The leaf function for graphs associated with Penrose tilings

Carole Porrier
Laboratoire de Combinatoire
et d'Informatique Mathématique (LaCIM)
Université du Québec à Montréal
Montréal (Québec), Canada

Alexandre Blondin Massé
Laboratoire de Combinatoire
et d'Informatique Mathématique (LaCIM)
Université du Québec à Montréal
Montréal (Québec), Canada

***Abstract*—In graph theory, the question of fully leafed induced subtrees has recently been investigated by Blondin Massé *et al.* in regular tilings of the Euclidian plane and 3-dimensional space. The function L_G that gives the maximum number of leaves of an induced subtree of a graph G of order n, for any $n \in \mathbb{N}$, is called leaf function. This article is a first attempt at studying this problem in non-regular tilings, more specifically Penrose tilings. We rely not only on geometric properties of Penrose tilings, that allow us to find an upper bound for the leaf function in these tilings, but also on their links to the Fibonacci word, which give us a lower bound. In particular, we show that $2\varphi n/(4\varphi + 1) \leq L_{kd}(n) \leq \lfloor n/2 \rfloor + 1$, for any $n \in \mathbb{N}$, where φ is the golden ratio and L_{kd} is the leaf function for *kites and darts* Penrose tilings. As a byproduct, a purely discrete representation of points in the tiling, using quadruples, is described.**

I. Introduction

Since Roger Penrose introduced them in 1974 [1] and described them with more details in [2] and [3], Penrose tilings have aroused great interest among many mathematicians, including Robert Ammann and Raphael M. Robinson whose tilings of the plane are also quite interesting (see [4], Sections 10.2 and 10.4). Martin Gardner, known for his popularization of mathematics by proposing recreational puzzles in written media, greatly contributed to making these tilings famous [5], and we owe much of the vocabulary specific to Penrose tilings as well as many of their known properties to John H. Conway. Penrose tilings are so remarkable that Grünbaum and Shephard devoted several sections of their book on tilings to the study of Penrose tilings based on the work of these four researchers and Roger Penrose himself, including unpublished notes and private communications [6], [4]. Furthermore, the algebraic study of Penrose tilings by de Bruijn [7] allowed for the discovery of quasicrystals, that is crystals whose structure is not periodic [8].

Initially, Penrose described three interconnected sets of tiles, called respectively P1, P2 and P3. The set P1 is composed of six prototiles, while the sets P2 and P3 both use only 2 prototiles. The tiles of P2 are called *kites* and *darts* and an example of a tiling by P2 is illustrated in Figure 1. The two tiles of P3 are called *rhombs*, one being thinner and the other thicker. Since those tilings all similar properties [6] but still differ enough to be studied separately, we shall only focus on "kites and darts" tilings in the remaining part of this article.

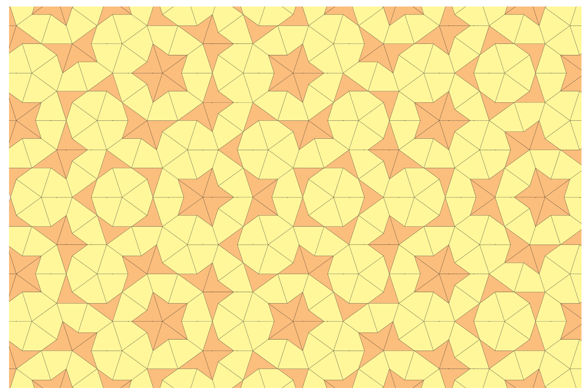


Fig. 1: A Penrose tiling by kites and darts.

Although they are intrinsically geometric objects, Penrose tilings can be viewed as simple graphs, whose vertices are the tiles and whose adjacency relation in the graph is the adjacency relation between tiles in the tiling. In particular, one might be interested in inspecting its subgraphs in order to gain some insight about its global structure. In a recent series of publications, Blondin Massé and his co-authors have studied the fully leafed induced subtrees, i.e. subtrees that maximize the number of leaves with respect to their order, appearing in different families of graphs ([9], [10], [11]). They looked at the four basic periodic lattices: the square, the triangle and the hexagonal lattices in 2D, and the cubic lattice in 3D. One of their results consisted in describing the *leaf function* L in each of these four lattices, i.e. the map associating with each integer $n \geq 2$ the number of leaves of a fully leafed induced subtree of n vertices [10].

