Séminaire Lotharingien de Combinatoire **87B** (2023)
Article #6, 20 pp.
Special issue for the 9^{th} *International Conference on Lattice Path Combinatorics and Applications*

RESTRICTED DYCK PATHS ON VALLEYS SEQUENCE

RIGOBERTO FLÓREZ[1], TOUFIK MANSOUR[2], JOSÉ L. RAMÍREZ[3], FABIO A. VELANDIA[4], AND DIEGO VILLAMIZAR[5]

[1]Department of Mathematical Sciences, The Citadel, Charleston, SC, U.S.A.; https://www.rigoflorez.com/

[2]Department of Mathematics, University of Haifa, 3498838 Haifa, Israel; https://math.haifa.ac.il/toufik/

[3]Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia; https://sites.google.com/site/ramirezrjl/

[4]Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia; https://orcid.org/0000-0002-9542-0782

[5]Escuela de Ciencias Exactas e Ingeniería, Universidad Sergio Arboleda, Bogotá, Colombia; https://sites.google.com/view/dvillami/

Abstract. In this paper we study a subfamily of a classic lattice path, the *Dyck paths*, called *restricted d-Dyck* paths, in short d-Dyck. A valley of a Dyck path P is a local minimum of P; if the difference between the heights of two consecutive valleys (from left to right) is at least d, we say that P is a restricted d-Dyck path. The *area* of a Dyck path is the sum of the absolute values of y-components of all points in the path. We find the number of peaks and the area of all paths of a given length in the set of d-Dyck paths. We give a bivariate generating function to count the number of the d-Dyck paths with respect to the semi-length and number of peaks. After that, we analyze in detail the case $d = -1$. Among other things, we give both the generating function and a recursive relation for the total area.

Keywords: Dyck path, d-Dyck path, generating function.

1. Introduction

A classic concept, the *Dyck paths*, has been widely studied. Recently, a subfamily of these paths, non-decreasing Dyck paths, has received a certain level of interest. It is because of some statistics are given by linear combinations of Fibonacci numbers and Lucas numbers. In this paper we keep studying a generalization of the non-decreasing Dyck paths. Other generalizations of non-decreasing Dyck paths have been given for Motzkin paths and for Łukasiewicz paths [14, 15].

We now give some definitions that we use in this paper. A Dyck path is a lattice path in the first quadrant of the xy-plane that starts at the origin, ends on the x-axis, and consists of (the same number of) North-East steps $U := (1, 1)$ and South-East steps $D := (1, -1)$. The *semi-length* of a path is the total number of U's that the path has.

A *valley* (*peak*) is a subpath of the form DU (UD) and the *valley vertex* of DU is the lowest point (a local minimum) of DU. The *level* of a valley is the y-component of its valley vertex. Following [16, 17] we define the *valley vertices vector* of a Dyck path P as the vector $\nu = (\nu_1, \nu_2, \dots, \nu_k)$ formed by all y-coordinates (listed from left to right) of all valley vertices of P.

For a fixed $d \in \mathbb{Z}$, a Dyck path P is called *restricted d-Dyck* or *d-Dyck* (for simplicity), if either P has at most one valley, or if its valley vertex vector ν satisfies that $\nu_{i+1} - \nu_i \geq d$, where $1 \leq i < k$. The set of all d-Dyck paths of semi-length n is denoted $\mathcal{D}_d(n)$, where $r_d(n)$ denotes its cardinality, and the set of all d-Dyck paths is denoted by $\mathcal{D}_d$.

The first well-known example of these paths is the set of 0-Dyck paths; in the literature, see [4, 6, 7, 9, 10, 12], this family is known as non-decreasing Dyck paths. The whole family of Dyck paths can be seen as a limit of d-Dyck and it occurs when $d \to -\infty$. Another example, from Figure 1 we observe that $\nu = (0, 1, 0, 3, 4, 3, 2)$ and that $\nu_{i+1} - \nu_i \geq -1$, for $i = 1, \dots, 6$, so the figure depicts a (-1)-Dyck path of length 28 (or semi-length 14).

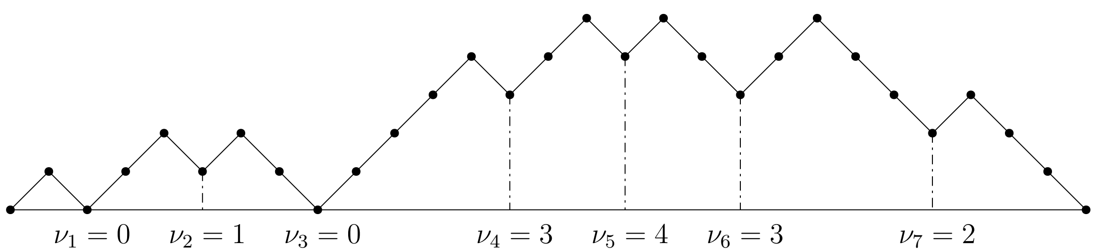

Figure 1. A (-1)-Dyck path of length 28.

The recurrence relations and/or the generating functions for d-Dyck when $d \geq 0$ have different behavior than the case $d < 0$. For example, generating functions accounting for the number of valleys, the number of peaks, and the area, for d-Dyck when $d \geq 0$, are all rational for all variables (see [4, 6, 7, 10, 12, 16, 17]). However, when we analyze in this paper several aspects for $d < 0$ (the number of paths, the area of the paths, and the number of peaks) we find that the generating functions are all algebraic (non-rational).

In this paper we give a bivariate generating function to count the number of paths in $\mathcal{D}_d(n)$, for $d \leq 0$, with respect to the number of peaks and semi-length. We also give a relationship between the total number of d-Dyck paths and the Catalan numbers. Additionally, we give an explicit symbolic expression for the generating function with respect to the semi-length. For the particular case $d = -1$ we give a combinatorial expression and a recursive relation for the total number of paths. We also analyze the asymptotic behavior for the sequence $r_{-1}(n)$.

It is well known that there are many bijections between Dyck paths and other combinatorial objects, we are wondering if there are other bijections between d-Dyck paths for $d < -1$ and other object of combinatorics.

The *area* of a Dyck path P is the sum of the values of y-components of all points in the path. That is, the area of P, denoted by $\texttt{area}(P)$, corresponds to the surface area under P and above of the x-axis. For example, if P is the path in Figure 1, then $\texttt{area}(P) = 70$. We use generating functions and recursive relations to analyze the distribution of the area of all paths in $\mathcal{D}_{-1}(n)$.

The problem of enumerating the area in directed lattice paths, in a general setting, was solved by Banderier and Gittenberger [3], building on the enumerative and asymptotics results from [2], where Dyck, Motzkin, and Łukasiewicz paths are particular cases.

A summary of notation used throughout the paper appears in Table 1 in the appendix.

2. Number of d-Dyck paths and Peaks Statistic

Given a family of lattice paths, a classic question is how many lattice paths are there of certain length, and a second classic question is how many peaks are there depending on the length of the path. These questions have been completely answered, for instance, for Dyck paths [8], d-Dyck paths for $d \geq 0$ [4, 17], and Motzkin paths [20] among others. In this section we give a bivariate generating function according to the semi-length and the number of peaks of the d-Dyck paths with $d < 0$.

Given a d-Dyck path P, we denote the semi-length of P by $\ell(P)$ and denote the number of peaks of P by $\rho(P)$. So, the bivariate generating function to count the number of paths and peaks of d-Dyck paths is defined by

$$L_d(x, y) := \sum_{P \in \mathcal{D}_d} x^{\ell(P)} y^{\rho(P)}.$$

2.1. Some facts known when $d \geq 0$.

These results can be found in [17].

- If $d \geq 0$, then the generating function $F_d(x, y)$ is given by

$$L_d(x, y) = 1 + \frac{xy(1 - 2x + x^2 + xy - x^{d+1}y)}{(1 - x)(1 - 2x + x^2 - x^{d+1}y)}.$$

- If $d \geq 1$,

$$r_d(n) = \sum_{k=0}^{\lfloor \frac{n+d-2}{d} \rfloor} \binom{n - (d-1)(k-1)}{2k}.$$

- If $n > d$, then we have the recursive relation

$$r_d(n) = 2r_d(n-1) - r_d(n-2) + r_d(n-d-1),$$

  with the initial values $r_d(n) = \binom{n}{2} + 1$, for $0 \leq n \leq d$.
- Let $p_d(n, k)$ be the number of d-Dyck paths of semi-length n, having exactly k peaks. If $d \geq 0$, then

$$p_d(n, k) = \binom{n + k - d(k-2) - 2}{2(k-1)}.$$

  For the whole set of Dyck paths, the number $p_{-\infty}(n, k)$, is given by the Narayana numbers $N(n, k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}$.

2.2. Peaks statistic for d a negative integer.

For the remaining part of the paper we consider only the case $d < 0$ and use e to denote $|d|$. A *pyramid* of semi-length $h \geq 1$ is a subpath of the form $X^h Y^h$; it is *maximal*, denote by Δ_h, if it can not be extended to a pyramid $X^{h+1} Y^{h+1}$.

