Paired patterns in lattice paths

Ran Pan
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112. USA
ran.pan.math@gmail.com

Jeffrey B. Remmel
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112. USA
remmel@math.ucsd.edu

Abstract

Let \( L_n \) denote the set of all paths from \([0,0]\) to \([n,n]\) which consist of either unit north steps \( N \) or unit east steps \( E \) or, equivalently, the set of all words \( L \in \{E, N\}^* \) with \( n \) \( E \)'s and \( n \) \( N \)'s. Given \( L \in L_n \) and a subset \( A \) of \([n] = \{1, \ldots, n\}\), we let \( ps_L(A) \) denote the word that results from \( L \) by removing the \( i \)th occurrence of \( E \) and the \( i \)th occurrence of \( N \) in \( L \) for all \( i \in [n] - A \), reading from left to right. Then we say that a paired pattern \( P \in L_k \) occurs in \( L \) if there is some \( A \subseteq [n] \) of size \( k \) such that \( ps_L(A) = P \). In this paper, we study the generating functions of paired pattern matching in \( L_n \).

1 Introduction

Let \( L_n \) denote the set of all paths from \([0,0]\) to \([n,n]\) which consist of either unit north \([0,1]\) steps or unit east \([1,0]\) steps. The six paths in \( L_2 \) are pictured at the top of Figure 2. Clearly,

\[
|L_n| = \binom{2n}{n}.
\]

We code elements in \( L_n \) as words over the alphabet \( \{N, E\} \) with \( n \) \( N \)'s and \( n \) \( E \)'s. Given \( L \in L_n \) and a subset \( A \) of \([n] = \{1, \ldots, n\}\), we let \( ps_L(A) \) denote the word that results from \( L \) by removing the \( i \)th occurrence of \( E \) and the \( i \)th occurrence of \( N \) in \( L \) for all \( i \in [n] - A \), reading from left to right. For example, suppose \( L = NEEENN \in L_3 \), then \( ps_L(\{1\}) = NE \), \( ps_L(\{2\}) = EN \), \( ps_L(\{3\}) = EN \), \( ps_L(\{1,2\}) = ps_L(\{1,3\}) = NEEN \), and \( ps_L(\{2,3\}) = EENN \). We shall think of a word in \( \{N, E\} \) with \( n \) \( N \)'s and \( n \) \( E \)'s as a paired pattern where the \( i \)th occurrence of \( E \) is paired with the \( i \)th occurrence of \( N \), reading from left to right, for \( i = 1, \ldots, n \).

Definition 1. Given a set of paired patterns \( \Gamma \subseteq L_k \) and word \( L \in L_n \), we say that

1. \( \Gamma \) occurs in \( L \) if there is an \( A \subseteq [n] \) of size \( k \) such that \( ps_L(A) \in \Gamma \),

2. there is a \( \Gamma \)-match in \( L \) starting at the \( j \)th paired step if \( ps_L(\{j, j+1, j+2 \cdots , j+k-1\}) \in \Gamma \) and

3. \( L \) avoids \( \Gamma \) if there is no \( \Gamma \)-matches in \( L \).

Alternatively, we can code a path \( L \) as a \( 2 \times n \) array \( T(L) \) where the bottom row of \( T \) consists of the positions of the east steps, reading from left to right, and the top row of \( T \) consists of the positions of the north steps, reading from left to right. We let \( T(L)_{k,1} \) denote the element
in the $k^{th}$ column of the bottom row of $T(L)$, and let $T(L)_{k,2}$ denote the element in the $k^{th}$ column of the top row. Given any $2 \times n$ array $S$ filled with pairwise distinct positive integers, let the reduction of $S$, $\text{red}(S)$, denote the $2 \times n$ array which results from $S$ by replacing the $i^{th}$ smallest integer in $S$ by $i$. An example of the reduction operation $\text{red}$ is pictured at the bottom of Figure 1.

![Diagram](image)

$T(L) = \begin{bmatrix} 3 & 5 & 6 & 7 & 9 & 11 \\ 1 & 2 & 4 & 8 & 10 & 12 \end{bmatrix}$

$p_{SL}([1, 4, 5]) = \text{ENNENE}$

$\text{red}\left( \begin{bmatrix} 3 & 7 & 9 \\ 1 & 8 & 10 \end{bmatrix} \right) = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 6 \end{bmatrix}$

Figure 1: The correspondence between paths and $2 \times n$ arrays.

It is then easy to see that given $L \in \mathcal{L}$ and $A \subseteq [n]$, the array associated with $p_{SL}(A)$ corresponds to the array obtained by taking the columns in $T(L)$ corresponding to $A$ and reducing. This process is pictured in Figure 1. This given, we can restate our pattern matching conditions in terms of $2 \times n$ arrays. That is, the $T_n$ denote the set of all $2 \times n$ arrays $T$ filled with the numbers $1, 2, \ldots, 2n$ such that the rows of $T$ are increasing reading from left to right. Given $T \in T_n$ and $A \subseteq [n]$, we let $T[A]$ be the array that results by removing the columns corresponding to elements in $[n] - A$. For example, if $T = T(L)$ is the array pictured in Figure 1 then $T[\{1, 4, 5\}]$ is pictured at the bottom left of Figure 1.

Then from the point of view of arrays in $T_n$, our paired pattern matching conditions can be stated as follows.

**Definition 2.** Given a set of $2 \times k$ arrays $\Gamma \subseteq T_k$ and a $2 \times n$ array $S \in T_n$, we say that

1. $\Gamma$ occurs in $S$ if there is an $A \subseteq [n]$ of size $k$ such that $\text{red}(S[A]) \in \Gamma$,

2. there is a $\Gamma$-match in $S$ starting at column $j$ if $\text{red}(S[\{j, j+1, j+2, \ldots, j+k-1\}]) \in \Gamma$ and

3. $S$ avoids $\Gamma$ if there is no $\Gamma$-matches in $S$.

Note that from this point of view, $\Gamma$-matches correspond naturally to consecutive patterns matches in $2 \times n$ arrays. Results about consecutive patterns in arrays can be found in [4]. We let $\text{mch}(L)$ denote the number of $\Gamma$-matches in $L$. If $\text{mch}(L) = 0$, then we will say that $L$ has no $\Gamma$-matches. If $\Gamma = \{P\}$ is a singleton, then we will write $P\text{-mch}(L)$ for $\text{mch}(L)$.

For example, there are six possible patterns of length four, as pictured in Figure 2 namely, $P_1 = \text{EENN}, P_2 = \text{ENEN}, P_3 = \text{NEEN}, P_4 = \text{ENNE}, P_5 = \text{NENE}, P_6 = \text{NNEE}$.

We note that paired patterns differ from classic consecutive patterns in words (e.g. [2] [10] [11]). Paired patterns actually describe relationships between paths and the diagonal $y = x$, the subdiagonal $y = x - 1$, and the superdiagonal $y = x + 1$. For our purposes, the set of Dyck
Figure 2: $L_2 = \{P_1, P_2, P_3, P_4, P_5, P_6\}$
The outline of this paper is as follows. In Section 2, we shall give the geometric interpretations of the number of $P_k$-matches in paths in $L$. In Section 3, we shall derive closed formulas for the generating functions $F_{P_k}(x, t)$ for $k = 1, \ldots, 6$ and explore some of the consequences of such formulas. In Section 4, we derive a number of formulas for $F_{\Delta}(x, t)$ for certain $\Delta \subseteq \{P_1, \ldots, P_6\}$. Finally, in Section 5, we discuss topics for future research such as finding bijections between paths with certain pattern matching condition and other known objects, extending the definition of paired patterns to Delannoy paths and finding generating functions $F_P(x, t)$ for paths $P$ of length greater than 4.

2 The geometric interpretation of the number of $P_k$-matches.

In this section, we shall give our geometric interpretations of $P_k$-matches for $k = 1, \ldots, 6$.

**Theorem 3.** Let $L \in L_n$. Then the number of $P_1$-matches in $L$ is the number of east steps below the subdiagonal $y = x - 1$. Hence, by symmetry, the number of $P_6$-matches in $L$ is the number of north steps above the superdiagonal $y = x + 1$.

**Proof.** Suppose that the $i$th east step in $L$ occurs below the subdiagonal $y = x - 1$ and that this step corresponds to the $j$th letter in the word $w_1 \ldots w_{2n}$ of $L$. Then it must be the case that the number of $E$’s in $w_1 \ldots w_j$ exceeds the number of $N$’s in $w_1 \ldots w_j$ by at least two. This means that when we restrict the diagram $T(L)$ to the letters $1, \ldots, j$, there are no elements in the $(i - 1)$th and $i$th columns of the bottom row. This means that $T(L)_{i-1,1} < T(L)_{i,1} < T(L)_{i-1,2} < T(L)_{i,2}$ so that red($T(L)$[$\{i - 1, i\}$]) matches the array for $P_1$. Hence each such east step represents a $P_1$-match in $L$.

On the other hand suppose red($T(L)$[$\{i - 1, i\}$]) matches the array for $P_1$. If $T(L)_{i,1} = j$, then in the word $w = w_1 \ldots w_{2n}$ of $L$, the $j$th $E$, reading from right to left, must be preceded by at most $i - 2$ north steps which means that the east step corresponding to $w_j$ is below the subdiagonal $y = x - 1$.

Given a path $L \in L_n$, we let $\text{bounce}^- (L)$ denote the number of points $[x, x]$ on $L$ such that the preceding step is a north step $N$ and the following step is an east step $E$. This means that the path bounces off the diagonal to the right. We let $\text{bounce}^+ (L)$ denote the number of points $[y, y]$ on $L$ which is preceded by an east step $E$ and followed by a north step $N$. This means that the path bounces off the diagonal to the left. For example, for the path $L$ pictured in Figure 3, the points $[6, 6]$, $[7, 7]$, and $[8, 8]$ are points preceded by a north step and followed by an east step so that $\text{bounce}^- (L) = 3$ and the point $[4, 4]$ is preceded by an east step and followed by a north step so that $\text{bounce}^+ (L) = 1$.

**Theorem 4.** Let $L \in L_n$. Then the number of $P_2$-matches in $L$ equals $\text{bounce}^- (L)$. Hence, by symmetry, the number of $P_3$-matches in $L$ equals $\text{bounce}^+ (L)$.

**Proof.** Consider the diagram $T(L)$ of $L$. Then a $P_2$-match in $L$ corresponds to a pair of columns $i - 1$ and $i$ such red($T(L)$[$\{i - 1, i\}$]) matches the array for $P_2$. This means that $T(L)_{i-1,1} < T(L)_{i-1,2} < T(L)_{i,1} < T(L)_{i,2}$. Now suppose that $T(L)_{i-1,2} = x$. It follows that all the elements in the columns to the right of $x$ must be greater than $x$ and all the elements in the columns to the left of $x$ must be less than $x$. Since $x > T(L)_{i-1,1}$ it follows that $x = 2(i - 1)$. Similarly, if $T(L)_{i,1} = y$, then all the elements in the columns to the right of $y$ must be greater than $y$ and all the elements in the columns to the left of $y$ must be less than $y$. Since $y < T(L)_{i-1,1}$ it
follows that \( y = 2(i - 1) + 1 \). This means that in the word of \( w = w_1 \ldots w_{2n} \) of \( L \), \( w_{2(i-1)} = N \) and is preceded by \( i \) east steps and \( i - 1 \) north steps so that point \( [i, i] \) is on the path of \( L \) and is preceded by a north step and followed by an east step.

