

December 2015

Generalized Catalan Numbers and Some Divisibility Properties

Jacob Bobrowski

University of Nevada, Las Vegas, bobrows2@unlv.nevada.edu

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>



Part of the [Mathematics Commons](#)

Repository Citation

Bobrowski, Jacob, "Generalized Catalan Numbers and Some Divisibility Properties" (2015). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 2518.

<https://digitalscholarship.unlv.edu/thesesdissertations/2518>

This Thesis is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

GENERALIZED CATALAN NUMBERS AND SOME DIVISIBILITY
PROPERTIES

By

Jacob Bobrowski

Bachelor of Arts - Mathematics
University of Nevada, Las Vegas
2013

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science - Mathematical Sciences

College of Sciences
Department of Mathematical Sciences
The Graduate College

University of Nevada, Las Vegas
December 2015

Thesis Approval

The Graduate College
The University of Nevada, Las Vegas

November 13, 2015

This thesis prepared by

Jacob Bobrowski

entitled

Generalized Catalan Numbers and Some Divisibility Properties

is approved in partial fulfillment of the requirements for the degree of

Master of Science – Mathematical Sciences
Department of Mathematical Sciences

Peter Shive, Ph.D.
Examination Committee Chair

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

Derrick DuBose, Ph.D.
Examination Committee Member

Arthur Baragar, Ph.D.
Examination Committee Member

David Beisecker, Ph.D.
Graduate College Faculty Representative

ABSTRACT

Generalized Catalan Numbers and Some Divisibility Properties

by

Jacob Bobrowski

Dr. Peter Shiue, Examination Committee Chair
Professor of Mathematical Sciences
University of Nevada, Las Vegas

I investigate the divisibility properties of generalized Catalan numbers by extending known results for ordinary Catalan numbers to their general case. First, I define the general Catalan numbers and provide a new derivation of a known formula. Second, I show several combinatorial representations of generalized Catalan numbers and survey bijections across these representation. Third, I extend several divisibility results proved by Koshy. Finally, I prove conditions under which sufficiently large primes form blocks of divisibility and indivisibility of the generalized Catalan numbers, extending a known result by Alter and Kubota.

ACKNOWLEDGEMENTS

I would like to thank each of my committee members, Dr. Derrick DuBose, Dr. Arthur Baragar, and Dr. Peter Shiue, as well as my fellow students and friends Minhwa Choi, Daniel Corral, and Katherine Yost.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
CHAPTER 1 GENERALIZED CATALAN NUMBERS AND LOBB'S PROOF .	1
CHAPTER 2 ALTERNATIVE REPRESENTATIONS OF GENERALIZED CATALAN NUMBERS	7
CHAPTER 3 GENERALIZING SOME RESULTS FROM KOSHY AND GAO .	12
CHAPTER 4 GENERALIZING A RESULT FROM ALTER AND KUBOTA ...	18
BIBLIOGRAPHY	28
CURRICULUM VITAE	29

CHAPTER 1

GENERALIZED CATALAN NUMBERS AND LOBB'S PROOF

The goal of this thesis is to describe a generalization of the Catalan numbers and some divisibility phenomenon exhibited by these numbers. Since the Catalan numbers appear in many contexts, see [9] for an extensive list, there are many ways one can choose to define them. One definition of the Catalan numbers is as an enumeration of certain sequences defined by Graham, Knuth, and Patashnik in [3]. We define the n th Catalan number, as the number of possible sequences of length $2n$ such that

1. Each terms of the sequence is either equal to 1 or equal to -1 .
2. Every partial sum of the sequence is nonnegative.
3. The total sum of the sequence is 0.

Conditions (1) and (3) ensure that the lengths of these sequences are always multiples of 2 as there must be exactly as many terms equal to 1 as are equal to -1 . One way to generalize the Catalan numbers is to alter this balance by applying a weight to one of the terms while keeping the other fixed. The generalization we consider fixes the terms equal to 1 and uses a parameter k to denote the difference between of the positive term and the negative term. Graham, Knuth, and Patashnik call such sequences k -Raney sequences, and describe the relationship between the

length of such sequences and the weight k . Suppose we have a sequence with m terms equal to 1 and n terms equal to $1 - k$. For such sequences to have a total sum of 0, we would need $m \cdot 1 + n \cdot (1 - k) = 0$, and hence $m + n = kn$. We can define a k -Raney sequence as any sequence whose length is a multiple of k and

1. Each terms of the sequence is either 1 or $1 - k$.
2. Every partial sum is nonnegative.
3. The total sum is 0.

We will define the n th k -Catalan number, denoted by $C_k(n)$, as the number of possible k -Raney sequences of length kn .

We will now seek a formula for the generalized Catalan numbers. In [8], Lobb establishes a proof for a common formula for the n th Catalan number using the 2-Raney definition. To do so, Lobb considers a different generalization of the Catalan numbers. He looks at sequences satisfying only conditions (1) and (2) and defines $L_{n,m}$ to be the number of such sequences of length $2n$ where $n + m$ of the terms equal 1 and $n - m$ of the terms equal -1 . These numbers satisfy the recurrence relation $L_{n+1,m} = L_{n,m+1} + 2L_{n,m} + L_{n,m-1}$, which makes short work of an inductive proof that $L_{n,m} = \frac{2m+1}{n+m+1} \binom{2n}{n+m}$. Since 2-Raney sequences are precisely those where the number of terms equal to 1 is the same as the number of terms equal to -1 , we have $C_n = L_{n,0} = \frac{1}{n+1} \binom{2n}{n}$.

Lobb's proof can be generalized for the n th k -Catalan number. To this end, we consider sequences that satisfy only conditions (1) and (2) of the definition of k -Raney

$n \backslash m$	0	1	2
1	1+1-2	1+1+1	
2	1+1+1+1-2-2 1+1+1-2+1-2 1+1-2+1+1-2	1+1+1+1+1-2 1+1+1+1-2+1 1+1+1-2+1+1 1+1-2+1+1+1	1+1+1+1+1+1
3	1+1+1+1+1+1-2-2-2 1+1+1+1+1-2-2+1-2 1+1+1+1+1-2+1-2-2 1+1+1+1-2-2+1+1-2 1+1+1+1-2+1+1-2-2 1+1+1-2+1-2+1+1-2 1+1+1-2+1-2+1-2 1+1-2+1+1+1-2-2 1+1-2+1+1-2+1-2 1+1-2+1+1-2+1+1-2	1+1+1+1+1+1+1-2-2 1+1+1+1+1+1-2+1-2 1+1+1+1+1-2+1+1-2 1+1+1+1-2+1+1+1-2 1+1+1-2+1+1+1+1-2 1+1-2+1+1+1+1-2 1+1+1+1+1-2-2+1 1+1+1+1+1-2+1-2+1 1+1+1+1-2+1+1-2+1 1+1+1-2+1+1+1-2+1 1+1-2+1+1+1+1-2+1 1+1+1+1+1-2-2+1+1 1+1+1+1-2+1-2+1+1 1+1+1-2+1+1-2+1+1 1+1-2+1+1-2+1+1 1+1-2+1+1+1-2+1+1 1+1+1+1-2-2+1+1+1 1+1+1-2+1-2+1+1+1 1+1-2+1+1-2+1+1+1	1+1+1+1+1+1+1+1-2 1+1+1+1+1+1+1-2+1 1+1+1+1+1+1-2+1+1 1+1+1+1+1-2+1+1+1 1+1+1+1+1+1+1+1-2 1+1+1+1+1+1+1-2+1 1+1+1+1+1+1-2+1+1 1+1+1+1+1-2+1+1+1 1+1+1+1-2+1+1+1+1 1+1+1-2+1+1+1+1+1 1+1-2+1+1+1+1+1+1

