Signed Sandra Li

Shushangxuan Li

Date 5/9/2016.

This Honors paper by Shushangxuan Li has been read and approved for Honors in the Department of Mathematics and Computer Science.

Dr. Eric Gottlieb
Project Advisor

Dr. Chris Seaton
Second Reader

Dr. Shubho Banerjee
Extra-Departmental Reader

Dr. Michael Sheard
Department Chair

Automorphism Groups of k-star n-path Saturated Graphs

Shushangxuan Li

Department of Mathematics and Computer Science
Rhodes College
Memphis, Tennessee

2016

Submitted in partial fulfillment of the requirements for the
Bachelor of Science degree with Honors in Mathematics

Contents

List of Figures

Abstract

In combinatorics, graph theory and Ramsey theory are two rich subjects of study. There are many directions and problems to pursue. The automorphism groups of graphs provide a connection between graph theory and group theory. In this project, we will study a Ramsey-theoretic aspect of graph theory. The main question is: how many edges must a tree contain in order to guarantee the presence of a k-star or an n-path? Another way to think about this problem is: what is the maximum number of edges that a tree with no k-star or n-path can contain? We will establish a general formula for the number of edges using induction and classify the trees that achieve this maximum. In addition, we will describe the automorphism group of these saturated trees. The same questions will be studied for connected graphs following a similar procedure. We will show that the number of edges for connected graphs has a close relation with the one for trees.

1 Background

1.1 Graph Theory

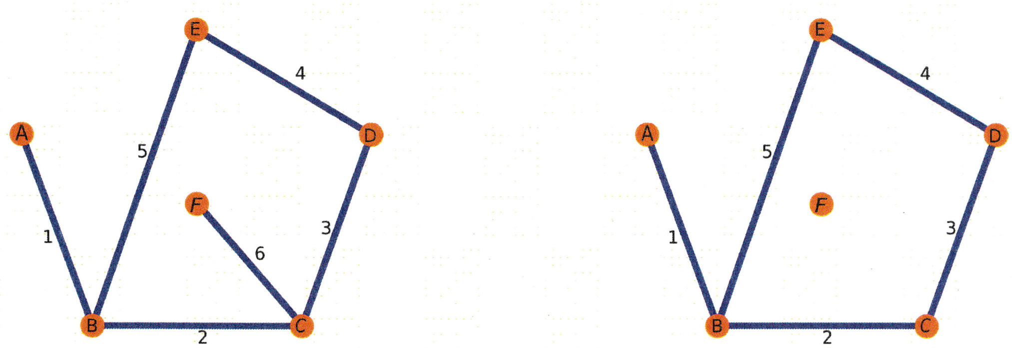
We will start by reviewing the fundamentals of graph theory. A *graph* G is a set of points, called *vertices*, $V(G)$, and a set of lines joining some pairs of distinct vertices, called *edges*, $E(G)$.

Some authors allow multiple edges joining the same pair of points, or edges that start and end in the same vertex, i.e. *loops*. [1] We do not allow multiple edges or loops in our graphs for this project.

Definition 1.1. *A sequence of distinct edges* $e_1 e_2 \ldots e_k$ *is called a trail if the endpoint of* e_i *is the starting point of* e_{i+1}*. In addition, a trail is called a closed trail if the end point of* e_k *is the starting point of* e_1*. A cycle is a closed trail that does not touch any vertex twice except the initial vertex, which must be the ending vertex. If a trail does not touch any vertex twice, then we call it a path. The number of edges connected to a vertex is called the degree of the vertex.* [1]

A path with n edges is denoted P_n.

In figure 1a, $e_2 e_3 e_4 e_5 e_1$ is a trail, $e_2 e_3 e_4 e_5$ is a cycle, $e_2 e_3 e_4$ is a path of length 3.



(a) connected graph (b) disconnected graph

Figure 1: Example graphs

Definition 1.2. *If the graph G has the property that for any two vertices x and y, one can find a path from x to y, then we say that G is a connected graph.* [1]

In Figure 1, figure 1a is a connected graph. Remove e_6 from it, we have figure 1b, which is not a connected graph. There is no path from F to the other vertices.

Definition 1.3. *A connected graph G is a tree if G does not contain any cycles.*

Theorem 1.1. *The number of edges in a tree is one less than the number of vertices.*

While $E(G), V(G)$ denote the set of edges, vertices of G, when no confusion will arise, we also use this to denote the number of edges, vertices of G.

For the tree T in Figure 2, there are 15 vertices and 14 edges, so $E(T) = V(T) - 1$.

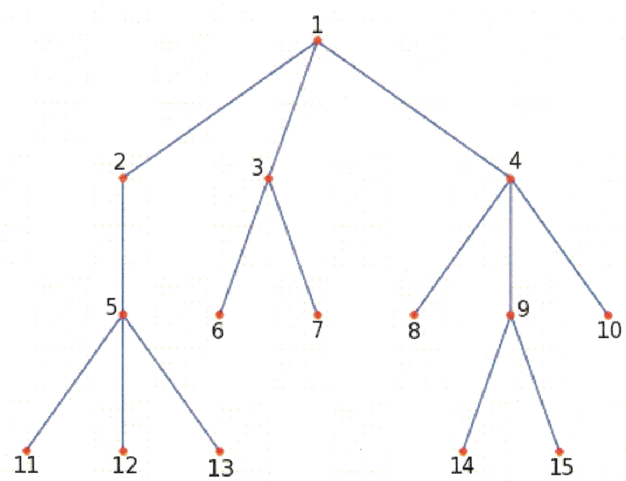


Figure 2: tree

Definition 1.4. *A star is a tree in which no path length is greater than 2.*

Theorem 1.2. *In a star, only one vertex can have degree greater than 1.*

We denote a star with k edges by S_k.

In Figure 2, vertex 5 together with vertices 2, 11, 12, 13 and the four edges form a 4-star, S_4. Vertex 3 together with vertices 1, 6, 7 and the three edges form a

3-star, S_3.

Definition 1.5. *Vertices of trees that have degree one are called leaves. Vertices that are not leaves are called internal nodes.*

In Figure 2, vertices 6, 7, 8, 10, 11, 12, 13, 14, 15 are leaves, while vertices 1, 2, 3, 4, 5, 9 are internal nodes.

Theorem 1.3. *(Handshake Lemma)* [1] *In a graph G without loops, the number of vertices of odd degree is even.*

This is easily seen by observing that $\sum_{v \in V} \deg(v) = 2|E|$.

In Figure 2, the sum of degree of the vertices is $28 = 2 * 14$.

1.2 Group Theory

Definition 1.6. *A group G is an ordered pair $(S, *)$ consisting of a nonempty set S and a binary operation $*$ on S, subject to the following three laws:* [2]

1. *For all $x, y, z \in S$, $x * (y * z) = (x * y) * z$. (associativity)*
2. *S contains an element e with the property that, for all $x \in S$, $x*e = e*x = x$. The element e is called the identity element of G.*
3. *For every $x \in S$ there exists an $x^{-1} \in S$ with the property that $x * x^{-1} = x^{-1} * x = e$. The element x^{-1} is called the inverse of x.*

Definition 1.7. *The symmetric group S_n of order $n!$ consists of the set of bijections from $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$.*

The group operation $*$ is composition of bijections. The elements of S_n are written in cycle notation; see Ramsey Theory [3] for details.

