AN ALGORITHM FOR DECIDING IF A POLYOMINO TILES THE PLANE

IAN GAMBINI\(^1\) AND LAURENT VUILLON\(^2\)

Abstract. For polyominoes coded by their boundary word, we describe a quadratic \(O(n^2)\) algorithm in the boundary length \(n\) which improves the naïve \(O(n^4)\) algorithm. Techniques used emanate from algorithmics, discrete geometry and combinatorics on words.

Mathematics Subject Classification. 68R15, 52C20.

1. Introduction

During the DMCCG conference held at the Institut Henri Poincaré (IHP) in July 2001, we discussed with Alberto del Lungo some problems about polyominoes. One of his concerns was the design of a fast algorithm for computing the number of polyominoes that tile the plane by translations. What he really had in mind was probably their enumeration according to some convenient parameter. The algorithmic approach, by providing computational evidence, is a convenient way to get some insight about the algebraicity or rationality of certain classes of polyominoes. Let us recall some achievements along these lines.

Tilings, regular or not, have puzzled lots of people from ancient times up to now; and even now, despite of the efforts of many mathematicians these objects remain mysterious. Indeed these objects reveal an incredible amount of simple to state problems that translate into very complex combinatorial ones [1, 4, 9], like for example the squaring of a square [10]. Golomb [11] in his book presents many aspects of polyominoes and in particular he searches how to tile a finite...
I. GAMBINI AND L. VUILLON

The enumeration of general polyominoes is a difficult problem and there is no closed formula to count them. However, recent work, conducted mainly by the Bordeaux school and its satellites, allowed to enumerate some very restrictive classes like the directed, parallelogram, convex ones according to various parameters such as the half-perimeter, area, height, width, and some other refinements [4, 8, 14, 15]. More precisely closed formula are known for parallelogram polyominoes [8], for symmetry classes of parallelogram polyominoes [15], polyominoes with notion of convexity [3, 5] and symmetry classes of convex polyominoes in the square lattice [14].

Nevertheless, from the algorithmic point of view Nivat and Beauquier found a characterization of polyominoes that tile the plane by translations [2]. We use this characterization to build our algorithm for deciding if a given polyomino tiles the plane by translation. The methods of this article use techniques from algorithmic, discrete geometry and combinatorics on words.

2. Definitions and notation

A polyomino is a simply connected union of unit squares, that is a union of unit squares without holes. Let \( P \) be a polyomino. A tiling by translations of \( P \) is a partition of the whole plane by translated images of \( P \). A polyomino that tiles the plane by translation is called a tile.

Let \( \Sigma = \{a, b, \bar{a}, \bar{b}\} \) be a four letter alphabet. A reduced word on \( \Sigma \) is a word on the free group over \( \Sigma \) where all cancellations are done (namely each occurrence of \( aa, \bar{a}a, bb \) and \( \bar{b}b \) is replaced by \( \epsilon \) the empty word). Let \( b(P) \) be the boundary word of \( P \) that is the reduced word in the free group on \( \{a, b\} \) where \( a \) represents a right step, \( b \) an up step, \( \bar{a} \) a left step and \( \bar{b} \) a down step that codes the boundary of the polyomino \( P \) in the following way. Starting from an origin on the boundary of \( P \), the boundary word \( b(P) \) is the concatenation of labels of boundary unit segments read in trigonometric order. The starting point is not meaningful. Thus the boundary word \( b(P) \) is a cyclic word.

We define the \( \overline{\alpha} \) operator on \( \Sigma^+ \) by

\[
\begin{align*}
(i) \quad & \overline{(a)} = \overline{\alpha} \text{ if } \alpha \in \Sigma = \{a, b, \bar{a}, \bar{b}\}; \\
(ii) \quad & \overline{\alpha.\nu} = \overline{\nu}.(\overline{\alpha}); \\
(iii) \quad & \overline{\overline{\alpha}} = \alpha \text{ and } \overline{\overline{\nu}} = \nu.
\end{align*}
\]

The following characterization of tiling polyominoes is due to Beauquier and Nivat [2]:

**Theorem 1** (Beauquier, Nivat). A polyomino \( P \) tiles the plane by translations if and only if the boundary word \( b(P) \) is equal up to a cyclic permutation of the symbols to \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} \) where one of the variables in the factorization may be empty.

If the boundary word is equal to \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} \) (resp. \( X \cdot Y \cdot \overline{X} \cdot \overline{Y} \)) such a polyomino is called pseudo-hexagon (resp. pseudo-square).
For example, the polyomino on the left in Figure 1 is a pseudo-hexagon and the boundary word is equal to $X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} = a \cdot ab \cdot \overline{ab} \cdot \overline{a} \cdot \overline{ba} \cdot ba$ (where $X = a, Y = ab, Z = \overline{ab}$).