Table 1.1: Enumerations of some sequences counted by $L_{n,m}^3$

sequences. We let $L_{n,m}^k$ denote the number of such sequences with $(k-1)n+m$ terms equal to 1 and $n-m$ terms equal to $1-k$. See Table 1.1 for an enumeration of some of these sequences with $k=3$. Before we establish a formula for $L_{n,m}^k$, we will need to generalize Lobb's recurrence relation.

Theorem 1. *If $k \geq 2$, $n \geq 1$, and $n \geq m \geq 0$, then $L_{n+1,m}^k = \sum_{j=0}^k \binom{k}{j} L_{n,m+k-1-j}^k$.*

Proof. We consider an arbitrary sequence counted by $L_{n+1,m}^k$. Such a sequence has length $k(n+1)$ and $(k-1)(n+1)+m$ terms equal to 1. Let j denote the number of

terms equal to 1 that appear in the final k terms of the sequence. Then in the first kn terms, there are $(k-1)(n+1) + (m-j) = (k-1)n + (m+k-1-j)$ terms equal to 1. Since the first kn terms are all equal to either 1 or $1-k$ and have every partial sum nonnegative, they form a sequence counted by $L_{n,m+k-1-j}^k$. Now we observe that there are $\binom{k}{j}$ such arrangement where j of the final k terms equal 1. Summing over every applicable value for j , we arrive at the recurrence relation

$$L_{n+1,m}^k = \sum_{j=0}^k \binom{k}{j} L_{n,m+k-1-j}^k$$

with the understanding that $\binom{a}{b} = 0$ if $b < 0$ or $b > a$. □

We can now prove the generalization of Lobb's formula, but we find it convenient to use a different form for the induction.

Theorem 2. *If $k \geq 2$, $n \geq 1$, and $n \geq m \geq 0$, then $L_{n,m}^k = \binom{kn}{n-m} - (k-1)\binom{kn}{n-m-1}$.*

Proof. We prove the theorem by induction on n . First fix an integer $k \geq 2$.

Consider $L_{1,m}^k$. If $m \neq 0, 1$, then the formula yields 0, matching the number of sequences in this case. When $m = 0$, we have $k-1$ terms equal to 1 and one term equal $1-k$. Clearly, to have every partial sum positive, the term equal to $1-k$ must be in the last position, and hence $L_{1,0}^k = 1$. Our formula yields $\binom{k}{1} - (k-1)\binom{k}{0} = k - (k-1) = 1$. When $m = 1$, we have all k terms as '+1' and hence $L_{1,1}^k = 1$. Our formula yields $\binom{k}{0} - (k-1)\binom{k}{-1} = 1 - (k-1) \cdot 0 = 1$

Having established the base case, we fix an integer $N > 1$ and suppose that for

all $l < N$

$$L_{N,l}^k = \binom{kN}{N-l} - (k-1) \binom{kN}{N-l-1}$$

Let $t = k - 1 + m$. We fix an integer $m \leq n$ and consider

$$\begin{aligned} L_{N+1,m}^k &= \sum_{j=0}^k \binom{k}{j} L_{N,t-j}^k \\ &= \sum_{j=0}^k \binom{k}{j} \left[\binom{kN}{N-(t-j)} - (k-1) \binom{kN}{N-(t-j)-1} \right] \\ &= \sum_{j=0}^k \binom{k}{j} \left[\binom{kN}{kN-(N-(t-j))} - (k-1) \binom{kN}{kN-(N-(t-j)-1)} \right] \end{aligned}$$

Hence, $L_{N+1,m}^k$ is equal to

$$\sum_{j=0}^k \binom{k}{j} \binom{kN}{(k-1)(N+1)+m-j} - (k-1) \sum_{j=0}^k \binom{k}{j} \binom{kN}{(k-1)(N+1)+m-j+1}$$

Recalling the Vandermonde identity $\sum_{i=0}^a \binom{a}{i} \binom{b}{c-i} = \binom{a+b}{c}$

$$\begin{aligned} L_{N+1,m}^k &= \binom{kN+k}{(k-1)(N+1)+m} - (k-1) \binom{kN+k}{(k-1)(N+1)+m+1} \\ &= \binom{k(N+1)}{(N+1)-m} - (k-1) \binom{k(N+1)}{(N+1)-m-1} \end{aligned}$$

as desired. This completes the induction and the proof. \square

We note that when $m = 0$, the n th k -Catalan number is then given by

$$C_k(n) = L_{n,0}^k = \binom{kn}{n} - (k-1) \binom{kn}{n-1}.$$

n	$C_2(n)$	$C_3(n)$	$C_4(n)$	$C_5(n)$	$C_6(n)$	$C_7(n)$
1	1	1	1	1	1	1
2	2	3	4	5	6	7
3	5	12	22	35	51	70
4	14	55	140	285	506	819
5	42	273	969	2530	5481	10472
6	132	1428	7084	23751	62832	141778
7	429	7752	53820	231880	749398	1997688
8	1430	43263	420732	2330445	9203634	28989675
9	4862	246675	3362260	23950355	115607310	430321633
10	16796	1430715	27343888	250543370	1478314266	6503352856

Table 1.2: Initial values of k -Catalan numbers with $k = 2, 3, 4, 5, 6, 7$

The numbers $L_{n,m}^k$ have another form that is useful for the generalized Catalan numbers and matches the form taken by Lobb's formula. Observe that

$$\begin{aligned}
\binom{kn}{n-m} - (k-1)\binom{kn}{n-m-1} &= \binom{kn}{n-m} - \frac{(k-1)(n-m)}{kn-n+m}\binom{kn}{n-m} \\
&= \left[1 - \frac{(k-1)(n-m)}{kn-n+m}\right]\binom{kn}{n-m} \\
&= \frac{kn-n+m - (k-1)(n-m)}{kn-n+m}\binom{kn}{n-m} \\
&= \frac{km+1}{kn-n+m}\binom{kn}{n-m}
\end{aligned}$$

so that

$$L_{n,m}^k = \frac{km+1}{(k-1)n+m+1}\binom{kn}{n-m}$$

This gives the n th k -Catalan number the formula $C_k(n) = L_{n,0}^k = \frac{1}{(k-1)n+1}\binom{kn}{n}$. A table of values is presented in Table 1.2.

