GENERALIZED CATALAN NUMBERS

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF MATHEMATICS

A project submitted in partial fulfillment of the requirement of the degree of master of science in mathematics

By: Eyerusalem Aseressie Abeje
Advisor: Zelealem Belaineh(PhD)

June, 2014

Addis Ababa, Ethiopia

Contents

Acknowledgement

Foremost, I express my sincere gratitude to my advisor Dr. Zelealem Belaineh to whom I am very grateful for all his support throughout this entire process. I also express my thanks to the Department of Mathematics of Addis Ababa University and to all my friends for their unmentioned cooperation on my project.

On top of this, my thanks goes to Wolaita University for its financial support in attending the degree of master of science.

Abstract

Ballot problem was introduced in the late nineteen century to determine the probability of counting of votes when there are two candidates A and B such that A receives a votes, and B receives b votes with $a > b$. This project focuses on the proof of Generalized Catalan numbers by using Ballot Theorem. We have introduced basic concepts and definitions in the first part. And in the second part, ballot problem is proved with three approaches such as proof by counting bad path, proof by induction and proof by cycle lemma. Then we used ballot theorem for the proof of generalized Catalan numbers formula.

Introduction

Catalan Numbers are a sequence of numbers that arise in various problems. These numbers were first discovered by Leonhard Euler in the 18th century while he was trying to see how many ways a polygon with $n+2$ sides can be divided into n triangles without any of the lines intersecting. However, according to a 1988 article by Chinese mathematician J. J. Luo, Chinese mathematician Antu Ming discovered them about 1730 through his geometric models. Mings work was published in Chinese, so it was not known in the West. Belgian mathematician Eugene Charles Catalan also discovered Catalan numbers in 1838, while studying well-formed sequences of parentheses. Even though Catalan numbers were discovered long time ago, they were named after Eugene Charles Catalan. Kreher and Stinson (1956) define the Catalan numbers in term of totally balanced sequence.

Catalan numbers, which appear in many counting problem such as monotonic paths, dividing polygons, and non crossing partitions, form a sequence of natural numbers. The terms of the sequence can be calculated by the formula: $C_n = \frac{1}{n+1}\binom{2n}{n}$. It grows rapidly: $C_n = 1, 1, 2, 5, 14, 42, 132, 429, 1430, ...$ for $n = 0, 1, 2, 3, 4, 5, 7, 8, ...$

Ballot problem was introduced in the late nineteen century to determine the probability of counting of votes when there are two candidates A and B such that candidate A receives a votes and candidate B receives b votes, where $a \geq kb$ for some positive integer k. In how many ways can the ballots be ordered so that A maintains more than k times as many votes as B throughout the counting of the ballots? In 1887 Joseph Bertrand introduced the ballot problem for $k = 1$, and gave a proof by induction. In the same year Emile Barbier stated a solution to the ballot problem for arbitrary k but with out proof. Still in the same year Désiré André produced a short combinatorial proof of the ballot problem for $k = 1$. In 1923, Aeppli gave the first proof of the ballot problem for $k \geq 1$.

This paper is organized into two chapters. The first chapter deals with preliminary concepts and definitions about lattice paths, generalized Catalan numbers and ballot theorem. The second chapter focuses on the proof of ballot problem and generalized Catalan numbers $C_{(n,k)} = \frac{1}{kn+1}\binom{kn+n}{n}$. Ballot problem is proved with three approaches such as proof by counting bad path, proof by induction and proof by cycle lemma. Then the ballot theorem is used to prove the formula of generalized Catalan numbers $\frac{1}{kn+1}\binom{kn+n}{n}$.

Chapter 1

Preliminary

1.1 Definition and Preliminary Concepts

This chapter discuss basic definitions and concepts on lattice paths, generalized Catalan numbers and Ballot theorem.

1.1.1 Catalan Numbers

They are a sequence of numbers that arise in various problems. The terms of the sequence can be calculated by the formula:

$$C_n = \tfrac{1}{n+1}\tbinom{2n}{n}$$

Notice how the terms of the sequence generated grow rapidly:
$C_n = 1, 1, 2, 5, 14, 42, 132, 429, 1430, ...$ There are other different variations of the formula but these are all equivalent, including:

$$C_n = \tbinom{2n}{n} - \tbinom{2n}{n+1}$$

Proposition 1.1. *Catalan numbers are natural numbers.*

Proof. We shall now see that the Catalan numbers are natural numbers. We shall do this by proving the previous formula: $\tbinom{2n}{n} - \tbinom{2n}{n+1}$

For $n = 0$ we can check using the above formula $\binom{2(0)}{0} - \binom{2(0)}{1}$

$$\begin{aligned}\binom{2(0)}{0} - \binom{2(0)}{1} &= \binom{0}{0} - \binom{0}{1}\\ &= 1 - 0 = 1 = C_0\end{aligned}$$

Now assume $n \geq 1$:

$$\begin{aligned}C_n = \frac{1}{n+1}\binom{2n}{n} &= \frac{(2n)!}{(n+1)!n!}\\ &= (2n)!\frac{1}{(n+1)!n!}\\ &= (2n)!\frac{(n+1)-n}{(n+1)!n!}\\ &= (2n)!\frac{n+1}{(n+1)!n!} - \frac{n}{(n+1)!n!}\\ &= (2n)!\frac{1}{n!n!} - \frac{1}{(n+1)!(n-1)!}\frac{(2n)!}{n!n!} - \frac{(2n!)}{(n+1)!(n-1)!}\\ &= \binom{2n}{n} - \binom{2n}{n+1}\end{aligned}$$

since $\binom{2n}{n}$ and $\binom{2n}{n+1}$ are elements of natural numbers also $\binom{2n}{n} > \binom{2n}{n+1}$, since $\binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1}\binom{2n}{n}$ is positive number. We can conclude that as C_n is the difference of two integers, $C_n = \binom{2n}{n} - \binom{2n}{n+1}$ an element of natural number. Expressing C_n as the difference of two binomial coefficients, we have thus proved that it is in fact a natural number.