Theorem 2.1. *If d is a negative integer and $e := |d|$, then the generating function $L_e(x, y)$ satisfies the functional equation*

$$L_e(x, y) = xy + xL_e(x, y) + xS_e(x, y)L_e(x, y), \tag{2.1}$$

where $S_e(x)$ satisfies the algebraic equation

$$(1 - xS_e(x, y))^e(y + (1 - y)xS_e(x, y)) - S_e(x, y)(1 - xS_e(x, y))^{e+1} - \frac{x^{e+2}y}{1-x}S_e(x, y) = 0.$$

Proof. We start this proof by introducing some notation. The set $\mathcal{Q}_{d,i} \subseteq \mathcal{D}_d$ denotes the family of non-empty paths where the last valley is at level i. We consider the generating function

$$Q_i^{(e)}(x, y) := \sum_{P \in \mathcal{Q}_{d,i}} x^{\ell(P)} y^{\rho(P)}.$$

It is convenient to consider the sum over the $Q_i^{(e)}(x, y)$. We also consider the generating function, with respect to the lengths and peaks, that counts the d-Dyck paths that have either no valleys or the last valley is at level less than e. That is,

$$S_e(x, y) = \frac{y}{1-x} + \sum_{j=0}^{e-1} Q_j^{(e)}(x, y). \tag{2.2}$$

A path P can be uniquely decomposed as either UD, UTD, or $UQDT$ (by considering the first return decomposition), where $T \in \mathcal{D}_d$ and Q is either a pyramid or is a path in $\cup_{i=0}^{e-1} \mathcal{Q}_{d,i}$ (see Figure 2, for a graphical representation of this decomposition). Notice that $\nu_{i+1} - \nu_i \geq d$ and the decomposition $UQDT$ ensures that Q holds as in the former line.

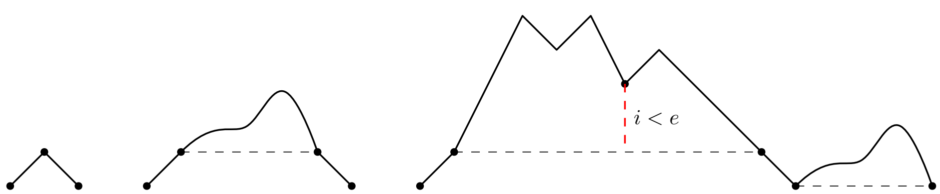

FIGURE 2. Decomposition of a d-Dyck path.

From the symbolic method we obtain the functional equation

$$L_e(x, y) = xy + xL_e(x, y) + xS_e(x, y)L_e(x, y).$$

Now we are going to obtain a system of equations for the generating functions $Q_i(x, y)$. Let Q be a path in the set $\mathcal{Q}_{d,i}$. If $i = 0$, then the path Q can be decomposed uniquely as either $UQ'D\Delta$ or $UQ'DR$, where Δ is a pyramid, R is a path in $\mathcal{Q}_{d,0}$, and Q' is either a pyramid or $Q' \in \cup_{i=0}^{e-1} \mathcal{Q}_{d,i}$. Therefore, we have the functional equation

$$Q_0^{(e)}(x, y) = xS_e(x, y)\frac{xy}{1-x} + xS_e(x, y)Q_0^{(e)}(x, y).$$

For $i > 0$, any path Q can be decomposed uniquely in one of these two forms UR_1D or $UQDR_2$, where $R_1 \in \mathcal{Q}_{d,i-1}, R_2 \in \mathcal{Q}_{d,i}$, and Q is either a pyramid or $Q \in \cup_{i=0}^{e-1} \mathcal{Q}_{d,i}$. So, we have the functional equation

$$Q_i^{(e)}(x, y) = xQ_{i-1}^{(e)}(x, y) + xS_e(x, y)Q_i^{(e)}(x, y). \tag{2.3}$$

Summarizing the discussion above, we obtain the system of equations:

$$\begin{cases} Q_0^{(e)}(x,y) & = xS_e(x,y)\frac{xy}{1-x} + xS_e(x,y)Q_0^{(e)}(x,y) \\ Q_1^{(e)}(x,y) & = xQ_0^{(e)}(x,y) + xS_e(x,y)Q_1^{(e)}(x,y) \\ & \vdots \\ Q_i^{(e)}(x,y) & = xQ_{i-1}^{(e)}(x,y) + xS_e(x,y)Q_i^{(e)}(x,y) \\ & \vdots \\ Q_{e-1}^{(e)}(x,y) & = xQ_{e-2}^{(e)}(x,y) + xS_e(x,y)Q_{e-1}^{(e)}(x,y). \end{cases} \tag{2.4}$$

Summing up the equations in (2.4), we obtain that

$$\sum_{j=0}^{e-1} Q_j^{(e)}(x,y) = xS_e(x,y)\left(\sum_{j=0}^{e-1} Q_j^{(e)}(x,y) + \frac{xy}{1-x}\right) + x\sum_{j=0}^{e-2} Q_j^{(e)}(x,y).$$

From this and (2.2) we have

$$\begin{aligned} S_e(x,y) - \frac{y}{1-x} = x\left(S_e(x,y) - \frac{y}{1-x} - Q_{e-1}^{(e)}(x,y)\right) \\ + xS_e(x,y)\left(S_e(x,y) - \frac{y}{1-x}\right) + \frac{x^2y}{1-x}S_e(x,y). \end{aligned} \tag{2.5}$$

Iterating (2.3), we have $Q_i^{(e)}(x,y)$, with $i \geq 0$, can be expressed as

$$Q_i^{(e)}(x,y) = \frac{x^{i+2}yS_e(x,y)}{(1-x)(1-xS_e(x,y))^{i+1}}. \tag{2.6}$$

Substituting (2.6) into (2.5) we obtain the desired functional equation. □

Solving (2.5) for $S_e(x,y)$ we have

$$S_e(x,y) = \frac{1-x+xy-\sqrt{1-2x+x^2-2xy-2x^2y+x^2y^2+4x^2Q_{e-1}^{(e)}(x,y)}}{2x}. \tag{2.7}$$

We observe that substituting (2.7) into (2.1), we have

$$\begin{aligned} L_e(x,y) &= \frac{xy}{1-x-xS_e(x,y)} \\ &= \frac{xy}{1-x-\dfrac{1-x+xy-\sqrt{1-2x+x^2-2xy-2x^2y+x^2y^2+4x^2Q_{e-1}^{(e)}(x,y)}}{2}}. \end{aligned}$$

Since $S_e(x,y)$ is a power series and by (2.6), we obtain that $Q_{e-1}^{(e)}(x,y) \to 0$ as $e \to \infty$, where here we assumed that $|x| < 1$ (for details on convergence of generating functions; see [11, p. 731]). Therefore,

$$\lim_{e\to\infty} L_e(x,y) = \frac{1-x-xy-\sqrt{1-2x+x^2-2xy-2x^2y+x^2y^2}}{2x}.$$

This last generating function is the distribution of the Narayana sequence. This corroborates with the fact that the restricted $(-\infty)$-Dyck paths coincide with the non-empty Dyck paths.

Theorem 2.2. *If $1 \leq k \leq |d| + 3$, then the k-th coefficient of the generating function $L_e(x, 1)$ coincides with the Catalan number $C_k = \frac{1}{k+1}\binom{2k}{k}$.*

Proof. We first observe that the shortest Dyck path that contains a forbidden sequence of valleys is $P = U^{e+2}DUD^{e+2}UD$ (clearly, $\ell(P) = e + 4$) with $e = |d|$. Therefore, if $d < 0$, then $r_d(n) = C_n$, for $n = 1, 2, \dots, |d| + 3$. $\square$

The first few values for the sequence $r_d(n)$, for $d \in \{-1, -2, -3, -4\}$ are

$$\begin{aligned}
\{r_{-1}(n)\}_{n\geq 1} &= \{\mathbf{1}, \mathbf{2}, \mathbf{5}, \mathbf{14}, 41, 123, 375, 1157, 3603, \dots\},\\
\{r_{-2}(n)\}_{n\geq 1} &= \{\mathbf{1}, \mathbf{2}, \mathbf{5}, \mathbf{14}, \mathbf{42}, 131, 419, 1365, 4511, \dots\},\\
\{r_{-3}(n)\}_{n\geq 1} &= \{\mathbf{1}, \mathbf{2}, \mathbf{5}, \mathbf{14}, \mathbf{42}, \mathbf{132}, 428, 1419, 4785, \dots\},\\
\{r_{-4}(n)\}_{n\geq 1} &= \{\mathbf{1}, \mathbf{2}, \mathbf{5}, \mathbf{14}, \mathbf{42}, \mathbf{132}, \mathbf{429}, 1429, 4850, \dots\}.
\end{aligned}$$

For example, there are 41 (-1)-Dyck paths out of the 42 Dyck paths of length 10. Figure 3 depicts the only Dyck path of length 10 that is not a (-1)-Dyck path.

