UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

UNAVOIDABLE PATTERNS, BALANCEABILITY AND AMOEBAS

DEFENSA DE TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:
DENAE VENTURA ARREDONDO

DIRECTORA DE LA TESIS

ADRIANA HANSBERG PASTOR
INSTITUTO DE MATEMÁTICAS UNIDAD JURIQUILLA UNAM

MIEMBROS DEL COMITÉ TUTOR

AMANDA MONTEJANO CANTORAL
FACULTAD DE CIENCIAS

CÉSAR HERNÁNDEZ CRUZ
INSTITUTO DE MATEMÁTICAS

CIUDAD DE MÉXICO, 17 DE MAYO 2023.

UNAM
Universidad Nacional Autónoma de México

Biblioteca Central
Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Contents

List of Figures

List of Tables

Introduction

Ramsey's theorem [54] from 1930 states that, no matter how we color the edges of a large enough complete graph, a monochromatic copy of a fixed graph is unavoidable. On the other hand, Turán's theorem from 1941 says that, there is minimum edge density such that any graph that satisfies it contains a copy of a fixed subgraph H. Both theorems set the start for Ramsey theory and extremal theory, correspondingly, in combinatorics.

In the last 20 years, the study of unavoidable patterns in colorings of a base structure has shown an increasing interest among the community in combinatorics. By Ramsey's theorem, an arbitrary 2-edge coloring of a large enough complete graph contains an unavoidable pattern: a monochromatic copy of some previously fixed graph G. If we tweak the problem a bit and consider 2-edge colorings with sufficient edges in each color, what patterns are we bound to find?

Sometime before 2007, Bollobás conjectured that, given an edge proportion, there is a large enough complete graph where any 2-edge coloring with at least this edge proportion in each color guarantees the existence of one of two unavoidable patterns [19]. Given a positive integer t, the first pattern consists of a complete graph on $2t$ vertices where one of the colors makes a complete graph on t vertices. The second pattern is also a complete graph on $2t$ vertices except that now one of the colors makes two disjoint complete graphs on t vertices each. This conjecture set the start of a series of research problems on unavoidable patterns and other natural concepts that derived in the way. The conjecture by Bollobás was confirmed by Cutler and Montágh [19] in 2007 and provided a bound on the size of the base graph. In 2019, Caro, Hansberg and Montejano [13] stated a Turán version of this result. They proved that, if one has a large enough complete graph, then there is a least amount of edges one can ask for in each color so that every 2-edge coloring of the complete graph guarantees one of the patterns defined above. If at least one of these patterns always appears in any 2-edge coloring with sufficient edges in each color, then any colored graph contained in both patterns will also appear. One type of colored graphs we will study in this context are the balanceable graphs. A 2-edge coloring of G contains a balanced copy of a graph H if there is a copy of H with half of its edges in each color (or with exactly the integral part of half of its edges in one color). If there is an integer k such that every 2-edge coloring of K_n with at least k edges in each color contains a balanced copy of H, then we say that H is balanceable.

Notice that if there is a graph H which is a subgraph of both of the above patterns in a way that it is balanced, then this graph is balanceable. Determining balanceability and the least amount of edges (the balancing number) one requires to have in any 2-edge coloring to guarantee the existence of a balanced copy of a fixed graph is one of the problems we approach in this thesis. Several open problems were stated and solved in previous research, such as the balanceability of some graph classes like complete graphs with even edge number, some cycles and all trees. However, many open problems are yet to be solved.

The objective of this dissertation work is to study and contribute to the area of unavoidable patterns in 2-edge colorings. This involves determining the balanceability of some graph classes for which balanceability is not known, and to find the balancing number of graphs known to be balanceable. For graphs that are not balanceable, we define an extension of the balancing number that allows us to measure to some extent the degree of "unbalanceability" of such graphs. We study a family of graphs called amoebas which contains a subfamily of balanceable graphs, called global amoebas. Amoeba graphs are catalogued as local and/or global amoebas. We also work with local amoebas and give a recursive construction of certain families of local amoebas. We also explore unavoidable patterns in 2-edge colorings, balanceability and omnitonality in the complete bipartite graph. This work has contributed two research papers so far, one [22] of which has been published in the journal Discrete Applied Mathematics and the other [20] has been accepted in the Electronic Journal of Combinatorics. There is also some work still in progress [21, 29, 46] and some open problems which will be mentioned in Section 5.1. We proceed to give an outline of the contents of this thesis work.

In Chapter 1, we present the preliminaries. We begin in Section 1.1 by stating basic definitions from graph theory that are used throughout this work, making special emphasis on edge coloring concepts which concern most of the open problems we solve. In Section 1.2, we give a brief state-of-the-art on Ramsey theory and extremal theory, paying close attention to the problem of forbidding a bipartite graph which has great relevance in Chapter 4. In Section 1.3, we discuss thoroughly the origin of the main problem and provide the proofs of the most important results. In Section 1.4, results concerning unavoidable patterns are stated. The concepts of omnitonality and balanceability are stated in terms of results concerning unavoidable patterns and are discussed in Section 1.4.1 and Section 1.4.2 correspondingly. Here we also investigate the balancing number and the omnitonality number of different graph families. In Section 1.5, we speak of some graph families that are balanceable, omnitonal or both. One of these is the family of amoebas which is defined in Section 1.5.1 and afterwards we explore stars and paths in 1.5.2. We end the chapter with the definition of the generalized balancing number in Section 1.6.

In Chapter 2, we discuss results on balanceability and the balancing number. Sections 2.1 and 2.3 include results published in the first research paper [22] and a few additional results [21]. Sections 2.4 and 2.5 contain results from the second research paper [20]. Section 2.1 contains different results concerning the sufficient conditions for balanceability of a graph. In Section 2.2, we discuss a necessary condition for balanceability. In Section 2.3, we discuss the balanceability of some graph families such as complete graphs with an odd number of edges, two disjoint copies of a complete graph, the d-cube, a special class of circulant graphs and grids. In Section 2.4, we study the balancing number of some graphs classes. We define the generalized balancing number as an extension of the balancing number in Section 2.5 and provide the balancing number and generalized balancing number of cycles in Section 2.5.1. Section 2.5.2 contains results about the generalized balancing number of K_5. Finally, in Section 2.6 we list some open problems related to this chapter.

Chapter 3 deals with the graph family of amoebas. It includes the definition of amoebas using algebraic tools and some preliminary results on local and global amoebas. In Section 3.1, we provide a construction method that serves as a practical tool to construct families of local amoebas recursively. In Section 3.1.1 and Section 3.1.2 we present two such types of constructions. In Section 3.3 we expose some open problems that are relevant to the topic of amoeba graphs.

In Chapter 4, we explore the complete bipartite graph as a base graph for finding unavoidable patterns in 2-edge colorings. In Section 4.1, we study which are the unavoidable patterns that can be found in 2-edge colorings of a large enough complete bipartite graph with sufficient edges in each color. Section 4.2 includes results concerning bipartite omnitonality. In Section 4.3, we go over some results on bipartite balanceability and the bipartite balancing number of paths and stars. In Section 4.4 we provide a list of open problems that remain for future work in the context of $K_{n,n}$.

In Chapter 5, we present the conclusions of this work, we state or recall some open problems left for future work, and we describe some research lines that could be also pursued.

Chapter 1

Preliminaries

1.1 Definitions

In this chapter we include basic definitions and some graph theory notation that will be used throughout this thesis work. Definitions and notation used only in a particular chapter is be stated within the corresponding chapter.

If k be a positive integer, let $[k] = \{1, 2, \cdots, k-1, k\}$. We work with simple finite graphs $G = (V, E)$, where V, or $V(G)$, is the vertex set and E, or $E(G)$, is the edge set. Given a graph G, let $n(G)$ or $|V(G)|$ denote the number of vertices (or *order*) of G and $e(G)$ or $|E(G)|$ denote the number of edges in G. If u and v are distinct vertices joined by an edge in a graph G, we say that u is a neighbor of v and $uv \in E(G)$. If $v \in V(G)$, let $d(v)$ denote the *degree* of v in G, which is the number of neighbors of v in G. If $v \in V(G)$, let $N_G(v)$ be the set of neighbors of v in the graph G. If no specification on the graph is needed we only write $N(v)$. If $U \subseteq V(G)$, let $N(U)$ (or $N_G(U)$) be the union of neighbors of each vertex in U, namely $N(U) = \cup_{v \in U} N(v)$. If X and Y are disjoint subsets of $V(G)$, we denote $e(X, Y)$ as the number of edges in G with one end in X and the other in Y. Given a graph G and a subset W of $V(G)$, let $G[W]$ be the graph induced by the set W and let $e(W)$ be the number of edges induced by W. If X is a set, we say that a *t-subset* of X is a subset of X with t elements. We say that a graph G is *H-free* if it does not contain the graph H as a subgraph.

We say that two graphs G and H are *isomorphic*, denoted as $G \cong H$, if there is a bijection f between $V(G)$ and $V(H)$ such that any two vertices u and v of G are adjacent in G if and only if $f(u)$ and $f(v)$ are adjacent in H. A *walk* is a graph which consists of a finite or infinite sequence of edges which joins a sequence of vertices. A *trail* is a walk in which all edges are distinct. A *path* P_k is a trail which consists of a finite sequence of k edges that join a sequence of $k+1$ distinct vertices and can be expressed as $P_k = v_1 v_2 \cdots v_k v_{k+1}$. Let vv_1P_k ($P_k v_{k+1} u$ respectively) be the path $vv_1v_2 \cdots v_k v_{k+1}$ ($v_1 v_2 \cdots v_k v_{k+1} u$ respectively). If $v_i, v_j \in V(P_k)$ with $i < j$, let $v_i P_k v_j$ be the path $v_i v_{i+1} \cdots v_{j-1} v_j$ and let $v_j P_k v_i$ be the path $v_j v_{j-1} \cdots v_{i+1} v_i$. A *cycle* C_k on k vertices is a non-empty trail in which only the first and last vertices coincide. The *complete graph* K_n is a graph on n vertices where every pair of distinct vertices is connected by an edge. The *complete bipartite graph* $K_{n,m}$ is a graph whose vertices can be partitioned into two subsets X and Y, called *partition sets*, where $|X| = n$ and $|Y| = m$ such that no edge has both endpoints in the same subset and every possible edge that could connect vertices in different subsets is part of the graph. Hence all $n \cdot m$ possible edges are present. If G is a graph and k a positive integer, we denote as kG as the graph made of the disjoint union of k copies of G. Hence, the graph $2K_n$ is the graph constructed by the disjoint union of two complete graphs K_n, where n is a positive integer.

1.1.1 Edge colorings

A 2-*edge coloring* (or simply 2-*coloring*) of $E(G)$ is a function $f : E(G) \to \{r, b\}$ which associates each edge to one of two colors, r or b. Once the edges of a graph G have been colored by some 2-coloring f, we simply say that G is *colored* if no specification on the coloring is needed. Notice that we can associate every 2-coloring of the edges of K_n to a partition $E(K_n) = R \sqcup B$, where we can define R (respectively B) as the chromatic class, or edge set, that contains edges e such that $f(e) = r$ (respectively $f(e) = b$). We call the edges in R *red* and the edges in B *blue*, and both notions of a 2-coloring will be used. We call the graph induced by the set R (respectively B) the *red graph* (respectively the *blue graph*). We call the number of red (respectively blue) edges incident to a vertex v the *red neighborhood of* v (respectively *blue neighborhood of* v) and denote it as $deg_R(v)$ (respectively $deg_B(v)$).

The following definitions employ a graph $G = (V, E)$ of order n and a 2-edge coloring f which can also be seen as a partition $E(G) = R \sqcup B$.

- We say that a set Y is *monochromatic* under the coloring f if $f(y) = c$ for all $y \in Y$, for some color c.
- If Y is a monochromatic set where all $y \in Y$ is assigned the color red, for instance, then we say that Y is red.
- We say that a 2-edge coloring of G has *edge density* ε if there are at least $\varepsilon\binom{n}{2}$ edges in each color.
- We say there is an (r, b)-*colored copy of* H if there is a copy of H with exactly r red edges and b blue edges, where $e(H) = r + b$.
- We say that a 2-coloring of a graph H contains a *balanced copy* of G if we can find a copy of G in the colored H such that E can be partitioned into two sets (E_1, E_2) with $E_1 \subseteq R$, $E_2 \subseteq B$ and such that $||E_1| - |E_2|| \leq 1$.

The symmetric group S_n is the set of all permutations of a set of n elements with the operation of composition. The stabilizer group of a group H for an element i is the group generated by elements of H that fix i and is denoted as $\mathrm{Stab}_H(i)$. All further definitions regarding the algebraic setting will be stated in each relevant section.

1.2 Preliminaries in Ramsey theory and extremal graph theory

Ramsey theory and extremal graph theory are two areas that have become important pillars in combinatorics. Many of the problems solved in this thesis work and in other related texts make use of powerful tools found in these two important branches of combinatorics. In the following sections, relevant results to this work in each area are discussed. In Section 1.2.1 we state classic results in Ramsey theory and in Section 1.2.2 we state classic results in the area of extremal combinatorics.

1.2.1 Ramsey theory: Complete chaos is impossible

Ramsey theory is an important and well-established area in combinatorics. For a deep insight into Ramsey theory, see [36, 58] Even though it is named after Frank Plumpton Ramsey, who established the finite and infinite versions of the famous Ramsey's theorem in 1930 [54], problems of this nature were solved as far as 1892 with Hilbert's cube lemma [48]. But what is Ramsey theory? In words of the mathematician Veselin Jungic,

"If mathematics is a science of patterns, then Ramsey theory is a science of the stubbornness of patterns."

In other words, Ramsey theory studies the existence of ordered substructures in arbitrarily ordered structures. We define $R(s,t)$ as the minimum n such that any 2-edge coloring of K_n with colors red and blue must contain either a red K_s or a blue K_t as a subgraph. We can simply write $R(s)$ instead of $R(s,t)$ if $s = t$. $R(s)$ is also called the *diagonal Ramsey number*. The well known value of $R(3,3)$ is 6, because any 2-edge coloring of K_6 contains a monochromatic triangle and there exists a 2-edge coloring of K_5, consisting of a red C_5 and a blue C_5, which does not contain a monochromatic K_3 as a subgraph. The number $R(n_1, n_2, \cdots, n_r)$ is the minimum number N such that any r-edge coloring of K_N contains an i-colored K_{n_i} as a subgraph, for some $i \in \{1, 2, \cdots, r\}$. $R(n_1, n_2, \cdots, n_r)$ is called the *Ramsey number for r colors* or simply the *Ramsey number*. Ramsey's theorem states that $R(n_1, n_2, \cdots, n_r)$ is finite [53]. To prove this result we make use of the following less general case.

Theorem 1.1 (Ramsey's theorem for two colors [53])**.** *The Ramsey number $R(s,t)$ is finite for all $s, t \geq 2$.*

Proof. We use induction on $s+t$. Let $s+t = 4$. Because any 2-edge coloring of K_2 contains a monochromatic edge and this does not hold for K_1, we have that $R(2,2) = 2$.

Suppose that $R(s,t)$ is finite when $s+t = n-1$. It is sufficient to show that

$$R(s,t) \leq R(s-1,t) + R(s,t-1)$$

when $s+t = n$ to achieve the result. Let $N = R(s-1,t) + R(s,t-1)$. We prove that any 2-edge coloring of K_N contains a red K_s or a blue K_t.

Take any 2-edge coloring of K_N. Let v be any vertex of K_N and focus on its neighboring vertices. Let A be the set of vertices adjacent to v via red edges and let B be the set of vertices adjacent to v via blue edges. Notice that

$$|A| + |B| = R(s-1,t) + R(s,t-1) - 1$$

so one of the following must hold: either $|A| \geq R(s-1,t)$ or $|B| \geq R(s,t-1)$.

If $|A| \geq R(s-1,t)$ then there is a red K_{s-1} within A such that along with v would make a red K_s, or there is a blue K_t within A. In either case, we achieve the result. The arguments are analogous if $|B| \geq R(s,t-1)$.

□

Very few values of Ramsey numbers are known. The exact values up to date are $R(3,3)=6$ [43], $R(3,4)=9$ [43], $R(3,5)=14$ [43], $R(3,6)=18$ [42], $R(3,7)=23$ [32], $R(3,8)=28$ [44], $R(3,9)=36$ [44], $R(4,4)=18$ [43], $R(4,5)=25$ [51]. A classical result [28] of Erdős and Szekeres states that

$$R(s) \leq (1+o(1))\frac{4^{s-1}}{\sqrt{\pi s}}.$$

Another well-known result which is an exponential lower bound by Erdős [24] is

$$R(s) \geq (1+o(1))\frac{s}{\sqrt{2}e}2^{s/2}.$$

The best known bounds for diagonal Ramsey numbers are

$$(1+o(1))\frac{\sqrt{2}s}{e}2^{s/2} \leq R(s) \leq s^{-(c\log s)/(\log\log s)}4^s$$

and they are due to Spencer [59] and Conlon [17] respectively.

We now state and prove a more general variant of Ramsey's theorem using r colors.

Theorem 1.2 (Ramsey's theorem 1930 [53])**.** *For every positive integer r, there is some integer $N = R(n_1, n_2, \cdots, n_r)$ such that if the edges of K_N are colored with r colors, then there is always a monochromatic K_{n_i} as a subgraph in color i, for some $i \in \{1, 2, \cdots, r\}$.*

Proof. We proceed by induction on the number of colors r. The fact that $R(n_1, n_2)$ is finite is given by Theorem 1.1.

The induction hypothesis states that $R(n_1, n_2, \cdots, n_{r-1})$ is finite. Therefore, showing that

$$R(n_1, n_2, \cdots, n_{r-1}, n_r) \leq R(n_1, n_2, \cdots, n_{r-2}, R(n_{r-1}, n_r))$$

is sufficient to prove the theorem's result. Let $T = R(n_1, n_2, \cdots, n_{r-2}, R(n_{r-1}, n_r))$. We prove that any r-edge coloring of K_T contains a monochromatic K_{n_i} as a subgraph in color i, for some $i \in \{1, 2, \cdots, r\}$.

Take any r-edge coloring of K_T. Suppose that the color $r-1$ is light blue and the color r is dark blue. If we look at all the light blue and dark blue edges as being in the same color class, we have an $(r-1)$-edge coloring of K_T which, by the induction hypothesis, means that there is a monochromatic K_t as a subgraph with $t \in \{n_1, \cdots, n_{r-2}, R(n_{r-1}, n_r)\}$. If $t \in \{n_1, \cdots, n_{r-2}\}$ the inequality clearly holds. If $t = R(n_{r-1}, n_r)$, then we can re-distinguish between light blue and dark blue edges, and by the induction hypothesis there is a light blue $K_{n_{r-1}}$ or a dark blue K_{n_r} as a subgraph of K_T and the inequality holds. □

Ramsey-type theorems state typically that any coloring of a large enough structure contains a monochromatic fixed substructure.

One of the first results in Ramsey theory is Schur's theorem.

Theorem 1.3 (Schur 1916, finitary version [56])**.** *For every positive integer r, there exists a positive integer $s = s(r)$ such that any coloring of $[s]$ with r colors contains a monochromatic solution to the equation $x + y = z$ with $x, y, z \in [s]$.*

Proof. Let r be a positive integer and let N be the Ramsey number for r colors $R(3, 3, \cdots, 3)$. Consider and arbitrary r-coloring $\chi : [N] \to [r]$. Color the edges of the complete graph K_{N+1} by assigning the edge $\{i, j\}$, with $i < j$, the color $\chi(j - i)$. By Ramsey's theorem (Theorem 1.2) there is a monochromatic triangle on some vertices, say $i < j < k$. Therefore, $\chi(j-i) = \chi(k-j) = \chi(k-i)$. If $x = j-i$, $y = k-j$ and $z = k-i$, then $x+y = z$ holds. □

Schur's theorem in its finitary version allows us to ask how big does $N(r)$ have to be as a function of r. This is a typical Ramsey theory question and as for most questions of this type, there is no concrete answer, except for a few values. The first three Schur numbers are $s(1) = 1$, $s(2) = 5$ and $s(3) = 14$, which can be deduced manually. The following two Schur numbers are $s(4) = 45$ [1, 40] and $s(5) = 161$ [47]. For instance, $s(r)$ is only known for $r \leq 5$ Schur's theorem is one of the first results in the area of additive combinatorics. Another fundamental Ramsey-type theorem which was also an important development in additive combinatorics is Van der Waerden's theorem, which is stated as follows.

Theorem 1.4 (Van der Waerden 1927 [62])**.** *For every $r \in \mathbb{N}$ there exists a positive integer $N(r) = N$ such that any coloring of $[N]$ with r colors contains a monochromatic solution to the equation $x + y = 2z$.*

In some sense, Ramsey theory can be seen as the area that studies the unavoidability of highly regular patterns (such as a monochromatic triangle or a monochromatic solution to an equation) in arbitrarily ordered structures. Results of this kind date back to 1892 and it is a very active research area. In this work we will use Ramsey theory to look for other unavoidable colored patterns which will be discussed in further chapters.

1.2.2 Extremal graph theory

The origins of extremal graph theory date back to 1907 with a theorem due to Willem Mantel [50], a Dutch mathematician who wondered about how many edges a graph could have with the property of being triangle-free.

Theorem 1.5 (Mantel)**.** *A triangle-free graph on n vertices has at most $\lfloor \frac{n^2}{4} \rfloor$ edges.*

We provide the latest proof of Mantel's theorem by Aharoni, De Vos, González, Montejano and Šámal from 2020. It was obtained as a byproduct by the authors when working on a rainbow version of Mantel's theorem.

We begin by proving the following Lemma.

Lemma 1.6 ([2]). *Let G be a graph and let P be the set of pairs of distinct vertices $\{x, y\} \subseteq V(G)$ such that $N(x) \cap N(y) \neq \emptyset$. If M is a maximal matching in G, then $|P| \geq |E(G)| - |M|$.*

Proof. Let $M = \{e_1, e_2, \cdots, e_k\}$ be a maximal matching of G. Because M is maximal, we know that every edge $e \in E(G)$ has at least one endpoint in common with an edge in M. Let $s(e)$ be the minimum integer such that $e \cap e_{s(e)} \neq \emptyset$. Let $f : E(G) \setminus M \to P$ with $f(e) = e \triangle e_{s(e)}$. It is easy to see that f is an injective function, and so the result follows. □

Proof of Theorem 1.5 [2]. Let G be a triangle-free graph and M a maximum matching of G. Because G has no triangles we can sum over every element in P and every edge to get the following inequality.

$$|P| + |E(G)| \leq \binom{n}{2}.$$

By Lemma 1.6, we have $|E(G)| - \frac{1}{2}n \leq |E(G)| - |M| \leq |P|$. If we combine both inequalities, we get $2|E(G)| \leq \binom{n}{2} + \frac{1}{2}n$ and so $|E(G)| \leq \frac{1}{4}n^2$.

□

There is a particular example of a triangle-free graph on n vertices with exactly $\lfloor \frac{n^2}{4} \rfloor$ edges and that is the complete bipartite graph with $\lfloor \frac{n}{2} \rfloor$ vertices in one part and $\lceil \frac{n}{2} \rceil$ in the other. This is called an *extremal graph* for Mantel's theorem.

The area of extremal theory is also commonly called Turán theory as its main iniciator is considered to be Pál Turán, who proved the generalization of Mantel's theorem for K_{r+1}-free graphs (see Theorem 1.7). Extremal theory asks questions such as "How many edges can a graph on n vertices have in order to avoid a fixed graph H as a subgraph?" or the equivalent question "What is the least amount of edges m that guarantees that any graph on n vertices and more than m edges will contain a fixed graph H?" Turán's theorem [60] dates back to 1941. Before we state it, we describe the extremal graph called the Turán graph.

The *Turán graph* $T_{n,r}$ is the complete r-partite graph on n vertices with parts of sizes $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$. Consider $n = qr + s$ for a positive integer s with $0 \leq s \leq r$, then $T_{n,r}$ is the complete r-partite graph with s parts of $q+1$ vertices and the remaining $r - s$ parts with q vertices. Hence, the exact number of edges in $T_{n,r}$ is

$$\left(1 - \frac{1}{r}\right) \frac{n^2 - s^2}{2} + \binom{s}{2}.$$

We proceed to state Turan's theorem along with a short proof that employs the probabilistic method.

Theorem 1.7 (Turán 1941 [60]). *If G is an n vertex K_{r+1}-free graph, then it has at most $e(T_{n,r})$ edges.*

Proof. We fix r and proceed by induction on n, with $n > r$ or else the result is trivial. Assume that Turán's theorem holds for all graphs on fewer than n vertices. Let G be a K_{r+1}-free graph with the maximum number of edges possible. Observe that G must contain K_r as a subgraph or else we could add an edge and G could remain K_{r+1}-free. Let A be the set of vertices of the subgraph K_r and let $B = V(G) \setminus A$. Because G is K_{r+1}-free, every vertex $v \in B$ can have at most $r-1$ neighbors in A. Therefore

$$e(G) \leq \binom{r}{2} + |B|(r-1) + e(B) \quad \leq \binom{r}{2} + (n-r)(r-1) + e(T_{n-r,r}) = e(T_{n,r})$$

The first inequality follows from the number of edges in A, the maximum number of edges going from B to A and the number of edges in B. The second inequality employs the cardinality of B and the induction hypothesis. Finally, the third inequality follows from the observation that removing one vertex from each of the r parts in $T_{n,r}$ would remove $\binom{r}{2} + (n-r)(r-1) + e(T_{n-r,r})$ edges.

$\square$

Let H be a graph, we say that $\text{ex}(n, H)$ is the maximum number of edges an H-free graph on n vertices can have. Another well-studied parameter in graph theory is the chromatic number. Let H be a graph and let $c : V(H) \to \{1, 2, \cdots, k\}$ be a coloring of $V(H)$ such that $c(v) \neq c(u)$ if $uv \in E(H)$. We call c a *proper coloring* of $V(H)$. The chromatic number $\chi(H)$ is the minimum number t such that there is a proper coloring of $V(H)$ with t colors. Finally, we state an important theorem in the area which establishes a direct connection between the extremal number and its chromatic number.

Theorem 1.8 (Erdős-Stone 1946 [27], Erdős-Simonovits 1966 [26]). *For a fixed graph H,* $\text{ex}(n, H) = (1 - \frac{1}{\chi(H)-1} + o(1))\frac{n^2}{2}$.

If we know the chromatic number of H, we know a lot of information about the growth rate of the function $\text{ex}(n, H)$. As long as $\chi(H) \geq 3$, we know the first order asymptotics from the Erdős-Stone-Simonovits theorem. But what do we know for bipartite graphs, i.e. for graphs H with $\chi(H) = 2$?

1.2.3 Forbidding a bipartite graph

If H is bipartite, the Erdős-Stone-Simonovits theorem only tells us that $\text{ex}(n, H) = o(n^2)$, but we will shortly see that this can be improved. Nevertheless, there are many open problems concerning the growth rate of the function $\text{ex}(n, H)$ when H is a bipartite graph. Another interesting aspect is to find the largest number of edges $z(m, n; s, t)$ (known as the *Zarankiewicz number*) in a bipartite graph, where one part has m vertices and the other has n vertices, with no copies of the complete bipartite graph $K_{s,t}$. If G is a bipartite graph and

$n = m$, the Zarankiewicz number can also be denoted as $z(n, G)$ and it refer to the maximum number of edges in a bipartite graph with n vertices in each part, having no copies of G. Determining $z(m, n; s, t)$ is a famous open problem called the Zarankiewicz problem.

Problem 1.9. *(Zarankiewicz problem) For integers m, n, s, t such that $m \geq s \geq 2$ and $n \geq t \geq 2$, determine the maximum number of edges a bipartite graph, with m and n vertices in each part respectively, can have without $K_{s,t}$ as a subgraph.*

Now we state and prove an important result in this area: the Kővári-Sós-Turán theorem.

Theorem 1.10 (Kővári-Sós-Turán 1954 [34])**.** *For integers m, n, s, t such that $n \geq s \geq 2$ and $m \geq t \geq 2$,*

$$z(m, n; s, t) \leq \left((t-1)^{\frac{1}{s}}(n-s+1)m^{1-\frac{1}{s}} + (s-1)m\right).$$

The Zarankiewicz number and the extremal number of a graph G are closely related as we can see in the following proposition.

Proposition 1.11. *Let G be a bipartite graph, then $2\,\mathrm{ex}(n, G) \leq z(n, G)$.*

Proof. Let H be a graph on n vertices, namely $\{v_1, \cdots, v_n\}$, with $\mathrm{ex}(n, G)$ edges that does not contain G as a subgraph. Let $V_1 = \{u_1, \cdots, u_n\}$ and $V_2 = \{w_1, \cdots, w_n\}$ be two disjoint copies of $V(H)$ where u_i and w_i are copies of v_i for $1 \leq i \leq n$. We construct a bipartite graph H' with partition sets V_1 and V_2, and $u_i w_j \in E(H')$ if and only if $v_i v_j \in E(H)$. Therefore if H does not contain G as a subgraph, then neither does H'. Because $e(H') = 2e(H) = 2\,\mathrm{ex}(n, H)$, then the inequality $2\,\mathrm{ex}(n, G) \leq z(n, G)$ holds. $\square$

Because every bipartite graph H with bipartition $V(H) = A \cup B$ is a subgraph of some $K_{|A|,|B|}$, then a natural lower bound can be given

$$ex(n, H) \leq ex(n, K_{|A|,|B|}).$$

The following theorem is a consequence of Theorem 1.10 and Proposition 1.11. However, we present it here with a proof. Now we state a theorem that gives an upper bound of $\mathrm{ex}(n, K_{s,t})$.

Theorem 1.12 (Kővári-Sós-Turán 1954 [34])**.** *For every pair of integers $1 \leq s, t$ with $s \leq t$,*

$$ex(n, K_{s,t}) \leq \frac{1}{2}\left((t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + \frac{1}{2}(s-1)n\right).$$

Proof. We use a double counting argument. Let G be a graph on n vertices and m edges that is $K_{s,t}$-free, with s and t fixed. Let k be the number of stars $K_{1,s}$ in G.

For the upper bound of k, we know that every set of s vertices has at most $t-1$ common neighbors because G is $K_{s,t}$-free. Hence

$$k \leq \binom{n}{s}(t-1).$$

To give a lower bound on k, we go over every vertex and count the $K_{1,s}$'s it can make and use an argument of convexity in the first inequality to obtain the following.

$$k = \sum_{v\in V(G)} \binom{d(v)}{s} \geq n\binom{\frac{1}{n}\sum_{v\in V(G)} d(v)}{s} = n\binom{\frac{2m}{n}}{s}.$$

We put both bounds together keeping in mind that s and t are fixed. The point is to see how m and n depend on each other as they get large.

$$n\binom{\frac{2m}{n}}{s} \leq \binom{n}{s}(t-1).$$

Since $\binom{n}{s} = (1+o(1))\frac{n^s}{s!}$ asymptotically for a fixed s. We now apply the identity on both sides and reach the desired result.

$$n\left(\frac{2m}{n}\right)^s \leq (1+o(1))n^s(t-1)$$

$$m \leq \frac{1}{2}(1+o(1))^{\frac{1}{s}}(t-1)^{\frac{1}{s}}n^{2-\frac{1}{s}}.$$

□

Hence, for a fixed s and t, the upper bound on $\mathrm{ex}(n, K_{s,t})$ grows like $n^{2-\frac{1}{s}}$. We state the explicit upper bound for $\mathrm{ex}(n, K_{t,t})$ as it will be used in Chapter 4.

Corollary 1.13 (Kövári-Sós-Turán 1954 [34])**.** *For every positive integer t,*

$$\mathrm{ex}(n, K_{t,t}) < \frac{1}{2}\left((t-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n\right).$$

The most natural question one can ask is if Corollary 1.13 is tight and in fact it is a major conjecture in extremal graph theory. Only a small number of values is known. There are some values of s and t for which we do know this theorem is tight, for example $s = t = 2$, $s = t = 3$, and s and t when t is sufficiently large compared to s. This means that there are constructions of graphs that are $K_{s,t}$-free according to these parameters and whose numbers of edges match the previous upper bound up to a constant factor. We state three important results, along with a brief description of the techniques employed or proof sketches that corroborate the conjecture for certain values of s and t. The conjecture remains open for all remaining values of s and t.

Theorem 1.14 (Erdős-Rényi-Sós 1966 [25])**.** $\mathrm{ex}(n, K_{2,2}) \geq (\frac{1}{2} - o(1))n^{3/2}$.

This result employs an algebraic construction of a $K_{2,2}$-free graph with many edges. Taking $n = p^2 - 1$ with p prime, the authors consider the *polarity graph* G where the vertex set is $\mathbb{F}_p^2 \setminus \{0, 0\}$ and two vertices (x, y) and (a, b) form an edge if $ax + by = 1$ in $\mathbb{F}_p$. Notice that for two distinct vertices (a, b) and (a', b'), there is at most one common neighbor $(x, y) \in V(G)$ satisfying both $ax + by = 1$ and $a'x + b'y = 1$, making G a $K_{2,2}$-free graph. Because every vertex has degree p or $p - 1$,

$$e(G) = \left(\frac{1}{2} - o(1)\right)p^3 = \left(\frac{1}{2} - o(1)\right)n^{\frac{3}{2}}.$$

Notice that if n is not $p^2 - 1$ for any prime p then we use the largest prime p such that $p^2 - 1 \leq n$ and construct the polarity graph on $p^2 - 1$ vertices and add $n - p^2 + 1$ isolated vertices.

For the next theorem, we discuss a construction for $K_{3,3}$-free graphs, which follows a similar idea as the previous construction.

Theorem 1.15 (Brown 1966 [7])**.** $\text{ex}(n, K_{3,3}) \geq (\frac{1}{2} - o(1))n^{5/3}$.

This proof takes n to be p^3 where p is a prime. The authors build the graph G with vertex set $\mathbb{F}_p^3$ and edge set $\{(x, y, z)(a, b, c) \mid (a - x)^2 + (b - y)^2 + (c - z)^2 = u \text{ in } \mathcal{F}_p\}$ where u is some fixed non-zero element that was carefully chosen in $\mathbb{F}_p$. The element u needs to be chosen so that the graph G is indeed $K_{3,3}$-free. If we had considered points in $\mathbb{R}^3$, the property of being $K_{3,3}$-free is equivalent to the statement that three unit spheres have at most two common points. This statement can be proved using algebraic tools that are similar to those employed to algebraically manipulate $\mathbb{F}_p$ and verify that G is $K_{3,3}$-free.

In G each vertex has degree around p^2 due to the distribution of $(a-x)^2+(b-y)^2+(c-z)^2$ is almost uniform across $\mathbb{F}_p$ as (x, y, z) varies randomly over $\mathbb{F}_p^3$. Hence a $\frac{1}{p}$ fraction of (x, y, z) is expected to satisfy $(a - x)^2 + (b - y)^2 + (c - z)^2 = u$. This is meant to give only an intuitive idea of the proof.

The cases of $K_{2,2}$ and $K_{3,3}$ are fully solved but the case of $K_{4,4}$ remains a central open problem in extremal graph theory.

Problem 1.16. *What is the growth rate of* $\text{ex}(n, K_{4,4})$*? Does it match the upper bound in Theorem 1.10?*

So far we have matched the Kővári-Sós-Turán bound up to a constant factor for $K_{2,2}$ and $K_{3,3}$, but many cases remain open. The next theorem presents a construction that matches the Kővári-Sós-Turán bound for $K_{s,t}$ when s and t are sufficiently far apart.

Theorem 1.17 (Kollár, Rónyai, Szabó 1996 [33], Alon, Rónyai, Szabó 1999 [4])**.** *If* $t \geq (s - 1)! + 1$ *then* $\text{ex}(n, K_{s,t}) = \Theta(n^{2-\frac{1}{s}})$.

The proof of this theorem is substantially more elaborate. We aim to provide the main idea of the overall proof by stating the propositions that lead to the main result.

The authors begin by proving a weaker version for $t \geq s! + 1$ which is later adjusted to achieve the desired bound. Let $n = p^s$ where p is a prime and $s \geq 2$. Consider the norm map $N : \mathbb{F}_{p^s} \to \mathbb{F}_p$ defined by

$$N(x) = x \cdot x^p \cdot x^{p^2} \cdot x^{p^{s-1}} = x^{\frac{p^s-1}{p-1}}.$$

They define the $NormGraph_{p,s} = (V, E)$ as follows

$$V = \mathbb{F}_{p^s}$$

and

$$E = \{ab \mid a \neq b, N(a + b) = 1\}.$$

In a series of propositions, the authors prove that $|E| \geq \frac{1}{2}n^{2-\frac{1}{s}}$ via counting arguments, and that the $NormGraph_{p,s}$ is $K_{s,s!+1}$-free by bounding from above the number of common neighbors to a set of s vertices and using results that employ algebraic geometry. They make an adjustment to achieve the bound $t \geq (s-1)! + 1$ in Theorem 1.17 by defining the graph $ProjNormGraph_{p,s} = (V, E_P)$ with

$$V = \mathbb{F}_{p^{s-1}} \times \mathbb{F}_p^{\times}$$

for $s \geq 3$. Taking $n = (p-1)p^{s-1}$, they define the edge relations as

$$(X, x) \sim (Y, y)$$

if and only if

$$N(X + Y) = xy.$$

By proving that $|E_P| = (\frac{1}{2} - o(1))n^{2-\frac{1}{s}}$, we know $ProjNormGraph_{p,s}$ has suffient edges and finally the authors prove that it is $K_{s,(s-1)!+1}$-free using algebraic geometry and the result is achieved.

If H is a bipartite graph, then it is always contained in $K_{s,t}$ for some s and t sufficiently large. Therefore, the Kővári-Sós-Turán theorem gives us an immediate upper bound for $\text{ex}(n, H)$. Nevertheless, if H is sparse, then this bound can be quite bad. We will explore a powerful technique that gives us a better upper bound on $\text{ex}(n, H)$ when H is a sparse bipartite graph.

Theorem 1.18 (Fűredi 1991 [37]. Alon, Krivelevich and Sudakov 2003 [3])**.** *Let H be a bipartite graph whose vertex set is $A \cup B$ where every vertex in A has degree at most r. Then there exists a constant $C = C(H)$ such that*

$$\text{ex}(n, H) \leq Cn^{2-\frac{1}{r}}.$$

This bound is tight due to Theorem 1.17 by taking $H = K_{r,t}$ for some $t \leq (r-1)! + 1$.

We discuss the proof of Theorem 1.18 as it makes use of a powerful probabilistic technique called dependent random choice which is also used further along in this thesis work. Broadly stated, the result, that is enounced in the following theorem, says the following:

If G has many edges, then there exists a large subset U of $V(G)$ such that every small subset in U has a large common neighborhood.

Before stating the theorem, we make use of some basic probability notation. Let $P(A)$ denote the probability that an event A can occur. Consider a random variable X with a finite list $x_1, \cdots, x_k$ of possible outcomes, where each x_i has probability p_i of occurring, for $1 \leq i \leq k$. Let $k \in \mathbb{Z}^+$ and $[k] = \{1, 2, \cdots, k\}$. The *expected value* of X is defined as $\mathbb{E}(X) = \sum_{i \in [k]} x_i p_i$.

Theorem 1.19 (Dependent random choice. Alon, Krivelevich and Sudakov 2003 [3])**.** *Let $u, n, r, m, t \in \mathbb{N}$ and $\alpha > 0$ satisfy the following inequality*

$$n\alpha^t - \binom{n}{r}\left(\frac{m}{n}\right)^t \geq u.$$

Then every graph G of order n and at least $\alpha n^2/2$ edges contains a subset U of vertices with size at least u such that every subset of r elements S in U has at least m common neighbors.

Proof. Let T be a list of t vertices chosen uniformly at random from $V(G)$ allowing repetition. Let A be the common neighborhood of T. We wish to know how large A can be, therefore the expected value of $|A|$ is

$$\begin{aligned}
\mathbb{E}|A| &= \sum_{v \in V} \mathbb{P}(v \in A) && (1.1)\\
&= \sum_{v \in V} \mathbb{P}(T \subseteq N(v)) && (1.2)\\
&= \sum_{v \in V} \left(\frac{d(v)}{n}\right)^t && (1.3)\\
&\geq n\left(\frac{1}{n}\sum_{v \in V}\frac{d(v)}{n}\right)^t && (1.4)\\
&\geq n\alpha^t && (1.5)\\
& && (1.6)
\end{aligned}$$

Eguality 1.3 is straight forward. The probability that each vertex in T is in $N(v)$ is precisely $\frac{d(v)}{n}$. Because T has t elements, the product of these t probabilities gives us the desired equality 1.4. The inequality 1.5 is given by convexity (Jensen's inequality) which states that the sum of the t powered average degrees is at least n times the t powered average degree. Hence, $\sum_{v \in V}(\frac{d(v)}{n})^t \geq n\left(\frac{1}{n}\sum_{v \in V}\frac{d(v)}{n}\right)^t$. The last equality is given by the hand-shake lemma and the hypothesis on the number of edges in G.

For every subset S of V with r elements, the event of A containing S holds if and only if T is contained in the common neighborhood of S. The probability of this event is $\left(\frac{c}{n}\right)^t$ where c is the number of common neighbors.

We say S is a *bad* set if it has less than m common neighbors. This implies that each bad set with r elements $S \subset V$ is contained in A with probability less than $(\frac{m}{n})^t$. By linearity of expectation and letting b be the number of bad subsets of A with r elements,

$$\mathbb{E}[b] < \binom{n}{r}\left(\frac{m}{n}\right)^t.$$

We are interested in not having any bad subsets, therefore we can eliminate one element in each bad subset. The number of remaining elements is at least $|A| - b$, whose expected value is at least

$$n\alpha^t - \binom{n}{r}\left(\frac{m}{n}\right)^t \geq u.$$

Therefore, there exists a set T such that there are at least u elements in A after destroying all the bad subsets. The set U with the remaining u elements satisfies the theorem's properties. $\square$

We now state a consequence of the dependent random choice theorem which will be useful in the proof of Theorem 4.3.

Corollary 1.20 (Consequence of dependent random choice)**.** *For every $m, t \in \mathbb{Z}^+$ there is a constant $C = C(m,t) > 0$ such that every graph G on n vertices with $e(G) \geq C(n^{2-\frac{1}{t}})$ contains a set $U \subset V(G)$ with m vertices such that every t-subset of U has at least m common neighbors.*

Proof. By Theorem 1.19 with $\alpha = Cn^{-\frac{1}{t}}$, $u = m$ and $r = t$, we can rewrite the inequality from Theorem 1.19 to obtain

$$n(Cn^{-\frac{1}{t}})^t - \binom{n}{t}\left(\frac{m}{n}\right)^t \geq m.$$

Now we have to find a constant C so that this inequality is true. Notice that the first term evaluates to C^t and the second term is $\theta(1)$. Therefore, we can easily find a big enough constant C to make the inequality hold. $\square$

1.3 Unavoidable patterns: Ramsey-type approach

Given a positive integer t, the family $\mathcal{F}_t$ contains all 2-edge colored copies of K_{2t} where either one color forms a copy of K_t or one color forms two disjoint copies of K_t. Bollobás conjectured that for every $\varepsilon > 0$ and positive integer t, there is a positive integer $n(t,\varepsilon)$ such that every 2-edge coloring of K_n with $n > n(t,\varepsilon)$ with at least $\varepsilon\binom{n}{2}$ edges in each color class contains an element of $\mathcal{F}_t$ (see [19]). The notion of the family $\mathcal{F}_t$ is also defined in [19] as copies of type A and type B, where a 2-edge-coloring of K_{2t} is a *type-A* if the edges of one of the colors induce a complete graph K_t and it is a *type-B* if the edges of one of the colors induce two disjoint K_t's (or equivalently one $K_{t,t}$).