Definition 1.8. *An automorphism of a graph G is a bijection ϕ from $V(G)$ to itself in which preserves edge–vertex connectivity, such that (x, y) is an edge of G if and only if $(\phi(x), \phi(y))$ is in $E(G)$.*

For example, the automorphism group of graph G in Figure 3 has six permuta-

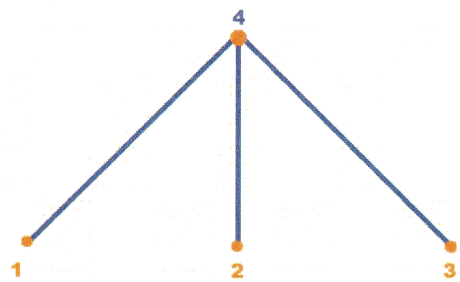


Figure 3: S_3

tions, $\mathrm{Aut}(G) = \{(12), (13), (23), (123), (132), e\}$. However, (134) is not in the automorphism group because $(1, 4)$ is an edge of G, but $(\pi(2), \pi(4)) = (2, 1)$ is not an edge of G.

With the understanding of necessary graph theory and group theory knowledge, we will present the proofs of the famous theorems mentioned in Chapter 1.

2 Proofs of some related theorems

We highlight the following theorems that have close relations to our project.

2.1 Erdös-Szekeres Theorem, 1935 [2]

We will introduce what a sequence is and a version of the pigeonhole principle before stating the theorem:

Definition 2.1. *A finite sequence* $S = \{a_1, \ldots, a_n\}$ *of real numbers is a function* $a : \{1, \ldots, n\} \to \mathbb{R}$, *for some* $n \in \mathbb{N}$, *defined by* $a(k) = a_k$ *for all* $k \in \mathbb{N}_n$. *A sequence* $\{b_1, \ldots, b_m\}$ *is a subsequence of* $\{a_1, \ldots, a_n\}$ *if there exists a one-to-one function* $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ *for which* $f(i) < f(j)$ *whenever* $i < j$, *and* $b(i) = a(f(i))$ *for all* $i \in \mathbb{N}_m$. *A sequence is monotonically increasing if* $a_i \leq a_j$ *for all* $i < j$, *monotonically decreasing if* $a_i \geq a_j$ *for all* $i < j$.

Theorem 2.1 (Pigeonhole Principle). *If* $f : A \to B$ *is a function, and* $|A| = |B| + 1$, *then there exists* $b \in B$ *with* $|f^{-1}(b)| \geq 2$. *In other words,* $f(a_1) = f(a_2)$ *for some distinct* $a_1, a_2 \in A$.

Now, use the pigeonhole principle, we will present and prove the theorem.

Theorem 2.2 (Erdös-Szekeres Theorem, 1935). *Let* $m, n \in \mathbb{N}$, *any sequence* $S = \{a_1, \ldots, a_{mn+1}\}$ *of* $mn + 1$ *real numbers contains a monotonically increasing subsequence of* $m + 1$ *terms or a monotonically decreasing subsequence of* $n + 1$ *terms (or both), and* $mn + 1$ *is the smallest value with this property.*

Proof. Assume that S does not contain a subsequence of either type. We define a function $f : \mathbb{N}_{mn+1} \to \mathbb{N}_m \times \mathbb{N}_n$ by setting $f(t) = (x_t, y_t)$, where x_t is the greatest number of terms in a monotonically increasing subsequence of S beginning with a_t. By the pigeonhole principle, there exist $j, k \in \{1, \ldots, mn + 1\}$, $j < k$, with $f(j) = f(k)$. It follows that if $a_j \leq a_k$, then $x_j > x_k$, while if $a_j \geq a_k$, then $y_j > y_k$. Both inequalities contradict $f(j) = f(k)$. Therefore, our original assumption is false, and S does not contain a monotonically increasing subsequence of $m + 1$

terms or a monotonically decreasing subsequence of $n+1$ terms.

The expression $mn+1$ is best possible in the sense that there exists a sequence of mn real numbers which contains neither a monotonically increasing subsequence of $m+1$ terms or a monotonically decreasing subsequence of $n+1$ terms. We form such a sequence by concatenating n sequences of m increasing terms in the following manner. For each $j \in \mathbb{N}_n$, let $S_j = \{a_{1j}, \ldots, a_{mj}\}$ be a monotonically increasing sequence of m real numbers, with $a_{hj} > a_{lk}$ whenever $j < k$. The sequence $S = \{a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{1n}, \ldots, a_{mn}\}$ contains no monotonic subsequence of the desired length. ■

2.2 Dilworth's Lemma [2]

We need to introduce chain, antichain and another version of the pigeonhole principle here:

A partial order $\leq$ on X is a total order if every two elements of X are comparable. If $\leq$ is a partial order on X, and Y is a subset of X in which every two elements are comparable, then Y is a chain. If no two distinct elements of Y are comparable, then Y is an antichain.

Theorem 2.3. *If $f : A \to B$ is a function, with A and B finite nonempty set, then there exists $b \in B$ with $|f^{-1}(b)| \geq |A|/|B|$.*

Theorem 2.4 (Dilworth's Lemma). *In any partial order on a set X of $mn+1$ elements, there exists a chain of length $m+1$ or an antichain of size $n+1$, and $mn+1$ is the smallest value with this property.*

Proof. Suppose there is no chain of size $m+1$. Then we may define a function $f : X \to \{1, \ldots, m\}$ with $f(x)$ equal to the greatest number of elements in a chain with greatest element x. By the pigeonhole principle, some $n+1$ elements of X have the same image under f. By the definition of f, these elements are incomparable; that is, they form an antichain of size $n+1$. ■

2.3 Ramsey's Theorem, 1930 [2]

We need to introduce a third version of the pigeonhole principle before introducing Ramsey's theorem:

Theorem 2.5. *If $f : A \to B$ is a function from a finite nonempty set A to a b-set $B = \{B_1, \ldots, B_b\}$, then the following two statements hold:*

If $|A| = a_1 + \cdots + a_b - b + 1$, then there exists B_i with $|f^{-1}(B_i)| \geq a_i$.

If $|A| = a_1 + \cdots + a_b + b - 1$, then there exists B_i with $|f^{-1}(B_i)| \leq a_i$.

Theorem 2.6 (Ramsey's Theorem, 1930). *Given $a, b \geq 2$, there exists a least integer $R(a, b)$ with the following property: Every green-red coloring of the edges of the complete graph on $R(a, b)$ vertices yields a green K_a or a red K_b. Furthermore, $R(a, b) \leq R(a-1, b) + R(a, b-1)$ for all $a, b \geq 3$.*

Proof. We employ induction on a and b. The basis of the induction consists of the statements $R(a, 2) = a$ and $R(2, b) = b$. These are trivial.

In the first assertion, if we two-color K_a and any edge is red, then we obtain a red K_2, while if no edge is red, then we obtain a green K_a. Thus $R(2, a) \leq a$. Equality follows from the fact that an all green colored K_{a-1} contains neither a green K_a nor a red K_2. The second assertion is proved similarly. Now, assuming the existence of $R(a-1, b)$ and $R(a, b-1)$, we will show that $R(a, b)$ exists.