The Catalan numbers appear within combinatorial problems in mathematics. A few of the main problems are;

▲ Catalan numbers in Pascals triangle

▲ The Grid problem

Catalan numbers in Pascal's triangle

Pascal's triangle is the binomial coefficients $\binom{n}{m}$, where $0 \leq m \leq n$; can be arranged in the form a triangular array, show in the figure below

$$\begin{array}{c}
\binom{0}{0} \\
\binom{1}{0} \quad \binom{1}{1} \\
\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2} \\
\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3} \\
\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4} \\
\binom{5}{0} \quad \binom{5}{1} \quad \binom{5}{2} \quad \binom{5}{3} \quad \binom{5}{4} \quad \binom{5}{5} \\
\binom{6}{0} \quad \binom{6}{1} \quad \binom{6}{2} \quad \binom{6}{3} \quad \binom{6}{4} \quad \binom{6}{5} \quad \binom{6}{6} \\
\binom{7}{0} \quad \binom{7}{1} \quad \binom{7}{2} \quad \binom{7}{3} \quad \binom{7}{4} \quad \binom{7}{5} \quad \binom{7}{6} \quad \binom{7}{7}
\end{array}$$

The Catalan series can be found in Pascal's triangle: Looking at the numbers in the central column and the adjacent column, we notice that the difference of these numbers produces the Catalan sequence.
$C_n = \binom{2n}{n} - \binom{2n}{n-1}$

$$\begin{array}{lc}
2-1 & 1 \\
 & 1 \quad 1 \\
6-4 & 1 \quad \boxed{2} \quad \boxed{1} \\
 & 1 \quad 3 \quad 3 \quad 1 \\
20-15 & 1 \quad 4 \quad \boxed{6} \quad \boxed{4} \quad 1 \\
 & 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\
70-56 & 1 \quad 6 \quad 15 \quad \boxed{20} \quad \boxed{15} \quad 6 \quad 1 \\
 & 1 \quad 7 \quad 21 \quad 35 \quad 35 \quad 21 \quad 7 \quad 1 \\
 & 1 \quad 8 \quad 28 \quad 56 \quad \boxed{70} \quad \boxed{56} \quad 28 \quad 8 \quad 1 \\
 & 1 \quad 9 \quad 36 \quad 84 \quad 126 \quad 126 \quad 84 \quad 36 \quad 9
\end{array}$$

Catalan numbers in Grids problem

This problem looks at the number of ways to cross an n by n grid, in the shortest way, starting in the bottom left corner and going to the top right, without crossing the diagonal line.

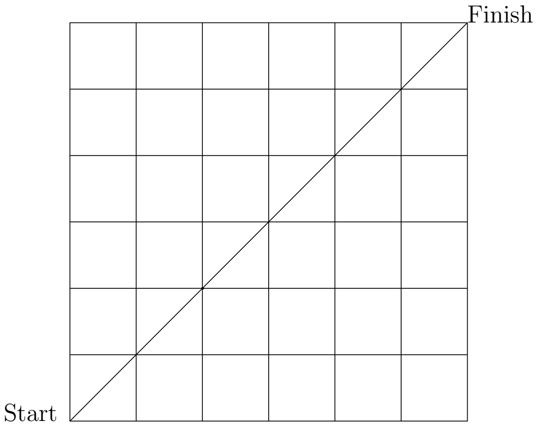


Figure 1.1 must make only below the diagonal

A typical path:

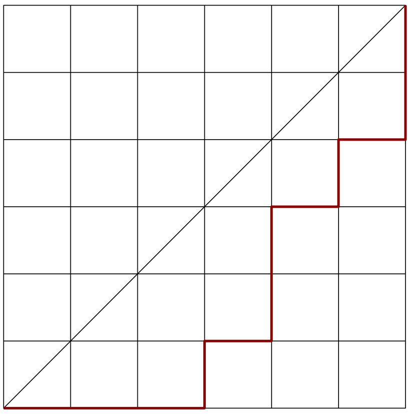


Figure 1.2

Example 1.1. *On a* 3*-by-*3 *integer grid, find the number of different possible integral walk allowing only* 3 *rightward moves and* 3 *upward moves the origin* $(0,0)$ *to the destination* $(3,3)$ *which do not cross the diagonal* $y = x$*. The number of paths not cross the diagonal are* 5*.*

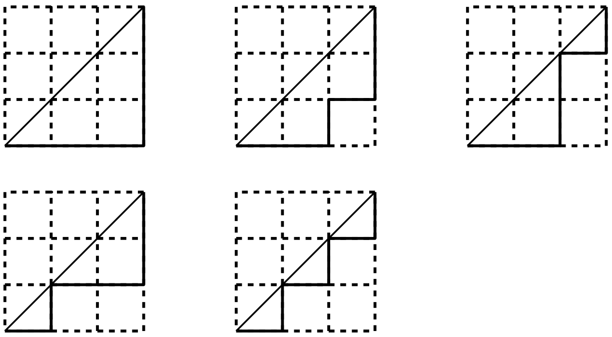


Figure 1.2: The solution is 5

These possible integral walks are called lattice path.

Definition 1.1.1. *A lattice path is a path in the plane consisting of unit up-steps and right-steps, whose ends are points with integer coordinates.*

Lattice paths from $(0,0)$ to (n,n) which never go above the diagonal line $y=x$ are good paths and lattice paths which cross the diagonal and/or above the diagonal are called bad paths.

Equivalently, we define lattice paths in the plane consisting of unit up steps and down steps such that no step ends on or below the x - axis. Paths that remain above the x-axis (after the origin) are good paths, while those with steps that end on or below the x-axis are bad. A down step that starts above the x-axis and ends on or below the x-axis is called a bad path.

Definition 1.1.2. *The Reflection Principle is that the set of paths from* $(0,0)$ *to* (n,n) *that cross the diagonal* $(y=x)$ *somewhere is in one-to-one correspondence with the set of all paths from* $(-1,1)$ *to* (n,n).

Theorem 1.1. *Let* n *be a non negative integer. Then the number of lattice paths from (0,0) to (n,n) which never go above the diagonal line* $y=x$ *is the Catalan number* $C_n = \frac{1}{n+1}\binom{2n}{n}$.

Proof by Reflection Principle.

