BOUNDING THE NUMBER OF GRAPHS CONTAINING VERY LONG INDUCED PATHS

by

Steven Kay Butler

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mathematics

Brigham Young University

April 2003

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Steven Kay Butler

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Wayne W. Barrett, Chair
Date	James W. Cannon
Date	Rodney W. Forcade

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Steven Kay Butler in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

____________________	____________________
Date	Wayne W. Barrett Chair, Graduate Committee

Accepted for the Department

____

Tyler J. Jarvis
Graduate Coordinator

Accepted for the College

____

G. Rex Bryce, Associate Dean
College of Physical and Mathematical Sciences

ABSTRACT

BOUNDING THE NUMBER OF GRAPHS CONTAINING VERY LONG INDUCED PATHS

Steven Kay Butler

Department of Mathematics

Master of Science

Induced graphs are used to describe the structure of a graph, one such type of induced graph that has been studied are long paths.

In this thesis we show a way to represent such graphs in terms of an array with two colors and a labeled graph. Using this representation and the techniques of Polya counting we will then be able to get upper and lower bounds for graphs containing a long path as an induced subgraph.

In particular, if we let $\mathcal{P}(n,k)$ be the number of graphs on $n+k$ vertices which contains P_n, a path on n vertices, as an induced subgraph then using our upper and lower bounds for $\mathcal{P}(n,k)$ we will show that for any fixed value of k that $\mathcal{P}(n,k) \sim 2^{(nk+\binom{k}{2})}/2k!$.

Contents

Introduction **1**

1 Graph theory **1**

1.1 Paths 2

1.2 Induced subgraphs 3

2 Counting our graphs **3**

2.1 Some special cases 4

3 Polya counting **5**

3.1 Counting colorings with Burnside's Lemma 6

3.2 Polya counting and graph theory 8

4 Coloring and our problem **8**

4.1 Multiple coloring representations for a fixed graph 9

4.2 Working with representative colorings 10

5 Counting our inequivalent colorings **11**

5.1 Stirling numbers of the first kind 11

5.2 Counting our inequivalent colorings 13

6 Upper bound **16**

7 Lower bound **17**

7.1 Having a long induced path in only one way 17

7.2 Colorings fixed only under the identity automorphism 19

7.3 The lower bound 22

8 Asymptotic behavior **24**

Conclusion **26**

References **27**

Introduction

In graph theory, important information about the structure of a graph can be obtained from knowing what induced subgraphs it does or does not contain. Some graphs are completely characterized in this manner. For instance, a forest is a graph that does not contain a cycle as an induced subgraph, and a bipartite graph is a graph which does not contain an odd cycle as an induced subgraph.

For the thesis we want to consider the graphs which contain paths as induced subgraphs, and more particularly very long paths as induced subgraphs (very long in the sense that the induced path will involve most of the vertices of the graph). Such graphs have been investigated in [1], [4] and [8].

Graphical enumeration is a branch of graph theory which counts the number of graphs with a given structure. Classical problems in graphical enumeration include counting the number of trees and the number of connected graphs; many of these results are available in [6]. Some problems in graphical enumeration prove unwieldy and the best results involve getting an order of magnitude on the growth.

For this thesis we will examine the problem of enumerating the number of graphs which contain very long paths as an induced subgraph. We will let $\mathcal{P}(n,k)$ represent the number of graphs on $n+k$ vertices which contain P_n, a path on n vertices, as an induced subgraph. Then we will establish upper and lower bounds for $\mathcal{P}(n,k)$ and from these bounds we will show that if we fix k then

$$\mathcal{P}(n,k) \sim \frac{2^{(nk+\binom{k}{2})}}{2k!}.$$

1 Graph theory

This section contains a brief overview of the elements of graph theory that will arise in this thesis. More detailed and complete information about graph theory can be found in [2] and [7].

A graph, $G=(V,E)$, is a pair of sets V and E where V is the vertex set and $E \subset V \times V$ is the edge set. We will follow the convention of denoting the elements of E without bracketing, that is we will use ab instead of (a,b) or $\{a,b\}$.

In this thesis we will assume that all graphs are simple graphs. This means that the elements of E are unordered (i.e., the edge ab is the same as the edge ba), E cannot have repeated elements (we cannot have multiple edges between two vertices) and no element of E has the form aa (we do not have loops).

The terminology for vertex and edge come from the way of representing a graph with a picture. Each vertex of V is associated with a point in the plane (or other appropriate space) and each edge of E is a line connecting the two vertices. An example of a graph and its corresponding picture in the plane are shown in Figure 1.

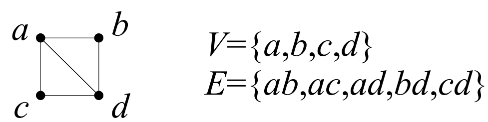

Figure 1: A graph

Two vertices in a graph are adjacent if there is an edge connecting them. For the graph in Figure 1 vertices a and d are adjacent while vertices b and c are not.

The degree of a vertex v, denoted by $\deg(v)$, is the number of times that v appears as an element in the unordered pairs of E. Graphically, it is the number of edges that connect with the vertex. For the graph in Figure 1 we have $\deg(a) = 3$ and $\deg(b) = 2$.

Two graphs, $G = (V, E)$ and $H = (W, F)$ are isomorphic if there exists a bijection $\phi : V \to W$ so that $ab \in E$ if an only if $\phi(a)\phi(b) \in F$. That is, two graphs are isomorphic if we can send the vertex set of G to the vertex set of H in a way that preserves adjacency.

1.1 Paths

Many of the commonly encountered graphs have been given special names. These names are given to either describe the structure of the graph (such as the claw, a wheel, or a cocktail party graph) or are named after mathematicians who used the graph in a new and important way (such as the Petersen graph).

The graph that we will focus on is the path. A path with n vertices, denoted by P_n, is a graph that is isomorphic with the graph $G = (V, E)$ where

$$\begin{aligned} V &= \{a_1, a_2, a_3, \ldots, a_{n-1}, a_n\} \\ E &= \{a_1a_2, a_2a_3, \ldots, a_{n-1}a_n\}. \end{aligned}$$

Graphical examples of paths are shown in Figure 2.

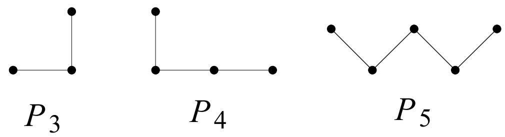

Figure 2: Examples of paths

1.2 Induced subgraphs

A subgraph $G' = (V', E')$ of a graph $G = (V, E)$ is a graph with $V' \subset V$ and $E' \subset E$. If E' contains all of the edges of E that connect the vertices of V' then the graph is an induced subgraph. A subgraph and an induced subgraph of the graph shown in Figure 1 are shown in Figure 3.

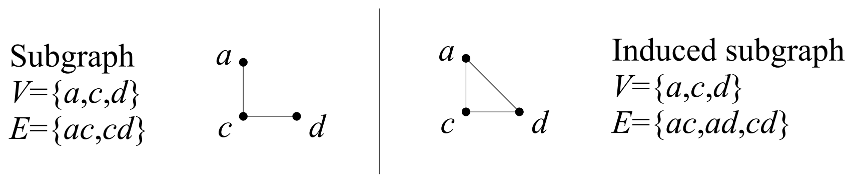

Figure 3: A subgraph and an induced subgraph of the graph in Figure 1

As stated in the introduction, the induced subgraphs can give important information about the structure of a graph and are often used to describe classes of graphs.

