ *be a word-representant of graph* G *and* $w_H = v_1v_2\dots v_m$ *be a word-representant of graph* H*. Define a function* $g : V(G)^* \times V(H)^* \to (V(G) \times V(H))^*$ *as,*

$$g^{w_H}(w_G) = (u_1^{v_1}u_1^{v_2}\dots u_1^{v_m})\dots(u_n^{v_1}u_n^{v_2}\dots u_n^{v_m}).$$

We denote (u_i, v_j) *as* $u_i^{v_j}$*, and* $g(w_G, w_H)$ *as* $g^{w_H}(w_G)$ *for convenience.*

Definition 2.3.9. *Let* $w_G = u_1u_2\dots u_n$ *be a word-representant of graph* G *and* $w_H = v_1v_2\dots v_m$ *be a word-representant of graph* H*. Define a function* $J : V(G)^* \times V(H)^* \to (V(G) \times V(H))^*$ *as,*

$$J^{w_H}(w_G) = (u_1^{v_1}u_2^{v_1}\dots u_n^{v_1})\dots(u_1^{v_m}u_2^{v_m}\dots u_n^{v_m}).$$

We denote (u_i, v_j) *as* $u_i^{v_j}$*, and* $J(w_G, w_H)$ *as* $J^{w_H}(w_G)$ *for convenience.*

Proposition 5. Let G and H be two word-representable graphs. Let $i \in V(H)$ be a letter and $w_1, w_2 \in V(G)^*$ be two words. Then,

$$g^i(w_1)J^i(w_2) = J^i(w_1w_2) = g^i(w_1w_2)$$

Proof. Let $w_1 = u_1u_2\dots u_n$ and $w_2 = v_1v_2\dots v_m$. Clearly,

$$g^i(w_1)J^i(w_2) = (u_1^i)\dots(u_n^i)(v_1^i\dots v_m^i) = (u_1^i\dots u_n^i)(v_1^i\dots v_m^i) = J^i(w_1w_2) = g^i(w_1w_2)$$

□

As mentioned at the beginning of this section, the following results provide an upper bound for the minimum length of the word-representants of the Cartesian product of two word-representable graphs in terms of the minimum length of their word-representants. Hence, finding a word of a certain length that represents the Cartesian product suffices to establish an upper bound for the minimum length. However, finding the word-representant is a challenging task. Thus, the main proof idea is to find a word using the morphism functions $g(w)$ and $J(w)$ as defined above. The proof ends by verifying whether the above-formed word represents the Cartesian product.

Theorem 2.3.10. *Let G be a word-representable graph with the minimum length of its word-representant $\ell(G)$. Then, the minimum length of the word-representants of the graph $G \,\square\, K_2$,*

$$\ell(G \,\square\, K_2) \leq 2\ell(G) + 3|G| - 2.$$

Proof. Let w_G be a minimum length word-representant of the graph G. Let $V(G) = \{u_1, u_2, \ldots, u_m\}$, $V(K_2) = \{1, 2\}$ and $u_j^i \in V(G) \times V(K_2)$. Let $\pi(w_G) = u_1 u_2 \ldots u_m$ be the initial permutation of the word w_G. Let $w_{K_2} = 12$ be a word-representant of the graph K_2. We claim that the word

$$w_{G \,\square\, K_2} = g^{r(w_{K_2})}(\pi(w_G))J^2(\pi(w_G))g^{w_{K_2}}(w_G)$$

represents the graph $G \,\square\, K_2$. By definition of the Cartesian product of graphs, it suffices to show that two vertices, u_j^i and u_k^l alternate in $w_{G \,\square\, K_2}$ if and only if either $j = k$ and i and l alternate in w_{K_2} or $i = l$ and u_j and u_k alternate in w_G.

Suppose $j = k$. Then, $w_{G \,\square\, K_2}|_{\{u_j^1,u_j^2\}} = u_j^2u_j^1u_j^2g^{w_{K_2}}(w_G)|_{\{u_j^1,u_j^2\}}$. Hence, u_j^1 and u_j^2 alternate in $w_{G \,\square\, K_2}$ as $g^{w_{K_2}}(w_G)|_{\{u_j^1,u_j^2\}} = u_j^1u_j^2u_j^1u_j^2\ldots$.

Suppose $i = l$. Then, by *Proposition* 5, we get $w_{G \,\square\, K_2}|_{\{u_j^i,u_k^i\}}$ as

$$(g^i(\pi(w_G))J^2(\pi(w_G))g^i(w_G))|_{\{u_j^i,u_k^i\}} = \begin{cases} g^1(\pi(w_G)w_G|_{\{u_j,u_k\}}) & \text{if } i = 1,\\ J^2(\pi(w_G)\pi(w_G)w_G|_{\{u_j,u_k\}}) & \text{if } i = 2.\end{cases}$$

Hence, by *Observation* 1.5.8, u_j^i and u_k^i alternate in $w_{G \,\square\, K_2}$ if and only if u_j and u_k alternate in w_G.

Suppose $j \neq k$ and $i \neq l$. Without loss of generality, suppose $j < k$ and $i = 1$, $l = 2$. Then $w_{G \,\square\, K_2}|_{\{u_j^1,u_k^2\}} = u_j^1u_k^2u_k^2g^{w_{K_2}}(w_G)|_{\{u_j^1,u_k^2\}}$. Hence, u_j^1 and u_k^2 do not alternate in $w_{G \,\square\, K_2}$. Moreover, if $i = 2$ and $l = 1$, we have $w_{G \,\square\, K_2}|_{\{u_j^2,u_k^1\}} = u_j^2u_k^1u_j^2g^{w_{K_2}}(w_G)|_{\{u_j^2,u_k^1\}}$. As $j < k$, $g^{w_{K_2}}(w_G)|_{\{u_j^2,u_k^1\}} = u_j^2u_k^1\ldots$. Therefore, $w_{G \,\square\, K_2}|_{\{u_j^2,u_k^1\}} = u_j^2u_k^1u_j^2u_j^2u_k^1\ldots$. Hence, u_j^2 and u_k^1 do not alternate in $w_{G \,\square\, K_2}$. As a result, $w_{G \,\square\, K_2}$ represents the graph $G \,\square\, K_2$. Therefore, $|w_{G \,\square\, K_2}| = 2\ell(G) + 3m$.

Consider $w_{G \,\square\, K_2} = (u_1^2u_1^1u_2^2u_2^1\ldots u_m^2u_m^1)(u_1^2\ldots u_m^2)g^{w_{K_2}}(w_G)$. Since $g^{w_{K_2}}(w_G) = u_1^1u_1^2\ldots$, by two consecutive usage of *Lemma* 2.3.6, we get

$$w'_{G\,\square\,K_2} = (u_2^2u_2^1u_3^2u_3^1\dots u_m^2u_m^1)(u_1^2\dots u_m^2)g^{w_{K_2}}(w_G).$$

Hence, $w'_{G\,\square\,K_2}$ also represents $G \,\square\, K_2$. Therefore,

$$\ell(G \,\square\, K_2) \leq |w'_{G\,\square\,K_2}| = 2\ell(G) + 3m - 2.$$

□

The following result gives us a tighter bound for the graph product $K_n \,\square\, K_2$.

Theorem 2.3.11. *The minimum length of the word-representants of the graph* $K_n \,\square\, K_2$ *for all* $n \geq 2$,

$$\ell(K_n \,\square\, K_2) \leq 5n - 4.$$

Proof. We know that, from *Theorem* 2.3.10,

$$w_{K_n\,\square\,K_2} = (2^b2^a3^b3^a\dots n^bn^a)(1^b\dots n^b)(1^a1^b\dots n^an^b)$$

represents the graph $K_n \,\square\, K_2$, where $w_{K_n} = 12\dots n$, $w_{K_2} = ab$ and $i^k \in V(K_n) \times V(K_2)$. By two consecutive usage of *Lemma* 2.3.7, we get

$$w'_{K_n\,\square\,K_2} = (2^b2^a3^b3^a\dots n^bn^a)(1^b\dots n^b)(1^a1^b\dots(n-1)^a(n-1)^b).$$

Hence, $w'_{K_n\,\square\,K_2}$ also represents the graph $K_n \,\square\, K_2$. Therefore, as $\ell(K_n) = n$,

$$\ell(K_n \,\square\, K_2) \leq |w'_{K_n\,\square\,K_2}| = 5n - 4.$$

□

The example below justifies the upper bound's tightness established in the above theorem.

Example 2.3.12. *We know that* $\ell(K_2 \,\square\, K_2) = \ell(C_4) = 6$. *Based on Theorem 2.3.11, we have* $\ell(K_2 \,\square\, K_2) \leq 5(2) - 4 = 6$.

Theorem 2.3.13. *Let G be a word-representable graph with the minimum length of its word-representant $\ell(G)$. Then, for all $n \geq 3$, the minimum length of the word-representants of the graph $G \,\square\, K_n$,*

$$\ell(G \,\square\, K_n) \leq n\ell(G) + (n^2 - 1)|G|.$$

Proof. Let w_G be a minimum length word-representant of the graph G. Let $V(G) = \{u_1, u_2, \ldots, u_m\}$, $V(K_n) = \{1, 2, \ldots, n\}$ and $u_j^i \in V(G) \times V(K_n)$. Let $\pi(w_G) = u_1u_2\ldots u_m$ be the initial permutation of the word w_G. Let $\pi(w_{K_n}, i) = i(i+1)\ldots n12\ldots(i-1)$, where $w_{K_n} = 123\ldots n$ is a word-representant of the graph K_n. Define a function $h^i(w_G) = g^{\pi(w_{K_n}, i)}(w_G)$. Let,

$$w_{G\,\square\,K_n} = h^2(\pi(w_G))J^2(\pi(w_G))h^3(\pi(w_G))\ldots J^{i-1}(\pi(w_G))h^i(\pi(w_G))\ldots J^n(\pi(w_G))h^1(w_G)$$

We claim that the word, $w_{G\,\square\,K_n}$, represents the graph $G \,\square\, K_n$. By the definition of the Cartesian product of graphs, it suffices to show that two vertices, u_j^i and u_k^l alternate in $w_{G\,\square\,K_n}$ if and only if either of the following cases holds.

(i) $j = k$ and i, l alternate in w_{K_n}.

(ii) $i = l$ and u_j, u_k alternate in w_G.

Case (i): Suppose $j = k$, and without loss of generality, assume $i < l$. Suppose $i > 1$ and $l < n$. Then,

$$w_{G\,\square\,K_n}|_{\{u_j^i,u_j^l\}} = [(\prod_{r=2}^{i} h^r(\pi(w_G)))J^i(\pi(w_G))(\prod_{r=i+1}^{l} h^r(\pi(w_G)))J^l(\pi(w_G))(\prod_{r=l+1}^{n} h^r(\pi(w_G)))h^1(w_G)]|_{\{u_j^i,u_j^l\}}.$$

Here,

$$h^r(\pi(w_G))|_{\{u_j^i,u_j^l\}} = \begin{cases} u_j^iu_j^l & \text{if } r \leq i, \\ u_j^lu_j^i & \text{if } i < r \leq l, \\ u_j^iu_j^l & \text{if } r > l. \end{cases}$$

Moreover, $h^1(w_G)|_{\{u_j^i,u_j^l\}} = u_j^iu_j^l\ldots$. Hence,

$$w_{G\,\square\,K_n}|_{\{u_j^i,u_j^l\}} = (\prod_{r=2}^{i} u_j^iu_j^l)u_j^i(\prod_{r=i+1}^{l} u_j^lu_j^i)u_j^\ell(\prod_{r=l+1}^{n} u_j^iu_j^l)u_j^iu_j^l\ldots$$

. As a result, u_j^i and u_j^l alternate in $w_{G\,\square\,K_n}$ as i and l alternate in w_{K_n}.

Suppose $j = k$ and $i = 1$ and $l < n$. Then,

$$w_{G \,\square\, K_n}|_{\{u_j^1, u_j^l\}} = (\prod_{r=2}^{l} u_j^l u_j^1) u_j^\ell (\prod_{r=l+1}^{n} u_j^1 u_j^l) u_j^1 u_j^l \ldots$$

Hence, u_j^1 and u_j^l alternate in $w_{G \,\square\, K_n}$.

Suppose $j = k$ and $i > 1$ and $l = n$. Then,

$$w_{G \,\square\, K_n}|_{\{u_j^i, u_j^n\}} = (\prod_{r=2}^{i} u_j^i u_j^n) u_j^i (\prod_{r=i+1}^{n} u_j^n u_j^i) u_j^n u_j^i u_j^n \ldots$$

Hence, u_j^i and u_j^n alternate in $w_{G \,\square\, K_n}$.

Suppose $j = k$, $i = 1$ and $l = n$. Then, $w_{G \,\square\, K_n}|_{\{u_j^1, u_j^n\}} = (\prod_{r=2}^{n} u_j^n u_j^1) u_j^n u_j^1 u_j^n \ldots$ Hence, u_j^1 and u_j^n alternate in $w_{G \,\square\, K_n}$.

Case (ii): Suppose $i = l$, and without loss of generality, assume $j < k$. Then,

$$w_{G \,\square\, K_n}|_{\{u_j^i, u_k^i\}} = [(\prod_{r=2}^{i} h^r(\pi(w_G))) J^i(\pi(w_G)) (\prod_{r=i+1}^{n} h^r(\pi(w_G))) h^1(w_G)]|_{\{u_j^i, u_k^i\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^i, u_k^i\}} = u_j^i u_k^i$. Hence,

$$w_{G \,\square\, K_n}|_{\{u_j^i, u_k^i\}} = (\prod_{r=2}^{i} u_j^i u_k^i) u_j^i u_k^i (\prod_{r=i+1}^{n} u_j^i u_k^i) g^i(w_G|_{\{u_j, u_k\}}).$$

Hence, u_j^i and u_k^i alternate in $w_{G \,\square\, K_n}$ if and only if u_j and u_k alternate in w_G.

Case (iii): Suppose $i \neq l$ and $j \neq k$. We assume without loss of generality $i < l$ and $j < k$ throughout this case. In the first part, we prove that u_j^i and u_k^l do not alternate in $w_{G \,\square\, K_n}$, and in the second part, we prove that u_j^l and u_k^i do not alternate in $w_{G \,\square\, K_n}$.

First Part:

First, we consider

$$w_{G \,\square\, K_n}|_{\{u_j^i, u_k^l\}} = [(\prod_{r=2}^{i} h^r(\pi(w_G))) J^i(\pi(w_G)) (\prod_{r=i+1}^{l} h^r(\pi(w_G))) J^\ell(\pi(w_G)) (\prod_{r=l+1}^{n} h^r(\pi(w_G))) h^1(w_G)]|_{\{u_j^i, u_k^l\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^i,u_k^l\}} = u_j^i u_k^l$. Hence, for all $1 < i < l < n$,

$$w_{G\,\square\,K_n}|_{\{u_j^i,u_k^l\}} = (\prod_{r=2}^{i} u_j^i u_k^l)u_j^i(\prod_{r=i+1}^{l} u_j^i u_k^l)u_k^\ell(\prod_{r=l+1}^{n} u_j^i u_k^l)u_j^i u_k^l \dots$$

Hence, u_j^i and u_k^l do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i = 1$ and $l < n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^1,u_k^l\}} = (\prod_{r=2}^{l} u_j^1 u_k^l)u_k^\ell(\prod_{r=l+1}^{n} u_j^1 u_k^l)u_j^1 u_k^l \dots$ Hence, u_j^1 and u_k^l do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i > 1$ and $l = n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^i,u_k^n\}} = (\prod_{r=2}^{i} u_j^i u_k^n)u_j^i(\prod_{r=i+1}^{n} u_j^i u_k^n)u_k^n u_j^i u_k^n \dots$ Hence, u_j^i and u_k^n do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i = 1$ and $l = n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^1,u_k^n\}} = (\prod_{r=2}^{n} u_j^1 u_k^n)u_k^n u_j^1 u_k^n \dots$ Hence, u_j^1 and u_k^n do not alternate in $w_{G\,\square\,K_n}$.

Second Part:

Now consider,

$$w_{G\,\square\,K_n}|_{\{u_j^l,u_k^i\}} = [(\prod_{r=2}^{i} h^r(\pi(w_G)))J^i(\pi(w_G))(\prod_{r=i+1}^{l} h^r(\pi(w_G)))J^\ell(\pi(w_G))(\prod_{r=l+1}^{n} h^r(\pi(w_G)))h^1(w_G)]|_{\{u_j^l,u_k^i\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^i,u_k^i\}} = u_j^l u_k^i$. Hence for all $1 < i < l < n$,

$$w_{G\,\square\,K_n}|_{\{u_j^l,u_k^i\}} = (\prod_{r=2}^{i} u_j^l u_k^i)u_k^i(\prod_{r=i+1}^{l} u_j^l u_k^i)u_j^\ell(\prod_{r=l+1}^{n} u_j^l u_k^i)u_j^l u_k^i \dots$$

Hence, u_j^l and u_k^i do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i = 1$ and $l < n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^l,u_k^1\}} = (\prod_{r=2}^{l} u_j^l u_k^1)u_j^\ell(\prod_{r=l+1}^{n} u_j^l u_k^1)u_j^l u_k^1 \dots$ Hence, u_j^l and u_k^1 do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i > 1$ and $l = n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^n,u_k^i\}} = (\prod_{r=2}^{i} u_j^n u_k^i)u_k^i(\prod_{r=i+1}^{n} u_j^n u_k^i)u_j^n u_j^n u_k^i \dots$ Hence, u_j^n and u_k^i do not alternate in $w_{G\,\square\,K_n}$.

Suppose $i = 1$ and $l = n$. Then, $w_{G\,\square\,K_n}|_{\{u_j^n,u_k^1\}} = (\prod_{r=2}^{n} u_j^n u_k^1)u_j^n u_j^n u_k^1 \dots$ Hence, u_j^n and u_k^1 do not alternate in $w_{G\,\square\,K_n}$.

Therefore, $w_{G\,\square\,K_n}$ represents the graph $G \,\square\, K_n$. Hence, $\ell(G \,\square\, K_n) \leq |w_{G\,\square\,K_n}|$. As

$|h^r(w_G)| = n|w_G|$ for all r,

$$\begin{aligned}\ell(G \,\square\, K_n) &\leq (n-1)|h^r(\pi(w_G))| + (n-1)|J^r(\pi(w_G))| + |h^1(w_G)| \\ &= (n-1)nm + (n-1)m + n\ell(G) \\ &= n\ell(G) + (n^2-1)m.\end{aligned}$$

□

Remark 4. By *Theorem* 2.3.5, we know that if G is $k(\geq 2)$-representable, then $\ell(G \,\square\, K_n) \leq n(k+n-1)|G|$. Thus, a natural question arises: Is the bound provided in *Theorem* 2.3.10 and *Theorem* 2.3.13 tighter or not?

For example, let us consider $Pr_3 \,\square\, K_2$.

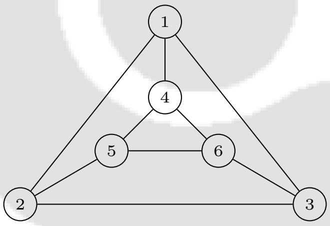


Figure 2.4: Minimum length word-representant of Pr_3, $w = 52634561425$

Since $\mathcal{R}(Pr_3) = 3$, $\ell(Pr_3 \,\square\, K_2) \leq n(k+n-1)|Pr_3| = 2\ .\ 4\ .\ 6 = 48$.

As shown in *Figure* 2.4, $\ell(Pr_3) = 11$, and $3|Pr_3| - \ell(Pr_3) = 3\ .\ 6 - 11 = 7$. Therefore, by *Theorem* 2.3.10, we get $\ell(Pr_3 \,\square\, K_2) \leq 2\ .\ 11 + 3\ .\ 6 - 2 = 38 < 48$.

In general, let us consider

$$n(k+n-1)|G| - (n\ell(G) + (n^2-1)|G|) = n(k|G| - \ell(G)) - (n-1)|G|$$

Therefore, if $k|G| - \ell(G) > (n-1)|G|/n$, the bound provided by *Theorem* 2.3.13 will be tighter than the existing bound. However, until now, there are no known results regarding the parameter $k|G| - \ell(G)$.

Theorem 2.3.14. *Let G and H be two word-representable graphs such that $|G| \geq |H|$, with the minimum length of their word-representants $\ell(G)$ and $\ell(H)$, respectively. Then, the minimum length of the word-representants of the graph $G \,\square\, H$,*

$$\ell(G \,\square\, H) \leq |H|\ell(G) + |G|\ell(H) + (|H|^2 - 1)|G|.$$

Proof. Let w_G and w_H be minimum length word-representants of the graph G and H, respectively. Let $V(G) = \{u_1, u_2, \dots, u_m\}$, $V(H) = \{v_1, v_2, \dots, v_n\}$ and $u_j^{v_i} \in V(G) \times V(H)$. Let $\pi(w_G) = u_1u_2\dots u_m$ be the initial permutation and $\sigma(w_G) = t_1t_2\dots t_m$ be the final permutation of the word w_G. Let $\pi(w_H) = v_1v_2\dots v_n$ be the initial permutation of the word w_H. Let $\pi(w_H, i) = v_iv_{i+1}\dots v_nv_1\dots v_{i-1}$. Let us define a function $h^i(w_G) = g^{\pi(w_H,i)}(w_G)$. Let

$$w_{G\,\square\,H} = h^2(\pi(w_G))J^{v_2}(\pi(w_G))h^3(\pi(w_G))\dots J^{v_i}(\pi(w_G))h^{i+1}(\pi(w_G))\dots J^{v_n}(\pi(w_G))h^1(w_G)J^{w_H}(\sigma(w_G)).$$

We claim that the word $w_{G\,\square\,H}$, represents the graph $G \,\square\, H$. By definition of the Cartesian product of graphs, it suffices to prove that two vertices, $u_j^{v_i}$ and $u_k^{v_l}$ alternate in $w_{G\,\square\,H}$ if and only if either of the following cases holds.

(i) $j = k$ and v_i, v_l alternate in w_H.

(ii) $i = l$ and u_j, u_k alternate in w_G.

Case (i): Suppose $j = k$. Without loss of generality, assume $i < l$. Suppose that $i > 1$ and $l < n$. Then, $w_{G\,\square\,H}|_{\{u_j^{v_i},u_j^{v_l}\}}$ is

$$[(\prod_{r=2}^{i} h^r(\pi(w_G)))J^{v_i}(\pi(w_G))(\prod_{r=i+1}^{l} h^r(\pi(w_G)))J^{v_l}(\pi(w_G))(\prod_{r=l+1}^{n} h^r(\pi(w_G)))h^1(w_G)J^{w_H}(\sigma(w_G))]|_{\{u_j^{v_i},u_j^{v_l}\}}.$$

Here,

$$h^r(\pi(w_G))|_{\{u_j^{v_i},u_j^{v_l}\}} = \begin{cases} u_j^{v_i}u_j^{v_l} & \text{if } r \leq i, \\ u_j^{v_l}u_j^{v_i} & \text{if } i < r \leq l, \\ u_j^{v_i}u_j^{v_l} & \text{if } r > l. \end{cases}$$

Further, $h^1(w_G)|_{\{u_j^{v_i},u_j^{v_l}\}} = u_j^{v_i}u_j^{v_l}\dots$ and $J^{w_H}(\sigma(w_G))|_{\{u_j^{v_i},u_j^{v_l}\}} = J^{w_H|_{\{v_i,v_l\}}}(u_j)$. Hence,

$$w_{G\,\square\,H}|_{\{u_j^{v_i},u_j^{v_l}\}} = (\prod_{r=2}^{i} u_j^{v_i}u_j^{v_l})u_j^{v_i}(\prod_{r=i+1}^{l} u_j^{v_l}u_j^{v_i})u_j^{v_l}(\prod_{r=l+1}^{n} u_j^{v_i}u_j^{v_l})(u_j^{v_i}u_j^{v_l}\dots)J^{w_H|_{\{v_i,v_l\}}}(u_j).$$

Hence, $u_j^{v_i}$ and $u_j^{v_l}$ alternate in $w_{G\,\square\,H}|_{\{u_j^{v_i},u_j^{v_l}\}}$ if and only if v_i and v_l alternate in w_H.

Suppose $i = 1$ and $l < n$. Then,

$$w_{G\,\square\,H}|_{\{u_j^{v_1},u_j^{v_l}\}} = (\prod_{r=2}^{l} u_j^{v_l}u_j^{v_1})u_j^{v_l}(\prod_{r=l+1}^{n} u_j^{v_1}u_j^{v_l})(u_j^{v_1}u_j^{v_l}\dots)J^{w_H|_{\{v_1,v_l\}}}(u_j).$$

Hence, $u_j^{v_1}$ and $u_j^{v_l}$ alternate in $w_{G \,\square\, H}|_{\{u_j^{v_1}, u_j^{v_l}\}}$ if and only if v_1 and v_l alternate in w_H.

Suppose $i > 1$ and $l = n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_i}, u_j^{v_n}\}} = (\prod_{r=2}^{i} u_j^{v_i} u_j^{v_n}) u_j^{v_i} (\prod_{r=i+1}^{n} u_j^{v_n} u_j^{v_i}) u_j^{v_n} (u_j^{v_i} u_j^{v_n} \ldots) J^{w_H|_{\{v_i, v_n\}}}(u_j).$$

Hence, $u_j^{v_i}$ and $u_j^{v_n}$ alternate in $w_{G \,\square\, H}|_{\{u_j^{v_i}, u_j^{v_n}\}}$ if and only if v_i and v_n alternate in w_H.

Suppose $i = 1$ and $l = n$. Then, $w_{G \,\square\, K_n}|_{\{u_j^{v_1}, u_j^{v_n}\}} = (\prod_{r=2}^{n} u_j^{v_n} u_j^{v_1}) u_j^{v_n} (u_j^{v_1} u_j^{v_n} \ldots)$ $J^{w_H|_{\{v_1, v_n\}}}(u_j)$. Hence, $u_j^{v_1}$ and $u_j^{v_n}$ alternate in $w_{G \,\square\, H}|_{\{u_j^{v_1}, u_j^{v_n}\}}$ if and only if v_1 and v_n alternate in w_H.

Case (ii): Suppose $i = l$ and consider $j < k$ without loss of generality. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_i}, u_k^{v_i}\}} = [(\prod_{r=2}^{i} h^r(\pi(w_G))) J^{v_i}(\pi(w_G)) (\prod_{r=i+1}^{n} h^r(\pi(w_G))) h^1(w_G) J^{w_H}(\sigma(w_G))]|_{\{u_j^{v_i}, u_k^{v_i}\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^{v_i}, u_k^{v_i}\}} = u_j^{v_i} u_k^{v_i}$. Further, $h^1(w_G)|_{\{u_j^{v_i}, u_k^{v_i}\}} = g^{v_i}(w_G|_{\{u_j, u_k\}})$ and $J^{w_H}(\sigma(w_G))|_{\{u_j^{v_i}, u_k^{v_i}\}}$
$= J^{w_H|_{v_i}}(\sigma(w_G)|_{\{u_j, u_k\}})$. Hence,

$$w_{G \,\square\, H}|_{\{u_j^{v_i}, u_k^{v_i}\}} = (\prod_{r=2}^{i} u_j^{v_i} u_k^{v_i}) u_j^{v_i} u_k^{v_i} (\prod_{r=i+1}^{n} u_j^{v_i} u_k^{v_i}) g^{v_i}(w_G|_{\{u_j, u_k\}}) J^{w_H|_{v_i}}(\sigma(w_G)|_{\{u_j, u_k\}}).$$

Hence, $u_j^{v_i}$ and $u_k^{v_i}$ alternate in $w_{G \,\square\, H}|_{\{u_j^{v_i}, u_k^{v_i}\}}$ if and only if u_j and u_k alternate in w_G.

