A MOTZKIN VARIATION ON THE TENNIS BALL PROBLEM

NAIOMI T. CAMERON AND LYNNELL S. MATTHEWS

ABSTRACT. We consider a variation on the Tennis Ball Problem studied by Mallows-Shapiro and Merlini, et al. The solution to the original problem is the well known Catalan numbers, while the variations discussed in this paper yield the Motzkin numbers and other related sequences. For this variation, we present a generating function for the sum of the labels on the balls.

1. Introduction

The s-tennis ball problem is the following: At first turn, you are given s balls labelled $1, 2, \dots, s$, where s is a fixed positive integer. You toss one of them out of the window onto the lawn. At the second turn, balls numbered $s+1, s+2, \dots, 2s$ are given to you and now you toss any of the $2s-1$ remaining balls onto the lawn. At the third turn, balls numbered $2s+1, 2s+2, \dots, 3s$ are received and then one of the remaining balls goes out on the lawn. This process continues for n turns. At that point, consider the combination of balls left on the lawn.

Question 1. *How many different combinations of balls on the lawn are possible after n turns?*

Question 2. *What is the sum of the balls on the lawn over all distinct possibilities after n turns?*

For the case when $s = 2$, balls labelled $1, 2, 3, 4, 5, 6, \dots$ are provided in sequence two at a time. We write the combinations on the lawn by order of nondecreasing labels, rather than order of arrival on the lawn. For example, the possible combination $1, 2, 4$ (in order of arrival after 3 turns) is the same as the combination $1, 4, 2$ and so we write it uniquely as 124. The 14 possibilities after $n = 3$ turns are:

123	124	125	126	134	135	136
145	146	234	235	236	245	246

When $s = 2$, the answer to Question 1 is the $(n+1)$-st Catalan number,

$$c_{n+1} = \frac{1}{n+2}\binom{2n+2}{n+1},$$

while the answer to Question 2 is

$$\frac{2n^2+5n+4}{n+2}\binom{2n+1}{n} - 2^{2n+1},$$

2000 *Mathematics Subject Classification.* Primary 05A19, 05A05 Secondary 11B39.
Key words and phrases. Catalan numbers, Motzkin numbers, combinatorial proofs, lattice paths.

a result due to Mallows and Shapiro [5]. In 2002, Merlini, Sprugnoli and Verri settled Questions 1 and 2 for all s in [6]. Jani and Zeleke provide a bijective proof of the answer to Question 1 in [4].

In the original statement of the problem, the s balls given at turn i are labelled consecutively $s(i-1)+1, \ldots, si-1, si$. In this paper, we consider the following variation. Label the s balls provided in turn i identically with the number i. For example, when $s=2$, after $n=3$ turns, the six balls received are labelled 1, 1, 2, 2, 3, 3 and there are four possible combinations left on the lawn:

112 113 122 123

It turns out that the answer to Question 1 for this variation of the problem when $s=2$ is given by the Motzkin numbers. We call this variation of the problem s-MTBP for Motzkin Tennis Ball problem. In section 2, we address Question 1 for s-MTBP and in section 3, we provide an answer for Question 2.

2. s-Motzkin Numbers

The Motzkin numbers $1, 1, 2, 4, 9, 21, \ldots$ arise in several settings related to the Catalan sequence. The generating function for the Motzkin numbers is

$$\frac{1-x-\sqrt{1-2x-3x^2}}{2x^2}.$$

They have been the subject of several papers, [1, 8, 11], and have a number of interesting interpretations (see exercise 6.39 of [10]). There is one of particular interest here. The n-th Motzkin number, m_n, is the number of *Motzkin paths*, that is, lattice paths from $(0,0)$ to $(n,0)$ which use the step set $\{U(1,1), D(1,-1), L(1,0)\}$ and never go below the x-axis.

Let us consider 2-MTBP. To show that m_n is the answer to Question 1, that is, the number of distinct combinations of balls on the lawn after n turns, we exhibit a one-to-one correspondence between Motzkin paths of length n and balls on the lawn after n turns. To each combination of balls on the lawn, we associate a path of length n in the following way. Each label in the combination is assigned a step in the corresponding path. Reading a combination from left to right, if the label appears twice, we associate an up step $U(1,1)$. If the label appears exactly once, we associate a level step $L(1,0)$. If the label does not appear, then its corresponding step is a down step $D(1,-1)$. See Figure 1.

