

A WALK IN MY LATTICE PATH GARDEN

HELMUT PRODINGER

Dedicated to Neil Peart.

ABSTRACT. Various lattice path models are reviewed. The enumeration is done using generating functions. A few bijective considerations are woven in as well. The kernel method is often used. Computer algebra was an essential tool. Some results are new, some have appeared before.

The lattice path models we treated, are: Hoppy's walks, the combinatorics of sequence A002212 in [47] (skew Dyck paths, Schröder paths, Hex-trees, decorated ordered trees, multi-edge trees, etc.) Weighted unary-binary trees also occur, and we could improve on our old paper on Horton-Strahler numbers [17], by using a different substitution. Some material on ternary trees appears as well, as on Motzkin numbers and paths (a model due to Retakh), and a new concept called amplitude that was found in [6]. Some new results on Deutsch paths in a strip are included as well. During the Covid period, I spent much time with this beautiful concept that I dare to call Deutsch paths, since Emeric Deutsch stands at the beginning with a problem that he posted in the American Mathematical Monthly some 20 years ago.

CONTENTS

1. Introduction	2
2. Hoppy walks	2
3. Combinatorics of sequence A002212 in [47]	7
4. Binary trees and Horton-Strahler numbers	8
5. Marked ordered trees	14
6. A bijection between multi-edge trees and 3-coloured Motzkin paths	17
7. The combinatorics of skew Dyck paths	18
8. Counting ternary trees according to the number of middle edges and factorizing into $(3/2)$ -ary trees	33
9. More about Motzkin paths	36
10. The amplitude of Motzkin paths	40
11. Oscillations in Dyck paths revisited	45
12. Deutsch-paths in a strip	49
13. Publication status of the problems treated in this walk in the garden	54
References	54

Date: November 24, 2021.

Key words and phrases. Skew Dyck paths, decorated Dyck paths, generating functions, Motzkin paths, kernel method.

1. INTRODUCTION

Around 20 years ago I published a collection of examples about applications of the kernel method in the present journal. The success of this enterprise was unexpected and came as a very pleasant surprise. My current plan is to present again a collection of subjects, loosely related as they all have a lattice path flavour (trees are also allowed in my private book when lattice paths are mentioned). The subjects cover my last 2 or 3 years of research; some results were only posted on the arxiv, some were submitted but no evaluations were available up to this day, some have been appearing, and some are completely new. In each instance I try to make clear the current status.

As in the predecessor paper, the kernel method plays a role again, but also analytic techniques like singularity analysis and Mellin transform, as well as bijective results.

2. HOPPY WALKS

Deng and Mansour [7] introduce a rabbit named Hoppy and let him move according to certain rules. While the story about Hobby is charming and entertaining, we do not need this here and move straight ahead to the enumeration issues. Eventually, the enumeration problem is one about k -Dyck paths. The up-steps are $(1, k)$ and the down-steps are $(1, -1)$. The model that has $(1, 1)$ as up-step and the down-step are $(1, -k)$ will also be called k -Dyck paths.

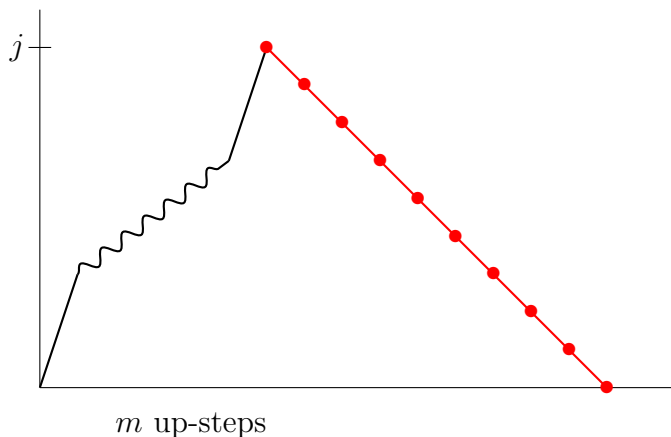


FIGURE 1. The number of final down-steps

The question is about the length of the sequence of down-steps printed in red. Or, phrased differently, how many k -Dyck paths end on level j , after m up-steps, the last step being an up-step. The recent paper [31] contains similar computations, although without the restriction that the last step must be an up-step.

Counting the number of up-steps is enough, since in total, there are $m + km = (k + 1)m$ steps. The original description of Deng and Mansour is a reflection of this picture, with up-steps of size 1 and down-steps of size $-k$, but we prefer it as given here, since we are going to use the adding-a-new-slice method, see [16, 30]. A slice is here a run of down-steps, followed by an up-step. The first up-step is treated separately, and then $m - 1$ new slices are added. We keep track of the level after each slice, using a variable u . The variable z is used to count the number of up-steps.

Deng and Mansour work out a formula which comprises $O(m)$ terms. Our method leads only to a sum of $O(j)$ terms.

We will use the adding a new slice technique [30] together with the kernel method to find useful bivariate generating functions.

The following substitution is essential for adding a new slice (one up-step followed by a maximal sequence of down-steps):

$$u^j \longrightarrow z \sum_{0 \leq h \leq j} u^{h+k} = \frac{zu^k}{1-u}(1-u^{j+1}).$$

Now let $F_m(z, u)$ be the generating function according to m runs of down-steps. The substitution leads to

$$F_{m+1}(z, u) = \frac{zu^k}{1-u}F_m(z, 1) - \frac{zu^{k+1}}{1-u}F_m(z, u), \quad F_0(z, u) = zu^k.$$

Let $F = \sum_{m \geq 0} F_m$, so that we don't care about the number m anymore; then

$$F(z, u) = zu^k + \frac{zu^k}{1-u}F(z, 1) - \frac{zu^{k+1}}{1-u}F(z, u),$$

or

$$F(z, u) \frac{1-u+zu^{k+1}}{1-u} = zu^k + \frac{zu^k}{1-u}F(z, 1).$$

The equation $1-u+zu^{k+1} = 0$ is famous when enumerating $(k+1)$ -ary trees (or k -Dyck paths). Its relevant combinatorial solution (also the only one being analytic at the origin) is

$$\bar{u} = \sum_{\ell \geq 0} \frac{1}{1+\ell(k+1)} \binom{1+\ell(k+1)}{\ell} z^\ell.$$

Since $u - \bar{u}$ is a factor of the LHS, it must also be a factor of the RHS, and we can compute (by dividing out the factor $(u - \bar{u})$) that

$$\frac{zu^k(1-u+F(z, 1))}{u-\bar{u}} = -zu^k.$$

Thus

$$F(z, u) = zu^k \frac{\bar{u} - u}{1 - u + zu^{k+1}}.$$

The first factor has even a combinatorial interpretation, as a description of the first step of the path. It is also clear from this that the level reached is $\geq k$ after each slice. We don't care about the factor zu^k anymore, as it produces only a simple shift. The main interest is now how to get to the coefficients of

$$\frac{\bar{u} - u}{1 - u + zu^{k+1}}$$

in an efficient way. There is also the formula

$$1 - u + zu^{k+1} = (\bar{u} - u) \left(1 - z \frac{u^{k+1} - \bar{u}^{k+1}}{u - \bar{u}} \right),$$

but it does not seem to be useful here.

First we deal with the denominators

$$S_j := [u^j] \frac{1}{1-u+zu^{k+1}} = \sum_{0 \leq m \leq j/k} (-1)^m \binom{j-km}{m} z^m.$$

One way to see this formula is to prove by induction that the sums S_j satisfy the recursion

$$S_j - S_{j-1} + zS_{j-k-1} = 0$$

and initial conditions $S_0 = \dots = S_k = 1$. In [31] such expressions also appear as determinants. Summarizing,

$$\frac{1}{1-u+zu^{k+1}} = \sum_{m \geq 0} (-1)^m z^m \sum_{j \geq km} \binom{j-km}{m} u^j.$$

Now we read off coefficients:

$$[u^j] \frac{\bar{u}}{1-u+zu^{k+1}} = \sum_{0 \leq m \leq j/k} (-1)^m \binom{j-km}{m} z^m \sum_{\ell \geq 0} \frac{1}{1+\ell(k+1)} \binom{1+\ell(k+1)}{\ell} z^\ell$$

and further

$$\begin{aligned} [z^n][u^j] \frac{\bar{u}}{1-u+zu^{k+1}} \\ = \sum_{0 \leq m \leq j/k} (-1)^m \binom{j-km}{m} \frac{1}{1+(n-m)(k+1)} \binom{1+(n-m)(k+1)}{n-m}. \end{aligned}$$

The final answer to the Deng-Mansour enumeration (without the shift) is

$$\begin{aligned} \sum_{0 \leq m \leq j/k} (-1)^m \binom{j-km}{m} \frac{1}{1+(n-m)(k+1)} \binom{1+(n-m)(k+1)}{n-m} \\ - (-1)^n \binom{j-1-kn}{n}. \end{aligned}$$

If one wants to take care of the factor zu^k as well, one needs to do the replacements $n \rightarrow n+1$ and $j \rightarrow j+k$ in the formula just derived. That enumerates then the k -Dyck paths ending at level j after n up-steps, where the last step is an up-step.

An application. The encyclopedia of integer sequences [47] has the sequences A334680, A334682, A334719, (with a reference to [1]) which is the total number of down-steps of the last down-run, for $k = 2, 3, 4$. So, if the path ends on level j , the contribution to the total is j .

All we have to do here is to differentiate

$$F(z, u) = zu^k \frac{\bar{u} - u}{1-u+zu^{k+1}}.$$

w.r.t. u , and then replace u by 1. The result is

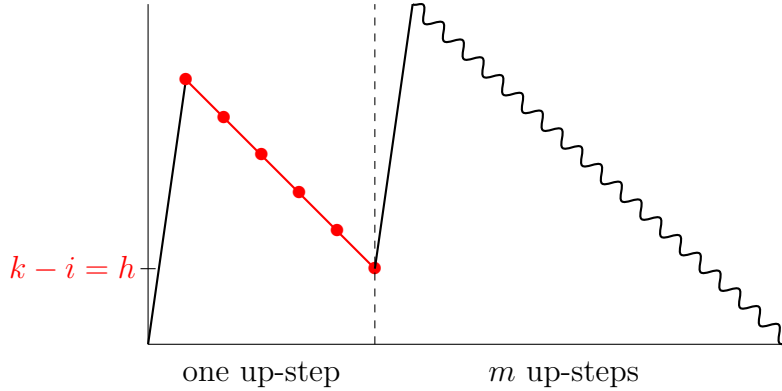
$$\frac{\bar{u}}{z} - \bar{u} - \frac{1}{z},$$

and the coefficient of z^m therein is

$$\frac{1}{1+(m+1)(k+1)} \binom{1+(m+1)(k+1)}{m+1} - \frac{1}{1+m(k+1)} \binom{1+m(k+1)}{m}.$$

The bivariate generating function does this enumeration cleanly and quickly.

Hoppy's early adventures. Now we investigate what Hoppy does after his first up-step; he might follow with $0, 1, \dots, k$ down-steps. Eventually, we want to sum all these steps (red in the picture).



A new slice is now an up-step, followed by a sequence of down-steps. The substitution of interest is:

$$u^i \rightarrow z \sum_{0 \leq h \leq i+k} u^h = \frac{z}{1-u} - \frac{zu^{i+k+1}}{1-u}.$$

Furthermore

$$F_{h+1}(z, u) = \frac{z}{1-u} F_h(z, 1) - \frac{zu^{k+1}}{1-u} F_h(z, u),$$

and $F_0 = u^h$, the starting level.

We have

$$H(z, u) = \sum_{h \geq 0} F_h(z, u) = u^h + \frac{z}{1-u} H(z, 1) - \frac{zu^{k+1}}{1-u} H(z, u)$$

or

$$H(z, u)(1 - u + zu^{k+1}) = u^h(1 - u) + zH(z, 1).$$

Plugging in \bar{u} into the RHS gives 0:

$$zH(z, 1) = -\bar{u}^h(1 - \bar{u}),$$

and

$$H(z, u) = \frac{u^h(1 - u) - \bar{u}^h(1 - \bar{u})}{1 - u + zu^{k+1}}.$$

But we only need $H(z, 0)$, since we return to the x -axis at the end:

$$H(z, 0) = [h = 0] + \bar{u}^{h+1} - \bar{u}^h.$$

The total contribution of red steps is then

$$k + \sum_{h=0}^k (k - h)(\bar{u}^{h+1} - \bar{u}^h) = \sum_{h=1}^k \bar{u}^h;$$

the coefficient of z^m in this is the total contribution. Since $\bar{u} = 1 + z\bar{u}^{k+1}$, there is the further simplification

$$-1 + \frac{1}{z} + \frac{1}{1 - \bar{u}} = \sum_{m \geq 1} \frac{k}{m+1} \binom{(k+1)m}{m} z^m.$$

The proof of this is as follows. Let $m \geq 1$, then

$$\begin{aligned} [z^m] \left(-1 + \frac{1}{z} + \frac{1}{1 - \bar{u}} \right) &= -[z^m] \frac{1}{z\bar{u}^{k+1}} \\ &= -[z^{m+1}] \sum_{\ell \geq 0} \frac{-(k+1)}{(k+1)\ell - (k+1)} \binom{(k+1)\ell - (k+1)}{\ell} z^\ell \\ &= [z^{m+1}] \sum_{\ell \geq 0} \frac{(k+1)}{(k+1)(\ell - 1)} \binom{(k+1)(\ell - 1)}{\ell} z^\ell \\ &= \frac{(k+1)}{(k+1)m} \binom{(k+1)m}{m+1} = \frac{k}{m+1} \binom{(k+1)m}{m}. \end{aligned}$$

We did not expect such a simple answer $\frac{k}{m+1} \binom{(k+1)m}{m}$ to this question about Hoppy's early adventures!

This analysis of Hoppy's early adventures covers sequences A007226, A007228, A124724 of [47], with references to [1].

Hoppy walks into negative territory. Hoppy is now adventurous and allows himself to go to level -1 as well, but not deeper. The setup with generating functions is the same, but the u -variable counts the level relative to the -1 level, so this has to be corrected later.

Hoppy, after some initial frustration discovers that he can now start with an up-step or a down-step!

First, let us start Hoppy with an up-step:

$$F(z, u) = zu^{k+1} + \frac{zu^k}{1-u} F(z, 1) - \frac{zu^{k+1}}{1-u} F(z, u),$$

or

$$F(z, u) \frac{1-u+zu^{k+1}}{1-u} = zu^{k+1} + \frac{zu^k}{1-u} F(z, 1).$$

Since $u - \bar{u}$ is a factor of the LHS, it must also be a factor of the RHS, and we can compute that

$$\bar{u}(1 - \bar{u}) + F(z, 1) = 0.$$

The RHS is

$$\frac{zu^k}{1-u} \left(u(1-u) - \bar{u}(1-\bar{u}) \right)$$

and

$$F(z, u) = \frac{zu^k}{1-u+zu^{k+1}} \left(u(1-u) - \bar{u}(1-\bar{u}) \right).$$

But Hoppy can also start with a downstep! So we have to add the result of the previous computation, and get finally

$$G(z, u) = \frac{zu^k}{1-u+zu^{k+1}} \left(u(1-u) - \bar{u}(1-\bar{u}) \right) + \frac{zu^k}{1-u+zu^{k+1}} (\bar{u} - u)$$

or better

$$G(z, u) = \frac{zu^k}{1 - u + zu^{k+1}} (\bar{u}^2 - u^2)$$

Now we need

$$\frac{\partial}{\partial u} G(z, 1) - G(z, 1).$$

This subtraction is necessary, since the contribution of u^j is not j as before but only $j - 1$. The result is

$$\frac{\bar{u}^2}{z} - 2\bar{u}^2 - \frac{1}{z}.$$

Hoppy knows that \bar{u}^d has beautiful coefficients:

$$\bar{u}^d = \sum_{\ell \geq 0} \binom{d-1+(k+1)\ell}{\ell} \frac{d}{k\ell+d}$$

and he inserts $k = 2$ (A030983):

$$3z + 16z^2 + 83z^3 + 442z^4 + 2420z^5 + \dots$$

$k = 3$ (A334608):

$$5z + 34z^2 + 236z^3 + 1714z^4 + 12922z^5 + \dots$$

$k = 4$ (A334610):

$$7z + 58z^2 + 505z^3 + 4650z^4 + 44677z^5 + \dots$$

In general,

$$\frac{\bar{u}^2}{z} - 2\bar{u}^2 - \frac{1}{z} = \sum_{\ell \geq 0} \left[\binom{1+(k+1)(\ell+1)}{\ell+1} \frac{2}{k(\ell+1)+2} - 2 \binom{1+(k+1)\ell}{\ell} \frac{2}{k\ell+2} \right] z^\ell.$$

Happy Hoppy decides to stop this line of computations here.

3. COMBINATORICS OF SEQUENCE A002212 IN [47]

The following sections gives some (mostly new) results about the sequence

$$1, 1, 3, 10, 36, 137, 543, 2219, 9285, 39587, 171369, 751236, 3328218, 14878455, \dots,$$

which is A002212 in [47].

Here is the plan about the structures enumerated by this sequence:

Hex-trees [23]; they are identified as weighted unary-binary trees, with weight one. Apart from left and right branches, as in binary trees, there are also unary branches, and they can come in different colours, here in just one colour. Unary-binary trees played a role in the present authors scientific development, as documented in [17], a paper written with the late and great Philippe Flajolet, about the register function (Horton-Strahler numbers) of unary-binary trees. Here, we can offer an improvement, using a “better” substitution than in [17]. The results can now be made fully explicit. As a by-product, this provides a definition and analysis of the Horton-Strahler numbers of Hex-trees. An introductory section (about binary trees) provides all the basics.

Then we move to skew Dyck paths [9]. They are like Dyck paths, but allow for an extra step $(-1, -1)$, provided that the path does not intersect itself. An equivalent model, defined and described using a bijection, is from [9]: Marked ordered trees. They are like ordered trees, with an additional feature, namely each rightmost edge (except

for one that leads to a leaf) can be coloured with two colours. Since we find this class of trees to be interesting, we analyze two parameters of them: number of leaves and height. While the number of leaves for ordered trees is about $n/2$, it is only $n/10$ in the new model. For the height, the leading term $\sqrt{\pi n}$ drops to $\frac{2}{\sqrt{5}}\sqrt{\pi n}$. Of course, many more parameters of this new class of trees could be investigated, which we encourage to do.

More about skew Dyck paths appears in a later section.

The next two classes of structures are multi-edge trees. Our interest in them was already triggered in an earlier publication [21]. They may be seen as ordered trees, but with weighted edges. The weights are integers ≥ 1 , and a weight a may be interpreted as a parallel edges. The other class are 3-Motzkin paths. They are like Motzkin paths (Dyck paths plus horizontal steps); however, the horizontal steps come in three different colours. A bijection is described. Since 3-Motzkin paths and multi-edge trees are very much alike (using a variation of the classical rotation correspondence), all the structures that are discussed in this paper can be linked via bijections.

4. BINARY TREES AND HORTON-STRAHLER NUMBERS

This section is classical and serves as the basis of some new developments about weighted unary-binary trees. A full account can be found here: [32].