Case (iii): Suppose $i \neq l$ and $j \neq k$. We assume without loss of generality $i < l$ and $j < k$ throughout this case. In the first part, we prove that $u_j^{v_i}$ and $u_k^{v_l}$ do not alternate in $w_{G \,\square\, K_n}$, and in the second part, we prove that $u_j^{v_l}$ and $u_k^{v_i}$ do not alternate in $w_{G \,\square\, K_n}$.

First Part:

Firstly we consider $w_{G \,\square\, H}|_{\{u_j^{v_i}, u_k^{v_l}\}}$, which is,

$$[(\prod_{r=2}^{i} h^r(\pi(w_G))) J^{v_i}(\pi(w_G)) (\prod_{r=i+1}^{l} h^r(\pi(w_G))) J^{v_l}(\pi(w_G)) (\prod_{r=l+1}^{n} h^r(\pi(w_G))) h^1(w_G) J^{w_H}(\sigma(w_G))]|_{\{u_j^{v_i}, u_k^{v_l}\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^{v_i}, u_k^{v_l}\}} = u_j^{v_i} u_k^{v_l}$. Hence for all $1 < i < l < n$,

$$w_{G \,\square\, H}|_{\{u_j^{v_i}, u_k^{v_l}\}} = (\prod_{r=2}^{i} u_j^{v_i} u_k^{v_l}) u_j^{v_i} (\prod_{r=i+1}^{l} u_j^{v_i} u_k^{v_l}) u_k^{v_l} (\prod_{r=l+1}^{n} u_j^{v_i} u_k^{v_l}) g^{v_i v_l}(w_G|_{\{u_j, u_k\}}) J^{w_H|_{\{v_i, v_l\}}}(\sigma(w_G)|_{\{u_j, u_k\}}).$$

Hence, $u_j^{v_i}$ and $u_k^{v_l}$ do not alternate in $w_{G \,\square\, H}$.

Suppose $i = 1$ and $l < n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_1},u_k^{v_l}\}} = (\prod_{r=2}^{l} u_j^{v_1}u_k^{v_l})u_k^{v_l}(\prod_{r=l+1}^{n} u_j^{v_1}u_k^{v_l})g^{v_1v_l}(w_G|_{\{u_j,u_k\}})J^{w_H|_{\{v_1,v_l\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_1}$ and $u_k^{v_l}$ do not alternate in $w_{G \,\square\, H}$.

Suppose $i > 1$ and $l = n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_i},u_k^{v_n}\}} = (\prod_{r=2}^{i} u_j^{v_i}u_k^{v_n})u_j^{v_i}(\prod_{r=i+1}^{n} u_j^{v_i}u_k^{v_n})u_k^{v_n}g^{v_iv_n}(w_G|_{\{u_j,u_k\}})J^{w_H|_{\{v_i,v_n\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_i}$ and $u_k^{v_n}$ do not alternate in $w_{G \,\square\, H}$.

Suppose $i = 1$ and $l = n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_1},u_k^{v_n}\}} = (\prod_{r=2}^{n} u_j^{v_1}u_k^{v_n})u_k^{v_n}g^{v_1v_n}(w_G|_{\{u_j,u_k\}})J^{w_H|_{\{v_1,v_n\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_1}$ and $u_k^{v_n}$ do not alternate in $w_{G \,\square\, H}$.

Second Part:

Now let us consider $w_{G \,\square\, H}|_{\{u_j^{v_l},u_k^{v_i}\}}$, which is,

$$[(\prod_{r=2}^{i} h^r(\pi(w_G)))J^{v_i}(\pi(w_G))(\prod_{r=i+1}^{l} h^r(\pi(w_G)))J^{v_l}(\pi(w_G))(\prod_{r=l+1}^{n} h^r(\pi(w_G)))h^1(w_G)J^{w_H}(\sigma(w_G))]|_{\{u_j^{v_l},u_k^{v_i}\}}.$$

For all r, $h^r(\pi(w_G))|_{\{u_j^{v_l},u_k^{v_i}\}} = u_j^{v_l}u_k^{v_i}$. Hence, for all $1 < i < l < n$,

$$w_{G \,\square\, H}|_{\{u_j^{v_l},u_k^{v_i}\}} = (\prod_{r=2}^{i} u_j^{v_l}u_k^{v_i})u_k^{v_i}(\prod_{r=i+1}^{l} u_j^{v_l}u_k^{v_i})u_j^{v_l}(\prod_{r=l+1}^{n} u_j^{v_l}u_k^{v_i})g^{v_iv_l}(w_G|_{\{u_j,u_k\}})J^{w_H|_{\{v_i,v_l\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_l}$ and $u_k^{v_i}$ do not alternate in $w_{G \,\square\, H}$

Suppose $i = 1$ and $l < n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_l},u_k^{v_1}\}} = (\prod_{r=2}^{l} u_j^{v_l}u_k^{v_1})u_j^{v_l}(\prod_{r=l+1}^{n} u_j^{v_l}u_k^{v_1})g^{v_1v_l}(w_G|_{\{u_j,u_k\}})J^{w_H|_{\{v_1,v_l\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_l}$ and $u_k^{v_1}$ do not alternate in $w_{G \,\square\, H}$.

Suppose $i > 1$ and $l = n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_n}, u_k^{v_i}\}} = (\prod_{r=2}^{i} u_j^{v_n} u_k^{v_i}) u_k^{v_i} (\prod_{r=i+1}^{n} u_j^{v_n} u_k^{v_i}) u_j^{v_n} g^{v_i v_n}(w_G|_{\{u_j,u_k\}}) J^{w_H|_{\{v_i,v_n\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_n}$ and $u_k^{v_i}$ do not alternate in $w_{G \,\square\, H}$.

Suppose $i = 1$ and $l = n$. Then,

$$w_{G \,\square\, H}|_{\{u_j^{v_n}, u_k^{v_1}\}} = (\prod_{r=2}^{n} u_j^{v_n} u_k^{v_1}) u_j^{v_n} g^{v_1 v_n}(w_G|_{\{u_j,u_k\}}) J^{w_H|_{\{v_1,v_n\}}}(\sigma(w_G)|_{\{u_j,u_k\}}).$$

Hence, $u_j^{v_n}$ and $u_k^{v_1}$ do not alternate in $w_{G \,\square\, H}$ since $u_j^{v_n}$ occurs to the left of $u_k^{v_1}$ in $g^{v_1 v_n}(w_G|_{\{u_j,u_k\}})$.

Therefore, $w_{G \,\square\, H}$ represents the graph $G \,\square\, H$. Hence, $\ell(G \,\square\, H) \leq |w_{G \,\square\, H}|$. As $|h^r(w_G)| = n\ell(G)$ for all r and $|J^{w_H}(\sigma(w_G))| = m\ell(H)$,

$$\begin{aligned}
\ell(G \,\square\, H) &\leq (n-1)|h^r(\pi(w_G))| + \sum_{i=2}^{n} |J^{v_i}(\pi(w_G))| + |h^1(w_G)| + |J^{w_H}(\sigma(w_G))| \\
&= (n-1)nm + (n-1)m + n\ell(G) + \ell(H)m \\
&= n\ell(G) + m\ell(H) + (n^2-1)m.
\end{aligned}$$

□

Remark 5. If we calculate an upper bound for $G \,\square\, K_n$ using *Theorem* 2.3.14, then $\ell(G \,\square\, K_n) \leq n\ell(G) + mn + (n^2-1)m$. However, using *Theorem* 2.3.13, we get $\ell(G \,\square\, K_n) \leq n\ell(G) + (n^2-1)m$, which is a much tighter bound than the former one. Hence, we stated three different theorems for $G \,\square\, K_2$, $G \,\square\, K_n$ and $G \,\square\, H$.

Corollary 2. Let G and H be two word-representable graphs with representation numbers k_1 and k_2, respectively. Then, $G \,\square\, H$ is $(k_1 + k_2 + \min\{|G|, |H|\})$-representable.

Proof. Let w_G and w_H be minimum length word representants of the graphs G and H, respectively. Without loss of generality, let us assume that $|G| \geq |H|$. Let $V(G) = \{u_1, u_2, \ldots, u_m\}$, $V(H) = \{v_1, v_2, \ldots, v_n\}$ and $u_j^{v_i} \in V(G) \times V(H)$. Let $\pi(w_G) = u_1 u_2 \ldots u_m$ be the initial permutation and $\sigma(w_G) = t_1 t_2 \ldots t_m$ be the final permutation of the word w_G. Let $\pi(w_H) = v_1 v_2 \ldots v_n$ be the initial permutation of the word w_H. Let $\pi(w_H, i) =$

$v_iv_{i+1}\dots v_nv_1\dots v_{i-1}$. Let us define a function $h^i(w_G) = g^{\pi(w_H,i)}(w_G)$. Let,

$$w_{G\,\square\,H} = h^2(\pi(w_G))\dots J^{v_i}(\pi(w_G))h^{i+1}(\pi(w_G))\dots J^{v_n}(\pi(w_G))h^1(w_G)J^{w_H}(\sigma(w_G)).$$

Then by *Theorem* 2.3.14, $w_{G\,\square\,H}$ represents the graph $G \,\square\, H$. Here, $O_{h^r(\pi(w_G))}(u_j^{v_i}) = 1$ for all r and $u_j^{v_i}$. Similarly, $O_{J^{v_i}(\pi(w_G))}(u_j^{v_i}) = 1$ for all i and $u_j^{v_i}$. Since the representation number of the graph G is k_1, $O_{h^1(w_G)}(u_j^{v_i}) \leq k_1$ for any $u_j^{v_i}$ and since the representation number of the graph H is k_2, $O_{J^{w_H}(\sigma(w_G))}(u_j^{v_i}) \leq k_2$ for any $u_j^{v_i}$. Hence for any $u_j^{v_i}$,

$$\begin{aligned} O_{w_{G\,\square\,H}}(u_j^{v_i}) &\leq \sum_{r=2}^{n} O_{h^r(\pi(w_G))}(u_j^{v_i}) + O_{J^{v_i}(\pi(w_G))}(u_j^{v_i}) + O_{h^1(w_G)}(u_j^{v_i}) + O_{J^{w_H}(\sigma(w_G))}(u_j^{v_i}) \\ &\leq n - 1 + 1 + k_1 + k_2 \\ &= k_1 + k_2 + n. \end{aligned}$$

□

2.3.3 Rooted Products

Another graph product that preserves word-representability is the rooted product. In this section, we establish an upper bound on the minimum length of word-representants for the rooted product of a word-representable graph G with K_2, K_n, and an arbitrary word-representable graph H.

Definition 2.3.15. *The rooted product of a graph G and a rooted graph H is the graph $G \circ H$, which is defined as follows: take $|V(G)|$ copies of H and for every vertex v_i of G, identify v_i with the root vertex of the ith copy of H.*

Figure 2.5 depicts an example of the rooted product of two graphs.

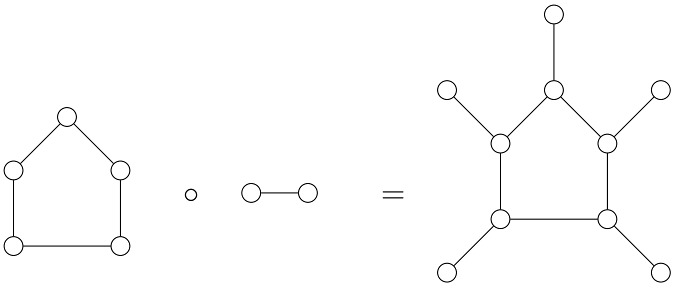


Figure 2.5: Rooted product of two graphs

Theorem 2.3.16 ([32]). *Let G and H be two word-representable graphs. Then, the rooted product $G \circ H$ is also word-representable. Moreover, $G \circ H$ is $k(\geq 2)$-representable, where $k = \max\{\mathcal{R}(G), \mathcal{R}(H)\}$.*

Definition 2.3.17. *Let V and V' be (possibly different) alphabets. The labeling function of a word over V is defined as $H : V^* \rightarrow V \times [k]$, where the i-th occurrence of each letter x is mapped to the pair (x, i), and k satisfies the property that every symbol occurs at most k times in w. The word $H(w)$ is called the labeled version of w. An occurrence-based function is defined as applying a string homomorphism $h : V \times [k] \rightarrow (V')^*$ to an already labelled version of a word. As a shorthand, we will write $h(w)$ instead of $h(H(w))$.*

Theorem 2.3.18. *Let G be a word-representable graph with the minimum length of its word-representants $\ell(G)$. Then, the minimum length of the word-representants of the graph $G \circ K_2$,*

$$\ell(G \circ K_2) \leq 2\ell(G) + \kappa_G.$$

where κ_G is the size of the maximum clique of the graph G.

Proof. Let w_G be a minimum length word-representant of the graph G. Let $V(K_2) = \{r, 1\}$, where r is the rooted vertex. Let $x^r, x^1 \in V(G) \times V(K_2)$, where $x \in V(G)$. We claim that the word $w_{G \circ K_2} = h(w_G)$ represents the graph $G \circ K_2$, where

If $O_{w_G}(x) = 1$

$$h(x, 1) = x^1\ x^r\ x^1$$

If $O_{w_G}(x) \geq 2$,

$$h(x, i) = \begin{cases} x^r & \text{if } i = 1, \\ x^1\ x^r\ x^1 & \text{if } i = 2, \\ x^r\ x^1 & \text{if } i \geq 3. \end{cases}$$

By the definition of the rooted product of graphs, it suffices to show that two vertices, x^i and y^j alternate in $w_{G \circ K_2}$ if and only if either $i = j = r$ and x and y alternate in w_G or $x = y$ and i and j alternate in w_{K_2}.

Suppose $i = j = r$. Then,

$$w_{G\circ K_2}|_{\{x^r,y^r\}} = h(w_G)|_{\{x^r,y^r\}}$$

Since for all $x \in V(G)$, x^r occurs exactly once for every occurrence of x in w_G, x^r and y^r alternate in $w_{G\circ K_2}$ if and only if x and y alternate in w_G.

Suppose $i \neq j$ and $x = y$. Then,

$$w_{G\circ K_2}|_{\{x^r,x^1\}} = h(w_G)|_{\{x^r,x^1\}} = \begin{cases} x^1\ x^r\ x^1 & \text{if } O_{w_G}(x) = 1, \\ x^r\ x^1\ x^r\ x^1\ \dots & \text{if } O_{w_G}(x) \geq 2. \end{cases}$$

Hence, x^r and x^1 alternate in $w_{G\circ K_2}$ as r and 1 alternate in w_{K_2}.

Suppose $i \neq j$ and $x \neq y$. Without loss of generality, assume that x occurs to the left of y in $\pi(w_G)$. Suppose $i = r$ and $j = 1$. Then,

$$w_{G\circ K_2}|_{\{x^r,y^1\}} = h(w_G)|_{\{x^r,y^1\}} = \begin{cases} \dots\ x^r\ y^1\ y^1\ \dots & \text{if } O_{w_G}(y) = 1, \\ x^r\ x^r\ y^1\ y^1\ \dots & \text{if } O_{w_G}(y) \geq 2. \end{cases}$$

Hence, x^r and y^1 do not alternate in $w_{G\circ K_2}$. Now suppose $i = 1$ and $j = r$. Then,

$$w_{G\circ K_2}|_{\{x^1,y^r\}} = h(w_G)|_{\{x^1,y^r\}} = \begin{cases} x^1\ x^1\ y^r\ \dots & \text{if } O_{w_G}(x) = 1, \\ y^r\ x^1\ x^1\ \dots & \text{if } O_{w_G}(x) \geq 2. \end{cases}$$

Hence, x^1 and y^r do not alternate in $w_{G\circ K_2}$.

Suppose $i = j = 1$ and $x \neq y$. Without loss of generality, assume that x occurs to the left of y in $\pi(w_G)$. Then,

$$w_{G\circ K_2}|_{\{x^1,y^1\}} = h(w_G)|_{\{x^1,y^1\}} = x^1\ x^1\ y^1\ y^1\ \dots \quad \text{or} \quad y^1\ y^1\ x^1\ x^1\ \dots$$

Hence, x^1 and y^1 do not alternate in $w_{G\circ K_2}$.

Therefore, $w_{G\circ K_2}$ represents the graph $G \circ K_2$. Hence,

$$\ell(G \circ K_2) \quad \leq \quad |w_{G\circ K_2}|$$

$$\begin{aligned}
&= \sum_{i=2}^{\mathcal{R}(G)} 2i|O(w_G, i)| + 3|O(w_G, 1)| \\
&= 2\sum_{i=1}^{\mathcal{R}(G)} i|O(w_G, i)| + |O(w_G, 1)| \\
&= 2\ell(G) + |O(w_G, 1)|.
\end{aligned}$$

where $\mathcal{R}(G)$, the representation number, is the least k for which the graph G is k-representable. Hence from *Lemma* 2.2.1, we get

$$\ell(G \circ K_2) \leq 2\ell(G) + \kappa_G$$

where κ_G is the size of the maximum clique of G. □

Theorem 2.3.19. *Let G be a word-representable graph with the minimum length of its word-representant $\ell(G)$. Then, the minimum length of the word-representants of the graph $G \circ K_n$,*

$$\ell(G \circ K_n) \leq n\ell(G) + (n-1)\kappa_G.$$

where κ_G is the size of the maximum clique of the graph G.

Proof. Let w_G be a minimum length word-representant of the graph G. Let $V(K_n) = \{r, 1, 2, \ldots, n-1\}$, where r is the rooted vertex. Therefore, $w_{K_n} = r12\ldots(n-1)$ represents the graph K_n. Let $x^i \in V(G) \times V(K_n)$, where $x \in V(G)$ and $i \in V(K_n)$. We claim that the word $w_{G \circ K_n} = h(w_G)$ represents the graph $G \circ K_n$, where

If $O_{w_G}(x) = 1$

$$h(x, 1) = x^1\ x^2\ \ldots\ x^{n-1}\ x^r\ x^1\ x^2\ \ldots\ x^{n-1}$$

If $O_{w_G}(x) \geq 2$,

$$h(x, i) = \begin{cases} x^r & \text{if } i = 1, \\ x^1\ x^2\ \ldots\ x^{n-1}\ x^r\ x^1\ x^2\ \ldots\ x^{n-1} & \text{if } i = 2, \\ x^r\ x^1\ \ldots\ x^{n-1} & \text{if } i \geq 3. \end{cases}$$

By definition of the rooted product of graphs, it suffices to show that two vertices, x^i

and y^j alternate in $w_{G\circ K_n}$ if and only if either $i = j = r$ and x and y alternate in w_G or $x = y$ and i and j alternate in w_{K_n}.

Suppose $i = j = r$. Then,

$$w_{G\circ K_n}|_{\{x^r,y^r\}} = h(w_G)|_{\{x^r,y^r\}}$$

Since for all $x \in V(G)$, x^r occurs exactly once for every occurrence of x in w_G, x^r and y^r alternate in $w_{G\circ K_n}$ if and only if x and y alternate in w_G.

Suppose $i \neq j$ and $x = y$. Without loss of generality, assume that $i < j$. Then,

If $i \neq j \neq r$,

$$w_{G\circ K_n}|_{\{x^i,x^j\}} = h(w_G)|_{\{x^i,x^j\}} = \begin{cases} x^i\ x^j\ x^i\ x^j & \text{if } O_{w_G}(x) = 1, \\ x^i\ x^j\ x^i\ x^j\ \ldots & \text{if } O_{w_G}(x) \geq 2. \end{cases}$$

If $i \neq j = r$,

$$w_{G\circ K_n}|_{\{x^i,x^r\}} = h(w_G)|_{\{x^i,x^r\}} = \begin{cases} x^i\ x^r\ x^i & \text{if } O_{w_G}(x) = 1, \\ x^r\ x^i\ x^r\ x^i\ x^r \ldots & \text{if } O_{w_G}(x) \geq 2. \end{cases}$$

Hence, x^i and x^j alternate in $w_{G\circ K_n}$ as i and j alternate in w_{K_n}.

Suppose $i \neq j$ and $x \neq y$. Without loss of generality, assume that x occurs to the left of y in $\pi(w_G)$. Then,

$$w_{G\circ K_n}|_{\{x^i,y^j\}} = h(w_G)|_{\{x^i,y^j\}} = x^i\ x^i\ y^j\ y^j\ \ldots$$

Hence, x^i and y^j do not alternate in $w_{G\circ K_n}$.

Suppose $i = j \neq r$ and $x \neq y$. Without loss of generality, assume that x occurs to the left of y in $\pi(w_G)$. Then,

$$w_{G\circ K_n}|_{\{x^i,y^i\}} = h(w_G)|_{\{x^i,y^i\}} = x^i\ x^i\ y^i\ y^i\ \ldots$$

Hence, x^i and y^i do not alternate in $w_{G\circ K_n}$.

Therefore, $w_{G\circ K_n}$ represents the graph $G \circ K_n$. Hence,

$$\begin{aligned}
\ell(G\circ K_n) &\leq |w_{G\circ K_n}| \\
&= \sum_{i=2}^{\mathcal{R}(G)} in|O(w_G,i)| + (2n-1)|O(w_G,1)| \\
&= n\sum_{i=1}^{\mathcal{R}(G)} i|O(w_G,i)| + (n-1)|O(w_G,1)| \\
&= n\ell(G) + (n-1)|O(w_G,1)|.
\end{aligned}$$

Hence, from *Lemma* 2.2.1, we get

$$\ell(G\circ K_n) \leq n\ell(G) + (n-1)\kappa_G.$$

where κ_G is the size of the maximum clique of G. □

Remark 6. By *Theorem* 2.3.16, we know that if G is $k(\geq 2)$-representable, then $\ell(G\circ K_n) \leq nk|G|$. Thus, similar to the Cartesian product, we compare whether the bound provided by *Theorem* 2.3.19 is tighter than the previous bound.

As we look into the proof of the *Theorem* 2.3.19, the bound obtained is

$$\ell(G\circ K_n) \leq n\ell(G) + (n-1)|O(w_G,1)|$$

where w_G is a minimum length word-representant of G.

Let us consider

$$nk|G| - (n\ell(G) + (n-1)|O(w_G,1)|) = n(k|G| - \ell(G)) - (n-1)|O(w_G,1)|$$

Since $k \geq 2$, it is evident that $k|G| - \ell(G) \geq |O(w_G,1)|$. Therefore,

$$nk|G| > n\ell(G) + (n-1)|O(w_G,1)|$$

Hence, the bound provided in *Theorem* 2.3.19 is tighter than the bound $\ell(G\circ K_n) \leq nk|G|$. We are using *Lemma* 2.2.1, $|O(w_G,1)| \leq \kappa_G$, to get the bound as a function of the size of the maximal clique of the graph.

Theorem 2.3.20. *Let G and H be two word-representable graphs, with minimum length of their word-representants $\ell(G)$ and $\ell(H)$, respectively. Then, the minimum length of the word-representants of the graph $G \circ H$,*

$$\ell(G \circ H) \leq |H|\ell(G) + |G|(\ell(H) + |H| - 1) + |H|\kappa_G$$

where κ_G is the size of the maximum clique of the graph G.

Proof. Let w_G and w_H be minimum length word-representants of the graphs G and H, respectively. Let $V(H) = \{v_1, v_2, \ldots, v_r, \ldots, v_n\}$, where v_r is the rooted vertex. Let $\pi(w_H) = v_1v_2 \ldots v_r \ldots v_n$ be the initial permutation of the word w_H. Let $x^{v_i} \in V(G) \times V(H)$, where $x \in V(G)$. Let u be the word obtained from w_G by removing all the copies of the letters appearing in w_G more than one time.

Claim: uw_G also represents the graph G.

By the definition of the word u, it contains the letters from the set $O(w_G, 1)$. Therefore, for all x, $y \in V(G) \setminus \{O(w_G, 1)\}$, x and y alternate in uw_G if and only if they alternate in w_G. If x, $y \in O(w_G, 1)$, x occurs to the left of y in u if and only if x occurs to the left of y in w_G. Therefore, x and y alternate in uw_G. Now, without loss of generality, assume that $x \in O(w_G, 1)$ and $y \in V(G) \setminus O(w_G, 1)$. Hence, $O_{w_G}(y) \geq 2$. If x and y alternate in w_G, then $uw_G|_{\{x,y\}} = xyxy$. If x and y do not alternate in w_G, they do not also alternate in uw_G. Therefore, uw_G also represents the graph G.

Let us define an occurrence-based function as follows:

$$h(x, i) = \begin{cases} x^{v_{r+1}} \ldots x^{v_n} x^{v_1} \ldots x^{v_r} \ldots x^{v_n} x^{v_1} \ldots x^{v_{r-1}} & i = 1, \\ x^{v_r} \ldots x^{v_n} & i = 2, \\ x^{v_1} \ldots x^{v_r} \ldots x^{v_n} & i \geq 3. \end{cases}$$

Therefore,

$$\begin{aligned} |h(uw_G)| &= n|O(w_G, 1)| + \sum_{i=1}^{\mathcal{R}(G)} ((i+1)n - r)|O(w_G, i)| \\ &= n|O(w_G, 1)| + (n - r)|G| + n\ell(G). \end{aligned}$$

Since $r \geq 1$,

$$|h(uw_G)| \leq n|O(w_G, 1)| + n\ell(G) + |G|(n-1).$$

We claim that the word $w_{G\circ H} = h(uw_G)J^{w_H}(\sigma(w_G))$, represents the graph $G \circ H$. By the definition of the rooted product of graphs, it suffices to show that two vertices, x^{v_i} and y^{v_j} alternate in $w_{G\circ H}$ if and only if either $i = j = r$ and x and y alternate in w_G or $x = y$ and v_i and v_j alternate in w_H.

Suppose $i = j = r$. Then,

$$w_{G\circ H}|_{\{x^{v_r},y^{v_r}\}} = h(uw_G)J^{w_H}(\sigma(w_G))|_{\{x^{v_r},y^{v_r}\}} = h(uw_G)|_{\{x^{v_r},y^{v_r}\}}J^{w_H|_{v_r}}(\sigma(w_G)|_{\{x,y\}})$$

For all $x \in V(G)$, x^{v_r} occurs exactly once in $h(w_G)$ in place of every occurrence of x in w_G. Hence, $h(uw_G)|_{\{x^{v_r},y^{v_r}\}} = J^{v_r}(uw_G|_{\{x,y\}})$. Therefore,

$$w_{G\circ H}|_{\{x^{v_r},y^{v_r}\}} = J^{v_r}(uw_G|_{\{x,y\}})J^{w_H|_{v_r}}(\sigma(w_G)|_{\{x,y\}})$$

Hence, x^{v_r} and y^{v_r} alternate in $w_{G\circ H}$ if and only if x and y alternate in w_G.

