Séminaire Lotharingien de Combinatoire **49** (2004), Article B49h

Binary strings without zigzags

Emanuele Munarini - Norma Zagaglia Salvi

Summary. *We study several enumerative properties of the set of all binary strings without zigzags, i.e., without substrings equal to* 101 *or* 010 *. Specifically we give the generating series, a recurrence and two explicit formulas for the number* $w_{m,n}$ *of these strings with* m 1*'s and* n 0*'s and in particular for the numbers* $w_n = w_{n,n}$ *of central strings. We also consider two matrices generated by the numbers* $w_{m,n}$ *and we prove that one is a Riordan matrix and the other one has a decomposition* LTL^t *where* L *is a lower triangular matrix and* T *is a tridiagonal matrix, both with integer entries. Finally, we give a combinatorial interpretation of the strings under consideration as binomial lattice paths without zigzags. Then we consider the more general case of Motzkin, Catalan, and trinomial paths without zigzags.*

AMS Classification: 05A15, 05A05, 05A10, 05A16.

Keywords: Binary strings, binary words, zigzags, Riordan matrix, secondary structures, Catalan-like paths, lattice paths, trinomial paths.

1 Introduction

We say that a binary string has a *zigzag* when 010 or 101 occur as substrings. Here σ is a substring, or a factor, of a string α when there exist two strings α_1 and α_2 such that $\alpha = \alpha_1 \sigma \alpha_2$. For instance, the strings 011001 and 1100111 are without zigzags, while the strings 110100 and 010011 both present a zigzag.

Our aim is to obtain enumerative and combinatorial properties of the set $\mathcal{W}$ of all binary strings without zigzags. Specifically we obtain the generating series, a recurrence and two explicit formulas for the numbers $w_{m,n}$ of all strings in $\mathcal{W}$ with m 0's and n 1's, and in particular for the numbers $w_n = w_{n,n}$ of central strings in $\mathcal{W}$, i.e., strings with an equal number of zeros and ones.

Then we consider the infinite matrix $W = [w_{i,j}]_{i,j\geq 0}$ and we prove that it admits a decomposition LTL^t where L is a lower triangular matrix and T is a tridiagonal matrix. Both the matrices L and T have nonnegative integer entries and those of L have a combinatorial interpretation as connection constants between two suitable persistent polynomial sequences.

We also consider the matrix $R = [r_{i,j}]_{i,j\geq 0}$, where $r_{i,j} = w_{i,i-j}$ for $i \geq j$ and $r_{i,j} = 0$ otherwise, and we prove that it is a Riordan matrix [16]. This means that the generating series of the columns of such a matrix have the form $\sum_{n\geq 0} r_{n,k}\, t^n = g(t)f(t)^k$ for suitable series $g(t)$ and $f(t)$. In our case $g(t)$ is the generating series for the numbers w_n and $f(t)$ is the generating series for the numbers of irreducible secondary structures [19].

Our interest in binary strings without zigzags stems out also from the fact they can be interpreted as particular lattice paths considering 0 as the step $(1,-1)$ and 1 as the step $(1,1)$. This interpretation leads us to consider also the case of Motzkin, Catalan (or Dyck) and trinomial paths without zigzags [1, 2, 6]. In particular we give the generating series and a recurrence for the numbers u_n of all central trinomial paths without zigzags ending on the x-axis at $(n,0)$. Finally we give the first-order asymptotic formula for w_n and u_n.

The analogous problem for circular strings has been studied in a separate work [14]. Both the problems, for linear and circular strings, were posed by Jie Wu in the particular case in which the number of 1's equals the number of 0's. In [21] the problem for linear strings is solved algorithmically, but no explicit formula is given.

2 Generating series

We say that a binary string has a *zigzag* when it has a substring equal to 101 or 010. Let $\mathcal{W}$ be the set of all binary strings without zigzags and let $\mathcal{W}_{m,n}$ be the set of all strings in $\mathcal{W}$ with m 1's and n 0's. Let v_n be the number of all strings in $\mathcal{W}$ of length n and let $w_{m,n} = |\mathcal{W}_{m,n}|$. Then consider the generating series

$$v(t) = \sum_{n\geq 0} v_n\, t^n\,, \qquad w(x,y) = \sum_{m,n\geq 0} w_{m,n}\, x^m y^n\,.$$

A closed form for these series can be easily obtained in the following way (see [10, 15] for a general approach to this kind of problems). Let $\mathcal{W}_0$ ($\mathcal{W}_1$) be the set of all strings in $\mathcal{W}$ starting with 0 (with 1). Then $\mathcal{W} = \varepsilon + \mathcal{W}_0 + \mathcal{W}_1$, where ε is the empty string. The sets $\mathcal{W}_0$ and $\mathcal{W}_1$ can be easily decomposed. Indeed, any string $\alpha \in \mathcal{W}_0$ is exactly of one of the following forms: i) $\alpha = 0\alpha'$ with $\alpha' = \varepsilon$ or $\alpha' \in \mathcal{W}_0$, ii) $\alpha = 01\alpha''$ with $\alpha'' = \varepsilon$ or $\alpha'' \in \mathcal{W}_1$. Dually, any string $\alpha \in \mathcal{W}_1$ is exactly of one of the following forms: i) $\alpha = 1\alpha'$ with $\alpha' = \varepsilon$ or $\alpha' \in \mathcal{W}_1$, ii) $\alpha = 10\alpha''$ with $\alpha'' = \varepsilon$ or $\alpha'' \in \mathcal{W}_0$. Hence

$$\begin{cases} \mathcal{W}_0 = 0 + 0\,\mathcal{W}_0 + 01 + 01\,\mathcal{W}_1 \\ \mathcal{W}_1 = 1 + 1\,\mathcal{W}_1 + 10 + 10\,\mathcal{W}_0\,. \end{cases} \tag{1}$$

