A NEW WAY OF COUNTING THE COLUMN-CONVEX POLYOMINOES BY THE PERIMETER

Svjetlan Feretić, Šetalište Joakima Rakovca 17, 51000 Rijeka, Croatia

Abstract. We derive in a simple way the perimeter generating function for the column-convex polyominoes.

Résumé. Nous retrouvons de façon facile la série génératrice des polyominos verticalement convexes, comptés suivant le périmètre.

1. Introduction

Besides its purely mathematical interest, the computation of the *self-avoiding polygon* (*SAP,* Fig.1) perimeter and area generating functions would have a significant bearing on the study of physical problems like fusion and evaporation, the configuration of polymer molecules and gel formation. But despite strenuous efforts over the past 40 years, so far only some restricted classes of the SAP's have been enumerated. Further, in all the known enumerations two SAP's are identified iff they can be transformed one into the other by a translation (the reflections and rotations are not allowed).

An important restricted class of the SAP's arises if we impose convexity in the direction of one of the lattice axes. When this axis is the y-axis, we speak of the *column-convex polyominoes* (*cc-polyominoes,* Fig.2).

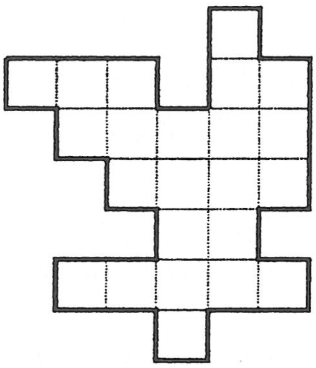

Figure 1. A self-avoiding polygon (SAP)

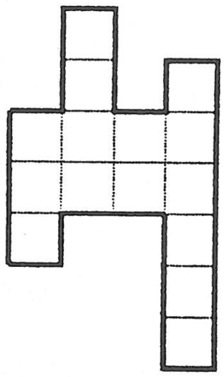

Figure 2. A column-convex (cc-) polyomino

Notation 1. Let P be a cc-polyomino. We shall write $He(P)$ for the number of horizontal edges of P, and $Ve(P)$ for the number of vertical edges of P. ∎

Definition 1. Let Ω be some family of cc-polyominoes. By the perimeter generating function (gf) for Ω we mean the formal sum $\sum_{P\in\Omega} x^{He(P)}y^{Ve(P)}$. ∎

The cc-polyominoes were introduced by Temperley [12] in 1956. The area gf of this model was found on the spot [12]. On the contrary, the perimeter gf of the cc-polyominoes $G(x,y)$ (and not to speak of their perimeter+ area gf) remained unknown for a long time after Temperley's paper had been published. At last Delest [5] applied the DSV-methodology [2, 6, 11, 13, 14] and the computer algebra program MACSYMA to obtain a formula for $G(x,x)$. Subsequently Brak, Guttmann and Enting [4] rederived the function

$G(x,x)$ using the Temperley methodology and *Mathematica*. Thus it turned out that the formula given in [5] can be written in a simpler form. The result of Brak *et al.* was generalized to the case $x \neq y$ by Lin [9].

In the course of preparation of their paper [7], Feretić and Svrtan were firstly using the DSV-methodology. So they encoded the cc-polyominoes and set up a system of four nonlinear equations. Some manipulation of this system left them with a single degree-four algebraic equation satisfied by $G(x,y)$. Wishing to calculate the then unknown Taylor coefficients of $G(x,y)$ by the Lagrange inversion, they factored that algebraic equation. The result was that

$$H = \frac{x^2y^2(1-H)^4}{(1-y^2)^2(1-2H)[(1-3H)^2 - x^2(1-H)^2]}, \tag{1}$$

where

$$H = \frac{G}{1-y^2}. \tag{2}$$

The equation (1) made it possible to express $< x^{2c}y^{2v} > G$ as a certain threefold sum of binomial coefficients. But the final surprise was still lurking nearby. Namely, after a while one of the authors of [7] found it out that the division of (1) by $1-H$ leads to a biquadratic equation satisfied by the function $L = (1-3H)/(1-H)$. Solving that biquadratic equation and using $H = 1 - 2/(3-L)$, the following unexpectedly simple formula for $G(x,y)$ was obtained:

$$G(x,y) = (1-y^2)\left[1 - \frac{2\sqrt{2}}{3\sqrt{2} - \sqrt{1 + x^2 + \sqrt{(1-x^2)^2 - \frac{16x^2y^2}{(1-y^2)^2}}}}\right]. \tag{3}$$