Binary trees may be expressed by the following symbolic equation, which says that they include the empty tree and trees recursively built from a root followed by two subtrees, which are binary trees:

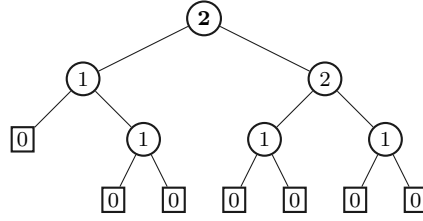
$$\mathcal{B} = \square + \begin{array}{c} \circ \\ \diagdown \quad \diagup \\ \mathcal{B} \quad \mathcal{B} \end{array}$$

Binary trees are counted by Catalan numbers and there is an important parameter reg , which in Computer Science is called the register function. It associates to each binary tree (which is used to code an arithmetic expression, with data in the leaves and operators in the internal nodes) the minimal number of extra registers that is needed to evaluate the tree. The optimal strategy is to evaluate difficult subtrees first, and use one register to keep its value, which does not hurt, if the other subtree requires less registers. If both subtrees are equally difficult, then one more register is used, compared to the requirements of the subtrees. This natural parameter is among Combinatorialists known as the Horton-Strahler numbers, and we will adopt this name throughout this paper.

There is a recursive description of this function: $\text{reg}(\square) = 0$, and if tree t has subtrees t_1 and t_2 , then

$$\text{reg}(t) = \begin{cases} \max\{\text{reg}(t_1), \text{reg}(t_2)\} & \text{if } \text{reg}(t_1) \neq \text{reg}(t_2), \\ 1 + \text{reg}(t_1) & \text{otherwise.} \end{cases}$$

The recursive description attaches numbers to the nodes, starting with 0's at the leaves and then going up; the number appearing at the root is the Horton-Strahler number of the tree.



Let \mathcal{R}_p denote the family of trees with Horton-Strahler number = p , then one gets immediately from the recursive definition:

$$\mathcal{R}_p = \begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \mathcal{R}_{p-1} \quad \mathcal{R}_{p-1} \end{array} + \begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \mathcal{R}_p \quad \sum_{j < p} \mathcal{R}_j \end{array} + \begin{array}{c} \circ \\ \swarrow \quad \searrow \\ \sum_{j < p} \mathcal{R}_j \quad \mathcal{R}_p \end{array}$$

In terms of generating functions, these equations read as

$$R_p(z) = zR_{p-1}^2(z) + 2zR_p(z) \sum_{j < p} R_j(z);$$

the variable z is used to mark the size (i. e., the number of internal nodes) of the binary tree.

A historic account of these concepts, from the angle of Philippe Flajolet, who was one of the pioneers is [37], compare also [36].

Amazingly, the recursion for the generating functions $R_p(z)$ can be solved explicitly! The substitution

$$z = \frac{u}{(1+u)^2}$$

that de Bruijn, Knuth, and Rice [8] also used, produces the nice expression

$$R_p(z) = \frac{1-u^2}{u} \frac{u^{2^p}}{1-u^{2^{p+1}}}.$$

Of course, once this is *known*, it can be proved by induction, using the recursive formula. For the readers benefit, this will be sketched now.

We start with the auxiliary formula

$$\sum_{0 \leq j < p} \frac{u^{2^j}}{1-u^{2^{j+1}}} = \frac{u}{1-u} - \frac{u^{2^p}}{1-u^{2^p}},$$

which we will prove now by induction: For $p = 0$, the formula $0 = \frac{u}{1-u} - \frac{u}{1-u}$ is correct, and then

$$\begin{aligned} \sum_{0 \leq j < p+1} \frac{u^{2^j}}{1-u^{2^{j+1}}} &= \frac{u}{1-u} - \frac{u^{2^p}}{1-u^{2^p}} + \frac{u^{2^p}}{1-u^{2^{p+1}}} \\ &= \frac{u}{1-u} - \frac{u^{2^p}(1+u^{2^p})}{1-u^{2^{p+1}}} + \frac{u^{2^p}}{1-u^{2^{p+1}}} = \frac{u}{1-u} - \frac{u^{2^{p+1}}}{1-u^{2^{p+1}}}. \end{aligned}$$

Now the formula for $R_p(z)$ can also be proved by induction. First, $R_0(z) = \frac{1-u^2}{u} \frac{u}{1-u^2} = 1$, as it should, and

$$zR_{p-1}^2(z) + 2zR_p(z) \sum_{j < p} R_j(z)$$

$$\begin{aligned}
&= \frac{u}{(1+u)^2} \frac{(1-u^2)^2}{u^2} \frac{u^{2p}}{(1-u^{2p})^2} + \frac{2u}{(1+u)^2} R_p(z) \sum_{j < p} \frac{1-u^2}{u} \frac{u^{2j}}{1-u^{2^{j+1}}} \\
&= \frac{u^{2p-1}(1-u)^2}{(1-u^{2p})^2} + \frac{2(1-u)}{(1+u)} R_p(z) \sum_{j < p} \frac{u^{2j}}{1-u^{2^{j+1}}}.
\end{aligned}$$

Solving

$$R_p(z) = \frac{u^{2p-1}(1-u)^2}{(1-u^{2p})^2} + \frac{2(1-u)}{(1+u)} R_p(z) \left[\frac{u}{1-u} - \frac{u^{2p}}{1-u^{2p}} \right]$$

leads to

$$R_p(z) \frac{1-u}{1+u} \left[1 + 2 \frac{u^{2p}}{1-u^{2p}} \right] = \frac{u^{2p-1}(1-u)^2}{(1-u^{2p})^2},$$

or, simplified

$$R_p(z) = \frac{u^{2p-1}(1-u^2)}{(1-u^{2p})(1+u^{2p})} = \frac{1-u^2}{u} \frac{u^{2p}}{(1-u^{2^{p+1}})},$$

which is the formula that we needed to prove. \square

Weighted unary-binary trees and Horton-Strahler numbers. The family of unary-binary trees \mathcal{M} might be defined by the symbolic equation

$$\mathcal{M} = \square + \begin{array}{c} \circ \\ | \\ \mathcal{M} \setminus \{\square\} \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{M} \quad \mathcal{M} \end{array}$$

The equation for the generating function is

$$M = 1 + z(M-1) + zM^2$$

with the solution

$$M = M(z) = \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2z} = 1 + z + 3z^2 + 10z^3 + 36z^4 + \dots;$$

the coefficients form again sequence A002212 in [47] and enumerate Schröder paths, among many other things. We will come to equivalent structures a bit later.

In the instance of unary-binary trees, we can also work with a substitution: Set $z = \frac{u}{1+3u+u^2}$, then $M(z) = 1+u$. Unary-binary trees and the register function were investigated in [17], but the present favourable substitution was not used. Therefore, in this previous paper, asymptotic results were available but no explicit formulae.

This works also with a weighted version, where we allow unary edges with a different colours. Then

$$\mathcal{N} = \square + a \cdot \begin{array}{c} \circ \\ | \\ \mathcal{N} \setminus \{\square\} \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{N} \quad \mathcal{N} \end{array}$$

and with the substitution $z = \frac{u}{1+(a+2)u+u^2}$, the generating function is beautifully expressed as $N(z) = 1+u$. For $a=0$, this covers also binary trees.

We will consider the Horton-Strahler numbers of unary-binary trees in the sequel. The definition is naturally extended by

$$\text{reg}\left(\begin{array}{c} \circ \\ | \\ t \end{array}\right) = \text{reg}(t).$$

Now we can move again to $R_p(z)$, the generating function of (generalized) unary-binary trees with Horton-Strahler number $= p$. The recursion (for $p \geq 1$) is

$$\mathcal{R}_p = \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{R}_{p-1} \quad \mathcal{R}_{p-1} \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{R}_p \quad \sum_{j < p} \mathcal{R}_j \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \sum_{j < p} \mathcal{R}_j \quad \mathcal{R}_p \end{array} + a \cdot \begin{array}{c} \circ \\ | \\ \mathcal{R}_p \end{array}$$

In terms of generating functions, these equations read as

$$R_p(z) = zR_{p-1}^2(z) + 2zR_p(z) \sum_{j < p} R_j(z) + azR_p(z), \quad p \geq 1; \quad R_0(z) = 1.$$

Amazingly, with the substitution $z = \frac{u}{1+(a+2)u+u^2}$, formally we get the *same* solution as in the binary case:

$$R_p(z) = \frac{1-u^2}{u} \frac{u^{2p}}{1-u^{2p+1}}.$$

The proof by induction is as before. One sees another advantage of the substitution: On a formal level, many manipulations do not need to be repeated. Only when one switches back to the z -world, things become different.

Now we move to Hex-trees.

$$\mathcal{H} = \square + \circ + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{H} \setminus \{\square\} \quad \mathcal{H} \setminus \{\square\} \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{H} \setminus \{\square\} \quad \square \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \square \quad \mathcal{H} \setminus \{\square\} \end{array} + \begin{array}{c} \circ \\ / \quad \backslash \\ \square \quad \square \end{array}$$

Hex trees. Hex trees either have two non-empty successors, or one of 3 types of unary successors (called left, middle, right). The author has seen this family first in [23], but one can find older literature following the references and the usual search engines.

The generating function satisfies

$$\begin{aligned} H(z) &= 1 + z(H(z) - 1)^2 + z + 3z(H(z) - 1) = \frac{1 - z - \sqrt{(1-z)(1-5z)}}{2z} \\ &= 1 + z + 3z^2 + 10z^3 + 36z^4 + 137z^5 + 543z^6 + 2219z^7 + 9285z^8 + 39587z^9 + \dots \end{aligned}$$

The same generating function also appears in [21], and it is again sequence A002212 in [47]. One can rewrite the symbolic equation as

$$\mathcal{H} = \square + \begin{array}{c} \circ \\ / \quad \backslash \\ \mathcal{H} \quad \mathcal{H} \end{array} + \begin{array}{c} \circ \\ | \\ \mathcal{H} \setminus \{\square\} \end{array}$$

and sees in this way that the Hex-trees are just unary-binary trees (with parameter $a = 1$).

Continuing with enumerations. First, we will enumerate the number of (generalized) unary-binary trees with n (internal) nodes. For that we need the notion of generalized trinomial coefficients, viz.

$$\binom{n; 1, a, 1}{k} := [z^k](1 + az + z^2)^n.$$

Of course, for $a = 2$, this simplifies to a binomial coefficient $\binom{2n}{k}$. We will use contour integration to pull out coefficients, and the contour of integer, in whatever variable, is a small circle (or equivalent) around the origin. The desired number is

$$\begin{aligned}
[z^n](1+u) &= \frac{1}{2\pi i} \oint \frac{dz}{z^{n+1}}(1+u) \\
&= \frac{1}{2\pi i} \oint \frac{du(1-u^2)(1+(a+2)u+u^2)^{n+1}}{(1+(a+2)u+u^2)^2 u^{n+1}}(1+u) \\
&= [u^{n+1}](1-u)(1+u)^2(1+(a+2)u+u^2)^{n-1} \\
&= \binom{n-1; 1, a+2, 1}{n+1} + \binom{n-1; 1, a+2, 1}{n} \\
&\quad - \binom{n-1; 1, a+2, 1}{n-1} - \binom{n-1; 1, a+2, 1}{n-2}.
\end{aligned}$$

Then we introduce $S_p(z) = R_p(z) + R_{p+1}(z) + R_{p+2}(z) + \dots$, the generating function of trees with Horton-Strahler number $\geq p$. Using the summation formula proved earlier, we get

$$S_p(z) = \frac{1-u^2}{u} \frac{u^{2p}}{1-u^{2p}} = \frac{1-u^2}{u} \sum_{k \geq 1} u^{k2^p}.$$

Further,

$$[z^n]S_p(z) = \sum_{k \geq 1} \frac{1}{2\pi i} \oint \frac{dz}{z^{n+1}} \frac{1-u^2}{u} u^{k2^p}.$$

Asymptotics. We start by deriving asymptotics for the number of (generalized) unary-binary trees with n (internal) nodes. This is a standard application of singularity analysis of generating functions, as described in [14] and [15].

We start from the generating function

$$N(z) = \frac{1 - az - \sqrt{1 - 2(a+2)z + a(a+4)z^2}}{2z}$$

and determine the singularity closest to the origin, which is the value making the square root disappear: $z = \frac{1}{a+4}$. After that, the local expansion of $N(z)$ around this singularity is determined:

$$N(z) \sim 2 - \sqrt{a+4} \sqrt{1 - (a+4)z}.$$

The translation lemmas given in [14] and [15] provide the asymptotics:

$$\begin{aligned}
[z^n]N(z) &\sim [z^n] \left(2 - \sqrt{a+4} \sqrt{1 - (a+4)z} \right) \\
&= -\sqrt{a+4} (a+4)^n \frac{n^{-3/2}}{\Gamma(-\frac{1}{2})} = (a+4)^{n+1/2} \frac{1}{2\sqrt{\pi} n^{3/2}}.
\end{aligned}$$

Just note that $a = 0$ is the well-known formula for binary trees with n nodes.

Now we move to the generating function for the average number of registers. Apart from normalization it is

$$\sum_{p \geq 1} p R_p(z) = \sum_{p \geq 1} S_p(z) = \frac{1-u^2}{u} \sum_{p \geq 1} \sum_{k \geq 1} u^{k2^p}$$

$$= \frac{1-u^2}{u} \sum_{n \geq 1} v_2(n) u^n,$$

where $v_2(n)$ is the highest exponent k such 2^k divides n .

This has to be studied around $u = 1$, which, upon setting $u = e^{-t}$, means around $t = 0$. Eventually, and that is the only thing that is different here, this is to be retranslated into a singular expansion of z around its singularity, which depends on the parameter a .

For the reader's convenience, we also repeat the steps that were known before. The first factor is elementary:

$$\frac{1-u^2}{u} \sim 2t + \frac{1}{3}t^3 + \dots$$

For

$$\sum_{p \geq 1} \sum_{k \geq 1} e^{-k2^p t},$$

one applies the Mellin transform, with the result

$$\frac{\Gamma(s)\zeta(s)}{2^s - 1}.$$

Applying the inversion formula, one finds

$$\sum_{p \geq 1} \sum_{k \geq 1} e^{-k2^p t} = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} t^{-s} \frac{\Gamma(s)\zeta(s)}{2^s - 1} ds.$$

Shifting the line of integration to the left, the residues at the poles $s = 1$, $s = 0$, $s = \chi_k = \frac{2k\pi i}{\log 2}$, $k \neq 0$ provide enough terms for our asymptotic expansion.

$$\frac{1}{t} + \frac{\gamma}{2 \log 2} - \frac{1}{4} - \frac{\log \pi}{2 \log 2} + \frac{\log t}{2 \log 2} + \frac{1}{\log 2} \sum_{k \neq 0} \Gamma(\chi_k) \zeta(\chi_k) t^{-\chi_k}.$$

Combined with the elementary factor, this leads to

$$2 + \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log t}{\log 2} \right) t + \frac{2t}{\log 2} \sum_{k \neq 0} \Gamma(\chi_k) \zeta(\chi_k) t^{-\chi_k} + O(t^2 \log t).$$

Now we want to translate into the original z -world. Since $z = \frac{u}{1+(a+2)u+u^2}$, $u = 1$ translates into the singularity $z = \frac{1}{4+a}$. Further,

$$t \sim \sqrt{4+a} \cdot \sqrt{1-z(4+a)},$$

let us abbreviate $A = 4+a$, then for singularity analysis we must consider

$$\begin{aligned} & \frac{\sqrt{A} \cdot \sqrt{1-zA} \log(1-zA)}{2 \log 2} \\ & + \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2 \log 2} \right) \sqrt{A} \cdot \sqrt{1-zA} \\ & + \frac{2}{\log 2} \sum_{k \neq 0} \Gamma(\chi_k) \zeta(\chi_k) A^{\frac{1-\chi_k}{2}} (1-zA)^{\frac{1-\chi_k}{2}}. \end{aligned}$$

The formula that is perhaps less known and needed here is [14]

$$[z^n] \log(1-z) \sqrt{1-z} \sim \frac{n^{-3/2} \log n}{2\sqrt{\pi}} + \frac{n^{-3/2}}{2\sqrt{\pi}} (-2 + \gamma + 2 \log 2);$$

furthermore we need

$$[z^n](1-z)^\alpha \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}.$$

We start with the most complicated term:

$$\begin{aligned} \frac{[z^n] \frac{\sqrt{A} \sqrt{1-zA} \log(1-zA)}{2 \log 2}}{[z^n] N(z)} &\sim \frac{\sqrt{A}}{2 \log 2} \frac{A^n \left(\frac{n^{-3/2} \log n}{2\sqrt{\pi}} + \frac{n^{-3/2}}{2\sqrt{\pi}} (-2 + \gamma + 2 \log 2) \right)}{A^{n+1/2} \frac{1}{2\sqrt{\pi n^{3/2}}}} \\ &= \log_4 n + 1 + \frac{\gamma}{2 \log 2} - \frac{1}{\log 2}. \end{aligned}$$

The next term we consider is

$$\begin{aligned} \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2 \log 2} \right) \sqrt{A} \frac{[z^n] \sqrt{1-zA}}{[z^n] N(z)} \\ \sim \left(\frac{\gamma}{\log 2} - \frac{1}{2} - \frac{\log \pi}{\log 2} + \frac{\log A}{2 \log 2} \right) \sqrt{A} \frac{[z^n] \sqrt{1-zA}}{-\sqrt{A} [z^n] \sqrt{1-zA}} \\ = -\frac{\gamma}{\log 2} + \frac{1}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2 \log 2}. \end{aligned}$$

The last term we consider is

$$\begin{aligned} \frac{2}{\log 2} \Gamma(\chi_k) \zeta(\chi_k) A^{\frac{1-\chi_k}{2}} \frac{[z^n] (1-zA)^{\frac{1-\chi_k}{2}}}{-\sqrt{A} [z^n] \sqrt{1-zA}} \\ \sim -\frac{4\sqrt{\pi}}{\log 2} \frac{\Gamma(\chi_k) \zeta(\chi_k)}{\Gamma\left(\frac{\chi_k-1}{2}\right)} A^{\frac{1-\chi_k}{2}} n^{\chi_k/2}. \end{aligned}$$

Eventually we have evaluated the average value of the Horton-Strahler numbers:

Theorem 1. *The average Horton-Strahler of weighted unary-binary trees with n nodes is given by the asymptotic formula*

$$\begin{aligned} \log_4 n - \frac{\gamma}{2 \log 2} - \frac{1}{\log 2} + \frac{3}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2 \log 2} - \frac{4\sqrt{\pi A}}{\log 2} \sum_{k \neq 0} \frac{\Gamma(\chi_k) \zeta(\chi_k)}{\Gamma\left(\frac{\chi_k-1}{2}\right)} A^{\frac{-\chi_k}{2}} n^{\chi_k/2} \\ = \log_4 n - \frac{\gamma}{2 \log 2} - \frac{1}{\log 2} + \frac{3}{2} + \frac{\log \pi}{\log 2} - \frac{\log A}{2 \log 2} + \psi(\log_4 n), \end{aligned}$$

with a tiny periodic function $\psi(x)$ of period 1.

5. MARKED ORDERED TREES

In [9] we find the following variation of ordered trees: Each rightmost edge might be marked or not, if it does not lead to an endnode (leaf). We depict a marked edge by the red colour and draw all of them of size 4 (4 nodes):

Now we move to a symbolic equation for the marked ordered trees:

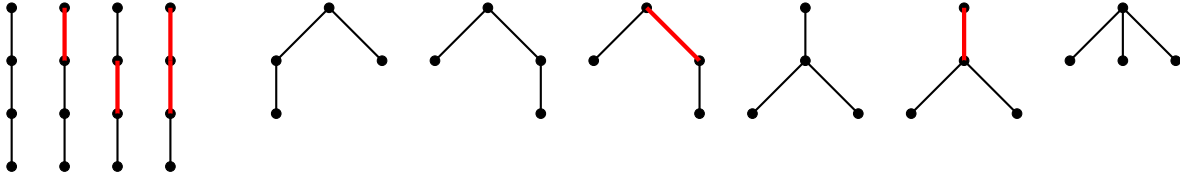


FIGURE 2. All 10 marked ordered trees with 4 nodes.

