Journal of Integer Sequences, Vol. 17 (2014),
Article 14.10.7

A Weighted Interpretation for the Super Catalan Numbers

Emily Allen and Irina Gheorghiciuc
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
USA
eaallen@andrew.cmu.edu
gheorghi@andrew.cmu.edu

Abstract

The super Catalan numbers $T(m,n) = (2m)!(2n)!/2m!n!(m+n)!$ are integers that generalize the Catalan numbers. With the exception of a few values of m, no combinatorial interpretation is known for $T(m,n)$. We give a weighted interpretation for $T(m,n)$ and develop a technique that converts this weighted interpretation into a conventional combinatorial interpretation in the case $m = 2$.

1 Introduction

As early as 1874 Eugène Catalan observed that the numbers

$$S(m,n) = \frac{\binom{2m}{m}\binom{2n}{n}}{\binom{m+n}{n}} = \frac{(2m)!(2n)!}{m!n!(m+n)!}$$

are integers. This can be proved algebraically by showing that, for every prime number p, the power of p which divides $m!n!(m+n)!$ is at most the power of p which divides $(2m)!(2n)!$. No combinatorial interpretation of $S(m,n)$ is yet known.

Interest in the subject in the modern era was reignited by Gessel [5]. He noted that, except for $S(0,0)$, the numbers $S(m,n)$ are even. Gessel refers to

$$T(m,n) = \frac{(2m)!(2n)!}{2(m!n!(m+n)!)}$$

as the super Catalan numbers. The super Catalan numbers defined by Gessel should not be confused with the little Schröder numbers, which are sometimes also called super Catalan numbers.

$m\backslash n$	0	1	2	3	4	5	6	7
0	na	1	3	10	35	126	462	1716
1	1	1	2	5	14	42	132	429
2	3	2	3	6	14	36	99	286
3	10	5	6	10	20	45	110	286
4	35	14	14	20	35	70	154	364
5	126	42	36	45	70	126	252	546
6	462	132	99	110	154	252	462	924
7	1716	429	286	286	364	546	924	1716

Table 1: A table for $T(m,n)$.

Clearly $T(0,n) = \binom{2n-1}{n}$, whilst $T(1,n) = C_n$ giving the Catalan numbers, a well-known sequence with over 66 combinatorial interpretations [10].

An interpretation of $T(2,n)$ in terms of blossom trees has been found by Schaeffer [9], and another in terms of cubic trees by Pippenger and Schleich [8]. An interpretation of $T(2,n)$ in terms of pairs of Dyck paths with restricted heights has been found by Gessel and Xin [6]. They have also provided a description of $T(3,n)$. An interpretation of $T(m,m+s)$ for $0 \leq s \leq 3$ in terms of restricted lattice paths has been given by Chen and Wang [3].

A weighted interpretation of $S(m,n)$ based on von Szily's identity has been given by Georgiadis, Munemasa and Tanaka [4]. Their interpretation is in terms of lattice paths of length $2m+2n$ with a condition on the y-coordinate of the end-point of the $2m^{th}$ step.

In Section 2 we provide a weighted interpretation of $T(m,n)$ for $m,n \geq 1$ in terms of 2-Motzkin paths of length $m+n-2$, or Dyck paths of length $2m+2n-2$. Since the lattice paths in [4] are not Dyck paths, our interpretation is different from the one by Georgiadis, Munemasa and Tanaka. In Section 3 we are able to use our weighted interpretation to re-derive a result by Gessel and Xin [6], which we were then able to generalize for super Catalan polynomials [2].

2 2-Motzkin paths

A 2-Motzkin path of length n starts at the origin, ends at the point $(n,0)$, never goes below the x-axis, and consists of unit steps that are diagonally *up*, diagonally *down*, *straight level* and *wavy level*. A Dyck path of length $2n$ is a 2-Motzkin path of length $2n$ with no *level* steps.

Given a 2-Motzkin path, the level of a point is defined to be its y-coordinate. The height of a path is the maximum y-coordinate which the path attains. The height of a path π will be denoted $h(\pi)$.

For a fixed $m \geq 0$, we call a 2-Motzkin path π m-positive if the m^{th} step begins on an even level, otherwise π is m-negative. Let $P(m,n)$ be the number of m-positive 2-Motzkin paths of length $m+n-2$, and $N(m,n)$ be the number of m-negative 2-Motzkin paths of length $m+n-2$.