This Ramsey-type conjecture was proved by Cutler and Montágh in 2007 [19] where they showed that $n(t,\epsilon) < 4^{\frac{t}{\varepsilon}}$.

In 2008, Fox and Sudakov [35] provided a much simpler proof of the conjecture and determined that $n(t,\varepsilon) = \varepsilon^{-ct}$ for some constant c. The bound is tight up to the constant factor in the exponent for all t and ε. We provide the proof of the better bound which makes use of two lemmas, one of which uses the probabilistic technique dependent random choice, given in Theorem 1.19.

Theorem 1.21 (Fox, Sudakov 2008 [35]). *If $n \geq (16/\varepsilon)^{2t+1}$, then every 2-edge coloring of K_n with at least $\varepsilon\binom{n}{2}$ edges in each color contains a member of $\mathcal{F}_t$.*

We prove two lemmas that are used in the proof of Theorem 1.21. The first lemma shows that any 2-edge coloring with edge density at least ε in each color has a large set of vertices with many common neighbors in both colors.

Lemma 1.22. *In every 2-edge coloring of K_n ($n \geq 4$) in which there are at least $\varepsilon\binom{n}{2}$ edges in each color, there is a subset $S \subset V(K_n)$ of cardinality at least $\frac{\varepsilon}{2}n$ and whose vertices have degree at least $\frac{\varepsilon}{4}$ in both colors.*

Proof. We proceed by contradiction assuming that $|S| < \frac{\varepsilon}{2}n$. Observe that $\varepsilon \leq 1/2$ by the hypothesis on the density of each color class. Let R be the set of vertices in K_n with blue degree less than $\frac{\varepsilon}{4}n$ and let B be those vertices with red degree less than $\frac{\varepsilon}{4}n$. Then $S \cup R \cup B$ is a partition of $V(K_n)$. Without loss of generality assume that $|R| \geq |B|$, then

$$|R| \geq \frac{1}{2}(n - |S|) > \frac{1}{2}n(1 - \frac{\varepsilon}{2}) \geq \frac{3}{8}n.$$

The number of red edges between R and B is less than $\frac{\varepsilon}{4}n|B|$ because each vertex in B has red degree less than $\frac{\varepsilon}{4}n$. This same number is also greater than $|R|(|B| - \frac{\varepsilon}{4}n)$ because every vertex in R has blue degree less than $\frac{\varepsilon}{4}n$. Hence,

$$\frac{\varepsilon}{4}n|B| > |R|(|B| - \frac{\varepsilon}{4}n) > \frac{3}{8}(|B| - \frac{\varepsilon}{4}n).$$

Using the fact that $\varepsilon \leq 1/2$ and rearranging the terms, we get

$$\frac{\varepsilon}{4}n \cdot \frac{3}{8}n > |B|(\frac{3}{8}n - \frac{3}{4}n) > |B|\frac{n}{4}$$

$$\frac{3}{8}\varepsilon n > |B|$$

Now we give an upper bound on the total number of blue edges. Observe that the vertices in R participate in at most $\frac{\varepsilon}{4}n|R|$ blue edges. The remaining edges are those in $B \cup S$ which are a total of $\binom{|B|+|S|}{2}$. Notice that $|B| + |S| < (\frac{3}{8} + \frac{1}{2})\varepsilon n < \varepsilon n$. Let $x = |B| + |S|$. The number of blue edges is at most

$$\begin{aligned}\frac{\varepsilon}{4}n|R| + \binom{|B|+|S|}{2} &= \frac{\varepsilon}{4}n(n-x) + \binom{x}{2}\\ &< \frac{\varepsilon}{4}n^2 + \binom{x}{2}\\ &< \frac{\varepsilon}{4}n^2 + \binom{\varepsilon n}{2}\\ &< \frac{\varepsilon}{4}n^2 + \frac{\varepsilon^2 n^2}{2}\\ &\leq \varepsilon\binom{n}{2}\end{aligned}$$

which contradicts that the number of blue edges is at least $\varepsilon\binom{n}{2}$. □

The next lemma uses the technique of dependent random choice.

Lemma 1.23. *For a 2-edge coloring of K_n, let S denote the vertex subset such that every vertex in S has degree at least αn in each color, and $s = |S|$. If $\beta \leq s^{-t/h}$, then there is a subset $T \subset S$ with $|T| \geq \frac{1}{2}\alpha^h(1-\alpha)^h s - 2$ such that every t-tuple in T has at least βn common neighbors in each color.*

Proof. Let $\alpha \leq \frac{1}{2}$ and $n \geq 4^{h+1}$, otherwise taking T as the empty set works. Let $U_1 = \{x_1, \cdots, x_h\}$, $U_2 = \{y_1, \cdots, y_h\}$ be two sets of h vertices taken uniformly at random with repetitions from $V(K_n)$. Let W be the set of vertices of S that are adjacent in red to every vertex in U_1 and adjacent in blue to every vertex in U_2. Hence

$$W = N_R(U_1) \cap N_B(U_2) \cap S.$$

Because every vertex in S has degree at least αn in each color and the inequality $\alpha \leq 1/2$ implies that $1 - \alpha - 1/n \geq 1 - \alpha - 2/n + 2\alpha/n$, then observe that

$$|N_R(v)||N_B(v)| \geq \alpha n((1-\alpha)n - 1) \geq \alpha(1-\alpha)(1-\frac{2}{n})n^2$$

for each $v \in S$. Now we take the expected value of $|W|$ using that $n \geq 4^{h+1}$.

$$\begin{aligned}\mathbb{E}[|W|] = \sum_{v\in S} Pr(v \in N_R(U_1))Pr(v \in N_B(U_2)) &= \sum_{v\in S}\left(\frac{|N_R(v)|}{n}\right)^h\left(\frac{|N_B(v)|}{n}\right)^h\\ &\geq \alpha^h(1-\alpha)^h(1-2/n)^h s\\ &\geq \frac{1}{2}\alpha^h(1-\alpha)^h s.\end{aligned}$$

The second inequality is given because the probability that $v \in N_R(U_1)$ equals the probability that v is a red neighbor to every member of U_1. The probability that a given set

$D \subset S$ of vertices is adjacent to U_1 by red edges is $\left(\frac{|N_R(D)|}{n}\right)^h$. Let Z denote the number of t-tuples in W with less than βn common red neighbors.

$$\mathbb{E}[Z] = \sum_{D \subset S, |D|=t, |N_R(D)|<\beta n} Pr(D \subset N_R(U_1)) \leq \binom{s}{t}\beta^h \leq s^t \beta^h \leq 1.$$

We similarly obtain that the expected number, denoted as Z', of t-tuples in W with less than βn common blue neighbors is at most 1. Due to linearity, the expectation of $|W| - Z - Z'$ is at least $\frac{1}{2}\alpha^h(1-\alpha)^h s - 2$ and this means that there are choices for U_1 and U_2 such that the corresponding value of $|W| - Z - Z'$ is at least $\frac{1}{2}\alpha^h(1-\alpha)^h s - 2$. We now delete a vertex from every t-tuple $D \subset W$ with less than βn red common neighbors or with less than βn blue common neighbors. Let T be the resulting set and notice that it satisfies the desired properties.

□

We can now prove Theorem 1.21. Recall that the Ramsey number $R(t)$ is the least positive integer n such that every 2-edge coloring of the complete graph K_n contains a monochromatic K_t. We use the bound $R(t) < 4^t$ mentioned in Section 1.2.1.

Proof of Theorem 1.21. We take a 2-edge coloring of K_n with $n \geq (16/\varepsilon)^{2t+1}$ in which both colors have denisity at least ε. By Lemma 1.22 there is a set S with at least $\frac{\varepsilon}{2}n$ vertices in which every vertex has at least $\frac{\varepsilon}{4}n$ neighbors in each color. We apply Lemma 1.23 with $\alpha = \frac{\varepsilon}{4}$, $\beta = \frac{R(t)}{n}$, $h = 2t$, and $s = |S| \geq \frac{\varepsilon}{2}n$ because

$$\beta = \frac{R(t)}{n} < 4^t/n < n^{-1/2} \leq s^{-t/h}.$$

There is a set $T \subset S$ with

$$|T| \geq \frac{1}{2}(\alpha(1-\alpha))^{2t}s - 2 \geq \left(\frac{\varepsilon}{8}\right)^{2t+1} n \geq 4^t \geq R(t)$$

and so, every t-tuple D in T has at least $\beta n = R(t)$ common neighbors in each color, and because $|T| \geq R(t)$, there is a monochromatic t-clique D in T. If the color of this t-clique is red, we know that $|N_B(D)| \geq \beta n = R(t)$ so there is a monochromatic t-clique X in $N_B(D)$. This gives us an element from $\mathcal{F}_t$ because the t-clique D is red, the edges between D and X are blue, and X is monochromatic. □

As mentioned before, the bound in Theorem 1.21 is tight up to the constant factor in the exponent. We omit the proof of this result.

Proposition 1.24 (Fox, Sudakov 2008 [35])**.** *If $\varepsilon \leq 1/2$ and $k \geq 2$, then $n(k,\varepsilon) > \varepsilon^{-(k-1)/2}$.*

1.4 Unavoidable patterns: Turán-type approach

In 2019, without awareness of Bollobás conjecture, which me mentioned at the beginning of Section 1.3, Yair Caro, Adriana Hansberg and Amanda Montejano considered maximizing the number of edges in the smallest color class in any 2-edge coloring of K_n which does not contain an element from $\mathcal{F}_t$, giving the problem a Turán taste. They proved that, for all sufficiently large n and any positive integer t, there exists a positive integer $m = m(t)$ and a number $\varphi(n,t) = O(n^{2-\frac{1}{m}})$ such that any 2-edge coloring of K_n with at least $\varphi(n,t)$ edges in each color class contains an element from $\mathcal{F}_t$ [13]. Later, Girão and Narayanan [39] showed that $\varphi(n,t) = \Omega(n^{2-\frac{1}{t}})$ and, conditional on the Kővari-Sós-Turán conjecture for the Turán number for complete bipartite graphs, they showed that the bound is sharp up to the involved constants. In [6], a much simpler proof of this result is given using the structure of the proof by Caro, Hansberg and Montejano and the dependent random choice technique. Bowen, Hansberg, Montejano and Müyesser [6] extend the result to an arbitrary number of colors, although the bound is not tight.

Recall that the family $\mathcal{F}_t$, defined in Section 1.3, which contains all 2-edge colored copies of K_{2t} where either one color forms a K_t or one color forms two disjoint copies of K_t. In [14], the authors view the elements of the family $\mathcal{F}_t$ as monochromatic induced copies of $K_{t,t}$ or $S_{t,t}$ (the split graph, which is made of a set of s vertices that induce a clique, another set of t indepenent vertices and all possible st edges between both sets). They also give a more general definition of this family which is used in 1.26. Let s, t be positive integers with $s \leq t$ and let $\mathcal{F}'_{s,t} = \{K_{s,t}, S_{s,t}\}$. When $s = t$, we write $\mathcal{F}'_t$ instead of $\mathcal{F}'_{t,t}$. For a positive integer n and a family of graphs $\mathcal{F}$ not containing complete graphs, let $\mathrm{ex}_2(K_n, \mathcal{F})$ be the minimum integer m (if it exists) such that, every 2-edge coloring of K_n having more than m edges in each color contains a member of $\mathcal{F}$ as an induced monochromatic subgraph. If there is no such m, we set $\mathrm{ex}_2(K_n, \mathcal{F}) = \infty$.

We state and prove Theorem 1.26 using the following variant of Corollary 1.20.

Lemma 1.25. *For all $K, s \in \mathbb{N}$, there exists a constant $C = C(K,s)$ such that any graph with at least $Cn^{2-1/s}$ edges contains a set S of K vertices in which each subset $X \subseteq S$ with s vertices has a common neighborhood of size at least K.*

We also use the 2-color Ramsey number $R(k)$ and the bipartite Ramsey number $BR(t)$, which is the smallest integer for which every 2-edge coloring of $K_{n,n}$, with $n \geq BR(t)$, contains a monochromatic $K_{t,t}$.

Theorem 1.26 ([14])**.** *For every pair of positive integers s and t with $1 \leq s \leq t$, there is a constant $C = C(s,t)$ such that, if n is large enough, then every 2-edge coloring $E(K_n) = R \sqcup B$ with at least $Cn^{2-\frac{1}{s}}$ edges in each color contains an induced monochromatic member from $\mathcal{F}'_{s,t}$.*

Proof. Given s and t, positive integers with $s \leq t$, let C be the constant given by 1.25 for s and $K = R(BR(t)) + 1$. For every n large enough, any 2-edge coloring of K_n with $Cn^{2-1/s}$ edges in each color, we can find K-sets of vertices S_R and S_B with the property that each s-subset

$X \subset S_R$ (respectively, each s-subset $X \subseteq S_B$) has a common red-neighborhood (respectively, a common blue neighborhood) of size at least K. Since $K = R(BR(t)) \geq R(s) + 1$, there are sets of vertices $X_1 \subset S_R$ and $X_2 \subset S_B$ inducing both monochromatic complete graphs of order s, and satisfying that $|N_R(X_1)| \geq K$ and $|N_B(X_2)| \geq K$. Moreover, there are sets $Y_1 \subset N_R(X_1)$ and $Y_2 \subset N_B(X_2)$ inducing monochromatic complete graphs of order $BR(t)$. In summary, we have both a red and a blue complete bipartite graph $K_{s,BR(t)}$ where both parts of the bipartition are monochromatic cliques. It remains to analyze all cases concerning the colors of those cliques. If Y_1 is blue, depending on the color of X_1, we get either an induced red copy of $K_{s,t}$, or an induced red copy of $S_{s,t}$. Thus, we can assume that Y_1 is red and, with similar arguments, that Y_2 is blue. Recall now that both Y_1 and Y_2 have $BR(t) + 1$ vertices and, because they are monochromatic cliques of different colors, they can intersect in at most one vertex. Thus, neglecting the intersection, if necessary, we can find a monochromatic copy of $K_{t,t}$ between Y_1 and Y_2. That is, there are t-sets of vertices Z_1 and Z_2 contained in Y_1 and Y_2 respectively, that induce a monochromatic complete bipartite graph. If such complete bipartite graph is blue, take any subset of s vertices in Z_2 to get an induced blue copy of $S_{s,t}$. In the other case, just take a subset of vertices in Z_1 to have an induced red copy of $S_{s,t}$. The proof is complete. $\square$

Theorem 1.26 proves that $\mathrm{ex}_2(K_n, \mathcal{F}'_{s,t}) \leq Cn^{2-\frac{1}{s}}$ where c is the constant in Theorem 1.26. The next theorem states that the above bound is tight for infinite values of s and t. Its proof relies strongly on Theorem 1.17.

Theorem 1.27 ([14])**.** *For every n sufficiently large, and integers s, t such that either $1 \leq s \leq 3 \leq t$, or $s \geq 3$ and $t \geq (s-1)! + 1$, we have $\mathrm{ex}_2(K_n, \mathcal{F}'_{s,t}) = \Theta(n^{2-1/s})$.*

In [14], the authors note that taking $s = 1$ in Theorem 1.26, i.e. considering a linear amount (on n) of edges in each color, forces the existence of an induced monochromatic star $K_{1,t}$. The following is an immediate corollary of Theorem 1.26.

Theorem 1.28 ([14])**.** *For every positive integer t, there is a constant $C = C(t)$ such that, if n is large enough, then every 2-edge coloring $E(K_n) = R \sqcup B$ with at least Cn edges in each color contains an induced monochromatic star $K_{1,t}$.*

Finally, they worked with edge colorings that required only a constant amount of edges in each color.

Theorem 1.29 ([14])**.** *For any integer $t \geq 1$, there is a constant $C = C(t)$ such that, for n sufficiently large, every 2-edge coloring of K_n with at least C edges in each color and without a monochromatic induced tK_2 contains an induced monochromatic star $K_{1,t}$.*

1.4.1 Omnitonality

Omnitonal graphs are those graphs that appear in all possible tonal variations of two colors in every 2-edge coloring of a large enough complete graph with sufficient edges in each color. Formally, we state the following definition.

Definition 1.30. *For a given graph G, let* $\mathrm{ot}(n,G)$ *be the minimum integer, if it exists, such that any* 2*-edge coloring* $E(K_n) = R \sqcup B$ *with more than* $\mathrm{ot}(n,G)$ *edges in each color contains an* (r,b)*-colored copy of* G *for any* $r \geq 0$ *and* $b \geq 0$ *such that* $r+b=e(G)$*. If* $\mathrm{ot}(n,G)$ *exists for every sufficiently large* n*, we say that* G *is* omnitonal.

This concept was introduced in [13] where the authors were interested in finding omnitonal graphs and determining or approximating their number $\mathrm{ot}(n,G)$, if possible. To prove that a graph G is not omnitonal, one has to exhibit infinitely many values of n for which there is a 2-edge coloring of K_n with the same number of red and blue edges (or differing by one) without an (r,b)-colored copy of G for some $r \geq 0$ and $b \geq 0$ such that $r+b=e(G)$.

Recall that Ramsey's theorem guarantees that, for a large enough n and a given graph G, there is a $(0,e(G))$-colored copy or a $(e(G),0)$-colored copy of G in every 2-edge coloring of K_n. To force the existence of other patterns, there must be enough edges in each color. The following definition provides the parameter we need to assure the existence of copies of G in other tonal proportions.

Definition 1.31. *Let* G *be a graph and* r *an integer with* $0 < r \leq \lfloor e(G)/2 \rfloor$*. Let* $\mathrm{bal}_r(n,G)$ *be the minimum integer, if it exists, such that every* 2*-edge coloring* $E(K_n) = R \sqcup B$ *with more than* $\mathrm{bal}_r(n,G)$ *edges in each color contains either an* $(r, e(G)-r)$*-colored copy of* G*, or an* $(e(G)-r, r)$*-colored copy of* G*. If* $\mathrm{bal}_r(n,G)$ *exists for every* n *sufficiently large, we say that* G *is* r-tonal.

Note that if $\mathrm{bal}_r(n,G)$ exists, then $\mathrm{bal}_r(n,G) \leq \frac{1}{2}\binom{n}{2}$. This also holds for $\mathrm{ot}(n,G) \leq \frac{1}{2}\binom{n}{2}$. In [13], Theorem 1.26 is used to give a characterization of r-tonal graphs and omnitonal graphs. They use the following lemma that follows directly from Lemmas 3.1 and 3.2 given in [10].

Lemma 1.32 ([10]). *For infinitely many positive integers* n*, we can choose* $t = t(n)$ *in a way that the* 2*-edge coloring of* K_n *where one color forms a* K_t *has exactly* $\lfloor \frac{1}{2}\binom{n}{2} \rfloor$ *edges of one color. Also, for infinitely many positive integers* n*, we can choose* $t = t(n)$ *in a way that the* 2*-edge coloring of* K_n *where one color forms two disjoint* K_t*'s has exactly* $\lfloor \frac{1}{2}\binom{n}{2} \rfloor$ *edges of one color.*

Theorem 1.33 ([13]). *Let* G *be a graph and let* r *be an integer with* $0 < r \leq \lfloor e(G)/2 \rfloor$*. The graph* G *is* r*-tonal if and only if* G *has both a partition* $V(G) = X \sqcup Y$ *and a set of vertices* $W \subseteq V(G)$ *such that* $e(X,Y), e(G[W]) \in \{r, e(G)-r\}$*.*

Lemma 1.32 and Theorem 1.33 are strongly used in the proof of the characterization of omnitonal graphs which is stated as follows.

Theorem 1.34 ([13]). *A graph* G *is omnitonal if and only if, for every integer* r *with* $0 \leq r \leq e(G)$*,* G *has both a partition* $V(G) = X \sqcup Y$ *and a set of vertices* $W \subseteq V(G)$ *such that* $e(X,Y) = e(G[W]) = r$*.*

In [13], the authors conclude that omnitonal graphs are bipartite, which is a consequence of Theorem 1.26. A result relevant to this thesis work states that all trees are omnitonal graphs. In Chapter 4, we present a corresponding theorem in the setting of 2-colorings of a complete bipartite graph $K_{n,n}$ (see Theorem 4.10).

Theorem 1.35 ([13])**.** *Every tree is omnitonal.*

Proof. Let T be a tree. We employ the characterization given by 1.34, therefore we verify that for every integer r with $0 \leq r \leq e(T)$, T has both a partition $V(T) = X \sqcup Y$ and a set of vertices $W \subseteq V(T)$ such that $e(X,Y) = e(G[W]) = r$. We use induction on $e(T)$. If $e(T) = 1$, both conditions are satisfied for $r = 0, 1$. Let T be a tree with $e(T) = m$ and let $v \in V(T)$ be a leaf where u is the only vertex of T adjacent to v. By the induction hypothesis, the tree $T' = T - \{v\}$ satisfies that for every $0 \leq r \leq m-1 = e(T')$, there is a partition $V(T') = X' \sqcup Y'$ and a set of vertices $W \subseteq V(T') \subseteq V(T)$ such that $e(X',Y') = e(T'[W]) = r$. Note that for every $0 \leq r \leq m-1$ the subset $W \subseteq V(T') \subseteq V(T)$ satisfies $e(T[W]) = r$. In the same way, for every $0 \leq r \leq m-1$ we can obtain a partition $V(T) = X \sqcup Y$ with $e(X,Y) = r$ by taking $X = X' \cup \{v\}$ and $Y = Y'$ if $u \in X'$, or $X = X'$ and $Y = Y' \cup \{v\}$ if $u \in Y'$. To show that there are both a partition $V(T) = X \sqcup Y$ and a set of vertices $W \subseteq V(T)$ such that $e(X,Y) = e(T[W]) = m = e(T)$ is trivial. □

It is a simple matter to see that the disjoint union of two omnitonal graphs is again an omnitonal graph. Hence, we can conclude with Theorem 1.35.

1.4.2 Balanceability

The concept of balanceability was introduced in [13]. In an informal setting, we ask the following question.

How many blue and red edges must there be in a 2-edge coloring of a large complete graph in order to assure a copy of a fixed graph H with exactly half of its edges colored blue and the rest red?

Caro, Hansberg and Montejano considered and studied this problem in [13].

The existence of a balanced copy of a graph G in any 2-edge coloring, even having as many edges as possible in each color, is not true in general for any graph G. Evidently, we need at least $\lfloor \frac{e(G)}{2} \rfloor$ edges in each color class. The following concept will help us characterize the graphs of our interest.

Definition 1.36. *If there exists an integer $k = k(n)$ such that, for n large enough, every 2-edge coloring $R \sqcup B$ of the edges of K_n with more than k edges in each color class contains a balanced copy of G, then we say G is* balanceable*. The smallest such k is called the* balancing number of G *and it is denoted as bal(n, G). For a balanceable graph G, let* $\mathrm{Bal}(n, G)$ *be the family of graphs with exactly* $\mathrm{bal}(n, G)$ *edges such that a 2-edge coloring $E(K_n) = R \sqcup B$ with exactly* $\mathrm{bal}(n, G)$ *edges in one color contains no balanced copy of G if and only if the graph induced by the red edges or by the blue edges is isomorphic to some $H \in$* $\mathrm{Bal}(n, G)$*.*

The two base questions in balanceability are:

i) Given a graph G, is G balanceable? (Does said k exist?)

ii) In the case of an affirmative answer to the first question, how small can k be? In other words, how big can k be such that we can explicitly show an edge coloring without balanced copies of G?

A very simple example of a balanceable graph is the path on 2 edges, P_2. Clearly, if we use at least one edge of each color in a 2-edge coloring of K_n with n large enough, the balanced path will always be obtained. There are graphs for which it is easy to determine if they are balanceable or not, but this is not the case in general. Observe that all omnitonal graphs are balanceable. A structural characterization [13] of balanceable graphs was given by Caro, Hansberg and Montejano.

Theorem 1.37 ([13]). *A graph $G(V,E)$ is balanceable if and only if G has a partition of its vertices $V = X \sqcup Y$ and a set of vertices $W \subseteq V$ such that $e(X,Y), e(G[W]) \in \{\lfloor \frac{1}{2}e(G) \rfloor, \lceil \frac{1}{2}e(G) \rceil\}$.*

In other words, a graph is balanceable if and only if there is an edge-cut of $\lfloor \frac{1}{2}e(G) \rfloor$ or $\lceil \frac{1}{2}e(G) \rceil$ edges and an induced subgraph with $\lfloor \frac{1}{2}e(G) \rfloor$ or $\lceil \frac{1}{2}e(G) \rceil$ of the edges. This theorem was proved by showing that, for every integer t and n sufficiently large, there exists an integer $m = m(n,t)$ such that any 2-coloring of K_n with more than m edges in each color class contains an element of $\mathcal{F}_t$. These copies can be used to find a balanced copy of a graph G or else to show that no balanced copy of G exists.

Beyond the computational question of deciding whether a given graph is balanceable or not, there is also the theoretical problem of providing exact values or good bounds for the Turán-type parameter $\mathrm{bal}(n,G)$. Observe that $\mathrm{bal}(n,G)$ can also be seen as the maximum number $k(n)$ such that there is a 2-edge coloring of K_n with one of the colors having precisely $k(n)$ edges, and such that there is no balanced copy of G. The family of such colorings would be the extremal configurations of this Turán-type parameter. The problem of determining the balancing number of certain families of graphs has been tackled in [8, 12, 13, 20].

1.5 Balanceable and omnitonal graphs

In this section, we discuss some graph families that are balanceable, omnitonal or both. We also determine or bound the numbers $\mathrm{bal}(n,G)$ and $\mathrm{ot}(n,G)$ of certain graph families. Some of the results stated in this section will be recalled and discussed in further chapters. We prove only those results which are relevant to our work.

1.5.1 Amoebas

As part of the work in [13], the authors describe a class of graphs which they called amoebas. They grew interest in this family because they are balanceable and contain a large family of omnitonal graphs. The definition they give in [13] for global amoebas evolved into a more algebraic definition in a later work [12] which we discuss profoundly in Chapter 3. In fact, the authors referred to these global amoebas, just as amoebas in [13]. This is due to the fact that amoebas were categorized as global amoebas and local amoebas until later in [12]. Because the results in this section can be argued in a purely combinatorial setting, we keep the first definition of amoebas and discuss the algebraic approach in Chapter 3.

Given a graph G of order $n(G)$ embedded in a complete graph K_n, where $n \geq n(G)$, we say that H (also embedded in K_n) is obtained from G by a *feasible edge-replacement*, if for some $e_1 \in E(G)$ and $e_2 \in E(K_n) \setminus E(G)$, $E(H) = (E(G) \setminus \{e_1\}) \cup \{e_2\}$. We consider only graphs with no isolated vertices.

Definition 1.38. *A graph H is a* local amoeba *if for any two isomorphic copies, H' and H'', of H on the same vertex set $V = V(H)$, there is a chain $H' = H_1, H_2, \cdots, H_r = H''$, such that for every $1 \leq i \leq r$, $H_i \cong H$ and H_i is obtained from H_{i-1} by a feasible edge-replacement. A graph G is a* global amoeba *if there is $t_0 \geq 0$ such that $G \cup tK_1$ is a local amoeba for every $t \geq t_0$*

As a simple example, one can convince oneself (even though it is not that simple to prove) that the path P_k is a global amoeba for every $k \geq 1$, and that a cycle C_k with $k \geq 3$ is not a global amoeba.

The following proposition provides an important property of local amoebas with minimum degree 0 or 1.

Proposition 1.39 ([12])**.** *Let G be a local amoeba of order n and minimum degree 0 or 1, then $G \cup K_1$ is a local amoeba, and therefore G is a global amoeba.*

Two important aspects of global amoebas is that all bipartite global amoebas are omnitonal and that all global amoebas are balanceable. We state and prove both statements.

We state a basic interpolation lemma for amoebas and a remark before proving the main statement.

Lemma 1.40 ([13])**.** *Let G be a global amoeba and consider a 2-edge coloring $E(K_n) = R \sqcup B$ where $n \geq n_0(G)$. Let $\alpha, \beta, \alpha', \beta'$ be integers such that $\alpha + \beta = \alpha' + \beta' = e(G)$ and $0 \leq \alpha \leq \alpha'$ and $0 \leq \beta \leq \beta'$. If there is an (α, β)-colored copy of G as well as an (α', β')-colored copy of G, then there is an (r, b)-colored copy of G for all integers r and b such that $r + b = e(G)$, $\alpha \leq r \leq \alpha'$ and $\beta' \leq b \leq \beta$.*

Remark 1.41. *From the Kővari-Sós-Turán theorem, stated in Theorem 1.10, it follows that $\mathrm{ex}(n, G) = o(n^2)$ for any bipartite graph G. This implies that, for large enough n, the inequality*

$$2(\mathrm{ex}(n, G) + 1) \leq \binom{n}{2}$$

holds. This means that we can consider 2-edge colorings $E(K_n) = R \sqcup B$ with at least $\mathrm{ex}(n, G) + 1$ edges in each color given that n is large enough.

Lemma 1.40 states that, for a given global amoeba G on n_0 vertices and a given a 2-edge coloring $E(K_n) = R \sqcup B$, where $n \geq n_0$, if there is a $(0, e(G))$-colored copy and a $(e(G), 0)$-colored copy of G, then the colored K_n contains the a copy of G in every possible (r, b)-color pattern for r and b with $r + b = e(G)$, with $0 \leq r \leq e(G)$ and $0 \leq b \leq e(G)$. Using Lemma 1.40 and Remark 1.41, we can prove the following theorem.

Theorem 1.42 ([13]). *Every bipartite global amoeba G is omnitonal with $\mathrm{ot}(n, G) = \mathrm{ex}(n, G)$, provided n is large enough to fulfill $\binom{n}{2} \geq 2\,\mathrm{ex}(n, G) + 1$ and $n \geq n_0(G)$.*

Proof. Let G be a bipartite global amoeba. By Remark 1.41 we can consider, for sufficiently large n, 2-edge colorings of $E(K_n)$ with $n \geq n_0(G)$ and at least $\mathrm{ex}(n, G) + 1$ edges in each color. Notice that any 2-edge coloring $E(K_n) = R \sqcup B$ with at least $\mathrm{ex}(n, G) + 1$ edges in each color contains a $(0, e(G))$-colored copy of G and an $(e(G), 0)$-colored copy of G, and by 1.40, there is an (r, b)-colored copy of G for all integers r and b such that $0 \leq r,\ b \leq e(G)$ and $r + b = e(G)$. Therefore, G is omnitonal and $\mathrm{ot}(n, G) \leq \mathrm{ex}(n, G)$. In order to see that $\mathrm{ex}(n, G) \leq \mathrm{ot}(n, G)$, notice that we can give a 2-edge coloring of $E(K_n)$ with at least $\mathrm{ex}(n, G)$ edges in each color such that there are no $(e(G), 0)$-colored copies of G and so G cannot be omnitonal. □

Because the balanceable property is not as restrictive as the omnitonal property, Caro, Hansberg and Montejano showed that the bipartite condition can be dropped to prove that every global amoeba is balanceable.

Theorem 1.43 ([13]). *Every global amoeba is balanceable.*

Proof. Let G be a global amoeba. We make use of an old argument of Erdős which states that every graph G has a bipartition $V(G) = X \sqcup Y$ such that $e(X, Y) \geq \lceil \frac{e(G)}{2} \rceil$ (see Lemma 2.14 in [38]). Therefore, we may consider a bipartite subgraph B of G with exactly $e(B) = \lceil \frac{e(G)}{2} \rceil$ edges. Let $E(K_n) = R \sqcup B$ be a 2-edge coloring with at least $\mathrm{ex}(n, B) + 1$ edges in each color. This is possible for n large enough by Remark 1.41. Therefore, K_n contains a $(0, e(B))$-colored copy of B and a $(e(B), 0)$-colored copy of B. We proceed to arbitrarily complete these copies of B into copies of G to get an (α, β)-colored copy of G, and an (α', β')-colored copy of G with $\lceil e(G)/2 \rceil \leq \beta$ and $\lceil e(G)/2 \rceil \leq \alpha'$. Since $\alpha + \beta = \alpha' + \beta' = e(G)$ this implies that $\alpha \leq \lfloor e(G)/2 \rfloor$ and $\beta' \leq \lfloor e(G)/2 \rfloor$. Bringing the bounds together, we get $\alpha \leq \lfloor e(G)/2 \rfloor \leq \lceil e(G)/2 \rceil \leq \alpha'$ and $\beta' \leq \lfloor e(G)/2 \rfloor \leq \lceil e(G)/2 \rceil \leq \beta$. By Lemma 1.40, we conclude that K_n contains a $(\lfloor e(G)/2 \rfloor, \lceil e(G)/2 \rceil)$-copy and a $(\lceil e(G)/2 \rceil, \lfloor e(G)/2 \rfloor)$-copy of G. □

1.5.2 Stars and Paths

We know from Theorem 1.35 that trees are omnitonal and therefore balanceable. We describe the work in [13] where the authors determine $\mathrm{bal}(n, G)$ and $\mathrm{Bal}(n, G)$ for the cases where G is a star or a path with an even number of edges. For the general case when G is a tree, the best upper bound they obtained emerged as a corollary from the bound for $\mathrm{ot}(n, G)$. Recall that a split graph $S_{s,t}$, or the (s, t)-split graph, is made of a set of s vertices that induce a clique, another set of t indepenent vertices and all possible st edges between both sets.

Theorem 1.44 ([13]). *Let k and n be integers with $k \geq 2$ even and such that $n \geq \max\{3, \frac{k^2}{4}+1\}$. Then $\mathrm{bal}(n, K_{1,k}) = (\frac{k-2}{2})n - \frac{k^2}{8} + \frac{k}{4}$ and $\mathrm{Bal}(n, K_{1,k})$ contains only one graph, namely the complete $(\frac{k-2}{2}, n - \frac{k-2}{2})$-split graph.*

Proof. Let

$$h(n,k) = \left(\frac{k-2}{2}\right) n - \frac{k^2}{8} + \frac{k}{4}.$$

First observe that a 2-edge coloring $E(K_n) = R \sqcup B$ with more than $h(n,k)$ edges in each color is satisfiable. We need to prove that $e(K_n) = \frac{n(n-1)}{2} \geq 2h(n,k) + 2$ holds true for all $n \geq \max\{3, \frac{k^2}{4} + 1\}$. If $k = 2$, then $h(n,k) = 0$ and the condition is satisfied for $n \geq 3$. If $k \geq 4$, we have to verify the mentioned inequality. Notice that $2h(n,k) + 2 = n(k-2) - \frac{k^2}{4} + \frac{k}{2} + 2 \leq \frac{n(n-1)}{2}$ if and only if $n^2 - (2k-3)n + \frac{k^2}{2} - k - 4 \geq 0$ which is indeed the case for $n \geq \frac{k^2}{4} + 1$ and $k \geq 4$.

Let H be the complete $(\frac{k-2}{2}, n - \frac{k-2}{2})$-split graph. Note that H has exactly $h(n,k)$ edges:

$$\begin{aligned}
e(H) &= \frac{1}{2}\left(\frac{k-2}{2}\right)\left(\frac{k-2}{2} - 1\right) + \left(\frac{k-2}{2}\right)\left(n - \frac{k-2}{2}\right) \\
&= \left(\frac{k-2}{2}\right) n + \left(\frac{k-2}{2}\right)\left(\frac{k-2}{4} - \frac{1}{2} - \frac{k-2}{2}\right) \\
&= \left(\frac{k-2}{2}\right) n + \left(\frac{k}{2} - 1\right)\left(\frac{k}{4} - \frac{k}{2}\right) \\
&= \left(\frac{k-2}{2}\right) n - \frac{k^2}{8} + \frac{k}{4}.
\end{aligned}$$

Observe that any 2-edge coloring $E(K_n) = R \sqcup B$ where the red graph or the blue graph is isomorphic to H contains no balanced copy of $K_{1,k}$, because in this coloring there are two types of vertices $v \in V(K_n)$, the ones for which $\{deg_R(v), deg_B(v)\} = \{0, n-1\}$ and the ones for which $\{deg_R(v), deg_B(v)\} = \{\frac{k}{2} - 1, n - \frac{k}{2}\}$. In any case, there is no way to have a balanced $K_{1,k}$.

We have proved that $\mathrm{bal}(n, K_{1,k}) \geq h(n,k)$ and that $H \in \mathrm{Bal}(n, K_{1,k})$. To prove that $\mathrm{bal}(n, K_{1,k}) \leq h(n,k)$ and that $\mathrm{Bal}(n, K_{1,k}) = \{H\}$ we will show that any coloring $E(K_n) = R \sqcup B$ with at least $h(n,k)$ edges in each color and where neither the red nor the blue graph is isomorphic to H contains a balanced copy of $K_{1,k}$. We define the following sets

$$V_R = \left\{ v \in V(K_n) \mid deg_R(v) \geq \frac{k}{2} \right\},$$

and

$$V_B = \left\{ v \in V(K_n) \mid deg_B(v) \geq \frac{k}{2} \right\}.$$

Let $E(K_n) = R \sqcup B$ be a 2-edge coloring with at least $h(n,k)$ edges in each color and such that the red graph nor the blue graph is isomorphic to H. If there is a vertex $v \in V_R \cap V_B$ then we are done as there would be a balanced $K_{1,k}$. So we may assume that $V_R \cap V_B = \emptyset$. Note that, since every vertex in K_n has degree $n-1 \geq k$ then $V(K_n) = V_R \sqcup V_B$, hence $|V_R| + |V_B| = n$. Assume without loss of generality that $|V_R| \leq |V_B|$.

Case 1: Suppose $|V_R| \leq \frac{k}{2} - 1$. Thus,

$$\begin{aligned} 2e(G) = \sum_{v \in V(K_n)} deg_R(v) &= \sum_{v \in V_R} deg_R(v) + \sum_{v \in V_B} deg_R(v) \\ &\leq |V_R|(n-1) + |V_B|(\frac{k-2}{2}) \\ &= |V_R|(n-1) + (n - |V_R|)(\frac{k-2}{2}) \\ &\leq (\frac{k}{2} - 1)(n-1) + (n - \frac{k}{2} + 1)(\frac{k-2}{2}) \\ &= 2(\frac{k-2}{2})n - \frac{k^2}{4} + \frac{k}{2} = 2h(n,k) \end{aligned}$$

Consequently, the number of red edges is at most $h(n,k)$. By the assumption on the number of edges in each color we know that the number of red edges must be exactly $h(n,k)$. By the inequalities before mentioned $|V_R| = (\frac{k}{2} - 1)$ and the red graph is isomorphic to H, which is a contradiction to the original assumption.

Case 2: Suppose now that $|V_R| \geq \frac{k}{2}$. Denote by $e'(R)$ the number of red edges between V_R and V_B. Since a vertex $v \in V_R$ satisfies $deg_B(v) < \frac{k}{2}$ then each vertex in V_R contributes to $e'(R)$ with at least $|V_B| - \frac{k}{2} + 1$ edges, thus

$$e'(R) \geq |V_R|(|V_B| - \frac{k}{2} + 1) \geq \frac{k}{2}(|V_B| - \frac{k}{2} + 1).$$

Now note that each vertex in V_B contributes to $e'(R)$ with no more than $\frac{k}{2}$ edges, and so

$$e'(R) \leq (\frac{k}{2} - 1)|V_B|.$$

Joining both inequalities, we obtain

$$\frac{k}{2}(|V_B| - \frac{k}{2} + 1) \leq (\frac{k}{2} - 1)|V_B|,$$

from which by the fact that $|V_R| + |V_B| = n$ and that $|V_R| \geq \frac{k}{2}$, it follows that

$$-\frac{k^2}{4}+\frac{k}{2}\leq -|V_B| = |V_R| - n \leq \frac{k}{2} - n.$$

This yields $n \leq \frac{k^2}{4}$, a contradiction to the hypothesis.

□

The balancing number of paths $\mathrm{bal}(n, P_k)$ and $\mathrm{Bal}(n, P_k)$ has also been determined by the authors of [13] for $k \geq 2$ even and n sufficiently large.

Theorem 1.45 ([13])**.** *Let $k \geq 2$ and n be integers with k even and such that $n \geq \frac{9}{32}k^2+\frac{1}{4}k+1$. Then*

$$\mathrm{bal}(n, P_k) = \begin{cases} (\frac{k-2}{4})n - \frac{k^2}{32} + \frac{1}{8}, & \textit{for } k \equiv 2 \pmod 4, \\ (\frac{k-4}{4})n - \frac{k^2}{32} + \frac{k}{8} + 1, & \textit{for } k \equiv 0 \pmod 4, \end{cases}$$

and $\mathrm{Bal}(n, P_k)$ *contains only one graph, namely the complete* $(\frac{k-2}{4}, n - \frac{k-2}{4})$*-split graph, if* $k \equiv 2 \pmod 4$*, and the complete* $(\frac{k-4}{4}, n - \frac{k-4}{4})$*-split graph plus one edge, if* $k \equiv 0 \pmod 4$*.*

1.6 Generalized balancing number

Although the question of the existence of the balancing number is still open for many graph classes, we are also interested in gauging how we may obtain balanceable copies of a non-balanceable graph, under a relaxation of the 2-edge coloring. In this work, we also generalize the notion of balancing number by extending the class of colorings under consideration to edge coverings. In this case, each edge receives a nonempty list of colors, and we may choose one among them as needed in order to construct a balanced copy of a graph G.

More formally, a 2-*edge covering* of K_n is a function $L : E(K_n) \to \{\{r\}, \{b\}, \{r, b\}\}$, that induces two sets R and B, called its *color classes*, which are defined as follows: $R = \{e \in E(K_n) \mid r \in L(e)\}$ and $B = \{e \in E(K_n) \mid b \in L(e)\}$. As we can see, $R \cup B = E(K_n)$, but the two color classes do not necessarily form a partition of the edges of K_n. The edges in $R \cap B$ are called *bicolored edges*, in the sense that we can choose their color as needed when looking for a balanced copy of a graph. This leads to a new definition of a balanced copy of a graph, which is a generalization of the previous one:

Definition 1.46. *Let $L : E(K_n) \to \{\{r\}, \{b\}, \{r, b\}\}$ be a 2-edge covering of K_n inducing color classes R and B. For a given graph $G(V, E)$, a balanced copy of G is a copy of G whose edge-set has a partition $E = E_1 \sqcup E_2$ such that $E_1 \subseteq R$, $E_2 \subseteq B$ and $||E_1| - |E_2|| \leq 1$.*

Note that some of the edges in the copy of G may be in $R \cap B$; for example, if $e \in E_1$ and $e \in R \cap B$ then we say that we *choose* color r for the bicolored edge e, if we want the color red to be counted for this edge.

Furthermore, for every finite graph $G(V, E)$, if $n \geq |V(G)|$, then K_n has a 2-edge covering where we may find a balanced copy of G. Indeed, consider the 2-edge covering L of K_n such that $L(e) = \{r, b\}$ for every edge e. Observe that, in any copy of G, we may choose the color r for half its edges and the color b for the rest. This allows us to define the generalized balancing number of a graph, as follows:

Definition 1.47. *Let $G(V, E)$ be a finite, simple graph. For $n \geq |V(G)|$, the* generalized balancing number of G*, denoted by* $\mathrm{bal}^*(n, G)$*, is the smallest integer k such that every 2-edge covering L of K_n inducing the color classes R and B with $|R|, |B| > k$ contains a balanced copy of G.*

As we can see, the generalized balancing number is a natural extension of the balancing number. Indeed, every 2-coloring, represented by a partition $E(K_n) = R \sqcup B$, corresponds to a 2-covering where every list $L(e)$ has exactly one element. It is important to observe that this extension does not add more complexity to the problem when $\mathrm{bal}(n, G)$ exists, and in which case satisfies $\mathrm{bal}(n, G) = \mathrm{bal}^*(n, G) < \frac{1}{2}\binom{n}{2}$ (for details, see Proposition 2.26 in Section 2.5).