Let G be the complete graph on $R(a-1, b) + R(a, b-1)$ vertices, and let v be an arbitrary vertex of G. By the pigeonhole principle, at least $R(a-1, b)$ green edges or at least $R(a, b-1)$ red edges emanate from v. Without loss of generality, suppose that v is joined by green edges to a complete subgraph on $R(a-1, b)$ vertices. By definition of $R(a-1, b)$, this subgraph must contain a red K_b or a green K_{a-1}. In the latter case, the green K_{a-1} and v, and all the edges between the two, constitute a green K_a. We have shown that G contains a green K_a or a red K_b. Therefore, $R(a, b)$ exists and satisfies $R(a, b) \leq R(a-1, b) + R(a, b-1)$. ∎

2.4 Van der Waerden's Theorem [3]

Theorem 2.7 (Van der Waerden's Theorem). *For all positive integers k and r, there exists an integer $W(k,r)$ so that, if the set of integers $\{1, 2, \ldots, W(k,r)\}$ is partitioned into r classes, then at least one class contains a k-term arithmetic progression.*

Proof. We define $m+1$ *l-equivalence classes of* $[0,l]^m$. For $0 \leq i \leq m$ the set of $(x_1, \ldots, x_m) \in [0,l]^m$ in which l appears in the i rightmost positions and nowhere else forms an l-equivalence class. The l-equivalence classes are disjoint. They partition a proper subset of $[0,l]^m$; the remaining sequences are not used. ∎

3 Number of Edges in Trees

3.1 Discovering the Lower Bound

Our objective is to answer the following question: *How many edges must a tree contain in order to guarantee the presence of a k-star or an n-path?*

In proving the Erdös-Szekeres Theorem, (i.e. that $mn+1$ is as small as possible), it was necessary to give an example of length mn where there is no increasing subsequence of length $m+1$ or a decreasing subsequence of length $n+1$. Then, with the addition of another real number, one of the two subsequences always exists.

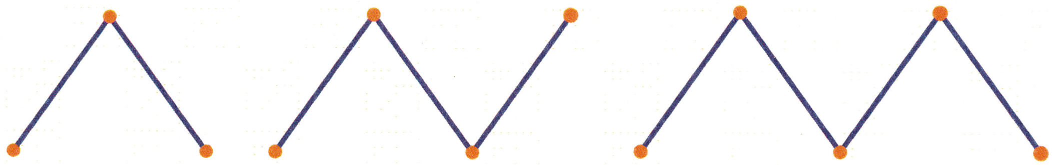
Similarly, for our problem, we will start by looking for the maximum "mn". As with the Erdös-Szekeres Theorem, the answer for minimum will be "$mn+1$".

We approach the problem by looking for the maximum number of edges in a tree with no k-star and no n-path. Our goal is to define a class of trees $T_{k,n}$ with this property. By "no k-star", we mean each vertex is of degree at most $k-1$. By "no n-path", we mean each path has length at most $n-1$. As we have two parameters, we shall fix one of them, k, and increase the value of n. The maximum number of edges a tree with no k-star, n-path can have is $E(T_{k,n})$.

We look for the pattern with small k, n value examples.

When $k = 2$, there can only be one edge, regardless of the value of n. Similarly, when $n = 2$, the maximum path length is one, there can only be one edge, regardless of the value of k. Therefore, $T_{2,n} = T_{k,2} = P_1$, hence, $E(T_{2,n}) = E(T_{k,2}) = 1$. In addition to these easy and obvious cases, we construct trees for $k \geq 3, n \geq 3$.

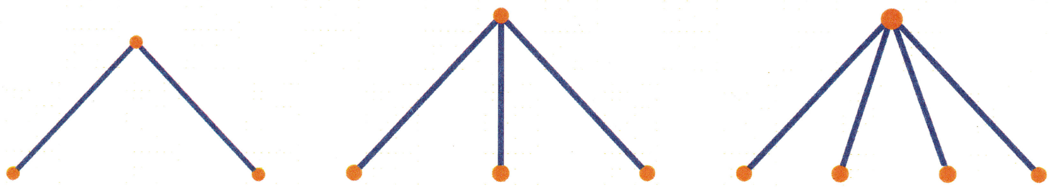
When $k = 3$, we allow S_2 in trees. As we increase n, we can add one edge to one of the two outermost vertices each step, as shown in Figure 4. Therefore, $T_{3,n} = P_{n-1}$, hence, $E(T_{3,n}) = n - 1$.



(a) $T_{3,3}$ (b) $T_{3,4}$ (c) $T_{3,5}$

Figure 4: Tree examples for $k = 3$

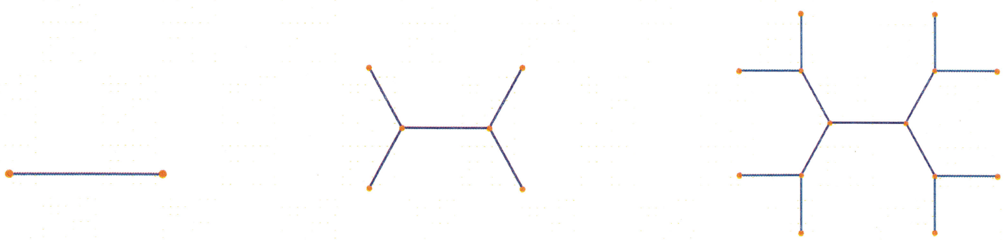
When $n = 3$, a tree without a 3-path is a star, as shown in Figure 5. Therefore, $T_{k,3} = S_{k-1}$, hence, $E(T_{k,3}) = k - 1$.



(a) $T_{3,3}$ (b) $T_{4,3}$ (c) $T_{5,3}$

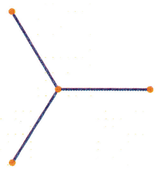
Figure 5: Tree examples for $n = 3$

The above examples are straightforward. Now moving to more complicated trees, we have the following examples when $k > 3, n > 3$:

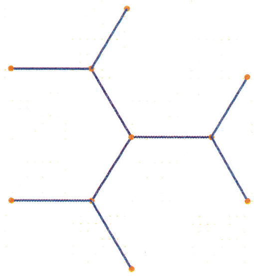


(a) $T_{4,2}$ (b) $T_{4,4}$ (c) $T_{4,6}$

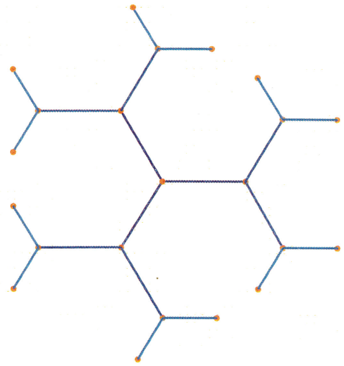
Figure 6: $k = 4$, n is even



(a) $T_{4,3}$



(b) $T_{4,5}$



(c) $T_{4,7}$

Figure 7: $k = 4$, n is odd

A pattern can be seen here:

$$E(T_{4,2}) = 1$$
$$E(T_{4,3}) = 1 + 2$$
$$E(T_{4,4}) = 1 + 2 + 2$$
$$E(T_{4,5}) = 1 + 2 + 2 + 2^2$$
$$E(T_{4,6}) = 1 + 2 + 2 + 2^2 + 2^2$$
$$E(T_{4,7}) = 1 + 2 + 2 + 2^2 + 2^2 + 2^3$$
$$E(T_{4,8}) = 1 + 2 + 2 + 2^2 + 2^2 + 2^3 + 2^3$$
$$E(T_{4,9}) = 1 + 2 + 2 + 2^2 + 2^2 + 2^3 + 2^3 + 2^4$$

For each step, to avoid the existence of a k-star while adding edges to an existing leaf, we should add $k-2$ edges to create $k-1$ star. If we add $k-2$ edges to all leaves, then the path length will increase by 2.

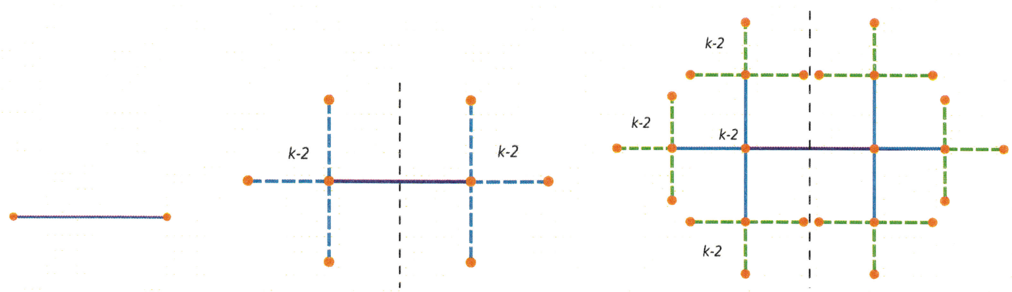
Now we recursively construct the family of $\{T_{k,n}\}_{n\geq 2}$ for fixed $k \geq 3$.