We want to count the number of good paths (not crossing and above the diagonal $y=x$) from $O(0,0)$ to $N(n,n)$. This is equivalent to counting the total number of paths from $O(0,0)$ to $N(n,n)$ minus the total number of bad paths. The total of $n+n$ moves, consisting of n moves up and n moves to the right in any order and there are $\binom{n+n}{n}$ ways to choose which of the n moves are up (or equivalently, $\binom{n+n}{n}$ ways to choose which n of them are to the right). So the total number of paths from $O(0,0)$ to $N(n,n)$ is $\binom{2n}{n}$. Let's consider a bad path which crosses the main diagonal $y = x$ and first intersects the line $y = x+1$ at the point Q. We then reflect the sub-path OQ (denoted by P_1) about the line $y = x+1$ to get a new path $O'N$, comprising of the new reflected sub-path P_1 and the sub-path QN (denoted by P_2). By reflection principle definition each bad path has a one-to-one correspondence with the new path $P_1'P_2$. Since each new path starts from $O'(-1,1)$ and ends at $N(n,n)$, the number of bad paths is $\binom{n-1+n+1}{n-1}$.

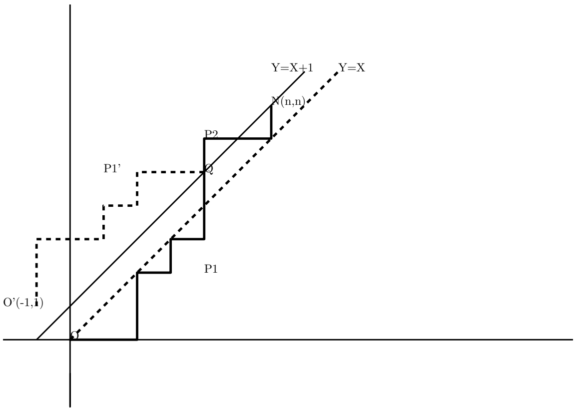


Figure 1.3:

Hence, the number of good paths are

$$\begin{aligned}
\binom{2n}{n} - \binom{n-1+n+1}{n-1} &= \binom{2n}{n} - \binom{2n}{n-1} \\
&= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!} \\
&= \frac{2n!}{n!(n-1)!}(\frac{1}{n} - \frac{1}{n+1}) \\
&= \frac{(2n)!}{n!(n-1)!}(\frac{1}{n(n+1)}) \\
&= \frac{(2n)!}{(n+1)n!n!} \\
&= \frac{1}{n+1}\binom{2n}{n} = C_n
\end{aligned}$$

Let $P_{n,k}$ be the number of paths from $(0,0)$ to (kn,n) for $k \geq 0$ and $n \geq 0$, then $P_{n,k} = \frac{1}{kn+1}\binom{(k+1)n}{n}$.
$P_{n,k}$ can be found by counting good lattice paths.

Example 1.2. *The number of lattice paths from* $(0,0)$ *to* $(2n,n)$ *which never go above the diagonal line* $y = \frac{1}{2}x$.

For $n = 1$

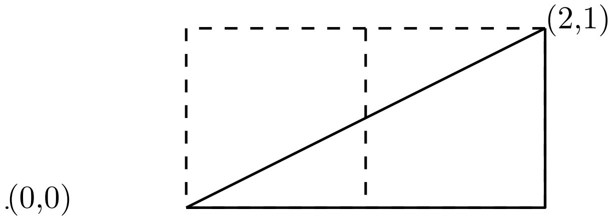


Figure 1.4: Number of good path is one, $1 = \frac{1}{3}\binom{3}{1}$

For $n = 2$

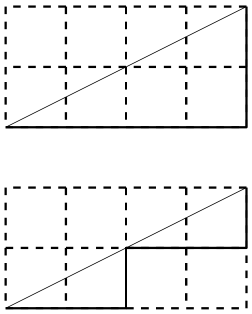


Figure 1.5: Number of good paths are three, $3 = \frac{1}{5}\binom{6}{2}$

For $n = 3$

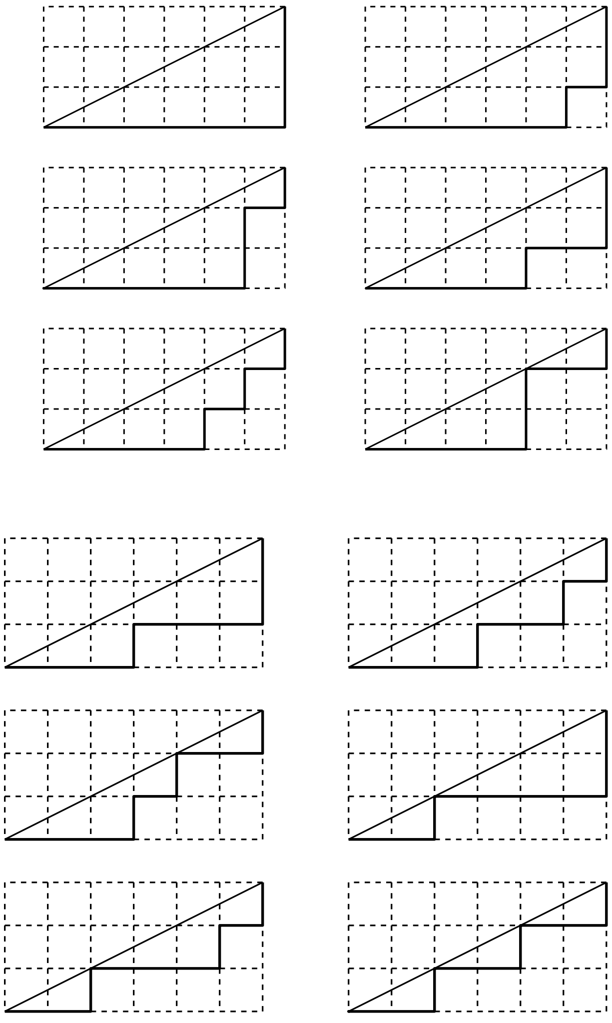


Figure 1.6: number of good paths are 12, $12 = \frac{1}{7}\binom{9}{3}$

From the above example, the number of lattice paths from $(0,0)$ to $(2n,n)$ which never go above the diagonal line $y=\frac{1}{2}x$ generates a sequence of numbers $1,3,12,55,...,\frac{1}{2n+1}\binom{3n}{n}$.

