Skew Dyck paths

Emeric Deutsch[*] Emanuele Munarini[†] Simone Rinaldi[‡]

In the discrete plane we consider paths using steps of three types: *up* steps $u = (1,1)$, *down* steps $d = (1,-1)$, and *left* steps $l = (-1,-1)$. A *Dyck path* of length $2n$ is a sequence of steps running from $(0,0)$ to $(2n,0)$, using up and down steps and remaining weakly above the x-axis (see Figure 1 (a)). It is well know [4] that the number of Dyck paths of length $2n$ is the nth *Catalan number* $c_n = \binom{2n}{n}\frac{1}{n+1}$. A *skew Dyck path* (briefly, *skew path*) is a sequence of steps running from $(0,0)$ to $(2m,0)$, using up, down, and left steps, remaining weakly above the x-axis, and where the up and left steps do not overlap (see Figure 1 (b)). We say that the *length* of the path is the number of its steps. Let $\mathbb{S}_n$ denote the class of skew paths of length $2n$. Differently from Dyck paths, a skew path ending in $(2m,0)$ has not necessarily length $2m$; on the contrary the set $\mathbb{S}^m$ of skew paths ending in $(2m,0)$ is infinite. Each skew path can be encoded by a *skew word* in the alphabet $\{u,d,l\}$. For example the skew path in Figure 1 (b) is encoded by the skew word *uuduudlduuuddl*.

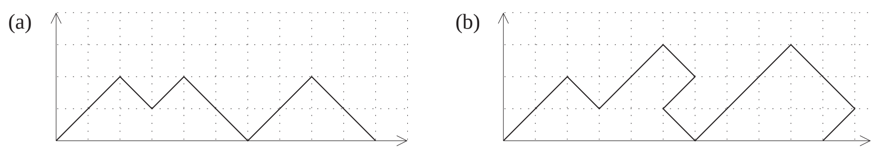

Figure 1: (a) a Dyck path; (b) a skew Dyck path.

It is easy to see that every non-empty skew path γ can be uniquely decomposed as

$$\gamma = u\,\gamma'\,d\,\gamma'' \qquad \text{or} \qquad \gamma = u\,\gamma'''\,l\,, \tag{1}$$

where $\gamma', \gamma'', \gamma'''$ are skew paths, γ''' being non empty. Hence, the generating function $S(x)$ of skew paths according to semi-length satisfies the following functional equation:

$$S(x) = 1 + x\,S^2(x) + x\,(S(x)-1), \tag{2}$$

which leads to the solution:

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

The first values of the coefficients s_n defined by $S(x)$ are $1,1,3,10,36,137,\dots$ (sequence A002212 in [3]). From (3) we can derive the recurrence:

$$(n+1)s_n - 3(2n-1)s_{n-1} + 5(n-2)s_{n-2} = 0 \qquad (\text{for } n \geq 2)\,.$$

Moreover, writing (2) as $S(x)^2 = (1-x)(S(x)-1)/x$, we have the identity

$$\sum_{i=0}^{n} s_i s_{n-i} = s_{n+1} - s_n \qquad (\text{for } n \geq 1)\,.$$

Applying the Lagrange Inversion Theorem to (2), we have the following formula:

$$s_{n+1} = [x^{n+1}]\,S(x) = \frac{1}{n+1}\sum_{i=1}^{n+1}\binom{n+1}{i}\binom{2i}{i-1} = \sum_{i=0}^{n}\binom{n}{i}c_{i+1}\,.$$

[*]Polytechnic University, Six Metrotech Center, Brooklyn, NY 11201, (deutsch@duke.poly.edu).
[†]Politecnico di Milano, Dipartimento di Matematica, Piazza Leonardo da Vinci 32, 20133 Milano, Italy. (emanuele.munarini@polimi.it).
[‡]Università di Siena, Dipartimento di Matematica, via del Capitano 15, 53100, Siena, Italy. (rinaldi@unisi.it).

