Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia

COPPE UFRJ

TWO PROBLEMS IN COMBINATORICS: ROLLER COASTER PERMUTATIONS & THE ERDŐS-SÓS CONJECTURE

Bruno Ramos Lima Netto

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientador: Fábio Happ Botler

Rio de Janeiro
Março de 2022

TWO PROBLEMS IN COMBINATORICS: ROLLER COASTER PERMUTATIONS & THE ERDŐS-SÓS CONJECTURE

Bruno Ramos Lima Netto

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Orientador: Fábio Happ Botler

Aprovada por: Prof. Fábio Happ Botler
Prof. Guilherme Oliveira Mota
Prof. Maycon Sambinelli

RIO DE JANEIRO, RJ – BRASIL
MARÇO DE 2022

Netto, Bruno Ramos Lima

Two Problems in Combinatorics: Roller Coaster Permutations & the Erdős-Sós Conjecture/Bruno Ramos Lima Netto. – Rio de Janeiro: UFRJ/COPPE, 2022.

X, 61 p.: il.; $29,7$cm.

Orientador: Fábio Happ Botler

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2022.

Referências Bibliográficas: p. 58 – 61.

1. Combinatorics. 2. Roller Coaster Permutations. 3. Extremal Graph Theory. 4. Tree Embeddings. I. Botler, Fábio Happ. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

To π, or not to π.

Acknowledgements

The content of this part is written in Portuguese, my native language.

Primeiramente, agradeço especialmente a minha familia. Aos meus pais, Sergio Lima Netto e Luciana Ramos da Silva, por serem fontes infindáveis de inspiração e carinho. Também incluo aqui meus padrastos, Isabela Chaves Santos e Thiago Nogueira, por todo amor e cuidado. Estendo esse agradecimento aos meus avós, partenos, maternos e madristinos[1], e a todos que considero tios e primos. Não posso deixar de lado todos os meus irmãos: Carol, Renata, Manuela, Daniel e Clara. Amo todos vocês.

Além disso, aqui faço um agradecimento muito especial aos meus amigos e companheiros de banda, João Gabriel Lopes e João Camargo Mello: **Acabou a banda**. Separo aqui um carinho especial ao meu amigo João Paulo Coelho pela companhia durante as madrugadas.

Agradeço também a todos os meus preciosos amigos da faculdade: Aloizio Macedo, Gabriel Sanfins, Leonardo Gama, Matheus Fontoura, Patrick Silva, Rodrigo Lima, Thiago Holleben e Vitor Luiz. Bem como, agradeço imensamente a todos aqueles que fazem ou fizeram parte do *Powerpointers*: Ricardo Turano, Pedro Xavier, Iago Leal, Alexandre Moreira, Tiago Vital, Gabriel Pessoa, e Ivani Ivanova.

Além disso, agradeço aos meus professores que marcaram minha jornada. Bernardo Freitas, Hamidreza Anbarlooei e Thiago Hartz.

Agradeço sinceramente aos membros da banca, professores Guilherme Oliveira Mota e Maycon Sambinelli.

Agradeço ao meu orientador, Fábio Botler, por ter estado sempre presente durante o meu mestrado e nossa pesquisa. Por todos os encontros semanais durante a pandemia que serviram de suporte em meio ao caos, e, por toda ajuda com a confecção desta dissertação.

Por fim, agradeço à Maria Eduarda Luporini, por todo apoio, carinho, paciência, inspiração e amor. E por ter sido minha maior companheira durante esse tempo de pesquisa. Obrigado por tudo.

O presente trabalho foi realizado com apoio da Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo nº 133347/2020-6.

[1]Pais da minha madrasta.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

DOIS PROBLEMAS EM COMBINATÓRIA: PERMUTAÇÕES MONTANHAS-RUSSAS & A CONJECTURA DE ERDŐS-SÓS

Bruno Ramos Lima Netto

Março/2022

Orientador: Fábio Happ Botler

Programa: Engenharia de Sistemas e Computação

Combinatória, como o nome sugere, é a ciência das combinações. Durante este trabalho, exploramos duas áreas principais da Combinatória: Permutações e Grafos. Embora muito próximas, tratamos as áreas de maneira separada.

Em Permutações, abordamos o problema das *permutações Montanhas-Russas*, que são permutações que maximizam, junto às suas subsequências, o quanto elas alternam entre subidas e descidas. Essa classe especial de permutações foi introduzida em 2013 por T. Ahmed e H. Snevily que, além de sua definição, levantaram diversas conjecturas sobre sua estrutura. Neste trabalho, apresentamos uma definição alternativa para essas permutações, bem como um modelo de Programa Linear Inteiro associado para encontrá-las. Através desse modelo, conseguimos obter novos exemplos de Montanhas-Russas, e, através de um modelo que adota restrições com base em certas conjecturas estruturais, obtivemos novas candidatas para Montanhas-Russas. Por fim, motivamos o estudo desse problema sob outras óticas, apresentando conjecturas relacionadas a outras representações de permutações.

Em Grafos, apresentamos uma vasta coleção de resultados a respeito da famosa Conjectura de Erdős-Sos presentes na literatura. Em 1962, P. Erdős e V. Sós conjecturaram que, para inteiros positivos n, k, todo grafo com n vértices e pelo menos $n(k-2)/2+1$ arestas contém, como subgrafo, todas as árvores com k vértices. Neste trabalho, dividimos tais resultados em quatro direções principais, cada uma representando um enfraquecimento diferente dessa conjectura, com o objetivo de apontar possiveis direções para contribuições ao estado da arte com respeito a esse problema.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

TWO PROBLEMS IN COMBINATORICS: ROLLER COASTER PERMUTATIONS & THE ERDŐS-SÓS CONJECTURE

Bruno Ramos Lima Netto

March/2022

Advisor: Fábio Happ Botler

Department: Systems Engineering and Computer Science

Combinatorics is, as its name suggests, the science of combinations. Throughout this work, we focus on two main sub-areas of Combinatorics: Permutations and Graphs. Albeit two very related fields, we deal with them in a separate matter.

In Permutations, we study the problem of the Roller Coaster permutations, which are permutations that maximize, along with its subsequences, the number of ascents and descents. This special class of permutations was introduced in 2013 by T. Ahmed and H. Snevily which, besides its definition, conjectured many properties on its structure. In this work, we present an alternative and equivalent definition for the Roller Coaster permutations, together with an Integer Linear Programming model to find such permutaions. With this model, we obtained new examples for Roller Coasters, and, with an extended version of this model, based on certain structural conjectures, we obtained new candidates for Roller Coasters. Lastly, we motivated the study of this problem from another point of view, presenting conjectures related to other representations of permutations.

In Graphs, we present an extensive collection of results with respect to the famous Erdős-Sós Conjecture, which are found in the literature. In 1962, P. Erdős and V. Sós conjectured that for positive integers n, k, every graph on n vertices and at least $n(k-2)/2+1$ edges, contains every tree on k vertices. In this work, we divided the partial results in four main directions, each representing a different weakening of this conjecture, with the objective of pointing out possible directions to contribute to the state of the art with respect to this problem.

Contents

List of Figures

List of Tables

Prologue

The present work is a study on two different problems in Combinatorics.

Our research began with the study of the problem concerning Roller Coaster permutations, which the main objective is to find a permutation that maximizes a certain function of “alternation”. We obtained new examples of such permutations as well as an equivalent and alternative definition that allowed us to find it much faster then with the classical definition.

After exploring the field of Extremal Graph Theory, we began studying the classical theorems and ideas, in preparation to understand the main results on the Erdős–Sós Conjecture. Our main objective was to find similar patterns in the proofs to point out directions to contribute to the state of the art with respect to this problem. We hope that this work might be of good use for future students and researchers interested in either of those combinatorial problems.

This work is structured as follows. In Chapter 1, we motivate both of the abstract objects that are the main subject of each of the problems. For the first one, Permutations, besides some formalities as notation and simple properties, we present a remarkable enumeration problem, known to be solved since 1879, concerning *alternating permutations*, which are permutations that switches between ascents and descents, and are believed to be closely related with the permutations studied in the following chapter. For the second object, Graphs, we present its historical origin and classical definitions of Graph Theory. Since its notation is not consistent across the literature, it is important that we specify the notation used in this work. Together with these formalities, we present some enumerative results on special classes of graphs, known as *caterpillars* and *spiders*, which are special cases treated in the last chapter of this work. Lastly, we present a section on representations of permutations, where we explain different possible representation of permutations as graphs, which is explored in the final part of Chapter 2.

In Chapter 2, we explore a problem concerning *Roller Coaster permutations*. Informally, they are permutations that, along with all of its subsequences, maximizes the amount of changes between ascents and descents. We give the formal definition, known results in the literature, as well as important structural conjectures. Our main result is an alternative and equivalent definition that, besides yielding to a faster algorithm to determine Roller Coaster permutations, motivated the development of an Integer Linear Programming (ILP) model to find them. We present the results obtained with said model, showing an improvement to the known Roller Coaster permutations in the literature. We also present an extended version of this ILP model that considers constraints based on a conjectured structure of the Roller Coasters, which improved the performance of the ILP model. With the extended model, we obtained candidates for even bigger permutations, improving known lower bounds. Finally, we conclude this chapter presenting the Roller Coaster permutation problem in different points of view that arises when one consider different representations of a permutation.

In Chapter 3, we transition to the second problem by presenting some classical results on Extremal Graph Theory, as a preparation to the study of one of the main conjectures in this field, the Erdős–Sós Conjecture. We present a motivational problem concerning the maximal cardinality of a *sum-free* set, and then proceed to the historical result of W. Mantel from 1907, which determines the maximum number of edges in a graph that avoid triangles. We proceed to other theorems on the number of edges of a graph that cannot avoid different structures, for example, complete graphs, even cycles, paths and, finally, trees. The problem of avoiding trees is the main motivation of the Erdős-Sós Conjecture, which states that a graph on n vertices and $n(k-2)/2+1$ edges contains all trees on k vertices. We present different positive results on this conjecture for specific graphs or specific trees, which are present in the literature. We divide these results in four different directions, each representing a different weakening of the conjecture, with the objective of pointing out possible directions to contribute to the state of the art with respect to this problem. For every statement in each direction, we present a sketch of the proof, whenever we found one in the literature. We conclude the chapter presenting a tentative extension of one of the present proofs, that we, unfortunately, were not able to finish.

Lastly, In Chapter 4, we present the concluding remarks as well as possible future work on both problems studied in this dissertation.

1 Introduction

The purpose of this chapter is to make this dissertation as self-contained as possible, and to provide a quick reference for the readers that are not familiar with some concepts in Combinatorics.

In Section 1.1, we explore the concept of Permutations, presenting the main notation and definitions that are studied in Chapter 2, together with an enumerative result on alternating permutations due to D. André [6].

In Section 1.2, we present some important, yet introductory, concepts from Graph Theory, as well as some important classes of graphs, that are explored further in the context of Extremal Graph Theory in Chapters 3. We also present known enumerative results of some special classes of graphs.

Finally, in Section 1.3, we explore some representations of permutations through different kinds of graphs, relating these two objects in attempt to exhibit special properties from permutations that would otherwise be unapparent.

Throughout the text, we denote the set of the positive integers $\{1, 2, 3, \dots\}$ by $\mathbb{N}$. For a given $n \in \mathbb{N}$, the finite set $\{1, 2, \dots n\}$ is denoted by $[n]$. Finally, we denote the empty set $\{\}$ by $\varnothing$.

1.1 Permutations

Permutations arise in many different fields in mathematics, as one of the classical abstractions that express in how many ways one may rearrange the elements of a particular set. Formally, fixed a set S, usually finite, a *permutation* π is a bijective function $\pi \colon S \to S$. A representation, used by A. Cauchy [15], denoted a permutation in two lines: the first containing the elements of S; and the second containing the image through π. For example, for the set $S = [4]$, a particular permutation π can be written as

$$\pi = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

This means that π satisfies $\pi(1) = 2$, $\pi(2) = 4$, $\pi(3) = 1$ and $\pi(4) = 3$. Note that the same permutation may be represented in different ways by applying π on the elements of S but in a different order. For example, the same permutation π can be written as

$$\pi = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 3 & 1 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

If there is a natural order of the elements in S, say $s_1, s_2, \ldots, s_n$, and π is defined as

$$\pi = \begin{pmatrix} s_1 & s_2 & s_3 \\ \pi(s_1) & \pi(s_2) & \pi(s_3) \end{pmatrix},$$

then one may omit the first row of elements and write simply

$$\pi = \pi(s_1)\pi(s_2)\pi(s_3),$$

which is known as the *One-Line* notation of a permutation. This notation motivates the understanding of a permutation as a word. For this reason, we shall denote the i-th element of a permutation π as π_i and, therefore, the permutation π is written as $(\pi_1, \pi_2, \pi_3, \ldots, \pi_n)$. We omit commas and parentheses whenever doing so produces no ambiguity. Given a set S, we denote by $|S|$ the number of elements in S. The *length* of a permutation $\pi\colon S \to S$, denoted by $|\pi|$ is equal to $|S|$.

A *subsequence* τ of a permutation π, denoted by $\tau \subseteq \pi$, is a sequence obtained from π by removing some (maybe none) of the elements of π, while keeping the order of the remaining elements.

Let S_n denote the set of all permutations of the set $[n]$. Observe that $|S_n| = n! = n(n-1)(n-2)\cdots(2)(1)$. Since to build a permutation $\pi \in S_n$, we can choose n entries to be π_1, then $n-1$ entries to be π_2 and so on. Therefore, there are $n(n-1)(n-2)\cdots(2)(1)$ possible permutations of length n.

Fixed a permutation $\pi = \pi_1\pi_2\pi_3\ldots\pi_n$, for $1 < i < n$, we call the subsequence $\pi_{i\text{-}1}\pi_i\pi_{i+1}$ a *peak* (resp. *valley*) if $\pi_i > \pi_{i\text{-}1}, \pi_{i+1}$ (resp. $\pi_i < \pi_{i\text{-}1}, \pi_{i+1}$). Now, for any subsequence $\tau \subseteq \pi$, we denote by $\mathrm{p}(\tau)$ (resp. $\mathrm{v}(\tau)$) the number of peaks (resp. valleys) of τ. For example $\mathrm{p}(42135) = 0$ but $\mathrm{v}(42135) = 1$ since 213 is a valley. We now generalize the idea of peaks and valleys to other subsequences, since, for example, in $\pi = 42135$, it may be interesting to look at the subsequence 215 and call it a valley by itself.

For $1 \leq i < j < k \leq n$, we say that the triple (π_i, π_j, π_k) is a *triangle*, if $\pi_i\pi_j\pi_k$ is either a peak or a valley. For this reason, we define the indicator function $\Delta\colon (\pi_i, \pi_j, \pi_k) \to \{0, 1\}$ that indicates whether the triple (π_i, π_j, π_k) is a triangle, which can be written as $\Delta(\pi_i, \pi_j, \pi_k) = \mathrm{p}(\pi_i\pi_j\pi_k) + \mathrm{v}(\pi_i\pi_j\pi_k)$.

Given a permutation π and a triangle $\pi_i, \pi_j, \pi_k \subseteq \pi$, we call the *basis* of the triangle the number $k - i + 1$. Moreover, for a given basis b and permutation π, we denote by $\Delta_b(\pi)$ the set of all the triangles in π with basis equal to b.

A permutation π is called *alternating* if it consists of consecutive triangles of base 3. That is, either $\pi_1 > \pi_2 < \pi_3 > \pi_4 < \cdots$ or $\pi_1 < \pi_2 > \pi_3 < \pi_4 > \cdots$.

Given a permutation π, we call a pair of consecutive elements (π_i, π_{i+1}) an *ascent* (resp. a *descent*) if $\pi_{i+1} > \pi_i$ (resp. $\pi_{i+1} < \pi_i$). We call a permutation π an *ascending alternating permutation* if π is an alternating permutation that starts with an ascent, i.e., $\pi_1 < \pi_2 > \pi_3 < \cdots$. In the same manner, we call π a *descending alternating permutation*, if π is an alternating permutation that start with a descent, i.e., $\pi_1 > \pi_2 < \pi_3 > \cdots$. We denote the set of all ascending (respectively, descending) alternating permutation of length n as Asc_n (respectively, Desc_n), and let E_n denote the size of the set Asc_n.

For any permutation $\pi \in S_n$, we define the permutation π^c as the *complement permutation*, and the permutation π^r as the *reverse permutation*, respectively, as the permutation whose i-th term is defined as:

$$(\pi^c)_i = n + 1 - \pi_i.$$
$$(\pi^r)_i = \pi_{n+1-i}.$$

Clearly they are injective operations, which means that for $\pi, \pi' \in S_n$ if $\pi \neq \pi'$ then $\pi^c \neq (\pi')^c$ and $\pi^r \neq (\pi')^r$. And since they are involutions, which means that $(\pi^c)^c = \pi$ and $(\pi^r)^r = \pi$, they are bijections. Note that for every permutation $\pi \in \mathrm{Asc}_n$, the permutation $\pi^c \in \mathrm{Desc}_n$, which implies that $|\mathrm{Asc}_n| = E_n = |\mathrm{Desc}_n|$.

Before we present the next result, we first introduce the concept of Generating Functions. Given a sequence $\mathcal{A} = \{a_k\}_{k\geq 0}$, the *generating function* A is the *formal power series* $A(x) = \sum_{k\geq 0}^{\infty} a_k x^k$, where the coefficient of x^k, denoted by $[x^k]A(x)$, is a_k. Note that:

$$A(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots$$

Assuming that A is k times differentiable, taking k derivatives with respect to x yields

$$D^k A(x) = k! a_k + \frac{(k+1)!}{1} a_{k+1} x + \frac{(k+2)!}{2!} a_{k+2} x^2 + \frac{(k+3)!}{(3!)} a_{k+3} x^3 + \cdots .$$

Evaluating at the point $x = 0$ results in

$$D^k A(0) = k! a_k.$$

Therefore, assuming that A is k times differentiable at 0, we have

$$[x^k]A(x) = a_k = \frac{1}{k!} D^k A(0).$$

Given a sequence $\mathcal{A} = \{a_k\}_{k \geq 0}$, the *exponential generating function* (EGF) of $\mathcal{A}$ is the *formal power series* $A(x) = \sum\limits_{k \geq 0}^{\infty} a_k \frac{x^k}{k!}$. Enumeration of permutations is mostly approached through EGFs, since the generating function $P(x) = \sum\limits_{k \geq 0} k! x^k$ diverges for all $x \neq 0$, which means that it has radius of convergence 0 and is not an analytical function. On the other hand, the exponential generating function $P(x) = \sum\limits_{k \geq 0} k! \frac{x^k}{k!}$ has radius of convergence 1, since it converges for $0 \leq x < 1$. Naturally, for an EGF $A(x) = \sum\limits_{k \geq 0} a_k \frac{x^k}{k!}$, one may write $a_k = [\frac{x^k}{k!}]A(x) = k![x^k]A(x)$.

This abstraction allows us to extract properties from the sequence even though we might not know its closed formula. For more on Analytic Combinatorics and Generating Functions, we refer to P. Flajolet and R. Sedgewick in [25] and H. Wilf in [48].

In Chapter 2, we study a special kind of permutation that is believed to be closely related to alternating permutations. The following theorem, which is a remarkable enumerative result of alternating permutations, shall serve as a motivation to the study of the structure of such permutations.