2 Counting our graphs

For small values of n and k we can determine $\mathcal{P}(n, k)$ by examining all of the graphs on $n + k$ vertices. For example, we can examine the eleven graphs on four vertices and determine that $\mathcal{P}(3, 1) = 6$, the graphs on four vertices are shown in Figure 4 with the graphs containing P_3 as an induced subgraph boxed.

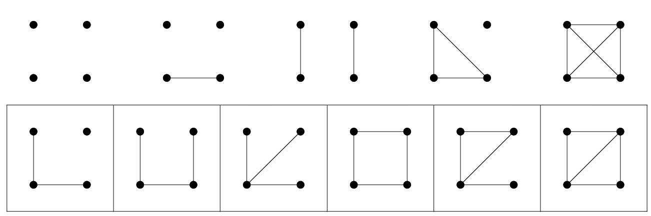

Figure 4: $\mathcal{P}(3, 1) = 6$

This method of examination quickly becomes undesirable as $n + k$ gets large because of the high growth rate of the number of graphs. For example, in order to determine $\mathcal{P}(19, 4)$ by inspection we would have to examine the

$$559, 946, 939, 699, 792, 080, 597, 976, 380, 819, 462, 179, 812, 276, 348, 458, 981, 632$$

different graphs on 23 vertices (information about counting the number of graphs on a given number of vertices is found in [6]).

Examining all of these graphs is not a desirable approach. As an alternative, we will approach our enumeration problem by using Polya counting techniques.

2.1 Some special cases

Before beginning into Polya counting techniques, we will find the value of $\mathcal{P}(n, k)$ in several special cases.

For example, P_1 consists of a graph with a single vertex and so every graph on n vertices contains P_1 as an induced subgraph. Similarly P_2 consists of two vertices and an edge connecting the vertices. So every graph on n vertices which has an edge will contain P_2 as an induced subgraph, in particular every graph except one (i.e., the graph with n vertices and no edges) will contain P_2 as an induced subgraph.

To count the number of graphs which contain P_3 as an induced subgraph we first introduce some more graph theory terms. A graph is connected if between any two vertices, a and b, of the graph there is a series of edges where the first edge is incident with a, the last edge is incident with b and any two consecutive edges share a common vertex. Pictorially, a graph is connected if for any two vertices we can connect them by tracing over edges of our graph without lifting our pencil.

The distance between two vertices in a connected graph is the smallest number of edges that we have to use to connect the two vertices. The diameter of a connected graph is the greatest distance between two vertices of the graph. The only graph with a diameter of 0 is the graph which consists of a single vertex and is called the complete graph on one vertex. The graphs that have a diameter of 1 are also complete graphs and are graphs with two or more vertices where all of the vertices are mutually adjacent.

We claim that any connected graph that has a diameter greater than 1 must contain P_3 as an induced subgraph. This can be seen by noting that since the diameter is greater than 1 there must be two vertices, c_0 and c_k, and a sequence of edges $c_0c_1, c_1c_2, \ldots, c_{k-1}c_k$ with $k \geq 2$ so that no shorter sequence of edges connects c_0 and c_k. The induced subgraph on the vertices $\{c_0, c_1, c_2\}$ must be P_3 since if c_0c_2 is an edge of the graph then the sequence of edges $c_0c_2, c_2c_3, \ldots, c_{k-1}c_k$ is a shorter sequence of edges that connects c_0 and c_k.

The maximal connected induced subgraphs of a graph are called the components of the graph. From the previous argument it follows that the only way that a graph can not contain P_3 as an induced subgraph is if all of the components of the graph are complete graphs.

Hence, all of the information about a graph which does not contain P_3 as an induced subgraph is contained in the size (i.e., the number of vertices) of the components. The sequence of numbers denoting the size of the components is a partition of the number of vertices. Further, every partition of the number of vertices corresponds to a graph which does not contain P_3 as an induced subgraph.

So let $p(n)$ denote the number of partitions of n. Then the number of graphs on n vertices that do not contain P_3 as an induced subgraph is $p(n)$. This can be seen in Figure 4 where the five graphs which do not contain P_3 as an induced subgraph correspond to the five partitions of four, i.e.,

$$1 + 1 + 1 + 1 = 2 + 1 + 1 = 2 + 2 = 3 + 1 = 4.$$

There is no easy characterization of graphs on n vertices which do not contain P_4 or any longer path as an induced subgraph until we get to P_n. In particular, there is only one graph on n vertices which contains P_n as an induced subgraph, the graph P_n.

If we let $\mathcal{U}(n)$ denote the number of non-isomorphic graphs on n vertices then we can summarize what we have so far in the following,

$$\begin{array}{rclcrcl} \mathcal{P}(1, n-1) &=& \mathcal{U}(n), & \quad & \mathcal{P}(3, n-3) &=& \mathcal{U}(n) - p(n), \\ \mathcal{P}(2, n-2) &=& \mathcal{U}(n) - 1, & & \mathcal{P}(n, 0) &=& 1. \end{array}$$

To get bounds for the rest of the values of $\mathcal{P}(n, k)$ we will use the methods of Polya counting.

3 Polya counting

Polya counting is used to count colorings of objects which have symmetry.

The perennial introduction to Polya counting is coloring the corners of a square with 2 different colors. Since there are four corners and each can be colored in one of two ways there are a total of $2^4 = 16$ different colorings of the corners.

However, some of these colorings are the same, or are equivalent, in the sense that we can get from one to the other by rotating or flipping the square (in other words, by using the symmetry of the square). We can group the 16 colorings into 6 equivalence classes of colorings as shown in Figure 5, where q and r are the two different colors. So there are only 6 ways to color the corners of a square with two different colors when we account for symmetry.

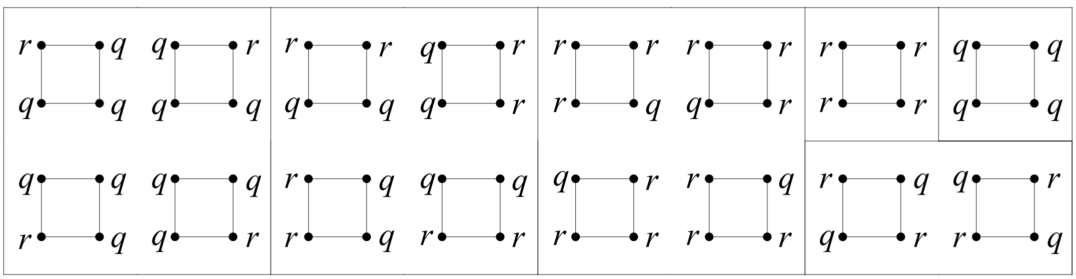

Figure 5: Colorings of the square

This example is easy to work through by inspection, but as the complexity of the object changes, inspection become more difficult.

3.1 Counting colorings with Burnside's Lemma

To count the number of inequivalent colorings we will use Burnside's Lemma, but first we need to introduce some terminology.

Let $\mathcal{C}$ be the collection of all of the colorings of our object and let A be a group that acts on $\mathcal{C}$ (it is the structure of this group which contains information about the symmetry of the object). Information about how groups act on sets can be found in [5].

Two colorings $c_1, c_2 \in \mathcal{C}$ are equivalent if there exists $\pi \in A$, such that $c_2 = \pi \cdot c_1$ where $\pi \cdot c_1$ denotes the action of π on the coloring c_1. This action then induces an equivalence relationship on the colorings.