There is a well-known bijection between 2-Motzkin paths of length $n-1$ and Dyck paths of length $2n$ [7]. Given a 2-Motzkin path, read the steps from left to right and do the following replacements: replace an *up* step with two *up* steps, a *down* step with two *down* steps, a *straight* step with an *up* step followed by a *down* step, and a *wavy* step with a *down* step followed by an *up* step. The resulting path may touch level -1, thus, in addition, add an *up* step to the beginning of the resulting path and a *down* step to the end to obtain a Dyck path.

Theorem 1. *For $m, n \geq 1$, the super Catalan number $T(m,n)$ counts the number of m-positive 2-Motzkin paths of length $m+n-2$ minus the number of m-negative 2-Motzkin paths of length $m+n-2$. That is,*

$$T(m,n) = P(m,n) - N(m,n).$$

Proof. The super Catalan numbers satisfy the following identity, attributed to Dan Rubenstein [5],

$$4T(m,n) = T(m+1,n) + T(m,n+1). \tag{1}$$

Note that (1) can be viewed as a recurrence for $T(m,n)$ on m if written as

$$T(m+1,n) = 4T(m,n) - T(m,n+1).$$

Given a 2-Motzkin path π of length $m+n-2$, define the weight of π to be 1 if π is m-positive and -1 if π is m-negative.

Let $F(m,n)$ be the sum of the weights of all 2-Motzkin paths of length $m+n-2$, that is, $F(m,n) = P(m,n) - N(m,n)$. To prove $F(m,n) = T(m,n)$, we will check the initial condition

$$F(1,n) = C_n$$

and the recurrence given by (1),

$$4F(m,n) = F(m+1,n) + F(m,n+1).$$

For $m = 1$, the weight of any 2-Motzkin path of length n is 1 because the first step always starts at the (even) level $y = 0$. Hence $F(1,n) = C_n$, giving the number of 2-Motzkin paths of length $n-1$.

Next we consider the sum of the weights counted by $F(m,n+1) + F(m+1,n)$. If a 2-Motzkin path of length $m+n-1$ has an *up* or *down* step at step m, it will be counted once as a m-positive path and once as a m-negative path, and will not contribute to this sum.

Paths of length $m+n-1$ with a *level* step at step m will be counted twice. Let π be such a 2-Motzkin path. By contracting the m^{th} step in π, we obtain a 2-Motzkin path of length $m+n-2$; furthermore, every 2-Motzkin path of length $m+n-2$ can be obtained by contracting exactly two 2-Motzkin paths of length $m+n-1$, one with a *wavy* step at step m and one with a *straight* step at step m.

Thus the sum of the weights counted by $F(m,n+1)+F(m+1,n)$ is twice the sum of the weights of 2-Motzkin paths of length $m+n-1$ with *level* steps at step m; which is four times the sum of the weights of 2-Motzkin paths of length $m+n-2$, that is, $4F(m,n)$. $\square$

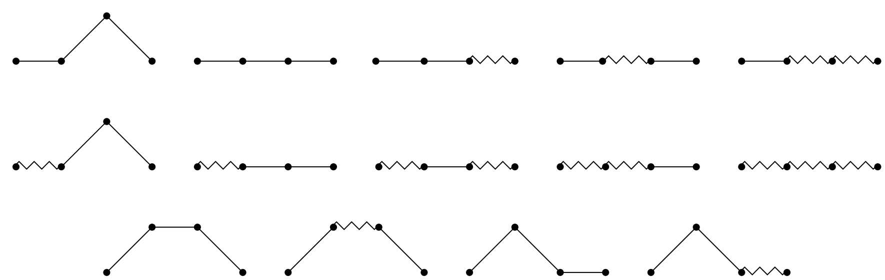

Figure 1: When $m=2$, there are ten m-positive 2-Motzkin paths and four m-negative 2-Motzkin paths of length 3. $T(2,3)=P(2,3)-N(2,3)=6$.

This weighted interpretation can be used to prove combinatorially that $T(m,n)=T(n,m)$. Let π be a path of length $m+n-2$ counted by $T(m,n)$. Consider the reverse of a path to be that path read from right to left. Since the m^{th} step of π and the n^{th} step of the reverse of π start at the same point, mapping a path to its reverse is a weight preserving involution between the 2-Motzkin paths counted by $T(m,n)$ and the 2-Motzkin paths counted by $T(n,m)$.