Theorem 1.1 (D. André, 1879)**.** *Let W_n be the set of all the alternating permutations of length n. We have that*

$$|W_n| = 2E_n = 2\left[\frac{x^n}{n!}\right] \sum_{k \geq 0} E_k \frac{x^k}{k!} = 2\left[\frac{x^n}{n!}\right] (\sec x + \tan x).$$

Proof. The first equality comes from the fact that $W_n = \mathrm{Asc}_n \cup \mathrm{Desc}_n$, and $\mathrm{Asc}_n \cap \mathrm{Desc}_n = \varnothing$, which implies that $|W_n| = 2E_n$. Now, let $0 \leq k \leq n$. There are $\binom{n}{k}$ ways to choose a k-subset S of $[n]$, and set $\bar{S} = [n] - S$. Let $u \colon S \to S$, and $v \colon \bar{S} \to \bar{S}$ be, respectively, a permutation of the elements of S and a permutation of the elements of $\bar{S}$. Note that there are E_k ways to choose u from Desc_k and E_{n-k} ways to choose v from Asc_{n-k}. Let w be the concatenation $(u^r, n{+}1, v)$ where u^r denotes the reverse of u. When $n \geq 2$, this yields each of the $2E_{n+1}$ permutations exactly once. Hence

$$2E_{n+1} = \sum_{k=0}^{n} \binom{n}{k} E_k E_{n-k}.$$

Multiplying the both sides by $\frac{x^n}{n!}$ and summing for all $n \geq 1$ yields

$$2\sum_{n\geq 1} 2E_{n+1}\frac{x^n}{n!} = \sum_{n\geq 1}\sum_{k=0}\binom{n}{k}E_kE_{n-k}. \tag{1.1}$$

Note that, for the left hand side (LHS), we have

$$2\sum_{n\geq 1} 2E_{n+1}\frac{x^n}{n!} = (2\sum_{n\geq 1} E_{n+1}\frac{x^n}{n!}) + 2E_0 - 2E_0 = 2\sum_{n\geq 0} E_{n+1}\frac{x^n}{n!} - 2E_0.$$

Let $F(x) = \sum\limits_{n\geq 0} E_n\frac{x^n}{n!}$, then we have $\frac{dF(x)}{dx} = \sum\limits_{n\geq 1} E_n\frac{x^{n-1}}{(n-1)!}$. Setting $m = n-1$, we have $\frac{dF(x)}{dx} = \sum\limits_{m\geq 0} E_{m+1}\frac{x^m}{m!}$. Defining $E_0 = E_1 = 1$, implies that the LHS is

$$2\sum_{n\geq 0} E_{n+1}\frac{x^n}{n!} = F'(x) - 2.$$

For the right-hand side (RHS) of Equation 1.1, we have:

$$\sum_{n\geq 1}\sum_{k=0}\binom{n}{k}E_kE_{n-k}.$$

By summing and subtracting the term with $n = 0$, which evaluates to 1, we obtain the convolution formula, resulting in:

$$\sum_{n\geq 0}\sum_{k=0}\binom{n}{k}E_kE_{n-k} - 1 = F^2(x) - 1$$

Therefore, we rewrite equation 1.1 as the following differential equation:

$$2F'(x) - 2 = F^2(x) - 1,$$

whose unique solution is $F(x) = sec(x) + tan(x)$.

Recall that $F(x) = \sum\limits_{n\geq 0} E_n\frac{x^n}{n!}$, which means that $E_k = [\frac{x^k}{k!}]F(x)$. Moreover, since $|W_n| = 2E_n$, we have:

$$|W_n| = 2\left[\frac{x^n}{n!}\right]F(x) = 2\left[\frac{x^n}{n!}\right](\sec x + \tan x)$$

as desired. □

Since F is a sum of two analytical functions, and F has a radius of convergence of $R = \frac{\pi}{2}$, F is analytical and we obtain the following asymptotic approximation for $k \to \infty$:

$$[x^k]F(x) \approx \left(\frac{2}{\pi}\right)^{k+1},$$

which implies that $|W_n| \approx 2\left(\frac{2}{\pi}\right)^{n+1} n!$.

1.2 Graphs

In this section we present a short history of Graph Theory, as well as definitions and notations in order to familiarize the reader with the terminology that we use in this dissertation.

The origin of Graph Theory can be traced back to the first half of the 18th Century, when L. Euler presented his solution to the Königsberg problem in [21]. The *Königsberg problem* asks wether there exists a continuous walk that crosses each of the seven bridges of Königsberg exactly once. In his proof, even though not drawing, L. Euler glanced on some important properties of a structure that today is known as a graph.

In Figure 1.1 we present a representation of the Königsberg bridges. In his proof, Euler showed when such walk could exist, noting that if there are more than two areas to which an odd number of bridges lead, then such a journey is impossible. Moreover, L. Euler noted that if the number of bridges is odd for precisely two areas, then the walk is possible if it starts in either of these two areas; and if there are no areas to which an odd number of bridges lead, then the journey can be accomplished starting anywhere. In conclusion, L. Euler proved that there are no walks that satisfy the Königsberg problem.

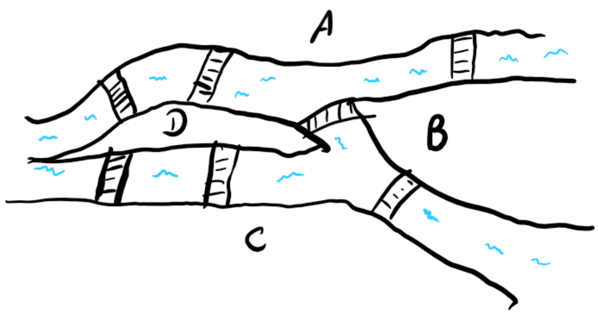


Figure 1.1: An illustration of the seven Königsberg bridges.

Throughout the text, an ordered pair of elements is denoted by (e_1, e_2), while an unordered pair of elements is denoted by $\{e_1, e_2\}$. This means that the pair $\{e_1, e_2\} = \{e_2, e_1\}$ while $(e_1, e_2) \neq (e_2, e_1)$.

Formally, a *graph* G is an ordered pair (V, E) such that V is the set of *vertices* and E is a *multiset*, which is a set that allows repetitions, of *edges*, which are unordered pairs of vertices. A graph is called *simple* if there are no repeated edges in E, i.e., E is a set, and there are no *loops* in E, which is an edge of the form $\{v, v\}$, for $v \in V(G)$. A *directed graph* is defined similarly, but taking E as a multiset of ordered pairs, which means that $(x, y) \neq (y, x)$.

For a given graph $G = (V, E)$, the set $V = V(G)$ is the vertex set of G, and $E = E(G)$ is its edge set. The *order* of G is the number $|V(G)|$, often represented by n, and the *size* of G is the number $|E(G)|$, often represented by m. If x, y are vertices, then an edge $\{x, y\}$ is said to *join* the vertices x, y and is often denoted simply by xy; the vertices x and y are *end-vertices* of xy. If $\{x, y\} \in E(G)$, then we say that x and y are *adjacent* vertices of G or *neighbors*. Two edges are *adjacent* if they have exactly one common end-vertex. For a vertex $v \in V(G)$ we define the set $N_G(v)$ as the set with all the vertices $u \in V(G)$ that are adjacent to v, also known as the *neighborhood* of v.

Naturally, we can represent a graph G with a figure consisting of the vertices joined by a simple line if such edge exists in $E(G)$. In the case of a directed graph, an arrow is drawn instead of a simple line, to represent the difference between the edges (u, v) and (v, u). For example, the graph with vertex set $\{A, B, C, D\}$ and edge set $\{AB, AD, AD, BC, BD, CD, CD\}$ is represented in Figure 1.2. Note that this figure captures the structure related to the Königsberg problem, where the masses of land are the vertices and the bridges are the edges.

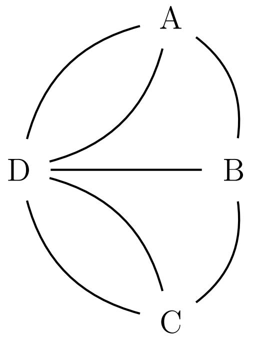


Figure 1.2: A Graph representation of the Königsberg bridges.

Fix a graph G, for a given vertex $v \in V(G)$, we call the *degree* of v, denoted by $\mathrm{d}_G(v)$, the number of edges $e \in E(G)$ that contains v. A vertex $u \in V(G)$ with $\mathrm{d}(u) = 1$ is called a *leaf* of G, and a vertex u with $\mathrm{d}(u) = |V(G)| - 1$ is called a *universal vertex*. We shall denote the maximum degree of a graph G by $\Delta(G)$ and the minimum degree by $\delta(G)$. Let $\overline{\mathrm{d}}(G)$ be the *average degree* of a given graph G, that is:

$$\overline{\mathrm{d}}(G) = \frac{1}{|V(G)|} \sum_{v \in V(G)} \mathrm{d}_G(v).$$

A very simple statement is the following lemma, which introduces the double counting argument.

Lemma 1.2. *(Handshaking lemma) Given a graph $G = (V, E)$, we have that:*

$$\sum_{v \in V(G)} \mathrm{d}_G(v) = 2|E(G)|. \tag{1.2}$$

Proof. Consider the pair (v, e) such that $v \in V(G)$ and $e \in E(G)$, with $v \in e$. We can count the number of such pairs in two different ways. Note that, by definition, there are $\mathrm{d}_G(v)$ edges containing v, and hence, every vertex v is contained in $\mathrm{d}_G(v)$ pairs. On the other hand, every edge contains two different end-vertices, which means that there are $2|E(G)|$ of such pairs. Therefore, these quantities must be equal, leading to the desired result. $\square$

As a corollary, we have that for any graph G, the number of vertices $v \in V(G)$ with $\mathrm{d}(v) \equiv 1 \pmod 2$ is even; otherwise, the left hand side of the Equation (1.2) would be odd, contradicting the lemma.

We now proceed to some important definitions that formalize the notion of finding a structure inside a graph. A graph G is *isomorphic* to a graph H, denoted by $G \simeq H$, if there exists a bijective function $\varphi : V(G) \to V(H)$ such that two vertices $u, v \in V(G)$ are adjacent in G if and only if $\varphi(u)$ and $\varphi(v)$ are adjacent in H. In this case, we say that the function φ is an *isomorphism*. Note that Figure 1.1 and 1.2 are isomorphic by the identity function $\varphi(x) = x$, for $x \in \{A, B, C, D\}$.

Given a graph G, we call a graph H a *subgraph* of G, or simply say that G *contains* H, denoted by $H \subseteq G$, a graph for which $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. For graphs T and G, if there is an isomorphism φ of T into $H \subseteq G$, we say that there is an *embedding* of T into G, or simply say that there is a *copy* of T in G. If a graph G does not contain a copy of T, we say that G is T-free.

Another important idea is to decompose the graph in substructures, usually disjoint, that presents some properties. A graph G is *bipartite* if we can write $V(G) = V_1 \cup V_2$ for some sets V_1, V_2, such that $V_1 \cap V_2 = \varnothing$, $V_i \neq \varnothing$, and for $u, v \in V_i$, u and v are not adjacent, for $i \in \{1, 2\}$. A graph G is *k-partite* if we can write $V(G) = \bigcup_{i=1}^{k} V_i$, for some sets V_i, for $i \in [k]$, such that $V_i \cap V_j = \varnothing$, $V_i \neq \varnothing$, for $i \neq j$ and for each part V_i, any pair of vertices $u, v \in V_i$ is not adjacent.

We extend the definition of partition to coloring of either edges or vertices of a graph, understanding that assigning a color is equivalent to assign the object to a part of the partition. In this work, we consider only the case of vertex coloring.

Given a graph $G = (V, E)$, the *chromatic number* of G, denoted by $\chi(G)$, is the minimum number of colors for which the vertices in V can be assigned such that no vertices of the same color is adjacent. It can be verified that $\chi(G) = 2$ for every bipartite graph G, by coloring each partition V_i with a different color. More generally, $\chi(G) \leq k$ for all k-partite graphs G.

1.2.1 Special Graphs

Some special classes of graphs ought to be highlighted and defined. A graph G is called *empty* if $E(G) = \varnothing$, and a graph G is called the *null graph* if $V(G) = \varnothing$. These graphs do not seem very interesting by themselves, but it is important to observe that every graph contains the null graph, and every graph on n vertices, contains 2^n empty graphs, including the null graph.

A graph is called *complete* if its edge set contains every possible edge. The simple complete graph with n vertices is denoted by K_n. Since for each vertex $v \in V(K_n)$ there are $n-1$ possible choices for an edge, excluding the loops, we have that $2|E(K_n)| = n(n-1)$ and $\delta(K_n) = \Delta(K_n) = n-1$. In Figure 1.3 we illustrate K_{12}.

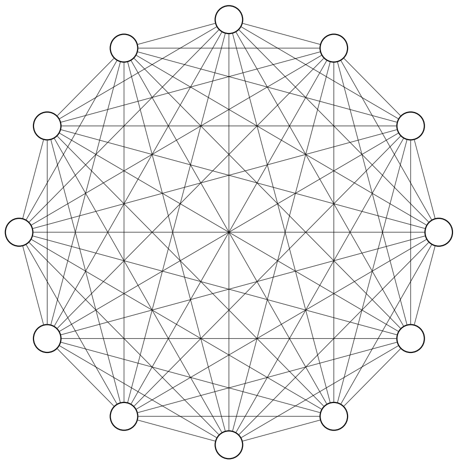


Figure 1.3: The complete graph K_{12}.

Given a complete graph K_n, we create a *tournament* by assigning a direction for each edge, thus transforming the undirected graph into a directed graph.

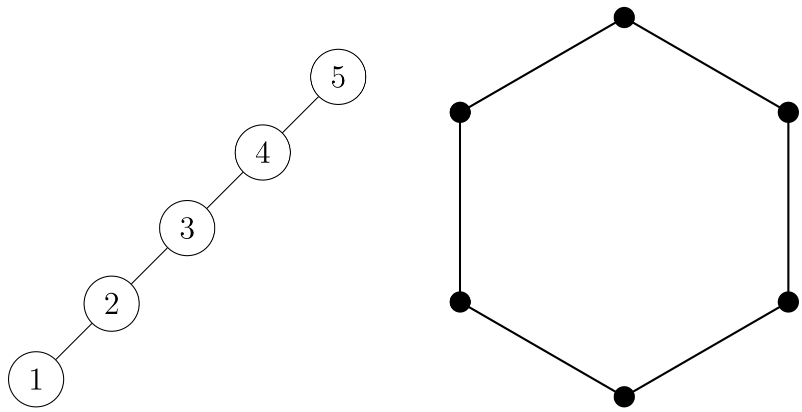
Given a graph G on n vertices, we denote by G^c the *complement graph of* G which is defined as a graph on the same vertices as G but two vertices $u, v \in V(G^c)$ are adjacent in G^c if and only if they are not adjacent in G.

A graph G of size ℓ is called a *path*, denoted by P_ℓ, if there is an ordering $\{v_1, \dots, v_{\ell+1}\}$, of its vertices, for which its edges are $\{v_i, v_{i+1}\}$, for $i \in [\ell]$. This means that G can be illustrated as straight line, as seen in Figure 1.4a. Recall that the size ℓ represents the number of edges in G, not to be confused with its number of vertices.

Given a graph $G = (V, E)$, the *distance* between two vertices $u, v \in V(G)$, denoted by $dist(u, v)$ is the number of edges in a shortest path connecting them. The *diameter* of a graph G, denoted by $D(G)$ is the greatest distance between any pair $u, v \in V(G)$. If there is not a path connecting the two vertices, we define $dist(u, v)$ as being ∞.

A graph G is called *connected* if for every pair of vertices $u, v \in V(G)$, there is a path from u to v using edges in $E(G)$. A graph is, therefore, called *disconnected*, if it is not connected, i.e., if there is at least one pair $u, v \in V(G)$ for which there is no path that begins in u and ends in v.

A graph is called a *cycle* on n vertices, denoted by C_n, if it consists of a path P_{n-1}, with $V(P_\ell) = \{v_1, v_2, \cdots, v_n\}$, together with the edge $\{v_1, v_n\}$. This is illustrated in Figure 1.4b. A graph G is called *acyclic* if it contains no cycle. We call the *girth* of a graph G the length of the smallest cycle in G.



(a) Representation of a labeled P_4

(b) Representation of an unlabeled C_6

Figure 1.4: Illustration of a path and a cycle, as well as a labeled and an unlabeled graph.

A graph is called a *tree* if it is connected and acyclic. We denote the family of all trees on k vertices by $\mathcal{T}_k$. The problem of enumerating trees is a well known problem since the 19th Century. A solution for the case of *labeled* trees, in which you assign for each vertex a different label, was proved by C. Borchardt [11], to be k^{k-2} in 1861. But a closed formula for the *unlabeled* case, in which you do not distinguish its vertices, is not known. Still, there are certain special classes of trees that can be enumerated explicitly even in the unlabeled case.

Given a tree T, the *derived tree* ∂T is the tree obtained from T by deleting all of its leaves. We call a tree T a *caterpillar* if ∂T is a path. In 1973, F. Harary and J. Schwenk proved the following enumeration theorem in [28].

Theorem 1.3. *The number of non-isomorphic caterpillars with $n \in \mathbb{N}$ vertices is given by*

$$C(n) = [x^n]\frac{x^3(1-3x^2)}{(1-2x)(1-2x^2)} = 2^n + 2^{\lfloor n/2 \rfloor}.$$

Different proofs for this theorem can be found in [28].

Another well studied class of trees is the following. A tree S is a *spider* if it has at most one vertex of degree greater than 2, which we call the *center* of S (if no vertex has degree greater than 2, then any vertex can be the center). A *leg* of a spider is a path from its center to one of its leaves. Therefore, a path is a spider with one or two legs.

Let $P(n,k)$ denote the *partition number* of n in k summands, which is defined as the number of ways of writing n as a sum of k positive integers. In this case, two sums that only differ in the order of the summands are considered the same partition. For example, we have that $P(4,2) = 2$ since $4 = 3+1 = 3+1$ and $4 = 2+2$.

The following theorem is an enumerative result concerning spider trees.

Theorem 1.4. *The number of non-isomorphic spiders with $n \in \mathbb{N}$ vertices is given by*

$$S(n) = 1 + \sum_{k \geq 3} P(n-1,k).$$

Proof. Consider a partition of $n-1$ in k positive integers $p_1, p_2, \ldots, p_k$. We can build a spider by starting with the center vertex and including a leg of size p_i for $i \in [k]$. Note that for $k = 2$ this construction is not injective since a spider with two legs is just a path. Therefore, we consider only partitions with $k \geq 3$, in which different partitions results in non-isomorphic spiders. Moreover, since we can reverse the construction to obtain the same partition of $n-1$ by removing the legs from the center vertex, this construction implies in a bijection between both quantities. All that is left is to include the spider which is isomorphic to the path, leading to the desired result. □

1.3 Representations of Permutations

In this section we explore three different ways to represent permutations. Throughout this work, we mainly represent a permutation π, graphically, as a sequence of the points (i, π_i), for $i \in [n]$, in the cartesian coordinate systems, and joining two consecutive points.

Recall that for any permutation $\pi \in S_n$ the complement permutation π^c, and the reverse permutation π^r are defined, respectively, by having their i-th term defined as

$$(\pi^c)_i = n + 1 - \pi_i.$$
$$(\pi^r)_i = \pi_{n+1-i}.$$

Note that you can also combine both transformations to obtain the reverse complement $(\pi^c)^r$, which is equal to the complement reverse $(\pi^r)^c$, and, for this reason, we denote it by π^{rc}. For example, given the permutation $\pi = 1, 2, 4, 3$, we have that $\pi^c = 4, 3, 1, 2$, $\pi^r = 3, 4, 2, 1$, and $\pi^{rc} = 2, 1, 3, 4$.

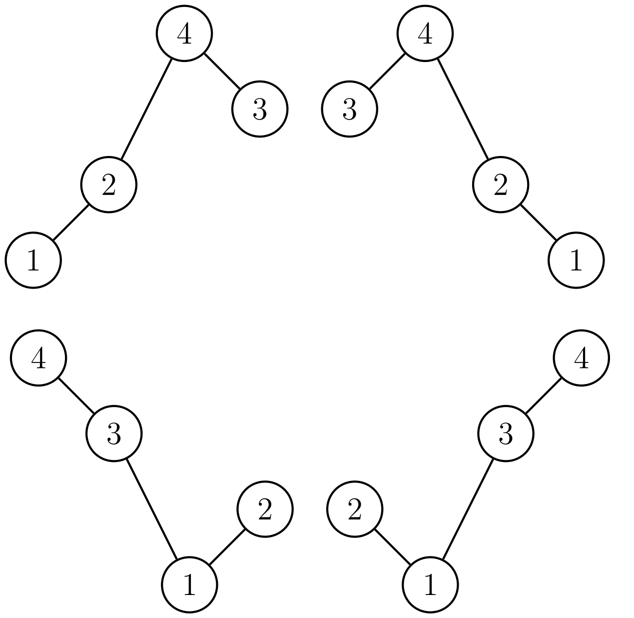
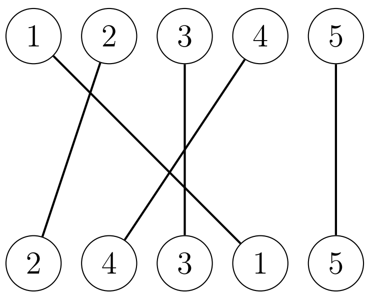


Figure 1.5: For $\pi = 1243$, the permutations π (top left), π^r (top right), π^c (bottom left), π^{rc} (bottom right).