Suppose $x = y$ and $i \neq j$. Without loss of generality, assume that $i < j$. Then,

$$w_{G\circ H}|_{\{x^{v_i},x^{v_j}\}} = h(uw_G)J^{w_H}(\sigma(w_G))|_{\{x^{v_i},x^{v_j}\}} = h(uw_G)|_{\{x^{v_i},x^{v_j}\}}J^{w_H|_{\{v_i,v_j\}}}(x)$$

If $i < j < r$ or $r < i < j$,

$$h(uw_G)|_{\{x^{v_i},x^{v_j}\}} = x^{v_i}\ x^{v_j}\ x^{v_i}\ x^{v_j}\ \dots\ x^{v_i}\ x^{v_j}.$$

If $i < r < j$,

$$h(uw_G)|_{\{x^{v_i},x^{v_j}\}} = x^{v_j}\ x^{v_i}\ x^{v_j}\ x^{v_i}\ \dots\ x^{v_i}\ x^{v_j}.$$

If $i < j = r$,

$$h(uw_G)|_{\{x^{v_i},x^{v_r}\}} = x^{v_i}\ x^{v_r}\ x^{v_i}\ x^{v_r}\ \dots\ x^{v_i}\ x^{v_r}.$$

If $r = i < j$,

$$h(uw_G)|_{\{x^{v_r},x^{v_j}\}} = x^{v_j}\ x^{v_r}\ x^{v_j}\ x^{v_r}\ \dots\ x^{v_r}\ x^{v_j}.$$

Therefore, in all the cases mentioned above, $w_{G\circ H}|_{\{x^{v_i},x^{v_j}\}} = (\dots\ x^{v_i}\ x^{v_j}\ x^{v_i}\ x^{v_j}\ )J^{w_H|_{\{v_i,v_j\}}}(x)$. Since $i < j$, $J^{w_H|_{\{v_i,v_j\}}}(x)$ starts with x^{v_i}. Therefore, x^{v_i} and x^{v_j} alternate in $w_{G\circ H}$ if and only if v_i and v_j alternate in w_H.

Suppose $i \neq j$ and $x \neq y$. Without loss of generality, assume that $i < j$, and x occurs to the left of y in $\pi(w_G)$. Then,

$$w_{G\circ H}|_{\{x^{v_i},y^{v_j}\}} = h(uw_G)|_{\{x^{v_i},y^{v_j}\}} J^{w_H|_{\{v_i,v_j\}}}(\sigma(w_G)|_{\{x,y\}})$$

If $i < j < r$ or $r < i < j$ or $i < r < j$,

$$h(uw_G)|_{\{x^{v_i},y^{v_j}\}} = x^{v_i}\ x^{v_i}\ y^{v_j}\ y^{v_j}\ \ldots$$

If $i < j = r$,

$$h(uw_G)|_{\{x^{v_i},y^{v_r}\}} = x^{v_i}\ x^{v_i}\ y^{v_r}\ \ldots$$

If $r = i < j$,

$$h(uw_G)|_{\{x^{v_r},y^{v_j}\}} = x^{v_r}\ y^{v_j}\ y^{v_j}\ \ldots$$

Therefore, in all the cases mentioned above, x^{v_i} and y^{v_j} do not alternate in $w_{G\circ H}$.

Suppose $i = j \neq r$ and $x \neq y$. Without loss of generality, assume that x occurs to the left of y in $\pi(w_G)$. Then,

$$w_{G\circ H}|_{\{x^{v_i},y^{v_i}\}} = h(uw_G)|_{\{x^{v_i},y^{v_i}\}} J^{w_H|_{v_i}}(\sigma(w_G)|_{\{x,y\}}) = (x^{v_i}\ x^{v_i}\ y^{v_i}\ y^{v_i}\ \ldots) J^{w_H|_{v_i}}(\sigma(w_G)|_{\{x,y\}})$$

Therefore, x^{v_i} and y^{v_i} do not alternate in $w_{G\circ H}$.

Hence, $w_{G\circ H}$ represents the graph $G \circ H$. Therefore, as $|J^{w_H}(\sigma(w_G))| = |G|\ell(H)$,

$$\begin{aligned}\ell(G\circ H) &\leq \ell(w_{G\circ H})\\ &= |h(uw_G)| + |J^{w_H}(\sigma(w_G))|\\ &= n|O(w_G,1)| + n\ell(G) + |G|(n-1) + |G|\ell(H)\end{aligned}$$

Hence from *Lemma* 2.2.1, we get

$$\ell(G\circ H) \leq n\kappa_G + n\ell(G) + |G|(n - 1 + \ell(H))$$

where κ_G is the size of the maximum clique of the graph G. □

Remark 7. The bound provided in the *Theorem* 2.3.16 is not sharp when compared to the bound $\ell(G \circ H) \leq k|G||H|$, where $k = \max\{\mathcal{R}(G), \mathcal{R}(H)\}$. Here are the following reasons:

- As we saw in the *Definition* 2.3.15, rooted product $G \circ H$ involves gluing two graphs G and H in a root vertex $|G|$ times. The minimum length word-representant of a graph obtained by gluing two graphs in a vertex depends on the number of occurrences of the glued vertices in their respective minimum length word-representants. In the case of $G \circ K_2$ and $G \circ K_n$, the number of occurrences of the root vertex in a minimum length word-representant of K_2 or K_n is precisely 1, which is more than the number of occurrences of every vertex of the graph G in its minimum length word-representant. Hence, we could give a tighter upper bound for $\ell(G \circ K_2)$ and $\ell(G \circ K_n)$. However, in the case of $G \circ H$, the number of occurrences of the rooted vertex in a minimum length word-representant of the graph H can be more or less than the number of occurrences of the vertices of G in its minimum length word-representant. Therefore, in order to give a general upper bound of $\ell(G \circ H)$, we obtained the bound given in *Theorem* 2.3.20.

- Moreover, in [32], the proof of the *Theorem 5.4.2* uses cyclic shift property (see *Proposition* 3) of uniform word-representants of graphs G and H to establish the representation number of the graph obtained by gluing G and H in a vertex, through which the bound $\ell(G \circ H) \leq k|G||H|$ is obtained. However, we aim to use the minimum length word-representants of the graphs G and H to find an upper bound of $\ell(G \circ H)$. We know that the cyclic shift property is applicable only for uniform words. Therefore, in order to give an upper bound on $\ell(G \circ H)$ as a function of $\ell(G)$ and $\ell(H)$, we obtained the bound given in *Theorem* 2.3.20.

3

On semi-transitive orientability of circulant graphs

3.1 Introduction

Among the various classes of graphs, *circulant graphs* occupy a prominent position owing to their high degree of symmetry, regularity, and diverse applications in network design, coding theory, and parallel computing. The study of circulant graphs dates back to Foster (1932), with the term itself derived from circulant matrices. Circulant graphs can be viewed as Cayley graphs on cyclic groups Z_n; consequently, they are vertex-transitive, as every pair of vertices can be mapped to each other via an automorphism. These graphs are defined by a set of "jump" distances, and their structural properties have been thoroughly examined. Despite substantial progress in understanding their connectivity, planarity, and factorization, the questions of their word-representability and semi-transitive orientability remain only partially resolved. Circulant graphs also form a notable subclass of Toeplitz graphs, for which several results concerning semi-transitivity have been reported in [9].

Kitaev and Pyatkin [34] established that all 4-regular circulant graphs are semi-transitive, but left open the question of whether circulant graphs with consecutive jumps (e.g., $C(n; t, t+1, \ldots, k)$) are semi-transitive.

This chapter contributes to the theory of word-representable and semi-transitive circulant graphs in several significant ways. First, regarding negative results on semi-transitivity, we prove that $C(n; t, t+1, \ldots, 2t)$ is not semi-transitive for $2 < \frac{n+1}{5} \le t < \frac{n-1}{4}$ (*Theorem* 3.3.1), thereby resolving a problem posed by Kitaev and Pyatkin in [34]. This result demonstrates that circulant graphs with consecutive jumps need not be semi-transitive.

Second, we present constructive results identifying classes of semi-transitive circulant graphs. In particular, we show that $C(n; a_1, a_2, \ldots, a_k)$ is semi-transitive whenever $a_1 \ge \frac{n+1}{4}$ (*Theorem* 3.3.2), and that $C(n; t, t+1, \ldots, \lfloor \frac{n}{2} \rfloor)$ is semi-transitive for all $1 \le t < \frac{n}{2}$ (*Theorem* 3.3.3). These results partially characterize the family of semi-transitive circulant graphs, significantly narrowing the range of unresolved cases.

Third, we investigate the representation numbers of k-regular circulant graphs. For 3-regular circulant graphs, we establish that $\mathcal{R}(G) \le 3$ (*Theorem* 3.4.2) and show that most such graphs are not 2-word-representable (*Theorem* 3.4.3). For 4-regular circulant graphs, we derive an upper bound $\mathcal{R}(G) \le 4$ when $\frac{n}{3} \le a < \frac{n}{2}$ (*Corollary* 5). These bounds contribute to the broader problem of determining the representation numbers of regular graphs. The chapter concludes with several open problems that outline promising directions for future research.

The remainder of this chapter is structured as follows. Section 3.2 reviews the definition and basic properties of circulant graphs. Section 3.3 presents our main results concerning the semi-transitive orientability of circulant graphs, and Section 3.4 establishes bounds on the representation numbers of k-regular circulant graphs.

3.2 Circulant graph

A *circulant graph* $C(n;R)$, for a set $R=\{a_1,a_2,\ldots,a_k\}$, is defined as the graph with vertex set $\{0,1,\ldots,n-1\}$ and edge set

$$E=\{\, ij \mid (i-j) \bmod n \text{ or } (j-i) \bmod n \in \{a_1,a_2,\ldots,a_k\}\,\},$$

where $0<a_1<a_2<\cdots<a_k<(n+1)/2$.

The circulant graph $C(n;a_1,a_2,\ldots,a_k)$ is regular of degree d, where

$$d=\begin{cases} 2k, & \text{if } a_k\neq n/2,\\ 2k-1, & \text{otherwise.}\end{cases}$$

Let n and r be positive integers with $n\geq 2$ and $r<n/2$. Then $C(n;r)$ consists of a collection of disjoint cycles. If $d=$

$gcd(n,r)$, there are exactly d such cycles, each of length n/d. We say that each of these cycles has *period* r, *length* n/d, and *rotation* r/d.

In this chapter, we make use of the following established results on circulant graphs.

Theorem 3.2.1 ([2, 46]). *A circulant graph $C(n;R)$, where $R=\{a_1,a_2,\ldots,a_k\}$, is connected if and only if $gcd(n,a_1,a_2,\ldots,a_k)=1$. Moreover, if $d=gcd(n,a_1,a_2,\ldots,a_k)$, then*

$$C(n;a_1,a_2,\ldots,a_k)\equiv d\,C\Big(\frac{n}{d};\frac{a_1}{d},\frac{a_2}{d},\ldots,\frac{a_k}{d}\Big),$$

that is, $C(n;a_1,a_2,\ldots,a_k)$ is isomorphic to d disjoint copies of $C\big(\frac{n}{d};\frac{a_1}{d},\frac{a_2}{d},\ldots,\frac{a_k}{d}\big)$.

Analogous to integers, circulant graphs can also be factorized uniquely into a Cartesian product of prime graphs. This result was established by Vilfred in [47]. In this chapter, we make use of the factorization of circulant graphs in some of our results. Hence, we require the following theorem.

Theorem 3.2.2 ([47]). *For $n\in\mathbb{N}$ and a set $R=\{a_1,a_2,\ldots,a_k\}$, the Cartesian product $P_2\,\square\,C(2n+1;R)\equiv C(2(2n+1);2R\cup\{2n+1\})\equiv C(2(2n+1);2dR\cup\{2n+1\})$, where* $\gcd(2(2n+1),d)=1$.

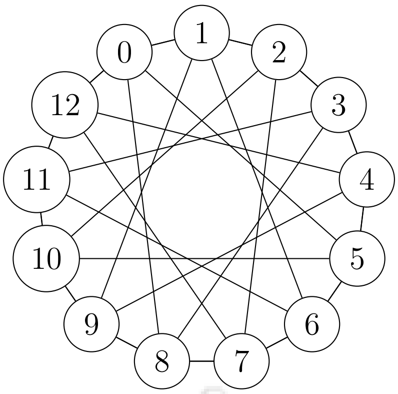


Figure 3.1: A semi-transitively orientable circulant graph $C(13; 1, 5)$

The word-representability of circulant graphs was studied by Kitaev and Pyatkin in [34]. They established several results concerning 4-regular circulant graphs. An example of a semi-transitive circulant graph is illustrated in *Figure* 3.1.

Theorem 3.2.3 ([34]). *The circulant graph* $C(13; 1, 5)$ *is semi-transitive.*

Lemma 3.2.4 ([34]). *A circulant graph* $C(n; 1, 2)$ *is semi-transitive for each* $n \geq 6$.

Theorem 3.2.5 ([34]). *Each* 4*-regular circulant graph is semi-transitive.*

Furthermore, Kitaev and Pyatkin provided an example of a non-semi-transitive circulant graph, namely $C(14; 1, 3, 4, 5)$. This example demonstrates that a circulant graph need not be semi-transitive. The following section presents results concerning both semi-transitive and non-semi-transitive circulant graphs.

3.3 Semi-transitive orientability of circulant graphs

The semi-transitive orientability of circulant graphs was first investigated by Kitaev and Pyatkin in [34]. They proved that every 4-regular circulant graph is semi-transitive. Furthermore, they posed an open problem asking whether $C(n; t, t+1, \ldots, k)$, for integers t and k satisfying $k - t > 1$, is semi-transitive.

In this section, we establish conditions under which $C(n; t, t+1, \ldots, k)$ is not semi-transitive (*Theorem* 3.3.1). In addition, with the aim of characterizing semi-transitive circulant graphs, we provide partial characterizations of this class in *Theorem* 3.3.2 and *Theorem* 3.3.3.

Theorem 3.3.1. *$C(n; t, t+1, \ldots, 2t)$ is not semi-transitive for* $2 < \frac{n+1}{5} \le t < \frac{n-1}{4}$.

Proof. By the definition of a circulant graph, two vertices i and j are adjacent if and only if $(i-j) \pmod n$ or $(j-i) \pmod n$ are in $\{t, t+1, \ldots, 2t\}$. Thus, i and j are adjacent if and only if either of the following cases holds for all $i > j$.

1. $t \le i - j \le 2t$.
2. $n - 2t \le i - j \le n - t$.

Consider an induced subgraph, H, with vertices $\{0, t-1, t, 2t-1, 2t+1, n-t\}$. We claim that H is isomorphic to W_5. According to *Remark* 1, this implies that the graph $C(n; t, t+1, \ldots, 2t)$ is not semi-transitive.

Consider the following table: $M = \{m_{ij}\}$ for all $i, j \in V(H)$, where $m_{ij} = |i - j|$.

i/j	**0**	**t-1**	**t**	**2t - 1**	**2t +1**	**n-t**
0	0	$t - 1$	t	$2t - 1$	$2t + 1$	$n - t$
t-1	$t - 1$	0	1	t	$t + 2$	$n - 2t + 1$
t	t	1	0	$t - 1$	$t + 1$	$n - 2t$
2t-1	$2t - 1$	t	$t - 1$	0	2	$n - 3t + 1$
2t +1	$2t + 1$	$t + 2$	$t + 1$	2	0	$n - 3t - 1$
n-t	$n - t$	$n - 2t + 1$	$n - 2t$	$n - 3t + 1$	$n - 3t - 1$	0

Since $\frac{n+1}{5} \le t < \frac{n-1}{4}$, we have $t < n - 3t - 1 < n - 3t + 1 \le 2t$. Hence, the adjacency matrix, $\{h_{ij}\}$, of the induced subgraph H can be written as follows:

$$\{h_{ij}\} = \begin{cases} 1 & \text{if } t \le m_{ij} \le 2t \text{ or } n-2t \le m_{ij} \le n-t, \\ 0 & \text{otherwise.} \end{cases} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Hence, the induced subgraph, H, formed by the above adjacency matrix, is isomorphic to W_5, as shown in *Figure* 3.2. □

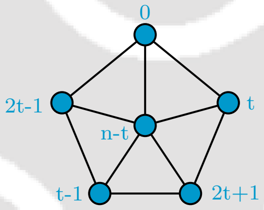


Figure 3.2: Induced subgraph $H \equiv W_5$

The previous result shows that $C(n; t, t+1, \ldots, 2t)$ is not semi-transitive for $2 < \frac{n+1}{5} \le t < \frac{n-1}{4}$. Hence, a natural question arises regarding semi-transitive orientability of $C(n; t, t+1, \ldots, 2t)$ for $t \ge \frac{n-1}{4}$. Interestingly, the next result positively answers this question for a more general case.

Theorem 3.3.2. *$C(n; a_1, a_2, \ldots, a_k)$ is semi-transitive for all $a_1 \ge \frac{n+1}{4}$*

Proof. Consider a circulant graph $G = C(n; a_1, a_2, \ldots, a_k)$ with the vertex set $\{0, 1, \ldots, n-1\}$. Orient the edge $ij \in E(G)$ as $i \to j$ for all $i < j$, where $i, j \in V(G)$. We claim that the given orientation is semi-transitive. It is easy to see that the orientation is acyclic. Suppose that there is a shortcut $v_0 \to v_1 \to \ldots \to v_m$ with a shortcutting arc $v_0 \to v_m$ where $m \ge 3$.

Note that if $i \to j$, then $j > i$ and $j \in \{i+a_1, \ldots, i+a_k, i+n-a_k, \ldots, i+n-a_1\}$. Since $a_1 \ge \frac{n+1}{4}$ and $a_k < \frac{n+1}{2}$, we have either $i + \frac{n+1}{4} \le j < i + \frac{n+1}{2}$ or $i + \frac{n-1}{2} < j \le i + \frac{3n-1}{4}$.

Now since $v_l \to v_{l+1}$ for $l \in \{0, 1, \ldots, m-1\}$, we have either $v_l + \frac{n+1}{4} \leq v_{l+1} < v_l + \frac{n+1}{2}$ or $v_l + \frac{n-1}{2} < v_{l+1} \leq v_l + \frac{3n-1}{4}$.

From this, we can see that $v_{l+1} \geq v_l + \frac{n+1}{4}$ for $l \in \{0, 1, \ldots, m-1\}$. By solving this recursion, we get $v_l \geq v_0 + \frac{l(n+1)}{4}$. Since $m \geq 3$, we have $v_m \geq v_0 + \frac{3(n+1)}{4}$.

Since $v_0 \to v_m$, we have $v_m \leq v_0 + \frac{3n-1}{4} < v_0 + \frac{3(n+1)}{4} \leq v_m$, which is a contradiction. Hence, the given orientation is semi-transitive. □

Theorem 3.3.3. *$C(n; t, t+1, \ldots, \lfloor \frac{n}{2} \rfloor)$ is semi-transitive for any $t \in \{1, 2, \ldots, \lfloor \frac{n}{2} \rfloor\}$*

Proof. Consider a circulant graph $G = C(n; t, t+1, \ldots, \lfloor \frac{n}{2} \rfloor)$ with the vertex set $\{0, 1, \ldots, n-1\}$. Orient the edge $ij \in E(G)$ as $i \to j$ for all $i < j$, where $i, j \in V(G)$. We claim that the given orientation is semi-transitive. It is easy to see that the orientation is acyclic. Suppose that there is a shortcut $v_0 \to v_1 \to \ldots \to v_m$ with a shortcutting arc $v_0 \to v_m$ where $m \geq 3$.

Note that if $i \to j$, then $j > i$ and $j \in \{i + t, \ldots, i + n - t\}$. Now since $v_l \to v_{l+1}$ for $l = \{0, 1, \ldots, m-1\}$, we have $v_l + t \leq v_{l+1} \leq v_l \ + n - t$. Hence, by solving this recursion, we get $v_0 + lt \leq v_l \leq v_0 + l(n-t)$.

Consider v_x and v_y, where $0 \leq x < y \leq m$. Since $v_0 \to v_m$, we have $v_0 + t \leq v_m \leq v_0 + n - t$. Hence, we have, $v_x + t < v_{x+1} \leq v_y < v_m \leq v_0 + n - t < v_x + n - t$. As a result, $v_x \to v_y$ for all $0 \leq x < y \leq m$. Hence, the vertices $\{v_0, v_1, \ldots, v_m\}$ induce a transitively oriented clique, which is a contradiction. Therefore, the given orientation is semi-transitive. □

3.4 Representation number for some circulant graphs

In this section, we determine the representation numbers of certain circulant graphs. In [34], Kitaev and Pyatkin posed a problem asking whether $C(n; 1, 2, \ldots, k)$ is semi-transitive. The following result provides a positive answer to this problem and establishes the corresponding representation number.

Throughout this section, for any positive integer i, we denote i mod n by $(i)_n$. Moreover, for any two words A and u, the notation $\underbrace{A}_{u}$ indicates that u is a subword of A.

Theorem 3.4.1. *$C(n; 1, 2, \ldots, k)$ is 2-word-representable.*

Proof. Consider a circulant graph $G = C(n; 1, 2, \ldots, k)$ with the vertex set $V(G) = \{0, 1, \ldots, n-1\}$. Define a morphism $h : V(G)^* \to V(G)^*$ as follows:

$$h(i) = \ (i)_n \ (i-k)_n, \ \textit{for } i \in V(G).$$

We claim that the word $w = h(u)$ represents G, where $u = 0\ 1\ 2\ \ldots\ (n-1)$. The word is of the form $w = 0\ (n-k)_n\ 1\ (n-k+1)_n\ 2\ \ldots\ (n-1)_n\ k\ 0\ (k+1)_n\ 1\ \ldots\ (n-1)_n\ (n-k-1)_n$. By the definition of a circulant graph, i is adjacent to j if and only if $(j-i) \equiv t \pmod{n}$, where $1 \leq t \leq k$. Let $u_i = (i)_n\ (i+1)_n\ \ldots\ (n-1)\ 0\ 1\ \ldots\ (i-1)_n$, where $0 \leq i \leq n-1$. By *Proposition* 3, the word $h(u_i)$ represents the graph G for all $0 \leq i \leq n-1$. As a result, i and j alternate in w if and only if they alternate in $h(u_i)$ for all $0 \leq i \leq n-1$. Therefore, it is enough to prove that i and $j \in \{(i-k)_n, (i-k+1)_n, \ldots, (i-1)_n, (i+1)_n, (i+2)_n, \ldots, (i+k)_n\}$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

By the definition of morphism h, we have

$$h(u_i) = (i)_n\ (i-k)_n\ (i+1)_n\ (i-k+1)_n\ \ldots\ (i-k-1)_n\ (i-1)_n\ (i+k)_n\ (i)_n \ldots (i-1)_n\ (i-k-1)_n.$$

Note that the letters occurring exactly once between the two occurrences of the letter i are $\{(i-k)_n, (i-k+1)_n, \ldots, (i-1)_n, (i+1)_n, (i+2)_n, \ldots, (i+k)_n\} = N_G(i)$. Since w is 2-uniform, i and $j \in N_G(i)$ alternate in $h(u_i)$. Therefore, w represents the graph G. □

Every 2-regular circulant graph, i.e., cycle graph, is 2-word-representable. Is this true for k-regular circulant graphs, where $k \geq 3$? Hence, the following problem is posed:

Problem 3.4.1. Are word-representable k-regular circulant graphs k-word-representable?

3.4.1 3-regular circulant graphs

In this subsection, we will answer *Problem* 3.4.1 positively for $k = 3$.

Theorem 3.4.2. *Let* $G \equiv C(2n; a, n)$ *be a* 3*-regular connected circulant graph such that* $gcd(a, 2n) = 1$. *Then the representation number of* G *satisfies* $\mathcal{R}(G) \leq 3$.

Proof. Consider a 3-regular connected circulant graph $G \equiv C(2n; a, n)$ with $\gcd(a, 2n) = 1$. Let the vertex set of the graph be $V(G) = \{0, 1, \ldots, 2n-1\}$. By the definition of a circulant graph, a vertex i is adjacent to $j \in \{(i+a)_{2n}, (i-a)_{2n}, (i+n)_{2n}\}$. Define a morphism $h : V(G)^* \to V(G)^*$ as follows:

$$h(i) = (i)_{2n}\ (i-a)_{2n}\ (i+n)_{2n},\ \textit{for } i \in V(G).$$

We claim that the word $w = h(u)$ represents G, where $u = 0\ (a)_{2n}\ (2a)_{2n}\ (3a)_{2n}\ \ldots ((2n-1)a)_{2n}$. The word is of the form $w = 0\ (2n-a)_{2n}\ (n)_{2n}\ (a)_{2n}\ 0\ (a+n)_{2n}\ \ldots (2n-a)_{2n}\ (2n-2a)_{2n}\ (n-a)_{2n}$. Note that a vertex i occurs once in every $h(i)$, $h((i+a)_{2n})$ and $h((i+n)_{2n})$. Therefore, each letter occurs three times in w.

Let $u_i = (i)_{2n}\ (i+a)_{2n}\ (i+2a)_{2n}\ \ldots\ (i+(2n-1)a)_{2n}$, where $0 \le i \le 2n-1$. By *Proposition* 3, the word $h(u_i)$ represents the graph G for all $0 \le i \le 2n-1$. As a result, i and j alternate in w if and only if they alternate in $h(u_i)$ for all $0 \le i \le 2n-1$. Therefore, it is enough to prove that i and $j \in \{(i+a)_{2n}, (i-a)_{2n}, (i+n)_{2n}\}$ alternate in $h(u_i)$ for all $0 \le i \le 2n-1$.

By the definition of morphism h, we have

$$h(u_i) = (i)_{2n}\ (i-a)_{2n}\ (i+n)_{2n}\ (i+a)_{2n}\ (i)_{2n}\ \ldots\ (i-a)_{2n}\ (i-2a)_{2n}\ (i-a+n)_{2n}.$$

Note that i and $j \in V(G) \setminus \{(i+a)_{2n}, (i-a)_{2n}, (i+n)_{2n}\}$ do not alternate in w. Now it is enough to prove that i and $j \in \{(i+a)_{2n}, (i-a)_{2n}, (i+n)_{2n}\}$ alternate in w. Recall that we say $x \prec_u y$ if x occurs to the left of y in the word u.

Assuming $(i+n)_{2n} \prec_{u_i} (i+2a)_{2n}$, it follows that $(i+a)_{2n} \equiv (i+n)_{2n}$ since $(i)_{2n} \not\equiv (i+n)_{2n}$. This leads to $n = a$, which is a contradiction. Similarly, if $(i+n)_{2n} \equiv (i+2a)_{2n}$, then $n = 2a$, which also results in a contradiction. Therefore, for all $0 \le i \le 2n-1$, we must have $(i+2a)_{2n} \prec_{u_i} (i+n)_{2n}$.

Now, suppose $(i-a)_{2n} \equiv (i+n)_{2n}$. This would imply $n = a$, again yielding a contradiction. Hence, $(i+n)_{2n} \prec_{u_i} (i-a)_{2n}$.