Similarly, for a fixed positive integer k, the number of lattice paths from $(0,0)$ to (kn,n) which never go above the diagonal line $y=\frac{1}{k}x$ generates a sequence of numbers:

$k=1:\ 1,1,2,5,14,...,\frac{1}{n+1}\binom{2n}{n}$

$k=2:\ 1,1,3,12,55,...,\frac{1}{2n+1}\binom{3n}{n}$

$k=3:\ 1,1,4,22,140,...,\frac{1}{3n+1}\binom{4n}{n}$ where, $n=0,1,2,3,...$

Thus, $P_{k,n}=\frac{1}{kn+1}\binom{kn+n}{n}$.

Chapter 2

Generalized Catalan Numbers

In this chapter, we intend to prove the formula of generalized Catalan numbers. Before we proceed directly to the proof of generalized Catalan numbers, it is important to prove ballot theorem.

2.1 Generalized Catalan Numbers

Proposition 2.1. *Every generalized Catalan number $C_{(n,k)}$ is an integer.*

Proof.

$$\begin{aligned}\binom{kn+n}{n}-\binom{kn+n}{n-1} &= \frac{(kn+n)!}{n!(kn)!}-\frac{(kn+n)!}{(n-1)(kn+1)!}\\ &= \frac{(kn+n)!}{n!(kn)!}-\frac{n(kn+n)!}{n(n-1)!(kn+1)(kn)!}\\ &= \frac{(kn+n)!}{n!(kn)!}-\frac{n(kn+n)}{n!(kn+1)(kn)!}\\ &= \frac{(kn+n)!}{n!(kn)!}(1-\frac{n}{kn+1})\end{aligned}$$

$$= \frac{(kn+n)!}{n!(kn)!}\frac{kn+1-n}{kn+1} \tag{2.1}$$
$$= \frac{1}{kn+1}\binom{kn+n}{n}[(k-1)n+1]$$

Because the LHS is an integer, it follows that the RHS is also an integer. But $\gcd(kn+1,(k-1)n+1) = 1$ (from prime factorization, the proof is in appendix) which implies $kn+1$ and $(k-1)n+1$ are relatively prime.
It follows that $\frac{1}{kn+1}\binom{kn+n}{n}$ is an integer since RHS is an integer and $\gcd(kn+1,(k-1)n+1) = 1$.
Hence, $C_{(n,k)}$ is an integer for every $n, k \geq 0$.

□

Definition 2.1.1. *A sequence $P_1P_2...P_j$ of A's and B's is called good sequence if for every position i, $1 \leq i \leq j$ the number of A's in $P_1P_2...P_i$ is more than k times the number of B's.*

Definition 2.1.2. *The Pigeonhole Principle version: If $k+1$ or more pigeons are distributed among k pigeonholes, then at least one pigeonhole contains two or more pigeons.*

Example 2.1. *In any group of n people there are at least two persons having the same number friends. (It is assumed that if a person x is a friend of y then y is also a friend of x.)*

Proof. The number of friends of a person x is an integer k with $0 \leq k \leq n-1$. If there is a person y whose number of friends is $n-1$, then everyone is a friend of y, that is, no one has 0 friend. This means that 0 and $n-1$ can not be simultaneously the numbers of friends of some people in the group. The pigeonhole principle tells us that there are at least two people having the same number of friends.

□

2.2 The Ballot Problem.

Suppose that an election, candidate A receives a votes and candidate B receives b votes, where $a \geq kb$ for some positive integer k. Then the number of ways the ballots can be ordered so that A maintains more than k times as many votes throughout the counting of the ballot is $\frac{a-kb}{a+b}\binom{a+b}{a}$.

Note: a permutation of the ballots good if A maintains more than k times as many votes throughout the counting of the ballot, and bad otherwise.

Example 2.2. *Suppose the ballots are candidate A gets* 5 *votes and candidate B gets* 3 *votes. Then we wish to count ballot permutations such as*

$$AABABAAB$$

but not

$$AABABBAA.$$

2.2.1 Proof by Counting the Bad Ballot Permutations

1. **André Method**
André was the first to solve the ballot problem by subtracting the number of bad ballot permutations from the number of all ballot permutations for the case $k = 1$. André interchanges two portions of a ballot permutation in his proof.
First, André observes that every ballot permutation starting with a B must be bad. The number of bad permutations starting with B equals the number of all permutations which one can form with a letters A and $b-1$ letters B, because it is enough to suppress the initial letter B to obtain the remaining letters. And there are $\binom{a+b-1}{a}$.
Next, André counts the number of bad ballot permutations starting with A.
Claim: number of bad permutations starting with A = number of all permutations with a A's and $b-1$ B's .

Proof. ($\longrightarrow$) Consider a bad ballot permutation starting with A. Reading from left to right, find the first B that causes the number of A's to equal the number of B's. Remove that B, causing the permutation to split into two parts. Interchange the parts and join them together to create a permutation with a A's and $b-1$ B's.
($\longleftarrow$) Consider any ballot permutation with a A's and $b-1$ B's. Reading from right to left, find the first A that causes the tail of the permutation (starting with that A) to contain one more A than B; this must occur since $a > b-1$. Take off the tail of the permutation, put it in front, and insert a B between the two parts. This creates a bad ballot permutation starting with A. $\square$

Therefore, number of bad permutation starting with A is equal to number of all permutations with a A's and $b-1$ B's. Hence $\binom{a+b-1}{a}$ start with A.

Then bad permutations start with A and start with B are $2\binom{a+b-1}{a}$.