Given $\pi \in A$ let $\mathcal{C}_\pi = \{c \in \mathcal{C} \,:\, \pi \cdot c = c\}$. That is, $\mathcal{C}_\pi$ is the set of colorings that are invariant under the action π.

In the example using the square if we label the corners 1, 2, 3 and 4 as in Figure 6 then the group acting on the colorings is the dihedral group of order eight, i.e., D_8. The action of the elements of the dihedral group are contained in the following permutations of the four corners:

Rotations:	(1)(2)(3)(4),	(1243),	(14)(23),	(1342)
Flips:	(1)(23)(4),	(12)(34),	(14)(2)(3),	(13)(24)

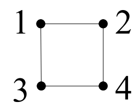

Figure 6: Labeling the corners of the square

If our two colors are q and r and we use $\pi = (13)(24)$ then we have

$$\mathcal{C}_\pi = \{qqqq,\, qrqr,\, rqrq,\, rrrr\}.$$

In this example we have $|\mathcal{C}_\pi| = 4 = 2^2$. In general, if π is composed of i disjoint cycles and we are coloring with m different colors then we will have that $|\mathcal{C}_\pi| = m^i$ since each cycle must be uniform in its choice of the m colors.

With our notation in place we can now state Burnside's Lemma which will allow us to count the number of inequivalent colorings of $\mathcal{C}$. Proofs of Burnside's Lemma can be found in [3] and [7].

Theorem 1 (Burnside's Lemma). *Let N be the number of equivalence classes of our colorings induced by the action of A. Then*

$$N = \frac{1}{|A|}\sum_{\pi \in A} |\mathcal{C}_\pi|.$$

Stated differently, the number of equivalence classes is the average of the number of colorings that are fixed by the permutations of A.

As a verification we return once again to our example of coloring the corners of the square. We can use the permutations already given and get,

$$N = \frac{1}{8}\big(2^4 + 2(2^3) + 3(2^2) + 2(2^1)\big) = 6,$$

which we have already confirmed through examination.

3.2 Polya counting and graph theory

For Polya counting to be useful for graph theory we will abstract this process of counting colorings.

For example, to use Polya counting to count the number of non-isomorphic graphs on n vertices we consider the colorings on the $\binom{n}{2}$ possible edges of the graph with the two colors 'yes' and 'no.' The action of the edges are the ones that result from the $n!$ permutations of the vertices.

We will use a similar approach to our problem by rephrasing the problem into one which uses coloring.

4 Coloring and our problem

In order to use the techniques of Polya counting for counting the number of graphs containing paths as induced subgraphs we need to introduce the idea of coloring into our problem.

Suppose that G is a graph on $n + k$ vertices and that the induced subgraph on the vertices of $N \subset V$ is P_n. Then there are three important structures that are present in G, namely P_n (which is the induced subgraph on the vertex set N), the induced subgraph on the vertex set $V \setminus N$ and the edges that connect the two vertex sets N and $V \setminus N$.

To represent the graph in terms of these three structures we introduce the P_n coloring representation of G with respect to N.

Definition 1. Let $G = (V, E)$ be a graph on $n + k$ vertices with vertex set $N \subset V$ which induces P_n. Then a P_n coloring representation of G with respect to N is a $k \times n$ array and a graph on k labeled vertices where

(i) each vertex of N corresponds to a column of the array and two columns are adjacent if and only if the corresponding vertices are adjacent.

(ii) the set $V \setminus N$ is labeled with the labels $\{1, 2, \ldots k\}$ where the ith vertex corresponds to the ith row of the array, the corresponding labeled induced subgraph of the set $V \setminus N$ is the graph on k labeled vertices.

(iii) each entry of the array is colored 'yes' if the vertex that corresponds to the column is adjacent to the vertex that corresponds to the row, otherwise the entry is colored 'no.'

When representing a P_n coloring representation of G graphically we will use dark for 'yes' and white for 'no.' We can also suppress the labeling on the graph with k vertices by putting each of the k vertices next to its corresponding row. An example of a P_n coloring representation of a graph is shown in Figure 7.

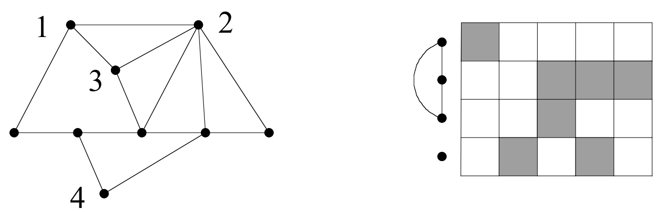

Figure 7: A graph and one of its P_n coloring representations

4.1 Multiple coloring representations for a fixed graph

Given a P_n coloring representation of a graph there is a unique graph that corresponds to that coloring. However, given a graph which contains P_n as an induced subgraph there are possibly many P_n coloring representations of G. This occurs for two reasons.

The first reason that this occurs is that there was some liberty in how we constructed our P_n coloring representation of G when we had a fixed N. In the definition there were two places where we were allowed a choice, namely how we corresponded vertices to the columns and the rows.

First note that when associating the vertices of the path with the columns, that in a path there are two vertices with degree one. When we associate a vertex of N with degree one it will have to go to either the first or last column. Once we have positioned the vertex, the rest of the positions for the vertices of the path are determined. In particular, we have only two ways to associate vertices of N with the columns of the array, the only difference being that we reverse the order of the columns.

The labeling of the vertices of the induced subgraph on the vertices of $V \setminus N$ was arbitrary. So there were $k!$ ways of labeling these vertices. The effect on our array caused by a change of labeling will appear as a permutation of the rows in our array.

Let $A = S_k \times \mathbb{Z}_2$ and let an element of A act on a coloring of the array by permuting the rows by the element of S_k and reversing the order of the columns if it is not the identity element of $\mathbb{Z}_2$. Then for two colorings c_1, c_2 of the array that are associated with P_n coloring representations with respect to the same N, there exists $\pi \in A$ so that $c_2 = \pi \cdot c_1$. This follows by noting that the difference in choices corresponds to a permutation of the rows and possibly reversing the order of the columns, which corresponds to an action of the array by one of the elements of A.

Conversely, if c_1 is the coloring of an array in a P_n coloring representation with respect to N and $\pi \in A$, then $c_2 = \pi \cdot c_1$ is also the coloring of an array in a P_n coloring representation with respect to the same N. This follows by noting that the action of π corresponds to making different choices in how we associate vertices with the rows and columns of the array, and so by changing our choices as dictated by π we can get a P_n coloring representation with respect to N that uses the coloring c_2.

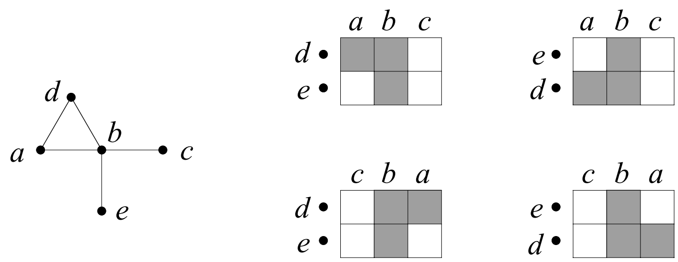

Figure 8: Different arrays arising from the P_n coloring representation

This establishes the following lemma.

Lemma 1. Let N be a vertex set that induces P_n in the graph G. Then the collection of all of the colorings of the arrays of the P_n coloring representations with respect to N are an equivalence class induced by the action of A on all of the colorings of the array.