Combining these results, we conclude that for every $0 \le i \le 2n-1$, the following ordering holds:

$$(i)_{2n} \prec_{u_i} (i+a)_{2n} \prec_{u_i} (i+2a)_{2n} \prec_{u_i} (i+n)_{2n} \prec_{u_i} (i-a)_{2n}.$$

By the definition of morphism h, we get the following cases:

If $(i+2a)_{2n} \not\equiv (i+n-a)_{2n}$ and $(i+n+a)_{2n} \not\equiv (i-a)_{2n}$,

$$h(u_i) = \underbrace{h(i)}_{(i)_{2n}\,(i-a)_{2n}\,(i+n)_{2n}} \underbrace{h((i+a)_{2n})}_{(i+a)_{2n}\,(i)_{2n}} \underbrace{h((i+2a)_{2n})}_{(i+a)_{2n}} \ldots \underbrace{h((i+n-a)_{2n})}_{(i-a)_{2n}} \underbrace{h((i+n)_{2n})}_{(i+n)_{2n}\,(i)_{2n}} \underbrace{h((i+n+a)_{2n})}_{(i+n)_{2n}\,(i+a)_{2n}} \ldots \underbrace{h((i-a)_{2n})}_{(i-a)_{2n}}.$$

If $(i+2a)_{2n} \equiv (i+n-a)_{2n}$ and $(i+n+a)_{2n} \not\equiv (i-a)_{2n}$,

$$h(u_i) = \underbrace{h(i)}_{(i)_{2n}\,(i-a)_{2n}\,(i+n)_{2n}} \underbrace{h((i+a)_{2n})}_{(i+a)_{2n}\,(i)_{2n}} \underbrace{h((i+2a)_{2n})}_{(i+a)_{2n}\,(i-a)_{2n}} \underbrace{h((i+n)_{2n})}_{(i+n)_{2n}\,(i)_{2n}} \underbrace{h((i+n+a)_{2n})}_{(i+n)_{2n}\,(i+a)_{2n}} \ldots \underbrace{h((i-a)_{2n})}_{(i-a)_{2n}}.$$

If $(i+2a)_{2n} \not\equiv (i+n-a)_{2n}$ and $(i+n+a)_{2n} \equiv (i-a)_{2n}$,

$$h(u_i) = \underbrace{h(i)}_{(i)_{2n}\,(i-a)_{2n}\,(i+n)_{2n}} \underbrace{h((i+a)_{2n})}_{(i+a)_{2n}\,(i)_{2n}} \underbrace{h((i+2a)_{2n})}_{(i+a)_{2n}} \ldots \underbrace{h((i+n-a)_{2n})}_{(i-a)_{2n}} \underbrace{h((i+n)_{2n})}_{(i+n)_{2n}\,(i)_{2n}} \underbrace{h((i-a)_{2n})}_{(i-a)_{2n}\,(i+n)_{2n}\,(i+a)_{2n}}.$$

If $(i+2a)_{2n} \equiv (i+n-a)_{2n}$ and $(i+n+a)_{2n} \equiv (i-a)_{2n}$,

$$h(u_i) = \underbrace{h(i)}_{(i)_{2n}\,(i-a)_{2n}\,(i+n)_{2n}} \underbrace{h((i+a)_{2n})}_{(i+a)_{2n}\,(i)_{2n}} \underbrace{h((i+2a)_{2n})}_{(i+a)_{2n}\,(i-a)_{2n}} \underbrace{h((i+n)_{2n})}_{(i+n)_{2n}\,(i)_{2n}} \underbrace{h((i-a)_{2n})}_{(i-a)_{2n}\,(i+n)_{2n}\,(i+a)_{2n}}.$$

Hence, for each i, i and $j \in \{(i+a)_{2n}, (i-a)_{2n}, (i+n)_{2n}\}$ alternate in $h(u_i)$. Therefore, w represents the graph G, which implies $\mathcal{R}(G) \leq 3$. ☐

A natural question is whether the circulant graph $C(2n; a, n)$ with $gcd(a, 2n) = 1$ is 2-word-representable. Since $C(4; 1, 2) \equiv K_4$ is 1-word-representable, it suffices to consider the case $n > 2$.

Theorem 3.4.3. *Let $G \equiv C(2n; a, n)$ be a 3-regular connected circulant graph with* $\gcd(a, 2n)$ $= 1$. *Then G is not 2-word-representable for all $a \geq 2$. Moreover, if $a = 1$, then G is not 2-word-representable for all $n > 3$.*

Proof. Consider the graph $G \equiv C(2n; a, n)$ with $\gcd(a, 2n) = 1$. Since G is connected, by *Theorem* 3.2.1, $\gcd(2n, a, n) = \gcd(a, n) = 1$. Let the vertex set $V(G) = \{0, 1, \ldots, 2n-1\}$. By the definition of a circulant graph, a vertex i is adjacent to $j \in \{(i+a)_{2n}, (i-a)_{2n}, (n+i)_{2n}\}$. Note that the set $\{(i+a)_{2n}, (i-a)_{2n}, (n+i)_{2n}\}$ is an independent set because of the following reasons:

- $(i+a) - (i-a) \equiv 2a \pmod{2n}$. Since $\gcd(a, 2n) = 1$, $2a \notin \{a, n\}$. Therefore, $i+a$ and $i-a$ are not adjacent in G.

- $(n+i)-(i-a) \equiv n+a \pmod{2n}$. Since $\gcd(a,2n)=1$, $n+a \notin \{a,n\}$. Therefore, $i-a$ and $n+i$ are not adjacent in G.
- $(n+i)-(i+a) \equiv n-a \pmod{2n}$. Since $\gcd(a,2n)=1$, $n-a \notin \{a,n\}$. Therefore, $i+a$ and $n+i$ are not adjacent in G.

Suppose that G is 2-word representable for $n > 2$. Let w be a 2-uniform word-representant of the graph G with a letter i such that no other letter occurs twice between the two copies of i. The word mentioned above always exists, as its absence implies that the graph is disconnected, which is a contradiction. By *Observation* 2, the letters between the copies of i are $\{i+a, i-a, n+i\}$. Without loss of generality, consider $i=0$. Using the fact that the set $\{(2n-a), a, n\}$ is an independent set by *Proposition* 1 and *Proposition* 3, we only need to consider three cases.

- w is of the form $0\ (2n-a)\ a\ n\ 0 \ldots n \ldots a \ldots (2n-a) \ldots$. Consider the vertex $n-a$.
  - $n-(n-a) \equiv a \pmod{2n}$. Therefore, n and $n-a$ are adjacent in G.
  - $(2n-a)-(n-a) \equiv n \pmod{2n}$. Therefore, $2n-a$ and $n-a$ are adjacent in G.
  - $(n-a)-a \equiv n-2a \pmod{2n}$. Suppose $a \geq 2$. Since $\gcd(a,n)=1$, $n-2a \notin \{a,n\}$. If $a=1$ and $n>3$, $a=1<n-2a<n$.

    Therefore, a and $n-a$ are not adjacent in G.

  If we introduce $n-a$ in the word w, we get either

  $$w = 0\ (2n-a)\ a\ n\ 0 \ldots (n-a) \ldots n \ldots (n-a) \ldots a \ldots (2n-a) \ldots$$

  where $n-a$ is not alternating with $2n-a$, which is a contradiction, or

  $$w = 0\ (2n-a)\ a\ n\ 0 \ldots (n-a) \ldots n \ldots a \ldots (n-a) \ldots (2n-a) \ldots$$

  where $n-a$ is alternating with a, which is also a contradiction.

- w is of the form $0\ a\ (2n-a)\ n\ 0 \ldots n \ldots (2n-a) \ldots a \ldots$ Consider the vertex $(n+a)$.
  - $(n+a)-n \equiv a \pmod{2n}$. Therefore, n and $n+a$ are adjacent in G.

  - $(n+a) - a \equiv n \pmod{2n}$. Therefore, a and $n+a$ are adjacent in G.
  - $(2n-a) - (n+a) \equiv n - 2a \pmod{2n}$. Suppose $a \geq 2$. Since $\gcd(a,n) = 1$, $n - 2a \notin \{a, n\}$. If $a = 1$ and $n > 3$, $a = 1 < n - 2a < n$.

    Therefore, $2n - a$ and $n + a$ are not adjacent in G.

  If we introduce $n+a$ in the word w, we get either

$$0\ a\ (2n-a)\ n\ 0 \ldots (n+a) \ldots n \ldots (n+a) \ldots (2n-a) \ldots a \ldots$$

  where $n+a$ is not alternating with a, which is a contradiction, or

$$0\ a\ (2n-a)\ n\ 0 \ldots (n+a) \ldots n \ldots (2n-a) \ldots a \ldots (n+a) \ldots$$

  where $n+a$ is alternating with $2n-a$, which is also a contradiction.

- w is of the form $0\ (2n-a)\ n\ a\ 0 \ldots a \ldots n \ldots (2n-a) \ldots$. Let us introduce $(n+a)$ in w. We get,

$$w = 0\ (2n-a)\ n\ a\ 0 \ldots (n+a) \ldots a \ldots n \ldots (n+a) \ldots (2n-a) \ldots$$

  If we introduce $n-a$ in w, we get

$$w = 0\ (2n-a)\ n\ a\ 0 \ldots (n+a) \ldots a \ldots (n-a) \ldots n \ldots (n+a) \ldots (2n-a) \ldots (n-a) \ldots$$

  Here, $n+a$ and $n-a$ are alternating. But $(n+a) - (n-a) \equiv 2a \pmod{2n}$, which is a contradiction.

Therefore, G is not 2-word representable for $n > 2$. □

Theorem 3.4.4. *Let $G \equiv C(2n; a, n)$ be a 3-regular connected circulant graph with* $\gcd(a, 2n) \neq 1$. *Then,* $\mathcal{R}(G) = 3$ *for all* $n > 2$.

Proof. Consider a 3-regular connected circulant graph $G \equiv C(2n; a, n)$ with $\gcd(a, 2n) \neq 1$. Since G is connected, by *Theorem* 3.2.1, we have $\gcd(a, n) = 1$. Since $\gcd(a, 2n) \neq 1$, a is even, and n is odd. Therefore, by *Theorem* 3.2.2, $C(2n; a, n) \equiv P_2 \,\square\, C(n; a/2)$. Since $\gcd(a, n) = 1$, we have $\gcd(a/2, n) = 1$. Therefore, $C(2n; a, n) \equiv P_2 \,\square\, C_n = Pr_n$

where Pr_n is a prism graph on n vertices. Since the representation number of prism is 3, $\mathcal{R}(G) = 3$. □

Corollary 3. Let $G \equiv C(2n; a, n)$ be a 3-regular circulant graph. Then, $\mathcal{R}(G) \leq 3$.

Proof. The statement is true for connected 3-regular circulant graph by *Theorem* 3.4.3 and *Theorem* 3.4.4. Suppose $G \equiv C(2n; a, n)$ is not connected. Then, by *Theorem* 3.2.1, $C(2n; a, n) \equiv dC(2n/d; a/d, n/d)$, where $d = \gcd(2n, a, n)$. The representation number of a disconnected graph is bounded above by the maximum representation number of its connected components. Since $\mathcal{R}(C(2n/d; a/d, n/d)) \leq 3$ by *Theorem* 3.4.3 and *Theorem* 3.4.4, we have $\mathcal{R}(G) \leq 3$. □

The Möbius ladder of order n on $2n$ vertices, denoted by M_n, is the simple graph obtained by connecting diametrically opposite vertices in the cycle C_{2n}. It is easy to see that M_n is isomorphic to the circulant graph $C(2n; 1, n)$. The following result follows directly from *Corollary* 3.

Corollary 4. $\mathcal{R}(M_n) = 3$ for all $n > 3$.

3.4.2 4-regular circulant graphs

In this subsection, we partially address *Problem* 3.4.1 for the case $k = 4$. In [49], Yu et al. classified the 4-regular circulant graphs using a Diophantine equation approach. We begin by recalling the definition of Cartesian graph bundles.

Definition 3.4.5. *Let B, F be graphs. A graph G is a Cartesian graph bundle with fibre F over the base graph B if there is a graph map $p : G \to B$ such that for each vertex $v \in V(B), p^{-1}(v) \equiv F$, and for each edge $e \in E(B), p^{-1}(e) \equiv K_2 \square F$. Let $\varphi : E(B) \to Aut(F)$ be a mapping that assigns an automorphism of the graph F to any edge of B. The bundle G is denoted by $G = B \square^{\varphi} F$.*

A Cartesian graph bundle can also be understood as a graph obtained from a base graph by replacing each of its vertices with a copy of a fibre graph, and each of its edges with a matching between the corresponding copies of the fibre at the endpoints of the edge. The edges of the matching define an isomorphism between the copies of the fibre. The following theorem provides a classification of 4-regular circulant graphs.

Theorem 3.4.6 ([49]). For the circulant graph $C(n; k_1, k_2)$, we have

$$C(n; k_1, k_2) \cong \begin{cases} dC(n/d_1; 1, k) & 1 \leq d_1 \leq d_2 \leq n/2, \quad d_1 \mid d_2, \\ d(C_{d_1/\gcd(d_1,d_2)} \square^{\varphi} C_{n/d_1}) & 1 < d_1 < d_2 < n/2, \quad d_1 \nmid d_2, \\ d(K_2 \square C_{n/d_1}) & 1 < d_1 < d_2 = n/2, \quad d_1 \nmid d_2. \end{cases}$$

where $d = \gcd(d_1, d_2)$, $d_i = \gcd(n, k_i)$ *for* $i = 1, 2$, $k = x_0$, φ *is a cyclic* x_0*-shift, and* (x_0, y_0) *is a solution of the linear Diophantine equation*

$$\frac{k_1}{d_1}x - \frac{n}{d_1}y = \frac{k_2}{\gcd(d_1, d_2)}$$

such that $0 \leq x_0 \leq n/d_1 - 1$*. Especially,* φ *is trivial if and only if* $\frac{d_1 d_2}{\gcd(d_1,d_2)} \equiv 0 \pmod{n}$.

Thus, by the above theorem, it suffices to show the following:

- $C(n/d_1; 1, k)$ is 4-representable.
- $C_{d_1/\gcd(d_1,d_2)} \square^{\varphi} C_{n/d_1}$ is 4-representable
- $K_2 \square C_{n/d_1}$ is 4-representable.

The cases in which $d_2 = n/2$ are resolved by *Theorem* 3.4.2 and *Theorem* 3.4.4. It therefore remains to consider the cases where $d_2 < n/2$. The following results provide a partial solution for the 4-regular circulant graph $C(n/d_1; 1, k)$.

Theorem 3.4.7. *Let* $G \equiv C(n; 1, a)$ *be a connected* 4*-regular circulant graph with* $\frac{n}{3} \leq a < \frac{n-1}{2}$ *and* $n > 6$*. Then* $\mathcal{R}(G) \leq 4$.

Proof. Consider a connected 4-regular circulant graph $G \equiv C(n; 1, a)$ with $\frac{n}{3} \leq a < \frac{n-1}{2}$. Let the vertex set $V(G) = \{0, 1, \ldots, n-1\}$. By the definition of a circulant graph, a vertex i is adjacent to $j \in \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$. Define a morphism $h : V(G)^* \to V(G)^*$ as follows:

$$h(i) = (i)_n \ (i-1)_n \ (i+a)_n \ (i-a)_n, \ \textit{for } i \in V(G).$$

We claim that the word $w = h(u)$ represents G, where $u = 012\ldots(n-1)$. Note that a vertex i occurs once in every $h(i)$, $h((i+1)_n)$, $h((i+a)_n)$, and $h((i-a)_n)$. Therefore, each letter occurs exactly four times in w.

Let $u_i = (i)_n\ (i+1)_n\ \ldots\ (n-1)\ 0\ 1\ \ldots\ (i-1)_n$, where $0 \le i \le n-1$. By *Proposition* 3, the word $h(u_i)$ represents the graph G for all $0 \le i \le n-1$. As a result, i and j alternate in w if and only if they alternate in $h(u_i)$ for all $0 \le i \le n-1$.

By the definition of morphism h, we have

$$h(u_i) = (i)_n\ (i-1)_n\ (i+a)_n\ (i-a)_n\ (i+1)_n\ (i)_n\ \ldots\ (i-1)_n\ (i-2)_n\ (i+a-1)_n\ (i-a-1)_n.$$

For any vertex $i \in V(G)$, $h(i)\ h(i+1)$ factor of $h(u_i)$ makes sure that i does not alternate with $j \ \in V(G) \setminus \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$. Therefore, it is enough to prove that i and $j \in \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$ alternate in $h(u_i)$ for all $0 \le i \le n-1$. Recall that we say $x \prec_u y$, if x occurs to the left of y in the word u.

Case (i): In this case, we will prove that $(i)_n$ and $(i-1)_n$ alternate in $h(u_i)$. Since $\frac{n}{3} \le a < \frac{n-1}{2}$ and $n > 6$, $(i+1)_n \prec_{u_i} (i+a-1)_n$ and $(i-a)_n \prec_{u_i} (i-1)_n$. Assuming $(i-a-1)_n \prec_{u_i} (i+a)_n$, it follows that $(n+i-a-1) < (i+a)$. This leads to $n-1 < 2a$, which is a contradiction. Similarly, if $(i-a-1)_n \equiv (i+a)_n$, then $n-1 = 2a$, which also results in a contradiction. Combining these results, we conclude that for every $0 \le i \le n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+a-1)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i-a-1)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i-1)_n}\ \underbrace{h((i+1)_n)}_{(i)_n}\ldots\underbrace{h((i+a-1)_n)}_{(i-1)_n}\underbrace{h((i+a)_n)}_{(i)_n}\ldots\underbrace{h((i-a-1)_n)}_{(i-1)_n}\underbrace{h((i-a)_n)}_{(i)_n}\ldots\underbrace{h((i-1)_n)}_{(i-1)_n}$$

Hence, $(i)_n$ and $(i-1)_n$ alternate in $h(u_i)$ for all $0 \le i \le n-1$.

Case (ii): In this case, we will prove that $(i)_n$ and $(i+1)_n$ alternate in $h(u_i)$. Since $\frac{n}{3} \le a < \frac{n-1}{2}$ and $n > 6$, $(i+2)_n \prec_{u_i} (i+a)_n$. Assuming $(i-a)_n \prec_{u_i} (i+a+1)_n$, it follows that $(n+i-a) < (i+a+1)$. This leads to $n-1 < 2a$, which is a contradiction. Similarly, if $(i-a)_n \equiv (i+a+1)_n$, then $n-1 = 2a$, which also results in a contradiction. Combining these results, we conclude that for every $0 \le i \le n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+2)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i+a+1)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-a+1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n}\ \underbrace{h((i+1)_n)}_{(i+1)_n\ (i)_n}\ \underbrace{h((i+2)_n)}_{(i+1)_n}\dots\underbrace{h((i+a)_n)}_{(i)_n}\ \underbrace{h((i+a+1)_n)}_{(i+1)_n}\dots\underbrace{h((i-a)_n)}_{(i)_n}\ \underbrace{h((i-a+1)_n)}_{(i+1)_n}\dots$$

Hence, $(i)_n$ and $(i+1)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

Case (iii): In this case, we will prove that $(i)_n$ and $(i-a)_n$ alternate in $h(u_i)$. Assuming $(i-2a)_n \prec_{u_i} (i+1)_n$, it follows that $(n+i-2a) < (i+1)$. This leads to $n-1 < 2a$, which is a contradiction. Similarly, if $(i-2a)_n \equiv (i+1)_n$, then $n-1 = 2a$, which also results in a contradiction.

Now, suppose $(i+a)_n \prec_{u_i} (i-2a)_n$, it follows that $(i+a) < (n+i-2a)$. This would imply $n > 3a$, again yielding a contradiction.

Combining these results and some results from previous cases, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i-2a)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-a+1)_n$$

By the definition of the morphism h, we get the following cases:

If $(i+a)_n \not\equiv (i-2a)_n$,

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i-a)_n}\ \underbrace{h((i+1)_n)}_{(i)_n}\dots\underbrace{h((i-2a)_n)}_{(i-a)_n}\dots\underbrace{h((i+a)_n)}_{(i)_n}\dots\underbrace{h((i-a)_n)}_{(i-a)_n\ (i)_n}\underbrace{h((i-a+1)_n)}_{(i-a)_n}\dots$$

If $(i+a)_n \equiv (i-2a)_n$,

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i-a)_n}\ \underbrace{h((i+1)_n)}_{(i)_n}\dots\underbrace{h((i+a)_n)}_{(i-a)_n\ (i)_n}\dots\underbrace{h((i-a)_n)}_{(i-a)_n\ (i)_n}\ \underbrace{h((i-a+1)_n)}_{(i-a)_n}\dots$$

Hence, $(i)_n$ and $(i-a)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

Case (iv): In this case, we will prove that $(i)_n$ and $(i+a)_n$ alternate in $h(u_i)$. Since $\frac{n}{3} \leq a < \frac{n-1}{2}$ and $n > 6$, $(i+1)_n \prec_{u_i} (i+a)_n$. Assuming $(i+2a)_n \prec_{u_i} (i-a)_n$, it follows that $(i+2a) < (n+i-a)$. This leads to $n > 3a$, which is a contradiction. Combining this result and some results from previous cases, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+a)_n \ \prec_{u_i} (i+a+1)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i+2a)_n$$

By the definition of the morphism h, we get the following cases:

If $(i+2a)_n \not\equiv (i-a)_n$,

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i+a)_n}\ \underbrace{h((i+1)_n)}_{(i)_n}\ldots\underbrace{h((i+a)_n)}_{(i+a)_n\ (i)_n}\underbrace{h((i+a+1)_n)}_{(i+a)_n}\ldots\underbrace{h((i-a)_n)}_{(i)_n}\ldots\underbrace{h((i+2a)_n)}_{(i+a)_n}\ldots$$

If $(i+2a)_n \equiv (i-a)_n$,

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i+a)_n}\ \underbrace{h((i+1)_n)}_{(i)_n}\ldots\underbrace{h((i+a)_n)}_{(i+a)_n\ (i)_n}\ \underbrace{h((i+a+1)_n)}_{(i+a)_n}\ldots\underbrace{h((i-a)_n)}_{(i)_n\ (i+a)_n}\ldots$$

Hence, $(i)_n$ and $(i+a)_n$ alternate in $h(u_i)$ for all $0 \le i \le n-1$. Therefore, w represents the graph G. □

Theorem 3.4.8. *Let $G \equiv C(n;1,a)$ be a connected 4-regular circulant graph with $a = \frac{n-1}{2}$ and $n > 6$. Then $\mathcal{R}(G) \le 4$.*

Proof. Consider a connected 4-regular circulant graph $G \equiv C(n;1,a)$ with $a = \frac{n-1}{2}$. Let the vertex set $V(G) = \{0,1,\ldots,n-1\}$. By the definition of a circulant graph, a vertex i is adjacent to $j \in \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$. Define a morphism $h : V(G)^* \to V(G)^*$ as follows:

$$h(i) = i\ (i-1)_n\ (i-a)_n\ (i+a)_n,\ \textit{for } i \in V(G).$$

We claim that the word $w = h(u)$ represents G, where $u = 012\ldots(n-1)$. Note that a vertex i occurs once in every $h(i)$, $h((i+1)_n)$, $h((i+a)_n)$, and $h((i-a)_n)$. Therefore, each letter occurs exactly four times in w.

Let $u_i = (i)_n\ (i+1)_n\ \ldots\ (n-1)\ 0\ 1\ \ldots\ (i-1)_n$, where $0 \le i \le n-1$. By *Proposition* 3, the word $h(u_i)$ represents the graph G for all $0 \le i \le n-1$. As a result, i and j alternate in w if and only if they alternate in $h(u_i)$ for all $0 \le i \le n-1$.

By the definition of morphism h, we have

$$h(u_i) = (i)_n\ (i-1)_n\ (i+a)_n\ (i-a)_n\ (i+1)_n\ (i)_n\ \ldots\ (i-1)_n\ (i-2)_n\ (i+a-1)_n\ (i-a-1)_n.$$

For any vertex $i \in V(G)$, $h(i)\ h((i+1)_n)$ factor of $h(u_i)$ makes sure that i does not alternate with $j \ \in V(G) \setminus \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$. Therefore, it is enough to prove that i and $j \in \{(i+1)_n, (i-1)_n, (i+a)_n, (i-a)_n\}$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$. Recall that we say $x \prec_u y$, if x occurs to the left of y in the word u.

Case (i): In this case, we will prove that $(i)_n$ and $(i-1)_n$ alternate in $h(u_i)$. Since $a = \frac{n-1}{2}$, and $n > 6$, $(i+1)_n \prec_{u_i} (i+a-1)_n$ and $(i-a)_n \prec_{u_i} (i-1)_n$. Moreover, if $(i-a-1)_n \not\equiv (i+a)_n$, then $n-1 \neq 2a$, which also results in a contradiction. Hence, $(i-a-1)_n \equiv (i+a)_n$. Combining these results, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \ \prec_{u_i} (i+a-1)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n\ (i-1)_n} \quad \underbrace{h((i+1)_n)}_{(i)_n} \ldots \underbrace{h((i+a-1)_n)}_{(i-1)_n} \ \underbrace{h((i+a)_n)}_{(i)_n\ (i-1)_n} \ \underbrace{h((i-a)_n)}_{(i)_n} \ldots \underbrace{h((i-1)_n)}_{(i-1)_n}.$$

Hence, $(i)_n$ and $(i-1)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

Case (ii): In this case, we will prove that $(i)_n$ and $(i+1)_n$ alternate in $h(u_i)$. Since $a = \frac{n-1}{2}$ and $n > 6$, $(i+2)_n \prec_{u_i} (i+a)_n$. Moreover, if $(i-a)_n \not\equiv (i+a+1)_n$, then $n-1 \neq 2a$, which also results in a contradiction. Hence, $(i-a)_n \equiv (i+a+1)_n$. Combining these results, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+2)_n \ \prec_{u_i} (i+a)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-a+1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n} \ \underbrace{h((i+1)_n)}_{(i+1)_n\ (i)_n} \ \underbrace{h((i+2)_n)}_{(i+1)_n} \ldots \underbrace{h((i+a)_n)}_{(i)_n} \ \underbrace{h((i-a)_n)}_{(i+1)_n\ (i)_n} \ \underbrace{h((i-a+1)_n)}_{(i+1)_n} \ldots$$

Hence, $(i)_n$ and $(i+1)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

Case (iii): In this case, we will prove that $(i)_n$ and $(i-a)_n$ alternate in $h(u_i)$. If $(i-2a)_n \not\equiv (i+1)_n$, then $n-1 \neq 2a$, which results in a contradiction. Hence, $(i-2a)_n \equiv (i+1)_n$. Since $a = \frac{n-1}{2}$ and $n > 6$, $(i+1)_n \prec_{u_i} (i+a)_n$. Combining these results and some results from previous cases, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-a+1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n \ (i-a)_n} \ \underbrace{h((i+1)_n)}_{(i)_n \ (i-a)_n} \ldots \underbrace{h((i+a)_n)}_{(i)_n} \ \underbrace{h((i-a)_n)}_{(i-a)_n \ (i)_n} \ \underbrace{h((i-a+1)_n)}_{(i-a)_n} \ldots$$

Hence, $(i)_n$ and $(i-a)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$.

Case (iv): In this case, we will prove that $(i)_n$ and $(i+a)_n$ alternate in $h(u_i)$. Since $a = \frac{n-1}{2}$ and $n > 6$, $(i+1)_n \prec_{u_i} (i+a)_n$ and $(i-a)_n \prec_{u_i} (i-1)_n$. Moreover, if $(i+2a)_n \not\equiv (i-1)_n$, then $n-1 \neq 2a$, which results in a contradiction. Hence, $(i+2a)_n \equiv (i-1)_n$. Combining these results and some results from previous cases, we conclude that for every $0 \leq i \leq n-1$, the following ordering holds:

$$(i)_n \prec_{u_i} (i+1)_n \prec_{u_i} (i+a)_n \prec_{u_i} (i+a+1)_n \prec_{u_i} (i-a)_n \prec_{u_i} (i-1)_n$$

By the definition of the morphism h, we get

$$h(u_i) = \underbrace{h(i)}_{(i)_n \ (i+a)_n} \ \underbrace{h((i+1)_n)}_{(i)_n} \ldots \underbrace{h((i+a)_n)}_{(i+a)_n \ (i)_n} \ \underbrace{h((i+a+1)_n)}_{(i+a)_n} \ldots \underbrace{h((i-a)_n)}_{(i)_n} \ldots \underbrace{h((i-1)_n)}_{(i+a)_n}.$$

Hence, $(i)_n$ and $(i+a)_n$ alternate in $h(u_i)$ for all $0 \leq i \leq n-1$. Therefore, w represents the graph G. □

By combining the above two results, we get the following Corollary.