CHAPTER 2

ALTERNATIVE REPRESENTATIONS OF GENERALIZED CATALAN NUMBERS

The Catalan numbers are well known for counting Dyck paths, binary trees, and applications of binary operations. In this section, we will discuss how these objects are counted by Catalan numbers, and how they can be generalized so that they are counted by generalized Catalan Numbers.

In [2], Deutsche defines a Dyck Path as a series of $2n$ steps $(1, 1)$ or $(1, -1)$ such that the series starts at $(0, 0)$, ends at $(0, 2n)$, and never goes below the x -axis. Associated with each path is its Dyck word, the sequence of letters obtained by labeling each step of a path with U for an upward step and D for a downward step. Deutsche shows that the sequence of numbers of Dyck paths of $2n$ steps satisfy a well known recurrence relation for the 2-Catalan numbers, as well as its initial values, and so the two are the same. Another proof comes from an obvious bijection between Dyck words and 2-Raney sequences. Given a Dyck path of length $2n$, one can replace U with $+1$ and D with -1 to find its associated 2-Raney sequence. Since the Catalan numbers enumerate 2-Raney sequences, they also enumerate Dyck paths. For an enumeration of all Dyck paths with six steps and their corresponding Dyck words and 2-Raney sequences, see Table 2.1.


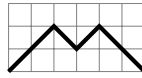
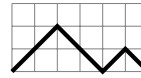

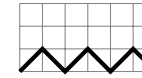
				
$U U U D D D$	$U U D U D D$	$U U D D U D$	$U D U U D D$	$U D U D U D$
$1+1+1-1-1-1$	$1+1-1+1-1-1$	$1+1-1-1+1-1$	$1-1+1+1-1-1$	$1-1+1-1+1$

Table 2.1: Enumeration of all Dyck paths of length 6

The generalization of 2-Raney sequences we considered applied a weight to the subtraction terms in the sequences. In [4], Heubach, Li, and Mansour define a corresponding generalization to Dyck paths that can be slightly altered to apply a weight to the downward step. These paths have a series of steps that start at $(0, 0)$, end at $(kn, 0)$, have each step as either $(1, 1)$ or $(1, 1 - k)$, and no step goes below the x -axis. They are enumerated by the k -Catalan numbers. This is clearly true by reasoning in the same way as with the Catalan number case. We label the steps with U for upward steps and D for downward steps, then replace U with $+1$ and D with $+(1 - k)$.

Having no steps go below the x -axis corresponds to having all of the partial sums nonnegative and having the path end on the y -axis corresponds to having the total sum equal 0. For an enumeration of such paths of length 3 for $k = 3$, see Table 2.2.

We now discuss the application of a k -ary operation. The number of objects required for applying a k -ary operation n times can be determined inductively. For a single application, we require k objects and to form a new expression of $n + 1$ applications from n applications, we must replace one object with k objects. Thus, for n applications of a k -ary operation, we require $(k - 1)n + 1$ objects. In [5], Hilton and Pedersen describe the following bijection. Given an expression of n applications

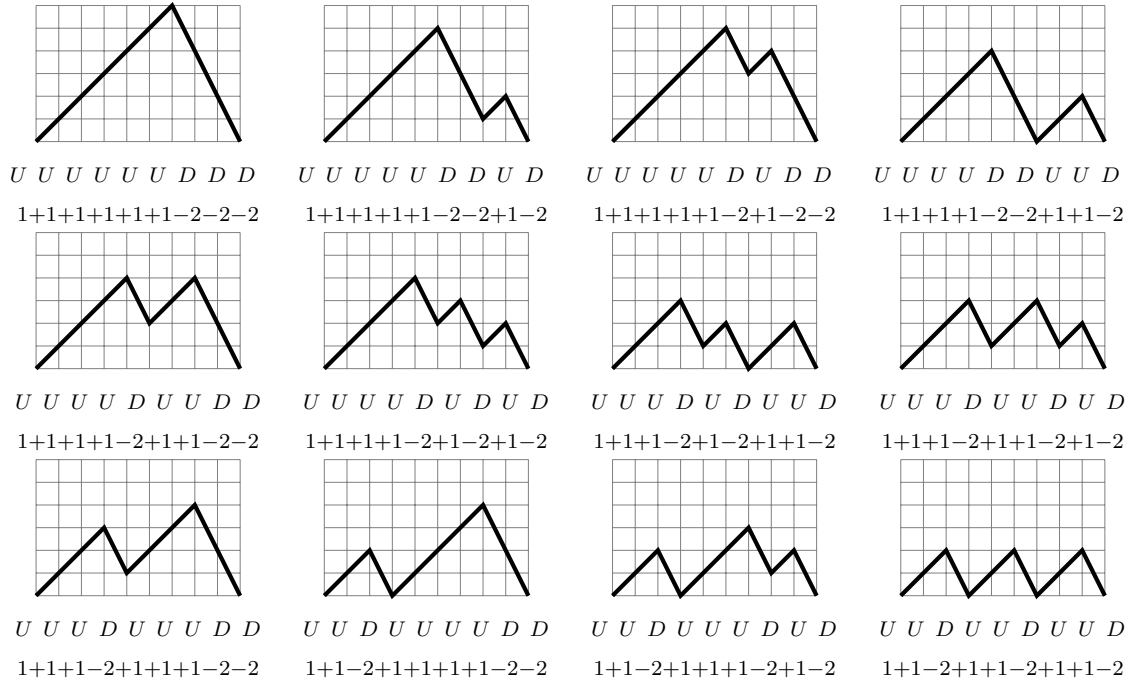


Table 2.2: Enumeration of generalized Dyck paths of length 3 with $k=3$

of a k -ary operation, delete all left parentheses and reverse the expression. Then we replace every object with a 1 and every parenthesis with a -1 . The resulting sequence is a k -Raney sequence.

Finally, we consider binary trees. Recall that a full k -ary tree is a tree such that every node has exactly k or exactly zero leaves. We call the nodes of such a tree with k leaves a source node and the nodes with zero leaves end nodes. Further, we recall that if a k -ary tree has n source nodes, then it has $(k-1)n+1$ end nodes and $kn+1$ total nodes. It is well known that the n th 2-Catalan counts the number of binary trees with n source nodes. We again look to [5], where Hilton and Pedersen suggest a bijection from the set of k -ary trees with n source nodes to the set

of possible applications of a k -ary operation on n objections. We make this bijection explicit here.