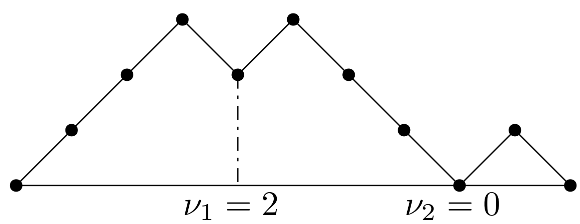

FIGURE 3. The only Dyck path of length 10 that is not a (-1)-Dyck path.

Recall that d is a negative integer and that $e := |d|$. Then by Theorem 2.1, we have

$$\begin{aligned}
(L_e(x,y)+y)^e &\left(xL_e^2(x,y) + (xy+x-1)L_e(x,y) + xy\right)\\
&- \frac{x}{1-x}((1-x)L_e(x,y) - xy)(L_e(x,y))^{e+1} = 0.
\end{aligned}$$

This implies that

$$\begin{aligned}
\sum_{j=2}^{e+1} x\binom{e}{j-2} y^{e+2-j}(L_e(x,y))^j &+ \sum_{j=1}^{e+1}(xy+x-1)\binom{e}{j-1} y^{e+1-j}(L_e(x,y))^j\\
&+ \sum_{j=0}^{e} x\binom{e}{j} y^{e+1-j}(L_e(x,y))^j + \frac{x^2y}{1-x}(L_e(x,y))^{e+1} = 0.
\end{aligned}$$

Hence, by taking $y = 1$ and collecting powers of $L_e(x, 1)$, we have

$$L_e(x,1) = Z\left(a_0 + \sum_{j=2}^{e+1} a_j(x)(L_e(x,1))^j\right),$$

where $Z = 1$, and

$$a_0 = \frac{x}{1-(e+2)x},$$
$$a_j = \frac{1}{1-(e+2)x}\left(x\binom{e+2}{j} - \binom{e}{j-1}\right), \quad j = 2, 3, \dots, e,$$
$$a_{e+1} = \frac{(e+2)x(1-x) - 1 + x(1+x)}{(1-x)(1-(e+2)x)}.$$

Hence, by the Lagrange inversion formula, we expand the generating function $L_e(x, 1)$ as a power series in Z to obtain

$$L_e(x,1) = \sum_{n\geq 1} \frac{[Z^{n-1}]}{n} \sum_{i_0+i_2+i_3+\cdots+i_{e+1}=n} \frac{n!}{i_0!i_2!\cdots i_{e+1}!} a_0^{i_0} Z^{2i_2+\cdots+(e+1)i_{e+1}} \prod_{j=2}^{e+1} a_j^{i_j},$$

that leads to the following result.

Theorem 2.3. *We have*

$$L_e(x,1) = \sum_{n\geq 1} \frac{\sum_{2i_2+\cdots+(e+1)i_{e+1}=n-1} \binom{n}{i_2,\dots,i_{e+1}} x^{n-i_2-\cdots-i_{e+1}} t^{i_{e+1}} \prod_{j=2}^{e}\left(x\binom{e+2}{j} - \binom{e}{j-1}\right)^{i_j}}{n(1-(e+2)x)^n},$$

where

$$\binom{n}{i_2,\dots,i_{e+1}} = \frac{n!}{i_2!\cdots i_{e+1}!(n-i_2-\cdots-i_{e+1})!} \text{ and } t = \frac{(e+2)x(1-x)-1+x(1+x)}{1-x}.$$

For example, Theorem 2.3 with $e = 2$ gives

$$L_2(x,1) = \sum_{n\geq 1} \frac{\sum_{2i_2+3i_3=n-1} \binom{n}{i_2,i_3} x^{n-i_2-i_3}(6x-2)^{i_2}(\frac{-3x^2+5x-1}{1-x})^{i_3}}{n(1-4x)^n}.$$

Thus,

$$L_2(x,1) = \frac{x}{1-4x} + \frac{x^2(6x-2)}{(1-4x)^3} + \frac{x^3t}{(1-4x)^4} + \frac{2x^3(6x-2)^2}{(1-4x)^5} + \frac{5x^4t(6x-2)}{(1-4x)^6}$$
$$+ \frac{5x^4(6x-2)^3 + 3x^5t^2}{(1-4x)^7} + \frac{21x^5(6x-2)^2t}{(1-4x)^8} + \frac{28x^6(-2+6x)t^2 + 14x^5(-2+6x)^4}{(1-4x)^9} + \cdots,$$

where $t = (-3x^2 + 5x - 1)/(1 - x)$.

3. Some results for the case $d = -1$

In this section we keep analyzing the bivariate generating function given in the previous section for the particular case $d = -1$. For this case, we provide more detailed results. We denote by $\mathcal{Q}$ the set of all non-empty paths in $\mathcal{D}_{-1}$ having at least one valley, where the last valley is at ground level. We denote by $\mathcal{Q}_n$ the subset of $\mathcal{Q}$ formed by all paths of semi-length n and denote by q_n the cardinality of $\mathcal{Q}_n$. For simplicity, when $d = -1$ (or $e = 1$) we use $L(x, y)$ instead of $L_1(x, y)$. As a consequence of Theorem 2.1, taking $d = -1$, we obtain this theorem.

Theorem 3.1. *The bivariate generating function $L(x,y)$ is given by*

$$L(x,y)=\frac{(x-1)y\left(1-x(2+y)-\sqrt{(1-x-2xy-2x^2y+x^2y^2-x^3y^2)/(1-x)}\right)}{2(1-2x+x^2-2xy+x^2y)}.$$

Notice that a path $Q\in\mathcal{Q}$ can be uniquely decomposed as either $U\Delta DU\Delta' D$, $U\Delta DR$, UR_1DR_2, or $URDU\Delta D$, where Δ,Δ' are pyramids, and $R,R_1,R_2\in\mathcal{Q}$ (see Figure 4 for a graphical representation of this decomposition).

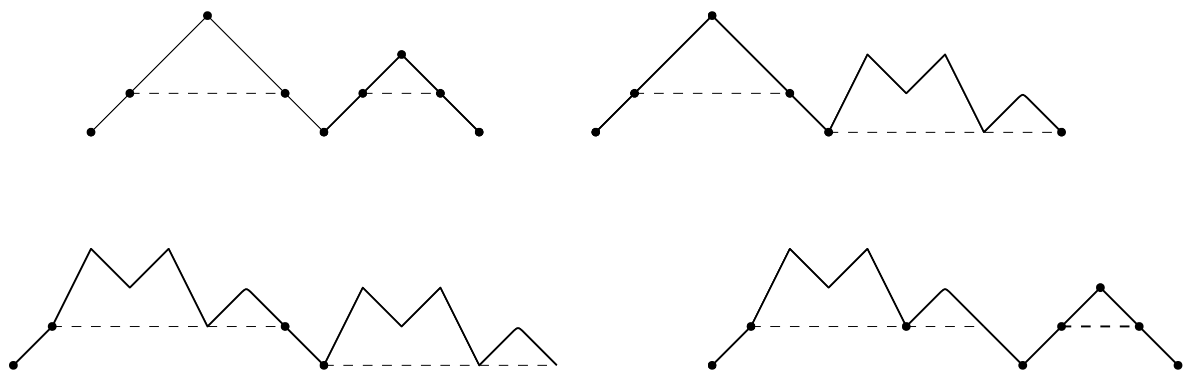

FIGURE 4. Decomposition of a (-1)-Dyck path in $\mathcal{Q}$.

Therefore, if

$$Q(x,y):=\sum_{Q\in\mathcal{Q}}x^{\ell(Q)}y^{\rho(P)},$$

then

$$Q(x,y)=x^2\left(\frac{y}{1-x}\right)^2+x\left(\frac{y}{1-x}\right)Q(x,y)+x(Q(x,y))^2+x^2\left(\frac{y}{1-x}\right)Q(x,y).$$

Solving the equation above for $Q(x,y)$, we find that

$$Q(x,y)=\frac{1-x-xy-x^2y-\sqrt{(1-x)(1-x-2xy-2x^2y+x^2y^2-x^3y^2)}}{2(1-x)x}. \quad (3.1)$$

Expressing $L(x,y)$ as a series expansion we obtain these first few terms:

$$\begin{aligned}L(x,y)&=xy+x^2\left(y^2+y\right)+x^3\left(y^3+3y^2+y\right)+x^4\left(y^4+\mathbf{6}y^3+6y^2+y\right)\\&+x^5\left(y^5+10y^4+19y^3+10y^2+y\right)+x^6\left(y^6+15y^5+46y^4+45y^3+15y^2+y\right)+\cdots.\end{aligned}$$

Figure 5 depicts all six paths in $\mathcal{D}_{-1}(4)$ with exactly 3 peaks. Notice that this is the coefficient of x^4y^3, in boldface type, in the above series.