Corollary 5. Let $G \equiv C(n; 1, a)$ be a connected 4-regular circulant graph with $\frac{n}{3} \leq a < \frac{n}{2}$ and $n > 6$. Then $\mathcal{R}(G) \leq 4$.

4

Forbidden induced subgraph characterization of word-representable split graphs

Let $\mathcal{E}_{i,j}$ denote the class of graphs whose vertex set can be partitioned into at most i independent sets and j cliques. For instance, $\mathcal{E}_{2,0}$ is the class of bipartite graphs, and $\mathcal{E}_{3,0}$ is the class of 3-colorable graphs. It is known that $\mathcal{E}_{i,j}$ is word-representable for $(i, j) \in \{(1, 0), (0, 1), (2, 0), (3, 0)\}$.

Furthermore, for $\mathcal{E}_{1,1}$, the class of split graphs, a necessary and sufficient condition for word-representability was obtained by Kitaev et al. [31]. Similarly, for $\mathcal{E}_{0,2}$, the class of co-bipartite graphs, a necessary and sufficient condition for word-representability was established by Das and Ramesh [14]. Consequently, for these two classes, a complete characterization of word-representability in terms of forbidden induced subgraphs is obtained in the following chapters.

4.1 Introduction

A *split graph*—a graph whose vertex set can be partitioned into a clique and an independent set—forms an important subclass for studying word-representability. The investigation of word-representable split graphs is motivated by both theoretical and practical considerations. From a theoretical standpoint, split graphs occupy a central position in graph theory: they are hereditary, meaning that every induced subgraph of a split graph is also a split graph, and they constitute a subclass of chordal graphs. Their simple definition, coupled with rich structural properties, makes them a natural framework for testing conjectures and exploring graph-theoretic phenomena. From a practical standpoint, split graphs serve as models for real-world systems consisting of a densely connected core interacting with a sparsely connected periphery. Examples include social networks, where a close-knit community engages with occasional outsiders, biological interaction networks, and database architectures optimized for hierarchical queries. These characteristics render split graphs a valuable object of study in both combinatorics and applications.

A hereditary graph class can often be characterized by a family of forbidden induced subgraphs: that is, a graph belongs to the class if and only if it contains no graph from a specified family $\mathcal{F}$ as an induced subgraph. The set $\mathcal{F}$ is called the collection of *minimal forbidden induced subgraphs* for the class. For example, *cographs* are precisely the graphs that do not contain an induced path on four vertices (P_4) [12]. Similarly, *trivially perfect graphs* are those that forbid both the path P_4 and the cycle C_4 [21].

The study of hereditary graph classes and their forbidden subgraph characterizations is central to algorithm design and complexity theory. When the set of minimal forbidden induced subgraphs $\mathcal{F}$ is finite, testing membership in the class reduces to checking for the presence of finitely many induced subgraphs, often enabling the design of efficient, polynomial-time recognition algorithms. Not every hereditary graph class, however, admits such a finite forbidden subgraph characterization, which poses additional challenges for both structural analysis and algorithm development.

Since word-representable graphs are hereditary, a line of research has focused on identifying their set of minimal forbidden induced subgraphs, which remains unknown. As a result, the problem of determining the forbidden induced subgraphs for word-representable graphs can be approached by restricting attention to smaller graph classes, such as split graphs.

The word-representability of split graphs was first studied by Kitaev et al. [31], where the authors characterized word-representable split graphs in terms of forbidden subgraphs in which vertices in the independent set have degree at most 2, or the size of the clique is 4. Moreover, they provided necessary and sufficient conditions for an orientation of a split graph to be semi-transitive. Subsequently, in [8], Chen et al. computationally characterized word-representable split graphs with clique size 5 in terms of forbidden subgraphs. More recently, in [42], Roy and Hariharasubramanian characterized word-representable split graphs with independent set size 4, in terms of forbidden subgraphs as well.

Recent work by Kitaev and Pyatkin [35] made significant progress by characterizing semi-transitive split graphs using matrix-theoretic tools. They showed that the adjacency structure of a split graph can be encoded as a $(0, 1)$-matrix and analyzed via the *circular-ones property*. Specifically, they proved that a split graph is semi-transitive if and only if its adjacency matrix satisfies the circular-ones property for rows, together with an additional structural constraint on the placement of ones.

Among various matrix properties, the circular-ones property has found applications in several graph-theoretic characterizations. The circular-ones property (see *Definition* 1.4.4) generalizes the consecutive-ones property. In [45], Tucker provided a forbidden submatrix characterization for the consecutive-ones property (see *Definition* 1.4.2). Later, Booth and Lueker [3] presented a linear-time algorithm for efficiently recognizing the consecutive-ones property. In [39], Safe provided an analogous characterization of the circular-ones property and presented a linear-time algorithm to recognize it. Furthermore, in [40], Safe introduced a matrix property called the *D-circular property*, a special case of the circular-ones property in which both the rows and the set difference between any two rows form circular intervals (see *Definition* 1.4.3) under some linear ordering of columns. In the same work, he also established a forbidden submatrix characterization for the D-circular property and presented a linear-time recognition algorithm for it.

In this chapter, we introduce a matrix property that extends the D-circular property, which we call the *I-circular property*. This property requires that the rows of a matrix, as well as all pairwise intersections of these rows, form circular intervals under some linear ordering of the columns. This refinement serves as a key tool for characterizing semi-transitive split graphs and offers a natural matrix-theoretic framework for identifying forbidden configurations.

The main objective of this work is to establish a *minimal forbidden submatrix characterization* for the I-circular property, thereby shedding light on the minimal forbidden induced subgraph characterization of semi-transitive split graphs. To achieve this, we develop a unified framework linking the I-circular property of matrices to the combinatorial properties of split graphs. This approach culminates in the identification of a complete set of minimal forbidden induced subgraphs.

This chapter is organized as follows. Section 4.2 introduces the I-circular property, explores its relationship with the circular-ones property, and identifies the set of minimal forbidden submatrices for the I-circular property. Subsections 4.2.1, 4.2.2, and 4.2.3 present several matrix constructions and lemmas that support the main results. Section 4.3 extends these matrix-theoretic results to the graph-theoretic domain, establishing a forbidden induced subgraph characterization of semi-transitive split graphs.

4.2 The I-Circular Property

In this section, we introduce the I-circular property, which plays a central role in our characterization of semi-transitive split graphs via minimal forbidden induced subgraphs. The I-circular property is a matrix property closely related to the well-known D-circular property introduced by Safe [40]. We formally define the I-circular property and establish a minimal forbidden submatrix characterization for it.

Definition 4.2.1. *A matrix M is said to have the* I-circular property *if there exists a linear ordering $\preccurlyeq_c$ of its columns such that every row of M forms a circular interval with respect to $\preccurlyeq_c$, and for every pair of rows r and s of M, the intersection $r \cap s$ is also a circular interval with respect to $\preccurlyeq_c$. In this case, the ordering $\preccurlyeq_c$ is called an* I-circular order *of M.*

Definition 4.2.2. *We define Λ to be an operator that associates with each matrix M a matrix $\Lambda(M)$ obtained by appending additional rows to M as follows. For every pair of nontrivial rows r and s of M such that the complement of r is properly contained in s, a new row equal to $r \cap s$ is appended. The order in which these rows are appended is left unspecified, since it is irrelevant for our purposes.*

Lemma 4.2.3. *A matrix M has the I-*circular property *if and only if the matrix $\Lambda(M)$ satisfies the* circular-ones property*.*

Proof. By definition, every I-circular order of M is a circular-ones order of $\Lambda(M)$. Consequently, if M has the I-circular property, then $\Lambda(M)$ has the circular-ones property.

Conversely, suppose that $\Lambda(M)$ has the circular-ones property, and let $\preccurlyeq_c$ be a circular-ones order of $\Lambda(M)$. Consider any two rows r and s of M. Since r and s are also rows of $\Lambda(M)$, each forms a circular interval with respect to $\preccurlyeq_c$. Therefore, their intersection $r \cap s$ is a circular interval with respect to $\preccurlyeq_c$, except possibly in the case where the complement of r is properly contained in s.

If either r or s is trivial, then $r \cap s$ is trivial, equal to r, or equal to s, and hence is a circular interval with respect to $\preccurlyeq_c$. Thus, we may assume without loss of generality that both r and s are nontrivial. In the remaining case, where the complement of r is properly contained in s, the row $r \cap s$ is explicitly added in the construction of $\Lambda(M)$ and therefore belongs to $\Lambda(M)$. Hence, $r \cap s$ is again a circular interval with respect to $\preccurlyeq_c$.

We conclude that, for every pair of rows r and s of M, the intersection $r \cap s$ is a circular interval with respect to $\preccurlyeq_c$. By definition, this implies that $\preccurlyeq_c$ is an I-circular order of M. Therefore, whenever $\Lambda(M)$ has the circular-ones property, M has the I-circular property. This completes the proof of the lemma. ☐

$$M_{VI} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

Figure 4.1: One of the matrix in $\mathcal{F}_{\text{IcircR}}$, M_{VI}.

The family of minimal forbidden submatrices characterizing the I-circular property is the set $\mathcal{F}_{\text{IcircR}}$, defined as follows.

Definition 4.2.4. *We denote*

$$\mathcal{F}_{\text{IcircR}} = \{M_I^*(k), M_{III}(k), M_{II}(k{+}1) \colon k \geq 3\} \cup \{0101 \oplus M_I^*(4),\ 0100 \oplus M_{II}(4),\ M_{IV},\ M_V,\ M_{VI}\}.$$

(Refer to Figures 1.1 and 4.1.)

We now state the following lemmas, which help us understand the relationship between the set $\mathcal{F}_{\text{IcircR}}$ and the set $\mathcal{F}_{\text{circR}}$.

Lemma 4.2.5. *Let* $a = a_1a_2 \ldots a_{k-2}\ 0\ 0$ *be a binary sequence of length* k *for some* $k \geq 4$. *If* $a \neq 0100$, *then* $a \oplus M_{II}(k)$ *contains some matrix in* $\mathcal{F}_{\text{IcircR}}$ *as a configuration.*

Proof. Let $a = a_1a_2 \ldots a_{k-2}\,0\,0$ be a binary sequence of length k for some $k \geq 4$. If $a_i = 0$ for all $1 \leq i \leq k-2$, then $a \oplus M_{II}(k) = M_{II}(k)$. Now suppose that $k \geq 5$ and $a_i = 1$ for some $2 \leq i \leq k-3$. It can be verified that $(a \oplus M_{II}(k))_{\langle i,\, k-1,\, k\rangle,\, \langle k-1,\, 1,\, i,\, k\rangle} = M_{VI}$.

Next, suppose that $a_i = 0$ for all $2 \leq i \leq k-3$, and that either $a_1 = 1$ or $a_{k-2} = 1$.

- If $a_1 = 1$ and $a_{k-2} = 0$, then

$$(a \oplus M_{II}(k))_{\langle 2,\, 3,\, \ldots,\, k-1,\, 1\rangle,\, \langle 2,\, 3,\, \ldots,\, k\rangle} = M_{II}(k-1).$$

- If $a_1 = 0$ and $a_{k-2} = 1$, then

$$(a \oplus M_{II}(k))_{\langle 1,\, 2,\, 3,\, \ldots,\, k-2,\, k\rangle,\, \langle 1,\, 2,\, 3,\, \ldots,\, k-2,\, k\rangle} = M_{II}(k-1).$$

- If $a_1 = a_{k-2} = 1$ and $k \geq 6$, then it can be verified that

$$(a \oplus M_{II}(k))_{\langle 2,\, 3,\, \ldots,\, k-2,\, 1\rangle,\, \langle 2,\, 3,\, \ldots,\, k-2,\, k\rangle} = M_{II}(k-2).$$

- If $a_1 = a_{k-2} = 1$ and $k = 5$, then it can be verified that

$$(10100 \oplus M_{II}(5))_{\langle 2,\, 1,\, 4,\, 3\rangle,\, \langle 3,\, 2,\, 1,\, 4\rangle} = 0100 \oplus M_{II}(4).$$

Now, the only remaining cases are when $a = 1100$ and $a = 1000$:

- If $a = 1100$, then

$$(1100 \oplus M_{II}(4))_{\langle 1,\, 2,\, 4,\, 3\rangle,\, \langle 3,\, 4,\, 1,\, 2\rangle} = M_{II}(4).$$

- If $a = 1000$, then

$$(1000 \oplus M_{II}(4))_{\langle 2,\, 1,\, 4,\, 3\rangle,\, \langle 3,\, 2,\, 1,\, 4\rangle} = 0100 \oplus M_{II}(4).$$

This completes the proof of the lemma. □

Lemma 4.2.6. *For every $F \in \mathcal{F}_{\text{IcircR}}$, the matrix $\Lambda(F)$ contains a matrix from $\mathcal{F}_{\text{circR}}$ as a configuration. Moreover, F is a minimal forbidden submatrix for the I-circular property.*

Proof. If $F = M_I^*(k)$ for some $k \geq 3$, or if $F = M_{III}(k)$ for some $k \geq 3$, or if $F = M_{IV}$, then the lemma holds immediately, as $\Lambda(F) = F$ and $F \in \mathcal{F}_{\text{circR}}$. Moreover, since, by *Theorem 1.4.6*, $\Lambda(F)$ is a minimal forbidden submatrix for the circular-ones property, F is a minimal forbidden submatrix for the I-circular property.

If $F = 0101 \oplus M_I^*(4)$, then $\Lambda(F)_{\langle 1,2,3,4\rangle, id_5} = F$ and $F \in \mathcal{F}_{\text{circR}}$. Furthermore, $\Lambda(F)_{\langle 1,2,3,4\rangle, id_5}$ is the only submatrix of $\Lambda(F)$ that lacks the circular-ones property. By *Theorem 1.4.6*, $\Lambda(F)$ is minimal for the circular-ones property, and thus F is minimal for the I-circular property.

Now consider the remaining cases. For each, note that every row and column of F is required to obtain a submatrix of $\Lambda(F)$ that has the same configuration as some matrix in $\mathcal{F}_{\text{circR}}$ and is also minimal for the circular-ones property by *Theorem 1.4.6*. Thus, in each case, F itself is minimal for the I-circular property.

- If $F = M_{II}(k)$ for some $k \geq 4$, let row $k+1$ of $\Lambda(F)$ be obtained by intersecting rows $k-1$ and k of F. With $\rho = \langle 1, 2, \ldots, k-2, k+1\rangle$, we have $\Lambda(F)_{\rho, id_k} = 00\ldots01 \oplus M_I^*(k-1)$.

- If $F = M_V$, $\Lambda(F)$ adds a sixth row from rows 2 and 4 of F, and $\Lambda(F)_{\langle 1,4,3,6\rangle, id_5} = 0100 \oplus M_I^*(4)$.

- If $F = 0100 \oplus M_{II}(4)$, $\Lambda(F)$ adds a fifth row from rows 3 and 4, and $\Lambda(F)_{\langle 1,2,6\rangle, id_4} = M_I^*(3)$.

- If $F = M_{VI}$, $\Lambda(F)$ adds rows 4, 5, and 6 from intersections of rows 1 and 2, 1 and 3, and 2 and 3, respectively, yielding $\Lambda(F)_{\langle 4,5,6\rangle, id_4} = 111 \oplus M_I^*(3)$.

This completes the proof that $F \in \mathcal{F}_{\text{IcircR}}$ is a minimal forbidden submatrix for the I-circular property in all cases. □

Corollary 6. No matrix in $\mathcal{F}_{\text{IcircR}}$ satisfies the I-circular property.

Proof. Let $F \in \mathcal{F}_{\text{IcircR}}$. By Lemma 4.2.6, the matrix $\Lambda(F)$ contains a matrix from $\mathcal{F}_{\text{circR}}$ as a configuration. It then follows from Theorem 1.4.6 that $\Lambda(F)$ does not satisfy the circular-ones property. Consequently, by Lemma 4.2.3, the matrix F does not have the I-circular property. This completes the proof of the corollary. ☐

Lemma 4.2.7. *Every matrix $F \in \mathcal{F}_{\text{circR}}$ contains a matrix belonging to $\mathcal{F}_{\text{IcircR}}$ as a configuration.*

Proof. Let F be some matrix in $\mathcal{F}_{\text{circR}}$. If F is M_{IV}, then $F \in \mathcal{F}_{\text{IcircR}}$. If F is M_V^*, then $F_{id_5,\langle 1,2,3,4,5\rangle} = M_V$. If F is $\overline{M_{IV}}$, then $F_{\langle 2,3,4\rangle,\langle 1,2,3,5\rangle} = M_{VI}$. If F is $\overline{M_V^*}$, then $F_{\langle 1,3,4\rangle,\langle 2,3,5,6\rangle} = M_{VI}$. Therefore, it only remains to consider the case where $F = a \oplus M_I^*(k)$ for some binary sequence $a = a_1a_2\dots a_k$ such that $a \in A_k$. If a is empty, then $F = M_I^*(k)$. Therefore, assume that 1 occurs at least once in a.

Suppose that $k \geq 5$ and that $1xy1$ occurs circularly in a at position i for some $i \in [k]$ and for some $x, y \in \{0,1\}$ and $(x,y) \neq (0,0)$. If we let $\rho = \langle i, i+1+y, i+4\rangle$ and $\sigma = \langle i, i+2, i+4, k+1\rangle$ (where the sums involving i are modulo k), then $a_\rho = 111$,

$$F_{\rho,\sigma} = (a \oplus M_I^*(k))_{\rho,\sigma} = 111 \oplus M_I^*(k)_{\rho,\sigma},$$

and, consequently, $F_{\rho,\sigma}$ has the same configuration as M_{VI}.

Now suppose that $k \geq 4$ and that 10^m1 occurs circularly in a at position i for some $i \in [k]$ and some $m \geq 2$. If we let $\rho = \langle i+1, i+2, \dots, i+m+1, i\rangle$ and $\sigma = \langle i+1, i+2, \dots, i+m+1, k+1\rangle$ (where the sums involving i are modulo k), then $F_{\rho,\sigma}$ has the same configuration as $M_{II}(k)$.

Now suppose that $k \geq 5$ and that 10101 occurs circularly in a at position i for some $i \in [k]$. If we let $\rho = \langle i, i+2, i+4\rangle$ and $\sigma = \langle i+1, i+3, i+4, k+1\rangle$ (where the sums involving i are modulo k), then $F_{\rho,\sigma}$ has the same configuration as M_{VI}.

Now suppose that $k \geq 4$ and that 1 occurs twice in a. First, assume that the pattern 11 occurs circularly in a at position i for some $i \in [k]$. If we let $\rho = \langle i+2, \dots, k, i, i+1\rangle$ and $\sigma = \langle i+2, \dots, k, 1, 2, \dots, i, k+1\rangle$ (where the sums involving i are taken modulo k), then $F_{\rho,\sigma}$ has the same configuration as $M_{II}(k)$.

Now assume that $k \geq 5$ and that the pattern 101 occurs circularly in a at position i for some $i \in [k]$. If we let $\rho = \langle i+3, i+4, \dots, k, 1, 2, \dots, i, i+2\rangle$ and $\sigma = \langle i+3, i+

$4, \ldots, k, 1, 2, \ldots, i, k+1\rangle$ (where the sums involving i are taken modulo k), then $F_{\rho,\sigma}$ has the same configuration as $M_{II}(k-1)$.

Now suppose that $k \geq 4$ and that 1 occurs once in a at position i for some $i \in [k]$. If we let $\rho = \langle i+1, i+2, \ldots, k, 1, 2, \ldots, i\rangle$ and $\sigma = \langle i+1, i+2, \ldots, k, 1, 2, \ldots, i, k+1\rangle$ (where the sums involving i are taken modulo k), then $F_{\rho,\sigma}$ has the same configuration as $M_{III}(k)$.

Now suppose that $k \geq 5$ and that 1 occurs thrice in a. Then, the only case left is when 111 occurs circularly in a at position i for some $i \in [k]$. If we let $\rho = \langle i+3, i+4, \ldots, k, 1, 2, \ldots, i\rangle$ and $\sigma = \langle i+3, i+4, \ldots, k, 1, 2, \ldots, i, k+1\rangle$ (where the sums involving i are modulo k), then $F_{\rho,\sigma}$ has the same configuration as $M_{II}(k-1)$.

Now suppose that $k = 4$. Since a is a binary bracelet of length 4, the only cases left are when $a = 0101$, $a = 0111$ and $a = 1111$. If $a = 0101$, then $F \in \mathcal{F}_{\text{IcircR}}$. If $a = 0111$, then $F_{\langle 1,2,3,4\rangle,\langle 1,2,3,5\rangle} = 0100 \oplus M_{II}(4)$. If $a = 1111$, then $F_{\langle 1,2,3,4\rangle,\langle 3,5,1,2\rangle} = 0100 \oplus M_{II}(4)$.

Finally, suppose that $k = 3$. By *Theorem* 1.4.6, we may assume that $a = 111$. Then, it can be verified that $111 \oplus M_I^*(3) = M_{III}(3)$. This completes the proof of the lemma. □

We state the following lemma which will be used in the proof of the main result.

Lemma 4.2.8. *[39] If a is any binary sequence of length 4, then $a \oplus M_V^*$ represents the same configuration as one of the matrices M_{IV}, $\overline{M_{IV}}$, M_V^*, and $\overline{M_V^*}$. Conversely, each of the matrices M_{IV}, $\overline{M_{IV}}$, M_V^*, and $\overline{M_V^*}$ represents the same configuration as $a \oplus M_V^*$ for some binary sequence a of length 4. Moreover, the four matrices M_{IV}, $\overline{M_{IV}}$, M_V^*, and $\overline{M_V^*}$ represent pairwise different configurations.*

4.2.1 Matrices Q and R

To each quaternary sequence b of length at least 3, we associate a matrix denoted by $R(b)$. As a preliminary step toward proving the theorem, we establish the following lemma, which asserts that, for all but finitely many such quaternary sequences b, the matrix $R(b)$ contains a matrix from $\mathcal{F}_{\text{IcircR}}$ as a configuration. We now introduce the definitions required to formalize this statement.

For each integer $k \geq 3$ and each $i \in [k]$, we define the following matrices, where throughout $i+1$ is interpreted modulo k:

- $Q_0(i,k)$ is the $1 \times (k+1)$ matrix whose only row has 1's at columns i and $i+1$ and 0's at the remaining ones;
- $Q_1(i,k)$ is the complement of $Q_0(i,k)$;
- $Q_2(i,k)$ is the $2 \times (k+1)$ matrix whose first row has a 0 at column $k+1$ and 1's at the remaining columns and whose second row has 1's at columns i, $i+1$, and $k+1$ and 0's at the remaining columns;
- $Q_3(i,k)$ is the $2 \times (k+1)$ matrix whose first row has a 0 at column i and 1's at the remaining columns and whose second row has 0's at columns $i+1$ and 1's at the remaining columns.

Given a quaternary sequence $b = b_1b_2\dots b_k$ of length k for some $k \geq 3$, we define $R(b)$ to be the matrix with $k+1$ columns whose rows consist of the rows of $Q_{b_1}(1,k)$, followed by those of $Q_{b_2}(2,k)$, then $Q_{b_3}(3,k)$, and so on, up to the rows of $Q_{b_k}(k,k)$. For example,

$$R(013102) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Lemma 4.2.9. *Let $b = b_1b_2\cdots b_k$ be a quaternary sequence of length $k \geq 3$. In the case $k = 3$, assume additionally that either $b_1, b_2, b_3 \in \{1,3\}$ or $b_1, b_2, b_3 \in \{0,2\}$. Then the matrix $R(b)$ contains, as a configuration, a matrix belonging to $\mathcal{F}_{\text{IcircR}}$.*

Proof. First, we show that for any quaternary sequence c obtained from b by a sequence of shift operations, the matrix $R(c)$ represents the same configuration as $R(b)$. To prove this, it suffices to show that $R(c)$ represents the same configuration as $R(b)$ when c is a single shift of b (since the general case then follows by induction). Let m denote the number of rows of $R(b)$, and suppose that c is a shift of b. Define the row permutation ρ as $\langle 2,3,\ldots,m,1\rangle$ if $b_1 \in \{0,1\}$, and as $\langle 3,4,\ldots,m,1,2\rangle$ if $b_1 \notin \{0,1\}$. Also, define the column permutation σ as $\langle 2,3,\ldots,k,1,k+1\rangle$. Then, we have $R(c) = R(b)_{\rho,\sigma}$.

Let $b = b_1 b_2 \ldots b_k$ be a quaternary sequence of length k, where $k \geq 3$. For each $i \in [k]$, let f_i denote the row index of $R(b)$ corresponding to the first row of $Q_{b_i}(i,k)$. If $b_i \notin \{0,1\}$, then let $s_i = f_i + 1$ denote the row index of $R(b)$ corresponding to the second row of $Q_{b_i}(i,k)$.

Case 1: Let b be a binary sequence. By construction, we have $R(b) = b \oplus M_I^*(k)$ for some $k \geq 3$. By applying a sequence of shift operations (if necessary), we may assume, without loss of generality, that b is a bracelet. Therefore, $R(b) \in \mathcal{F}_{\mathrm{IcircR}}$.

Case 2: Suppose that 3 occurs more that once in b. Let $b_i = b_j = 3$, where $i,j \in [k]$. Without loss of generality, we may assume that $i < j$. It can be verified that $R(b)_{\langle f_i,s_i,f_j\rangle,\langle i,i+1,j+1,k+1\rangle} = M_{VI}$, when $i+1 \neq j$, and $R(b)_{\langle f_i,s_i,s_j\rangle,\langle i,i+1,j+1,k+1\rangle} = M_{VI}$, when $i+1 = j$, where all additions involving i and j are taken modulo k.

Case 3: Suppose that both 2 and 3 occur in b. Let $b_i = 2$ and $b_j = 3$. It can be verified that $R(b)_{\langle f_i,s_i,f_j\rangle,\langle i,i+1,i+2,k+1\rangle} = M_{VI}$, where all additions involving i and j are taken modulo k.

Case 4: Suppose 3 occurs in b. By Case 2, it is evident that 3 occurs exactly once in b. By applying a sequence of shift operations (if necessary), we may assume, without loss of generality, that $b_k = 3$, and that $b' = b_1 b_2 \ldots b_{k-1}$ is a binary string. Therefore, $R(b)$ has the same configuration as $a \oplus M_{II}(k+1)$, where a is obtained by appending 00 to b'.

If $a \neq 0100$, then by *Lemma 4.2.5*, $R(b)$ contains some matrix in $\mathcal{F}_{\mathrm{IcircR}}$ as a configuration. If $a = 0100$, then $R(b) = 0100 \oplus M_{II}(4)$.