An example of what happens is shown in Figure 8.

The other reason that we can have multiple P_n coloring representations for the same graph is that we can have P_n as an induced subgraph of G in more than one way. An example of this situation is shown in Figure 9.

We will wait until we get to the lower bound to deal with this second problem.

4.2 Working with representative colorings

From Lemma 1 we only need to work with representative elements of our equivalence classes. So let $\mathcal{C}$ be a collection of representative colorings of the $k \times n$

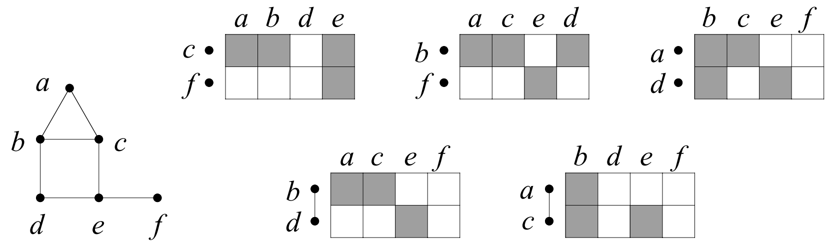

Figure 9: A graph with multiple P_4 representations

array with two colors, i.e., every coloring of our $k \times n$ array is equivalent with exactly one coloring of $\mathcal{C}$.

So if G is a graph with a P_n coloring representation with respect to N then the coloring of the array is equivalent to exactly one of the colorings of $\mathcal{C}$. In particular, by modifying our choices in our P_n coloring representation with respect to N we get a P_n coloring representation with respect to N of G so that the array has one of the colorings of $\mathcal{C}$.

It is important to note that we have still not overcome the arbitrariness of how we choose to associate vertices with the rows and the columns. Ambiguity can still arise when there is a non-identity element $\pi \in A$ that the coloring is invariant under. Again, we will examine this in more detail when we come to dealing with the lower bounds.

5 Counting our inequivalent colorings

The elements of $\mathcal{C}$ will form the basis for our upper and lower bounds, so we need to determine the size of $\mathcal{C}$. This can be most easily achieved by use of Theorem 1 (Burnside's Lemma).

Before finding the size of $\mathcal{C}$ it is useful to examine the structure and properties of the permutation group on k elements, i.e., S_k. The most important information for us will be given by the Stirling numbers of the first kind.

5.1 Stirling numbers of the first kind

The Stirling numbers of the first kind, denoted by $s(k, i)$ with i and k non-negative integers, are defined recursively by

$$s(k, i) = s(k-1, i-1) + (k-1)s(k-1, i)$$

with the boundary conditions

$$s(k,0)=0 \quad (k \geq 1), \qquad\qquad s(k,k)=1.$$

Applying the recursive definition and the boundary conditions to $s(k,k)$ we have for $k \geq 2$ that

$$s(k,k)=s(k-1,k-1)+(k-1)s(k-1,k) \quad \text{which yields} \quad s(k-1,k)=0.$$

This process can be repeated to show that $s(k,i)=0$ whenever $k<i$.

The Stirling numbers of the first kind, $s(k,i)$, count the number of permutations of S_k that consist of i cycles in the permutations cycle representation. This is a consequence of the following theorem found in [3].

Theorem 2. *The Stirling number of the first kind, $s(k,i)$, counts the number of arrangements of k objects into i non-empty circular permutations.*

This combined with the following theorem will aid us in counting the number of inequivalent colorings of $\mathcal{C}$.

Theorem 3. *Let $s(k,i)$ denote the Stirling numbers of the first kind. Then*

$$\sum_{i=1}^{k} s(k,i)x^i = \prod_{i=0}^{k-1}(x+i).$$

Proof. Let $g_k(x)=\sum_{i=1}^{k} s(k,i)x^i$. Then the recurrence relationship for the Stirling numbers of the first kind, along with the boundary conditions, translate into a recurrence relationship for the functions $g_k(x)$. Namely, we have the following.

$$\begin{aligned} g_k(x) &= \sum_{i=1}^{k} s(k,i)x^i \;=\; \sum_{i=1}^{k}(s(k-1,i-1)+(k-1)s(k-1,i))x^i \\ &= x\sum_{i=2}^{k} s(k-1,i-1)x^{i-1}+(k-1)\sum_{i=1}^{k} s(k-1,i)x^i \\ &= x\sum_{i=1}^{k-1} s(k-1,i)x^i+(k-1)\sum_{i=1}^{k-1} s(k-1,i)x^i \;=\; (x+(k-1))g_{k-1}(x) \end{aligned}$$

From the boundary conditions for the Stirling numbers of the first kind we have that $g_1(x) = x$ and so we have,

$$\begin{aligned}\sum_{i=1}^{k} s(k,i)x^i &= g_k(x) = (x+(k-1))g_{k-1}(x) \\ &= \cdots = (x+(k-1))(x+(k-2))\cdots(x+1)g_1(x) \\ &= (x+(k-1))(x+(k-2))\cdots(x+1)x = \prod_{i=0}^{k-1}(x+i).\end{aligned}$$

□

5.2 Counting our inequivalent colorings

Before counting our inequivalent colorings we need to introduce some notation that will be used in the theorem.

Recall that any element of S_k can be written uniquely up to rearrangement as the disjoint product of cycles; this is referred to as the cycle representation. So if $\sigma \in S_k$ then we will let $e(\sigma)$ denote the number of even cycles in the cycle representation of σ and $o(\sigma)$ denote the number of odd cycles in the cycle representation of σ.

Theorem 4. *Let $|\mathcal{C}|$ denote the number of representative elements of the colorings of the array under the action A, i.e., the number of equivalence classes under the action A. Then*

$$\begin{aligned}|\mathcal{C}| &= \frac{1}{2k!}\left(\sum_{\sigma\in S_k}\left[2^{n(e(\sigma)+o(\sigma))} + 2^{ne(\sigma)+\lfloor (n+1)/2\rfloor o(\sigma)}\right]\right) \\ &= \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + \sum_{\sigma\in S_k} 2^{ne(\sigma)+\lfloor (n+1)/2\rfloor o(\sigma)}\right) \\ &\leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2\rfloor k}\right)\end{aligned}$$

Proof. To apply Burnside's Lemma we need to determine the number of colorings of the array that are fixed under the action of π for each $\pi \in A$. Let $\pi = \sigma \times a$ where $\sigma \in S_k$ and $a \in \mathbb{Z}_2$.

Then, we will determine the number of colorings of the array that are fixed under π by looking at the behavior of the cycles in the cycle representation of σ in three cases.

(i) *A cycle in σ where a is the identity, i.e., we do not reverse the order of the columns.* We can arbitrarily choose the coloring for the row that corresponds to the first element of our cycle. Then in order for our coloring to remain fixed every other row corresponding to the elements of the cycle must have the same coloring. (Imagine that the rows were sequential, then we would color the first row and then copy the coloring down to every row that corresponded to our cycle.)

In particular, for every cycle we get to make n choices.

(ii) *An even cycle in σ where a is not the identity, i.e., we do reverse the order of the columns.* We choose an arbitrary coloring of the row that corresponds to the first element of the cycle. Then as we go through the elements of the cycle we reverse the order and fill in the rows as we go. (Imagine that the rows were sequential, then we would color the first row and then we would move down and reverse the row to color the next one and continue the process of moving down and reversing the row to color in the rest of the rows that corresponded to our cycle.)