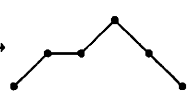
11233 $\longleftrightarrow$ 

FIGURE 1. If 11233 are the balls left on the lawn after $n=5$ turns, then the corresponding Motzkin path of length 5 is $ULUDD$.

Since there are exactly n balls on the lawn after n turns, whenever some label i is repeated on the lawn, there must be some label $j > i$ which does not appear on the lawn. Thus, in the corresponding path, we guarantee that for every U there is a corresponding D. Furthermore, since at least one ball must be tossed out at each turn, the number of D's will never exceed the number of preceding U's. Hence, the path ends at $(n,0)$ and never goes below the x-axis. In other words, the path is a

Motzkin path. Conversely, by the same scheme, every Motzkin path corresponds uniquely to a combination of balls on the lawn.

Now, we can generalize this correspondence to determine the number of balls on the lawn after n turns for any fixed positive number s of balls received at each turn. We do this in the following way. In general, when reading each label in the combination from left to right, if the label appears exactly l times, then associate the up step $U(1, l-1)$. If a label that does not appear, we associate the down step $D(1,-1)$. If a label which appears just once, we associate a level step $L(1,0)$. To any label which appears more than once, we associate an up step of the form $U(1, l-1)$. By the same argument as above, we have a one-to-one correspondence between the set of all combinations of balls on the lawn and the set M_n^s of paths from $(0,0)$ to $(n,0)$ which use the step set $\{(1,-1),(1,0),(1,1),(1,2),\ldots,(1,s-1)\}$ and never go below the x-axis. We shall call these paths *s-Motzkin paths* [see [3],[12]]. Therefore, by answering Question 1 we have the following thereom:

Theorem 1. *The set possible sequences for s-MTBP is in bijective correspondence with the set M_n^s of s-Motzkin paths. Furthermore, the generating function for the s-Motzkin paths, $M_s(z) = \sum_{n=0}^{\infty} M_n^s z^n$ satisfies the following recurrence:*

$$M_s(z) = 1 + zM_s(z) + z^2 M_s^2(z) + \ldots + z^s M_s^{s+1}(z). \tag{1}$$

□

Notice that 2-Motzkin paths are the same as the standard Motzkin paths.

3. A Formula for the Sum Sequence

The sum of the possible sequences for the 2-MTBP after n turns begins 1, 5, 20, 74, 259 as verified by hand calculation. A generating function for this sequence is given in this section. We first present the following theorem which relates the balls on the lawn sequence for 2-MTBP to the area under the associated Motzkin path.

Theorem 2. *Let α be a Motzkin path of length n and let $a(\alpha)$ be the area under α. Then*

$$a(\alpha) = \sum_{i=1}^{n} (i - \alpha_i), \tag{2}$$

where α_i is the i^{th} entry of the balls on the lawn sequence corresponding to α.

For example, in Figure 1 recall that the path α corresponds to sequence 11233. Using (2) to compute $a(\alpha)$ for this path we have the following:

$$a(\alpha) = (1-1) + (2-1) + (3-2) + (4-3) + (5-3) = 5$$

Proof. We proceed by induction on n. When $n = 1$, the theorem is true since the only Motzkin path α of length one has balls on the lawn sequence 1, and $(1-1) = 0$ is the area under α. We assume the statement is true for Motzkin paths of length k or less, where $k \geq 1$, and show that it is also true for Motzkin paths of length $k+1$. Suppose α has length $k+1$. Either α ends with a horizontal step or α ends with a down step.

Case 1: α ends with a horizontal step

Suppose α ends with a horizontal step. Then the final label in its balls on the lawn sequence is $k+1$, that is, $\alpha_{k+1} = k+1$. Let β be the path of length k obtained

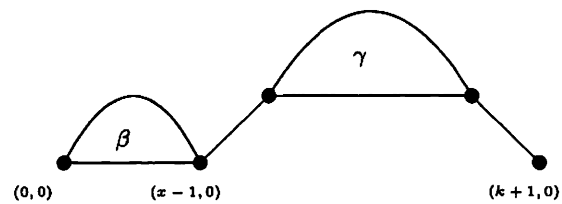


FIGURE 2. A Motzkin path α of length $k+1$ ending with a down step and partitioned into subpaths β and γ.