We can reformulate the result of Theorem 1 in terms of Dyck paths. In this case $P(m,n)$ is the number of Dyck paths of length $2m+2n-2$ whose $2m-1^{st}$ step ends on level 1 (mod 4), and $N(m,n)$ is the number of Dyck paths of length $2m+2n-2$ whose $2m-1^{st}$ step ends on level 3 (mod 4).

Similar to a Dyck path, a ballot path starts at the origin, uses a finite number of diagonally *up* and diagonally *down* steps, and does not go below the x-axis. A ballot path ends on or above the x-axis. Let $B(n,r)$ be the number of ballot paths that end at the point $(2n-1,2r-1)$. It is well known that $B(n,r)=\frac{r}{n}\binom{2n}{n+r}$. Then

$$T(m,n)=\sum_{r=1}^{\min\{m,n\}}(-1)^{r-1}B(m,r)B(n,r) \tag{2}$$

and

$$T(m,n)=\sum_{r=1}^{\min\{m,n\}}(-1)^{r-1}\frac{r^2}{nm}\binom{2m}{m+r}\binom{2n}{n+r}. \tag{3}$$

Equation (3) is a new identity for the super Catalan number $T(m,n)$. A q-analog of this identity is given in [2], and its algebraic proof appears in [1].

3 Combinatorial techniques

We define the total length of an ordered pair of Dyck paths (π, ρ) to be the sum of the lengths of the paths π and ρ. The height of the empty Dyck path is zero. In [6] Gessel and Xin use an inclusion-exclusion argument to prove the following result.

Theorem 2 (Gessel, Xin)**.** *For $n \geq 1$, the number $T(2,n)$ counts the ordered pairs of Dyck paths (π, ρ) of total length $2n$ with $|h(\pi) - h(\rho)| \leq 1$. Here π and ρ are allowed to be the empty path.*

Our goal in this section is to derive a similar result using Theorem 1 and some direct Dyck paths subtraction techniques that will be easier to generalize for larger values of m. We already were able to generalize this result to super Catalan Polynomials in [2].

Let $\mathcal{D}_n$ denote the set of Dyck paths of length $2n$. For a path $\pi \in \mathcal{D}_n$, let R be the rightmost highest point on π. We define the X-point of π to be the last, from left to right, level one point on the portion of π before and including R. In other words, if $h(\pi) > 1$, then the X-point is the last, from left to right, level one point before R. If $h(\pi) = 1$, then the X-point and R coincide. See Figure 2.

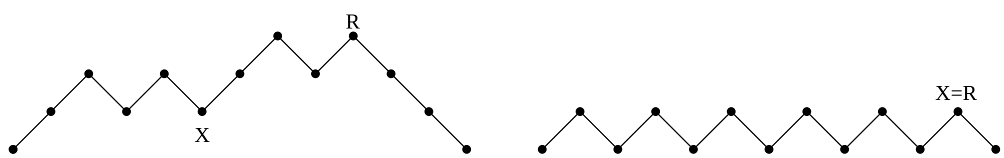

Figure 2: The X-point of two Dyck paths.

Let $h_{-}(\pi)$ denote the maximum level that the path π reaches from its beginning until and including the X-point, and $h_{+}(\pi)$ denote the maximum level that the path π reaches after and including the X-point. Obviously $h_{-}(\pi) \leq h_{+}(\pi) = h(\pi)$.

Theorem 3. *Let $n \geq 1$. The super Catalan number $T(2,n)$ counts Dyck paths π of length $2n$ such that $h_{+}(\pi) \leq h_{-}(\pi) + 2$, the path of height one counting twice.*

Proof. Let $\mathcal{A}_n$ denote the set of Dyck paths of length $2n$ that start with *up, down, up*, $\mathcal{B}_n$ denote the set of Dyck paths of length $2n$ that start with *up, up, down*, and $\mathcal{N}_n$ denote the set of Dyck paths of length $2n$ that start with *up, up, up*.