In [7] there is also an alternative proof for (3), which was found later. This second proof uses the Temperley recurrences [12], but the way of solving the recurrences is different than in [4, 9].

The aim of the present paper is to give an explanation for the "magic" behaviour of the functions like H and L. In section 2 we introduce two new classes of plane figures, whose abbreviated names are tapoes and stapoes. It is established by inspection that $H = J/(1+J)$, where $J =$(the perimeter gf for the stapoes)$/(1-y^2)$. In section 3 we use the DSV-methodology to derive the function J and thereby the functions H and G. The computations are easy, because we have to solve just one quadratic equation instead of a system of quadratic equations.

2. Two new objects

Let P be a cc-polyomino. The upper left corner of the first column of P is called the north-west pole of P and is denoted by $NW(P)$. The lower right corner of the last column of P is said to be the south-east pole of P (notation: $SE(P)$).

Imagine a plane figure T obtained by appending a vertical segment of $d \in \mathbb{N}_0$ lattice units to the south-east pole of a cc-polyomino P. We say that T is a *tailed polyomino* (for short: *tapo*). Naturally, the appended segment is termed the *tail* of T. By the columns of a tapo T we mean the columns of the underlying cc-polyomino P. The north-west pole of T is defined by $NW(T) = NW(P)$, while the south-east pole $SE(T)$ is defined to be the lower endpoint of the tail of T. See Fig. 3.

Now let us define the second new object. Suppose that, for some $n \in \mathrm{N}$, $n-1$ arbitrary tapoes $T_1, \ldots, T_{n-1}$ and a tapo with a null tail T_n are given. Let $T_1, \ldots, T_n$ be disposed in a way that, for $2 \leq i \leq n$, the north-west pole of T_i coincides with the south-east pole of T_{i-1}.

In a situation like this we say that the union $S = \cup_{1 \leq i \leq n} T_i$ is a *stapo* (short for: *a sequence of tailed polyominoes*). The tapoes $T_1, \ldots, T_n$ are called the *parts* of S. By the columns of a stapo we mean the columns of its parts. See Fig.4. Observe that the one-part stapoes are cc-polyominoes.

It is useful to adopt the following convention:

Convention. Let a tapo T be obtained by appending a segment of length d to a cc-polyomino which has $2v$ vertical edges. Then T has $2v + 2d$ vertical edges. ∎

Naturally, by the vertical perimeter of a stapo we mean the sum of the vertical perimeters of its parts. With these conventions, in the sequel we shall apply Notation 1 and Definition 1 not only to the cc-polyominoes, but also to the tapoes and stapoes.

Let H_d be the perimeter *gf* for the tapoes whose tail is exactly d units long. It is easy to see that $H_d = y^{2d} G$, where G is the perimeter *gf* for the cc-polyominoes. By this remark and (2), the perimeter *gf* for all the tapoes is

$$\sum_{d \geq 0} y^{2d} G = \frac{G}{1-y^2} = H. \tag{4}$$

An n-part stapo is , in substance, a sequence of $n-1$ tapoes and one cc-polyomino. Hence the perimeter *gf* for the n-part stapoes is $H^{n-1}G$. Let I be the perimeter *gf* for all the stapoes. We have

$$I = \sum_{n \geq 1} H^{n-1} G = \frac{G}{1-H}. \tag{5}$$

Further, it is convenient to put

$$J = I/(1-y^2). \tag{6}$$

The function J can be interpreted as the perimeter *gf* for the generalized stapoes, whose last part, too, is allowed to have a tail. From (2) and (5) it follows that

$$J = \frac{H}{1-H}, \tag{7}$$

so that

$$H = 1 - \frac{1}{1+J}. \tag{8}$$

3. The DSV-computation of the function G

3.1. Preliminaries on words and languages

Mostly due to the papers of the Bordeaux group for enumerative combinatorics [5, 6, ...], the algebraic language (*i.e.* DSV-) methodology is today a popular counting technique.