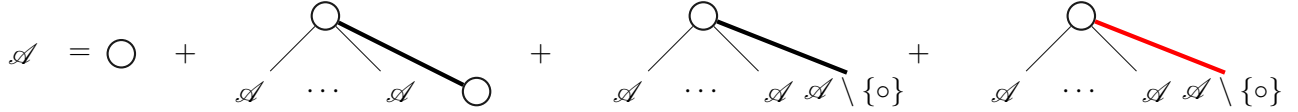


FIGURE 3. Symbolic equation for marked ordered trees.
 $\mathcal{A} \cdots \mathcal{A}$ refers to ≥ 0 copies of \mathcal{A} .

In terms of generating functions,

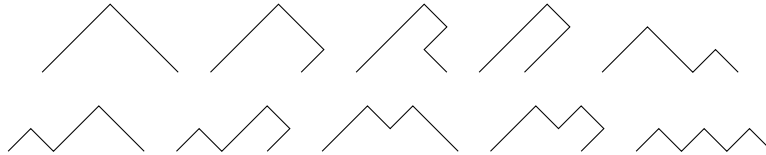
$$A = z + \frac{z}{1 - A}z + \frac{z}{1 - A}2(A - z),$$

with the solution

$$A(z) = \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2} = z + z^2 + z^3 + 3z^3 + 10z^4 + 36z^5 + \dots$$

The importance of this family of trees lies in the bijection to skew Dyck paths, as given in [9]. One walks around the tree as one usually does and translates it into a Dyck path. The only difference are the red edges. On the way down, nothing special is to be reported, but on the way up, it is translated into a skew step $(-1, -1)$. The present author believes that trees are more manageable when it comes to enumeration issues than skew Dyck paths.

The 10 trees of Figure 1 translate as follows:



Skew Dyck paths and a dual version (reading the paths from right-to-left) will be discussed in a later section.

Parameters of marked ordered trees. There are many parameters, usually considered in the context of ordered trees, that can be considered for marked ordered trees. Of course, we cannot be encyclopedic about such parameters. We just consider a few parameters and leave further analysis to the future.

The number of leaves. To get this, it is most natural to use an additional variable u when translating the symbolic equation, so that $z^n u^k$ refers to trees with n nodes and k leaves. One obtains

$$F = zu + \frac{z}{1 - F}(zu2(F - zu)),$$

with the solution

$$\begin{aligned} F(z, u) &= -z + \frac{zu}{2} + \frac{1}{2} - \frac{1}{2} \sqrt{4z^2 - 4z + z^2u^2 - 2zu + 1} \\ &= zu + z^2u + (2u + u^2)z^3 + (4u + 5u^2 + u^3)z^4 + \dots \end{aligned}$$

The factor $4u + 5u^2 + u^3$ corresponds to the 10 trees in Figure 1.

Of interest is also the average number of leaves, when all marked ordered trees of size n are considered to be equally likely. For that, we differentiate $F(z, u)$ w. r. t. u , followed by $u := 1$, with the result

$$\frac{z}{2} + \frac{z - z^2}{2\sqrt{1 - 6z + 5z^2}} = \frac{z}{1 - v}, \quad \text{with the usual } z = \frac{v}{1 + 3v + v^2}.$$

Since $F(z, 1) = z(1 + v)$, it follows that the average is asymptotic to

$$\begin{aligned} \frac{[z^{n+1}] \frac{z}{1-v}}{[z^{n+1}] z(1+v)} &= \frac{[z^n] \frac{1}{1-v}}{[z^n](1+v)} = \frac{[z^n] \frac{1}{\sqrt{5}} \frac{1}{\sqrt{1-5z}}}{5^{n+\frac{1}{2}} \frac{1}{2\sqrt{\pi}} n^{3/2}} \\ &= \frac{\frac{n^{-1/2}}{\Gamma(\frac{1}{2})}}{5^{n+\frac{1}{2}} \frac{1}{2\sqrt{\pi}} n^{3/2}} = \frac{n}{10}. \end{aligned}$$

Note that the corresponding number for ordered trees (unmarked) is $\frac{n}{2}$, so we have significantly less leaves here.

The height. As in the seminal paper [8], we define the height in terms of the longest chain of nodes from the root to a leaf. Further, let $p_h = p_h(z)$ be the generating function of marked ordered trees of height $\leq h$. From the symbolic equation,

$$p_{h+1} = z + \frac{z^2}{1 - p_h} + \frac{2z^2(p_h - z)}{1 - p_h} = -z + \frac{2z - z^2}{1 - p_h}, \quad h \geq 1, \quad p_1 = z.$$

By some creative guessing, separating numerator and denominator, we find the solution

$$p_h = z(1 + v) \frac{(1 + 2v)^{h-1} - v^h(v + 2)^{h-1}}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}},$$

which can be prove now by induction.

We have $p_1 = z(1 + v) \frac{1-v}{1-v^2} = z$. Furthermore, the induction step is best checked using a computer.

The limit of p_h for $h \rightarrow \infty$ is $z(1 + v)$, the generating function of *all* marked ordered trees, as expected. Taking differences, we get the generating functions of trees of height $> h$:

$$\begin{aligned} & z(1 + v) - z(1 + v) \frac{(1 + 2v)^{h-1} - v^h(v + 2)^{h-1}}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}} \\ &= z(1 + v) \frac{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1} - (1 + 2v)^{h-1} + v^h(v + 2)^{h-1}}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}} \\ &= z(1 - v^2) \frac{(v + 2)^{h-1} v^h}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}}. \end{aligned}$$

From this, the average height can be worked out, as in the model paper [21]. We sketch the essential steps. For the average height, one needs

$$\sum_{h \geq 0} z(1 - v^2) \frac{(v + 2)^{h-1} v^h}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}}$$

and its behaviour around $v = 1$, viz.

$$2z(1 - v) \sum_{h \geq 0} \frac{3^{h-1} v^h}{3^{h-1} - v^{h+1} 3^{h-1}} \sim 2z(1 - v) \sum_{h \geq 1} \frac{v^h}{1 - v^h}.$$

The behaviour of the series can be taken straight from [21].

We find there

$$\sum_{h \geq 1} \frac{v^h}{1 - v^h} = -\frac{\log(1 - v)}{1 - v},$$

and further

$$\sum_{h \geq 0} z(1 - v^2) \frac{(v + 2)^{h-1} v^h}{(1 + 2v)^{h-1} - v^{h+1}(v + 2)^{h-1}} \sim -2z \log(1 - v),$$

so that the coefficient of z^{n+1} is asymptotic to $-2[z^n] \log(1 - v)$. Since $1 - v \sim \sqrt{5}\sqrt{1 - 5z}$,

$$-2z \log(1 - v) \sim -2z \log \sqrt{1 - 5z} = -z \log(1 - 5z),$$

and the coefficient of z^{n+1} in it is asymptotic to $\frac{5^n}{n}$. This has to be divided (as derived earlier) by

$$5^{n+\frac{1}{2}} \frac{1}{2\sqrt{\pi n}^{3/2}},$$

with the result

$$2 \frac{5^n}{n} \frac{1}{5^{n+\frac{1}{2}}} \sqrt{\pi n}^{3/2} = \frac{2}{\sqrt{5}} \sqrt{\pi n}.$$

Note that the constant in front of $\sqrt{\pi n}$ for ordered trees is $\frac{2}{\sqrt{4}} = 1$, so the average height for marked ordered trees is indeed a bit smaller thanks to the extra markings.

6. A BIJECTION BETWEEN MULTI-EDGE TREES AND 3-COLOURED MOTZKIN PATHS

Multi-edge trees are like ordered (planar, plane, ...) trees, but instead of edges there are multiple edges. When drawing such a tree, instead of drawing, say 5 parallel edges, we just draw one edge and put the number 5 on it as a label. These trees were studied in [11, 21]. For the enumeration, one must count edges. The generating function $F(z)$ satisfies

$$F(z) = \sum_{k \geq 0} \left(\frac{z}{1 - z} F(z) \right)^k = \frac{1}{1 - \frac{z}{1 - z} F(z)},$$

whence

$$F(z) = \frac{1 - z - \sqrt{1 - 6z + 5z^2}}{2z} = 1 + z + 3z^2 + 10z^3 + 36z^4 + 137z^5 + 543z^6 + \dots$$

The coefficients form once again sequence A002212 in [47].

A Motzkin path consists of up-steps, down-steps, and horizontal steps, see sequence A091965 in [47] and the references given there. As Dyck paths, they start at the origin and end, after n steps again at the x -axis, but are not allowed to go below the x -axis.

A 3-coloured Motzkin path is built as a Motzkin path, but there are 3 different types of horizontal steps, which we call *red*, *green*, *blue*. The generating function $M(z)$ satisfies

$$M(z) = 1 + 3zM(z) + z^2M(z)^2 = \frac{1 - 3z - \sqrt{1 - 6z + 5z^2}}{2z^2}, \quad \text{or} \quad F(z) = 1 + zM(z).$$

So multi-edge trees with N edges (counting the multiplicities) correspond to 3-coloured Motzkin paths of length $N - 1$.

The purpose of this note is to describe a bijection. It transforms trees into paths, but all steps are reversible.

The details. As a first step, the multiplicities will be ignored, and the tree then has only n edges. The standard translation of such tree into the world of Dyck paths, which is in every book on combinatorics, leads to a Dyck path of length $2n$. Then the Dyck path will be transformed bijectively to a 2-coloured Motzkin path of length $n - 1$ (the colours used are red and green). This transformation plays a prominent role in [10], but is most likely much older. I believe that people like Viennot know this for 40 years. I would be glad to get a proper historic account from the gentle readers.

The last step is then to use the third colour (blue) to deal with the multiplicities.

The first up-step and the last down-step of the Dyck path will be deleted. Then, the remaining $2n - 2$ steps are coded pairwise into a 2-Motzkin path of length $n - 1$:



The last step is to deal with the multiplicities. If an edge is labelled with the number a , we will insert $a - 1$ horizontal blue steps in the following way: Since there are currently $n - 1$ symbols in the path, we have n possible positions to enter something (in the beginning, in the end, between symbols). We go through the tree in pre-order, and enter the multiplicities one by one using the blue horizontal steps.

To make this procedure more clear, we prepared a list of 10 multi-edge trees with 3 edges, and the corresponding 3-Motzkin paths of length 2, with intermediate steps completely worked out:

Connecting unary-binary trees with multi-edge trees. This is not too difficult: We start from multi-edge trees, and ignore the multiplicities at the moment. Then we apply the classical rotation correspondence (also called: natural correspondence). Then we add vertical edges, if the multiplicity is higher than 1. To be precise, if there is a node, and an edge with multiplicity a leads to it from the top, we insert $a - 1$ extra nodes in a chain on the top, and connect them with unary branches. The following example with 10 objects will help to understand this procedure. After that, all the structures studied in this paper are connected with bijections.

7. THE COMBINATORICS OF SKEW DYCK PATHS

Skew Dyck are a variation of Dyck paths, where additionally to steps $(1, 1)$ and $(1, -1)$ a south-west step $(-1, -1)$ is also allowed, provided that the path does not intersect itself. Otherwise, like for Dyck path, it must never go below the x -axis and

Multi-edge tree	Dyck path	2-Motzkin path	blue edges added

TABLE 1. First row is a multi-edge tree with 3 edges, second row is the standard Dyck path (multiplicities ignored), third row is cutting off first and last step, and then translated pairs of steps, fourth row is inserting blue horizontal edges, according to multiplicities.

end eventually (after $2n$ steps) on the x -axis. Here are a few references: [9, 23, 2, 33]. The enumerating sequence is

$$1, 1, 3, 10, 36, 137, 543, 2219, 9285, 39587, 171369, 751236, 3328218, 14878455, \dots,$$

which is A002212 in [47].

Skew Dyck appeared very briefly in a previous section; here we want to give a more thorough analysis of them, using generating functions and the kernel method. Here is a list of the 10 skew paths consisting of 6 steps:

We prefer to work with the equivalent model (resembling more traditional Dyck paths) where we replace each step $(-1, -1)$ by $(1, -1)$ but label it red. Here is the list of the 10 paths again (Figure 2):

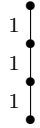











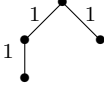








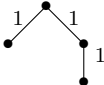



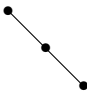




Multi-edge tree	Binary tree (rotation)	vertical edges added
		
		
		
		
		
		
		
		
		
		

TABLE 2. First row is a multi-edge tree with 3 edges, second row the corresponding binary tree, according to the classical rotation correspondence, ignoring the unary branches. Third row is inserting extra horizontal edges when the multiplicities are higher than 1.

The rules to generate such decorated Dyck paths are: each edge $(1, -1)$ may be black or red, but \nearrow and \searrow are forbidden.

Our interest is in particular in *partial* decorated Dyck paths, ending at level j , for fixed $j \geq 0$; the instance $j = 0$ is the classical case.

The analysis of partial skew Dyck paths was recently started in [2] (using the notion ‘prefix of a skew Dyck path’) using Riordan arrays instead of our kernel method. The latter gives us *bivariate* generating functions, from which it is easier to draw

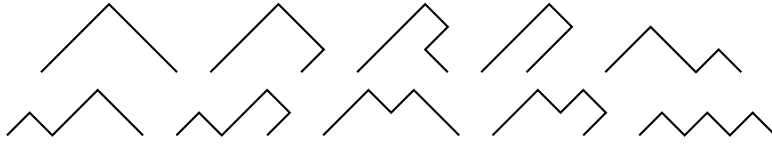


FIGURE 4. All 10 skew Dyck paths of length 6 (consisting of 6 steps).

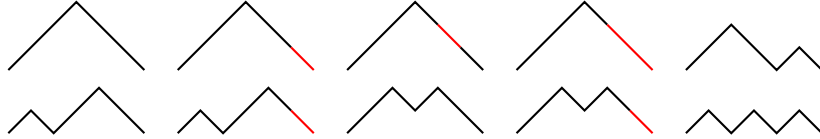


FIGURE 5. The 10 paths redrawn, with red south-east edges instead of south-west edges.

conclusions. Two variables, z and u , are used, where z marks the length of the path and j marks the end-level. We briefly mention that one can, using a third variable w , also count the number of red edges.

Again, once all generating functions are explicitly known, many corollaries can be derived in a standard fashion. We only do this in a few instances. But we would like to emphasize that the substitution

$$x = \frac{v}{1 + 3v + v^2},$$

which was used in [21, 33] allows to write *explicit enumerations*, using the notion of a (weighted) trinomial coefficient:

$$\binom{n; 1, 3, 1}{k} := [t^k](1 + 3t + t^2)^n.$$

The second part of this section deals with a dual version, where the paths are read from right to left.

Generating functions and the kernel method. We catch the essence of a decorated Dyck path using a state-diagram:

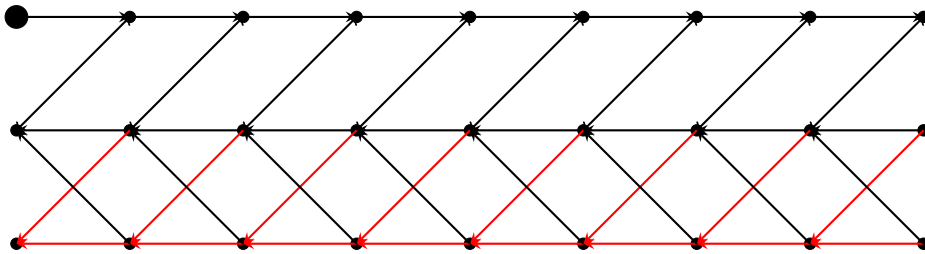


FIGURE 6. Three layers of states according to the type of steps leading to them (up, down-black, down-red).

It has three types of states, with j ranging from 0 to infinity; in the drawing, only $j = 0..8$ is shown. The first layer of states refers to an up-step leading to a state, the second layer refers to a black down-step leading to a state and the third layer

refers to a red down-step leading to a state. We will work out generating functions describing all paths leading to a particular state. We will use the notations f_j, g_j, h_j for the three respective layers, from top to bottom. Note that the syntactic rules of forbidden patterns \nearrow and \searrow can be clearly seen from the picture. The functions depend on the variable z (marking the number of steps), but mostly we just write f_j instead of $f_j(z)$, etc.