By Theorem 1, $T(2,n) = P(2,n) - N(2,n)$, where $P(2,n)$ is the number of 2-Motzkin paths of length n that start with a *level* step, and $N(2,n)$ is the number 2-Motzkin paths of length n that start with an *up* step. The canonical bijection between 2-Motzkin paths and Dyck paths leads to the following interpretation:

$$T(2,n) = |\mathcal{A}_{n+1}| + |\mathcal{B}_{n+1}| - |\mathcal{N}_{n+1}|.$$

Note that $\mathcal{A}_{n+1}$, $\mathcal{B}_{n+1}$ and $\mathcal{N}_{n+1}$ are subsets of $\mathcal{D}_{n+1}$. By contracting the second and third steps in the paths in $\mathcal{A}_{n+1}$ and $\mathcal{B}_{n+1}$ we get twice $\mathcal{D}_n$, so $|\mathcal{A}_{n+1}| = |\mathcal{B}_{n+1}| = C_n$.

We consider all paths π in $\mathcal{N}_{n+1}$ that do not attain level one between the third step of π and the rightmost highest point R on π. The set of all such paths will be denoted by $\mathcal{N}^*_{n+1}$. Let $\mathcal{N}^{**}_{n+1} = \mathcal{N}_{n+1} - \mathcal{N}^*_{n+1}$. Then

$$T(2,n) = 2|\mathcal{D}_n| - |\mathcal{N}^*_{n+1}| - |\mathcal{N}^{**}_{n+1}|. \tag{4}$$

First we establish an injection f from $\mathcal{N}^*_{n+1} \subset \mathcal{D}_{n+1}$ to $\mathcal{D}_n$. For $\pi \in \mathcal{N}^*_{n+1}$, let RQ be the *down* step that follows the rightmost highest point R of π. We define $f(\pi)$ to be the path obtained by removing the second and third steps in π, both of which are *up* steps, and then substituting the *down* step RQ by an *up* step. See Figure 3. Since π does not attain level one between its third step and R, $f(\pi)$ is a Dyck path of length $2n$. Note that Q is the leftmost highest point on $f(\pi)$. Also, since at least two *up* steps precede Q on $f(\pi)$, the height of $f(\pi)$ is at least two. Thus the Dyck path of height one and length $2n$ is not in the image of f.

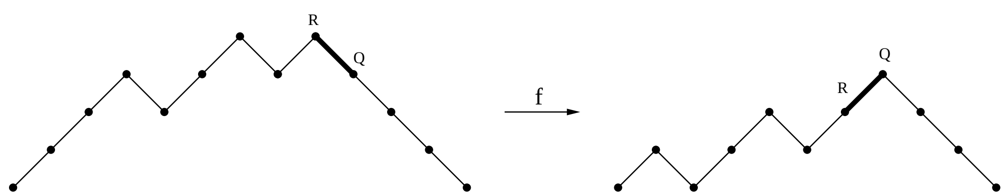

Figure 3: f removes the 2^{nd} and 3^{rd} steps, substitutes the *down* step RQ by an *up* step.

We will show that f is an injection and that the only path in $\mathcal{D}_n$ that is not in the image of f is the Dyck path of height one. Let ρ be in $\mathcal{D}_n$ of height $h(\rho) > 1$. Let Q be the leftmost highest point on ρ and RQ be the *up* step that precedes Q. Insert two *up* steps after the first step of ρ, then substitute the *up* step RQ by a *down* step, which makes R the rightmost highest point of the resulting path π. The path π is in $\mathcal{N}^*_{n+1}$ and $f(\pi) = \rho$.

It follows that $|\mathcal{D}_n| - |\mathcal{N}^*_{n+1}|$ counts only one path, the Dyck path of length $2n$ and height one.

Next we establish an injection g from $\mathcal{N}^{**}_{n+1} \subset \mathcal{D}_{n+1}$ to $\mathcal{D}_n$. A path π in $\mathcal{N}^{**}_{n+1}$ attains level one between its third step and the rightmost highest point R on π. Let Y be the first point between the third step of π and R at which π attains level one. The segment XY that consists of two *down* steps precedes Y. We remove the second and third steps of π and substitute the two *down* steps XY by two *up* steps. See Figure 4. The resulting path is a ballot path of length $2n$ that ends at level two. From left to right, X is the last level one