by deleting the final horizontal step of α. Then

$$\begin{aligned} a(\alpha) &= a(\beta) \\ &= \sum_{i=1}^{k}(i-\beta_i) \\ &= \sum_{i=1}^{k}(i-\alpha_i), \quad \text{since } \forall\ 1 \le i \le k,\ \beta_i = \alpha_i \\ &= \left(\sum_{i=1}^{k}(i-\alpha_i)\right) + (k+1) - \alpha_{k+1}, \quad \text{since } \alpha_{k+1} = k+1 \\ &= \sum_{i=1}^{k+1}(i-\alpha_i) \end{aligned}$$

Case 2: α ends with a down step

Suppose α ends with a down step. Then there exists an up step in α which goes from height zero to height one. Let x denote the position of the last such up step in α. Let β denote the subpath of α starting at the origin and ending at position $x-1$. Let γ denote the Motzkin path associated with the subpath starting at position $x+1$ and ending at position k of α. See Figure 2. Clearly, $a(\alpha) = a(\beta) + a(\gamma) + k + 1 - x$.

We will use the induction hypothesis to determine the area under β and γ. Since $\beta_i = \alpha_i$ for all $i = 1, \ldots, x-1$, computing $a(\beta)$ is straightforward. To determine $a(\gamma)$, we need its balls on the lawn sequence $\gamma_1\gamma_2\cdots\gamma_{k-x}$.

First, consider the (elevated) Motzkin path γ^* obtained by deleting β from α. To obtain the balls on the lawn sequence for this path, subtract $(x-1)$ from each of the labels α_i, for $i = x, \ldots, k+1$.

$$\alpha_x - (x-1), \alpha_{x+1} - (x-1), \alpha_{x+2} - (x-1), \ldots, \alpha_{k+1} - (x-1)$$

To obtain the balls on the lawn sequence for γ, remove the first two labels 11 from the balls on the lawn sequence for γ^*, then subtract 1 from each remaining label in γ^*. So we have

$$\alpha_{x+2} - (x-1) - 1, \alpha_{x+3} - (x-1) - 1, \ldots, \alpha_{k+1} - (x-1) - 1$$

or equivalently,

$$\alpha_{x+2} - x, \alpha_{x+3} - x, \ldots, \alpha_{k+1} - x$$

Hence, $\gamma_i = \alpha_{x+1+i} - x$, for $i = 1, \ldots, k-x$.

Now,

$$\begin{aligned}
a(\alpha) &= a(\beta)+a(\gamma)+k+1-x \\
&= \left(\sum_{i=1}^{x-1} i-\alpha_i\right)+\left(\sum_{i=1}^{k-x} i-(\alpha_{x+1+i}-x)\right)+k+1-x \\
&= \left(\sum_{i=1}^{x-1} i\right)+\left(\sum_{i=1}^{k-x} i+x\right)+k+1-x-\left(\sum_{i=1}^{x-1}\alpha_i\right)-\left(\sum_{i=1}^{k-x}\alpha_{x+1+i}\right) \\
&= \left(\sum_{i=1}^{x-1} i\right)+\left(\sum_{i=1}^{k-x} i+x\right)+k+1-x-\left[\left(\sum_{i=1}^{k+1}\alpha_i\right)-\alpha_x-\alpha_{x+1}\right] \\
&= \left(\sum_{i=1}^{x-1} i\right)+\left(\sum_{i=1}^{k-x} i+x\right)+k+1+x-\left(\sum_{i=1}^{k+1}\alpha_i\right), \text{ since } \alpha_x=\alpha_{x+1}=x \\
&= \left(\sum_{i=1}^{x} i\right)+\left(\sum_{i=1}^{k-x+1} i+x\right)-\left(\sum_{i=1}^{k+1}\alpha_i\right)
\end{aligned}$$

To complete the proof, it suffices to show that

$$\left(\sum_{i=1}^{x} i\right)+\left(\sum_{i=1}^{k-x+1} i+x\right)=\sum_{i=1}^{k+1} i$$