Given a full rooted k -ary tree with n source nodes, we construct an expression in polish notation of n applications of a k -ary operation. Beginning with an object y corresponding to the root. If the root is a source node, we order its children $1, 2, \dots, k$ and replace x with the expression $K(x_1, x_2, \dots, x_k)$ where $x_i = x$ if the i th child is an end node and $x_i = y$ if the i th child is a source node. After the first stage, if any y exists we replace it in the same way. We order its children $1, 2, \dots, k$ and replace y with the expression $K(x_1, x_2, \dots, x_k)$ where $x_i = x$ if the i th child of y is an end node and $x_i = y$ if the i th child of y is a source node. We then repeat this process until every y is replaced.

In this chapter, we have shown that the set of generalized Dyck paths of length kn with weight k , the set of possible applications of a k -ary operation on $(k-1)n+1$ objects, the set of k -ary trees with n source nodes, and the set of k -Raney sequences of length kn all have the same size, and hence are all counted by the n th k -Catalan number. This was achieved by demonstrating bijections across these set. In Figure 2.1, we demonstrate the corresponding representations of one one element in each set.

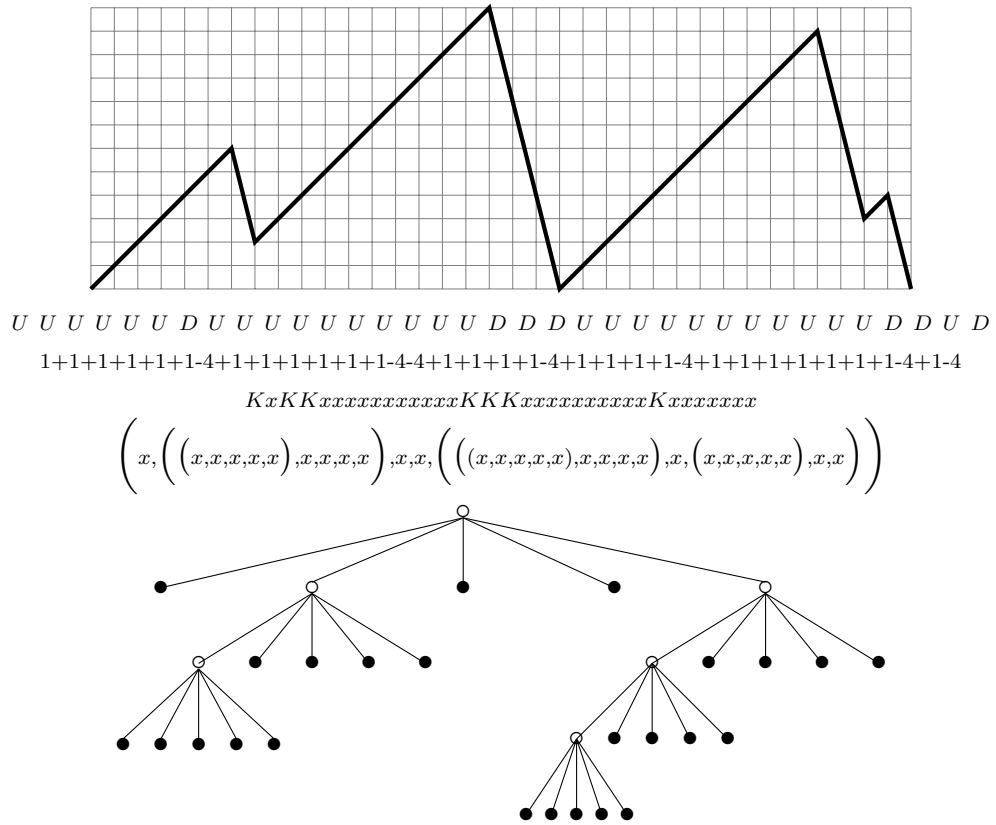


Figure 2.1: An example of a generalized Dyck path of length 7 with $k = 5$ and its corresponding representations

CHAPTER 3

GENERALIZING SOME RESULTS FROM KOSHY AND GAO

We now use the generalized Lobb numbers to prove some divisibility results about the general Catalan numbers. In [6], Koshy and Gao use Lobb's formula for $L_{n,m}$ to uncover some divisibility properties of the 2-Catalan Numbers. Using methods we will generalize shortly, they find that $L_{3t+1,1}^2 = \frac{3t+1}{t+1}C_2(3t+1)$, suggesting that $t+1$ divides $C_2(3t+1)$. Koshy and Gao later use a recurrence relation to prove this result, but their earlier approach allows some conclusions in the general case. Following Koshy and Gao's approach, we observe

$$\begin{aligned}
 L_{n,m}^k &= \frac{km+1}{(k-1)n+m+1} \binom{kn}{n-m} \\
 &= (km+1) \frac{(kn)!}{[n-m]! [(k-1)n+m+1]!} \\
 &= (km+1) \frac{(kn)!}{n! [(k-1)n+1]!} \prod_{j=1}^m \frac{n-j+1}{[(k-1)n+j+1]} \\
 &= (km+1) C_k(n) \prod_{j=1}^m \frac{n-j+1}{[(k-1)n+j+1]}
 \end{aligned}$$

Using this identity, we observe that

$$\begin{aligned}
L_{(k+1)t+1,1}^k &= C_k((k+1)t+1) \frac{(k+1)[(k+1)t+1]}{(k-1)[(k+1)t+1]+2} \\
&= C_k((k+1)t+1) \frac{(k+1)[(k+1)t+1]}{(k-1)(k+1)t+(k-1)+2} \\
&= C_k((k+1)t+1) \frac{(k+1)[(k+1)t+1]}{(k-1)(k+1)t+(k+1)} \\
&= C_k((k+1)t+1) \frac{(k+1)t+1}{(k-1)t+1}
\end{aligned}$$

Substituting $k = 2$ yields Koshy and Gao's original identity. Since the LHS is an integer, $(k-1)t+1$ must divide $C_k((k+1)t+1)[(k+1)t+1]$. We observe that $(k+1)t+1 = (k-1)t+1 + 2t$, and hence $\gcd((k+1)t+1, (k-1)t+1) = \gcd(2t, (k-1)t+1)$. Clearly any divisor of t won't be a divisor of $(k-1)t+1$, so the gcd must be either 1 or 2. This is sufficient to show that for any integers $k \geq 2$ and $t \geq 0$, $(k-1)t+1$ divides $2C_k((k+1)t+1)$. For the case $k = 2$, Koshy is able to drop the factor of 2, but the factor is required in the general case. For instance, when $k = 6$ and $t = 15$, we have $(k-1)t+1 = 76$ and $(k+1)t+1 = 91$, but $C_6(106)$ is not divisible by 76. These examples however are rare. For values of k and t between 1 and 100, for instance, there are only 48 cases where $(k-1)t+1$ does not divide $C_k((k+1)t+1)$.