We remark that all the previous identities can be explained in combinatorial terms. Moreover, we study the class of skew paths according to the following features:

1. **Bijections with other combinatorial objects**. The sequence s_n also enumerates the 3-*colored Motzkin paths* according to length [4], and the *restricted hexagonal polyominoes* according to number of cells [1]. We give explicit bijections between these three classes.

2. **An involution on skew paths**. We define a simple involution on skew paths, which directly follows from the recursive decomposition (1). Let $(-)^* : \mathbb{S}_n \to \mathbb{S}_n$ be the function defined by $\varepsilon^* = \varepsilon$, $(u\,d)^* = u\,d$, and $(u\,d\,\gamma)^* = u\,\gamma^*\,l$, $(u\,\gamma\,l)^* = u\,d\,\gamma^*$, $(u\,\gamma\,d\,\delta)^* = u\,\gamma^* d\,\delta^*$, where γ and δ are skew paths, with $\gamma \neq \varepsilon$. For instance we have $(uuduudld)^* = uuududld$. Easily one verifies that $(-)^*$ is an involution, i.e. that it is a bijection on skew paths and that $\gamma^{**} = \gamma$ for any skew path γ. We give a characterization of the fixed points of $(-)^*$, and prove that the number of these fixed points of length n is the $(n-1)$th Motzkin number, for $n \geq 1$.

3. **Statistics on skew paths.** We consider several statistics on skew paths, among which, the following are worth mentioning:

   (a) *Number of returns.* A *return* in a skew path is any occurrence of a down or left step at level zero. The number of skew paths of length $2n$ with exactly k returns is given by

   $$\frac{1}{n-k}\sum_{h=1}^{n-k}\binom{n-k}{h}\binom{2h+k-1}{h-1}\frac{h-k+3hk+k^2}{2h+k-1}\,.$$

   (b) *Number of left steps.* The number of skew paths of length $2n$ and having k left steps is given by $\binom{n-1}{k}c_{n-k}$.

   (c) *Number of doublerises.* A *doublerise* in a skew path is any occurrence of two consecutive up steps. The number of skew paths of length $2n$ having k doublerises is given by

   $$\frac{1}{n}\binom{n}{k}\sum_{h=0}^{\min\{k,n-k-1\}}\binom{k}{h}\binom{n-k}{h+1}2^{k-h}\,.$$

   (d) *Length of the base.* A skew path has *base* of length k when it ends at the point $(0,2k)$. The number of skew paths of length $2n$ and base of length k is given by $\binom{n-1}{k-1}c_k$.

   (e) *Number of peaks.* The number of skew paths of length $2n$ with k peaks is

   $$\sum_{j=0}^{n-k}\binom{j+1}{n-j-k}\binom{2j+k}{2j}(-1)^{n-j-k}c_j\,.$$

   (f) *Number of right returns.* A *right return* in a skew path is any occurrence of a down step at level zero. Using the factorization (1) it follows that the generating series of all skew paths according to semi-length and the number of right returns is given by $G(x,t) = (1-x+xG(x))/(1-xtG(x))$. This implies that the associated matrix $G = [g_{nk}]_{n,k\geq 0}$ is a Riordan matrix [2], and more precisely $G = (1-x+xS(x), xS(x))$. Moreover we have the recurrence: $g_{n+2,k+2} = g_{n+2,k+1} - g_{n+1,k+1} - g_{n+1,k} + g_{nk}$.

   (g) *Number of hills.* The generating series $H(x,t)$ of skew paths according to semi-length and number of hills (i.e. a peak at level zero) satisfies the identity $H(x,t) = 1 + xtH(x,t) + x(S(x)-1)H(x,t) + x(S(x)-1)$ from which $H(x,t) = (1-x+xS(x))/(1+x-xS(x)-xt)$. This implies that the associated matrix $H = [h_{nk}]_{n,k\geq 0}$ is the Riordan matrix $H = \left(\frac{1-x+xS(x)}{1+x-xS(x)}, \frac{x}{1+x-xS(x)}\right)$. We can also prove the recurrence:

   $$h_{n+1,k+1} = h_{nk} + \sum_{i=1}^{n-k} 3^{i-1}h_{n,k+i}\,.$$

4. **Enumeration according to area.** We define the *area* of a skew path starting from the case of Dyck paths. Here we consider the region defined by the path and the x-axis. This region can be uniquely decomposed into right triangles of side one, and we say that the number of these unit triangles is the *area* of the Dyck path (see Figure 2 (a)). Similarly in a skew path we define the area as the number of unit triangles in the region below the path and above the x-axis (see Figure 2 (b)). Let $\mathbb{S}^n$ be the set of all skew paths ending at the point $(0, 2n)$.