A very interesting symmetry arises from these transformed permutations: the fact that they can be obtained from the original permutation π either by a reflection over the x axis, obtaining the complement permutation, by a reflection over the y axis, obtaining the reverse permutation, or both, resulting in the reverse complement permutation.

Another way to represent a permutation π is through what is called a *permutation graph*. In order to draw the permutation graph, one first write π in its two-line notation. Then, for each number in the first line, we draw a line to itself in the second line. We define the permutation graph to be a graph whose vertices represents this lines and in which two vertices are adjacent if their lines intersect. For the permutation $\pi = 24315$, we have:



Since the line between the vertices with the label 1 intersected the lines for the vertices $2, 3$, and 4, the vertex 1 has degree 3 in the permutation graph. Doing this analysis for every other vertex results in the following permutation graph:

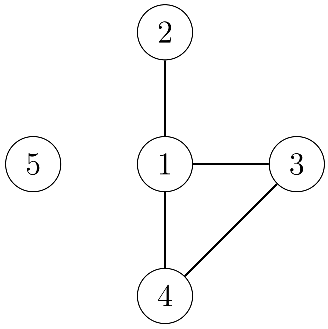


Figure 1.6: Permutation Graph of $\pi = 24315$.

In other words, fixed a permutation $\pi = \pi_1\pi_2\pi_3 \ldots \pi_n$, with $\pi \in S_n$, the associated permutation graph $G = (V, E)$ has $V = [n]$, and $ij \in E$ if and only if $i < j$ and $\pi_i > \pi_j$, for all $i \neq j$.

2 | Roller Coaster Permutations

The concept of Roller Coaster permutations was introduced by T. Ahmed and H. Snevily in [2]. Informally, a *Roller Coaster permutation* is a permutation that, along with all of its subsequences, maximizes the amount of changes between ascents and descents. The authors provided a number of conjectures concerning enumerative and structural properties of Roller Coaster permutations as well as showing its connections with the problem of partition numbers of a permutation and forbidden subpermutations.

To formally introduce Roller Coaster permutations, we need some auxiliary definitions. Let $\mathrm{i}(\tau)$ (resp. $\mathrm{d}(\tau)$) be the number of maximal increasing (resp. decreasing) subsequences of consecutive numbers in τ, where such a sequence consists of at least two numbers. Let $\mathrm{id}(\tau) = \mathrm{i}(\tau) + \mathrm{d}(\tau)$, and let $X(\pi)$ denote the set of the subsequences $\tau \subseteq \pi$ with at least three elements. Note that the size of the set $X(\pi)$ can be obtained from enumerating all the subsequences of π and removing subsequences of length $0, 1$, and 2, hence

$$|X(\pi)| = 2^n - n - 1 - \binom{n}{2}. \tag{2.1}$$

Finally, we define the function $\mathrm{t} : S_n \to \mathbb{N}$ as

$$\mathrm{t}(\pi) = \sum_{\tau \in X(\pi)} \mathrm{id}(\tau). \tag{2.2}$$

For illustration purposes, we evaluate it on two permutations of S_4:

$$\begin{aligned}
\mathrm{t}(3412) &= \mathrm{id}(3412) + \mathrm{id}(341) + \mathrm{id}(342) + \mathrm{id}(312) + \mathrm{id}(412) \\
&= \quad 3 \quad + \quad 2 \quad + \quad 2 \quad + \quad 2 \quad + \quad 2 \quad = 11. \\
\mathrm{t}(1234) &= \mathrm{id}(1234) + \mathrm{id}(123) + \mathrm{id}(124) + \mathrm{id}(134) + \mathrm{id}(234) \\
&= \quad 1 \quad + \quad 1 \quad + \quad 1 \quad + \quad 1 \quad + \quad 1 \quad = 5.
\end{aligned}$$

In a certain way, t measures the “alternation” of a given permutation, by counting the ascents and descents of the permutation and of all of its meaningful subsequences.

For this reason, we refer to t as the *measure of alternation* of a permutation, even though it is not precisely a mathematical measure. Seldom, we refer to $\mathrm{t}(\pi)$ simply as the t-value of π.

Recall that S_n is the set of all permutations of a set of n elements. Since the set S_n is finite, there is a maximum value of t for any given n, hence we define $\mathrm{t}_{max}(n) = \max_{\pi \in S_n} \mathrm{t}(\pi)$, and say that a permutation π is a *Roller Coaster* if $\mathrm{t}(\pi) = \mathrm{t}_{max}(n)$. Lastly, we define $RC(n)$ as the set of all the Roller Coasters of length n. In the example above, the permutation 1234 is not a Roller Coaster because $\mathrm{t}(1234) < \mathrm{t}(3412)$. On the other hand, it can be verified that 3412 is indeed a Roller Coaster by checking that $\mathrm{t}_{max}(4) = 11$.

Recall that for a permutation π, the reverse permutation and the complement permutation are denoted by π^r and π^c respectively. A trivial fact is that for $\pi \in S_n$, the function t satisfies:

$$\mathrm{t}(\pi) = \mathrm{t}(\pi^c) = \mathrm{t}(\pi^r) = \mathrm{t}(\pi^{rc}),$$

which implies that $|RC(n)| \geq 4$ for all $n \geq 3$.

T. Ahmed and H. Snevily obtained experimental results for t_{max} for $n \leq 13$, and a construction that provides lower bounds for $n \geq 14$ in [2]. For $n = 3, 4, \cdots, 24$ the authors presented the following values for t_{max} and cadidates for t_{max} (in italic):

$$\mathrm{t}_{max} = [2, 11, 37, 106, 270, 653, 1523, 3480, 7768, 17123, 37405, \mathit{81350}, \mathit{174954}, \mathit{374409}, \mathit{798471}, \mathit{1700036}, \mathit{3596124}, \mathit{7588303}, \mathit{15970785}, \mathit{33596706}, \mathit{70310126}, \mathit{146867861}]$$

The authors also provided examples of Roller Coasters for $n \leq 13$. In Figure 2.1, we present graphical examples for each permutation in $RC(n)$ for $3 \leq n \leq 8$.

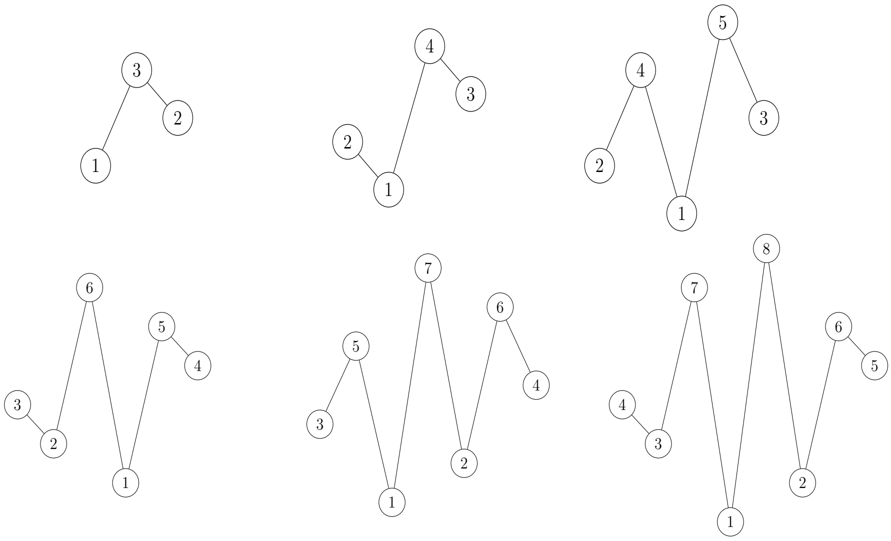


Figure 2.1: Roller Coaster permutations for $3 \leq n \leq 8$.

Based on these results, T. Ahmed and H. Snevily conjectured the following properties regarding the structure of the Roller Coaster permutations in [2].

Conjecture 2.1. *For a given $n \in \mathbb{N}$, if $\pi \in RC(n)$, then π is an alternating permutation.*

The authors also conjectured the following values for π_1 and π_n.

Conjecture 2.2. *Given a positive integer n, there is $\pi \in RC(n)$ such that $\pi_1 = \lfloor n/2 \rfloor$ and $\pi_n = \lfloor n/2 \rfloor + 1$*

This conjecture motivates a constructive approach to find Roller Coaster permutations, in which one choose π_1 and π_n as the numbers close to $n/2$ and choose π_i, for i close to $n/2$ as the extremes 1 and n, to locally maximize the alternation of the permutation π. Unfortunately, after several attempts we were unable to obtain satisfactory results with such algorithms since for every pattern that we thought we spotted, the next Roller Coaster permutation would avoid it, proving to be quite elusive objects.

In 2017, W. Adamczak [1] published a tentative proof for Conjecture 2.1 based on the argument that for any permutation $\pi = \tau, \pi_i, \pi_{i+1}, \pi_{i+2}, \sigma,$ in which (a) the subsequence σ is an alternating sequence; and (b) $\pi_i, \pi_{i+1}, \pi_{i+2}$ is a increasing or decreasing sequence, the operation of switching π_i with π_{i+1} always yields a permutation with a greater t-value. After a careful review of his proof, we found a counterexample for this argument. Consider the permutation $\pi = 7, 3, 5, 6, 4, 8, 1$. One may compute and verify that $\mathrm{t}(\pi) = 240$. If we consider the permutation $\pi' = 7, 3, 6, 5, 4, 8, 1$, by switching the position of the 5 with the 6, one may compute that $\mathrm{t}(\pi') = 236$. This implies that

$$240 = t(\pi) > t(\pi_2) = 236.$$

Alongside this proof attempt, the author attempted to prove other theorems, but, unfortunately, used the same argument. For this reason, we present these results as conjectures.

Conjecture 2.3. *There is a permutation $\pi \in RC(n)$, such that:*

(i) $\{\pi_i : \ i \text{ is odd}\} = \{\frac{n}{2}+2, \ldots, n\}$

(ii) $\{\pi_i : \ i \text{ is even}\} = \{1, \ldots, \frac{n}{2}-1\}$

Conjecture 2.3 explores the structure of the permutation by characterizing the subset of the odd-indexed terms and even-indexed terms. This motivates a greedy algorithm that tries to maximize the measure of alternation locally for each of these subsequences and then concatenates them. One example would be to pick $\tau, \sigma \in RC(n)$, then shifting the permutation τ, obtaining $\tau' = \tau_1 + n, \tau_2 + n, \ldots, \tau_n + n$, then merging together τ' and σ, building the permutation $\pi = \tau_1', \sigma_1, \tau_2', \sigma_2 \cdots$. Unfortunately, this method yields permutations for which its t-value is less than $\mathrm{t}_{max}(2n)$.

To present the next conjecture, we first need to define some auxiliary terms. Given a permutation $\pi = \pi_1\pi_2\ldots\pi_n$, the subsequence π_{ODD} is the sequence $\{\pi_i\}_{i=2k+1}$, for $1 \leq k \leq \lfloor\frac{n-3}{2}\rfloor$. Analogously, the subsequence π_{EVEN} is the sequence $\{\pi_i\}_{i=2k}$ for $1 \leq k \leq \lfloor\frac{n-1}{2}\rfloor$, for example, let $\pi = (1, 2, 3, 4, 5, 6, 7, 8)$, we have $\pi_{\mathrm{ODD}} = (3, 5, 7)$ and $\pi_{\mathrm{EVEN}} = (2, 4, 6)$, skipping the first and last terms of the permutation. A permutation π is called *recursively alternating* if either $|\pi| \leq 2$ or both of the following holds.

(i) π is alternating;

(ii) π_{ODD} and π_{EVEN} are recursively alternating;

This means that for a permutation π that is recursively alternating, the subsequence $\pi_{\mathrm{ODD,ODD}}$ is alternating, as well as the $\pi_{\mathrm{EVEN,ODD}}$, $\pi_{\mathrm{EVEN,EVEN}}$, and $\pi_{\mathrm{ODD,EVEN}}$, $\pi_{\mathrm{ODD,ODD,ODD}}$, et cetera. The second conjecture we present, from W. Adamczak [1], is the following:

Conjecture 2.4. *If $\pi \in RC(n)$, then π is* recursively alternating.

This conjecture is an even stronger version of Conjecture 2.1, since it implies that the permutation is alternating and also some of its subsequences are alternating.

Lastly, we present the following conjecture, due to T. Ahmed and H. Snevily [2] regarding the behavior of $\mathrm{t}_{max}(n)$ for $n \to \infty$. Based on the exact and conjectured values of t_{max}, one can define $L(n) = \frac{\mathrm{t}_{max}(n+1)}{\mathrm{t}_{max}(n)}$, which for $4 \leq n \leq 23$, gives us:

$$L : [3.364, 2.865, 2.547, 2.418, 2.332, 2.285, 2.232, 2.204, 2.184, 2.175$$
$$2.151, 2.140, 2.133, 2.129, 2.115, 2.110, 2.105, 2.104, 2.093, 2.089].$$

Question 2.5. *Does the limit*

$$\lim_{n\to\infty} L(n) = \lim_{n\to\infty} \frac{\mathrm{t}_{max}(n+1)}{\mathrm{t}_{max}(n)}$$

exist and is it 2?

If the answer to Question 2.5 is affirmative, this would imply that for large values of n, the number of peaks and valleys for a permutation $\pi \in RC(n+1)$ would be essentially the double of the number of peaks and valleys of $\pi \in RC(n)$. Moreover, the existence of the limit $\lim_{n\to\infty} L(n)$ suggests some regularity in the behavior of the function t.

This chapter is structured in the following way: In Section 2.1, we present the main result of this chapter, an alternate definition that allows us to redefine the measure of alternation of a Roller Coaster permutation, simplifying the method to find them.

In Section 2.2, we present an Integer Linear Programming (ILP) model based on the reformulation of the function t, as well as an extended version that includes constraints based on Conjectures 2.1 and 2.3. We also present new examples of Roller Coaster permutations when $14 \leq n \leq 17$, together with Roller Coaster candidates when $18 \leq n \leq 40$, which were obtained with the extended ILP. This results provide empirical evidence in concordance with Conjecture 2.4 and Question 2.5.

Moreover, in Section 2.3, we explore some alternative problems concerning Roller Coaster permutations based on the number of triangles in a permutation. We also motivate the inverse problem of determining a permutation given how many triangles it contains.

In Section 2.4, we present some properties of the representation of a Roller Coaster permutation through tournaments, motivating the search to reduce the problem of evaluating the function t to the problem of matrix multiplication. Moreover, we present some properties of the associated permutation graph as well as a structural conjecture regarding the connectivity of such graphs.

Finally, in Section 2.5, we conclude the chapter, restating the contributions achieved to this problem, as well as presenting a table containing improvements on the lower bounds for $14 \leq n \leq 40$, found with the proposed extended ILP model.

It is worth mentioning that part of this work was published and presented at the "VI Encontro de Teoria da Computação" in 2021 [12].

2.1 Fast computation of t

Recall that $\mathrm{t}(\pi) = \sum_{\tau\in X(\pi)} \mathrm{id}(\tau)$. This definition suggests a $\theta(2^n)$ algorithm to compute t, thereby taking $O(2^n n!)$ time to explore every permutation of length n in order to obtain $\mathrm{t}_{max}(n)$.

The main result of this section is the following identity.

Theorem 2.6. *Let $\pi \in S_n$. Then:*

$$t(\pi) = |X(\pi)| + \sum_{1\le i<j<k\le n} 2^{n-(k-i+1)}\Delta(\pi_i, \pi_j, \pi_k). \tag{2.3}$$

In order to prove Theorem 2.6, we first prove the following lemma, that states an alternative way to evaluate $\mathrm{id}(\pi)$.

Lemma 2.7. *Let $\pi \in S_n$. Then* $\mathrm{id}(\pi) = 1 + \mathrm{p}(\pi) + \mathrm{v}(\pi)$.

Proof. Let $r = \mathrm{id}(\pi)$, and let $\pi_1, \pi_2, \ldots, \pi_r$ be the maximal increasing and decreasing contiguous subsequences of π, in the order that they appear in π, where $\pi_i = \pi^i_1, \ldots, \pi^i_{s_i}$. Fix $i \in \{2, \cdots, r\}$. Note that $\pi^{i-1}_{s_i-1} = \pi^i_1$. By the maximality of π_i, we have that π_i is increasing if and only if π_{i-1} is decreasing. This implies that $\pi^{i-1}_{s_i-1}\pi^i_1\pi^i_2$ is either a peak or a valley. Moreover, these are the only peaks and valleys of π. Therefore, $\mathrm{p}(\pi) + \mathrm{v}(\pi) = r - 1 = \mathrm{id}(\pi) - 1$. □

This lemma gives us another way to define the measure of alternation, i.e., the function t. The proof of Theorem 2.6 goes as following.

Proof. First, by Lemma 2.7, we have

$$\mathrm{t}(\pi) = \sum_{\tau\in X(\pi)} \mathrm{id}(\tau) = \sum_{\tau\in X(\pi)} (1 + \mathrm{p}(\tau) + \mathrm{v}(\tau)) = |X(\pi)| + \sum_{\tau\in X(\pi)} \mathrm{p}(\tau) + \mathrm{v}(\tau).$$

Now, we double count the cardinality of the following set:

$$E = \{(\tau, \sigma) : \tau = \tau_1 \cdots \tau_r \in X(\pi), \sigma = \tau_i\tau_{i+1}\tau_{i+2}, 1 \le i \le r-2 \text{ such that } \Delta(\sigma) = 1\}.$$

Let $d_1(\tau) := |\{(x, y) \in E : x = \tau\}|$ and $d_2(\sigma) := |\{(x, y) \in E : y = \sigma\}|$. Clearly $\sum_{\tau\in X(\pi)} d_1(\tau) = |E| = \sum_{\sigma\in X(\pi),|\sigma|=3} d_2(\sigma)$, and $d_1(\tau)$ is the number of triangles of τ, i.e., $d_1(\tau) = \mathrm{p}(\tau) + \mathrm{v}(\tau)$. Note that, for any subsequence $\sigma = \pi_i\pi_j\pi_k$ of π for which $\Delta(\sigma) = 1$, the pair $(w_1\sigma w_2, \sigma) \in E$, for every $w_1 \subseteq \pi_1 \cdots \pi_{i-1}$ and $w_2 \subseteq$

$\pi_{k+1}\cdots\pi_n$. Therefore, $d_2(\sigma) = 2^{n-(k-i+1)}$, which is the number of permutations $w_1\sigma w_2$. Consequently,

$$\sum_{\sigma\in X(\pi),|\sigma|=3} d_2(\sigma) = \sum_{i<j<k} \Delta(\pi_i,\pi_j,\pi_k) d_2(\pi_i\pi_j\pi_k) = \sum_{1\le i<j<k\le n} 2^{n-(k-i+1)}\Delta(\pi_i,\pi_j,\pi_k).$$

Therefore, we have:

$$\sum_{\tau\in X(\pi)} (\mathrm{p}(\tau)+\mathrm{v}(\tau)) = |E| = \sum_{1\le i<j<k\le n} 2^{n-(k-i+1)}\Delta(\pi_i,\pi_j,\pi_k), \tag{2.4}$$

which leads to the desired result. □

We can rewrite Equation 2.3 by setting $b = k-i+1$, the basis of the triangle π_i,π_j,π_k, and considering the function $\Delta_b(\pi)$ that counts the number of triangles in π with basis b:

$$\mathrm{t}(\pi) = \sum_{b=3}^{n} 2^{n-b}\Delta_b(\pi) + |X(\pi)|. \tag{2.5}$$

Note that Equation 2.5, shows that a triangle with a smaller basis contributes more to $\mathrm{t}(\pi)$ than a triangle with larger basis. This supports Conjecture 2.1 since triangles with basis 3 refers to the peaks and valleys of a permutation, and an alternating permutation has the maximum number of both peaks and valleys.

In the same way, Equation 2.5 supports Conjecture 2.2, since a naive way to maximize the number of triangles of length $n-3$ is to set $\pi_1 = \lfloor\frac{n}{2}\rfloor$ and $\pi_n = \lfloor\frac{n}{2}\rfloor \pm 1$, which would mean that every other number π_i would make a triangle with π_1 and π_n, since there is not an integer between them.