We'll now discuss some further divisibility results. We first establish some recur-

rence relations for k -Catalan numbers. For $n \geq 2$:

$$\begin{aligned}
\frac{C_k(n)}{C_k(n-1)} &= \frac{[kn]!}{n![(k-1)n+1]!} \frac{(n-1)![(k-1)(n-1)+1]!}{[k(n-1)]!} \\
&= \frac{(n-1)!}{n!} \cdot \frac{[(k-1)(n-1)+1]!}{[(k-1)n+1]!} \cdot \frac{(kn)!}{[k(n-1)]!} \\
&= \frac{(n-1)!}{n!} \cdot \frac{[(k-1)(n-1)+1]!}{[(k-1)(n-1)+k]!} \cdot \frac{(kn)!}{(kn-k)!} \\
&= \frac{1}{n} \prod_{j=2}^k \frac{1}{[(k-1)(n-1)+j]} \prod_{j=0}^{k-1} (kn-j) \\
&= \frac{kn}{n} \prod_{j=1}^{k-1} \frac{1}{[(k-1)(n-1)+j+1]} \prod_{j=1}^{k-1} (kn-j) \\
&= k \prod_{j=1}^{k-1} \frac{kn-j}{[(k-1)(n-1)+j+1]}
\end{aligned}$$

We will refer to this relation as the general recurrence relation. Substituting $k = 2$ yields the well known relation

$$C_2(n) = \frac{4n-2}{n+1} C_2(n-1)$$

In [7], Koshy and Gao use this recurrence relation to prove a divisibility result of the 2-Catalan numbers with Mersenne number subscript. We recall that the j th Mersenne number is $M_j = 2^j - 1$, and that the sequence of these numbers satisfy the recurrence relation $M_{j+1} = 2M_j + 1$. Koshy and Gao substitute $n = M_{j+1}$ in the above recurrence relation to get

$$C_2(M_{j+1}) = \frac{4M_{j+1}-2}{M_{j+1}+1} C_2(M_{j+1}-1) = \frac{2^{j+2}-3}{2^j} C_2(2M_j).$$

Since $2^{j+2} - 3$ and 2^j are coprime, Koshy and Gao conclude that $2^j \mid C_2(2M_{k+1})$ and $2^{j+2} - 3 \mid C_2(M_{j+1})$.

We consider similar results for $k = 3$. Using the general recurrence relation in this case, we get

$$C_3(n) = \frac{3n(3n-1)(3n-2)}{n(2n+1)(2n)} C_3(n-1) = \frac{3(3n-1)(3n-2)}{2n(2n+1)} C_3(n-1)$$

We make the same substitution of $n = M_j$ to obtain

$$C_3(M_j) = \frac{3(3M_j-1)(3M_j-2)}{2M_j(2M_j+1)} C_3(M_j-1) = \frac{3(3M_j-1)(3M_j-2)}{2M_j M_{j+1}} C_3(M_j-1)$$

Whereas it was obvious that 2^j and $2^{j+2} - 3$ are coprime, it is unclear, and even false for some j , that $3(3M_j-1)(3M_j-2)$ and $2M_j M_{j+1}$ are coprime. To make a similar conclusion as were to made by Koshy and Gao, we seek conditions under which $\gcd(M_j M_{j+1}, (3M_j-1)(3M_j-2)) = 1$.

Since M_j and M_{j+1} are coprime, any divisor of $M_j M_{j+1}$ is a divisor of exactly one of M_j and M_{j+1} , and since the only common divisor of M_j and $(3M_j-1)(3M_j-2)$ is 1, we need only consider $\gcd(M_{j+1}, (3M_j-1)(3M_j-2))$. We write $M_{j+1} = 2M_j + 1$ and consider $\gcd(2M_j + 1, 3M_j - 1)$ and $\gcd(2M_j + 1, 3M_j - 2)$ separately.

First, let $d = \gcd(2M_j + 1, 3M_j - 1)$. If we let $2M_j + 1 = kd$ and $3M_j - 1 = ld$, then $6M_j + 3 = 3kd$ and $6M_j - 2 = 2ld$. Subtracting we get $5 = (3k - 2l)d$, and hence $d = 1$ or $d = 5$.

Now let $d = \gcd(2M_n + 1, 3M_n - 2)$. If we let $2M_j + 1 = kd$ and $3M_j - 2 = ld$

then $6M_j + 3 = 3kd$ and $6M_j - 4 = 2ld$. Subtracting we get $7 = (3k - 2l)d$, and hence $d = 1$ or $d = 7$.

Thus, in order to ensure $\gcd(M_j M_{j+1}, (3M_j - 1)(3M_j - 2)) = 1$, we need conditions on j to ensure that neither 5 nor 7 divide M_j .

Note that if $j = 4t + k$, then

$$M_{j+1} \equiv 2^{j+1} - 1 \equiv (2^4)^t 2^{k+1} - 1 \equiv 16^t 2^{k+1} - 1 \equiv 2^{k+1} - 1 \pmod{5}$$

By considering the cases of $k = 0, 1, 2$, or 3 , we see the only value of k for which $M_{j+1} \equiv_5 0$ is 3 .

Similarly, if $j = 3t + k$, then

$$M_{j+1} \equiv 2^{j+1} - 1 \equiv (2^3)^t 2^{k+1} - 1 \equiv 8^t 2^{k+1} - 1 \equiv 2^{k+1} - 1 \pmod{7}$$

By considering the cases of $k = 0, 1$, or 2 , we see the only value of k for which $M_{j+1} \equiv_7 0$ is 2 .

From the above two cases, we have:

Theorem 3. *If $j \not\equiv 3 \pmod{4}$ and $j \not\equiv 1 \pmod{3}$, then $M_j M_{j+1}$ divides $C_3(M_j - 1)$ and $(3M_j - 1)(3M_j - 2)$ divides $C_3(M_j)$.*

Proof. We know that under these conditions, $\gcd(M_j M_{j+1}, (3M_j - 1)(3M_j - 2)) = 1$. The theorem follows then from the equation $2M_j M_{j+1} C_3(M_j) = 3(3M_j - 1)(3M_j - 2) C_3(M_j - 1)$ □

I would like to thank Dr. Peter Shiue for suggesting this theorem.

CHAPTER 4

GENERALIZING A RESULT FROM ALTER AND KUBOTA

The 2-Catalan numbers have an interesting divisibility property described by Alter and Kubota in [1]. When a prime $p > 3$ divides a 2-Catalan number $C_2(n-1)$, but fails to divide $C_2(n)$, it fails to divide every successive 2-Catalan number until $C_2(n + \frac{p+3}{2})$. Further, in this case p divides $n + 1$. To prove these results, Alter and Kubota look at the recurrence relation

$$C_2(n) = \frac{4n - 2}{n + 1} C_2(n - 1)$$

that we derived in Chapter 3. If p divides $C_2(n - 1)$, but not $C_2(n)$, the $n + 1$ in the denominator must cancel a factor of p in $C_2(n - 1)$, and hence $n + 1$ must have a factor of p . Once we know $n + 1$ is divisible by p , we repeat applications of the recurrence relation to see

$$\begin{aligned} C_2(n + 1) &= \frac{4n + 2}{n + 2} C_2(n) \\ C_2(n + 2) &= \frac{4n + 6}{n + 3} \frac{4n + 2}{n + 2} C_2(n) \\ &\vdots \\ C_2(n + l) &= \prod_{j=1}^l \frac{4(n + j) - 2}{n + j + 1} C_2(n) \end{aligned}$$

Alter and Kubota observe that

$$4(n + j) - 2 \equiv_p 4(n + 1) + 4j - 6 \equiv_p 4j - 6$$

so that p divides $4(n + j) - 2$ if and only if $j \equiv_p \frac{p+3}{2}$, and hence p does not divide the numerator until $l = \frac{p+3}{2}$. Using the facts that $\frac{p+3}{2} < p$ when $p > 3$ and that p divides $n + 1$, we know that no term appearing in the denominator has a factor of p , and hence $C_2(n + \frac{p+3}{2})$ is divisible by p .