The generating function for the (-1)-Dyck paths is given by

$$L(x):=L(x,1)=\frac{-1+4x-3x^2+\sqrt{1-4x+2x^2+x^4}}{2(1-4x+2x^2)}. \quad (3.2)$$

Thus,

$$L(x)=x+2x^2+5x^3+14x^4+41x^5+123x^6+375x^7+1157x^8+\cdots.$$

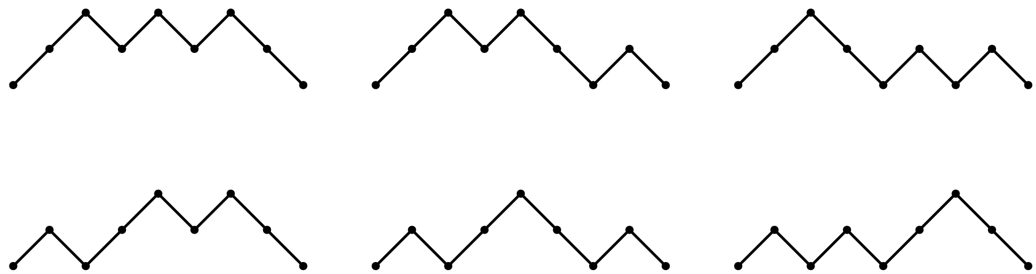

FIGURE 5. All six paths in $\mathcal{D}_{-1}(4)$ with exactly 3 peaks.

For the sake of simplicity, if there is not ambiguity, for the remaining part of the paper we use $r(n)$ instead of $r_{-1}(n)$. Our interest here is to give a combinatorial formula for this sequence. First of all, we give some preliminary results. Let $b(n)$ be the number of (-1)-Dyck paths of semi-length n that either have no valleys or the last valley is at ground level. Note that $b(n)-1$ is the n-th coefficient of the generating function $Q(x,1)$; see (3.1), or equivalently

$$\sum_{n\geq 0} b(n)x^n = Q(x,1) + \frac{1}{1-x} = \frac{1-x^2-\sqrt{1-4x+2x^2+x^4}}{2(1-x)x}$$
$$= 1 + x + 2x^2 + 4x^3 + 9x^4 + 22x^5 + 57x^6 + 154x^7 + 429x^8 + \cdots .$$

This generating function coincides with the generating function of the number of Dyck paths of semi-length n that avoid the subpath $UUDU$. From Proposition 5 of [19] and [5, p. 10] we conclude the following proposition.

Proposition 3.2. *For all $n \geq 0$ we have*

$$b(n) = 1 + \sum_{j=0}^{\lfloor\frac{n-1}{2}\rfloor} \frac{(-1)^j}{n-j}\binom{n-j}{j}\binom{2n-3j}{n-j+1} = \sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\sum_{j=0}^{n-k}\binom{n-k}{j}N(j,k),$$

where $N(n,k)=\frac{1}{n}\binom{n}{k}\binom{n}{k-1}$ are the Narayana numbers, with $N(0,0)=1$.

We tried to find a combinatorial proof of the previous proposition. However, we were not able to do it. This remark summarizes our observations toward a potential proof. It will be interesting to see such a combinatorial proof.

Remark 3.3. Let $\mathcal{B}(n) = \mathcal{Q}_n \cup \{\Delta_n\}$ denote the set of (-1)-Dyck paths having either no valleys or the last valley is at ground level. That is, $b(n) = |\mathcal{B}(n)|$. A South-East step in $P \in \mathcal{B}(n)$, satisfies one of these two conditions.

- The step belongs to a pyramid.
- The step is part of a valley, say for example, the m-th valley, such that $\nu_m - \nu_{m-1} = -1$. In this case, the valley with height ν_m is called (-1)-*valley*.

We denote by $\mathcal{B}_{j,k}(n)$ the set of paths in $\mathcal{B}(n)$ with exactly j valleys, where k of them are (-1)-valley. Now, a path $P \in \mathcal{B}_{j,k}(n)$ can be decomposed as

$$P = U^{s_0}\Delta_{t_1}U^{s_1}\Delta_{t_2}D^{r_1}U^{s_2}\cdots\Delta_{t_{j-1}}D^{r_{j-2}}U^{s_{j-1}}\Delta_{t_j}D^{r_{j-1}}\Delta_{t_{j+1}},$$

where $r_i \in \{0,1\}$, $t_i \geq 1$, $s_i \geq 0$, and with the additional property that there are exactly k indices i for which $r_i = 1$.

There are $n-k$ South-East steps that belong to one of the $j+1$ pyramids in the path. So, we can represent the possible choices of the t_i as an integer composition of $n-k$ into $j+1$ parts in $\binom{n-k-1}{j+1-1} = \binom{n-k-1}{j}$ ways. This means that setting $t_i = 1$ for all i, in the spirit of Proposition 3.2, the Narayana numbers $N(j, k+1)$ should correspond to the number of (-1)-Dyck paths of semi-length $j+1+k$ containing exactly j valleys and $k \leq j-1$ (-1)-valleys where the last valley is at ground level. That is $|\mathcal{B}_{j,k}(j+k+1)| = N(j, k+1)$ for $j > 0$.

Theorem 3.4. *The total number of paths in $\mathcal{D}_{-1}(n)$ is given by*

$$r(n) = \sum_{k=0}^{n}\sum_{i=0}^{n-k-1}\binom{n-k-1}{i}q^{(i)}(k),$$

where

$$q^{(i)}(n) = \sum_{n_1+n_2+\cdots+n_i=n} b(n_1)b(n_2)\cdots b(n_i).$$

Proof. Recall that $\ell(P)$ denotes the semi-length of a path P. Let us denote by $\mathcal{Q}^{(i)}(n)$ the set of i-tuples $(P_1,\dots,P_i)$ of paths $P_j \in \mathcal{B} = \bigcup_{m\geq 0}\mathcal{B}(m)$, such that $\ell(P_1)+\cdots+\ell(P_i) = n$. It is clear that $\left|\mathcal{Q}^{(i)}(n)\right| = q^{(i)}(n)$. Note that the empty path $\lambda \in \mathcal{B}$. Let $\mathcal{C}_i(n)$ be the set of integer compositions of n with i parts. The cardinality of this set is given by the binomial coefficient $\binom{n-1}{i-1}$.
Let $\mathcal{QC}(n)$ be the set of $(2i+1)$-tuples $(c_1, P_1, \dots, c_i, P_i, c_{i+1})$ such that the element $((P_1,\dots,P_i),(c_1,\dots,c_{i+1}))$ is in $\mathcal{Q}^{(i)}(j) \times \mathcal{C}_{i+1}(n-j)$. (Note that $\mathcal{QC}(n)$ is isomorphic to $\bigcup_{i,j}(\mathcal{Q}^{(i)}(j) \times \mathcal{C}_{i+1}(n-j))$.) We now consider the function

$$\varphi : \mathcal{QC}(n) \longrightarrow \mathcal{D}_{-1}(n),$$

defined by

$$\varphi\left((c_1, P_1, c_2, P_2, \dots, c_i, P_i, c_{i+1})\right) = U^{c_1}M_1U^{c_2}\cdots M_iU^{c_{i+1}}D^g,$$

where the integer $g \geq c_{i+1}$ is the number of necessary down-steps to reach the x-axis, and M_j is given by

$$M_j = \begin{cases} D^{c_j}, & \text{if } P_j = \lambda; \\ P_j, & \text{if } P_j = \Delta; \\ P_jD, & \text{otherwise.} \end{cases}$$

Figure 6 depicts two examples on the application of the function φ.

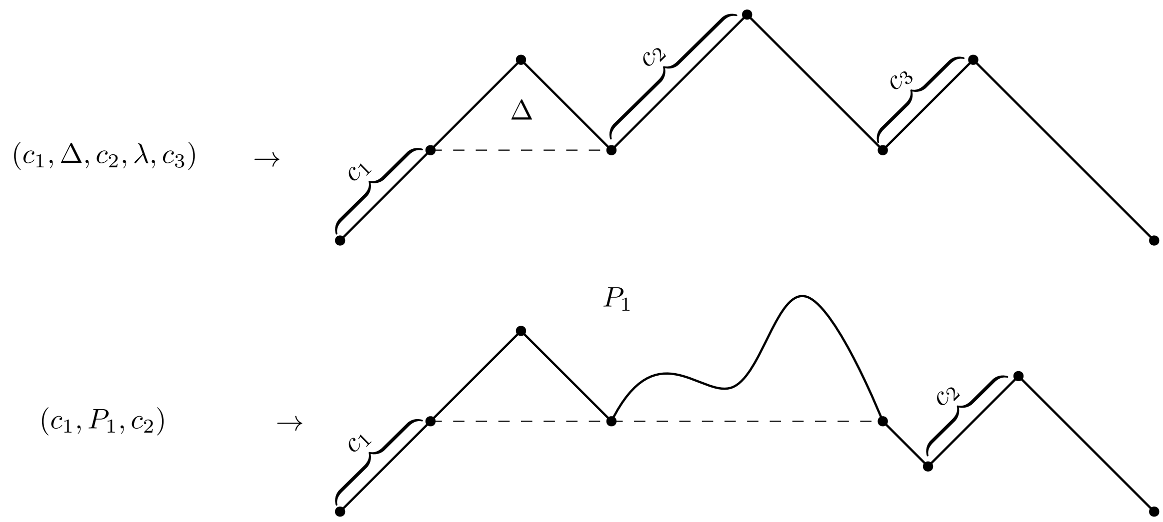

FIGURE 6. Function φ applied to the vectors $(c_1, \Delta, c_2, \lambda, c_3)$ and (c_1, P_1, c_2).