The following recursions can be read off immediately from the diagram:

$$\begin{aligned} f_0 &= 1, & f_{i+1} &= z f_i + z g_i, & i &\geq 0, \\ g_i &= z f_{i+1} + z g_{i+1} + z h_{i+1}, & i &\geq 0, \\ h_i &= z h_{i+1} + z g_{i+1}, & i &\geq 0. \end{aligned}$$

And now it is time to introduce the promised *bivariate* generating functions:

$$F(z, u) = \sum_{i \geq 0} f_i(z) u^i, \quad G(z, u) = \sum_{i \geq 0} g_i(z) u^i, \quad H(z, u) = \sum_{i \geq 0} h_i(z) u^i.$$

Again, often we just write $F(u)$ instead of $F(z, u)$ and treat z as a ‘silent’ variable. Summing the recursions leads to

$$\begin{aligned} \sum_{i \geq 0} u^i f_{i+1} &= \sum_{i \geq 0} u^i z f_i + \sum_{i \geq 0} u^i z g_i, \\ \sum_{i \geq 0} u^i g_i &= \sum_{i \geq 0} u^i z f_{i+1} + \sum_{i \geq 0} u^i z g_{i+1} + \sum_{i \geq 0} u^i z h_{i+1}, \\ \sum_{i \geq 0} u^i h_i &= \sum_{i \geq 0} u^i z h_{i+1} + \sum_{i \geq 0} u^i z g_{i+1}. \end{aligned}$$

This can be rewritten as

$$\begin{aligned} \frac{1}{u}(F(u) - 1) &= zF(u) + zG(u), \\ G(u) &= \frac{z}{u}(F(u) - 1) + \frac{z}{u}(G(u) - G(0)) + \frac{z}{u}(H(u) - H(0)), \\ H(u) &= \frac{z}{u}(G(u) - G(0)) + \frac{z}{u}(H(u) - G(0)). \end{aligned}$$

This is a typical application of the kernel method. For a gentle example-driven introduction to the kernel method, see [35]. First,

$$\begin{aligned} F(u) &= \frac{z^2 u G(0) + z^2 u H(0) + z^2 u - u - z^3 + 2z}{-z^3 - u + 2z + z u^2 - z^2 u}, \\ G(u) &= \frac{z(H(0) - u z H(0) + z^2 + G(0) - z u G(0) - z u)}{-z^3 - u + 2z + z u^2 - z^2 u}, \\ H(u) &= \frac{z(-u z H(0) - z^2 - z u G(0) + G(0) - z^2 H(0) + H(0) - z^2 G(0))}{-z^3 - u + 2z + z u^2 - z^2 u}. \end{aligned}$$

The denominator factors as $z(u - r_1)(u - r_2)$, with

$$r_1 = \frac{1 + z^2 + \sqrt{1 - 6z^2 + 5z^4}}{2z}, \quad r_2 = \frac{1 + z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2z}.$$

Note that $r_1 r_2 = 2 - z^2$. Since the factor $u - r_2$ in the denominator is “bad,” it must also cancel in the numerators. From this we conclude as a first step

$$G(0) = \frac{1 - 2z^2 H(0) - 3z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2z^2},$$

and by further simplification

$$H(0) = \frac{1 - 4z^2 + z^4 + (z^2 - 1)\sqrt{1 - 6z^2 + 5z^4}}{2 - z^2}.$$

Thus (with $W = \sqrt{1 - 6z^2 + 5z^4} = \sqrt{(1 - z^2)(1 - 5z^2)}$)

$$\begin{aligned} F(u) &= \frac{-1 - z^2 - W}{2z(u - r_1)} = \frac{1 + z^2 + W}{2zr_1(1 - u/r_1)}, \\ G(u) &= \frac{-1 + z^2 + W}{2z(u - r_1)} = \frac{1 - z^2 - W}{2zr_1(1 - u/r_1)}, \\ H(u) &= \frac{-1 + 3z^2 + W}{2z(u - r_1)} = \frac{1 - 3z^2 - W}{2zr_1(1 - u/r_1)}. \end{aligned}$$

The total generating function is

$$S(u) = F(u) + G(u) + H(u) = \frac{3 - 3z^2 - W}{2zr_1(1 - u/r_1)}.$$

The coefficient of $u^j z^n$ in $S(u)$ counts the partial paths of length n , ending at level j . We will write $s_j = [u^j]S(u)$. Furthermore

$$\begin{aligned} f_j &= [u^j]F(u) = [u^j] \frac{1 + z^2 + W}{2zr_1(1 - u/r_1)}, \\ g_j &= [u^j]G(u) = [u^j] \frac{1 - z^2 - W}{2zr_1(1 - u/r_1)}, \\ h_j &= [u^j]H(u) = [u^j] \frac{1 - 3z^2 - W}{2zr_1(1 - u/r_1)}. \end{aligned}$$

At this stage, we are only interested in

$$s_j = f_j + g_j + h_j = [u^j] \frac{3 - 3z^2 - W}{2zr_1(1 - u/r_1)} = \frac{3 - 3z^2 - W}{2zr_1^{j+1}},$$

which is the generating function of all (partial) paths ending at level j . Parity considerations give us that only coefficients $[z^n]s_j$ are non-zero if $n \equiv j \pmod{2}$. To make this more transparent, we set

$$P(z) = zr_1 = \frac{1 + z^2 + \sqrt{1 - 6z^2 + 5z^4}}{2},$$

and then

$$s_j = f_j + g_j + h_j = z^j \frac{3 - 3z^2 - W}{2P^{j+1}}.$$

Now we read off coefficients. We do this using residues and contour integration. The path of integration, in both variables x resp. v is a small circle or an equivalent contour.

$$\begin{aligned}
[z^{2m+j}]s_j &= [z^{2m}] \frac{3 - 3z^2 - W}{2P^{j+1}} = [x^m] \frac{3 - 3x - \sqrt{1 - 6x + 5x^2}}{2 \left(\frac{1+x-\sqrt{1-6x+5x^2}}{2} \right)^{j+1}} \\
&= [x^m] \frac{3 - 3 \frac{v}{1+3v+v^2} - \frac{1-v^2}{1+3v+v^2}}{2 \left(\frac{v(v+2)}{1+3v+v^2} \right)^{j+1}} \\
&= [x^m] \frac{(1+v)(1+2v)}{v^{j+1}(v+2)^{j+1}} (1+3v+v^2)^j \\
&= \frac{1}{2\pi i} \oint \frac{dx}{x^{m+1}} \frac{(1+v)(1+2v)}{v^{j+1}(v+2)^{j+1}} (1+3v+v^2)^j \\
&= \frac{1}{2\pi i} \oint \frac{dv}{v^{m+1}} \frac{(1+v)(1+2v)(1-v^2)}{v^{j+1}(v+2)^{j+1}} (1+3v+v^2)^{m-1+j} \\
&= [v^{m+j+1}] \frac{(1+v)^2(1+2v)(1-v)}{(v+2)^{j+1}} (1+3v+v^2)^{m-1+j}.
\end{aligned}$$

Note that

$$(1+v)^2(1+2v)(1-v) = -9 + 27(v+2) - 29(v+2)^2 + 13(v+2)^3 - 2(v+2)^4;$$

consequently

$$\begin{aligned}
[v^k] \frac{(1+v)^2(1+2v)(1-v)}{(v+2)^{j+1}} \\
&= -9 \frac{1}{2^{j+1+k}} \binom{-j-1}{k} + 27 \frac{1}{2^{j+k}} \binom{-j}{k} - 29 \frac{1}{2^{j-1+k}} \binom{-j+1}{k} \\
&\quad + 13 \frac{1}{2^{j-2+k}} \binom{-j+2}{k} - 2 \frac{1}{2^{j-3+k}} \binom{-j+3}{k} =: \lambda_{j;k}.
\end{aligned}$$

With this abbreviation we find

$$[v^{m+j+1}] \frac{(1+v)^2(1+2v)(1-v)}{(v+2)^{j+1}} (1+3v+v^2)^{m-1+j} = \sum_{k=0}^{m+j+1} \lambda_{j;k} \binom{m-1+j; 1, 3, 1}{m+j+1-k}.$$

This is not extremely pretty but it is *explicit* and as good as it gets. Here are the first few generating functions:

$$\begin{aligned}
s_0 &= 1 + z^2 + 3z^4 + 10z^6 + 36z^8 + 137z^{10} + 543z^{12} + \dots \\
s_1 &= z + 2z^3 + 6z^5 + 21z^7 + 79z^9 + 311z^{11} + 1265z^{13} + \dots \\
s_2 &= z^2 + 3z^4 + 10z^6 + 37z^8 + 145z^{10} + 589z^{12} + 2455z^{14} + \dots \\
s_3 &= z^3 + 4z^5 + 15z^7 + 59z^9 + 241z^{11} + 1010z^{13} + 4314z^{15} + \dots
\end{aligned}$$

We could also give such lists for the functions f_j , g_j , h_j , if desired. We summarize the essential findings of this section:

Theorem 2. *The generating function of decorated (partial) Dyck paths, consisting of n steps, ending on level j , is given by*

$$S(z, u) = \frac{3 - 3z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2zr_1(1 - u/r_1)},$$

with

$$r_1 = \frac{1 + z^2 + \sqrt{1 - 6z^2 + 5z^4}}{2z}.$$

Furthermore

$$[u^j]S(z, u) = \frac{3 - 3z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2zr_1^{j+1}}.$$

Open ended paths. If we do not specify the end of the paths, in other words we sum over all $j \geq 0$, then at the level of generating functions this is very easy, since we only have to set $u := 1$. We find

$$\begin{aligned} S(1) &= -\frac{(z+1)(z^2+3z-2) + (z+2)\sqrt{1-6z^2+5z^4}}{2z(z^2+2z-1)} \\ &= 1 + z + 2z^2 + 3z^3 + 7z^4 + 11z^5 + 26z^6 + 43z^7 + 102z^8 + 175z^9 + 416z^{10} + \dots \end{aligned}$$

Counting red edges. We can use an extra variable, w , to count additionally the red edges that occur in a path. We use the same letters for generating functions. Eventually, the coefficient $[z^n u^j w^k]S$ is the number of (partial) paths consisting of n steps, leading to level j , and having passed k red edges. The endpoint of the original skew path has then coordinates $(n - 2k, j)$. The computations are very similar, and we only sketch the key steps.

$$f_0 = 1, \quad f_{i+1} = zf_i + zg_i, \quad i \geq 0,$$

$$g_i = zf_{i+1} + zg_{i+1} + zh_{i+1}, \quad i \geq 0,$$

$$h_i = wzh_{i+1} + wzg_{i+1}, \quad i \geq 0;$$

$$\frac{1}{u}(F(u) - 1) = zF(u) + zG(u),$$

$$G(u) = \frac{z}{u}(F(u) - 1) + \frac{z}{u}(G(u) - G(0)) + \frac{z}{u}(H(u) - H(0)),$$

$$H(u) = \frac{wz}{u}(G(u) - G(0)) + \frac{wz}{u}(H(u) - H(0));$$

$$F(u) = \frac{z^2uG(0) + z^2uH(0) + z^2u - u - wz^3 + z + wz}{-wz^3 - u + z + wz + zu^2 - wz^2u},$$

$$G(u) = \frac{z(H(0) - uzH(0) + wz^2 + G(0) - zuG(0) - zu)}{-wz^3 - u + z + wz + zu^2 - wz^2u},$$

$$H(u) = \frac{wz(-uzH(0) - z^2 - zuG(0) + G(0) - z^2H(0) + H(0) - z^2G(0))}{-wz^3 - u + z + wz + zu^2 - wz^2u}.$$

The denominator factors as $z(u - r_1)(u - r_2)$, with

$$r_1 = \frac{1 + wz^2 + \sqrt{1 - (4 + 2w)z^2 + (4w + w^2)z^4}}{2z},$$

$$r_2 = \frac{1 + wz^2 - \sqrt{1 - (4 + 2w)z^2 + (4w + w^2)z^4}}{2z}.$$

Note the factorization $1 - (4 + 2w)z^2 + (4w + w^2)z^4 = (1 - z^2w)(1 - (4 + w)z^2)$. Since the factor $u - r_2$ in the denominator is “bad,” it must also cancel in the numerators. From this we eventually find, with the abbreviation $W = \sqrt{1 - (4 + 2w)z^2 + (4w + w^2)z^4}$

$$F(u) = \frac{-1 - wz^2 - W}{2z(u - r_1)},$$

$$G(u) = \frac{-1 + wz^2 + W}{2z(u - r_1)},$$

$$H(u) = \frac{-1 + (2 + w)z^2 + W}{2z(u - r_1)}.$$

The total generating function is

$$S(u) = F(u) + G(u) + H(u) = \frac{-2 - w + z^2(w + w^2) + wW}{2z(u - r_1)}.$$

The special case $u = 0$ (return to the x -axis) is to be noted:

$$S(0) = \frac{-2 - w + z^2(w + w^2) + wW}{-2zr_1} = \frac{1 - wz^2 - W}{2z^2}.$$

Since there are only even powers of z in this function, we replace $x = z^2$ and get

$$S(0) = \frac{1 - wx - \sqrt{1 - (4 + 2w)x + (4w + w^2)x^2}}{2x}$$

$$= 1 + x + (w + 2)x^2 + (w^2 + 4w + 5)x^3 + (w^3 + 6w^2 + 15w + 14)x^4 + \dots$$

Compare the factor $(w^2 + 4w + 5)$ with the earlier drawing of the 10 paths.

There is again a substitution that allows for better results:

$$z = \frac{v}{1 + (2 + w)v + v^2}, \quad \text{then} \quad S(0) = 1 + v.$$

Reading off coefficients can now be done using modified trinomial coefficients:

$$\binom{n; 1, 2 + w, 1}{k} = [t^k](1 + (2 + w)t + t^2)^n.$$

Again, we use contour integration to extract coefficients:

$$[x^n](1 + v) = \frac{1}{2\pi i} \oint \frac{dx}{x^{n+1}}(1 + v)$$

$$= \frac{1}{2\pi i} \oint \frac{dx}{v^{n+1}} \frac{1 - v^2}{(1 + (2 + w)v + v^2)^2} (1 + (2 + w)v + v^2)^{n+1} (1 + v)$$

$$= [v^n](1 - v)(1 + v)^2(1 + (2 + w)v + v^2)^{n-1}$$

$$= \binom{n-1; 1, 2+w, 1}{n} + \binom{n-1; 1, 2+w, 1}{n-1} - \binom{n-1; 1, 2+w, 1}{n-2} - \binom{n-1; 1, 2+w, 1}{n-3}.$$

Now we want to count the average number of red edges. For that, we differentiate $S(0)$ w.r.t. w , followed by $w := 1$. This leads to

$$\frac{-1 + 6x - 5x^2 + (1 + 3x)\sqrt{1 - 6x + 5x^2}}{2(1-x)(1-5x)}.$$

A simple application of singularity analysis leads to

$$\frac{\frac{1}{2\sqrt{5}}[x^n]_{\sqrt{1-5x}}}{-\sqrt{5}[x^n]_{\sqrt{1-5x}}} \sim \frac{n}{5}.$$

So, a random path consisting of $2n$ steps has about $n/5$ red steps, on average.

For readers who are not familiar with singularity analysis of generating functions [14, 15], we just mention that one determines the local expansion around the dominating singularity, which is at $z = \frac{1}{5}$ in our instance. In the denominator, we just have the total number of skew Dyck paths, according to the sequence A002212 in [47].

In the example of Figure 2, the exact average is $6/10$, which curiously is exactly the same as $3/5$.

We finish the discussion by considering fixed powers of w in $S(0)$, counting skew Dyck paths consisting of zero, one, two, three, ... red edges. We find

$$\begin{aligned} [w^0]S(0) &= \frac{1 - \sqrt{1 - 4x}}{2x}, \\ [w^1]S(0) &= \frac{1 - 2x - \sqrt{1 - 4x}}{2\sqrt{1 - 4x}}, \\ [w^2]S(0) &= \frac{x^3}{(1 - 4x)^{3/2}}, \\ [w^3]S(0) &= \frac{x^4(1 - 2x)}{(1 - 4x)^{5/2}}, \\ [w^4]S(0) &= \frac{x^5(1 - 4x + 5x^2)}{(1 - 4x)^{7/2}}, \quad \&c. \end{aligned}$$

The generating function $[w^0]S(0)$ is of course the generating function of Catalan numbers, since no red edges just means: ordinary Dyck paths. We can also conclude that the asymptotic behaviour is of the form $n^{k-3/2}4^n$, where the polynomial contribution gets higher, but the exponential growth stays the same: 4^n . This is compared to the scenario of an *arbitrary* number of red edges, when we get an exponential growth of the form 5^n .

Dual skew Dyck paths. The mirrored version of skew Dyck paths with two types of up-steps, $(1, 1)$ and $(-1, 1)$ are also cited among the objects in A002212 in [47]. We call them dual skew paths and drop the ‘dual’ when it isn’t necessary. When the paths come back to the x -axis, no new enumeration is necessary, but this is no longer true for paths ending at level j .

Here is a list of the 10 skew paths consisting of 6 steps:

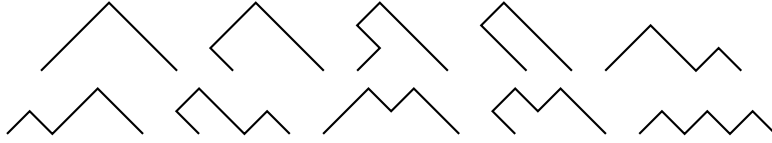


FIGURE 7. All 10 dual skew Dyck paths of length 6 (consisting of 6 steps).

We prefer to work with the equivalent model (resembling more traditional Dyck paths) where we replace each step $(-1, -1)$ by $(1, -1)$ but label it blue. Here is the list of the 10 paths again (Figure 2):

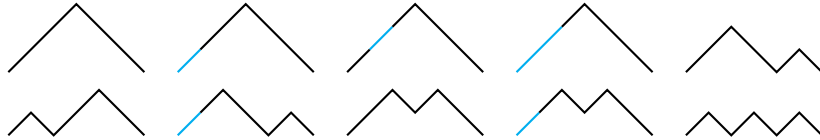


FIGURE 8. All 10 dual skew Dyck paths of length 6 (consisting of 6 steps).

The rules to generate such decorated Dyck paths are: each edge $(1, -1)$ may be black or blue, but \searrow and \swarrow are forbidden.

Our interest is in particular in *partial* decorated Dyck paths, ending at level j , for fixed $j \geq 0$; the instance $j = 0$ is the classical case.

As before, two variables, z and u , are used, where z marks the length of the path and j marks the end-level. We briefly mention that one can, using a third variable w , also count the number of blue edges.

The substitution

$$x = \frac{v}{1 + 3v + v^2},$$

is again the key to the success.

Generating functions and the kernel method. We catch the essence of a decorated (dual skew) Dyck path using a state-diagram:

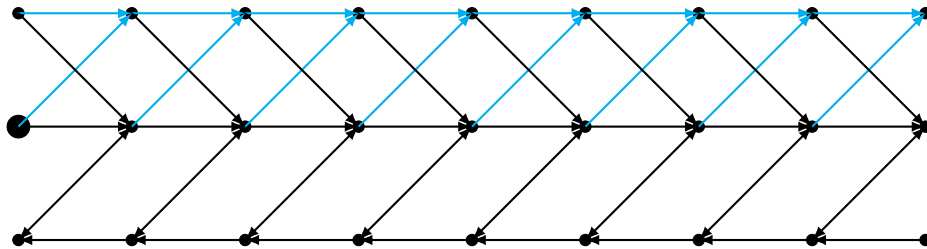


FIGURE 9. Three layers of states according to the type of steps leading to them (down, up-black, up-blue).

It has three types of states, with j ranging from 0 to infinity; in the drawing, only $j = 0..8$ is shown. The first layer of states refers to an up-step leading to a state, the second layer refers to a black down-step leading to a state and the third layer

refers to a blue down-step leading to a state. We will work out generating functions describing all paths leading to a particular state. We will use the notations c_j, a_j, b_j for the three respective layers, from top to bottom. Note that the syntactic rules of forbidden patterns \nearrow and \searrow can be clearly seen from the picture. The functions depend on the variable z (marking the number of steps), but mostly we just write a_j instead of $a_j(z)$, etc.