To obtain the series $v(t)$ consider the morphism $\varphi : \{0,1\}^* \to \mathbb{Z}[\![t]\!]$ defined by $\varphi(1) = \varphi(0) = t$. Then system (1) becomes the following system in $\mathbb{Z}[\![t]\!]$

$$\begin{cases} (1-t)\, w_0(t) - t^2\, w_1(t) = t + t^2 \\ -t^2\, w_0(t) + (1-t)\, w_1(t) = t + t^2 \end{cases}$$

whose solution is

$$w_0(t) = w_1(t) = \frac{t+t^2}{1-t-t^2} = \frac{1}{1-t-t^2} - 1 = \sum_{n\geq 0} f_n\, t^n - 1$$

where the f_n's are the Fibonacci numbers (with $f_0 = f_1 = 1$). Then

$$v(t) = 1 + w_0(t) + w_1(t) = \frac{1+t+t^2}{1-t-t^2} = \frac{2}{1-t-t^2} - 1$$

and hence $v_n = 2f_n - \delta_{n,0}$, where $\delta_{i,j}$ is the usual Kronecker delta. Since $\mathcal{W}_0$ and $\mathcal{W}_1$ are sets of complementary strings, there is a bijection between them. Then system (1) allow to prove directly that f_n is the number of all strings in $\mathcal{W}$ of length n not starting with 0 (or, equivalently, not starting with 1) and the identity $v_n = 2f_n - \delta_{n,0}$. This last identity can also be explained combinatorially in the following way. It is well known that the Fibonacci numbers f_{n+1} count all binary strings of length n not containing the pattern 11. For $n \geq 1$ to any binary string $\alpha = a_1 \cdots a_n$ associate the binary string $\beta = b_1 \cdots b_{n-1}$ where $b_k = \mathrm{xor}(a_k, a_{k+1})$. This is a 2-1-mapping where complementary strings are mapped onto the same string. Moreover α does not contain the patterns 101 and 010 if and only if β does not contain the pattern 11.

To obtain the series $w(x,y)$ consider the morphism $\psi : \{0,1\}^* \to \mathbb{Z}[\![x,y]\!]$ defined by $\psi(1) = x$ and $\psi(0) = y$. In this case system (1) becomes the following system in $\mathbb{Z}[\![x,y]\!]$

$$\begin{cases} (1-y)\, w_0(x,y) - xy\, w_1(x,y) = y + xy \\ -xy\, w_0(x,y) + (1-x)\, w_1(x,y) = x + xy \end{cases}$$

whose solution is

$$w_0(x,y) = \frac{y + x^2y^2}{1-x-y+xy-x^2y^2}\,, \quad w_1(x,y) = \frac{x + x^2y^2}{1-x-y+xy-x^2y^2}\,.$$

Since $w(x,y) = 1 + w_0(x,y) + w_1(x,y)$, we have

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

Notice that the form of this series implies the linear recurrence

$$w_{m+2,n+2} = w_{m+1,n+2} + w_{m+2,n+1} - w_{m+1,n+1} + w_{m,n} + \delta_{m,0}\delta_{n,0}\,. \tag{3}$$

Finally, since $\mathcal{W}_{m,n}$ and $\mathcal{W}_{n,m}$ are sets of complementary strings, it immediately follows the symmetry $w_{m,n} = w_{n,m}$.

3 Explicit formulas

In this section we will give two explicit formulas for the numbers $w_{m,n}$. One is obtained formally by expanding the generating series (2) while the other one is obtained combinatorially giving an explicit canonical decomposition of the strings in $\mathcal{W}$.

First formula. The first formula is obtained expanding series (2) in the following way:

$$
\begin{aligned}
w(x,y) &= \frac{1+xy+x^2y^2}{(1-x)(1-y)} \frac{1}{1-\dfrac{x^2y^2}{(1-x)(1-y)}} \\
&= (1+xy+x^2y^2)\sum_{k\geq 0} \frac{x^{2k}}{(1-x)^{k+1}} \frac{y^{2k}}{(1-y)^{k+1}} \\
&= \sum_{k\geq 0} \frac{x^{k-\lfloor k/3\rfloor}}{(1-x)^{\lfloor k/3\rfloor+1}} \frac{y^{k-\lfloor k/3\rfloor}}{(1-y)^{\lfloor k/3\rfloor+1}}\,.
\end{aligned}
$$

Since

$$
\frac{t^r}{(1-t)^{s+1}} = \sum_{n\geq 0} \binom{n-r+s}{s} t^n\,,
$$

we have the expansion

$$
\begin{aligned}
w(x,y) &= \sum_{k\geq 0} \left(\sum_{m\geq 0} \binom{m-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} x^m \right) \left(\sum_{n\geq 0} \binom{n-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} y^n \right) \\
&= \sum_{m,n,k\geq 0} \binom{m-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} \binom{n-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} x^m y^n
\end{aligned}
$$

which yields the identity

$$
w_{m,n} = \sum_{k\geq 0} \binom{m-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} \binom{n-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor} \tag{4}
$$

where the index k in the summation is at most $\min(m+\lfloor m/2\rfloor, n+\lfloor n/2\rfloor)$.