Case 5: Suppose that 2 occurs in b. By applying a sequence of shift operations (if necessary), we may assume, without loss of generality, that $b_1 = 2$. If $b_2 = 1$, then it can be verified that $R(b)_{\langle 1,2,3\rangle,\langle 1,2,k,k+1\rangle} = M_{VI}$. If $b_2 = 2$, then it can be verified that $R(b)_{\langle 2,3,4\rangle,\langle 1,2,3,k+1\rangle} = M_{VI}$. If $b_k = 1$, then $R(b)_{\langle 1,2,k\rangle,\langle 3,k+1,1,2\rangle} = M_{VI}$. If $b_k = 2$, then $R(b)_{\langle 1,2,k+1\rangle,\langle k,2,k+1,1\rangle} = M_{VI}$. Now suppose that $b_2 = b_k = 0$. If $k \geq 4$, then it can be easily verified that $R(b)_{\langle 1,2,3,k\rangle,\langle 1,k,3,2,k+1\rangle} = M_V$. If $k = 3$, then by the assumption of the lemma, $b_1, b_2, b_3 \in \{0,2\}$. By applying a sequence of shift operations (if necessary), we may assume, without loss of generality, that $b = 200$. Then, $R(b)_{\langle 3,4,1,2\rangle,\langle 3,2,4,1\rangle} = 0100 \oplus M_{II}(4)$. This completes the proof of the lemma. $\square$

4.2.2 Matrices U and W

To each quaternary sequence b of length 4, we associate a matrix denoted by $W(b)$. We establish the following Lemma, which states that for certain quaternary sequences b, the matrix $W(b)$ contains some matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration. We now introduce the necessary definitions to formalize this.

We first define, for each $i \in [4]$, the following matrices:

- $U_0(i)$ whose only rows coincides with row i of M_V^*;
- $U_1(i)$ is the complement of $U_0(i)$.

For each $i \in [3]$, we define the following matrix:

- $U_2(i)$ is the 2×6 matrix whose first row has 0's in columns $5 - 2i$ and $6 - 2i$, and 1's in the remaining columns, and whose second row has 0's in columns $7 - 2i$ and $8 - 2i$, and 1's in the remaining columns, where all subtractions involving i are taken modulo 6.

For $i = 2$, we define the following matrix:

- $U_3(i)$ is the 2×6 matrix whose first row has 0 in column 5 and 1's in the remaining columns, and whose second row has 0 in column 6 and 1's in the remaining columns.

For each quaternary sequence $b = b_1b_2b_3b_4$ of length 4 such that $b_4 = 0$, we define $W(b)$ as the matrix having six columns and whose rows are those of $U_{b_1}(1)$, followed by those of

$U_{b_2}(2)$, followed by those of $U_{b_3}(3)$, followed by those of $U_0(4)$. For instance,

$$W(2310) = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Lemma 4.2.10. *If $b = b_1b_2b_3b_4$ is a quaternary sequence such that $b_4 = 0$ and $b_1, b_3 \neq 3$, then $W(b)$ contains some matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration.*

Proof. Let $b = b_1b_2b_3b_4$ be a quaternary sequence of length 4. For each $i \in [4]$, let f_i denote the row index of $W(b)$ corresponding to the first row of $U_{b_i}(i)$. If $b_i \notin \{0, 1\}$, then let $s_i = f_i + 1$ denote the row index of $W(b)$ corresponding to the second row of $U_{b_i}(i)$.

Case 1: Let b be a binary sequence. By construction, we have $W(b) = b \oplus M_V^*$ which, by *Theorem* 1.4.6, represents the same configuration as some matrix in $\mathcal{F}_{\text{circR}}$. By *Lemma* 4.2.7, $W(b)$ contains some matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration.

Case 2: Suppose that 2 occurs in b, and let $b_i = 2$ for some $i \in [k]$. By construction, $b_4 = 0$. If we define $\sigma_1 = \langle 1, 2, 4, 5 \rangle$, $\sigma_2 = \langle 1, 4, 5, 6 \rangle$, and $\sigma_3 = \langle 1, 3, 4, 5 \rangle$, then it can be verified that $W(b)_{\langle f_i, s_i, f_4 \rangle, \sigma_i}$ has the same configuration as M_{VI}.

Case 3: Suppose that 3 occurs in b. By construction, $b_4 = 0$ and $b_2 = 3$. If $b_1 = b_3 = 0$, then it can be verified that $W(b)_{\langle 1,2,4,5 \rangle, \langle 1,2,3,4,5 \rangle} = M_V$. If $b_1 = 1$, then it can be verified that $W(b)_{\langle 1,2,5 \rangle, \langle 1,3,4,5 \rangle}$ has the same configuration as M_{VI}. If $b_3 = 1$, then it can be verified that $W(b)_{\langle 2,3,4 \rangle, \langle 2,3,5,6 \rangle}$ has the same configuration as M_{VI}. This completes the proof of the lemma. □

4.2.3 Matrices H and G

For each binary sequence $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$ of length 4 and each $i \in [3]$, we define $H_i(\alpha)$ to be the 6×6 matrix obtained from M_V^* by appending two additional rows (the fifth and sixth rows), both having 1's in columns $3-2i$ and $4-2i$. In the fifth row, the entries in columns $5-2i$, $6-2i$, $1-2i$, and $2-2i$ (with all subtractions taken modulo 6) are given by α_1, α_2, α_3, and α_4, respectively. In the sixth row, the corresponding entries are $\overline{\alpha_1}$, $\overline{\alpha_2}$, $\overline{\alpha_3}$, and $\overline{\alpha_4}$, respectively. For instance,

$$H_1(\alpha_1\alpha_2\alpha_3\alpha_4) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & \alpha_1 & \alpha_3 & \alpha_3 & \alpha_4 \\ 1 & 1 & \overline{\alpha_1} & \overline{\alpha_2} & \overline{\alpha_3} & \overline{\alpha_4} \end{pmatrix}.$$

For each binary sequence $\gamma = \gamma_1\gamma_2\gamma_3$ of length 3, we define

$$G(\gamma) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & \gamma_1 & \gamma_2 & 1 & 1 & \gamma_3 \\ 1 & \overline{\gamma_1} & \overline{\gamma_2} & 1 & 1 & \overline{\gamma_3} \end{pmatrix}.$$

Lemma 4.2.11. *Let α be a binary sequence of length four, and let $i \in [3]$. If $\alpha \notin \{0000, 0011, 1100, 1111\}$, then the matrix $H_i(\alpha)$ contains, as a configuration, a matrix in $\mathcal{F}_{\text{circR}}$ with fewer than six columns.*

Proof. Let $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$ be a binary sequence of length 4. Since $H_i(\overline{\alpha})_{\langle 1,2,3,4,6,5\rangle, id_6} = H_i(\alpha)$, the lemma holds for α if and only if it holds for $\overline{\alpha}$. In particular, by complementing α if necessary, we may assume that $\alpha_1 = 0$. We now consider the following cases.

Case 1: $\alpha = 0001$. In this case, $H_1(\alpha)_{\langle 2,5,4\rangle,\langle 6,5,3,1\rangle} = H_2(\alpha)_{\langle 3,5,4\rangle,\langle 6,1,3,4\rangle} = \overline{M_I^*(3)}$, and $H_3(\alpha)_{\langle 1,5,4\rangle,\langle 1,2,4,6\rangle} = M_I^*(3)$.

Case 2: $\alpha = 0010$. In this case, $H_1(\alpha)_{\langle 2,4,5\rangle,\langle 2,4,5,6\rangle} = H_2(\alpha)_{\langle 4,3,5\rangle,\langle 5,4,3,2\rangle} = M_I^*(3)$, and $H_3(\alpha)_{\langle 1,4,5\rangle,\langle 5,3,2,1\rangle} = \overline{M_I^*(3)}$.

Case 3: $\alpha = 0100$. In this case, $H_1(\alpha)_{\langle 3,4,5\rangle,\langle 5,2,3,4\rangle} = H_3(\alpha)_{\langle 2,5,4\rangle,\langle 6,5,2,4\rangle} = \overline{M_I^*(3)}$, and $H_2(\alpha)_{\langle 1,5,4\rangle,\langle 1,2,5,3\rangle} = M_I^*(3)$.

Case 4: $\alpha = 0101$. In this case, $H_1(\alpha)_{\langle 2,5,4\rangle,\langle 6,5,3,1\rangle} = H_2(\alpha)_{\langle 1,4,6\rangle,\langle 4,3,2,1\rangle} = \overline{M_I^*(3)}$, and $H_3(\alpha)_{\langle 1,4,2,5\rangle,\langle 1,2,6,5,4\rangle} = 01111 \oplus M_I^*(4)$.

Case 5: $\alpha = 0110$. In this case, $H_1(\alpha)_{\langle 3,6,5,2\rangle,\langle 3,4,5,6,2\rangle} = 0111 \oplus M_I^*(4)$, $H_2(\alpha)_{\langle 1,6,3,4\rangle,\langle 1,2,3,4,6\rangle} = 0100 \oplus M_I^*(4)$, and $H_3(\alpha)_{\langle 1,6,4\rangle,\langle 1,2,4,6\rangle} = M_I^*(3)$.

Case 6: $\alpha = 0111$. In this case, $H_1(\alpha)_{\langle 6,3,4\rangle,\langle 1,3,4,6\rangle} = H_3(\alpha)_{\langle 2,6,4\rangle,\langle 1,3,5,6\rangle} = M_I^*(3)$, and $H_2(\alpha)_{\langle 1,6,4\rangle,\langle 6,4,2,1\rangle} = \overline{M_I^*(3)}$. This completes the poof of the lemma. ☐

Lemma 4.2.12. *Let γ be a binary sequence of length three. If γ is nonconstant, then the matrix $G(\gamma)$ contains, as a configuration, a matrix in $\mathcal{F}_{\text{circR}}$ with fewer than six columns.*

Proof. Let $\gamma = \gamma_1\gamma_2\gamma_3$ be a binary sequence of length 3. Since $G(\overline{\gamma})_{\langle 1,2,3,4,6,5\rangle,id_6} = G(\gamma)$, the lemma holds for γ if and only if it holds for $\overline{\gamma}$. In particular, by complementing γ if necessary, we may assume that $\gamma_1 = 0$. We now consider the following cases.

Case 1: $\gamma = 001$. In this case, $G(\gamma)_{\langle 1,4,3,5\rangle,\langle 2,1,4,3,6\rangle} = 0001 \oplus M_I^*(4)$.

Case 2: $\gamma = 010$. In this case, $G(\gamma)_{\langle 1,5,2,4\rangle,\langle 1,2,6,5,3\rangle} = 0110 \oplus M_I^*(4)$.

Case 3: $\gamma = 011$. In this case, $G(\gamma)_{\langle 3,4,2,6\rangle,\langle 3,4,5,6,2\rangle} = 0011 \oplus M_I^*(4)$. This completes the proof of the lemma. ☐

Theorem 4.2.13. *A matrix M has I-circular property if and only if M contains no matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration.*

Proof. Suppose that M has the I-circular property. Assume, for the sake of contradiction, that M contains some matrix $F \in \mathcal{F}_{\text{IcircR}}$ as a configuration. By *Corollary* 6, the matrix F does not possess the I-circular property. This, in turn, implies that M itself cannot have the I-circular property, leading to a contradiction.

Suppose that M contains no matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration. Assume, for the sake of contradiction, that M does not have the I-circular property. Let k_M denote the number of rows of M. By Lemma 4.2.3, the matrix $\Lambda(M)$ does not satisfy the circular-ones property. It then follows from Theorem 1.4.6 that $\Lambda(M)$ contains a matrix from $\mathcal{F}_{\text{circR}}$ as a configuration.

Among all such matrices, let $F \in \mathcal{F}_{\text{circR}}$ be one with the minimum possible number of columns that occurs as a configuration in $\Lambda(M)$. Let ρ and σ be a row map and a column map, respectively, such that

$$\Lambda(M)_{\rho,\sigma} = F.$$

Case 1: Firstly, suppose that $F = a \oplus M_I^*(k)$ for some $k \geq 3$ and some $a \in A_k$. Assume that $a_i = 0$, and let $u = \rho(i)$ for each $i \in [k]$. If $u \in [k_M]$, then

$$F_{\langle u \rangle} = \Lambda(M)_{\langle u \rangle,\sigma} = M_{\langle u \rangle,\sigma} = Q_0(i,k).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r \rangle,\sigma}$ is properly contained in $M_{\langle s \rangle,\sigma}$ and

$$\Lambda(M)_{\langle u \rangle,\sigma} = M_{\langle r \rangle,\sigma} \cap M_{\langle s \rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u \rangle,\sigma}$ has 1's in columns i and $i+1$ and 0's in the remaining columns, both $M_{\langle r \rangle,\sigma}$ and $M_{\langle s \rangle,\sigma}$ contain 1's in columns i and $i+1$. Since $\overline{M}_{\langle r \rangle,\sigma}$ is properly contained in $M_{\langle s \rangle,\sigma}$, if $M_{\langle s \rangle,\sigma}$ has a 0 in any column, then $M_{\langle r \rangle,\sigma}$ has a 1 in that column, and if $M_{\langle r \rangle,\sigma}$ has a 0 in any column, then $M_{\langle s \rangle,\sigma}$ has a 1 in that column. Therefore, $M_{\langle r \rangle,\sigma}$ and $M_{\langle s \rangle,\sigma}$ do not coincide in any columns other than i and $i+1$. Without loss of generality, suppose that $M_{\langle r \rangle,\sigma}$ has a 0 and $M_{\langle s \rangle,\sigma}$ has a 1 in column $k+1$.

Claim: *$M_{\langle r \rangle,\sigma}$ has all $0'$s or all $1'$s in the columns different from $i, i+1$, and $k+1$.*

Suppose, for the sake of contradiction, that there is some 0 and some 1 in columns other than i, $i+1$, and $k+1$. Therefore, there exist $x, y \in [k]$ such that $M_{\langle r \rangle,\sigma}$ has 0's in all the columns in $\{x, x+1, \ldots, y\}$, and 1's in columns $x-1$ and $y+1$ (where addition and subtraction are modulo k), with $(x-1, y+1) \neq (i+1, i)$.

Hence, if $\rho' = \langle \rho(x-1), \rho(x), \ldots, \rho(y), s \rangle$, $\sigma' = \langle x-1, x, \ldots, y+1, k+1 \rangle$, and a' is the sequence obtained from $a_{\rho \circ \langle x-1, x, \ldots, y \rangle}$ by appending a 0, then $\Lambda(M)_{\rho',\sigma'} = a' \oplus M_I^*(|a'|)$ is a matrix representing the same configuration as some matrix in $\mathcal{F}_{\text{circR}}$ that is contained in $\Lambda(M)$ as a configuration and has fewer columns than F, contradicting the choice of F. Therefore, $M_{\langle r \rangle,\sigma}$ must have all 0's or all 1's in the columns other than i, $i+1$, and $k+1$.

If $M_{\langle r\rangle,\sigma}$ has 0's in the columns other than i, $i+1$, and $k+1$, then $M_{\langle r\rangle,\sigma}$ coincides with the only row of $Q_0(i,k)$. If $M_{\langle r\rangle,\sigma}$ has 1's in the columns other than i, $i+1$, and $k+1$, then $M_{\langle s\rangle,\sigma}$ has 0's in the columns other than i, $i+1$, and $k+1$. Therefore, $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ coincide with the rows of $Q_2(i,k)$.

Now assume that $a_i = 1$, and let $u = \rho(i)$ for each $i \in [k]$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = Q_1(i,k).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 0's in columns i and $i+1$, and 1's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in the columns other than i and $i+1$, and $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in columns i and $i+1$.

Let (α,β) and $(\overline{\alpha},\overline{\beta})$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns i and $i+1$, respectively. If $(\alpha,\beta) = (0,0)$ (respectively, $(\alpha,\beta) = (1,1)$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $Q_1(i,k)$. If $(\alpha,\beta) = (0,1)$ or $(\alpha,\beta) = (1,0)$, then $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ coincide with the rows of $Q_3(i,k)$.

Therefore, for each $i \in [k]$ and $j \in \{0,1,2,3\}$, each row of $Q_j(i,k)$ coincides with some row of $M_{id_{k_M},\sigma}$. Thus, there exists a quaternary sequence $b = b_1b_2\ldots b_k$ such that $R(b)$ is contained in M as a configuration. Since, by *Lemma* 4.2.9, $R(b)$ contains some matrix $F' \in \mathcal{F}_{\text{IcircR}}$ as a configuration, it follows that M contains F' as a configuration. This contradicts the assumption that M contains no matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration. Therefore, Case 1 cannot occur.

Case 2: Now suppose that $F = a \oplus M_V^*$ for some $a \in \{0000, 0100, 0010, 1010\}$. By *Lemma* 4.2.8, we can say that F has the same configuration as some matrix in $\{M_{IV}, \overline{M_{IV}}, M_V^*, \overline{M_V^*}\}$.

Subcase 1: Assume that $a_i = 0$, and let $u = \rho(i)$ for each $i \in \{1,3\}$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = U_0(i).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 1's in columns i and $i+1$, and 0's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in columns i and $i+1$. Moreover, since $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$, the rows $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in any columns other than i and $i+1$.

Let $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $(\overline{\alpha}_1, \overline{\alpha}_2, \overline{\alpha}_3, \overline{\alpha}_4)$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns $i+2$, $i+3$, $i+4$, and $i+5$, respectively (with all additions involving i taken modulo 6). Let $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$. Note that $H_i(\alpha)$ is contained in $\Lambda(M)$ as a configuration. Suppose, for the sake of contradiction, that $\alpha \notin \{0000, 0011, 1100, 1111\}$. By *Lemma* 4.2.11, $H_i(\alpha)$ contains some matrix $F' \in \mathcal{F}_{\text{circR}}$ as a configuration that has fewer columns than F. Hence, $\Lambda(M)$ contains F' as a configuration with fewer columns than F, contradicting the choice of F. Therefore, $\alpha \in \{0000, 0011, 1100, 1111\}$. If $\alpha = 0000$ (respectively, $\alpha = 1111$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $U_0(i)$. If $\alpha = 0011$ or $\alpha = 1100$, then $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ coincide with the rows of $U_2(i)$.

Now assume that $a_i = 1$, and let $u = \rho(i)$ for each $i \in \{1, 3\}$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = U_1(i).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 0's in columns i and $i+1$, and 1's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in the columns other than i and $i+1$. Moreover, since $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$, the rows $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in the columns i and $i+1$.

Let (α, β) and $(\overline{\alpha}, \overline{\beta})$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns i and $i+1$, respectively. If $(\alpha, \beta) = (0, 0)$ (respectively, $(\alpha, \beta) = (1, 1)$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $U_1(i)$. If $(\alpha, \beta) = (0, 1)$ or $(\alpha, \beta) = (1, 0)$, then $M_{\langle r\rangle,\sigma}$ and

$M_{\langle s\rangle,\sigma}$ coincide with the rows of $U_3(i)$.

Subcase 2: Assume that $a_2 = 0$, and let $u = \rho(2)$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = U_0(2).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 0's in the columns 5 and 6, and 1's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in the columns other than 5 and 6. Moreover, since $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$, the rows $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in the columns 5 and 6.

Let (α, β) and $(\overline{\alpha}, \overline{\beta})$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns 5 and 6, respectively. If $(\alpha, \beta) = (0, 0)$ (respectively, $(\alpha, \beta) = (1, 1)$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $U_0(2)$. If $(\alpha, \beta) = (0, 1)$ or $(\alpha, \beta) = (1, 0)$, then $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ coincide with the rows of $U_3(2)$.

Now assume that $a_2 = 1$, and let $u = \rho(2)$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = U_1(2).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 1's in columns 5 and 6, and 0's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in the columns 5 and 6. Moreover, since $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$, the rows $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in the columns 1, 2, 3 and 4.

Let $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $(\overline{\alpha}_1, \overline{\alpha}_2, \overline{\alpha}_3, \overline{\alpha}_4)$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns 1, 2, 3, and 4, respectively. Let $\alpha = \alpha_1\alpha_2\alpha_3\alpha_4$. Note that $H_2(\alpha)$ is contained in $\Lambda(M)$ as a configuration. Based on the argument of the previous subcase, we have

$\alpha = \{0000, 0011, 1100, 1111\}$. If $\alpha = 0000$ (respectively, $\alpha = 1111$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $U_1(2)$. If $\alpha = 0011$ or $\alpha = 1100$, then $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ coincide with the rows of $U_2(2)$.

Subcase 3: By construction, $a_4 = 0$. Let $u = \rho(4)$. If $u \in [k_M]$, then

$$F_{\langle u\rangle} = \Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle u\rangle,\sigma} = U_0(4).$$

Otherwise, let $r, s \in [k_M]$ be such that $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$ and

$$\Lambda(M)_{\langle u\rangle,\sigma} = M_{\langle r\rangle,\sigma} \cap M_{\langle s\rangle,\sigma}.$$

Since $\Lambda(M)_{\langle u\rangle,\sigma}$ has 1's in columns 1, 4 and 5, and 0's in the remaining columns, both $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ contain 1's in the columns 1, 4 and 5. Moreover, since $\overline{M}_{\langle r\rangle,\sigma}$ is properly contained in $M_{\langle s\rangle,\sigma}$, the rows $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ do not coincide in the columns 2, 3 and 6.

Let $(\gamma_1, \gamma_2, \gamma_3)$ and $(\overline{\gamma}_1, \overline{\gamma}_2, \overline{\gamma}_3)$ be the entries of $M_{\langle r\rangle,\sigma}$ and $M_{\langle s\rangle,\sigma}$ at columns 2, 3, and 6, respectively. Let $\gamma = \gamma_1\gamma_2\gamma_3$. Note that $G(\gamma)$ is contained in $\Lambda(M)$ as a configuration. Suppose, for the sake of contradiction, that γ is nonconstant. By *Lemma* 4.2.12, $G(\gamma)$ contains some matrix $F' \in \mathcal{F}_{\text{circR}}$ as a configuration that has fewer columns than F. Hence, $\Lambda(M)$ contains F' as a configuration with fewer columns than F, contradicting the choice of F. Therefore, γ must be constant. If $\gamma = 000$ (respectively, $\gamma = 111$), then $M_{\langle r\rangle,\sigma}$ (respectively, $M_{\langle s\rangle,\sigma}$) coincides with the only row of $U_0(4)$.

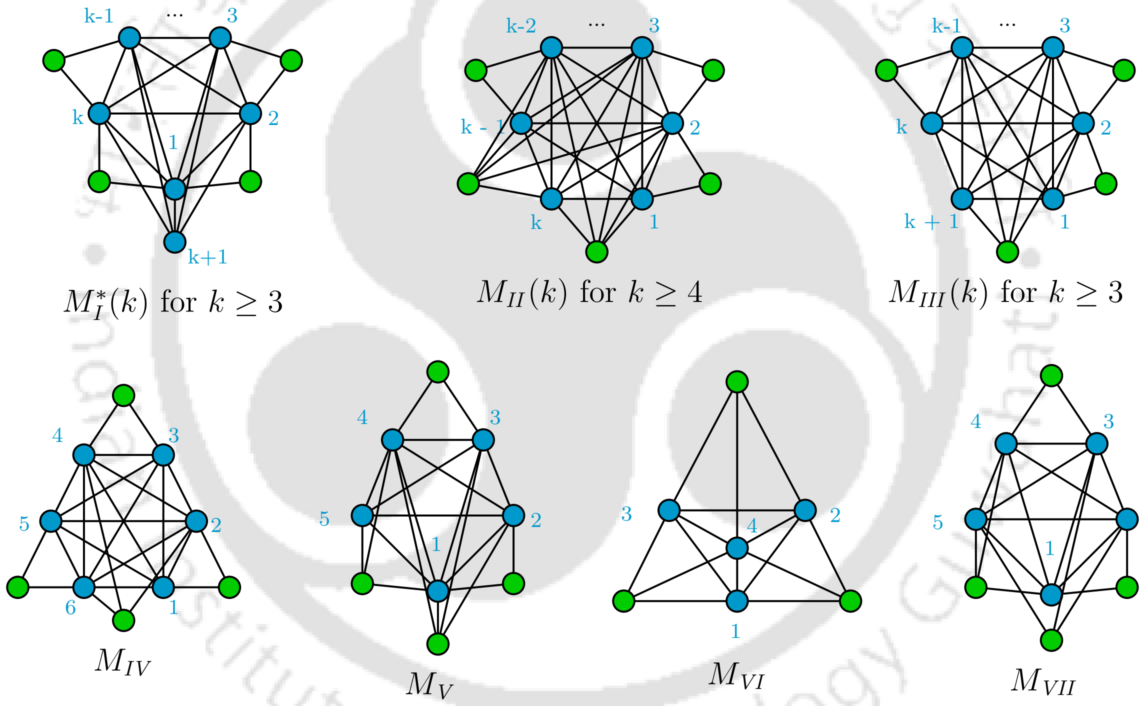
Therefore, for each $i \in [4]$ and $j \in \{0, 1, 2, 3\}$, each row of $U_j(i)$ coincides with some row of $M_{id_{k_M},\sigma}$. Thus, there exists a quaternary sequence $b = b_1b_2b_3b_4$ such that $W(b)$ is contained in M as a configuration. Since, by *Lemma* 4.2.10, $W(b)$ contains some matrix $F' \in \mathcal{F}_{\text{IcircR}}$ as a configuration, it follows that M contains F' as a configuration. This contradicts the assumption that M contains no matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration. Therefore, Case 2 cannot occur.

The fact that Case 1 and Case 2 cannot occur contradicts $F \in \mathcal{F}_{\text{circR}}$. This contradiction completes the proof of the theorem. □

4.3 Forbidden Induced Subgraphs of Semi-Transitive Split Graphs

In this section, we establish the connection between the I-circular property of matrices and semi-transitive split graphs, thereby characterizing the latter through forbidden induced subgraphs. The following theorem follows from the proofs of *Theorem 4* and *Theorem 5* in [35].

Theorem 4.3.1. *A split graph G is semi-transitive if and only if the matrix $A(G)$ has I-circular property.*



$M_I^*(k)$ for $k \geq 3$ $M_{II}(k)$ for $k \geq 4$ $M_{III}(k)$ for $k \geq 3$

M_{IV} M_V M_{VI} M_{VII}

Figure 4.2: Minimal Forbidden Induced Subgraphs for Semi-Transitive Split Graphs $\mathcal{G}_{\text{split}}$.

For a $(0,1)$-matrix M of size $m \times n$, we define $SG(M)$ as the split graph whose independent set corresponds to the rows of M, clique corresponds to its columns, and where a vertex i from the independent set is adjacent to a vertex j from the clique if and only if $m_{ij} = 1$.

Lemma 4.3.2. *If $F \in \mathcal{F}_{\text{IcircR}}$, then $SG(F)$ contains an induced subgraph isomorphic to some graph in $\mathcal{G}_{\text{split}}$.*

Proof. Let F be a matrix in $\mathcal{F}_{\text{IcircR}}$. If $F \in \{M_I^*(k), M_{III}(k), M_{II}(k+1) \colon k \geq 3\} \cup \{0101 \oplus$

$M_I^*(4)$, M_{IV}, M_V, $M_{VI}\}$, then $SG(F)$ is isomorphic to some graph in $\mathcal{G}_{\text{split}}$. If $F = 0100 \oplus M_{II}(4)$, then $SG(F_{\text{id}_4,\langle 1,2,4\rangle})$ is isomorphic to $M_I^*(3)$. This completes the proof of the lemma. ☐

Lemma 4.3.3. *All graphs in $\mathcal{G}_{\text{split}}$ are minimal non-semi-transitive.*

Proof. Let G be a graph in $\mathcal{G}_{\text{split}}$. Observe that $A(G) \in \mathcal{F}_{\text{IcircR}}$. Therefore, by *Corollary 6*, *Lemma 4.2.6*, and *Theorem 4.3.1*, G is a minimal non-semi-transitive graph. ☐

Theorem 4.3.4. *Let G be a split graph. Then, G is semi-transitive if and only if G contains none of the graphs in $\mathcal{G}_{\text{split}}$ (depicted in Figure 4.2) as induced subgraph.*

Proof. Assume that G is semi-transitive. Suppose, for the sake of contradiction, that G contains some graph in $\mathcal{G}_{\text{split}}$ as an induced subgraph. Then, by *Lemma 4.3.3* and the hereditary nature of semi-transitive graphs, G would be non-semi-transitive, which is a contradiction.