The main interest of the generalized balancing number is the study of non-balanceable graphs, where we interpret having $\mathrm{bal}^*(n, G)$ close to $\frac{1}{2}\binom{n}{2}$ as G being close to balanceable, in the sense that a few more than $\frac{1}{2}\binom{n}{2}$ edges from each color (implying that there must be a few bicolored edges) are sufficient to guarantee a balanced copy of G. We refer to the needed number of edges that exceeds $\frac{1}{2}\binom{n}{2}$ in each color as the *color excess* of edges in each color. For example, we prove in Section 2.5.1, that, for cycles C_{4k+2} of length congruent to 2 modulo 4, which are not balanceable, a color excess of just 1 edge in each color is sufficient to guarantee a balanced copy of C_{4k+2}. In other words, we show that $\mathrm{bal}^*(n, C_{4k+2}) = \frac{1}{2}\binom{n}{2}$, which is the smallest possible value for the generalized balancing number of a non-balanceable graph. On the other hand, there are graphs for which a much larger color excess in each color is necessary to guarantee a balanced copy of them. Such is the case of K_5 where the color excess is of order $\Theta(n^{\frac{3}{2}})$; see Theorem 2.37.

In Chapter 2, we identify the balanceability of some graph classes, find the balancing number of balanceable cycles and define the generalized balancing number as an extension of the balancing number. In Chapter 3, we also work with a special type of balanceable graphs called amoebas and provide a way to recursively generate them using group theory tools. In Chapter 4, we seek to translate these problems to a $K_{n,n}$ setting where we use the complete bipartite graph as our base graph. The final chapter of this dissertation work is Chapter 5, where we state a conclusion statement and mention our ambition to continue working on the solution of other relevant problems as a project for future work in Section 5.1.

Chapter 2

Balanceability and balancing number in a K_n setting

2.1 Balanceability

The problem of determining whether a graph is balanceable or not has been studied by several authors [10, 13–15]. In this section, we discuss some results concerning balanceability and the graph families whose balanceability or non-balanceability has been established up to date. Then we define the balancing number and generalized balancing number in order to establish a connection between them. Recall Bollobás's patterns. Given a positive integer t, the first pattern consists of a complete graph on $2t$ vertices where one of the colors makes a complete graph on t vertices. The second pattern is also a complete graph on $2t$ vertices except that now one of the colors makes two disjoint complete graphs on t vertices each. In the Introduction, we noted that if there is a graph H which is a subgraph of both of Bollobás's patterns in a way that it is balanced, then this graph is balanceable. This turns out to be a characterization of balanceable graphs which we discuss in the next section.

2.1.1 Sufficient conditions for balanceability

Balanceability aims to answer the two important questions mentioned in Chapter 1. In this section, we will focus on answering the first one, namely: *Given a graph G, is it balanceable?*

In the process of studying the characterization from Theorem 1.37, various results arose about sufficient or necessary conditions that a graph can satisfy in order to be balanceable. We intend to expose them briefly in this section. The first result states that the existence of an independent set with just enough neighbors implies that the graph is indeed balanceable.

Proposition 2.1. *Let $G(V, E)$ be a graph. If there exists a subset $I \subset V$ of independent vertices such that $\sum_{v \in I} d(v) = \left\lfloor \frac{|E|}{2} \right\rfloor$, then G is balanceable.*

Proof. Let $X = I$ and $Y = W = V \setminus I$. Due to the condition on I, we have

$$e(X, Y) = \left\lfloor \frac{|E|}{2} \right\rfloor = e(G[W])$$

and thus, by Theorem 1.37, G is balanceable. □

Using this proposition, one can easily prove balanceability for many graphs. For example, it yields easily that cycles C_{4k} for $k \geq 1$ are balanceable.

Corollary 2.2. *Let ℓ be a positive integer. The cycle $C_{4\ell}$ is balanceable.*

Proof. We denote the vertices of $C_{4\ell}$ by $u_0, u_1, \ldots, u_{4\ell-1}$. By setting $I = \{u_{4i} \mid 1 \leq i \leq \ell - 1\}$, we can apply Proposition 2.1 and get the result. □

Furthermore, we can set I to be a singleton, in which case we obtain the following:

Proposition 2.3. *Let $G(V, E)$ be a graph. If there is a vertex $v \in V$ with $d(v) = \left\lfloor \frac{|E|}{2} \right\rfloor$, then G is balanceable.*

Proof. Let I be the singleton containing the vertex v with $d(v) = \left\lfloor \frac{|E|}{2} \right\rfloor$ and apply Proposition 2.1. □

In addition, this proposition proves that the wheel graphs are balanceable by taking the vertex of degree $n - 1$ as the only member of the independent set. Recall that the wheel graphs W_n of order n are made up of a cycle on $n - 1$ vertices and a vertex that is adjacent to all other vertices.

Corollary 2.4. *Wheels are balanceable.*

Proof. The wheel W_n contains $2n$ edges, and the center of W_n has degree n, hence Proposition 2.3 applies. □

Finally, we state the following property about the balanceability of some bipartite regular graphs:

Proposition 2.5. *Let d be a positive integer. If G is a bipartite, d-regular graph on $4n$ vertices, then G is balanceable.*

Proof. Let G be a bipartite, d-regular graph of order $4n$, and let $A \cup B$ be its bipartition. Since G is regular, $|A| = |B| = 2n$, and $|E(G)| = 2nd$.

Let I be a subset of n vertices in A. Clearly, I is an independent set, and by regularity $\sum_{v \in I} d(v) = nd = \frac{|E(G)|}{2}$. Thus, by Proposition 2.1, G is balanceable. □

2.2 A necessary condition for balanceability

We now focus on more global properties, such as regularity or whether the graph is eulerian or not, to state a necessary condition for balanceability. Afterwards, we expose some graph families that satisfy the condition. Recall that a connected graph is *eulerian* if all its vertices have even degree.

Proposition 2.6. *Let $G(V, E)$ be an eulerian graph with even number of edges. If G is balanceable, then $\frac{|E|}{2}$ is even.*

Proof. Let G be balanceable and assume by contradiction that $\frac{|E|}{2}$ is odd. Then, using Theorem 1.37, there is a partition of V in two sets X and $Y = V \setminus X$ such that half the edges are between X and Y. Let us denote by X_{odd} the set of vertices of X that have an odd number of neighbours in Y. The fact that $\frac{|E|}{2}$ is odd implies that $|X_{odd}|$ is odd. This in turn implies that the vertices of X_{odd} have odd degree in $G[X]$, since they have even degree in G. Thus, $G[X]$ has an odd number of vertices with odd degree, which is impossible. This contradiction yields the result. □

Even though being eulerian is not a common characteristic, they constitute a wide and interesting family of graphs that are characterized as those connected graphs whose vertices have all even degree. By means of the previous proposition, one can prove that C_{4k+2} is not balanceable for $k \geq 1$ and that K_n is also not balanceable if $n \equiv 5 \pmod 8$. While these propositions may seem simple, they played a predominant role in determining the balanceability of some infinite graph families.

Corollary 2.7. *Let ℓ be a positive integer. The cycle $C_{4\ell+2}$ is not balanceable.*

Proof. The cycle $C_{4\ell+2}$ is eulerian, and has $4\ell + 2$ edges, so $\frac{|E(C_{4\ell+2})|}{2} = 2\ell + 1$ is odd and we can apply Proposition 2.6. □

We also study the balanceability of some specific regular graphs. First, Proposition 2.6 allows us to prove that some regular graphs are not balanceable:

Corollary 2.8. *Let d be an even positive integer, and let G be a d-regular graph of order n with an even number of edges. G is not balanceable in the following cases:*

1. *If $d, n \equiv 2 \bmod 4$;*
2. *If $d = 4a$ with a odd, and $n \equiv 1, 3 \bmod 4$.*

Note that this is a sufficient condition for non-balanceability, so a regular graph that verifies neither conditions may still be non-balanceable.

Proof. If G is d-regular of order n and size e, then $e = \frac{dn}{2}$. Furthermore, since d is even, G is eulerian. We will study the two cases:

1. Given $d = 4a + 2$ and $n = 4b + 2$, then $\frac{e}{2} = \frac{(4a+2)(4b+2)}{4} = 4ab + 2a + 2b + 1$ is odd. Hence, Proposition 2.6 implies that G is not balanceable.

2. Given $d = 4a$ and $n = 4b + c$ with a odd and $c \in \{1, 3\}$, then, $\frac{e}{2} = \frac{4a(4b+c)}{4} = 4ab + ac$. Since $c \in \{1, 3\}$, this has the same parity than a, hence it is odd and Proposition 2.6 implies that G is not balanceable.

$\square$

Note that Corollary 2.8 easily implies that complete graphs such as K_{8k+5} are not balanceable. In [13], the authors determined which complete graphs with even edge number are balanceable. We state this result in the next section (Theorem 2.9).

In the next section, we will exhibit balanceability results of several graph families.

2.3 Interesting graph families

There are several graph families for which balanceability has been determined. Table 2.1 includes some of these graph families. Further results concerning balanceability will be discussed in this section in detail.

Graph family	**Balanceability**
K_4	balanceable [11]
K_5	not balanceable [10]
K_m with $m > 2$ and $m \neq 4$, $m \equiv 0, 1 \pmod 4$	not balanceable [11]
Trees	balanceable [13]
Complete bipartite graphs $K_{t,t}$	balanceable [8]
Specific circulant graphs $C_{k,\ell}$ with k even	balanceable [22]
Specific circulant graphs $C_{k,\ell}$ with k odd	not balanceable [22]
Rectangular grids $G_{k,\ell}$ with k and ℓ having the same parity	balanceable [22]
Triangular grids T_h with $h \pmod 8 \in \{0, 1\}$	balanceable [22]
Amoebas	balanceable [13]

Table 2.1: Known graph balanceability.

Caro, Hansberg and Montejano proved in [11] by means of number theory tools that the only nontrivial balanceable complete graph of even size (edge number) is K_4. We cite the equivalent theorem that appears in [10] because it employs more relevant notation and jargon.

Theorem 2.9 ([10]). *(i) For any positive integer $m \geq 2$, $m \neq 4$, $m \cong 0, 1 \pmod 4$ the complete graph K_m is not balanceable.*

(ii) The complete graph K_4 is balanceable with $\mathrm{bal}(n, K_4) = n$, *if* $n \cong 0 \pmod 4$, *and* $\mathrm{bal}(n, K_4) = n - 1$, *else. Moreover,* $\mathrm{Bal}(n, K_4) = \{J \cup \bigcup_{i=1}^{q} C_4\}$, *where* $J \in \{\emptyset, K_1, K_2, P_2\}$, *depending on the residue of* $n \pmod 4$, *and* $q = \lfloor \frac{n}{4} \rfloor$.

The balanceability of K_5 is easily determined by using Theorem 1.37.

Proposition 2.10 ([10])**.** *The complete graph on 5 vertices K_5 is not balanceable.*

Dailly, Hansberg, Talon and Ventura proved in [21] that the complete graphs on n vertices where $n \in \{2, 3, 7, 11, 14, 38, 62, 79, 359, 43262\}$ with an odd edge number are balanceable. The problem of determining if these are the only values of n (including 1 and 4) for which K_n is balanceable is still an open problem. We proceed to state our result and its computer-assisted proof.

Proposition 2.11. *Let n be a positive integer such that $n \equiv 2, 3 \bmod 4$ and $n \leq 2,303,999,904,000,003$. The graph K_n is balanceable if and only if $n \in \{2, 3, 7, 11, 14, 38, 62, 79, 359, 43262\}$.*

Proof. The proof is computer-assisted. By Theorem 1.37, in order for K_n to be balanceable, it must have a partition of its vertices $V(K_n) = X \sqcup Y$ and a set of vertices $W \subseteq V$ such that $e(X, Y), e(G[W]) \in \{\lfloor \frac{1}{2} e(G) \rfloor, \lceil \frac{1}{2} e(G) \rceil\}$. For this to happen K_n needs to verify two equations.

$$\frac{x(x-1)}{2} = \frac{n(n-1)}{4} \pm \frac{1}{2}$$

and

$$a(n-a) = \frac{n(n-1)}{4} \pm \frac{1}{2}$$

with a and x two positive integers.

The first equation implies $2x^2 - n^2 - 2x + n \pm 2 = 0$. This has two possible forms, which have solutions constructed recursively.

The first form is $2x^2 - n^2 - 2x + n - 2 = 0$, which has the following recursive systems for solutions:

$$\begin{cases} x_{k+1} = 3x_k + 2n_k - 2 \\ n_{k+1} = 4x_k + 3n_k - 3 \end{cases}$$

$$\begin{cases} x_{k+1} = 3x_k - 2n_k \\ n_{k+1} = -4x_k + 3n_k + 1 \end{cases}$$

Each of those systems can be started with those two pairs of initial values: $(x_0, n_0) = (2, 2)$ and $(x_0, n_0) = (-1, -1)$. Let S_x^- be the set of n's that are constructed from those systems.

The second form is $2x^2 - n^2 - 2x + n + 2 = 0$, which has the following recursive systems for solutions:

$$\begin{cases} x_{k+1} = 3x_k + 2n_k - 2 \\ n_{k+1} = 4x_k + 3n_k - 3 \end{cases}$$

$$\begin{cases} x_{k+1} = 3x_k - 2n_k \\ n_{k+1} = -4x_k + 3n_k + 1 \end{cases}$$

Each of those systems can be started with those four pairs of initial values: $(x_0, n_0) = (1, 2)$, $(x_0, n_0) = (0, 2)$, $(x_0, n_0) = (0, -1)$ and $(x_0, n_0) = (1, -1)$. Let S_x^+ be the set of n's that are constructed from those systems.

Let $S_x = S_x^- \cup S_x^+$. The set S_x is the set of all values of n that are solutions of the first equation.

The second equation implies $4a^2 - 4an + n^2 - n \pm 2$. This has two possible forms, which have explicit solutions.

The first form is $4a^2 - 4an + n^2 - n - 2$, which has the following explicit solutions, for every nonnegative integer k:

$$\begin{cases} a = 8k^2 + 2k - 1 \\ n = 16k^2 - 2 \end{cases}$$

$$\begin{cases} a = 8k^2 + 6k \\ n = 16k^2 + 8k - 1 \end{cases}$$

$$\begin{cases} a = 8k^2 + 10k + 2 \\ n = 16k^2 + 16k + 2 \end{cases}$$

$$\begin{cases} a = 8k^2 + 14k + 5 \\ n = 16k^2 + 24k + 7 \end{cases}$$

Let S_a^- be the set of n's that are obtained from those systems.

The second form is $4a^2 - 4an + n^2 - n + 2$, which has the following explicit solutions, for every nonnegative integer k:

$$\begin{cases} a = 8k^2 + 2k + 1 \\ n = 16k^2 + 2 \end{cases}$$

$$\begin{cases} a = 8k^2 + 6k + 2 \\ n = 16k^2 + 8k + 3 \end{cases}$$

$$\begin{cases} a = 8k^2 + 10k + 4 \\ n = 16k^2 + 16k + 6 \end{cases}$$

$$\begin{cases} a = 8k^2 + 14k + 7 \\ n = 16k^2 + 24k + 11 \end{cases}$$

Let S_a^+ be the set of n's that are obtained from those systems.

Let $S_a = S_a^- \cup S_a^+$. The set S_a is the set of all values of n that are solutions of the second equation.

Now, let $S = S_a \cap S_x$. The set S is the set of all values of n that are solutions of both equations. Hence, K_n is balanceable if and only if $n \in S$. Hence, we implemented a generic two integer variable equation solver program [31] to construct the solutions, which gave us the result in the statement of Proposition 2.11. □

We also worked on the balanceability of the graph made of two disjoint copies of K_n where n is any positive integer.

Theorem 2.12. *Let n be a positive integer. The graph $2K_n$ is balanceable if and only if there are two nonnegative integers a and b such that $n = a^2 + b^2$.*

The sum of two squares theorem by Dudley [61] implies the following:

Corollary 2.13. *Let n be a positive integer. The graph $2K_n$ is balanceable if and only the prime decomposition of n contains no factor p^k where prime $p \equiv 3 \bmod 4$ and k is odd.*

Proof of Theorem 2.12. We use Theorem 1.37. First, note that $2K_n$ trivially has an induced subgraph containing half of its edges: K_n. Hence, we are looking for the existence of a cut of $2K_n$ containing half of its edges. Such a cut can only be of the form, without loss of generality, $(K_k \cup K_\ell, K_{n-k} \cup K_{n-\ell})$, with $k, \ell \in \{0, \ldots, n\}$ and $k(n-k) + \ell(n-\ell) = \frac{n(n-1)}{2}$. This equation yields:

$$n = \frac{1}{2}\sqrt{8k\ell - 4k^2 - 4\ell^2 + 4k + 4\ell + 1} + k + \ell + \frac{1}{2}$$

Since n is a positive integer, the quantity under the square root has to be a square: there is an x such that $8k\ell - 4k^2 - 4\ell^2 + 4k + 4\ell + 1 = x^2$. Let us reformulate the equation, and by denoting $b = k - \ell$:

$$\begin{aligned} x^2 &= 8k\ell - 4k^2 - 4\ell^2 + 4k + 4\ell + 1 \\ &= -4(k-\ell)^2 + 4(k-\ell) + 1 \\ &= -4(k-\ell)^2 + 4(k-\ell) + 8\ell + 1 \\ &= -4b^2 + 4b + 8\ell + 1 \\ &= -(2b-1)^2 + 8\ell + 2 \end{aligned}$$

Hence, $8\ell + 2 = x^2 + (2b-1)^2$. Since $8\ell + 2$ is even, x and $2b-1$ have to be of the same parity, and thus they are both odd integers. Let us denote $x = 2a - 1$. Note that we have $\ell = \frac{(2a-1)^2+(2b-1)^2-2}{8}$. Thus, we have:

$$\begin{aligned} x^2 &= -(2b-1)^2 + 8\ell + 2 \\ &= -(2b-1)^2 + (2a-1)^2 + (2b-1)^2 - 2 + 2 \\ &= (2a-1)^2 \end{aligned}$$

By going back to the definition of n, we have:

$$\begin{aligned} n &= \frac{\sqrt{x^2}}{2} + k + \ell + \frac{1}{2} \\ &= \frac{2a-1}{2} + (b+\ell) + \ell + \frac{1}{2} \\ &= a + b + 2\ell \\ &= a + b + \frac{(2a-1)^2 + (2b-1)^2 - 2}{4} \\ &= a + b + \frac{4a^2 - 4a + 1 + 4b^2 - 4b + 1 - 2}{4} \\ &= a + b + a^2 - a + b^2 - b \\ &= a^2 + b^2 \end{aligned}$$

This proves the statement of the theorem.

□

The graph Q_d, better known as the *d-cube* or *cube*, has the 2^d d-dimensional binary vectors (vectors with only 0s and 1s in their coordinates) as its vertex set and two vertices are adjacent if they differ in exactly one coordinate. Notice that $|V(Q_d)| = 2^d$ and $|E(Q_d)| = d2^{d-1}$.

Theorem 2.14. *The d-cube Q_d is balanceable for $d \geq 1$.*

Proof. By proposition 2.1, it is sufficient to find an independent set of vertices U such that $\sum_{u \in U} deg(u) = d2^{d-2}$. Because of the regularity of Q_d, the set U must have 2^{d-2} vertices.

Let A be the set of vertices with an even number of 1s in their coordinates and let B be the rest of the vertices in Q_d. Notice that A and B are each independent sets because of how the edge set is defined in Q_d. One of the sets A or B has at least $\frac{|V|}{2} = 2^{d-1}$ vertices. Therefore we can take 2^{d-2} vertices from that set and define the set U to conclude that Q_d is balanceable. □

In the next two subsections, we will characterize the balanceability of other graph classes, starting with a special class of circulant graphs.

2.3.1 Balanceability of a special class of circulant graphs

Circulant graphs are defined as Cayley graphs of cyclic groups. The balanceability of some subfamilies of such graphs has already been mentioned, such is the case of cycles and complete graphs. We discuss the balanceability of a specific class of circulant graphs with even number of edges. We can define this class as graphs $C_{k,\ell}$ that are made up of a cycle C_k of order k with vertices $u_0, \cdots, u_{k-1}$ where the chords $u_i u_{i+\ell}$ (addition modulo k) are part of the edge set, taking k and ℓ as positive integers. This family includes the antiprisms and Möbius ladders.

The following statement fully characterizes which of the graphs of the form $C_{k,\ell}$ are balanceable:

Theorem 2.15. *Let k and ℓ be two integers such that $k > 3$ and $\ell \in \{2, \ldots, k-2\}$. The graph $C_{k,\ell}$ is balanceable if and only if k is even and $(k, \ell) \neq (6, 2)$.*

Due to the many cases we have to consider, the proof of Theorem 2.15 will be divided into several lemmas. The statement of Theorem 2.15 and the different cases and lemmas that prove them are summarized in Table 2.2. For the remainder of this section, we will assume that $\ell \leq \frac{k}{2}$, since if this is not the case then we can apply the same reasoning with $\ell' = k - \ell$.

$k \backslash \ell$	$\frac{k}{2}$	$< \frac{k}{2}$		
		odd	2	even, > 2
odd		Not balanceable (Lemma 2.18)		
$\equiv 0 \bmod 4$	Balanceable (Lemma 2.16)	Balanceable (Lemma 2.19)	Balanceable (Lemma 2.20)	
$\equiv 2 \bmod 4$	Balanceable (Lemma 2.17)	Balanceable (Lemma 2.21)	Balanceable except $C_{6,2}$ (Lemma 2.22)	Balanceable (Lemma 2.23)

Table 2.2: The balanceability of the graph $C_{k,\ell}$. If $\ell > \frac{k}{2}$, then refer to $C_{k,k-\ell}$.

First, we consider the case of $C_{k,\frac{k}{2}}$, which only exist if k is even. For $k \geq 6$, those graphs are exactly the Möbius ladders. We will distinguish two cases according to the parity of the edge number of $C_{k,\frac{k}{2}}$, which happens to be even for $k \equiv 0 \bmod 4$, and odd for $k \equiv 2 \bmod 4$.

Lemma 2.16. *Let k be a multiple of 4. The graph $C_{k,\frac{k}{2}}$ is balanceable.*

Proof. Let $k = 4a$ with $a > 1$. Note that, in $C_{k,\frac{k}{2}}$, every vertex has degree 3. Furthermore, let e be the number of edges in $C_{k,\frac{k}{2}}$, then $e = \frac{3k}{2} = 6a$, which implies that $\frac{e}{2} = 3a$. Denote the vertices of $C_{k,\frac{k}{2}}$ by $u_0, u_1, u_2, \ldots, u_{k-1}$, and let $I = \{u_0, u_2, \ldots, u_{2a-2}\}$. It is easy to see that I is an independent set of size a, and thus $\sum_{v \in I} d(v) = 3a = \frac{e}{2}$. Thus, by Proposition 2.1, $C_{k,\frac{k}{2}}$ is balanceable. $\square$

Lemma 2.17. *Let k be an integer such that $k \equiv 2 \bmod 4$. The graph $C_{k,\frac{k}{2}}$ is balanceable.*

Proof. Let $k = 4a + 2$ with $a \geq 1$.

Let e be the number of edges in $C_{4a+2,2a+1}$, then $e = 6a + 3$ and thus $\lfloor \frac{e}{2} \rfloor = 3a + 1$ and $\lceil \frac{e}{2} \rceil = 3a + 2$. Note also that all vertices have degree 3.

We will begin by proving that we can partition the vertices of $C_{4a+2,2a+1}$ in two sets X and Y such that $e(X, Y) = 3a + 1$. First, let $X_1 = \{u_0, u_2, \dots, u_{2a-4}\}$ (if $a = 1$ then X_1 is empty). Then, let $X_2 = \{u_{2a-2}, u_{2a-1}\}$. Now, let $X = X_1 \cup X_2$ and $Y = V(C_{4a+2,2a+1}) \setminus X$. Clearly, the vertices in X_1 are an independent set, and are independent from the vertices in X_2. Thus, each vertex in X_1 has three neighbors in Y, and each vertex in X_2 has two neighbors in Y (since they are adjacent). This implies that we have $e(X, Y) = 3(a - 1) + 4 = 3a + 1$.

Now, we will prove that there is a subset of vertices W such that $e(G[W]) \in \{3a+1, 3a+2\}$. The idea will be to take vertices along the cycle, thus taking a path and some chords within $G[W]$. There are two cases to consider:

- If a is even, then we set $W := \{u_0, u_1, \dots, u_{2a+\frac{a}{2}+1}\}$. The edges within $G[W]$ are of two sorts: there are $2a + \frac{a}{2} + 1$ along the cycle, and there are $\frac{a}{2} + 1$ chords (from u_0u_{2a+1} to $u_{\frac{a}{2}}u_{2a+\frac{a}{2}+1}$). Adding those, we have $e(G[W]) = 2a + \frac{a}{2} + 1 + \frac{a}{2} + 1 = 3a + 2$.

- If a is odd, then we set $W := \{u_0, u_1, \dots, u_{2a+\frac{a-1}{2}+1}\}$. The edges within $G[W]$ are of two sorts: there are $2a + \frac{a-1}{2} + 1$ along the cycle, and there are $\frac{a-1}{2} + 1$ chords (from u_0u_{2a+1} to $u_{\frac{a-1}{2}}u_{2a+\frac{a-1}{2}+1}$). Adding those, we have $e(G[W]) = 2a + \frac{a-1}{2} + 1 + \frac{a-1}{2} + 1 = 3a + 1$.

Those two cases allow us to conclude that $C_{4a+2,2a+1}$ verifies the conditions of Theorem 1.37, and thus $C_{4a+2,2a+1}$ is balanceable. □

In all future cases, we assume $\ell < \frac{k}{2}$, thus every vertex of $C_{k,\ell}$ has degree 4. Furthermore, let e be the number of edges in $C_{k,\ell}$, then $e = 2k$. We will also denote the vertices of $C_{k,\ell}$ by $u_0, u_1, u_2, \dots, u_{k-1}$.

Lemma 2.18. *Let k be an odd integer. The graph $C_{k,\ell}$ is not balanceable.*

Proof. Since k is odd, $\frac{e}{2} = \frac{2k}{2} = k$ is odd. Since $C_{k,\ell}$ is eulerian, Proposition 2.6 implies that $C_{k,\ell}$ is not balanceable. □

Lemma 2.19. *Let k be an integer such that $k \equiv 0 \bmod 4$, and let ℓ be an odd integer. The graph $C_{k,\ell}$ is balanceable.*

Proof. It is easy to see that $C_{k,\ell}$ is bipartite (the vertices with an even index being one part, and the vertices with an odd index being the other part). Furthermore, it has order $4a$ and is 4-regular. Thus, by Proposition 2.5, $C_{k,\ell}$ is balanceable. □

Lemma 2.20. *Let k be an integer such that $k \equiv 0 \bmod 4$, and let ℓ be an even integer. The graph $C_{k,\ell}$ is balanceable.*

Proof. We denote $k = 4a$ and $\ell = 2b$. The idea is to select an independent set I of size a, allowing us to invoke Proposition 2.1. We start by adding to I the vertices $u_0, u_2, \dots, u_{2b-2}$, thus a set of b independent vertices. However, we cannot add u_{2b} because of the edge u_0u_{2b}.

Instead, we can add to I the vertices $u_{2b+1}, u_{2b+3}, \ldots, u_{4b-1}$, thus a set of b new vertices that are independent from each other as well as from the first ones. Again, we have to jump the vertex u_{4b+1} and start from u_{4b+2}. By applying this, we will select $\lfloor \frac{a}{b} \rfloor$ such sets of b vertices, and then we can apply the same construction and add the $a - \lfloor \frac{a}{b} \rfloor b$ last vertices we need to have $|I| = a$ (those last vertices will be called *leftover vertices* in the remainder of the proof). This is depicted in Figure 2.1.

We now need to prove that I is an independent set. Note that the only thing that we need to prove is that the index of the second-biggest-index neighbour of u_0 is greater than the index of the last vertex that we selected. Indeed, we selected the sets in such a way that all vertices are independent from each other going forward.

The last vertex that is selected in I with our construction will have index:

$$i_{max} = \left\lfloor \frac{a}{b} \right\rfloor 2b + \left\lfloor \frac{a}{b} \right\rfloor + 2\left(a - \left\lfloor \frac{a}{b} \right\rfloor b\right) - 1 - 1 - r$$

with $r = 1$ if $\lfloor \frac{a}{b} \rfloor = \frac{a}{b}$ and $r = 0$ otherwise.

The $\lfloor \frac{a}{b} \rfloor 2b$ are the vertices selected in the $\lfloor \frac{a}{b} \rfloor$ sets that are themselves separated from each other by one supplementary vertex (thus the $\lfloor \frac{a}{b} \rfloor$); then the $2(a - \lfloor \frac{a}{b} \rfloor b)$ are the leftover vertices; and then we have to subtract 1 for the last vertex (which does not count) and again substract 1 for the fact that the indices start at 0. Finally, if $\lfloor \frac{a}{b} \rfloor = \frac{a}{b}$, then there are no leftover vertices and thus we can substract 1 from the total.

Thus, we have $i_{max} = 2a + \frac{a}{b} - 2 - r \leq 2a + \frac{a}{b} - 2$. We only have to prove that $i_{max} < 4a - 2b$ since this would prove that the last vertex that we selected has an index smaller than the second-biggest-index neighbour of u_0. There are two cases to consider:

1. If $2b \geq a$, then we have $\frac{a}{b} \leq 2$. Since $\lfloor \frac{a}{b} \rfloor \leq \frac{a}{b}$, we have $i_{max} \leq 2a + \lfloor \frac{a}{b} \rfloor - 2 \leq 2a + 2 - 2 = 2a$. Furthermore, since $2b < \frac{k}{2} = 2a$ we have $4a - 2b > 2a$. This implies that $i_{max} < 4a - 2b$, proving that I is an independent set.

2. If $2b < a$, then $i_{max} \leq 3a - 2$ since $\lfloor \frac{a}{b} \rfloor \leq a$. Furthermore, $4a - 2b > 4a - a = 3a$, and thus $i_{max} < 4a - 2b$, proving that I is an independent set.

Thus, I is an independent set of size a, and since every vertex has degree 4 we have $\sum_{v \in I} d(v) = 4a = \frac{e}{2}$, and Proposition 2.1 implies that $C_{k,\ell}$ is balanceable. $\square$

For the next three lemmas, we cannot construct an independent set I such that $\sum_{v \in I} d(v) = \frac{e}{2}$, since all the degrees are 4 and $\frac{e}{2}$ is not a multiple of 4. We will instead prove that the vertices of $C_{k,\ell}$ can be partitioned in such a way that we can apply Theorem 1.37.

Lemma 2.21. *Let k be an integer such that $k \equiv 2 \bmod 4$, and let ℓ be an odd integer with $\ell < \frac{k}{2}$. The graph $C_{k,\ell}$ is balanceable.*

Proof. Let $k = 4a + 2$ and ℓ be an odd integer. We have two cases to consider.

First, we will prove that we can partition the vertices of $C_{k,\ell}$ in two sets X and Y such that $e(X,Y) = \frac{e}{2} = 4a + 2$. We begin by setting $X := \{u_0, u_1\}$, which puts 6 edges between X and Y as long as no neighbour of u_0 or u_1 is in Y. We now select $a - 1$ independent

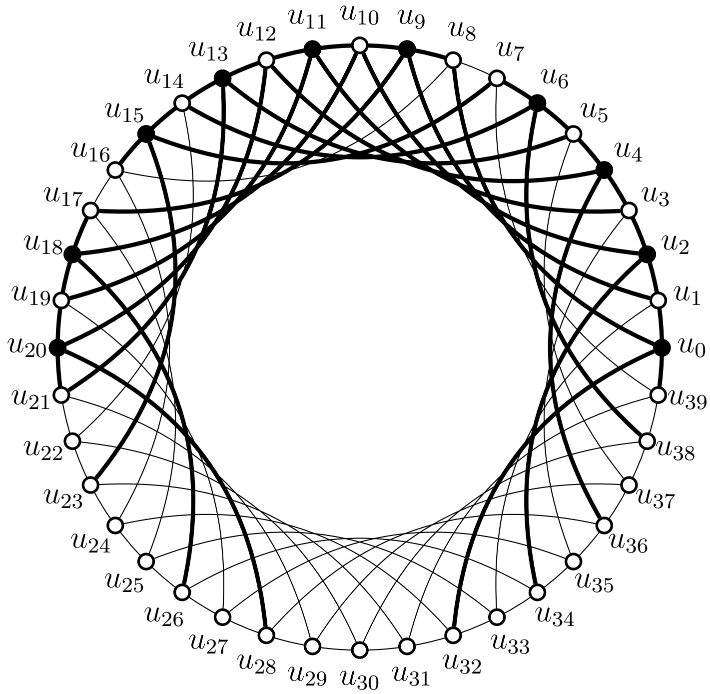


Figure 2.1: A depiction of the proof of Lemma 2.20 on $C_{40,8}$. The vertices that we selected in I are in black, and the out-edges of I are bolded.

vertices that are not neighbours of u_0 and u_1 and put them in X. Note that there are $4a-6$ vertices not neighbours of u_0 and u_1: $2a-3$ with an even index and $2a-3$ with an odd index. Thus, we can select $a-1$ vertices of even index (without loss of generality), which is always possible. Indeed, assume by contradiction that $a-1 > 2a-3$; then $a < 2$, *i.e.* $k < 10$, *i.e.* $k = 6$, which is a contradiction since ℓ is odd, but $\ell > 1$ and $\ell < \frac{k}{2} = 3$ so this case cannot occur. Since the $a-1$ vertices we just added to X are independent, we have $e(X,Y) = 6 + 4(a-1) = 4a+2$. This construction is depicted in Figure 2.2.

Now, denote $C_{k,\ell}$ by G, its vertex-set by V and its edge-set by E. We will prove that we can partition V in two sets W and $V \setminus W$ such that $e(G[W]) = \frac{e}{2} = 4a+2$. Note that two adjacent vertices in $V \setminus W$ independent from all other vertices in $V \setminus W$ put 7 edges in $E \setminus E(G[W])$. We will construct $V \setminus W$ by selecting two pairs of adjacent vertices independent from each other, and $a-3$ independent vertices that will not be neighbours of the four vertices previously selected. Thus, we will have $e - e(G[W]) = 2 \times 7 + 4(a-3) = 4a+2$, and thus $e(G[W]) = 4a+2$. There are three cases to consider.

First, assume that $k = 10$, the only graph to consider is $C_{10,3}$. In this case, we cannot construct the sets as explained above (since $2 \times 7 = 14 > 10 = \frac{e}{2}$). However, by setting $W = \{u_0, u_1, u_2, u_3, u_4, u_5, u_6\}$, there are 10 edges in $G[W]$, so this case is covered.

Then, assume that $\ell = 3$ and $k > 10$. We put u_0, u_1, u_5 and u_6 in $V \setminus W$. There are $k - 13 = 4a - 11$ vertices that are neither those nor neighbours of those: $2a-5$ with an even index and $2a-6$ with an odd index. Thus, we can select $a-3$ independent vertices with an even index, which is always possible since $a - 3 > 2a - 5$ if and only if $a < 2$, *i.e.* $k < 10$, which cannot occur as discussed previously. This implies that we have $e - e(G[W]) = 4a + 2$. This construction is depicted on the left-hand side of Figure 2.3.

Finally, assume that $\ell > 3$ and $k > 10$. We put u_0, u_1, u_3 and u_4 in $V \setminus W$. There are $k - 15 = 4a - 13$ vertices that are neither those nor neighbours of those: $2a - 6$ with an even index and $2a - 7$ with an odd index. Thus, we can select $a - 3$ independent vertices with an even index, which is always possible. Indeed, assume by contradiction that $a - 3 > 2a - 6$, then $a < 3$, *i.e.* $k < 14$; the case $k = 6$ has been discussed previously, and the case $k = 10$ cannot occur either since this would imply $\ell = 5 = \frac{k}{2}$, a contradiction. This implies that we have $e - e(G[W]) = 4a + 2$. This construction is depicted on the right-hand side of Figure 2.3.

Altogether, this allows us to invoke Theorem 1.37, and thus to conclude that the circulant graph $C_{4a+2,\ell}$ is balanceable when ℓ is odd and $\ell < 2a + 1$. $\square$

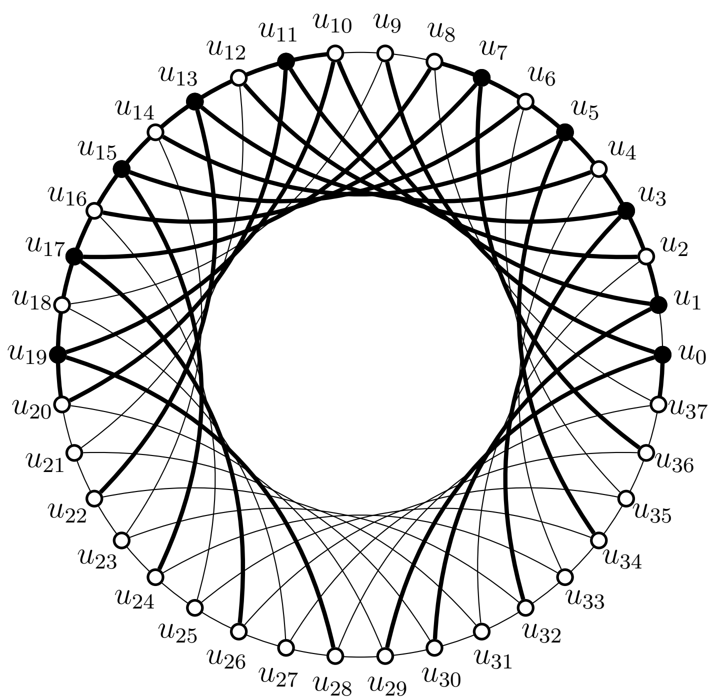


Figure 2.2: A depiction of the first case of the proof of Lemma 2.21 on $C_{38,9}$. The vertices that we selected in X are in black, and the edges between X and Y are bolded.

Lemma 2.22. *Let k be an integer such that $k \equiv 2 \bmod 4$. The graph $C_{k,2}$ is balanceable if and only if $k \neq 6$.*

Proof. This proof contains two parts: first, we will prove that $C_{6,2}$ is not balanceable; then, we will prove that $C_{4a+2,2}$ is balanceable when $a > 1$.

First, assume that $k = 6$. We will prove that there is no subset of vertices W such that $e(G[W]) = 6$. First, note that W cannot possibly be empty or all the vertices. Taking this

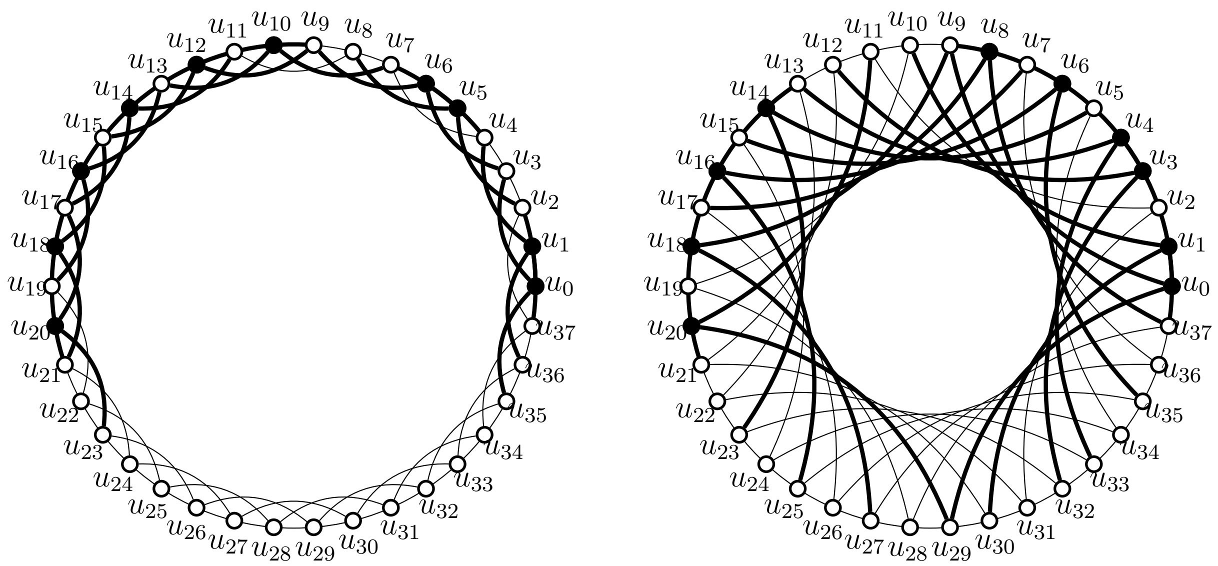


Figure 2.3: A depiction of the second case of the proof of Lemma 2.21 on $C_{38,3}$ and $C_{38,9}$. The vertices that we selected in $V \setminus W$ are in black, and the edges outside of $G[W]$ are bolded.

into account, Table 2.3 shows possible sets for different sizes of W as well as $e(G[W])$ in each case (the possible sets are up to renaming vertices). Since no set W gives $e(G[W]) = 6$, Theorem 1.37 implies that $C_{6,2}$ is not balanceable.

$\lvert W\rvert$	1	2	3	4	5
Possible vertices in W $\to e(G[W])$	$u_0 \to 0$	$u_0, u_1 \to 1$ $u_0, u_2 \to 1$ $u_0, u_3 \to 0$	$u_0, u_1, u_2 \to 3$ $u_0, u_1, u_3 \to 2$ $u_0, u_2, u_4 \to 3$	$u_0, u_1, u_2, u_3 \to 5$ $u_0, u_1, u_2, u_4 \to 5$ $u_0, u_1, u_3, u_4 \to 4$	$u_0, \ldots, u_4 \to 8$

Table 2.3: Possible sets W of vertices of $C_{6,2}$ (up to renaming the vertices) and the value of $e(G[W])$ for each of them.

Now, assume that $a > 1$, we will prove that $C_{4a+2,2}$ is balanceable. The proof is similar to the proof of Lemma 2.21.

First, we will prove that we can partition the vertices of $C_{4a+2,2}$ in two sets X and Y such that $e(X, Y) = 4a + 2$. For this, we set $X_1 = \{u_0, u_3, \ldots, u_{3(a-2)}\}$, $X_2 = \{u_{4a-1}, u_{4a}\}$, and $X = X_1 \cup X_2$. It is easy to see that X_1 is an independent set and that no vertex in X_2 is adjacent to a vertex in X_1 (since we have $a > 1 \Rightarrow 4a - 3 > 3a - 2 > 3(a - 2)$). Thus, we have $e(X, Y) = e(X_1, Y) + e(X_2, Y) = 4(a - 1) + 6 = 4a + 2$.

Then, as in the proof of Lemma 2.21, assume that $a = 2$, *i.e.* $k = 10$. If we set $W = \{u_0, u_1, u_2, u_3, u_4, u_6, u_8\}$, then we have $e(G[W]) = 10$. This, with the previous point (that applies if $a = 2$), proves that $C_{10,2}$ is balanceable. Assume now that $a > 2$. Let $V_1 = \{u_0, u_3, \ldots, u_{3(a-4)}\}$ (if $a = 3$ then we set $V_1 = \emptyset$), $V_2 = \{u_{4a-6}, u_{4a-5}\}$ and $V_3 =$

$\{u_{4a-2}, u_{4a-1}\}$; then set $W = V \setminus (V_1 \cup V_2 \cup V_3)$. It is easy to see that V_1 is an independent set and that no vertex in V_2 (resp. V_3) is adjacent to a vertex in V_1 or V_3 (resp. V_1 or V_2), since we have $a > 2 \Rightarrow 4a-8 > 3a-6 > 3(a-4)$. Thus, we have $e - e(W) = 4(a-3)+14 = 4a+2$, which implies $e(G[W]) = 4a + 2$.