When we return to the row that corresponds to the first element of the cycle we will have made an even number of reversals and so we will match up with what we started with. In particular, for every cycle we get to make n choices.

(iii) *An odd cycle in σ where a is not the identity, i.e., we do reverse the order of the columns.* We choose an arbitrary coloring of the row that corresponds to the first element of the cycle. Then as we go through the elements of the cycle we reverse the order and fill in the rows as we go, as in the previous case.

When we return to the row that corresponds to the first element of the cycle we will have made an odd number of reversals and in particular will be the reverse of what we started with. In order to match up with what we started with (i.e., in order for the coloring to remain fixed) the coloring of the first row has to be symmetric. In particular, for every cycle we get to make $\lfloor (n+1)/2 \rfloor$ choices, that is we get to color half of one row which will then determine the rest of the coloring for the cycle.

From (i) we have that when π is of the form $\sigma \times a$ and a is the identity then for every cycle in σ we get n choices and so there are $2^{n(e(\sigma)+o(\sigma))}$ colorings fixed by π.

Combining (ii) and (iii) we have that when π is of the form $\sigma \times a$ and a is not the identity, then for every even cycle in σ we get n choices and for every odd cycle in σ we get $\lfloor (n+1)/2 \rfloor$ choices and so there are $2^{ne(\sigma)+\lfloor (n+1)/2 \rfloor o(\sigma)}$ colorings fixed by π.

Since $A = S_k \times \mathbb{Z}_2$ we have that $|A| = |S_k| \cdot |\mathbb{Z}_2| = 2k!$. We will now apply Burnside's Lemma and get that

$$|\mathcal{C}| = \frac{1}{2k!} \left(\underbrace{\sum_{\sigma \in S_k} 2^{n(e(\sigma)+o(\sigma))}}_{(I)} + \underbrace{\sum_{\sigma \in S_k} 2^{ne(\sigma)+\lfloor (n+1)/2 \rfloor o(\sigma)}}_{(II)} \right)$$

where the first sum is all elements of π of the form $\sigma \times a$ where a is the identity and the second sum is all elements of π of the from $\sigma \times a$ where a is not the identity. Combining these sums gives us our first equality in the conclusion of the theorem.

In (I) note that $e(\sigma) + o(\sigma)$ is the number of cycles in σ. If we group the permutations of S_k according to how many cycles the permutations have then we will have k groups (for $1, 2, \ldots, k$) each with $s(k, i)$ elements, where i is the number of cycles. So we have that

$$\sum_{\sigma \in S_k} 2^{n(e(\sigma)+o(\sigma))} = \sum_{i=1}^{k} s(k, i) 2^{ni}.$$

Now applying Theorem 3 with $x = 2^n$ we have

$$\sum_{i=1}^{k} s(k, i) 2^{ni} = \prod_{i=0}^{k-1} (2^n + i).$$

Putting this in for (I) we have

$$|\mathcal{C}| = \frac{1}{2k!} \left(\prod_{i=0}^{k-1} (2^n + i) + \sum_{\sigma \in S_k} 2^{ne(\sigma)+\lfloor (n+1)/2 \rfloor o(\sigma)} \right)$$

which gives us our second equality in the conclusion of the theorem.

For (II) note that when $\sigma \in S_k$ that $2e(\sigma) + o(\sigma) \le k$ and so

$$ne(\sigma) + \lfloor (n+1)/2 \rfloor o(\sigma) \le \lfloor (n+1)/2 \rfloor (2e(\sigma) + o(\sigma)) \le \lfloor (n+1)/2 \rfloor k.$$

As an immediate consequence we have

$$\sum_{\sigma \in S_k} 2^{ne(\sigma)+\lfloor (n+1)/2 \rfloor o(\sigma)} \le \sum_{\sigma \in S_k} 2^{\lfloor (n+1)/2 \rfloor k} = k! 2^{\lfloor (n+1)/2 \rfloor k}$$

Putting this in for (II), along with what we have already done for (I), we have

$$|\mathcal{C}| \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2\rfloor k}\right)$$

which gives us our final inequality for the theorem. $\square$

Although we have two exact formulas for $|\mathcal{C}|$ it is most convenient to use the bound

$$|\mathcal{C}| \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2\rfloor k}\right).$$

While this bound is not sharp it can be calculated without knowing anything about the elements of S_k.

In addition, asymptotically this bound behaves like the exact formulas. This is because the first part of the bound (the $\left(\prod_{i=0}^{k-1}(2^n+i)\right)/(2k!)$) is exact and as we shall see later behaves like $2^{nk}/(2k!)$ as n gets large while the second term is not sharp and behaves like $2^{\lfloor (n+1)/2\rfloor k}/2$ which comparatively becomes insignificant as n gets large.

6 Upper bound

With our bound for the size of $\mathcal{C}$ in place we can now give an upper bound for $\mathcal{P}(n,k)$. This is done by bounding the number of P_n coloring representations that are possible for graphs on $n+k$ vertices as shown in the following theorem.

Theorem 5. *We have*

$$\mathcal{P}(n,k) \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2\rfloor k}\right)2^{\binom{k}{2}}.$$

Proof. Let G be any graph on $n+k$ vertices which contains P_n as an induced subgraph. Then by the discussion that followed Lemma 1 there is a P_n coloring representation of G that uses one of the colorings of $\mathcal{C}$ along with some labeled graph on k vertices.

In particular, all graphs on $n+k$ vertices which contain P_n as an induced subgraph are represented at least once (possibly several times) in the combinations of all $k \times n$ arrays which are colored with a coloring of $\mathcal{C}$ and all graphs on k labeled vertices. Since each one of these combinations corresponds to only one graph then the number of graphs on $n+k$ vertices containing P_n as an induced subgraph is at most the number of these combinations.

There are a total of $2^{\binom{k}{2}}$ graphs on k labeled vertices and so there are a total of $|\mathcal{C}|2^{\binom{k}{2}}$ combinations of colorings of the array with a coloring of $\mathcal{C}'$ with labeled graphs. Using Theorem 4 we have,

$$\mathcal{P}(n,k) \leq |\mathcal{C}|2^{\binom{k}{2}} \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2 \rfloor k}\right)2^{\binom{k}{2}}.$$

□

7 Lower bound

To get our upper bound we created a combination of colorings of the array and labeled graphs where each graph on $n+k$ vertices which contained P_n as an induced subgraph was represented *at least once*. To get our lower bound we will use a similar approach in that we will create a combination of colorings and labeled graphs where each graph on $n+k$ vertices which contains P_n as an induced subgraph will be represented *at most once*.

In order to obtain a lower bound we will show that by making restrictions on our colorings that we can overcome two problems that were inherent in our upper bounds. The first problem is having P_n as an induced subgraph in multiple ways. The second problem is the colorings of our array which are invariant under multiple automorphisms.

7.1 Having a long induced path in only one way

In Figure 9 we saw that one graph can have P_n as an induced subgraph in multiple ways. In looking at all combinations of colorings of $\mathcal{C}$ and labeled graphs this caused some graphs to be counted multiple times. To overcome this problem we will use the following lemmas.

Lemma 2. Let $G=(V,E)$ and let $N \subseteq V$ induce P_n as a subgraph then for all $v \in N$ we have $\deg(v) \leq |V| - n + 2$.