Observe that given $b \ge 3$, and a permutation with length n, there are at most $(b-2)(n-(b-1))$ triangles with basis b in π. Since for every triangle of basis b, there are $b-2$ numbers between the endpoints of the basis and there are $(n-(b-1))$ possible choices for these endpoints. This gives us the following upper bound for $\mathrm{t}_{max}(n)$:

$$\begin{aligned}\mathrm{t}_{max}(n) &\le |X(\pi)| + \sum_{b=3}^{n} 2^{n-b}(b-2)(n-(b-1)) = \\ &= 2^{n-1}(n-4) + n + 2 + 2^n - n - 1 - \binom{n}{2} = \\ &= \frac{1}{2}(2^n-(n+1))(n-2).\end{aligned}$$

Note that no permutation $\pi \in RC(5)$ contains every triangle of basis 4. For example, in the permutation $\pi = 24153$, the triple $2,4,5$ is not a triangle. Since the number of triangles of different bases is already maximized, it cannot swap a triangle of a different basis for a triangle of basis 4, which means that π, is in a certain way, triangle saturated. Similarly, we notice that any permutation $\pi \in RC(6)$ contains two of such triples. For example, the permutation $\pi = 326154$ contains the triples 321 and 654 that are not triangles. For permutations of $RC(7)$ there are three of such triples and for any permutation of $RC(8)$, there are seven triples that are not triangles. As we can see in the permutation 43718265, triples: 431, 378, 126, 865, of basis 4, are not triangles, and there are 478, and 125 of basis 5 and 432 of basis 6. This idea could further be explored to improve the upper bound for t_{max}, but due to time constraints we shall leave it for future work.

Recall that $L(n) = \frac{\mathrm{t}_{max}(n+1)}{\mathrm{t}_{max}(n)}$, and let $M(n) = \max_{\pi \in S_n} \sum_{b=3}^{n} \Delta_b(\pi) 2^{n-b}$. We have:

$$L(n) = \frac{\mathrm{t}_{max}(n+1)}{\mathrm{t}_{max}(n)} = \frac{2^{n+1} - n - 2 - \binom{n+1}{2} + M(n+1)}{2^n - n - 1 - \binom{n}{2} + M(n)}$$

Multiplying and dividing by $1/2^n$, we have:

$$L(n) = \frac{2^n}{2^n} \frac{\mathrm{t}_{max}(n+1)}{\mathrm{t}_{max}(n)} = \frac{2 + \frac{-n-2-\binom{n+1}{2}}{2^n} + M(n+1)/2^n}{1 + \frac{-n-1-\binom{n}{2}}{2^n} + M(n)/2^n}$$

If we prove that $M(n+1) \in o(2^n)$, then we have that the limit $\lim_{n\to\infty} L(n) = 2$. Unfortunately, our best upper bound gives $M(n+1) \in O(n2^n)$, which implies that $\lim_{n\to\infty} L(n) \leq \infty$. This motivates further research for better upper bounds for $M(n)$.

2.2 An Integer Linear Program

A Linear Program is a mathematical modelling technique to write a problem as to find the maximum of a linear function which is subject to linear constraints. The linear function that we try to maximize is called the *objective function* of the model. Consider a function $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$, as an objective function. If the domain X is restricted to the integers, we call the model a *Integer Linear Program* (ILP). In this section we present an ILP model to obtain Roller Coasters of a given size n.

The model's objective function is derived from function t in Theorem (2.3) and the main integer variable $\mathbf{x} = (x_1, x_2, ..., x_n)$ represents the permutation itself. We use auxiliary binary variables $p_{i,j,k}, v_{i,j,k}$ with $1 \leq i < j < k \leq n$, and $w_{i,j}$ with $1 \leq i < j \leq n$, where $p_{i,j,k}$ (resp. $v_{i,j,k}$) indicates whether x_i, x_j, x_k is a peak (resp. a valley), and $w_{i,j}$ indicates whether $x_i > x_j$. That is, the binary variables are equal to 1 if their associated property is satisfied and 0 otherwise.

For $(x_1, \ldots, x_n)$ to be a permutation, we must have $x_i \neq x_j$, for every $i \neq j$. This constraint is expressed by Equations (2.6b) and (2.6c). For $x_i x_j x_k$ to be a triangle, $x_i x_j x_k$ must be either a peak, for which we have $x_j > x_i$ and $x_j > x_k$, and can be expressed by equations $x_j \geq x_i - n(1 - p_{i,j,k}) + 1$ and $x_j \geq x_k - n(1 - p_{i,j,k}) + 1$; or a valley, for which we have $x_j < x_i$ and $x_j < x_k$, and can be expressed by equations $x_j \leq x_k + n(1 - v_{i,j,k}) - 1$ and $x_j \leq x_i + n(1 - v_{i,j,k}) - 1$. These constraints are denoted by $PV_{i,j,k}$ (see Equation (2.6d)).

$$\begin{aligned}
\max \quad & \mathrm{t}(\mathbf{x}) = \sum_{1 \leq i < j < k \leq n} 2^{-(k-i+1)} \left(p_{i,j,k} + v_{i,j,k}\right) && && (2.6a)\\
\text{s.t.} \quad & w_{i,j} + w_{j,i} = 1, && \forall\, i \neq j, && (2.6b)\\
& x_i \geq x_j + n(w_{i,j} - 1) + 1, && \forall\, i \neq j, && (2.6c)\\
& PV_{i,j,k}, && \forall\, i < j < k. && (2.6d)
\end{aligned}$$

Model 2.2: An Integer Programming Model for finding Roller Coasters.

Unfortunately, we were not able to run Model 2.2 for $n \geq 18$. But we were able to extend it by including new constraints based on Conjectures 2.1, and 2.3, which characterize the structure of the Roller Coaster permutations and reduce the feasible region of the ILP.

Observe that assuming Conjecture 2.1 holds, we can obtain an upper bound for the number of Roller Coaster permutations, since Theorem 1.1 gives an approximation for $|W_n|$, the set of the alternating permutations. By Theorem 1.1, we can approximate the probability that a permutation of length n, obtaining

$$\frac{|W_n|}{|S_n|} \approx 2\left(\frac{2}{\pi}\right)^{n+1} \frac{n!}{n!} = 2\left(\frac{2}{\pi}\right)^{n+1},$$

which converges to 0 when $n \to \infty$. Therefore, a model with the additional constraint that the permutation is alternating tends to converge faster as n grows, since the feasible region is smaller. If Conjecture 2.1 holds, it also means that the Roller Coaster permutations become really scarce for large n.

Additionally, using Conjecture 2.3 as a set of additional constraints, which is naturally translated to $x_i \geq n/2$ when i is even, and $x_i < n/2$ when i is odd, as in Equations 2.7d and 2.7e, we reduce the feasible region further. This corresponds to the following model.

$$
\begin{aligned}
\max \quad & \mathrm{t}(\mathbf{x}) = \sum_{1\le i<j<k\le n} 2^{-(k-i+1)}\left(p_{i,j,k}+v_{i,j,k}\right) && && (2.7a)\\
\text{s.t.} \quad & w_{i,j}+w_{j,i}=1, && \forall\, i\neq j, && (2.7b)\\
& x_i \ge x_j + n(w_{i,j}-1)+1, && \forall\, i\neq j, && (2.7c)\\
& x_i \ge n/2 && \forall\, i \text{ even}, && (2.7d)\\
& x_i < n/2 && \forall\, i \text{ odd}, && (2.7e)\\
& PV_{i,j,k}, && \forall\, i<j<k. && (2.7f)
\end{aligned}
$$

Model 2.3: An Integer Programming Model for finding Roller Coasters with additional constraints based on Conjecture 2.3.

2.2.1 New Bounds on Roller Coaster Permutations

Running Model 2.2, we obtained the following Roller Coaster permutations for $14 \le n \le 17$, confirming the construction presented by T. Ahmed and H. Snevily in [2] for $n = 14, 15, 16$, and improving their construction in the case $n = 17$.

Table 2.1: Permutations found with Model 2.2 for n = 14,. . . ,17.

N	
14	[7, 11, 3, 13, 5, 9, 1, 14, 6, 10, 2, 12, 4, 8]
15	[7, 12, 3, 14, 5, 10, 1, 15, 6, 9, 2, 13, 4, 11, 8]
16	[8, 12, 4, 14, 2, 10, 6, 16, 1, 11, 7, 15, 3, 13, 5, 9]
17	[8, 14, 3, 15, 6, 10, 2, 17, 7, 12, 1, 16, 5, 11, 4, 13, 9]

These permutations respectively correspond to the following values of t_{max}:

$$\mathrm{t}_{max} = \{81350, 174954, 374409, \mathbf{798783}\}$$

Since Model 2.2 does not include the additional constraints based on Conjectures 2.1 and 2.3, these results are indeed optimal. In particular, the value $\mathrm{t}_{max}(17) = 798783$ shows an improvement on the lower bound given by T. Ahmed and H. Snevily [2]. Model 2.3, on the other hand, contains additional constraints based on Conjecture 2.3 which makes the feasible region smaller. For this reason, it yielded new permutations for n up to 40 (see Table 2.5), with the same computational power. These new permutations improved some of the lower bounds on $\mathrm{t}_{\max}$ known so far (see Table 2.3).

Note that if Conjecture 2.3 holds, then the permutations found are indeed Roller Coasters, and their respective values of t are $\mathrm{t}_{\max}$. It's worth mentioning that the experiments were coded on Sagemath [42] and we also used the Gurobi solver [27].

In Table 2.2, we present the Roller Coaster candidates obtained for $18 \le n \le 24$.

Table 2.2: Permutations found with Model 2.3 for $18 \leq n \leq 24$.

N	
18	[9, 14, 4, 16, 7, 11, 2, 18, 6, 13, 1, 17, 8, 12, 3, 15, 5, 10]
19	[9, 15, 5, 17, 2, 12, 8, 19, 3, 13, 6, 16, 1, 11, 7, 18, 4, 14, 10]
20	[10, 15, 5, 18, 2, 12, 8, 20, 4, 14, 7, 17, 1, 13, 9, 19, 3, 16, 6, 11]
21	[10, 17, 4, 19, 8, 13, 1, 21, 6, 15, 3, 18, 9, 12, 2, 20, 7, 14, 5, 16, 11]
22	[11, 17, 5, 20, 8, 14, 2, 22, 10, 16, 4, 19, 7, 13, 1, 21, 9, 15, 3, 18, 6, 12]
23	[11, 17, 6, 21, 3, 15, 9, 23, 1, 16, 7, 19, 4, 13, 10, 22, 2, 14, 8, 20, 5, 18, 12]
24	[12, 18, 6, 21, 3, 15, 9, 23, 1, 17, 11, 20, 5, 14, 8, 24, 2, 16, 10, 22, 4, 19, 7, 13]

For $n \geq 25$, the Roller Coaster candidates are presented in the end of the chapter. The bounds for t_{max} obtained with Model 2.3 are presented in Table 2.3.

The t-values obtained from the construction given by T. Ahmed and H. in Snevily [2] are in the lines indicated by 'AS' while the t-values obtained by Model 2.3 are present in the lines indicated by 'BN'. Bold values represent improvements to the known t-values and known lower bounds.

Table 2.3: Table containing the t-value obtained as in [2] and by Model 2.3, for n = 14,...,40. Improved lower bounds are presented with bold text.

	14	15	16	17	18
AS	81350	174954	374409	798471	1700036
BN	81350	174954	374409	**798783**	1700036
	19	20	21	22	23
AS	3596124	7588303	15970785	33596706	70310126
BN	**3597020**	7588303	15970785	33596706	70310126
	24	25	26	27	28
AS	146867861	306492900	639129568	1327542841	2755084935
BN	146867861	**306500899**	**639198976**	**1328781760**	**2758443963**
	29	30	31	32	33
AS	5720021634	11863992638	24524469439	50593221917	104565405932
BN	**5720153893**	11863992638	**24525731250**	**50650675297**	**104569114183**
	34	35	36	37	38
AS	215826275292	444271587981	914139811651	1881877624386	3872524536090
BN	**215844113148**	**444587412964**	**914999923559**	**1882036116393**	**3872525917922**
	39	40			
AS	7948257224143	16292370258569			
BN	**7949294221494**	**16308000242795**			

Observe that this shows that the construction given by T. Ahmed and H. Snevily is a very good approximation for this values of n, even finding an optimal value in the case of $n = 30$.

2.3 On the Number of Triangles in a Permutation

Recall that by Theorem 2.6, given a permutation $\pi \in S_n$, we can evaluate $t(\pi)$ with

$$\mathrm{t}(\pi) = \sum_{b=3}^{n} 2^{n-b}\Delta_b(\pi) + |X(\pi)|.$$

This gives us a natural way to decompose a permutation according to the number of triangles it contains with each basis b. For example, the permutation 312 contains a single triangle of basis 3. And the permutation 2143, contains two triangles of basis 3 and two triangles of basis 4. Let $\mathbb{T}$ denote the total number of triangles in a permutation π, that is:

$$\mathbb{T}(\pi) = \sum_{b} \Delta_b(\pi).$$

Question 2.8. *What is the expected behavior for* $\mathbb{T}(\pi)$ *?*

Moreover, we can define a decomposition $T\colon S_n \to \mathbb{N}^{n-2}$ given by

$$T(\pi) = (\Delta_3(\pi), \Delta_4(\pi), \cdots, \Delta_n(\pi)).$$

We were able to compute T for every permutation $\pi \in RC(n)$, for $n \leq 12$, given by T. Ahmed and H. Snevily [2], and for each n, they all have the same decomposition. The following table presents the evaluation of T over the Roller Coaster permutations for $3 \leq n \leq 12$.

Table 2.4: Table containing $T(\pi)$ for $\pi \in RC(n)$ for $3 \leq n \leq 12$.

n	$T(\pi)$
3	(1)
4	(2,2)
5	(3,3,3)
6	(4,4,6,4)
7	(5,5,8,7,5)
8	(6,6,10,10,10,6)
9	(7,7,15,9,12,9,7)
10	(8,8,18,12,16,12,12,8)
11	(9,9,20,15,22,14,18,13,9)
12	(10,10,22,18,28,16,24,18,16,10)

Observe that if $\Delta_3(\pi) = n-2$, for all $\pi \in RC(n)$, for all n, then π is alternating and hence Conjecture 2.1 holds, since it represents an alternating sequence. Moreover, proving that $\Delta_n(\pi) = n$, for $\pi \in RC(n)$, for all n, is equivalent to proving that $\pi_1 = \pi_n \pm 1$.

A very difficult problem arises with the inverse operation, that is, how can we construct a permutation $\pi \in S_n$ that has a given triangle decomposition $T' = (\Delta_3, \Delta_4, \cdots, \Delta_n)$. A sequence of integer is a *realizable triangle sequence*, if there is a permutation $\pi \in S_n$ such that $T(\pi) = T'$. An analogous problem to the Graph

Realization problem, which asks which sequences of integers can be represented as the degree sequence of a finite simple graph on n vertices, arises:

Question 2.9. *Which sequences of integers are realizable triangle sequences?*

2.4 Alternative Representations of Roller Coasters

In this section, we present some natural representations of Roller Coaster permutations that could be explored further in light of different ideas. We nevertheless try to motivate them by relating to other well studied areas of Combinatorics.

2.4.1 Roller Coaster as Tournaments

The first representation is called the *tournament representation* of a permutation, defined in the following way: given a permutation $\pi \in S_n$, we build a directed graph that has a vertex for each entry π_i and a directed edge from π_i to π_j for all $j > i$. We illustrate the tournament representation of $\pi = 24153$ In Figure 2.4.

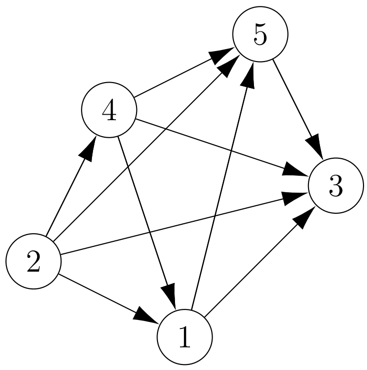


Figure 2.4: Digraph Representation of $\pi = 24153$.

We can further explore this definition by giving a weight to each edge related to the difference of its endpoints. If we simply have that the weight w of an edge (v_1, v_2) is $w((v_1, v_2)) = v_2 - v_1$ we would have that every edge in the tournament with a positive weight represents an ascent and every edge with a negative weight represents a descent. Then, we can write the *adjacency matrix* associated to the tournament, denoted by M, in which each entry $M_{i,j}$ equals to $\pi_j - \pi_i$, representing the weight of the edge (π_i, π_j). This matrix has some interesting properties, namely:

- For all i, $M_{i,i} = 0$, since there are no loops.
- Let $M(k) = M^k$, each entry $M(k)_{i,j}$ represents the sum of the weights of all the edges of all paths of length k between π_i and π_j in the tournament.

We can build a normalized adjacency matrix N, where $N_{i,j} \in \{-1, 0, 1\}$, from M by having each entry $N_{i,j} = \frac{M_{i,j}}{|M_{i,j}|}$. Observe that in the tournament, as well as in the normalized tournament, every triangle is a path of length 2 that has different signs in the weights of its edges. In other words, a triangle consists of two 2 consecutive edges e_1, e_2 for which $w(e_1)w(e_2) = -1$. This encourages, using fast algorithms for matrix multiplication, the search for an algorithm with time complexity $O(n^{\log_2 7})$, or even smaller, as show by J. Alman and V. Williams [4], for cosmically large values of n, to evaluate t_π and could imply in a much faster way to find Roller Coaster permutations.

2.4.2 Permutation Graphs of Roller Coasters

In this section we exhibit some interesting properties of the Roller Coaster permutations when represented as permutation graphs.

It is known that the permutation graph of a permutation π is isomorphic to the permutation graph of π^{rc}, and the permutation graph G of a permutation π is the complement G^c of the permutation graph of π^c. For each $\pi \in RC(n)$, we can select one representative π of this equivalence class, and, removing π^{rc}, π^c and π^r from $RC(n)$, construct the set $SRC(n)$, that contains only one representative for each symmetry through the Permutation Graphs.

For example, from T. Ahmed and H. Snevily [2], we have:

$$RC(5) = \{24153, 31524, 25143, 32514, 41523, 34152, 35142, 42513\}$$

We can simplify this set selecting the representative $\{24153, 25143\}$. These permutations have the following permutation graphs:

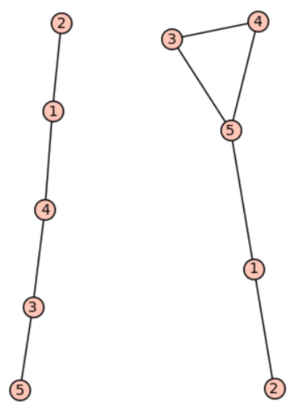


Figure 2.5: Permutation graphs for 24153 (left) and 25143 (right).

Unfortunately, we were not able to relate in any way the permutation graphs between these two representative nor the measure of alternation. Note that the permutation graph associated with the permutation $\pi = n(n-1)\cdots 1$ is the complete graph, and the permutation graph associated with the permutation $\pi^r = 1, 2, 3, \ldots, n$ is the empty graph. This means that $\mathrm{t}(\pi)$ is not related with the density of the permutation graph of π, since $\mathrm{t}(\pi) = \mathrm{t}(\pi^r)$.

Further exploration and analysis through the permutation graphs for $\pi \in RC(n)$, motivates the following conjecture.

Conjecture 2.10. *For $n \geq 4$, if G is a permutation graph of a permutation $\pi \in RC(n)$, then G is connected.*

2.5 Conclusion

In this chapter, we presented an alternative and fast algorithm to calculate t, based on the number of triangles in a permutation. This motivates a further study of the decomposition of a permutation in triangles by reducing the problem to graphs decomposition or other representations.

We also provided an Integer Linear Programming model to find Roller Coasters, in which we were able to find new Roller Coasters for $n \leq 17$, as well as an extended ILP model that provided us with new lower bounds for $\mathrm{t}_{max}(n)$ for $n \leq 40$. It is worth mentioning that if proven, Conjecture 2.1, and 2.3, would imply that these lower bounds are indeed optimal values for t_{max}, and hence that the candidates presented are indeed Roller Coaster permutations for $n \leq 40$.

Finally, we showed the problem of Roller Coaster permutations through different points of view, stablishing connections with different fields of Combinatorics, encouraging further study that may result in advancements in both Roller Coaster permutations and other fields.