Vice versa, if \([i, i]\) is on the path of \( L \) and is preceded by a north step and followed by an east step, then it is easy to see that in the array \( T(L) \) of \( L \), we must have that \( T(L)_{i-1,2} = 2(i - 1) \) and \( T(L)_{i,1} = 2(i - 1) + 1 \) so that \( \text{red}(T(L)[\{i - 1, i\}]) \) must match \( P_2 \).

Given a path \( L \in \mathcal{L}_n \), we let \( \text{cross}^h(L) \) denote the number of points \([x, x]\) on \( L \) such that the preceding step is an east step \( E \) and the following step is an east step \( E \). This means that the path crosses the diagonal horizontally. We let \( \text{cross}^v(L) \) denote the number of points \([y, y]\) on \( L \) which is preceded by a north step \( N \) and followed by a north step \( N \). This means that the path crosses the diagonal vertically. For example, for the path \( L \) pictured in Figure 3, there is a horizontal crossing of the diagonal at the point \([5, 5]\) so that \( \text{cross}^h(L) = 1 \) and there are vertical crossings at the points \([4, 4]\) and \([9, 9]\) so that \( \text{cross}^v(L) = 2 \).

**Theorem 5.** Let \( L \in \mathcal{L}_n \). Then the number of \( P_3 \)-matches in \( L \) equals \( \text{cross}^h(L) \). Hence, by symmetry, the number of \( P_4 \)-matches in \( L \) equals \( \text{cross}^v(L) \).

**Proof.** Consider the diagram \( T(L) \) of \( L \). Then a \( P_3 \)-match in \( L \) corresponds to a pair of columns \( i - 1 \) and \( i \) such that \( \text{red}(T(L)[\{i - 1, i\}]) \) matches the array for \( P_3 \). This means that \( T(L)_{i-1,2} < T(L)_{i-1,1} < T(L)_{i,1} < T(L)_{i,2} \). Now suppose that \( T(L)_{i-1,1} = x \). It follows all the elements in the columns to right of \( x \) must be greater than \( x \) and all elements in the columns to the left of \( x \) must be less than \( x \). Since \( x > T(L)_{i-1,2} \) it follows that \( x = 2(i - 1) \). Similarly, if \( T(L)_{i,1} = y \). It follows all the elements in the columns to the right of \( y \) must be greater than \( y \) and all the elements in the columns to the left of \( y \) must be less than \( y \). Since \( y < T(L)_{i-1,1} \) it follows that \( y = 2(i - 1) + 1 \). This means that in the word of \( w = w_1 \ldots w_{2n} \) of \( L \), \( w_{2(i-1)} = E \) and is preceded by \( i - 1 \) east steps and \( i \) north steps so that point \([i, i]\) is on the path of \( L \) and is preceded by an east step and followed by an east step.

Vice versa, if \([i, i]\) is on the path of \( L \) and is preceded by an east step and followed by an east step, then it is easy to see that in the array \( T(L) \) of \( L \), we must have that \( T(L)_{i-1,1} = 2(i - 1) \) and \( T(L)_{i,1} = 2(i - 1) + 1 \) so that \( \text{red}(T(L)[\{i - 1, i\}]) \) must match \( P_3 \).
3 Generating functions

Let \( F_i(x, t) = F_{P_i}(x, t) \) for \( i = 1, \ldots, 6 \). The goal of this section is to compute the generating functions \( F_k(x, t) \) for \( k = 1, \ldots, 6 \).

To obtain a recurrence for Dyck paths, the usual way is to factorize Dyck paths based on where it returns to the diagonal for the first time. Application of this decomposition can be found in many papers focused on lattice path enumeration such as [2], [3] and [11]. We shall show that similar ideas allow us to obtain recurrences for the number of \( P_k \)-matches.

3.1 Pattern \( P_1 \)

For pattern \( P_1 \), consider the ordinary generating function \( F_1(x, t) \) as follows,

\[
F_1(x, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x^{P_1\text{-mch}(L)}.
\]

We know for a path \( L \), \( P_1\text{-mch}(L) \) is equal to the number of east steps below subdiagonal \( y = x - 1 \). By our observation in the introduction, \( F_1(x, t) = F_6(x, t) \).

We split the analysis of \( P_1\text{-mch}(L) \) into two cases. Case 1 is when \( P_1\text{-mch}(L) = 0 \), that is, path \( L \) stays above \( y = x - 1 \). It is easy to see that the number of paths in \( \mathcal{L}_n \) above \( y = x - 1 \) is \( C_{n+1} = \frac{1}{n+2} \binom{2n+2}{n+1} \), the \((n+1)^{th}\) Catalan number. Case 2 is when \( P_1\text{-mch}(L) > 0 \), that is, path \( L \) has at least one east step below \( y = x - 1 \). Now consider the first time the path touches \( y = x - 1 \) and the first time after that where the path hits a point \([i, i]\) on the diagonal. It is easy to see that the two steps preceding the point \([i, i]\) must be north steps. An example of recurrence is pictured in Figure 4, where two boxes are the two positions mentioned above and three diagonal dots stand for a whatever path follows the second box.

Suppose the position of the first box has coordinates \([j, j-1]\), \( j \geq 1 \), clearly there are \( C_j \) ways to choose steps before reaching \([j, j-1] \). Similarly, suppose the position of the second box has coordinates \([i+j, i+j]\), \( i \geq 1 \), clearly there are \( C_i \) ways to choose steps between \([j, j-1] \) and \([i+j, i+j] \).

Since the ordinary generating function for Catalan numbers is

\[
C(x) = \sum_{n \geq 0} C_n x^n = \frac{1 - \sqrt{1 - 4x}}{2x} = 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 \cdots,
\]

Figure 4: An example of recurrence based on \( P_1 \).
it follows that
\[
F_1(x, t) = \sum_{n \geq 0} C_{n+1} t^n + \sum_{i \geq 1} \sum_{j \geq 1} C_i C_j x^i t^{i+j} F_1(x, t)
\]
\[
= \frac{C(t) - 1}{t} + \sum_{i \geq 1} C_i (xt)^i \sum_{j \geq 1} C_j t^j F_1(x, t)
\]
\[
= \frac{C(t) - 1}{t} + \sum_{i \geq 1} C_i (xt)^i (C(t) - 1) F_1(x, t)
\]
\[
= \frac{C(t) - 1}{t} + (C(xt) - 1)(C(t) - 1) F_1(x, t)
\]
and therefore by Equation 4,
\[
F_1(x, t) = \frac{C(t) - 1}{t (1 - (C(xt) - 1)(C(t) - 1))} = \frac{2x}{x \sqrt{1 - 4t + \sqrt{1 - 4xt + x - 1}}}
\]
Some initial terms of \(F_1(x, t)\) are
\[
F_1(x, t) = 1 + 2t + (x + 5)t^2 + (2x^2 + 4x + 14)t^3 + (5x^3 + 9x^2 + 14x + 42)t^4
\]
\[+(14x^4 + 24x^3 + 34x^2 + 48x + 132)t^5\]
\[+(42x^5 + 70x^4 + 95x^3 + 123x^2 + 165x + 429)t^6 + \cdots.\]
The number of paths in \(L_n\) avoiding \(P_1\) is \(C_{n+1}\), \((n+1)\)th Catalan number. In general, the number of paths having exactly \(k\) \(P_1\)-matches has the generating function as follows,
\[
\left. \frac{1}{k!} \frac{\partial^k F_1(x, t)}{\partial x^k} \right|_{x=0} \text{ or } \left. \frac{1}{k!} \frac{\partial^k F_1(x, t)}{\partial x^k} \right|_{x=0}.
\]
We evaluate the derivative at \(x = 0\) or when \(x = 0\) is a singularity of the derivative, we take the limit as \(x\) approaches zero. For example,
\[
\left. \frac{\partial F_1(x, t)}{\partial x} \right|_{x=0} = \left( \frac{-1 + \sqrt{1 - 4t + 2t^2}}{2t} \right)^2
\]
\[= t^2 + 4t^3 + 14t^4 + 48t^5 + 165t^6 + 572t^7 + 2002t^8 + \cdots.\]
The sequence 1, 4, 14, 48, 165, 572, 202, \ldots is sequence A002057 in the OEIS [14]. It has a number of combinatorial interpretations including the number of standard tableaux of shape \((n+2, n-1)\) and, with an offset of 4, the number of 123-avoiding permutations on \(\{1, 2, \cdots, n\}\) for which the integer \(n\) is in the fourth spot. It follows from the hook length formula for the number of standard tableaux that the number of paths \(L\) in \(L_n\) with exactly one east below the subdiagonal \(y = x-1\) equals \(4((2n-1)!)\/(n-2)!n+2)!)\) and is equal to the number of 123-avoiding permutations on \(\{1, 2, \cdots, n+2\}\) for which the integer \(n\) is in the fourth spot.
Similarly, one can obtain the generating function for the number of paths having exactly two east steps below the subdiagonal as follows,
\[
\left. \frac{1}{2!} \frac{\partial^2 F_1(x, t)}{\partial x^2} \right|_{x=0} = \frac{-(-1 + \sqrt{1 - 4t + 2t}) (-1 + \sqrt{1 - 4t + 2t})^2}{8t^2}
\]
\[= 2t^3 + 9t^4 + 34t^5 + 123t^6 + 440t^7 + 1573t^8 + 5642t^9 + \cdots.\]
The sequence 2, 9, 34, 123, 440, 1573, 5642, · · · is sequence A120989 in the OEIS [14]. The $n^{th}$
term in this sequence counts the level of the first leaf in preorder of a binary tree, summed over
all binary trees with $n - 2$ edges. Thus the number of paths $L$ in $\mathcal{L}_n$ with exactly two east steps
below the subdiagonal $y = x - 1$ equals the sum of the level of the first leaf in preorder over all
binary trees with $n - 2$ edges. We leave open the problem of giving a bijective proof of this fact.

Next, we shall answer the following question, for a random path $L \in \mathcal{L}_n$, what is the
expectation of $P_1$-mch($L$), or in other words, on average how many east steps of $L$ are below
$y = x - 1$? Consider that

$$\frac{\partial F_1(x, t)}{\partial x} \bigg|_{x=1} = -\frac{1 + \sqrt{1 - 4t + 2t}}{2(1 - 4t)^{3/2}}$$

$$= t^2 + 8t^3 + 47t^4 + 244t^5 + 1186t^6 + 5536t^7 + \cdots .$$

(6)

For example, a random $L \in \mathcal{L}_7$, expectation of $P_1$-mch($L$)

$$E[P_1\text{-mch}(L) : L \in \mathcal{L}_7] = \frac{5536}{14} \approx 1.63,$$

which implies in average there are roughly 1.63 east steps below $y = x - 1$.

In general, by the OEIS, the coefficient of $t^n$ in Equation (6) has formula

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

Using Stirling’s formula to approximate $n!$, we have

$$E[P_1\text{-mch}(L) : L \in \mathcal{L}_n] = \frac{(n + 1)(2n) - 4n}{2(2n)} \sim \frac{n + 1}{2} - \sqrt{\pi n},$$

(7)

which implies when $n$ is large, for a random path $L \in \mathcal{L}_n$, the expected number of east steps
that lie below $y = x - 1$ is $\frac{n + 1}{2} - \sqrt{\pi n}$.

The sequence 1, 8, 47, 244, 1186, 5536, · · · from Equation (6) is sequence A029760 and A139262
in the OEIS [14]. A029760 and A139262 count the total area under all the Dyck paths from
[0, 0] to [n, n], the total number of inversions in all 132-avoiding permutations of length $n$
and also total number of two-element anti-chains over all ordered trees on $n$ edges. Again we leave
open the problem of finding a bijective proof of these facts. We suspect that finding a bijective
proof is a challenge because Dyck paths, 132-avoiding permutations and ordered trees are all
Catalan objects while lattice paths in $\mathcal{L}_n$ are not.