Second formula. A second expression for $w_{m,n}$ can be obtained by keeping track of the number of factors 01 and 10 ("switches") in words without zigzags. Note that occurrences of these factors cannot overlap, which leads to a simple description in terms of regular expressions. Indeed, let A be the set of words denoted by the regular expression 0^+1^+10. We then have $\mathcal{W}_0 = \sum_{k\geq 0} \mathcal{W}_0^{(k)}$, where

$$
\mathcal{W}_0^{(0)} = 0^+\,, \quad \mathcal{W}_0^{(2k)} = A^k 0^* \ \ (k>0)\,, \quad \mathcal{W}_0^{(2k+1)} = A^k 0^+ 1^+ \ \ (k\geq 0)\,,
$$

and similarly for $\mathcal{W}_1$. The generating function for A w.r.t. the number of zeros and ones is $x^2y^2/((1-x)(1-y))$, so that by binomial expansion of the obvious generating functions for the $\mathcal{W}_i^{(k)}$ $(i=0,1)$ one finds the following expressions for the numbers $w_{m,n}^{(k)}$ of strings without zigzags with m ones and n zeros and with precisely k occurrences of "switches":

$$\begin{aligned}
w_{m,n}^{(0)} &= \delta_{m,0}+\delta_{n,0}-\delta_{m,0}\delta_{n,0} \\
w_{m,n}^{(2k)} &= \binom{m-k}{k}\binom{n-k-1}{k-1}+\binom{m-k-1}{k-1}\binom{n-k-1}{k} \\
&= \frac{k\,(m+n-2k)}{(m-k)(n-k)}\binom{m-k}{k}\binom{n-k}{k}\ (k>0) \\
w_{m,n}^{(2k+1)} &= 2\binom{m-k-1}{k}\binom{n-k-1}{k} = 2\,\frac{(m-2k)(n-2k)}{(m-k)(n-k)}\binom{m-k}{k}\binom{n-k}{k}\ (k\geq 0)\,.
\end{aligned}$$

This leads to

$$w_{m,n}=\sum_{k\geq 0}\binom{m-k}{k}\binom{n-k}{k}\frac{2mn-3(m+n)k+6k^2}{(m-k)(n-k)} \qquad (m,n>0)\,. \tag{5}$$

The generating function

$$\sum_{m,n,k\geq 0} w_{m,n}^{(k)}\,x^m y^n z^k = \frac{(1-xyz)^2-xy}{(1-x)(1-y)-(xyz)^2}$$

can be obtained by routine calculation from

$$\mathcal{W}=\varepsilon+\sum_{k\geq 0}\left(\mathcal{W}_0^{(k)}+\mathcal{W}_1^{(k)}\right)$$

or from an obvious generalization of the system of equations given in Section 2.

4 Central strings

We say that a binary string is *central* when the number of 1's is equal to the number of 0's. Let $w_n = w_{n,n}$ be the number of all *central binary strings without zigzags*. The first values of w_n are 1, 2, 4, 8, 18, 42, 100, 242, 592, 1460, 3624, 9042, 22656, 56970, 143688 (sequence #A078678 in [17]).

Of course the explicit formulas (4) and (5) for the numbers $w_{m,n}$ yield explicit formulas for the numbers w_n. Specifically

$$w_n=\sum_{k=0}^{n+\lfloor n/2\rfloor}\binom{n-k+2\lfloor k/3\rfloor}{\lfloor k/3\rfloor}^2, \qquad w_n=\sum_{k\geq 0}^{\lfloor n/2\rfloor}\binom{n-k}{k}^2\frac{2n^2-6nk+6k^2}{(n-k)^2}\,.$$

The generating function for the numbers w_n is the diagonal of the double series $w(x,y)$. By Cauchy's integral theorem [5, 11, 19] it is given by

$$w(t)=\sum_{n\geq 0} w_n\,t^n=\frac{1}{2\pi\mathrm{i}}\oint w\left(z,\frac{t}{z}\right)\frac{\mathrm{d}z}{z}=\frac{1}{2\pi\mathrm{i}}\oint\frac{1+t+t^2}{-z^2+(1+t-t^2)z-t}\,\mathrm{d}z$$

where the integral is taken over a simple contour containing all the singularities $s(t)$ of the series such that $s(t) \to 0$ as $t \to 0$. The polynomial $z^2 - (1 + t - t^2)z + t$ at the denominator has roots

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

Notice that under the radical we have the polynomial

$$(1 + t - t^2)^2 - 4t = 1 - 2t - t^2 - 2t^3 + t^4 = (1 + t + t^2)(1 - 3t + t^2).$$