The following recursions can be read off immediately from the diagram:

$$\begin{aligned} a_0 &= 1, & a_{i+1} &= za_i + zb_i + zc_i, & i &\geq 0, \\ b_i &= za_{i+1} + zb_{i+1}, & i &\geq 0, \\ c_{i+1} &= za_i + zc_i, & i &\geq 0. \end{aligned}$$

And now it is time to introduce the promised *bivariate* generating functions:

$$A(z, u) = \sum_{i \geq 0} a_i(z)u^i, \quad B(z, u) = \sum_{i \geq 0} b_i(z)u^i, \quad C(z, u) = \sum_{i \geq 0} c_i(z)u^i.$$

Again, often we just write $A(u)$ instead of $A(z, u)$ and treat z as a ‘silent’ variable. Summing the recursions leads to

$$\begin{aligned} \sum_{i \geq 0} u^i a_i &= 1 + u \sum_{i \geq 0} u^i (za_i + zb_i + zc_i) \\ &= 1 + uzA(u) + uzB(u) + uzC(u), \\ \sum_{i \geq 0} u^i b_i &= \sum_{i \geq 0} u^i (za_{i+1} + zb_{i+1}) \\ &= \frac{z}{u} \sum_{i \geq 1} u^i a_i + \frac{z}{u} \sum_{i \geq 1} u^i b_i, \\ \sum_{i \geq 1} u^i c_i &= uz \sum_{i \geq 0} u^i a_i + uz \sum_{i \geq 0} u^i c_i. \end{aligned}$$

This can be rewritten as

$$\begin{aligned} A(u) &= 1 + uzA(u) + uzB(u) + uzC(u), \\ B(u) &= \frac{z}{u}(A(u) - a_0) + \frac{z}{u}(B(u) - b_0), \\ C(u) &= c_0 + uzA(u) + uzC(u). \end{aligned}$$

Note that $a_0 = 1, c_0 = 0$. Simplification leads to

$$C(u) = \frac{uzA(u)}{1 - uz}$$

and

$$B(u) = \frac{z(A(u) - 1 - B(0))}{u - z},$$

leaving us with just one equation

$$A(u) = \frac{(z - u + uz^2 + uz^2 B(0))(uz - 1)}{u^2 z^3 + uz^2 - 2u^2 z - z + u}.$$

This is again a typical application of the kernel method.

$$u^2 z^3 + uz^2 - 2u^2 z - z + u = z(z^2 - 2)(u - s_1)(u - s_2)$$

The denominator factors as $2z(z^2 - 2)(u - s_1)(u - s_2)$, with

$$s_1 = \frac{1 + z^2 + \sqrt{1 - 6z^2 + 5z^4}}{2z(2 - z^2)}, \quad s_2 = \frac{1 + z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2z(2 - z^2)}.$$

Note that $s_1 s_2 = \frac{1}{2 - z^2}$. Since the factor $u - s_2$ in the denominator is “bad,” it must also cancel in the numerators. From this we conclude (again with the abbreviation $W = \sqrt{1 - 6z^2 + 5z^4}$)

$$B(0) = \frac{zs_2}{1 - 2zs_2},$$

and further

$$A(u) = \frac{(1 - uz)(1 + z^2 + W)}{2z(z^2 - 2)(u - s_1)},$$

$$B(u) = \frac{1 - 2z^2 - W}{z(2 - z^2)(u - s_1)},$$

$$C(u) = \frac{1 + z^2 + W}{2(z^2 - 2)} \frac{u}{u - s_1},$$

and for the function of main interest

$$G(u) = A(u) + B(u) + C(u) = \frac{3z^2 - 3 + W}{2z(2 - z^2)(u - s_1)}.$$

Note that

$$\frac{1}{s_1} = \frac{1 + z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2z} = zS,$$

$$\frac{1}{s_2} = \frac{1 + z^2 + \sqrt{1 - 6z^2 + 5z^4}}{2z}.$$

Then

$$\begin{aligned} [u^j]G(u) &= [u^j] \frac{3z^2 - 3 + W}{2z(z^2 - 2)s_1(1 - u/s_1)} \\ &= \frac{3z^2 - 3 + W}{2z(z^2 - 2)s_1^{j+1}} = \frac{3z^2 - 3 + W}{2(z^2 - 2)} z^j S^{j+1}. \end{aligned}$$

So $[u^j]G(u)$ contains only powers of the form z^{j+2N} . Now we continue

$$\begin{aligned} [z^{j+2N}u^j]G(u) &= [z^{2N}] \frac{3z^2 - 3 + W}{2(z^2 - 2)} S^{j+1} \\ &= [x^N] \frac{3x - 3 + \sqrt{1 - 6x + 5x^2}}{2(x - 2)} \left(\frac{1 + x - \sqrt{1 - 6x + 5x^2}}{2x} \right)^{j+1} \\ &= [x^N](v + 1)(v + 2)^j \end{aligned}$$

which is the generating function of all (partial) paths ending at level j .

Now we read off coefficients. We do this using residues and contour integration. The path of integration, in both variables x resp. v is a small circle or an equivalent contour;

$$\begin{aligned} [z^{j+2N}u^j]G(u) &= [x^N](v + 1)(v + 2)^j \\ &= \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} (v + 1)(v + 2)^j \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi i} \oint \frac{dv}{v^{N+1}} (1+3v+v^2)^{N+1} \frac{(1-v^2)}{(1+3v+v^2)^2} (v+1)(v+2)^j \\
&= [v^N] (1+3v+v^2)^{N-1} (1-v)(1+v)^2 (v+2)^j.
\end{aligned}$$

Note that

$$(1-v)(1+v)^2 = 3 - 7(v+2) + 5(v+2)^2 - (v+2)^3;$$

consequently

$$[z^{j+2N}u^j]G(u) = [v^N](1+3v+v^2)^{N-1} \left[3 - 7(v+2) + 5(v+2)^2 - (v+2)^3 \right] (v+2)^j.$$

We abbreviate:

$$\begin{aligned}
\mu_{j;k} &= [v^k] \left[3(v+2)^j - 7(v+2)^{j+1} + 5(v+2)^{j+2} - (v+2)^{j+3} \right] \\
&= 3 \binom{j}{k} 2^{j-k} - 7 \binom{j+1}{k} 2^{j+1-k} + 5 \binom{j+2}{k} 2^{j+2-k} - \binom{j+3}{k} 2^{j+3-k}.
\end{aligned}$$

With this notation we get

$$[z^{j+2N}u^j]G(u) = \sum_{0 \leq k \leq N-1} \mu_{j;k} \binom{N-1; 1, 3, 1}{N-k}.$$

Here are the first few generating functions:

$$\begin{aligned}
G_0 &= 1 + z^2 + 3z^4 + 10z^6 + 36z^8 + 137z^{10} + 543z^{12} + 2219z^{14} + \dots \\
G_1 &= 2z + 3z^3 + 10z^5 + 36z^7 + 137z^9 + 543z^{11} + 2219z^{13} + 9285z^{15} + \dots \\
G_2 &= 4z^2 + 8z^4 + 29z^6 + 111z^8 + 442z^{10} + 1813z^{12} + 7609z^{14} + 32521z^{16} + \dots \\
G_3 &= 8z^3 + 20z^5 + 78z^7 + 315z^9 + 1306z^{11} + 5527z^{13} + 23779z^{15} + 103699z^{17} + \dots
\end{aligned}$$

We could also give such lists for the functions a_j , b_j , c_j , if desired. We summarize the essential findings of the rest of this section:

Theorem 3. *The generating function of decorated (partial) dual skew Dyck paths, consisting of n steps, ending on level j , is given by*

$$G(z, u) = \frac{3z^2 - 3 + \sqrt{1 - 6z^2 + 5z^4}}{2z(2 - z^2)(u - s_1)},$$

with

$$s_1 = \frac{2z}{1 + z^2 - \sqrt{1 - 6z^2 + 5z^4}}.$$

Furthermore

$$[u^j]G(z, u) = \frac{3z^2 - 3 + \sqrt{1 - 6z^2 + 5z^4}}{2(z^2 - 2)} z^j S^{j+1},$$

with

$$S = \frac{1 + z^2 - \sqrt{1 - 6z^2 + 5z^4}}{2z^2}.$$

Open ended paths. If we do not specify the end of the paths, in other words we sum over all $j \geq 0$, then at the level of generating functions this is very easy, since we only have to set $u := 1$. We find

$$\begin{aligned} G(1) &= \frac{(1+z)(1-3z)}{2z(z^2+2z-1) - \sqrt{1-6z^2+5z^4}} \\ &= 1 + 2z + 5z^2 + 11z^3 + 27z^4 + 62z^5 + 151z^6 + 354z^7 + 859z^8 + 2036z^9 + \dots \end{aligned}$$

Counting blue edges. We can use an extra variable, w , to count additionally the blue edges that occur in a path. We use the same letters for generating functions. Eventually, the coefficient $[z^n u^j w^k]S$ is the number of (partial) paths consisting of n steps, leading to level j , and having passed k blue edges. The endpoint of the original skew path has then coordinates $(n - 2k, j)$. The computations are very similar, and we only sketch the key steps.

$$\begin{aligned} a_0 &= 1, & a_{i+1} &= za_i + zb_i + zc_i, & i &\geq 0, \\ b_i &= za_{i+1} + zb_{i+1}, & i &\geq 0, \\ c_{i+1} &= wza_i + wzc_i, & i &\geq 0. \end{aligned}$$

This leads to

$$\begin{aligned} A(u) &= 1 + uzA(u) + uzB(u) + uzC(u), \\ B(u) &= \frac{z}{u}(A(u) - a_0) + \frac{z}{u}(B(u) - b_0), \\ C(u) &= c_0 + wuzA(u) + wuzC(u). \end{aligned}$$

Solving,

$$S(u) = A(u) + B(u) + C(u) = \frac{u - wuz^2 - zA(0) - zB(0) + uwz^2A(0) + uwz^2B(0)}{u^2z^3w + u - wu^2z - u^2z - z + wuz^2}.$$

The denominator factors as $-z(1+w-z^2w)(u-s_1)(u-s_2)$, with

$$\begin{aligned} s_1 &= \frac{1 + z^2w + \sqrt{1 - 2z^2w + z^4w^2 - 4z^2 + 4z^4w}}{2z(1+w-z^2w)}, \\ s_2 &= \frac{1 + z^2w - \sqrt{1 - 2z^2w + z^4w^2 - 4z^2 + 4z^4w}}{2z(1+w-z^2w)}. \end{aligned}$$

Note the factorization $1 - (4+2w)z^2 + (4w+w^2)z^4 = (1-z^2w)(1-(4+w)z^2)$. Since the factor $u - r_2$ in the denominator is ‘‘bad,’’ it must also cancel in the numerators. From this we eventually find, with the abbreviation $W = \sqrt{1 - (4+2w)z^2 + (4w+w^2)z^4}$)

$$G(0) = \frac{1 - z^2w - W}{2z^2},$$

and further

$$G(u) = \frac{w - z^2w^2 - wW + 2 - 2z^2w}{2z(-w - 1 + z^2w)(u - s_1)}.$$

The special case $u = 0$ (return to the x -axis) is to be noted:

$$G(0) = 1 + z^2 + (w+2)z^4 + (w^2+4w+5)z^6 + (w+2)(w^2+4w+7)z^8 + \dots$$

Compare the factor $(w^2 + 4w + 5)$ with the earlier drawing of the 10 paths. There is again a substitution that allows for better results:

$$z = \frac{v}{1 + (2 + w)v + v^2}, \quad \text{then } G(0) = 1 + v.$$

Since $S(u) = G(u)$ with $S(u)$ from the first part of the paper, as it means the same objects, read from left to right resp. from right to left, no new analysis is required.

8. COUNTING TERNARY TREES ACCORDING TO THE NUMBER OF MIDDLE EDGES AND FACTORIZING INTO $(3/2)$ -ARY TREES

The recent preprint [4] triggered my interest in the sequence A120986 in [47]. The double-indexed sequence enumerates ternary trees according to the number of edges and the number of middle edges. We consider here $T(n, k)$, the number of ternary trees with n nodes and k middle edges. The difference is marginal, but we want to compare/relate our analysis with [26], and there it is also the number of nodes that is considered. Let $G = G(x, u) = \sum_{n, k \geq 0} T(n, k)x^n u^k$. Then it is easy to see (decomposition at the root) that

$$G = 1 + xG^2(1 - u + uG).$$

The substitution

$$x = \frac{t(1 - t)^2}{(1 - t + ut)}$$

makes the cubic equation manageable and also allows, as in [26], to introduce a (refined) version of the $(3/2)$ -ary trees.

Here is a small table of these numbers and a ternary tree:

$n \backslash k$	0	1	2	3	4	5
0	1					
1	1					
2	2	1				
3	5	6	1			
4	14	28	12	1		
5	42	120	90	20	1	
6	132	495	550	220	30	1

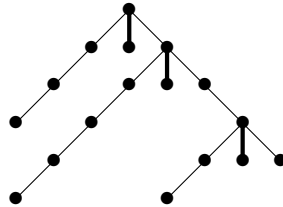


FIGURE 10. Ternary tree with 17 nodes and 3 middle edges

Analysis of the cubic equation. The cubic equation has the following solutions:

$$\begin{aligned} r_1 &= \frac{1}{1-t}, \\ r_2 &= \frac{-t+t^2-t^2u+\sqrt{t(1-t+ut)(4u+t-4ut-t^2+t^2u)}}{2ut(1-t)}, \\ r_3 &= \frac{-t+t^2-t^2u-\sqrt{t(1-t+ut)(4u+t-4ut-t^2+t^2u)}}{2ut(1-t)}. \end{aligned}$$

Note that

$$r_2 r_3 = -\frac{1-t+ut}{ut(1-t)}.$$

The root with the combinatorial significance is r_1 . But it is the explicit form of the two other roots that makes everything here interesting and challenging.

We extract coefficients of r_1 using contour integration, which is closely related to the Lagrange inversion formula. The path of integration is a small circle in the x -plane which is then transformed into a small circle in the t -plane.

$$\begin{aligned} [x^n]r_1 &= \frac{1}{2\pi i} \oint \frac{dx}{x^{n+1}} \frac{1}{1-t} \\ &= \frac{1}{2\pi i} \oint \frac{dt(1-t)(1-3t+2t^2-2t^2u)}{(1-t+tu)^2} \frac{(1-t+tu)^{n+1}}{t^{n+1}(1-t)^{2n+2}} \frac{1}{1-t} \\ &= [t^n](1-3t+2t^2-2t^2u) \frac{(1-t+tu)^{n-1}}{(1-t)^{2n+2}}. \end{aligned}$$

Furthermore

$$\begin{aligned} [x^n u^k]r_1 &= [t^n][u^k](1-3t+2t^2-2t^2u) \frac{(1-t+tu)^{n-1}}{(1-t)^{2n+2}} \\ &= [t^n] \binom{n-1}{k} \frac{t^k(1-2t)}{(1-t)^{n+k+2}} - 2[t^n] \binom{n-1}{k-1} \frac{t^{k+1}}{(1-t)^{n+k+2}} \\ &= \binom{n-1}{k} [t^{n-k}] \frac{(1-2t)}{(1-t)^{n+k+2}} - 2 \binom{n-1}{k-1} [t^{n-k-1}] \frac{1}{(1-t)^{n+k+2}} \\ &= \binom{n-1}{k} \binom{2n+1}{n-k} - 2 \binom{n-1}{k} \binom{2n}{n-k-1} - 2 \binom{n-1}{k-1} \binom{2n}{n-k-1} \\ &= \frac{1}{n} \binom{n}{k} \binom{2n}{n-1-k}. \end{aligned}$$

For $u = 1$, which means that the middle edges are not especially counted, we get

$$\sum_k \frac{1}{n} \binom{n}{k} \binom{2n}{n-1-k} = \frac{1}{n} \binom{3n}{n-1},$$

the number of ternary trees with n nodes.

Factorizing the solution of the cubic equation. For $u = 1$, Knuth [26] was able to factor the generating function r_1 into two factors, for which he coined the catchy name (3/2)-ary trees. For this factorization, see also [34, 3]. The goal in this section is to perform this factorization in the context of counting middle edges, i. e., for the

generating function with the additional variable u . In Knuth's instance, the generating function was expressible as a generalized binomial series (in the sense of Lambert [18]), but that does not seem to be an option here.

Note that

$$\frac{1}{r_2} = \frac{t}{2} - \frac{\sqrt{t}\sqrt{t(1-t) + u(2-t)^2}}{2\sqrt{1-t+tu}}$$

and

$$\frac{1}{r_3} = \frac{t}{2} + \frac{\sqrt{t}\sqrt{t(1-t) + u(2-t)^2}}{2\sqrt{1-t+tu}}.$$

From the cubic equation we deduce that

$$r_1 = -\frac{1}{uxr_2r_3},$$

which is the desired factorization. The factor ux will be fairly split as $\sqrt{ux} \cdot \sqrt{ux}$, whereas the minus sign goes to the factor $1/r_2$. In the following we work out how this factorization can be obtained. To say it again, it is not as appealing as in the original case.

Let us write

$$t = x\Phi(t), \quad \text{with} \quad \Phi(t) = \frac{1-t+tu}{(1-t)^2},$$

so that we can use the Lagrange inversion formula to get

$$[x^n]t^\ell = \frac{\ell}{n}[t^{n-\ell}]\Phi(t)^n$$

and

$$\begin{aligned} [x^n u^k]t^\ell &= \frac{\ell}{n}[t^{n-\ell}][u^k] \frac{(1-t+tu)^n}{(1-t)^{2n}} \\ &= \frac{\ell}{n}[t^{n-\ell-k}] \binom{n}{k} \frac{1}{(1-t)^{n+k}} = \frac{\ell}{n} \binom{n}{k} \binom{2n-\ell-1}{n-\ell-k}. \end{aligned}$$

In particular,

$$t = \sum_{n \geq 1} x^n \sum_{0 \leq k \leq n} \frac{1}{n} \binom{n}{k} \binom{2n-2}{n-1-k} u^k;$$

this series expansion may be used in the following developments whenever needed.

To proceed further, we set $u = 1 + U$ and $\tau = t/u$:

$$\begin{aligned} \frac{1}{r_2} &= \frac{t}{2} - \frac{\sqrt{x}}{2(1-t)} \sqrt{4-3t+U(2-t)^2}, \\ \frac{1}{r_3} &= \frac{t}{2} + \frac{\sqrt{x}}{2(1-t)} \sqrt{4-3t+U(2-t)^2} \end{aligned}$$

Since the first term is well understood, we concentrate on the second:

$$\begin{aligned} &\frac{\sqrt{x}}{2(1-t)} \sqrt{4-3t+U(2-t)^2} \\ &= \sqrt{ux} \left(1 + \frac{1}{8} (5+4U) \tau + \frac{1}{128} (71+136U+64U^2) \tau^2 \right) \end{aligned}$$

$$+ \frac{1}{1024} (541 + 1596U + 1568U^2 + 512U^3) \tau^3 + \dots \Big) =: \sqrt{ux} \cdot \Xi.$$

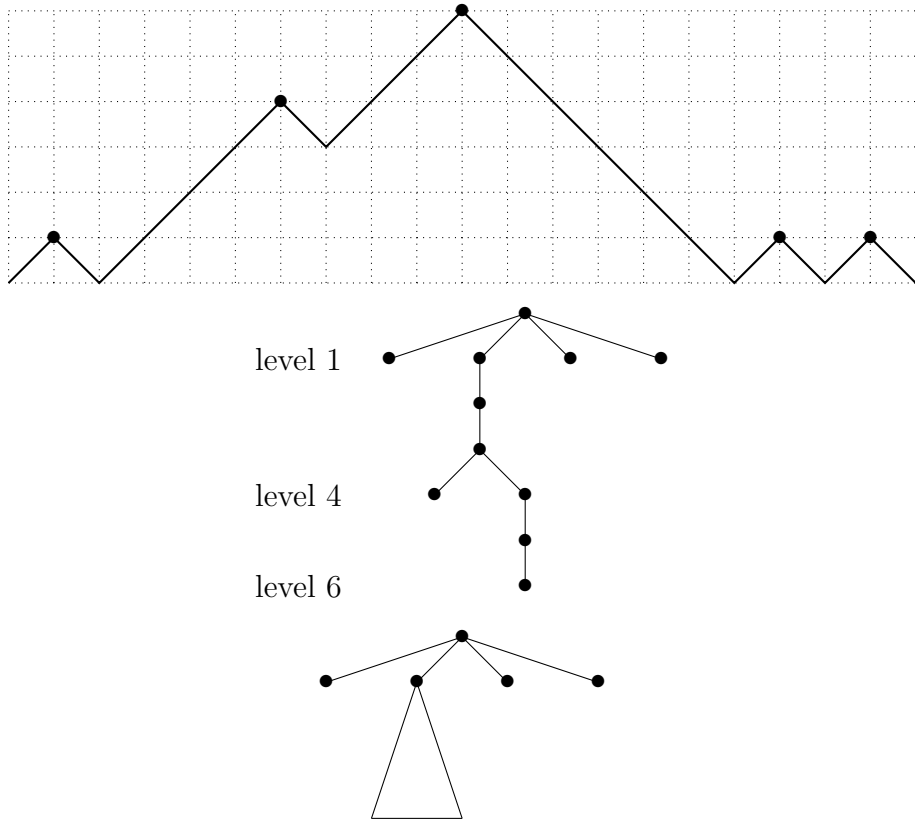
With this expanded form Ξ , we have now our final formula, the expansion of r_1 into two factors:

$$r_1 = \frac{1}{1-t} = -\frac{1}{ux} \frac{1}{r_2} \frac{1}{r_3} = \left(-\frac{1}{\sqrt{ux}} \frac{t}{2} + \Xi \right) \left(\frac{1}{\sqrt{ux}} \frac{t}{2} + \Xi \right).$$

These two factors do not have a combinatorial meaning, as it seems, but we can still stick to the $(3/2)$ -ary tree notation, with the additional counting of middle edges.

