Journal of Integer Sequences, Vol. 19 (2016),
Article 16.6.1

Returns and Hills on Generalized Dyck Paths

Naiomi T. Cameron
Department of Mathematical Sciences
Lewis & Clark College
Portland, OR 97219
USA
ncameron@lclark.edu

Jillian E. McLeod
Department of Mathematics
U.S. Coast Guard Academy
New London, CT 06320
USA
jillian.e.mcleod@uscga.edu

Abstract

In 2009, Shapiro posed the following question: "What is the asymptotic proportion of Dyck paths having an even number of hills?" In this paper, we answer Shapiro's question, as well as a generalization of the question to ternary paths. We find that the probability that a randomly chosen ternary path has an even number of hills approaches 125/169 as the length of the path approaches infinity. Our strategy relies on properties of the Fine number sequence and extends certain relationships between the Catalan and Fine number generating functions.

1 Introduction

Much has been written about patterns of returns and hills on Dyck paths, which are enumerated by the Catalan numbers. Here we begin by extending some of those results to the setting of the sequence,

$$t_n = \frac{1}{2n+1}\binom{3n}{n},$$

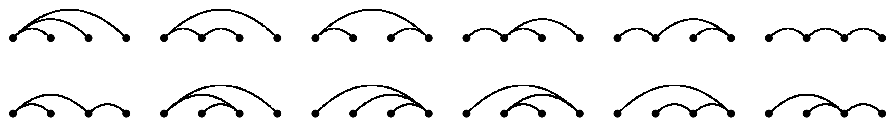

Figure 1: There are $t_3 = 12$ non-crossing trees with 4 vertices.

which we will henceforth refer to as the *ternary numbers*.[1]

The ternary numbers (A001764) are a well-known sequence, appearing often throughout the literature and having many interpretations. For example, Stanley [12, Exercise 5.45–46] documents the ternary numbers counting noncrossing trees with n edges (see Figure 1), recursively labeled forests on $[n]$, and plane ternary trees with $3n+1$ vertices. Bearing these combinatorial interpretations in mind, it is not difficult to see that the generating function for the ternary numbers $T(z) = \sum_{n=0}^{\infty} t_n z^n$ satisfies the cubic equation

$$T(z) = 1 + zT^3(z). \tag{1}$$

For example, equation (1) can be explained in terms of non-crossing trees by noting that every noncrossing tree is either empty or can be pictorially decomposed as

where each of $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$ is a noncrossing tree. Since each edge of a noncrossing tree is associated with a z in $T(z)$ and $T(z)$ is the generating function for $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$, we obtain the relationship in (1).

The ternary numbers can be considered an analogue of the Catalan number sequence A000108, $c_n = \frac{1}{n+1}\binom{2n}{n}$, whose generating function $C(z) = \sum_{n=0}^{\infty} c_n z^n$ satisfies $C(z) = 1 + zC^2(z)$ and has the closed form

$$C(z) = \frac{1 - \sqrt{1-4z}}{2z^2}.$$

Among many other interpretations, the Catalan numbers count *Dyck paths*, which are paths that start at $(0,0)$, end at $(2n, 0)$, use up steps U of the form $(1,1)$, down steps D of the form $(1,-1)$ and never go below the x-axis. For example, $UUDUDD$ is a Dyck path of length 6 and there are a total of $c_3 = 5$ possible Dyck paths of length 6. (See Figure 2.)

One way to generalize Dyck paths is to hold the length of the down step D as a parameter.

Definition 1. Let $t \geq 0$. A *t-Dyck path* is a path from $(0,0)$ to $(tn, 0)$ where $n \geq 0$ with step set $\{U(1,1)\} \cup \{D(1, -t+1)\}$ which never goes below the x-axis.

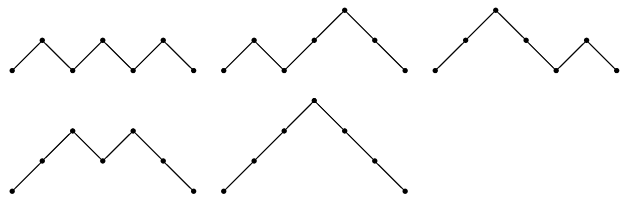

Figure 2: Of the $c_3 = 5$ Dyck paths of length 6 there are two with no hills.

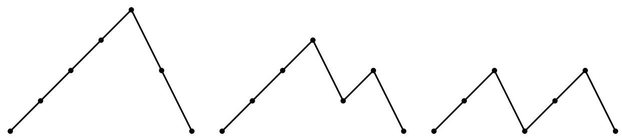

Figure 3: Of the $t_2 = 3$ ternary paths of length 6, there are two with no hills.

In the case where $t = 3$, we obtain so-called *ternary paths* which are counted by the ternary numbers. For example, $UUUDUD$ is one of the 3 possible ternary paths of length 6. (See Figure 3.) Using a decomposition argument similar to that used for noncrossing trees, one can show that $T(z)$ is the generating function for the number of ternary paths of length $3n$.

In general, it is known that the number of t-Dyck paths of length tn is the generalized Catalan number

$$\frac{1}{(t-1)n+1}\binom{tn}{n}.$$

Let $C_t(z)$ denote the generating function for the generalized Catalan numbers, that is,

$$C_t(z) = \sum_{n=0}^{\infty} \frac{1}{(t-1)n+1}\binom{tn}{n} z^n.$$

The fact that

$$C_t(z) = 1 + z(C_t(z))^t$$

follows immediately from the fact that every t-Dyck path is either the empty path or has the form $U\mathcal{P}_1 U\mathcal{P}_2 U\mathcal{P}_3 \cdots U\mathcal{P}_{t-1} D\mathcal{P}_t$ where each $\mathcal{P}_i$ is a t-Dyck path. Note that $C_0(z) = 1$, $C_1(z) = 1/(1-z)$, $C_2(z) = C(z)$ is the generating function for the Catalan numbers and

$$C_3(z) = T(z) \tag{2}$$

is the generating function for the ternary numbers $t_n = \frac{1}{2n+1}\binom{3n}{n}$. We also want to point out a fact that we use often in this paper. Adopting the conventional notation $[z^n]A(z)$ to

[1] The reader should not confuse our use of the term "ternary numbers" with any unrelated synonym for base 3 numbers.

denote the coefficient of z^n in a formal power series $A(z)$, we note that Graham et al. [7] showed via the cycle lemma that

$$[z^n]C_t^s(z) = \frac{s}{tn+s}\binom{tn+s}{n}. \tag{3}$$

In this more general context, we want to consider a few of the commonly studied statistics on Dyck paths. One common statistic for Dyck paths is the number of returns. A *return* on a t-Dyck path is a non-origin point on the path with ordinate 0. An *elevated t-Dyck path* is a t-Dyck path with exactly one return. Notice that an elevated t-Dyck path has the form $U\mathcal{P}_1U\mathcal{P}_2U\mathcal{P}_3\cdots U\mathcal{P}_{t-1}D$ where each $\mathcal{P}_i$ is a t-Dyck path. Therefore, we know that the generating function for the number of elevated t-Dyck paths is $z(C_t(z))^{t-1}$. In Section 2.1, we provide a limiting distribution for the number of returns on t-Dyck paths.

Another common statistic associated with Dyck paths is the number of peaks at a certain height. A *peak* in a t-Dyck path is an occurrence of the sequence UD and the *height* of a peak is the ordinate of the intersection of its two steps. For Dyck paths, a peak of height one is also known as a hill. More generally though, for $t \geq 1$, we define a *hill* in a t-Dyck path as a ground level subpath of the form $U^{t-1}D$. Note that a hill in a t-Dyck path is not generally the same as a peak of height $t-1$. For instance, there are two 3-Dyck paths of length 6 having no hills, while there is just one 3-Dyck path of length 6 having no peaks at height 2. (See Figure 3.) In this paper, we are more concerned with hills than peaks.

Any discussion of hills among Dyck paths will surely include references to the Fine numbers (A000957). The Fine number sequence, which begins

$$1, 0, 1, 2, 6, 18, 57, 186, 622, 2120, 7338, 25724, 91144, \ldots$$

has a number of interesting connections to the Catalan sequence. In this paper, f_n will denote the nth Fine number and $F(z) = \sum_{k=0}^{\infty} f_n z^n$ will denote the generating function for the Fine numbers, which in its closed form, is

$$F(z) = \frac{1-\sqrt{1-4z}}{z(3-\sqrt{1-4z})}.$$

The Fine number sequence is well documented in the literature, notably as the subject of some surveys [1, 2, 3, 4]. Among their numerous applications, the Fine numbers count Dyck paths of length $2n$ having no hills. For example, of the five Dyck paths of length 6, there are $f_3 = 2$ having no hills. See Figure 2. In Sections 3.2 and 4 of this article, we describe and interpret a natural analogue of the Fine number sequence for t-Dyck paths.