Next, by manipulating $F_1(x, t)$ we can also find the number of paths having even number
many east steps below the subdiagonal $y = x - 1$. The generating function is as follows,

$$\frac{1}{2}(F_1(1, t) + F_1(-1, t))$$

$$= 1 + 2t + 5t^2 + 16t^3 + 51t^4 + 180t^5 + 622t^6 + 2288t^7 + \cdots .$$

Similarly, the generating function for the number of paths having odd number many east steps
below the subdiagonal $y = x - 1$ is

$$\frac{1}{2}(F_1(1, t) - F_1(-1, t))$$

$$= t^2 + 4t^3 + 19t^4 + 72t^5 + 302t^6 + 1144t^7 + 4643t^8 + \cdots .$$

Neither of the series correspond to entries in the OEIS [14].
3.2 Pattern $P_2$

For pattern $P_2$, $P_2$-mch($L$) counts the number of times $L$ bounces off the diagonal $y = x$ to the right, in other words, $P_2$-mch($L$) = bounce$^-$($L$). We shall study

$$F_2(x, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x^{P_2\text{-mch}(L)}. \quad (8)$$

As we observed in the introduction, $F_2(x, t) = F_5(x, t)$.

We shall consider two cases. Case 1 are the paths that start with an east step and Case 2 are the paths that start with a north step. We define

$$G_2(x, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } E} x^{P_2\text{-mch}(L)}$$

and

$$H_2(x, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } N} x^{P_2\text{-mch}(L)}.$$

Clearly, $F_2(x, t) = 1 + G_2(x, t) + H_2(x, t)$. For $H_2(x, t)$, we consider where is the first time a path starting with a north step crosses the diagonal $y = x$ horizontally. In the middle diagram of Figure 5 the three dots stand for a path starting with ‘$E$’ or an empty path. 

$$H_2(x, t) = \sum_{j \geq 1} C_j t^j (G_2(x, t) + 1) = (C(t) - 1)(G_2(x, t) + 1). \quad (9)$$

Similarly, for $G_2(x, t)$, we consider where is the first time a path starting with an east step crosses the diagonal $y = x$ vertically. In the right diagram of Figure 5 three dots stand for a path starting with ‘$N$’ or an empty path. Since we want to keep track of $P_2$-matches, here we need to introduce Catalan’s triangle $C_{i,j}$, which is the number of Dyck paths in $\mathcal{L}_{2j}$ with $i$ returns to the diagonal [2]. By [2], $C_{i,j}$ has generating function as follows,

$$C(x, t) = \sum_{i \geq 0} \sum_{j \geq 0} C_{i,j} x^i t^j = 1 + \frac{1 - \sqrt{1 - 4t}}{(\sqrt{1 - 4t} - 1)x + 2} \quad (10)$$
Then
\[ G_2(x, t) = \sum_{i \geq 0} \sum_{j \geq 1} C_{i,j} x^i t^j (H_2(x, t) + 1) = \frac{1 - \sqrt{1 - 4t}}{(\sqrt{1 - 4t} - 1)x + 2}(H_2(x, t) + 1). \] (11)

By Equation (9),
\[ G_2(x, t) = 1 - \sqrt{1 - 4t}((C(t) - 1)(H_2(x, t) + 1) + 1). \] (12)

We can then solve \( G_2(x, t) \) to obtain that
\[ G_2(x, t) = 1 - \sqrt{1 - 4t}((C(t) - 1)(H_2(x, t) + 1) + 1). \]

Then we have
\[ F_2(x, t) = 1 + G_2(x, t) + C(x, t) \]
\[ = 1 + G_2(x, t) + (C(t) - 1)(G_2(x, t) + 1) \]
\[ = (G_2(x, t) + 1)(C(t) - 1) \]
\[ = \frac{-(\sqrt{1 - 4t} - 1)((\sqrt{1 - 4t} - 1) + 1)}{2(\sqrt{1 - 4t}(t(x - 1) + 1) - t(x - 5) - 1)}. \]

A few initial terms of \( F_2(x, t) \) are
\[ F_2(x, t) = 1 + 2t + (x + 5)t^2 + (x^2 + 4x + 15)t^3 + (x^3 + 5x^2 + 16x + 48)t^4 \]
\[ + (x^4 + 6x^3 + 23x^2 + 62x + 160)t^5 + \cdots. \]

\( F_2(0, t) \) is the generating function for the number of paths that do not bounce off the diagonal to the right. One can compute that
\[ F_2(0, t) = \frac{2(t + \sqrt{1 - 4t} - 1)}{(\sqrt{1 - 4t} - 5) - \sqrt{1 - 4t} + 1} \]
\[ = 1 + 2t + 5t^2 + 15t^3 + 48t^4 + 160t^5 + 548t^6 + 1914t^7 + \cdots. \]

The sequence 1, 2, 5, 15, 48, 160, 548, 1914 · · · does not appear in the OEIS [14].

Similarly, we can compute the generating function of the number of paths that bounce at diagonal to right exactly one time. That is,
\[ \frac{\partial F_2(x, t)}{\partial x} \bigg|_{x=0} = \left( \frac{-1 + \sqrt{1 - 4t} + 2t}{1 - \sqrt{1 - 4t} + (-5 + \sqrt{1 - 4t})t} \right)^2 \]
\[ = t^2 + 4t^3 + 16t^4 + 62t^5 + 238t^6 + 910t^7 + \cdots. \]

The sequence 1, 4, 16, 62, 238, 910 · · · does not appear in the OEIS [14].
Also we could ask, for a random path \( L \in \mathcal{L}_n \), what is the expectation of \( P_2\text{-mch}(L) \), or in other words, on average how many times do \( L \) bounce at \( y = x \) to right? Consider that

\[
\frac{\partial F_2(x, t)}{\partial x} \bigg|_{x=1} = \left( -1 + \sqrt{1 - 4t + 2t} \right)^2 \quad (13)
\]

\[
= t^2 + 6t^3 + 29t^4 + 130t^5 + 562t^6 + 2380t^7 + \cdots . \quad (14)
\]

Coefficient of \( t^n \) in Equation (14) agrees with sequence A008549 of the OEIS [14] which counts the total area of all the Dyck excursions of length \( 2n - 2 \). By OEIS [14], the coefficient of \( t^n \) is given by the formula \( 4^{n-1} - \binom{2n-1}{n-1} \). Using Stirling’s formula to approximate \( n! \), one finds that

\[
\mathbb{E}[P_2\text{-mch}(L) : L \in \mathcal{L}_n] = \sum_{L \in \mathcal{L}_n} \text{bounce}^{-}(L) \bigg|_{L \in \mathcal{L}_n} = 4^{n-1} - \frac{\binom{2n-1}{n-1}}{2^n} \sim \frac{\sqrt{\pi n}}{4} - \frac{1}{2} \approx 0.443\sqrt{n},
\]

which implies when \( n \) is large, the expected number of times a random path \( L \in \mathcal{L}_n \) bounces off the diagonal to the right is roughly \( 0.443\sqrt{n} \).

Next, by manipulating \( F_2(x, t) \) we can also find the number of paths having even number of bounces off the diagonal to the right. The generating function is as follows,

\[
\frac{1}{2} (F_2(1, t) + F_2(-1, t)) = 1 + 2t + 5t^2 + 16t^3 + 53t^4 + 184t^5 + 654t^6 + 2368t^7 \cdots .
\]

Similarly, the generating function for the number of paths having odd number of bounces off the diagonal to the right is

\[
\frac{1}{2} (F_2(1, t) - F_2(-1, t)) = t^2 + 4t^3 + 17t^4 + 68t^5 + 270t^6 + 1064t^7 + 4181t^8 \cdots .
\]

Again, neither of the series correspond to sequences in the OEIS [14].

### 3.3 Pattern \( P_3 \)

For pattern \( P_3 \), as discussed in Section 2, \( P_3\text{-mch}(L) \) counts the number of times \( L \) crosses the diagonal \( y = x \) horizontally. We shall study

\[
F_3(x, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x^{P_3\text{-mch}(L)} . \quad (15)
\]

By our observation in the introduction \( F_3(x, t) = F_4(x, t) \).

Similar to the discussion of \( P_2 \), we consider two cases. Case 1 are the paths that start with a north step and Case 2 are the paths that start with an east step. We define

\[
G_3(x, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } E} x^{P_3\text{-mch}(L)}
\]

and

\[
H_3(x, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } N} x^{P_3\text{-mch}(L)} .
\]
Clearly, $F_3(x,t) = 1 + G_3(x,t) + H_3(x,t)$. Essentially, the way we shall decompose the paths in this case is the same as how we decomposed the paths for pattern $P_2$. For paths starting with a north step, we consider where is the first the path crosses the diagonal $y = x$ from left to right and then it is followed by a path starting with an east step or an empty path. Then

$$H_3(x,t) = \sum_{j \geq 1} C_j t^j (xG_3(x,t) + 1) = (C(t) - 1)(xG_3(x,t) + 1) \quad (16)$$

Similarly, for paths starting with an east step, we consider where is the first time that the path crosses the diagonal vertically. Then

$$G_3(x,t) = \sum_{j \geq 1} C_j t^j (H_3(x,t) + 1) = (C(t) - 1)(H_3(x,t) + 1). \quad (17)$$

Then by Equation (16),

$$H_3(x,t) = (C(t) - 1)(x(C(t) - 1)(H_3(x,t) + 1) + 1)$$

and we solve the formula above for $H_3(x,t)$

$$H_3(x,t) = \frac{(1 - C(t))(x(C(t) - 1) + 1)}{x(C(t) - 1)^2 - 1}$$

$$= -\frac{(2t + \sqrt{1 - 4t} - 1)(2t(x - 1) + (\sqrt{1 - 4t} - 1)x)}{2(2t^2(x - 1) + 2(\sqrt{1 - 4t} - 2)(tx - \sqrt{1 - 4tx} + x))}$$

Therefore,

$$F_3(x,t) = 1 + G_3(x,t) + H_3(x,t)$$

$$= 1 + C(t)H_3(x,t) + C(t) - H_3(x,t) - 1 + h_N(x,t)$$

$$= (H_3(x,t) + 1)C(t)$$

$$= \frac{2}{(2t(x - 1) + (\sqrt{1 - 4t} - 1)x + \sqrt{1 - 4t} + 1)}$$

A few initial terms of $F_3(x,t)$ are

$$F_3(x,t) = 1 + 2t + (x + 5)t^2 + (6x + 14)t^3 + (x^2 + 27x + 42)t^4 + (10x^2 + 110x + 132)t^5 + \cdots.$$ 

Next, we shall find the generating function of the number of paths crossing the diagonal horizontally exactly once.

$$\frac{\partial F_3(x,t)}{\partial x} \bigg|_{x=0} = -\frac{2(1 + \frac{1 + \sqrt{1 - 4t} + 2t}{1 + \sqrt{1 - 4t} - 2t})}{\sqrt{1 - 4t} + 2t}$$

$$= t^2 + 6t^3 + 27t^4 + 110t^5 + 429t^6 + 1638t^7 + \cdots.$$

The sequence $1, 6, 27, 110, 429, 1638, \cdots$ is sequence A003517 on OEIS [14]. This sequence has several combinatorial interpretations such as the number of standard tableaux of shape $(n+3, n-2)$ and the number of permutations of $\{1, \ldots, n+1\}$ with exactly one increasing subsequence of
length 3. It follows from the hook length formula for the number of standard tableaux that the
number of paths $L$ in $\mathcal{L}_n$ with exactly one horizontal crossing equal $6((2n+1)!)/(n-2)!(n+4)!$.