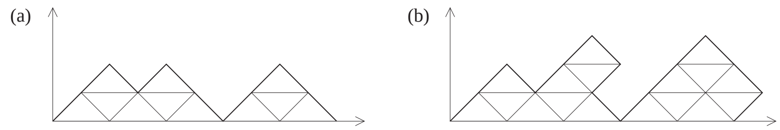

Figure 2: (a) a Dyck path of area 11; (b) a skew Dyck path of area 17.

Then let $A_n(q) = \sum_{m\geq 0} a_{mn} q^m$ be the generating series for the paths in $\mathbb{S}^n$ according to their area. Using the main decomposition (1) we obtain the recurrence

$$A_{n+1}(q) = \sum_{k=0}^{n} q^{2k+1} A_k(q) A_{n-k}(q) + q^{2n+2} A_{n+1}(q)\,.$$

Since the initial value is $A_0(q) = 1$, it is straightforward to conjecture and then to prove that $A_n(q) = \frac{q^n}{(1-q^2)^n}$ for every $n \in \mathbb{N}$. This implies that the number of skew paths ending in $(0, 2n)$ with area m is $\binom{(m-n)/2+n-1}{n-1}$ if $m = n \pmod 2$ and 0 otherwise. Moreover we have the generating series

$$A(q;x) = \sum_{n\geq 0} A_n(q)x^n = \frac{1-q^2}{1-q^2-qx} \qquad \text{and} \qquad A(q;1) = 1 + \frac{q}{1-q-q^2}\,.$$

Hence the number of all skew paths of area n is the nth Fibonacci number, F_n, for every $n \geq 1$. We also give a combinatorial proof of this result, as sketched here: the skew paths of area n are obtained from those of area $n-1$ by prepending a ud and from those of area $n-2$ by placing a square of side 1 over the first d step (Figure 3). Hence, with $n \geq 3$, the number a_n of skew paths having area n satisfies the recurrence $a_n = a_{n-1} + a_{n-2}$ and consequently $a_n = F_n$. This results is quite surprising, since the generating function of Dyck paths and of many other known lattice paths, according to area, leads to non holonomic generating functions.

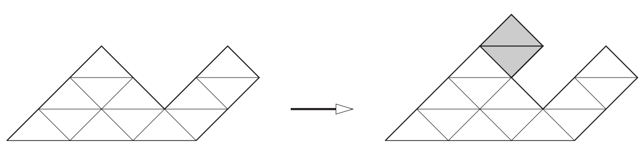

Figure 3: A skew path with area 15 obtained in a unique way from one with area 13.

5. **Enumeration of skew bargraphs.** To any skew path γ from $(0,0)$ to $(2n,0)$ we can associate a superdiagonal bargraph $B(\gamma)$ whose boundary is defined by γ itself, rotated counterclockwise by $\pi/4$, and by the lines $y = 0$ and $x = n$ (see Figure 4). This bargraph has the property that the highest side of each column is placed above the diagonal $y = x$. This class of bargraphs is itself an interesting combinatorial class to be studied according to various parameters. The generating function of skew bargraphs according to *semi-perimeter* (i.e. half the length of their boundary) is given by

$$\frac{1-x-\sqrt{1-2\,x-3\,x^2+4\,x^3}}{2x^2} = 1 + x^2 + x^3 + 3x^4 + 5x^5 + 12x^6 + 24x^7 + 55x^8 + \ldots$$

Its coefficients generate the sequence A090345 in [3], and also are the number of Motzkin paths of length n with no level steps at even level.