Let $k \geq 3$ be a fixed integer. We treat even and odd values of $n \geq 2$ separately.

For the base case, we already showed that $T_{k,2} = P_1$, and let $T_{k,3} = S_{k-1}$.

Using the recursion relation, we add $k-2$ pendant edges to the leaves of $T_{k,n}$ to obtain $T_{k,n+2}$.

To find the general expression for $E(T_{k,n})$, we need to summarize the recursion relation here, with the help of $T_{5,n}$ in Figure 8.



(a) $T_{5,2}$ (b) $T_{5,4}$ (c) $T_{5,6}$

Figure 8: Recursion relation, n is even

When n is even, obviously, $E(T_{k,2}) = 1$.

Observe that to obtain $T_{k,4}$ from $T_{k,2}$, we add $(k-2)$ edges to each leaf of $T_{k,2}$. As shown in figure 8b, we add 3 edges (dotted blue lines) to each of the 2 leaves in figure 8a. Therefore, we get $E(T_{k,4}) = E(T_{k,2}) + 2(k-2) = 1 + 2(k-2)$.

Similarly, we get $E(T_{k,6}) = E(T_{k,4}) + 2(k-2)^2 = 1 + 2(k-2) + 2(k-2)^2$.

More generally, we find that

$$E(T_{k,n}) = E(T_{k,n-2}) + 2(k-2)^{\frac{n-2}{2}} \tag{1}$$

Theorem 3.1. *When n is even, the number of edges in $T_{k,n}$ is*
$E(T_{k,n}) = \dfrac{2(k-2)^{\frac{n}{2}} - (k-1)}{k-3}$.

Proof. For the base step, when $n = 2$, we have $\dfrac{2(k-2)^{\frac{2}{2}} - (k-1)}{k-3} = 1 = E(T_{k,2})$.

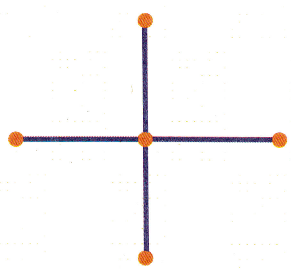
For the induction step, we apply the recursion relation:

$$
\begin{aligned}
E(T_{k,n}) &= E(T_{k,n-2}) + 2(k-2)^{\frac{n-2}{2}} \\
&= \frac{2(k-2)^{\frac{n-2}{2}} - (k-1)}{k-3} + 2(k-2)^{\frac{n-2}{2}} \\
&= \frac{2(k-2)^{\frac{n-2}{2}} - (k-1) + 2(k-2)^{\frac{n-2}{2}}(k-3)}{k-3} \\
&= \frac{2(k-2)^{\frac{n}{2}} - (k-1)}{k-3}
\end{aligned}
$$

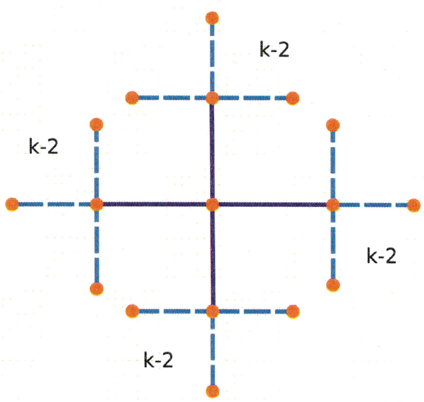
Thus, we conclude that

$$
E_e(T_{k,n}) = \frac{2(k-2)^{\frac{n}{2}} - (k-1)}{k-3} \tag{2}
$$

(e stands for even) ∎



(a) $k = 5, n = 3$



(b) $k = 5, n = 5$

Figure 9: Recursion relation, n is odd

When n is odd, $E(T_{k,3}) = k - 1$.

To obtain $T_{k,5}$ from $T_{k,3}$, we add $k-2$ edges to each leaf $T_{k,3}$. As shown in figure 9b, we add 3 edges (blue dotted lines) to each of the 4 leaves in figure 9a. Therefore, we get $E(T_{k,5}) = E(T_{k,3}) + (k-1)(k-2)$.

Similarly, we get $E(T_{k,7}) = E(T_{k,5}) + (k-1)(k-2)^2$.

More generally, we find that

$$E(T_{k,n}) = E(T_{k,n-2}) + (k-1)(k-2)^{\frac{n-3}{2}}$$

Theorem 3.2. *When n is odd, the number of edges in $T_{k,n}$ is* $E(T_{k,n}) = \dfrac{(k-1)\left[(k-2)^{\frac{n-1}{2}}-1\right]}{k-3}$.

Proof. For the base step, when $n=3$, we have $\dfrac{(k-1)\left[(k-2)^{\frac{3-1}{2}}-1\right]}{k-3} = k-1 = E(T_{k,3})$.

For the induction step, we can apply the recursion relation:

$$\begin{aligned}
E(T_{k,n}) &= E(T_{k,n-2}) + (k-1)(k-2)^{\frac{n-3}{2}} \\
&= \frac{(k-1)\left[(k-2)^{\frac{n-3}{2}}-1\right]}{k-3} + (k-1)(k-2)^{\frac{n-3}{2}} \\
&= \frac{(k-1)\left[(k-2)^{\frac{n-3}{2}}-1\right] + (k-1)(k-2)^{\frac{n-3}{2}}(k-3)}{k-3} \\
&= \frac{(k-1)(k-2)^{\frac{n-3}{2}}(k-2)-(k-1)}{k-3} \\
&= \frac{(k-1)\left[(k-2)^{\frac{n-1}{2}}-1\right]}{k-3}
\end{aligned}$$

Thus, we conclude that

$$E_o(T_{k,n}) = \frac{(k-1)\left[(k-2)^{\frac{n-1}{2}}-1\right]}{k-3}$$

(o stands for odd) ■

We obtain a general formula that works for both even and odd values of n.

Corollary 3.2.1. *The maximum number of edges in a tree with no k-star and no n-path is* $E(T_{k,n}) = \dfrac{(k-1)\left[(k-2)^{\lceil \frac{n}{2}\rceil -1}-1\right]}{k-3} + \dfrac{(1+(-1)^n)}{2}(k-2)^{\frac{n}{2}-1}$.

Let $F(k,n)$ be the minimum number of edges required to guarantee the existence of a k-star or an n-path in a tree. The tree $T_{k,n}$ shows that $F(k,n) \geq E(T_{k,n}) + 1$.

3.2 Proof of the Upper Bound

We have found a lower bound, $E(T_{k,n})$ for $F(k,n)$. We will show that $E(T_{k,n}) + 1$ is an upper bound for $F(k,n)$. In other words, $E(T_{k,n}) + 1$ is sufficient and no more edges are needed.

We approach the proof by induction.

Let T be any tree with no k-star and at least $E(T_{k,n}) + 1$ edges. We will show that T has an n-path. The degree of each vertex is less than k, so $\deg(v) \leq k - 1$ for all $v \in V$.

Let T' be T without its leaves, $E(T')$ be the number of edges tree T' has.

Let $L(T)$ represent the set of leaves of T, and $I(T)$ represent the set of internal nodes of T, so $I(T) = V(T) - L(T)$.

Observe that the maximum length of a path in T is two more than the maximum length path in T'.