The Motzkin numbers (A001006) form another combinatorial sequence with deep connections to the Catalan numbers. The Motzkin numbers $\{m_n\}$ count *Motzkin paths*, which are paths from $(0,0)$ to $(n,0)$ with step set $\{U(1,1), L(1,0), D(1,-1)\}$ which never go below the x-axis. If we modify Motzkin paths by requiring that the level step take one of two possible colors, then $C^2(z)$ is the generating function for the number of resulting paths. With the further restriction of no level steps on the x-axis, the generating function is $F(z)$. In the effort to describe a Fine analogue for t-Dyck paths, there is a natural encounter with a Motzkin analogue for t-Dyck paths, as we will see in Section 4.

At various points in the paper, we will encounter Riordan arrays. A *Riordan array* $\mathcal{R} = (g(z), f(z))$ is an infinite lower triangular matrix whose kth column has a generating function given by $g(z)f^k(z)$ with $g(0) \neq 0$ and $f(0) = 0$. It is sometimes helpful to use Riordan arrays to summarize statistics on paths and other combinatorial objects and we will do so here when convenient. For example, the entry in row n and column k (where $n, k \geq 0$) of the Riordan array (A033184)

$$\big(C(z), zC(z)\big) = \begin{pmatrix} 1 & & & & & & \\ 1 & 1 & & & & & \\ 2 & 2 & 1 & & & & \\ 5 & 5 & 3 & 1 & & & \\ 14 & 14 & 9 & 4 & 1 & & \\ 42 & 42 & 28 & 14 & 5 & 1 & \\ 132 & 132 & 90 & 48 & 20 & 6 & 1 \\ & & & \cdots & & & \end{pmatrix}, \tag{4}$$

is $[z^n]z^kC^{k+1}(z)$, the coefficient of z^n in $z^kC^{k+1}(z)$. Note also that $[z^n]z^kC^{k+1}(z)$ is the number of Dyck paths of length $2(n+1)$ having exactly $k+1$ returns to the x-axis. Riordan arrays form a group under matrix multiplication known as the Riordan group, which has served as both a tool and object of combinatorial study [11]. The product of two Riordan arrays is given by

$$(g_1(z), f_1(z)) * (g_2(z), f_2(z)) = (g_1(z)g_2(f_1(z)), f_2(f_1(z)))\,.$$

One can also express the product of a Riordan array and an infinite column vector as

$$(g_1(z), f_1(z)) * h(z) = g_1(z)h(f_1(z)),$$

where $h(z)$ is the generating function for the entries of an infinite column vector.

In the study of returns and hills among t-Dyck paths, we will encounter the *negative binomial* distribution, which often shows up in combinatorial statistics [5], and we wish to remind the reader about the characteristics of a negative binomial distribution. In general, we say Y has the negative binomial distribution with parameters r, p and write $Y \sim \text{negbin}(r, p)$, when

$$P(Y = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r},$$

or equivalently, Y is the random variable for the trial at which the rth success occurs in a sequence of Bernoulli trials with success probability p. The distribution negbin(r, p) is known to have expected value $E(Y) = \frac{r(1-p)}{p}$ and variance $\text{Var}(Y) = \frac{r(1-p)}{p^2}$.

The article is organized as follows. In Section 2, we consider the asymptotic distribution for the number of returns and the number of hills among t-Dyck paths. In Section 3, we compute exact probabilities for the number of hills among Dyck paths of length $2n$ in order to determine the asymptotic proportion of Dyck paths having an even number of hills. We also introduce an analogue of the Fine numbers for the ternary numbers and determine the asymptotic probability that a ternary path has an even number of hills. In Section 4, we describe a Fine analogue for $C_t(z)$ in terms of colored Motzkin-like paths and we present a combinatorial identity involving its generating function.

2 Returns and hills via the Riordan group

In this section, we use the Riordan group to study the number of returns and hills among ternary paths. The ultimate goal is to determine the limiting distribution for the number of returns and for the number of hills among t-Dyck paths.

2.1 The number of returns among ternary paths, t-Dyck paths

First, we describe the generating function for the number of ternary paths having a given number of returns. Recall from equation (2) that $C_3(z) = T(z)$ is the generating function for the number of ternary paths.

Proposition 2. *Let $T(z)$ be the generating function for the number of ternary paths. For $0 \le m \le n$, $z^m T^{2m}(z)$ is the generating function for the number of ternary paths of length $3n$ having exactly m returns to the x-axis.*

Proof. Any ternary path of length $3n$ where $n \ge 1$ has at least one return, i.e., the empty path is the only path with no returns. For $n \ge 0$, any nonempty ternary path of length $3n$ with exactly m returns must has the form

$$(UP_1UQ_1D)(UP_2UQ_2D)(UP_3UQ_3D)\cdots(UP_mUQ_mD)$$

where each P_i, Q_i is a (possibly empty) ternary path. When counting ternary paths with $T(z)$, each D is associated with a z, so the generating function for the number of ternary paths with exactly m returns is $z^m T^{2m}(z)$. $\square$

If $r_{n,m}$ denotes the number of ternary paths of length $3n$ having m returns to the x-axis, then the following array (A109971)

$$(r_{n,m})_{n,m\ge 0} = \begin{pmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 0 & 2 & 1 & & & & \\ 0 & 7 & 4 & 1 & & & \\ 0 & 30 & 18 & 6 & 1 & & \\ 0 & 143 & 88 & 33 & 8 & 1 & \\ 0 & 728 & 455 & 182 & 52 & 10 & 1 \\ & & & \cdots & & & \end{pmatrix} = \big(1, zT^2(z)\big), \tag{5}$$

is a Riordan array whose row sums are the ternary numbers.

Theorem 3. *The probability that a randomly chosen ternary path has m returns approaches $\frac{4m}{3^{m+1}}$ as path length approaches infinity. The asymptotic distribution for the (nonzero) number of returns among ternary paths is the negative binomial distribution with parameters 2 and $2/3$. Hence, the expected number of returns among ternary paths approaches 2 with variance $3/2$.*

Proof. We proceed by calculating the associated probabilities. Let $Y_n \geq 0$ denote the number of returns for a ternary path of length $3n$. Let $p_m(n)$ be the probability that a randomly chosen ternary path of length $3n$ has m returns, that is, $p_m(n) = P(Y_n = m)$. We know from Proposition 2 that the generating function for the number of paths with m returns is $z^m T^{2m}(z)$. Making use of (3), we obtain

$$\begin{aligned} p_m(n) &= \frac{[z^n]z^m T^{2m}(z)}{[z^n]T(z)} \\ &= \frac{\frac{2m}{3(n-m)+2m}\binom{3(n-m)+2m}{n-m}}{\frac{1}{2n+1}\binom{3n}{n}} \\ &= \frac{2m(2n+1)(n-m+1)(n-m+2)\cdots(n-1)}{3(3n-m)(3n-m+1)(3n-m+2)\cdots(3n-1)} \\ &= \frac{2m(2n+1)(n-1)^{\underline{m-1}}}{3(3n-1)^{\underline{m}}} \end{aligned}$$

where $x^{\underline{j}} = x(x-1)(x-2)\cdots(x-j+1)$. It follows immediately that as $n \to \infty$,

$$p_m(n) \to \frac{4m}{3^{m+1}} = \binom{(m+1)-1}{1}\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^{(m+1)-2} = P(Y = m+1)$$

where $m \geq 0$ and $Y \sim \text{negbin}(2, \frac{2}{3})$.

So, we see that $p_m(n) \to P(Y = m+1)$ where $Y \sim \text{negbin}(2, \frac{2}{3})$. Notice that $Y_n = 0$ if and only if $n = 0$. So if we wish to restrict to the case where the number of returns is nonzero, we may consider $Y_n + 1$. Since $E(Y) = \frac{2(1/3)}{2/3} = 1$, we can conclude that $E(Y_n+1) = E(Y_n) + 1 \to E(Y) + 1 = 2$ and $\text{Var}(Y_n+1) = \text{Var}(Y_n) \to \text{Var}(Y) = \frac{E(Y)}{2/3} = \frac{3}{2}$.

We can also use the Riordan group to determine (exactly) $E(Y_n)$, the expected number of returns among ternary paths of length $3n$. Observe that the total number of returns among all ternary paths of length $3n$ is given by the generating function for the infinite column vector,

$$\begin{pmatrix} 1 \\ 0 & 1 \\ 0 & 2 & 1 \\ 0 & 7 & 4 & 1 \\ 0 & 30 & 18 & 6 & 1 \\ 0 & 143 & 88 & 33 & 8 & 1 \\ 0 & 728 & 455 & 182 & 52 & 10 & 1 \\ & & & \cdots \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ \vdots \end{pmatrix},$$

where the matrix on the left is given in equation (5). In terms of Riordan group multiplication, this product is

$$\begin{aligned} (1, zT^2(z)) * \frac{z}{(1-z)^2} &= \frac{zT^2(z)}{(1-zT^2(z))^2} \\ &= z\left(\frac{T}{1-zT^2}\right)^2 \\ &= zT^4. \end{aligned}$$

The last equality follows from the fact that

$$\frac{T}{1-zT^2} = \frac{T^2}{T-zT^3} = \frac{T^2}{1} = T^2. \tag{6}$$

Hence, the expected number of returns among ternary paths of length $3n$ is

$$\begin{aligned} E(Y_n) &= \frac{[z^n]zT^4}{[z^n]T} \\ &= \frac{\frac{4}{3(n-1)+4}\binom{3(n-1)+4}{n-1}}{\frac{1}{2n+1}\binom{3n}{n}} \\ &= \frac{4(3n+1)n}{(3n+1)(2n+2)} \\ &= \frac{2n}{n+1}. \end{aligned}$$

And we have verified that indeed, the expected number of returns for ternary paths approaches 2.