The above constructions allow us to invoke Theorem 1.37, which implies that $C_{4a+2,2}$ is balanceable. □

Lemma 2.23. *Let k be an integer such that $k \equiv 2 \bmod 4$, and let ℓ be an even integer such that $\ell > 2$. The graph $C_{k,\ell}$ is balanceable.*

Proof. Let $k = 4a + 2$ and $\ell = 2b$. The proof for this lemma is a mix of the proofs for Lemmas 2.20 and 2.22: we will use the structure we constructed in the proof for Lemma 2.20 and add either one or two independent edges to it, modifying the structure to keep everything independent from each other.

First, we construct X in two steps. We begin by creating a set X_1 by applying the same construction than in the proof of Lemma 2.20 (so several sets of $\frac{\ell}{2}$ vertices at distance 2 along the outer cycle from each other, each set being separated from the others by another vertex): a total of $a-1$ such vertices are added to X_1. Then, let $X_2 := \{u_{4a-1}, u_{4a}\}$. Now, we need X_1 and X_2 to be independent from each other, so if a vertex in X_1 is adjacent to a vertex in X_2 (this can happen to at most one vertex), we remove it from X_1 and add to this set the next vertex in the construction described in the proof of Lemma 2.20 (we may start a new set this way). This is depicted on the left-hand side of Figure 2.4. The last vertex that is selected in X_1 with our construction will have index:

$$i_{max} \leq 2a + \left\lfloor \frac{a}{b} \right\rfloor - 2 + 3 - 2.$$

That is, the same maximum index than in the proof of Lemma 2.20, but with two corrections: $+3$ may happen since we could start a new set by shifting the neighbour of either u_{4a-1} or u_{4a} (this gives us $+2$, and may give us an additional $+1$ if the vertex we shift creates a new set), and -2 since we only need $a-1$ vertices in X_1 (instead of the a from the proof of Lemma 2.20). Now, we need to prove that this last index is less than $4a - 1 - \ell$.

If $\ell \geq a$, then we can check that we will always have $i_{max} = 2a-3 < 2a-2 < 4a-1-\ell$ since $\ell < \frac{k}{2} = 2a+1$. Indeed, we will put in X_1 first the $b-1$ vertices $u_0, u_2, \ldots, u_{\ell-4}$, then the $a-b$ vertices $u_{\ell-1}, u_{\ell+1}, \ldots, u_{\ell+2(a-b)-3}$ (which is always possible since $\ell \geq a$). The last index will thus always be $\ell + 2(a-b) - 3 = 2a - 3$.

Assume now that $\ell < a$. Since $\ell > 2$, we have $b \geq 2$ and thus $\lfloor \frac{a}{b} \rfloor \leq \lfloor \frac{a}{2} \rfloor \leq \frac{a}{2} + 1$. Thus, $i_{max} \leq 2a + \lfloor \frac{a}{b} \rfloor - 1 \leq 2a + \frac{a}{2} + 1 - 1 = \frac{5a}{2} < 3a \leq 4a - 1 - \ell$.

Hence, in this construction, X_1 and X_2 are independent from each other, and by setting $X = X_1 \cup X_2$ we have $e(X, Y) = 6 + 4(a-1) = 4a + 2$.

Now, as in the proof of Lemma 2.21, we have to deal with the case of $C_{10,4}$. In this case, by setting $W := \{u_0, u_1, u_2, u_4, u_5, u_6, u_8\}$, we have $e(G[W]) = 10$.

Finally, for $k \geq 14$, we construct $V \setminus W$ by applying the same construction than in the proof of Lemma 2.20. Let V_1 be a set of $a-3$ vertices constructed this way, then let either

$V_2 = \{u_{4a-4,4a-3}\}$ (if $\ell > 4$) or $V_2 = \{u_{4a-7}, u_{4a-6}\}$ (if $\ell = 4$), and $V_3 = \{u_{4a-1}, u_{4a}\}$. Again, we shift the potential vertices in V_1 adjacent to a vertex in V_2 or V_3 (at most 2 such vertices), and thus the highest index we can reach is:

$$i_{max} \leq 2a + \left\lfloor \frac{a}{b} \right\rfloor - 2 + 6 - 6.$$

The $+6$ comes from the two potential shifts, and the -6 from the fact that we select $a-3$ vertices instead of a. We now need to verify that $i_{max} < 4a - i - \ell$ for $i \in \{4, 7\}$ (depending on the value of ℓ). We have three cases to check:

1. If $\ell = 4$, then $i_{max} \leq 2a + \lfloor \frac{a}{2} \rfloor - 2$; and $4a - 7 - \ell = 4a - 11$. Now, if $a > 6$ then since $\lfloor \frac{a}{2} \rfloor \leq \frac{a}{2}$ it is easy to check that $i_{max} \leq 2a + \frac{a}{2} - 2 < 4a - 11$. We need to check that $i_{max} < 4a - 11$ in the remaining cases:

   (a) If $a = 3$, then we have $V_1 = \emptyset$ so no contradiction arises;

   (b) If $a = 4$, then we have $i_{max} = 0$ and $4a - 11 = 5$ so no contradiction arises;

   (c) If $a = 5$, then we have $i_{max} = 3$ and $4a - 11 = 9$ so no contradiction arises;

   (d) If $a = 6$, then we have $i_{max} = 5$ and $4a - 11 = 13$ so no contradiction arises.

   Thus, if $\ell = 4$, then $i_{max} < 4a - 7 - \ell$.

2. If $\ell \geq a$, then we can check that we will always have $i_{max} = 2a - 5$ and $4a - 4 - \ell \geq 2a - 4$ since $\ell < \frac{k}{2} = 2a + 1$. Indeed, we will put in V_1 first the $b - 3$ vertices $u_0, u_2, \ldots, u_{\ell-8}$ as well as $u_{\ell-4}$, then the $a - b - 1$ vertices $u_{\ell-1}, u_{\ell+1}, \ldots, u_{\ell+2(a-b)-5}$ (which is always possible since $\ell \geq a$). The last index will thus always be $\ell + 2(a - b) - 5 = 2a - 5$.

3. If $\ell > 4$ (thus $b > 2$) and $\ell < a$, then since $\lfloor \frac{a}{b} \rfloor < \frac{a}{2}$ we have $i_{max} \leq 2a + \lfloor \frac{a}{b} \rfloor - 2 < \frac{5a}{2} - 2$; and $4a - 4 - \ell > 3a - 4$. Now, we know that $a > \ell > 4$, so it is easy to check that $\frac{5a}{2} - 2 < 3a - 4$, and thus, that $i_{max} < 4a - 4 - \ell$.

All those cases prove that V_2 and V_3 are independent from V_1. By setting $V \setminus W = V_1 \cup V_2 \cup V_3$, we have $e - e(G[W]) = 14 + 4(a - 3) = 4a + 2$, and thus $e(G[W]) = 4a + 2$.

Those two constructions, depicted in Figure 2.4, allow us to invoke Theorem 1.37, which implies that $C_{4a+2,\ell}$ is balanceable when ℓ is even and $\ell > 2$. □

Together, Lemmas 2.16 to 2.23 prove the validity of Theorem 2.15, which fully characterizes, among this special class of circulant graphs, which are balanceable and which are not. In particular, note that every Möbius ladder is balanceable (by Lemmas 2.16 and 2.17) and that the 3-antiprism graph is the only non-balanceable antiprism graph (by Lemma 2.22 and a special case of Lemma 2.20).

With this special class of circulant graphs, we get one step closer towards a full classification of the balanceability of circulant graphs, which had been informally started by the study of complete graphs and cycles. This full classification is an interesting open problem for future studies.

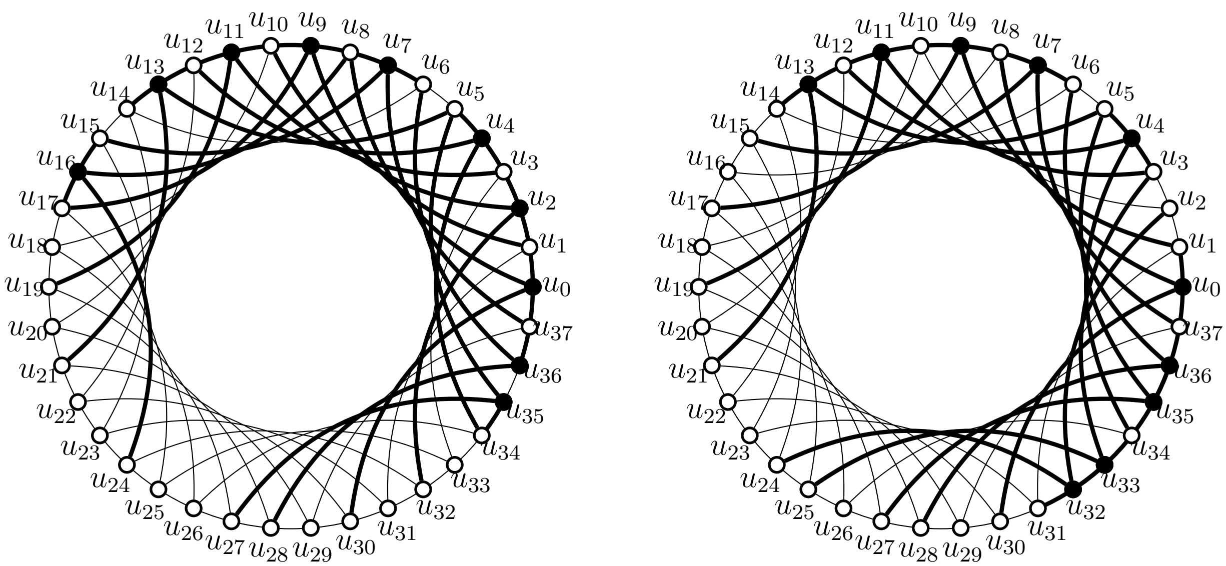


Figure 2.4: A depiction of the proof of Lemma 2.23 on $C_{38,8}$. On the left-hand side, the vertices in X are bolded, as well as the edges between X and Y. On the right-hand side, the vertices in $V \setminus W$ are bolded, as well as the edges outside $G[W]$.

2.3.2 Balanceability of grids

In this section, we study the balanceability of grid graphs with an even edge number. Particularly, we consider rectangular and triangular grids.

Rectangular grids

Let $G_{k,\ell}$ be the rectangular grid graph with k vertices per row and ℓ vertices per column. It is easy to see that $G_{k,\ell}$ has $k(\ell-1)+(k-1)\ell = 2k\ell-(k+\ell)$ edges, and this number is even if and only if k and ℓ have the same parity.

Theorem 2.24. *Let k and ℓ be two integers such that $k, \ell > 1$. If k and ℓ have the same parity, then $G_{k,\ell}$ is balanceable.*

Proof. In the grid graph $G_{k,\ell}$ with vertex-set V and edge-set E, vertices can have degree two, three or four. The repartition is as follows:

- 4 vertices of degree two (the corners);
- $2(k-2)+2(\ell-2) = 2(k+\ell)-8$ vertices of degree three (the sides);
- $k\ell-2(k+\ell)+4$ vertices of degree four (the inside).

It is well-known that $\sum_{v \in V} d(v) = 2|E|$. We want to find an independent set of vertices I such that $\sum_{v \in I} d(v) = \frac{|E|}{2}$. To do this, we can select one fourth of the vertices in every degree set. There are several cases.

Case 1: If k and ℓ are even, then we can select 1 vertex of degree two, $\frac{k+\ell}{2} - 2$ vertices of degree three, and $\frac{k\ell}{4} - \frac{k+\ell}{2} + 1$ vertex of degree four. An example is depicted on Figure 2.5. It is always possible to select those vertices such that they induce an independent set, since there is an independent set containing half the vertices of $G_{k,\ell}$, and in particular half the corners, half of the sides and half of the inside. By applying Proposition 2.1, $G_{k,\ell}$ is balanceable.

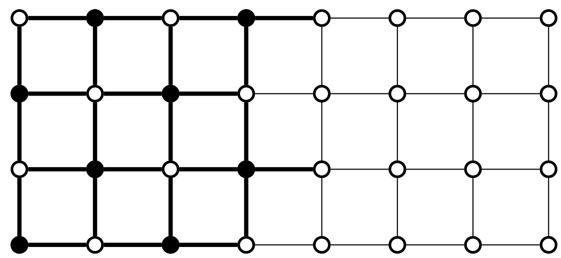


Figure 2.5: The independent set I such that $\sum_{v \in I} d(v) = \frac{|E(G)|}{2}$ for $G_{4,8}$. Vertices in I are bolded, as well as the out-edges of I.

Case 2: If k and ℓ are odd, and $k + \ell$ is not a multiple of 4, then we can select 1 vertex of degree two, $\frac{k+\ell}{2} - 3$ vertices of degree three, and $\frac{k\ell+7}{4} - \frac{k+\ell}{2}$ vertices of degree four. An example is depicted on Figure 2.6. Again, it is always possible to select those vertices such that they induce an independent set (by the same argument than the previous case). Furthermore, $k\ell + 7$ is a multiple of 4: by noting $k = 2a + 1$ and $\ell = 2b + 1$, we have $k\ell + 7 = 4ab + 2a + 2b + 1 + 7 = 4ab + 8 + (2a + 2b) = 4ab + 8 + (k + \ell - 2)$, and the fact that $k + \ell$ is not a multiple of 4 implies that $(k + \ell - 2)$ is. We will thus have:

$$\begin{aligned}
\sum_{v \in I} d(v) &= 2 + 3\left(\frac{k+\ell}{2} - 3\right) + 4\left(\frac{k\ell+7}{4} - \frac{k+\ell}{2}\right) \\
&= 2 + 3\frac{k+\ell}{2} - 9 + k\ell + 7 - 4\frac{k+\ell}{2} \\
&= k\ell - \frac{k+\ell}{2} \\
&= \frac{|E|}{2}
\end{aligned}$$

Proposition 2.1 then implies that $G_{k,\ell}$ is balanceable.

Case 3: If k and ℓ are odd, and $k + \ell$ is a multiple of 4, then we can select 2 vertices of degree two, $\frac{k+\ell}{2} - 3$ vertices of degree three, and $\frac{k\ell+5}{4} - \frac{k+\ell}{2}$ vertices of degree four. An example is depicted on Figure 2.7. Again, it is always possible to select those vertices

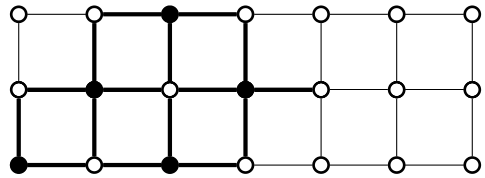


Figure 2.6: The independent set I such that $\sum_{v\in I} d(v) = \frac{|E(G)|}{2}$ for $G_{3,7}$. Vertices in I are bolded, as well as the out-edges of I.

such that they induce an independent set (by the same argument than the previous case). Furthermore, $k\ell + 5$ is a multiple of 4: by noting $k = 2a + 1$ and $\ell = 2b + 1$, we have $k\ell + 5 = 4ab + 2a + 2b + 1 + 5 = 4ab + (2a + 2b + 6) = 4ab + (k + \ell + 4)$. We will thus have:

$$\begin{aligned}
\sum_{v\in I} d(v) &= 4 + 3\left(\frac{k+\ell}{2} - 3\right) + 4\left(\frac{k\ell+5}{4} - \frac{k+\ell}{2}\right) \\
&= 4 + 3\frac{k+\ell}{2} - 9 + k\ell + 5 - 4\frac{k+\ell}{2} \\
&= k\ell - \frac{k+\ell}{2} \\
&= \frac{|E|}{2}
\end{aligned}$$

Proposition 2.1 then implies that $G_{k,\ell}$ is balanceable.

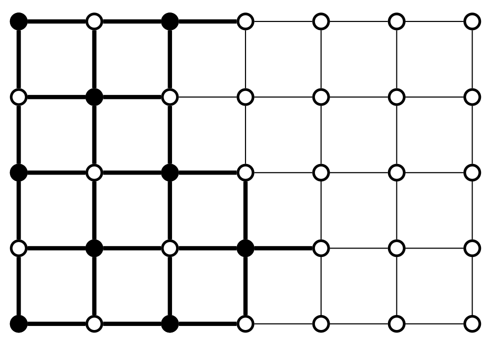


Figure 2.7: The independent set I such that $\sum_{v\in I} d(v) = \frac{|E(G)|}{2}$ for $G_{5,7}$. Vertices in I are bolded, as well as the out-edges of I.

All possible cases have been covered, and thus if k and ℓ have the same parity, then the rectangular grid $G_{k,\ell}$ is balanceable. □

Triangular grids

Let T_h be the (equilateral) triangular grid with h vertices on each side. It is easy to see that T_h has $\frac{3(h-1)h}{2}$ edges, and that this is even if and only if $h \bmod 8 \in \{0, 1, 4, 5\}$. We will prove that some triangular grids are not balanceable, while others are.

Theorem 2.25. *Let h be a positive integer such that $h \bmod 8 \in \{0, 1, 4, 5\}$. The triangular grid T_h is balanceable if and only if $h \bmod 8 \in \{0, 1\}$.*

Proof. We prove two statements here: the non-balanceability of T_{8k+4} and T_{8k+5}; as well as the balanceability of T_{8k} and T_{8k+1}.

We will consider the vertices of T_h row by row, starting from a single vertex at the top of the grid. The vertex u_i^j will be the ith vertex (starting from the left) in the jth row (starting from the top), so the top vertex is u_1^1, the second row contains u_1^2 and u_2^2, and so on. Note that the three corner vertices have degree 2, the vertices on the sides of the grid have degree 4, and the vertices in the middle have degree 6; thus T_h is eulerian.

First, assume that $h = 8k + 4$. Then, $\frac{|E(G)|}{2} = \frac{3(8k+3)(8k+4)}{4} = 48k^2 + 42k + 9$ and thus is odd. Since T_h is eulerian, Proposition 2.6 implies that it is not balanceable. The reasoning is the same with $h = 8k + 5$, with $\frac{|E(G)|}{2} = 48k^2 + 54k + 15$.

Now, assume that $h \in \{8k, 8k + 1\}$. We will prove that there is an independent set I such that $\sum_{v \in I} d(v) = \frac{|E(G)|}{2}$, and apply Proposition 2.1 to complete the proof. We define, for every row except the first, second and last ones, two kinds of independent sets: we call *A-set* of the ith row the independent set containing all vertices u_{2j+1}^i for $j \geq 0$; and we call *B-set* of the ith row the independent set containing all vertices u_{2j}^i for $j \geq 1$. Note that, if i is odd, then the A-set of the ith row contains two vertices of degree 4 and $\frac{i-3}{2}$ vertices of degree 6; and the B-set of the ith row contains $\frac{i-1}{2}$ vertices of degree 6. In the following, we will call degree of an A-set (resp. B-set) the sum of the degrees of the vertices it contains.

Case 1: $h = 8k$. Note that in this case, $\frac{|E(G)|}{2} = 48k^2 - 6k$. We take the following vertices in I:

1. u_1^1:
2. B-sets on rows $3 + 2i$ for $i \in \{0, \ldots, k - 1\}$;
3. A-sets on rows $3 + 2k, 3 + 2k + 2, \ldots, h - 1$.

This is depicted on the left-hand side of Figure 2.8.

Thus, I contains u_1^1 which has degree 2, k B-sets which have degree $6(i + 1)$ for $i \in \{0, \ldots, k - 1\}$, and $3k - 1$ A-sets which have degree $8 + 6i$ for $i \in \{k, \ldots, 4k - 2\}$. Thus, we have:

$$\begin{aligned}
\sum_{v \in I} d(v) &= 2 + \sum_{i=0}^{k-1} 6(i + 1) + \sum_{i=k}^{4k-2} (8 + 6i) \\
&= 2 + \frac{6k(k + 1)}{2} + 8(4k - 2) + \frac{6(4k - 2)(4k - 1)}{2} - 8(k - 1) - \frac{6(k - 1)k}{2} \\
&= 48k^2 - 6k
\end{aligned}$$

Case 2: $h = 8k + 1$. Note that in this case, $\frac{|E(G)|}{2} = 48k^2 + 6k$. If $h = 1$ then the result trivially holds since G is the trivial graph. Otherwise, we take the following vertices in I:

1. u_1^1;

2. A-sets on rows $3, 5, \ldots, h-2$, from which we remove k vertices of degree 6 (this is always possible since those A-sets will contain $\sum_{i=0}^{4k-2} i = 8k^2 - 6k + 1$ vertices of degree 6, and $k \geq 1 \Rightarrow 8k^2 - 6k + 1 > k$);

3. $u_1^h, u_3^h, \ldots, u_h^h$.

This is depicted on the right-hand side of Figure 2.8.

Thus, I contains u_1^1 which has degree 2, $4k-1$ A-sets which have degree $8+6i$ for $i \in \{0, \ldots, 4k-2\}$, from which we remove k vertices of degree 6 thus removing $6k$, and the vertices selected on the last row ($4k-1$ of degree 4 and two of degree 2). Thus, we have:

$$\begin{aligned}
\sum_{v \in I} d(v) &= 2 + \sum_{i=0}^{4k-2} (8+6i) - 6k + 4(k-1) + 2 + 2 \\
&= 2 + 8(4k-1) + \frac{6(4k-2)(4k-1)}{2} - 6k + 16k \\
&= 48k^2 + 6k
\end{aligned}$$

Thus, we have proved that if $h \bmod 8 \in \{0, 1\}$, then T_h is balanceable; and that if $h \bmod 8 \in \{4, 5\}$, then T_h is not balanceable. This completes the proof of Theorem 2.25. $\square$

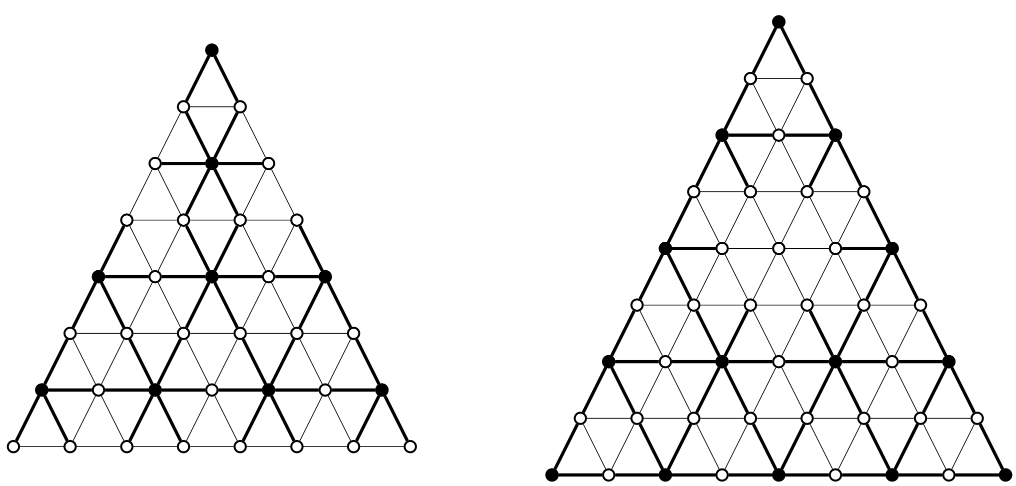


Figure 2.8: The independent set I such that $\sum_{v \in I} d(v) = \frac{|E(G)|}{2}$ for T_8 (on the left) and T_9 (on the right). Vertices in I are bolded, as well as the out-edges of I.

To summarize the results exposed in this section, the table 2.1 contains all of the known balanceable and non-balanceable graph families up to date (amoeba graphs are defined in Chapter 3). The results previously presented were published in [22] that was sent to the journal Discrete Applied Mathematics and was accepted in September of 2020.

2.4 Balancing Number

Recall that, if a graph G is balanceable with parameters as in Definition 1.36, then the smallest such k is called the *balancing number of* G and is denoted as $bal(n, G)$. Even though the concept is quite new, there has been some work done in order to find or improve bounds for the balancing number of certain graphs. For example, Caro, Lauri and Zarb [15] studied the balancing number of graphs of at most four edges. In [6], colorings using arbitrary many colors are studied and the 3-balancing number for paths is determined up to a constant factor. Caro, Hansberg and Montejano determined all graphs with constant balancing number in [14]. In table 2.4, we can see the exact results and bounds for the balancing number of certain graph families. The balancing number of paths and stars was discussed in 1.5.2 which will serve as a comparison to analogous results in a $K_{n,n}$ setting in 4. In this section we determine exact values of $bal(n, G)$ for $C_{4k\pm1}$ and tight (up to the first term) lower and upper bounds for C_{4k}, which are all balanceable cycles.

Graph G	**Balancing number** $\mathrm{bal}(n, G)$
K_4	n if $n \equiv 0 \pmod 4$ and $n-1$ in any other case [11]
Trees T on k edges	$\leq (k-1)n$ [13]
Paths P_k	$(\frac{k-2}{4})n - \frac{k^2}{32} + \frac{k}{8}$ if $k \equiv 2 \pmod 4$ and $(\frac{k-2}{4})n - \frac{k^2}{32} + \frac{k}{8} + 1$ if $k \equiv 0 \pmod 4$ [13]
Stars $K_{1,k}$	$(\frac{k-2}{2})n - \frac{k^2}{8} + \frac{k}{4} \leq (k-1)n$ [13]
Paths $P_{4k+\alpha-1}$ with $\alpha \in \{-1, 1\}$	$(k-1)n - \frac{1}{2}(k^2 - k - 1 - \alpha)$ [13]
Cycles $C_{4k+\alpha}$ with $\alpha \in \{-1, 1\}$	$(k-1)n - \frac{1}{2}(k^2 - k - 1 - \alpha)$ [20]
Cycles C_{4k}	$(k-1)n - (k-1)^2 \leq \mathrm{bal}(n, C_{4k}) < (k-1)n + 12k^2 + 3k$ [20]

Table 2.4: Balancing numbers $\mathrm{bal}(n, G)$ of relevant graphs.

2.5 Generalized balancing number

The problem of determining the balancing number is still open for many graph classes. It is clear that, if a graph is balanceable, then determining its balancing number is a relevant problem, yet not necessarily a simple one. On the other hand, if a graph G is not balanceable, we manage to obtain balanced copies of G by allowing a relaxation of the 2-coloring. In this research work, we extend the notion of balancing number by considering *edge coverings*. In an edge covering, each edge receives a nonempty set of colors. Hence, if an edge has more than one color in its set, then we may consider the color we wish in order to find the balanced copy of the graph in question.

In this section, we provide general bounds for the generalized balancing number, which are relevant for graphs where the balancing number does not exist. In particular, Proposition 2.28 says that looking at the number of bicolored edges suffices to get upper bounds; which, in turn, leads to consider the Turán number for a particular class of graphs described in Theorem 2.29. Before proceeding to these results, we provide a proof of the fact mentioned in the introduction: if $\mathrm{bal}(n, G)$ exists for a graph G, then $\mathrm{bal}^*(n, G) = \mathrm{bal}(n, G)$.

Proposition 2.26. *Let G be a graph and n be an integer. If $\mathrm{bal}(n, G)$ exists, then $\mathrm{bal}(n, G) = \mathrm{bal}^*(n, G) < \frac{1}{2}\binom{n}{2}$.*

Proof. Consider a graph $G(V, E)$ and n for which $\mathrm{bal}(n, G)$ exists; in particular, $n \geq |V|$. As we previously discussed, the class considered for $\mathrm{bal}^*(n, G)$ extends the class of 2-colorings; this implies that $\mathrm{bal}(n, G) \leq \mathrm{bal}^*(n, G)$. On the other hand, it is clear by definition that $\mathrm{bal}(n, G) < \frac{1}{2}\binom{n}{2}$.

To establish the equality we prove that every 2-edge covering $E(K_n) = R \cup B$ satisfying $|R|, |B| > \mathrm{bal}(n, G)$ has a balanced copy of G. Suppose that there is a 2-coloring $E(K_n) = R' \cup B'$ with $R' \subset R$, $B' \subset B$ and $|R'|, |B'| > \mathrm{bal}(n, G)$, then a balanced copy of G under the 2-coloring $R' \cup B'$ also corresponds to a balanced copy of G under the 2-edge covering $R \cup B$. Hence, it remains to show that we may construct a coloring $R' \cup B'$ with such properties. If $|R \setminus B| > \mathrm{bal}(n, G)$, then we simply let $R' = R \setminus B$ and $B' = B$. Otherwise, let $R' \subset R$ be an arbitrary subset such that $R \setminus R' \subset B$ and $|R'| = \mathrm{bal}(n, G) + 1$, then let $B' = B \setminus R'$. In either case we have $R' \subset R$ and $B' \subset B$. The constraint on the size of R' and B' is clearly satisfied also; for B' in the latter case, observe that $\mathrm{bal}(n, G) < \frac{1}{2}\binom{n}{2}$ implies $|B'| = \binom{n}{2} - |R'| \geq \mathrm{bal}(n, G)$ completing the proof that $\mathrm{bal}(n, G) = \mathrm{bal}^*(n, G)$. Hence, every 2-edge covering $E(K_n) = R \cup B$ satisfying $|R|, |B| > \mathrm{bal}(n, G)$ has a balanced copy of G. □

Proposition 2.26 immediately implies the following statement:

Corollary 2.27. *Let $G(V, E)$ be a graph and $n \geq |V|$ be an integer. If $\mathrm{bal}(n, G)$ does not exist, then $\frac{1}{2}\binom{n}{2} \leq \mathrm{bal}^*(n, G) < \binom{n}{2}$.*

The next result, which is the key for the main theorem of the section, uses a simple relation between the sizes of the color classes and the necessary number of bicolored edges.

Proposition 2.28. *Let G be a graph and let b be a positive integer. If every 2-list edge coloring with at least b bicolored edges has a balanced copy of G then $\mathrm{bal}^*(n, G) \leq \frac{1}{2}\binom{n}{2} + \left\lceil \frac{b}{2} \right\rceil - 1$.*

Proof. First, observe that an inclusion-exclusion argument gives that if $R \cup B$ are the color classes of a 2-edge covering of K_n satisfying $|R| = |B| = \frac{1}{2}\binom{n}{2} + m$, then there are exactly $2m$ bicolored edges. Consequently, for any 2-edge covering inducing color classes R and B with $|R|, |B| \geq \frac{1}{2}\binom{n}{2} + \left\lceil \frac{b}{2} \right\rceil$, there are at least b bicolored edges, and so, by hypothesis, there is a balanced copy of G. □

In the remainder of the paper we say that a 2-edge covering has a *color excess* of b edges, referring to the number of edges in each color by which $\frac{1}{2}\binom{n}{2}$ is at least surpassed. More precisely, we say that a 2-edge covering $E(K_n) = R \cup B$ has a color excess of b edges, if $|R|, |B| \geq \frac{1}{2}\binom{n}{2} + b$. In such a case, since $|R \cap B| = |R| + |B| - |R \cup B| = |R| + |B| - \binom{n}{2} \geq 2b$, it clearly follows that there have to be at least $2b$ bicolored edges.

The following theorem uses the idea of exploiting the flexibility of bicolored edges. Once we have a copy of G we have to choose the color of the bicolored edges, and we do so according to the edges which have a unique color. In particular, if more than half the edges of such a copy are bicolored, then we may distribute these between the two color classes to balance the copy, regardless of the color of the rest of the edges. With this perspective, a general bound on the generalized balancing number may be reduced to guaranteeing that the 2-edge coverings have enough bicolored edges. To this aim, we need to consider for a graph G, the family of all its subgraphs $H \leq G$, where the graph H satisfies that $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, having half of the edges,

$$\mathcal{H}(G) = \left\{ H : H \leq G, e(H) = \left\lfloor \frac{e(G)}{2} \right\rfloor, H \text{ has no isolated vertices} \right\};$$

recall that an *isolated* vertex is a vertex of degree 0. We note at this point that this family was already used in [9] to gain a similar insight in studying the existence of balanced copies of spanning subgraphs of a 2-colored K_n.

Theorem 2.29. *Let $G(V, E)$ be a graph and $n \geq |V|$ be an integer. Then we have*

$$\mathrm{bal}^*(n, G) \leq \frac{1}{2}\binom{n}{2} + \left\lceil \frac{\mathrm{ex}(n, \mathcal{H}(G))}{2} \right\rceil.$$

Proof. Let $G(V, E)$ be a graph, and let R and B be the color classes induced by a 2-edge covering of K_n with $|R|, |B| > \frac{1}{2}\binom{n}{2} + \lceil \frac{\mathrm{ex}(n, \mathcal{H}(G))}{2} \rceil$. This implies that there are at least $\mathrm{ex}(n, \mathcal{H}(G)) + 1$ bicolored edges. In particular, since K_n is of order n, the subgraph of K_n induced by the bicolored edges contains a graph in $\mathcal{H}(G)$, say H. Starting from H, we can complete with other edges to construct a copy of G. This copy has at least half its edges that are bicolored, and thus we can make it balanced. Hence, we can find a balanced copy of G, proving the upper bound on $\mathrm{bal}^*(n, G)$. □

Theorem 2.29 gives a general upper bound on the generalized balancing number of graphs which, by Proposition 2.26, is only relevant when the balancing number does not exist. This fairly general bound may be tight (up to the order of the second term $\mathrm{ex}(n, \mathcal{H}(G))$), as is the case of K_5. In Section 2.5.2, we will use this result to give an upper bound on the generalized balancing number of K_5, which will be matched with a lower bound of the same order up to a relevant term. However this general upper bound can also be far from the exact value of the generalized balancing number, as can be noticed already in Section 2.5.1, where we determine the generalized balancing number of the unbalanceable cycles.

2.5.1 The balancing number and generalized balancing number of cycles

In Corollaries 2.2 and 2.7, it was proved that the cycle C_{4k} is balanceable while the cycle C_{4k+2} is not. We note here that all odd cycles are also balanceable, and provide exact values for their balancing number. Moreover, we give tight bounds, up to the first order term, for the balancing number of cycles of length $4k$, for $k \geq 1$. Finally, we determine the generalized balancing number of C_{4k+2}, for $k \geq 1$.

Balanceable cycles

We now explore the family of balanceable cycles. Their balanceability has been discussed in Section 2.1. Theorem 2.31 provides the balancing number for odd size-cycles. We can observe this result to be a direct consequence of the balanceability of paths of even length. We denote the path on ℓ edges by P_ℓ; the exact values of $\mathrm{bal}(n, P_\ell)$ are found in [13, Theorem 3.7].

Theorem 2.30. *[13] Let $k \geq 2$ and n be integers with k even and such that $n \geq \frac{9}{32}k^2 + \frac{1}{4}k + 1$. Then*

$$\mathrm{bal}(n, P_{k+1}) = \mathrm{bal}(n, P_k) = \begin{cases} (\frac{k-2}{4})n - \frac{k^2}{32} + \frac{1}{8}, & \textit{for } k \equiv 2 \pmod 4, \\ (\frac{k-4}{4})n - \frac{k^2}{32} + \frac{k}{8} + 1, & \textit{for } k \equiv 0 \pmod 4, \end{cases}$$

Theorem 2.31 states the exact values of $\mathrm{bal}(n, C_{4k+\alpha})$ as a direct consequence of the previous theorem.

Theorem 2.31. *Let k be a positive integer, let $n \geq \frac{9}{2}k^2 + \frac{13}{4}k + \frac{49}{32}$, and let $\alpha \in \{-1, 1\}$. We have the following:*

$$\mathrm{bal}^*(n, C_{4k+\alpha}) = \mathrm{bal}(n, C_{4k+\alpha}) = \mathrm{bal}(n, P_{4k+\alpha-1}) = (k-1)n - \frac{1}{2}(k^2 - k - 1 - \alpha).$$

Proof. Let k be a positive integer. We first prove that the balancing number of odd cycles is equal to the balancing number of paths with one edge less, proving at the same time that odd cycles are balanceable. We will demonstrate the result only for C_{4k+1} (i.e for $\alpha = 1$) since the exact same arguments can be made for C_{4k-1}.

First, we prove that $\mathrm{bal}(n, C_{4k+1}) \leq \mathrm{bal}(n, P_{4k})$. Assume that we have a 2-coloring $R \sqcup B$ of the edges of K_n with $|R|, |B| > \mathrm{bal}(n, P_{4k})$. This implies that there is a balanced copy of P_{4k}; that is, a path with an equal number of edges in R and in B. Regardless of the color of the edge connecting the endpoints of the path, the addition of this edge to the path creates a balanced cycle. Hence, we obtain the claimed upper bound on $\mathrm{bal}(n, C_{4k+1})$.

Now, we prove that $\mathrm{bal}(n, P_{4k}) \leq \mathrm{bal}(n, C_{4k+1})$. Assume that we have a 2-coloring $R \sqcup B$ of the edges of K_n with $|R|, |B| > \mathrm{bal}(n, C_{4k+1})$. This implies that there is a balanced copy of C_{4k+1}; without loss of generality, assume that this cycle contains $2k$ red edges and $2k+1$ blue edges. The path obtained from the cycle by deleting one of the blue edges is a balanced path of length $4k$, completing the proof that $\mathrm{bal}(n, C_{4k+1}) = \mathrm{bal}(n, P_{4k})$.

Finally, since we just proved $C_{4k+\alpha}$ is balanceable, the equality $\mathrm{bal}^*(n, C_{4k+\alpha}) = \mathrm{bal}(n, C_{4k+\alpha})$ is clear from Proposition 2.26. $\square$

The problem of finding the exact value of the balancing number of cycles of length $4k$ is more challenging; Theorem 2.32 below gives an upper and a lower bound for $\mathrm{bal}(n, C_{4k})$ which are tight up to the first term, $(k-1)n$; note that, contrary to the case of odd-length cycles, we need additional edges in each color class (of the order of k^2) in order to give an upper bound of $\mathrm{bal}(n, C_{4k})$ and therefore guarantee a balanced copy of C_{4k}.

Theorem 2.32. *Let k be a positive integer. For $n \geq \frac{9}{2}k^2 + \frac{13}{4}k + \frac{49}{32}$, we have the following:*

$$(k-1)n - (k-1)^2 \leq \mathrm{bal}^*(n, C_{4k}) = \mathrm{bal}(n, C_{4k}) < (k-1)n + 12k^2 + 3k.$$

The next two lemmas directly prove the bounds of Theorem 2.32 for $\mathrm{bal}(n, C_{4k})$. Those bounds show that the cycles C_{4k} are balanceable, which, in turn, allows us to invoke Proposition 2.26 in order to have the equality $\mathrm{bal}^*(n, C_{4k}) = \mathrm{bal}(n, C_{4k})$.

First, we show that there is a natural 2-coloring avoiding any balanced cycle of length $4k$ which provides us with a lower bound for $\mathrm{bal}(n, C_{4k})$.

Lemma 2.33. *For any $n \geq 4k$, we have* $\mathrm{bal}(n, C_{4k}) \geq (k-1)n - (k-1)^2$.

Proof. We will give a 2-coloring $R \sqcup B$ of the edges of K_n with $|B| \geq |R| = (k-1)n - (k-1)^2$ and without a balanced copy of C_{4k}. To this aim, let $V(K_n) = V_1 \sqcup V_2$, where $|V_1| = k-1$ and $|V_2| = n-k+1$, and we color the edges in $E(V_1, V_2)$ with red, and the remaining edges get the color blue. This coloring satisfies $|R| = (k-1)(n-k+1) = (k-1)n - (k-1)^2$, and it is not difficult to verify that $|B| \geq |R|$ for any $k \geq 1$ and $n \geq 4k$. Furthermore, any $4k$-cycle in this coloring can have at most $k-1$ vertices in V_1 and thus have at most $2k-2$ red edges. It follows that we cannot get a balanced copy of C_{4k}, which implies that $\mathrm{bal}(n, C_{4k}) \geq (k-1)(n-k+1)$. $\square$

Similar to the proof idea of Theorem 2.31, an upper bound for $\mathrm{bal}(n, C_{4k})$ can be given by constructing a balanced copy of C_{4k} from a balanced copy of C_{4k-1}. We will show that, if this construction is not possible, then some structure for the 2-coloring of the edges of K_n is forced, in which case we are able to find, in turn, a balanced $4k$-cycle by means of a long red path that is glued together with a long blue path, together with some extra edges that close the cycle. To guarantee the existence of the long red path, we make use of the extremal number for paths. Note that the additional edges (namely $(k-1)n + 12k^2 + 3k - \mathrm{bal}(n, C_{4k-1})$) are necessary to guarantee this extremal number is exceeded.

Lemma 2.34. *For $n \geq \frac{9}{2}k^2 + \frac{13}{4}k + \frac{49}{32}$, we have* $\mathrm{bal}(n, C_{4k}) \leq (k-1)n + 12k^2 + 3k$.

Proof. For $k \geq 1$ we have that $\frac{9}{2}k^2 + \frac{13}{4}k + \frac{49}{32} \geq 10k - 1$; thus we may assume that we can apply Theorem 2.31 and that $n \geq 10k - 1$, which is a sufficient assumption on n for all forthcoming arguments.

We first verify that the condition $\min\{|R|, |B|\} \geq (k-1)n + 12k^2 + 3k$ is satisfiable; this is $\binom{n}{2} \geq 2(k-1)n + 24k^2 + 6k$. Rearranging the terms, it suffices to verify that $n(n - 4k + 3) \geq 48k^2 + 12k$. Indeed, since $n \geq 10k - 1$ we have,

$$n(n - 4k + 3) \geq (10k - 1)(6k + 2) = 60k^2 + 14k - 2 \geq 48k^2 + 12k;$$

so we may consider any 2-coloring of the edges of K_n with $|R|, |B| > (k-1)n + 12k^2 + 3k$.

We now prove the lemma by contradiction. Let $R \sqcup B$ be a 2-coloring of the edges of K_n with $|R|, |B| \geq (k-1)n + 12k^2 + 3k$. Assume that this coloring has no copy of a balanced $4k$-cycle. By Theorem 2.31 there is a balanced copy C of C_{4k-1} in this coloring. Without loss of generality, we may assume that C consists of a cycle with $2k$ blue edges and $2k - 1$ red edges. This implies that C has, at some place, a red edge followed by two blue edges: say C has consecutive vertices u_0, u_1, u_2, u_3 where $u_0u_1 \in R$ and $u_1u_2, u_2u_3 \in B$.

Let $V = V(K_n)$ and $W = V(C)$. In what follows we will infer a structure among the vertices in $V \setminus W$ which will lead to a contradiction to the initial assumption that there is no balanced $4k$-cycle. The next three claims stem from the fact that some specific structure outside of C would give a balanced $4k$-cycle. Let X (resp. Y) correspond to the sets of vertices $v \in V \setminus W$ such that $u_1v \in R$ (resp. $u_1v \in B$). Note that $V \setminus W = X \cup Y$, though either X or Y may be empty. We will now strengthen the structure with three claims.