Lemma 3.3.

$$\begin{aligned} 2|E| &= \sum_{v \in V} deg\ v \\ &\leq \sum_{v \in L} 1 + \sum_{v \in I} deg\ v \\ &\leq L(T) + (k-1)I(T) \end{aligned}$$

Since $E(T) = V(T) - 1$, *and* $I(T) = V(T') = E(T') + 1$, *we have*

$$\begin{aligned}
2(L(T) + I(T) - 1) &\leq L(T) + (k-1)I(T) \\
(k-1)I(T) &\geq L(T) + I(T) + I(T) - 2 \\
(k-2)I(T) &\geq E(T) + 1 - 2 \\
E(T') + 1 &\geq \frac{E(T) - 1}{k-2} \\
E(T') &\geq \frac{E(T_{k,n})}{k-2} - 1
\end{aligned}$$

We will prove that $E(T_{k,n-2}) + 1$ is also the upper bound for $E(T')$, so we need to prove that $E(T_{k,n-2}) + 1 \geq \dfrac{E(T_{k,n})}{k-2} - 1$.

Proof. By induction, we can apply the lemma when n is odd,

$$\begin{aligned}
E_o(T_{k,n-2}) - \frac{E_o(T_{k,n})}{k-2} + 2 &= \frac{(k-1)\left[(k-2)^{\frac{n-3}{2}} - 1\right]}{k-3} - \frac{(k-1)\left[(k-2)^{\frac{n-1}{2}} - 1\right]}{(k-3)(k-2)} + 2 \\
&= \frac{(k-1)\left[(k-2)^{\frac{n-1}{2}} - (k-2)\right] - (k-1)\left[(k-2)^{\frac{n-1}{2}} - 1\right]}{(k-3)(k-2)} + 2 \\
&= \frac{(k-1)(-k+3)}{(k-3)(k-2)} + 2 \\
&= 2 - \frac{k-1}{k-2} \\
&= \frac{k-3}{k-2} \geq 0
\end{aligned}$$

Thus, $E_o(T_{k,n-2}) + 1 \geq \dfrac{E_o(T_{k,n})}{k-2} - 1$.

Similarly, we can apply the lemma when n is even,

$$\begin{aligned}
E(T_{k,n-2}) - \frac{E(T_{k,n})}{k-2} + 2 &= \frac{2(k-2)^{\frac{n-2}{2}} - (k-1)}{k-3} - \frac{2(k-2)^{\frac{n}{2}} - (k-1)}{(k-2)(k-3)} + 2 \\
&= \frac{2(k-2)^{\frac{n}{2}} - (k-1)(k-2) - 2(k-2)^{\frac{n}{2}} + (k-1)}{(k-2)(k-3)} + 2 \\
&= \frac{(k-1)(-k+3)}{(k-2)(k-3)} + 2 \\
&= \frac{k-3}{k-2} \geq 0
\end{aligned}$$

Thus, $E_e(T_{k,n-2}) + 1 \geq \dfrac{E_e(T_{k,n})}{k-2} - 1$.

We conclude that $E(T') \geq E(T_{k,n-2}) + 1$, for $k > 3, n > 2$.

Thus, as T' will have more than $E(k, n-2) + 1$ edges, T' has a path length of $n-2$. Therefore, T has n-path. We conclude that $E(T_{k,n}) + 1$ is an upper bound for the number of edges. ∎

Therefore, the minimum number of edges required to have either a k-star or an n-path is exactly $F(k,n) = \dfrac{(k-1)\left[(k-2)^{\lceil \frac{n}{2} \rceil - 1} - 1\right]}{k-3} + \dfrac{(1+(-1)^n)}{2}(k-2)^{\frac{n}{2}-1} + 1$.

3.3 Classification of k, n-Saturated Trees

We say that a tree T is k, n-saturated if it has no k-star or n-path but the addition of another edge would create either a k-star or a n-path.

Theorem 3.4. *The only tree with $E(T_{k,n})$ edges and no k-star or n-path is $T_{k,n}$.*

Let R be any k, n-saturated tree with $E(k, n)$ edges. We will show that R is $T_{k,n}$. Let R' be R without its leaves, so R' has no k-star and no $(n-2)$-path. We will show that $E(R') = E(T_{k,n-2})$. Then, by induction, $R' = T_{k,n-2}$.

Note: $E(R') = E(R) - L(R)$

Proof. From the above proof of the upper bound, we can use the relation between $E(R)$ and $E(R')$: $E(R') = \dfrac{E(R) - k + 1}{k-2}$.

For the base case, when $n = 3$, R' is a vertex. Thus, R is a star, which is $T_{k,3}$.

By induction, as we know that the number of edges in R is $E_o(T_{k,n})$, when n is odd,

$$\begin{aligned}
E_o(R') &= \frac{E_o(T_{k,n}) - k + 1}{k-2} \\
&= \frac{\dfrac{(k-1)\left[1-(k-2)^{\frac{n-1}{2}}\right]}{-k+3} - (k-1)}{k-2} \\
&= \frac{(k-1)\left[1+k-3-(k-2)^{\frac{n-1}{2}}\right]}{(-k+3)(k-2)} \\
&= \frac{(k-1)\left[1-(k-2)^{\frac{n-3}{2}}\right]}{-k+3} \\
&= E_o(T_{k,n-2})
\end{aligned}$$

Thus, by induction hypothesis, R' is $T_{k,n-2}$.

It follows from $E_o(T_{k,n}) = E_o(T_{k,n-2}) + (k-1)(k-2)^{\frac{n-3}{2}}$ that the number of leaves is $(k-1)(k-2)^{\frac{n-5}{2}}$ in R', and we need to add $(k-2)$ edges to each leaf of R' to recover R, which is $(k-1)(k-2)^{\frac{n-3}{2}} = L(R)$.

For the base case, when $n = 4$, R' is an edge, $T_{k,2}$.

By induction, when n is even,

$$\begin{aligned}
E_e(R') &= \frac{E_e(T_{k,n}) - k + 1}{k-2} \\
&= \frac{\dfrac{(k-1) - 2(k-2)^{\frac{n}{2}}}{-k+3} - (k-1)}{k-2} \\
&= \frac{(k-1)(k-2) - 2(k-2)^{\frac{n}{2}}}{k-2} \\
&= \frac{(k-1) - 2(k-2)^{\frac{n-2}{2}}}{-k+3} \\
&= E_e(T_{k,n-2})
\end{aligned}$$

Thus, by induction hypothesis, R' is $T_{k,n-2}$.

It follows from $E_e(T_{k,n}) = E_e(T_{k,n-2}) + 2(k-2)^{\frac{n-2}{2}}$ that the number of leaves is

$2(k-2)^{\frac{n-4}{2}}$ in R', and we need to add $(k-2)$ edges to each leaf of R' to recover R, which is $2(k-2)^{\frac{n-2}{2}} = L(R)$.

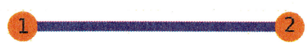
Therefore, as $R' = T_{k,n-2}$, we can conclude that $R = T_{k,n}$. ■

Thus, the only k, n-saturated trees are the family of trees we constructed.

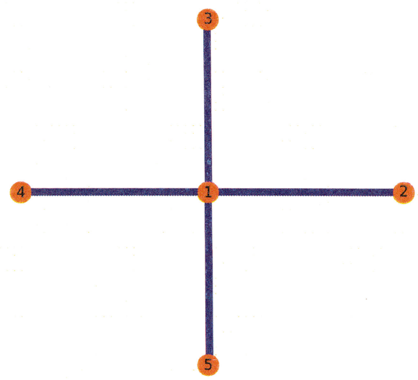
4 Automorphism Groups of k, n-Saturated Trees

From last chapter, we can see a nice symmetry in the trees we constructed. Here is how we relate graph theory with group theory: study the automorphism groups of the graphs. We would like to find a general expression for all the saturated trees automorphism groups.