Suppose that G contains none of the graphs in $\mathcal{G}_{\text{split}}$ as an induced subgraph. Assume, for the sake of contradiction, that G is not semi-transitive. Then, by *Theorem 4.3.1*, $A(G)$ does not have the I-circular property. Moreover, by *Theorem 4.2.13*, $A(G)$ contains some matrix in $\mathcal{F}_{\text{IcircR}}$ as a configuration. Therefore, by *Lemma 4.3.2*, G contains an induced subgraph isomorphic to some graph in $\mathcal{G}_{\text{split}}$, which is a contradiction. This concludes the proof. ☐

5

Forbidden induced subgraph characterization of word-representable co-bipartite graphs

5.1 Introduction

A *split graph* is a graph whose vertex set can be partitioned into an independent set and a clique. Similarly, if the vertex set can be partitioned into two cliques, the graph is called a co-bipartite graph. More formally, a *co-bipartite* graph is a graph whose complement is a bipartite graph. The word-representability of co-bipartite graphs was first studied by Chen et al. [7], where the authors characterized word-representable co-bipartite graphs in terms of forbidden induced subgraphs when one of the cliques has size at most four. A complete characterization in terms of necessary and sufficient conditions was later provided by Das and Hariharasubramanian [14]. Furthermore, they provided an upper bound for the representation number of word-representable co-bipartite graphs [13].

Building on this growing body of research, the present work concentrates on a detailed structural and algorithmic characterization of word-representable co-bipartite graphs. In particular, we provide a *forbidden induced subgraph characterization* for this class of graphs, grounded in the theory of circularly compatible ones matrices. This result unifies two previously separate lines of research: the geometric theory of circle graphs, developed notably by Bouchet [4], and the matrix-theoretic framework for circular-ones properties, introduced by Tucker [43, 44] and later refined by Safe [39, 40].

The central theorem of this chapter establishes that a co-bipartite graph is semi-transitive if and only if its associated $(0,1)$-matrix satisfies the *circularly compatible ones property*. This equivalence provides not only a minimal forbidden induced subgraph characterization but also a linear-time recognition algorithm for semi-transitive co-bipartite graphs.

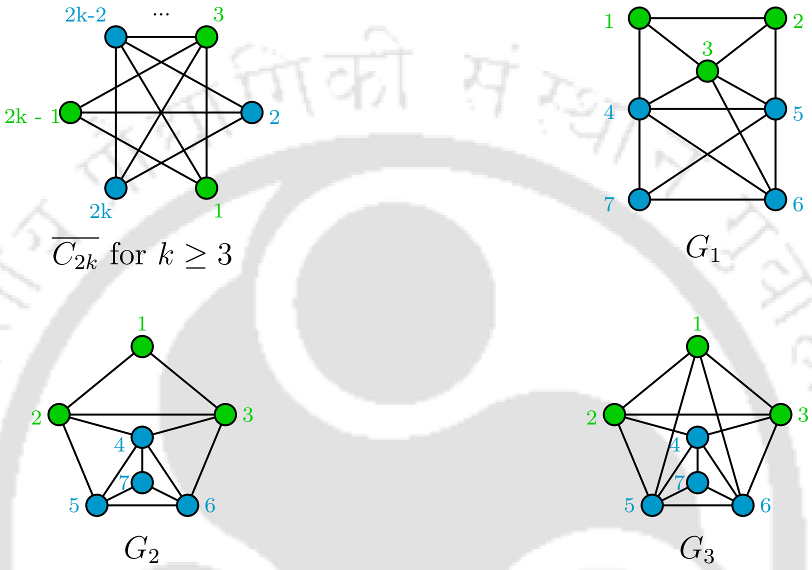
Using an algebraic characterization of circle graphs, Bouchet [4] proved that if a bipartite graph G is the complement of a circle graph, then G itself is also a circle graph. Later, Esperet and Stehlík [16] provided an elementary proof of this result. In Section 5.2, we present an even more elementary and concise proof of a stronger version of this theorem: if a bipartite graph G is the complement of a circle graph, then both G and $\overline{G}$ are permutation graphs. This result can also be viewed as a direct consequence of the proof by Esperet and Stehlík [16].

The remainder of this chapter is organized as follows. Section 5.2 presents the forbidden induced subgraph characterization of co-bipartite circle graphs and establishes their equivalence with co-bipartite permutation graphs. Section 5.3 contains the main theorem, which proves the equivalence between semi-transitivity and the circularly compatible ones property, and further provides a complete forbidden induced subgraph characterization along with a linear-time recognition algorithm.

5.2 Forbidden induced subgraph characterization of co-bipartite circle graphs

In this section, we establish the forbidden induced subgraph characterization of circle graphs within the class of co-bipartite graphs. Our main result demonstrates that the class of co-bipartite circle graphs coincides with the class of co-bipartite permutation graphs.

Consequently, by applying Gallai's forbidden induced subgraph characterization of comparability graphs [19], we obtain the corresponding characterization for co-bipartite circle graphs. The family of minimal forbidden induced subgraphs of permutation graphs within co-bipartite graphs, denoted by $\mathcal{G}^p_{co}$, is illustrated in *Figure* 5.1. Before presenting the proof of the main theorem, we recall several auxiliary lemmas and preliminaries on circle graphs that will be used in the argument.

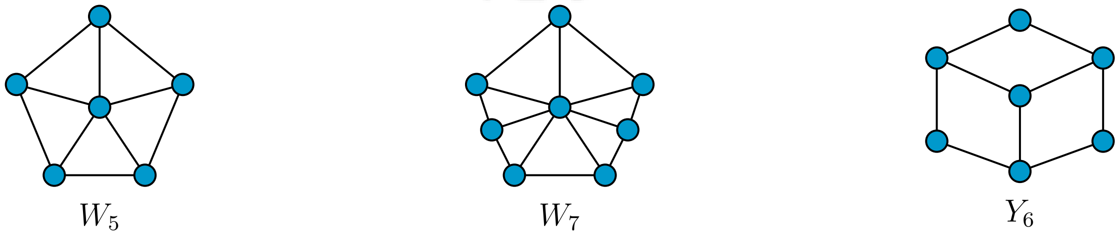


$\overline{C_{2k}}$ for $k \geq 3$ G_1 G_2 G_3

Figure 5.1: Minimal forbidden induced subgraphs for co-bipartite permutation graphs, $\mathcal{G}_c$, where vertices of the same color induce a clique.

The following characterization of circle graphs using local complementation will be used in the proof of the results of this section.

Theorem 5.2.1 ([4]). *A graph G is a circle graph if and only if G has no ℓ-reduction isomorphic to W_5, W_7, or Y_6.*



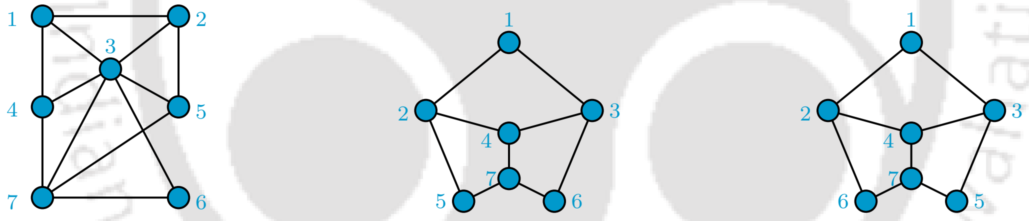
W_5 W_7 Y_6

Figure 5.2: Circle graph obstructions.

The following result shows that no graph in $\mathcal{G}^p_{co}$ is a circle graph. The graphs G_1, G_2, and G_3 are excluded by noting that each admits an l-reduction isomorphic to one of the graphs W_5 or Y_6. For $G \cong \overline{C_{2k}}$, we argue by contradiction: assuming the existence of a 2-uniform word-representant, we restrict it to a suitable closed neighborhood and obtain a permutation graph. We then show that the corresponding permutation word is essentially unique. This rigidity enforces strict alternation constraints, leaving no admissible positions for the remaining vertices. The resulting contradiction implies that $\overline{C_{2k}}$, and hence every graph in $\mathcal{G}^p_{co}$, is not a circle graph.

Lemma 5.2.2. *If $G \in \mathcal{G}^p_{co}$, then G is not a circle graph.*

Proof. Let G be a graph in $\mathcal{G}^p_{co}$. Consider the vertex set of G as provided in *Figure* 5.1.
Case 1 (local complementation of G_1 leading to W_5). Suppose $G \cong G_1$. It follows that $G * 7\ 6$ contains an induced subgraph isomorphic to W_5.
Case 2 (local complementation of G_2 leading to Y_6). Suppose $G \cong G_2$. It follows that $G * 7\ 1$ contains an induces subgraph isomorphic to Y_6.
Case 3 (local complementation of G_3 leading to Y_6). Suppose $G \cong G_3$. It follows that $G * 1\ 2\ 3$ contains an induces subgraph isomorphic to Y_6. Therefore, for the above three cases, by *Theorem 5.2.1*, it follows that G is not a circle graph. Hence, suppose that $G \cong \overline{C_{2k}}$ for some $k \geq 3$. Assume, for the sake of contradiction, that G is a circle graph. This implies that there exists a 2-uniform word-representant of G, say w_G. Consider the graph $H \cong G[N_G[1]]$. By *Theorem 1.6.4* and the hereditary nature of circle graphs, it follows that H is a permutation graph. Let w_H denote the subword of w_G obtained by restricting its alphabet to the vertex set of H. Then it is evident that w_H is a 2-permutation word-representant of H. By *Proposition 3*, we may assume that the first letter of w_H is 3. We claim that



$G_1 * 7\ 6$ $G_2 * 7\ 1$ $G_3 * 1\ 2\ 3$

Figure 5.3: Local complementations of G_1, G_2, and G_3 yield W_5 or Y_6.

$$w_H = 3\ 5\ 4\ 7\ 6\ \cdots\ (2k-1)\ (2k-2)\ 1\ 4\ 3\ 6\ 5\ \cdots\ (2k-2)\ (2k-3)\ (2k-1)\ 1$$

is the unique permutation word-representant of H, up to cyclic shifts and reversals.

The proof of the claim proceeds by induction on k. For the base case, consider $k = 3$. Since the first letter of w_H is 3, the two occurrences of 4 in w_H must appear immediately after the first occurrence and immediately before the second occurrence of 3, respectively. Similarly, the two occurrences of 5 in w_H must appear between the first occurrences and after the second occurrences of 3 and 4, respectively. Finally, the two occurrences of 1 are placed after the first and second occurrences of 3, 4, and 5, respectively. Therefore,

$$w_H = 3\ 5\ 4\ 1\ 4\ 3\ 5\ 1.$$

This establishes the base case of the induction. Assume that the claim holds for $k = i - 1$ for some $i > 4$. Then, we have

$$w_H = 3\ 5\ 4\ 7\ 6\ \cdots\ (2i-3)\ (2i-4)\ 1\ 4\ 3\ 6\ 5\ \cdots\ (2i-4)\ (2i-5)\ (2i-3)\ 1.$$

Now, consider $k = i$. It suffices to show that there exists a unique way, up to cyclic shifts and reversals, to place the vertices $2i - 2$ and $2i - 1$ in w_H. The vertex $2i - 2$ is adjacent to all vertices except $2i - 3$ and $2i - 1$. Therefore, if the first occurrence of $2i - 2$ is placed before the first occurrence of $2i - 3$, then, to maintain the non-alternation between $2i - 2$ and $2i - 3$, the second occurrence of $2i - 2$ must be placed after the second occurrence of $2i - 3$. However, this placement would result in a non-alternation between $2i - 2$ and $2i - 4$, yielding a contradiction. Hence, the unique placement of vertex $2i - 2$ is as follows:

$$w_H = 3\ 5\ 4\ 7\ 6\ \cdots\ (2i-3)\ (2i-4)\ (2i-2)\ 1\ 4\ 3\ 6\ 5\ \cdots\ (2i-4)\ (2i-5)\ (2i-2)\ (2i-3)\ 1.$$

Similarly, the vertex $2i - 1$ is adjacent to all vertices except $2i - 2$ and $2i$. Therefore, if the first occurrence of $2i - 1$ is placed after the first occurrence of $2i - 2$, then, to maintain the non-alternation between $2i - 1$ and $2i - 2$, the second occurrence of $2i - 1$ must be placed before the second occurrence of $2i - 2$. However, this placement would result in a non-alternation between $2i - 1$ and $2i - 3$, yielding a contradiction. Hence, the first occurrence of $2i - 1$ must be placed before the first occurrence of $2i - 2$.

Furthermore, if the first occurrence of $2i - 1$ is placed before the first occurrence of $2i - 4$, then, to maintain the alternation between $2i - 1$ and $2i - 4$, the second occurrence of $2i-1$ must be placed before the second occurrence of $2i-4$. This, however, would lead to an alternation between $2i-1$ and $2i-2$, again a contradiction. Therefore, the unique placement of the vertex $2i-1$ is as follows:

$$w_H = 3\ 5\ 4\ 7\ 6\ \cdots\ (2i-4)\ (2i-1)\ (2i-2)\ 1\ 4\ 3\ 6\ 5\ \cdots\ (2i-2)\ (2i-3)\ (2i-1)\ 1.$$

Hence, the placement of the vertices $2i-2$ and $2i-1$ in w_H is unique up to cyclic shifts and reversals, thereby proving the claim of the induction hypothesis.

Now consider w_G. Since w_H is a subword of w_G, we can obtain w_G by appropriately placing the vertices 2 and $2k$ in w_H. The vertex 2 is adjacent to all vertices except 1 and 3. Suppose the first occurrence of 2 is placed before the first occurrence of 3. Then, to maintain the non-alternation between 2 and 3, the second occurrence of 2 must appear after the second occurrence of 3. Furthermore, to preserve the non-alternation between 2 and 1, the second occurrence of 2 must also appear after the second occurrence of 1. However, such a placement would create a non-alternation between 2 and another vertex, leading to a contradiction. Hence, the first occurrence of 2 must be placed after the first occurrence of 3.

If, instead, the first occurrence of 2 is placed after the first occurrence of 5, then to maintain the alternation between 2 and 5, the second occurrence of 2 must also appear after the second occurrence of 5. This, however, results in a non-alternation between 2 and 4, yielding a contradiction. Therefore, the first occurrence of 2 must lie between the first occurrences of 3 and 5. Consequently, to maintain the non-alternation between 2 and 1, and alternation between 2 and all other vertices except 1 and 3, the second occurrence of 2 must be placed between the first occurrence of $2i-2$ and 1. Hence,

$$w_G|_{V(G)\setminus\{2k\}} = 3\ 2\ 5\ 4\ 7\ 6\ \cdots\ (2k-1)\ (2k-2)\ 2\ 1\ 4\ 3\ 6\ 5\ \cdots\ (2k-3)\ (2k-1)\ 1.$$

The vertex $2k$ is adjacent to all vertices except $2k-1$ and 1. Suppose that the first occurrence of $2k$ appears before the first occurrence of $2k-1$. Then, to maintain the non-alternation between $2k$ and $2k-1$, either the second occurrence of $2k$ must appear before the first occurrence of $2k-1$, or after the second occurrence of $2k-1$. However, in both cases, a non-alternation arises between $2k$ and $2k-2$, yielding a contradiction. Therefore, assume that the first occurrence of $2k$ appears after the first occurrence of $2k-1$. In this case, to preserve the non-alternation between $2k$ and $2k-1$, and the alternation between $2k$ and $2k-3$, the second occurrence of $2k$ must occur between the second occurrences of $2k-3$ and $2k-1$. Furthermore, to maintain the non-alternation between 1 and $2k$, the first occurrence of $2k$ must appear after the first occurrence of 1. This, however, introduces a non-alternation between $2k$ and 2, again leading to a contradiction. Hence, vertex $2k$ cannot appear anywhere in w_G, contradicting the assumption that G is 2-word-representable. Therefore, G is not a circle graph. This completes the proof of the lemma. □

Theorem 5.2.3. *Let G be a co-bipartite graph. Then G is a circle graph if and only if it is a permutation graph.*

Proof. Since permutation graphs form a subclass of circle graphs, the converse follows trivially. Thus, it suffices to prove the forward direction. Suppose that G is a circle graph. For the sake of contradiction, assume that G is not a permutation graph. Then, by the characterization of permutation graphs within co-bipartite graphs in terms of minimal forbidden induced subgraphs, G contains some graph in $\mathcal{G}^p_{co}$. However, by *Lemma* 5.2.2 and the hereditary property of circle graphs, it follows that G is not a circle graph, a contradiction. □

Corollary 7. Let G be a co-bipartite graph. Then G is a circle graph if and only if it contains no graph in $\mathcal{G}^p_{co}$ as an induced subgraph.

Having established a forbidden induced subgraph characterization of co-bipartite circle graphs, we now turn our attention to the broader class of word-representable co-bipartite graphs. Instead of working directly with graph orientations, we adopt a matrix-theoretic perspective, which enables us to leverage existing characterizations and algorithms for circular-ones-type properties.

5.3 Forbidden induced subgraph characterization of word-representable co-bipartite graphs

For a co-bipartite graph $G = (X, Y)$, we define $M(G) = \{m_{ij}\}$ to be a $(0,1)$-matrix whose rows correspond to the partition X and whose columns correspond to the partition Y, where $m_{ij} = 1$ if and only if the corresponding row vertex and column vertex are adjacent in G. In this section, we establish a connection between the circularly compatible ones property of matrices, introduced by Tucker in [43], and semi-transitive co-bipartite graphs, thereby providing a forbidden induced subgraph characterization of the latter. Since the characterization of word-representable co-bipartite graphs given by Das and Ramesh in [13] involves monotonicity, we are naturally motivated to consider the circularly compatible ones property, which also incorporates a monotonicity condition.

In [40], Safe provided a minimal forbidden submatrix characterization and a linear-time recognition algorithm for the circularly compatible ones property. Consequently, by establishing a connection between semi-transitive co-bipartite graphs and the circularly compatible ones property of matrices, we obtain a minimal forbidden induced subgraph characterization for semi-transitive co-bipartite graphs.Throughout this section, for any two vertices x and y in an oriented graph, we write $x \to y$ if there exists an arc between x and y in the given orientation.

For a $(0,1)$-matrix M of size $m \times n$, we define $CG(M) = (X, Y)$ to be the co-bipartite graph in which the bipartition (X, Y) corresponds to the rows and columns of M, respectively. We take $X = \{1, \ldots, m\}$ and $Y = \{1, \ldots, n\}$. A vertex $i \in X$ is adjacent to a vertex $j \in Y$ if and only if the corresponding entry of M is 1. By definition, $CG(M)$ is isomorphic to $CG(M^t)$.

Lemma 5.3.1. *If $F \in \mathcal{F}^{\infty}_{\text{CCO}}$, then $CG(F)$ is not semi-transitive.*

Proof. Let $F \in \mathcal{F}^{\infty}_{\text{CCO}}$. Since $CG(F)$ is isomorphic to $CG(F^t)$, it suffices to prove that $CG(F)$ is not semi-transitive for $F \in \{M_I^*(k), \overline{M_I^*(k)} \colon k \geq 3\} \cup \{Z_2^*, Z_3^*, Z_4^*, Z_5, \overline{Z_2^*}, \overline{Z_4^*}\}$. If $F = \overline{M_I^*(k)}$ for some $k \geq 3$, then one can verify that $CG(F) \cong \overline{C_{2k}} \times K_1$. When $F \in \{\overline{Z_2^*}, \overline{Z_4^*}\}$, we have $CG(F) \cong G_1 \times K_1$, and when $F \in \{Z_3^*, Z_4^*\}$, we have $CG(F) \cong G_2 \times K_1$. In all these cases, $CG(F) \cong H \times K_1$ for some $H \in \mathcal{G}^p_{co}$. Hence, by *Theorem 1.6.4*, it follows that $CG(F)$ is not semi-transitive. Finally, if $F \in \{Z_2^*, Z_5\}$, then by *Theorem 9* of [7], $CG(F)$ is not semi-transitive.

Suppose that $F = M_I^*(k)$ for some $k \geq 3$. Assume, for the sake of contradiction, that $CG(F)$ admits a semi-transitive orientation. Let the rows and columns of F be denoted by $\{r_1, r_2, \ldots, r_k\}$ and $\{c_1, c_2, \ldots, c_{k+1}\}$, respectively. Define linear orders $\preccurlyeq_r$ and $\preccurlyeq_c$ on the rows and columns of F as follows: for any two vertices r_i and r_j, we write $r_i \preccurlyeq_r r_j$

if $r_i \to r_j$ in the semi-transitive orientation; and for any two vertices c_i and c_j, we write $c_i \preccurlyeq_c c_j$ if $c_i \to c_j$ in the semi-transitive orientation. By *Theorem 1.4.6*, it follows that F does not have the circular-ones property. Hence, there exists at least one row-vertex, say r_i, that is not a circular interval of $\preccurlyeq_c$. Consequently, there exist four column-vertices $c_t \preccurlyeq_c c_u \preccurlyeq_c c_v \preccurlyeq_c c_w$ such that either $c_t, c_v \notin N_{CG(F)}(r_i)$ and $c_u, c_w \in N_{CG(F)}(r_i)$, or $c_t, c_v \in N_{CG(F)}(r_i)$ and $c_u, c_w \notin N_{CG(F)}(r_i)$. Since reversing all edges in a semi-transitive orientation also yields a semi-transitive orientation, there are four possible configurations, as illustrated in *Figure 5.4*. In each case, the configuration induces either a shortcut or a directed cycle, which contradicts the assumption that the orientation is semi-transitive. This completes the proof of the lemma. ☐

Figure 5.4: Shortcut possibilities

Theorem 5.3.2. *Let G be a co-bipartite graph. G is semi-transitive if and only if $M(G)$ has circularly compatible ones property.*

Proof. Let G be a co-bipartite graph, and let (X, Y) be the bipartition of the vertex set $V(G)$. Suppose that G is semi-transitive. Assume, for the sake of contradiction, that $M(G)$ does not have the circularly compatible ones property. Then, by Theorem 1.4.14, $M(G)$ contains some matrix in $\mathcal{F}^{\infty}_{\mathrm{CCO}}$ as a configuration. However, by Lemma 5.3.1 and the hereditary nature of semi-transitive graphs, it follows that G is not semi-transitive, a contradiction.

Now suppose that $M(G)$ has the circularly compatible ones property. The idea of the proof is to orient the graph according to the circularly compatible ones biorder and to show that the resulting orientation is semi-transitive. Observe that the matrix $M(G)$ falls into one of the following cases:

1. $M(G)$ has either no trivial rows or no trivial columns.

2. $M(G)$ has an all-0s row and an all-0s column.

3. $M(G)$ has an all-1s row and an all-1s column.

Case 1: Assume that $M(G)$ has either no trivial rows or no trivial columns. Without loss of generality, we may assume that $M(G)$ has no trivial rows (since $CG(M) \cong CG(M^t)$). Therefore, by Theorems 1.4.14 and 1.4.11, $M(G)$ admits a monotone circular biorder ($\preccurlyeq_r, \preccurlyeq_c$). Let $\{r_1, r_2, \ldots, r_m\}$ and $\{c_1, c_2, \ldots, c_n\}$ denote the sets of rows and columns of M, listed in ascending order with respect to $\preccurlyeq_r$ and $\preccurlyeq_c$, respectively. By definition, we then have $X = \{r_1, r_2, \ldots, r_m\}$ and $Y = \{c_1, c_2, \ldots, c_n\}$. We define the orientation $\prec$ of the graph induced by the biorder ($\preccurlyeq_r, \preccurlyeq_c$) as follows:

- For all $r_i, r_j \in X$, $r_i \prec r_j$ if and only if $r_i \preccurlyeq_r r_j$.
- For all $c_i, c_j \in Y$, $c_i \prec c_j$ if and only if $c_i \preccurlyeq_c c_j$.
- If $r_i = [d_i, e_i]_{\preccurlyeq_c}$ with $d_i \preccurlyeq_c e_i$, then $r_i \prec x$ for all x such that $d_i \preccurlyeq_c x \preccurlyeq_c e_i$.
- If $r_i = [d_i, e_i]_{\preccurlyeq_c}$ with $e_i \prec_c d_i$, then $x \prec r_i$ for all $x \preccurlyeq_c e_i$ and $r_i \prec x$ for all x with $d_i \preccurlyeq_c x$.

Before proceeding with the proof, we introduce some definitions. Let r_i be a vertex that belongs to X. We call r_i a *linear-row vertex* if $r_i = [d_i, e_i]_{\preccurlyeq_c}$ with $d_i \preccurlyeq_c e_i$, and a *circular-row vertex* if $r_i = [d_i, e_i]_{\preccurlyeq_c}$ with $e_i \prec_c d_i$.

Claim 1. If r_i is a linear-row vertex and r_j is a circular-row vertex, then $r_i \prec r_j$. Moreover, if $r_i = [d_i, e_i]_{\preccurlyeq_c}$ and $r_j = [d_j, e_j]_{\preccurlyeq_c}$, then $e_j \preccurlyeq_c e_i$.

Suppose, for the sake of contradiction, that $r_j \prec r_i$. This implies $r_j \prec_r r_i$. Let $r_i = [d_i, f_i]^+_{\preccurlyeq_c}$ and $r_j = [d_j, f_j]^+_{\preccurlyeq_c}$ be the unwrapped circular intervals, where $f_i = e_i$ and $f_j = e_j^+$ (see *Definition* 1.4.9). By the definition of the monotone circular property, we have $f_j \preccurlyeq^+_c f_i$, which implies $e_j^+ \preccurlyeq^+_c e_i$, a contradiction. Therefore, $r_i \prec r_j$, as claimed.

Hence, by the definition of the orientation, it follows that r_q is a linear-row vertex for all $1 \leq q \leq i$, and a circular-row vertex for all $j \leq q \leq m$. This implies that $f_1 = e_1$ and $f_m = e_m^+$. Therefore, by the definition of the monotone circular property (see (*iii*) of *Definition 1.4.10*), it follows that $f_m \preccurlyeq^+_c e_1^+$. Consequently, $e_m^+ \preccurlyeq^+_c e_1^+$, which in turn implies that $e_m \preccurlyeq_c e_1$. Hence, we obtain the following ordering:

$$e_j \preccurlyeq_c \cdots \preccurlyeq_c e_m \preccurlyeq_c e_1 \preccurlyeq_c \cdots \preccurlyeq_c e_i.$$

Therefore, $e_j \preccurlyeq_c e_i$. This completes the proof of the claim.

Claim 2. Let $v_1 \to v_2 \to \cdots \to v_p$, *denoted by* $\vec{S}$, *be a directed path with respect to the orientation* $\prec$. *If* $\vec{S}$ *does not contain any linear-row vertex, then* $\vec{S}$ *induces a transitive tournament.*

The claim follows directly if $\vec{S}$ does not contain any vertex from X. Therefore, suppose that $\vec{S}$ contains at least one circular-row vertex. Since $\vec{S}$ trivially induces a transitive tournament when $p < 3$, assume that $p \geq 3$. Consider any three vertices v_a, v_b, v_c in $\vec{S}$ such that $v_a \to v_b \to v_c$ for some $1 \leq a < b < c \leq p$. Suppose, for the sake of contradiction, that $v_a \not\to v_c$. Let v_a be a circular-row vertex, where $v_a = [d_a, e_a]_{\preccurlyeq_c}$. This implies that $v_c \in Y$. If v_b is also a circular-row vertex, where $v_b = [d_b, e_b]_{\preccurlyeq_c}$, then by the monotone circular property, we have $d_a \preccurlyeq_c d_b \preccurlyeq_c v_c$. This implies that $v_a \to v_c$, a contradiction. If $v_b \in Y$, then by the definition of the orientation, we again have $v_a \to v_c$, a contradiction.