Claim 1. *For each $v \in Y$, $u_iv \in B$ for all $1 \leq i \leq 3$.*

Proof of Claim 1. Let v be a vertex in Y. By definition $u_1v \in B$. If $u_2v \in R$, then we may extend C by replacing the edge u_1u_2 with the path u_1vu_2 to obtain a balanced $4k$-cycle, a contradiction (see Figure 2.9a). It follows that $u_2v \in B$. Now, applying a similar argument (replacing u_2u_3 with the path u_2vu_3), we can conclude that $u_3v \in B$. $\triangleright$

Claim 2. *For each $v \in X$, $u_0v \in B$ and $u_2v \in R$.*

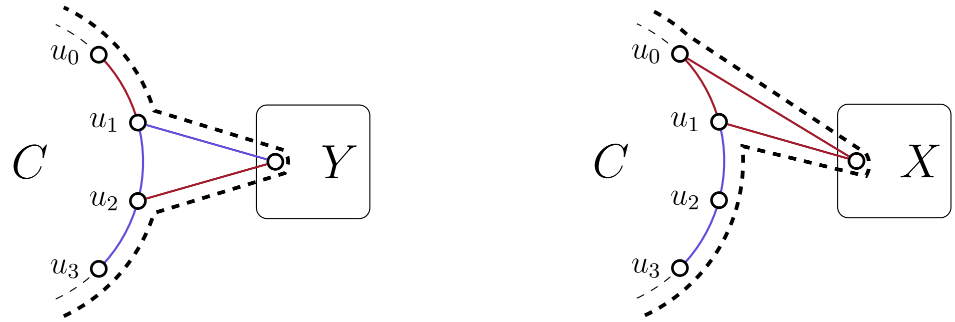
Proof of Claim 2. Let v be a vertex in X. By definition, $u_1v \in R$. The same argument than in the proof of Claim 1 gives that $u_2v \in R$. Now, assume by contradiction that $u_0v \in R$. Then, we may extend C by replacing the edge u_0u_1 with the path u_0vu_1, and thus obtain a balanced $4k$-cycle, a contradiction (see Figure 2.9b). $\triangleright$

We now have a more constrained structure, which is depicted on Figure 2.10.

Claim 3. *We have $E(X, Y) \subseteq R$, and $E(X) \cup E(Y) \subseteq B$.*

Proof of Claim 3. First, assume by contradiction that there are $v, v' \in X$ such that $vv' \in R$. By Claim 2, the path $u_0vv'u_2$ consists of two red edges and one blue edge; thus we may extend C by replacing $u_0u_1u_2$ with the path $u_0vv'u_2$ to obtain a balanced $4k$-cycle, a contradiction (see Figure 2.11a). It follows that $vv' \in B$ for all $v, v' \in X$.

Next, assume by contradiction that there are $v, v' \in Y$ such that $vv' \in R$. By Claim 1, the path $u_1vv'u_3$ consists of two blue edges and one red edge; thus we may extend C by



(a) The proof of Claim 1: if, for any $v \in Y$, $u_2v \in R$, then we can alter C to construct a balanced $4k$-cycle.

(b) The proof of Claim 2: if, for any $v \in X$, $u_0v \in R$, then we can alter C to construct a balanced $4k$-cycle.

Figure 2.9: Strengthening the structure: Claims 1 and 2.

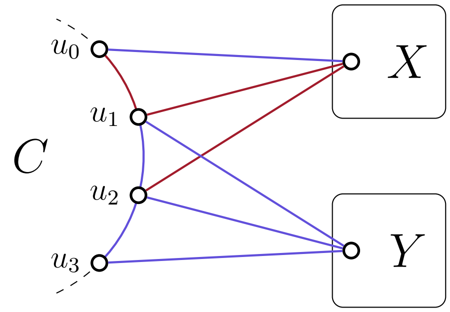


Figure 2.10: The structure after Claims 1 and 2. All the edges from the u_is to vertices in X and Y follow this structure.

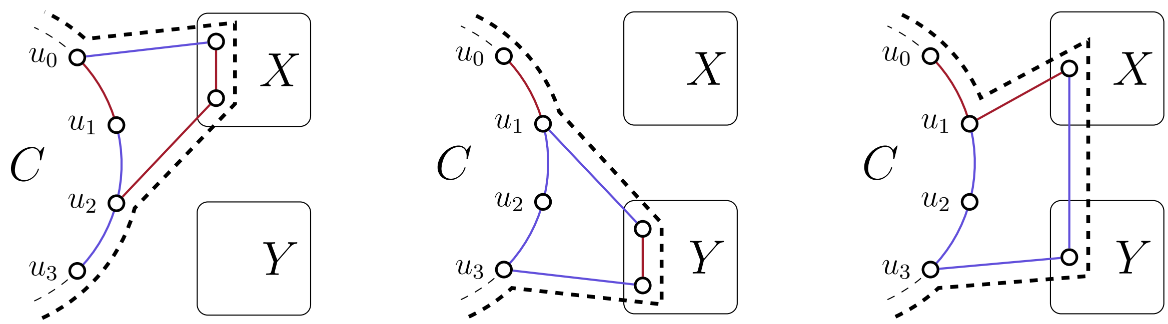
replacing $u_1u_2u_3$ with the path $u_1vv'u_3$ to obtain a balanced $4k$-cycle, a contradiction (see Figure 2.11b). It follows that $vv' \in B$ for all $v, v' \in Y$.

Finally, assume by contradiction that there are $v \in X$ and $v' \in Y$ such that $vv' \in B$. Since $u_1v \in R$ (by definition) and $u_3v' \in B$ (by Claim 1), we can replace the path $u_1u_2u_3$ by the path $u_1vv'u_3$, and obtain a balanced $4k$-cycle, a contradiction (see Figure 2.11c). It follows that for each $v \in X$ and $v' \in Y$, $vv' \in R$. ▷

We now use the structure we found in Claim 3 to find a contradiction. Recall that $n \geq 10k-1$ which implies that $|X \cup Y| = |V \setminus W| \geq n-4k+1 \geq 6k$ and so $\max\{|X|, |Y|\} \geq 3k$.

For the remainder of the proof, we have two possibilities: either $|X| \leq |Y|$ or $|X| > |Y|$. However, note that those two cases are symmetrical since we will not care about the specific colors of the edges between the vertices u_0, u_1, u_2, u_3 and $X \cup Y$, but rather more in general within and between the sets W, X, and Y.

Hence, without loss of generality, we assume that $|X| \leq |Y|$. This condition will imply $|X| < k$. Indeed, assume by contradiction that $|X| \geq k$. Then, we can obtain a balanced $4k$-cycle by taking a blue path of length k within Y, and then a red path of length k closing the cycle by going back and forth $2k$ times between X and Y (which is possible since $|Y| \geq 3k$ and $|X| \geq k$). This contradiction implies that $|X| < k$.



(a) There are no red edges in X.

(b) There are no red edges in Y.

(c) There are no blue edges between X and Y.

Figure 2.11: Strengthening the structure: Claim 3. In every case, we can use C to get a balanced $4k$-cycle, a contradiction.

Thus, we have a partition $V = Y \sqcup (W \cup X)$ where all the edges within Y are blue and $|Y| \geq n - 5k + 2$ (since $|X \cup Y| = n - 4k + 1$ and $|X| \leq k - 1$). We will now study the number of red edges in $E(Y, W \cup X)$. Let H be the bipartite graph induced by the set of red edges contained in $E(Y, W \cup X)$. We start by giving a lower bound on the number of edges in H, which is the number of red edges in K_n minus the number of red edges in $E(W \cup X)$; recall that $E(X) \cup E(Y) \subseteq B$ by Claim 3. Hence, we have:

$$|E(W \cup X) \cap R| \leq e(W) + e(W, X) < \binom{4k-1}{2} + k(4k-1) = (4k-1)(3k-1)$$

and so, we have:

$$\begin{aligned} e(H) = |R| - |E(W \cup X) \cap R| &\geq (k-1)n + 12k^2 + 3k - (4k-1)(3k-1) \\ &\geq (k-1)n + 10k - 1. \end{aligned}$$

However, $\mathrm{ex}(n, P_{2k-1}) \leq (k-1)n$ (see Theorem 5.5 in [38]), which means that, in a graph with n vertices and at least $(k-1)n$ edges, there is a path on $2k-1$ edges. As a consequence, there is a path P of length $2k-1$ edges in H. Since P has an even number of vertices, we may assume that $P = v_1 w_1 v_2 \ldots w_{k-1} v_k w_k$ with all $v_i \in Y$ and all $w_i \in W \cup X$.

Let $H' = (Y', X')$ be the subgraph of H induced by $Y' = Y \setminus \{v_1, \ldots, v_k\}$ and $X' = (W \cup X) \setminus \{w_1, \ldots, w_{k-1}\}$. Observe that $|X'| = |W \cup X| - (k-1) \leq 4k-1$ and using the lower bound on $e(H)$, we get

$$\begin{aligned} e(H') &= e(H) - e(Y, \{w_1, \ldots, w_{k-1}\}) - e(W \cup X, \{v_1, \ldots, v_k\}) \\ &\geq e(H) - (n - 5k + 2)(k-1) - (4k-1)k \\ &= e(H) - (k-1)n + k^2 - 6k + 2 \\ &\geq (k-1)n + 10k - 1 - (k-1)n + k^2 - 6k + 2 \\ &> 4k. \end{aligned}$$

It follows that $e(H') > |X'|$ and, by the pigeonhole principle, there is a vertex $w \in X'$ that has two neighbors v and v' in Y'.

However, this allows us to construct a balanced $4k$-cycle. Indeed, start from v_1 and take the path P all the way to v_k (this gives us $2k-2$ red edges, then go to v and take the path vwv' (this gives one blue edge and two red edges), and finally take a path of $2k-1$ blue edges that ends back in v_1 and using vertices $x_1, \ldots, x_{2k-2}$ in Y'. Note that we can select the x_i's as distinct from the v_i's and from v and v', since $|Y| \geq 3k$. This cycle, shown on Figure 2.12, has $2k$ edges in each color class, thus we have a contradiction.

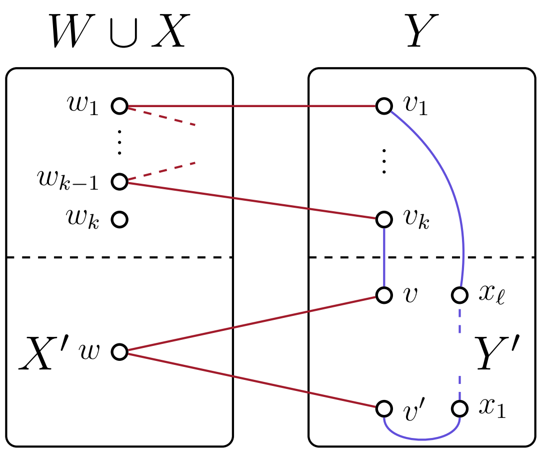


Figure 2.12: Constructing a balanced $4k$-cycle by using the structure between $W \cup X$ and Y (we have $\ell = 2k-2$).

This contradiction proves the lemma.

$\square$

Non-balanceable cycles

In this section, we obtain the exact value of the generalized balancing number for C_{4k+2}, for $k \geq 1$, which represents the class of non-balanceable cycles. This case is remarkable because it suffices that each color class covers at least half the edges in K_n plus one additional edge (*i.e.*, the coloring has a color excess of 1), which implies the existence of at least two bicolored edges regardless of the coloring. Moreover, the construction of the balanced cycle in Theorem 2.36 uses the existence of at most one bicolored edge, justifying the heuristic that the generalized balancing number (when it is at least $\frac{1}{2}\binom{n}{2}$) provides a measure of how close the graph is to being balanceable.

We note that the general upper bound from Theorem 2.29 gives us a sufficient condition for the existence of $\frac{1}{2}\binom{n}{2} + \frac{kn}{2} + \mathcal{O}(1)$ bicolored edges to guarantee a balanced C_{4k+2}. Indeed, consider $\mathcal{H} = \mathcal{H}(C_{4k+2})$, that is, the family of linear forests (i.e. unions of disjoint paths) on $2k+1$ edges. Noting that $\mathrm{ex}(n, \mathcal{H})$ is equal to the extremal number of the family containing all linear forests on n vertices and $2k+1$ edges, Theorem 1.5 in [52] states that, for n sufficiently large:

$$\mathrm{ex}(n,\mathcal{H}) = \binom{n}{2} - \binom{n-k}{2} = kn - \frac{k(k+1)}{2}$$

Hence, Theorem 2.29 yields

$$\mathrm{bal}^*(n, C_{4k+2}) \leq \frac{1}{2}\binom{n}{2} + \left\lceil \frac{kn}{2} - \frac{k(k+1)}{4} \right\rceil = \frac{1}{2}\binom{n}{2} + \frac{kn}{2} + \mathcal{O}(1).$$

We base the construction of the balanced cycle on the existence of one of two substructures called unavoidable patterns, which were discussed in Section 1.4, particularly in Theorem 1.26. They are also closely related to the characterization of balanceable graphs; see [13, Theorem 2.4]. In fact, one of these substructures, if it is large enough, naturally contains a balanced C_{4k+2}; that is, the color of the bicolored edges may be established before looking for the balanced cycle. However, for the second substructure, the construction of the balanced cycle uses a bicolored edge to leverage the fact that C_{4k+2} is not balanceable; that is, in such case the balanced copy always includes a bicolored edge.

Theorem 2.35. *[13] Let t be a positive integer. For n sufficiently large, there exists a positive integer $m = m(t)$ such that*

$$\varphi(n,t) = \mathcal{O}(n^{2-\frac{1}{m}}).$$

Recall that $\mathcal{F}_t$ is the family of all 2-edge colored copies of K_{2t} where either one color class induces a copy of K_t or one color class induces two disjoint copies of K_t. Also recall that a 2-edge-coloring of K_{2t} is a *type-A* if the edges of one of the colors induces a complete graph K_t and it is a *type-B* if the edges of one of the colors induces two disjoint K_t's (or equivalently one $K_{t,t}$).

Theorem 2.36. *Let k be a positive integer. For n sufficiently large, we have* $\mathrm{bal}^*(n, C_{4k+2}) = \frac{1}{2}\binom{n}{2}$.

Proof. Since C_{4k+2} is not balanceable, by Corollary 2.27, we have $\mathrm{bal}^*(n, C_{4k+2}) \geq \frac{1}{2}\binom{n}{2}$. To prove the equality, we simply have to consider a 2-edge covering of K_n inducing color classes R and B where $|R|, |B| \geq \frac{1}{2}\binom{n}{2} + 1$ and find a balanced copy of C_{4k+2}. Note that there are at least 2 bicolored edges in the 2-edge covering. Let t be an integer verifying $t \geq 3k+1$.

For the first step, let us ignore the fact that we have bicolored edges: every bicolored edge is set to a fixed color, making sure that both color classes remain balanced and thus contain half (± 1) the edges of K_n. This allows us to apply Theorem 2.35, which ensures that, within K_n, there is a copy H of K_{2t} such that there is a partition of its vertex set $V(H) = X \sqcup Y$ such that $|X| = |Y| = t$, and one of the following hold:

- *H is of type-A: $E(X) \subseteq R$, and $E(Y) \cup E(X,Y) \subseteq B$ (or vice-versa).*
- *H is of type-B: $E(X) \cup E(Y) \subseteq R$, and $E(X,Y) \subseteq B$ (or vice-versa).*

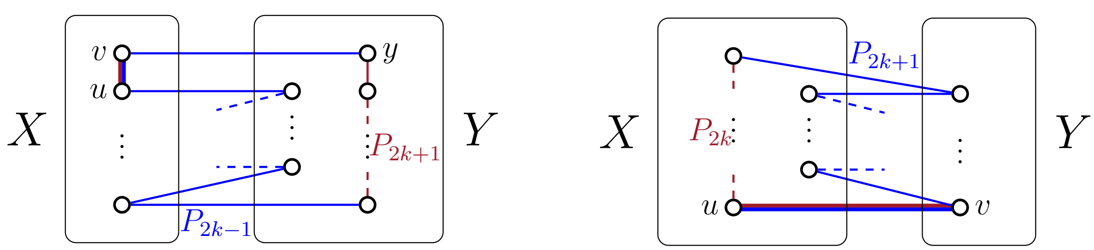
Now, let $e \in R \cap B$ be one of the bicolored edges (the other one will not be needed at all). We prove that whichever type of copy of K_{2t} exists and wherever the bicolored edge e is in K_n, we can find a balanced copy of C_{4k+2}. There are four cases to consider.

Case 1: *H is of type-A.* In this case, it is possible to construct the following balanced $(4k+2)$-cycle: follow a red path of length $2k+1$ in X, then go to Y through a blue edge, follow a blue path of length $2k-1$ in Y, and finally close the cycle by going back to the first vertex that we used in X (using another blue edge). Note that we did not make use of any bicolored edge in this case.

Case 2: *H is of type-B and $e \in E(H)$.* Then either $e \in E(X)$ (or $e \in E(Y)$, but this is symmetric), or $e \in E(X, Y)$. Let $e = uv$. Both subcases are depicted on Figure 2.13.

- *Subcase 2.1: The bicolored edge $e \in E(X)$.* Let $y \in Y$. We construct a cycle of length $4k+2$ starting with the path uvy, following with a red path of length $2k+1$ with all its vertices in Y, and then we alternate between Y and X passing through $2k-1$ blue edges and closing the cycle in u. This cycle has $2k$ blue edges, $2k+1$ red edges, as well as the bicolored edge e. By considering the bicolored edge as being blue, we have a balanced copy of C_{4k+2}. This is depicted on Figure 2.13a.

- *Subcase 2.2: The bicolored edge $e \in E(X, Y)$.* Say $u \in X$ and $v \in Y$. We construct the following cycle: starting from vertex u, we go to v through the bicolored edge, then alternate between Y and X following a blue path of length $2k+1$, and complete the cycle with a red path of length $2k$ inside Y that ends in u. This cycle has $2k+1$ blue edges, $2k$ red edges, and the bicolored edge e. By considering the bicolored edge as being red, we have a balanced copy of C_{4k+2}. This is depicted on Figure 2.13b.



(a) Subcase 2.1: we consider the bicolored edge uv as having the color b.

(b) Subcase 2.2: we consider the bicolored edge uv as having the color r.

Figure 2.13: Illustration of **Case 2** of the proof. The bicolored edge is depicted thick and with both colors.

Case 3: *H is of type-B, and $e = uv$ with $u \in V(H)$ and $v \in V \setminus V(H)$.* Assume, without loss of generality, that $u \in X$. We construct the following path of length $4k$: starting from u, take a red path of length $2k$ in X, and complete it with a blue path of length $2k$ alternating vertices between X and Y. Let $w \in X$ be the last vertex of this path, we close the cycle with the path wvu. Now, if $vw \in R$ (resp. $vw \in B$), then we consider the bicolored edge as being in B (resp. in R). The cycle we constructed has $2k+1$ edges of each color class, and thus it is a balanced copy of C_{4k+2}.

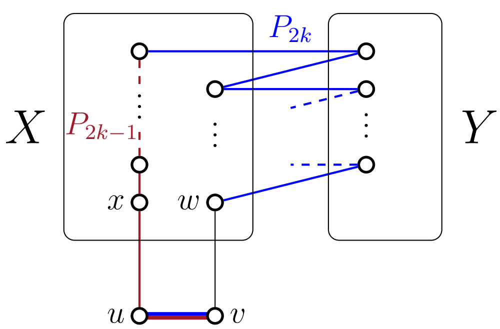
Case 4: *H is of type-B, and $e = uv$ with $u, v \in V \setminus V(H)$.* There are two possible subcases to study. Both subcases are depicted on Figure 2.14.

- *Subcase 4.1: There is a red edge ux (or vx) for some $x \in X$.* We construct the following cycle: for any vertex $w \in X \setminus \{x\}$, take the 3-path $wvux$, then follow with a red path of length $2k-1$ in X, and, alternating between X and Y, close with a blue path of length $2k$ that ends in w.

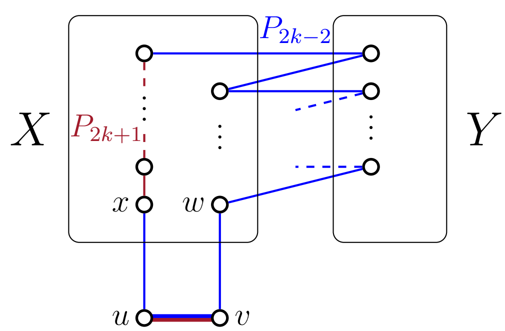
  This cycle has $2k$ red edges and $2k$ blue edges between that are different from uv and vw. If vw is red (resp. blue), then we consider uv as being blue (resp. red) and have a balanced copy of C_{4k+2}. This is depicted on Figure 2.14a.

- *Subcase 4.2: All edges ux and vx are blue for every $x \in X$.* For any two vertices $x, w \in X$, we construct the following cycle: take the 3-path $wvux$, continue with a red path of length $2k+1$ in X, then alternate vertices between X and Y building a blue path of length $2k-2$ that finishes in w and closes the cycle (if $k = 1$, just take w as the last vertex of the red path).

  This cycle has $2k+1$ red edges, $2k$ blue edges and the bicolored edge e, that can be considered as being blue. Hence, we have a balanced copy of C_{4k+2}. This is depicted on Figure 2.14b.



(a) Subcase 4.1: we consider the bicolored edge uv as being in a color class different from wv.



(b) Subcase 4.2: we consider the bicolored edge uv as having the color b.

Figure 2.14: Illustration of **Case 4** of the proof on C_{14}. The bicolored edge is depicted thick and with both colors.

All the cases have been covered: if there is a bicolored edge in K_n and there is a copy H of K_{2t} of type A or B, then we can find a balanced copy of C_{4k+2}, which proves the result. □

2.5.2 The generalized balancing number of K_5

Using the characterization of balanceable graphs, it was proved that K_5 is not balanceable in Proposition 2.10. In this section, we provide lower and upper bounds for the generalized balancing number of K_5; surprisingly, these bounds are matching up to the relevant term. To clarify, trivially, $\mathrm{bal}^*(n, K_5) \geq \frac{1}{2}\binom{n}{2}$ and the estimates we obtain in the next theorem have an additional term of order $\mathrm{ex}(n, \{C_3, C_4, C_5\}) = \Theta(n^{\frac{3}{2}})$. This implies that guaranteeing a balanced copy of K_5 requires a remarkably high color excess, implying that we always need a very high amount of bicolored edges.

Theorem 2.37. *Let $c = 2\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)^{\frac{5}{2}}$. For any $\varepsilon > 0$ and n sufficiently large, we have*

$$\frac{1}{2}\binom{n}{2} + (1-\varepsilon)cn^{\frac{3}{2}} \leq \mathrm{bal}^*(n, K_5) \leq \frac{1}{2}\binom{n}{2} + (1+\varepsilon)\frac{1}{4\sqrt{2}}n^{\frac{3}{2}}.$$

Observe that $c \approx 0.016$ while $\frac{1}{4\sqrt{2}} \approx 0.177$. The proof of Theorem 2.37 follows directly from Corollary 2.40 and Lemma 2.41 that we state and prove below. For both arguments we focus on the structure of the graph induced by the bicolored edges, where we take into account the girth and the edge number. Recall that, for a graph G, the length of a smallest cycle in G is called the *girth* and is denoted by $g(G)$; if G has no cycles, then its girth is defined to be infinity. Throughout this section we rely on $\mathrm{ex}(n, \{C_3, C_4, C_5\})$, the extremal number for graphs of girth at least 6; more precisely, we exploit that $\mathrm{ex}(n, \{C_3, C_4, C_5\})$ is strictly increasing on n and that

$$\mathrm{ex}(n, \{C_3, C_4, C_5\}) = (1 + o(1))\frac{1}{2\sqrt{2}}n^{\frac{3}{2}}, \tag{2.1}$$

where the asymptotic expression is given in Theorem 4.5 of [38] also stated in the following theorem.

Theorem 2.38. *[38] For $k = 2, 3$ and 5 as $n \to \infty$ we have*

$$\mathrm{ex}(\{C_3, C_4, \ldots, C_{2k+1}\}) = (1 + o(1))\frac{1}{2^{1+(1/k)}}n^{1+(1/k)}.$$

To verify that $\mathrm{ex}(n, \{C_3, C_4, C_5\})$ is strictly increasing on n, let $m_n = \mathrm{ex}(n, \{C_3, C_4, C_5\})$, and take a graph G on $n-1$ vertices and $m_{n-1} - 1$ edges with girth at least 6. Then we may construct a graph G' on n vertices and m_{n-1} edges with girth at least 6 by just adding to G a new vertex connected by an edge to any of the vertices in G. This proves that $m_{n-1} + 1 \leq m_n$.

By iteratively applying this argument, it follows more generally that $\mathrm{ex}(n - k, \{C_3, C_4, C_5\}) \leq \mathrm{ex}(n, \{C_3, C_4, C_5\}) - k$ $(m_{n-k} \leq m_n - k)$.

For the upper bound in Theorem 2.37, we use Theorem 2.29, which boils down to analysing $\mathrm{ex}(n, \mathcal{H}(K_5))$; this is done in the following theorem,

where we show that $\mathrm{ex}(n, \mathcal{H}(K_5)) = \mathrm{ex}(n, \{C_3, C_4, C_5\})$.

Theorem 2.39. *For $n \geq 5$, we have* $\mathrm{ex}(n, \mathcal{H}(K_5)) = \mathrm{ex}(n, \{C_3, C_4, C_5\})$.

Proof. Let $\mathcal{H} = \mathcal{H}(K_5)$, that is, the family of subgraphs of K_5 that have 5 edges and no isolates. Observe that $\mathcal{H}$ contains precisely six graphs; namely, the 5-cycle, the 4-pan[1] (also called P, or the banner), its complementary $\overline{P}$, the bull, the cricket and the diamond. Those are depicted in Figure 2.15.

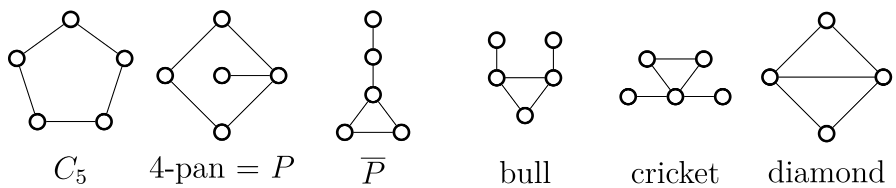


Figure 2.15: The family $\mathcal{H}(K_5)$.

Observe that every graph from $\mathcal{H}$ has either a C_3, a C_4, or a C_5. Hence, the class of graphs of order n having girth at least 6 is contained in the class of the $\mathcal{H}$-free graphs of order n. This implies directly that $\mathrm{ex}(n, \mathcal{H}(K_5)) \geq \mathrm{ex}(n, \{C_3, C_4, C_5\})$.

We will prove now the other inequality, that is, that every graph on n vertices and more than $\mathrm{ex}(n, \{C_3, C_4, C_5\})$ edges contains a subgraph from $\mathcal{H}$.

We use an induction argument. First, let us observe that $m_5 = 4$, $m_6 = 6$, $m_7 = 7$ and $m_8 = 9$ since the maximal graphs on n vertices of girth at least 6 are, respectively: spanning trees ($n = 5$), C_6 ($n = 6$), C_7 and the 6-pan ($n = 7$), and finally, the graph that consists of vertices a, b, c, d, e, f, e', f' that are arranged in two cycles $abcdefa$ and $abcde'f'a$ ($n = 8$).

Consider $n \geq 5$ and any graph F on n vertices and with at least $m_n + 1$ edges; that is, F has girth at most 5. We will prove that it contains a subgraph in $\mathcal{H}$. We start with the base cases $n \in \{5, 6, 7, 8\}$. Let F' be a subgraph of F on exactly $m_n + 1$ edges. We will prove that F', which has girth at most 5, contains a subgraph in $\mathcal{H}$ (and hence F does):

1. If $n = 5$, then F' has 5 edges, and so must have a subgraph in $\mathcal{H}$.

2. If $n = 6$, then F' has 7 edges. Suppose every set of 5 vertices in F' induces a graph of at most 4 edges. Since $e(F') = 7$, the vertex not contained in a given 5-set has to have degree at least 3. But this happens to every set of 5 vertices. Hence, $2e(F') \geq 6 \cdot 3 = 18$, implying that $e(F') \geq 9$, a contradiction. Hence, there is a 5-vertex set inducing a graph on at least 5 edges in F' and thus F' contains a subgraph from $\mathcal{H}$.

3. If $n = 7$, then F' has 8 edges and contains at least an induced cycle of length at most 5. If F' contains an induced C_5, then it trivially contains a subgraph from $\mathcal{H}$. If F' contains an induced C_4, then since there are at least four remaining edges and only three remaining vertices, this implies that at least one vertex from the 4-cycle has a neighbour among the other three vertices, which in turn implies that F' contains a 4-pan, which is in $\mathcal{H}$. If F' contains a triangle, then there are three cases: first, there are at least two edges between the triangle and the remaining vertices, and F' contains

[1] The n-pan is an n-cycle with a pendant edge attached to a vertex of the cycle.

a bull, a cricket, or a diamond, which are in $\mathcal{H}$; second, there is no edge between the triangle and the 4 remaining vertices, which implies that they must induce a diamond, which is in $\mathcal{H}$; finally, if there is exactly one edge between the triangle and one of the remaining vertices, say u, then u has to have a neighbour in the other remaining vertices (since otherwise there would be four edges among three vertices, which is impossible), and F' contains the complement of a 4-pan, which is in $\mathcal{H}$.

4. If $n = 8$, then F' has 10 edges and contains an induced cycle of length at most 5. If F contains an induced C_5, then it trivially contains a subgraph from $\mathcal{H}$. If F' contains an induced C_4, then there are two cases: either there is at least one edge between the 4-cycle and the remaining vertices, and thus F' contains a 4-pan, which is in $\mathcal{H}$; or the four remaining vertices have to induce a diamond, which is in $\mathcal{H}$. If F' contains a triangle, then there are two cases: either there are at least 2 edges between the triangle and the remaining vertices, and thus F' contains either a bull or a cricket, which are in $\mathcal{H}$; or the five remaining vertices have at least 6 edges, and by the argument in the case $n = 5$ implies that F' contains a subgraph in $\mathcal{H}$.

For the induction step, we will use the following general argument. Suppose that F is a graph on n vertices and at least $m_n + 1$ edges; if F' may be constructed from F by removing k vertices and k edges, then F' has (also) girth at most 5 since it has at least $m_n - k \geq m_{n-k}$ edges (recall that $m_n = \text{ex}(n, \{C_3, C_4, C_5\})$ is strictly increasing).

If $n \geq 9$ and $k \leq 4$ we may apply the induction hypothesis and infer that F' contains a subgraph in $\mathcal{H}$ (and so does F). In what follows we refer to this argument as the *removal induction hypothesis*.

Now, assume that $n \geq 9$. First, if F contains a vertex of degree 1, then we can remove it and apply the removal induction hypothesis. Thus, we may assume that F has minimum degree at least 2. Recall that F has girth at most 5. There are three cases to consider:

Case 1: *F has girth 5.* Naturally, F contains a subgraph in $\mathcal{H}$; namely, C_5.

Case 2: *F has girth 4.* We may consider a C_4 in F. If all four vertices have degree 2, then we can remove them from F and apply the removal induction hypothesis. Otherwise, at least one of them has a third neighbour; the cycle together with such neighbor forms a 4-pan, that is F contains a subgraph of $\mathcal{H}$.

Case 3: *F has girth 3.* We may consider a C_3 in F. If all three vertices have degree 2, then likewise we can apply the removal induction hypothesis. Otherwise, at least one of them has a third neighbour u. However, u has degree at least 2, so u itself has another neighbour; we then obtain either the diamond or the complement of the 4-pan as a subgraph, both of which are in $\mathcal{H}$.

This proves that F contains a subgraph in $\mathcal{H}$ whenever it has more than $m_n = \text{ex}(n, \{C_3, C_4, C_5\})$ edges. Hence $\text{ex}(n, \mathcal{H}) \leq \text{ex}(n, \{C_3, C_4, C_5\})$. □

By combining Theorems 2.29 and 2.39, we obtain the desired upper bound on $\text{bal}^*(n, K_5)$.

Corollary 2.40. *For any $\varepsilon > 0$ and n sufficiently large,*

$$\text{bal}^*(n, K_5) \leq \frac{1}{2}\binom{n}{2} + (1 + \varepsilon)\frac{1}{4\sqrt{2}}n^{\frac{3}{2}}.$$

Proof. Let $\varepsilon > 0$. For n sufficiently large, we have with Theorem 2.29, Theorem 2.39 and (2.1) that

$$\begin{aligned}\operatorname{bal}^*(n, K_5) &\leq \frac{1}{2}\binom{n}{2} + \left\lceil \frac{1}{2}\operatorname{ex}(n, \mathcal{H}(K_5)) \right\rceil \\ &= \frac{1}{2}\binom{n}{2} + \left\lceil \frac{1}{2}\operatorname{ex}(n, \{C_3, C_4, C_5\}) \right\rceil \\ &\leq \frac{1}{2}\binom{n}{2} + (1+\varepsilon)\frac{1}{4\sqrt{2}}n^{\frac{3}{2}}.\end{aligned}$$

□

We will now obtain a lower bound for the generalized balancing number of K_5. In Lemma 2.41, we provide a 2-edge covering of K_n where the subgraph induced by the bicolored edges is of girth at least 6. By analyzing all possible overlaps of a copy of K_5 and the bicolored edges, we prove that this 2-edge covering does not contain a balanced copy of K_5.

Lemma 2.41. *Let $c = 2\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\right)^{\frac{5}{2}}$. For any $\varepsilon > 0$ and n sufficiently large,*

$$\operatorname{bal}^*(n, K_5) \geq \frac{1}{2}\binom{n}{2} + (1-\varepsilon)cn^{\frac{3}{2}}.$$

Proof. Suppose that there are integers k, k' and m such that $k \leq k' \leq n$ and that there exists a graph H on k vertices, m edges and girth at least 6. The precise values for these integers will be specified, in terms of n and $\varepsilon > 0$ further on. First, using the assumptions above, we construct a 2-edge covering of K_n and prove that it does not contain a balanced copy of K_5.

Let us partition the vertices of K_n in two parts X and Y such that $|Y| = k'$ (and thus $|X| = n - k'$); assign the list $\{r\}$ to every edge within X; assign the list $\{r, b\}$ to m edges in Y inducing a copy of H; and finally assign the list $\{b\}$ to every other edge within Y and to every edge between X and Y.

We claim that no copy of K_5 can be balanced in this covering. First, any copy of K_5 with all its vertices in X has no blue edges and, thus, it cannot be balanced. Now, let G be a copy of K_5 with at least one vertex in Y and let x and y be the number of vertices of G in X and Y, respectively; note that $y \geq 1$. Recall that bicolored edges form a graph of girth at least 6 and so G has at most $y-1$ bicolored edges in G and precisely $\binom{x}{2}$ red (non-bicolored) edges. To conclude the proof that there is no balanced copy of K_5 observe that if $x \geq 4$ then G has at most 4 blue edges, including bicolored ones. Whereas if $x \leq 3$, then G has at most $\binom{x}{2} + y - 1 \leq x + y - 1 = 4$ red edges, including bicolored ones; thus, G may not be balanced.

It remains to prove that, given $\varepsilon > 0$, we may choose k, k' and m so that the color classes of the covering above have size at least $\frac{1}{2}\binom{n}{2} + (1-\varepsilon)cn^{\frac{3}{2}}$; as this would establish the lemma.

Fix $\varepsilon > 0$ and let $\alpha = 1 - \frac{1}{\sqrt{2}}$, $\beta = \left(1 - \frac{\varepsilon}{2}\right)\left(\frac{\alpha}{2}\right)^{\frac{3}{2}}$; then let $k = \lceil \alpha n \rceil$, $k' = \left\lceil \alpha n + \beta n^{\frac{1}{2}} \right\rceil$

and $m = \left\lfloor \beta n^{\frac{3}{2}} \right\rfloor$. Observe that

$$m = \left\lfloor \beta n^{\frac{3}{2}} \right\rfloor \leq \left(1 - \frac{\varepsilon}{2}\right)\left(\frac{\alpha}{2}\right)^{\frac{3}{2}} n^{\frac{3}{2}} \leq \left(1 - \frac{\varepsilon}{2}\right)\left(\frac{k}{2}\right)^{\frac{3}{2}} \leq \mathrm{ex}(k, \{C_3, C_4, C_5\});$$

where the last inequality holds for n large enough since $\mathrm{ex}(k, \{C_3, C_4, C_5\}) = (1 + o(1))(\frac{k}{2})^{\frac{3}{2}}$ by Theorem 2.38. This establishes the existence of a graph H with girth at least 6, as desired. Moreover, we have clearly $k \leq k' \leq n$. Next, we will show that $\max\{|R|, |B|\} > \frac{n^2}{4} + \left(1 - \frac{\varepsilon}{2}\right)\alpha\beta n^{\frac{3}{2}}$.

In the following expressions we avoid cumbersome notation by assuming that n is large enough that we may omit rounding to integers; in particular we will simply write $n - k' = (1-\alpha)n - \beta n^{\frac{1}{2}}$ (to clarify, considering the precise expression of $n - k'$ would only add, to $|R|$ and $|B|$, terms of order $O(n)$ which may be neglected).

We clearly have $|R| = \binom{n-k'}{2} + m$ and $|B| = \binom{k'}{2} + k(n - k')$. First, we consider the size of R; using that $m = \beta n^{\frac{3}{2}}$, we obtain

$$\begin{aligned}
\binom{n-k'}{2} + m &= \frac{1}{2}\left((1-\alpha)n - \beta n^{\frac{1}{2}})^2 - (1-\alpha)n + \beta n^{\frac{1}{2}}\right) + \beta n^{\frac{3}{2}} \\
&= \frac{(1-\alpha)^2 n^2}{2} + \alpha\beta n^{\frac{3}{2}} + \frac{(\beta^2 + \alpha - 1)n}{2} + \frac{\beta n^{\frac{1}{2}}}{2} \\
&> \frac{n^2}{4} + \alpha\beta n^{\frac{3}{2}} - \frac{n}{2};
\end{aligned}$$

where in the last inequality we used $1 - \alpha = \frac{1}{\sqrt{2}}$ and removed lower order positive terms. In addition, we have $\varepsilon\alpha\beta n^{\frac{1}{2}} \geq 1$ for n large enough, and so

$$\alpha\beta n^{\frac{3}{2}} - \frac{n}{2} = \left(1 - \frac{\varepsilon}{2}\right)\alpha\beta n^{\frac{3}{2}} + \frac{n}{2}\left(\varepsilon\alpha\beta n^{\frac{1}{2}} - 1\right) > \left(1 - \frac{\varepsilon}{2}\right)\alpha\beta n^{\frac{3}{2}};$$

which in turn implies that $|R| > \frac{n^2}{4} + (1 - \frac{\varepsilon}{2})\alpha\beta n^{\frac{3}{2}}$ for n sufficiently large. Similar computations for the size of B, in particular, using that $\frac{\alpha^2}{2} + \alpha(1-\alpha) = \frac{1}{4}$, yield

$$\begin{aligned}
\binom{k'}{2} + k'(n - k') &= \frac{1}{2}\left((\alpha n + \beta n^{\frac{1}{2}})^2 - (\alpha n + \beta n^{\frac{1}{2}})\right) + \left(\alpha n + \beta n^{\frac{1}{2}}\right)\left((1-\alpha)n - \beta n^{\frac{1}{2}}\right) \\
&= \frac{n^2}{4} + \beta(1-\alpha)n^{\frac{3}{2}} - \frac{(\beta^2 + \alpha)n}{2} - \frac{\beta}{2}n^{\frac{1}{2}}.
\end{aligned}$$

In this case we use that $2 - 4\alpha = \frac{4}{\sqrt{2}} - 2 > 0$, and so for n large enough we have

$$\beta(1-\alpha)n^{\frac{3}{2}} - \frac{\beta^2 n}{2} - \beta n^{\frac{1}{2}} = \alpha\beta n^{\frac{3}{2}} + \frac{\beta n^{\frac{3}{2}}}{2}(2 - 4\alpha - (\beta + \alpha\beta^{-1})n^{-\frac{1}{2}} - n^{-1}) > \alpha\beta n^{\frac{3}{2}};$$

in particular, $|B| > \frac{n^2}{4} + (1 - \frac{\varepsilon}{2})\alpha\beta n^{\frac{3}{2}}$. Finally, by definition of β, we have that

$$\left(1 - \frac{\varepsilon}{2}\right)\alpha\beta = 2\left(1 - \frac{\varepsilon}{2}\right)^2\left(\frac{\alpha}{2}\right)^{\frac{5}{2}} > (1 - \varepsilon)c;$$

which, together with $\frac{n^2}{4} \geq \frac{1}{2}\binom{n}{2}$, yields, for n large enough,

$$\max\{|R|, |B|\} > \frac{n^2}{4} + \left(1 - \frac{\varepsilon}{2}\right) \alpha\beta n^{\frac{3}{2}} \geq \frac{1}{2}\binom{n}{2} + (1-\varepsilon)cn^{\frac{3}{2}}.$$

□

2.6 Open problems

In this chapter, we stated some results concerning balanceability, balancing number and generalized balancing number. Nevertheless, there are always problems left to solve. In this section we state some problems that arose in relevance to the topics covered in this chapter.

1. Considering the complexity of the problem of determining whether a graph is balanceable or not. This problem boils down to finding an edge cut with $\lfloor \frac{1}{2}e(G) \rfloor$ or $\lceil \frac{1}{2}e(G) \rceil$ edges and an induced subgraph with $\lfloor \frac{1}{2}e(G) \rfloor$ or $\lceil \frac{1}{2}e(G) \rceil$ edges. Particularly, the problem of finding an edge cut with half of the edges is a variant of the problem EXACT-CUT (which studies the complexity of finding an edge cut with exactly k edges and it is NP-complete). The other problem deals with the existence of an induced subgraph with exactly half of the edges. We conjecture that the balancerability problem is NP-complete.

2. Determine the balanceability of K_n for $n > 2,303,999,904,000,003$ when $n \equiv 2,3 \pmod 4$.

3. What other graph families satisfy being balanceable? We are interested in determining the balanceability of graphs such as planar graphs, outerplanar graphs, chordal graphs, k-trees and other circulant graphs.

4. Study the generalized balancing number of other graphs such as K_n with $n \geq 5$, $2K_n$, circulant graphs, wheel graphs and grids.

Chapter 3

Amoebas

Amoeba graphs were first introduced in [13] by means of a graph theoretic definition. In that work, Caro, Hansberg, and Montejano determined that all amoebas are balanceable (see Theorem 1.43). Afterwards, in [12], the authors were able to endow these graphs with algebraic properties in order to achieve a better handling of them. We are interested in solving some of the open questions that the authors left in [12] and [13]. We now move on to state the relevant concepts and notation for this chapter.

When two groups P and Q are isomorphic, we will write $P \cong Q$. Let X be a finite set and let S_X be the symmetric group which consists of all permutations of elements of X. As usual, $S_n = S_{[n]}$, where $[n] = \{1, 2, \cdots, n\}$. The automorphism group of a graph G is denoted as $\mathrm{Aut}(G)$, and so any graph G of order n satisfies that $\mathrm{Aut(G)} \cong S$ for some $S \leqslant S_n$.

Lemma 3.1 states a useful property of a k-cycle in S_n.

Lemma 3.1 ([55]). *For a k-cycle $(i_1 i_2 \dots i_k)$ in S_n and an arbitrary $\sigma \in S_n$,*

$$\sigma(i_1 i_2 \dots i_k)\sigma^{-1} = (\sigma(i_1)\sigma(i_2) \dots \sigma(i_k)).$$

The next lemma on generating sets of S_n is relevant later statements results.

Lemma 3.2 ([55]). *For $n \geq 2$, S_n is generated by any of the following sets of transpositions:*

1. $\{(ij) : i, j \in \{1, 2, \dots, n\}\}$,
2. $\{(1j) : j \in \{1, 2, \dots, n\}\}$,
3. $\{(ij) : j = i + 1, i \in \{1, 2, \dots, n-1\}\}$;

or any of the following pairs of transposition and cycle:

4. $\{(12), (12 \dots n)\}\}$,

5. $\{(12), (23\ldots n)\}\}$.