Roller Coaster candidates running Model 2.3

Table 2.5: Permutations found with Model 2.3 for $14 \leq n \leq 24$.

N	
14	[7, 11, 3, 13, 5, 9, 1, 14, 6, 10, 2, 12, 4, 8]
15	[7, 12, 3, 14, 5, 10, 1, 15, 6, 9, 2, 13, 4, 11, 8]
16	[8, 12, 4, 14, 2, 10, 6, 16, 1, 11, 7, 15, 3, 13, 5, 9]
17	[8, 14, 3, 15, 6, 10, 2, 17, 7, 12, 1, 16, 5, 11, 4, 13, 9]
18	[9, 14, 4, 16, 7, 11, 2, 18, 6, 13, 1, 17, 8, 12, 3, 15, 5, 10]
19	[9, 15, 5, 17, 2, 12, 8, 19, 3, 13, 6, 16, 1, 11, 7, 18, 4, 14, 10]
20	[10, 15, 5, 18, 2, 12, 8, 20, 4, 14, 7, 17, 1, 13, 9, 19, 3, 16, 6, 11]
21	[10, 17, 4, 19, 8, 13, 1, 21, 6, 15, 3, 18, 9, 12, 2, 20, 7, 14, 5, 16, 11]
22	[11, 17, 5, 20, 8, 14, 2, 22, 10, 16, 4, 19, 7, 13, 1, 21, 9, 15, 3, 18, 6, 12]
23	[11, 17, 6, 21, 3, 15, 9, 23, 1, 16, 7, 19, 4, 13, 10, 22, 2, 14, 8, 20, 5, 18, 12]
24	[12, 18, 6, 21, 3, 15, 9, 23, 1, 17, 11, 20, 5, 14, 8, 24, 2, 16, 10, 22, 4, 19, 7, 13]
25	[12, 19, 5, 16, 9, 23, 2, 18, 10, 25, 1, 14, 7, 21, 4, 17, 11, 24, 3, 15, 8, 22, 6, 20, 13]
26	[13, 20, 6, 23, 10, 16, 3, 25, 8, 18, 1, 22, 12, 15, 5, 26, 9, 19, 2, 24, 11, 17, 4, 21, 7, 14]
27	[13, 20, 7, 24, 3, 17, 11, 23, 5, 19, 9, 27, 1, 15, 12, 22, 4, 18, 8, 26, 2, 16, 10, 25, 6, 21, 14]
28	[14, 21, 7, 25, 3, 17, 11, 27, 5, 19, 9, 23, 1, 16, 13, 28, 6, 20, 10, 24, 2, 18, 12, 26, 4, 22, 8, 15]
29	[14, 23, 7, 26, 11, 18, 3, 28, 9, 20, 5, 24, 13, 16, 1, 29, 8, 21, 4, 25, 12, 17, 2, 27, 10, 19, 6, 22, 15]
30	[15, 23, 7, 27, 11, 19, 3, 29, 13, 21, 5, 25, 9, 17, 1, 30, 14, 22, 6, 26, 10, 18, 2, 28, 12, 20, 4, 24, 8, 16]
31	[15, 24, 8, 27, 11, 20, 3, 29, 13, 18, 5, 25, 9, 22, 1, 31, 14, 17, 6, 26, 10, 21, 2, 30, 12, 19, 4, 28, 7, 23, 16]
32	[16, 24, 8, 28, 4, 20, 12, 30, 2, 22, 14, 26, 6, 18, 10, 32, 1, 23, 15, 27, 7, 19, 11, 31, 3, 21, 13, 29, 5, 25, 9, 17]
33	[16, 26, 8, 29, 11, 20, 3, 31, 13, 22, 5, 27, 10, 18, 1, 33, 15, 24, 6, 28, 9, 19, 2, 32, 14, 23, 4, 30, 12, 21, 7, 25, 17]
34	[17, 26, 8, 30, 13, 21, 4, 32, 11, 23, 2, 28, 15, 19, 6, 34, 10, 25, 1, 29, 16, 20, 7, 33, 12, 24, 3, 31, 14, 22, 5, 27, 9, 18]
35	[17, 27, 9, 31, 4, 22, 14, 29, 6, 23, 12, 34, 2, 20, 16, 30, 7, 25, 10, 35, 1, 19, 15, 28, 5, 24, 11, 33, 3, 21, 13, 32, 8, 26, 18]
36	[18, 27, 9, 32, 4, 22, 14, 34, 7, 24, 12, 29, 2, 20, 16, 36, 6, 26, 11, 31, 1, 21, 17, 35, 8, 25, 13, 30, 3, 23, 15, 33, 5, 28, 10, 19]
37	[18, 29, 9, 24, 14, 34, 3, 21, 11, 31, 6, 26, 17, 36, 2, 22, 12, 32, 7, 27, 15, 37, 1, 20, 10, 30, 5, 25, 16, 35, 4, 23, 13, 33, 8, 28, 19]
38	[19, 29, 9, 34, 14, 24, 5, 36, 17, 27, 7, 31, 11, 21, 2, 38, 16, 26, 6, 33, 13, 23, 1, 37, 18, 28, 8, 32, 12, 22, 3, 35, 15, 25, 4, 30, 10, 20]
39	[19, 30, 10, 35, 14, 24, 4, 38, 17, 22, 7, 31, 11, 27, 1, 37, 16, 23, 6, 33, 13, 28, 3, 39, 18, 21, 8, 32, 12, 26, 2, 36, 15, 25, 5, 34, 9, 29, 20]
40	[20, 30, 11, 35, 5, 25, 15, 38, 2, 27, 18, 32, 7, 22, 13, 40, 3, 29, 17, 34, 9, 24, 12, 37, 1, 28, 19, 33, 8, 23, 14, 39, 4, 26, 16, 36, 6, 31, 10, 21]

3 | Extremal Graph Theory

The purpose of this chapter is to present classical results in Extremal Graph Theory as well as some of the techniques used in their proofs, and its objective is to familiarize the reader and serve as a foundation for Section 3.1, in which we cover one of the main conjectures of this field.

The main goal of Extremal Combinatorics is to study problems where one should determine the maximum or minimum possible cardinality of a finite collection of objects that satisfies certain property.

We start with the following problem, in the context of Set Theory. A set A is called *sum-free* if there exists no triple $a, b, c \in A$ such that $a + b = c$. Note that there is no restriction of a, b, c being different, therefore, the triple $(1, 1, 2)$ counts as a sum, since $1 + 1 = 2$. As an example and motivational problem, we have the following question

Question 3.1. *What is the maximum cardinality of a sum-free set $A \subseteq [n]$?*

Some simple sets that satisfies such property are

- $\{1, 3, 5, 7, \dots, \lfloor \frac{n+1}{2} \rfloor - 1\}$ – odd numbers;
- $\{\lfloor \frac{n}{2} \rfloor + 1, \lfloor \frac{n}{2} \rfloor + 2, \lfloor \frac{n}{2} \rfloor + 3, \cdots, n\}$ – the numbers greater than $n/2$.

Therefore, we can conclude that $|A| \geq \lceil \frac{n}{2} \rceil$. But it is not clear that actually $|A|$ cannot be bigger than $\lceil \frac{n}{2} \rceil$. The following results shows that $\lceil n/2 \rceil$ is indeed tight.

Theorem 3.2. *Let $A \subseteq \{1, 2, 3, \dots, n\}$, if $|A| \geq \lceil \frac{n}{2} \rceil + 1$, then A is not sum-free.*

Proof. Suppose that A is a sum-free set with $|A| \geq \lceil \frac{n}{2} \rceil + 1$, and let $c = \max_{x \in A} x$. Consider the set $B = c - A$, defined as the set $\{c - a : a \in A\}$. Take $B^* = B \setminus \{0\}$. Note that $|A| = |B^*| + 1 \geq \lceil \frac{n}{2} \rceil + 1$. Therefore, we have that $|A| + |B^*| \geq 2\lceil \frac{n}{2} \rceil + 1$, which implies that $|A \cap B^*| \neq \varnothing$ by the Pigeonhole Principle. Let $x \in B^* \cap A$, this means that $x = c - a$ for a, c in A, a contradiction. $\square$

This simple problem motivates the following question: which other properties become unavoidable for sets greater than a certain size? Determining this threshold became one of the main themes studied in Extremal Combinatorics. In the following sections, we present some classical results in Extremal Graph Theory, that deals with this question in the context of graphs.

One of the first results in Extremal Graph Theory is Mantel's theorem [**?**], proved in 1907, which says that any graph $G = (V, E)$ with $|V| = n$ and $|E| \leq \lfloor n/2 \rfloor \lceil n/2 \rceil$ edges, contains at least one copy of K_3.

Theorem 3.3 (Mantel, 1907)**.** *Let G be a graph on n vertices. If G does not contain a copy of K_3, then $|E(G)| \leq \lfloor n/2 \rfloor \lceil n/2 \rceil$.*

This is clearly best possible, as one may partition the set of vertices into two sets of size $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ and form the complete bipartite graph between them, which has precisely $\lfloor n/2 \rfloor \lceil n/2 \rceil$, and has no copies of K_3.

Proof. Consider a K_3-free graph $G = (V, E)$ with n vertices. Choose an edge $uv \in E$. Since G is K_3-free, there are at most $n - 1$ edges adjacent to the edge uv. Consider the graph $G' = G - \{u, v\}$. Since $G' \subseteq G$, G' is also K_3-free. By induction, we have

$$|E(G')| \leq \left(\left\lfloor \frac{n-2}{2} \right\rfloor \right) \left(\left\lceil \frac{n-2}{2} \right\rceil \right) = \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil - n + 1.$$

Therefore, we obtain

$$|E(G)| \leq |E(G')| + n - 1 \leq \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil,$$

as desired. □

This problem motivates the definition of the *extremal number*, denoted by $ex(n, \mathcal{H})$, which is the maximum number of edges in a graph G on n vertices that is H-free for every graph $H \in \mathcal{H}$. When $\mathcal{H} = \{H\}$, we simply write $ex(n, H)$. Finally, we call a graph G an *H-extremal graph* if $|E(G)| = ex(n, H)$. This gives the following reformulation of Theorem 3.3.

Theorem 3.4 (Mantel, 1907)**.** *For $n \in \mathbb{N}$, $ex(n, K_3) = \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil - n + 1$.*

Consider graphs G and H with v_G, e_G (resp., v_H, e_H) denoting the number of vertices and edges in G (resp., in H) and suppose $v_G \geq v_H$. Observe that if $e_G = \binom{v_G}{2}$, then G must contain a copy of H and if $e_G < e_H$, then G is H-free. In fact, containing a graph as subgraph is a *monotonic property*, which means that if we add edges to G, the graph H will not disappear. Therefore, we can conclude that $ex(v_G, H)$ is well defined for all graphs H with $0 < e_H < \binom{v_G}{2}$.

Even though such a number must exist, proving bounds for $ex(n, H)$ can be difficult. The next result was only proved in 1941, by P. Turán [44], generalizing Mantel's theorem.

Theorem 3.5 (Turán, 1941)**.** *For $m, n \in \mathbb{N}$, if $m \leq n$, we have*

$$ex(n, K_m) = \left(1 - \frac{1}{m-1}\right)\frac{n^2}{2},$$

in the same work, P. Turán characterized the structure of such extremal graphs, which are the complete $(m-1)$-partite graphs for whose parts has sizes as equal as possible, also known as *Turán's Graphs*, and denoted by $T_m(n)$, (see Figure 3.1). A classical proof for Theorem 3.5, that can be found in [3], goes as follows.

Proof. Let G be a K_m-free graph on n vertices and $ex(n, K_m)$ edges. Note that if $n \leq m-1$, then G cannot contain any copy of K_m and has at most $\binom{n}{2}$ edges. So in this case, we only have to prove that

$$\binom{n}{2} = \frac{n(n-1)}{2} \leq \left(1 - \frac{1}{m-1}\right)\frac{n^2}{2}.$$

In which dividing by $n^2/2$, we obtain

$$1 - \frac{1}{n} \leq \left(1 - \frac{1}{m-1}\right),$$

which holds, because $n \leq m-1$.

Thus we assume that $n \geq m$. We may assume that G has the maximum number of edges possible without containing a copy of K_m. This implies that G contain at least one copy of K_{m-1}, since otherwise we can add edges to get one while avoiding a K_m. Let us denote the copy of K_{m-1} in G by A and define $B = G - A$. Observe that A has $e_A = \binom{m-1}{2}$ edges. Let us denote the number of edges in B by e_B and the number of edges that join A and B by $e_{A,B}$.

Since G is K_m-free, every vertex of B is adjacent to at most $m-2$ vertices in A. Therefore we have $e_{A,B} \leq (m-2)(n-m+1)$. By induction, we have

$$e_B \leq \frac{1}{2}\left(1 - \frac{1}{m-1}\right)(n-m+1)^2.$$

Therefore, we have

$$\begin{aligned}
|E(G)| &\leq e_A + e_B + e_{A,B} \\
&\leq \binom{m-1}{2} + \frac{1}{2}\left(1 - \frac{1}{m-1}\right)(n-m+1)^2 + (m-2)(n-m+1) \\
&= \left(1 - \frac{1}{m-1}\right).
\end{aligned}$$

Concluding the proof. □

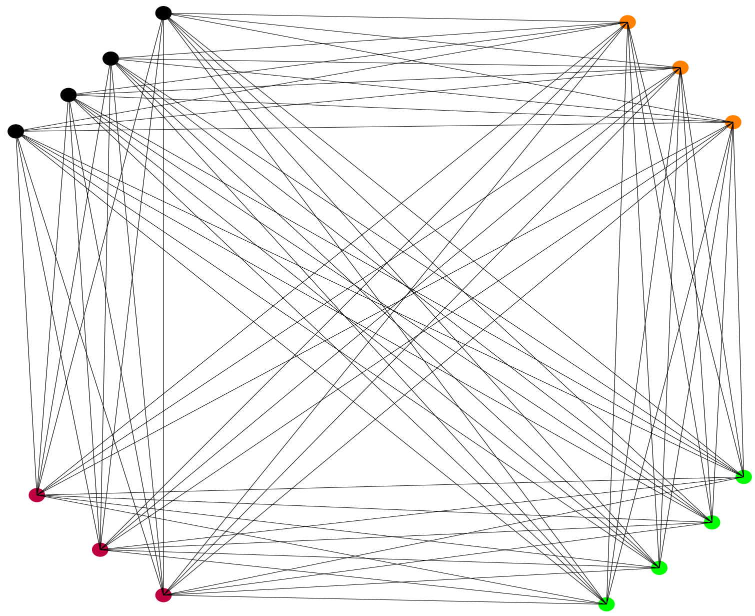


Figure 3.1: An illustration of the Turán's graph $t_4(13)$.

Note that for every complete graph K_m and K_n, if $m > n$, then $K_n \subseteq K_m$. Let $\mathcal{K}^L = \{K_3, K_4, \cdots, K_L\}$. Then $ex(n, \mathcal{K}^L) = ex(n, K_3) = \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil$, for all $L \geq 3$, given by Theorem 3.3. This is the first result for a whole family of graphs, even though it is a trivial result, and leads us to the following statement.

Proposition 3.6. *Let $\mathcal{H} = \{H_1, H_2, \ldots, H_r\}$ be a family of graphs and G be a graph on n vertices. We have*

$$ex(n, \mathcal{H}) \leq \min_{H \in \mathcal{H}} ex(n, H).$$

Because if G satisfies $E(G) > \min_{H \in \mathcal{H}} ex(n, H)$, then there is an index $i \in [r]$ such that $H_i \subseteq G$, meaning that G is not $\mathcal{H}$-free.

In 1946, P. Erdős and A. Stone [20] found the following asymptotic bound for the extremal number for non-bipartite graphs H.

Theorem 3.7 (Erdős–Stone, 1946)**.** *If H is a non-empty graph with $\chi(H) > 2$, then*

$$ex(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \frac{n^2}{2}.$$

Unfortunately, for the case of bipartite graphs, since $\chi(H) = 2$, Theorem 3.7 does not provide any useful bound for $ex(n, H)$. Different results exist concerning the extremal number for bipartite graphs. In 1954, T. Kővári, V. Sós e P. Turán [33] showed the following bound to the extremal number to complete bipartite graphs, denoted by $K_{s,t}$.

Theorem 3.8 (Kővári–Sós–Turán, 1954)**.** *Let $s, t \in \mathbb{N}$, with $s \leq t$. There exists $C = C(s, t) > 0$ such that*

$$ex(n, K_{s,t}) \leq Cn^{2-1/s}.$$

The proof of this result relies in counting a specific structure called an *s-cherry*, which consists of a copy of $K_{1,s}$ and can be found in [33].

For different classes of bipartite graphs, the first result for the extremal number of C_{2k} was found by J. Bondy and M. Simonovits [10] in 1974.

Theorem 3.9 (Bondy–Simonovits, 1974)**.** *There exists a positive integer n_0 such that for integers $n \geq n_0$, we have*

$$ex(n, C_{2k}) \leq 20kn^{1+1/k}.$$

This bound was later improved to $ex(n, C_{2k}) \leq 80n^{1+1/k}\sqrt{k \log k} + 10k^2 n$, by B. Bukh and Z. Jiang [14], in 2017.

Despite many decades of intense interest [45], constructions of C_{2k}-free graphs with $\Omega(n^{1+1/k})$ edges are known only for $k = 2, 3, 5$. An algebraic construction that yields n-vertex graphs with no C_4, C_6, and C_{10} with a close to the maximum number of edges was found by R. Wenger [47] in 1991, but only recently, in 2021, a geometrical interpretation of this construction was given by D. Conlon [17].

Another special case for bipartite graphs is when $H = P_\ell$ is a path, as shown by P. Erdős and T. Gallai, in 1959, to have the following extremal number.

Theorem 3.10 (Erdős–Gallai, 1959)**.** *Let P_ℓ denote a path of length (number of edges) ℓ. Then, for $n \geq \ell$*

$$ex(n, P_\ell) = \frac{(\ell - 1)n}{2}.$$

In order to prove this theorem, we need the following lemma.

Lemma 3.11. *Every connected graph on n vertices contains a copy of P_k, where*

$$k = \min\{2\delta(G), n - 1\}.$$

Proof. Given a connected graph $G = (V, E)$, let $P = p_1 p_2 \cdots p_{\ell+1}$ be the longest path of G. Note that if there is a vertex $v \in V(G)$ such that $v \notin V(P)$ and v is adjacent to either p_1 and $p_{\ell+1}$, then it contradicts the maximality of P.

We claim that if $\ell < k \leq 2\delta(G)$, then one can find a cycle C in G using the vertices of $V(P)$. Indeed, by the Pigeonhole Principle, there is an $i \in [\ell]$ such that $\{p_1 p_i, p_{i-1} p_{\ell+1}\} \subseteq E(G)$.

Since G is connected, if there is a vertex $v \notin V(C)$, then there is an vertex p_i that is adjacent to v, which contradicts the maximality of P.

Therefore, either $\ell = n - 1$, which is a path containing all the vertices of G, or $\ell = 2\delta(G)$, as desired. □

With this lemma at hand the proof for Theorem 3.10 goes as follows.

Proof. The proof proceeds by induction on n. Let G be a P_ℓ-free graph on n vertices. If $n \leq \ell$, then

$$|E(G)| \leq \binom{n}{2} = \frac{n(n-1)}{2} \leq \frac{n(\ell-1)}{2}.$$

Therefore, we may assume that $n \geq \ell+1$, and that the theorem holds for graphs with $n' < n$ vertices. First, suppose that G is not connected. This implies that $G = \bigcup_i G_i$ where G_i is a connected component of G and G_i, G_j are disjoint graphs. Since G is P_ℓ-free, we have that G_i is P_ℓ-free for all i. Therefore, by the induction hypothesis, we have

$$|E(G)| = \sum_{i=1}^{m} |E(G_i)| \leq \frac{(\ell-1)}{2} \sum_{i=1}^{m} |V(G_i)| = \frac{n(\ell-1)}{2}.$$

as desired.