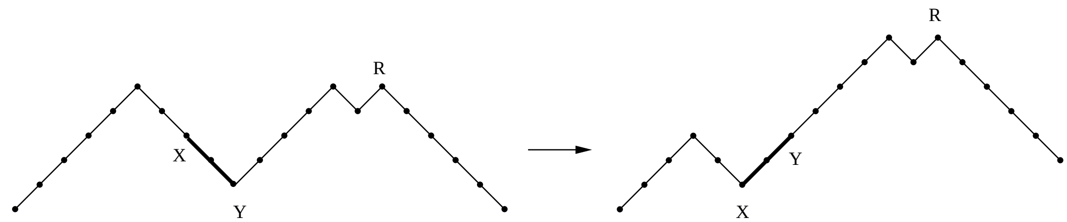

Figure 4: First part of g action is removing the 2^{nd} and 3^{rd} steps, substituting the two *down* steps XY by two *up* steps.

point on this ballot path. The maximum level that this path reaches up to and including point X is less than the maximum level it reaches after and including point X by at least 4.

Let L be the leftmost highest point of this ballot path and ML be the *up* step that precedes L. Substitute the *up* step ML by a *down* step. See Figure 5. The resulting path $g(\pi)$ is in $\mathcal{D}_n$ and M is its rightmost highest point. Note that X is the last level one point on $g(\pi)$ before its rightmost highest point M and $h_+(g(\pi)) \geq h_-(g(\pi)) + 3$.

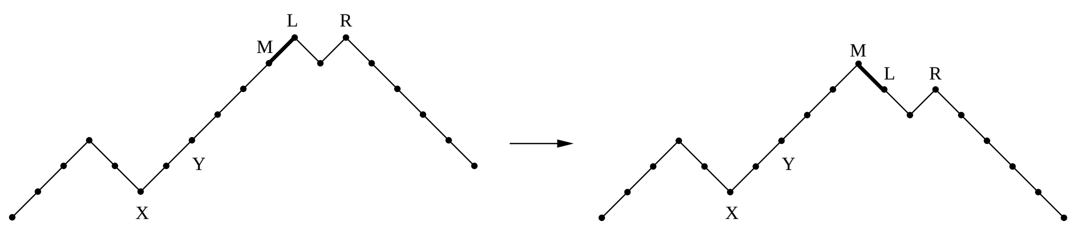

Figure 5: Second part of g action is substituting the *up* step ML with a *down* steps.

We will show that g is an injection and that the only paths in $\mathcal{D}_n$ that are not in the image of g are the Dyck paths σ that satisfy $h_+(\sigma) \leq h_-(\sigma) + 2$. Let ρ be in $\mathcal{D}_n$ and $h_+(\rho) \geq h_-(\rho) + 3$. Let M be the rightmost highest point on ρ and ML be the *down* step that follows M. Let X be the X-point of ρ, that is the last level one point, from left to right, before and including M. Substitute the *down* step ML by an *up* step. The result is a ballot path of length $2n$ that ends at level two. Note that L is the leftmost highest point on this ballot path. Let R denote the rightmost highest point on this ballot path. From left to right, X is the last level one point on this ballot path. The maximum level that this path reaches up to and including point X is less than the maximum level it reaches after and including point X by at least 4. Since X is the last level one point, it is followed by the segment XY that consists of two *up* steps. Next we insert two *up* steps after the first step of this ballot path and then substitute the two *up* steps XY by two *down* steps. The resulting path is a Dyck path of length $2n+2$, we denote it by π. Point Y is the first level one point after the third step of π. Note that the maximum level that this Dyck path reaches after Y is at least the maximum level that this Dyck path reaches up to and including Y, which means that the rightmost highest point R is to the right of Y. If follows that $p \in \mathcal{N}^{**}_{n+1}$ and $g(\pi) = \rho$.

Thus $|\mathcal{D}_n| - |\mathcal{N}^{**}_{n+1}|$ counts Dyck paths π of length $2n$ that satisfy $h_+(\pi) \leq h_-(\pi) + 2$. Note that the Dyck path of length $2n$ and height one is among these paths.

Equation (4) can be re-written as

$$T(2, n) = (|\mathcal{D}_n| - |\mathcal{N}^{*}_{n+1}|) + (|\mathcal{D}_n| - |\mathcal{N}^{**}_{n+1}|).$$

Hence $T(2, n)$ counts Dyck paths π of length $2n$ such that $h_+(\pi) \leq h_-(\pi) + 2$, the path of height one counting twice. □

We will now show a simple bijection from the objects described in Theorem 3 to those in Theorem 2.