This article aims to provide similar results for Penrose tilings, a first step in studying non-periodic tilings, where the situation is more intricate. Although we are not sure whether the induced subtrees built in Section VI are fully leafed, they present remarkable properties and are closely linked to the Fibonacci word. Another interesting observation is that they are caterpillar graphs, i.e. trees such that the removal of all their leaves yields a chain graph. In particular, since they have a linear structure, they might be interpreted as words embedded in graphs. For instance, in [12], the second author

978-1-7281-4129-9/19/$31.00 ©2019 IEEE
DOI 10.1109/GC.2019.00014

of this article and his co-authors established a link between the leaf function of caterpillar graphs and prefix-normal words, a family of words closely related to the bubble sort [13].

The next sections are divided as follows. In Section II, we present the basic definitions and notations on tilings, graphs and words, while Section III is devoted to the known properties of "kites and darts" Penrose tilings. Section IV concerns computational aspects of generating these tilings and their graphs, and describes a discrete representation of the vertices using quadruples. Section V then presents an upper bound for their leaf function and finally, in Section VI, we exhibit special caterpillars having many leaves, therefore providing a lower bound for the leaf function.

II. Preliminaries

We now recall some useful definitions and notations about tilings, graphs and words.

A *tiling* of the plane is a countable family $\mathcal{T} = \{T_n\}_{n\in\mathbb{N}}$ of closed sets of $\mathbb{R}^2$ which is both a *covering* and a *packing* of the plane, that is (1) tiles $T_1, T_2, \ldots$ verify $\bigcup_{n\in\mathbb{N}} T_n = \mathbb{R}^2$ and (2) for all $i \neq j$, $\mathrm{int}(T_i) \cap \mathrm{int}(T_j) = \emptyset$. The intersection of any two tiles is either a set of edges (line segments) and/or vertices (points), or the empty set. Two tiles are called *adjacent* if they share an edge. Two tilings $\mathcal{T}_1$ and $\mathcal{T}_2$ are *congruent* if there is an isometry σ of the plane such that $\mathcal{T}_1 = \sigma(\mathcal{T}_2)$. We say that two tilings $\mathcal{T}_1$ and $\mathcal{T}_2$ are *equal* if there is a similarity τ of the plane (i.e. an isometry followed by a scaling) such that $\mathcal{T}_1 = \tau(\mathcal{T}_2)$. A tiling $\mathcal{T}$ is called *monohedral* if any tile of $\mathcal{T}$ is congruent with a given set T, that is if all tiles have the same size and shape. This set T is called *prototile* and we say that the prototile T *admits* the tiling $\mathcal{T}$. A tiling $\mathcal{T}$ is *dihedral* if any of its tiles T_i is congruent with any of two (non-congruent) prototiles T and T'. This article focuses on Penrose tilings which are dihedral, where the prototiles are a kite and a dart.

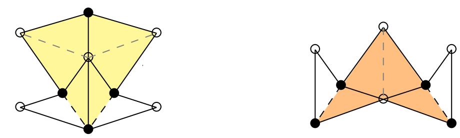
If $\mathcal{T}$ is a tiling, an isometry σ is a *symmetry* of $\mathcal{T}$ if the image through σ of any tile of $\mathcal{T}$ is a tile of $\mathcal{T}$, which also preserves markings and colors when there are any. A tiling is called *symmetric* if its symmetry group is non trivial, that is if it contains at least one symmetry in addition to the identity. A tiling is *periodic* if its symmetry group contains at least two translations in non-parallel directions. If a tiling is not periodic but its symmetries include rotations about a fixed point, then this point is unique and is called the *center* of the tiling. Penrose tilings by kites and darts are all non-periodic but two of them have a center. A *patch* in a tiling is any set of tiles (not prototiles) whose union forms a connected set without hole. The theory of tilings is extensively described in [4], including the study of Penrose tilings.