We may also compute (exactly) the variance $\operatorname{Var} Y_n$ by considering the square of the number of returns. Notice that

$$\begin{pmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 0 & 2 & 1 & & & & \\ 0 & 7 & 4 & 1 & & & \\ 0 & 30 & 18 & 6 & 1 & & \\ 0 & 143 & 88 & 33 & 8 & 1 & \\ 0 & 728 & 455 & 182 & 52 & 10 & 1 \\ & & & \cdots & & & \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 4 \\ 9 \\ 16 \\ 25 \\ 36 \\ \vdots \end{pmatrix} \begin{aligned} &= (1, zT^2(z)) * \frac{z+z^2}{(1-z)^3} \\ &= \frac{zT^2}{(1-zT^2)^3} + \frac{z^2T^4}{(1-zT^2)^3} \\ &= \frac{zT^4}{1-zT^2} + z^2T^7 \\ &= zT^5 + z^2T^7 \end{aligned}$$

which implies that

$$\begin{aligned} \operatorname{Var} Y_n &= EY_n^2 - (EY_n)^2 \\ &= \frac{[z^n]zT^5 + z^2T^7}{[z^n]T} - \left(\frac{2n}{n+1}\right)^2 \\ &= \frac{\frac{4}{3(n-1)+4}\binom{3(n-1)+4}{n-1}}{\frac{1}{2n+1}\binom{3n}{n}} \end{aligned}$$

and Var $Y_n \to \frac{3}{2}$ as $n \to \infty$.

$\square$

Now in the case of Dyck paths, the expected number of returns approaches 3 with variance 4 and the limiting distribution is negbin$(2, \frac{1}{2})$. Naturally, it would be nice to know the corresponding statistics for generalized Dyck paths, and it is possible do so given the fact that the generating function for the number t-Dyck paths having exactly m returns is $z^m C_t^{(t-1)m}(z)$. The necessary calculations that follow proceed almost exactly as above and we obtain the following result.

Theorem 4. *Let Y_n denote the random variable for the number of returns among t-Dyck paths of length tn and let $Y = \lim_{n\to\infty} Y_n$. Then*

$$Y \sim \text{negbin}\left(2, \frac{t-1}{t}\right)$$

Hence, the expected number of returns among nontrivial t-Dyck paths approaches $\frac{t+1}{t-1}$ with variance $\frac{2t}{(t-1)^2}$.

Proof. If Y_n denotes the number of returns among t-Dyck paths of length tn, then

$$P(Y_n = m) = \frac{(t-1)m((t-1)n+1)n^{\underline{m}}}{tn(tn-1)^{\underline{m}}}$$

which approaches

$$\frac{m(t-1)^2}{t^{m+1}} = \binom{m+1-1}{2-1}\left(\frac{t-1}{t}\right)^2\left(\frac{1}{t}\right)^{m-1}$$

as $n \to \infty$. $\square$

2.2 The number of hills among ternary paths, t-Dyck paths

We begin this subsection with the observation that the formula for the number of ternary paths of length $3n$ having exactly k hills is not as nice as the formula we saw for the number of returns in the previous subsection. This is due to the fact that counting hills among ternary paths requires an analogue to the Fine numbers.

Let $F_t(z)$ be the generating function for the number of t-Dyck paths of length tn having no hills. (In the case where $t = 2$, $F_t(z)$ is the generating function for the Fine numbers.) It follows that the generating function for the number of t-Dyck paths of length tn having exactly m hills is $z^m F_t^m(z)$ and the nth row of the Riordan array $(F_t(z), zF_t(z))$ gives a distribution of the t-Dyck paths of length tn by number of hills. Though there is no closed formula for the entries of $(F_t(z), zF_t(z))$, there is much that can be said about this Riordan array and in Sections 3.2 and 4, we indulge the reader with more detail in the case when $t = 3$. The following identity extends naturally from the relationship between the Catalan and Fine numbers:

$$C_t(z) = \frac{F_t(z)}{1 - zF_t(z)} \iff F_t(z) = \frac{C_t(z)}{1 + zC_t(z)} \tag{7}$$

Because this identity will come in handy now, we will use it here without proof. However, the interested reader is directed to Theorem 14 in Section 4 for further exploration and justification of (7).

Because there is no nice closed formula for the number of hills among t-Dyck paths, some care must be taken to give a closed formula for the exact probability that a ternary path has a given number of hills. Again, we defer to Section 3.2 for the necessary details regarding this computation. At present, however, we can use the Riordan group to compute the moments for the asymptotic distribution of hills among ternary paths, and in fact more generally, among t-Dyck paths.

Theorem 5. *The expected number of hills among t-Dyck paths approaches $\frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2}$ with variance $\frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2}\left(\frac{t^{t-1}+(t-1)^{t-2}}{t^{t-1}}\right)$ as path length approaches infinity. In particular, the expected number of hills among ternary paths approaches $\frac{4}{9}$ with variance $\frac{44}{81}$.*

Proof. Let X_n denote the random variable for the number of hills on a t-Dyck path of length tn. It is relatively straightforward to compute the first moment $E(X_n)$. The generating function for the total number of hills among t-Dyck paths is simply $zC_t^2(z)$ — this is true because every hill is preceded and followed by a t-Dyck path of some length. Hence,

$$\begin{aligned}
E(X_n) &= \frac{[z^n]zC_t^2}{[z^n]C_t} = \frac{[z^{n-1}]C_t^2}{[z^n]C_t} \\
&= \frac{\frac{2}{t(n-1)+2}\binom{t(n-1)+2}{n-1}}{\frac{1}{(t-1)n+1}\binom{tn}{n}} \\
&= \frac{2n(tn-(t-1))!((t-1)n+1)!}{(tn)!((t-1)n-(t-3))!} \\
&= \frac{2((t-1)n+1)^{\underline{t-2}}}{t(tn-1)^{\underline{t-2}}},
\end{aligned}$$

where $x^{\underline{j}} = x(x-1)(x-2)\cdots(x-j+1)$. Consequently, we see that $E(X_n) \to \frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2}$ as $n \to \infty$.

To compute the second moment $E(X_n^2)$, we need to compute the sum of the squares of the number of hills among t-Dyck paths of length tn. Here, we will make use of Riordan group algebra. Given that the generating function for the sequence of squares is $\frac{z+z^2}{(1-z)^3}$, we want to consider the sequence obtained by the Riordan product

$$\begin{aligned}
(F_t, zF_t) * \frac{z+z^2}{(1-z)^3} &= \frac{zF_t^2 + z^2F_t^3}{(1-zF_t)^3} \\
&= \frac{zF_t^2}{(1-zF_t)^3} + z^2C_t^3 \\
&= \frac{zC_t^3}{\frac{C_t}{1+zC_t}} + z^2C_t^3 \\
&= zC_t^2 + 2z^2C_t^3,
\end{aligned}$$

which gives us the sum of the squares of the number of hills among t-Dyck paths. Hence, the second moment is

$$\begin{aligned}
E(X_n^2) &= \frac{[z^n]zC_t^2 + 2z^2C_t^3}{[z^n]C_t} \\
&= E(X_n) + \frac{[z^{n-2}]2C_t^3}{[z^n]C_t} \\
&= E(X_n) + 2\frac{\frac{3}{t(n-2)+3}\binom{t(n-2)+3}{n-2}}{\frac{1}{(t-1)n+1}\binom{tn}{n}} \\
&= \frac{2((t-1)n+1)^{\underline{t-2}}}{t(tn-1)^{\underline{t-2}}} + \frac{6(n-1)n}{(t(n-2)+3)(t(n-2)+4)}\frac{((t-1)n+1)^{\underline{2t-4}}}{(tn)^{\underline{2t-4}}}
\end{aligned}$$

and $E(X_n^2) \to \frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2} + \frac{6}{t^2}\left(\frac{t-1}{t}\right)^{2t-4} = \frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2}\left(\frac{t^{t-1}+3(t-1)^{t-2}}{t^{t-1}}\right)$

Now, we may compute the variance of X_n as follows

$$\begin{aligned}\mathrm{Var}(X_n) &= E(X_n^2) - (E(X_n))^2\\ &= \frac{2((t-1)n+1)^{\underline{t-2}}}{t(tn-1)^{\underline{t-2}}} + \frac{6(n-1)n}{(t(n-2)+3)(t(n-2)+4)}\frac{((t-1)n+1)^{\underline{2t-4}}}{(tn)^{\underline{2t-4}}}\\ &- \left(\frac{2((t-1)n+1)^{\underline{t-2}}}{t(tn-1)^{\underline{t-2}}}\right)^2\end{aligned}$$

and notice that $\mathrm{Var}(X_n) \to \frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2} + \frac{6}{t^2}\left(\frac{t-1}{t}\right)^{2t-4} - \frac{4}{t^2}\left(\frac{t-1}{t}\right)^{2t-4} = \frac{2}{t}\left(\frac{t-1}{t}\right)^{t-2}\left(\frac{t^{t-1}+(t-1)^{t-2}}{t^{t-1}}\right)$ as $n \to \infty$. □

The relationship between the asymptotic mean and variance for the number of hills provides the initial evidence for the following conjecture.