Similarly, the number of paths $L$ in $\mathcal{L}_n$ with exactly 2 horizontal crossings has the following
generating function:

$$\frac{1}{2!} \frac{\partial^2 F_3(x, t)}{\partial x^2} \bigg|_{x \to 0} = \frac{4 \left( -1 + \sqrt{1 - 4t + 2t} \right)}{(1 + \sqrt{1 - 4t - 2t})^2} = t^4 + 10t^6 + 65t^8 + 350t^{10} + 1700t^{12} + 7752t^{14} \cdots,$$

The sequence 1, 10, 65, 350, 1700, $\cdots$ is sequence A003519 on OEIS [14]. It counts the number
of standard tableaux of shape $(n-5,n-4)$ from which it follows that the number of paths $L$
in $\mathcal{L}_n$ with exactly 2 horizontal crossings equals $\frac{10}{n+6} \binom{2n+1}{n-4}$.

Also we could ask, for a random path $L \in \mathcal{L}_n$, what is the expectation of $P_3$-mch($L$), or in
other words, on average how many times does $L$ cross $y = x$ from left to right? In this case, we
have computed that

$$\frac{\partial}{\partial x} F_3(x, t) \bigg|_{x = 1} = \frac{-1 + \sqrt{1 - 4t + 2t}}{-2 + 8t} = t^2 + 6t^3 + 29t^4 + 130t^5 + 562t^6 + 2880t^7 + 9949t^8 + \cdots,$$

which means the total number of $P_3$-matches in paths in $\mathcal{L}_n$ is equal to the total number of
$P_2$-matches paths in $\mathcal{L}_n$.

Next we give a bijection that shows this fact. Since the total number of $P_3$-matches in paths
in $\mathcal{L}_n$ is half of total $\{P_3, P_3\}$-matches in paths in $\mathcal{L}_n$ and the total number of $P_2$-matches in
paths in $\mathcal{L}_n$ is half of total $\{P_2, P_3\}$-matches in paths in $\mathcal{L}_n$, we only need to show that the total
number of $\{P_2, P_3\}$-matches in paths in $\mathcal{L}_n$ is equal to the total number of $\{P_3, P_3\}$-matches in
paths in $\mathcal{L}_n$. In other words, we only need to show that the total number of times that all
the paths in $\mathcal{L}_n$ bounce off the diagonal is equal to the total number of times that all the paths in
$\mathcal{L}_n$ cross the diagonal.

By the reflection principle, we can construct a bijection between the set of paths in $\mathcal{L}_n$
crossing the diagonal $k$ times, denoted by $\mathcal{C}_{n,k}$ and the set of paths in $\mathcal{L}_n$ bouncing off
the diagonal $k$ times, denoted by $\mathcal{B}_{n,k}$. The procedure of the bijection is as follows. For any path
$L \in \mathcal{C}_{n,k}$, $L$ crosses the diagonal $k$ times and suppose $L$ touches the diagonal $j$ times at positions
$p_1, p_2, \cdots , p_j$, $j \geq k$. First we retain the part between $[0,0]$ and $p_1$ of the path and flip the
path between $p_1$ and $[n,n]$ along the diagonal, then we can get a new path $L_1$. At the second
step, we retain the part between $[0,0]$ and $p_2$ of the path $L_1$ and flip the part between $p_2$ and
$[n,n]$ along the diagonal, then we can get a new path $L_2$. We repeat the process above until we
acquire $L_j$. $L_j$ is a path in $\mathcal{B}_{n,k}$ because the procedure above transforms a crossing of $L$ into a
bouncing of $L_j$ and a bouncing of $L$ into a crossing of $L_j$. An example is pictured in Figure 6.
$L \in \mathcal{C}_{5,2}$ is mapped to $L_3 \in \mathcal{B}_{5,2}$ under the bijection.

Therefore,

$$\mathbb{E}[P_3$-mch($L$)] = \mathbb{E}[P_2$-mch($L$)] \sim \frac{\sqrt{n \pi}}{4} - \frac{1}{2} \approx 0.443 \sqrt{n}.$$

Next, by manipulating $F_3(x, t)$ we can also find the number of paths having even number
many horizontal crossings. The generating function is as follows,

\[
\frac{1}{2} (F_3(1, t) + F_3(-1, t)) = 1 + 2t + 5t^2 + 14t^3 + 43t^4 + 142t^5 + 494t^6 + 1780t^7 \cdots ,
\]

The sequence 1, 2, 5, 14, 43, 142, 494, \ldots is sequence A005317 in the OEIS [14] where no combinatorial interpretation is given. Thus we have given a combinatorial interpretation to this sequence.

Similarly, the generating function for the number of paths having odd number many horizontal crossings is

\[
\frac{1}{2} (F_3(1, t) - F_3(-1, t)) = t^2 + 6t^3 + 27t^4 + 110t^5 + 430t^6 + 1652t^7 + 6307t^8 \cdots ,
\]

in which coefficient of \( t^n \) also counts number of unordered pairs of distinct length-\( n \) binary words having the same number of 1’s according to A108958 in the OEIS [14]. We leave open the problem of giving a bijective proof of this fact.

4 Multivariate generating functions

In this section, we shall study multivariate generating functions for \( \Delta \)-matches for certain \( \Delta \subseteq \{P_1, \ldots, P_6\} \). Our choices for the \( \Delta \) that we consider are motivated by picking those pattern matching conditions which have the clearest geometric interpretations. Let

\[
F_\Delta(x, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} \left( \prod_{j \in \Delta} x_j^{P_j-\text{mch}(L)} \right),
\]

where \( \Delta \) is a subset of \{1, 2, 3, 4, 5, 6\}. We start by looking at the two elements sets that have symmetry, namely, \( \Delta = \{1, 6\} \), \( \Delta = \{2, 5\} \), and \( \Delta = \{3, 4\} \).

4.1 \( P_1 \) and \( P_6 \)

Pattern \( P_1 \) has one east step below \( y = x - 1 \) and \( P_6 \) has one east step above \( y = x + 1 \). We know that for a path \( L \in \mathcal{L}_n \), \( P_1\text{-mch}(L) \) and \( P_6\text{-mch}(L) \) are the numbers of east steps below \( y = x - 1 \) and above \( y = x + 1 \), respectively.
Clearly, if the derivative cannot be evaluated at zero, we take the limit. Paths avoiding $P$ generating function for coefficients of $P$ account. Case 1 are the paths that have no pattern matching of either $P_1$ or $P_6$ in path $L$ is $P_1$ and Case 3 are the paths $L$ such that the first pattern matching of either $P_1$ or $P_6$ in path $L$ is $P_6$. Then we have

$$F_{1,6}(x_1, x_6, t) = 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x_1^{P_1 \cdot \text{mch}(L)} x_6^{P_6 \cdot \text{mch}(L)}.$$\\

We use essentially the same ideas as in Section 3.1 to decompose the paths in $\mathcal{L}_n$ to obtain recurrences that will allow us to compute $F_{1,6}(x_1, x_6, t)$. In this case, we take three cases into account. Case 1 are the paths that have no $P_1$-match or $P_6$-match. In addition, we can see paths avoiding $P_1$ and $P_6$ must stay between $y = x - 1$ and $y = x + 1$. It is easy to see that if the word of such as path is $u_1 \ldots u_{2n}$, then either $u_{2i-1}u_{2i} = EN$ or $u_{2i-1}u_{2i} = NE$ for all $i$. Thus the number of paths in $\mathcal{L}_n$ bounded by $y = x - 1$ and $y = x + 1$ is $2^n$. Case 2 are the paths $L$ such that the first pattern matching of either $P_1$ or $P_6$ in path $L$ is $P_1$ and Case 3 are the paths $L$ such that the first pattern matching of either $P_1$ or $P_6$ in $L$ is $P_6$. Then we have

$$F_{1,6}(x_1, x_6, t) = \sum_{n \geq 0} 2^n t^n + \sum_{i \geq 1} \sum_{j \geq 1} (C_i 2^{-1} x_1^{i+j} + C_i C 2^{j-1} x_6^{i+j}) F_{1,6}(x_1, x_6, t)$$

$$= \frac{1}{1 - 2t} + \frac{t}{1 - 2t}(C(x_1 t) + C(x_6 t) - 2) F_{1,6}(x_1, x_6, t).$$\\

Then solving above equation for $F_{1,6}$, we have

$$F_{1,6}(x_1, x_6, t) = \frac{2x_1 x_6}{(-1 + \sqrt{1 - 4x_1 t}) x_6 + (1 - \sqrt{1 - 4x_6 t}) x_1 + 2x_1 x_6}$$

$$= 1 + 2t + (x_1 + x_6 + 4) t^2 + (2x_1^2 + 4x_1 + 2x_6^2 + 4x_6 + 8) t^3$$

$$+(5x_1^3 + 9x_1^2 + 12x_1 + 5x_6^3 + 9x_6^2 + 12x_6 + 2x_1 x_6 + 16) t^4 + \cdots.$$\\

Clearly, $F_{1,6}(x, 1, t) = F_{1,6}(1, x, t) = F_1(x, t)$. Next, we discuss coefficients of $x_1 t^n$ and $x_1 x_6 t^n$ in $F_{1,6}(x_1, x_6, t)$ which count the number of paths in $\mathcal{L}_n$ having exactly one $P_1$ pattern and avoiding $P_6$ and the number of paths in $\mathcal{L}_n$ having exactly one $P_1$ and exactly one $P_6$. In general, the generating function for coefficients of $x_1^j x_6^k$ is

$$\left. \frac{1}{j!k!} \frac{\partial^{j+k} F_{1,6}(x_1, x_6, t)}{\partial x_1^j \partial x_6^k} \right|_{x_1=0, x_6=0}, \quad (18)$$\\

Where if the derivative cannot be evaluated at zero, we take the limit.

By the symmetry of $P_1$ and $P_6$, the coefficient of $x_1 t^n$ in $F_{1,6}(x_1, x_6, t)$ equals the coefficient of $x_6 t^n$ in $F_{1,6}(x_1, x_6, t)$. By Equation (18), the generating function for the coefficients of $x_1 t^n$ in $F_{1,6}(x_1, x_6, t)$ equals

$$\frac{t^2}{(1 - 2t)^2} = t^2 + 4t^3 + 12t^4 + 32t^5 + 80t^6 + 192t^7 + 448t^8 \cdots.$$
The sequence 1, 4, 12, 32, 80, 192, ⋯ is A001787 in the OEIS [14]. The $n^{th}$ term of this sequence is $n2^{n-1}$ which means that the number of paths $L \in \mathcal{L}_n$ with exactly one east step below the subdiagonal $y = x - 1$ and no east step above the superdiagonal $y = x + 1$ equals $(n - 1)2^{n-2}$. This is easy to prove directly. That is, if $L$ is such a path, the one east that occurs below the subdiagonal $y = x - 1$ must arise by starting at a point $[a, a]$ on the diagonal where $0 \leq a \leq n - 2$ followed by a sequence $EENN$. If we remove this sequence from the word of $L$, we will end up with the word $u_1 \ldots u_{2n-4}$ of path $L' \in \mathcal{L}_{n-2}$ which has no east steps either below the subdiagonal $y = x - 1$ or above the superdiagonal $y = x + 1$. It is easy to see that in such a path $L'$ either $u_{2i-1}u_{2i} = EN$ or $u_{2i-1}u_{2i} = NE$ for all $i$. Hence there are $2^{n-2}$ such paths $L'$ so that the number of paths $L \in \mathcal{L}_n$ with exactly one east step below the subdiagonal $y = x - 1$ and no east step above the superdiagonal $y = x + 1$ equals $(n - 1)2^{n-2}$.