Thus we may assume that G is connected. Note that by Lemma 3.11, if $\delta(G) \geq \ell/2$, then $P_\ell \subseteq G$. On the other hand, if $\delta(G) < \ell/2$, then there is a vertex $u \in V(G)$ for which $\mathrm{d}(u) \leq \frac{(\ell-1)}{2}$. Let $G' = G - \{u\}$, and note that G' is P_ℓ-free. Then, by the induction hypothesis, we have

$$|E(G)| \leq |E(G')| + \frac{(\ell-1)}{2} = \frac{(\ell-1)(n-1)}{2} + \frac{(\ell-1)}{2} = \frac{(\ell-1)n}{2},$$

as desired. $\square$

Recall that $\mathcal{T}_k$ denotes the family of trees on k vertices. Since $P_{k-1} \subseteq \mathcal{T}_k$, from Proposition 3.6 we have for a positive integer $n \in \mathbb{N}$

$$\frac{(k-2)n}{2} = ex(n, P_{k-1}) \geq \min_{T \in \mathcal{T}_k} ex(n, T) \geq ex(n, \mathcal{T}_k).$$

This observation motivated the statement of the following conjecture from 1963, due to P. Erdős and V. Sós [12], which generalizes Theorem 3.10 for all trees on k vertices, and is the main subject of Section 3.1.

Conjecture 3.12 (Erdős–Sós)**.** *Let $n, k \in \mathbb{N}$. For all $T \in \mathcal{T}_k$, we have*

$$ex(n, T) \leq \frac{n(k-2)}{2}.$$

In other words, Conjecture 3.12 states that if G is a graph on n vertices with more than $n(k-2)/2$ edges, then G contains every tree on k vertices as a subgraph. This conjecture, if true, gives a bound for the number of unlabeled trees, as the number of subgraphs of G.

In the early 90s, M. Ajtai, J. Komlós, M. Simonovits and E. Szemerédi announced a proof for Conjecture 3.12, but, unfortunately, it was never published. For this reason, the scientific community kept working in many particular cases. A weaker version of this statement can be proved with the help of the following lemmas.

Lemma 3.13. *Every non-empty graph G contains a subgraph H with*

$$\delta(H) > \frac{|E(G)|}{|V(G)|}.$$

Proof. Consider the following graph sequence

$$G = G_0 \supset G_1 \supset \cdots \supset G_m = H,$$

in which for each $i \in [n]$ the graph G_i is obtained from G_{i-1} by removing a vertex with degree at most $|E(G)|/|V(G)|$ in G_{i-1}. Suppose that $m \geq |V(G)|$. Let v_i be the vertex removed from the graph G_{i-1} to obtain G_i. Since $\mathrm{d}(v_i) \leq |E(G)|/|V(G)|$ in G_{i-1} and since all edges of G would be removed in some step, we have that

$$|E(G)| = \sum_{i=1}^{m} \mathrm{d}_{G_{i-1}}(v_i) \leq (|V(G)| - 1)\frac{|E(G)|}{|V(G)|} < |E(G)|,$$

which is a contradiction.

Therefore, $m < |V(G)|$, and this process stops before removing every vertex $v \in V(G)$, which results in a subgraph H as desired. □

The next lemma gives us necessary conditions to find the desired tree in H.

Lemma 3.14. *For $k \in \mathbb{N}$, if G is a graph for which $\delta(G) \geq k - 1$, then $T \subseteq G$, for all $T \in \mathcal{T}_k$.*

Proof. The proof is by induction on k. If $k = 2$, then T is a pair of adjacent vertices, and, therefore, it is contained in every graph with at least one edge. Let T be a tree with $k \geq 3$ vertices, let be G a graph for which $\delta(G) \geq k - 1$, and suppose that the statement holds for $k' < k$.

First, since T is a tree with at least three vertices, it contains at least one leaf v. Let $T' = T - \{v\}$. Since T' is a tree, by the induction hypothesis, we have that $T' \subseteq G$. Let u be the unique vertex in T for which $uv \in E(T)$. Since $\mathrm{d}_G(u) \geq \delta(G) \geq k - 1$, there exists a vertex $x \in G$ such that x is not in the copy of T' in G. For this reason, we can extend this copy by setting the image of v as x, obtaining a copy of T in G, as desired. □

With these lemmas, we can prove the following statement, which is a weaker version of the Erdős–Sós conjecture.

Theorem 3.15. *Let $n, k \in \mathbb{N}$, and let T be a tree on k vertices, then*

$$ex(n, T) \leq (k - 2)n.$$

Proof. Suppose that there is a graph G with $|E| > (k-2)n$ that is T-free. By Lemma 3.13 there exists a subgraph $H \subseteq G$ with

$$\delta(H) > \frac{|E(G)|}{|V(G)|} = k-2,$$

and, therefore, $\delta(H) \geq k-1$.

By Lemma 3.14, we have that $T \subseteq H \subseteq G$, as desired. $\square$

3.1 Erdős–Sós Conjecture

In this section, we display a plethora of partial results concerning Conjecture 3.12 to identify possible directions one may follow to contribute to the state of the art with respect to this problem. We divided the results in four different directions, each representing a weakening of Conjecture 3.12 that consider a different kind of hypothesis. For some of these theorems, we show a sketch of its proof and some considerations, as an attempt to bring clarity to the proofs.

It is worth restating Conjecture 3.12 in terms of the average degree $\overline{\mathrm{d}}$ of the graph G, which is defined as $\overline{\mathrm{d}}(G) = \frac{1}{|V(G)|}\sum_{v \in V(G)} \mathrm{d}(v)$.

Conjecture 3.16 (Erdős–Sós)**.** *Let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, then $T \subseteq G$.*

Suppose that $|V(G)| = n$, note that the hypothesis $\overline{\mathrm{d}}(G) > k-2$ implies the following inequality.

$$|E(G)| = \frac{n\overline{\mathrm{d}}(G)}{2} > \frac{n(k-2)}{2}. \tag{3.1}$$

Which implies that $|E(G)| \geq \lfloor \frac{n(k-2)}{2} \rfloor + 1$ Thus, if n is even, then $|E(G)| \geq \frac{n(k-2)}{2} + 1$; and, if n is odd, we have $|E(G)| \geq \frac{n(k-2)}{2} + \frac{1}{2}$.

3.1.1 Large Trees

In this section, we present some results found in the literature concerning Conjecture 3.16 in the case that the graph G has $n = k + c$ vertices, where k is the number of vertices of the tree T, and c is a nonnegative constant.

The first theorem in this direction was proved by B. Zhou in 1984 [51] and considers the case $c = 0$, in which $n = k$.

Theorem 3.17 (Zhou, 1984)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$ and $n = k$, then $T \subseteq G$.*

The main technique of the proof is to first find an isomorphism φ' of a tree $T' = T - \{v\}$, where v is a leaf of T, into $G' = G - u$, and u is a universal vertex of G, which is illustrated in Figure 3.2. It is worth mentioning that this idea is the same as the idea of the proof of Lemma 3.14.

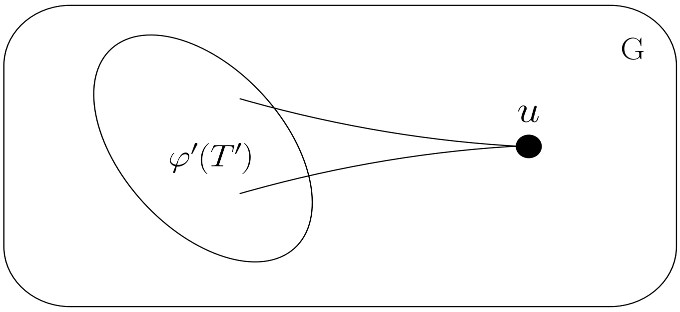


Figure 3.2: An illustration of the isomorphism φ' of T' into $G-\{u\}$ for the universal vertex $u \in V(G)$.

Thus, one may extend the isomorphism φ' into φ by setting $\varphi(v) = u$ and $\varphi(x) = \varphi'(x)$, for $x \neq v$. Note that since u is universal, every vertex in $V(\varphi'(T'))$ is adjacent to u in G, which implies that φ is an isomorphism of T into G.

A simplified version of the proof Theorem 3.17 goes as follows.

Proof. For $n = 3$, the theorem clearly holds since if G has $\overline{\mathrm{d}}(G) > 1$ it has a vertex v with $\mathrm{d}(v) = 2$, and, therefore, contains all trees on three vertices. The proof follows by induction on n, so suppose that the assertion holds for all $m \leq n-1$.

Observe that since $\overline{\mathrm{d}}(G)$ is greater than $k-2$, there must be at least one vertex $u \in V(G)$ such that $\mathrm{d}(u) = k-1$. First, suppose that n is even, by Equation 3.1, we may suppose that G satisfies $2|E(G)| \geq n(n-2)+2$.

Given a tree T on k vertices. If T is a path, then the statement holds due to Theorem 3.10. Therefore, one may suppose that T is not a path. Consider the tree $T' = T - \{v\}$, for a leaf v of T, consider the graph G' obtained by deleting u of G, that is $G' = G - \{u\}$. Then

$$2|E(G')| \geq n(n-2) + 2 - 2\mathrm{d}(u) = n^2 - 4n + 4 > (n-1)(n-3).$$

Which implies that $\overline{\mathrm{d}}(G') > n-3$. Therefore, by the induction hypothesis, we have $T' \subseteq G'$. Note that we can extend the isomorphism φ' of T' into G' by setting the image of $\varphi(v) = u$, since $\mathrm{d}(u) = n-1$ and, thus is adjacent to every other vertex of G. This implies that $T \subseteq G$, as desired.

Now, if n is odd, and T is a star, the fact that $u \in V(G)$, with $\mathrm{d}(u) = n-1$, implies that $T \subseteq G$. Therefore, one may suppose that T is not a star. Recall that ∂T is the tree obtained from T by deleting all leaves of T. Let L denote the set of all leaves in T, and L' denote the set of all leaves in ∂T. Let a *branch* of T be a set consisting of one vertex v of L' and the leaves of L that are adjacent to v. Since

$|L| \geq 2$, there must be one branch B of T such that $2 \leq |V(B)| = p \leq \lfloor n/2 \rfloor$. Let $A = \{v_1, v_2, \ldots, v_{p-1}\}$.

$$\sum_{v \in A} \mathrm{d}(v) \leq (p-1)(n-2).$$

If for every $v_i \in A$ and $v_j \in G - A$, v_i and v_j are adjacent in G, delete all the leaves of T that belong to B and denote the remaining tree of order $n - p + 1$ by T_2. There are $(p-1)(n-p+1)$ edges between A and $G - A$. Therefore, we have

$$\begin{aligned} 2|E(G-A)| &\geq n(n-2) + 1 - 2(p-1)(n-p+1) \\ &\quad - [(p-1)(n-2) - (p-1)(n-p+1)] \\ &= (n-p)^2 \\ &= (n-p-1)(n-p+1) + 1. \end{aligned}$$

Again, by the inductive hypothesis, we have that $T_2 \subseteq G - A$ and, we can extend T_2 to T, proving that $T \subseteq G$. The final case is if there is a pair $v_i \in A$, and $v_j \in G - A$ such that v_i and v_j are not adjacent. In this case, we proceed to delete all of the vertices of T that belong to B and denote the remaining tree of order $n - p$ as T_3. There are at most $(p-1)(n-p+1) - 1$ edges between A and $G - A$. Consider the set $G' = G - A - \{u\}$. Thus, we have

$$\begin{aligned} 2|E(G - A - \{u\})| &\geq (n-1)^2 - 2[(p-1)(n-p+1) - 1] \\ &\quad - [(p-1)(n-2) - (p-1)(n-p+1) + 1] - 2n + 2p \\ &= (n-p)(n-p-2) + 1 \end{aligned}$$

Once more, by the inductive hypothesis we have that $T_3 \subseteq G - A - \{u\}$, which can be extended to $T \subseteq G$, concluding the proof. □

The second theorem on this direction was shown by P. Slater, S. Teo and H. Yap in 1985 [40], in the context of graph packing. Given graphs G and H, if $H \subseteq G^c$, which is the complement of G, then we say that G and H are *packable*. We can rewrite the Erdős–Sós Conjecture as a graph packing problem as follows.

Conjecture 3.18. *For positive integers n and k, let G be a graph on n vertices and T any tree on k vertices, if $|E(G)| < \frac{1}{2}n(n-k+1)$, then T and G are packable.*

P. Slater, S. Teo and H. Yap proved the following general theorem on graph packing.

Theorem 3.19 (Slater–Teo–Yap, 1985)**.** *For $n, k \in \mathbb{N}$, let G a graph on n vertices and $n-1$ edges, and T be a tree on $k \geq 5$ vertices. If $n = k+1$, and neither T nor G is a star, then T and G are packable.*

Note that if $k = n$ or $k = n - 1$ in Conjecture 3.18, then $|E(G)| \leq n - 1$, which, by Theorem 3.19, implies that G and T pack, assuming that they are not stars. This means that Theorem 3.19 also gives a proof of Conjecture 3.16 in the case that $n = k$. Unfortunately, its proof technique fails for the cases $k = n - c$, for $c \geq 2$. This means that Conjecture 3.16 holds for graphs G with $|V(G)| = k + 1$ when G nor T are stars. If T is a star, then T is contained in G, because the hypothesis of $\overline{\mathrm{d}}(G) > k - 2$ implies the existence of a vertex with degree at least $k - 1$, which would be set as the pre-image of the center of the star T.

The case where G is a tree on n vertices was proven in 1981, by P. Slater et al. in [40]. For this reason, one may start the proof supposing that G is not a tree. We present a sketch of the proof of Theorem 3.19

Sketch of the proof: Suppose that G is a graph on n vertex that is not a tree. The proof follows by induction on n. The case where $n \in \{5, 6\}$ can be done by verifying that G^c for G with n vertices and $n - 1$ edges contains every tree T that is not a star. The main idea is similar to the proof of Theorem 3.17 but instead of finding an isomorphism of T' in G', one must find an isomorphism of T' in $(G^c)'$, for special trees T' and graphs G' that are considered in different cases.

In the first case, consider the case where T is an extended star S'_n, which is a tree that is obtained from a star S_{n-1} by subdividing one of its edges once, i.e., by replacing an edge by a path with two edges. We denote the center of the star by c and the leaf of the subdivided edge by y. Since G is not a tree, either G has an isolated vertex, say v_p, or G has at least two vertices of degree 1. Assuming G has an isolated vertex v_p, since $n \geq 7$, $G' = G - \{v_p\}$ has two vertices that are not adjacent, say v_1 and v_2. Thus, one may find an isomorphism φ of T into G^c, since $\mathrm{d}_{G^c}(v_p) = n - 1$, by setting $\varphi(c) = v_p$, $\varphi(x) = v_1$, and $\varphi(y) = v_2$. Otherwise, if G has at least two vertices of degree 1, then let u be one such vertex and v its unique neighbor. Thus, one find an isomorphism φ of T into G^c by setting $\varphi(c) = u$, $\varphi(x) = t$ and $\varphi(y) = v$. Henceforth, assume that $T \neq S'_n$, which means that T has two leaves t_1 and t_2 such that $\mathrm{d}(t_1, t_2) \geq 3$ and $T - \{t_1, t_2\}$ is not a star.

In a similar way, proceed to the case that G has an isolated vertex or if G has two leaves u and v such that $dist(u, v) \geq 3$, by finding an isomorphism φ' of $T' = T - \{t_1, t_2\}$ into $(G^c)' \subseteq G^c$ and then extending to an isomorphism φ of T into G^c by defining $\varphi(x) \in (G^c)'$, for $x \in T'$ and by finding low degree vertices in G which are good candidates for $\varphi(t_1)$ and $\varphi(t_2)$.

If G does not have this structure, since G is not a tree, it must have at least two components, and at least one of which is acyclic, because $|E(G)| = k - 1$. This means that G contains exactly one acyclic component and every other component is a cycle. Furthermore, this acyclic component must be a star S_n, where $n \geq 2$. The next case to prove is the case where one of the components of G is a cycle C_m

with $m \geq 4$ and follows the same argument from the previous case.

All that remains is the case that $G = rC_3 \cup S_j$ for $j \geq 2$, i.e., G consists of r disjoint copies of C_3 and a star S_t First consider $r \geq 2$, which for $n \geq 8$ means that T has at least three vertices, say t_1, t_2, and t_3, with degree 1 such that $T' = T - \{t_1, t_2, t_3\}$ is not a star S_{n-3}. By induction, there is an isomorphism φ' of T' into $(G^c)' = (G^c) - v_1, v_2, v_3$, where v_1, v_2, v_3 are three vertices of G that form a cycle, and then proceed to extend φ' to an isomorphism φ of T into G^c. All that remains is the case that $G = C_3 \cup S_k$ for $k \geq 4$, which is the final case. One can find three independent vertices t_1, t_2, t_3 in T for which t_1 is adjacent to a vertex with degree 1, then finding an isomorphism φ' of $T' = T - \{t_1, t_2, t_3\}$ into $(G^c)' = G^c - \{v_1, v_2, v_3\}$, where v_1, v_2, and v_3 are the vertices that form a cycle in G, and then extending φ' to an isomorphism φ of T into G^c, concluding the demonstration. $\square$

The third theorem on this direction was shown by M. Woźniak in 1996 [49], which still uses the packing reformulation of the Erdős–Sós conjecture.

Theorem 3.20 (Woźniak, 1996). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $2|E(G)| < 3n$ then T and G are packable.*

Notice that this represents the case that $k = n - 2$ on Conjecture 3.18. This proof follows the main idea as the proof of Theorem 3.19. First the author finds an isomorphism φ' of a tree $T' \subseteq T$ into $G' \subseteq G$ and then extends such isomorphism. The main difference is the number of cases that are considered, since it becomes harder to find an isomorphism of T into G^c when T is a tree on fewer vertices. A sketch of the proof goes as follows.

Sketch of the proof: The proof is by induction on n. The author divides the proof in three main cases in terms of the minimum degree of G. The author considers a vertex $y \in V(G)$ such that $\mathrm{d}(y) = \Delta(G)$ and also considers the longest path P of T, denoting by $a_1 a_2 \cdots a_r$. The first case is if G has an isolated vertex x. By induction, there is an isomorphism φ' of $T' = T - \{a_2, v\}$ into $(G^c)' = G^c - \{x, y\}$, where $v \in V(T)$ is a leaf. Then one may extend φ' to an isomorphism φ of T into G^c by setting $\varphi(a_2) = x$ and $\varphi(v) = y$.

The second case is when G has a leaf v, which is further divided depending whether n is even or odd. If n is even, the isomorphism can be constructed analogously to the first case. But in the case that n is odd, it considers u, which the vertex that is adjacent to v, and the case is divided into the "subsubcases" $\mathrm{d}(u) \geq 3$, or $\mathrm{d}(u) = 1$ or $\mathrm{d}(u) = 2$. In the first two, the isomorphism is constructed in a direct and similar manner to the first case, but in the case that $\mathrm{d}(y) = 2$, one must proceed more carefully, leading to more cases, which considers the vertex z which is the vertex of G adjacent to u. If $\mathrm{d}(z) \geq 4$, $\mathrm{d}(z) = 1$, $\mathrm{d}(z) = 2$, the packing is constructed similarly to the first case. But in the case that $\mathrm{d}(z) = 3$, the case is further explored, by further dividing the problem into smaller cases. In the end, one divide the case into two cases that have a packing that can be constructed similarly to the first case.

The third and final case is when $\mathrm{d}(v) \geq 2$, for all $v \in V(G)$, repeating the same strategy in the last case by dividing the problem in smaller subcases in which a packing can be constructed and extended easily, by considering different trees T' and host graphs G'. $\square$

We continue with the next result on this direction which is due to G. Tiner in 2010 [43].

Theorem 3.21 (Tiner, 2010)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and $n = k+3$, then $T \subseteq G$.*

Unfortunately, we are not able to access this publication as of today, but we believe that the methods employed follow a similar strategy of the other results. The fifth theorem on this direction, shown by L. Yuan and X. Zhang in 2015 [50] is the following.

Theorem 3.22 (Yuan–Zhang, 2015)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and $n = k+4$, then $T \subseteq G$.*

Theorem 3.22 is the last result for trees on $k = n - c$ vertices that uses the same proof strategy as Theorem 3.17. The main reason for this is the increase on the complexity of the proof compared to the other cases, showing the limitations of this strategy. Instead of a couple special cases to be considered for the isomorphism φ' of a tree $T' = T - X$ into a graph $G' = G - Y$, for set of special vertices $X \subseteq V(T)$ and $Y \subseteq V(G)$, consider forty six different subcases which require different sets X and Y. The authors return to the average degree formulation of Conjecture 3.16, instead of the graph packing formulation. Here, we present a sketch of the proof of Theorem 3.22.