Proof. Since the maximum degree of a vertex in P_n is 2, if $v \in N$ then v can be adjacent to at most 2 other vertices of N. In particular v is not adjacent to $n-2$ of the vertices of V lying in N. Thus, the maximum degree that v can have is $|V| - (n-2)$. □

Lemma 3. If the $k \times n$ array in a P_n coloring representation of a graph has at least $k+3$ entries in each row colored 'yes' then the graph contains P_n as an induced subgraph in exactly one way.

Proof. Each row corresponds to a vertex in the graph, and by our assumption each vertex corresponding to a row has degree at least $k+3$. By Lemma 2, with $|V| = n+k$, it follow that none of the k vertices that correspond to the rows can lie in an induced subgraph which is P_n. Thus only n of the $n+k$ vertices can lie in an induced subgraph which is P_n and so we have no more than one way to have P_n as an induced subgraph.

Since the columns correspond to a P_n which is an induced subgraph we have at least one way that the graph contains P_n as an induced subgraph.

Combining these two we have exactly one way that the graph contains P_n as an induced subgraph. □

If we add the restriction that our colorings have at least $k+3$ or more elements colored 'yes' then from Lemma 3 all of the corresponding graphs with such colorings can have P_n as an induced subgraph in only one way. In order for the restriction to be useful we will get a bound on the number of graphs that do not satisfy this restriction. This will be done with the following lemma.

Lemma 4. Let $\mathcal{D}$ be a maximal collection of inequivalent colorings with two colors of the $k \times n$ array under the action of A, such that each coloring of $\mathcal{D}$ contains at least one row with $k+2$ or fewer elements colored 'yes.' Then

$$|\mathcal{D}| \leq \left(\sum_{i=0}^{k+2} \binom{n}{i}\right) \left(\frac{1}{2(k-1)!} \left(\prod_{i=0}^{k-2} (2^n + i) + (k-1)! 2^{\lfloor (n+1)/2 \rfloor (k-1)}\right)\right).$$

Proof. Let $\mathcal{E}$ denote a maximal collection of inequivalent colorings of the $(k-1) \times n$ array under the action of the group $A' = S_{k-1} \times \mathbb{Z}_2$. Now consider the collection of $k \times n$ arrays where the first row contains $k+2$ or fewer elements colored 'yes' and the remaining $k-1$ rows correspond to a coloring of $\mathcal{E}$.

We claim that every coloring of $\mathcal{D}$ is equivalent with at least one coloring in this collection. To see this start with $d \in \mathcal{D}$. Then d is equivalent to a coloring of the $k \times n$ array, call it d', where the first row has $k+2$ or fewer elements colored 'yes' by an action which permutes the first row with any row which contains $k+2$ or fewer elements colored 'yes.'

The resulting rows 2 through k are a coloring of the $(k-1) \times n$ array and in particular there is an automorphism in A' which acts on these $k-1$ rows which takes it to a coloring of $\mathcal{E}$. This automorphism in A' can be extended to an automorphism in A that acts on d' where the first row does not change position (but possibly might reverse order) and the remaining $(k-1)$ rows becomes one of the colorings of $\mathcal{E}$.

In particular we have that $d \in \mathcal{D}$ is equivalent to a coloring d' which in turn is equivalent to a coloring which has $k+2$ or fewer elements colored 'yes' in the first row and the remaining $k-1$ rows corresponds to a coloring of $\mathcal{E}$.

So we can bound $|\mathcal{D}|$ by bounding the number of colorings in this collection. First note that there are a total of

$$\sum_{i=0}^{k+2} \binom{n}{i}$$

different ways that the first row can have $k+2$ or fewer elements colored yes.

From Theorem 4 we have that

$$|\mathcal{E}| \leq \frac{1}{2(k-1)!} \left(\prod_{i=0}^{k-2} (2^n + i) + (k-1)! 2^{\lfloor (n+1)/2 \rfloor (k-1)} \right).$$

Combining these we have that

$$|\mathcal{D}| \leq \left(\sum_{i=0}^{k+2} \binom{n}{i} \right) \left(\frac{1}{2(k-1)!} \left(\prod_{i=0}^{k-2} (2^n + i) + (k-1)! 2^{\lfloor (n+1)/2 \rfloor (k-1)} \right) \right).$$

□

7.2 Colorings fixed only under the identity automorphism

A second source for overcounting when we did our upper bound arose because some colorings of $\mathcal{C}$ are invariant under multiple actions. This can cause a graph to have several distinct labeled graphs associated with a fixed coloring of the array. An example of this situation is shown in Figure 10.

This occurs because the different actions for which the graph is invariant corresponds to different sets of choices in how we make our P_n coloring

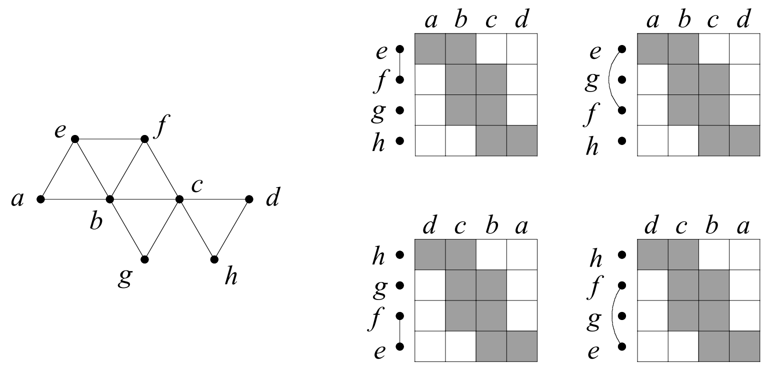

Figure 10: Representations with a fixed coloring and multiple labeled graphs

representation. In particular, we can have the same coloring of the array but have changed the labeling of the graph on k vertices.

If we add the restriction that our colorings remain invariant only under the identity automorphism then we can overcome this problem. In order for this restriction to be useful we need to get a bound on the number of colorings which satisfy this restriction. This will be done by breaking the problem into two pieces and then examining each piece in turn as done in the following lemmas.

Lemma 5. Let $\mathcal{D}$ be a maximal collection of inequivalent colorings with two colors of the $k \times n$ array under the action of A such that for each coloring of $\mathcal{D}$ there is a non-identity element $\pi \in A$ where $\pi = \sigma \times a$ and a is the identity (so we do not reverse the order of the columns) and for which the coloring is invariant under the action by π. Then

$$|\mathcal{D}| \leq (k-1)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i)+(k-1)!2^{\lfloor (n+1)/2\rfloor (k-1)}\right)\right).$$

Proof. Let $\mathcal{E}$ denote a maximal collection of inequivalent colorings of the $(k-1)\times n$ array under the action of the group $A' = S_{k-1}\times \mathbb{Z}_2$. Now start with any coloring of $\mathcal{E}$ and put the coloring into rows 2 through k of a $k \times n$ array. For row 1 put a duplicate of each of the rows 2 through k in turn, making a total of $k-1$ different colorings (with some possible repetitions) of the $k \times n$ array. An example of this construction is shown in Figure 11.

For every coloring of $\mathcal{E}$ we have now constructed $k-1$ colorings of the $k \times n$ array, and so we have constructed at most $(k-1)|\mathcal{E}|$ colorings. By application of Theorem 4 we have that the number of colorings that we have constructed is

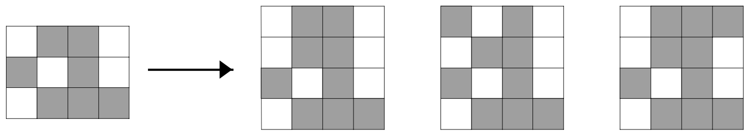

Figure 11: An example of the construction in Lemma 5

bounded above by

$$(k-1)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i)+(k-1)!2^{\lfloor (n+1)/2\rfloor (k-1)}\right)\right).$$

All that remains in the proof is to show that each coloring of $\mathcal{D}$ is equivalent to at least one of the colorings that we have constructed.