Thus, the number of good ballot permutations are

$$\begin{aligned}\binom{a+b}{a} - 2\binom{a+b-1}{a} &= \frac{(a+b)!}{a!b!} - \frac{2(a+b-1)!}{a!(b-1)!}\\ &= \frac{(a+b)!}{a!b!} - \frac{2b(a+b)!}{a!b!(a+b)}\\ &= \frac{(a+b)!}{a!b!}(1-\frac{2b}{a+b})\\ &= \frac{(a+b)!}{a!b!}\frac{a+b-2b}{a+b}\\ &= \frac{a-b}{a+b}\binom{a+b}{a}\end{aligned}$$

Example:
($\rightarrow$) Given a bad permutation starting with A;

$$AABBABAA$$

then, find the first bad B

$$AAB\boxed{\text{B}}ABAA$$

remove it,

$$AAB\ ABAA,$$

exchange the two parts:

$$ABAAAAB.$$

This creates ballot permutation with a A's and $b-1$ B's.

(←) Given any ballot permutation with a A's and $b-1$ B's where $a > b-1$

$$ABAAAAB$$

Identify the smallest tail of the permutation containing one more A than B.

$$ABAA\boxed{\text{AAB}}$$

Remove the tail, place at the beginning of the permutation.

$$AAB\ ABAA$$

Insert a B between the two parts

$$AABBABAA$$

The result is a bad ballot permutation with a $A's$, b $B's$, and it starts with A.

2. **Proof by rotation** 180^0

For simplicity, we consider lattice paths in the plane with votes for A as up steps and votes for B as down steps. Hence, ballot permutations are considered as lattice paths in the plane with votes for A as up steps and votes for B as down steps such that no step ends on or below the x-axis. Paths that remain above the x-axis (after the origin) are good, while those with steps that end on or below the x-axis are bad.

For example; the lattice paths in the plane with votes for A = 8 and votes for B = 6.

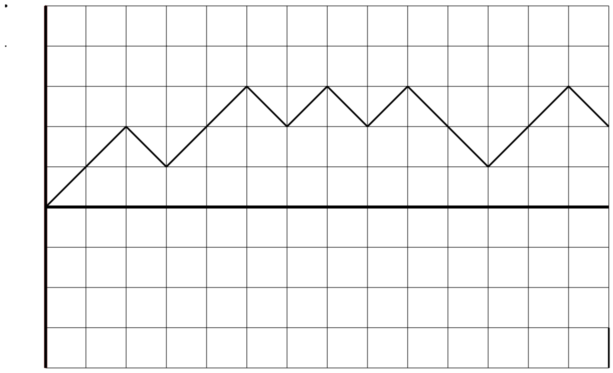


Figure 2.1: good ballot permutation

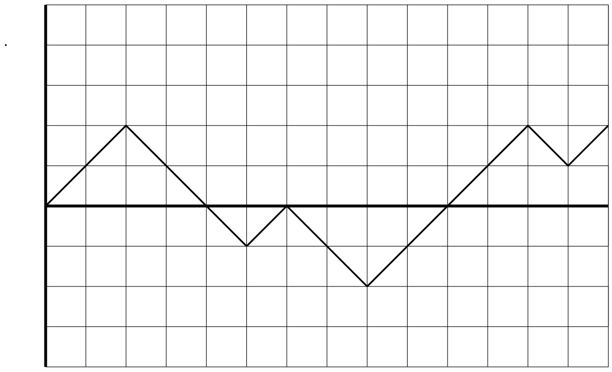


Figure 2.2: bad ballot permutation

Note: the ballot problem is to find the number of ways ballot permutations such that candidate A maintain votes more than k times the votes of candidate B.
Here let us examine the solution of ballot problem in a lattice path. In this case we count lattice paths with a one unit upsteps and b k unit downsteps. Example; $k = 3$ $a = 8$ $b = 2$.

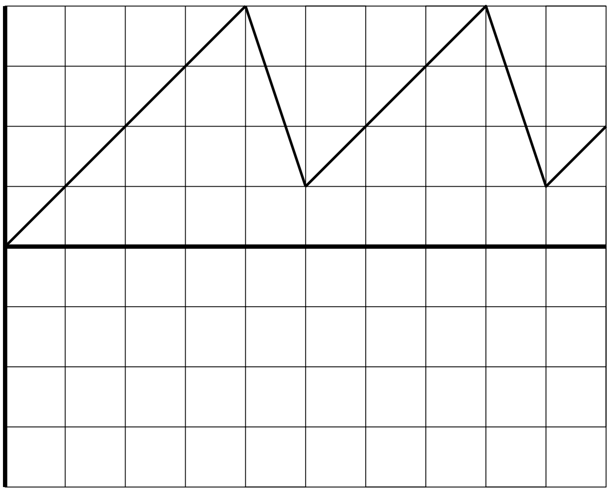


Figure 2.3: good path with $k = 3, a = 8$ and $b = 2$

Bad path with first bad step end 0 unit below the x-axis.

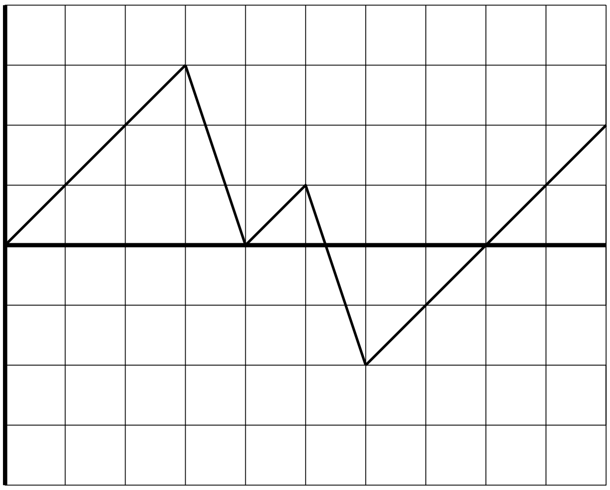


Figure 2.4: bad path with $k = 3, a = 8$ and $b = 2$

Proof. (Proof of ballot problem by rotation)
$k \geq 1$. For $0 \leq i \leq k$, let B_i denote the set of bad paths whose first bad step ends i units below the x-axis. Clearly these $k+1$ sets are disjoint and their union is the set of all bad paths. The paths in B_k are exactly those paths that start with a down step. Because the first bad step is k unit below the

x-axis and k is the maximum down steps. Hence, by André method $B_k = \binom{a+b-1}{a}$. We now show that for any $i \neq k$ we actually have $|B_i| = |B_k|$. Let P be a path in B_i, $(i \neq k)$, and identify the first step of P that ends i units below the x-axis. Let X be the initial segment of P that ends with that step and write $P = XY$. Let X' denote the path that results from rotating X by 180^0 exchanging its endpoints. Since X ends with a down step, X' starts with a down step, and consequently $X'Y \in B_k$.