For the remaining part of this proof we use $\Delta_0 = \lambda$. We define ϕ from $\mathcal{D}_{-1}(n)$ to $\mathcal{QC}(n)$ via the Algorithm 1 below.

Algorithm 1 Function ϕ

---

(1) Let $i = 1$.

(2) If there are (-1)-valleys, go to step (3), else, the path is non-decreasing, and for some integers $s_m \geq 0$, $g \geq 0$, and $t_m > 0$, it can be decomposed as

$$P = U^{s_0}\Delta_{t_1}U^{s_1}\Delta_{t_2}\cdots U^{s_{j-1}}\Delta_{t_j}\Delta_{t_{j+1}}D^g.$$

- If no valleys, that is $j = s_0 = g = 0$, then **return** the vector (t_1).
- If there is just one valley, that is $j = 1$, set

$$(s_0', t_1') = \begin{cases} (s_0, t_1), & \text{if } s_0 > 0; \\ (t_1, 0), & \text{otherwise;} \end{cases}$$

  and then **return** the vector $(s_0', \Delta_{t_1'}, t_2)$.
- In the general case, set

$$(s_m', t_{m+1}') = \begin{cases} (s_m, t_{m+1}), & \text{if } s_m > 0; \\ (t_{m+1}, 0), & \text{otherwise;} \end{cases}$$

  for $m < j$ and then **return** the vector $(s_0', \Delta_{t_1'}, s_1', \ldots, \Delta_{t_j'}, t_{j+1})$.

(3) Find the **rightmost** occurrence of an (-1)-valley, that is, a subpath of the form $D\Delta_k DU$, with $k > 0$. Denote the height of this valley as h_i. Decompose the path P as $P = \widetilde{P}_0 P_i D \widetilde{P}_1$, where P_i is the maximal subpath that is a Dyck path to the left of the aforementioned (-1)-valley. Notice that P_i ends on $D\Delta_k$ and so, it belongs to $\mathcal{B}$. Let $\hat{P}_0 = \widetilde{P}_0 D^{h_i+1}$ and $\hat{P}_1 = \widetilde{P}_1 D^{-h_i}$, where D^{-r} means deleting the last r South-East steps of the path. By assumption, the path $\hat{P}_1$ is non-decreasing: go to step (2) using $\hat{P}_1$ and call the returned tuple τ_1. Increase the value of i by one and go to step (2) with the path $\hat{P}_0$ and call the returning tuple τ_0. **Return** (τ_0, P_i, τ_1).

We now give an example of the application of the Algorithm 1. Consider, for instance, the path given in Figure 7. First of all, search for the rightmost (-1)-valley (in the first case, it is denoted by dashed circle after P_1), they are decorated by a red circle around them, the algorithm applied to $\hat{P}_1 = UD$ (where $\widetilde{P}_1 = UDD$) gives $\tau_1 = (1)$. We extract the path P_1 and we locate the next (-1)-valley, the right part on this instance, given by $UDUDU$, corresponds to $\tau_1 = (1, \lambda, 1, \lambda, 1)$, and the left part of P_2 corresponds to $(1, \Delta_1, 1)$, and so the whole path is encoded by the vector $(1, \Delta_1, 1, P_2, 1, \lambda, 1, \lambda, 1, P_1, 1)$.

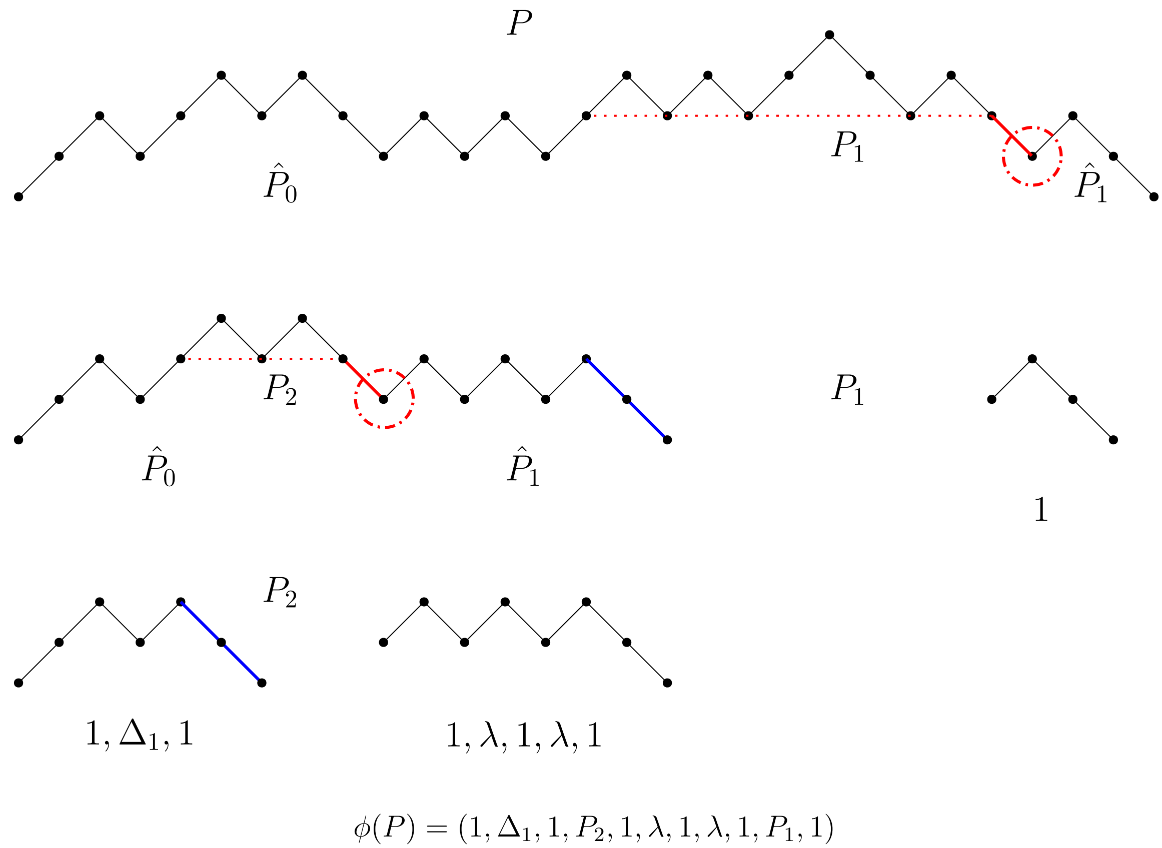

FIGURE 7. Example inverse function.

Using these decompositions, we show by induction that $\phi \circ \varphi|_{\mathcal{QC}_k(n)} = id_{\mathcal{QC}_k(n)}$ and $\varphi \circ \phi|_{\mathcal{B}_k(n)} = id_{\mathcal{B}_k(n)}$ for every $k \geq 0$. These equalities and the functionality of ϕ, given by choosing the paths P_i in a **maximal** way, imply that φ is a bijection and ϕ is its inverse. Let $S(k)$ be the statement

$$\phi \circ \varphi|_{\mathcal{QC}_k(n)} = id_{\mathcal{QC}_k(n)} \quad \text{and} \quad \varphi \circ \phi|_{\mathcal{B}_k(n)} = id_{\mathcal{B}_k(n)}.$$

For the basis step, $S(0)$, first notice that if $P \in \mathcal{B}_0(n)$, then the Algorithm 1 Part (2) guarantees that $\phi(P)$ contains only paths of the form Δ_t for $t \geq 0$. On the other hand, φ returns a path without (-1)-valley when given a tuple $\tau \in \mathcal{QC}_0(n)$ by definition of the M_j's, and so the functions and their compositions are well defined when restricted to $\mathcal{B}_0(n)$ and $\mathcal{QC}_0(n)$. Let $P = U^{s_0}\Delta_{t_1} \cdots \Delta_{t_i}\Delta_{t_{i+1}}D^g \in \mathcal{B}_0(n)$, using Algorithm 1 we get $\phi(P) = (s'_0, \Delta_{t'_1}, \ldots, \Delta_{t'_i}, t_{i+1})$.