Since only $z^- \to 0$ as $t \to 0$, by the residue theorem, we have

$$w(t) = \lim_{z \to z^-} \frac{1 + t + t^2}{-(z - z^+)} = \frac{1 + t + t^2}{z^+ - z^-}$$

that is

$$w(t) = \frac{1 + t + t^2}{\sqrt{1 - 2t - t^2 - 2t^3 + t^4}} = \sqrt{\frac{1 + t + t^2}{1 - 3t + t^2}}. \tag{6}$$

Since

$$\frac{1 + t + t^2}{1 - 3t + t^2} = 1 + \frac{4t}{1 - 3t + t^2} = 1 + 4t \sum_{n \geq 0} f_{2n+1}\, t^n,$$

one has, from identity (6), a convolution representation of the odd-indexed Fibonacci numbers in terms of the w_n

$$4 f_{2n-1} = \sum_{k=0}^{n} w_k w_{n-k} \qquad (n > 0)$$

which is very much reminiscent of the familiar

$$4^n = \sum_{k=0}^{n} \binom{2k}{k} \binom{2(n-k)}{n-k}.$$

Finally, taking the logarithmic derivative of $w(t)$, we obtain the identity

$$(1 - 2t - t^2 - 2t^3 + t^4)\, w'(t) = 2(1 - x^2)\, w(t)$$

which implies the following linear recurrence for the numbers w_n

$$(n + 4)\, w_{n+4} - 2(n + 4)\, w_{n+3} - (n + 2)\, w_{n+2} - 2\, n\, w_{n+1} + n\, w_n = 0.$$

5 The matrix W

We now turn to consider the infinite matrix

$$W = [w_{i,j}]_{i,j\geq 0} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 2 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 5 & 8 & 11 & 14 & 17 & 20 & 23 \\ 1 & 2 & 6 & 11 & 18 & 26 & 35 & 45 & 56 \\ 1 & 2 & 7 & 14 & 26 & 42 & 62 & 86 & 114 \\ 1 & 2 & 8 & 17 & 35 & 62 & 100 & 150 & 213 \\ 1 & 2 & 9 & 20 & 45 & 86 & 150 & 242 & 367 \\ 1 & 2 & 10 & 23 & 56 & 114 & 213 & 367 & 592 \\ \dots \end{bmatrix}.$$

We will prove that it can be decomposed as LTL^t where L is the lower triangular matrix

$$L = [l_{i,j}]_{i,j\geq 0} = \begin{bmatrix} 1 \\ 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 3 & 2 & 1 \\ 1 & 1 & 1 & 4 & 3 & 3 & 1 \\ 1 & 1 & 1 & 5 & 4 & 6 & 3 & 1 \\ 1 & 1 & 1 & 6 & 5 & 10 & 6 & 4 & 1 \\ \dots \end{bmatrix}$$

with entries given by

$$l_{i,0} = 1\,, \qquad l_{i+1,j+1} = \binom{i - \lceil j/2 \rceil}{i-j} = \binom{i - \lceil j/2 \rceil}{\lfloor j/2 \rfloor}\,,$$

L^t is the transpose of L and T is the tridiagonal matrix

$$T = [t_{i,j}]_{i,j\geq 0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 & 0 \\ & 0 & 2 & 1 \\ & & 1 & 2 & 0 \\ & & & 0 & 2 & 1 \\ & & & & 1 & 2 & 0 \\ & & & & & 0 & 2 & 1 \\ & & & & & & & \dots \end{bmatrix}$$

with entries

$$t_{i,j} = [j \bmod 2 = 0, j \neq 0]\,\delta_{i,j+1} + (2 - [i = 0,1])\,\delta_{i,j} + [i \bmod 2 = 0, i \neq 0]\,\delta_{i+1,j}$$

where the square brackets denote the Iverson notation for the characteristic function of a proposition [9] and $\delta_{i,j} = [i = j]$ is the usual Kronecker delta.

The decomposition $W = LTL^t$ is equivalent to the identity

$$\sum_{h,k\geq 0} l_{i,h} t_{h,k} l_{j,k} = w_{i,j} \,. \tag{7}$$

To prove such an identity we consider the series

$$S(x,y) = \sum_{i,j\geq 0} \left(\sum_{h,k\geq 0} l_{i,h} t_{h,k} l_{j,k} \right) x^i y^j \,.$$

Substituting $t_{h,k}$ with its explicit expression and using the formulas

$$\sum_{i\geq 0} l_{i,0}\, x^i = \frac{1}{1-x}\,, \qquad \sum_{i\geq 0} l_{i,k}\, x^i = \frac{x^k}{(1-x)^{\lceil k/2 \rceil}}\,,$$

after some straightforward computations we obtain $S(x,y) = w(x,y)$, proving identity (7).

The stated decomposition for the matrix W also holds for the finite matrix $W_n = [w_{i,j}]_{i,j=0}^n$. More precisely $W_n = L_n T_n L_n^t$ where $L_n = [l_{i,j}]_{i,j=0}^n$ and $T_n = [t_{i,j}]_{i,j=0}^n$. Hence $\det W_n = \det T_n$ and $\det T_n = \det T'_{n-2}$ where T'_{n-2} is the matrix obtained deleting the first two rows and columns of T_n. Moreover, if $d_n = \det T'_n$ it is easy to see that $d_{n+2} = 3d_n$ and $d_0 = 2$, $d_1 = 3$. Hence it follows that $d_n = \frac{5-(-1)^n}{2}\, 3^{\lfloor n/2 \rfloor}$. So $\det W_n = 1$ for $n = 0, 1$ and $\det W_n = d_{n-2}$ for $n \geq 2$. Since the entries of the matrix W count certain lattice paths (see Section 7) it could be interesting to prove the above result combinatorially in the style of Gessel-Viennot theorem for the number of configurations of non-intersecting lattice paths [8].