Here we shall dispense with giving an introduction to this method. However, we shall give some non-standard definitions concerning the free monoid $\{x, y, \bar{y}\}^*$, which is the one relevant to our forthcoming proof.

For $v \in \{x, y, \bar{y}\}^*$, we put $\delta(v) = |v|_y - |v|_{\bar{y}}$ and say that $\delta(v)$ is the *rank* of v.

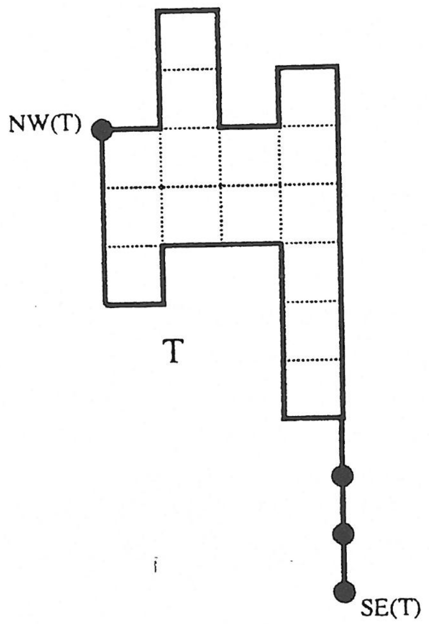

Figure 3. A tailed polyomino (tapo) with four columns and 24 vertical edges

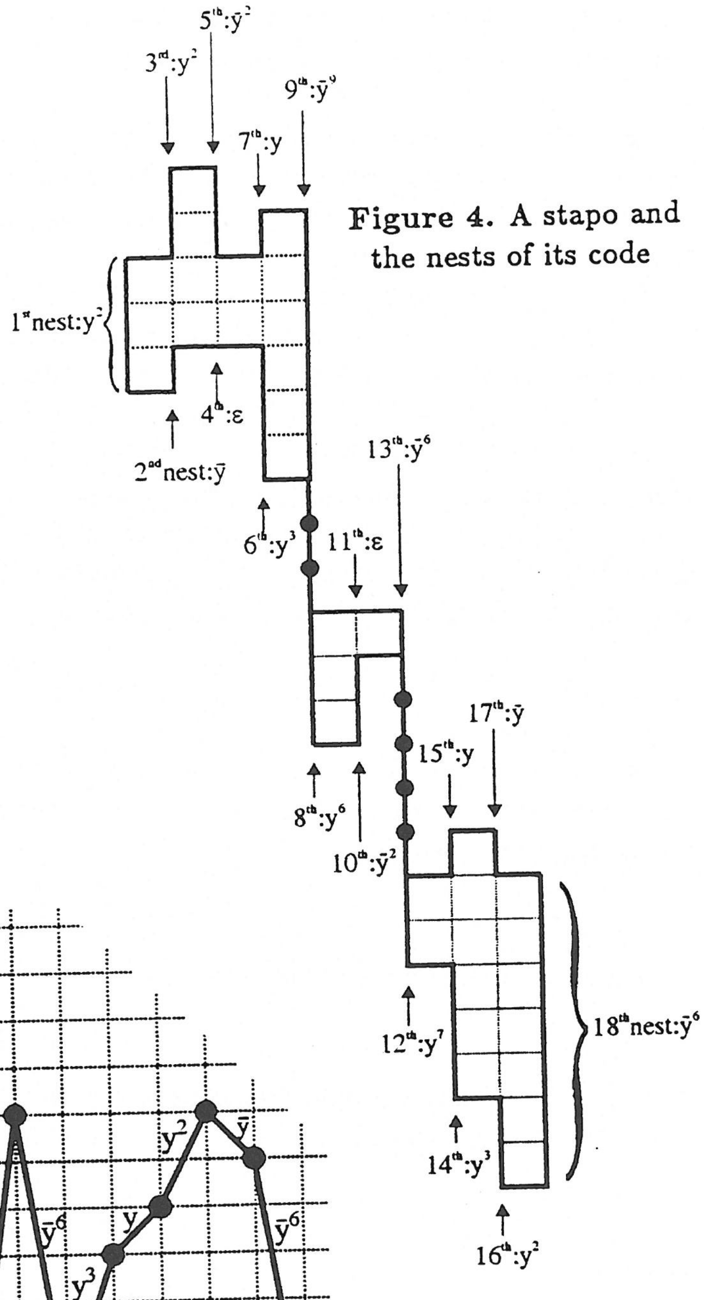

Figure 4. A stapo and the nests of its code

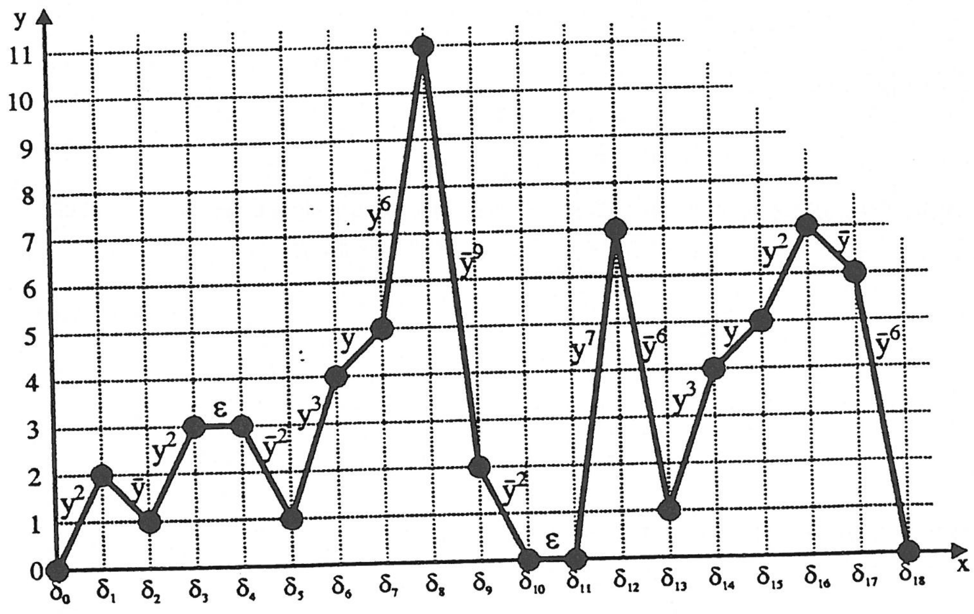

Figure 5. The "partial ranks" of the word which encodes the stapo of Fig.4, *i.e.* the numbers $\delta_i(\psi(S))$, $(i = 0, 1, \ldots, 18)$.

Let $w \in \{x, y, \bar{y}\}^*$ and let $|w|_x = n$. Clearly, w can be written as $u_1 \cdot x \cdot u_2 \cdot x \cdots u_n \cdot x \cdot u_{n+1}$, where $u_i \in \{y, \bar{y}\}^*$, for every i. The word u_i will be called the i^{th} *nest* of w. Now, we put $\delta_0(w) = 0$, $\delta_1(w) = \delta(u_1)$ and

$$\delta_i(w) = \delta(u_1 \cdot x \cdot u_2 \cdots x \cdot u_i) \quad (i = 2, \ldots, n+1). \tag{9}$$

Also, we define $\tau(w)$ to be the word obtained from w by swapping the nests u_i and u_{i+1}, for every $i = 2, 4, \ldots, 2\lfloor n/2 \rfloor$.

We say that w is a *Motzkin word* if the rank of w is zero and the ranks of all the left factors of w are nonnegative.