The vertices and edges of a tiling form an undirected graph, whose dual is the graph we are interested in. Let $G = (V, E)$ be a simple graph. For any $u \in V$ and $U \subseteq V$, the set of *neighbors* of u in G is denoted by $N_G(u)$, which is naturally extended to U by defining $N_G(U) = \{N_G(u') \mid u' \in U\}$. The *subgraph of G induced by U* is $G[U] = (U, E \cap \mathcal{P}_2(U))$, where $\mathcal{P}_2(U)$ is the set of 2-element subsets of U. In this article, the vertices of G are tiles composing a tiling of the plane by Penrose kites and darts, and E is given by the adjacency relation between tiles. The graph thus obtained is infinite but we only consider its finite subgraphs. More precisely, we focus on induced subtrees of fixed order n, that is subgraphs of G which are trees, for all $n \in \mathbb{N}$. We denote $\mathcal{T}_G(n)$ the set of all induced subtrees of order n of G, and $|T|_{\leafNE}$ the number of leaves of any tree T. The leaf function of G is then defined for all $n \in \{1, 2, \cdots, |G|\}$ by

$$L_G(n) = \max\{|T|_{\leafNE} : T \in \mathcal{T}_G(n)\}.$$

An induced subtree T of G of order n is said *fully leafed* if $|T|_{\leafNE} = L_G(n)$. A caterpillar graph C is a tree whose derived tree, obtained from C by removing its leaves, is a path [14].

To conclude this section, we briefly recall some definitions on words. An alphabet $\mathcal{A}$ is a set of symbols (or characters). A word on $\mathcal{A}$ is a sequence of symbols of $\mathcal{A}$, and the set of all words on $\mathcal{A}$ is denoted $\mathcal{A}^*$. We say that a word y is a subword or factor of a word w if there exist words x, z such that $w = xyz$. A binary word is a word on $\{0, 1\}$ or any other 2-element alphabet. The Fibonacci word is an infinite binary word, and the structure of Penrose tilings is guided by "musical sequences", that is factors of the Fibonacci word. More information about words can be found in [15].



(a) Kite (b) Dart

Fig. 2: Decomposition of kites and darts.

III. Penrose Tilings

Tilings by kites and darts were first described by Roger Penrose in [2]. The dart is actually a non-convex kite and in each tile, the longer side is φ times longer than the shorter one, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. The tiles must be arranged according to specific assembly rules, for instance using a marking of the tiles, or arrows on the sides, or two colors for the corners of the tiles (opposite corners then have the same color). Following such rules, there are infinitely many ways in which the tiles can be arranged, so that there are infinitely many tilings by Penrose kites and darts, and as many graphs corresponding to them. Most knowledge on Penrose tilings is compiled in [4]. Many properties of these tilings rely on triangular decomposition: any tile cut along its reflection line gives two isosceles triangles called A-tiles – a kite is divided into two large tiles L_A and a dart into two short tiles S_A. A-tiles can in turn be decomposed into smaller tiles of the same shape, with which smaller kites and darts can be recomposed, as shown in Figure 2. Provided a scaling by $\varphi : 1$, the new kites and darts have the same size as the original ones.

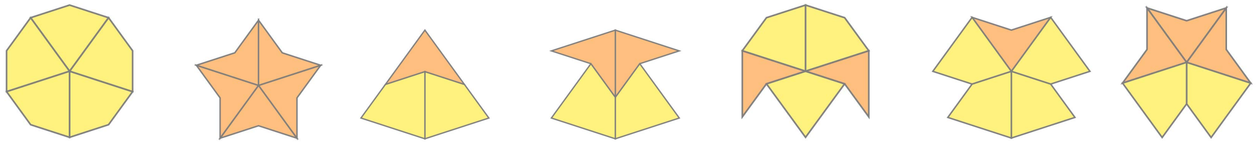
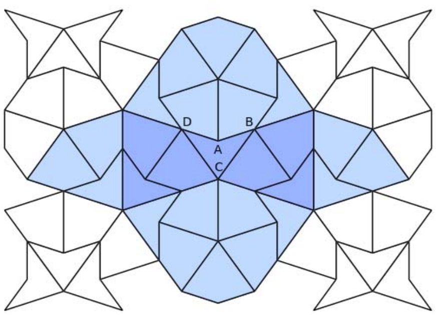
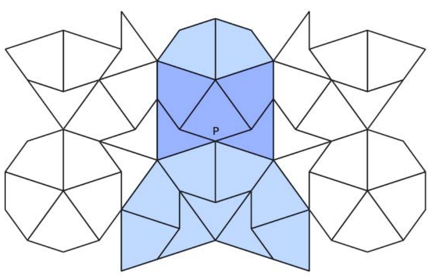


Fig. 3: The seven vertex configurations in a Penrose tiling by kites and darts.