With straightforward algebraic manipulation, indeed, we see that

$$\begin{aligned}
\left(\sum_{i=1}^{x} i\right)+\left(\sum_{i=1}^{k-x+1} i+x\right) &= \frac{x(x+1)}{2}+\frac{(k-x+1)(k-x+2)}{2}+x(k-x+1) \\
&= \frac{(k-x+1)(k+x+2)+x(x+1)}{2} \\
&= \frac{k^2+3k+2}{2} \\
&= \frac{(k+1)(k+2)}{2} \\
&= \sum_{i=1}^{k+1} i
\end{aligned}$$

□

Using Theorem 2, we rewrite Eq. (2) as

(3) $$a(\alpha)=\sum_{i=1}^{n} i-\sum_{i=1}^{n}\alpha_i$$

and to find the the cumulative area of all Motzkin paths of n steps, A_n, we find $a(\alpha)$ for each path then sum over the m_n paths. Thus, the left side will correspond to the difference between m_n copies of $\sum_{i=1}^{n} i$ and the cumulative sum of the balls on the lawn, S_n. Solving for S_n, we have

(4) $$S_n=m_n\cdot\left(\sum_{i=1}^{n} i\right)-A_n,$$

which leads to the following theorem.

Theorem 3. *For the 2-MTBP, let S_n be the sum of the balls on the lawn over all distinct possibilities after n turns, then*

$$S_n = m_n \cdot \binom{n+1}{2} - A_n, \tag{5}$$

where m_n is the n^{th} Motzkin number and A_n is the total area under all Motzkin paths with n steps.

In [2], Banderier provided the generating function for A_n, whose first few terms are $0, 1, 4, 16, 56, 190, 624, 2014, \ldots$ In [9], we find that

$$A(z) = \sum_{n=0}^{\infty} A_n z^n = \frac{x^2 + 2x - 1 + (1-x)\sqrt{(x+1)(1-3x)}}{2x^2(3x-1)(x+1)}.$$

We can now determine the generating function for the sum of the balls on the lawn $S(z) = \sum_{n=0}^{\infty} S_n z^n$, since

$$\begin{aligned}
S(z) &= \sum_{n=0}^{\infty} m_n \cdot \binom{n+1}{2} z^n - \sum_{n=0}^{\infty} A_n z^n \\
&= \frac{d}{dz}\left(\frac{z^2 M'(z)}{2}\right) - A(z) \\
&= \frac{z^2 M''(z)}{2} + \frac{2zM'(z)}{2} - A(z) \\
&= \frac{1}{2}\left(\frac{1}{x^2} - \frac{1-3x-3x^2+3x^3}{x^2(1-2x-3x^2)^{3/2}}\right) - \frac{x^2+2x-1+(1-x)\sqrt{(x+1)(1-3x)}}{2x^2(3x-1)(x+1)} \\
&= \frac{\sqrt{1-2x-3x^2} - 1}{(x+1)(3x-1)\sqrt{1-2x-3x^2}} \\
&= x + 5x^2 + 20x^3 + 74x^4 + 259x^5 + 881x^6 + 2932x^7 + 9614x^8 + O(x^9)
\end{aligned}$$

It is cumbersome, but not difficult, to show that Theorem 3 extends to s-Motzkin paths as well. While we omit the details, we note that the proof follows the same format as that of Theorem 3. Since the relationship between area under s-Motzkin paths and the sum of the labels of the balls for s-MTBP is exactly the same as it is for the case when $s = 2$, one can always produce a generating function for the sum sequence in s-MTBP given a generating function for the area under s-Motzkin paths.

4. Conclusions and Open Questions

In [6], Merlini, et al. treat a variant of the s-TBP called the $(4, 2)$-TBP which supplies 4 balls at each turn but now throws out 2 at a time. In [7], Mier and Noy, provide a solution to the generalized (s, t)-TBP. This leads to a question not addressed in this paper, e.g., what is the (s, t)-MTBP analagous result?

Furthermore, this paper examines a family of sequences related to the Motzkin numbers. A natural question is: "What other pertubations can be made to the original Tennis Ball Problem that will lead to other well known lattice path sequence bijections, in particular the small and large Schroder numbers?"