We say that w is a word with *pure* nests if every nest of w belongs either to $\{y\}^*$ or to $\{\bar{y}\}^*$.

The letter $\mathcal{B}$ will denote the language formed by the Motzkin words with pure nests.

For example, let $w = y \cdot x \cdot yy \cdot x \cdot x \cdot \bar{y} \cdot x \cdot y \cdot x \cdot \bar{y}\bar{y}\bar{y}$. This w is an element of $\mathcal{B}$, the numbers $\delta_0(w), \ldots, \delta_6(w)$ are $0, 1, 3, 3, 2, 3, 0$ and $\tau(w) = y \cdot x \cdot x \cdot yy \cdot x \cdot y \cdot x \cdot \bar{y} \cdot x \cdot \bar{y}\bar{y}\bar{y}$.

3.2. A coding for the stapoes

Let S be a stapo with c columns and $2v$ vertical edges ($c, v \in \mathbb{N}$). Let y_i and Y_i be the minimal and the maximal ordinate of the i^{th} column of S. Observe that for all i, we have $Y_i > y_i$ and $Y_i > y_{i+1}$.

We define the *code* of S to be the word $w = \psi(S)$ having the following properties:

i) w is a Motzkin word with pure nests;

ii) $|w|_x = 2c - 1$;

iii) $\delta_{2i-1}(w) = Y_i - y_i - 1 \quad (i \in \underline{c})$,[1] and $\delta_{2i}(w) = Y_i - y_{i+1} - 1 \quad (i \in \underline{c-1})$.

(To be sure, there is only one such w.) Essentially, we encode the stapoes similarly as Delest [5] encoded the cc-polyominoes. An example for our coding is shown in Figures 4 and 5.

Let u_i denote the i^{th} nest of w. Owing to the purity of these nests, we have $|u_i| = |\delta(u_i)|$, for every $i \in \underline{2c}$. This fact and the property iii) imply

$$|u_1| = Y_1 - y_1 - 1, \quad |u_{2c}| = Y_c - y_c - 1,$$

and for $i \in \underline{c-1}$,

$$|u_{2i}| = |y_i - y_{i+1}|, \quad |u_{2i+1}| = |Y_{i+1} - Y_i|.$$

Assume, without loss of generality, that the minimal abscissa of S is zero. Clearly, S has $Y_1 - y_1$ vertical edges with abscissa zero and $Y_c - y_c$ vertical edges with abscissa c. Further, whether a tail with abscissa $i \in \underline{c-1}$ exist or not, there are always $|y_i - y_{i+1}| + |Y_{i+1} - Y_i|$ vertical edges with abscissa i. (This statement may be checked by examining Fig.4). Since S has $2v$ vertical edges, putting our remarks together we find

$$|w|_y + |w|_{\bar{y}} = \sum_{1 \le j \le 2c} |u_j| = 2v - 2. \tag{10}$$

[1]The symbol $\underline{c}$ denotes the set of integers $\{1, 2, \ldots, c\}$.

Let

$$\mathcal{B}_{cv} = \{\, w \in \mathcal{B} \;:\; |w|_x = 2c-1,\ |w|_y + |w|_{\bar{y}} = 2v-2 \,\}.$$

We have proved the following result:

Proposition 1. We have $\psi(S) \in \mathcal{B}_{cv}$. ▮

It is of interest to make some considerations about the word $\tau(w)$, where w is code of the stapo S. Quite obviously, the nests of this word are pure, and we have

$$|\tau(w)|_x = 2c-1, \quad |\tau(w)|_y + |\tau(w)|_{\bar{y}} = 2v-2. \tag{11}$$

Having determined the ranks of the nests of w with the aid of the above property iii), from the definition of $\tau(w)$ we obtain

$$\delta_{2i-1}(\tau(w)) = Y_i - y_i - 1 \quad (i \in \underline{c}) \tag{12}$$

and

$$\delta_{2i}(\tau(w)) = Y_{i+1} - y_i - 1 \quad (i \in \underline{c-1}). \tag{13}$$

Thus, in the general situation we don't know whether a number $\delta_{2i}(\tau(w))$ be nonnegative or negative. But let us see what happens in the special case of S being a cc-polyomino. Then $Y_{i+1} > y_i$ for every $i \in \underline{c-1}$, so that (12) and (13) give

$$\delta_j(\tau(w)) \geq 0 \quad (\forall j \in \underline{2c-1}). \tag{14}$$

On account of the purity of nests, (14) proves that in this case $\tau(w)$ *is* a Motzkin word.