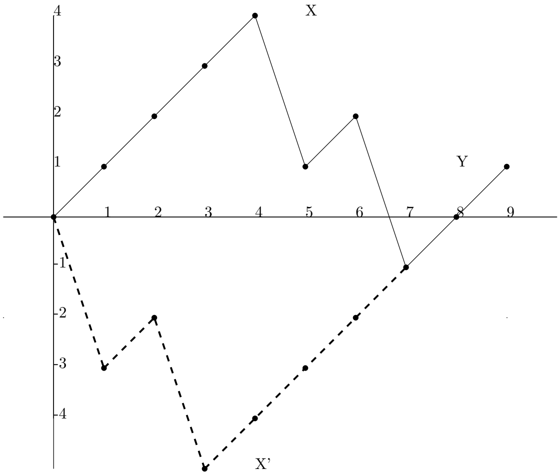


Figure 2.5: *Rotation of X by 180^0 exchanging the endpoints*

The same process converts a path in B_k into a path in B_i, $(i \neq k)$. If $P \in B_k$, then identify the first step that ends i units below the x-axis. Let X denote the initial segment of P that ends with that step and write $P = XY$. Since X necessarily ends with an up step, we have $X'Y \in B_i$.

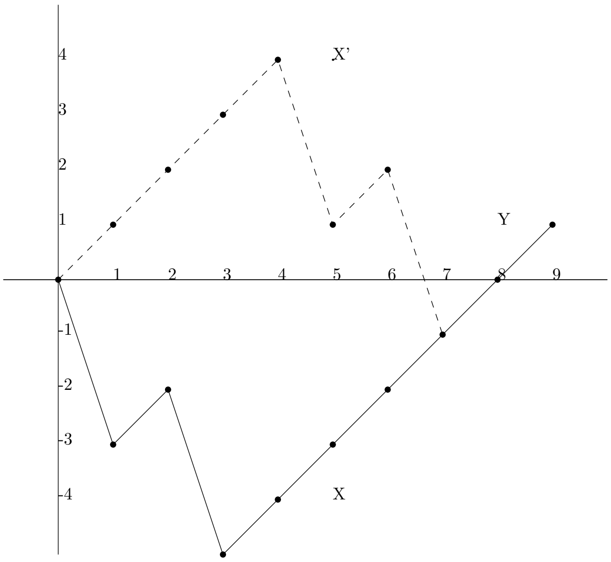


Figure 2.6: *Rotation of X by* 180^0 *exchanging the endpoints*

Thus each of the $k+1$ sets B_i have cardinality $\binom{a+b-1}{a}$, and the number of good paths is

$$\begin{aligned}\binom{a+b}{a} - (k+1)\binom{a+b-1}{a} &= \frac{(a+b)!}{a!b!} - (k+1)\frac{(a+b-1)!}{a!(b-1)!} \\ &= \frac{(a+b)!}{a!b!} - (k+1)b\frac{(a+b)!}{(a+b)a!b!} \\ &= \frac{(a+b)!}{a!b!}(1-(k+1)\frac{b}{a+b}) \\ &= \binom{a+b}{a}(1-\frac{(kb+b)}{a+b}) \\ &= \binom{a+b}{a}\frac{a+b-kb-b}{a+b} \\ &= \binom{a+b}{a}\frac{a-kb}{a+b} \\ &= \frac{a-kb}{a+b}\binom{a+b}{a}\end{aligned}$$

2.2.2 Proof By Induction

Let $N_k(a,b)$ denote the number of ways the $a+b$ ballots $(a \geq kb)$ can be ordered so that candidate A maintains more than k times as many votes as B throughout the counting of the ballots.
For the condition $b = 0$, $a + 0$ ballots $(a \geq k0)$ can only be ordered in one way such that candidate A maintains more than k times as many votes as B throughout the counting of the ballots. i.e

$$\begin{aligned} N_k(a,0) &= 1 \\ &= \frac{a-k(0)}{a+0}\binom{a+0}{a} \\ &= \frac{a}{a}\binom{a}{a} \end{aligned}$$

for all $a > 0$
For the condition $a = kb$, since candidate A maintains more than k times as many votes as B throughout the counting of the ballot $a > kb \Rightarrow kb > kb$ $(\rightarrow\leftarrow)$ i.e there is no way $a = kb > kb$ can be ordered such that candidate A maintains more than k times as many votes as B throughout the counting of the ballots. i.e

$$\begin{aligned} N_k(kb,b) &= 0 \\ &= \frac{kb-kb}{kb+kb}\binom{kb+kb}{kb} \\ &= \frac{0}{2kb}\binom{2kb}{kb} \end{aligned}$$

for all $b > 0$.
Hence the above two conditions satisfy $N_k(a,b) = \frac{a-kb}{a+b}\binom{a+b}{a}$.

Inductive assumption: assume it is true both when A receives $a-1$ votes and B receives b votes, and when A receives a votes and B receives $b-1$ votes,

i.e $N_k(a-1,b) = \frac{a-1-kb}{a+b-1}\binom{a+b-1}{b}$ and

$N_k(a, b-1) = \frac{a+k(b-1)}{a+b-1}\binom{a+b-1}{a}$

We must prove the formula is true for A receives a votes and B receives b votes.