Conjecture 6. Let X_n denote the random variable for the number of hills among t-Dyck paths of length tn and let $X = \lim_{n\to\infty} X_n$. Then

$$X \sim \mathrm{negbin}\left(2, \frac{t^{t-1}}{t^{t-1}+(t-1)^{t-2}}\right)$$

The evidence for this conjecture is based on verification of the result in the case where $t \leq 3$ (see Section 3) and the computation of the moments $E(X_n^m)$. To compute $E(X_n^m)$, we consider the mth power of the number of hills. The generating function for the power sequence $0^m, 1^m, 2^m, 3^m, \ldots$ is $\frac{e_m(z)}{(1-z)^{m+1}}$, where

$$e_m(z) = \sum_{k=1}^{m} e_{m,k} z^k$$

is the Eulerian polynomial of degree m and

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

are the Eulerian numbers (A008292). To find the mth moment for the number of hills, we consider the Riordan product

$$(F_t, zF_t) * \frac{e_m(z)}{(1-z)^{m+1}} = \sum_{k=1}^{m} e_{m,k} \frac{z^k F_t^{k+1}}{(1-zF_t)^{m+1}}$$

Now we may compute the expected value of X_n^m, making use of identity (7), as follows:

$$\begin{aligned}
E(X_n^m) &= \frac{[z^n]\displaystyle\sum_{k=1}^{m} e_{m,k}\frac{z^k F_t^{k+1}}{(1-zF_t)^{m+1}}}{[z^n]C_t} \\
&= \frac{\displaystyle\sum_{k=1}^{m} e_{m,k}\ [z^n]\ z^k C_t^{k+1}\left(\frac{C_t}{F_t}\right)^{m-k}}{[z^n]C_t} \\
&= \sum_{k=1}^{m} e_{m,k}\ \frac{[z^{n-k}]\ C_t^{k+1}(1+zC_t)^{m-k}}{[z^n]C_t} \\
&= \sum_{k=1}^{m}\left(\sum_{j=0}^{m-k} e_{m,k}\binom{m-k}{j}\frac{\frac{j+k+1}{t(n-(j+k))+j+k+1}\binom{t(n-(j+k))+j+k+1}{n-(j+k)}}{\frac{1}{tn+1}\binom{tn+1}{n}}\right) \\
&= \frac{1}{[z^n]C_t(z)}\sum_{k=1}^{m}\left(\sum_{j=0}^{m-k} e_{m,k}\binom{m-k}{j}[z^n]z^{j+k}C_t^{j+k+1}\right)
\end{aligned}$$

Using Mathematica, this quantity has been verified to match the moments of

$$\text{negbin}\left(2, \frac{t^{t-1}}{t^{t-1}+(t-1)^{t-2}}\right)$$

for $t \le 10$ and $m \le 20$.

3 Dyck paths and ternary paths with an even number of hills

In this section, we provide a response and extension to the following question posed by L. Shapiro [2] in a private communication: "What is the asymptotic proportion of Dyck paths having an even number of hills?" Our approach relies on properties of the Fine numbers, though it is also possible to use the techniques of the last section to answer the question. Along the way we uncover some interesting combinatorial relationships between the Fine and Catalan numbers sequences and one of their many generalizations.

3.1 Even number of hills among Dyck paths

Recall that $F(z) = \sum_{n=0}^{\infty} f_n z^n$ is the generating function for the Fine numbers. Our first observation is that the number of Dyck paths of length $2n$ having exactly k hills is given by the (n,k)-th entry of th Riordan array (A065600)

$$(F(z), zF(z)) = \begin{pmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 1 & 0 & 1 & & & & \\ 2 & 2 & 0 & 1 & & & \\ 6 & 4 & 3 & 0 & 1 & & \\ 18 & 13 & 6 & 4 & 0 & 1 & \\ 57 & 40 & 21 & 8 & 5 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{8}$$

where $F(z)$ is the generating function for the Fine numbers.

It has been shown that the proportion of Dyck paths having no hills approaches $4/9$. This result, which appears in [4], follows from something described there as the "hill-killer" identity:

$$C(z) + 1 = (z+2)F(z) \tag{9}$$

which implies that

$$c_n = f_{n-1} + 2f_n \tag{10}$$

for $n \geq 1$. Dividing both sides of (10) by c_n and using the fact that $c_{n-1}/c_n \sim 1/4$, one can reveal that

$$f_n/c_n \sim 4/9. \tag{11}$$

To find the proportion of Dyck paths having exactly one hill, we can make use of the following useful identity:

$$(2z + z^2)F^2(z) - (1+2z)F(z) + 1 = 0, \tag{12}$$

a result proven combinatorially by Cheon et al. [2]. If a_n is the number of Dyck paths of length $2n$ having one hill then $zF^2(z)$ is the generating function for a_n (A065601). Identity (12) implies that

$$2a_n + a_{n-1} = f_n + 2f_{n-1} \tag{13}$$

for $n \geq 1$. Using the fact that $f_n/c_n \sim 4/9$, $c_{n-1}/c_n \sim 1/4$ and (13), we find the proportion of Dyck paths having exactly one hill approaches $8/27$, that is,

$$a_n/c_n \sim 8/27. \tag{14}$$

This can be further generalized by multiplying both sides of (12) by $z^{k-1}F^{k-1}$, as

$$z^{k+1}F^{k+1} + 2z^kF^{k+1} - z^{k-1}F^k - 2z^kF^k + z^{k-1}F^{k-1} = 0 \tag{15}$$

to obtain a more general "hill-killer" recurrence.

Proposition 7. *Let $f_{n,k}$ denote the number of Dyck paths of semilength n having exactly k hills. That is, $f_{n,k} = [z^n]z^k F^{k+1}$. Then*

$$f_{n-1,k} + 2f_{n,k} - f_{n,k-1} - 2f_{n-1,k-1} + f_{n-1,k-2} = 0 \tag{16}$$

Proof. Follows directly from (15). □

We may now use the recurrence in Proposition 7 to answer Shapiro's question.

Theorem 8. *The proportion of Dyck paths having an even number of hills approaches* 5/8 *as the length of the paths approaches infinity.*

Proof. Dividing both sides of equation (16) by c_n leads to the following:

$$\frac{f_{n-1,k}}{c_{n-1}} \cdot \frac{c_{n-1}}{c_n} + 2\frac{f_{n,k}}{c_n} - \frac{f_{n,k-1}}{c_n} - 2\frac{f_{n-1,k-1}}{c_{n-1}} \cdot \frac{c_{n-1}}{c_n} + \frac{f_{n-1,k-2}}{c_{n-1}} \cdot \frac{c_{n-1}}{c_n} = 0 \tag{17}$$

Let $L_k := \lim_{n\to\infty} \frac{f_{n,k}}{c_n}$ be the asymptotic proportion of Dyck paths having k hills. Using the fact that $\lim_{n\to\infty} \frac{c_{n-1}}{c_n} = \frac{1}{4}$, equation (17) implies

$$\frac{1}{4}L_k + 2L_k - L_{k-1} - 2L_{k-1} \cdot \frac{1}{4} + L_{k-2} \cdot \frac{1}{4} \quad = \quad 0$$

which simplifies to

$$\frac{9}{4}L_k \quad = \quad \frac{3}{2}L_{k-1} - \frac{1}{4}L_{k-2} \tag{18}$$

for $k \geq 2$. Note that $L_0 = 4/9$ and $L_1 = 8/27$, as we saw in equations (11) and (14).

Solving the linear recurrence in (18) gives,

$$L(z) := \sum_{k=0}^{\infty} L_k z^k = \frac{4}{9 - 6z + z^2} = \frac{4/9}{(1 - z/3)^2},$$

the (asymptotic) probability distribution for the number of hills. Hence,

$$[z^k]L(z) = (k+1) \cdot \frac{4}{9} \left(\frac{1}{3}\right)^k$$

which is the negative binomial distribution with index 2 and parameter 2/3. (Incidentally, this is the same as the asymptotic distribution for the number of returns among ternary paths, as we saw in Section 2.1.)