Sketch of the proof: The proof is by induction on n. For $n \leq 5$, it is easily verifiable that the conjecture holds. Therefore, assume that the statement holds for all of the graphs with less than n vertices. Let T be a tree on $k = n-4$ vertices with a longest path $P = a_0 a_1 a_2 \cdots a_r$, and let $N_T(a_1) \setminus \{a_2\} = \{b_0, b_1, \ldots, b_s\}$.

Since $\overline{\mathrm{d}}(G) > k-2$, the only cases that need to be considered are the cases $\Delta(G) = k+3, k+2, k+1, k$, and $k-1$. These are the five main cases of the proof and are analogous to the cases of $\delta(G) = 0, 1, 2, 3, 4$ when considering the graph packing formulation, which would mean that $\Delta(G^c) = n-1, n-2, n-3, n-4$ and $n-5$.

In the case $\Delta(G) = k+3$, let $u \in V(G)$ be a vertex such that $\mathrm{d}(u) = k+3$. We consider the graph graph $G' = G - \{u\}$ and the tree $T' = T - \{a_1, b_1, \ldots, b_s\}$. This implies that

$$\overline{\mathrm{d}}(G') > \frac{k^2 - 2k - 2}{k+2} > k - 4.$$

By the induction hypothesis, there is an isomorphism φ' of T' into G'. Then, one may extend this embedding by setting $\varphi(a_1) = u$ and $\varphi(x) = \varphi'(x)$ for all $x \in V(T)$ such that $x \notin a_1, b_1, \dots, b_s$. Since u is universal, u is adjacent to at least s vertices in $V(G) \setminus \varphi'(V(T'))$, which means that φ is an isomorphism of T into G.

The proof follows considering the cases for which $\Delta(G) = k+2, k+1, k$, and $k-1$, in a similar, yet more complicated, way as in first case. □

The main takeaway of this proof is that every time that it seems hard to find an isomorphism of T' into G', one tries to divide the case into an easy case and a case slightly less complicated than the general case. This strategy finally converges to a case that when divided, the construction of the isomorphism of T into G is analogous to one of the other, already solved, cases.

Lastly, a more general result was proven in this direction by A. Görlich and A. Żak in 2016 [26].

Theorem 3.23 (Görlich–Żak, 2016)**.** *Let c be a positive integer and let $k_0(c) = \gamma c^{12} log^4(c)$, where γ is a sufficiently large constant. Then, for every $t \in [c]$ and for every integer $k \geq k_0(c)$ the following holds. Let G be a graph on n vertices, and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$ and $n = k+t$, then $T \subseteq G$.*

The proof strategy employed by A. Görlich and A. Żak differ entirely from the previous proofs. This time, the proof relies in a probabilistic argument by refining the approach presented by N. Alon and R. Yuster in [5] for another problem. For important definitions and results in Probability Theory, we refer the reader to S. Ross in [37] and M. DeGroot and M. Schervish in [18].

Let X be a random variable that follows a binomial distribution with $n \in \mathbb{N}$ trials and success probability $p \in [0,1]$, denoted by $X \sim \mathrm{Bin}(n,p)$. We denote by $\mathbb{E}[X]$ the expected value of X. A well known result from probability theory states that if $X \sim \mathrm{Bin}(n,p)$, then $\mathbb{E}[X] = np$. The author uses the following version of the Chernoff bounds for the probabilities of two events associated with the random variable X.

If $\mu \geq \mathbb{E}[X] = np$, then

$$Pr\left[X \geq 2\mu\right] \leq e^{-\mu/3}. \tag{3.2}$$

On the other hand, if $\mu \leq \mathbb{E}[X] = np$, then

$$Pr\left[X \leq \mu/2\right] \leq e^{-\mu/8}. \tag{3.3}$$

Given a graph $G = (V, E)$, for a subset of vertices $W \subseteq V$, we define the neighborhood of W, denoted by $N_G(W)$, as $N_G(W) = \bigcup_{w \in W} N(w) \setminus W$.

The authors also provide the following lemma, which is an adaptation of a result from [5] and the main tool in their proof.

Lemma 3.24. *Let G be a graph with n vertices and at most m edges. Let $V(G) = \{v_1, \dots, v_n\}$ with $\mathrm{d}(v_1) \geq \mathrm{d}(v_2) \geq \cdots \geq \mathrm{d}(v_n)$. Let A_i, for $i \in [n]$, be any subset of $V(G)$ with the additional requirement that if $u \in A_i$, then $\mathrm{d}(u) < a$, for a positive integer a. Consider, for $i \in [n]$, the random subset B_i of A_i, where each vertex of A_i is independently selected to B_i with probability $p < 1/a$. Consider the sets*

$$C_i = \left(\bigcup_{j=1}^{i-1} B_j \right) \cap N(v_i),$$

$$D_i = B_i \setminus \left(\bigcup_{j=1}^{i-1} (N_G(B_j) \cup B_j) \right).$$

Then we have

$$Pr\left[|C_i| \geq 4mp\right] \leq e^{-2mp/3}, \text{ for } i \in [n], \tag{3.4}$$

$$Pr\left[|D_i| \leq \frac{2|A_i|}{2e}\right] \leq e^{-p|A_i|/8e}, \text{ for } i \in [\lfloor (ap)^{-1} \rfloor]. \tag{3.5}$$

The main technique of the proof of Theorem 3.23 is to choose A_i and p carefully so that the bounds given by equations 3.4 and 3.5 are both greater than $1/2$, which implies in both events occurring simultaneously with a positive probability.

We now give a proof sketch for Theorem 3.23.

Sketch of the proof: The proof is by induction on t. By Theorem 3.17, the result holds for $t = 0$. Thus, fix a $t \in [c]$ and assume that the statement holds for $t - 1$.

Consider the set $S_i \subseteq V(G) \setminus N(v_i)$, for $i \in [n]$, with the assumption that if $u \in S_i$, then $\mathrm{d}(u) < 5c$ and that $\Delta(T) < 60cn^{3/4}$. By construction, we have that $|S_i| \geq \frac{n}{4} + t$. Moreover, for $i \in [n]$, pick a random subset B_i of S_i where each vertex of S_i is independently selected to B_i with probability

$$p = \frac{n^{-3/4}}{1.5 \cdot 10^2 c}.$$

This implies, by Lemma 3.24, that the events where $|C_i| \leq \frac{n^{1/4}}{750c^2}$, for $i \in [n]$, and $|D_i| \geq 3$, for $i \in \left[\lfloor (300c^2 n^{3/4}) \rfloor\right]$ are bounded by Equations (3.4) and (3.5), respectively. Moreover, by the union bound, one can prove that both events happen with a positive probability. The next step of the proof is to construct an isomorphism of T into G in three steps. At each point of the construction, some vertices of T are matched to some vertices of G, while the other vertices remain unmatched. Initially, all vertices are unmatched. The main idea of step 1 is to match certain vertices of G that have the largest degrees since they are the easier to match. The main idea of step 2 is to find an independent set J on the set of the vertices that are not matched in T and expand the matching to with the vertices in $T \setminus J$ to the vertices in G. Lastly, in step 3, one may proceed to extend the matching to the vertices in J, finding an isomorphism of T into G.

The rest of the proof is a variation of the same argument for the case that $\Delta(T) \geq 60cn^{3/4}$, but starting with a different set of vertices. The idea is still to match T with a subset of G, but since T has vertices with high degrees, this becomes slightly harder and needs more preparation. Therefore, there are two preparatory steps, in which one start matching the maximum degree vertices of T and G then proceed to their neighbors. Then, in the second preparatory step, try to match as much as possible of the vertices in G with the lowest degrees and the vertices in T. Then, complete the matching with a step analogous to the previous case in steps $1, 2$, and 3. $\square$

This was the last result found in the direction of weakening the Conjecture 3.16 by considering only large trees and brings some new light to this case of the problem, motivating the search for better constants or analogous results for other cases of Conjecture 3.16.

3.1.2 Special Trees

In this section, we present some of the theorems that were proved by introducing an hypothesis to Conjecture 3.16 that specifies either some properties of T, or a special tree class in which all trees in a class share a special property.

According to W. Moser and J. Pach in [35], M. Perles showed that Conjecture 3.16 holds for all trees T that are caterpillars but, unfortunately, we are unable to find his proof.

In 1989, A. Sidorenko in [39] proved the following version of Conjecture 3.16.

Theorem 3.25 (Sidorenko, 1989)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and T has a vertex that is adjacent to at least $k/2 - 1$ leaves, then $T \subseteq G$.*

The proof of Theorem 3.25 follows the same idea as the proof of Theorem 3.17. In which one finds an isomorphism of a tree $T' = T - X$ into $G' = G - Y$, and then the idea is to extend it to T by setting the image of the remaining vertices of T in the remaining vertices of G.

Following this direction, the next result obtained was proved by M. Woźniak in 1996 [49].

Theorem 3.26 (Woźniak, 1996). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and T is a spider in which in each leg has at most two edges, then $T \subseteq G$.*

It is worth mentioning that the author proved this theorem in the context of graph packing, seen in Subsection 3.1.1, by constructing a packing of T and G^c. Its is very similar to the proof of Theorem 3.25 and, hence, we omit from this dissertation.

Later, in 2004, A. McLennan [34] proved the following result. Recall that, for a given tree T, the diameter of T is denoted by $D(T)$, and is the greatest distance between any pair of vertices $u, v \in V(T)$.

Theorem 3.27 (McLennan, 2004). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and $D(T) \leq 4$ then $T \subseteq G$.*

The proof of Theorem 3.27 relies in proving the following lemma:

Lemma 3.28. *Let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and $D(T) \leq 4$, then there exist $u \in V(G)$ and distinct $v_1, \dots, v_p \in N(u)$ such that:*

$$|N(v_i) \setminus \{u, v_1, \dots, v_p\}| \geq k - 2p + i - 1, \text{ for } i = 1, \dots, p.$$

Note that the hypothesis $\overline{\mathrm{d}}(G) > k - 2$, implies the existence of a vertex with $k - 1$ neighbors, but Lemma 3.28 requires a vertex whose neighbors have many neighbors. To prove such lemma, A. McLennan prove the following proposition.

Proposition 3.29. *Suppose that $1 \leq q \leq p$, $u \in V(G)$ and $A \subset N(u)$ is a set such that $\mathrm{d}(v) \geq k - p$ for all $v \in A$ and*

$$\sum_{v \in A} \left(1 - \frac{k - p - 1}{\mathrm{d}(v)}\right) > p - q.$$

Then there are distinct $v_q, \dots, v_p \in A$ such that

$$|N(v_i) \setminus \{u, v_q,\ dots, v_p\}| \geq k - 2p + 1 - 1, \text{ for } i \in \{q, \dots, p\}.$$

Note that the case where $q = 1$ in Proposition 3.29 is equivalent to Lemma 3.28. We now present a proof sketch for Theorem 3.27.

Sketch of the proof: Fix a tree T of diameter at most 4. If the diameter of T is either 2 or 3 then Theorem 3.25 implies that there is an isomorphism of T into G. Therefore, you may assume that T has diameter exactly 4. Let a_0 and a_4 be the vertices in T whose distance is 4. Let a_1, a_2, a_3 be the vertices on the path between

a_0 and a_4. Then, the distance between a_2 and any other vertex is not greater than 2. The tree T can now be characterized by the number of neighbors of a_2, which is denoted by p. Also, denote by α_i the number of neighbors of v_i, other than a_2, for which v_i is adjacent to a_2, for $i \in [p]$. Then $\alpha_1 + \cdots + \alpha_p = k - p - 1$.

For $i = 1, \dots, p$, let $\beta_i := \alpha_1 + \cdots + \alpha_i$. If there is a vertex $u \in V(G)$ and distinct vertices $v_1, \dots, v_p \in N(u)$ such that

$$|N(v_1) \setminus \{u, v_1, \dots, v_p\}| \geq B_1, \dots, |N(v_p) \setminus \{u, v_1, \dots, v_p\}| \geq \beta_p,$$

then we can find an isomorphism of T into G by choosing $S_1 \subseteq N(v_1) \setminus \{u, v_1, \dots, v_p\}$, with α_1 elements, $S_2 \subseteq N(v_2) \setminus (\{u, v_1, \dots, v_p\} \cup S_1)$, with α_2 elements and so on. It turns out that one is able to find a satisfactory tuple of vertices $(u, v_1, \dots, v_p)$ when $\beta_1, \dots, \beta_p$ satisfy the following condition:

$$\beta_i \leq \max\{0, (k - 2p + i - 1)\}.$$

An application of Lemma 3.28 concludes the proof. □

In the same work, McLennan gives an alternative proof of Theorem 3.25, and shows that this argument alone does not work for trees with diameter at most 5.

The next result was proved, in 2007, by G. Fan and L. Sun in [23], and strengthens Woźniak's results on Spiders with diameter at most 4.

Theorem 3.30 (Fan–Sun, 2007)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and T is either a spider with three legs or a spider whose legs have length at most 4, then $T \subseteq G$.*

The proof of the case that T is a spider with 3 legs relies in the following observation.

Observation 3.31. *Let G be a graph on n vertices, and let T be a spider with k vertices and three legs of lengths ℓ_1, ℓ_2, ℓ_3. Without loss of generality, assume that $1 \leq \ell_1 \leq \ell_2 \leq \ell_3$. If there is a path $P = x_0 x_1 \cdots x_p$ with $p \geq l_1$ such that $G - \{x_1, x_2, \dots, x_p\}$ has a cycle D containing x_0 with $|E(D)| \geq \ell_2 + \ell_3 + 1$, then G has a copy of T having x_0 as its center.*

In 2016, G. Fan and Z. Huo used the same strategy to prove an extension of Theorem 3.30 in [22].

Theorem 3.32 (Fan–Huo, 2016)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and T is a spider of four legs, then $T \subseteq G$.*

The proof strategy is the same as in the Theorem 3.30 but needs to consider different subcases analogously to the extensions of the proofs in Section 3.1.1. Finally, in 2019, G. Fan, Y. Hong, and Q. Liu showed that Conjecture 3.16 holds for all spiders in [24].

Theorem 3.33 (Fan–Hong–Liu, 2019)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and T is a spider, then $T \subseteq G$.*

The strategy of the proof is to start with an isomorphism φ' of a spider $T' \subseteq T$ into $G' \subseteq G$ and extend it to an isomorphism φ of the desired spider T into G. The authors proved five different extension lemmas that shows how one could construct an extension φ to the isomorphism φ'. Similarly to the proof of Theorem 3.22, the proof is divided in several different subcases and is very extensive. While Theorem 3.33 shows that Conjecture 3.16 holds for a large family of trees, there still a large number of trees remaining.

3.1.3 Special Graphs

In this section, we present some theorems that were proven by including an hypothesis to the Erdős–Sós conjecture that requires an extra property on the graph G, besides $\overline{\mathrm{d}}(G) > k - 2$.

The first theorem on this direction was proven by S. Brandt and E. Dobson in 1996 in [13], combining the average degree hypothesis with a condition on the girth of the graph. Recall that the girth of a graph G is the length of the shortest cycle contained in G, and is denoted by $girth(G)$.

Theorem 3.34 (Brandt–Dobson, 1996)**.** *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and $girth(G) \geq 5$, then $T \subseteq G$.*

The authors achieve this result as a consequence of the following lemma, that gives sufficient degree conditions for G to contain a graph T when G has girth at least 5 and is a variation of Lemma 3.14.

Lemma 3.35. *Let G be a graph on n vertices with girth at least 5 and T be a tree with k vertices. If $\delta(G) \geq k/2$ and $\Delta(G) \geq \Delta(T)$, then $T \subseteq G$.*

Given a proof of Lemma 3.35, the proof of Theorem 3.34 goes as follows.

Proof. Take a subgraph $H \subseteq G$, such that H has the minimum number of vertices which satisfies $|E(H)| > |H|\frac{(k-2)}{2}$. Clearly, $\Delta(H) \geq k$ and, since $|E(H - v)| \leq (|H| - 1)\frac{(k-2)}{2}$ for every vertex v, we have that $\delta(H) \geq k/2$. So H satisfies the requirements of Lemma 3.35, and, therefore, G contains every tree T on k vertices. □

Later, in 1997, J. Saclé and M. Woźniak [38] proved a version of Conjecture 3.16, for graphs G without C_4, which is a direct extension of Theorem 3.34. by including graphs that contains C_3.

Theorem 3.36 (Saclé–Woźniak, 1997). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and G is a C_4-free graph, then $T \subseteq G$.*

The proof of Theorem 3.36 follows the same strategy as the proofs for Theorems 3.17. The authors divide the problem in several cases and, for each case, constructs an isomorphism φ' of $T' \subseteq T$ into $G' \subseteq G$ and then extends it to the whole tree T into G.

In 2000, M. Wang, G. Li, and A. Liu [46] showed the following theorem, which unfortunately, we were not able to access its proof.

Theorem 3.37 (Wang–Li–Liu, 2000). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and G^c has girth greater than 4, then $T \subseteq G$.*

In 2009, N. Eaton and G. Tiner proved the following case for the Conjecture 3.16, extending it to trees with large minimum degrees.

Theorem 3.38 (Eaton–Tiner, 2009). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, $k \geq 8$, and $\delta(G) \geq k-4$ then $T \subseteq G$.*

Alongside Theorem 3.38, N. Eaton and G. Tiner improved Theorem 3.25 by proving the following theorem.

Theorem 3.39 (Eaton–Tiner, 2009). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and T has a vertex that is adjacent to at least $\lceil \frac{k}{2} - 2 \rceil$ leaves, then $T \subseteq G$.*

Alongside this result, they proved the following generalization:

Theorem 3.40 (Eaton–Tiner, 2009). *For $n, k, d \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, $\delta(G) \geq d$ and T has a vertex that is adjacent to at least $(k-1) - d$ leaves, then $T \subseteq G$.*

The main idea of the proof is similar as in the other constructive proofs but instead of considering each case in a separate manner, N. Eaton and G. Tiner provide some important generalizations for when an isomorphism φ of $T' \subseteq T$ into $G' \subseteq G$ can be extended to an isomorphism φ of T into G.

Lastly, N. Eaton and G. Tiner, in 2013, proved one last result in this direction.

Theorem 3.41 (Eaton–Tiner, 2013). *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k-2$, and G is P_{k+5} free, then $T \subseteq G$.*

3.1.4 Approximated Results

In this section, we present some of the most recent theorems. These theorems are based on an approximation of the conjecture by including hypotheses that depends on other parameters.

In 2020, G. Besomi, M. Pavez-Signé and M. Stein proved the following special case of Conjecture 3.16 in [9].

Theorem 3.42 (Besomi–Pavez-Signé–Stein, 2020)**.** *For all $\delta > 0$ amd $\Delta \in \mathbb{N}$, there is $n_0 \in \mathbb{N}$ such that for each $k, n \in \mathbb{N}$ with $n \geq n_0$ and $\delta n \leq k \leq n$, and for each graph G on n vertices and tree T on k vertices the following holds. If $\overline{\mathrm{d}}(G) > k-2$ and $\Delta(T) \leq \Delta$, then $T \subseteq G$.*

To prove this theorem, G. Besomi, M. Pavez-Signé and M. Stein also prove the following auxiliary lemma that helps to find an isomorphism for trees T in graphs G with a specific structure.

Lemma 3.43. *For each $k, \Delta \in \mathbb{N}$ and each graph G on n vertices with $\overline{\mathrm{d}}(G) < k-2$ and $\delta(G) \geq \frac{k}{2}$ the following holds.*

(a) If $k \geq 10^6$, and $n \leq (1 + 10^{-11})k$, then G contains each tree T with k vertices and $\Delta(T) \leq \frac{\sqrt{(k)}}{1000}$;

(b) If $k \geq 8\Delta^2$, G is bipartite, and there is a partition $V(G) = A \cup B$ such that $|E(A)| + |E(B)| \leq \beta |E(G)|$ for $\beta = \frac{1}{50\Delta^2}$, with $|A|, |B| \leq \left(1 + \frac{1}{25\Delta^2}\right) k$, then G contains each tree T with k vertices and $\Delta(T) \leq \Delta$.

Here we give a short sketch of the proof of Theorem 3.42.