The generating function of the coefficients of $x_1x_6t^n$ in $F_{1,6}(x_1,x_6,t)$ equals

$$\frac{2t^4}{(1-2t)^3} = 2t^4 + 12t^5 + 48t^6 + 160t^7 + 480t^8 + 1344t^9 + \cdots.$$ 

The sequence 2, 12, 48, 160, 480, ⋯ is sequence A001815 in the OEIS [14]. We can show directly that the number of paths $L \in \mathcal{L}_n$ with exactly one east step below the subdiagonal $y = x - 1$ and exactly step above the superdiagonal $y = x + 1$ equals $(n - 2)(n - 3)2^{n-4}$. That is, if $L$ is such a path, then the one east that occurs below the subdiagonal $y = x - 1$ must arise by starting at a point $[a, a]$ on the diagonal where $0 \leq a \leq n - 2$ followed by a sequence $EENN$ and the one east that occurs above the superdiagonal $y = x + 1$ must arise by starting at a point $[b, b]$ on the diagonal where $0 \leq a \leq n - 2$ followed by a sequence $NNEE$. We have $n - 1$ choices for the point $[a, a]$. But these $n - 1$ choices lead to different situations according to different values of $a$. If $a = 0$ or $a = n - 2$, we have $n - 3$ choices to choose a point $[b, b]$ on the diagonal followed by a sequence $NNEE$. If $0 < a < n - 2$, there are $n - 4$ choices to choose a point $[b, b]$ on the diagonal followed by a sequence $NNEE$. So the total ways to choose positions of one $P_1$-match and one $P_6$-match is equal to $2(n - 3) + (n - 3)(n - 4) = (n - 2)(n - 3)$. We remove sequence $EENN$ and $NNEE$ from the word of $L$, we will end up with the word $u_1 \ldots u_{2n-8}$ of path $L' \in \mathcal{L}_{n-4}$ which has no east steps either below the subdiagonal $y = x - 1$ or above the superdiagonal $y = x + 1$. Hence there are $2^{n-4}$ such paths $L'$ so that the number of path $L \in \mathcal{L}_n$ with exactly one east step below the subdiagonal $y = x - 1$ and exactly east step above the superdiagonal $y = x + 1$ equals $(n - 2)(n - 3)2^{n-4}$.

If we are interested in counting lattice paths by the number of east steps below $y = x - 1$ or above $y = x + 1$, then we consider the following generating function:

$$F_{1,6}(x,x,t) = \frac{x}{-1 + x + \sqrt{1 - 4xt}}.$$ 

And also clearly,

$$\frac{\partial F_{1,6}(x,x,t)}{\partial x} \bigg|_{x=1} = 2 \frac{\partial F_{1}(x,t)}{\partial x} \bigg|_{x=1}$$

because the symmetry of $P_1$ and $P_6$. Then by Equation [12],

$$\mathbb{E} \{ \{P_1, P_6\}-\text{mch}(L) : L \in \mathcal{L}_n \} = 2\mathbb{E} \{ P_1-\text{mch}(L) : L \in \mathcal{L}_n \} \sim n + 1 - 2\sqrt{n\pi}.$$ 

Next, by manipulating $F_{1,6}(x_1,x_6,t)$ we can also find the number of paths having even
number many east steps below $y = x - 1$ or above $y = x + 1$. The generating function equals

$$\frac{1}{2} (F_{1,6}(1, 1, t) + F_{1,6}(-1, -1, t))$$

$$= 1 + 2t + 4t^2 + 12t^3 + 36t^4 + 132t^5 + 456t^6 + 1752t^7 \cdots .$$

Similarly, the generating function for the number of paths having odd number many east steps below the subdiagonal $y = x - 1$ or above $y = x + 1$ is

$$\frac{1}{2} (F_{1,6}(1, 1, t) - F_{1,6}(-1, -1, t))$$

$$= 2t^2 + 8t^3 + 34t^4 + 120t^5 + 468t^6 + 1680t^7 + 6530t^8 \cdots .$$

Neither of the two sequences above is recorded in the OEIS [14].

4.2 $P_2$ and $P_5$

In this subsection, we shall study

$$F_{2,5}(x_2, x_5, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x_2^{P_2 \text{-mch}(L)} x_5^{P_5 \text{-mch}(L)}.$$

$F_{2,5}(x_2, x_5, t)$ is the generating function which keeps track of the number of lattice paths by the number of times it bounces off the diagonal to the right or to the left. By the symmetry induced by reflecting paths about the diagonal discussed in the introduction, it is easy to see that $F_{2,5}(x_2, x_5, t)$ is a symmetric function in $x_2$ and $x_5$. It is also clear that $F_{2,5}(x, x, t)$ is the generating function which counts number of times a lattice path in $\mathcal{L}_n$ bounces off the diagonal $y = x$.

First we define

$$G_{2,5}(x_2, x_5, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } E} x_2^{P_2 \text{-mch}(L)} x_5^{P_5 \text{-mch}(L)} t^n$$

and

$$H_{2,5}(x_2, x_5, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } N} x_2^{P_2 \text{-mch}(L)} x_5^{P_5 \text{-mch}(L)} t^n.$$

Clearly,

$$F_{2,5}(x_2, x_5, t) = 1 + G_{2,5}(x_2, x_5, t) + H_{2,5}(x_2, x_5, t).$$

Here we employ the decomposition of paths used in Section 3.2 then we have

$$G_{2,5}(x_2, x_5, t) = \sum_{i \geq 0} \sum_{j \geq 1} C_{i,j} x_2^i t^j (H_{2,5}(x_2, x_5, t) + 1)$$

$$= (C(x_2, t) - 1)(H_{2,5}(x_2, x_5, t) + 1)$$

17
Solving the above formula for $H_{2,5}(x_2, x_5, t)$, we can obtain the generating functions of the coefficients of $G_{2,5}(x_2, x_5, t) + 1$

\[
H_{2,5}(x_2, x_5, t) = \sum_{i \geq 0} \sum_{j \geq 1} C_{i,j} x_5^i t^j (G_{2,5}(x_2, x_5, t) + 1)
\]

Therefore,

\[
G_{2,5}(x_2, x_5, t) = (C(x_2, t) - 1) ((C(x_5, t) - 1) (G_{2,5}(x_2, x_5, t) + 1) + 1).
\]

Solving the above formula for $G_{2,5}$ we have,

\[
G_{2,5}(x_2, x_5, t) = -\frac{(C(x_2, t) - 1) C(x_5, t)}{C(x_2, t)(C(x_5, t) - 1) - C(x_5, t)} 2(1 - x_5)t + (x_5 - 2) \left( 1 - \sqrt{1 - 4t} \right)
\]

\[
= \frac{-1}{1 + \sqrt{1 - 4t} + 2x_5 - 1} \left( 1 - \sqrt{1 - 4t} \right) - x_5 + \sqrt{1 - 4t} x_5 + 2(1 - x_2 x_5)t.
\]

Therefore,

\[
F_{2,5}(x_2, x_5, t) = 1 + G_{2,5}(x_2, x_5, t) + H_{2,5}(x_2, x_5, t)
\]

\[
= 1 + G_{2,5}(x_2, x_5, t) + (C(x_5, t) - 1)(G_{2,5}(x_2, x_5, t) + 1)
\]

\[
= C(x_5, t)(G_{2,5}(x_2, x_5, t) + 1)
\]

\[
= \left( 1 + \frac{1 - \sqrt{1 - 4t}}{2 - x_5 (1 - \sqrt{1 - 4t})} \right) \left( 1 - \frac{2(1 - x_5)t + (x_5 - 2) \left( 1 - \sqrt{1 - 4t} \right)}{1 + \sqrt{1 - 4t} + 2x_5 - 1} \left( 1 - \sqrt{1 - 4t} \right) - x_5 + \sqrt{1 - 4t} x_5 + 2(1 - x_2 x_5)t \right)
\]

A few initial terms of $F_{2,5}(x_2, x_5, t)$ are

\[
F_{2,5}(x_2, x_5, t) = 1 + 2t + (x_2 + x_5 + 4)t^2 + (x_2^2 + 4x_2 + x_5^2 + 4x_5 + 10)t^3
\]

\[
+ (x_2^3 + 5x_2^2 + 14x_2 + x_5^3 + 5x_2^2 + 14x_5 + 2x_2 x_5 + 28)t^4
\]

\[
+ (x_2^4 + 6x_2^3 + 21x_2^2 + 48x_2 + x_5^4 + 6x_2^3 + 21x_2^2 + 48x_5 + 2x_2 x_5 + 2x_2 x_5^2 + 12x_2 x_5 + 84)t^5
\]

\[
+ \cdots.
\]

By Equation (18), we can obtain the generating functions of the coefficients of $x_2 t^n$ in $F_{2,5}(x_2, x_5, t)$ which equals

\[
\left. \frac{\partial F_{2,5}(x_2, 0, t)}{\partial x_2} \right|_{x_2=0} = \frac{1 - \sqrt{1 - 4t} + 2t(-2 + \sqrt{1 - 4t} + t)}{2t^2}
\]

\[
= t^2 + 4t^4 + 14t^4 + 48t^5 + 165t^6 + 572t^7 + 7072t^8 + \cdots
\]

\[
= \left. \frac{\partial F_1(x, t)}{\partial x} \right|_{x=0}.
\]

This implies there exists a bijection between paths having exactly one $P_2$-match but no $P_3$-matches and paths having exactly one step below $y = x - 1$. We leave this as an open problem.
Similarly, we can get coefficients of \(x_2x_5t^n\),

\[
\frac{\partial^2 F_{2,5}(x_2, x_5, t)}{\partial x_2 \partial x_5} \bigg|_{x_2=x_5=0} = 2t^4 + 12t^5 + 56t^6 + 236t^7 + 948t^8 + 3712t^9 \cdots.
\]

The sequence 2, 12, 56, 236, 948, \cdots is not in the OEIS [14].

It is also the case that \(F_{2,5}(1, x, t) = F_5(x, t) = F_2(x, t) = F_{2,5}(x, 1, t)\) and

\[
F_{2,5}(0, 0, t) = 1 + 2t + 4t^2 + 10t^3 + 28t^4 + 84t^5 + 264t^6 \cdots
\]

\[
= 1 + 2C_1t + 2C_2t^2 + 2C_3t^3 + 2C_4t^4 + 2C_5t^5 + \cdots,
\]

where \(C_k\) is the \(k\)th Catalan number. \(F_{2,5}(x, x, t)\) is the generating function over paths \(L\) in \(L_n\) by the number of times \(L\) bounces off the diagonal.

\[
F_{2,5}(x, x, t) = \frac{1 - \sqrt{1 - 4t - t - x + x^2t}}{-x + (1 + x^2)t}
\]

\[
= 1 + 2t + 2(x + 2)t + 2(x + 2)t^2 + 2(x^2 + 4x + 5)t^3
\]

\[
+ 2(x^3 + 6x^2 + 14x + 14)t^4 + 2(x^4 + 8x^3 + 27x^2 + 48x + 42)t^5 \cdots.
\]

We take partial derivative of \(F_{2,5}(x, x, t)\) with respect \(x\) and evaluate at \(x = 1\),

\[
\left. \frac{\partial F_{2,5}(x, x, t)}{\partial x} \right|_{x=1} = \frac{\sqrt{1 - 4t}}{-1 + 4t} - \frac{1 - 2t}{-1 + 4t}
\]

\[
= \sum_{n \geq 2} \left( 2n \left( \frac{4^n}{2} - 2 \left( \frac{2n - 1}{n - 1} \right) \right) \right) t^n
\]

\[
= 2t^2 + 12t^3 + 58t^4 + 260t^5 + 1124t^6 + 4760t^6 + 19898t^8 + \cdots
\]

\[
= 2 \left. \frac{\partial F_{2,5}(x, t)}{\partial x} \right|_{x=1}.
\]