9. MORE ABOUT MOTZKIN PATHS

Retakh's Motzkin paths. V. Retakh [12] introduced the following restricted class of Dyck paths: Peaks are only allowed on level 1 and on even-numbered levels. Here is an example, and the corresponding plane tree using the standard bijection.



Ekhad and Zeilberger [12] proved recently that these restricted paths are enumerated by Motzkin numbers. Recall that the generating function of the Motzkin numbers $M(z)$ according to length satisfies $M = 1 + zM + z^2M^2$ and thus

$$M(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2}.$$

Here, I want to present a few additional observations, also including the height of the paths (or the associated plane trees). First, we are going to confirm the connection to Motzkin paths:

Since the level 1 is somewhat special, we only consider trees as symbolized by the triangle. We will use two generating functions, to deal with the odd/even situation. We have

$$F = \frac{zG}{1-G} \quad \text{and} \quad G = \frac{z}{1-F}.$$

This is not too difficult to see, since the family \mathcal{F} of trees as symbolized by the triangle does not contain a single node, so $F = zG + zG^2 + \dots$. However, the next generation \mathcal{G} can contain a single node, and thus $G = z + zF + zF^2 + \dots$.

Solving this (best by a computer) we find $F(z) = z^2M(z)$ and the total generating function (allowing sequences of single nodes between copies of \mathcal{F}) is

$$\frac{z}{1-z} \sum_{r \geq 0} \left(\frac{F}{1-z} \right)^r = zM(z),$$

as predicted. Recall that the number of nodes in trees is always one more than the half-length of the corresponding Dyck path.

We will compute the average height of such restricted paths, using singularity analysis of generating functions, as in [14, 15]. Whether we define the height in terms of the maximal chain of edges resp. nodes only makes a difference of one, and we will only compute the average height according to the leading term of order \sqrt{n} . For readers who wish to see how more terms could be computed, at least in principle, we cite [38].

The average height of planted plane trees (and subclasses of them) has been of central interest to combinatorialists and theoretical computer scientists alike since the seminal paper [8]. The number of leaves (endnodes) is also one of the key parameters since Narayana [27]. We investigate it in the last section of this paper to see how the restrictions according to Retakh influence this parameter.

The height. Now we will use the substitution $z = \frac{v}{1+v+v^2}$, which occurred for the first time in [39], but has been used more recently in different models where Motzkin numbers are involved [40, 20, 2]. Motzkin numbers appear in [47] as sequence A001006. For example, the generating function $M(z)$ of Motzkin numbers has the very simple form $M(z) = 1 + v + v^2$ using this auxiliary variable. We define

$$G_{k+1} = \frac{z}{1 - \frac{zG_k}{1-G_k}}, \quad \text{with} \quad G_1 = z.$$

There is a simple formula, viz.

$$G_k = \frac{v}{1+v} \frac{1-v^{2k}}{1-v^{2k+1}}.$$

This is easy to prove by induction, which we will do for the convenience of the reader. The start is

$$G_1 = \frac{v}{1+v} \frac{1-v^2}{1-v^3} = \frac{v}{1+v} \frac{1+v}{1+v+v^2} = \frac{v}{1+v+v^2} = z.$$

And now

$$G_{k+1} = \frac{z}{1 - \frac{zG_k}{1-G_k}} = \frac{z(1-G_k)}{1 - (1+z)G_k} = \frac{v}{1+v+v^2} \frac{1 - \frac{v}{1+v} \frac{1-v^{2k}}{1-v^{2k+1}}}{1 - \frac{(1+v)^2}{1+v+v^2} \frac{v}{1+v} \frac{1-v^{2k}}{1-v^{2k+1}}}$$

$$\begin{aligned}
&= \frac{v}{1+v} \frac{1+v - v \frac{1-v^{2k}}{1-v^{2k+1}}}{1+v+v^2 - v(1+v) \frac{1-v^{2k}}{1-v^{2k+1}}} = \frac{v}{1+v} \frac{(1+v)(1-v^{2k+1}) - v(1-v^{2k})}{(1+v+v^2)(1-v^{2k+1}) - v(1+v)(1-v^{2k})} \\
&= \frac{v}{1+v} \frac{1-v^{2k+2}}{1-v^{2k+3}},
\end{aligned}$$

as claimed. From this we also get

$$F_k = \frac{zG_k}{1-G_k} = \frac{v^2}{1+v+v^2} \frac{1-v^{2k}}{1-v^{2k+2}}.$$

For $k \geq 1$, F_k is the generating function of trees (like in the triangle) of height $\leq 2k$.

Note that the height is currently counted in terms of nodes;

$$F_1 = \frac{z^2}{1-z},$$

which describes a root with $\ell \geq 1$ leaves attached to the root.

Now we incorporate the irregular beginning of the tree and compute

$$\frac{z}{1-z} \sum_{r \geq 0} \left(\frac{F_h}{1-z} \right)^r = \frac{z}{1-z} \frac{1}{1 - \frac{F_h}{1-z}} = v \frac{1-v^{2h+2}}{1-v^{2h+4}}.$$

From here onwards it seems to be more natural to define the height of the whole tree in terms of the number of *edges*, and then the quantity we just derived is the generating function of all trees with height $\leq 2h$, for $h \geq 1$. Note that the limit $h \rightarrow \infty$ gives us simply $v = zM(z)$, which is consistent. There is also a contribution of trees of height ≤ 1 , namely $\frac{z}{1-z} = \frac{v}{1+v^2}$, but this term is, when we compute the average height, irrelevant and only contributes to the error term, as we only compute the leading term, which is of order \sqrt{n} .

So, apart from normalization, we are led to investigate

$$\begin{aligned}
&\sum_{h \geq 1} 2h \left[v \frac{1-v^{2h+2}}{1-v^{2h+4}} - v \frac{1-v^{2h}}{1-v^{2h+2}} \right] = 2v(1-v^{-2}) \sum_{h \geq 1} h \left[\frac{v^{2h+4}}{1-v^{2h+4}} - \frac{v^{2h+2}}{1-v^{2h+2}} \right] \\
&= 2v(1-v^{-2}) \sum_{h \geq 0} h \frac{v^{2h+4}}{1-v^{2h+4}} - 2v(1-v^{-2}) \sum_{h \geq 0} (h+1) \frac{v^{2h+4}}{1-v^{2h+4}} \\
&= -2v + \frac{2(1-v^2)}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}}.
\end{aligned}$$

Note that we could get explicit coefficients from here, using trinomial coefficients, $\binom{n,3}{k} = [v^k](1+v+v^2)^n$ (notation from [5]). To show the reader how this works, we compute

$$\begin{aligned}
[z^{n+1}] \frac{1-v^2}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} &= \frac{1}{2\pi i} \oint \frac{dz}{z^{n+2}} \frac{1-v^2}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} \\
&= \frac{1}{2\pi i} \oint dv (1-v^2)^2 \frac{(1+v+v^2)^n}{v^{n+3}} \sum_{h \geq 1} \sum_{k \geq 1} v^{2hk} \\
&= [v^{n+2}](1-2v^2+v^4) \sum_{h \geq 1} d(h) v^{2h} (1+v+v^2)^n
\end{aligned}$$

$$= \sum_{h \geq 1} d(h) \left[\binom{n, 3}{n+2-2h} - 2 \binom{n, 3}{n-2h} + \binom{n, 3}{n-2-2h} \right].$$

Note that $d(h)$ is the number of divisors of h . We will, however, not use this explicit form. The expression as derived before,

$$-2v + \frac{2(1-v^2)}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}},$$

has to be expanded around $v = 1$, which is a standard application of the Mellin transform. Details are worked out in [21], for example:

$$\sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} = \sum_{k \geq 1} d(k)v^{2k} \sim -\frac{\log(1-v^2)}{1-v^2} \sim -\frac{\log(1-v)}{2(1-v)}.$$

Note again that $d(k)$ is the number of divisors of k . Consequently

$$-2v + \frac{2(1-v^2)}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} \sim -2 \log(1-v).$$

We have $1-v \sim \sqrt{3}\sqrt{1-3z}$, and $z = \frac{1}{3}$ is the relevant singularity when discussing Motzkin numbers. We can continue

$$-2v + \frac{2(1-v^2)}{v} \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} \sim -\log(1-3z).$$

The coefficient of z^n in this is $\frac{3^n}{n}$. This has to be divided by

$$[z^n]zM(z) = [z^{n-1}]M(z) \sim \frac{3^{n+\frac{1}{2}}}{2\sqrt{\pi n^{3/2}}},$$

with the final result for the average height of restricted Dyck paths (à la Retakh):

$$\sim 2\sqrt{\frac{\pi n}{3}}.$$

Recall [39] that the average height of Motzkin paths of length n is asymptotic to

$$\sqrt{\frac{\pi n}{3}}.$$

The number of leaves. We can use a second variable, u , to count the number of leaves. Then we have

$$F(z, u) = \frac{zG(z, u)}{1-G(z, u)} \quad \text{and} \quad G(z, u) = zu + \frac{zF(z, u)}{1-F(z, u)},$$

which leads to

$$F(z, u) = \frac{1 - zu - z^2 + z^2u - \sqrt{1 - 2zu - 2z^2 - 2z^2u + z^2u^2 - 2z^3u + 2z^3u^2 + z^4 - 2z^4u + z^4u^2}}{2(1 - zu + z)}.$$

Bringing the irregular beginning also into the game leads to

$$\frac{z}{1-zu} \sum_{r \geq 0} \left(\frac{F}{1-zu} \right)^r + zu - z.$$

This is an ugly expression that we do not display here. However, we can compute the average number of leaves, by differentiation w.r.t. u , followed by setting $u = 1$:

$$R := \frac{v(1+v)(1-v+2v^2-v^3)}{(1-v)(1+v+v^2)}.$$

The coefficient of z^n in this can be expressed in terms of trinomial coefficients, if needed. However, we only compute an asymptotic formula, to keep this section short. Expanding around $v = 1$, we find

$$R \sim \frac{2}{3} \frac{1}{1-v} \sim \frac{2}{3} \frac{1}{\sqrt{3}\sqrt{1-3z}},$$

and thus

$$[z^n]R \sim \frac{2}{3} \frac{1}{\sqrt{3}} 3^n \frac{1}{\sqrt{\pi n}}.$$

We divide this again by

$$\frac{3^{n+\frac{1}{2}}}{2\sqrt{\pi}n^{3/2}}$$

with the result

$$\frac{4}{9}n,$$

which is the asymptotic number of leaves in a Retakh tree of size n . Recall that for unrestricted planar trees, the result is $\frac{n}{2}$, which is a folklore result using Narayana numbers. So the constant in the restricted case, $\frac{4}{9}$, is a bit smaller than $\frac{1}{2}$.

With some effort, more precise approximations could be obtained, as well as the variance. This might be a good project for a student.

10. THE AMPLITUDE OF MOTZKIN PATHS

A Motzkin path consists of up-steps, down-steps, and horizontal steps, see sequence A091965 in [47] and the references given there. As Dyck paths, they start at the origin and end, after n steps again at the x -axis, but are not allowed to go below the x -axis. The height of a Motzkin path is the highest y -coordinate that the path reaches. The average height of such random paths of length n was considered in an early paper [39], it is asymptotically given by $\sqrt{\frac{\pi n}{3}}$.

In the recent paper [6] an interesting new concept was introduced: the *amplitude*. It is basically twice the height, but with a twist. If there exists a horizontal step on level h , which is the height, the amplitude is $2h + 1$, otherwise it is $2h$. To clarify the concept, we created a list of all 9 Motzkin paths of length 4 with height and amplitude given.

The goal of this paper is to investigate this new parameter; in the next section, generating functions will be given, in the following section explicit enumerations, involving trinomial coefficients $\binom{n,3}{k} = [t^k](1+t+t^2)^n$ (notation following Comtet's book [?]). In the last section, the intuitive result that the average amplitude is about twice the

Motzkin path	horizontal on maximal level	height	amplitude
	Yes	0	1
	No	1	2
	No	1	2
	Yes	1	3
	Yes	1	3
	Yes	1	3
	No	1	2
	No	1	2
	No	2	4

average height, is confirmed, and then it will be shown, that, asymptotically, there are about as many Motzkin paths with/without horizontal steps on the maximal level.

Generating functions. Let

$$M^{\leq h}(z) = \sum_{n \geq 0} [\text{number of Motzkin paths of length } n \text{ and height } \leq h] z^n.$$

For the computation, let $f_i = f_i(z)$ be the generating function of Motzkin-like paths, bounded by height h , but ending at level i . Distinguishing the last step, we get

$$f_0 = 1 + z f_0 + z f_1, \quad f_i = z f_{i-1} + z f_i + z f_{i+1} \quad \text{for } 0 < i < h, \quad f_h = z f_{h-1} + z f_h.$$

This system is best written in matrix form:

$$\begin{pmatrix} 1-z & -z & 0 & \dots & & & \\ -z & 1-z & -z & 0 & \dots & & \\ 0 & -z & 1-z & -z & 0 & \dots & \\ \vdots & \ddots & \ddots & & & & \\ & & & & -z & 1-z & \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ f_h \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Let D_n be the determinant of the system matrix, with n rows and columns. Using Cramer's rule to solve a linear system, one finds

$$M^{\leq h}(z) = f_0 = \frac{D_h}{D_{h+1}}.$$

Expanding the determinant along the first row, we get the recursion $D_n = (1-z)D_{n-1} - z^2 D_{n-2}$, and $D_1 = 1 - z$, $D_0 = 1$. Solving,

$$D_n = \frac{1}{\sqrt{1-2z-3z^2}} \left[\left(\frac{1-z+\sqrt{1-2z-3z^2}}{2} \right)^{n+1} - \left(\frac{1-z-\sqrt{1-2z-3z^2}}{2} \right)^{n+1} \right].$$

If one deals with enumeration of Motzkin-like objects, the substitution $z = \frac{v}{1+v+v^2}$ makes the expressions prettier:

$$D_n = \frac{1}{1-v^2} \frac{1-v^{2n+2}}{(1+v+v^2)^n}$$

and further

$$M^{\leq h}(z) = f_0 = \frac{D_h}{D_{h+1}} = (1 + v + v^2) \frac{1 - v^{2h+2}}{1 - v^{2h+4}}.$$

Now let $N^{\leq h}(z)$ be the generating function of Motzkin paths of height $\leq h$, but where horizontal steps on level h are forbidden. The system to compute this is quite similar:

$$\begin{pmatrix} 1-z & -z & 0 & \dots & & \\ -z & 1-z & -z & 0 & \dots & \\ 0 & -z & 1-z & -z & 0 & \dots \\ \ddots & \ddots & \ddots & & & \\ & & & -z & \mathbf{1} & \end{pmatrix} \begin{pmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_h \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

The only difference in the matrix is the entry in the last row, written in boldface. Let D_n^* be the determinant of this system matrix, with n rows and columns. Again,

$$N^{\leq h}(z) = g_0 = \frac{D_h^*}{D_{h+1}^*}.$$

Expanding the matrix along the last row, we find

$$D_n^* = D_{n-1} - z^2 D_{n-2} = \frac{1}{1-v} \frac{1 - v^{2n+1}}{(1 + v + v^2)^n}$$

and

$$N^{\leq h}(z) = g_0 = \frac{D_h^*}{D_{h+1}^*} = (1 + v + v^2) \frac{1 - v^{2h+1}}{1 - v^{2h+3}}.$$

Now let us consider

$$M^{\leq h}(z) - N^{\leq h}(z).$$

There is obviously a lot of cancellation going on. The objects which are still counted, have height $= h$, and *have* horizontal steps on level h . That is one of the two quantities that we wanted to compute, and we get

$$\begin{aligned} \text{Horiz}_h(z) &= (1 + v + v^2) \frac{1 - v^{2h+2}}{1 - v^{2h+4}} - (1 + v + v^2) \frac{1 - v^{2h+1}}{1 - v^{2h+3}} \\ &= (1 + v + v^2)(1 - v^{-2}) \left[\frac{v^{2h+4}}{1 - v^{2h+4}} - \frac{v^{2h+3}}{1 - v^{2h+3}} \right]. \end{aligned}$$

Similarly, considering $N^{\leq h}(z) - M^{\leq h-1}(z)$, we find that only objects are counted that have height $= h$, and *no* horizontal steps on level h . Thus

$$\begin{aligned} \text{No-Horiz}_h(z) &= (1 + v + v^2) \left[\frac{1 - v^{2h+1}}{1 - v^{2h+3}} - \frac{1 - v^{2h}}{1 - v^{2h+2}} \right] \\ &= (1 + v + v^2)(1 - v^{-2}) \left[\frac{v^{2h+3}}{1 - v^{2h+3}} - \frac{v^{2h+2}}{1 - v^{2h+2}} \right]. \end{aligned}$$

As a check, we get

$$\begin{aligned} \text{Horiz}_h(z) + \text{No-Horiz}_h(z) &= (1 + v + v^2)(1 - v^{-2}) \left[\frac{v^{2h+4}}{1 - v^{2h+4}} - \frac{v^{2h+2}}{1 - v^{2h+2}} \right] \\ &= M^{\leq h}(z) - M^{\leq h-1}(z) = M^=h(z). \end{aligned}$$

Explicit enumerations. All our generating functions contain the term

$$(1 + v + v^2)(1 - v^{-2}) \frac{v^{2h+a}}{1 - v^{2h+a}} = (1 + v + v^2)(1 - v^{-2}) \sum_{k \geq 1} v^{k(2h+a)}$$

for various values of a . We show how to compute the coefficient of z^n in this. It will be done using contour integration. The contour is a small circle in the z -plane or v -plane.

$$\begin{aligned} [z^n](1 + v + v^2)(1 - v^{-2}) \frac{v^{2h+a}}{1 - v^{2h+a}} &= \frac{1}{2\pi i} \oint \frac{dz}{z^{n+1}} (1 + v + v^2)(1 - v^{-2}) \sum_{k \geq 1} v^{k(2h+a)} \\ &= \frac{1}{2\pi i} \oint \frac{dv(1 - v^2)}{(1 + v + v^2)^2} \frac{(1 + v + v^2)^{n+1}}{v^{n+1}} (1 + v + v^2)(1 - v^{-2}) \sum_{k \geq 1} v^{k(2h+a)} \\ &= - \sum_{k \geq 1} [v^{n+2-k(2h+a)}] (1 - v^2)^2 (1 + v + v^2)^n \\ &= - \sum_{k \geq 1} \left[\binom{n, 3}{n+2-k(2h+a)} - 2 \binom{n, 3}{n-k(2h+a)} + \binom{n, 3}{n-2-k(2h+a)} \right]. \end{aligned}$$