In fact, a polyomino $P$ may have many factorizations of its contour word. For example, in Figure 1 the boundary word of the right polyomino has the factorizations $\overline{ba} \cdot aba \cdot b \cdot \overline{ab} \cdot \overline{a} \cdot \overline{b}$ and $ba \cdot a \cdot bab \cdot \overline{ab} \cdot \overline{a} \cdot \overline{b}$.

A regular tiling is a tiling by translations of a polyomino $P$ such that each tile in the tiling has the same surrounding by translated copies of the tile $P$ according to a given factorization of its contour word (such tilings are also called in the literature lattice tilings); see Figure 2 for two regular tilings from the two factorizations of the contour word mentioned above. Each factorization leads to a regular tiling of the plane by translations as follows. If $P$ is a pseudo-hexagon, the factorization $b(P) = X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z}$ defines 6 sides of the tile where the sides in correspondence are identified by the pairings $X, \overline{X}, Y, \overline{Y}, Z, \overline{Z}$. The translations corresponding to these pairings allow then to tile the whole plane in a regular way. In the case of pseudo squares the construction with 4 sides is similar. Observe that two distinct factorizations of the boundary word of $P$ give two distinct regular tilings of the plane.
3. Algorithm

Let $n = |b(P)|$ be the length of the boundary word of $P$. In the following algorithm the indices of the boundary word $b(P)$ (or $b$ for simplicity from here on when there are indices) are taken modulo $n$. For example, the letter $b[-1]$ is of course the letter $b[n-1]$.

PS=0; PH=0;
// Step 1: Searching correspondence between $b[0]$ and $b[i]
for i = 1 to n − 1
    if $b[0] = b[i]$ then
        // Step 2: Propagation
        // Searching the largest $b[x_1..y_1] = b[x_2..y_2]$ containing 0 and $b[x_2..y_2]$ containing $i$
        $x_1 = 0; y_1 = 1; x_2 = i; y_2 = i + 1$
        while ($b[x_1 - 1] = b[y_2]$)  \{ $x_1 = x_1 - 1, y_2 = y_2 + 1$\}
        while ($b[y_1] = b[x_2 - 1]$)  \{ $y_1 = y_1 + 1, x_2 = x_2 - 1$\}
        // end of Step 2
        $U = b[y_1 + 1..x_2 - 1]; V = b[y_2 + 1..x_1 - 1];$
        if $|U| = |V|$ then
            if $U = \overline{V}$ then PS=PS+1 // Step 3: Pseudo-square
            PH=PH+ KMP($U, VV$) // Step 4: Pseudo-hexagon
        end if
    end if
end for

The end of this section explains the algorithm step by step.

**Instance:** the boundary word $b[0..n - 1]$ of length $n$ of a polyomino $P$.

**Answer:** the number of factorizations in pseudo-squares and in pseudo-hexagons tiling the plane by translation.

**Step 1.** For each position $i$ from 1 to $n - 1$, we try to match with the complementary letter of value $b[0]$.

**Step 2.** We make a propagation (scanning back and forth the boundary words) in order to have two sides of maximal length in correspondence. In other words, for each position of value $b[0]$, by propagation we find two complementary words $X$ and $\overline{X}$ on the boundary word starting from $X = b[0]$ and $\overline{X} = b[0]$ and extending the pair of complementary words $(X, \overline{X})$ in order to find the longest $X$. Then by this method we find a factorization of $b(P)$ by $X \cdot U \cdot \overline{X} \cdot V$.

**Step 3.** We now check that the remaining sides $U$ and $V$ have same length. And we answer that the polyomino is a pseudo-square if $U = \overline{V}$ that is if we have found a factorization on $X \cdot Y \cdot \overline{X} \cdot \overline{Y}$ with $Y = U$.

**Step 4.** We check if the polyomino is a pseudo-hexagon by searching four more sides in two-by-two correspondence, that is the factorizations $U = Y \cdot Z$ and $V = \overline{Y} \cdot \overline{Z}$. We use the following property: if such factorization exists, it is provided by an occurrence of the word $U = \overline{Z} \cdot \overline{Y}$ in $VV = \overline{Y} \cdot Z \cdot \overline{Y} \cdot Z$. 
This part can be done for instance by the KMP algorithm of Knuth et al. [13] or by the algorithm of Boyer-Moore [6].