In the case of the small Schroder numbers, $s_n \leftrightarrow 1, 3, 11, 45, \ldots$, we proceed as we did in the original statement of the problem in section 1 with odd numbered balls being two sided, i.e., one side is heads and the other side tails. So at each turn two consecutive balls are added one of which will be two sided and one ball is thrown on the lawn. The 11 possibilities after $n = 2$ turns are as follows:

$$12 \quad \tilde{1}2 \quad 13 \quad 1\tilde{3} \quad \tilde{1}3 \quad \tilde{1}\tilde{3}$$
$$14 \quad \tilde{1}4 \quad 23 \quad 2\tilde{3} \quad 24$$

To show this pertubation yields the Schröder numbers, i.e. answers Question 1, consider the following Schröder Dyck paths: The n-th Schroder number, s_n, is the number of lattice paths from $(0,0)$ to $(2n,0)$ which use the step set $\{U(1,1), D(1,-1)\}$ and never go below the x-axis and up steps at even heights can be one of two colors, i.e. black and red. We establish a o bijection between these paths and the Schröder Tennis Ball Problem (STBP) in the following way. Each label $1, ..., 2n$ is assigned a step in the correpsonding path of length $2n+2$. Each path begins with an up step then the labels on the balls dictate the next $2n$ steps in the path and the path ends with a down step. For each STBP sequence, reading from left to right, the label will indicate an up step. The type of up step is determined by whether the label is odd or even as follows:

(1) If odd, the label will be heads or tails, associate black with heads and green with tails to decide which color up step to use.
(2) If even, it is not a two sided ball, we associate a black up step.

If the label does not appear, then its corresponding step is a down step. See Figure 3

$2\tilde{3}4$ ⟷

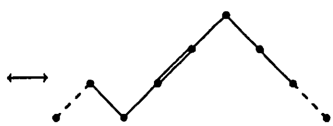


FIGURE 3. If $2\tilde{3}4$ are the balls left on the lawn after $n = 3$ turns, then the corresponding Schröer path of length 8 is $UDU\tilde{U}UDDD$.

At each turn, there are n balls on the lawn so for some label i, there must be some label $j > i$ which does not appear on the lawn. Thus, in the corresponding path of $2n+2$ for each U there is a corresponding D. So conversely, by the same scheme each path corresponds uniquely to a STBP sequence.

In STBP, the relationship between the labels of a valid combination and the area under the corresponding path does not appear to mimic that which we see in the MTBP, as it is no longer even obvious how to "sum" labels. However, one wonders if there is a way to characterize the relationship between area and labels in such a way that one can quantify the expected "value" of the labels for arbitrary n in STBP.

References

[1] M. Aigner, Motzkin numbers, *European J. Combin.*, **19** (1998), no. 6, 663–675.
[2] C Banderier, B. Gittenberger, *Analytic Combinatorics of Lattice Paths: Enumeration and Asymptotics for the Average Area*, Proceedings of the 4th Colloquium of Mathematics and Computer Science, DMTCS, AG, (2006) pp. 345-355.
[3] C. Banderier, D. Merlini, *Lattice Paths with an Infinite Set of Jumps* 2002.
[4] M. Jani, M. Zeleke, A bijective proof of a tennis ball problem, *Bull. Inst. Combin. Appl.*, 41 (2004), 89–95.
[5] C. Mallows, L. Shapiro, Balls on the lawn, *J. Integer Seq.*, **2** (1999), Article 99.1.5.

[6] D. Merlini, R. Sprugnoli, M. C. Verri, The tennis ball problem, *J. Combin. Theory Ser. A*, **99** (2002), 307–344.
[7] A. Mier, M. Noy, *A Solution to the Tennis Ball Problem*, Formal Power Series and Algebraic Combinatorics, (2004).
[8] E. Pergola, R. Pinzani, S. Rinaldi, R. A. Sulanke, A bijective approach to the area of generalized Motzkin paths, *Adv. in Appl. Math.*, **28** (2002), no. 3-4, 580–591.
[9] Sloan, On-line Integer Sequences, http://www.research.att.com/ njas/sequences/, #A05785.
[10] R. Stanley, *Enumerative Combinatorics, Volume 2*, Cambridge University Press, 1999.
[11] R. Sulanke, Moments of generalized Motzkin paths, *J. Integer Seq.*, **3** (2000), no. 1, Article 00.1.1.
[12] L. Takacas, Enumeration of Rooted Trees and Forrests, *Math. Scientist*, **18** (1993), 1–10.

LEWIS & CLARK COLLEGE
E-mail address: `ncameron@lclark.edu`

GETTYSBURG COLLEGE
E-mail address: `lmatthew@gettysburg.edu`