Let X be a set and let G be a group with identity e. A *(left) group action* is a function $\theta : G \times X \to X$ such that for every $g \in G$ and $x \in X$ $gx = \theta(g, x) \in X$ and the following group action axioms hold:

1. For every $g, h \in G$ and $x \in X$, $g(hx) = (gh)x$
2. For every $x \in X$, $ex = x$.

In this case the group G is said to act on the set X (from the left). Consider a group G acting on a set X. The *orbit* of an element x in X is the set of elements in X to which x can be moved by elements of G. The orbit of x is denoted as Gx, where $Gx = \{gx \mid g \in G\}$. For every x in X, the *stabilizer subgroup* G_x of G with respect to x is the set of all elements in G that fix x. The action of G on X is called *transitive* if for any $x, y \in X$ there is an element $g \in G$ such that $gx = y$.

The following lemma from group theory will be of great importance in the proof of Lemma 3.10.

Lemma 3.3 ([55]). *If $x \in X$ and $S \subseteq S_{X\setminus\{x\}}$ is a set of permutations that act transitively on $X \setminus \{x\}$, and if $\varphi \in S_X$ with $\varphi(x) \neq x$, then the set $S \cup \{\varphi\}$ generates S_X.*

The next definition involves the pasting of two functions.

Definition 3.4. *Let X, Y, A, B be sets and $f : X \to Y$ and $g : A \to B$ be two functions such that $f(x) = g(x)$ for all $x \in X \cap A$, then $f \cup g : X \cup A \to Y \cup B$ is defined as*

$$f \cup g(x) = \begin{cases} f(x), & \textit{if } x \in X \\ g(x), & \textit{if } x \in A \setminus X. \end{cases}$$

Along this chapter, we will consider graphs $G = G(V, E)$ equipped with a labeling on their vertex set $\lambda : V \to X$, which will always be a bijection. We define $v_x = \lambda^{-1}(x)$, for each $x \in X$, and $L_G = \{ij \mid v_iv_j \in E(G)\}$ with no distinction between ij and ji. For each $\sigma \in S_X$, let G_σ be the copy of G on the same vertex set V defined by $E(G_\sigma) = \{v_iv_j \mid \sigma(i)\sigma(j) \in L_G\}$. Notice that each labeled copy of G on the vertex set V corresponds to a permutation $\sigma \in S_X$ and vice versa. For every graph G' on V isomorphic to G, there are $|\mathrm{Aut}(G)|$ different copies G_σ that correspond to G' and, furthermore, the group $A_G = \{\sigma \in S_X \mid G_\sigma = G\}$ is isomorphic to $\mathrm{Aut}(G)$.

Notice that the set of labels

$$L_{G_\sigma} = \{\sigma(i)\sigma(j) \mid v_iv_j \in E(G)\}$$

of the edges of G_σ is the same for all $\sigma \in S_X$. The corresponding copies of the vertices and edges of G in G_σ are given by their labels: the copy of vertex v_i of G is the vertex of G_σ

having label i, while the copy of an edge $v_iv_j \in E(G)$ is the edge of G_σ having label ij.

Given $e \in E(G)$ and $e' \in E(\overline{G})$, the graph $G - e + e'$ is obtained from G by performing the *edge-replacement* that substitutes e by e'. If $G - e + e'$ is a graph isomorphic to G, we say that the edge-replacement is *feasible*. We consider also the so-called *neutral edge-replacement* $\emptyset \to \emptyset$ as a feasible edge-replacement, which is given when no edge is replaced at all.

Let

$$R_G = \{rs \to kl \mid G - v_rv_s + v_kv_l \cong G\} \cup \{\emptyset \to \emptyset\}$$

be the set of all feasible edge-replacements of G given by their labels and let $R_G^* = R_G \backslash \{\emptyset \to \emptyset\}$. We will use sometimes the notation $e \to e' \in R_G$ when we do not require to specify the labels of the vertices involved in the edge-replacement. Notice that $R_{G_\rho} = R_G$ for any $\rho \in S_X$, because any $e \to e' \in R_G$ also represents a feasible edge-replacement of any copy G_ρ with $\rho \in S_X$.

Moreover, the set $\mathrm{Fer}_G(e \to e')$ consists of all permutations of labels that correspond to the feasible edge-replacement $e \to e'$. More precisely, for an edge-replacement $rs \to k\ell \in R_G^*$, a permutation of the labels σ is an element of $\mathrm{Fer}_G(rs \to k\ell)$ if and only if $G_\sigma = G - v_rv_s + v_kv_\ell$. For the neutral edge-replacement, we set $A_G = \mathrm{Fer}_G(\emptyset \to \emptyset) \cong \mathrm{Aut}(G)$.

We denote by $\mathrm{Fer}(G)$ the group generated by the permutations associated to all feasible edge-replacements, that is, $\mathrm{Fer}(G)$ is generated by the set

$$\mathcal{E}_G = \bigcup_{e \to e' \in R_G} \mathrm{Fer}_G(e \to e').$$

The group $\mathrm{Fer}(G)$ acts on the set $\{G_\rho \mid \rho \in S_X\}$ by $(\sigma, G_\rho) \mapsto G_{\sigma\rho}$ where $\sigma \in \mathrm{Fer}(G)$. Note that employing this action exhibits what happens when a series of edge-replacements, associated to σ, is applied on a copy G_ρ of G which results in $G_{\sigma\rho}$. Being able to go from any copy G_ρ to any other copy $G_{\rho'}$ by following a series of feasible edge replacements means that for any $\rho, \rho' \in S_X$, there exists $\sigma \in \mathrm{Fer}(G)$ such that $\rho' = \sigma\rho$, meaning that $\mathrm{Fer}(G) = S_X$. We also recall Proposition 1.39, which states that if $G \cup t_0K_1$ is a local amoeba for some $t_0 > 0$, then $G \cup tK_1$ is a local amoeba for any $t \geq t_0$. Therefore, Definition 1.38 can also be stated by means of the group $\mathrm{Fer}(G)$.

In this context, a graph G of order n is called a *local amoeba* if $\mathrm{Fer}(G) \cong S_n$, and a *global amoeba* if $\mathrm{Fer}(G \cup tK_1) \cong S_{n+t}$, where t is large enough. In words, if G is a local amoeba, then any other copy of G on the same vertex set can be reached by a chain of feasible edge-replacements. If G is a global amoeba, then there is a large enough integer $t \geq 1$ such that any copy of G embedded on a vertex set of cardinality $n + t$ can be reached from any other copy on the same vertex set by a chain of feasible edge-replacements. In fact, Theorem 3.8 states that it suffices if t is only 1.

Lemma 3.5 ([12])**.** *Let G be a graph provided with a labeling $\lambda : V(G) \to X$ on its vertices. Then* $\mathrm{Fer}(G) = \mathrm{Fer}(\overline{G})$.

Proof. It is sufficient to show that, for every feasible edge-replacement $e \to e' \in R_G$ and $\sigma \in \mathrm{Fer}_G(e \to e')$, there is a feasible edge-replacement in $\overline{G}$ given by σ, and vice-versa. If

$e \to e' = \emptyset \to \emptyset$, then it is clear that $\overline{G} \cong \overline{G_\sigma}$ for any $\sigma \in \mathrm{Fer}_G(\emptyset \to \emptyset) = A_G$. Suppose that $rs \to kl \in R^*_G$, and $\sigma \in \mathrm{Fer}_G(rs \to kl)$. Observe that

$$\overline{G} \cong \overline{G_\sigma} = \overline{G - v_r v_s + v_k v_l} = \overline{G} - v_k v_l + v_r v_s.$$

This implies that $\sigma \in \mathrm{Fer}_{\overline{G}}(kl \to rs)$ and the proof is complete. □

The next proposition is a direct consequence from Lemma 3.5.

Proposition 3.6 ([12]). *A graph G is a local amoeba if and only if its complementary graph $\overline{G}$ is a local amoeba.*

The next proposition states a characterization of local and global amoebas when the neutral edge-replacement is the only feasible edge-replacement of the graph in question.

Proposition 3.7 ([12]). *Let G be a graph of order $n \geq 1$ having the neutral edge-replacement as its only feasible edge-replacement. Then*

(i) G is a local amoeba if and only if either $G = K_n$, or $G = \overline{K_n}$;

(ii) G is a global amoeba if and only if either $G = \frac{n}{2}K_2$, for even n, or $G = \overline{K_n}$.

The following theorem provides a simpler definition for global amoeba, which will be used from now on.

Theorem 3.8 ([12]). *Let G be a non-empty graph. Let $\lambda : V(G) \to X$ be a labeling on its vertices, and let $\Gamma = \mathrm{Fer}(G)$. For each $x \in X$, let $v_x = \lambda^{-1}(x)$. The following statements are equivalent:*

(i) G is a global amoeba.

(ii) $G \cup K_1$ is a local amoeba.

(iii) For each $x \in X$, there is a $y \in \Gamma x$ such that $\deg_G(v_y) = 1$.

(iv) For each $x \in X$ such that $\deg_G(v_x) \geq 2$, there is a $\sigma \in \Gamma$ such that $\deg_G(v_{\sigma(x)}) = \deg_G(v_x) - 1$.

3.1 Recursive local amoebas

In [12], different ways of constructing local or global amoebas are presented. Further on, in [45], the authors work with recursive constructions. In particular, a general method to construct families of global amoebas is developed. Using this method, the authors give an alternative proof of the fact that certain family of Fibonacci-like trees defined in [12] consists of global amoebas. We will deal with this family in this section, too, and we will demonstrate that they are local amoebas as well, solving a problem stated in [12]. Moreover, we present a general construction that can be used to generate families of local amoebas recursively.

For a graph G provided with a labeling $\lambda : V(G) \to X$ on its vertices, consider the set $\mathcal{E}_G^i$ of all permutations associated to edge replacements in R_G that fix the label $i \in X$, that is,

$$\mathcal{E}_G^i = \mathcal{E}_G \cap \mathrm{Stab}_{\mathrm{Fer}(G)}(i).$$

Let $\mathrm{Fer}^i(G)$ be the subgroup of $\mathrm{Stab}_{\mathrm{Fer}(G)}(i)$ generated by the set $\mathcal{E}_G^i$.

Lemma 3.9. *Let H and J be two vertex disjoint graphs provided with their corresponding disjoint sets of labels X and Y. Consider vertices $v_x \in V(H)$, $v_y \in V(J)$ with labels $x \in X$ and $y \in Y$, respectively, and the graph $G = (H \cup J) + v_x v_y$ with the inherited set of labels $X \cup Y$. If $\alpha \in \mathcal{E}_H^x$, then $\alpha \cup \mathrm{id}_{\mathrm{Fer}(J)} \in \mathcal{E}_G^x$.*

Proof. Given $\alpha \in \mathcal{E}_H^x$, then there is a feasible edge replacement $e \to e' \in R_H$ such that $\alpha \in \mathrm{Fer}_H(e \to e')$. Since $\alpha(x) = x$, the role v_x is playing in H is the same as in H_α, and so the edge replacement $e \to e'$ can also be applied on G. It follows that

$$G_{\alpha \cup \mathrm{id}_{\mathrm{Fer}(J)}} = (H_\alpha \cup J) + v_x v_y \cong (H \cup J) + v_x v_y = G,$$

implying that $e \to e'$ is a feasible edge-replacement in G and $\alpha \cup \mathrm{id}_{\mathrm{Fer}(J)} \in \mathcal{E}_G^x$. $\square$

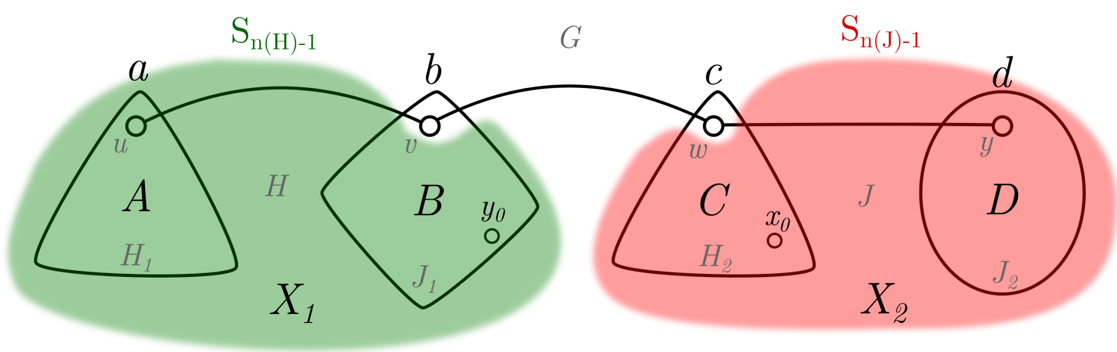


Figure 3.1: General diagram of Lemma 3.10.

Lemma 3.10. *Let H_1, J_1, H_2, J_2 be non-empty graphs that are pairwise vertex disjoint and such that there is an isomorphism $\varphi : H_1 \to H_2$. Let $u \in V(H_1), v \in V(J_1), w \in V(H_2), y \in V(J_2)$ be vertices such that $\varphi(u) = w$. We define the following graphs:*

$$H = (H_1 \cup J_1) + uv, J = (H_2 \cup J_2) + wy, \text{ and } G = (H \cup J) + vw.$$

Let $\lambda_1 : V(H) \to X_1$ and $\lambda_2 : V(J) \to X_2$ be labelings of H and J, and consider $\lambda = \lambda_1 \cup \lambda_2$ as the labeling of G. Let $\lambda(u) = a$, $\lambda(v) = b$, $\lambda(w) = c$, and $\lambda(y) = d$. If

$$\mathrm{Fer}^b(H) \cong S_{n(H)-1} \text{ and } \mathrm{Fer}^c(J) \cong S_{n(J)-1},$$

it follows that $\mathrm{Fer}^b(G) \cong S_{n(G)-1}$.

Proof. Let $\lambda(V(H_1)) = A$, $\lambda(V(J_1)) = B$, $\lambda(V(H_2)) = C$, and $\lambda(V(J_2)) = D$, and $X = X_1 \cup X_2$ (clearly, $X_1 = A \cup B$, and $X_2 = C \cup D$). See Figure 3.1 for a general diagram of the proof. We take a fixed element $y_0 \in B \setminus \{b\}$ and we will prove that $(x\ y_0) \in \mathrm{Fer}^b(G)$ for every $x \in X \setminus \{b\}$, which would yield that $\mathrm{Fer}^b(G) \cong S_{n(G)-1}$.

To this aim, let $\omega : A \to C$ be the bijection induced by $\varphi : H_1 \to H_2$. Then we have $\omega(a) = c$. Note that $cd \to ad$ is a feasible edge-replacement of G that yields the permutation $\rho = \prod_{x \in A}(x\ \omega(x)) \in R_G(cd \to ad)$ which interchanges the elements in A and C by means of ω and fixes everything else. In particular, $\rho \in \mathrm{Fer}^b(G)$.

Because $\mathrm{Fer}^b(H) \cong S_{n(H)-1}$, we know that $(x\ y_0) \in \mathrm{Fer}^b(H)$ for any $x \in (A \cup B) \setminus \{b\}$. Hence, by Lemma 3.9, $(x\ y_0) \in \mathrm{Fer}^b(G)$ for any $x \in (A \cup B) \setminus \{b\}$.

Take now any $x \in C$ and consider $\omega(x)^{-1} \in A$. By what we showed above, the permutations ρ and $(\omega^{-1}(x)\ y_0)$ are contained in $\mathrm{Fer}^b(G)$, and thus $(x\ y_0) = \rho(\omega^{-1}(x)\ y_0)\rho^{-1} \in \mathrm{Fer}^b(G)$ for any $x \in C$.

Finally, we consider an arbitrary $x \in D$. Take now a fixed $x_0 \in C \setminus \{c\}$, and observe that $(x\ x_0) \in \mathrm{Fer}^c(J) \cong S_{n(J)-1}$. By means of Lemma 3.9, we can even say that $(x\ x_0) \in \mathrm{Fer}^b(G)$. Since $(x_0\ y_0) \in \mathrm{Fer}^b(G)$ by what we showed above, we conclude that $(x\ y_0) = (x\ x_0)(x_0\ y_0)(x\ x_0) \in \mathrm{Fer}^b(G)$.

Hence, we have considered all possibilities for $x \in X \setminus \{b\}$, and we can assert that $(x\ y_0) \in \mathrm{Fer}^b(G)$ for every $x \in X \setminus \{b\}$. Since such a set of permutations generates the symmetric group on $X \setminus \{b\}$, it follows that $\mathrm{Fer}^b(G) \cong S_{n(G)-1}$, and we are done. □

Lemma 3.11. *Let G be a graph with label set X. If $b \in X$ is such that*

$$\mathrm{Fer}^b(G) \cong S_{n(G)-1} \text{ and } \mathrm{Fer}(G) \setminus \mathrm{Stab}_G(b) \neq \emptyset,$$

then G is a local amoeba.

Proof. The statement follows from Lemma 3.3 setting $x = b$ and $S = \mathrm{Fer}^b(G)$. □

Combining Lemmas 3.10 and 3.11, we obtain the following corollary.

Corollary 3.12. *Let H and J be graphs satisfying the conditions of Lemma 3.10 and such that $\mathrm{Fer}(G) \setminus \mathrm{Stab}_G(b) \neq \emptyset$. Then the graph $G = (H \cup J) + vw$ is a local amoeba.*

Proof. By Lemma 3.10, we know that $\mathrm{Fer}^b(G) \cong S_{n(G)-1}$. Since we also have that $\mathrm{Fer}(G) \setminus \mathrm{Stab}_G(b) \neq \emptyset$, Lemma 3.11 implies that G is a local amoeba. □

Corollary 3.12 suggests that one can design recursive constructions of local amoebas via this method. Such a construction of a family of global amoeba trees that have a Fibonacci-like structure is given in [12]. We will show in Section 3.1.1 that these trees are in fact local amoebas, too. Moreover, in Section 3.1.2, we will handle another recursive construction that was defined in [30].

Observe also that the direct application of this method in a recursive manner results in the construction of a family of quite sparse graphs as we are adding each time a single edge connecting two smaller graphs. However, as complements of local amoebas are again local amoebas by Lemma 3.5, it is evident that there are also families of dense local amoebas that can be constructed recursively by these tools, too.

3.1.1 Fibonacci-type trees

The first family we discuss is a family of global amoeba-trees originally defined in [12]. It is constructed using a Fibonacci recursion that gives rise to an infinite family of global-amoeba trees $\mathcal{T} = \{T_k \mid k \geq 1\}$. In this section, we prove that each Fibonacci-type tree T_k is a local amoeba as well, for every $k \in \mathbb{Z}^+$.

The Fibonacci-type trees T_k, with $k \geq 1$, are defined recursively in the following way. Let T_1 and T_2 be graphs be both isomorphic to K_2. For $k \geq 3$, T_k is built from a copy of T_{k-1} and a copy of T_{k-2} by adding an edge between a pair of vertices of maximum degree in each tree. That is, if $H \cong T_{k-1}$, and $J \cong T_{k-2}$, and u and v are a vertices of maximum degree in H and J, respectively, then $G = (H \cup J) + uv \cong T_k$. Observe that, for $k \geq 4$, the tree T_k has a unique vertex of maximum degree, which is in fact u, while, for $k \leq 3$, there are two vertices of maximum degree that are similar (i.e., there is an automorphism sending the one into the other).

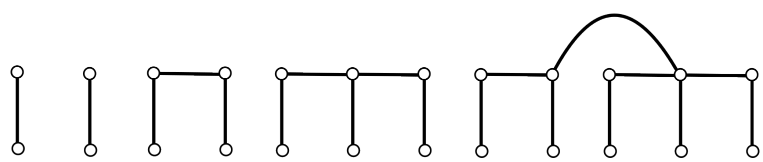


Figure 3.2: The first five Fibonacci-type trees T_k for $1 \leq k \leq 5$.

In [12], the authors proved that all Fibonacci-type trees T_k are global amoebas and that, for $1 \leq k \leq 5$, T_k is a local amoeba, too. They left as an open problem to prove if they are also local amoebas for $k \geq 6$. The next theorem states a solution to this problem using Lemma 3.10 and Lemma 3.11.

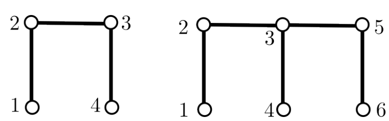


Figure 3.3: A labeling of T_3 and T_4.

Theorem 3.13. *For every $k \in \mathbb{Z}^+$, the Fibonacci-type tree $T_k \in \mathcal{T}$ is a local amoeba.*

Proof. Let $k \geq 1$ and let X be the set of labels for T_k. Let $b \in X$ be the label of a vertex of maximum degree in T_k. We use induction over k and Lemma 3.10 to prove that $\mathrm{Fer}^b(T_k) \cong S_{n(T_k)-1}$, then employ Lemma 3.11 to prove that $\mathrm{Fer}(T_k) \cong S_{n(T_k)}$. Let v_i be the vertex of T_k with label $i \in X$. Then v_b is a vertex of maximum degree. For the base of the induction, it is easy to see that T_1 and T_2 are local amoebas such that $\mathrm{Fer}^b(T_1)$ and $\mathrm{Fer}^b(T_2)$ are isomorphic to the symmetric group S_1. For the cases when $k = 3, 4$ we proceed manually because these trees are too small to satisfy the theorem. Notice that T_3 is a path of length 3 therefore we can view the tree as $T_3 = v_1v_2v_3v_4$ with label set $\{1, 2, 3, 4\}$ (see Figure 3.3). We use the permutations $(1432) \in \mathrm{Fer}_{T_3}(12 \to 14)$ and $(01) \in \mathrm{Fer}_{T_3}(02 \to 12)$ to generate S_4.

Now let $k = 4$. Let $\{1, 2, 3, 4, 5, 6\}$ be the set of labels of T_4 arranged as in Figure 3.3. We use the permutations $(12) \in \mathrm{Fer}_{T_4}(23 \to 13)$ and $(146235) \in \mathrm{Fer}_{T_4}(35 \to 26)$ to generate S_4. Let now $k \geq 5$. By construction, T_k consists of subtrees $H \cong T_{k-1}$ and $J \cong T_{k-2}$ and an edge $v_b v_c$ joining their vertices $v_b \in V(H)$ and $v_c \in V(J)$ of maximum degree. As $H \cong T_{k-1}$, there are subtrees $H_1 \cong T_{k-3}$ and $J_1 \cong T_{k-2}$ of H, with v_a and v_b the vertices of maximum degree in H_1 and J_1, respectively, such that $H = (H_1 \cup J_1) + v_a v_b$. Similarly, there are subtrees $H_2 \cong T_{k-3}$ and $J_2 \cong T_{k-4}$ of J, with v_c and v_d the vertices of maximum degree in H_2 and J_2, respectively, such that $J = (H_2 \cup J_2) + v_c v_d$. Clearly, there is an isomorphism from H_1 to H_2 that sends v_a to v_c. Moreover, by the induction hypothesis, $\mathrm{Fer}^b(H) \cong S_{n(H)-1}$ and $\mathrm{Fer}^c(J) \cong S_{n(J)-1}$. Hence, all conditions of Lemma 3.10 are satisfied, and it follows that $\mathrm{Fer}^b(T_k) \cong S_{n(T_k)-1}$.

Now consider the feasible edge replacement $ab \to ac$ of T_k and consider the permutation $\varphi \in \mathrm{Fer}_{T_k}(ab \to ac)$ induced by an isomorphism from $J_1 \cong T_{k-2}$ to $J \cong T_{k-2}$ that sends v_b to v_c, i.e. we have $\varphi(b) = c \neq b$. It follows by Lemma 3.11 that T_k is a local amoeba.

We conclude that every Fibonacci-type tree $T_k \in \mathcal{T}$ is a local amoeba for every $k \in \mathbb{Z}^+$. $\square$

The proof of Proposition 3.14 follows directly from Proposition 1.39.

Proposition 3.14. *For every $k \in \mathbb{Z}^+$, the Fibonacci-type tree $T_k \in \mathcal{T}$ is a global amoeba.*

3.1.2 Other recursive constructions

We define recursively two families of trees $\mathcal{A} = \{A_k \mid k \geq 2\}$, and $\mathcal{A}' = \{A'_k \mid k \geq 2\}$ the following way. Let $A_1 \cong K_2$ and $A'_1 \cong K_1$. For $k \geq 2$, consider a copy H of A_{k-1} and a copy J of A'_{k-1}. Let u be a vertex of maximum degree in H, and let v be a vertex of maximum degree in J. Then we define $A_k = (H \cup J) + uv$. To define A'_k, consider two copies of A'_{k-1}, say H' and J', and let w and z be vertices of maximum degree in H' and J', respectively. Then we define $A'_k = (H' \cup J') + wz$. To show that this is well defined, notice first that A_1 has two vertices of maximum degree which are similar, while A'_1 has trivially only one vertex of maximum degree. Moreover, for $k \geq 2$, A_k has, by construction, a unique vertex of maximum degree, and A'_k has two vertices of maximum degree which are similar to each other.

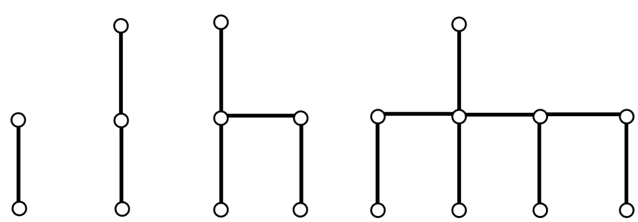


Figure 3.4: The first four A_k trees for $1 \leq k \leq 4$.

Theorem 3.15. *For every $k \geq 1$, $A_k \in \mathcal{A}$ and $A'_k \in \mathcal{A}'$ are local amoebas.*

Proof. Let $k \geq 1$ and let X be the set of labels for A_k, and Y the set of labels for A'_k. Let $b \in X$ and $b' \in Y$ be the labels of a vertex of maximum degree in A_k and A'_k, respectively.

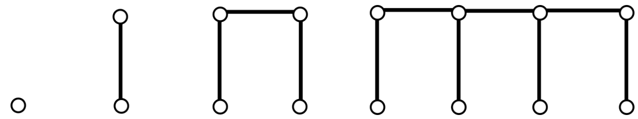


Figure 3.5: The first four A'_k trees for $1 \le k \le 4$.

We start by proving that $\mathrm{Fer}^b(A_k) \cong S_{n(A_k)-1}$, and $\mathrm{Fer}^{b'}(A'_k) \cong S_{n(A'_k)-1}$. Finally, we use Lemma 3.11 to conclude that A_k and A'_k are local amoebas for every $k \ge 2$.

To prove that that $\mathrm{Fer}^b(A_k) \cong S_{n(A_k)-1}$ and $\mathrm{Fer}^{b'}(A'_k) \cong S_{n(A'_k)-1}$, we proceed by induction on k. For $k = 1, 2$, the result is given trivially.

For $k = 3$ we proceed manually as this case is too small to satisfy the theorem's hypothesis. Let A_3 have a label set $\{1, 2, 3, 4, 5\}$ such that it is arranged as in **??**. Note that the permutations $(24)(35) \in \mathrm{Fer}_{A_3}(12 \to 14)$, $(45) \in \mathrm{Fer}_{A_3}(24 \to 25)$ and $(14) \in \mathrm{Fer}_{A_3}(45 \to 15)$ satisfy $\langle (24)(35), (45), (14) \rangle = \langle (2435), (14) \rangle \cong S_5$ by Lemma 3.2. Let now $k \ge 4$ and assume that, for every j with $1 \le j < k$, this property is satisfied for A_j and A'_j.

We consider first A'_k. Let $A'_k = (H' \cup J') + v'w'$, where both H' and J' are isomorphic to A'_{k-1}, and v' and w' are vertices of maximum degree in H' and J', respectively. As v' and w' are also vertices of maximum degree in A_k, we can assume that v' has label b'. Now let H'_1, J'_1, H'_2, J'_2 be all graphs isomorphic to A'_{k-2} and $u' \in V(H'_1)$, $x' \in V(J'_2)$ such that $H' = (H'_1 \cup J'_1) + u'v'$ and $J' = (H'_2 \cup J'_2) + w'x'$. Let c' be the label of w'. By induction hypothesis, $\mathrm{Fer}^{b'}(H') \cong S_{n(H')-1}$, and $\mathrm{Fer}^{c'}(J') \cong S_{n(J')-1}$. As the conditions of Lemma 3.10 are satisfied, we can conclude that $\mathrm{Fer}^{b'}(A'_k) \cong S_{n(A'_k)-1}$.

The proof of $\mathrm{Fer}^b(A_k) \cong S_{n(A_k)-1}$ works similarly. Let $A_k = (H \cup J) + vw$, where $H \cong A_{k-1}$ and $J \cong A'_{k-1}$, and v and w are vertices of maximum degree in H and J, respectively. Observe first that, by construction of A_k, there is a leaf z adjacent to v such that $H - z \cong A'_{k-1}$. Now consider graphs H_1, J_1, H_2, J_2 such that H_1, H_2 and J_2 are all isomorphic to A'_{k-2} and $J_1 \cong A_{k-2}$, and vertices $u \in V(H_1)$, $x \in V(J_2)$ such that $H = (H_1 \cup J_1) + uv$ and $J = (H_2 \cup J_2) + wx$. Let c and d be the labels of w and z, respectively. By the induction hypothesis, $\mathrm{Fer}^b(H) \cong S_{n(H)-1}$, and $\mathrm{Fer}^c(J) \cong S_{n(J)-1}$. Hence, we can apply again Lemma 3.10 to deduce that $\mathrm{Fer}^b(A_k) \cong S_{n(A_k)-1}$.

Now we show that both A_k and A'_k are local amoebas. Consider first a permutation $\varphi' \in \mathrm{Fer}(A'_k)$ that interchanges the labels among the sets $V(H')$ and $V(J')$ induced by an isomorphism that sends b' to c'. Since $\varphi'(b') = c' \neq b'$, it follows by Lemma 3.11 that A'_k is a local amoeba. Finally, notice that $db \to dc$ is a feasible edge-replacement in A_k, and take a permutation $\varphi \in \mathrm{Fer}(A_k)$ induced by an isomorphism between $H - z$ and J that takes v to w. Hence, $\varphi(b) = c \neq b$ and by Lemma Lemma 3.11, it follows that A'_k is a local amoeba. $\square$

The proof of Proposition 3.16 follows from Proposition 1.39.

Proposition 3.16. *For every $k \ge 1$, $A_k \in \mathcal{A}$ and $A'_k \in \mathcal{A}'$ are global amoebas.*

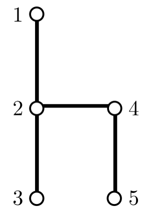


Figure 3.6: A labeling of A_3.

3.2 Balancing number of global amoebas

In this section we approach the balancing number of global amoebas through the Turán number of a class of graphs, defined before in 2.5, re-stated in (3.1), and provide a lower bound for such number in terms of cuts defined in (3.2). We then apply these results to the Fibonacci-type trees and the A_k-trees defined in Section 3.1.2.

Let G be a graph on m edges. Let

$$\mathcal{H}_G = \left\{ H : H \subset G,\ H \text{ has } \left\lfloor \frac{m}{2} \right\rfloor \text{ edges and no isolated vertices} \right\} \tag{3.1}$$

be the class of subgraphs of G that contain half the number of edges of G (rounding down when m is odd).

For a graph $G = (V, E)$ and partition $V = S \cup T$, let

$$e(S, T) = |\{uv \in E(G) : u \in S, v \in T\}|; \tag{3.2}$$

we say that $e(S, T)$ is the size of the cut (S, T) and we say that the cut has order $\min\{|S|, |T|\}$.

The next theorem establishes the equivalence between the balancing number of a global amoeba and the Turán number of the class $\mathcal{H}_G$.

Theorem 3.17. *Let G be a global amoeba with m edges. Then* $\mathrm{bal}(n, G) = \mathrm{ex}(n, \mathcal{H}_G)$.

Proof. First we show that $\mathrm{ex}(n, \mathcal{H}_G) \leq \mathrm{bal}(n, G)$. Consider any 2-edge-coloring $E(K_n) = R \cup B$ satisfying $\mathrm{bal}(n, G) < \min\{|R|, |B|\}$. By definition of balancing number, there is a balanced copy of G, which is the union of G_R and G_B where each graph is monochromatic red and blue respectively. Setting $m = e(G)$, it follows that G_R and G_B have each $\lfloor \frac{m}{2} \rfloor$ edges and, thus, each of the classes R and B have a copy of a graph in $\mathcal{H}_G$. Since each color class in the 2-edge coloring has more than $\mathrm{bal}(n, G)$ edges, we conclude that $\mathrm{ex}(n, \mathcal{H}_G) \leq \mathrm{bal}(n, G)$.

In what follows we consider a 2-edge coloring of K_n satisfying $\mathrm{ex}(n, \mathcal{H}_G) < \min\{|R|, |B|\}$. We show that the coloring contains a balanced copy of G from which we conclude that $\mathrm{bal}(n, G) \leq \mathrm{ex}(n, \mathcal{H}_G)$.

Since $\mathrm{ex}(n, \mathcal{H}_G) < |R|$, there is a red copy $H_R \subset R$ of a graph in $\mathcal{H}_G$. Let $G_R \subset K_n$ be a copy of G such that $H_R \subset G_R$ and let ℓ_R be the number of red edges in G_R. Similarly, we may consider a blue copy $H_B \subset B$ of a graph in $\mathcal{H}_G$. Let $G_B \subset K_n$ be a copy of G such that $H_B \subset G_B$ and let ℓ_B be the number of blue edges in G_B. By the definition of $\mathcal{H}_G$, we have that $\min\{\ell_R, \ell_B\} \geq \lfloor \frac{m}{2} \rfloor$, where $m = e(G)$.

By the property of global amoebas, we find a sequence of subgraphs $G_0 = G_R, G_1, \dots G_k = G_B$ of K_n such that, for each $1 \le i \le k$, G_i may be obtained from G_{i-1} by a feasible edge replacement. Let $\ell_{R,i}$ be the number of red edges in G_i. Clearly, $|\ell_{R,i} - \ell_{R,i-1}| \le 1$; on the other hand, $\ell_{R,0} = \ell_R \ge \lfloor \frac{m}{2} \rfloor$ and $\ell_{R,k} = m - \ell_B \le m - \lfloor \frac{m}{2} \rfloor = \lceil \frac{m}{2} \rceil$. It follows that there is some $0 \le i \le k$ for which G_i is a balanced copy of G. Hence, $\mathrm{bal}(n, G) \le \mathrm{ex}(n, \mathcal{H}_G)$, as desired. $\square$

The following theorem gives a lower bound on $\mathrm{ex}(n, \mathcal{H}_G)$ in terms of the maximum size of a cut of order at most ℓ.

Theorem 3.18. *Let $G = (V, E)$ be a graph of size m and let $\mathcal{H}_G$ be defined as in (3.1). If $\ell \in \mathbb{N}$ satisfies*

$$\max\{e(S, V \setminus S) : S \subset V, |S| \le \ell\} < \left\lfloor \frac{m}{2} \right\rfloor,$$

then $\mathrm{ex}(n, \mathcal{H}_G) \ge \ell(n - \ell)$.

Proof. Let $\ell \in \mathbb{N}$ be as stated above. We claim that the complete bipartite graph $K_{\ell,n-\ell}$ does not contain any $H \in \mathcal{H}_G$ as a subgraph and so $\mathrm{ex}(n, \mathcal{H}_G) \ge \ell(n - \ell)$.

Suppose to the contrary that there is a copy of $H \in \mathcal{H}_G$ in $K_{\ell(n-\ell)}$. It follows that H is bipartite and, furthermore, one of its parts has order at most ℓ. Thus, there is $S' \subset V$ such that $\lfloor \frac{m}{2} \rfloor = e(H) \le e(S', V \setminus S') \le \max\{e(S, V \setminus S) : S \subset V, |S| \le \ell\} < \left\lfloor \frac{m}{2} \right\rfloor$; which is a contradiction. $\square$

The following proposition states a property of global amoebas that will be used in the proof of Corollary 3.20.

Proposition 3.19. *[12] Let $d_1 \ge d_2 \ge \cdots \ge d_n$ be the degree sequence of a global amoeba G of order n, and let $D = \{d_i \mid i \in [n]\}$. Then*

i) $D = \{0\} \cup [\Delta]$, where Δ is the maximum degree of G, and

ii) for every $i \in [n]$, we have $d_i \le n + 1 - i$.

Corollary 3.20. *Let $G = (V, E)$ be a global amoeba with k vertices, m edges and degree sequence $d_1 \ge \cdots \ge d_k$. For any $\ell \ge 1$ satisfying $\sum_{i=1}^{\ell} d_i < \left\lfloor \frac{m}{2} \right\rfloor$, we have* $\mathrm{bal}(n, G) \ge \ell(n - \ell)$. *In particular,*

$$\mathrm{bal}(n, G) \ge \left\lfloor \frac{m-1}{2k+1} \right\rfloor \left(n - \left\lfloor \frac{m-1}{2k+1} \right\rfloor\right).$$

Proof. The first statement follows from Theorems 3.17 and 3.18. Indeed, let $\ell \ge 1$ be as stated above. Then

$$\max\{e(S, V \setminus S) : S \subset V, |S| \le \ell\} \le \max\left\{\sum_{v \in S} \deg(v) : S \subset V, |S| \le \ell\right\} \le \sum_{i=1}^{\ell} d_i < \left\lfloor \frac{m}{2} \right\rfloor,$$

which implies that $\mathrm{bal}(n, G) = \mathrm{ex}(n, \mathcal{H}_G) \geq \ell(n-\ell)$.

We now use properties of global amoebas to bound $\sum_{i=1}^{\ell} d_i$. By Proposition 3.19, since G is a global amoeba, $d_i \leq k+1-i$ for $1 \leq i \leq k$. In particular, $\sum_{i=1}^{\ell} d_i \leq \sum_{i=1}^{\ell}(k+1-i) = (k+1)\ell - \ell(\ell+1)/2$. We show that for $\ell = \left\lfloor \frac{m-1}{2k+1} \right\rfloor$, we have

$$(k+1)\ell - \frac{\ell(\ell+1)}{2} \leq \frac{m-1}{2} < \left\lfloor \frac{m}{2} \right\rfloor. \tag{3.3}$$

The solutions $x \in \mathbb{R}$ to the quadratic inequality $2(k+1)x - x(x+1) \leq m-1$ satisfy $x \leq x_-$ or $x \geq x_+$, where

$$x_{\pm} = \frac{2k+1}{2}\left(1 \pm \sqrt{1 - \frac{4(m-1)}{(2k+1)^2}}\right).$$

Using[1] that $\sqrt{1-y} \leq 1 - y/2$ for $y \in \mathbb{R}$, we have that ℓ satisfies (3.3) since

$$x_- \geq \frac{2k+1}{2}\left(\frac{2(m-1)}{(2k+1)^2}\right) = \frac{m-1}{2k+1} \geq \ell.$$

□

We note that $x_+ \geq k > \ell$ and so, using this approach, the bound obtained in Corollary 3.20 may only be improved by increasing the value of a feasible $\ell \leq x_-$.

3.2.1 Fibonacci-type trees

For a non-negative integer k, let F_k denote the k-th Fibonacci number (so that $F_1 = F_2 = 1$, $F_3 = 2$ and so on). Recall that T_k is the k-th Fibonacci-type tree.

Theorem 3.21. *We have* $\mathrm{bal}(n, T_4) = \mathrm{bal}(n, A_3) = 1$ *and* $\mathrm{bal}(n, T_5) = \mathrm{bal}(n, A_4) = 6$.

Proof. Notice that $\mathcal{H}_{T_4} = \mathcal{H}_{A_3}$ consists of exactly two non-isomorphic graphs: $2K_2$ and P_2. Thus $\mathrm{ex}(n, \mathcal{H}_{T_4}) = 1$ and by Theorem 3.17 we are done.

Similarly, we note that $\mathcal{H}_{T_5} = \mathcal{H}_{A_4}$. To save on words, call a graph containing no copies of elements of $\mathcal{H}_{T_5}$ "extremal". Now observe that $\mathcal{H}_{T_5}$ consists of all graphs on four edges except $C_4, K_3 + e$ (where e is an edge disjoint from K_3) and $K_4 - P_2$. Thus, any extremal graph on four or more edges must contain one of the these three. As a consequence of this fact, it is straightforward to show, but cumbersome to write down, that the only extremal graph on five edges is K_4 minus an edge. It follows that K_4 is the only extremal graph on six edges. Similarly, any extremal graph on seven edges must contain K_4. Since both the star of four edges and the star of three edges plus a disjoint edge are elements of $\mathcal{H}_{T_5}$, no such extremal graph exists. Hence,

$$\mathrm{bal}(n, A_4) = \mathrm{bal}(n, T_5) = \mathrm{ex}(n, \mathcal{H}_{T_5}) = 6.$$

□

[1]The function $f(y) = \sqrt{1-y} - 1 + y/2$ is concave and its maximum is attained at $f(0) = 0$.

The following theorem, which we prove in two parts, provides us with an asymptotically tight bound on the balancing number of these Fibonacci-type trees.

Theorem 3.22. *For $k \geq 6$,*

$$F_{k-4}(n - F_{k-4}) \leq \text{bal}(n, T_k) \leq (F_{k-2} - 2)n(1 + o(1)).$$

The following technical lemma greatly simplifies the calculations ahead.

Lemma 3.23. *For k and i, two positive integers, let us denote by V_i^k the set of vertices of T_k of degree i and $t_i^k := |V_i^k|$. Then the following holds for all $k \geq 4$.*

(a) $t_1^k = F_k$;

(b) $t_2^k = F_{k-1}$;

(c) $t_{k-1}^k = 1$, *if* $k \geq 5$, *then* $t_{k-2}^k = 1$, *and if* $k \geq 6$, *then* $t_{k-3}^k = 1$; *and*

(d) if $2 < i < k - 1$, $t_i^k = F_{k-i-1}$;

(e) T_k *has a unique vertex of maximum degree* $k - 1$.

Proof. We prove the lemma by a simple inductive argument. Indeed the statement is clearly true for $k \in \{4, 5\}$. Furthermore, if it holds for $k - 1$ and $k - 2$, notice that for any $i < k - 3$, by the construction of T_k and (e) applied to $k - 2$, we get that V_i^k is the disjoint union of V_i^{k-1} and V_i^{k-2}. Applying (a), (b) and (c) for $k - 1$ and $k - 2$, we immediately have (a) and (b) for k:

$$t_1^k = F_k; \qquad t_2^k = F_{k-1};$$

and

$$t_i^k = t_i^{k-1} + t_i^{k-2} = F_{k-i-1}.$$

That (e) for k is true follows directly from the construction of T_k and the inductive hypothesis, which incidentally implies that (d) must be true for k as well. We now focus on the case when $k \in \{k - 2, k - 3\}$. Using (e) for $k - 1$ and $k - 2$, we know that $T_{k-1} \cup T_{k-2}$ has exactly two vertices of degree $k - 3$, namely the vertex of maximum degree of T_{k-2} and of the copy of T_{k-2} inside of T_{k-1}. T_k has have exactly one vertex of degree $k - 3$ since one of them is joined via an edge to the only vertex of degree $k - 2$ in T_{k-1}. Again by (e), neither T_{k-1} nor T_{k-2} have any vertices of degree $k - 2$ and so the only such vertex in T_k is produced by the new edge (and only one of the vertices has degree $k - 3$). Consequently, $t_i^{k-3} = t_i^{k-2} = 1$ and so we are done. □

Lemma 3.24. *For $k \geq 6$, $F_{k-4}(n - F_{k-4}) \leq \text{bal}(n, T_k)$.*

Proof. First define, for all $i < k$, $t_i := t_i^k$ as in Lemma 3.23. Let $\ell := F_{k-4}$. By properties of the Fibonacci sequence, $\sum_{i=1}^{m} F_i = F_{m+2} - 1$, so by Lemma 3.23,

$$\ell = 1 + \sum_{i=1}^{k-6} F_i = \sum_{i=5}^{k-1} t_i.$$

Then, the sum of the first ℓ elements of the degree sequence of T_k is equal to $\sum_{i=5}^{k-1} it_i$. In the spirit of Corollary 3.20, we want this latter number to be smaller than $\left\lfloor \frac{m}{2} \right\rfloor = \left\lfloor \frac{2F_n - 1}{2} \right\rfloor$.