To find the general formula, we look at trees with $k > 3, n > 1$. Using the example when $k = 5$, the graph automorphism shows a similar pattern as the number of edges in trees. When $k = 5$, it is obvious that

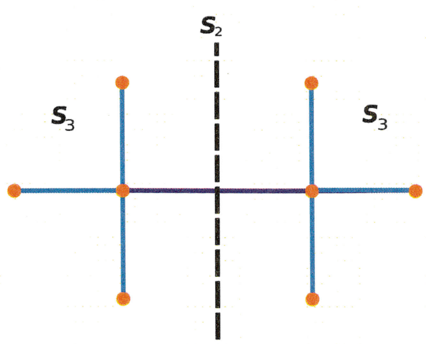


(a) $\text{Aut}(T_{5,2}) = \mathcal{S}_2$

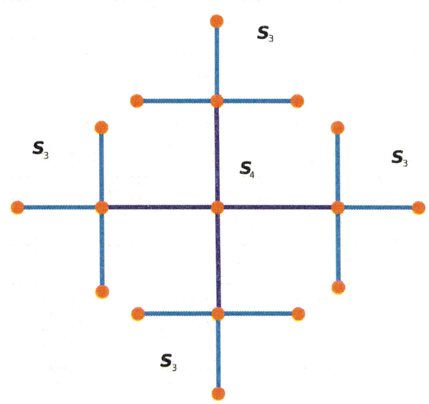


(b) $\text{Aut}(T_{5,3}) = \mathcal{S}_4$

Figure 10: Aut $(T_{5,2})$, Aut $(T_{5,3})$



(a) Aut $(T_{5,4})$



(b) Aut $(T_{5,5})$

Figure 11: Aut $(T_{5,4})$, Aut $(T_{5,5})$

Observe that in addition to the permutations of the two vertices in $T_{5,2}$, the left three vertices in figure 11a permute with each other, and so do the right three vertices. Thus, we have $\mathrm{Aut}(T_{5,4}) = \mathcal{S}_2 \times \mathcal{S}_3^2$

Similarly, in figure 11b, the left three vertices permute, so do the up, bottom and right three vertice, so we have $\mathrm{Aut}(T_{5,5}) = \mathcal{S}_4 \times \mathcal{S}_3^4$.

For $\mathrm{Aut}(T_{5,6}) : \mathcal{S}_2 \times \mathcal{S}_3^2 \times \mathcal{S}_3^6$, we give an example of the permutations in Figure 12 to illustrate why we had the symmetry group for the automorphism: $(13\,13')(10\,12')(11\,11')(12\,10')(1\,7')(2\,8')(3\,9')(7\,1')(8\,2')(9\,3')(4\,5')(5\,6')(6\,4')$

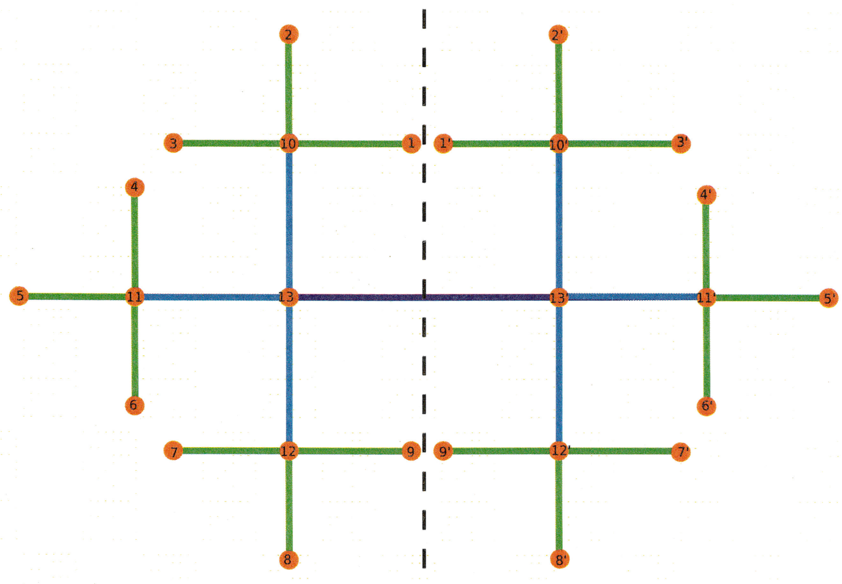


Figure 12: Aut $(T_{5,6})$

We can see that 13 and 13′ permute, and the internal nodes and leaves connected with them would permute correspondingly.

As we increase n,

$$\mathrm{Aut}(T_{5,7}) : \mathcal{S}_4 \times \mathcal{S}_3^4 \times \mathcal{S}_3^{12}$$
$$\mathrm{Aut}(T_{5,8}) : \mathcal{S}_2 \times \mathcal{S}_3^2 \times \mathcal{S}_3^6 \times \mathcal{S}_3{}^{18}$$
$$\mathrm{Aut}(T_{5,9}) : \mathcal{S}_4 \times \mathcal{S}_3^4 \times \mathcal{S}_2^{12} \times \mathcal{S}_2^{36}$$

It appears that $\text{Aut}(T_{5,n}) = \begin{cases} \mathcal{S}_2 \times [\mathcal{S}_3]^{3^{\frac{n-2}{2}}-1} & n \text{ even} \\ \mathcal{S}_4 \times [\mathcal{S}_3]^{2(3^{\frac{n-3}{2}}-1)} & n \text{ odd} \end{cases}$

Replacing 4 in $\text{Aut}(T_{5,n})$ with $k-1$ for $\text{Aut}(T_{k,n})$, and replacing 3 with $k-2$, we can see a general pattern here:

$$\begin{aligned}
\text{Aut}(T_{k,4}) &= \mathcal{S}_2 \times \mathcal{S}_{k-2}^2 \\
&= \text{Aut}(T_{k,2}) \times \mathcal{S}_{k-2}^2 \\
\text{Aut}(T_{k,5}) &= \mathcal{S}_{k-1} \times \mathcal{S}_{k-2}^{k-1} \\
&= \text{Aut}(T_{k,3}) \times \mathcal{S}_{k-2}^{k-1} \\
\text{Aut}(T_{k,6}) &= \mathcal{S}_2 \times \mathcal{S}_{k-2}^2 \times \mathcal{S}_{k-2}^{2(k-2)} \\
&= \text{Aut}(T_{k,4}) \times \mathcal{S}_{k-2}^{2(k-2)} \\
\text{Aut}(T_{k,7}) &= \mathcal{S}_{k-1} \times \mathcal{S}_{k-2}^{k-1} \times \mathcal{S}_{k-2}^{(k-1)(k-2)} \\
&= \text{Aut}(T_{k,5}) \times \mathcal{S}_{k-2}^{(k-1)(k-2)} \\
&\dots
\end{aligned}$$

Recall that $E_e(T_{k,n}) = E_e(T_{k,n-2}) + 2(k-2)^{\frac{n-2}{2}}$, as the number of edges being added is $(k-2) \times L_e(T_{k,n})$, then $L_e(T_{k,n}) = 2(k-2)^{\frac{n-2}{2}-1}$.

Similarly, we have $E_o(T_{k,n}) = E_o(T_{k,n-2}) + (k-1)(k-2)^{\frac{n-3}{2}}$, so
$L_o(T_{k,n}) = (k-1)(k-2)^{\frac{n-3}{2}-1}$.

Observe that as n grows, the number of symmetry group S_{k-2} being added to $T_{k,n-2}$ is the number of leaves in $T_{k,n-2}$. Therefore, we have a recursion relation here:
$\text{Aut}(T_{k,n}) = \text{Aut}(T_{k,n-2}) \times [S_{k-2}]^{L(T_{k,n-2})}$.