Let

$$\mathcal{B}^+_{cv} = \{\, w \in \mathcal{B}_{cv} \;:\; \tau(w) \text{ also lies in } \mathcal{B}_{cv} \,\}.$$

We have:

Proposition 2. If S is a cc-polyomino, then $\psi(S) \in \mathcal{B}^+_{cv}$. ▮

Notation 2. For $c, v \in \mathbf{N}$, $\mathcal{S}_{cv}$ will denote the family of stapoes having c columns and $2v$ vertical edges. Also, we put

$$\mathcal{P}_{cv} = \{\, P \in \mathcal{S}_{cv} \;:\; P \text{ is a cc-polyomino} \,\}. ▮$$

To be fair, so far we have only proved that $\psi(\mathcal{S}_{cv}) \subseteq \mathcal{B}_{cv}$ and $\psi(\mathcal{P}_{cv}) \subseteq \mathcal{B}^+_{cv}$. But it is just a technical matter to arrive at a stronger conclusion:

Proposition 3. ψ is a bijection between $\mathcal{S}_{cv}$ and $\mathcal{B}_{cv}$, and also a bijection between $\mathcal{P}_{cv}$ and $\mathcal{B}^+_{cv}$. ▮

Now, the absence of the awkward requirement "$\tau(w) \in \mathcal{B}_{cv}$" indicates that it will probably be easier to enumerate the family $\mathcal{B}_{cv}$ than the family $\mathcal{B}^+_{cv}$. In fact, it was right for this reason that the stapoes have been introduced.

3.3. The grammar and the algebra

We define a power series $B(x, y)$ as follows. For $i, j \in \mathbf{N}_0$, the coefficient of $x^i y^j$ in B, usually written as $< x^i y^j > B$, is

$$\text{card}\{\, w \in \mathcal{B} \;:\; |w|_x = i,\ |w|_y + |w|_{\bar{y}} = j \,\}.$$

Next, let $D(x,y) = [B(x,y) - B(-x,y)]/2$. The definitions and Proposition 3 imply

$$< x^{2c}y^{2v} > I = |\mathcal{S}_{cv}| = |\mathcal{B}_{cv}| =< x^{2c-1}y^{2v-2} > B =$$
$$=< x^{2c-1}y^{2v-2} > D =< x^{2c}y^{2v} > xy^2D \quad (c,v \in \mathrm{N}). \tag{15}$$

Since in the power series I and xy^2D all the powers of x and y are even, (15) implies

$$I = xy^2D. \tag{16}$$

It is readily seen that the language $\mathcal{B}$ has the unambiguous grammar

$$\mathcal{B} = \epsilon + x\mathcal{B} + y(\mathcal{B} - \epsilon)\bar{y}(\epsilon + x\mathcal{B}). \tag{17}$$

Letting the letters in (17) commute and putting $y = \bar{y}$, we find that the *gf* $B(x,y)$ satisfies the quadratic equation

$$B = 1 + xB + y^2(B-1)(1+xB),$$

or equivalently

$$xy^2B^2 - (1-x)(1-y^2)B + 1 - y^2 = 0. \tag{18}$$

Solving the equation (18) we obtain

$$B = \frac{(1-x)(1-y^2) - \Delta_-^{1/2}}{2xy^2} \tag{19}$$

and

$$D = \frac{2(1-y^2) - \Delta_-^{1/2} - \Delta_+^{1/2}}{4xy^2}, \tag{20}$$

where $\Delta_\pm = (1 \pm x)^2(1-y^2)^2 \pm 4xy^2(1-y^2)$. Using (16) to obtain I, (6) to obtain J, (8) to obtain H and (2) to obtain G, we get the following theorem:

Theorem 1. The perimeter generating function for the column-convex polyominoes is given by

$$G(x,y) = (1-y^2)\left[1 - \frac{4}{6 - \sqrt{(1-x)^2 - \frac{4xy^2}{1-y^2}} - \sqrt{(1+x)^2 + \frac{4xy^2}{1-y^2}}}\right]. \blacksquare \tag{21}$$

Let $\delta_-^{1/2}$ and $\delta_+^{1/2}$ denote the first and the second of the square roots which appear in the denominator of (21). The fact that the formulas (3) and (21) determine the same function is due to the possibility to write $\delta_-^{1/2} + \delta_+^{1/2}$ as $[\delta_- + \delta_+ + 2(\delta_-\delta_+)^{1/2}]^{1/2}$. Next, it follows from (21) that

$$\delta_-^{1/2} = 2L + \delta_+^{1/2}, \tag{22}$$

where $L = (1-3H)/(1-H)$. When suitably squared two times, (22) turns into the biquadratic equation

$$L^4 - (1+x^2)L^2 + x^2\left[\frac{1+y^2}{1-y^2}\right]^2 = 0, \tag{23}$$

the same one which appeared unexpectedly in [7]. Finally, rewriting (23) in the form

$$1 - L = \frac{4x^2y^2}{(1-y^2)^2(L+1)(L^2-x^2)} \tag{24}$$

and then multiplying by $(1-H)/2$, we obtain the equation (1).

References

[1] M.Bousquet-Mélou, *A method for the enumeration of various classes of column-convex polygons*, rapport LaBRI $n^o 578-93$ (1993), Université de Bordeaux I.

[2] M.Bousquet-Mélou, *Convex polyominoes and algebraic languages*, J. Phys. A: Math. Gen. 25 (1992), 1935-1944.

[3] R.Brak and A.J.Guttmann, *Exact solution of the staircase and row-convex polygon perimeter and area generating function*, J. Phys. A: Math. Gen. 23 (1990), 4581-4588.

[4] R.Brak, A.J.Guttmann and I.G.Enting, *Exact solution of the row-convex polygon perimeter generating function*, J.Phys.A: Math.Gen. 23 (1990), 2319-2326.

[5] M.P.Delest, *Generating functions for column-convex polyominoes*, J. Combin. Theory Ser.A 48 (1988), 12-31.

[6] M.P.Delest and X.G.Viennot, *Algebraic languages and polyominoes enumeration*, Theor.Comp.Sci. 34 (1984), 169-206.

[7] S.Feretić and D.Svrtan, *On the number of column-convex polyominoes with given perimeter and number of columns*, in: A.Barlotti, M.Delest, R.Pinzani editors, 5^{th} FPSAC proceedings, Firenze 1993, 201-214.

[8] D.A.Klarner, *Some results concerning polyominoes*, Fibonacci Quart. 3 (1965), 9-20.

[9] K.Y.Lin, *Perimeter generating function for row-convex polygons on the rectangular lattice*, J.Phys.A.: Math.Gen. 23 (1990), 4703-4705.

[10]K.Y.Lin and W.J.Tzeng, *Perimeter and area generating functions of the staircase and row-convex polygons on the rectangular lattice*, Int. J. Mod. Phys. B 5 (1991), 1913-1925.

[11]M.P.Schützenberger, *Context-free languages and pushdown automata*, Information and Control 6 (1963), 246-264.

[12]H.N.V.Temperley, *Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules*, Phys.Rev. 103 (1956), 1-16.

[13]X.G.Viennot, *A survey of polyominoes enumeration*, in: P. Leroux, C. Reutenauer editors, 4^{th} FPSAC proceedings, Montréal 1992, 399-420.

[14]X.G.Viennot, *Enumerative combinatorics and algebraic languages*, in: L. Budach editor, Proceedings FCT'85, Lecture Notes in Computer Science 199, Springer - Verlag, Berlin, 1985.