Suppose that $v_a \in Y$. This implies that v_c is a circular-row vertex, where $v_c = [d_c, e_c]_{\preccurlyeq_c}$. If v_b is also a circular-row vertex, where $v_b = [d_b, e_b]_{\preccurlyeq_c}$, then by the monotone circular property, we have $v_a \preccurlyeq_c e_b \preccurlyeq_c e_c$. This implies that $v_a \to v_c$, a contradiction. If $v_b \in Y$, then by the definition of the orientation, we again have $v_a \to v_c$, a contradiction. Therefore, for any three vertices v_a, v_b, v_c in $\vec{S}$ such that $v_a \to v_b \to v_c$ for some $1 \leq a < b < c \leq p$, we have $v_a \to v_c$. Hence, $v_1 \to v_2 \to \cdots \to v_p$ induces a transitive tournament. This completes the proof of the claim.

We claim that the given orientation $\prec$ is semi-transitive. Suppose, for the sake of contradiction, that the given orientation is not semi-transitive. Let $\vec{S}$ be a shortcut path of minimal length, $\vec{S} = v_1 \to v_2 \to \cdots \to v_p$, where $v_1 \to v_p$ is the shortcut edge and $p \geq 4$. By *Claim 2*, it is evident that $\vec{S}$ contains at least one linear-row vertex.

Suppose that k_1 is the greatest integer in $[p]$ such that v_{k_1} is a linear-row vertex. Then, by *Claim 1* and by the definition of the orientation, v_i is also a linear-row vertex for all $1 \leq i \leq k_1$. Similarly, let k_2 be the smallest integer in $[p]$ such that $v_{k_2} \in Y$. Therefore, by *Claim 1* and the definition of the orientation, v_j is a circular-row vertex for all $k_1 < j < k_2$.

Now suppose, for the sake of contradiction, that $k_2 > 3$. We claim that the path $v_1 \to v_{k_2-1} \to \cdots \to v_p$, denoted by $\vec{S'}$, induces a shortcut path with $v_1 \to v_p$ being the shortcut edge. Assume that $\vec{S'}$ contains more than three vertices. Then, for the sake of contradiction, suppose that $\vec{S'}$ induces a transitive tournament. Let v_t be a vertex such that $k_2 \leq t \leq p$. If $v_t \in X$, then $v_l \to v_t$ for all $1 \leq l < k_2$. Suppose $v_t \in Y$. Since $v_1 \to v_t$ and $v_{k_2-1} \to v_t$ for all $k_2 \leq t \leq p$, by the monotone circular property, we obtain the following ordering:

$$d_1 \preccurlyeq_c \cdots \preccurlyeq_c d_{k_2-1} \preccurlyeq_c v_t \preccurlyeq_c e_1 \preccurlyeq_c \cdots \preccurlyeq_c e_{k_1}.$$

By the definition of the orientation, $v_l \to v_t$ for all $1 \leq l < k_2$ and $k_2 \leq t \leq p$. This implies that $\vec{S}$ induces a transitive tournament, a contradiction. Suppose that $\vec{S'}$ contains at most three vertices. Then $v_{k_2} = v_p$. Since $p = k_2 \geq 4$, and $v_1 \to v_p$ as well as $v_{k_2-1} \to v_p$, by the monotone circular property, we have $v_l \to v_p$ for all $1 \leq l < k_2$. This again implies that $\vec{S}$ induces a transitive tournament, a contradiction. Therefore, $\vec{S'}$ must induce a shortcut path. However, since $k_2 > 3$, $\vec{S'}$ contains fewer vertices than $\vec{S}$, contradicting the choice of $\vec{S}$. Hence, $k_2 \leq 3$.

Now suppose that k_3 is the greatest integer in $[p]$ such that $v_{k_3} \in Y$. Assume, for the sake of contradiction, that $k_3 < p$. If $k_3 = k_2$, then by the definition of the monotone circular property and the second part of *Claim 1*, we obtain the following ordering:

$$d_1 \preccurlyeq_c \cdots \preccurlyeq_c d_{k_2-1} \preccurlyeq_c v_{k_2} \preccurlyeq_c e_{k_1+1} \preccurlyeq_c \cdots \preccurlyeq_c e_p \preccurlyeq_c e_1 \preccurlyeq_c \cdots \preccurlyeq_c e_{k_1}.$$

This implies that $\vec{S}$ induces a transitive tournament, which is a contradiction.

Now assume that $k_3 > k_2$. Then, by *Claim 1*, it follows that v_l is a circular-row vertex for all $k_3 < l \leq p$. Therefore, by the definition of the monotone circular property and the second part of *Claim 1*, we have $d_1 \preccurlyeq_c v_{k_3} \preccurlyeq_c e_1$. Hence, $v_1 \to v_{k_3}$. We claim that the path $v_1 \to v_2 \to \cdots \to v_{k_3}$, denoted by $\vec{S'}$, induces a shortcut path with $v_1 \to v_{k_3}$ being the shortcut edge. Assume that $k_3 \geq 4$. Suppose, for the sake of contradiction, that $\vec{S'}$ induces a transitive tournament. Let v_t be a vertex for some $1 < t \leq k_3$. If $v_t \in X$, then $v_t \to v_l$ for all $k_3 < l \leq p$. Suppose $v_t \in Y$. Since $v_{k_3} \to v_{k_3+1}$, by the definition of the orientation and the monotone circular property, we obtain $v_t \to v_l$ for all $k_3 < l \leq p$. This implies that $\vec{S}$ induces a transitive tournament, a contradiction.

Now suppose that $k_3 \leq 3$. This implies that v_1 is a linear-row vertex, $v_2, v_3 \in Y$, and v_l is a circular-row vertex for all $4 \leq l \leq p$. By the definition of the monotone circular property and the second part of *Claim 1*, it follows that $\vec{S}$ induces a transitive tournament, which is a contradiction. Therefore, $\vec{S'}$ induces a shortcut path. However, since $k_3 < p$, $\vec{S'}$ contains fewer vertices than $\vec{S}$, contradicting the choice of $\vec{S}$. Hence, $v_p \in Y$.

As stated in *Claim 2*, the path $v_3 \rightarrow v_4 \rightarrow \cdots \rightarrow v_p$ induces a transitive tournament, since it does not contain any linear-row vertex. Let v_t be a vertex for some $3 \leq t \leq p$. If $v_2 \in Y$ or if v_2 is a circular-row vertex, then by the same argument, it follows that $v_2 \rightarrow v_t$. Since v_1 is a linear-row vertex, if $v_t \in X$, then $v_1 \rightarrow v_t$. Suppose $v_t \in Y$. Since $v_1 \rightarrow v_2$ and $v_1 \rightarrow v_p$, by the definition of the monotone circular property, we have $d_1 \preccurlyeq_c v_t \preccurlyeq_c e_1$. This implies that $v_1 \rightarrow v_t$, which in turn implies that $\vec{S}$ induces a transitive tournament, a contradiction.

Now suppose that v_2 is a linear-row vertex. Since $v_1 \rightarrow v_p$ and $v_2 \rightarrow v_3$, by the definition of the monotone circular property, we have $d_1 \preccurlyeq_c d_2 \preccurlyeq_c v_t \preccurlyeq_c e_1 \preccurlyeq_c e_2$. Therefore, $v_1 \rightarrow v_t$ and $v_2 \rightarrow v_t$. This implies that $\vec{S}$ induces a transitive tournament, a contradiction. Since we have exhausted all possible cases, this contradicts the assumption that the orientation $\prec$ is not semi-transitive. Hence, the given orientation $\prec$ is semi-transitive. This completes the proof of this case.

Case 2: Assume that $M(G)$ has an all-0s row and an all-0s column. Then, by the definition of the circularly compatible ones property, it follows that $M(G)$ has the circular-ones property for both rows and columns. Moreover, since it has an all-0s row and an all-0s column, by *Theorem 1.4.5*, it follows that $M(G)$ has the consecutive-ones property for both rows and columns. Let $(\preccurlyeq_r, \preccurlyeq_c)$ be a circularly compatible-ones biorder such that $\{r_1, r_2, \ldots, r_m\}$ is an ascending order of $\preccurlyeq_r$, with r_m being the all-0s row, and $\{c_1, c_2, \ldots, c_n\}$ is an ascending order of $\preccurlyeq_c$, with c_n being the all-0s column. We define the orientation $\prec$ of the graph induced by the biorder $(\preccurlyeq_r, \preccurlyeq_c)$ as described in *Case 1*.

We claim that the given orientation $\prec$ is semi-transitive. Suppose, for the sake of contradiction, that the given orientation is not semi-transitive. Let $\vec{S}$ be a shortcut path of minimal length, $\vec{S} = v_1 \rightarrow v_2 \rightarrow \cdots \rightarrow v_p$, where $v_1 \rightarrow v_p$ is the shortcut edge and $p \geq 4$. If $\vec{S}$ does not contain either r_m or c_n, then by *Case 1*, it follows that $\vec{S}$ induces a transitive tournament. Therefore, assume that $\vec{S}$ contains either r_m or c_n. By the definition of the orientation, it is evident that $\vec{S}$ cannot contain both r_m and c_n, as there are no outgoing edges from either of these vertices. Moreover, whichever of the two $\vec{S}$ contains must be its sink vertex.

Suppose $v_p = c_n$. Then, by the definition of the orientation, it follows that $v_1 \in Y$. As there are no outgoing edges from the partition Y, it is evident that $\vec{S}$ induces a transitive tournament, a contradiction. Now suppose $v_p = r_m$. Then, by the definition of the orientation, we have $v_1 \in X$. Assume that there exists a vertex $v_i \in Y$ for some $2 \leq i < p$. Since there are no outgoing edges from vertices in Y, it follows that every vertex v_j with $i \leq j \leq p$ also belongs to Y. However, this contradicts the fact that $v_p \in X$. Therefore, $\vec{S}$ does not contain any vertex from the partition Y. Consequently, $\vec{S}$ induces a transitive tournament, a contradiction. Hence, the orientation $\prec$ is semi-transitive.

Case 3: Assume that $M(G)$ has an all-1 row and an all-1 column. This implies that G contains a universal vertex, say v. Therefore, by *Theorem 1.6.4*, G is semi-transitive if and only if $G - v$ is a permutation graph. Suppose, for the sake of contradiction, that $G - v$ is not a permutation graph. Then, by *Corollary 7*, it follows that $G - v$ contains some graph $H \in \mathcal{G}^p_{co}$ as an induced subgraph. Let H' denote the subgraph of G induced by $V(H) \cup \{v\}$.

If $H \cong \overline{C_{2k}}$ for some $k \geq 3$, then it can be easily verified that $M(H')$ contains the same configuration as $M_I^*(k)$. If $H \cong G_1$, then $M(H')$ contains the same configuration as $\overline{Z_2^*}$. If $H \cong G_2$, then $M(H')$ contains the same configuration as Z_3^*. Finally, if $H \cong G_3$, then $M(H')$ contains $M_I^*(3)$ as a configuration.

Therefore, from all the above cases, we conclude that $M(G)$ contains some matrix in $\mathcal{F}_{\text{CCO}}$ as a configuration, which is a contradiction. This completes the proof of the theorem. ☐

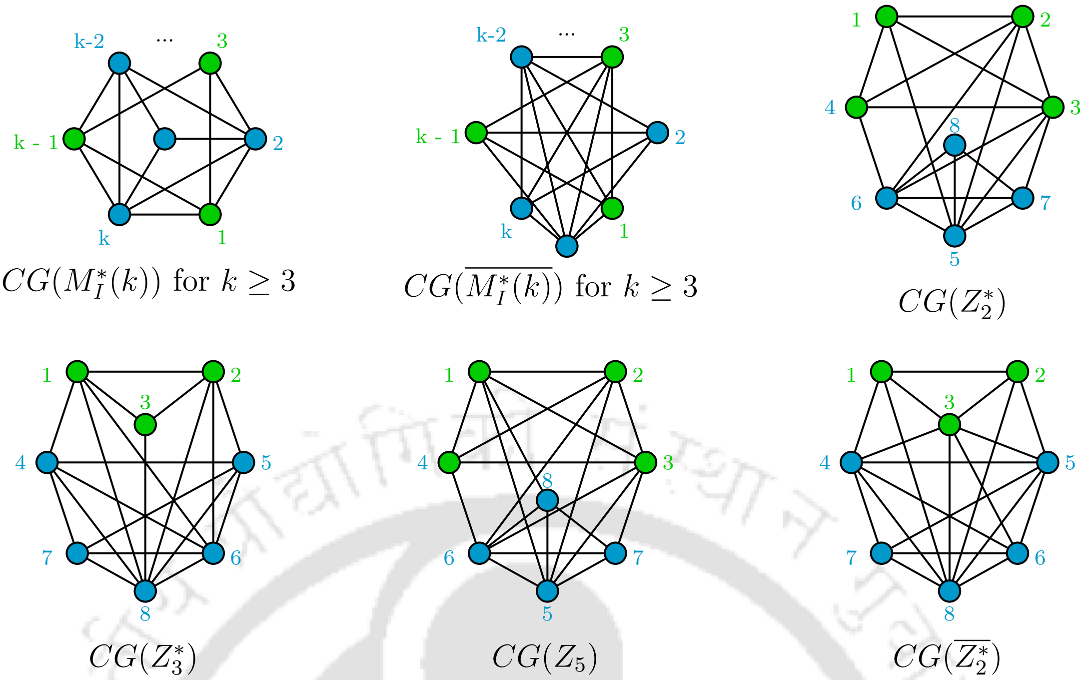
Note that, in the above figure, vertices of the same colour induce a clique.

Lemma 5.3.3. *If $F \in \mathcal{F}^{\infty}_{\text{CCO}}$, then $CG(F)$ contains an induced subgraph isomorphic to some graph in $\mathcal{G}_{co}$.*

Proof. Let F be a matrix in $\mathcal{F}^{\infty}_{\text{CCO}}$. If $F \in \{M_I^*(k), \overline{M_I^*(k)} \colon k \geq 3\} \cup \{Z_2^*, Z_3^*, \overline{Z_2^*}, Z_5\}$, then $CG(F)$ is isomorphic to a graph in $\mathcal{G}_{co}$. If $F \in \{Z_4^*, \overline{Z_4^*}\}$, then $CG(F)$ is isomorphic to $CG(Z_3^*)$. This completes the proof of the lemma. ☐

Lemma 5.3.4. *All graphs in $\mathcal{G}_{co}$ are minimal non-semi-transitive.*

Proof. Let G be a graph in $\mathcal{G}_{co}$. Observe that $M(G) \in \mathcal{F}^{\infty}_{\text{CCO}}$. Therefore, by *Theorems 1.4.14* and *5.3.2*, G is a minimal non-semi-transitive graph. ☐



$CG(M_I^*(k))$ for $k \geq 3$ $CG(\overline{M_I^*(k)})$ for $k \geq 3$ $CG(Z_2^*)$

$CG(Z_3^*)$ $CG(Z_5)$ $CG(\overline{Z_2^*})$

Figure 5.5: Minimal forbidden induced subgraphs for semi-transitive co-bipartite graphs, $\mathcal{G}_{co}$, where vertices of the same color induce a clique.

Theorem 5.3.5. *Let G be a co-bipartite graph. Then G is semi-transitive if and only if it contains none of the graphs in $\mathcal{G}_{co}$ (depicted in Figure 5.5) as an induced subgraph. Moreover, there exists a linear-time algorithm which, given a co-bipartite graph, determines whether it is semi-transitive.*

Proof. Assume that G is semi-transitive. Suppose, for the sake of contradiction, that G contains some graph in $\mathcal{G}_{co}$ as an induced subgraph. Then, by *Lemma 5.3.4* and the hereditary property of semi-transitive graphs, G would be non-semi-transitive—a contradiction.

Conversely, suppose that G contains none of the graphs in $\mathcal{G}_{co}$ as an induced subgraph. Assume, for the sake of contradiction, that G is not semi-transitive. Then, by *Theorem 5.3.2*, the matrix $M(G)$ does not have the circularly compatible ones property. Moreover, by *Theorem 1.4.14*, $M(G)$ contains a matrix $F \in \mathcal{F}_{\text{CCO}}^{\infty}$ as a configuration. Hence, by *Lemma 5.3.3*, G contains an induced subgraph isomorphic to a graph in $\mathcal{G}_{co}$, a contradiction.

Finally, note that for any co-bipartite graph G, $M(G)$ can be obtained in linear time. Hence, by *Theorems 1.4.15* and *5.3.2*, the algorithmic part of the statement follows. This completes the proof. $\square$

6

Scope of future work

In Chapter 2, we obtained upper bounds on the minimum length of word-representants under several graph operations, including edge addition, vertex identification, Cartesian products, and rooted products. An interesting direction for future work is to further tighten these upper bounds for the minimum-length word-representants arising from these operations. Moreover, the problem of determining whether a given word-representant of a graph is of minimum length remains open.

A major open problem arising from Chapter 3 is to obtain a complete characterization of word-representable graphs within broader graph classes, such as circulant graphs of higher regularity. In particular, it remains open whether every k-regular word-representable circulant graph is k-representable for $k \geq 4$.

Chapter 4 and Chapter 5 focus on the word-representability of split graphs and co-bipartite graphs, respectively. Together, these results yield a complete minimal forbidden induced subgraph characterization for the classes $\mathcal{E}_{i,j}$ with $i + j \leq 2$. It is already known that $\mathcal{E}_{3,0}$, the class of 3-colorable graphs, is word-representable. A natural direction for further research is to investigate word-representability in graph classes $\mathcal{E}_{i,j}$ with $i + j \geq 3$ and to obtain minimal forbidden induced subgraph characterizations for these classes.

An interesting open problem is whether the matrix-theoretic framework introduced in these chapters—most notably the circularly compatible ones property, or suitable generalizations thereof—can be extended to yield structural and algorithmic characterizations for these broader graph classes.

The explicit characterization of minimal forbidden induced subgraphs obtained in these chapters opens the door to an enumerative study of word-representable split graphs and word-representable co-bipartite graphs. Such an investigation may yield new combinatorial insights and uncover connections with known integer sequences or well-established combinatorial structures.

The successful application of the circularly compatible ones property to the characterization of semi-transitivity suggests that other matrix properties may play a similar role in the study of word-representability for broader graph classes, such as chordal or weakly chordal graphs.

Bibliography

[1] A. Basu, S. Das, S. Ghosh, and M. Sen. Circular-arc bigraphs and its subclasses. *J. Graph Theory*, 73(4):361–376, 2013. `doi:10.1002/jgt.21681`.

[2] F. Boesch and R. Tindell. Circulants and their connectivities. *J. Graph Theory*, 8(4):487–499, 1984.

[3] K. S. Booth and G. S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using pq-tree algorithms. *J. Comput. System Sci.*, 13(3):335–379, 1976.

[4] A. Bouchet. Circle graph obstructions. *J. Combin. Theory Ser. B*, 60(1):107–144, 1994. `doi:10.1006/jctb.1994.1008`.

[5] B. Broere. Word-representable graphs, 2018. Master's thesis, Radboud Univ. Nijmegen.

[6] B. Broere and H. Zantema. The k-dimensional cube is k-representable. *J. Autom. Lang. Comb.*, 24(1):3–12, 2019.

[7] H. Z. Q. Chen, H. Hameed, and S. Kitaev. On the word-representability of $K_m - K_n$ graphs. *Discuss. Math. Graph Theory*, 2025.

[8] H. Z. Q. Chen, S. Kitaev, and A. Saito. Representing split graphs by words. *Discuss. Math. Graph Theory*, 42(4):1263–1280, 2022.

[9] G.-S. Cheon, J. Kim, M. Kim, and S. Kitaev. Word-representability of Toeplitz graphs. *Discrete Appl. Math.*, 270:96–105, 2019.

[10] I. Choi, J. Kim, and M. Kim. On operations preserving semi-transitive orientability of graphs. *J. Comb. Optim.*, 37(4):1351–1366, 2019.

[11] A. Collins, S. Kitaev, and V. Lozin. New results on word-representable graphs. *Discrete Appl. Math.*, 216:136–141, 2017.

[12] D. G. Corneil, H. Lerchs, and L. S. Burlingham. Complement reducible graphs. *Discrete Appl. Math.*, 3:163–174, 1981.

[13] B. Das and H. Ramesh. Representation number of word-representable co-bipartite graphs. *arXiv Preprint*, 2025. `arXiv:arXiv:2509.03064`.

[14] B. Das and H. Ramesh. Word-representability of co-bipartite graphs. *arXiv Preprint*, 2025. `arXiv:arXiv:2501.10112`.

[15] J. Enright and S. Kitaev. Polygon-circle and word-representable graphs. *Electron. Notes Discrete Math.*, 71:3–8, 2019.

[16] L. Esperet and M. Stehlík. Bipartite complements of circle graphs. *Discrete Math.*, 343(6):111834, 2, 2020. `doi:10.1016/j.disc.2020.111834`.

[17] P. Fleischmann, L. Haschke, T. Löck, and D. Nowotka. Word-representable graphs from a word's perspective. *Acta Inform.*, 61(4):383–400, 2024. `doi:10.1007/s00236-024-00462-y`.

[18] M. Gaetz and C. Ji. Enumeration and extensions of word-representants. *Discrete Appl. Math.*, 284:423–433, 2020.

[19] T. Gallai. Transitiv orientierbare Graphen. *Acta Math. Acad. Sci. Hungar.*, 18:25–66, 1967. `doi:10.1007/BF02020961`.

[20] M. Glen, S. Kitaev, and A. Pyatkin. On the representation number of a crown graph. *Discrete Appl. Math.*, 244:89–93, 2018.

[21] M. C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Academic Press, 1980.

[22] M. M. Halldórsson, S. Kitaev, and A. Pyatkin. Graphs capturing alternations in words. In *Proceedings of the 14th International Conference on Developments in Language Theory*, pages 436–437. Springer, 2010.

[23] M. M. Halldórsson, S. Kitaev, and A. Pyatkin. Alternation graphs. In *Proc. 37th Int. Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, volume 6986 of *LNCS*, pages 191–202. Springer, 2011.

[24] M. M. Halldórsson, S. Kitaev, and A. Pyatkin. Semi-transitive orientations and word-representable graphs. *Discrete Appl. Math.*, 201:164–171, 2016.

[25] Z. Hefty, P. Horn, C. Muir, and A. Owens. Word-representable graphs: orientations, posets, and bounds. *Electron. J. Combin.*, 31(4):Paper No. 4.2, 26, 2024. `doi:10.37236/12806`.

[26] S. Huang, S. Kitaev, and A. Pyatkin. An embedding technique in the study of word-representability of graphs. *Discrete Appl. Math.*, 346:170–182, 2024.

[27] K. Iamthong. Word-representability of split graphs generated by morphisms. *Discrete Appl. Math.*, 314:284–303, 2022.

[28] K. Iamthong and S. Kitaev. Semi-transitivity of directed split graphs generated by morphisms. *J. Combin.*, 14(1), 2023.

[29] S. Kitaev. On graphs with representation number 3. *J. Autom. Lang. Comb.*, 18(2):97–112, 2013.

[30] S. Kitaev. A comprehensive introduction to the theory of word-representable graphs. In *Developments in Language Theory: 21st International Conference, DLT 2017, Liège, Belgium, August 7-11, 2017, Proceedings*, pages 36–67. Springer, 2017.

[31] S. Kitaev, Y. Long, J. Ma, and H. Wu. Word-representability of split graphs. *J. Combin.*, 12(4):725–746, 2021.

[32] S. Kitaev and V. Lozin. *Words and Graphs.* Springer, 2015.

[33] S. Kitaev and A. Pyatkin. On representable graphs. *J. Autom. Lang. Comb.*, 13(1):45–54, 2008.

[34] S. Kitaev and A. Pyatkin. On semi-transitive orientability of triangle-free graphs. *Discuss. Math. Graph Theory*, 43(2):533–547, 2023.

[35] S. Kitaev and A. Pyatkin. On semi-transitive orientability of split graphs. *Inform. Process. Lett.*, 184:106435, 2024.

[36] S. Kitaev and A. Saito. On semi-transitive orientability of kneser graphs and their complements. *Discrete Math.*, 343(8):111909, 2020.

[37] S. Kitaev and S. Seif. Word problem of the perkins semigroup via directed acyclic graphs. *Order*, 25(3):177–194, 2008.

[38] A. V. Petyuk. On word-representability of simplified de Bruijn graphs. *arXiv preprint arXiv:2210.14762*, 2022.

[39] M. D. Safe. Characterization and linear-time detection of minimal obstructions to concave-round graphs and the circular-ones property. *J. Graph Theory*, 93:268–298, 2020.

[40] M. D. Safe. Circularly compatible ones, d-circularity, and proper circular-arc bigraphs. *SIAM J. Discrete Math.*, 35(2):707–751, 2021.

[41] J. Sawada. Generating bracelets in constant amortized time. *SIAM J. Comput.*, 31(1):259–268, 2001.

[42] R. Suchanda and H. Ramesh. Word-representability of split graphs with independent set of size 4. *arXiv preprint arXiv:2507.08483*, 2025.

[43] A. C. Tucker. *Two Characterizations of Proper Circular-arc Graphs*. ProQuest LLC, Ann Arbor, MI, 1969. Thesis (Ph.D.)–Stanford University. URL: `http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:7010537`.

[44] A. C. Tucker. Matrix characterizations of circular-arc graphs. *Pacific J. Math.*, 39:535–545, 1971. URL: `http://projecteuclid.org/euclid.pjm/1102969574`.

[45] A. C. Tucker. A structure theorem for the consecutive 1's property. *J. Combin. Theory Ser. B*, 12(2):153–162, 1972.

[46] V. Vilfred. σ-labelled graphs and circulant graphs [ph.d. thesis]. *Univ. of Kerala, Thiruvananthapuram, India*, 1994.

[47] V. Vilfred. A theory of Cartesian product and factorization of circulant graphs. *Journal of Discrete Mathematics*, 2013(1):163740, 2013. `doi:10.1155/2013/163740`.

[48] D. B. West. *Introduction to Graph Theory.* Prentice Hall, 1996.

[49] X. Yu, Z. Shao, and Z. Li. On the classification and dispersability of circulant graphs with two jump lengths. *Discrete Appl. Math.*, 355:268–286, 2024.

Publications

Publications from Thesis work

1. Eshwar Srinivasan and Ramesh Hariharasubramanian, *Minimum length word-representants of word-representable graphs*, Discrete Appl. Math., **343** (2024), 149–158.
2. Eshwar Srinivasan and Ramesh Hariharasubramanian, *Minimum length wird-representants of graph products*, Discrete Appl. Math., **358** (2024), 91–104.
3. Eshwar Srinivasan and Ramesh Hariharasubramanian, *On semi-transitive orientability of circulant graphs*, Discrete Appl. Math., **377** (2025), 498–509.
4. Eshwar Srinivasan and Ramesh Hariharasubramanian, *Forbidden induced subgraph characterization of word-representable split graphs*, (2025), 22 pages, Communicated.
5. Eshwar Srinivasan and Ramesh Hariharasubramanian, *Forbidden induced subgraph characterization of word-representable co-bipartite graphs*, (2025), 24 pages, Communicated.