It follows that

\[
\mathbb{E}[\{P_2, P_5\}\text{-mch}(L) : L \in L_n] = 2\mathbb{E}[P_5\text{-mch}(L) : L \in L_n] \approx \frac{\sqrt{\pi n}}{2} - 1 \approx 0.886\sqrt{n}
\]

gives the expected number of times a path in \(L_n\) bounces off the diagonal.

\[
\frac{1}{2}(F_{2,5}(1, 1, t) + F_{2,5}(-1, -1, t))
\]

is the generating function of the number of lattice paths in \(L_n\) that bounce off the diagonal an even number of times. We have computed that

\[
\frac{1}{2}(F_{2,5}(1, 1, t) + F_{2,5}(-1, -1, t))
\]

\[
= 1 - \sqrt{1 - 4t + (-6 + 4\sqrt{1 - 4t}) t + 4t^2}
\]

\[
= 1 + 2 \sum_{n \geq 1} \left( \frac{2n - 2}{n - 1} \right) t^n
\]

\[
= 1 + 2t + 4t^2 + 12t^3 + 40t^4 + 140t^5 + 504t^6 + \cdots.
\]
The sequence 2, 4, 12, 40, 140, · · · is sequence A028329 in the OEIS [14]. It would be nice to have a direct combinatorial proof that the number of lattice paths in \( L_n \) that bounce off the diagonal an even number of times equals \( 2^{(2n-2)} \).

\[
\frac{1}{2} (F_{2,5}(1, 1, t) - F_{2,5}(-1, -1, t))
\]

is the generating function of the number of lattice paths \( L \) in \( L_n \) that bounce off the diagonal an odd number of times. We have computed that

\[
\frac{1}{2} (F_{2,5}(1, 1, t) - F_{2,5}(-1, -1, t)) = 2 \sum_{n \geq 2} \left( \frac{2n - 2}{n - 2} \right) t^n
\]

\[
= 2t^2 + 8t^3 + 30t^4 + 112t^5 + 420t^6 + 1584t^7 + \cdots.
\]

The sequence 2, 8, 30, 112, 420, 1584, · · · is sequence A162551 in the OEIS [14]. It would be nice to have a direct combinatorial proof of that the number of lattice paths in \( L_n \) that bounce off the diagonal an odd number of times equals \( 2^{(2n-2)} \).

4.3 \( P_3 \) and \( P_4 \)

We define

\[
F_{3,4}(x_3, x_4, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in L_n} x_3^{P_{3 \text{-mch}}(L)} x_4^{P_{4 \text{-mch}}(L)},
\]

where \( x_3 \) is used to keep track of the number of horizontal crossings and \( x_4 \) is used to keep track of the number of vertical crossings. Clearly, \( F_{3,4}(x_3, x_4, t) \) is symmetric in \( x_3 \) and \( x_4 \).

![Pattern P3 and P4](image)

Figure 9: Pattern \( P_3 \) and \( P_4 \).

We also define

\[
G_{3,4}(x_3, x_4, t) := \sum_{n \geq 1} t^n \sum_{L \in L_n \text{ starting with } E} x_3^{P_{3 \text{-mch}}(L)} x_4^{P_{4 \text{-mch}}(L)},
\]

and

\[
H_{3,4}(x_3, x_4, t) := \sum_{n \geq 1} t^n \sum_{L \in L_n \text{ starting with } N} x_3^{P_{3 \text{-mch}}(L)} x_4^{P_{4 \text{-mch}}(L)}.
\]

We employ the same decomposition of paths used in Section 4.2 for \( P_3 \) and \( P_4 \). Then

\[
H_{3,4}(x_3, x_4, t) = \sum_{j \geq 1} C_j t^j (x_3 G_{3,4}(x_3, x_4, t) + 1) = (C(t) - 1)(x_3 G_{3,4}(x_3, x_4, t) + 1)
\]

and

\[
G_{3,4}(x_3, x_4, t) = \sum_{j \geq 1} C_j t^j (x_4 H_{3,4}(x_3, x_4, t) + 1) = (C(t) - 1)(x_4 H_{3,4}(x_3, x_4, t) + 1).
\]
Combining the two equations above, we can then solve for $G_{3,4}$ to obtain that

$$G_{3,4}(x_3, x_4, t) = \frac{(1 - C(t))((x_4 C(t) - 1) + 1)}{x_3 x_4 (C(t) - 1)^2 - 1}$$

$$= -\frac{(-1 + \sqrt{1 - 4t} + 2t)(2t(1 - x_4) + (-1 + \sqrt{1 - 4t})x_4)}{2(-1 + \sqrt{1 - 4t})x_3 x_4 + 4(-2 + \sqrt{1 - 4t})x_3 x_4 t + 4(x_3 x_4 - 1)t^2}.$$  

Then

$$F_{3,4}(x_3, x_4, t) = 1 + G_{3,4}(x_3, x_4, t) + H_{3,4}(x_3, x_4, t)$$
$$= 1 + G_{3,4}(x_3, x_4, t) + (C(t) - 1)((x_3 G_{3,4}(x_3, x_4, t) + 1)$$
$$= (x_3 C(t) - x_3 + 1) G_{3,4}(x_3, x_4, t) + C(t)$$

$$= 1 - \sqrt{1 - 4t} \frac{2t}{-1 + \left(-1 + \frac{\sqrt{1 - 4t}}{2t}\right)^2 x_3 x_4}.$$  

A few initial terms of $F_{3,4}(x_3, x_4, t)$ are

$$F_{3,4}(x_3, x_4, t) = 1 + 2t + (x_3 + x_4 + 4)t^2 + (4x_3 + 4x_4 + 2x_3 x_4 + 10)t^3$$
$$+ (14x_3 + 14x_4 + x_3^2 + x_3 x_4 + 2x_3 x_4 + 12x_3 x_4 + 28)t^4$$
$$+ (48x_3 + 48x_4 + 2x_3^2 x_4 + 8x_3^2 x_4 + 8x_3 x_4 + 54x_3 x_4 + 84)t^5$$
$$+ \cdots.$$  

By symmetry, $F_{3,4}(1, x, t) = F_{4}(x, t) = F_3(x, t) = F_{3,4}(x, 1, t)$. It is also clear that $F_{3,4}(0, 0, t) = F_{2,5}(0, 0, t) = 2C(t)$, where $C(t)$ is the generating function of Catalan numbers, since if a path in $\mathcal{L}_n$ has no vertical or horizontal crossings, then the path either stays on or below the diagonal or on and above the diagonal.

By Equation \eqref{equation}, we see that coefficients of $x_3 t^n$ in $F_{3,4}(x_3, x_4, t)$ yield the generating function of the number of paths in $\mathcal{L}_n$ that have exactly one horizontal crossing and no vertical crossings. We have computed that

$$\frac{\partial F_{3,4}(x_3, 0, t)}{\partial x_3} \bigg|_{x_3=0} = \frac{1 - \sqrt{1 - 4t} + 2t(-2 + \sqrt{1 - 4t} + t)}{2t^2}$$
$$= t^2 + 4t^3 + 14t^4 + 48t^5 + 165t^6 + 572t^7 + 7072t^8 + \cdots$$

$$= \frac{\partial F_1(x, t)}{\partial x} \bigg|_{x=0}$$
$$= \frac{\partial F_{2,5}(x_2, 0, t)}{\partial x_2} \bigg|_{x_2=0},$$

which implies the number of paths in $\mathcal{L}_n$ having exactly one $P_3$-match and avoiding $P_4$ is equal to the number of paths in $\mathcal{L}_n$ having exactly one $P_2$-match and avoiding $P_5$. This can be verified by the bijection defined in Section \ref{section}. However, coefficients of $x_3 x_4 t^n$ in $F_{3,4}(x_3, x_4, t)$ is not equal to the coefficient of $x_2 x_5 t^n$ in $F_{2,5}(x_2, x_5, t)$. This is due to the fact that a path in $\mathcal{L}_n$
cannot cross the diagonal horizontally twice without crossing the diagonal vertically. We have computed that
\[
\frac{\partial^2 F_{3,4}(x_3, x_4, t)}{\partial x_3 \partial x_4} \bigg|_{x_3 = x_4 = 0} = 2 \frac{\partial F_3(x, t)}{\partial x} \bigg|_{x=0}
\]
\[
= -\frac{8t^2 \left( -1 + \sqrt{1 - 4t} + 2t \right)}{(\sqrt{1 - 4t} \left( 1 + \sqrt{1 - 4t} - 2t \right) )^3}
\]
\[
= 2t^2 + 12t^3 + 54t^4 + 220t^5 + 858t^6 + 3276t^7 + \cdots.
\]
The sequence 2, 12, 54, 220, 858, 3276, \cdots is Column 2 in A118920 and the exactly same interpretation is given by Emeric Deutsch in the OEIS [14].

For \( F_{3,4}(x, x, t) \), we can show that \( F_{3,4}(x, x, t) = F_{2,5}(x, x, t) \) by the bijection defined in Section 3.3, which gives us that for a random \( L \in \mathcal{L}_n \), the expectation of the number of crossings has asymptotic approximation as follows,
\[
\mathbb{E}[\{P_3, P_4\} \text{-mch}(L)] \sim \frac{\sqrt{\pi n}}{2} - 1 \approx 0.886\sqrt{n},
\]
and also
\[
\frac{1}{2} (F_{3,4}(1, 1, t) + F_{3,4}(-1, -1, t)) = \frac{1}{2} (F_{2,5}(1, 1, t) + F_{2,5}(-1, -1, t)).
\]

4.4 \( P_2 \) and \( P_4 \)

Due to space limitations, we shall consider only one more set of patterns of size 2, namely, \( \Delta = \{2, 4\} \). First, we define
\[
F_{2,4}(x_2, x_4, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x_2^{P_2 \text{-mch}(L)} x_4^{P_4 \text{-mch}(L)}.
\]
\( F_{2,4}(x_2, x_4, t) \) counts the number of lattice paths by the number of times it bounces off the diagonal to the right and by the the number of times it crosses the diagonal vertically. It follows that \( F_{2,4}(x, x, t) \) is the generating function over lattice paths \( L \in \mathcal{L}_n \) by the number of times \( L \) touches the diagonal with a north step. By symmetry, \( F_{3,5}(x, x, t) \) is also the generating function over lattice paths \( L \in \mathcal{L}_n \) by the number of times \( L \) touches the diagonal with an east step.

First we define
\[
G_{2,4}(x_2, x_4, t) := \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } E} x_2^{P_2 \text{-mch}(L)} x_4^{P_4 \text{-mch}(L)}
\]
We have computed that bounce off the diagonal to the right exactly one time but do not cross the diagonal vertically.

Combining the two equations above, we can solve them for interpretations such as the number of permutations on \( \{ 1, 2, \ldots, n \} \), and

\[
\frac{\partial F_{2,4}(x_2, x_4, t)}{\partial x_2} \bigg|_{x_2=0} = \frac{(-1 + \sqrt{1 - 4t})^3}{8t} = t^2 + 3t^3 + 9t^4 + 28t^5 + 90t^6 + 297t^7 + 1001t^8 + \cdots.
\]

The sequence 1, 3, 9, 28, 90, 297, \ldots is sequence A000245 in the OEIS [14] which has several interpretations such as the number of permutations on \( \{ 1, 2, \ldots, n + 2 \} \) that are 123-avoiding.
and for which the integer \( n \) is in the third spot, the number of lattice paths in \( L_{n-1} \) which touch but do not cross the \( y = x - 1 \) and the number of Dyck paths in \( L_n \) that start with ‘EE’.