To obtain the asymptotic probability that a Dyck path has an *even* number of hills, take the sum of the even indexed coefficients of $L(z)$, which is

$$\sum_{k=0}^{\infty} [z^{2k}]L(z) = \sum_{n=0}^{\infty} (2k+1) \cdot \frac{4}{9} \left(\frac{1}{9}\right)^k = \frac{5}{8}$$

□

3.2 Even number of hills among ternary paths

Here we extend our approach in Section 3.1 to answering Shapiro's question about Dyck paths and determine the asymptotic probability that a randomly chosen ternary path has an even number of hills. In this extension, we encounter and interpret the appropriate analogue of the Fine numbers for the ternary numbers. Recall from (2) that $C_3(z) = T(z)$ denotes the generating function for the ternary numbers

$$t_n = \frac{1}{2n+1}\binom{3n}{n}$$

The Riordan array (A110616)

$$(T(z), zT(z)) = \begin{pmatrix} 1 \\ 1 & 1 \\ 3 & 2 & 1 \\ 12 & 7 & 3 & 1 \\ 55 & 30 & 12 & 4 & 1 \\ 273 & 143 & 55 & 18 & 5 & 1 \\ 1428 & 728 & 273 & 88 & 25 & 6 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{19}$$

counts ternary paths ending at height $k \geq 0$. The (n,k)th entry is the number of ternary paths with length $3n+k$ and terminal height k.

Let $t_{n,k}$ denote the (n,k)th entry of (19), so that $t_{n,0} = t_n$. Then $t_{n,k} = [z^n]z^kT^{k+1} = [z^{n-k}]T^{k+1}$, which as a consequence of (2) and (3) is

$$t_{n,k} = \frac{k+1}{2n-k+1}\binom{3n-2k}{n-k}. \tag{20}$$

This fact can be used to show that $\lim_{n\to\infty} t_{n-1}/t_n = 4/27$.

The generating function $T^2(z)$ for the nonzero terms in the second column of (19) corresponds to A006013 which counts

- elevated ternary paths of length $3(n+1)$,
- ground level points on all ternary paths of length $3n$,
- leaves at the root of all noncrossing trees with $n+2$ vertices,
- even trees with $2(n+1)$ edges where the degree of the root is two, and
- a third of the leaves at the root of all ternary trees with $3(n+1)$ edges.

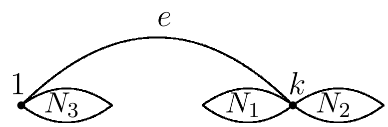

Figure 4: A nonempty noncrossing tree N with no leaf at its root must have the form above.

It is not hard to show, using (20), that the average number of ground level points on all ternary paths approaches 3, that is $\lim_{n\to\infty} t_{n+1,1}/t_n = 3$. Similarly, since the number of leaves at the root for all noncrossing trees of order $n+1$ is given by $t_{n,1} = [z^n]zT^2(z) = [z^{n-1}]T^2(z) = \frac{1}{n}\binom{3n-2}{n-1}$, one can show that the average number of leaves at the root $t_{n,1}/t_n$ approaches $4/9$.

Next we offer an analogue to the Fine number sequence for the ternary numbers. We will refer to a ternary path having no hills as a *ternary-Fine path*. Let $F_3(z)$ denote the generating function for the number of ternary-Fine paths of length $3n$. The sequence corresponding to $F_3(z)$ begins $1, 0, 2, 7, 34, 171, \ldots$ (A023053).

Theorem 9. *If $T(z)$ is the generating function for the number of ternary paths of length $3n$ and $F_3(z)$ is the generating function for the number of ternary paths of length $3n$ having no hills, then $T(z) = \dfrac{F_3(z)}{1 - zF_3(z)}$ and equivalently, $F_3(z) = \dfrac{T(z)}{1 + zT(z)}$.*

Proof. Every ternary path can be decomposed according to its number of hills. A ternary path with k hills would have the form

$$(P_1UUD)(P_2UUD)\cdots(P_kUUD)P_{k+1}$$

where each P_i is a (possibly empty) ternary path with no hills. The associated generating function is $z^kF_3^{k+1}(z)$ and summing over all $k \geq 0$ provides the result. $\square$

In addition to ternary-Fine paths of length $3n$, $F_3(z)$ also counts

- paths from $(0,0)$ to $(n,0)$ which never go below the x-axis and use step set $\{U(1,1)\} \cup \{L_k(1,0) : k = 1,2,3\} \cup \{D_k(1,-1) : k = 1,2,3\} \cup \{D(1,-2)\}$ but L_1, L_2, L_3, D_3 are not allowed to touch the x-axis,
- noncrossing trees with $n+1$ vertices having no leaves at the root,
- 2-plane trees with $n+1$ vertices, root labeled 2 and no leaves at the root

A detailed justification of the first bullet above will be presented in Section 4.

Remark 10. We may refer to the sequence associated with $F_3(z)$ as the *ternary-Fine numbers*. It is worth noting that using the Lindstrom-Gessel-Viennot nonintersecting path system argument for determinants [6, 9], one may conclude that the Hankel transform of the ternary-Fine numbers is equal to the Hankel transform of the ternary numbers (A051255).

Next we discuss the limiting distribution for the number of hills among ternary paths. Before proceeding, we present a few generating function identities for F_3 which prove useful in accomplishing this.

Theorem 11. *Let $F_3(z)$ be the generating function for the number of ternary paths of length $3n$ having no hills. Then*

$$zF_3(z)T^2(z) = F_3(z) + zF_3(z) - 1$$

Proof. First, rewrite this equation as

$$F_3 = 1 + zF_3(T^2 - 1).$$

We shall prove this last statement combinatorially in terms of noncrossing trees. F_3 counts noncrossing trees having no leaves at the root. Observe that any noncrossing tree N with no leaf at the root is either empty or it consists of the following components: (A) the "longest" edge e from the root (that is, the edge e from 1 to k, where k is as large as possible); (B) subtrees N_1, N_2 (not both empty) attached to the left and right of k respectively, and; (C) a (possibly empty) subtree N_3 with no leaves at its root attached to the root 1. See Figure 4. Since, z counts (A), $T^2 - 1$ counts (B) and F_3 counts (C), it is clear that $1 + zF_3(T^2 - 1)$ also counts the number of noncrossing trees having no leaf at the root. □

The two preceding theorems can be used to give the following recursive identity for the $F_3(z)$. Note that Theorem 12 can be thought of as an analogue of the Fine number identity (12) referenced in [2]. In Section 4, we will consider the generalization of this identity for t-Dyck paths.

Theorem 12. *Let $F_3(z)$ be the generating function for the number of ternary paths of length $3n$ having no hills. Then*

$$F_3^3(z - z^2 - z^3) + F_3^2(2z + 3z^2) - F_3(1 + 3z) + 1 = 0 \tag{21}$$

Proof. First, we establish the following preliminary result:

$$T^2 = F_3^2(1 - z - z^2) + F_3(2 + 3z) - 2 \tag{22}$$

We begin with the observation from Theorem 9 that

$$zT^2 = \frac{zF_3^2}{(1 - zF_3)^2}$$

which implies

$$\begin{aligned} zF_3^2 &= zT^2(1 - 2zF_3 + z^2F_3^2) \\ zF_3^2 &= zT^2 - 2z^2T^2F_3 + z^3T^2F_3^2) \\ zT^2 &= zF_3^2 + 2z^2T^2F - z^3T^2F_3^2. \end{aligned}$$

Now, using Theorem 11, we obtain

$$\begin{aligned} zT^2 &= zF_3^2 + 2z^2T^2F - z^3T^2F_3^2 \\ zT^2 &= zF_3^2 + 2zF_3 + 2z^2F_3 - 2z - z^2F_3^2 - z^3F_3^2 + z^2F_3 \\ T^2 &= F_3^2(1 - z - z^2) + F_3(2 + 3z) - 2. \end{aligned}$$

Multiplying both sides of equation (22) by zF and using Theorem 11 provides the result. □

The number of ternary paths of length $3n$ having exactly k hills is given by the (n,k)th entry of the Riordan matrix (A101371)

$$(F_3(z), zF_3(z)) = \left(\begin{array}{ccccccc} 1 & & & & & & \\ 0 & 1 & & & & & \\ 2 & 0 & 1 & & & & \\ 7 & 4 & 0 & 1 & & & \\ 34 & 14 & 6 & 0 & 1 & & \\ 171 & 72 & 21 & 8 & 0 & 1 & \\ 905 & 370 & 114 & 28 & 10 & 0 & 1 \\ & & \cdots & & & & \end{array}\right). \tag{23}$$

Notice that the row sums of (23) are the ternary numbers. Let $\tilde{f}_{n,k} := [z^n]z^kF_3^{k+1}$ denote the (n,k)th entry of (23), so that $\tilde{f}_{n,0} = \tilde{f}_n$.