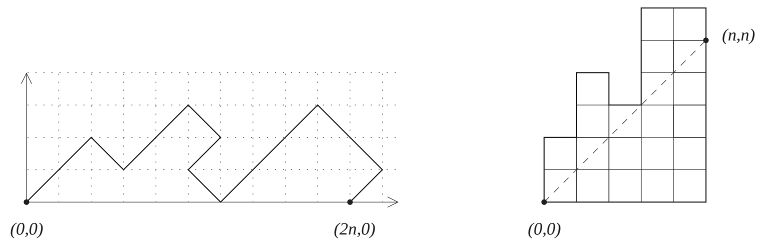

Figure 4: A skew path and the associated skew bargraph.

To treat the enumeration according to *area* (i.e. the number of cells), let $\mathbb{B}_n$ be the set of skew bargraphs having area n (and b_n its cardinality), and $\mathbb{B}_{n,k}$ the set of skew bargraphs having area n and k columns (and $b_{n,k}$ its cardinality), with $n, k \geq 1$. For $k \geq 1$, the minimal area n for which $b_{n,k} \neq 0$ is $n = k(k+1)/2$. Moreover, with $n > k(k+1)/2$ we can divide $\mathbb{B}_{n,k}$ into two disjoint subsets:

(a) the bargraphs whose first column is made exactly of one cell. Each of these bargraphs can be obtained in a unique way by adding a row made of k cells at the bottom of a bargraph of $\mathbb{B}_{n-k,k-1}$; thus the cardinality of this set is $b_{n-k,k-1}$ (see Figure 5 (a));

(b) the remaining bargraphs. Each of these bargraphs can be obtained in a unique way by adding a cell on the top of the first column of a bargraph of $\mathbb{B}_{n-1,k}$; then the cardinality of this set is given by $b_{n-1,k}$ (see Figure 5 (b)).

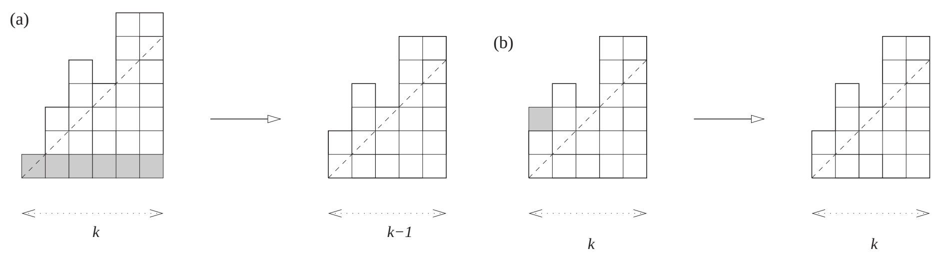

Figure 5: Bargraphs in the above classes.

Hence we have $b_{n,k} = b_{n-1,k} + b_{n-k,k-1}$. Passing to generating functions $B(x) = \sum_{n\geq 0} b_n x^n$ and $B_k(x) = \sum_{n\geq 0} b_{n,k} x^n$, we have the recurrence $B_k(x) = xB_k(x) + x^k B_{k-1}$ with the initial value $B_1(x) = x/(1-x)$. Consequently

$$B_k(x) = \frac{x^k}{1-x}\, B_{k-1}(x) = \frac{x^{\frac{k(k+1)}{2}}}{(1-x)^k} \qquad \text{and} \qquad B(x) = \sum_{k\geq 1} \frac{x^{\frac{k(k+1)}{2}}}{(1-x)^k}.$$

Finally we have $b_n = \sum_{k=0}^{m} \binom{n-\binom{k}{2}}{k}$, where $m = (\sqrt{8n+1}-1)/2$, and the first few values: $1, 2, 3, 5, 8, 12, 18, 27, 40, 58, 83, 118, 167, \ldots$ (sequence A06978 in [3]).

References

[1] F. Harary, R. C. Read, *The enumeration of tree-like polyhexes*, Proc. Edinb. Math. Soc. (2) **17** (1970), 1-13.

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

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

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