There are seven vertex configurations, that is seven ways in which tiles can be arranged around a vertex of the tiling. These are given in Figure 3. Penrose tilings have the local isomorphism property: any patch of tiles in a tiling by kites and darts (for instance the vertex configurations) is congruent to infinitely many patches in any tiling composed with the same prototiles. Figure 4 shows two small particular patches (darker colored tiles), called (*long* and *short*) *bowties*. A sequence of bowties placed side by side, as the colored tiles in Figure 5, is called a *Conway worm*. Any kites and darts tiling contains arbitrarily long finite worms, that cross each other.



(a) Long bowtie in the queen's kingdom.



(b) Short bowtie in the jack's kingdom.

Fig. 4: Long and short bowties (darker colored tiles) respectively in the queen's and the jack's kingdoms (all colored tiles).

A special marking of the tiles gives three sets of parallel straight lines, in three different directions, called *Ammann bars* (see [4] for more details). Figure 5 shows one of these sets (red lines). The interval distance between any two consecutive bars in the same set can take only two values, such that the longer one, denoted by L, is φ times the smaller one, denoted by S. For a given set of Ammann bars, the sequence of L's and S's is called a *musical sequence*, and is a factor of the Fibonacci word. One significant property is that in a musical sequence we cannot have two S's in a row, nor three L's. Also, for a given length of the sequence, there are only two possible values for the number of L intervals (see Proposition 4). This is why, starting with a given patch, Ammann bars often *force* the placement of some tiles outside the patch if one wishes to extend the tiling in order to cover the whole plane. In particular, each vertex configuration induces Ammann bars that force a (sometimes infinite) number of tiles in the plane and the set of all such tiles, along with the ones in the vertex configuration, is called an *empire*. We call *kingdom* the largest connected subset of an empire, which includes the vertex configuration. For instance, Figure 4 shows the queen's kingdom and the jack's kingdom.

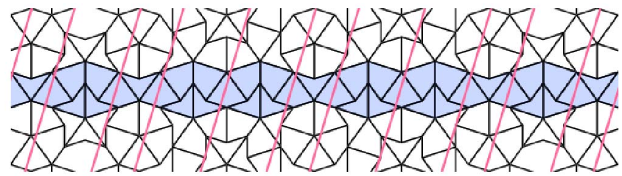


Fig. 5: A set of parallel Ammann bars along a Conway worm.

IV. Computing the Graphs Corresponding to Penrose Tilings

As previously said, the graph corresponding to a Penrose tiling has tiles as its vertices, and a set of edges given by the adjacency relation between tiles (see Figure 6). We call *Penrose tree* an induced subtree in such a graph. If the tree is a caterpillar, we shall simply call it a *Penrose caterpillar*.

There are a few ways to generate a Penrose tiling and several ways to encode the vertices of the tiling and of the graph. Though the projection method suggested by de Bruijn [7] is quite commonly used, we prefer the substitution method – subdividing the tiles like in Figure 2 and then scaling them – because it gives more insight on the structure and we believe it might help us find the leaf function. In the square and cubic lattices, the coordinates of the vertices are simply the tuples of integers (respectively in 2D and 3D) ; in the triangular and hexagonal lattices, the graphs are isomorphic to ones with only regular integer coordinates.