Similarly, the coefficient of \( x_1 t^n \) in \( F_{2,4}(x_2, x_4, t) \) is the number of paths in \( L_n \) which have exactly one vertical crossing but never bounce off the diagonal to the right. We have computed that

\[
\frac{\partial F_{2,4}(0, x_4, t)}{\partial x_4} \bigg|_{x_4=0} = -\frac{(-3 + \sqrt{1 - 4t})(-1 + \sqrt{1 - 4t} + 2t)^2}{8t^2} = t^2 + 5t^3 + 19t^4 + 68t^5 + 240t^6 + 847t^7 + 3003t^8 + \cdots .
\]

The sequence 1, 5, 19, 68, 240, \cdots is sequence A070857 in the OEIS [14] which has no combinatorial interpretation. Thus we have given a combinatorial interpretation to this sequence.

The coefficient of \( x_2 x_4 t^n \) in \( F_{2,4}(x_2, x_4, t) \) is the number of paths in \( L_n \) which bounce off the diagonal to the right exactly once and cross the diagonal vertically exactly one. The corresponding generating function equals

\[
\frac{\partial^2 F_{2,4}(x_2, x_4, t)}{\partial x_2 \partial x_4} \bigg|_{x_2=x_4=0} = -\frac{(-1 + \sqrt{1 - 4t})^3(-2 + \sqrt{1 - 4t})(-1 + \sqrt{1 - 4t} + 2t)}{16t^2} = t^3 + 7t^4 + 32t^5 + 129t^6 + 495t^7 + 1859t^8 + \cdots ,
\]

which has no matches in the OEIS [14].

As we mentioned, \( F_{2,4}(x, x, t) \) is the generating function for the times of paths touching the diagonal \( y = x \) with a north step,

\[
F_{2,4}(x, x, t) = \frac{1 - \sqrt{1 - 4t} - t - x + x^2t}{-x + (1 + x)^2t} = 1 + 2t + (2x + 4)t^2 + (2x^2 + 8x + 10)t^3 + (2x^3 + 12x^2 + 28x + 28)t^4 + \cdots = F_{2,5}(x, x, t) = F_{3,4}(x, x, t).
\]

This fact can be also shown by constructing a bijection. Let \( C_{n,k} \) denote the set of paths in \( L_n \) that cross the diagonal \( k \) times and \( T_{n,k} \) denote the set of paths in \( L_n \) that touch the diagonal with a north step \( k \) times.

Next, we shall construct a bijection between \( T_{n,k} \) and \( C_{n,k} \), which is similar to the bijection defined in Section 3.3. For any path \( L \in T_{n,k} \), assume \( L \) touches the diagonal \( j \) times and these positions are denoted by \( \{p_1, p_2, \cdots , p_j\} \). We let \( p_0 = [0, 0] \) and \( p_{j+1} = [n, n] \). If \( p_i \) is a bouncing right position or \( p_i \) is a horizontal crossing position, we flip the part between \( p_{i-1} \) and \( p_i \) along the diagonal. Then we obtain a new path \( L' \). In this bijection, we can see that the number of crossings of \( L \) is equal to the number of north-touchings of \( L' \), and the number of north-touchings of \( L \) is equal to the number of crossings of \( L' \). For example, in Figure 11, \( L \) is mapped to \( L' \) and \( \{P_2, P_4\} - \text{mch}(L) = \{P_3, P_4\} - \text{mch}(L') = 3 \) and \( \{P_3, P_4\} - \text{mch}(L) = \{P_2, P_4\} - \text{mch}(L') = 2 \).

Because \( F_{2,4}(x, x, t) = F_{3,4}(x, x, t) \), for a random \( L \in L_n \),

\[
\mathbb{E}[\{P_2, P_4\} - \text{mch}(L)] \sim \frac{\pi n}{2} - 1 \approx 0.886\sqrt{n}
\]

24
Figure 11: \( L \) is mapped to \( L' \) by the bijection.

4.5 \( P_2, P_3, P_4 \) and \( P_5 \)

The last example of this section is a generating functions of a subset of \( \{1, \ldots, 6\} \) of size 4. That is, we shall study the generating function

\[
F_{2,3,4,5}(x_2, x_3, x_4, x_5, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n} x_2^{P_2-mch(L)} x_3^{P_3-mch(L)} x_4^{P_4-mch(L)} x_5^{P_5-mch(L)}.
\]

For convenience, in this subsection we use \( F_{2,3,4,5} \) to denote \( F_{2,3,4,5}(x_2, x_3, x_4, x_5, t) \), \( G_{2,3,4,5} \) to denote \( G_{2,3,4,5}(x_2, x_3, x_4, x_5, t) \) and \( H_{2,3,4,5} \) to denote \( H_{2,3,4,5}(x_2, x_3, x_4, x_5, t) \) where

\[
G_{2,3,4,5} := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } E k=2} \prod_{k=2}^5 x_k^{P_k-mch(L)}
\]

and

\[
H_{2,3,4,5} := 1 + \sum_{n \geq 1} t^n \sum_{L \in \mathcal{L}_n \text{ starting with } N k=2} \prod_{k=2}^5 x_k^{P_k-mch(L)}.
\]

Similar to the recurrences used in previous subsections, we have

\[
G_{2,3,4,5} = \sum_{i \geq 0} \sum_{j \geq 1} C_{i,j} x_2^i t^j (x_4 H_{2,3,4,5} + 1) \\
= (C(x_2, t) - 1)(x_4 H_{2,3,4,5} + 1)
\]

and

\[
H_{2,3,4,5} = \sum_{i \geq 0} \sum_{j \geq 1} C_{i,j} x_2^i t^j (x_3 G_{2,3,4,5} + 1) \\
= (C(x_5, t) - 1)(x_3 G_{2,3,4,5} + 1)
\]

Combining the two equations above, we can solve them for \( G_{2,3,4,5} \),

\[
G_{2,3,4,5} = \frac{(C(x_2, t) - 1)(x_4(C(x_5, t) - 1) + 1)}{x_3 x_4 (C(x_2, t) - 1)(C(x_5, t) - 1) - 1}.
\]
Then

\[
F_{2,3,4,5} = 1 + G_{2,3,4,5} + H_{2,3,4,5}
\]

\[
= 1 + G_{2,3,4,5} + (C(x_5, t) - 1) (x_3G_{2,3,4,5} + 1)
\]

\[
= C(x_5, t) (x_3G_{2,3,4,5} + 1) + (1 - x_3)G_{2,3,4,5}
\]

\[
= P(x_2, x_3, x_4, x_5, t)
\]

\[
Q(x_2, x_3, x_4, x_5, t),
\]

where

\[
P(x_2, x_3, x_4, x_5, t) =
\]

\[
(-1 + \sqrt{1 - 4t + 2t})x_3(-1 + x_4) + x_4 - \sqrt{1 - 4tx_4 - 2tx_4 + x_2(-(-1 + \sqrt{1 - 4t}(-2 + x_5) - 2t(-1 + x_5)) + 2\sqrt{1 - 4tx_5} + 2tx_5 - 2(-2 + \sqrt{1 - 4t + x_5})
\]

and

\[
Q(x_2, x_3, x_4, x_5, t) =
\]

\[
2 + (-1 + \sqrt{1 - 4t + 2t})x_3x_4 + (-1 + \sqrt{1 - 4t})x_5 + x_2(-1 + \sqrt{1 - 4t} - (-1 + \sqrt{1 - 4t + 2t})x_5).
\]

One can imagine that even a few initial terms of \(F_{2,3,4,5}(x_2, x_3, x_4, x_5, t)\) are very long so that we will not list them here. However, it is easy to verify that the constant coefficients of \(t^n\) is just \(2C_n\) because there are two sets of Dyck paths, namely, the ones that stay on or above the diagonal and the ones that stay on or below the diagonal.

By manipulating \(F_{2,3,4,5}(x_2, x_3, x_4, x_5, t)\), one is able to answer certain complicated enumerative problems, such as how many paths in \(\mathcal{L}_n\) are there that cross the diagonal vertically exactly once and horizontally exactly twice, and bounce off the diagonal to the right once but not to the left. The answer to this question has the generating function as follows,

\[
\frac{1}{2!} \frac{\partial^4 F_{2,3,4,5}(x_2, x_3, x_4, 0, t)}{\partial x_2 \partial x_3 \partial x_4^2} \bigg|_{x_2 = x_3 = 0} = \frac{(1 - \sqrt{1 - 4t})^5}{16}
\]

\[
= 2t^5 + 10t^6 + 40t^7 + 150t^8 + 550t^9 + 2002t^{10} + 7280t^{11} \cdots.
\]

Amazingly, the sequence 2, 10, 40, 150550, 2002, \cdots is twice the sequence A000344 in the OEIS [14], which has interpretations such as the number of paths in \(\mathcal{L}_{n-3}\) that touch but do not cross \(y = x - 2\) and the number of standard tableaux of shape \((n - 1, n - 5)\). We leave open the problem of finding a bijective proofs of these facts.

Next, we consider the formula \(F_{2,3,4,5}(x, x, x, x, t)\) which gives us the generating functions for the times of touching the diagonal,

\[
F_{2,3,4,5}(x, x, x, x, t) = \frac{1 + (x - 1) (-1 + \sqrt{1 - 4t})}{1 + (-1 + \sqrt{1 - 4t}) x}
\]

\[
= 1 + 2t + (4x + 2)t^2 + (8x^2 + 8x + 4)t^3 + (16x^3 + 24x^2 + 20x + 10)t^4 + (32x^4 + 64x^3 + 72x^2 + 56x + 28)t^5 + \cdots.
\]
Next, we want to ask for a random \( L \in \mathcal{L}_n \) how many times in average that \( L \) touches the diagonal. Applying the same idea that we used in previous sections, we see that

\[
\frac{\partial F_{2,3,4,5}(x,x,x,x,t)}{\partial x} \bigg|_{x=1} = \frac{(\sqrt{1-4t} - 1)^2}{4t - 1} = 4t^2 + 24t^3 + 116t^4 + 520t^5 + 2248t^6 + 9530t^7 + \cdots
\]

So for a random \( L \in \mathcal{L}_n \), the expectation of times \( L \) touches the diagonal is that

\[
E[\{P_2, P_3, P_4, P_5\}-\text{mch}(L)] = \frac{4^n - 4(2n-1)}{(2n)^n} \sim \sqrt{\pi n} - 2 \approx 1.772\sqrt{n}.
\]

Similarly, we can also obtain the generating functions for the number of paths touching the diagonal an even number of times or an odd number of times. We have computed that

\[
\frac{1}{2} \left( F_{2,3,4,5}(1,1,1,1,t) + F_{2,3,4,5}(-1,-1,-1,-1,t) \right)
\]

\[
= \frac{4t + \sqrt{1-4t}}{4t + 2\sqrt{1-4t} - 1}
\]

\[
= 1 + 2t + 2t^2 + 12t^3 + 34t^4 + 132t^5 + 468t^6 + 1752t^7 + 6530t^8 + \cdots.
\]

and

\[
\frac{1}{2} \left( F_{2,3,4,5}(1,1,1,1,t) - F_{2,3,4,5}(-1,-1,-1,-1,t) \right)
\]

\[
= -\frac{2 \left( -1 + \sqrt{1-4t} + 2t \right)}{-1 + 2\sqrt{1-4t} + 4t}
\]

\[
= 4t^2 + 8t^3 + 36t^4 + 120t^5 + 456t^6 + 1680t^7 + 6340t^8 + \cdots.
\]

Neither of the two series have matches in the OEIS [13].