We will show that when n is even, $\text{Aut}_e(T_{k,n}) = \mathcal{S}_2 \times \mathcal{S}_{k-2}^{\alpha}$, where $\alpha = \dfrac{2[(k-2)^{\frac{n}{2}-1}-1]}{k-3}$.

Proof. For the base case, when $n = 2, \text{Aut}_e(T_{k,2}) = \mathcal{S}_2$. By induction, we can apply the recursion relation when n is even,

$$\begin{aligned}\mathrm{Aut}_e(T_{k,n}) &= \mathrm{Aut}_e(T_{k,n-2}) \times S_{k-2}{}^{L_e(T_{k,n-2})} \\ &= \mathcal{S}_2 \times [\mathcal{S}_{k-2}]^{\frac{2[(k-2)^{\frac{n-2}{2}-1}-1]}{k-3}} \times [\mathcal{S}_{k-2}]^{2\cdot(k-2)^{\frac{n-2}{2}-1}} \\ &= \mathcal{S}_2 \times [\mathcal{S}_{k-2}]^{2\cdot\frac{(k-2)^{\frac{n}{2}-2}-1+(k-3)^{\frac{n}{2}-2}}{k-3}} \\ &= \mathcal{S}_2 \times [\mathcal{S}_{k-2}]^{\frac{2[(k-2)^{\frac{n}{2}-1}-1]}{k-3}} \\ &= \mathcal{S}_2 \times \mathcal{S}_{k-2}^{\alpha}\end{aligned}$$

■

Similarly, we will show that when n is odd, $\mathrm{Aut}_o(T_{k,n}) = \mathcal{S}_{k-1} \times \mathcal{S}_{k-2}^{\beta}$, where $\beta = \dfrac{k-1}{k-3}[(k-2)^{\frac{n-3}{2}} - 1]$.

Proof. For the base case, when $n = 3, \mathrm{Aut}(T_{k,3}) = S_{k-1}$. By induction, we apply the same recursion relation when n is odd:

$$\begin{aligned}\mathrm{Aut}_o(T_{k,n}) &= \mathrm{Aut}_o(T_{k,n-2}) \times S_{k-2}{}^{L_o(T_{k,n-2})} \\ &= \mathcal{S}_{k-1} \times [\mathcal{S}_{k-2}]^{\frac{k-1}{k-3}[(k-2)^{\frac{n-3}{2}-1}]} \times [\mathcal{S}_{k-2}]^{(k-1)(k-2)^{\frac{n-3}{2}-1}} \\ &= \mathcal{S}_{k-1} \times [\mathcal{S}_{k-2}]^{\frac{(k-1)(k-3+1)(k-2)^{\frac{n-5}{2}}}{k-3}} \\ &= \mathcal{S}_{k-1} \times [\mathcal{S}_{k-2}]^{\frac{k-1}{k-3}[(k-2)^{\frac{n-3}{2}}-1]} \\ &= \mathcal{S}_{k-1} \times \mathcal{S}_{k-2}^{\beta}\end{aligned}$$

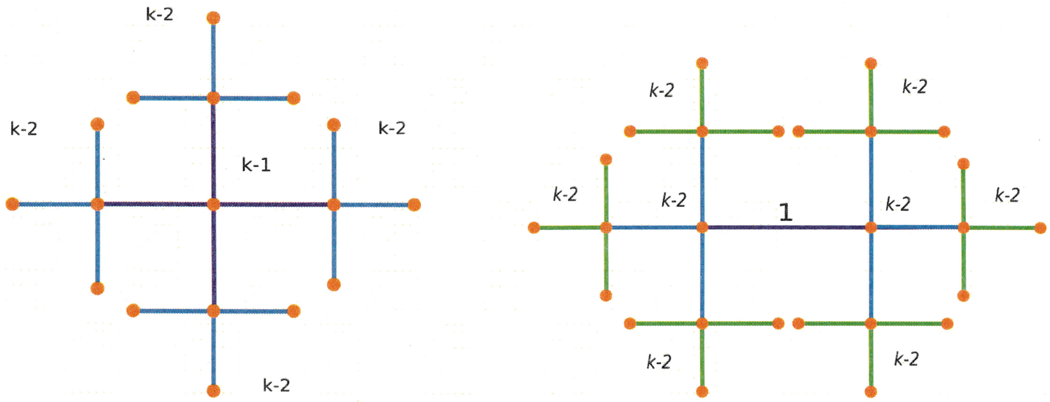
■

Thus, we conclude that
$\mathrm{Aut}_e(T_{k,n}) = \mathcal{S}_2 \times \mathcal{S}_{k-2}^{\alpha}$, where $\alpha = \dfrac{2[(k-2)^{\frac{n}{2}-1} - 1]}{k-3}$, when n is even
$\mathrm{Aut}_o(T_{k,n}) = \mathcal{S}_{k-1} \times \mathcal{S}_{k-2}^{\beta}$, where $\beta = \dfrac{k-1}{k-3}[(k-2)^{\frac{n-3}{2}} - 1]$, when n is odd.

Observe that the pattern of α, β looks similar to what we have for $E(T_{k,n})$, we want to study that: *how can we relate the automorphism group to the number of edges we found above?*

Focus on the number of S_{k-2} in $\mathrm{Aut}_e(T_{k,n})$,

$$\begin{aligned}\alpha = \frac{2[(k-2)^{\frac{n}{2}-1}-1]}{k-3} &= \frac{2(k-2)^{\frac{n}{2}}-2(k-2)}{k-3}\\ &= \frac{2(k-2)^{\frac{n}{2}}-(k-1)-(k-3)}{(k-2)(k-3)}\\ &= \frac{2(k-2)^{\frac{n}{2}}-(k-1)}{(k-2)(k-3)} - \frac{k-3}{(k-2)(k-3)}\\ &= \frac{E_e(k,n)-1}{k-2}\end{aligned}$$



(a) $k=5, n=5$ (b) $k=5, n=6$

Figure 13: Relation between $\mathrm{Aut}(T_{k,n})$ and $E(T_{k,n})$

Focus on the number of S_{k-2} in $\mathrm{Aut}_o(T_{k,n})$,

$$\begin{aligned}\beta = \frac{k-1}{k-3}\left((k-2)^{\frac{n-3}{2}}-1\right) &= \frac{(k-1)(k-2)^{\frac{n-1}{2}}-(k-1)(k-2)-(k-1)+(k-1)}{(k-2)(k-3)}\\ &= \frac{(k-1)\left((k-2)^{\frac{n-1}{2}}-1\right)}{(k-3)(k-2)} - \frac{(k-1)(k-2)-(k-1)}{(k-2)(k-3)}\\ &= \frac{E_o(k,n)-(k-1)}{k-2}\end{aligned}$$

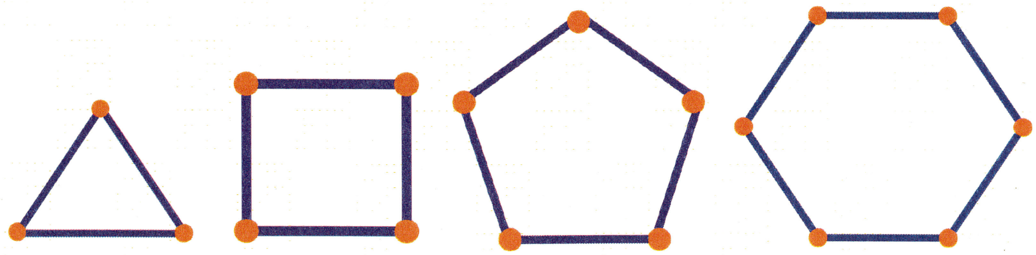
This is a nice property of our trees. We can see that the structure of the trees are fairly systematic. Now we look at connected graphs and see what properties they have.