In the case of Penrose tilings, identifying points with respect to a coordinate system is more intricate. Let us call a *Penrose point* any point of $\mathbb{R}^2$ that is the vertex of at least one Penrose tile. Assuming that the plane origin is the vertex of at least

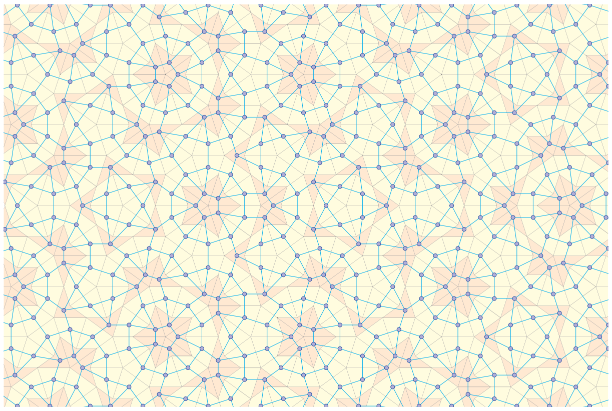


Fig. 6: A paler version of the Penrose tiling of Figure 1 superimposed with its underlying graph in blue.

one tile, it is easy to see that, for any Penrose point p, there exist 20 integers a_k and b_k, for $k = 0, 1, \ldots, 9$, such that

$$p = \sum_{k=0}^{9} (a_k + b_k \varphi) e^{ki\pi/5}. \quad (1)$$

Indeed, when moving along the edges of the tiles, there are 10 possible directions, given by $e^{ki\pi/5}$ for $k = 0, 1, \ldots, 9$, and 2 possible steps, either by a unit or a φ distance. Hence, from a computational perspective, the location of any vertex of a tile can be represented by 20 integers, and translations are simply obtained by coordinate-wise additions and substractions. This representation is not unique but a canonical representation can be derived in virtue of the following theorem.

THEOREM 1 – *For any Penrose point p, there exist four unique integers a_0, a_1, a_2 and a_3 such that*

$$p = \sum_{k=0}^{3} a_k e^{ki\pi/5}.$$

Proof: (Existence) Using Equation 1, we can write

$$p = \sum_{k=0}^{9} (a'_k + b'_k \varphi) e^{ki\pi/5},$$

for some integers a'_k, b'_k, with $k = 0, 1, \ldots, 9$. We claim that the integers a_0, a_1, a_2 and a_3 are given as follows:

$$\begin{aligned} a_0 &= (a'_0 + b'_9 + b'_1) - (a'_5 + b'_4 + b'_6) - a_4, \\ a_1 &= (a'_1 + b'_0 + b'_2) - (a'_6 + b'_5 + b'_7) + a_4, \\ a_2 &= (a'_2 + b'_1 + b'_3) - (a'_7 + b'_6 + b'_8) - a_4, \\ a_3 &= (a'_3 + b'_2 + b'_4) - (a'_8 + b'_7 + b'_9) + a_4 \end{aligned}$$

where $a_4 = (a'_4 + b'_3 + b'_5) - (a'_9 + b'_8 + b'_0)$.

In order to substantiate our claim, we first provide three identities that hold for any integer $k \in \{0, 1, \ldots, 9\}$, where the addition and substraction is taken modulo 10:

(1) $e^{ki\pi/5} = -e^{(k+5)i\pi/5}$, obtained by applying a rotation of angle π;

(2) $\varphi e^{ki\pi/5} = e^{(k+1)i\pi/5} + e^{(k-1)i\pi/5}$, since $\varphi = 2\cos(\pi/5)$ and

(3) $\sum_{k=0}^{4} (-1)^{k+1} e^{ki\pi/5} = 0$, using the relation $\varphi^2 = 1 + \varphi$.

Therefore,

$$\begin{aligned} \sum_{k=0}^{3} a_k e^{ki\pi/5} &= \left(\sum_{k=0}^{4} a_k e^{ki\pi/5}\right) - a_4 e^{4i\pi/5} \\ &= \sum_{k=0}^{4} (a'_k - a'_{k+5}) e^{ki\pi/5} \\ &\quad + \sum_{k=0}^{4} \left(b'_k e^{(k+1)i\pi/5} - b'_k e^{(k+4)i\pi/5}\right) \\ &\quad + \sum_{k=5}^{9} \left(b'_k e^{(k-1)i\pi/5} - b'_k e^{(k-4)i\pi/5}\right) \\ &\quad + \sum_{k=0}^{4} (-1)^{k+1} a_4 e^{ki\pi/5} \\ &= \sum_{k=0}^{4} a'_k e^{ki\pi/5} + \sum_{k=5}^{9} a'_k e^{ki\pi/5} \\ &\quad + \sum_{k=0}^{4} \left(b'_k e^{(k+1)i\pi/5} + b'_k e^{(k-1)i\pi/5}\right) \\ &\quad + \sum_{k=5}^{9} \left(b'_k e^{(k-1)i\pi/5} + b'_k e^{(k+1)i\pi/5}\right) + 0 \\ &= \sum_{k=0}^{9} (a'_k + b'_k \varphi) e^{ki\pi/5}, \end{aligned}$$

as claimed. Notice that the penultimate equality follows from Identities (1) and (3), and the last equality from Identity (2).