We conclude this section noticing that the coefficients $l_{i+1,j+1}$ are connection constants between two persistent polynomial sequences. Specifically

$$x^n = \sum_{k=0}^{n} \binom{n - \lceil k/2 \rceil}{n-k} x^{\lfloor k/2 \rfloor} (x-1)^{\lceil k/2 \rceil} \,.$$

6 A Riordan matrix

Let $r_{n,k} = w_{n,n-k}$ for $k \leq n$ and $r_{n,k} = 0$ otherwise. Then the matrix

$$R = [r_{n,k}]_{n,k\geq 0} = \begin{bmatrix} 1 & & & & & & & & \\ 2 & 1 & & & & & & & \\ 4 & 2 & 1 & & & & & & \\ 8 & 5 & 2 & 1 & & & & & \\ 18 & 11 & 6 & 2 & 1 & & & & \\ 42 & 26 & 14 & 7 & 2 & 1 & & & \\ 100 & 62 & 35 & 17 & 8 & 2 & 1 & & \\ 242 & 150 & 86 & 45 & 20 & 9 & 2 & 1 & \\ 592 & 367 & 213 & 114 & 56 & 23 & 10 & 2 & 1 \\ \cdots & & & & & & & & \end{bmatrix}$$

is a Riordan matrix [16, 12]. Indeed recurrence (3) yields the following recurrence for the coefficients $r_{n,k}$:

$$r_{n+2,k+1} = r_{n+1,k} + r_{n+2,k+2} - r_{n+1,k+1} + r_{n,k+1}\,. \tag{8}$$

Considering the series $r_k(t) = \sum_{n\geq k} r_{n,k}\, t^n$, recurrence (8) becomes

$$r_{k+2}(t) - (1+t-t^2)\, r_{k+1}(t) + t\, r_k(t) = 0\,.$$

Now suppose there exist two series $g(t)$ and $f(t)$ such that $r_k(t) = g(t)f(t)^k$ for all $k \in \mathbb{N}$. Then $g(t) = r_0(t) = w(t)$ and $f(t)$ is defined by the quadratic equation

$$f(t)^2 - (1+t-t^2)f(t) + t = 0$$

whose solution is

$$f(t) = \frac{1+t-t^2-\sqrt{1-2t-t^2-2t^3+t^4}}{2}\,.$$

Then

$$r_k(t) = \sqrt{\frac{1+t+t^2}{1-3t+t^2}}\left(\frac{1+t-t^2-\sqrt{1-2t-t^2-2t^3+t^4}}{2}\right)^k\,.$$

In conclusion, since $g_0 = 1$, $f_0 = 0$ and $f_1 \neq 0$, R is the Riordan matrix

$$\left(\sqrt{\frac{1+t+t^2}{1-3t+t^2}}\,,\ \frac{1+t-t^2-\sqrt{1-2t-t^2-2t^3+t^4}}{2}\right)\,.$$

The first few coefficients in the expansion of $f(t)$ are

$$f(t) = t + t^3 + t^4 + 2t^5 + 4t^6 + 8t^7 + 17t^8 + 37t^9 + 82t^{10} + \cdots$$

which suggests (looking at entry A004148, formerly M1141, of [17]) that these coefficients are the numbers of (irreducible) *secondary structures* in the sense of [20, 19]. Indeed, while $f(t)$ is the generating function for irreducible secondary structures, the generating function $s(t)$ for all secondary structures is simply the iteration of $f(t)$, i.e.,

$$s(t) = \sum_{n\geq 0} s_n\, t^n = \frac{1}{1-f(t)} = \frac{1-t+t^2-\sqrt{1-2t-t^2-2t^3+t^4}}{2t^2}$$

Consider now the row sums $r_n = \sum_{k=0}^{n} r_{n,k}$ of the matrix R. The first values of r_n are 1, 3, 7, 16, 38, 92, 225, 555, 1378, 3439, 8619, 21678, 54687, and their generating series is given by

$$r(t) = \sum_{n\geq 0} r_n\, t^n = \frac{w(t)}{1-f(t)} = w(t)s(t)\,.$$

This implies the convolution identity

$$r_n = \sum_{k=0}^{n} w_k s_{n-k}\,.$$

Finally we notice that the matrix R is completely determined by the recurrences

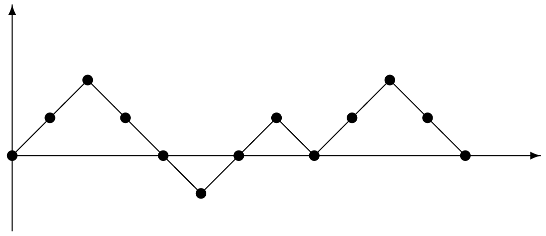
$$\begin{cases} r_{n+2,k+1} = r_{n+1,k} + r_{n,k+1} + r_{n,k+2} + \cdots + r_{n,n} \\ r_{n+2,0} = r_{n+1,0} + r_{n,0} + 2r_{n,1} + \cdots + 2r_{n,n} \end{cases}$$

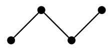
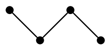
and by the initial condition $r_{0,0} = 1$.