Now, we have

$$\varphi(\phi(P)) = \varphi(s'_0, \Delta_{t'_1}, \dots, \Delta_{t'_i}, t_{i+1}) = U^{s'_0} M_1 \cdots M_i U^{t_{i+1}} D^{g'},$$

with $M_m = D^{s'_{m-1}} = D^{t_m}$ if $t'_m = 0$ and $s'_{m-1} = t_m$, or $M_m = \Delta_{t'_m} = \Delta_{t_m}$ if $s_{m-1} > 0$, for $1 \le m \le i$. This gives $\varphi(\phi(P)) = P$.
Let $\tau = (c_1, \Delta_{t_1}, \dots, c_i, \Delta_{t_i}, c_{i+1})$ with $c_m > 0$ and $t_m \ge 0$, then

$$\varphi(\tau) = U^{c_1} M_1 \cdots U^{c_i} M_i U^{c_{i+1}} D^g,$$

where

$$M_m = \begin{cases} D^{c_m}, & \text{if } t_m = 0; \\ \Delta_{t_m}, & \text{otherwise;} \end{cases}$$

and let $1 \le x_1 < x_2 < \cdots < x_q \le i$ be such that $t_{x_m} = 0$, that is the paths $\Delta_{t_{x_m}}$ in τ that are of the form $\Delta_0 = \lambda$. Notice that $c_m > 0$ and the definition of M_m imply that either $U^{c_m} M_m = \Delta_{c_m}$ exactly for $t_m = 0$ or $U^{c_m} \Delta_{t_m}$ for $t_m > 0$, which allows us to decompose $\varphi(\tau)$ as

$$\varphi(\tau) = \left(U^{c_1} \Delta_{t_1} \cdots U^{c_{x_1 - 1}} \Delta_{t_{x_1 - 1}}\right) U^0 \Delta_{c_{x_1}} \cdots U^0 \Delta_{c_{x_q}} \left(U^{c_{x_q+1}} \Delta_{t_{x_q+1}} \cdots \Delta_{t_i}\right) \Delta_{c_{i+1}} D^{g - c_{i+1}},$$

where every pyramid in the decomposition is non-empty and so the decomposition is unique. We now have that $\phi(\varphi(\tau)) = (c_1, \Delta_{t_1}, \dots, c_{x_1 - 1}, \Delta_{t_{x_i - 1}}, c_{x_1}, \lambda, \dots, c_{x_q}, \lambda, \dots, c_{i+1}) = \tau$.
For the inductive step $S(k)$, we assume that we have the desired equalities for $\ell < k$. Notice that any tuple $\tau \in \mathcal{QC}_k(n)$ can be decomposed as $\tau = (\tau_0, P_1, \tau_1)$ with τ_0 containing $\ell < k$ paths that are not pyramids, $P_1 \neq \Delta_t$ for any $t \ge 0$, and $\tau_1 \in \mathcal{QC}_0(n')$. Notice, further, that

$$\varphi((\tau_0, P_1, \tau_1)) = \varphi(\tau_0) D^{-} P_1 D \varphi(\tau_1) D^{g_2},$$

where $\varphi(\tau_0)D^{-}$ means deleting the South-East steps suffix of $\varphi(\tau_0)$. By the recursive step in the Algorithm 1, we have that $\phi(\varphi(\tau)) = \tau$ by using the inductive hypothesis. Analogously, we can decompose a path as in the recursive step of Algorithm 1, and the inductive hypothesis give $\varphi(\phi(P)) = P$. □

Proposition 3.5 is a direct consequence of the decomposition given in the proof of Theorem 3.1. The first result follows from Figure 4 and the second result uses the first part of this proposition and the decomposition UTD, $U\Delta DT$, or $UQDT$ as given in the proof of Theorem 3.1.

Proposition 3.5. *If $n > 1$, then these hold*

(1) If $q_n = |\mathcal{Q}_n|$, then

$$q_n = 2q_{n-1} + q_{n-2} + q_{n-3} + \sum_{i=2}^{n-4} q_i (q_{n-i-1} - q_{n-i-2}) + 1,$$

for $n > 3$, with the initial values $q_1 = 0$, $q_2 = 1$, and $q_3 = 3$.

(2) If $r(n) = |\mathcal{D}_d(n)|$, then

$$r(n) = 3r(n-1) - r(n-2) + q_{n-2} + \sum_{i=2}^{n-3} q_i (r(n-i-1) - r(n-i-2)),$$

for $n > 3$, with the initial values $r(1) = 1$, $r(2) = 2$, and $r(3) = 5$.

The generating function of the sequence $r(n)$ is algebraic of order two, then $r(n)$ satisfies a recurrence relation with polynomial coefficients; see [1, Proposition 4]. This can be automatically solved with implementation of Kauers in *Mathematica* [18]. In particular we obtain that $r(n)$ satisfies the recurrence relation:

$$2nr(n) - 4nr(n+1) + (12+5n)r(n+2) - 4(15+4n)r(n+3) \\ + 10(9+2n)r(n+4) - 2(21+4n)r(n+5) + (6+n)r(n+6) = 0, \quad \text{with } n \geq 6$$

and the initial values $r(0) = 0, r(1) = 1, r(2) = 2, r(3) = 5, r(4) = 14$, and $r(5) = 41$.

In Theorem 3.6 we give an asymptotic approximation for the sequence $r(n)$. To accomplish this goal, we use the singularity analysis method to find the asymptotes of the coefficients of a generating function (see, for example, [11] for the details).

We recall that in literature $f_n \sim g_n$ means that f_n and g_n are asymptotic equivalent. That is, $f_n/g_n \to 1$ as $n \to \infty$.

Theorem 3.6. *If ρ is the smallest real positive root of $1 - 4x + 2x^2 + x^4$, then the number of (-1)-Dyck paths has this asymptotic approximation*

$$r(n) \sim \frac{\rho^{-n}}{\sqrt{n^3\pi}} \cdot \frac{\sqrt{\rho(4-4\rho-4\rho^3)}}{4(-1+4\rho-2\rho^2)},$$

where ρ is called the dominant singularity of the generating function $L(x)$.

Proof. From a symbolic computation we find that

$$\rho = \frac{1}{3}\left(-1 - \frac{4\ 2^{2/3}}{\sqrt[3]{13+3\sqrt{33}}} + \sqrt[3]{2\left(13+3\sqrt{33}\right)}\right) \approx 0.295598.$$

From the expression given in (3.2) for $L(x)$ we have

$$L(x) = \frac{-1+4x-3x^2}{2(1-4x+2x^2)} + \frac{\sqrt{1-4x+2x^2+x^4}}{2(1-4x+2x^2)} \sim (\rho-x)^{1/2}\frac{\sqrt{\rho(4-4\rho-4\rho^3)}}{2(1-4\rho+2\rho^2)} \quad \text{as } x \to \rho^-.$$

Therefore,

$$r(n) \sim \frac{n^{-1/2-1}}{\rho^n(-2\sqrt{\pi})}\frac{\sqrt{\rho(4-4\rho-4\rho^3)}}{2(1-4\rho+2\rho^2)} = \frac{\rho^{-n}}{\sqrt{n^3\pi}}\frac{\sqrt{\rho(4-4\rho-4\rho^3)}}{4(-1+4\rho-2\rho^2)}.$$ □

4. The area of the (-1)-Dyck paths

In this section we use generating functions and recursive relations to analyze the distribution of the area of the paths in the set of restricted (-1)-Dyck paths. We recall that the *area* of a Dyck path is the sum of the absolute values of y-components of all points in the path. We use $\texttt{area}(P)$ to denote the area of a path P. From Figure 1 on Page 2, we can see that $\texttt{area}(P) = 70$. We use $a(n)$ to denote the total area of all paths in $\mathcal{D}_{-1}(n)$. In Theorem 4.1 we give a generating function for the sequence $a(n)$. We now introduce a bivariate generating function depending on this previous parameter and $\ell(P)$ (the semi-length of P). So,

$$A(x,q) := \sum_{P \in \mathcal{D}_{-1}} x^{\ell(P)} q^{\texttt{area}(P)}.$$

Let $\mathcal{Q} \subset \mathcal{D}_{-1}(n)$ be the set formed by all paths having at least one valley, where the last valley is at ground level; let $\mathcal{Q}_n \subset \mathcal{Q}$ be the set formed by all paths of semi-length n, and let $q_n = |\mathcal{Q}_n|$.

Theorem 4.1. *The generating function for the sequence $a(n)$ is given by*

$$V(x) = \sum_{n\geq 0} a(n)x^n = \frac{b(x) - c(x)\sqrt{1-4x+2x^2+x^4}}{(1-x)^2(1-4x+2x^2)^3(1-3x-x^2-x^3)},$$

where

$$b(x) = 2x - 23x^2 + 107x^3 - 262x^4 + 359x^5 - 256x^6 + 82x^7 - 5x^8 - 10x^9 + 6x^{10},$$
$$c(x) = x - 10x^2 + 41x^3 - 89x^4 + 108x^5 - 73x^6 + 18x^7 + 2x^8.$$

Proof. From the decomposition UD, UTD, $U\Delta DT$, or $UQDT$ given in the proof of Theorem 3.1 we obtain the functional equation

$$A(x,q) = xq + xqA(xq^2,q) + E(x,q)A(x,q) + xqB(xq^2,q)A(x,q), \tag{4.1}$$

where $E(x,q) := \sum_{j\geq 1} x^j q^{j^2}$ and $B(x,q) := \sum_{P\in\mathcal{Q}} x^{\ell(P)} q^{\texttt{area}(P)}$. Note that $E(x,q)$ corresponds to the generating function that counts the total number of non-empty pyramids in the given decomposition.
From the decomposition given in Figure 4, we obtain the functional equation

$$B(x,q) = E(x,q)^2 + E(x,q)B(x,q) + xqB(q^2x,q)B(x,q) + xqB(q^2x,q)E(x,q). \tag{4.2}$$

Let $M(x)$ be the generating function of the total area of the (-1)-Dyck paths in $\mathcal{Q}$. From the definition of $A(x,q)$ we have

$$V(x) = \left.\frac{\partial A(x,q)}{\partial q}\right|_{q=1}.$$

Substituting x by xq^2 in (4.2), and then differentiating with respect to q and taking $q=1$, we obtain

$$\begin{aligned} W(x) := \left.\frac{\partial B(xq^2,q)}{\partial q}\right|_{q=1} &= \frac{2(3-x)x^2}{(1-x)^4} + \frac{(3-x)x}{(1-x)^3}Q(x) + \frac{x}{1-x}\left(W(x) + 2x\frac{\partial Q(x)}{\partial x}\right) \\ &+ 3xQ(x)^2 + xQ(x)\left(W(x) + 4x\frac{\partial Q(x)}{\partial x}\right) + xQ(x)\left(W(x) + 2x\frac{\partial Q(x)}{\partial x}\right) \\ &+ \frac{3x^2}{1-x}Q(x) + \frac{x^2}{1-x}\left(W(x) + 4x\frac{\partial Q(x)}{\partial x}\right) + \frac{x^2(3-x)}{(1-x)^3}Q(x), \end{aligned} \tag{4.3}$$

where $Q(x) := Q(x,1)$ and $Q(x,y)$ is the generating function given in (3.1) on Page 8.