We here seek similar results for the general Catalan numbers. A computer search of values for k , p , and n shows that the result does not generalize fully. For example, 17 divides $C_5(12)$ and does not divide $C_5(13)$, but does divide $C_5(14)$. An analysis of the divisibility of general Catalan numbers by sufficiently large primes does, however, yield blocks of indivisibility of consistent length. The first such block in each sequence $\{C_k(n)\}_{n \geq 0}$ is formed by its initial terms.

Using formula the formula for the k -Catalan numbers derived in Chapter 1, we can write $C_k(n) = \frac{(kn)!}{n![(k-1)n+1]}$. Clearly, for any prime p larger than k , if $n \leq \lfloor \frac{p}{k} \rfloor$, then $kn < p$. Thus p cannot divide $(kn)!$, and hence cannot divide $C_k(n)$. Further, if $p > k^2$, then p does divide $C_k(\lfloor \frac{p}{k} \rfloor + 1)$. To see this, we divide p by k and find q, r such that $0 < r < k$ and $p = qk + r$. We calculate

$$C_k(\lfloor \frac{p}{k} \rfloor + 1) = C_k(q+1) = \frac{[k(q+1)]!}{(q+1)!} [(k-1)(q+1) + 1]! = \frac{1}{(q+1)!} \prod_{j=0}^{q-1} [k(q+1) - j]$$

Since $p > k^2$ and $k \nmid p$, we know $k \leq q$ and $r \geq 1$, and hence $k - r \leq q - 1$. We

observe first that the term in the product corresponding to the index $j = k - r$ is $k(q + 1) - (k - r) = kq + r = p$ and second that $p \nmid (q + 1)!$, and conclude that $p \mid C_k(\lfloor \frac{p}{k} \rfloor + 1)$. We will now find conditions under which block of length $\lfloor \frac{p}{k} \rfloor + 1$ that are indivisibility by a prime appear later in the sequence. To identify a block, we need to know first that a given prime divides $C_k(n)$ for some initial n , then that the prime does not divide the following $\lfloor \frac{p}{k} \rfloor + 1$ terms, and finally that the prime divides $C_k(n + \lfloor \frac{p}{k} \rfloor + 1)$. Our first result describes the starting place of possible blocks.

Theorem 4. *If $k \geq 3$ and p is a prime, then $p \mid C_k(mp - 1)$ for any $m > 0$.*

Proof. Starting with $C_k(mp - 1) = \frac{1}{(mp-1)!} \prod_{j=2}^{mp-1} [(k-1)(mp-1) + j]$ we first identify the factors of the numerator and of the denominator that are multiples of p . Clearly there are exactly $m - 1$ occurrences of p in the denominator corresponding to tp for $t = 1, 2, \dots, m - 1$.

For the numerator, first fix an index integer j and consider the corresponding term in the product. Since $(k - 1)(mp - 1) + j = (k - 1)mp - (k - 1) + j$, we have that

$$p \mid (k - 1)(mp - 1) + j \iff j \equiv (k - 1) \pmod{p}$$

We observe that $2 \leq (k - 1) + tp \leq mp - 1$ for $t = 0, 1, \dots, m - 1$, and conclude that there are m occurrences of p in the numerator corresponding to terms of the form $(k - 1)mp + tp = [(k - 1)m + t]p$.

We now separate the terms that are divisible by p from the terms that aren't.

Define $D = \frac{(mp-1)!}{\prod_{t=1}^{m-1} tp}$ and $E = \frac{\prod_{j=2}^{mp-1} [(k-1)(mp-1)+j]}{\prod_{t=0}^{m-1} [(k-1)m+t]p}$ so that

$$C_k(mp-1) = \frac{E}{D} \cdot \frac{\prod_{t=0}^{m-1} [(k-1)m+t]p}{\prod_{t=1}^{m-1} tp} = \frac{E}{D} \cdot \frac{\prod_{t=0}^{m-1} [(k-1)m+t]}{\prod_{t=1}^{m-1} t}$$

where neither E nor D are divisible by p . Recall that the p -adic valuation of an integer a is defined by $\nu_p(a) = \max\{j : p^j \mid a\}$. We can now determine if $C_k(mp-1)$ is divisible by p by comparing the p -adic valuation of $\prod_{t=0}^{m-1} [(k-1)m+t]$ and $\prod_{t=1}^{m-1} t$. Recall further that by Legendre's Formula, $\nu_p(a!) = \sum_j \lfloor \frac{a}{p^j} \rfloor$.

Divide m by p^j for all $j \geq 1$ to obtain q_j, r_j such that $m = q_j p^j + r_j$ where $0 \leq r_j \leq p^j$. Consider

$$\begin{aligned} \nu_p\left(C_k(mp-1)\right) &= \nu_p\left(\prod_{t=0}^{m-1} [(k-1)m+t]\right) - \nu_p\left(\prod_{t=1}^{m-1} t\right) \\ &= 1 + \nu_p\left(\frac{(km-1)!}{[(k-1)m]!}\right) - \nu_p((m-1)!) \\ &= 1 + \nu_p((km-1)!) - \nu_p((m-1)!) - \nu_p([(k-1)m]!) \\ &= 1 + \sum_{j=1} \left\lfloor \frac{km-1}{p^j} \right\rfloor - \sum_{j=1} \left\lfloor \frac{m-1}{p^j} \right\rfloor - \sum_{j=1} \left\lfloor \frac{(k-1)m-1}{p^j} \right\rfloor \\ &= 1 + \sum_{j=1} \left[\left\lfloor \frac{km-1}{p^j} \right\rfloor - \left\lfloor \frac{m-1}{p^j} \right\rfloor - \left\lfloor \frac{(k-1)m-1}{p^j} \right\rfloor \right] \end{aligned}$$