First note that since we do not reverse the order of the columns that the only way a coloring can remain invariant under the action of a non-identity element of A is for there to be a duplicate row in the array.

Let $d \in \mathcal{D}$ be a coloring. Then d must contain two rows which are duplicated and in particular we have that d is equivalent to a coloring, d', in which the first row is a duplicate of some other row by applying any action of A which takes one of the duplicate rows to the first row.

Rows 2 through k are a coloring of the $(k-1)\times n$ array and in particular there is an action in A' which acts on these $k-1$ rows which takes it to a coloring of $\mathcal{E}$. This action in A' can be extended to an action in A that acts on d' where the first row does not change position (but possibly might reverse order) and the remaining $k-1$ rows becomes one of the colorings of $\mathcal{E}$.

In particular we have that $d \in \mathcal{D}$ is equivalent to a coloring d' which in turn is equivalent to a coloring where rows 2 through k are a coloring in $\mathcal{E}$ and the first row is a duplicate of one of the rows of 2 through k. This completes the proof. □

Lemma 6. Let $\mathcal{D}$ be a maximal collection of inequivalent colorings of the $k \times n$ array with two colors under the action of the group A such that for each coloring of $\mathcal{D}$ there is an element $\pi \in A$ so that $\pi = \sigma \cdot a$ where a is not the identity (so the effect of the action reverses the order of the columns) so that the coloring is invariant under π. Then

$$|\mathcal{D}| \leq \left\lfloor \frac{k+2}{2} \right\rfloor 2^{\lfloor (n+1)/2\rfloor k}.$$

Proof. Let $d \in \mathcal{D}$ and consider any row of d. If the row of d is not symmetric then in order for the coloring to remain invariant under an automorphism which reverses the order of the columns there must be some row which has the coloring in reverse order. In particular, the rows of d are either symmetric or they can be placed in pairs which are the reverse order of each other.

Suppose that we have j rows that are paired together. Then the remaining $k-2j$ rows must be symmetric. For every pair of rows we get to make a full choice for one row and the other row will have its coloring determined, so we get a total of nj choices. For the symmetric rows we get to color half the row and then the other half must be colored in reverse order to be symmetric and so we get $(k-2j)\lfloor(n+1)/2\rfloor$ choices.

The number of pairs that we can have is between 0 and $\lfloor k/2\rfloor$, so we get

$$|\mathcal{D}| \leq \sum_{j=0}^{\lfloor(k/2)\rfloor} 2^{nj+(k-2j)\lfloor(n+1)/2\rfloor}.$$

Now note that $n \leq 2\lfloor(n+1)/2\rfloor$ so we have

$$\begin{aligned}\sum_{j=0}^{\lfloor k/2\rfloor} 2^{nj+(k-2j)\lfloor(n+1)/2\rfloor} &\leq \sum_{j=0}^{\lfloor k/2\rfloor} 2^{2j\lfloor(n+1)/2\rfloor+(k-2j)\lfloor(n+1)/2\rfloor} = \sum_{j=0}^{\lfloor k/2\rfloor} 2^{\lfloor(n+1)/2\rfloor k} \\ &= \left(\left\lfloor\frac{k}{2}\right\rfloor+1\right)2^{\lfloor(n+1)/2\rfloor k} = \left\lfloor\frac{k+2}{2}\right\rfloor 2^{\lfloor(n+1)/2\rfloor k}.\end{aligned}$$

Any coloring that is invariant under an action which reverses the columns will be equivalent to one of these, i.e., we permute the rows to put the pairs in order at the top of the array and the symmetric rows at the bottom. This concludes the proof. □

7.3 The lower bound

The lemmas will provide a way to create a large number of P_n coloring representations on $n+k$ vertices that correspond to non-isomorphic graphs. This will establish the lower bound for $\mathcal{P}(n,k)$ given by the next theorem.

Theorem 6. *Let*

$$\begin{aligned} S &= \left(\sum_{i=0}^{k+2}\binom{n}{i}\right)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i)+(k-1)!2^{\lfloor(n+1)/2\rfloor(k-1)}\right)\right) \\ T &= (k-1)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i)+(k-1)!2^{\lfloor(n+1)/2\rfloor(k-1)}\right)\right) \\ U &= \left\lfloor\frac{k+2}{2}\right\rfloor 2^{\lfloor(n+1)/2\rfloor k}. \end{aligned}$$

Then we have

$$\left(\frac{1}{2k!}\prod_{i=0}^{k-1}(2^n+i) - S - T - U\right)2^{\binom{k}{2}} \leq \mathcal{P}(n,k).$$

Proof. Let $\mathcal{D}$ be a maximal collection of inequivalent colorings of the $k \times n$ array where every coloring satisfies the following conditions:

(i) every row of the coloring has at least $k+3$ entries colored 'yes,'

(ii) the only action in A for which the coloring is invariant is the action induced by the identity element of A.

Then consider all combinations of colorings of the array with colorings of $\mathcal{D}$ along with all graphs on k labeled vertices.

We claim that any graph on $n+k$ vertices which contains P_n as an induced subgraph is represented at most once in this combination. To see this let G be a graph on $n+k$ vertices. Then if G has no P_n coloring representation that has the array colored with a coloring of $\mathcal{D}$ then it cannot show up in the combination.

So suppose that the graph does have a P_n coloring representation of G with a coloring of the array colored with a coloring of $\mathcal{D}$. Then by (i) and Lemma 3 we know that the graph contains P_n as an induced subgraph in exactly one way, so there is only one coloring of $\mathcal{D}$ which corresponds to a P_n coloring representations of the graph.

Now note that because of (ii) there is only one possible way of assigning the vertices to the columns and the rows so that the coloring matches with the coloring of $\mathcal{D}$. If there were two distinct ways of assigning the vertices to the columns and the rows then we could form an automorphism that is not the identity for which the coloring would be invariant, which is impossible by assumption. It follows that for G there is only one labeled graph that we can associate with the coloring of $\mathcal{D}$. In particular, G can only show up once in the combination of the colorings of $\mathcal{D}$ and all of the labeled graphs.

Since each graph on $n+k$ vertices which contains P_n as an induced subgraph is represented at most once in these combinations, and each one of these combinations corresponds to a graph on $n+k$ vertices which contains P_n as an induced subgraph, we can conclude that

$$|\mathcal{D}|2^{\binom{k}{2}} \leq \mathcal{P}(n,k)$$

All that remains is to bound $|\mathcal{D}|$.

To bound $|\mathcal{D}|$ we will start with a lower bound for the number of inequivalent colorings of the $k \times n$ array and then get an upper bound for the number of colorings that we need to throw out so that the colorings that remain will satisfy conditions (i) and (ii).