(Unicity) It is sufficient to prove that the set

$$\{e^{ki\pi/5} \mid k = 0, 1, 2, 3\}$$

is linearly independent over $\mathbb{Z}$, that is,

$$\sum_{k=0}^{3} a_k e^{ki\pi/5} = 0 \text{ implies } a_k = 0 \text{ for } k = 0, 1, 2, 3.$$

Since $\cos(3\pi/5) = -\cos(2\pi/5)$ and $\sin(3\pi/5) = \sin(2\pi/5)$, one first shows that

$$\begin{aligned} 0 &= \sum_{k=0}^{3} a_k e^{ki\pi/5} \\ &= a_0 + a_1 e^{i\pi/5} + (a_2 - a_3)\cos(2\pi/5) \\ &\quad + i(a_2 + a_3)\sin(2\pi/5). \end{aligned}$$

Next, using the double angle identities for both sin and cos, and after factoring and regrouping, we find

$$\begin{aligned} 0 &= a_0 + a_2 - a_3 + \cos(\pi/5)\,(a_1 + 2(a_2 - a_3)\cos(\pi/5)) \\ &\quad + i\sin(\pi/5)\,(a_1 + 2(a_2 + a_3)\cos(\pi/5)). \end{aligned}$$

But all a_k are integers, and we know that $\cos(\pi/5)$, $\sin(\pi/5)$ are not, which implies $a_2 = a_3$, $a_1 = 0$, $a_0 = 0$ and $a_2 = -a_3$, i.e. $a_0 = a_1 = a_2 = a_3 = 0$, concluding the proof. ■

Let $f : \mathbb{Z}^{20} \to \mathbb{Z}^{20}$ be the function that maps any 20-tuple integers onto $(a_0, a_1, a_2, a_3, 0, \ldots, 0)$, where a_0, a_1, a_2 and a_3 are defined in the proof of Theorem 1. Let p and p' be two Penrose points and $f(p)$, $f(p')$ be the unique 20-tuple obtained by applying f to any 20-tuple representation of p and p' respectively. Then the following problems are easily decidable:

- Are p and p' equal? We only have to check if $f(p) = f(p')$.
- Can we totally order Penrose points? It is sufficient to compute $f(p)$ and $f(p')$ and then use the lexicographical order on the obtained quadruples.
- Are two given tiles equal? This can be decided by checking if their origins are equal and if they are oriented in the same direction.
- Are two given tiles adjacent? Yes if they share two points but are not equal.

Moreover, except for the drawing parts, all computations on Penrose tilings and Penrose graphs can be performed exclusively on integers and are therefore not sensible to floating numerical errors. The reader is invited to look at [16], where similar ideas were used for representing Penrose points by four integers, although the representation is different and is for tilings by Penrose rhombs.

V. Upper bound for the leaf function of Penrose Trees

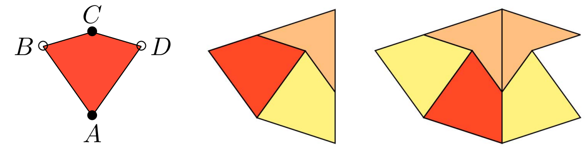
Due to the isomorphism property, graphs associated with kites and darts tilings all have the same leaf function, which we denote by L_{kd}. Let us begin with an observation on L_{kd}:

Proposition 1 – *The leaf function L_{kd} for graphs corresponding to Penrose tilings is non-decreasing.*