Notice that since $\sum_{i=1}^{k} it_i = 2(2F_k - 1)$, it is enough to argue that

$$\sum_{i=1}^{4} it_i \geq 3F_{k-1}.$$

But this computation is derived by properly expanding the terms of the Fibonacci sequence and using the fact that $3F_{k-4} + 4F_{k-5} \geq 2F_{k-2} - 1$.

$$\begin{aligned} \sum_{i=1}^{4} it_i &= F_k + 2F_{k-1} + 3F_{k-4} + 4F_{k-5} \\ &\geq F_k + 2F_{k-1} + 2F_{k-2} - 1 \\ &= 3F_k - 1. \end{aligned}$$

This proves the lower bound given in the lemma's statement. And a straightforward calculation shows that the bound may not be improved with the same techniques, i.e. by increasing the number of addends in the above computation. □

Lemma 3.25. *For $k \geq 6$, $\operatorname{bal}(n, T_k) \leq (F_{k-2} - 2)n(1 + o(1))$.*

Proof. Similarly to the previous lemma, we show that, for any $k \geq 5$, there exists a star forest S_k on F_{k-1} edges contained in T_k such that $\operatorname{ex}(n, S_k) = (F_{k-2} - 2)n(1 + o(1))$. Then, applying monotonicity, Theorem 3.17 and [49][Theorem 3], we conclude the proof.

For $k \in \{5, 6\}$, the construction of S_k and S_k^+ can be seen in Figure 3.7. We construct the subsequent forests recursively. Assume that S_{k-1} and S_{k-1}^+, and S_{k-2} and S_{k-2}^+ are star forests in T_{k-1} and T_{k-2} respectively. Then, in T_k, consider $S_k := S_{k-1}^+ \cup S_{k-2}^+$ and $S_k^+ := S_{k-1} \cup S_{k-2}^+$. Hence, S_k has $F_{k-1} + F_{k-2} + 1 = F_k - 1$ edges and, similarly, S_k^+ has F_k edges.

If we now let c_m be the number of components of S_m, a_m the number of vertices of degree 4 and b_m the number of vertices of degree 3, we have that, inductively,

$$\begin{aligned} a_5 = a_6 &= 1 & a_m &= a_{m-1} + a_{m-2} = F_{n-4} \\ b_5 = 0, \quad b_6 &= 1 & b_m &= b_{m-1} + b_{m-2} = F_{m-5} \\ c_5 = 1, \quad c_6 &= 2 & c_m &= c_{m-1} + c_{m-2} + 1 = F_{m-2} - 1 \end{aligned}$$

We conclude the proof by calculating $\operatorname{ex}(n, S_k)$. Indeed here we apply [49][Theorem 3]. In the case that $i \leq a_m$,

$$\operatorname{ex}(n, S_k) = \frac{2i + d_i - 3}{2} \leq \frac{2a_m + d_{a_m} - 3}{2} = \frac{2F_{n-4} + 1}{2}.$$

Similarly, when $a_m < i \leq a_m + b_m$,

$$\mathrm{ex}(n, S_k) = \frac{2i + d_i - 3}{2} \leq \frac{2(a_m + b_m) + 3 - 3}{2} = F_{m-3}.$$

And finally, when $a_m + b_m < i \leq c_m$, we have that

$$\mathrm{ex}(n, S_k) = \frac{2i + d_i - 3}{2} \leq \frac{2c_m + 1 - 3}{2} = F_{m-2} - 2.$$

Finally,

$$\mathrm{bal}(n, T_k) = \mathrm{ex}(n, \mathcal{H}_{T_k}) \leq \mathrm{ex}(n, S_k) \leq (F_{k-2} - 2)n(1 + o(1)).$$

□

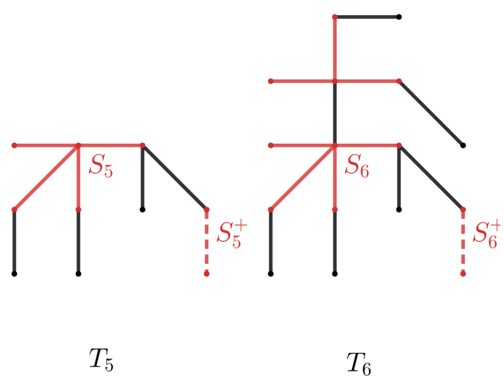


Figure 3.7: Star forests (in red) S_k. S_k^+ contains, in addition, the dotted red edges.

The following theorem states how the balancing number of A_k is related to the balancing number of A'_k. It is proven in a series of lemmas and propositions.

Theorem 3.26. *For $k \geq 5$,*

$$2^{k-5}(n - 2^{k-5}) \leq \mathrm{bal}(n, A'_k) \leq \mathrm{bal}(n, A_k) \leq (2^{k-3} - 1)(n - 2^{k-4}).$$

The following Lemma grants some the number of vertices in A_k with degree d for $1 \leq d \leq k$. This will be helpful when proving the analogous result for A'_k.

Lemma 3.27. *For k and i, two positive integers, let us denote by V_i^k the set of vertices of A_k of degree i and $t_i^k := |V_i^k|$. Then the following holds for all $k \geq 2$.*

(a) $n(A_k) = 2^{k-1} + 1$ *and* $e(A_k) = 2^{k-1}$*;*

(b) $t_1^k = 2^{k-2} + 1$*;*

(c) $t_i^k = 2^{k-i-1}$*, if* $2 \leq i \leq k-2$ *when* $k \geq 3$*;*

(d) $t_{k-1}^k = 1$ *for* $k \geq 3$*; and*

(e) $t_k^k = 1$;

Proof. We prove $e(A_k) = 2^{k-1}$ by induction. It is clear that $e(A_1) = 1$. By the construction of A_k and the inductive step, we can conclude that $e(A_k) = 2e(A_{k-1}) = 2(2^{k-2}) = 2^{k-1}$. The fact that $n(A_k) = 2^{k-1} + 1$ is straightforward by the property that states that $n(T) = e(T) + 1$ for every tree T.

It is clear that $t_1^1 = 2$ and $t_2^2 = 1$. To prove (b) we also proceed by induction on k. The base is clear as $t_1^2 = 2$. For $k \geq 2$ notice that A_k has double the number of vertices of degree 1 in A_k minus one. Hence, $t_1^k = 2(t_1^{k-1}) - 1 = 2^{k-2} + 1$.

We continue to prove (d) and (e). By the construction of A_k, it is clear that the maximum degree vertex is unique for $k \geq 2$ and the vertex of degree $k-1$ is unique for $k \geq 3$.

To prove (c), we use induction on k. Note that for $2 \leq i \leq k-2$ the number of vertices in A_k with degree i is double the number of vertices of degree i in A_{k-1}. Therefore, for each $2 \leq i \leq k-2$, we start with the base of the induction in A_{i+2} and we calculate t_i^{i+2}. Note that $t_i^{i+1} = 1$ by (d), therefore $t_i^{i+2} = 2t_i^{i+1} = 2$.We proceed with the induction step: and $t_i^k = 2 \cdot t_i^{k-1} = 2 \cdot 2^{k-i-2} = 2^{k-i-1}$. This concludes the proof. □

The balancing number of A_k is closely related to the balancing number of A'_k, as we can see in the following result.

Proposition 3.28. *For $k \geq 2$, we have that* $\mathrm{bal}(n, A'_k) \leq \mathrm{bal}(n, A_k)$.

Proof. Let $E(K_n) = R \sqcup B$ be a 2-edge coloring with at least $\mathrm{bal}(n, A_k)$ edges in each color. Therefore, there is a balanced copy of A_k, which contains a balanced copy of A'_k. This concludes the proof. □

We employ the following Lemma to provide a lower bound of $\mathrm{bal}(n, A'_k)$.

Lemma 3.29. *For k and i, two positive integers, let us denote by V_i^k the set of vertices of A'_k of degree i and $t_i^k := |V_i^k|$. Then the following holds for all $k \geq 2$.*

(a) $n(A'_k) = 2^{k-1}$ *and* $e(A'_k) = 2^{k-1} - 1$;

(b) $t_i^k = 2^{k-i-1}$, *if* $1 \leq i \leq k-2$ *for* $k \geq 3$;

(c) $t_{k-1}^k = 2$ *for* $k \geq 2$.

Proof. By Lemma 3.27, note that $n(A'_k) = n(A_k) - 1 = 2^{k-1}$ and $e(A'_k) = e(A_k) - 1 = 2^{k-1} - 1$. This proves (a). The proof of (b) is analogous to the proof of (c) from Lemma 3.27. The proof of (c) follows from the construction of A'_k and from (d) and (e) in Lemma 3.27. □

Proposition 3.30. *For $k \geq 5$, $2^{k-5}(n - 2^{k-5}) \leq \mathrm{bal}(n, A'_k)$.*

Proof. Let $d_1 \geq d_2 \geq \cdots \geq d_{2^{k-1}}$ be the degree sequence of A'_k. Let $S(k,q)$ be the sum of the first ℓ elements of the degree sequence of A'_k. By Lemma 3.29, $S(k,q)$ can be expressed in the following way.

$$\begin{aligned} S(k,q) &= 2(k-1) + 2(k-2) + 2^2(k-3) + \cdots + 2^{q-1}(k-q) \\ &= 2(k-1) + \sum_{i=1}^{q-1} 2^i(k-(i+1)) \end{aligned}$$

Therefore, $\ell = 2 + \sum_{i=1}^{q-1} 2^i$. We wish to find the largest value q can have so that

$$S(k,q) < \frac{2^{k-1}-1}{2} \tag{3.4}$$

holds. We use the identity $\sum_{i=0}^{n} r^i = \frac{1-r^{n+1}}{1-r}$, for $|r| < 1$, and expand $\sum_{i=0}^{n} 2^i$.

$$\sum_{i=0}^{n} 2^i = 2^n \sum_{i=0}^{n} 2^{i-n} = 2^n \sum_{i=0}^{n} \left(\frac{1}{2}\right)^{n-i} = 2^n \sum_{i=0}^{n} \left(\frac{1}{2}\right)^{i} = 2^n \cdot \frac{1-\left(\frac{1}{2}\right)^{n+1}}{1-\frac{1}{2}} = 2^{n+1} - 1 \tag{3.5}$$

This implies that

$$\ell = 2 + \sum_{i=1}^{q-1} 2^i = 2 + \left(\sum_{i=0}^{q-1} 2^i\right) - 1 = 2^q.$$

We proceed to express $S(k,q)$ in a simpler manner.

$$\begin{aligned} S(k,q) &= 2(k-1) + \sum_{i=1}^{q-1} 2^i(k-(i+1)) \\ &= 2(k-1) + (k-1)\sum_{i=1}^{q-1} 2^i - \sum_{i=1}^{q-1} i \cdot 2^i \\ &= 2(k-1) + (k-1)f(q) - g(q). \end{aligned}$$

We work with the newly defined functions $f(q)$ and $g(q)$. By eq. (3.5), $f(q)$ is expressed as follows.

$$f(q) = \sum_{i=1}^{q-1} 2^i = 2^q - 2.$$

We use eq. (3.5) again to find a simpler expression for $g(q)$.

$$
\begin{aligned}
g(q) &= \sum_{i=1}^{q-1} i \cdot 2^i = \sum_{i=1}^{q-1} 2^i + \sum_{i=2}^{q-1} 2^i + \cdots + \sum_{i=q-1}^{q-1} 2^i \\
&= \sum_{j=1}^{q-1} \sum_{i=j}^{q-1} 2^i = \sum_{j=1}^{q-1} 2^j \sum_{i=0}^{q-1-j} 2^i = \sum_{j=1}^{q-1} 2^j (2^{q-j} - 1) \\
&= \sum_{j=1}^{q-1} (2^q - 2^j) \\
&= 2^q(q-1) - 2^q + 2 \\
&= 2^q(q-2) + 2.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
S(k,q) &= 2(k-1) + (k-1)(2^q - 2) - (2^q(q-2) + 2) \\
&= 2^q(k-q+1) - 2.
\end{aligned}
$$

We solve the inequality $2^q(k-q+1) - 2 < \frac{2^{k-1}-1}{2}$ for q, which is equivalent to solving $2^{q+1}(k-q+1) - 2^2 < 2^{k-1} - 1$. Note that $q = k-5$ satisfies the inequality.

$$
\begin{aligned}
2^{k-4}(6) - 4 &< 2^{k-1} - 1 \\
2^{k-4}(6 - 2^3) &< 3 \\
-2^{k-3} &< 3
\end{aligned}
$$

If $q = k-4$, we arrive to $2^{k-3} < 3$, which is false for $k \geq 5$. Therefore, $q = k-5$ is the highest value that satisfies eq. (3.4). By Corollary 3.20, we conclude that

$$
2^{k-5}(n - 2^{k-5}) \leq \operatorname{bal}(n, A'_k)
$$

for $k \geq 5$.

$\square$

Lemma 3.31. *For $k \geq 5$,* $\operatorname{bal}(n, A_k) \leq (2^{k-3} - 1)(n - 2^{k-4})$.

Proof. We consider vertices of degree at least 5 and use the fact that A_k has 2^{k-1} edges.

For the upper bound, we find, for $k \geq 4$, a star forest contained in A_k with the following properties for any such k.

1. S_k is a forest of 2^{k-3} stars with $2^{k-1} + 1$ edges in total.

2. If v is a leaf of S_k, then $M_k v \in S_k$.

3. The maximum degree of the i-th star (in descending order) is

$$d_i = \begin{cases} 4 & i = 1 \\ 3 & 1 < i \leq 2^{k-4} \\ 1 & 2^{k-4} < i \leq 2^{k-3} \end{cases}.$$

For $k = 4$, S_k consists of the four edges incident on M_k and the remaining disjoint edge, and clearly the properties are satisfied. Assuming A_{k-1} has such a star forest, the construction of A_k requires two copies A_{k-1} and A'_{k-1}, removing the edge $M'_{k-1}v$ and joining the vertices M'_{k-1} and M_{k-1}. Say the star forests are S_{k-1} and S'_{k-1}. We claim $S_k := S'_{k-1} \setminus \{M_{k-1}v\} \cup S_{k-1}$ has all the sought properties. Now we apply monotonicity and [49][Theorem 3].

$$\mathrm{bal}(n, A_k) \leq \mathrm{ex}(n, S_k) = \max_{1 \leq i \leq 2^{k-3}} (i-1)(n-i+1) + \binom{i-1}{2} \left\lfloor \frac{d_i - 1}{2}(n-i+1) \right\rfloor.$$

For large enough n, these terms are dominated by the case when $i = 2^{k-3}$. Hence,

$$\mathrm{bal}(n, A_k) \leq (2^{k-3} - 1)(n - 2^{k-3} + 1) + \binom{2^{k-3} - 1}{2} = (2^{k-3} - 1)(n - 2^{k-4}).$$

□

3.3 Open problems

We finish this chapter stating some open problems that are left for future work concerning the topic of amoebas. It would be interesting to explore if there is a way to characterize global and local amoebas via constructions. However, this problem seems to be quite ambitious and maybe even impossible. The problems we state here concern constructions of amoebas that try to go into that direction.

1. Generalize the construction of Lemma 3.10.
2. Look for other general ways of constructing global or local amoebas.
3. Characterize local amoeba trees.
4. Characterize global amoeba trees.
5. Study the balancing number of other amoeba families by using constructions.

Chapter 4

The $K_{n,n}$ setting

All concepts studied so far, like the unavoidable patters, balanceability, and omnitonality, use the complete graph as a base graph. However, these concepts can also be explored on any other base graph. In this chapter we employ the complete bipartite graph $K_{n,n}$ as a base graph to find unavoidable patterns and discuss the concepts of balanceability and omnitonality.

4.1 Unavoidable patterns

In Section 1.2.3, the Zarankiewicz problem was discussed. and, Theorem 1.10 gives an upper bound for the Zarankiewickz number $z(m, n; s, t)$. For simplicity, we will use $z(n, t)$ when referring to $z(n, K_{t,t})$. The particular case of Theorem 1.10 for $m = n$ and $s = t$ is given below.

Corollary 4.1 (Kővári-Sós-Turán 1954 [34]). *For every positive integer t, the following is an upper bound of the Zarankiewickz number for $K_{t,t}$,*

$$z(n, t) < (t-1)^{\frac{1}{t}} n^{2-\frac{1}{t}} + \frac{1}{2}(t-1)n.$$

The following result is deduced by the Dependent random choice lemma given in Corollary 1.20.

Proposition 4.2 (Dependent random choice). *For all m and t positive integers, there exists a $C = C(m, t) > 0$ such that any graph G on $2n$ vertices with $e(G) \geq C(2n)^{2-\frac{1}{t}}$ edges contains a set $S \subset V(G)$ of m vertices where every subset $X \subset S$ such that $|X| = t$ has at least m common neighbors.*

We now state the main result of this section where we explore the base graph $K_{n,n}$ for n large enough to look for unavoidable patterns in 2-edge colorings with sufficient edges in each color.

Theorem 4.3. *Let t be a positive integer. For all large enough n and $T \geq t$, there exists a positive integer $z = z_2(n,t) = \mathcal{O}(n^{2-\frac{1}{t}})$ such that any coloring $E(K_{n,n}) = R \sqcup B$ with at least z edges in each color class contains a colored copy of $K_{T,t+T}$ where one color forms a graph isomorphic to $K_{T,T}$, or a copy of $K_{t,2T}$ where one color forms a graph isomorphic to $K_{t,T}$.*

Proof. Let $K_{n,n}$ be the complete bipartite graph with partition sets A and B, n large enough and let m and T be integers such that $n > m \geq T \geq t$ and

$$\frac{m^2}{2} \geq (T-1)^{\frac{1}{T}} n^{2-\frac{1}{T}} + \frac{1}{2}(T-1)n \geq (t-1)^{\frac{1}{t}} n^{2-\frac{1}{t}} + \frac{1}{2}(t-1)n \tag{4.1}$$

holds. Let $z = C(2n)^{2-\frac{1}{t}}$ where $C = C(m,t)$ is large enough. Let $E(K_{n,n}) = R \sqcup B$ be any 2-edge coloring with at least z edges in each color class. This is possible because $z_2(n,t) = o(n^2)$. By Proposition 4.2, the amount of blue edges guarantees that there is a set of vertices S_1 in A (without loss of generality) with $|S_1| = m$ and such that any subset $T_1 \subseteq S_1$ of order t must have a common blue neighborhood M_1 of at least m neighbors.

A similar argument can be made for the red edges. There is also a set of vertices S_2 with $|S_2| = m$ such that any subset $T_2 \subseteq S_2$ of order t must have a common red neighborhood M_2 of at least m neighbors. We will divide the proof in two cases, one being when S_2 is contained in A and the other when S_2 is contained in B. Notice that S_2 cannot intersect both A and B, because that would yield the existence of edges in one of the parts A or B.

If $S_2 \subset A$, take any subset $T_1 \subset S_1$ of order t. The set T_1 has a common blue neighborhood M_1 of order m in B. Now we will use the edges induced by S_2 and M_1 to find a monochromatic $K_{T,T}$. By the pigeonhole principle, there are at least $\frac{m^2}{2}$ edges induced by S_2 and M_1 in one color class. Because of eq. (4.1), the amount of these edges is an upper bound of $z(m,T)$ (by Corollary 4.1). Hence there is a monochromatic $K_{T,T}$ in the graph induced by $S_2 \cup M_1$. If this $K_{T,T}$ is red, then the graph induced by $T_1 \cup M_1 \cup S_2$ is a $K_{T,t+T}$, where one color forms a graph isomorphic to $K_{T,T}$. If the $K_{T,T}$ in the graph induced by $S_2 \cup M_1$ is blue and has partition sets X and Y where $X \subset S_2$ and $Y \subset M_1$ then we take a subset $X_1 \subseteq X$ with t vertices and by Proposition 4.2 X_1 has a red neighborhood M_2 of order m in B. Notice that the graph induced by $X_1 \cup Y \cup M_2$ is a $K_{t,2T}$ where one color forms a graph isomorphic to $K_{T,T}$.

If $S_2 \subset B$, there are at least $\frac{m^2}{2}$ edges of the same color, say red, between the sets S_1 and S_2 by the pigeonhole principle. Because $\frac{m^2}{2} \geq (T-1)^{\frac{1}{T}} n^{2-\frac{1}{T}} + \frac{1}{2}(T-1)n > z(m,T)$, there is a red $K_{T,T}$ contained in the graph induced by the sets S_1 and S_2. Let T_1 and T_2 be the partition sets of the red $K_{T,T}$, with $T_1 \subset S_1$ and $T_2 \subset S_2$. By Proposition 4.2 we know that any t-subset of T_1 has a blue common neighborhood M_1 of order m. The graph induced by $T_1 \cup T_2 \cup M_1$ is a $K_{t,2T}$ where one color forms a graph isomorphic to $K_{t,T}$. If the $K_{T,T}$ contained in the graph induced by the sets T_1 and T_2 is blue, we make a similar argument simply taking any t-subset from T_2 and its red common neighborhood in A. In this case, we also find a $K_{t,2T}$ where one color forms a graph isomorphic to $K_{t,T}$.

□

Notice that the parameter T allows the existence of large monochromatic complete bipartite graphs in every 2-edge coloring of $E(K_{n,n})$ with sufficient edges in each color. As a direct consequence, if $T = t$, we obtain only one colored pattern. This pattern will appear no matter the value of T, which is the reason we rather use the following corollary throughout this chapter.

Corollary 4.4. *Let t be a positive integer. For all large enough n, there exists a positive integer $z = z_2(n,t) = \mathcal{O}(n^{2-\frac{1}{t}})$ such that any coloring $E(K_{n,n}) = R \sqcup B$ with at least z_2 edges in each color class contains a colored copy of $K_{t,2t}$ where one color forms a graph isomorphic to $K_{t,t}$.*

The following proposition employs Theorem 1.17 to show some values for which $z_2(n,t)$ is tight. Note that any graph H on n vertices and $\mathrm{ex}(n,G)$ edges that does not contain G as a subgraph can be used to define a bipartite graph H' with n vertices in each part and $z(n,G)$ edges with no copies of G, by the proof of Proposition 1.11. With this in mind, one can give a 2-edge coloring of $E(K_{n,n})$ where one color, say red, forms the before mentioned graph H' and as a consequence, there are no red copies of G in this coloring. This can also be extended to a coloring of the edges of a complete bipartite graph, which is the idea of the proof of the lower bound of $z_2(n,t)$.

Proposition 4.5. *For sufficiently large n and integers s and t that satisfy $1 \leq s \leq 3 \leq t$ or $t \geq (s-1)! + 1$, Theorem 4.3 is tight.*

Proof. Recall Theorem 1.17 which states that for a sufficiently large n and integers s and t that satisfy $1 \leq s \leq 3 \leq t$ or $t \geq (s-1)! + 1$, the value of $\mathrm{ex}(n,G)$ is tight. The authors exhibit a graph H on at most n vertices which is $K_{s,t}$-free for the mentioned values of s and t. By Proposition 1.11, we can extend this graph H to a bipartite graph H' with at most n vertices in each part which is also $K_{s,t}$-free. We now give a 2-edge coloring $E(K_{n,n}) = R \sqcup B$ where one color, say blue, forms a copy of H' and the rest of the edges are red. Because $e(H') = \mathcal{O}(n^{2-\frac{1}{s}})$, the edge coloring has at least $\mathcal{O}(n^{2-\frac{1}{s}})$ edges in each color. Therefore, this coloring does not contain blue copies of $K_{s,t}$ when s and t satisfy $1 \leq s \leq 3 \leq t$ or $t \geq (s-1)! + 1$. Note that this coloring does not contain colored copies of $K_{T,s+T}$ where one color forms a graph isomorphic to $K_{s,T}$, or a copies of $K_{s,2T}$ where one color forms a graph isomorphic to $K_{s,T}$ for $T \geq t$ and the stated values of s and t. We have reached the desired result.

□

4.2 Bipartite omnitonality

We now define the concept of bipartite r-tonality and bipartite omnitonality on the base graph $K_{n,n}$, along with a characterization for bipartite r-tonal graphs and a proof that every tree is bipartite-omnitonal.

Definition 4.6. *Let G be a bipartite graph and r an integer with $0 \leq r \leq \lfloor \frac{e(G)}{2} \rfloor$. Let $\mathrm{bbal}_r(n,G)$ be the minimum integer, if it exists, such that every 2-edge coloring $E(K_{n,n}) = R \sqcup B$ with $\min\{|R|,|B|\} > \mathrm{bbal}_r(n,G)$ contains either an $(r, e(G)-r)-$colored copy of G,* *or an $(e(G)-r,r)-$colored copy of G. If* $\mathrm{bbal}_r(n,G)$ *exists for every n sufficiently large, we say that G is* bipartite r-tonal.

Now we can state the definition of bipartite omnitonality.

Definition 4.7. *Let G be a bipartite graph and let* $\mathrm{bot}(n,G)$ *be the minimum integer, if it exists, such that any 2-coloring $E(K_{n,n})=R\sqcup B$ with $\min\{|R|,|B|\}>\mathrm{bot}(n,G)$ contains an $(r,b)-$colored copy of G for any $r\geq 0$ and $b\geq 0$ such that $r+b=e(G)$. If* $\mathrm{bot}(n,G)$ *exists for a large enough n, we say that G is* bipartite-omnitonal. *For a bipartite-omnitonal graph G, let* $\mathrm{Bot}(n,G)$ *be the family of graphs $\mathcal{H}$ on exactly* $\mathrm{bot}(n,G)$ *edges such that there is a coloring $E(K_{n,n})=R\sqcup B$ with $|R|=\mathrm{bot}(n,G)$ and with no (r,b)-colored copy of G for some pair $r,b\geq 0$ with $r+b=e(G)$ and such that $G[R]$ is isomorphic to H.*

Observe that if G is bipartite r-tonal, then $\mathrm{bbal}_r(n,G)\leq\frac{n^2}{2}$. The same upper bound holds for $\mathrm{bot}(n,G)$ in the case of bipartite omnitonal graphs.

Theorem 4.8. *Let G be a bipartite graph and let r be a positive integer with $r\leq\lfloor\frac{e(G)}{2}\rfloor$. Then G is bipartite r-tonal if and only if there is a set of vertices $U\subseteq V(G)$ such that $e(U,N(U))=r$ and U is contained in one of the partition sets of G.*

Proof. Let G be bipartite r-tonal and n even and sufficiently large such that $\mathrm{bbal}_r(n,G)$ exists. Let the partition sets of $K_{n,n}$ be X and Y with $Y=Y_r\sqcup Y_b$ and $|Y_r|=\frac{n}{2}$. Let $E(K_{n,n})=R\sqcup B$ be a 2-edge coloring where the red edges are all possible edges between the sets X and Y_r and the blue edges are all possible edges between X and Y_b. Because G is bipartite r-tonal and $|R|=|B|=\frac{n^2}{2}$, we can guarantee the existence of a colored copy of G where one color class, say red, contains exactly r edges. Let $U=V(G)\cap Y_r$. This must be a set of independent vertices that satisfies that $e_G(U,N_G(U))=r$ and U is clearly contained in one of the partition sets of $K_{n,n}$, and hence, of G.

Conversely, let G be a bipartite graph that contains a set of vertices U such that $e_G(U,N_G(U))=r$ and U is contained in one of the partition sets of G. Let $t=n(G)$ and let $E(K_{n,n})=R\sqcup B$ be a 2-edge coloring, where n is large enough, with more than $z_2(n,t)$ edges in each color, just as in Theorem 4.3. By Corollary 4.4, there is a colored copy of a $K_{t,2t}$ where one color forms a $K_{t,t}$. Let $X=X_b\cup X_r$ and Y be the partition sets of this $K_{t,2t}$, where $|X_b|=|X_r|=|Y|=t$ and such that $E(X_b,Y)\subseteq B$ and $E(Y_r,X)\subseteq R$.

Let X_G and Y_G be the partition sets of $V(G)$ such that $U\subseteq X_G$. Now we can embed G in this $K_{t,2t}$ by placing the set U inside X_b, the set $X_G\setminus U$ inside X_r and the set Y_G inside Y. This copy of G has exactly r blue edges. Hence, G is bipartite r-tonal.

□

Similarly as in Theorem 4.8, we can prove a characterization of bipartite omnitonal graphs.

Theorem 4.9. *A bipartite graph G is bipartite omnitonal if and only if for every $0\leq r\leq e(G)$ there is a set of vertices $U_r\subseteq V(G)$ such that $e_G(U_r,N_G(U_r))=r$ and U_r is contained in one of the partition sets of G.*

Proof. Suppose G is bipartite omnitonal. Let n be even and large enough so that $\mathrm{bot}(n, G)$ exists. Consider an edge coloring $E(K_{n,n}) = R \sqcup B$ where one color forms a $K_{n,\frac{n}{2}}$. Since G is bipartite omnitonal and $\mathrm{bot}(n, G) \leq \frac{n^2}{2} = |R| = |B|$, there must be a copy of G in $K_{n,n}$ with r red edges for every $0 \leq r \leq e(G)$. This implies that, for every $0 \leq r \leq e(G)$, there is a set of vertices U_r such that $e_G(U_r, N_G(U_r)) = r$ and U_r is contained in one of the partition sets of G.

Conversely, suppose that for every $0 \leq r \leq e(G)$, there is a set of vertices $U_r \subseteq V(G)$ such that $e_G(U_r, N_G(U_r)) = r$ and U_r is contained in one of the partition sets of G. Let $E(K_{n,n}) = R \sqcup B$ be a 2-edge coloring with more than $z = z_2(n, t)$, where $t = n(G)$ and z is like in Corollary 4.4. Therefore, if n is sufficiently large, there is a copy of $K_{t,2t}$ (with partition sets X and Y) where one color, say red, forms a graph isomorphic to $K_{t,t}$ (with partition sets X and $Y_1 \subseteq Y$). Because there is a set of vertices $U_r \subseteq V(G)$ such that $e_G(U_r, N_G(U_r)) = r$ and U_r is contained in one of the partition sets of G, we can place U_r in Y_1 to find a copy of G with r red edges and $e(G) - r$ blue edges for every $0 \leq r \leq e(G)$. This concludes the proof. $\square$

As in the setting where K_n is the base graph (see Theorem 1.35) it turns out that all trees are bipartite-omnitonal.

Theorem 4.10. *Every tree T is bipartite-omnitonal.*

Proof. Let T be a tree with $e(T) = k$. In view of Theorem 4.8, we prove that, for every $0 \leq r \leq e(T)$, there is a set U_r in one of the partition sets of T such that $e_T(U_r, N_T(U_r)) = r$ by induction on k.

If $k = 1$, we may that U_0 as the empty set and U_1 as a single vertex. Both sets satisfy that $e_T(U_r, N_T(U_r)) = r$ for $r = 0, 1$. Let v be a leaf in $V(T)$ and let u be its only neighbor. By the induction hypothesis, the tree $T' = T - v$ is bipartite-omnitonal, which means that, for every $0 \leq r \leq k - 1$, there is a set U'_r contained in one of the partition sets of T' such that $e_{T'}(U'_r, N_{T'}(U'_r)) = r$.

For each r, there are two cases that can happen. If $u \in U'_r$, then $v \in V(T) \setminus U'_r$, and we may take $U_{r+1} = U'_r$ so that $e_T(U_{r+1}, N_T(U_{r+1})) = r + 1$. The second case is that $u \in V(T) \setminus U'_r$. In this case, we let $U_{r+1} = U'_r \cup \{v\}$. Hence for every $1 \leq r \leq k$, there is a set U_r that satisfies the theorem. The only case missing is when $r = 0$. In this case we can take U_0 as the empty set because $e_T(U_0, N_T(U_0)) = 0$. This concludes the proof.

$\square$

4.3 Bipartite balanceability and the bipartite balancing number

The bipartite balancing number is defined on the base graph $K_{n,n}$. We provide the bipartite-balancing number for paths and stars.

Definition 4.11. *If there exists an integer $k = k(n)$ such that, for n large enough every 2-edge coloring $R \sqcup B$ of $E(K_{n,n})$ with more than k edges in each color class contains a balanced copy of G, then we say G is* bipartite-balanceable. *The smallest such k is called the bipartite-balancing number of G and it is denoted as* $\mathrm{bbal}(n, G)$. *For a bipartite balanceable graph G, let* $\mathrm{Bbal}(n, G)$ *be the family of graphs with exactly* $\mathrm{bbal}(n, G)$ *edges such that a 2-edge coloring $E(K_{n,n}) = R \sqcup B$ with exactly* $\mathrm{bbal}(n, G)$ *edges in one color contains no balanced copy of G if and only if the graph induced by the red edges or by the blue edges is isomorphic to some $H \in$* $\mathrm{Bbal}(n, G)$.

The following corollary follows directly from Theorem 4.10.

Corollary 4.12. *Every tree T is bipartite-balanceable.*

We explore the value of $\mathrm{bbal}(n, G)$ for some families of trees. We will start with the bipartite-balancing number of paths. We prove the lower bound of $\mathrm{bbal}(P_k, n)$ by giving an explicit 2-edge coloring of $K_{n,n}$ with precisely $\lfloor \frac{k-2}{4} \rfloor n$ edges in one color and such that there is no balanced copy of P_k. The upper bound of $\mathrm{bbal}(P_k, n)$ is proven using induction on k. It is easy to see that applying an inductive step using one more edge trivially gives us a balanced path because the length of the path increases by one and hence the balanced property remains. Therefore, we will conduct the proof taking k to be even.

Theorem 4.13. *Let $n \geq 2$ and $k \geq 2$ be integers, where k is even. If $n > \frac{k-2}{2}\left(\frac{k}{2} - \left\lfloor \frac{k-2}{4} \right\rfloor\right)$, then*

$$\mathrm{bbal}(n, P_k) = \mathrm{bbal}(n, P_{k+1}) = \left\lfloor \frac{k-2}{4} \right\rfloor n + 0^{\alpha},$$

where $\alpha \in \{0, 2\}$ is such that $k \equiv \alpha \pmod 4$. Moreover, if $s = \left\lfloor \frac{k-2}{4} \right\rfloor$, then $\mathrm{Bbal}(n, P_k) = \mathrm{Bbal}(n, P_{k+1}) = \{K_{s,n}\}$, *for $k \equiv 2 \pmod 4$, and* $\mathrm{Bbal}(n, P_k) = \mathrm{Bbal}(n, P_{k+1}) = \{H_{s,n}\}$, *for $k \equiv 0 \pmod 4$, where $H_{s,n}$ is isomorphic to the graph $K_{s,n}$ with one extra pendant vertex adjacent to a vertex of the partite set with n vertices.*

Proof. Let $n \geq 2$ and $k \geq 2$ be integers, where k is even, and $n > \frac{k-2}{2}\left(\frac{k}{2} - \left\lfloor \frac{k-2}{4} \right\rfloor\right)$. Let $p(n, k) = \left\lfloor \frac{k-2}{4} \right\rfloor n + 0^{\alpha}$ where $\alpha \in \{0, 2\}$ is such that $k \equiv \alpha \pmod 4$. We determine the bipartite-balancing number of paths with an even number of edges. Once this is determined, the odd case is straightforward, and it is discussed at the end of the proof. We begin by proving that a 2-edge coloring $E(K_{n,n}) = R \sqcup B$ with more than $p(n, k)$ edges in each color is satisfiable. We proceed to prove that

$$e(K_{n,n}) = n^2 > 2p(n, k)$$

for all $n > \frac{k-2}{2}\left(\frac{k}{2} - \left\lfloor \frac{k-2}{4} \right\rfloor\right)$. If $k \equiv 0 \pmod 4$, then $2p(n, k) = 2(\left\lfloor \frac{k-2}{4} \right\rfloor n + 0^{\alpha}) = 2(\frac{k-4}{4} n + 1) < n^2$, because $n > \frac{k-4}{2} + 2$. If $k \equiv 2 \pmod 4$, then $2p(n, k) = 2(\left\lfloor \frac{k-2}{4} \right\rfloor n + 0^{\alpha}) = 2(\frac{k-2}{4}) n < n^2$, because $n > \frac{k-2}{2}$.

We proceed now to prove, by induction on k, that any 2-coloring of $K_{n,n}$ with at least $\mathrm{p}(n, k)$ edges in each color has either a balanced P_k or the edges in one of the colors induce the graph $B(k)$, where $B(k) = H_{s,n}$, if $\equiv 0 \pmod 4$, or $B(k) = K_{s,n}$, if $k \equiv 2 \pmod 4$. We assume there is no balanced P_k in order to arrive to one of these extremal colorings, each of which satisfies that every copy of P_k has at most $\frac{k}{2} - 1$ edges in one color class.

When $k = 2$, any 2-edge coloring of $K_{n,n}$ with at least one edge in each color is enough to find a balanced P_2, hence $\mathrm{bbal}(n, P_2) = 0$, while $K_{0,n}$ represents the extremal coloring which in this case is the empty graph. Let $k > 2$ be even, and assume the statement of the theorem is true for each even integer k' with $2 \leq k' < k$.

Suppose that $E(K_{n,n}) = R \sqcup B$ is a 2-edge coloring with at least $\mathrm{p}(n, k)$ edges in each color. It is a simple matter to check that the conditions for $k' = k - 2$ are satisfied. If $k \equiv 0 \pmod 4$ then $\mathrm{p}(n, k) = \lfloor \frac{k-2}{4} \rfloor n + 1 = \frac{k-4}{4} n + 1 = \lfloor \frac{k-4}{4} \rfloor n + 1 = \mathrm{p}(n, k')$, and if $k \equiv 2 \pmod 4$ then $\mathrm{p}(n, k) = \lfloor \frac{k-2}{4} \rfloor n = \frac{k-2}{4} n > \lfloor \frac{k-4}{4} \rfloor n = \mathrm{p}(n, k')$. We can assume that, by the induction hypothesis, there is a balanced copy of P_{k-2}, say $P = v_1 v_2 \cdots v_{k-1}$. Let $K_{n,n}$ have partite sets V and W, where $V_P = V \cap V(P)$ and $W_P = W \cap V(P)$. We will show that either there is a balanced P_k or we will have one of two extremal colorings depending on the value of k modulo 4.

We begin by analyzing the structure of the coloring outside of P. Note first that all of the edges that come out of $V(P)$ with an endpoint in $\{v_1, v_{k-1}\}$ must have the same color. Otherwise, if $v_1 v$ is red and $v_{k-1} u$ is blue with u, v being distinct vertices outside of $V(P)$, then $v v_1 P v_{k-1} u$ forms a balanced P_k. We assume, without loss of generality, that all such edges are blue, that is

$$E(\{v_1, v_{k-1}\}, (V \cup W) \setminus V(P)) \subseteq B. \tag{4.2}$$

By a similar argument, we can conclude that

$$E(V \setminus V_P, W \setminus W_P) \subseteq B. \tag{4.3}$$

Otherwise, if $xy \in E(V \setminus V_P, W \setminus W_P) \cap R$, then $x y v_1 P v_{k-1}$ makes a balanced P_k.

Claim 4. *Any edge with one endpoint outside of $V(P)$ that is incident to a blue edge in P, must be blue.*

Proof of Claim 4. We will show now that every time there is a blue edge in P, no red edges can come out of it. If $x \in V_P$, $y \in W_P$, $w \in W \setminus W_P$, $xy \in E(P) \cap B$ and $vy \in R$ for some $v \in V \setminus V_P$, then $x P v_1 w v y P v_{k-1}$ makes a balanced P_k (recall that $v_1 w$ and wv are blue). If now we take $x \in W_P$, $y \in V_P$, $xy \in E(P) \cap B$, $wy \in R$ for some $w \in W \setminus W_P$ and some $w' \in W \setminus W_P$ with $w' \neq w$, then $x P v_1 w y P v_{k-1} w'$ forms a balanced P_k. ▷

Let V_P^r be the set of vertices in V_P that make at least one red edge with a vertex in $W \setminus W_P$, and let W_P^r be the set of vertices in W_P that make at least one red edge with a vertex in $V \setminus V_P$. Therefore, by Claim 4 all vertices in $V_P^r \cup W_P^r$ are incident to two red edges from P. By definition,

$$e_R(v, W \setminus W_P) \neq 0, \; e_R(u, V \setminus V_P) \neq 0 \tag{4.4}$$

for all $v \in V_P^r$ and for all $u \in W_P^r$.

Note that every blue edge in P has one endpoint in $V_P \setminus V_P^r$ and the other endpoint in $W_P \setminus W_P^r$. Also every vertex in $V_P \setminus V_P^r$ and $W_P \setminus W_P^r$ is either incident to some blue edge in P, or it is incident to two red edges in P and makes only blue edges with $(V \setminus V_P) \cup (W \setminus W_P)$ (otherwise, it would be in $V_P^r \cup W_P^r$). Along with Claim 4, we can conclude that all vertices in $V_P^r \cup W_P^r$ are incident to two red edges from P, and

$$E(V_P \setminus V_P^r, W \setminus W_P) \cup E(W_P \setminus W_P^r, V \setminus V_P) \subseteq B. \tag{4.5}$$

Claim 5. *All red edges not in $E(P)$ incident to a red edge e in P must be incident to the same endpoint in e. Therefore, every red edge in $E(P)$ has an endpoint with only blue incident edges.*

Proof of Claim 5. If $xy \in E(P) \cap R$ and there are $u, v \in (V \cup W) \setminus V(P)$ such that $ux, vy \in R$, then $v_1PxuvyPv_{k-1}$ forms a balanced P_k since we already know that $uv \in B$. ▷

Notice that $|V_P^r \cup W_P^r| \leq \lfloor \frac{k-2}{4} \rfloor$ as from each vertex in $V_P^r \cup W_P^r$ there are two red edges from P and P has $\frac{k-2}{2}$ red edges. On the other hand, if $|V_P^r \cup W_P^r| < \lfloor \frac{k-2}{4} \rfloor$, say $|V_P^r \cup W_P^r| = |V_P^r| + |W_P^r| = \lfloor \frac{k-2}{4} \rfloor - q$ for some $1 \leq q \leq \left\lfloor \frac{k-2}{4} \right\rfloor$, then the number of red edges would not satisfy the hypothesis as we can see in the next bound that employs previous coloring observations.

$$\begin{aligned}
|R| &= e_R(V_P^r, W \setminus W_P) + e_R(W_P^r, V \setminus V_P) + e_R(V_P, W_P) \\
&\leq |V_P^r| \cdot |W \setminus W_P| + |W_P^r| \cdot |V \setminus V_P| + |V_P| \cdot |W_P| \\
&= |V_P^r| \cdot (n - |W_P|) + \left(\left\lfloor \frac{k-2}{4} \right\rfloor - |V_P^r| - q \right) (n - |V_P|) + |V_P| \cdot |W_P| \\
&\leq |V_P^r| \cdot (n - |W_P|) + \left(\left\lfloor \frac{k-2}{4} \right\rfloor - |V_P^r| - q \right) (n - |W_P|) + |V_P| \cdot |W_P| \\
&\leq (n - |W_P|) \left(|V_P^r| + \left\lfloor \frac{k-2}{4} \right\rfloor - |V_P^r| - q \right) + \frac{k}{2} \cdot \frac{k-2}{2} \\
&< n \left(\left\lfloor \frac{k-2}{4} \right\rfloor - q \right) + \frac{k^2}{4} \\
&\leq n \left(\left\lfloor \frac{k-2}{4} \right\rfloor - 1 \right) + \frac{k^2}{4}
\end{aligned}$$

considering Equations (4.3) and (4.5). However, this contradicts the fact that $|R| \geq \lfloor \frac{k-2}{4} \rfloor n$ due to the fact that $n > \frac{k-2}{2} \left(\frac{k}{2} - \left\lfloor \frac{k-2}{4} \right\rfloor \right)$ when $k \geq 4$.