Considering each term in the sum, we have

$$\begin{aligned} \lfloor \frac{km-1}{p^j} \rfloor &= \lfloor \frac{k(q_j p^j + r_j) - 1}{p^j} \rfloor = \lfloor kq_j + \frac{kr_j - 1}{p^j} \rfloor \\ \lfloor \frac{m-1}{p^j} \rfloor &= \lfloor \frac{(q_j p^j + r_j) - 1}{p^j} \rfloor = \lfloor q_j + \frac{r_j - 1}{p^j} \rfloor \\ \lfloor \frac{(k-1)m-1}{p^j} \rfloor &= \lfloor \frac{(k-1)(q_j p^j + r_j) - 1}{p^j} \rfloor = \lfloor (k-1)q_j + \frac{(k-1)r_j - 1}{p^j} \rfloor \end{aligned}$$

so that the sum is

$$\sum_{j=1} \left[kq_j - q_j - (k-1)q_j + \lfloor \frac{kr_j - 1}{p^j} \rfloor - \lfloor \frac{r_j - 1}{p^j} \rfloor - \lfloor \frac{(k-1)r_j - 1}{p^j} \rfloor \right],$$

and hence

$$\nu_p(C_k(mp-1)) = 1 + \sum_{j=1} \left[\lfloor \frac{kr_j - 1}{p^j} \rfloor - \lfloor \frac{r_j - 1}{p^j} \rfloor - \lfloor \frac{(k-1)r_j - 1}{p^j} \rfloor \right].$$

This sum is clearly nonnegative since $\frac{kr_j-1}{p^j} \geq \frac{(k-1)r_j-1}{p^j}$ and $p^j \geq r_j$, so the p -adic valuation of $C_k(mp-1)$ is positive, and hence $p \mid C_k(mp-1)$. \square

We can now prove that after mp th term, the next $\lfloor \frac{p}{k} \rfloor + 1$ terms are all divisible by p or all not divisible by p . In fact, each of these terms has the same p -adic valuation.

Theorem 5. *If p is a prime greater than k^2 , then $\nu_p(C_k(mp+l)) = \nu_p(C_k(mp))$ for all l such that $1 \leq l \leq \lfloor \frac{p}{k} \rfloor$*

Proof. It is sufficient to show that $\nu_p(C_k(mp+l)) = \nu_p(C_k(mp+l-1))$ for any l such that $1 \leq l \leq \lfloor \frac{p}{k} \rfloor$. Substituting $n = mp+l$ in the general recurrence relation, we get

$$\frac{C_k(mp+l)}{C_k(mp+l-1)} = k \prod_{j=1}^{k-1} \frac{k(mp+l)-j}{(k-1)(mp+l-1)+j+1}$$

We show that p does not divide $\prod_{j=1}^{k-1} [k(mp+l)-j]$ or $\prod_{j=1}^k (k-1)(mp+l-1)+j+1$. Pick any j such that $1 \leq j \leq k-1$. We note $k(mp+l)-j \equiv_p kl-j$, but since $l \leq \lfloor \frac{p}{k} \rfloor$ and $j \geq 1$, we know $kl-j \leq k \lfloor \frac{p}{k} \rfloor - 1 < p-1 < p$. Thus $p \nmid k(mp+l)-j$. Similarly, $(k-1)(mp+l-1)+j+1 \equiv_p (k-1)(l-1)+j+1$, but $(k-1)(l-1)+j+1 < kl-k+k < kl < p$. Thus $p \nmid (k-1)(mp+l-1)+j+1$. Since no factor of p is in the numerator or denominator of the RHS of the equation, we know $\nu_p(C_k(mp+l)) = \nu_p(C_k(mp+l-1))$. \square

Using Lucas' Theorem, we can show that the terms $C_k(mp+l)$ where $0 \leq l \leq \lfloor \frac{p}{k} \rfloor$ are all related to the initial terms $C_k(l)$. Recall that Lucas' Theorem states that for a prime p and integers a, b such that $a = \sum_{i=0}^t a_i p^i$ and $b = \sum_{i=0}^t b_i p^i$ with $0 \leq a_i, b_i < p$, we have that $\binom{a}{b} \equiv_p \prod_{i=0}^t \binom{a_i}{b_i}$.

Theorem 6. For $l : 0 \leq l \leq \lfloor \frac{p}{k} \rfloor$, $C_k(mp+l) \equiv_p \binom{km}{m} C_k(l)$.

Proof. We will use Lucas' Theorem in two ways. First, we will prove that $\binom{k(mp+l)}{mp+l} \equiv_p \binom{kmp}{mp} \binom{kl}{l}$.

Consider $\binom{k(mp+l)}{mp+l} = \binom{kmp+kl}{mp+l}$. We write the base- p expansion of $kmp+kl$ as $\sum_{i=0}^t c_i p^i$ and of $mp+l$ as $\sum_{i=0}^t d_i p^i$. Since $l \leq \lfloor \frac{p}{k} \rfloor$, we know $l < kl < p$, and hence $c_0 = kl$ and $d_0 = l$.

Note then that the base- p expansions of kmp and mp are $\sum_{i=1}^t c_i p^i$ and $\sum_{i=1}^t d_i p^i$.

Thus, by Lucas's Theorem,

$$\binom{k(mp+l)}{mp+l} \equiv_p \prod_{i=0}^t \binom{c_i}{d_i} = \binom{c_0}{d_0} \prod_{i=1}^t \binom{c_i}{d_i} \equiv_p \binom{kl}{l} \binom{kmp}{mp}$$

Using this, we conclude

$$C_k(mp+l) \equiv_p \frac{1}{(k-1)(mp+l)+1} \binom{k(mp+l)}{mp+l} \equiv_p \frac{1}{(k-1)l+1} \binom{kl}{l} \binom{kmp}{mp}$$

Our second use of the theorem will be to prove that $\binom{kmp}{mp} \equiv_p \binom{km}{m}$. Using the expansion above,

$$\binom{kmp}{mp} \equiv \prod_{i=0}^t \binom{c_i}{d_i} = \binom{0}{0} \prod_{i=1}^t \binom{c_i}{d_i} \equiv \binom{km}{m}$$

Using this and the above congruence, we conclude

$$C_k(mp+l) \equiv_p \binom{km}{m} C_k(l).$$

□

So far, we have found a candidate for the initial term of a block of indivisibility and we have shown that the terms following this initial term all divisible or indivisible by a prime together. We now need conditions under which they are all indivisible.

We consider a prime p such that $p > k^2$. Substituting $n = mp$ in the general recurrence relation, we have

$$\frac{C_k(mp)}{C_k(mp-1)} = k \prod_{j=1}^{k-1} \frac{km p - j}{(k-1)(mp-1) + j + 1}$$

Since p does not divide k or $\prod_{j=1}^{k-1} (km p - j)$, we have

$$\nu_p(C_k(mp)) = \nu_p(C_k(mp-1)) - \nu_p\left(\prod_{j=1}^{k-1} (k-1)(mp-1) + j + 1\right)$$

The term in $\prod_{j=1}^{k-1} (k-1)(mp-1) + j + 1$ corresponding to the index $j = k-2$ is $(k-1)(mp-1) + (k-2) + 1 = (k-1)mp$, so since the terms are consecutive and there are fewer than p terms, this is the only term divisible by p . Further, $k-1$ and p are coprime, so $\nu_p\left(\prod_{j=1}^{k-1} (k-1)(mp-1) + j + 1\right) = \nu_p(m) + 1$. Hence, $\nu_p(C_k(mp)) = \nu_p(C_k(mp-1)) - \nu_p(m) - 1$. By applying theorem 5, we arrive at the following equation. If p is a prime such that $p > k^2$ and l is an integer such that $0 \geq l \geq \lfloor \frac{p}{k} \rfloor$, then

$$\nu_p(C_k(mp+l)) = \nu_p(C_k(mp-1)) - \nu_p(m) - 1$$

We can now state the main theorem.