In this computation, we used again the notion of *trinomial* coefficients:

$$\binom{n, 3}{k} = [t^k](1 + t + t^2)^n.$$

The average amplitude. Here we compute:

$$\begin{aligned} &\sum_{h \geq 0} (2h+1) \text{Horiz}_h(z) + \sum_{h \geq 1} (2h) \text{No-Horiz}_h(z) \\ &= (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} (2h+1) \left[\frac{v^{2h+4}}{1 - v^{2h+4}} - \frac{v^{2h+3}}{1 - v^{2h+3}} \right] \\ &\quad + (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} (2h) \left[\frac{v^{2h+3}}{1 - v^{2h+3}} - \frac{v^{2h+2}}{1 - v^{2h+2}} \right] \\ &= -(1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} \frac{v^{2h+3}}{1 - v^{2h+3}} \\ &\quad + (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} (2h+1) \frac{v^{2h+4}}{1 - v^{2h+4}} \\ &\quad - (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} (2h+2) \frac{v^{2h+4}}{1 - v^{2h+4}} \\ &= -(1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} \frac{v^{2h+3}}{1 - v^{2h+3}} - (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 1} \frac{v^{2h+2}}{1 - v^{2h+2}} \\ &= -(1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 0} \frac{v^{2h+1}}{1 - v^{2h+1}} + (1 + v + v^2)(1 - v^{-2}) \frac{v}{1 - v} \\ &\quad - (1 + v + v^2)(1 - v^{-2}) \sum_{h \geq 1} \frac{v^{2h}}{1 - v^{2h}} + (1 + v + v^2)(1 - v^{-2}) \frac{v^2}{1 - v^2} \end{aligned}$$

$$\begin{aligned}
&= -(1+v+v^2)(1-v^{-2}) \sum_{h \geq 0} \frac{v^{2h+1}}{1-v^{2h+1}} \\
&- (1+v+v^2)(1-v^{-2}) \sum_{h \geq 1} \frac{v^{2h}}{1-v^{2h}} - \frac{(1+2v)(1+v+v^2)}{v} \\
&= -(1+v+v^2)(1-v^{-2}) \sum_{h \geq 1} \frac{v^h}{1-v^h} - \frac{(1+2v)(1+v+v^2)}{v}.
\end{aligned}$$

To find asymptotics from here, we need the local expansion of this generating function around $v \sim 1$, or, equivalently, $z \sim \frac{1}{3}$. See [13] for more explanations of how this method works; it also involves singularity analysis of generating functions [14]. While this combined approach of Mellin transform and singularity analysis has been used for more than 30 years, we would like to cite [21], where many technical details have been worked out in a similar instance. For example, the expansion of the sum that we need is derived there. We combine all the expansions:

$$\left(6(1-v) + \dots\right) \left(-\frac{\log(1-v)}{1-v} + \frac{\gamma}{1-v} + \dots\right) - 9 + 6(1-v) + \dots = -6 \log(1-v) + \dots$$

Translating this into the z -world means $1-v \sim \sqrt{3}\sqrt{1-3z}$, and then

$$[z^n] \left(-6 \log(1-v)\right) \sim [z^n] \left(-6 \log \sqrt{1-3z}\right) = 3 \frac{3^n}{n} = \frac{3^{n+1}}{n}.$$

The total number of Motzkin paths is

$$[z^n] \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2} \sim [z^n] \left(3-3\sqrt{3}\sqrt{1-3z}\right) \sim 3\sqrt{3} \frac{3^n}{2\sqrt{\pi n^{3/2}}}.$$

Taking quotients, we get the average amplitude of random Motzkin paths of length n , which is asymptotic to

$$2\sqrt{\frac{\pi n}{3}}.$$

This is about twice as much as the average height of random Motzkin paths, which is what we expected.

Now we consider

$$\begin{aligned}
\sum_{h \geq 0} \text{No-Horiz}_h(z) &= (1+v+v^2)(1-v^{-2}) \sum_{h \geq 0} \left[\frac{v^{2h+3}}{1-v^{2h+3}} - \frac{v^{2h+2}}{1-v^{2h+2}} \right] \\
&= (1+v+v^2)(1-v^{-2}) \sum_{h \geq 1} \left[\frac{v^{2h-1}}{1-v^{2h-1}} - \frac{v^{2h}}{1-v^{2h}} \right] + \frac{(1+v)(1+v+v^2)}{v} \\
&= (1+v+v^2)(1-v^{-2}) \sum_{h \geq 1} \frac{v^h}{1-v^h} (-1)^{h-1} + \frac{(1+v)(1+v+v^2)}{v}.
\end{aligned}$$

This needs to be expanded about $v = 1$:

$$\begin{aligned}
(1+v+v^2)(1-v^{-2}) &\sim -6(1-v) - 3(1-v)^2 + \dots; \\
\frac{(1+v)(1+v+v^2)}{v} &\sim 6 - 3(1-v) + \dots;
\end{aligned}$$

for the remaining sum we need the Mellin transform [13]. Set $v = e^{-t}$, and transform

$$\mathcal{M} \sum_{h \geq 1} \frac{v^h}{1 - v^h} (-1)^{h-1} = \mathcal{M} \sum_{h, k \geq 1} e^{-thk} (-1)^{h-1} = \Gamma(s) \zeta(s)^2 (1 - 2^{1-s}).$$

By the Mellin inversion formula:

$$\sum_{h \geq 1} \frac{v^h}{1 - v^h} (-1)^{h-1} = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s) \zeta(s)^2 (1 - 2^{1-s}) t^{-s} ds.$$

The line of integration will be shifted to the left, and the collected residues constitute the expansion about $t = 0$:

$$\sum_{h \geq 1} \frac{v^h}{1 - v^h} (-1)^{h-1} \sim \frac{\log 2}{t} - \frac{1}{4} + \frac{t}{48} \sim \frac{\log 2}{1-v} - \frac{\log 2}{2} - \frac{1}{4} + \dots$$

Combining,

$$\sum_{h \geq 0} \text{No-Horiz}_h(z) \sim 6 - 6 \log 2 - \frac{3}{2}(1-v) + \dots$$

But

$$\frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2} = 1 + v + v^2 \sim 3 - 3(1-v) + \dots$$

Comparing the coefficients of $1 - v$, we find that asymptotically about half of the Motzkin paths belong to the ‘No-horizontal’ and about half of the Motzkin paths belong to the ‘horizontal’ class. Again, this is intuitively clear, once one sees the generating functions of both families.

11. OSCILLATIONS IN DYCK PATHS REVISITED

Rainer Kemp’s paper [22] was unfortunately largely overlooked. An extension was published quickly [25], and then it fell into oblivion. We want to come back to this gem, with modern methods, in particular, the kernel method and singularity analysis. Kemp was interested in peaks and valleys of Dyck paths, which he called MAX-turns and MIN-turns, probably motivated by Computer Science applications. The peaks/valleys are enumerated from left to right, and the interest is the height of them, for a fixed number and the length of the Dyck path going to infinity. In the corresponding ordered (plane) tree, the peaks correspond to the leaves.

Very precise information is available for leaves of binary trees [24, 19, 28, 29] but the situation is a bit different for Dyck paths since the number of peaks/valleys isn’t directly related to the length of the Dyck path. (Naryana numbers enumerate them.) Kemp’s results in a nutshell are: The average height of the m -th peak/valley tends to a constant for n going to infinity (the length of the random Dyck path). This constant is $\sim 4\sqrt{2m/\pi}$, and the difference between the height of the peak and the valley is about 2, with more terms being available in principle.

A trivariate generating function for heights of valleys. The goal is to derive an expression for $\Phi(u) = \Phi(u; z, w)$, where z is used for the length of the path, w for the enumeration of the *valleys* (w^m corresponds to the m -th valley), and u is used to record the height of the m -th (and last) valley of a partial Dyck path (the path does not need to return to the x -axis). We could think about it continued in any possible fashion, as in the following figure.

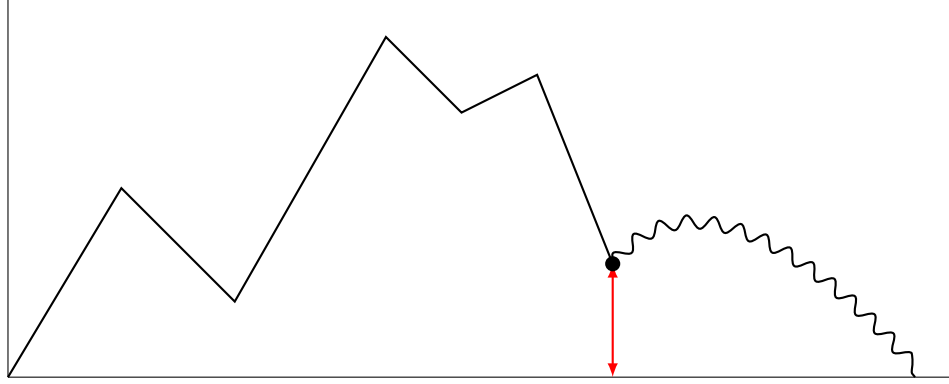


FIGURE 11. The third valley at level j .

Our goal is, as often, to use the adding-a-new-slice technique, namely adding another ‘mountain’ (a maximal sequence of up-steps, followed by a maximal sequence of down-steps), or going from the m -th valley to the $(m + 1)$ -st valley.

We investigate what can happen to u^j :

$$\sum_{l \geq 1} \sum_{i=1}^{j+l} z^l u^{j+l} z^i u^{-i}.$$

Working this out, the following substitution is essential for our problem:

$$u^j \longrightarrow \frac{z^2 u^k}{(u-z)(1-zu^k)} u^j - \frac{z^{k+2}}{(u-z)(1-z^{k+1})} z^j.$$

Working this into a generating function of the type

$$\Phi(u) = \sum_{m \geq 0} w^m \varphi_m(u),$$

where the variable w is used to keep track of the number of mountains, we find from the substitution

$$\Phi(u) = 1 + \frac{wz^2 u}{(u-z)(1-zu)} \Phi(u) - \frac{wz^3}{(u-z)(1-z^2)} \Phi(z),$$

where 1 stands for the empty path having no mountains. Rearranging,

$$\Phi(u) \frac{z(u-s_1)(u-s_2)}{(u-z)(zu-1)} = 1 - \frac{wz^3}{(u-z)(1-z^2)} \Phi(z),$$

and

$$\Phi(u) = \frac{(zu-1)}{z(u-s_1)(u-s_2)} \left[u-z - \frac{wz^3}{(1-z^2)} \Phi(z) \right].$$

Here,

$$s_2 = \frac{z^2 + 1 - wz^2 - \sqrt{z^4 - 2z^2 - 2z^4w + 1 - 2wz^2 + w^2z^4}}{2z} \quad \text{and} \quad s_1 = \frac{1}{s_2}.$$

In the spirit of the kernel method, the factor $u - s_2$ is ‘bad’ and must cancel out. That leads first to

$$\Phi(z) = \frac{(1 - z^2)(s_2 - z)}{wz^3}$$

and further to

$$\Phi(u) = \frac{(zu - 1)}{z(u - s_1)} = \frac{s_2(1 - zu)}{z(1 - us_2)}.$$

From this it is easy to read off coefficients:

$$[u^j]\Phi(u) = [u^j]\frac{s_2(1 - zu)}{z(1 - us_2)} = \frac{1}{z}s_2^{j+1} - s_2^j.$$

Note that setting $w = 1$ ignores the number of mountains, and the generating function would then be enumerating partial Dyck paths ending on level j with a down-step. The answer could then be derived by combinatorial means as well.

For Kemp’s problem, we need

$$S = \sum_{j \geq 0} j \left(\frac{1}{z}s_2^{j+1} - s_2^j \right) \cdot \left(\frac{1}{z}s_2^{j+1} - s_2^j \right) \Big|_{w=1}.$$

The factor j comes in because of the average value of the height of the valley, the first factor is what we just worked out, and the third factor is the rest, which, when read from right to left, is just what we discussed, since the number of mountains in the rest is irrelevant. Thanks to Computer Algebra (not available when Kemp worked on the oscillations), we get

$$S = 4 \frac{(-3z + W_1z - W_1 + 1)(-W_2 + wzW_2 + 1 + z^2w^2 - wz^2 - 2wz - z)}{z(-3z - W_1z + 1 - W_1 - wz + wzW_1 - W_2 + W_2W_1)^2}$$

with

$$W_1 = \sqrt{1 - 4z} \quad \text{and} \quad W_2 = \sqrt{z^2 - 2z - 2z^2w + 1 - 2wz + w^2z^2}.$$

Note carefully that z^2 was replaced by z , since Dyck paths (returning to the x -axis) must have an even number of steps. Their enumeration is classical:

$$D(z) = \frac{1 - \sqrt{1 - 4z}}{2z} \sim 2 - 2\sqrt{1 - 4z},$$

for z close to the (dominant) singularity $z = \frac{1}{4}$. We are in the regime of the subcritical case ([15], Section IX-3). The function S has a similar local expansion:

$$S \sim \text{constant}_1 - \text{constant}_2\sqrt{1 - 4z},$$

and the function $\frac{-\text{constant}_2}{-2}$ is the resulting generating function. Working out the details,

$$S \sim \frac{w + \sqrt{(1-w)(9-w)} - 3}{-1 + w} - \sqrt{1 - 4z} \left(\frac{w^2 + 2w - 3 + (1+w)\sqrt{(1-w)(9-w)}}{(1-w)^2} \right) + \dots$$

Eventually we are led to

$$\text{Valley}(w) := \frac{w^2 + 2w - 3 + (1+w)\sqrt{(1-w)(9-w)}}{2(1-w)^2}.$$

To say it again, the coefficient of w^m in this is the average value of the m -th valley in ‘very long’ Dyck path. To say more about it, we can use singularity analysis again and expand (this time around $w = 1$, which is dominant):

$$\text{Valley}(w) \sim \frac{2\sqrt{2}}{(1-w)^{3/2}} - \frac{2}{1-w} - \frac{7}{8} \frac{\sqrt{2}}{\sqrt{1-w}}.$$

The traditional translation theorems [14] lead to the average value of the height of the m -th valley:

$$4\sqrt{2}\sqrt{\frac{m}{\pi}} - 2 + \frac{5\sqrt{2}}{8\sqrt{\pi m}} + \dots$$

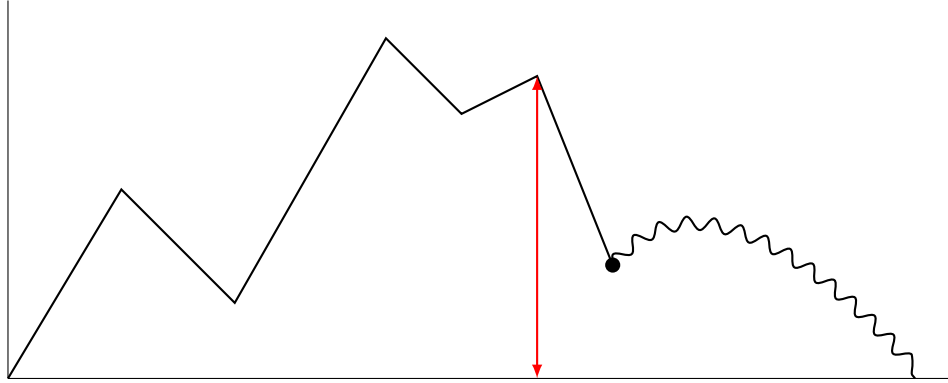


FIGURE 12. The third peak at level j .

From valleys to peaks. We don’t need too many new computations, as we can modify the previous results. If one adds an arbitrary non-empty number of up-steps after the m -th valley, one has reached the $(m+1)$ -st peak! This is basically a substitution! Start from

$$\Phi(u) = \frac{s_2(1-zu)}{z(1-us_2)}$$

and attach a sequence of up-steps: $u^j \rightarrow \frac{zu}{1-zu}u^j$. A factor w is also important, since the m -th valley corresponds to the $(m+1)$ -st peak. Now

$$\frac{zww}{1-zu} \frac{s_2(1-zu)}{z(1-us_2)} = \frac{us_2w}{1-us_2} = w \sum_{j \geq 1} u^j s_2^j.$$

The computation

$$w \sum_{j \geq 0} j s_2^j \cdot s_2^j \Big|_{w=1}$$

was basically done before, and the local expansion leads to

$$\frac{2w}{1-w} - \frac{2w\sqrt{(1-w)(9-w)}}{(1-w)^2} \sqrt{1-4z},$$

and the generating function of the average values of the m -th peak is

$$\text{Peak}(w) = \frac{w\sqrt{(1-w)(9-w)}}{(1-w)^2}.$$

A local expansion of this results in

$$\text{Peak}(w) \sim \frac{2\sqrt{2}}{(1-w)^{3/2}} - \frac{15}{8} \frac{\sqrt{2}}{\sqrt{1-w}}.$$

Taking differences:

$$\text{Peak}(w) - \text{Valley}(w) \sim \frac{2}{1-w} - \frac{\sqrt{2}}{\sqrt{1-w}},$$

and translating into asymptotics:

$$2 - \frac{\sqrt{2}}{\sqrt{\pi m}}.$$

The formula $2 + O(m^{-1/2})$ was already known to Kemp [22]. As Kemp stated in [22], which was confirmed in [25], the generating functions $\text{Peak}(w)$ and $\text{Valley}(w)$ can be expressed by Legendre polynomials at special values. This is a bit artificial and not too useful in itself.

12. DEUTSCH-PATHS IN A STRIP

Emeric Deutsch [41] had the idea to consider a variation of ordinary Dyck paths, by augmenting the usual up-steps and down-steps by one unit each, by down-steps of size $3, 5, 7, \dots$. This leads to ternary equations, as can be seen for instance from [43].

The present author started to investigate a related but simpler model of down-steps $1, 2, 3, 4, \dots$ and investigated it (named Deutsch paths in honour of Emeric Deutsch) in a series of papers, [42, 44, 45].

This section is a further member of this series: The condition that (as with Dyck paths) the paths cannot enter negative territory, is relaxed, by introducing a negative boundary $-t$. Here are two recent publications about such a negative boundary: [46] and [31].

Instead of allowing negative altitudes, we think about the whole system shifted up by t units, and start at the point $(0, t)$ instead. This is much better for the generating functions that we are going to investigate. Eventually, the results can be re-interpreted as results about enumerations with respect to a negative boundary.

The setting with flexible initial level t and final level j allows us to consider the Deutsch paths also from right to left (they are not symmetric!), without any new computations.

The next sections achieves this, using the celebrated kernel-method, one of the tools that is dear to our heart [35].

In the following section, an additional upper bound is introduced, so that the Deutsch paths live now in a strip. The way to attack this is linear algebra. Once everything has been computed, one can relax the conditions and let lower/upper boundary go to $\mp\infty$.