**Answer**

The variable PS (resp. PH) gives the number of factorizations of \( b(P) \) (the boundary word of \( P \)) by pseudo-squares (resp. pseudo-hexagons). If PS=0 and PH=0 then \( P \) does not tile the plane by translation.

### 3.1. Proof of the algorithm

By step 2 of the algorithm, we have the following property. For each position of \( \overline{b}(0) \), by propagation we find two complementary words \( X \) and \( \overline{X} \) on the boundary word starting from \( X = \overline{b}(0) \) and extending the pair of complementary words \((X, \overline{X})\) in order to find the longest \( X \). Then by this method we find a factorization of \( b(P) \) by \( X \cdot U \cdot \overline{X} \cdot V \).

By this method, for each couple \((\overline{b}(0), \overline{b}(i) = \overline{b}(0))\) the algorithm finds by propagation a unique couple \((X, \overline{X})\) with \( X \) of maximal length. Step 3 and Step 4 find a factorization if it exists.

Thus given a boundary word of a polyomino \( P \) there are 3 cases to consider, either there is factorization A) by pseudo-hexagon or B) a factorization by pseudo-square or C) no factorization. And for each case we have to prove that the algorithm finds it.

- **P** is a pseudo-hexagon.

In this case the boundary word can be factorized up to a cyclic permutation of letters on \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} \) and we may assume without loss of generality that \( X \) contains the letter \( \overline{b}(0) \) (otherwise we make a cyclic permutation of letters). As the algorithm propagates to the left and to the right for each position of \( \overline{b}(0) \), let be \( \ell \) the \( \ell \text{th} \) element of \( X \) corresponding to \( \overline{b}(0) \). \( X_\ell = \overline{b}(0) \), so \( \overline{X}_{n-\ell+1} = \overline{b}(0) \) where \( |X| = |\overline{X}| = n \). The Step 2 of the algorithm finds at most a couple of complementary words \((X', \overline{X'})\) containing respectively the words \( X \) and \( \overline{X} \).

a) If \( X' = X \) then with the help of the KMP-algorithm Step 4 produces the good factorization \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} \).

b) When \( |X'| > |X| \) there is a difficulty and we proceed by contradiction. Assuming that the algorithm finds such pair of complementary words \((X', \overline{X'})\) then \( X' = L X R \) and \( Y' \) does exist such that the factorization of \( P \) is equal to \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} = X \cdot R Y' \cdot Z \cdot R \cdot X \cdot L \cdot \overline{Y'} \cdot \overline{Z} \). By this equality we have \( Y = R Y' \cdot Z = Z' \cdot R \). Thus \( Y = \overline{Y'} \cdot R \cdot Z = R' \cdot Z' \). If we use this information in the factorization \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} \) we obtain \( X \cdot Y \cdot Z \cdot \overline{X} \cdot \overline{Y} \cdot \overline{Z} = R' \cdot R \cdot Z' \). We find a contradiction because all the letters of \( R' \cdot R \) cancel two by two \( (R' \cdot R = \overline{r}_p \cdots \overline{r}_2 \overline{r}_1 \overline{r}_2 \cdots \overline{r}_p = \overline{r}_p \cdots \overline{r}_2 \overline{r}_2 \cdots \overline{r}_p = \cdots = \epsilon) \). This means in particular that the boundary word of \( P \) is not a reduced word and by construction the boundary word of \( P \) is a reduced word.
• $P$ is a pseudo-square.

Here the boundary word can be factorized up to a cyclic permutation of letters on $X \cdot Y \cdot \overline{X} \cdot \overline{Y}$ and we assume also that $X$ contains the letter $b[0]$. As in the previous case, the algorithm finds at most a couple of complementary words $X'$ and $\overline{X'}$ containing respectively the words $X$ and $\overline{X}$.

a) If $X' = X$ then by step 3 it finds the good factorization in $X \cdot Y \cdot \overline{X} \cdot \overline{Y}$.

b) When $X' \neq X$ then there is another difficulty. We proceed by contradiction. Assume that the algorithm finds $X' = LXR$ and then it exists $Y'$ such that $X \cdot Y \cdot \overline{X} \cdot \overline{Y} = X \cdot RY' \cdot \overline{X} \cdot \overline{Y} L$. By this equality we have $Y = RY' \overline{R}$. Thus $\overline{Y} = RY' \overline{R}$. If we use this information in the factorization $X \cdot Y \cdot \overline{X} \cdot \overline{Y}$ we obtain $X \cdot RY' \overline{R} \cdot \overline{X} \cdot RY' \overline{R}$. If we replace $R$ by $r_1 r_2 \cdots r_p$ where $r_i$'s are letters then

$$X \cdot Y \cdot \overline{X} \cdot \overline{Y} = X r_1 r_2 \cdots r_p Y' \overline{r}_p \cdots \overline{r}_2 \overline{Y}' \overline{r}_1 \cdots \overline{r}_2 \overline{Y}'$$