Now Theorem 12 suggests a recursion for $\tilde{f}_{n,k}$. By multiplying both sides of (21) by $z^{k-1}F^{k-1}$, we have

$$F_3^{k+2}(z^k - z^{k+1} - z^{k+2}) + F_3^{k+1}(2z^k + 3z^{k+1}) - F_3^k(z^{k-1} + 3z^k) + z^{k-1}F_3^{k-1} = 0,$$

for all $k \geq 1$, and by taking the nth coefficient of this last expression, we obtain the following recursion for the triangle in (23):

$$\tilde{f}_{n+1,k+1} - \tilde{f}_{n,k+1} - \tilde{f}_{n-1,k+1} + 2\tilde{f}_{n,k} + 3\tilde{f}_{n-1,k} - \tilde{f}_{n,k-1} - 3\tilde{f}_{n-1,k-1} + \tilde{f}_{n-1,k-2} = 0 \tag{24}$$

for all $n \geq 1$ and $k \geq 2$.

We can now also answer Shapiro's question in the context of ternary paths.

Theorem 13. *The proportion of ternary paths having an even number of hills approaches* 125/169 *as the length of the paths approaches infinity.*

Proof. Let $\tilde{L}_k$ denote the asymptotic probability that a ternary path has k hills. That is, let

$$\tilde{L}_k = \lim_{n\to\infty} \frac{\tilde{f}_{n,k}}{t_n}.$$

We will use equation (22) to find a generating function for $\tilde{L}_k$. Divide both sides of equation (24) by t_n and let n tend to infinity.

Notice that as $n \to \infty$,

$$\frac{\tilde{f}_{n+1,k}}{t_n} = \frac{\tilde{f}_{n+1,k}}{t_{n+1}} \cdot \frac{t_{n+1}}{t_n} \to \frac{27}{4}\tilde{L}_k$$

and

$$\frac{\tilde{f}_{n-1,k}}{t_n} = \frac{\tilde{f}_{n-1,k}}{t_{n-1}} \cdot \frac{t_{n-1}}{t_n} \to \frac{4}{27}\tilde{L}_k.$$

With these two observations and using recursion (24), we have

$$\left(\frac{27}{4} - 1 - \frac{4}{27}\right)\tilde{L}_{k+1} + \left(2 + 3\cdot\frac{4}{27}\right)\tilde{L}_k - \left(1 + 3\cdot\frac{4}{27}\right)\tilde{L}_{k-1} + \frac{4}{27}\tilde{L}_{k-2} = 0 \tag{25}$$

$$\Rightarrow \frac{605}{108}\tilde{L}_{k+1} + \frac{22}{9}\tilde{L}_k - \frac{13}{9}\tilde{L}_{k-1} + \frac{4}{27}\tilde{L}_{k-2} = 0 \tag{26}$$

Using this recurrence, along with initial conditions $\tilde{L}_0, \tilde{L}_1$, and $\tilde{L}_2$, we can determine that the generating function for $\tilde{L}_k$ is

$$\tilde{L}(z) = \sum_{k=0}^{\infty} \tilde{L}_k z^k = \frac{324z + 405}{605 + 264z - 156z^2 + 16z^3} = \left(\frac{9}{11-2z}\right)^2 = \left(\frac{9/11}{1-2z/11}\right)^2 \tag{27}$$

(We refer the interested reader to Appendix A for a verification of the initial conditions $\tilde{L}_0, \tilde{L}_1$, and $\tilde{L}_2$.) Hence, the proportion of ternary paths having k hills is

$$[z^k]\tilde{L}(z) = (k+1)\cdot\frac{81}{121}\left(\frac{2}{11}\right)^k$$

which is the negative binomial distribution negbin$(2, 9/11)$ with mean $4/9$ and variance $44/81$. Finally, the proportion of ternary paths having an even number of hills is given by

$$\sum_{k=0}^{\infty}[z^{2k}]\tilde{L}(z) = \frac{81}{121}\sum_{k=0}^{\infty}(2k+1)\cdot\left(\frac{4}{121}\right)^k = \frac{125}{169}$$

$\square$

4 Generalized Fine and Motzkin analogues

In this section, we generalize the Fine analogue from the previous section and provide some interpretations. In particular, we give a generating function identity for which Theorem 12 is a special case. In connection with this generalized Fine analogue, we also introduce an analogue for the Motzkin numbers and discuss some of its interpretations.

Recall that a t-Dyck path is a path from $(0,0)$ to $(tn,0)$ with step set $\{U(1,1)\} \cup \{D(1,-t+1)\}$ which never goes below the x-axis and that $C_t(z)$ is the generating function for the number of t-Dyck paths of length tn, that is,

$$C_t(z) = \sum_{n=0}^{\infty} \frac{1}{(t-1)n+1}\binom{tn}{n} z^n.$$

Let's start with our Fine analogue.

Theorem 14. *Let $F_t(z)$ be the generating function for the number of t-Dyck paths having no hills. Then*

$$C_t(z) = \frac{F_t(z)}{1 - zF_t(z)}.$$

Proof. This follows from the fact that t-Dyck paths can be sorted according to their number of hills. A t-Dyck path with k hills would have the form

$$\mathcal{F}_0 U^{t-1} D \mathcal{F}_1 U^{t-1} D \mathcal{F}_2 \cdots U^{t-1} D \mathcal{F}_k,$$

where each $\mathcal{F}_i$ is a t-Dyck path with no hills. Hence, the generating function for the number of t-Dyck paths with k hills is $z^k(F_t(z))^{k+1}$ and summing from $k=0$ to infinity gives the result. □

Note that $F_1(z) = 1$, $F_2(z) = F(z)$ is the generating function for the Fine numbers and $F_3(z) = F_3(z)$ is the Fine analogue of Section 3.2. In this paper, we have seen non-obvious generating function identities for $F_2(z)$ and $F_3(z)$ which we want to understand in terms of $F_t(z)$. Recalling the approach to equation (12) which was used in [2], we first want to consider a different path interpretation of $F_t(z)$.

Consider the set of paths from $(0,0)$ to $(n,0)$ with step set $\{(1,1)\} \cup \{(1,-k+1) : k = 1,2,\dots t\}$ which never go below the x-axis. In the case when $t=2$, these are Motzkin paths. Hence, we will refer to these paths in general as $[t]$-Motzkin paths. There is a well-known correspondence between these $[t]$-Motzkin paths, plane $[t]$-trees and other combinatorial structures [12]. With an appropriate coloring, these objects become another manifestation of the generalized Catalan numbers.

Remark 15. The reader should be careful to distinguish the $[t]$-Motzkin paths described here from "t-Motzkin paths" that have been described elsewhere in the literature. For instance, in [2] the term "2-Motzkin path" is used to describe a path with step set

$$\{U(1,1), L_1(1,0), L_2(1,0), D(1,-1)\}.$$

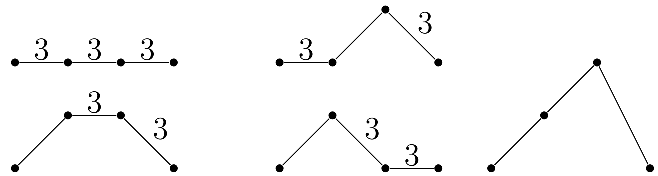

Figure 5: There are $27 + 9 + 9 + 9 + 1 = 55$ binomial [3]-Motzkin paths of length 3. Each up step of the form $(1, 1)$ can be just one color; each level step of the form $(1, 0)$ can be one of 3 colors; each down step of the form $(1, -1)$ can be one of 3 colors; each down step of the form $(1, -2)$ can be just one color.

However, here, the step set for $[t]$-Motzkin paths has cardinality $t+1$. If we let $M_t(z)$ denote the generating function for the number of $[t]$-Motzkin paths of length n, then $M_1(z) = C_1(z) = 1/(1-z)$, $M_2(z) = M(z)$ and

$$M_3(z) = 1 + z + 2z^2 + 5z^3 + 13z^4 + 36z^5 + 104z^6 + 309z^7 + \cdots,$$

is A036765. Based on a natural decomposition of $[t]$-Motzkin paths by initial step, one can easily see that

$$M_t(z) = \sum_{k=0}^{t} z^k (M_t(z))^k.$$

Definition 16. Let $t \geq 1$. A **binomial $[t]$-Motzkin path** is a path from $(0, 0)$ to $(n, 0)$ with step set

$$\bigcup_{k=0}^{t} \left\{ S_i(1, -k+1) : i = 1, 2, \ldots, \binom{t}{k} \right\}$$

$$\begin{aligned} &= \{S_1(1,1)\} \cup \{S_1(1,0), \ldots, S_{\binom{t}{1}}(1,0)\} \\ &\cup \{S_1(1,-1), \ldots, S_{\binom{t}{2}}(1,-1)\} \cup \ldots \\ \cdots \cup &\{S_1(1,-k+1), \ldots, S_{\binom{t}{k}}(1,-k+1)\} \cup \ldots \\ &\cup \{S_1(1,-t+3), \ldots, S_{\binom{t}{t-2}}(1,-t+3)\} \cup \\ \{S_1(1,-t+2), \ldots, S_{\binom{t}{t-1}}&(1,-t+2)\} \cup \{S_1(1,-t+1)\} \end{aligned}$$

which never goes below the x-axis. That is to say, there are 2^t possible steps to choose from. Each step has the form $S_i(1, -k+1)$ where $k = 0, 1, \ldots, t$. Each $S_i(1, -k+1)$ is distinguished with a color i where i is between 1 and $\binom{t}{k}$.