5 Number of Edges in Connected Graphs

Substituting "tree" with "connected graph", we look for the maximum number of edges that a connected graph can have to avoid both k-star and n-path.

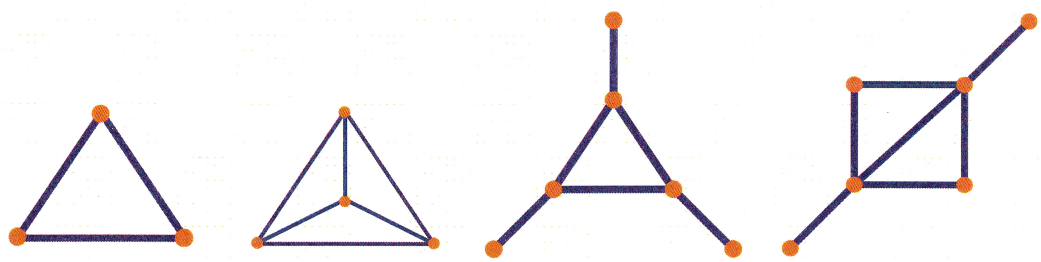
We construct a class of connected graphs, $C_{k,n}$, starting with small k values.

Let $k = 3$, in Figure 14, we can see that the maximum number of edges is n.



(a) $C_{3,3}$ (b) $C_{3,4}$ (c) $C_{3,5}$ (d) $C_{3,6}$

Figure 14: Connected graphs, $k = 3$



(a) $C_{4,3}$ (b) $C_{4,4}$ (c) $C_{4,5}$ (d) $C_{4,5}$

Figure 15: Connected graphs, $k = 4$

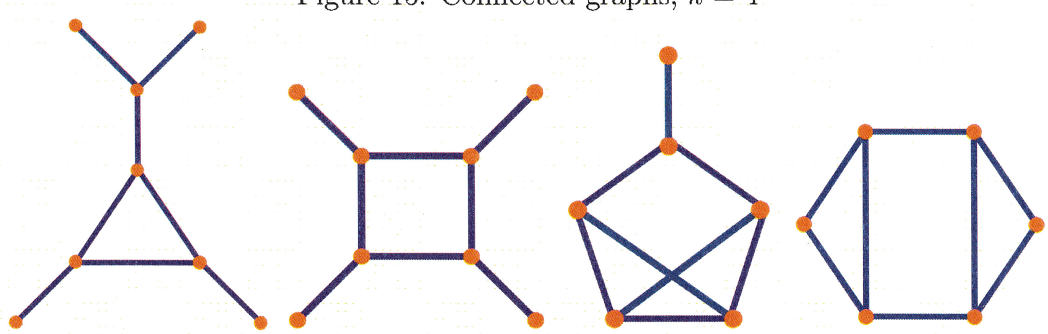


Figure 16: Connected graphs, $C_{4,6}$

Let $k = 4$, we start with the shapes we found for $k = 3$ and try to add as many edges as possible.

As shown in Figure 15 and 16, We have found all the possible connected graphs up to $k = 4, n = 6$. Comparing the number of maximum edges for the above connected graphs with the one of trees, we can see that in most cases, $E(T_{k,n})$ is greater than $E(C_{k,n})$.

$E(T_{k,n})$	# Edges
(4,3)	3
(4,4)	5
(4,5)	9
(4,6)	13
(4,7)	21
(4,8)	29

$E(C_{k,n})$	# Edges
(4,3)	3
(4,4)	6
(4,5)	7
(4,6)	8
(4,7)	...
(4,8)	...

Let's look at the complete graph on k vertices in Figure 17, which gives the maximum number of edges,

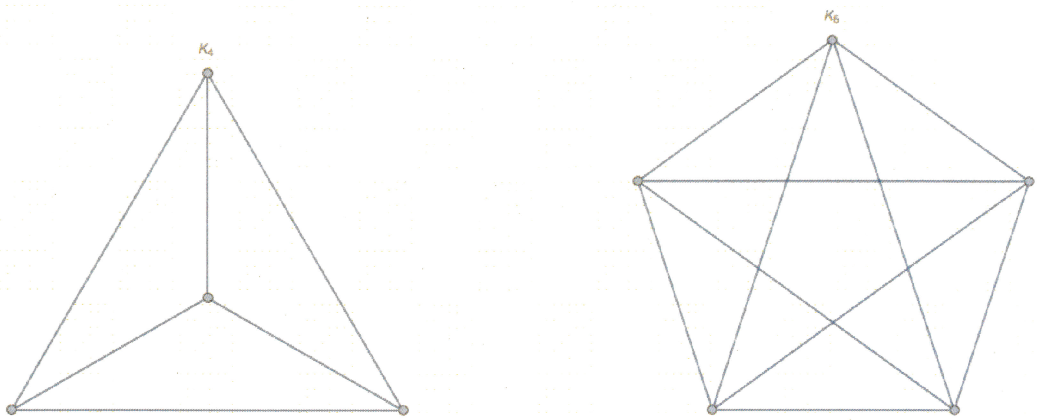


Figure 17: Complete Graphs

we can see that K_4 is a 4, 4-saturated connected graph and K_5 is a 5, 5-saturated connected graph. We can use the fact that when $k = n$, K_n gives the maximum number of edges for $C_{k,n}$, which is $\frac{k(k-1)}{2}$. Recall that in $E(T_{k,n})$, we have $(k-1)$ multiplied with $(k-2)$ to some power, which grows faster than complete connected graph, $(k-1)$ multiplied with $\frac{k}{2}$.

If we only allow connected graphs that are not trees, then the maximum number of edges in $C_{k,n}$ will always be smaller than $E(T_{k,n})$ except $E(C_{4,4})$.

Since we want to know the minimum number of edges required to have either a

k-star or an n-path, we should have as less edges as possible, which is $E(C_{k,n}) + 1$. However, with $E(C_{k,n}) + 1$ edges, if we rearrange the edges following a tree structure, then we will have no k-star or n-path given that $E(C_{k,n}) < E(T_{k,n})$ for most cases.

Thus, we have the conjecture that $E(C_{k,n}) + 1 = F(k,n) = E(T_{k,n}) + 1$, when $k > 3, n > 3$ except $k = 4, n = 4$.

In addition, observe that both figure 15c and figure 15d are $4,5$-saturated connected graphs. All the connected graphs in Figure 16 are $4,6$-saturated connected graphs. As a result, we have shown that k,n-saturated graphs will have multiple possible structures as the value of k or n increases. Therefore, with $E(T_{k,n}) + 1$, the connected graphs we constructed will not be the only k,n-saturated connected graphs. Thus, the classification of them is worthy further study in the future.

6 Acknowledgements

I would like to express my sincere gratitude to my advisor, Prof. Eric Gottlieb, for supervising this project. Prof. Gottlieb reviewed my work carefully and took efforts to correct my endless grammatical mistakes. I constantly benefited from his continuous guidance and support all along my college life.

I would also like to thank Prof. Shubho Banerjee for his advice on writing a mathematics paper to non-mathematics readers, and Prof. Seaton for his willingness to be the second reader.

I am furthermore appreciative to Prof. Michael Sheard and Prof. Chris Mouron for all the corrections to my thesis.

Finally, I would like to thank the department for giving me the opportunity to undertake this research and present at conferences.

References

[1] Miklós Bóna. *A walk through combinatorics: an introduction to enumeration and graph theory*. World scientific, 2011.

[2] Martin J Erickson. *Introduction to combinatorics*, volume 79. John Wiley & Sons, 2013.

[3] Ronald L Graham, Bruce L Rothschild, and Joel H Spencer. *Ramsey thoery*, volume 20. John Wiley & Sons, 1990.