There are many factorizations in pseudo-square and we will work on the following one $X r_1 \cdot Y'' \cdot \overline{r}_1 \cdot X r_2 \cdot Y'' \cdot \overline{r}_2$ where $Y'' = r_2 \cdots r_p Y' \overline{r}_p \cdots \overline{r}_2$, $Y'' = r_2 \cdots r_p \overline{Y}' \overline{r}_p \cdots \overline{r}_2$ in order to show that the contour word is not one of a polyomino.

We have $X r_1 Y'' \overline{r}_1 X r_2 \overline{r}_2$ and we will now show by an argument of discrete geometry that this is not a boundary word of a simply connected union of unit squares (i.e. of a polyomino). In this decomposition $r_1$ is just a letter then for the reasoning we will take $r_1 = a$ (the reasoning is the same with $r_1 = b, \overline{a}, \overline{b}$). We use tools from discrete geometry introduced by Daurat and Nivat [7]. Since a polyomino is a simply connected union of squares (the boundary word delimits squares inside the polyomino $P$ (noted I-squares) and squares outside the polyomino $P$ (noted O-squares)). A corner on the boundary of $P$ is called salient if it is surrounded by one I-square and three O-squares. A corner on the boundary of $P$ is called reentrant if it is surrounded by three I-squares and 1 O-square. Daurat and Nivat proved in [7] that for any polyomino $P$ the number $S(P)$ of its salient points and the number $R(P)$ of its reentrant points satisfy $S(P) = R(P) + 4$, see Figure 3. For example, if $P$ is a pseudo-square with boundary word $XYX \overline{Y}$ the number of salient points associated with $X$ is equal to the number of reentrant points associated with $\overline{X}$ (by this reasoning $S(X) = R(\overline{X})$, $S(Y) = R(\overline{Y})$, $R(X) = S(\overline{X})$ and $R(Y) = S(\overline{Y})$). Thus by Daurat-Nivat theorem, it follows that the four points where $X, \overline{X}, Y, \overline{Y}$ connect are salient.

In our case we have the factorization in $X a Y'' \overline{a} X a \overline{Y}' \overline{a}$ and on the plane we have to place 4 segments associated with $a, \overline{a}, a, \overline{a}$ according to the factorization. In fact each segment determines two points on the boundary. We find the same relation as in the previous example: $S(X) = R(\overline{X})$, $S(Y) = R(\overline{Y})$, $R(X) = S(\overline{X})$ and $R(Y) = S(\overline{Y})$ and we just have to consider the 8 remaining points. The factorization is $X a Y'' \overline{a} X a \overline{Y}' \overline{a}$.
AN ALGORITHM FOR DECIDING IF A POLYOMINO TILES THE PLANE

with $X = x_1 \cdots x_m$ and $Y'' = y_1 \cdots y_n$. Then we have to compute the difference between $S(x_m a) + S(ay_1) + S(y_n a) + S(ax_1) + S(y_1 a) + S(ay_m) + S(y_n a) + S(ax_1)$ and $R(x_m a) + R(ay_1) + R(y_n a) + R(ax_1)$. But if the point associated with two letters $uv$ is salient (resp. reentrant) then by construction the point associated with $vu$ is reentrant (resp. salient). If $S(uv) = 1$ then $R(vu) = 1$. By this property $S(x_m a) + S(ay_1) + S(y_n a) + S(ax_1) = R(x_m a) + R(ay_1) + R(y_n a) + R(ax_1)$. And globally for the polyomino $P$ associated with the boundary word $XaY''aXaY''a$ we have $S(P) = R(P)$. This is in contradiction with the result on salient and reentrant points. Thus $P$ is not simply connected and cannot be a polyomino.

* $P$ does not tile the plane.

In this case by the characterization of Beauquier and Nivat there is no factorization on $X \cdot Y \cdot X \cdot Y$ nor $X \cdot Y \cdot Z \cdot X \cdot Y \cdot Z$. Then the algorithm fails in steps 3 and 4 to find a characterization and answer that $P$ does not tile the plane by translation.

3.2. Complexity of the algorithm

Let $n$ be the length of the boundary word associated with $P$.

In the first step, the algorithm tries to find all the positions of value $b[0]$ in the boundary word with complexity $O(n)$.

In step 2, the propagation give complexity $O(n)$. Thus the total complexity for steps 1 and 2 is $O(n \times n)$.