For $b > 0$ and $a > kb$, consider the last vote in a ballot permutation. The number of ballots excluding the last vote becomes $a + b - 1$. Hence, A and B has votes either $a - 1$ and b or a and $b - 1$ respectively. To reach the total ballot $a + b$, the last vote should either be for A, meaning A will have a votes or the last vote should be for B meaning B will have b votes. Hence the total number of ways is the sum of $N_k(a - 1, b)$ and $N_k(a, b - 1)$ Therefore,

$$N_k(a, b) = N_k(a, b - 1) + N_k(a - 1, b)$$

by induction assumption

$$\begin{aligned}
N_k(a,b) &= \frac{a-k(b-1)}{a+b-1}\binom{a+b-1}{a} + \frac{a-1+kb}{a+b-1}\binom{a+b-1}{b} \\
&= \frac{a-k(b-1)}{a+b-1}\frac{(a+b-1)!}{a!(b-1)!} + \frac{a-1-kb}{a+b-1}\frac{(a+b-1)!}{b!(a-1)!} \\
&= \frac{a-kb+k}{a+b-1}\frac{b(a+b)(a+b-1)!}{b(a+b)a!(b-1)!} + \frac{a-1-kb}{a+b-1}\frac{a(a+b)(a+b-1)!}{a(a+b)b!(a-1)} \\
&= \frac{a-kb+k}{a+b-1}\frac{b(a+b)!}{(a+b)a!b!} + \frac{a-1-kb}{a+b-1}\frac{a(a+b)!}{a!b!} \\
&= \frac{(a+b)!}{a!b!}\left(\frac{(a-kb+k)b}{(a+b-1)(a+b)} + \frac{a(a-1+kb)}{(a+b-1)(a+b)}\right) \\
&= \frac{(a+b)!}{a!b!}\left(\frac{a^2-kab+ba-kb^2-a+kb}{(a+b-1)(a+b)}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \binom{a+b}{a}\frac{(a+b-1)(a-kb)}{(a+b-1)(a+b)} \qquad (2.2)\\
&= \binom{a+b}{a}\frac{a-kb}{a+b}\\
&= \frac{a-kb}{a+b}\binom{a+b}{a}
\end{aligned}
$$

2.2.3 Proof by the Cycle Lemma

Lemma 2.1. ***Cycle lemma:*** *For any sequence $P_1P_2...P_{m+n}$ of m A's and n B's, $m \geq kn$, there exist exactly $m-kn$ (out of $m+n$) cyclic permutations $P_iP_{i+1}...P_{m+n}P_1...P_{i-1}$, $1 \leq i \leq m+n$, that are good sequence.*

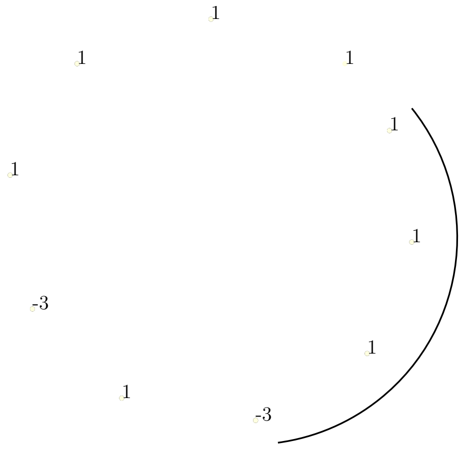
Proof. Arrange $m+n$ ballot permutation on a cycle. Removing a subsequence of k A's followed by one B from the cycle does not change the number of good permutations, since the $k+1$ ballot permutations have no net effect and no good permutation could have begun with any of the deleted ballot permutation. By the pigeon-hole principle: as long as $m \geq kn > 0$, there must be such a subsequence on the cycle; these subsequences removed one by one until only A's remain. The remaining $m-kn$ A's yield $m-kn$ good cyclic permutations.

Proof. (**Proof of ballot problem by cycle lemma**)
A ballot permutation is a sequence of $a+b$ terms where each term is either 1 or $-k$; votes for A correspond to the 1's and votes for B correspond to the $-k$'s.
A sequence is called good if every partial sum is positive, that means $a-kb > 0$ is good sequence and bad otherwise.
Let C be any circular arrangement of a 1's and b $-k$'s.
By cycle lemma: let C' be a circular arrangement created from C by removed k 1's and one $-k$. Then C' has $a' = a-k$ terms of 1's and $b' = b-1$ terms of $-k$'s. Thus $a' - kb' = a-k-k(b-1) = a-kb \geq 0$ so C' satisfies the

same hypotheses as C.

Let C" be a circular arrangement created from C' by removed k 1's and one $-k$. Then C" has $a'' = a' - k$ terms of 1's and $b'' = b' - 1$ terms of $-k$'s. Thus $a'' - kb'' = a' - k - k(b' - 1) = a' - kb' = a - kb \geq 0$ so C" satisfies the same hypotheses as C' and C. Thus, a term of C" and C' starts a good sequence if and only if the corresponding term in C starts a good sequence. Therefore, C", C' and C have exactly the same number of terms that start good sequences. Continuing in this manner, one removes the sequences until a circular arrangement consisting only of 1's remains. At this stage, there are $a - kb$ 1's, and every term starts a good sequence. Hence there are exactly $a - kb$ good sequences in C. If there is periodicity in C then not all $a - kb$ good sequences will be distinct that means there are duplicates to the extent that there is periodicity in the cycle. However, the ratio of good sequences to all sequences is $a - kb$ to $a + b$. Therefore, the number of good paths is $\frac{a-kb}{a+b}\binom{a+b}{a}$.

For any ballot sequence of a votes for A and b votes for B, exactly $a - kb$ of the $a + b$ cyclic permutations of the sequence are good. Consequently, a fraction $(a - kb)/(a + b)$ of all ballot permutations are good.



C

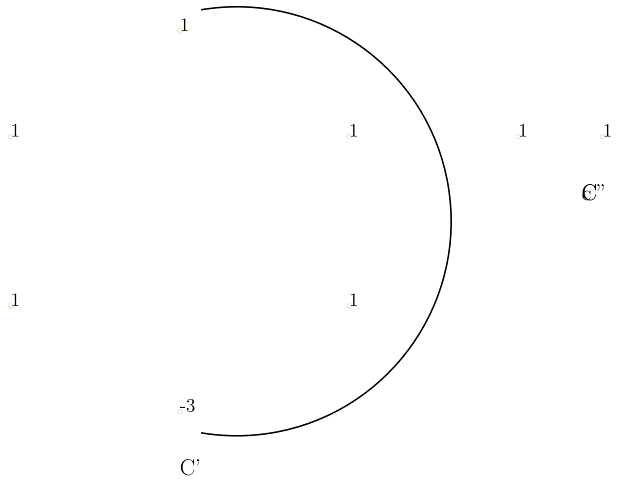


Figure 2.7: *Example of C, C’ and C” with* $a = 8$*,* $b = 2$*,* $k = 3$*.*

Thus the number of good sequence is $a - kb = 8 - 6 = 2$

2.3 Generalized Catalan numbers and Ballot problem

Theorem 2.1. *The generalized Catalan numbers:* $C_{(n,k)} = \frac{1}{kn+1}\binom{kn+n}{n} = P_{n,k}$.