Now, differentiating (4.1) with respect to q and then taking $q=1$ we obtain,

$$\begin{aligned} V(x) &= x + xL(x) + x\left(V(x) + 2x\frac{\partial L(x)}{\partial x}\right) + \frac{x(x+1)}{(1-x)^3}L(x) \\ &+ \frac{x}{1-x}V(x) + xQ(x)L(x) + xW(x)L(x) + xQ(x)V(x). \end{aligned} \tag{4.4}$$

Solving (4.3) for $W(x)$ and substituting into (4.4) and then solving the resulting expression for $V(x)$ we obtain the desired result. □

The first few values of the series of $V(x)$ are

$$V(x) = \sum_{n\geq 1} a(n)x^n = x + 6x^2 + 29x^3 + 130x^4 + 547x^5 + 2198x^6 + 8551x^7 + \cdots .$$

We now give a recursive relation for $a(n)$. Again for the sake of simplicity, the proof here is based on a geometric decomposition of the paths. So, we avoid some details. However, in [13] there are detailed proofs of Proposition 4.2 and Theorem 4.3. We recall that $q_n = |\mathcal{Q}_n|$ and that for simplicity we use $r(n)$ instead of $r_{-1}(n)$.

The following two results may follow as a direct application of (4.2). However, we include here a different combinatorial proof.

Proposition 4.2. *If A_n with $n \geq 1$ is the total area of all paths in $\mathcal{Q}_n$, then*

$$A_n = 2A_{n-1} + A_{n-2} + 2A_{n-3} + q_n - q_{n-1} + 2nq_{n-2} + 2(n-5)q_{n-3} + 4n^2 - 14n + 13 + \sum_{i=2}^{n-4} 2(A_i + iq_i + i(i+1))(q_{n-i-1} - q_{n-i-2}), \quad \text{with } n > 4,$$

and the initial values $A_1 = 0$, $A_2 = 2$, $A_3 = 13$, and $A_4 = 58$.

Proof. From Figure 4 we know that a path in $\mathcal{Q}_n$ can be decomposed in one of these four cases; $\Delta_i\Delta_{n-i}$, $\Delta_i Q$, $XQY\Delta_i$, $XQ'YQ$ where $Q, Q' \in \mathcal{Q}$

Case 1. The area of $\Delta_i\Delta_{n-i}$ is $i^2 + (n-i)^2$. Since for a fixed $i \in \{1, 2, \dots, n-1\}$, there is exactly one path of the form $\Delta_i\Delta_{n-i}$ in $\mathcal{Q}_n$, we have that the total area of this type of paths is $\sum_{i=1}^{n-1}(i^2 + (n-i)^2) = n(n-1)(2n-1)/3$.

Case 2. The area of $P_i := \Delta_i Q$ is $i^2 + A_{n-i}$. Since for every $i \in \{1, 2, \dots, n-2\}$ there are q_{n-i} paths of the form P_i, we have that the total area of all paths of the form P_i is given by $i^2 q_{n-i} + A_i$. Therefore, the total area of this type of paths is $\sum_{i=1}^{n-2} i^2 q_{n-i} + \sum_{j=2}^{n-1} A_j$.

Case 3. For a fixed i, the area of a path of the form $XQ'YQ''$ is given by $2i + 1 + A_i + A_{n-i-1}$, where $Q' \in \mathcal{Q}_i$, $Q'' \in \mathcal{Q}_{n-i-1}$ and $i \in \{2, 3, \dots, n-3\}$. Note that for a fixed i and a fixed $Q \in \mathcal{Q}_{n-i-1}$ there q_i paths of the form $XQ'YQ$ with $Q \in \mathcal{Q}_i$. This implies that for a fixed $i \in \{2, 3, \dots, n-3\}$ the total area of this type of paths is $A_{n-i-1}q_i + (2i+1)q_i q_{n-i-1} + A_i q_{n-i-1}$. We conclude for i varying from 2 to $n-3$, we obtain that the total area of this type of paths is

$$\sum_{i=2}^{n-3} A_{n-i-1}q_i + \sum_{i=2}^{n-3}((2i+1)q_i q_{n-i-1} + A_i q_{n-i-1}).$$

Case 4. The area of $H_i := XQ_\ell Y\Delta_i$ is given by area of Δ_i (which is i^2) plus the area of $XQ_\ell Y$ (this is given by A_ℓ, the area of Q_ℓ, plus $2i+1$ which is the area of the trapezoid generated by X and Y). Since for every $i \in \{1, 2, \dots, n-3\}$ there are q_{n-i-1} paths of the form H_i with $Q \in \mathcal{Q}_{n-i}$, we conclude that the total area of this type of paths is

$$\sum_{i=1}^{n-3} i^2 q_{n-i-1} + \sum_{i=2}^{n-2}((2i+1)q_i + A_i).$$

Adding the results from Cases 1-4, we obtain that the recursive relation for the area A_n is given by

$$A_n = \sum_{i=1}^{n-1}\left(i^2+(n-i)^2\right) + \sum_{i=1}^{n-2} i^2 q_{n-i} + \sum_{i=2}^{n-1} A_i + \sum_{i=2}^{n-3}(2i+1)q_i q_{n-(i+1)} + \sum_{i=2}^{n-3} A_i q_{n-(i+1)} + \sum_{i=2}^{n-3} A_i q_{n-(i+1)} + \sum_{i=2}^{n-2} A_i + \sum_{i=1}^{n-3} i^2 q_{n-(i+1)} + \sum_{i=2}^{n-2}(2i+1)q_i.$$

Subtracting A_n from A_{n+1} and simplifying we have

$$A_n = 2A_{n-1} + A_{n-2} + 2A_{n-3} + (2n-5)q_{n-3} + (2n-4)q_{n-2} + q_{n-1} + 4n^2 - 14n + 15 + \sum_{i=2}^{n-4}(2A_i + (2i+1)q_i)(q_{n-i-1} - q_{n-i-2}) + \sum_{i=2}^{n-3}\left(2i^2 - 2i + 1\right)(q_{n-i} - q_{n-i-1}).$$

We now rearrange this expression to obtain q_n (see the expression within brackets) given in Proposition 3.5

$$A_n = 2A_{n-1} + A_{n-2} + 2A_{n-3} + (2n-6)q_{n-3} + (2n-4)q_{n-2} - q_{n-1} + 4n^2 - 14n + 13 + \sum_{i=2}^{n-4} 2(A_i + iq_i)(q_{n-i-1} - q_{n-i-2}) + \sum_{i=2}^{n-3} 2\left(i^2 - i\right)(q_{n-i} - q_{n-i-1}) + [2q_{n-1} + q_{n-2} + q_{n-3} + \sum_{i=2}^{n-4} q_i(q_{-i+n-1} - q_{-i+n-2}) + 1].$$

After some simplifications we obtain the desired recursive relation. □

The proof of the following theorem is similar to the proof of Proposition 4.2. We recall that $r(i) = |\mathcal{D}_{-1}(i)|$ and $q_j = |\mathcal{Q}_j|$.

Theorem 4.3. *If $a(n)$ is the total area of all paths in $\mathcal{D}_{-1}(n)$, for $n \geq 1$, then $a(n)$ satisfies the recursive relation*

$$a(n) = 3a(n-1) - a(n-2) + A_{n-2} + 2(n-1)q_{n-2} + 2nr(n-1) + 2(3-n)r(n-2) - 4r(n-3) + (n-1)^2 + \sum_{i=3}^{n-2} q_{i-1}(a(n-i) - a(n-i-1)) + \sum_{i=3}^{n-2}\left(A_{i-1} + (2i-1)q_{i-1} + i^2\right)(r(n-i) - r(n-i-1)).$$

Proof. First of all, we note that a path in $\mathcal{D}_{-1}(n)$ can be decomposed as XQ_1Y, $\Delta_i Q_{n-i}$, and $XQ'YD$, where $Q_j, D \in \mathcal{D}_{-1}$, and $Q' \in \mathcal{Q}_j$. This decomposition gives these three cases to consider.