Example 17. See Figure 5 for a depiction of binomial [3]-Motzkin paths of length 3. There are 55 binomial [3]-Motzkin paths of length 3. Since each up step U is one color, each level step L is one of 3 colors, each down step D_1 of the form $(1,-1)$ is one of 3 colors and the down step D_2 of the form $(1,-2)$ can only be one color, we have 27 paths of the form LLL, 9 of the form LUD_1, 9 of the form ULD_1, 9 of the form UD_1L and 1 of the form UUD_2.

In general, binomial $[t]$-Motzkin paths are counted by generalized Catalan numbers.

Theorem 18. *The generating function for the number of binomial* $[t]$*-Motzkin paths of length* n *is* $(C_t(z))^t$*.*

Proof. To see this, suppose $M_t^*(z)$ is the generating function for binomial $[t]$-Motzkin paths. Each binomial $[t]$-Motzkin path is either (1) empty, (2) of the form $L\mathcal{P}_1$ where L is a level step and $\mathcal{P}_1$ is a colored t-Motzkin path, or (3) of the form $U\mathcal{P}_1U\mathcal{P}_2U\mathcal{P}_3\cdots U\mathcal{P}_{k-1}D_k\mathcal{P}_k$ where $k=2,\ldots,t$, U is an up step, each $\mathcal{P}_i$ is a binomial $[t]$-Motzkin path and D_k is a down step of the form $(1,-k+1)$. Since there are $\binom{t}{1}$ types of level step and $\binom{t}{k}$ types of down step of the form $(1,-k+1)$, we have

$$\begin{aligned}
M_t^*(z) &= 1+\binom{t}{1}zM_t^*(z)+\binom{t}{2}z^2(M_t^*(z))^2+\binom{t}{3}z^3(M_t^*(z))^3+\cdots \\
&= \sum_{k=0}^{t}\binom{t}{k}z^k(M_t^*(z))^k \\
&= (1+zM_t^*(z))^t,
\end{aligned}$$

which implies that

$$1+zM_t^* = (M_t^*)^{1/t}. \tag{28}$$

Since $(C_t(z))^t$ satisfies (28) and $M_t^*(0)=1=C_t(0)$, it must be that $M_t^*(z)=(C_t(z))^t=(C_t(z)-1)/z$. $\square$

There is of course a natural bijection ϕ between binomial $[t]$-Motzkin paths of length n and t-Dyck paths of length $t(n+1)$. Suppose P is a binomial $[t]$-Motzkin path. To create $\phi(P)$, first start with U^{t-1}. Then reading P from left to right, whenever we encounter a step $S_i(1,-k+1)$ with color i, we adjoin a unique subpath $P_{k,i}$ of length t consisting of k down steps of the form $D(1,-k+1)$ and $t-k$ U's. Since there are $\binom{t}{k}$ possibilities for the color i and $\binom{t}{k}$ possible paths $P_{k,i}$, we can indeed assign each color i with a unique path $P_{k,i}$. Notice that the adjoining of $P_{k,i}$ would add a net of $-t(k-1)$ to the height of the path. Complete $\phi(P)$ by adjoining the final step $D(1,-t+1)$. The resulting path $\phi(P)$ is a path of length $tn+t$ using $U(1,1)$'s and $D(1,-k+1)$'s, which never goes below the x-axis, i.e., a t-Dyck path. The inverse map ϕ^{-1} is achieved by reading a t-Dyck from left to right, t consecutive steps at a time (ignoring the first $t-1$ $U's$ and the last D) and then reversing the step to subpath assignment.

It is known that binomial [2]-Motzkin paths having no level steps on the x-axis are counted by the Fine numbers with generating function $F(z)=F_2(z)$. Given the bijection ϕ described above, it is clear that a 2-Dyck path with a hill would correspond to binomial

[2]-Motzkin paths with at least one level step on the x-axis. For $t \geq 3$, we want to give an interpretation of $F_t(z)$ in terms of binomial [t]-Motzkin paths which is consistent with the case when $t = 2$. Considering the map ϕ, one can imagine how a level step on the x-axis in a binomial [t]-Motzkin path could create a hill in a t-Dyck path, but clearly for $t \geq 3$, simply disallowing such level steps would not be sufficient to generate $F_t(z)$.

Theorem 19. *$F_t(z)$ is the generating function for the number of binomial [t]-Motzkin paths of length n with the following step restrictions:*

1. *no level steps allowed on the x-axis and*
2. *only $\binom{t-1}{k-1}$ colors are allowed for down steps of the form $(1, -k+1)$ which end on the x-axis.*

We may refer to the restricted binomial [t]-Motzkin paths described in Theorem 19 as **t-Fine paths.**

Proof. Let $H_t(z)$ be the generating function for the number of t-Fine paths of length n. First, we will show that $H_t(z)$ satisfies:

$$H_t(z) = 1 + zH_t(z)[(C_t(z))^{t-1} - 1] \tag{29}$$

To see this, observe that every t-Fine path is either empty or has a return. If the t-Fine path has a return, then it has the form

$$(U\mathcal{P}_1)(U\mathcal{P}_2)\cdots(U\mathcal{P}_{k-1})D_k\mathcal{F},$$

where each $\mathcal{P}_i$ is a binomial [t]-Motzkin path, D_k is the down step of the form $(1, -k+1)$ incident with the first return and $\mathcal{F}$ is a t-Fine path. Since D_k can be one of $\binom{t-1}{k-1}$ colors and $k = 2, 3, \ldots,$ or t, we have

$$\begin{aligned} H_t &= 1 + zH_t\left[\sum_{k=2}^{t}\binom{t-1}{k-1}(zC_t^t)^{k-1}\right] \\ &= 1 + zH_t\left[\sum_{k=1}^{t-1}\binom{t-1}{k}(C_t-1)^k\right] \\ &= 1 + zH_t[((C_t-1)+1)^{t-1} - 1] \\ &= 1 + zH_t[(C_t)^{t-1} - 1], \end{aligned}$$

which implies that $C_t^{t-1} = \dfrac{H_t - 1 + zH_t}{zH_t}$ or equivalently,

$$\frac{C_t - 1}{zC_t} = \frac{H_t + zH_t - 1}{zH_t}.$$

And on the other hand, we have from Theorem 14 that

$$\frac{C_t - 1}{zC_t} = \frac{\frac{F_t}{1-zF_t} - 1}{z\frac{F_t}{1-zF_t}} = \frac{F_t + zF_t - 1}{zF_t}.$$

Hence, given $H_t(0) = 1 = F_t(0)$, it must be that $H_t(z) = F_t(z)$. ☐

Finally, we observe that as a direct consequence of Theorem 14 and the fact that $C_t = 1 + zC_t^t$, we have the following generating function identity for $F_t(z)$.

Theorem 20.

$$1 + zF_t^t + \sum_{k=0}^{t-1}\left[\binom{t-1}{k} + \binom{t}{k+1}z\right](-1)^{k+1}z^kF_t^{k+1} = 0$$

Proof. Using Theorem 14, we obtain

$$\begin{aligned} C_t(z) &= 1 + zC_t^t(z) \\ \implies F_t(z) &= 1 - zF_t(z) + zC_t^{t-1}(z)F_t(z) \\ \implies F_t(z) &= \frac{(1 - zF_t(z))^t}{(1 - zF_t(z))^{t-1}} + \frac{z(F_t(z))^t}{(1 - zF_t(z))^{t-1}} \\ \implies F_t(z)(1 - zF_t(z))^{t-1} &= (1 - zF_t(z))^t + z(F_t(z))^t \end{aligned}$$

and the result follows from the Binomial Theorem. ☐

5 Conclusions and further work

Our study of returns and hills on t-Dyck paths has shown that, asymptotically speaking, the number of returns is easier to determine than the number of hills. We established that the asymptotic distribution for the number of returns among t-Dyck paths is negative binomial with parameters $(2, \frac{t-1}{t})$. Furthermore, in response to L. Shapiro's question, we have established that the probability that a randomly chosen Dyck path has an even number of hills approaches $\frac{5}{8}$ and the probability that a randomly chosen ternary path has an even number of hills approaches $\frac{125}{169}$. Hence, it is apparently more likely for a ternary path to have an even number of hills than for a Dyck path to have an even number of hills. In this case, we have used a recursive identity for the generalized Fine functions $F_2(z)$ and $F_3(z)$ to determine the exact probabilities needed to reach our conclusion. While such a recursive identity for $F_t(z)$ exists for all t (Theorem 20), the efficiency of its use decreases as t increases. In Section 2.2, we have provided strong evidence that the asymptotic distribution for the number of hills among t-Dyck paths is in fact negative binomial with parameters $\left(2, \frac{t^{t-1}}{t^{t-1}+(t-1)^{t-2}}\right)$, and it would be very desirable to make this evidence conclusive.