7 Lattice paths

A *Motzkin path* [2] of length n is a lattice path in $\mathbb{N} \times \mathbb{Z}$ which starts at $(0, 0)$ and ends at $(n, 0)$, has steps $(1, 1)$, $(1, 0)$, $(1, -1)$ and never falls below $y = 0$. A *Catalan*, or *Dyck*, *path* is a Motzkin path without horizontal steps. A *central trinomial path* is defined as a Motzkin path with the possibility to go below $y = 0$, while a *central binomial path* is a central trinomial path without horizontal steps.

Binary strings with m 1's and n 0's can be interpreted as lattice paths in $\mathbb{N} \times \mathbb{Z}$ from the origin $(0, 0)$ to the point $(m + n, m - n)$ considering 1 and 0 as the unitary diagonal steps $(1, 1)$ and $(1, -1)$, respectively. For instance the string $\alpha = 110001101100$ corresponds to the path



This implies that the strings in $\mathcal{W}$ can be interpreted as binomial paths without zigzags, i.e., without a sequence of consecutive steps of the form  and . In particular $w_{m,n}$ is the number of all binomial paths without zigzags ending at $(m + n, m - n)$, and w_n is the number of all central binomial paths without zigzags ending at $(2n, 0)$. In the above example α is central but is not without zigzags.

This interpretation leads to consider the more general problem to enumerate all Motzkin, Catalan and central trinomial paths without zigzags.

Motzkin paths. Let $\mathcal{Z}_n$ be the set of all Motzkin paths of length n without zigzags. To each horizontal step we assign a weight s. For any $\gamma \in \mathcal{Z}_n$ define the weight $w(\gamma)$ as the product of the weights of all its steps. Then let

$$Z_n^{(s)} = \sum_{\gamma \in \mathcal{Z}_n} w(\gamma) \qquad \text{and} \qquad Z^{(s)}(t) = \sum_{n \geq 0} Z_n^{(s)} \, t^n .$$

For $s = 1$ we have the Motzkin paths while for $s = 0$ we have the Catalan paths. Any Motzkin path can be uniquely decomposed in the product of horizontal steps on the x-axis and elevated Motzkin paths, i.e., paths of the form $(1, 1)\, \gamma\, (1, -1)$ where γ is any Motzkin path. Such a decomposition holds also for the Motzkin paths without zigzags, even though clearly not all the possible patterns are allowed. Let $\mathcal{Z}_{hs}$, $\mathcal{Z}_{hill}$ and $\mathcal{Z}_{ep}$ be the sets of all Motzkin paths without zigzags whose final factor is a horizonal step, a hill (i.e., $(1, 1)(1, -1)$) or an elevated path respectively. It is easy to see that

$$\begin{cases} Z^{(s)}(t) = 1 + Z_{hs}^{(s)}(t) + Z_{hill}^{(s)}(t) + Z_{ep}^{(s)}(t) \\ Z_{hs}^{(s)}(t) = stZ^{(s)}(t) \\ Z_{hill}^{(s)}(t) = t^2 + st^3 Z^{(s)}(t) \\ Z_{ep}^{(s)}(t) = (1 + Z_{hs}^{(s)}(t) + Z_{ep}^{(s)}(t))\, (Z^{(s)}(t) - 1)\, t^2 . \end{cases}$$

Solving this system it is easy to obtain the equation

$$(1-st^3)Z^{(s)}(t)^2-(1-st+t^2-st^3+t^4-st^5)Z^{(s)}(t)+1+t^2+t^4=0\,.$$

Then

$$Z^{(s)}(t)=\frac{1-st+t^2-st^3+t^4-st^5-\sqrt{\Delta}}{2t^2(1-st^3)}$$

where $\Delta=1-2st+(s^2-2)t^2-4st^3+(2s^2-1)t^4-2st^5+(3s^2-2)t^6+(2s^2+1)t^8+2st^9+s^2t^{10}$. Moreover we have the recurrence

$$Z^{(s)}_{n+5}=sZ^{(s)}_{n+4}-Z^{(s)}_{n+3}+sZ^{(s)}_{n+2}-Z^{(s)}_{n+1}+sZ^{(s)}_{n}+\sum_{k=0}^{n+3}Z^{(s)}_{k}Z^{(s)}_{n+3-k}-s\sum_{k=0}^{n}Z^{(s)}_{k}Z^{(s)}_{n-k}\,.$$

This result can be generalized to the case in which we assign a weight a to all horizontal steps on x-axis and a weight s to all other horizontal steps. In this way, for instance, for $a=0$ and $s=1$ we have all *Riordan paths* [2, 1, 4] without zigzags while for $a=0$ and $s=2$ we have all *Fine paths* [2, 1, 7] without zigzags. In this (generalized) case we have the system

$$\begin{cases} Z^{(a,s)}(t)=1+Z^{(a,s)}_{hs}(t)+Z^{(a,s)}_{hill}(t)+Z^{(a,s)}_{ep}(t)\\ Z^{(a,s)}_{hs}(t)=atZ^{(a,s)}(t)\\ Z^{(a,s)}_{hill}(t)=t^2+at^3Z^{(a,s)}(t)\\ Z^{(a,s)}_{ep}(t)=(1+Z^{(a,s)}_{hs}(t)+Z^{(a,s)}_{ep}(t))\,(Z^{(s)}(t)-1)\;t^2 \end{cases}$$

which yield the generating function

$$Z^{(a,s)}(t)=\frac{1-(2a-s)t+t^2-(2a-s)t^3+t^4-(2a-s)t^5-\sqrt{\Delta}}{s-a+(a^2-as+1)t+(s-a)t^2+a(a-s)t^3-at^4+a(a-s)t^5}\,.$$