Our lower bound for the number of inequivalent colorings of the $k \times n$ array comes from Theorem 4. Letting $|\mathcal{C}|$ denote the number of inequivalent colorings of the $k \times n$ array we have that

$$\frac{1}{2k!}\prod_{i=0}^{k-1}(2^n+i) \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + \sum_{\sigma \in S_k} 2^{ne(\sigma)+\lfloor (n+1)/2 \rfloor o(\sigma)}\right) = |\mathcal{C}|.$$

Our upper bound for the number of colorings that we need to throw out will come by combining Lemmas 4, 5 and 6. By Lemma 4, the number of graphs which do not satisfy (i) is bounded above by S. By Lemmas 5 and 6, the number of graphs which do not satisfy (ii) is bounded above by $T+U$. Combining these we have that the number of graphs which do not satisfy (i) or (ii) is bounded above by $S+T+U$.

Subtracting our upper bound for the number of colorings that we need to throw out from our lower bound for the number of inequivalent colorings we can conclude that

$$\frac{1}{2k!}\prod_{i=0}^{k-1}(2^n+i) - S - T - U \leq |\mathcal{D}|.$$

This concludes the proof. □

8 Asymptotic behavior

Comparing the upper and lower bound we see that as we fix k and let n get large the term which will dominate in both bounds is the same. In particular we have the following theorem.

Theorem 7. *Let k be fixed. Then*

$$\lim_{n\to\infty} \frac{\mathcal{P}(n,k)}{2^{nk}} = \frac{2^{\binom{k}{2}}}{2k!}.$$

Proof. By Theorem 5 we have

$$\mathcal{P}(n,k) \leq \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(2^n+i) + k!2^{\lfloor (n+1)/2 \rfloor k}\right)2^{\binom{k}{2}}.$$

Dividing both sides through by 2^{nk} and simplifying we have

$$\frac{\mathcal{P}(n,k)}{2^{nk}} \le \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(1+i2^{-n}) + k!2^{(\lfloor (n+1)/2\rfloor - n)k}\right)2^{\binom{k}{2}}.$$

Taking the limit as n goes to infinity on both sides we have

$$\limsup_{n\to\infty}\frac{\mathcal{P}(n,k)}{2^{nk}} \le \frac{1}{2k!}\left(\prod_{i=0}^{k-1}(1+0) + k!(0)\right)2^{\binom{k}{2}} = \frac{2^{\binom{k}{2}}}{2k!}.$$

By Theorem 6 if we let

$$\begin{aligned}
S &= \left(\sum_{i=0}^{k+2}\binom{n}{i}\right)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i) + (k-1)!2^{\lfloor (n+1)/2\rfloor (k-1)}\right)\right)\\
T &= (k-1)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(2^n+i) + (k-1)!2^{\lfloor (n+1)/2\rfloor (k-1)}\right)\right)\\
U &= \left\lfloor \frac{k+2}{2}\right\rfloor 2^{\lfloor (n+1)/2\rfloor k},
\end{aligned}$$

then we have

$$\mathcal{P}(n,k) \ge \left(\frac{1}{2k!}\prod_{i=0}^{k-1}(2^n+i) - S - T - U\right)2^{\binom{k}{2}}.$$

Dividing both sides by 2^{nk} and simplifying we have

$$\frac{\mathcal{P}(n,k)}{2^{nk}} \ge \left(\frac{1}{2k!}\prod_{i=0}^{k-1}(1+i2^{-n}) - \frac{S}{2^{nk}} - \frac{T}{2^{nk}} - \frac{U}{2^{nk}}\right)2^{\binom{k}{2}}.$$

Consider $S/2^{nk}$. Simplifying we have that

$$\frac{S}{2^{nk}} = \left(\frac{\sum_{i=0}^{k+2}\binom{n}{i}}{2^n}\right)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(1+i2^{-n}) + (k-1)!2^{(\lfloor (n+1)/2\rfloor - n)(k-1)}\right)\right).$$

Now note that the second term is similar to what we had before and will approach $1/(2(k-1)!)$. The top of the first term (the $\sum_{i=0}^{k+2}\binom{n}{i}$) corresponds to a polynomial of n with degree $k+2$ while the bottom of the first term (the 2^n) is an exponential. So as n gets large the bottom will dominate and drive the first term to zero. In particular we have that $S/2^{nk}$ goes to zero as n goes to infinity.

Now consider $T/2^{nk}$. Simplifying we have that

$$\frac{T}{2^{nk}} = \left(\frac{k-1}{2^n}\right)\left(\frac{1}{2(k-1)!}\left(\prod_{i=0}^{k-2}(1+i2^{-n}) + (k-1)!2^{(\lfloor (n+1)/2\rfloor - n)(k-1)}\right)\right).$$

As before the second terms will approach $1/(2(k-1)!)$. For the first term (the $(k-1)/2^n$) we know that the bottom will dominate and will drive the term to zero. In particular we have that $T/2^{nk}$ goes to zero as n goes to infinity.

Now consider $U/2^{nk}$. By the definition of U we have that

$$\frac{U}{2^{nk}} = \left\lfloor \frac{k+2}{2} \right\rfloor 2^{(\lfloor (n+1)/2 \rfloor - n)k}.$$

As n gets large the exponent becomes a large negative value and drives the term to zero. In particular we have that $U/2^{nk}$ goes to zero as n goes to infinity.

So we have

$$\liminf_{n\to\infty} \frac{\mathcal{P}(n,k)}{2^{nk}} \geq \lim_{n\to\infty} \left(\frac{1}{2k!} \prod_{i=0}^{k-1} (1 + i2^{-n}) - \frac{S}{2^{nk}} - \frac{T}{2^{nk}} - \frac{U}{2^{nk}} \right) 2^{\binom{k}{2}} = \frac{2^{\binom{k}{2}}}{2k!}.$$

Combining the two results concludes the proof. □

As an immediate consequence of the theorem we have that if k is fixed that

$$\mathcal{P}(n,k) \sim \frac{2^{(nk + \binom{k}{2})}}{2k!}.$$

Conclusion

By using the techniques of Polya counting on a $k \times n$ array with two colors we have been able to find upper and lower bounds for the number of non-isomorphic graphs on $n+k$ vertices which contain P_n as an induced subgraph. Further, these upper and lower bounds both grow at the same rate as n gets large for a fixed value of k.

We noted earlier that finding $\mathcal{P}(19,4)$ would be difficult to do by examining all of the graphs on 23 vertices. With the techniques of this thesis, we can now say that $(6.70802)10^{22} \leq \mathcal{P}(19,4) \leq (1.0075)10^{23}$. In this example we can also see that graphs which contain very long paths as induced subgraphs are rare. Comparing the upper bound that we have for $\mathcal{P}(19,4)$ with the number of graphs on 23 vertices (given earlier) we see that for every graph with contains P_{19} as an induced subgraph there are at least $(5.558)10^{30}$ graphs which don't.

References

[1] G. Bacsó and Z. Tuza, A characterization of graphs without long induced paths, Journal of Graph Theory 14 (1990) no.4, 455-464.

[2] B. Bollobás, Modern Graph Theory, Springer, 1998.

[3] R. Brualdi, Introductory Combinatorics (3rd ed), Prentice Hall, 1999.

[4] J. Dong, Some results on graphs without long induced paths, Journal of Graph Theory 22 (1996) no.1, 23-28.

[5] D. Dummit, R. Foote, Abstract Algebra (2nd ed), Wiley, 1999.

[6] F. Harary and E. Palmer, Graphical Enumeration, Academic Press, 1973.

[7] K. Harris, J. Hirst and M. Mossinghoff, Combinatorics and Graph Theory, Springer, 2000.

[8] G. Woeginger, J. Sgall, The complexity of coloring graphs without long induced paths, Acta Cybernetica 15 (2001) no. 1, 107-117.