6 Acknowledgments

The authors wish to thank an anonymous referee for valuable suggestions which improved the quality of the paper.

A Initial conditions for $\tilde{L}_n$

Here are the computations to determine $\tilde{L}_0$, $\tilde{L}_1$, and $\tilde{L}_2$ as initial conditions for the recurrence in (26).

Equation (22) implies that, for $n \geq 1$,

$$t_{n+1,1} = \tilde{f}_{n+1,1} - \tilde{f}_{n,1} - \tilde{f}_{n-1,1} + 2\tilde{f}_{n,0} + 3\tilde{f}_{n-1,0} \tag{30}$$

Dividing both sides of (30) by t_{n+1}, we obtain:

$$\begin{aligned}
\frac{t_{n+1,1}}{t_{n+1}} &= \frac{\tilde{f}_{n+1,1}}{t_{n+1}} - \frac{\tilde{f}_{n,1}}{t_n} \cdot \frac{t_n}{t_{n+1}} - \frac{\tilde{f}_{n-1,1}}{t_{n-1}} \cdot \frac{t_{n-1}}{t_n} \cdot \frac{t_n}{t_{n+1}} + 2\frac{\tilde{f}_{n,0}}{t_n} \cdot \frac{t_n}{t_{n+1}} + 3\frac{\tilde{f}_{n-1,0}}{t_{n-1}} \cdot \frac{t_{n-1}}{t_n} \cdot \frac{t_n}{t_{n+1}} \\
\Rightarrow \frac{4}{9} &= \tilde{L}_1 - \tilde{L}_1\left(\frac{4}{27}\right) - \tilde{L}_1\left(\frac{4}{27}\right)^2 + 2\tilde{L}_0\left(\frac{4}{27}\right) + 3\tilde{L}_0\left(\frac{4}{27}\right)^2 \\
\Rightarrow \frac{4}{9} &= \tilde{L}_1\left(1 - \frac{4}{27} - \left(\frac{4}{27}\right)^2\right) + \tilde{L}_0\left(2\left(\frac{4}{27}\right) + 3\left(\frac{4}{27}\right)^2\right) \\
\Rightarrow \frac{4}{9} &= \frac{605}{81}\tilde{L}_1 + \frac{264}{81}\tilde{L}_0 \\
&\Rightarrow 324 = 605\tilde{L}_1 + 264\tilde{L}_0 \\
&\Rightarrow 2^2 \cdot 3^4 = 5 \cdot 11^2\tilde{L}_1 + 2^3 \cdot 3 \cdot 11\tilde{L}_0
\end{aligned}$$

Now, if we multiply equation (22) by zT^2, we obtain

$$zT^4 = zF_3^2T^2(1 - z - z^2) + zF_3T^2(2 + 3z) - 2zT^2$$

Using Theorem 11, we can simplify this to

$$\begin{aligned}
zT^4 + 2zT^2 &= (F_3^2 + zF_3^2 - F_3)(1 - z - z^2) + (F_3 + zF_3 - 1)(2 + 3z) \\
\Rightarrow zT^4 + 2zT^2 &= F_3^2(1 - 2z^2 - z^3) + F_3(1 + 6z + 4z^2) - (2 + 3z)
\end{aligned}$$

and therefore, we have for $n \geq 2$,

$$t_{n+2,3} + 2t_{n,1} = \tilde{f}_{n+1,1} - 2\tilde{f}_{n-1,1} - \tilde{f}_{n-2,1} + \tilde{f}_n + 6\tilde{f}_{n-1} + 4\tilde{f}_{n-2} \tag{31}$$

Dividing both sides of (31) by t_{n+1}, we obtain

$$
\begin{aligned}
2\left(\frac{4}{27}\right)+2\left(\frac{4}{9}\right)\left(\frac{4}{27}\right) &= \tilde{L}_1-2\tilde{L}_1\left(\frac{4}{27}\right)^2-\tilde{L}_1\left(\frac{4}{27}\right)^3 \\
&\quad +\tilde{L}_0\left(\frac{4}{27}\right)+6\tilde{L}_0\left(\frac{4}{27}\right)^2+4\tilde{L}_0\left(\frac{4}{27}\right)^3 \\
\Rightarrow \frac{8\cdot 13}{3^5} &= \tilde{L}_1\left[1-2\left(\frac{4}{27}\right)^2-\left(\frac{4}{27}\right)^3\right]+\tilde{L}_0\left[\frac{4}{27}+6\left(\frac{4}{27}\right)^2+4\left(\frac{4}{27}\right)^3\right] \\
\Rightarrow 2^3\cdot 3^4\cdot 13 &= 5\cdot 11^2\cdot 31\cdot \tilde{L}_1+2^2\cdot 11\cdot 131\cdot \tilde{L}_0
\end{aligned}
$$

(Note that (20) can be used to conclude that $\lim_{n\to\infty} t_{n+2,3}/t_{n+1}=2$.)

With two equations and two unknowns, we can determine $\tilde{L}_0$ and $\tilde{L}_1$.

$$\boxed{\tilde{L}_0=\frac{2^3}{11^2}=\frac{81}{121}} \tag{32}$$

$$\boxed{\tilde{L}_1=\frac{2^2\cdot 3^4}{11^3}=\frac{324}{1331}} \tag{33}$$

Now, equation (21) can be used to determine $\tilde{L}_2$. Equation (21) implies that for $n\geq 1$,

$$\tilde{f}_n+3\tilde{f}_{n-1}=\tilde{f}_{n+1,2}-\tilde{f}_{n,2}-\tilde{f}_{n-1,2}+2\tilde{f}_{n,1}+3\tilde{f}_{n-1,1} \tag{34}$$

Dividing both sides by t_{n-1}, we obtain

$$
\begin{aligned}
\tilde{L}_0\left(\frac{27}{4}\right)+3\tilde{L}_0 &= \tilde{L}_2\left(\frac{27}{4}\right)^2-\tilde{L}_2\left(\frac{27}{4}\right)-\tilde{L}_2+2\tilde{L}_1\left(\frac{27}{4}\right)+3\tilde{L}_1 \\
\frac{39}{4}\tilde{L}_0 &= \frac{605}{16}\tilde{L}_2+\frac{66}{4}\tilde{L}_1
\end{aligned}
$$

$$\Rightarrow \boxed{\tilde{L}_2=\frac{2^2\cdot 3^5}{11^4}=\frac{972}{14641}} \tag{35}$$

References

[1] E. Barcucci, E. Pergola, R. Pinzani, and S. Rinaldi, ECO method and hill-free generalized Motzkin paths, *Sém. Lothar. Combin.* **46** (2001/2), Article B46b. Available at https://www.emis.de/journals/SLC/wpapers/s46rinaldi.html.

[2] G. Cheon, S. Lee, and L. Shapiro, The Fine numbers refined, *European J. Combin.* **31** (2010), 120–128.

[3] E. Deutsch, Dyck path enumeration, *Discrete Math.* **204** (1999), 167–202.

[4] E. Deutsch and L. Shapiro, A survey of the Fine numbers, *Discrete Math.* **241** (2001), 241–265.

[5] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

[6] I. Gessel and G. Viennot, Binomial determinants, paths, and hook length formulae, *Adv. in Math.* **58** (1985), 300–321.

[7] R. Graham, D. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994.

[8] N. Gu, H. Prodinger, and S. Wagner, Bijections for a class of labeled plane trees, *Europ. J. Combin.* **31** (2010), 720–732.

[9] B. Lindström, On vector representations of induced matroids, *Bull. London Math. Soc.* **5** (1973), 85–90.

[10] *The On-Line Encyclopedia of Integer Sequences*, published electronically at http://oeis.org, 2016.

[11] L. Shapiro, S. Getu, W.-J. Woan, and L. Woodson, The Riordan group, *Discrete Math.* **34** (1991), 229–239.

[12] R. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, 1999.

2010 *Mathematics Subject Classification*: Primary 05A15; Secondary 05A16, 11B83.
Keywords: Catalan number, ternary number, Dyck path, Fine number, Motzkin number, hill, return, lattice path statistics.

(Concerned with sequences A000108, A000957, A001006, A001764, A006013, A008292, A023053, A033184, A036765, A051255, A065600, A065601, A101371, A109971, and A110616.)

Received February 5 2016; revised versions received May 19 2016; June 13 2016. Published in *Journal of Integer Sequences*, June 15 2016.

Return to Journal of Integer Sequences home page.