Case 1. The area of XQY is $(2n-1) + a(n-1)$, where $a(n-1)$ is the area of $Q \in \mathcal{D}_{-1}(n-1)$ and $2n-1$ is the are of the trapezoid generated by X and Y. This gives that the total area of all paths of the form XQY with $Q \in \mathcal{D}_{-1}(n-1)$ is $(2n-1)r(n-1) + a(n-1)$.

Case 2. The area of $K_i := X^iY^iQ_\ell$ is $i^2 + a(n-i)$, where $Q_\ell \in \mathcal{D}_{-1}(n-i)$. Since for a fixed $i \in \{1, 2, \dots, n-1\}$ there are $r(n-i)$ paths of form K_i, we conclude that the total area of all these paths is $\sum_{i=1}^{n-1} i^2 r(n-i) + a(n-i)$.

Case 3. The area of $M_i := XQ'YD$ is $((2i+1)+A_i+a(n-i-1)$, where $Q' \in \mathcal{Q}_i$ and $D \in \mathcal{D}_{-1}(n-i-1)$. Note that for a given path $D \in \mathcal{D}_{-1}(n-i-1)$, there are as many paths of the form $XQ'YD$ as paths in $\mathcal{Q}_i$. It is easy to see that for a fixed $i \in \{2,3,\dots,n-2\}$ there are $r(n-i-1)$ subpaths of the form $XQ'Y$. Note that X and Y give rise to a trapezoid, where the two parallel sides have lengths $2i$ and $2i+2$, giving rise to an area of $2i+1$. So, the contribution to the area given by the first subpaths of the form $XQ'Y$ is equal to the area of the trapezoids plus the area of all paths of the form Q' (these are on top of the trapezoids). Thus, the area of a trapezoid multiplied by the total number of the paths of the form Q' plus the area of all paths of the form Q' and then all of these multiplied by the total number of paths of the form D. Thus, the contribution to the area given by the first subpaths of the form $XQ'Y$ (overall paths of the form M_i for a fixed i), is $((2i+1)q_ir(n-i-1)+A_ir(n-i-1))$.
We conclude that the total area of this type of paths is

$$\sum_{i=2}^{n-2} A_ir(n-i-1) + \sum_{i=2}^{n-2}(2i+1)q_ir(n-i-1).$$

Adding the results from Cases 1-3, we obtain that the recursive relation for the area $a(n)$ is given by

$$\begin{aligned}a(n) = a(n-1) + (2n-1)r(n-1) + \sum_{i=1}^{n-1} i^2r(n-i) + \sum_{i=1}^{n-1} a(n-i)\\ + \sum_{i=2}^{n-2} q_ia(n-i-1) + \sum_{i=2}^{n-2} A_ir(n-i-1) + \sum_{i=2}^{n-2}(2i+1)q_ir(n-i-1).\end{aligned}$$

Subtracting $a(n)$ from $a(n+1)$ and simplifying we have

$$\begin{aligned}a(n) = 3a(n-1)-a(n-2)+A_{n-2}+2(n-1)q_{n-2}+(2n-1)r(n-1)+(3-2n)r(n-2)+(n-1)^2\\ + \sum_{i=3}^{n-2} q_{i-1}(a(n-i)-a(n-i-1)) + \sum_{i=3}^{n-2} A_{i-1}(r(n-i)-r(n-i-1))\\ + \sum_{i=3}^{n-2}(2i-1)q_{i-1}(r(n-i)-r(n-i-1)) + \sum_{i=1}^{n-2} i^2(r(n-i)-r(n-i-1)).\end{aligned}$$

After some other simplifications we have that

$$\begin{aligned}a(n) = 3a(n-1) - a(n-2) + A_{n-2} + 2(n-1)q_{n-2} + 2nr(n-1)\\ + 2(3-n)r(n-2) - 4r(n-3) + (n-1)^2 + \sum_{i=3}^{n-2} q_{i-1}(a(n-i)-a(n-i-1))\\ + \sum_{i=3}^{n-2}\left(A_{i-1}+(2i-1)q_{i-1}+i^2\right)(r(n-i)-r(n-i-1)).\end{aligned}$$

This completes the proof. □

Notice that the total area of the Dyck paths (cf. [21]) is given by $4^n - \binom{2n+1}{n}$.

5. Appendix. Notation table

Concept	Notation
Set restricted d-Dyck paths	$\mathcal{D}_d$
Set restricted d-Dyck paths of length n	$\mathcal{D}_n$
Cardinality of $\mathcal{D}_d(n)$	$r_d(n)$
Cardinality of $\mathcal{D}_{-1}(n)$	$r_{-1}(n)$ or $r(n)$
Area of a path P	$\texttt{area}(\mathrm{P})$
Semi-length of P	$\ell(P)$
Number of peaks of P	$\rho(P)$
Number of paths in $\mathcal{D}_d(n)$ having exactly k peaks.	$p_d(n,k)$
Paths with the last valley at level i	$\mathcal{Q}_{d,i}$
General pyramid	Δ
Pyramid $(XY)^k$	Δ_k

Table 1. Summary of notation.

Acknowledgments. The authors are grateful to the referees and the editor for the helpful suggestions and comments. The first author was partially supported by The Citadel Foundation. The third author was partially supported by Universidad Nacional de Colombia. The last author thanks the ANTA group at Aalto University where part of the presented work was written.

References

[1] Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Combin. Probab. Comput.*, 24(1):1–53, 2015.

[2] Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theor. Comput. Sci.*, 281:37–80, 2002.

[3] Cyril Banderier and Bernhard Gittenberger. Analytic combinatorics of lattice paths: enumeration and asymptotics for the average area. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, pages 345–356. Discrete Math. Theor. Comput. Sci., 2006.

[4] Elena Barcucci, Alberto Del Lungo, Stefano Fezzi, and Renzo Pinzani. Nondecreasing Dyck paths and q-Fibonacci numbers. *Discrete Math.*, 170:211–217, 1997.

[5] Paul Barry. Riordan pseudo-involutions, continued fractions and Somos-4 sequences. *J. Integer Seq.*, 22:Article 19.6.1, 2019.

[6] Éva Czabarka, Rigoberto Flórez, and Leandro Junes. Some enumerations on non-decreasing Dyck paths. *Electron. J. Comb.*, 22(1):# P1.3, 1–22, 2015.

[7] Éva Czabarka, Rigoberto Flórez, Leandro Junes, and José L. Ramírez. Enumerations of peaks and valleys on non-decreasing Dyck paths. *Discrete Math.*, 341:2789–2807, 2018.

[8] Emeric Deutsch. Dyck path enumeration. *Discrete Math.*, 204:167–202, 1999.

[9] Emeric Deutsch and Helmut Prodinger. A bijection between directed column-convex polyominoes and ordered trees of height at most three. *Theor. Comput. Sci.*, 307:319–325, 2003.

[10] Sergi Elizalde, Rigoberto Flórez, and José L. Ramírez. Enumerating symmetric peaks in non-decreasing Dyck paths. *Ars Math. Contemp.*, 21(2):# P2.04, 2021.

[11] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[12] Rigoberto Flórez, Leandro Junes, and José L. Ramírez. Enumerating several aspects of non-decreasing Dyck paths. *Discrete Math.*, 342(11):3079–3097, 2019.

[13] Rigoberto Flórez, Toufik Mansour, José L. Ramírez, Fabio A. Velandia, and Diego Villamizar. Restricted Dyck paths on valleys sequence. *arXiv*, 2021.

[14] Rigoberto Flórez and José L. Ramírez. Some enumerations on non-decreasing Motzkin paths. *Australas. J. Comb.*, 72:138–154, 2018.

[15] Rigoberto Flórez and José L. Ramírez. Enumerations of rational non-decreasing Dyck paths with integer slope. *Graphs Comb.*, 37(6):2775–2801, 2021.

[16] Rigoberto Flórez, José L. Ramírez, Fabio A. Velandia, and Diego Villamizar. Some connections between restricted Dyck paths, polyominoes, and non-crossing partitions. In *Proceedings ot he 52th Southeastern International Conference on Combinatorics, Graph Theory & Computing*. To appear in Springer Proceedings in Mathematics & Statistics, 2021.

[17] Rigoberto Flórez, José L. Ramírez, Fabio A. Velandia, and Diego Villamizar. A refinement of Dyck paths: a combinatorial approach. *Discrete Math. Algorithms Appl.*, 14(7), 2022.

[18] Manuel Kauers. Guessing handbook. *Technical Report 09-07, RISC-Linz*, 2009.

[19] Aristidis Sapounakis, Ioannis Tasoulas, and Panagiotis Tsikouras. Ordered trees and the inorder traversal. *Discrete Math.*, 306(15):1732–1741, 2006.

[20] Aristidis Sapounakis and Panagiotis Tsikouras. Counting peaks and valleys in k-colored Motzkin paths. *Electron. J. Comb.*, 12:R#16, 2005.

[21] Wen-jin Woan. Area of Catalan paths. *Discrete Math.*, 226(15):439–444, 2001.