Generating functions and the kernel method. As discussed, we consider Deutsch paths starting at $(0, t)$ and ending at (n, j) , for $n, t, j \geq 0$. First we consider univariate generating functions $f_j(z)$, where z^n stays for n steps done, and j is the final destination. The recursion is immediate:

$$f_j(z) = \llbracket t = j \rrbracket + z f_{j-1}(z) + z \sum_{k>j} f_k(z),$$

where $f_{-1}(z) = 0$. Next, we consider

$$F(z, u) := \sum_{j \geq 0} f_j(z) u^j,$$

and get

$$\begin{aligned} F(z, u) &= u^t + zuF(z, u) + z \sum_{j \geq 0} u^j \sum_{k>j} f_k(z) \\ &= u^t + zuF(z, u) + z \sum_{k>0} f_k(z) \sum_{0 \leq j < k} u^j \\ &= u^t + zuF(z, u) + z \sum_{k \geq 0} f_k(z) \frac{1 - u^k}{1 - u} \\ &= u^t + zuF(z, u) + \frac{z}{1 - u} [F(z, 1) - F(z, u)] \\ &= \frac{u^t(1 - u) + zF(z, 1)}{z - zu + zu^2 + 1 - u}. \end{aligned}$$

Since the critical value is around $u = 1$, we write the denominator as

$$z(u - 1)^2 + (u - 1)(z - 1) + z = z(u - 1 - r_1)(u - 1 - r_2),$$

with

$$r_1 = \frac{1 - z + \sqrt{1 - 2z - 3z^2}}{2z}, \quad r_2 = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z}.$$

The factor $(u - 1 - r_2)$ is bad, so the numerator must vanish for $[u^t(1 - u) + zF(z, 1)]|_{u=1+r_2}$, therefore

$$zF(z, 1) = (1 + r_2)^t r_2.$$

Furthermore

$$F(z, u) = \frac{\frac{u^t(1-u) + zF(z,1)}{u - r_2}}{z(u - r_1)}.$$

The expressions become prettier using the substitution $z = \frac{v}{1+v+v^2}$; then

$$r_1 = \frac{1}{v}, \quad r_2 = v.$$

It can be proved by induction (or computer algebra) that

$$\frac{u^t(1 - u) + v(1 + v)^t}{u - 1 - v} = -v \sum_{k=0}^{t-1} (1 + v)^{t-1-k} - u^t.$$

Furthermore

$$\frac{1}{z(u - 1 - r_1)} = -\frac{1}{z(1 + r_1)(1 - \frac{u}{1+r_1})},$$

and so

$$f_j(z) = [u^j]F(z, u) = [u^j] \left[v \sum_{k=0}^{t-1} (1+v)^{t-1-k} u^k + u^t \right] \sum_{\ell \geq 0} \frac{u^\ell}{z(1+r_1)^{\ell+1}}.$$

Of interest are two special cases: The case that was studied before [42] is $t = 0$:

$$f_j = \frac{(1+v+v^2)v^j}{(1+v)^{j+1}}.$$

The other special case is $j = 0$ for general t , as it may be interpreted as Deutsch paths read from right to left, starting at level 0 and ending at level $t \geq 1$ (for $t = 0$, the previous formula applies):

$$\begin{aligned} f_0(z) &= [u^0] \left[v \sum_{k=0}^{t-1} (1+v)^{t-1-k} u^k + u^t \right] \sum_{\ell \geq 0} \frac{u^\ell}{z(1+r_1)^{\ell+1}} \\ &= v(1+v)^{t-1} \frac{1}{z(1+r_1)} = v(1+v+v^2)(1+v)^{t-2}. \end{aligned}$$

The next section will present a simplification of the expression for $f_j(z)$, which could be obtained directly by distinguishing cases and summing some geometric series.

Refined analysis: lower and upper boundary. Now we consider Deutsch paths bounded from below by zero and bounded from above by $m - 1$; they start at level t and end at level j after n steps. For that, we use generating functions $\varphi_j(z)$ (the quantity t is a silent parameter here). The recursions that are straight-forwarded are best organized in a matrix, as the following example shows.

$$\begin{pmatrix} 1 & -z & -z & -z & -z & -z & -z & -z \\ -z & 1 & -z & -z & -z & -z & -z & -z \\ 0 & -z & 1 & -z & -z & -z & -z & -z \\ 0 & 0 & -z & 1 & -z & -z & -z & -z \\ 0 & 0 & 0 & -z & 1 & -z & -z & -z \\ 0 & 0 & 0 & 0 & -z & 1 & -z & -z \\ 0 & 0 & 0 & 0 & 0 & -z & 1 & -z \\ 0 & 0 & 0 & 0 & 0 & 0 & -z & 1 \end{pmatrix} \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \varphi_4 \\ \varphi_5 \\ \varphi_6 \\ \varphi_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Bigg\} t$$

The goal is now to solve this system. For that the substitution $z = \frac{v}{1+v+v^2}$ is used throughout. The method is to use Cramer's rule, which means that the right-hand side has to replace various columns of the matrix, and determinants have to be computed. At the end, one has to divide by the determinant of the system.

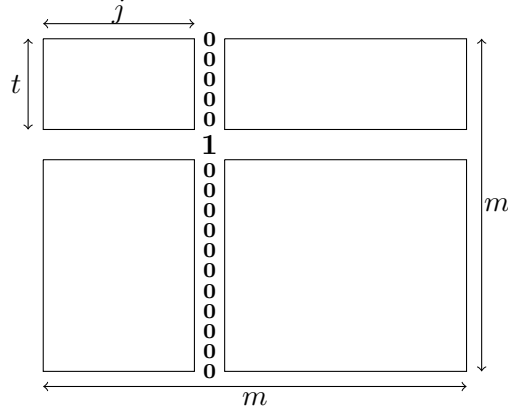
Let D_m be the determinant of the matrix with m rows and columns. The recursion

$$(1+v+v^2)^2 m_{n+2} - (1+v+v^2)(1+v)^2 D_{m+1} + v(1+v)^2 D_m = 0$$

appeared already in [42] and is not difficult to derive and to solve:

$$D_m = \frac{(1+v)^{m-1}}{(1+v+v^2)^m} \frac{1-v^{m+2}}{1-v}.$$

To solve the system with Cramer's rule, we must compute a determinant of the following type,



where the various rows are replaced by the right-hand side. While it is not impossible to solve this recursion by hand, it is very easy to make mistakes, so it is best to employ a computer. Let $D(m; t, j)$ the determinant according to the drawing.

It is not unexpected that the results are different for $j < t$ resp. $j \geq t$. Here is what we found:

$$D(m; t, j) = \frac{(1+v)^{t-j-3+m}(1-v^{j+1})v(1-v^{m-t})}{(1-v)^2(1+v+v^2)^{m-1}}, \quad \text{for } j < t,$$

$$D(m; t, j) = \frac{v^{j-t}(1-v^{t+2})(1-v^{1-j+m})}{(1-v)^2(1+v+v^2)^{m-1}(1+v)^{j-t+3-m}}, \quad \text{for } j \geq t.$$

To solve the system, we have to divide by the determinant D_m , with the result

$$\varphi_j = \frac{D(m; t, j)}{D_m} = \frac{(1+v)^{t-j-2}(1-v^{j+1})v(1-v^{m-t})(1+v+v^2)}{(1-v)(1-v^{m+2})}, \quad \text{for } j < t,$$

$$\varphi_j = \frac{D(m; t, j)}{D_m} = \frac{v^{j-t}(1-v^{t+2})(1-v^{1-j+m})(1+v+v^2)}{(1-v)(1+v)^{j-t+2}(1-v^{m+2})}, \quad \text{for } j \geq t.$$

We found all this using Computer algebra. Some critical minds may argue that this is only experimental. One way of rectifying this would be to show that indeed the functions φ_j solve the system, which consists of summing various geometric series; again, a computer could be helpful for such an enterprise.

Of interest are also the limits for $m \rightarrow \infty$, i.e., no upper boundary:

$$\varphi_j = \lim_{m \rightarrow \infty} \frac{D(m; t, j)}{D_m} = \frac{(1+v)^{t-j-2}(1-v^{j+1})v(1+v+v^2)}{(1-v)}, \quad \text{for } j < t,$$

$$\varphi_j = \frac{v^{j-t}(1-v^{t+2})(1+v+v^2)}{(1-v)(1+v)^{j-t+2}}, \quad \text{for } j \geq t.$$

The special case $t = 0$ appeared already in the previous section:

$$\varphi_j = \frac{v^j(1+v+v^2)}{(1+v)^{j+1}}.$$

Likewise, for $t \geq 1$,

$$\varphi_0 = v(1+v+v^2)(1+v)^{t-2}.$$

In particular, the formulæ show that the expression from the previous section can be simplified in general, which could have been seen directly, of course.

Theorem 4. *The generating function of Deutsch path with lower boundary 0 and upper boundary $m - 1$, starting at $(0, t)$ and ending at (n, j) is given by*

$$\frac{(1+v)^{t-j-2}(1-v^{j+1})v(1-v^{m-t})(1+v+v^2)}{(1-v)(1-v^{m+2})}, \quad \text{for } j < t,$$

$$\frac{v^{j-t}(1-v^{t+2})(1-v^{1-j+m})(1+v+v^2)}{(1-v)(1+v)^{j-t+2}(1-v^{m+2})}, \quad \text{for } j \geq t,$$

with the substitution $z = \frac{v}{1+v+v^2}$.

By shifting everything down, we can interpret the results as Deutsch walks between boundaries $-t$ and $m - 1 - t$, starting at the origin $(0, 0)$ and ending at $(n, j - t)$.

Theorem 5. *The generating function of Deutsch path with lower boundary $-t$ and upper boundary h , starting at $(0, 0)$ and ending at (n, i) with $-t \leq i \leq h$ is given by*

$$\frac{(1+v)^{i-2}(1-v^{i+t+1})v(1-v^{h+1})(1+v+v^2)}{(1-v)(1-v^{h+t+3})}, \quad \text{for } i < 0,$$

$$\frac{v^i(1-v^{t+2})(1-v^{2-i+h})(1+v+v^2)}{(1-v)(1+v)^{i+2}(1-v^{h+t+3})}, \quad \text{for } i \geq 0.$$

It is possible to consider the limits $t \rightarrow \infty$ and/or $h \rightarrow \infty$ resulting in simplified formulæ.

Theorem 6. *The generating function of Deutsch path with lower boundary $-t$ and upper boundary ∞ , starting at $(0, 0)$ and ending at (n, i) with $-t \leq i$ is given by*

$$\frac{(1+v)^{i-2}(1-v^{i+t+1})v(1+v+v^2)}{(1-v)}, \quad \text{for } i < 0,$$

$$\frac{v^i(1-v^{t+2})(1+v+v^2)}{(1-v)(1+v)^{i+2}}, \quad \text{for } i \geq 0.$$

Theorem 7. *The generating function of Deutsch path with lower boundary $-\infty$ and upper boundary h , starting at $(0, 0)$ and ending at (n, i) with $\leq i \leq h$ is given by*

$$\frac{(1+v)^{i-2}v(1-v^{h+1})(1+v+v^2)}{(1-v)}, \quad \text{for } i < 0,$$

$$\frac{v^i(1-v^{2-i+h})(1+v+v^2)}{(1-v)(1+v)^{i+2}}, \quad \text{for } i \geq 0.$$

Theorem 8. *The generating function of unbounded Deutsch path starting at $(0, 0)$ and ending at (n, i) is given by*

$$\frac{(1+v)^{i-2}v(1+v+v^2)}{(1-v)}, \quad \text{for } i < 0,$$

$$\frac{v^i(1+v+v^2)}{(1-v)(1+v)^{i+2}}, \quad \text{for } i \geq 0.$$

13. PUBLICATION STATUS OF THE PROBLEMS TREATED IN THIS WALK IN THE GARDEN

- The section on Hoppy walks is new; an earlier version appeared as arXiv:2009.13474
- The section on Combinatorics of sequence A002212 is new; an earlier version appeared as arXiv:2106.14782
- The section on Binary trees and Horton-Strahler numbers is classical; it was included since it can be used almost verbatim for unary-binary trees when employing the proper substitution
- The section on marked ordered trees is new and basically included in arXiv:2106.14782
- The bijection between multi-edge trees and 3-coloured Motzkin paths is new; early version in arXiv:2105.03350
- The section on combinatorics of skew Dyck paths was submitted to journal, and no answer was received; current version in arXiv:2108.09785
- The section on oscillations in honour of Rainer Kemp was written for this survey.
- The short analysis of Retakh's Motzkin paths appeared in ECA.
- The amplitude of Motzkin paths was submitted to a journal; no answer.
- The enumeration of Deutsch-paths in a strip was exclusively written for this survey article; arXiv:2108.12797

REFERENCES

- [1] Andrei Asinowski, Benjamin Hackl, Sarah J. Selkirk, Down-step statistics in generalized Dyck path, arXiv:2007.15562, 2020.
- [2] Jean-Luc Baril, José L. Ramírez, and Lina M. Simbaqueba. Counting Prefixes of Skew Dyck Paths. *Journal of integer sequences*, Article 21.8.2, 2021.
- [3] M. Bousquet-Mélou and M. Petkovšek. Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, 307(2):257–276, 2003. in: Random generation of combinatorial objects and bijective combinatorics.
- [4] Alexander Burstein. The Distribution of peak heights modulo k and double descents on k -Dyck paths-project. arXiv:2009.00760 2020.
- [5] Louis Comtet. *Advanced combinatorics*. D. Reidel Publishing Co., Dordrecht, enlarged edition, 1974. The art of finite and infinite expansions.
- [6] Julien Courtiel, Andrew Elvey Price, and Irène Marcovici. Bijections between walks inside a triangular domain and Motzkin paths of bounded amplitude. *The electronic journal of combinatorics*, 28(2)(2):#P2.6h, 2021.
- [7] Eva Deng and Toufik Mansour, Three Hoppy path problems and ternary paths, *Discrete Applied Mathematics* 156, 2008, 770–779.
- [8] N. G. De Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In R. C. Read, editor, *Graph Theory and Computing*, pages 15–22. Academic Press, 1972.
- [9] E. Deutsch, E. Munarini, S. Rinaldi. Skew Dyck paths. *J. Stat. Plann. Infer.*, 140 (8) (2010) 2191–2203.
- [10] Emeric Deutsch and Louis W. Shapiro. A bijection between ordered trees and 2-Motzkin paths and its many consequences. volume 256, pages 655–670. 2002. LaCIM 2000 Conference on Combinatorics, Computer Science and Applications (Montreal, QC).
- [11] M. Dziemiańczuk. Enumerations of plane trees with multiple edges and Raney lattice paths. *Discrete Math.*, 337:9–24, 2014.
- [12] Shalosh B. Ekhad, Doron Zeilberger, Automatic Counting of Restricted Dyck Paths via (Numeric and Symbolic) Dynamic Programming arXiv:2006.01961, 2020
- [13] Philippe Flajolet, Xavier Gourdon, and Philippe Dumas. Mellin transforms and asymptotics: harmonic sums. volume 144, pages 3–58. 1995. Special volume on mathematical analysis of algorithms.

- [14] P. Flajolet and A. M. Odlyzko. Singularity analysis of generating functions. 3:216–240, 1990.
- [15] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [16] P. Flajolet and H. Prodinger, Level number sequences for trees, *Discrete Mathematics* 65, 1987, 149–156.
- [17] P. Flajolet and H. Prodinger. Register allocation for unary-binary trees. *SIAM J. Comput.*, 15(3):629–640, 1986.
- [18] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, MA, 1999.
- [19] W. Gutjahr and G. C. Pflug. The asymptotic distribution of leaf heights in binary trees. *Graphs Combin.*, 8(3):243–251, 1992.
- [20] Benjamin Hackl, Clemens Heuberger, H. Prodinger, Ascents in Non-Negative Lattice Paths, arXiv:1801.02996, 2018.
- [21] C. Heuberger, H. Prodinger, and S. Wagner. The height of multiple edge plane trees. *Aequationes Math.*, 90(3):625–645, 2016.
- [22] R. Kemp. On the average oscillation of a stack. *Combinatorica*, 2(2):157–176, 1982.
- [23] H. Kim and R. P. Stanley. A refined enumeration of hex trees and related polynomials. *European J. Combin.*, 54:207–219, 2016.
- [24] Peter Kirschenhofer. On the height of leaves in binary trees. *J. Combin. Inform. System Sci.*, 8(1):44–60, 1983.
- [25] Peter Kirschenhofer and Helmut Prodinger. On the average hyperoscillations of planted plane trees. *Combinatorica*, 2(2):177–186, 1982.
- [26] D. Knuth. Donald Knuth’s 20th Annual Christmas Tree Lecture: (3/2)-ary Trees. <https://youtu.be/P4AaGQlo0HY>, Dec 2014.
- [27] T. V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, University of Toronto Press, 100–101 (1979).
- [28] Alois Panholzer and Helmut Prodinger. Descendants and ascendants in binary trees. *Discrete Math. Theor. Comput. Sci.*, 1(1):247–266, 1997.
- [29] Alois Panholzer and Helmut Prodinger. Moments of level numbers of leaves in binary trees. volume 101, pages 267–279. 2002. Special issue on lattice path combinatorics and applications (Vienna, 1998).
- [30] H. Prodinger. Analytic methods. In *Handbook of enumerative combinatorics*, Discrete Math. Appl. (Boca Raton), pages 173–252. CRC Press, Boca Raton, FL, 2015.
- [31] H. Prodinger, On k -Dyck paths with a negative boundary. *arXiv:1912.06930*
- [32] H. Prodinger. Some analytic techniques for the investigation of the asymptotic behaviour of tree parameters. *EATCS Bulletin*, 47 (1992), 180-199.
- [33] H. Prodinger. Weighted unary-binary trees, Hex-trees, marked ordered trees, and related structures. *arXiv:2106.14782*, 2021.
- [34] H. Prodinger. On some problems by Cameron about ternary paths — a linear algebra approach. *Rocky Mountain Journal of Mathematics*, 51 (2021), No. 2, 709–720.
- [35] H. Prodinger. *The Kernel Method: A Collection of Examples*, Séminaire Lotharingien de Combinatoire, B50f (2004), 19 pp.
- [36] H. Prodinger. Philippe Flajolet’s early work in combinatorics. 2:1 (2022) Article S1H1.:11 pages, 2022.
- [37] H. Prodinger. Introduction to Philippe Flajolet’s work on the register function and related topics. <https://www.math.tugraz.at/~prodinger/flajolet-introduction-register.pdf>, 7 pages, 2025.
- [38] H. Prodinger. The height of planted plane trees revisited. *Ars Combinatoria*, 16-B:51–55, 1983.
- [39] H. Prodinger. The average height of a stack where three operations are allowed and some related problems. *J. Combin. Inform. System Sci.*, 5(4):287–304, 1980.
- [40] H. Prodinger, Deutsch paths and their enumeration, submitted (2020). NEW citation.
- [41] Emeric Deutsch. Problem 10751. *American Mathematical Monthly*, 107 (2000); solution *American Mathematical Monthly*, 108 (2001).
- [42] Helmut Prodinger. Deutsch paths and their enumeration, *Open J. Discret. Appl. Math.* 2021, 4(1), 12–18.

- [43] Helmut Prodinger. Generating functions for a lattice path model introduced by Deutsch, *Special Matrices* 2021, 9: 217–225.
- [44] Helmut Prodinger. Enumeration of Deutsch paths by the adding-a-new-slice method and applications, *Matimyas Matematika, Journal of the Mathematical Society of the Philippines*, 44 (2021), 15–22.
- [45] Helmut Prodinger. Non-decreasing Deutsch paths, *The Fibonacci Quarterly* 59 (2021), 232–236.
- [46] Sarah J. Selkirk. *MSc-thesis: On a generalisation of k -Dyck paths. Stellenbosch University*. 2019.
- [47] Neil J. A. Sloane and The OEIS Foundation Inc. The on-line encyclopedia of integer sequences, 2021.

HELMUT PRODINGER, MATHEMATICS DEPARTMENT, STELLENBOSCH UNIVERSITY, 7602 STELLENBOSCH, SOUTH AFRICA.

Email address: `hproding@sun.ac.za`