Therefore,

$$|V_P^r \cup W_P^r| = \left\lfloor \frac{k-2}{4} \right\rfloor. \tag{4.6}$$

We can conclude that exactly $2 \left\lfloor \frac{k-2}{4} \right\rfloor$ red edges in P are incident to vertices from $V_P^r \cup W_P^r$. Therefore, we arrive to the following claim.

Claim 6. *Every red edge in P is incident to some vertex in $V_P^r \cup W_P^r$, except for one edge with endpoints in $V_P \setminus V_P^r$ and $W_P \setminus W_P^r$ in the case where $k \equiv 0 \pmod 4$.*

Now we can analyze the structure of the coloring within P to prove the next inclusion.

$$E(V_P \setminus V_P^r, W_P \setminus W_P^r) \setminus E(P) \subseteq B. \tag{4.7}$$

Proof of Equation (4.7) By Claim 6, recall that every vertex v in $(V_P \setminus V_P^r) \cup (W_P \setminus W_P^r)$ is either incident to some blue edge in P or it is incident to two red edges in P. Additionally, v makes only blue edges with $(V \setminus V_P) \cup (W \setminus W_P)$ by Equation (4.5). Let $v_i, v_j \in (V_P \cup W_P) \setminus (V_P^r \cup W_P^r)$ such that $i < j$. We go over all possible cases assuming $v_i v_j \in R$. In each case, we construct a balanced P_k.

Case 1: Let $v_i v_{i+1}, v_j v_{j+1} \in B$ with $v_i \in V$ and $v_j \in W$. Let $x, w \in V \setminus V_P$ and $y \in W \setminus W_P$. If $v_1, v_{k-1} \in V_P$, the path $v_1 P v_i v_j P v_{i+1} x y v_{k-1} P v_{j+1}$ makes a balanced P_k. If $v_1, v_{k-1} \in W_P$, the path $w v_1 P v_i v_j P v_{i+1} x v_{k-1} P v_{j+1}$ makes a balanced P_k.

This case is analogous by symmetry to the case where $v_{i-1} v_i, v_{j-1} v_j \in B$ with $v_i \in W$ and $v_j \in V$.

Case 2: Let $v_{i-1} v_i, v_{j-1} v_j \in B$ with $v_i \in V$ and $v_j \in W$. Let $x, y \in W \setminus W_P$ and $w \in V \setminus V_P$. If $v_1, v_{k-1} \in V_P$, the path $v_{i-1} P v_1 x v_{j-1} P v_i v_j P v_{k-1} y$ makes a balanced P_k. If $v_1, v_{k-1} \in W_P$, the path $v_{i-1} P v_1 w y v_{j-1} P v_i v_j P v_{k-1}$ makes a balanced P_k. This case is analogous by symmetry to the case where $v_i v_{i+1}, v_j v_{j+1} \in B$ with $v_i \in W$ and $v_j \in V$.

Case 3: Let $v_i v_{i+1}, v_{j-1} v_j \in B$. Let $x, w \in V \setminus V_P$ and $y \in W \setminus W_P$. We can assume that $v_{i-1} v_i, v_j v_{j+1} \in R$, otherwise, see Cases 1 and 2. Without loss of generality we can assume that, $v_i \in V$ and $v_j \in W$. If $v_1, v_{k-1} \in V_P$, the path $v_1 P v_i v_j P v_{k-1} y x v_{j-1} P v_{i+1}$ makes a balanced P_k. If $v_1, v_{k-1} \in W_P$, the path $w v_1 P v_i v_j P v_{k-1} x v_{j-1} P v_{i+1}$ makes a balanced P_k.

Case 4: Let $v_{i-1} v_i, v_j v_{j+1} \in B$. We can assume that $v_i v_{i+1}, v_{j-1} v_j \in R$, otherwise, see Cases 1 or 2. Without loss of generality, we can assume that $v_i \in V$ and $v_j \in W$. We can suppose $v_{i-1} v_{j+1} \in B$, otherwise, see Case 3. Suppose there is an edge $v_t v_{t+1} \in B$ with $i < t < j$. Let $x, y \in W \setminus W_P$ such that $x \neq y$ and $w \in V \setminus V_P$.

If $v_t \in V_P$ and $v_1, v_{k-1} \in V_P$, the path $v_{t+1} P v_j v_i P v_t x v_1 P v_{i-1} v_{j+1} P v_{k-1} y$ makes a balanced P_k. If $v_t \in V_P$ and $v_1, v_{k-1} \in W_P$, the path $v_{t+1} P v_j v_i P v_t x w v_1 P v_{i-1} v_{j+1} P v_{k-1}$ makes a balanced P_k.

If $v_t \in W_P$ and $v_1, v_{k-1} \in W_P$, the path $v_1 P v_{i-1} v_{j+1} P v_{k-1} w x v_{t+1} P v_j v_i P v_t$ makes a balanced P_k. Therefore, all edges in P between v_i and v_j must be red. Because there are no odd cycles in $K_{n,n}$, there are at least two vertices in $V(P)$ between v_i and v_j. By Claim 6, $v_{j-1} \in V_P^r$ or $v_{i+1} \in W_P^r$. Suppose $v_{j-1} \in V_P^r$. There are $x, y \in W \setminus W_P$, with $x \neq y$, and such that $v_{j-1} x \in R$. If $v_1, v_{k-1} \in V_P$, the path $v_{i-1} P v_1 x v_{j-1} P v_i v_j P v_{k-1} y$ makes a balanced P_k (recall $v_1 x, v_{k-1} y \in B$). If $v_1, v_{k-1} \in W_P$, the path $v_{i-1} P v_1 w x v_{j-1} P v_i v_j P v_{k-1}$ makes a balanced P_k. The case when $v_{i+1} \in W_P^r$ is analogous by symmetry.

Case 5: Let $v_i \in V_P \setminus V_P^r$, $v_j \in W_P \setminus W_P^r$. We consider four subcases. We prove each subcase assuming $v_{i-1} \in W_P \setminus W_P^r$ and $v_{i-1}v_i \in B$ or $v_{i+1} \in W_P \setminus W_P^r$ and $v_iv_{i+1} \in B$, when possible.

- *Subcase 5.1: Let $v_{j-1} \in V_P^r$ and $v_{j+1} \in V_P \setminus V_P^r$.* If $v_{i-1}v_i \in B$ for $i \geq 2$, then $v_1Pv_{i-1}xyv_{j-1}Pv_iv_jPv_{k-1}$ makes a balanced P_k for $x \in V \setminus V_P$, and $y \in W \setminus W_P$ such that $v_{j-1}y \in R$.

  If $v_iv_{i+1} \in B$ for $i \geq 1$ and $v_1, v_{k-1} \in V_P$, then $xv_1Pv_iv_jPv_{k-1}yv_{j-1}Pv_{i+1}$ makes a balanced P_k for $x, y \in W \setminus W_P$ and $v_{j-1}y \in R$. If $v_1, v_{k-1} \in W_P$, then $v_1Pv_iv_jPv_{k-1}yxv_{j-1}Pv_{i+1}$ is a balanced P_k for $x \in W \setminus W_P$ such that $v_{j-1}x \in R$, and $y \in V \setminus V_P$.

- *Subcase 5.2: Let $v_{j-1} \in V_P \setminus V_P^r$ and $v_{j+1} \in V_P^r$.* This implies that $v_{j-2}v_{j-1} \in B$. If $v_{i-1}v_i \in B$ for $i \geq 2$, then $v_1Pv_{i-1}xv_{j-2}Pv_iv_jv_{j-1}yv_{j+1}Pv_{k-1}$ is a balanced P_k for $x \in V \setminus V_P$ and $y \in W \setminus W_P$ such that $v_{j+1}y \in R$.

  If $v_iv_{i+1} \in B$ for $i \geq 1$, then $v_1Pv_iv_jv_{j-1}Pv_{i+1}xyv_{j+1}Pv_{k-1}$ is a balanced P_k for $x \in V \setminus V_P$ such that $v_{j-1}x \in R$ and $y \in W \setminus W_P$.

- *Subcase 5.3: Let $j = k-1$ and $v_{j-1} = v_{k-2} \in V_P^r$.* If $v_{i-1}v_i \in B$ for $i \geq 2$ and $v_1, v_{k-1} \in W_P$, then $xv_1Pv_{i-1}yv_{j-1}Pv_iv_j$ is a balanced P_k for $x \in V \setminus V_P$ and $y \in W \setminus W_P$ such that $v_{j-1}y \in R$. If $v_1, v_{k-1} \in V_P$, then $v_1Pv_{i-1}wyv_{j-1}Pv_iv_j$ is a balanced P_k for $w \in W \setminus W_P$ and $y \in V \setminus V_P$ such that $v_{j-1}y \in R$.

  If $v_iv_{i+1} \in B$, for $i \geq 1$ and $v_1, v_{k-1} \in W_P$, then $v_1Pv_iv_jyxv_{j-1}Pv_{i+1}$ is a balanced P_k, for $x \in W \setminus W_P$ such that $v_{j-1}x \in R$, and $y \in V \setminus V_P$. If $v_1, v_{k-1} \in V_P$, then $wv_1Pv_iv_jxv_{j-1}Pv_{i+1}$ is a balanced P_k where $w \in V \setminus V_P$.

- *Subcase 5.4: Let $i = 1$, $v_1v_2 \in R$, and $v_{j-1} \in V_P^r$.* There are vertices $x \in W \setminus W_P$ and $y \in (V \cup W) \setminus V(P)$ such that $xv_{j-1} \in R$ and $v_{k-1}y \in B$ by Equation (4.2). The path $xv_{j-1}Pv_1v_jPv_{k-1}y$ makes a balanced P_k.

  There are vertices $x, y \in W \setminus W_P$ such that $xv_{j-2} \in B$ and $v_{j+1}y \in R$ by Equation (4.5). The path $v_{j-1}v_jv_1Pv_{j-2}xyv_{j+1}Pv_{k-1}$ makes a balanced P_k.

$\triangleright$

By the previous observations, note that the edges in $E(V \setminus V_P^r, W \setminus W_P^r)$ are all blue if $k \equiv 2 \pmod 4$. If $k \equiv 0 \pmod 4$, the same holds except for one red edge which is in $E(V_P \setminus V_P^r, W_P \setminus W_P^r)$. Also note that, in both cases $E(V_P^r, W_P^r) \subseteq B$. We proceed to prove that one of the sets, V_P^r or W_P^r, must be empty. Recall that, $|V_P^r| + |W_P^r| = \lfloor \frac{k-2}{4} \rfloor$. Without loss of generality suppose that $|V_P^r| \geq |W_P^r|$. If $|W_P^r| \geq 1$, the number of red edges is bounded as follows.

$$
\begin{aligned}
|R| &= e_R(V_P^r, W) + e_R(W_P^r, V) \\
&\leq |V_P^r|(n - |W_P^r|) + |W_P^r|(n - |V_P^r|) \\
&\leq |V_P^r|(n - |W_P^r|) + |W_P^r|(n - |W_P^r|) \\
&= (|V_P^r| + |W_P^r|)(n - |W_P^r|) \\
&\leq \left\lfloor \frac{k-2}{4} \right\rfloor (n-1)
\end{aligned}
$$

This contradicts the fact that $|R| \geq \lfloor \frac{k-2}{4} \rfloor n$. Therefore, W_P^r must be empty. Thus, we can conclude that

i) If $k \equiv 0 \pmod 4$ then the red edges form a $K_{s,n}$ with $s = \lfloor \frac{k-2}{4} \rfloor$ and an extra pendant vertex adjacent to a vertex of the partite set with n vertices.

ii) If $k \equiv 2 \pmod 4$ then the red edges form a $K_{s,n}$, with $s = \lfloor \frac{k-2}{4} \rfloor$.

Observe that $\text{bbal}(n, P_{k+1}) = \text{bbal}(n, P_k)$ for $k \geq 2$ even. This holds in any 2-edge coloring of $E(K_{n,n})$ where there is a balanced copy of P_k, because every balanced P_k can be extended to a balanced P_{k+1} by considering one more edge adjacent to one of its end vertices. Regardless of its color, the extended path P_{k+1} has $\left\lfloor \frac{k}{2} \right\rfloor$ edges in one color and $\left\lceil \frac{k}{2} \right\rceil$ in the other. The fact that $\text{Bbal}(n, P_k) = \text{Bbal}(n, P_{k+1})$ holds, because the extremal coloring associated to $\text{Bbal}(n, P_k)$ also avoids copies of balanced P_{k+1}. Hence, the bipartite-balancing number of paths of odd length is also determined. This concludes the proof.

□

We now state the exact bipartite-balancing number for stars.

Theorem 4.14. *Let n and k be integers with k even, $k \geq 2$ and $n \geq \frac{k^2}{2} + k - 2$, then*

$$
\text{bbal}(n, K_{1,k}) = \text{bbal}(n, K_{1,k+1}) = (k-2)\left(n - \frac{k-2}{4}\right)
$$

and $\text{Bbal}(n, K_{1,k}) = \text{Bbal}(n, K_{1,k+1})$ *contains only one graph which is a bipartite graph H with partition sets A and B where $A = A_1 \sqcup A_2$ and $B = B_1 \sqcup B_2$ with $|A_1| = |B_2| = n - \frac{k-2}{2}$ and $|A_2| = |B_1| = \frac{k-2}{2}$. The edges of H are all possible edges between A_1 and B_1 and between A_2 and B.*

Proof. Let n and k be integers with k even, $k \geq 2$ and $n \geq \frac{k^2}{2} + k - 2$. We determine the bipartite-balancing number for stars of even number of edges. We can easily extend this result for stars of odd number of edges. It is simple to see that $\text{bbal}(n, K_{1,k}) = \text{bbal}(n, K_{1,k+1})$ when $k \geq 2$ is even. Note that, in every 2-edge coloring of $K_{n,n}$ where there is a balanced $K_{1,k}$, we can extend this copy to a balanced $K_{1,k+1}$ by considering one more edge adjacent to the vertex of maximum degree. Regardless of its color, the extended star $K_{1,k+1}$ has $\left\lfloor \frac{k}{2} \right\rfloor$ edges in one color and $\left\lceil \frac{k}{2} \right\rceil$ in the other. The fact that $\mathrm{Bbal}(n, K_{1,k}) = \mathrm{Bbal}(n, K_{1,k+1})$ holds, because the extremal coloring associated to $\mathrm{Bbal}(n, K_{1,k})$ also avoids copies of balanced $K_{1,k+1}$.

Therefore, we prove the result for stars of even number of edges. Let $a(n,k) := (k-2)(n - \frac{k-2}{4})$. We begin by proving that a 2-edge coloring $E(K_{n,n}) = R \sqcup B$ with more than $a(n,k)$ edges in each color is satisfiable. We proceed to prove that

$$e(K_{n,n}) = n^2 > 2a(n,k)$$

for all $n \geq \frac{k^2}{2} + k - 2$. Note that $n^2 > 2a(n,k) = 2(k-2)(n - \frac{k-2}{4})$ is equivalent to $(n-k+2)^2 - \frac{(k-2)^2}{2} > 0$ and this is true when $n > (k-2)(1 + \frac{1}{\sqrt{2}})$ and $k \geq 4$, which holds by the hypothesis.

Let H be the bipartite graph described in the theorem statement. Observe that H has exactly $|A_1||B_1| + |A_2||B| = (n - \frac{k-2}{2})\frac{k-2}{2} + \frac{k-2}{2}n = (k-2)(n - \frac{k-2}{4}) = a(n,k)$ edges.

Note that any 2-edge coloring $E(K_{n,n}) = R \sqcup B$, where the graph induced by one of the color classes is isomorphic to H, does not contain a balanced copy of $K_{1,k}$. If a vertex v in this coloring has at least $\frac{k}{2}$ neighbors in one color, it will have less than $\frac{k}{2}$ neighbors in the other color making the existence of a balanced $K_{1,k}$ impossible. Hence, $\mathrm{bbal}(n, K_{1,k}) \geq a(n,k)$ and $H \in \mathrm{Bbal}(n, K_{1,k})$ is proved.

To prove the upper bound $\mathrm{bbal}(n, K_{1,k}) \leq a(n,k)$ and $\mathrm{Bbal}(n, K_{1,k}) = \{H\}$, we must show that any 2-edge coloring $E(K_{n,n}) = R \sqcup B$, with at least $a(n,k)$ edges in each color and such that neither the red graph nor the blue graph is isomorphic to H, must contain a balanced copy of $K_{1,k}$.

Let $V(K_{n,n})$ have the bipartition $V \sqcup W$. We define the following sets

$$V_r = \{v \in V \mid \deg_R(v) \geq \frac{k}{2}\}, \qquad V_b = \{v \in V \mid \deg_B(v) \geq \frac{k}{2}\}.$$

$$W_r = \{v \in W \mid \deg_R(v) \geq \frac{k}{2}\}, \qquad W_b = \{v \in W \mid \deg_B(v) \geq \frac{k}{2}\}.$$

Let $E(K_{n,n}) = R \sqcup B$ be a 2-edge coloring with at least $a(n,k)$ edges in each color and such that neither the red graph nor the blue graph is not isomorphic to H. If there is a vertex in $V_r \cap V_b$ or in $W_r \cap W_b$, then the graph induced by this vertex and its neighbors contains a balanced $K_{1,k}$. Therefore, we will assume that $V_r \cap V_b = \emptyset$ and $W_r \cap W_b = \emptyset$ and consequently that $V = V_r \sqcup V_b$ and $W = W_r \sqcup W_b$. This partition and the fact that any vertex $v \in V_R$ satisfies that $\deg_B(v) \leq \frac{k}{2} - 1$ imply that $\deg_R(v) \geq n - \frac{k}{2} + 1$ for all $v \in V_r$. Analogously, $\deg_B(v) \geq n - \frac{k}{2} + 1$ for all $v \in V_b$, $\deg_R(w) \geq n - \frac{k}{2} + 1$ for all $w \in W_r$, and $\deg_B(w) \geq n - \frac{k}{2} + 1$ for all $w \in W_b$. We separate the proof in two cases.

Case 1. One of $\{V_r, V_b, W_r, W_b\}$ has at most $\frac{k}{2} - 1$ elements. Without loss of generality, suppose V_r to be such set, hence $|V_r| \leq \frac{k}{2} - 1$. Now we bound the number of red edges.

$$
\begin{aligned}
a(k,n) \leq |R| &= e_R\left(V_r, W\right) + e_R(V_b, W) \\
&\leq |V_r| \cdot n + |V_b|(\frac{k}{2} - 1) \\
&= |V_r| \cdot n + (n - |V_r|)\left(\frac{k}{2} - 1\right) \\
&= n \cdot \left(\frac{k}{2} - 1\right) + |V_r|\left(n - \frac{k}{2} + 1\right) \\
&\leq n \cdot \left(\frac{k}{2} - 1\right) + \left(\frac{k}{2} - 1\right)\left(n - \frac{k}{2} + 1\right) \\
&= a(k,n).
\end{aligned}
$$

Because the equality is reached, we can conclude the following:

- $\deg_R(v) = n$ for all $v \in V_r$, which implies that $\deg_B(v) = 0$ for all $v \in V_r$;
- $\deg_R(v) = \frac{k}{2} - 1$ for all $v \in V_b$, which implies that $\deg_B(v) = n - \frac{k}{2} + 1$ for all $v \in V_b$;
- and $|V_r| = \frac{k}{2} - 1$, which implies that $|V_b| = n - \frac{k}{2} + 1$.

This also implies that every vertex in W is connected to every vertex in V_r by a red edge. Note that there cannot be any red edges connecting V_b and W_b. Otherwise, if vw is a red edge with $v \in V_b$ and $w \in W_b$ then the vertex w would have more than $\frac{k}{2} - 1$ red neighbors. Therefore, all edges connecting V_b and W_b must be blue.

If $|W_r| \leq \frac{k}{2} - 1$, then

$$
\begin{aligned}
a(n,k) \leq |R| &= e_R(W_r, V) + e_R(W_b, V) \\
&\leq |W_r| \cdot n + \left(\frac{k}{2} - 1\right)|W_b| \\
&= |W_r| \cdot n + (n - |W_r|)\left(\frac{k}{2} - 1\right) \\
&= n \cdot \left(\frac{k}{2} - 1\right) + |W_r|\left(n - \frac{k}{2} + 1\right) \\
&\leq n \cdot \left(\frac{k}{2} - 1\right) + \left(\frac{k}{2} - 1\right)\left(n - \frac{k}{2} + 1\right) \\
&= a(n,k).
\end{aligned}
$$

This implies that $|W_r| = \frac{k}{2} - 1$, $\deg_R(w) = n$ for all $w \in W_r$ and $\deg_R(w) = \frac{k}{2} - 1$ for all $w \in W_b$, which means that the red graph is isomorphic to H, a contradiction to the assumption we made at the beginning.

Therefore we can assume that $|W_r| = \frac{k}{2}-1+s$, for some $s \geq 1$. Because $\deg_B(v) = n-\frac{k}{2}+1$ for all $v \in V_b$ and all edges vw with $v \in V_b$ and $w \in W_b$ are blue, then $e_B(v, W_r) = s$ for $v \in V_b$.

This implies that $e_B(W_r, V_b) = |V_b| \cdot s = (n - \frac{k}{2} + 1) \cdot s$. Recall that $\deg_B(w) \leq \frac{k}{2} - 1$ for all $w \in W_r$. We obtain the following inequality.

$$s\left(n - \frac{k}{2} + 1\right) = e_B(W_r, V_b) \leq |W_r|\left(\frac{k}{2} - 1\right) = \left(\frac{k}{2} - 1 + s\right)\left(\frac{k}{2} - 1\right)$$

Solving for n, we obtain

$$n \leq \left(\frac{k}{2} - 1\right)^2 \frac{1}{s} + 2\left(\frac{k}{2} - 1\right) \leq \left(\frac{k}{2} - 1\right)^2 + (k-2) = \frac{k^2}{4} - 1.$$

This is a contradiction to the fact that $n \geq \frac{k^2}{2} + k - 2$.

Case 2. Suppose that $|V_r|, |V_b|, |W_r|, |W_b| \geq \frac{k}{2}$. Without loss of generality, we can assume that $|W_b| = \max\{|V_r|, |V_b|, |W_r|, |W_b|\}$. This implies that $|W_r| = \min\{|V_r|, |V_b|, |W_r|, |W_b|\}$.

Let $|V_r| = |W_r| + t$ for some $t \geq 0$. We use the following two bounds on $|R|$.

$$|R| = e_R(V_r, W) + e_R(V_b, W) \geq |V_r|\left(n - \frac{k}{2} + 1\right) = (|W_r| + t)\left(n - \frac{k}{2} + 1\right).$$

$$|R| = e_R(W_r, V) + e_r(W_b, V) \leq n \cdot |W_r| + \left(\frac{k}{2} - 1\right)|W_b| = n \cdot |W_r| + \left(\frac{k}{2} - 1\right)(n - |W_r|).$$

Joining both bounds, we get

$$n\left(t - \frac{k}{2} + 1\right) \leq t\left(\frac{k}{2} - 1\right).$$

Solving for t, we get the following.

$$t \leq \frac{n\left(\frac{k}{2} - 1\right)}{n - \frac{k}{2} + 1} = \frac{k}{2} - 1 + \frac{\left(\frac{k}{2} - 1\right)^2}{n - \frac{k}{2} + 1} < \frac{k}{2}$$

The last inequality holds if $n - \frac{k}{2} + 1 > (\frac{k}{2} - 1)^2$, which is true due to the hypothesis that $n \geq \frac{k^2}{2} + k - 2$.

This means that $t \leq \frac{k}{2} - 1$. Note that $|V_b| = n - |V_r| = n - |W_r| - t$. We bound $e(V_b, W_r)$ by using the following.

$$e_B(W_r, V_b) \leq |W_r|\left(\frac{k}{2} - 1\right);$$

$$e_R(V_b, W_r) \leq |V_b|\left(\frac{k}{2} - 1\right).$$

The number of edges $e(V_b, W_r)$ is bounded as follows.

$$\begin{aligned}
(n - |W_r| - t)|W_r| &= |V_b| \cdot |W_r| \\
&= e(V_b, W_r) \\
&\leq (|V_b| + |W_r|)\left(\frac{k}{2} - 1\right) \\
&= (n-t)(\frac{k}{2} - 1).
\end{aligned}$$

This is equivalent to:

$$n\left(|W_r| - \frac{k}{2} + 1\right) \leq |W_r|^2 + t \cdot |W_r| - t \cdot (\frac{k}{2} - 1).$$

Using the fact that $|W_r| - \frac{k}{2} + 1 > 0$, we arrive at the next inequality for n. We leave a detailed discussion of the last inequality at the end of the proof.

$$n \leq \frac{|W_R|^2 + t(|W_r| - \frac{k}{2} + 1)}{|W_r| - \frac{k}{2} + 1} = \frac{|W_r|^2}{|W_r| - \frac{k}{2} + 1} + t \leq \frac{n}{2} + \frac{k^2}{4} - 1. \tag{4.8}$$

This means that $n \leq \frac{k^2}{2} - 2$, which is a contradiction to $n \geq \frac{k^2}{2} + k - 2$ for $k \geq 2$.

We conclude the proof by justifying the last inequality in Equation (4.8). Let $x = |W_r|$ and let f be the following function depending on x, $f(x) = \frac{x^2}{x - \frac{k}{2} + 1}$. The function $f(x)$ can be bounded as follows.

$$\begin{aligned}
f(x) = \frac{x^2}{x - \frac{k}{2} + 1} &= \frac{\left(x - \left(\frac{k}{2} - 1\right)\right)^2 + 2x\left(\frac{k}{2} - 1\right) - \left(\frac{k}{2} - 1\right)^2}{x - \frac{k}{2} + 1} \\
&= x - \frac{k}{2} + 1 + \frac{\left(\frac{k}{2} - 1\right)\left(x - \frac{k}{2} + 1\right) + x\left(\frac{k}{2} - 1\right)}{x - \frac{k}{2} + 1} \\
&= x + \left(\frac{k}{2} - 1\right)\frac{x}{x - \left(\frac{k}{2} - 1\right)} \\
&\leq \frac{n}{2} + \left(\frac{k}{2} - 1\right)\frac{\frac{k}{2}}{\frac{k}{2} - (\frac{k}{2} - 1)} \\
&= \frac{n}{2} + \frac{k^2}{4} - \frac{k}{2}.
\end{aligned}$$

The last inequality is due to the fact that $x \leq \frac{n}{2}$ and that $g(x) = \frac{x}{x - \left(\frac{k}{2} - 1\right)}$ is a decreasing function defined on $\frac{k}{2} \leq x \leq \frac{n}{2}$. Therefore, the maximum value of $g(x)$ is reached when $x = \frac{k}{2}$. This bound on $f(x)$ and the fact that $t \leq \frac{k}{2} - 1$ provide the last inequality in Equation (4.8).

$$n \leq \frac{|W_r|^2}{|W_r| - \frac{k}{2} + 1} + t \leq \frac{n}{2} + \frac{k^2}{4} - \frac{k}{2} + \frac{k}{2} - 1 = \frac{n}{2} + \frac{k^2}{4} - 1.$$

□

4.4 Open problems

We state here some open problems left for future work that are relevant to the topics discussed in the $K_{n,n}$ setting.

1. We would like to determine a tight bound for the bipartite balancing number of trees.
2. Determine $\mathrm{bot}(n, P_k)$ and $\mathrm{bot}(n, K_{1,k})$.
3. Search for other graphs that are bipartite-balanceable and/or bipartite-omnitonal.
4. Determine the bipartite-balancing number and/or bipartite-omnitonal number of other graph families.
5. Define and study amoebas in the context of $K_{n,n}$.
6. Study other base structures, such as complete multi-partite graphs, grids in different surfaces (like the plane, the torus, the projective plane, etc.) and hypergraphs.
7. Consider the problem of determining the unavoidable patterns in different base structures and arbitrary number of colors.

Chapter 5

Conclusion

The research work conducted during the elaboration of this dissertation has explored recent concepts in the area of combinatorics and graph theory in the context of unavoidable patterns. One of the objectives of this PhD program was to study balanceability and unavoidable patterns in 2-edge colorings of the complete graph and afterwards identify the balanceability of some graph classes, find the balancing number of balanceable cycles, and define a more general parameter as an extension of the balancing number. We were also able to give a recursive construction of a family of local amoebas, determine the unavoidable patterns in 2-edge colorings of the base graph $K_{n,n}$, characterize the balanceability and omnitonality of graphs in $K_{n,n}$ and determine (or bound) the bipartite balancing number in $K_{n,n}$ of trees in general and of paths and stars. Two research papers were produced and accepted for publication. One of them is already published. The first research paper, called *On the balanceability of some graph classes* [22], was published in the journal Discrete Applied Mathematics. The second research paper, called *The balancing number and generalized balancing number of some graph classes* [20], was recently accepted in the journal Electronic Journal of Combinatorics. There are two more research papers in progress regarding the results in Chapter 3 and Chapter 4 which will be sent for publication in 2023. Although the objective was acquired, there are still many questions remaining. We discuss such questions in the next section.

5.1 Future work

This section includes all of the problems stated in the Open Problems Sections of Chapters 2, 3 and 4 that remain unanswered and are left as future work. Other new problems came up in our research concerning distinct topics in combinatorics. These problems are discussed in the last part called Other research directions and will also be left to tackle in the future.

Problems that remain

1. Considering the complexity of the problem of determining whether a graph is balanceable or not. This problem boils down to finding an edge cut with $\lfloor\frac{1}{2}e(G)\rfloor$ or $\lceil\frac{1}{2}e(G)\rceil$ edges

and an induced subgraph with $\lfloor \frac{1}{2}e(G) \rfloor$ or $\lceil \frac{1}{2}e(G) \rceil$ edges. Particularly, the problem of finding an edge cut with half of the edges is a variant of the problem EXACT-CUT (which studies the complexity of finding an edge cut with exactly k edges and it is NP-complete). The other problem deals with the existence of an induced subgraph with exactly half of the edges. We conjecture that the balancerability problem is NP-complete.

2. Determine the balanceability of K_n for $n > 2,303,999,904,000,003$ when $n \equiv 2,3 \pmod 4$.

3. What other graph families satisfy being balanceable? We are interested in determining the balanceability of graphs such as planar graphs, outerplanar graphs, chordal graphs, k-trees and other circulant graphs.

4. Study the generalized balancing number of other graphs such as K_n with $n \geq 5$, $2K_n$, circulant graphs, wheel graphs and grids.

5. Generalize the construction of Lemma 3.10.

6. Look for other general ways of constructing global or local amoebas.

7. Characterize local amoeba trees.

8. Characterize global amoeba trees.

9. Study the balancing number of other amoeba families by using constructions.

10. Generalize the constructions given by Lemma 3.10 and Lemma 3.11, and the ones stated in [12].

11. Search for other characterizations of local and global amoebas by defining new recursive constructions.

12. Determine a characterization of local and/or global amoeba trees by using constructions.

13. Study the balancing number of other amoeba families by using constructions.

14. We would like to determine a tight bound for the bipartite balancing number of trees.

15. Determine $\mathrm{bot}(n, P_k)$ and $\mathrm{bot}(n, K_{1,k})$.

16. Search for other graphs that are bipartite-balanceable and/or bipartite-omnitonal.

17. Determine the bipartite-balancing number and/or bipartite-omnitonal number of other graph families.

18. Define and study amoebas in the context of $K_{n,n}$.

19. Study other base structures, such as complete multi-partite graphs, grids in different surfaces (like the plane, the torus, the projective plane, etc.) and hypergraphs.

20. Consider the problem of determining the unavoidable patterns in different base structures and arbitrary number of colors.

The following is a brief description of new open problems in other research directions.

Other research directions

There are new interesting questions that arise from the initial problem like what other unavoidable patterns can exist in different base structures. Given a base graph H, can we guarantee the existence of unavoidable patterns in any 2-edge coloring of H with sufficient edges in each color? Possible interesting base graphs could be simplicial complexes or complete hypergraphs. If the answer to this question is affirmative, it would be interesting to determine the unavoidable patterns, as well as the threshold of the amount of edges each color class requires in order to maintain their existence. We would also like to know what properties must the base graph H satisfy so that the existence of unavoidable patterns is assured. One could also study how certain properties such as density, regularity, vertex transitivity, and other structural properties on the base structure affect the existence of unavoidable patterns. Finally, all these questions concerning 2-colorings could be extended to a setting of an arbitrary number of colors.

5.2 Conclusion

We have achieved several results concerning the balanceability of some graph classes, as well as sufficient conditions for a graph to be balanceable. In terms of the balancing number, we have arrived at exact values of the balancing number of C_{4k-1} and C_{4k+1}, as well as tight bounds of the same order corresponding to $\mathrm{bal}(n, C_{4k})$. In addition, we defined the generalized balancing number, which turned out to be an extension of the traditional balancing number. Also, using this new concept, we were able to give some general bounds, as well as the exact value of $\mathrm{bal}^*(n, C_{4k+2})$ and bounds on the generalized balancing number of K_5. We worked with a family of graphs called local amoebas for which we described a recursive way to construct infinite families of them. We also exposed some examples of such infinite families. Finally, we determined the unavoidable patterns in 2-edge colorings of $K_{n,n}$ and in this same setting we defined the bipartite balancing number and the bipartite omnitonal number. We bounded the bipartite balancing number for trees in general and we determined the bipartite balancing number of paths and stars.

Agradecimientos

Alphabetical Index

Bibliography

[1] Abbott, H., & Hanson, D. (1972). A problem of Schur and its generalizations. *Acta Arithmetica*, 2(20), 175-187.

[2] Aharoni, R., DeVos, M., de la Maza, S. G. H., Montejano, A., & Šámal, R. (2018). A rainbow version of Mantel's Theorem. *arXiv preprint arXiv:1812.11872.*

[3] Alon, N., Krivelevich, M., & Sudakov, B. (2003). Turán numbers of bipartite graphs and related Ramsey-type questions. *Combinatorics, Probability and Computing.* 12(5-6), 477-494.

[4] Alon, Noga; Rónyai, Lajos; Szabó, Tibor Norm-graphs: variations and applications. *J. Combin. Theory Ser.* B 76 (1999), no. 2, 280–290.

[5] Bollobás, B. (2004). Extremal graph theory. *Courier Corporation.*

[6] Bowen, M., Hansberg, A., Montejano, A., & Müyesser, A. (2019). Colored unavoidable patterns and balanceable graphs. *arXiv preprint arXiv:1912.06302.*

[7] Brown, W. G. On graphs that do not contain a Thomsen graph. *Canad. Math. Bull.* 9 (1966), 281–285.

[8] Caro, Y., González I., Hansberg A., Jácome M., Matos T. & Montejano A. Graphs with constant balancing number. *Work in progress.*

[9] Caro, Y., Hansberg, A., Lauri, J., & Zarb, C. (2022). On Zero-Sum Spanning Trees and Zero-Sum Connectivity. *The Electronic Journal of Combinatorics,* P1-9..

[10] Caro, Y., Hansberg, A., & Montejano, A. (2019). Zero-Sum K_m Over $\mathbb{Z}$ and the Story of K_4. *Graphs and Combinatorics*, 35(4), 855-865.

[11] Caro, Y., Hansberg, A., & Montejano, A. (2019). Zero-sum subsequences in bounded-sum $\{-1, 1\}$-sequences. *Journal of Combinatorial Theory, Series A*, 161, 387-419.

[12] Caro, Y., Hansberg, A., & Montejano, A. (2020). Graphs isomorphisms under edge-replacements and the family of amoebas. *arXiv preprint arXiv:2007.11769.*

[13] Caro, Y., Hansberg, A., & Montejano, A. (2021). Unavoidable chromatic patterns in 2-colorings of the complete graph. *Journal of Graph Theory*, 97(1), 123-147.

[14] Caro, Y., Hansberg, A., & Montejano, A. (2022). The evolution of unavoidable bichromatic patterns and extremal cases of balanceability. *arXiv preprint arXiv:2204.04269*.

[15] Caro, Y., Lauri, J., & Zarb, C. (2020). On small balanceable, strongly-balanceable and omnitonal graphs. *Discussiones Mathematicae Graph Theory*, 42(4), 1219-1235.

[16] Chartrand, G., Lesniak, L., & Zhang, P. (2010). Graphs & digraphs (Vol. 39). CRC press.

[17] Conlon, D. (2009). A new upper bound for diagonal Ramsey numbers. *Annals of Mathematics*, 941-960.

[18] Conlon, D., Fox, J., & Sudakov, B. (2015). Recent developments in graph Ramsey theory. *Surveys in combinatorics*, 424(2015), 49-118.

[19] Cutler, J., & Montágh, B. (2008). Unavoidable subgraphs of colored graphs. *Discrete mathematics*, 308(19), 4396-4413.

[20] Dailly, A., Eslava, L., Hansberg, A., & Ventura, D. (2022). The balancing number and generalized balancing number of some graph classes. *The Electronic Journal of Combinatorics*, 29, P00.

[21] Dailly, A., Hansberg, A., Talon, A., & Ventura, D. (2022). On the balanceability of K_n. *Preprint*.

[22] Dailly, A., Hansberg, A., & Ventura, D. (2020). On the balanceability of some graph classes. *Discrete Applied Mathematics*, 291, 51-63.

[23] Erdős, P. (1984). Some combinatorial, geometric and set theoretic problems in measure theory. *In Measure Theory Oberwolfach 1983* (pp. 321-327). Springer, Berlin, Heidelberg.

[24] Erdős, P. (1947). Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53.

[25] Erdős, P.; Rényi, A.; Sós, V. T. On a problem of graph theory. *Studia Sci. Math. Hungar.* 1 (1966), 215–235.

[26] Erdős, P.; Simonovits, M. A limit theorem in graph theory. *Studia Sci. Math. Hungar. 1* (1966), 51–57.

[27] Erdős, P.; Stone, A. H. On the structure of linear graphs. *Bull. Amer. Math. Soc. 52* (1946), 1087–1091.

[28] Erdős, P.; Szekeres, G. (1935). A combinatorial problem in geometry. *Compositio mathematica*, 2, 463-470.

[29] Eslava, L., Hansberg, A., Matos, T., & Ventura, D. (2022). A recursive construction of local amoebas. *Preprint*.

[30] Espinosa Hernández, J. (2020) Gráficas inevitables en 2-coloraciones de la gráfica completa: el caso de las amoebas, Tesis de licenciatura, Facultad de Ciencias, UNAM.

[31] Generic two integer variable equation solver *https://www.alpertron.com.ar/QUAD.HTM*.

[32] Kalbfleisch, J. G. (1966). Chromatic graphs and Ramsey's theorem. Ph. D. Thesis, University of Waterloo.

[33] Kollár, János; Rónyai, Lajos; Szabó, Tibor. Norm-graphs and bipartite Turán numbers. *Combinatorica 16* (1996), no. 3, 399–406.

[34] Kővari, T.; Sós, V. T.; Turán, P. On a problem of K. Zarankiewicz. *Colloq. Math. 3.* (1954), 50–57.

[35] Fox, J., & Sudakov, B. (2008). Unavoidable patterns. *Journal of Combinatorial Theory, Series A*, 115, no. 8, 1561-1569.

[36] Fox, J., & Sudakov, B. (2009). Density theorems for bipartite graphs and related Ramsey-type results. *Combinatorica*, 29(2), 153-196.

[37] Fűredi, Z. (1991). On a Turán type problem of Erdős. *Combinatorica.* 11(1), 75-79.

[38] Fűredi, Z., & Simonovits, M. (2013). The history of degenerate (bipartite) extremal graph problems. *Erdös Centennial* (pp. 169-264). Springer, Berlin, Heidelberg.

[39] Girão, A., & Narayanan, B. (2022). Turán theorems for unavoidable patterns. *Mathematical Proceedings of the Cambridge Philosophical Society*, 172(2), 423-442. doi:10.1017/S030500412100027X.

[40] Golomb, S. W., & Baumert, L. D. (1965). Backtrack programming. *Journal of the ACM (JACM)*, 12(4), 516-524.

[41] Graham, R. L. (1990). Rothschild, and Spencer. Ramsey theory.

[42] Graver, J. E., & Yackel, J. (1968). Some graph theoretic results associated with Ramsey's theorem. *Journal of Combinatorial Theory*, 4(2), 125-175.

[43] Greenwood, R. E., & Gleason, A. M. (1955). Combinatorial relations and chromatic graphs. *Canadian Journal of Mathematics*, 7, 1-7.

[44] Grinstead, C. M., & Roberts, S. M. (1982). On the Ramsey numbers R (3, 8) and R (3, 9). *Journal of Combinatorial Theory, Series B*, 33(1), 27-51.

[45] Hansberg, A., Montejano, A., & Caro, Y. (2021). Recursive constructions of amoebas. *Procedia Computer Science*, 195, 257-265.

[46] Hansberg, A. & Ventura, D. (2022). Unavoidable patterns, balanceability and omnitonality in the complete bipartite graph. Preprint.

[47] Heule, M. (2018, April). Schur number five. *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).

[48] Hilbert, D. (1892). Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. *Journal für die reine und angewandte Mathematik*, 1892(110), 104-129.

[49] Lidicky, B., Liu, H., Palmer, C. (2012). On the Turán number of forests. *arXiv:1204.3102v1*

[50] Mantel, W. (1907). Problem 28. *Wiskundige Opgaven*, 10(60-61), 320.

[51] McKay, B. D., & Radziszowski, S. P. (1995). R(4, 5)=25, J. Graph Theory 19, 309-322.

[52] Ning, B., & Wang, J. (2020). The formula for Turán number of spanning linear forests. *Discrete Mathematics*, 343(8), 111924.

[53] Ramsey, F. P. (1929) On a Problem of Formal Logic. *Proc. London Math. Soc.* (2) 30, no. 4, 264–286.

[54] Ramsey, F. P. (2009). On a problem of formal logic. In *Classic Papers in Combinatorics* (pp. 1-24). Birkhäuser Boston.

[55] Rotman, J. J. (2012). An introduction to the theory of groups (Vol. 148). *Springer Science & Business Media.*

[56] Schur, I. (1917). Über Kongruenz $x^m + y^m \equiv z^m \pmod{p}$. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 25, 114-116.

[57] Simonovits, M., & Sós, V. T. (2001). Ramsey–Turán theory. *Discrete Mathematics*, 229(1-3), 293-340.

[58] Soifer, A. (Ed.). (2010). Ramsey Theory: Yesterday, today, and tomorrow (Vol. 285). *Springer Science & Business Media.*

[59] Spencer, J. (1975). Ramsey's theorem—a new lower bound. *Journal of Combinatorial Theory*, Series A, 18(1), 108-115.

[60] Turán, P. (1941). On an external problem in graph theory. *Mat. Fiz. Lapok*, 48, 436-452.

[61] Underwood, D. (1978). Elementary number theory. Second edition. *W.H. Freeman and Company New York.*

[62] Van der Waerden, B. L. (1927). Beweis einer baudetschen Vermutung. *Nieuw Arch. Wiskunde*, 15, 212-216.