Proof. Suppose that candidate A receives $m + 1$ votes and candidate B receives n votes, where $(m + 1) \geq kn$ for some positive integer k, and compute the number of ways the ballots can be ordered so that A always has at least k times as many votes as B throughout the counting of the ballots.
Requiring only that $m = kn$ produces the generalized Catalan numbers, also called the k-Catalan numbers. Since the ballots can be ordered A always has at least k times as many votes as B throughout the counting of the ballots, then, take the least $m = kn$.
The solution to the ballot problem in the above discussion is $\frac{m+1-kn}{m+1+n}\binom{m+1+n}{m+1}$.
$m = kn$,

$$
\begin{aligned}
\frac{m+1-kn}{m+1+n}\binom{m+1+n}{m+1} &= \frac{m+1-kn}{m+1+n}\frac{(m+1+n)!}{(m+1)!n!} \\
&= \frac{kn+1-kn}{kn+1+n}\frac{(kn+1+n)!}{(kn+1)!n!} \\
&= \frac{1}{kn+1+n}\frac{(kn+1+n)(kn+n)!}{(kn+1)(kn)!n!} \\
&= \frac{(kn+n)!}{(kn+1)(kn)!n!} \\
&= \frac{1}{kn+1}\binom{kn+n}{n}
\end{aligned}
$$

Hence, The Generalized Catalan Number is $C_{(n.k)} = P_{n,k} = \frac{1}{kn+1}\binom{kn+n}{n}$.

Bibliography

[1] T. Dolic, Perfect Matchings, Catalan Numbers, and Pascal's Triangle, Mathematics Magazine 80 (June 2007), 219226.

[2] I.P. Goulden and L.G.Serrano, Maintaining the spirit of the reflection principle when the boundary line has arbitrary integer slope, J. Comb. Theory, Ser. A 104 (2003) 317326.

[3] Hilton and J. Pedersen. (1991). Catalan numbers, their generalization and their uses, Math. Intelligencer 13, pp. 6475).

[4] P.A. MacMahon, Memoir on the theory of the partitions of numbers. part iv: on the probability that the successful candidate at an election by ballot may never at any time have fewer votes than the one who is unsuccessful; on a generalization of this question; and on its connexion with other questions of partition, permutation, and combination, Philosophical Transactions of the Royal Society of London, Series A 209 (1909) 153175. Also Collected Papers Vol. 1 (G.E. Andrews, ed.), MIT Press, Cambridge, Mass 1978, 12921314.

[5] Marc Renault, Four Proofs of the Bollot Theorem, Mathematics Magazine 80, December 2007, 345-352.

[6] M. S. Renault, Lost (and found) in translation: André's actual method and its application to the generalized ballot problem, Amer. Math. Shippensburg University, May 2007.

[7] A.D. Sands, On generalized Catalan numbers, Discr. Math. 21(2) (1978) 219-221.

[8] L. Takács, On the ballot theorems, Advances in Combinatorial Methods and Applications to Probability and Statistics, Birkhauser, 1997.

[9] J.H. Van Lint and R.M. Wilson, A Course in Combinatorics, Cambridge University Press, 2001.

[10] D. Zeilberger, André's reflection proof generalized to the many-candidate ballot problem, Discrete Mathematics 44 (1983) 325-326.

Appendix

Proposition 2.2. *For $n \geq 0$ and $k \geq 0$, then $kn+1$ and $(k-1)n+1$ are relative prime.*

Proof. By prime factorization,
Let $p_1, p_2, ..., p_r$ and $q_1, q_2, ..., q_t$ are prime numbers. And $\alpha 1, \alpha 2, ..., \alpha r$ and $\beta 1, \beta 2, ..., \beta t$ are positive integer.
Let $p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r}$ be the prime factorization of n where, $p_1, p_2, p_3, ..., p_r$ are arranged in ascending order.i.e $n = p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r}$ and
$q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta 3} ... q_t^{\beta t}$ be the prime factorization of k where, $q_1, q_2, q_3, ..., q_t$ are arranged in ascending order. i.e $k = q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta 3} ... q_t^{\beta t}$
Then,
$kn + 1 = q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta 3} ... q_t^{\beta t} p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r} + q_1^0$ since $1 = q_1^0 = p_1^0$
$= q_1^0 (q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta 3} ... q_t^{\beta t} p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r} + 1)$
similarly,
$(k-1)n + 1 = (q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta} ... q_t^{\beta t} - 1) p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r} + q_1^0$ since $1 = q_1^0 = p_1^0$
$= q_1^0 ((q_1^{\beta 1} q_2^{\beta 2} q_3^{\beta} ... q_t^{\beta t} - 1) p_1^{\alpha 1} p_2^{\alpha 2} p_3^{\alpha 3} ... p_r^{\alpha r} + 1)$ since $1 = q_1^0 = p_1^0$

q_i and p_j are not necessarily distinct where,$i = 1, ..., t$ and $j = 1, ...r$.
Therefore, $\gcd(kn+1, (k-1)n+1) = q_1^0 = p_1^0 = 1$.
implies $kn+1$and $(k-1)n+1$ are relative prime. ☐

Addis Ababa University
College Of Computational And Natural Sciences
School of Graduate Studies
Department of Mathematics

The undersigned hereby certify that they have read and recommend to the department of mathematics for acceptance of this project entitled “**Generalized Catalan numbers**” by Eyerusalem Aserssie in partial fulfillment of the requirements for the degree of master of science in mathematics.

Advisor: Dr. Zelealem Belaineh

Sign.________________ Date________________

Examiner 1: ________________

Sign.________________ Date________________

Examiner 2: ________________

Sign.________________ Date________________

By: Eyerusalem Aserssie

Sign.________________ Date________________