Sketch of the proof: The proof of this theorem is divided into two main cases. The first case considered is when G is connected and n is considerably larger than k. In this case, one may employ the Szemerédi's Regularity Lemma [41] to find an almost spanning subgraph $H \subseteq G$ that admits a *regular partition.* If this component is large enough, then one may show that either H is bipartite or it contains a useful matching structure that can be used to find the isomorphism of T into H, and, therefore, of T into G. Otherwise, H is a union of graphs that are almost complete and of size close to k or almost complete bipartite graphs of size close to $2k$. In this case, one can use an extra edge of G to connect two components and find an embedding of T into G.

If, on the other hand, n is very close to k, if G is close to being bipartite of order $2k$, or if G is the disjoint union of such graphs then the result follows from Lemma 3.43. $\square$

It is worth mentioning that the same authors obtained several other results on degree conditions for embedding trees [8], which is a very closely related problem to Conjecture 3.16.

Lastly, V. Rozhon, in 2020, independently proved the following result.

Theorem 3.44 (Rozhon, 2020). *For any $\eta > 0$, there is $n_0 \in \mathbb{N}$ and $\gamma > 0$ such that for every $n > n_0$ and $k > 0$, any graph G on n vertices with $\overline{\mathrm{d}}(G) \geq k + \eta n$ contains every tree T on k vertices with $\Delta(T) \geq \gamma k$.*

The proof of Theorem 3.44 also uses the Szemerédi's regularity lemma [41] to find a suitable subraph $H \subseteq G$ that serves as a host for the copy of T in G.

3.2 Related Problems

In this section, we exhibit some problems from Extremal Graph Theory that are closely related to the Conjecture 3.16. The first problem is a simple variation, formulated by M. Loebl, J. Komlós and V. Sós [19] in which one replaces the average degree $\overline{\mathrm{d}}(G)$ with the *median degree* of G, denoted by $d^*(G)$, which is defined as follows. Consider the ordered degree sequence of G, denoted by $\{d_i\}_{1\leq i\leq n}$. If $n \in \mathbb{N}$ is odd, then $d^*(G) = d_{(n+1)/2}$. Else, if n is even, then $d^*(G) = \left(\frac{1}{2}\right)(d_{n/2} + d_{n/2+1})$.

The Loebl–Komlós–Sós Conjecture is stated as follows.

Conjecture 3.45 (Loebl–Komlós–Sós). *Let G be a graph on n vertices and T be a tree on k vertices. If $d^*(G) > k - 2$, then $T \subseteq G$.*

Let σ be the variance of the degree sequence of a graph G. A well known relation between $\overline{\mathrm{d}}(G)$ and $d^*(G)$ is the following.

$$|\overline{\mathrm{d}}(G) - d^*(G)| \leq \sigma.$$

This implies that Conjecture 3.45 is indeed closely related to Conjecture 3.16 but it is not obvious when one implies the other. It is worth mentioning that many particular cases that are solved for Conjecture 3.45 follows the same directions of the particular cases of Conjecture 3.16.

The case when G is a graph on $n = k + c$ vertices, for $c = 0, 1, 2, 3$, was proved by C. Bazgan, H. Li, and M. Woźniak [7] in 2000. The case for which T is a tree of diameter 5 or a certain caterpillar, was proven by D. Piguet and M. Stein, in 2007, in [36]. Lastly, the approximated version of Conjecture 3.45 was proven by J. Hladký, J. Komlós, D. Piguet, M. Simonovits, M. Stein, and E. Szemerédi, in 2015, in a series of four papers [29, 30, 31, 32], in which the main result is the following.

Theorem 3.46 (Hladký–Komlós–Piguet–Simonovits–Stein–Szemerédi, 2015)**.** *For every $\alpha > 0$ there exists k_0 such that for any $k > k_0$ we have the following. Each graph G on n vertices with at least $\left(\frac{1}{2} + \alpha\right) n$ vertices of degree at least $(1 + \alpha)k$ contains each tree T of order k.*

We believe that the techniques and strategies used to prove certain special cases of Conjecture 3.16 can be used to prove other cases in Conjecture 3.45 due to their closeness and vice versa.

Another closely related problem is the problem concerning *panarboreal* graphs, which are graphs that contains all possible trees on n vertices. Let $s(n)$ denote the minimum number of edges that a graph G on n vertices can have such that any tree on n vertices is isomorphic to a subgraph $H \subseteq G$.

In 1983, F. Chung and R. Graham proved the following bounds for $s(n)$ in [16].

Theorem 3.47 (Chung–Graham, 1983)**.** *If $n \in \mathbb{N}$, then*

$$\frac{1}{2} n \log(n) \leq s(n) \leq \frac{5}{\log(4)} n \log(n) + O(n).$$

It is worth mentioning that F. Chung and R. Graham point out that the constant $\frac{5}{\log(4)}$ could be improved. Unfortunately, since 1983, no improvements has been done.

3.3 Conclusion

In this chapter, we presented a collection of results in Extremal Graph Theory, covering from Mantel's Theorem up to the Erdős–Sós Conjecture, to introduce some canonical results and techniques employed in this field. We also presented a plethora of results on Conjecture 3.16, divided in four different directions, pointing the main strategy and techniques of their proofs.

We finish this chapter by presenting a path for the proof of the case $n = k + 5$, following the strategy of the similar results on Section 3.1.1. The proposed statement is the following.

Statement 3.48. *For $n, k \in \mathbb{N}$, let G be a graph on n vertices and T be a tree on k vertices. If $\overline{\mathrm{d}}(G) > k - 2$, and $n = k + 5$, then $T \subseteq G$.*

Sketch of the proof: Following the idea for the proof of Theorem 3.22, the proof follows by induction on n. Since $\overline{\mathrm{d}}(G) > k - 2$, we only need to consider the cases $\Delta(G) = k + 4, k + 3, k + 2, k + 1, k,$ and $k - 1$. Let $u \in V(G)$ be the vertex such that $\mathrm{d}(u) = \Delta(G)$, and let $z \in V(G)$ be the vertex such that $\mathrm{d}(z) = \delta(G)$. By Theorem 3.38, we may assume that $\mathrm{d}(z) \leq k - 5$. Let T be a tree on k vertices with a longest path $P = a_0 a_1 \ldots a_{r-1} a_r$; and $N_T(a_1) \setminus \{a_2\} = \{b_1, \ldots, b_s\}$.

Suppose that $\Delta(G) = k+4$, which means that u is a universal vertex. Consider the graph $G' = G - \{u, z\}$, and $T' = T - \{a_1, b_1, \dots, b_s\}$, which means that

$$2|E(G')| = 2|E(G)| - 2\mathrm{d}(u) - 2\mathrm{d}(z) + 2 = (k-2)(k+5) - 2(k+4) - \mathrm{d}(z) + 2.$$

The factor $+2$ comes from the fact that u is a universal vertex, therefore we are counting the edge $\{u, z\}$ twice. Since $\mathrm{d}(z) \leq k-5$, we have

$$2|E(G')| \geq (k-2)(k+5) - 2(k+4) - 2(k-5) + 2 = k^2 - k - 6.$$

The assertion $\overline{\mathrm{d}}(G') > k-4$ implies that $2|E(G')| > (k-4)(k+3)$, therefore

$$2|E(G')| = k^2 - k - 6 > (k-4)(k+3) = k^2 - k - 12.$$

Which holds for all values of k. For this reason, by the induction hypothesis, we can find an isomorphism φ' of T' into G'. Then, we can extend φ' by setting $\varphi(x) = \varphi'(x)$, for all $x \in V(T')$ and letting $\varphi(a_1) = u$. Because $\mathrm{d}(u) = k+4 > s$. This implies that $T \subseteq G$.

The second case is if $\mathrm{d}(u) = \Delta(G) = k+3$. This means that there is a vertex $x \in V(G)$ such that u and x are not adjacent. Consider $G' = G - \{u, x\}$ and $T' = T - \{a_1, b_1, \dots, b_s\}$. We have

$$2|E(G')| = 2|E(G)| - 2(k+4) - 2\mathrm{d}(x) = (k-2)(k+5) - 2(k+4) - 2\mathrm{d}(x).$$

The assertion $\overline{\mathrm{d}}(G') > k-4$ requires that $2|E(G')| > (k-4)(k+3)$, therefore

$$2|E(G')| = k^2 + k - 18 - 2\mathrm{d}(u) > (k-4)(k+3) = k^2 - k - 12,$$

which means that $2\mathrm{d}(u) < 2k-6$. Therefore, if $\mathrm{d}(u) \leq k-2$, then we can use the induction hypothesis to find an isomorphism φ' of T' into G' and then proceed to extend it analogously to the previous case. Else, if $\mathrm{d}(u) \geq k$, we consider the graphs $G' = G - \{u, z, x\}$ and $T' = T - \{a_1, b_1, b_s, a_r\}$. We can find an isomorphism φ' of T' into G' but the extension to an isomorphism φ of T into G differs slightly depending whether x and z are adjacent. The missing case is if $\mathrm{d}(x) = k-1$. We first divide the case whether x is adjacent to z. If they are not adjacent, then the case is analogous to previous case. Otherwise, we cannot proceed with our proof.

The other cases proceeds very similarly, by finding some cases that are analogous, but ending in a subcase that requires a new strategy to find an extension of the isomorphism found by the induction hypothesis. □

Observe that this proofs a partial case in the case $n = k+5$ but it requires that the other cases hold to prove that the Erdős-Sós Conjecture holds for $n = k+5$, namely $\Delta(G) = k+3, k+2, k+1, k$ and $k-1$.

4 Conclusion

We were able to develop an alternative and equivalent definition to the function t, concerning the problem of Roller Coaster permutations, studied in Chapter 2, in terms of the number of triangles in a permutation. It is worth mentioning that this can be generalized to a sequence in any well ordered set, meaning that the numbers doesn't need to be integers, for example.

We presented an ILP model, in Model 2.2, and an extended ILP, in Model 2.3, that obtained new Roller Coasters and Roller Coaster candidates, improving known bounds for $\mathrm{t}_{max}(n)$, for $n \leq 40$. We leave for future work the study of Conjecture 2.1, and 2.4 since, if they hold, our new bounds are indeed optimal values of t_{max}.

Another interesting problem posed was the problem concerning the number of triangles in a permutation, which branches to the decomposition of a permutation in triangles and the realizable triangle sequence problem.

Finally, we present the problem in different point of views, depending on the representation of the permutation, since the problem might be reducible to other well studied problems in Combinatorics and Mathematics, like the Matrix Multiplication problem.

For the Erdős-Sós Conjecture, we presented a plethora of results divided in four different directions. Each such direction points to a different way to contribute to the state of the art in this problem. To illustrate this idea, we presented a tentative proof for the case $n = k+5$ as a natural extension of the argument for Theorem 3.17, but it fails at some singular cases. Analogously, one could try to extend the case for special trees for different classes of trees, or the case for special graphs, for different properties on these graphs. Moreover, one could try to contribute to a known result by improving the constants on the known bounds, for example, by improving the constant of Theorem 3.23, by employing a stronger probabilistic argument, or possibly a non-probabilistic argument.

Lastly, we presented two closely related problems that also offer similar proof strategies on their results. Note that one could also consider a version of the conjecture by including a hypothesis on the variance of the degree sequence of the graph, which would bring even closer to Conjecture 3.45.

References

[1] ADAMCZAK, W., 2016, "A Note on the Structure of Roller Coaster Permutations", *Journal of Mathematics Research 9*.

[2] AHMED, T., SNEVILY, H., 2013, "Some properties of roller coaster permutations", *Bulletin of the Institute of Combinatorics and its Applications*, v. 68 (01).

[3] AIGNER, M., ZIEGLER, G. M., 2009, "Proofs from THE BOOK". Springer Publishing Company, Incorporated. ISBN: 3642008550.

[4] ALMAN, J., WILLIAMS, V. V., 2021, "A refined laser method and faster matrix multiplication". In: *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 522–539. SIAM.

[5] ALON, N., YUSTER, R., 2013, "The Turán number of sparse spanning graphs", *Journal of Combinatorial Theory, Series B*, v. 103, n. 3, pp. 337–343.

[6] ANDRÉ, D., 1879, "Développements de sec x et de tang x", *CR Acad. Sci. Paris*, v. 88, pp. 965–967.

[7] BAZGAN, C., LI, H., WOŹNIAK, M., 2000, "On the Loebl–Komlós–Sós conjecture", *Journal of Graph Theory*, v. 34, n. 4, pp. 269–276.

[8] BESOMI, G., PAVEZ-SIGNÉ, M., STEIN, M., 2019, "Degree conditions for embedding trees", *SIAM Journal on Discrete Mathematics*, v. 33, n. 3, pp. 1521–1555.

[9] BESOMI, G., PAVEZ-SIGNÉ, M., STEIN, M., 2020, "On the Erdős-Sós conjecture for trees with bounded degree", *arXiv preprint arXiv:1906.10219.*

[10] BONDY, J. A., SIMONOVITS, M., 1974, "Cycles of even length in graphs", *Journal of Combinatorial Theory, Series B*, v. 16, n. 2, pp. 97–105.

[11] BORCHARDT, C. W., 1861, "Über eine Interpolationsformel für eine Art symmetrischer Functionen und über deren Anwendung." .

[12] BOTLER, F., NETTO, B. L., 2021, “New Bounds on Roller Coaster Permutations”. In: *Anais do VI Encontro de Teoria da Computação*, pp. 50–53. SBC, July.

[13] BRANDT, S., DOBSON, E., 1996, “The Erdős-Sós conjecture for graphs of girth 5”, *Discrete Mathematics*, v. 150, n. 1-3, pp. 411–414.

[14] BUKH, B., JIANG, Z., 2017, “A bound on the number of edges in graphs without an even cycle”, *Combinatorics, Probability and Computing*, v. 26, n. 1, pp. 1–15.

[15] CAUCHY, A.-L., 1815, “Mémoire sur le nombre des valeurs qu’une fonction peut acquérir, lorsqu’on y permute de toutes les manieres possibles les quantités qu’elle renferme”, *Journal de l’École Polytechnique*, v. 17, pp. 1–28.

[16] CHUNG, F. R., GRAHAM, R. L., 1983, “On universal graphs for spanning trees”, *Journal of the London Mathematical Society*, v. 2, n. 2, pp. 203–211.

[17] CONLON, D., 2021, “Extremal numbers of cycles revisited”, *The American Mathematical Monthly*, v. 128, n. 5, pp. 464–466.

[18] DEGROOT, M. H., SCHERVISH, M. J., 2012, “Probability and statistics”. Pearson Education.

[19] ERDŐS, P., FÜREDI, Z., LOEBL, M., et al., 1995, “Discrepancy of trees”, *Studia Scientiarum Mathematicarum Hungarica*, v. 30, n. 1-2, pp. 47–57.

[20] ERDÖS, P., STONE, A. H., 1946, “On the structure of linear graphs”, *Bulletin of the American Mathematical Society*, v. 52, n. 12, pp. 1087–1091.

[21] EULER, L., 1741, “Solutio problematis ad geometriam situs pertinentis”, *Commentarii academiae scientiarum Petropolitanae*, pp. 128–140.

[22] FAN, G., HUO, Z., 2016, “The Erdős–Sós conjecture for spiders of four legs”, *Journal of Combinatorics*, v. 7, n. 2, pp. 271–283.

[23] FAN, G., SUN, L., 2007, “The Erdös–Sós conjecture for spiders”, *Discrete mathematics*, v. 307, n. 23, pp. 3055–3062.

[24] FAN, G., HONG, Y., LIU, Q., 2018, “The Erdős-Sós Conjecture for Spiders”, *arXiv preprint arXiv:1804.06567*.

[25] FLAJOLET, P., SEDGEWICK, R., 2009, “Analytic combinatorics”. Cambridge University Press.

[26] GOERLICH, A., ŻAK, A., 2016, “On Erdős-Sós conjecture for trees of large size”, *The Electronic Journal of Combinatorics*, pp. P1–52.

[27] GUROBI OPTIMIZATION, L., 2021. “Gurobi Optimizer Reference Manual”. Disponível em: <`http://www.gurobi.com`>.

[28] HARARY, F., SCHWENK, A. J., 1973, “The number of caterpillars”, *Discrete Mathematics*, v. 6, n. 4, pp. 359–365.

[29] HLADKỲ, J., KOMLÓS, J., PIGUET, D., et al., 2017, “The approximate Loebl-Komlós-Sós Conjecture I: The sparse decomposition”, *SIAM Journal on Discrete Mathematics*, v. 31, n. 2, pp. 945–982.

[30] HLADKỲ, J., KOMLÓS, J., PIGUET, D., et al., 2017, “The approximate Loebl-Komlós-Sós Conjecture II: The rough structure of LKS graphs”, *SIAM Journal on Discrete Mathematics*, v. 31, n. 2, pp. 983–1016.

[31] HLADKỲ, J., KOMLÓS, J., PIGUET, D., et al., 2017, “The approximate Loebl-Komlós-Sós Conjecture III: The finer structure of LKS graphs”, *SIAM Journal on Discrete Mathematics*, v. 31, n. 2, pp. 1017–1071.

[32] HLADKỲ, J., KOMLÓS, J., PIGUET, D., et al., 2017, “The approximate Loebl-Komlós-Sós Conjecture IV: Embedding techniques and the proof of the main result”, *SIAM Journal on Discrete Mathematics*, v. 31, n. 2, pp. 1072–1148.

[33] KŐVÁRI, P., T SÓS, V., TURÁN, P., 1954, “On a problem of Zarankiewicz”. In: *Colloquium Mathematicum*, v. 3, pp. 50–57. Polska Akademia Nauk.

[34] MCLENNAN, A., 2005, “The Erdős-Sós Conjecture for Trees of Diameter Four”, *J. Graph Theory*, v. 49, n. 4 (aug), pp. 291–301. ISSN: 0364-9024.

[35] MOSER, W., PACH, J., 1993, “Recent developments in combinatorial geometry”. In: *New Trends in Discrete and Computational Geometry*, Springer, pp. 281–302.

[36] PIGUET, D., STEIN, M. J., 2007, “The Loebl-Komlós-Sós conjecture for trees of diameter 5 and for certain caterpillars”, *arXiv preprint arXiv:0712.3382*.

[37] ROSS, S. M., 2014, “A first course in probability”. Pearson.

[38] SACLÉ, J.-F., WOŹNIAK, M., 1997, "The Erdős–Sós Conjecture for Graphs without C4", *Journal of Combinatorial Theory, Series B*, v. 70, n. 2, pp. 367–372.

[39] SIDORENKO, A., 1989, "Asymptotic solution for a new class of forbiddenr-graphs", *Combinatorica*, v. 9, n. 2, pp. 207–215.

[40] SLATER, P., TEO, S., YAP, H., 1985, "Packing a tree with a graph of the same size", *Journal of graph theory*, v. 9, n. 2, pp. 213–216.

[41] SZEMERÉDI, E., 1975, "Regular partitions of graphs". Stanford Univ. Calif. Dept. of Computer Science.

[42] THE SAGE DEVELOPERS, 2020, *SageMath, the Sage Mathematics Software System (Version 9.2)*. `https://www.sagemath.org`.

[43] TINER, G. F., 2007, "On the Erdős-Sós Conjecture", *ETD Collection for University of Rhode Island*, p. AAI3277009.

[44] TURÁN, P., 1941, "Eine extremalaufgabe aus der Graphentheorie", *Mat. Fiz. Lapok*, v. 48, n. 436-452, pp. 61.

[45] VERSTRAËTE, J., 2016, "Extremal problems for cycles in graphs". In: *Recent trends in combinatorics*, Springer, pp. 83–116.

[46] WANG, M., LI, G., LIU, A., 2000, "A result of Erdős-Sós conjecture", *Ars Combinatorica*, v. 55, pp. 123–127.

[47] WENGER, R., 1991, "Extremal graphs with no C^4's, C^6's, or C^{10}'s", *Journal of Combinatorial Theory, Series B*, v. 52, n. 1, pp. 113–116.

[48] WILF, H. S., 2005, "generatingfunctionology". CRC press.

[49] WOŹNIAK, M., 1996, "On the Erdős–Sós conjecture", *Journal of Graph Theory*, v. 21, n. 2, pp. 229–234.

[50] YUAN, L.-T., ZHANG, X.-D., 2014, "On the Erdős-Sós Conjecture for graphs on n= k+ 4 vertices", *Ars Mathematica Contemporanea 13*.

[51] ZHOU, B., 1984, "ANOTE ON THE Erdős-Sós CONJECTURE", *Acta Mathematica Scientia*, v. 4, n. 3, pp. 287–289.