By observing Equation (20) and (22), we find that coefficient of \( t^k \) in Equation (26) is equal to

\[
\begin{cases}
\text{coefficient of } t^k \text{ in } \frac{1}{2}(F_{1,6}(1,1,1,t) - F_{1,6}(1,1,1,t)), \text{ if } k \text{ is even} \\
\text{coefficient of } t^k \text{ in } \frac{1}{2}(F_{1,6}(1,1,1,t) + F_{1,6}(1,1,1,t)), \text{ if } k \text{ is odd.}
\end{cases}
\]

This is because all the six patterns in \( \mathcal{L}_2 \) are mutually exclusive. For any path \( L \in \mathcal{L}_k \), \( \mathcal{L}_2\)-\text{mch}(L) = k - 1, which implies that

\[
\{P_1, P_6\}-\text{mch}(L) + \{P_2, P_3, P_4, P_5\}-\text{mch}(L) = k - 1.
\]

If \( k \) is odd, \( \{P_1, P_6\}-\text{mch}(L) \) and \( \{P_2, P_3, P_4, P_5\}-\text{mch}(L) \) have the same parity and otherwise, they do not.
5 Future research

In this paper, we computed the generating functions $F_{P_k}(x, t)$ for $k = 1, \ldots, 6$ and $F_{\Delta}(x, t)$ for certain selected $\Delta \subseteq \{1, \ldots, n\}$. In a subsequent paper, we will systematically compute $F_{\Delta}(x, t)$ for all sets of size two. There are only nine such generating functions up to symmetry and we have computed five of them since $F_{P_2, P_3}(x_2, x_3, t)$ is a specialization of $F_{2,3,4,5}$. The ones that we have not computed in the paper are represented by $F_{P_1, P_2}(x_1, x_2, t)$, $F_{P_1, P_3}(x_1, x_3, t)$, $F_{P_1, P_4}(x_1, x_4, t)$, and $F_{P_1, P_5}(x_1, x_5, t)$. As a special case for pattern $P_1, P_2$, $F_{1,2}(x, x, t)$ keeps track of the number of paths in $L_n$ that have $k$ steps below the diagonal. For any fixed $k$, the coefficient of $x^k t^n$ in $F_{1,2}(x, x, t)$ is also Catalan number $C_n$, shown by Chung and Feller [1].

We shall explore these generating functions in a subsequent paper where we will also add some additional parameters which keep track of both the area below the diagonal and the area above the diagonal in path in $L_n$.

There are many interesting bijective problems that arise from our results. For example, in Section 3.1, we find that the total east steps below $y = x - 1$ of all the paths in $L_n$ is equal to the total area under all Dyck paths in $L_n$. We take $L_3$ as an example, there are 6 paths having $P_1$-matches and there are 5 Dyck paths, pictured in Figure 12. The total east steps below $y = x - 1$ is equal to $2 + 2 + 1 + 1 + 1 + 1 = 8$ and the total area under all the Dyck paths is also equal to $0 + 1 + 2 + 2 + 3 = 8$. Although how to design the bijection is unknown, it is interesting to see paired pattern matching does have connection to other statistics for lattice paths.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{Total number of east steps below $y = x - 1$ in $L_n$ equals the total area below Dyck paths in $D_n$, $n = 3$ as an example.}
\end{figure}

Another direction for further research is to consider Delannoy paths. In this paper, we only consider paths consisting of north steps $[0, 1]$ and east steps $[1, 0]$. Naturally, we can extend our definitions to Delannoy paths which are paths consisting of east steps $[1, 0]$, north steps $[0, 1]$, and northeast steps $[1, 1]$ which start at $[0, 0]$ and end at $[n, n]$. We denote the steps $[1, 0]$, $[0, 1]$ and $[1, 1]$ by $E$, $N$ and $D$ respectively. The set of all the Delannoy paths from $[0, 0]$ to $[n, n]$ is denoted by $S_n$.

According to [12], a Schröder path is a path from $[0, 0]$ to $[n, n]$ consisting of east steps $[1, 0]$, north steps $[0, 1]$, and northeast steps $[1, 1]$ which never goes above the diagonal $y = x$. The number of Schröder paths from $[0, 0]$ to $[n, n]$ is counted by large Schröder number $D_n$ whose
ordinary generating function equals

\[ D(x) = \sum_{n \geq 0} D_n x^n = \frac{1 - x - \sqrt{1 - 6x + x^2}}{2x} = 1 + 2x + 6x^2 + 22x^3 + 90x^4 + 394x^5 + \cdots. \]

The \( n \)th little Schröder number \( \tilde{D}(n) \) counts the number of Schröder paths from \([0, 0]\) to \([n, n]\) without northeast steps on the diagonal \( y = x \) whose ordinary generating function equals

\[ \tilde{D}(x) = \sum_{n \geq 0} \tilde{D}_n x^n = \frac{1 + x - \sqrt{1 - 6x + x^2}}{4x} = 1 + x + 3x^2 + 11x^3 + 45x^4 + 197x^5 + \cdots. \]

Here, we adopt the same definition of paired pattern for Delannoy paths. For example, in Figure 13, \( L = EDNDDNNEDE \in S_7 \). \( ps_L(1, 2) = ENNE = P_4 \) and \( ps_L(2, 3) = NNEE = P_6 \), that is, \( P_4 \text{-mch}(L) = P_6 \text{-mch}(L) = 1 \). It matches our observation: \( L \) crosses the diagonal \( y = x \) ‘vertically’ once and there is one east step above \( y = x + 1 \).

We take pattern \( P_4 \) as example, \( P_4 \text{-mch}(L) \) is the number of times \( L \) crosses the diagonal \( y = x \) vertically. We shall study the ordinary generating function

\[ F_{S_4}(x, t) := 1 + \sum_{n \geq 1} t^n \sum_{L \in S_n} x^{P_4 \text{-mch}(L)}. \]

We split the discussion into two cases. Case 1 are the paths in \( S_n \) that start with a north step and Case 2 are the path in \( S_n \) that start with an east step or a northeast step. We define

\[ G_{S_4}(x, t) := \sum_{n \geq 1} t^n \sum_{L \in S_n \text{ starting with } E \text{ or } D} x^{P_4 \text{-mch}(L)} \]

and

\[ H_{S_4}(x, t) := \sum_{n \geq 1} t^n \sum_{L \in S_n \text{ starting with } N} x^{P_4 \text{-mch}(L)}. \]

Clearly,

\[ F_{S_4}(x, t) = 1 + G_{S_4}(x, t) + H_{S_4}(x, t). \]
We obtain following formulas based on recursion on where is the first time the path starting with ‘E’ or ‘D’ crosses the diagonal $y = x$ from bottom to top.

$$GS_4(x, t) = \left(D(t) - \frac{1}{1-t}\right) (xHS_4(x, t) + 1) + \frac{t}{1-t} (HS_4(x, t) + 1)$$

Similarly, we consider where is the first time a path starting with a north step and having no northeast steps on the diagonal crosses the diagonal ‘horizontally’.

$$HS_4(x, t) = \left(\tilde{D}(t) - 1\right) (GS_4(x, t) + 1)$$

Solving for $GS_4(x, t)$, we have

$$GS_4(x, t) = -\frac{(t-1)D(t)((\tilde{D}(t) - 1)x + 1) + (\tilde{D}(t) - 1)x - t + 1}{(D(t) - 1)x(D(t)(t - 1) + 1) - 2t + 1}$$

Then we have

$$FS_4(x, t) = 1 + GS_4(x, t) + HS_4(x, t)$$

$$= 1 + GS_4(x, t) + \left(\tilde{D}(t) - 1\right) (GS_4(x, t) + 1)$$

$$= \tilde{D}(t)(GS_4(x, t) + 1)$$

$$= 2$$

$$= 3 + \sqrt{1 - 6t + t^2} - \frac{2(x-1)}{x-1} + t(x-1) - 3x + \sqrt{1 - 6t + t^2}.$$ 

A few initial terms of $FS_4(x, t)$ are

$$FS_4(x, t) = 1 + 3t + (x + 12)t^2 + (11x + 52)t^3 + (x^2 + 84x + 236)t^4 + (19x^2 + 556x + 1108)t^5 + (x^3 + 220x^2 + 3428x + 5340)t^6 + \cdots.$$ 

By setting $x = 0$ in $FS_4(x, t)$, we obtain the generating function of the number of Delannoy paths that do not cross the diagonal vertically,

$$FS_4(0, t) = \frac{(t - 1)\left(-1 + 3t + \sqrt{1 - 6t + t^2}\right)}{t^2\left(3 - t + \sqrt{1 - 6t + t^2}\right)}$$

$$= 1 + 3t + 12t^2 + 52t^3 + 236t^4 + 1108t^5 + 5340t^6 + \cdots.$$ 

The sequence 1, 3, 12, 52, 236, · · · does not appear in the OEIS [14].

Finally, one can study pattern matching for paired patterns in both lattice paths and Delannoy paths for patterns $P$ of length $\geq 6$. For example, based on Definitions 1 and 2 and Theorem 3 and 4, one can obtain geometric interpretations for the number of $P$-matches in a path $L$.

For example, consider the two patterns $P_a$ and $P_b$ are pictured in Figure [14]. Note that $P_a$ has one east step below $y = x - 2$ and $P_b$ has a vertical crossing immediately followed by a horizontal crossing. For any path $L \in \mathcal{L}_n$, $P_{a\text{-mch}}(L)$ can be interpreted as the number of east steps of $L$ below $y = x - 2$ and $P_{b\text{-mch}}(L)$ can be interpreted as the number of such pairs of crossings of $L$. 

30
Figure 14: Examples of two patterns in $\mathcal{L}_3$

References

[1] K.L. Chung and W. Feller, On fluctuations in coin tossing, Proc. Natl. Acad. Sci. USA, 35 (1949), pp. 605608.

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

[3] J. Goldman and T. Sundquist, Lattice path enumeration by formal schema, Adv. Math. 13 (1992), 216–251.

[4] J. Harmse and J.B. Remmel, Patterns in column strict fillings of rectangular arrays, Pure Math. and Applications, 22 (2011), 131–171.

[5] K. Humphreys, A history and a survey of lattice path enumeration, J. Statist. Plann. and Infer. 140 (2010), no. 8.

[6] C. Krattenthaler, Lattice path enumeration. In Handbook of enumerative combinatorics, Discrete Math. Appl. (Boca Raton). CRC Press, Boca Raton, FL, (2015), 589–678.

[7] K. Manes, A. Sapounakis, I. Tasoulas and P. Tsikouras, Counting strings at height $j$ in Dyck paths, J. Statist. Plann. and Infer. 141 (2011), 2100-2107.

[8] T. Mansour, Statistics on Dyck paths, J. Integer Seq. 9 (2006), Article 06.1.5.

[9] S. G. Mohanty, Lattice path counting and applications, Academic Press, New York, 1979

[10] A. Sapounakis, P. Tsikouras, I. Tasoulas, K. Manes, Strings of length 3 in Grand-Dyck paths and the Chung-Feller property, Electronic J. Combinatorics, 19(2) (2011), #P2.

[11] A. Sapounakis, I. Tasoulas and P. Tsikouras, Counting strings in Dyck paths, Discrete Math. 307 (2007), 2909–2924.

[12] E. Schröder, Vier kombinatorische Probleme, Z. Math. Phys., 15 (1870), 361–376.

[13] K. Sen, J.L. Jain and J.M. Gupta, Lattice path approach to transient solution of M/M/1 with (0,k) control policy. Journal of Statistical Planning and Inference. 34 (1993). 259–267.

[14] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences. Published electronically at [http://oeis.org/](http://oeis.org/)