In step 3, we try to find a factorization by checking if $U = V$ and the complexity of this verification is $O(n)$. Remark also that step 3 makes just the continuation of step 2 and thus the complexities are added. Thus the total complexity for steps 1, 2 and 3 remains $O(n \times (n + n))$.

In step 4, we try to find a factorization by using the KMP algorithm and according to the complexity of KMP algorithm this step is on $O(n + k)$ where $n$ is the length of $VV$ and $k$ the length of $U$. Remark that step 4 just make the continuation of step 2 and step 3 and then we add the complexity of both parts. Thus the computation of the total complexity of the algorithm gives an algorithm on $O(n \times (n + n + (n + n))) = O(n^2)$. 
4. Enumeration of polyominoes by computer

We can use our algorithm to compute the number of polyominoes with only pseudo-square factorizations (only PS), with only pseudo-hexagon factorizations (only PH) and with both factorizations in pseudo-square and pseudo-hexagon. The last column is the number of polyominoes of length \( n \) that tile the plane by translations. In fact, factorizations exist only for even-length perimeter and of course there is no factorization for odd-length perimeter. In literature, authors use the half-perimeter in order to enumerate the polyominoes that is the length of the boundary word divided by 2. The enumeration of polyominoes in this table is just the sequence A002931 of the On-line Encyclopedia of Integer Sequences by Sloane.

We present the result for half-perimeters being between 2 and 18.

<table>
<thead>
<tr>
<th>Half-perimeter</th>
<th>polyominoes</th>
<th>only PS</th>
<th>only PH</th>
<th>both</th>
<th>tiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>0</td>
<td>20</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>124</td>
<td>1</td>
<td>82</td>
<td>17</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>588</td>
<td>8</td>
<td>298</td>
<td>46</td>
<td>352</td>
</tr>
<tr>
<td>8</td>
<td>2938</td>
<td>40</td>
<td>1007</td>
<td>103</td>
<td>1150</td>
</tr>
<tr>
<td>9</td>
<td>15268</td>
<td>170</td>
<td>3326</td>
<td>220</td>
<td>3716</td>
</tr>
<tr>
<td>10</td>
<td>81826</td>
<td>523</td>
<td>10394</td>
<td>513</td>
<td>11430</td>
</tr>
<tr>
<td>11</td>
<td>449572</td>
<td>1624</td>
<td>31918</td>
<td>1126</td>
<td>34668</td>
</tr>
<tr>
<td>12</td>
<td>2521270</td>
<td>4729</td>
<td>95767</td>
<td>2529</td>
<td>103025</td>
</tr>
<tr>
<td>13</td>
<td>14385376</td>
<td>13448</td>
<td>282816</td>
<td>5688</td>
<td>301952</td>
</tr>
<tr>
<td>14</td>
<td>83290424</td>
<td>37180</td>
<td>824720</td>
<td>12989</td>
<td>874889</td>
</tr>
<tr>
<td>15</td>
<td>488384528</td>
<td>102074</td>
<td>2383628</td>
<td>29630</td>
<td>251532</td>
</tr>
<tr>
<td>16</td>
<td>2895432660</td>
<td>276668</td>
<td>6828850</td>
<td>68569</td>
<td>7174087</td>
</tr>
<tr>
<td>17</td>
<td>17332874364</td>
<td>745724</td>
<td>19452798</td>
<td>159064</td>
<td>30357586</td>
</tr>
<tr>
<td>18</td>
<td>104653427012</td>
<td>1999420</td>
<td>55084940</td>
<td>371115</td>
<td>57455475</td>
</tr>
</tbody>
</table>

In the spirit of the works of Leroux et al. [14,15], we will complete this study by investigating symmetry classes of pseudo-hexagons and pseudo-squares. These results help to understand better the combinatorics of the polyominoes that tile the plane by translations and may be useful for deriving a closed formula or a recurrence relation for the number of pseudo-squares or pseudo-hexagons or regular tilings. In this direction Alberto del Lungo and co-authors obtained by using the so-called ECO method the enumeration of parallelogram polyominoes (polyominoes with two non-crossing paths from an origin to an end with only right and up steps) and convex polyominoes [1,8]. The counting of pseudo-square and pseudo-hexagon parallelogram polyominoes begs therefore for a closed formula and Alberto asked us this question in July 2001, but still now we don’t have the method to enumerate such classes of polyominoes.
AN ALGORITHM FOR DECIDING IF A POLYOMINO TILES THE PLANE

References


Communicated by C. Choffrut.
Received February 11, 2005. Accepted April 3, 2006.