Central trinomial paths. To study these paths when they avoid zigzags it seems more convenient to consider them as trinomial strings and use the results obtained on binary strings without zigzags. Specifically we see a *trinomial path* as a word on the alphabet $\{0,1,h\}$, considering 1 and 0 as unitary diagonal steps as before and h as the unitary horizontal step $(1,0)$. A zigzag is still one of the strings 101 or 010. Let $\mathcal{U}$ be the set of all ternary strings without zigzags. Since, in terms of regular expressions, $\mathcal{U}=\mathcal{W}(h\mathcal{W})^*$, we obtain from the morphism $\nu:\{0,1,h\}^*\to\mathbb{Z}[\![x,y,z]\!]$, defined by $\nu(1)=x$, $\nu(0)=y$ and $\nu(h)=z$, the generating function

$$u(x,y,z)=\sum_{n\geq 0}w(x,y)^{n+1}z^n=\frac{w(x,y)}{1-w(x,y)z}$$

that is

$$u(x,y,z)=\frac{1+xy+x^2y^2}{1-x-y+xy-x^2y^2-(1+xy+x^2y^2)z}$$

where the coefficient $u_{i,j,k}$ of $x^iy^jz^k$ in $u(x,y,z)$ is the number of all strings in $\mathcal{U}$ with i 1's, j 0's and k h's.

We say that a string in $\mathcal{U}$ is *central* when the number of 1's is equal to the number of 0's. Hence the central strings in $\mathcal{U}$ correspond to trinomial paths without zigzags ending on the x-axis. Let u_n be the number of such paths ending at $(n,0)$. To obtain the generating function for these numbers we consider first the series

$$u(x,y) = \sum_{i,j\geq 0} u_{i,i,j}\; x^i y^j\,.$$

By Cauchy's integral theorem we have

$$u(x,y) = \frac{1}{2\pi \mathrm{i}} \oint u\left(z, \frac{x}{z}, y\right)\,\frac{\mathrm{d}z}{z} = \frac{1}{2\pi \mathrm{i}} \oint \frac{1+x+x^2}{-z^2+(1+x-y-x^2-xy-x^2y)z-x}\,\mathrm{d}z\,.$$

The polynomial at the denominator has roots

$$z^{\pm} = \frac{1+x-y-x^2-xy-x^2y \pm \sqrt{(1+x-y-x^2-xy-x^2y)^2-4x}}{2}$$

of which only $z^- \to 0$ as $x \to 0$. Hence, by the residue theorem, we have

$$u(x,y) = \lim_{z\to z^-} \frac{1+x+x^2}{-(z-z^+)} = \frac{1+x+x^2}{z^+ - z^-}$$

that is

$$u(x,y) = \frac{1+x+x^2}{\sqrt{(1+x-y-x^2-xy-x^2y)^2-4x}}\,.$$

Then the generating function for the numbers u_n is given by $u(t) = u(t^2,t)$, that is

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

Taking the logarithmic derivative of this series we obtain the identity

$$(1-4t+4t^2-2t^3+t^6)\;u'(t) = (1+t-3t^2-t^3+t^4-t^5)\;u(t)$$

which yields the recurrence

$$(n+6)\,u_{n+6} - (4n+21)\,u_{n+5} + (4n+15)\,u_{n+4} - (2n+3)\,u_{n+3} + u_{n+2} - u_{n+1} + (n+1)\,u_n = 0\,.$$

The first few values of u_n are 1, 1, 3, 7, 17, 43, 111, 291, 771, 2059, 5533, 14943, 40523 (sequence #A078079 in [17]).

We can treat the case in which each horizontal step has weight s in the same way. We have only to define $\nu(h) = sz$. In particular

$$u^{(s)}(t) = \sqrt{\frac{1+t^2+t^4}{1-2st+(s^2-3)t^2-2st^3+(s^2+1)t^4+2st^5+s^2t^6}}\,.$$