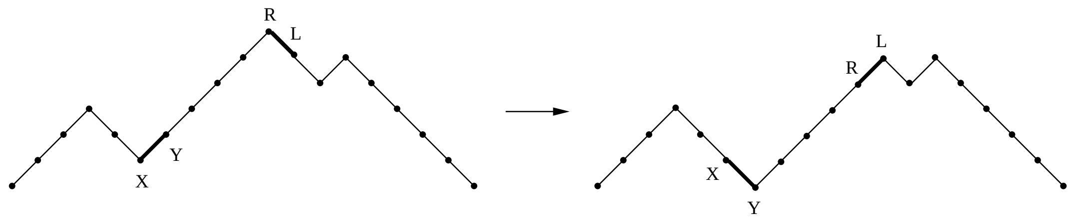

Figure 6: From Dyck paths described in Theorem 3 to pairs of Dyck path described in Theorem 2.

Let π be a Dyck path of length $2n$ and height $h(\pi) > 1$, such that $h_+(\pi) \leq h_-(\pi) + 2$. Let R be the rightmost highest point of π. Note that X is followed by an *up* step XY and R is followed by a *down* step RL. Substitute the *up* step XY with a down step, substitute the *down* step RL with an *up* step. See Figure 6. As a result, the portion of π between Y and R will be lowered by two levels. Since π does not attain level one between Y and R, the resulting path is a Dyck path with point Y on level zero.

Note that Y separates this Dyck path into a pair of Dyck paths (ρ, σ). The height of ρ is $h_-(\pi)$, the height of σ is $h_+(\pi) - 1$. Thus $|h(\rho) - h(\sigma)| \leq 1$. Since L is the leftmost highest point on σ, this mapping is reversible. Theorem 3 counts the Dyck path τ of height one twice. This corresponds to the pairs (τ, ϵ) and (ϵ, τ) in Theorem 2, where ϵ is the empty path.

4 Acknowledgments

We are very thankful to Ira Gessel for his helpful comments. The problem about the combinatorial interpretation of the super Catalan Numbers was first mentioned to us by the late Herb Wilf, to whom we are immensely grateful.

References

[1] E. Allen, *Combinatorial interpretations of generalizations of Catalan numbers and ballot numbers*, Ph.D. thesis, Carnegie Mellon University, 2014.

[2] E. Allen and I. Gheorghiciuc, On super Catalan polynomials, preprint, http://arxiv.org/abs/1403.5296, August 31 2014.

[3] X. Chen and J. Wang, The super Catalan numbers $S(m, m+s)$ for $s \leq 3$ and some integer factorial ratios, preprint, http://www.math.umn.edu/~reiner/REU/ChenWang2012.pdf, 2012.

[4] E. Georgiadis, A. Munemasa, and H. Tanaka, A note on super Catalan numbers, *Interdiscip. Inform. Sci.* **18** (2012), 23–24.

[5] I. Gessel, Super ballot numbers, *J. Symb. Comput.* **14** (1992), 179–194.

[6] I. Gessel and G. Xin, A combinatorial interpretation of the numbers $6(2n)!/n!(n+2)!$, *J. Integer Seq.* **8** (2005) Article 05.2.3.

[7] M. Pierre-Delest and G. Viennot, Algebraic languages and polyominoes enumeration, *Theoret. Comput. Sci.* **34** (1984), 196–206.

[8] N. Pippenger and K. Schleich, Topological characteristics of random triangulated surfaces, *Random Structures Algorithms* **28** (2006), 247–288.

[9] G. Schaeffer, A combinatorial interpretation of super-Catalan numbers of order two, unpublished manuscript, 2003.

[10] R. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, 1998.

2010 *Mathematics Subject Classification*: Primary 05A15; Secondary 05A19.
Keywords: super Catalan number, Dyck path, 2-Motzkin path, combinatorial interpretation, combinatorial identity.

(Concerned with sequences A000108, A007054, and A007272.)

Received August 26 2014; revised version received August 31 2014; September 11 2014. Published in *Journal of Integer Sequences*, November 5 2014.

Return to Journal of Integer Sequences home page.