Theorem 7. *If $k \geq 3$ and p is a prime such that $p > k^2$, then for any $m > 0$ such that $\nu_p(m) \geq \nu_p(C_k(mp-1)) - 1$:*

1. $p \mid C_k(mp-1)$
2. $p \nmid C_k(mp+l)$ for any l such that $0 \leq l \leq \lfloor \frac{p}{k} \rfloor$

$$3. p \mid C_k(mp + \lfloor \frac{p}{k} \rfloor + 1)$$

Proof. (1) has already been proven, and (2) is obvious from the above equation. That leaves (3). We write

$$\frac{C_k(mp + \lfloor \frac{p}{k} \rfloor + 1)}{C_k(mp + \lfloor \frac{p}{k} \rfloor)} = k \prod_{j=1}^{k-1} \frac{k(mp + \lfloor \frac{p}{k} \rfloor + 1) - j}{(k-1)(mp + \lfloor \frac{p}{k} \rfloor) + j + 1}$$

Since k is not divisible by p , we have

$$\nu_p(C_k(mp + \lfloor \frac{p}{k} \rfloor + 1)) = \nu_p\left(\prod_{j=1}^k (k(mp + \lfloor \frac{p}{k} \rfloor + 1) - j)\right) - \nu_p\left(\prod_{j=1}^{k-1} [(k-1)(mp + \lfloor \frac{p}{k} \rfloor) + j + 1]\right)$$

We divide p by k to get q and r such that $p = qk + r$ where $0 \leq r < k$. Now consider

$\prod_{j=1}^k (k(mp + \lfloor \frac{p}{k} \rfloor + 1) - j)$. We observe

$$k(mp + \lfloor \frac{p}{k} \rfloor + 1) - j \equiv_p k \lfloor \frac{p}{k} \rfloor + k - j = kq + k - j = p + k - j - r \equiv_p k - j - r$$

Thus we find that p divides a term in this product if and only if $j = k - r$. Since p is prime, and r is the remainder in the division of p by k , we know $1 \leq k - r \leq k - 1$.

Thus there is one term in the this product divisible by p , and hence $\nu_p\left(\prod_{j=1}^k (k(mp + \lfloor \frac{p}{k} \rfloor + 1) - j)\right)$ is positive.

Now consider $\prod_{j=1}^{k-1} [(k-1)(mp + \lfloor \frac{p}{k} \rfloor) + j + 1]$. We observe

$$(k-1)(mp + \lfloor \frac{p}{k} \rfloor) \equiv_p (k-1) \lfloor \frac{p}{k} \rfloor = (k-1)q = kq - q.$$

Since $kq = p - r$, We have

$$(k - 1)(mp + \lfloor \frac{p}{k} \rfloor) + j + 1 = p - r - q + j + 1 \equiv_p j + 1 - q - r$$

Thus p divides terms in the first product if and only if $j = q + r - 1$. We assumed $k^2 < p$ and that p is a prime, so $q > k$ and $r > 0$, and hence any such j is larger than $k - 1$. Thus p doesn't divide any of the terms in this product, and hence $\nu_p(\prod_{j=1}^{k-1} [(k - 1)(mp + \lfloor \frac{p}{k} \rfloor) + j + 1]) = 0$. The above considerations show that $\nu_p(C_k(mp + \lfloor \frac{p}{k} \rfloor + 1))$ is positive, and hence p does divide $C_k(mp + \lfloor \frac{p}{k} \rfloor + 1)$. \square

We now consider an example. We know by theorem 4 that $C_3(142)$ is divisible by 11 since $142 = 11 \times 13 - 1$. In fact, $\nu_{11}(C_3(142)) = 1$. Using equation the derived equation from above and the fact that 11 does not divide 13, we know that 11 does not divide $C_3(143)$, $C_3(144)$, $C_3(145)$, or $C_3(146)$, but does divide $C_3(147)$. We can continue to consider this example to see how the above equation applies when the condition in theorem 7 fails. It can be confirmed that $\nu_{13}(C_3(142)) = 2$, and hence since 13 does not divide 11, $C_3(143)$, $C_3(144)$, $C_3(145)$, $C_3(146)$, and $C_3(147)$ are divisible by 13 but not by 13^2 , where as $C_3(148)$ is divisible by 13^2 .

BIBLIOGRAPHY

- [1] R. Alter, K. Kubota, Prime and Prime Power Divisibility of Catalan Numbers, *Journal of Comb. Theory* **15**, 243–256 (1973).
- [2] E. Deutsch, Dyck Path Enumeration, *Disc. Math.* **204**, 167 – 202 (1999).
- [3] R. Graham, D. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley Publishing Co., Reading, MA (1989).
- [4] S. Heubach, N. Li, T. Mansour, A Garden of k -Catalan Structures, (2008) <http://web.calstatela.edu/faculty/sheubac/papers/k-Catalan%20structures.pdf>
- [5] P. Hilton, J. Pedersen, Catalan Numbers, Their Generalization, and Their Uses, *The Math. Intel.* **V13N2**, 64 – 75 (1991).
- [6] T. Koshy, Z, Gao, Some Divisibility Properties of the Catalan Numbers, *Math. Gaz.* **95**, 96 – 102 (2011).
- [7] T. Koshy, Z, Gao, Catalan Numbers with Mersenne Subscripts, *Math. Sci.* **38** **2**, 86 – 91 (2013).
- [8] A. Lobb, Deriving the n th Catalan Number, *Math. Gaz.* **83**, 109 – 110 (1999).
- [9] R.P. Stanley, *Catalan Addendum*, (05/25/13), <<http://www-math.mit.edu/~rstan/ec/catadd.pdf>>, (11/04/15).

CURRICULUM VITAE

University of Nevada, Las Vegas

Jacob Bobrowski

Home Address:

2948 Calle Grande
Las Vegas, Nevada 89120

Degrees:

Bachelor of Arts, Mathematics, 2013
University of Nevada, Las Vegas

Thesis Title: Generalized Catalan Numbers and Some Divisibility Properties

Thesis Examination Committee:

Chairperson, Dr. Peter Shiue, Ph.D.
Committee Member, Dr. Arthur Baragar, Ph. D.
Committee Member, Dr. Derrick DuBose, Ph. D.
Graduate Faculty Representative, Dr. David Beisecker, Ph. D.