8 Asymptotics

In this final section we give the first-order asymptotic formulas for w_n and u_n. To obtain such a formula we use the following theorem ([3], p. 252): given a complex number $\xi \neq 0$ and a complex function $f(t)$ analytic at the origin, if $f(t) = (1-t/\xi)^{-\alpha}\psi(t)$ where $\psi(t)$ is a series with radius of convergence $R > |\xi|$ and $\alpha \notin \{0, -1, -2, \ldots\}$, then

$$[t^n]f(t) \sim \frac{\psi(\xi)}{\xi^n}\,\frac{n^{\alpha-1}}{\Gamma(\alpha)}$$

where Γ is Euler's gamma function. For the series $w(t)$ we have

$$w(t) = \left(1 - \frac{t}{\xi}\right)^{-\alpha} \sqrt{\frac{1+t+t^2}{1-\xi t}}$$

with $\xi = (3-\sqrt{5})/2$ and $\alpha = 1/2$. Hence, since $\Gamma(1/2) = \sqrt{\pi}$, it follows that

$$w_n \sim \frac{2}{\sqrt{n\pi\sqrt{5}}}\left(\frac{3+\sqrt{5}}{2}\right)^n = \frac{2\,\varphi^{2n}}{\sqrt{n\pi\sqrt{5}}}$$

where $\varphi = (1+\sqrt{5})/2$ is the golden ratio. In particular

$$\lim_{n\to\infty} \frac{w_{n+1}}{w_n} = \frac{3+\sqrt{5}}{2} = \varphi^2\,.$$

It could be interesting to explain combinatorially the appearance of φ.

Since for the numbers of the secondary structures we have the asymptotic formula [20]

$$s_n \sim \sqrt{\frac{15+7\sqrt{5}}{8\pi}}\left(\frac{3+\sqrt{5}}{2}\right)^n \frac{1}{n^{3/2}}$$

it follows that

$$\lim_{n\to+\infty} \frac{n s_n}{w_n} = \frac{5+3\sqrt{5}}{8}\,.$$

Finally, the series $u(t)$ can be written as

$$u(t) = \left(1 - \frac{t}{\xi}\right)^{-1/2} \sqrt{\frac{1-t+t^2}{(1-t)(1-t/\xi_1)(1-t/\xi_2)}}$$

where

$$\xi = \frac{1}{6}\sqrt[3]{188+12\sqrt{249}} - \frac{4}{3}\,\frac{1}{\sqrt[3]{188+12\sqrt{249}}} - \frac{2}{3} \simeq 0.353\,.$$

Then $u_n \sim ab^n/\sqrt{n\pi}$, where $a = \psi(\xi) \simeq 0.944$ and $b = 1/\xi \simeq 2.831$.

Acknowledgments. The authors are grateful to Volker Strehl and Douglas Rogers for some useful suggestions.

References

[1] M. Aigner, Catalan-like Numbers and Determinants, *J. Combin. Theory* Ser. A **87** (1999), 33–51.

[2] M. Aigner, Catalan and other numbers: a recurrent theme, in *Algebraic Combinatorics and Computer Science*, 347–390, Springer Italia, Milan, 2001.

[3] F. Bergeron, G. Labelle, P. Leroux, *Combinatorial Species and Tree-like Structures*, Encyclopedia of Mathematics and Its Applications **67**, Cambridge University Press, Cambridge, 1998.

[4] F.R. Bernhart, Catalan, Motzkin, and Riordan numbers, *Discrete Math.* **204** (1999), 73–112

[5] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht-Holland, Boston, 1974.

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

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

[8] I.M. Gessel, X. Viennot, Binomial determinants, paths, and hook length formulae, *Advances in Mathematics* **58** (1985), 300-321.

[9] R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, MA, 1989.

[10] L.J. Guibas, A.M. Odlyzko, String Overlappings, Pattern Matching, and Nontransitive Games, *J. of Combin. Theory* A **30** (1981), 183–208.

[11] M.L.J. Hautus, D.A. Klarner, The Diagonal of a Double Power Series, *Duke Math. J.* **38** (1971), 229–235.

[12] D. Merlini, D. Rogers, R. Sprugnoli, M.C. Verri, On some alternative characterizations of Riordan arrays, *Canad. J. Math.* **49** (1997), 301–320.

[13] E. Munarini, N. Zagaglia Salvi, On the rank polynomial of the lattice of order ideals of fences and crowns, *Discrete Math.* **259** (2002), 163–177.

[14] E. Munarini, N. Zagaglia Salvi, Circular binary strings without zigzags, *Integers* **3** (2003), A19, 17 pages (electronic).

[15] A.M. Odlyzko, Enumeration of strings, in *Combinatorial Algorithms on Words*, A. Apostolico and Z. Galil (eds.), Springer 1985, 205–228 .

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

[17] N.J.A. Sloane, *On-Line Encyclopedia of Integer Sequences*, published electronically at `http://www.research.att.com/~njas/sequences/`.

[18] R.P. Stanley, *Enumerative Combinatorics*, Volume 1, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, Cambridge, 1997.

[19] R.P. Stanley, *Enumerative Combinatorics*, Volume 2, Cambridge Studies in Advanced Mathematics **62**, Cambridge University Press, Cambridge, 1999.

[20] P.R. Stein, M.S. Waterman, On some new sequences generalizing the Catalan and Motzkin numbers, *Discrete Math.* **26** (1979), 261–272.

[21] F. Dai, M. Dhawan, C. Wu, J. Wu, On Solutions to a Seating Problem. *2002 ISCA International Conference on Parallel and Distributed Computing Systems*, 2002.

Emanuele Munarini (munarini@mate.polimi.it), Dipartimento di Matematica, Politecnico di Milano, P.zza Leonardo da Vinci 32, 20133 Milano, Italy.

Norma Zagaglia Salvi (norzag@mate.polimi.it), Dipartimento di Matematica, Politecnico di Milano, P.zza Leonardo da Vinci 32, 20133 Milano, Italy.