Proof: The proof naturally relies on the fact that we consider infinite graphs, but this condition is not enough. Let n be an integer greater than 1 and T a fully leafed induced subtree of order n. Recall that the vertices of T are tiles (that is, subsets of the plane) and consider the convex hull C of the union of these tiles. Then at worst all angles of C have a measure of $4\pi/5$ and only a few configurations have to be examined considering a tile which lies in a corner. If the edge of the tile which lies on the boundary of C is a long one, then adding a tile adjacent to it that is not in C will not create a cycle (due to the angles). Otherwise we have a kite with its short side on the boundary of C, so we can add another kite exterior to C (like in the deuce configuration, without the darts) without creating a cycle. In both cases we obtain a tree of order $n + 1$ with at least as many leaves as T, so $L_{kd}(n + 1) \geq L_{kd}(n)$. ■

Since any tile in a kites and darts tiling is adjacent to four other tiles, the maximum degree of a vertex in a Penrose tree cannot be more than 4. We actually have:

Proposition 2 – *The maximal degree of a vertex in a Penrose tree is 3.*



(a) Kite (b) Case 1 (c) Case 2

Fig. 7: Adjacency constraints on kites in Penrose trees: for a kite $ABCD$ (red), when focusing on the side CD, only two cases are possible.

Proof: For any kite or dart in a Penrose tree T, at most 3 of its adjacent tiles can also be in T. Firstly, any dart lies in an ace for it is the only possible vertex configuration for the vertex in the reflex angle of the dart. It follows that if the dart is in T, then only one among two (adjacent) kites in the corresponding ace can be in T. As for the kites, two cases are to be considered, shown in Figure 7. If we focus on the side CD of the red kite (7a), there is exactly one way to arrange a dart adjacent to it (Case 1) and exactly one way to arrange a kite adjacent to it (Case 2). In Case 1, the yellow kite is forced so that if $ABCD$ is in T then the dart and the yellow kite cannot be both in T. In Case 2, vertex C can only be in a deuce, so that the darts are forced and then so is the kite adjacent to AB. Subsequently, if the red kite is in T then T cannot contain both the tiles adjacent to AB and BC. ■

This constraint on degrees turns out to be useful in finding an upper bound for L_{kd}, and more generally for the leaf function in any graph sharing the same constraint on degrees. To prove proposition 3, we need the following lemma:

Lemma 1 – *Let L be the leaf function of a graph G such that the maximal degree of a vertex in any induced subtree of G is at most 3. Then for any positive integer n, if $L(n) > L(n-1)$ we have $L(n + 1) \leq L(n)$.*

Proof: Let n a positive integer such that $L(n) > L(n - 1)$. Suppose $L(n) < L(n + 1)$ and let T_{n+1} be a fully leafed induced subtree of G of order $n + 1$. Since $L(n) < L(n + 1)$, by deleting a leaf f from T_{n+1} we get a tree T_n whose number of leaves $|T_n|_{\ell}$ is greater or equal to $L(n)$, so $|T_n|_{\ell} = L(n)$ by maximality of the leaf function. Let v be the vertex of T_n adjacent to f in T_{n+1}, then v cannot be a leaf in T_n because $|T_n|_{\ell} = L(n) < L(n+1) = |T_{n+1}|_{\ell}$. Suppose v has degree 0, then T_n is of order 1 so $|T_n|_{\ell} = L(1) = L(0)$, which contradicts the assumption $L(n) > L(n - 1)$. Hence v has degree at least 2 in T_n. Without loss of generality, we can choose f so that v is adjacent to at most one inner vertex of T_{n+1}. If v had degree 3 in T_n then it would have degree 4 in T_{n+1}, which is impossible. Finally if v has degree 2 in T_n, since at most one of its neighbors is an inner one, there is a leaf f' adjacent to v in T_n. Let T_{n-1} be the tree obtained by deleting f' from T_n, then v is a leaf in T_{n-1} so we have $|T_{n-1}|_{\ell} = |T_n|_{\ell} = L(n) > L(n - 1) \geq |T_{n-1}|_{\ell}$, which is absurd. ■

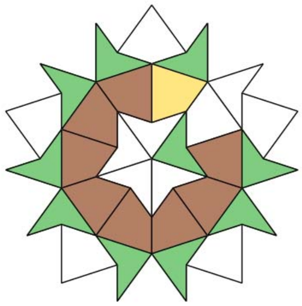
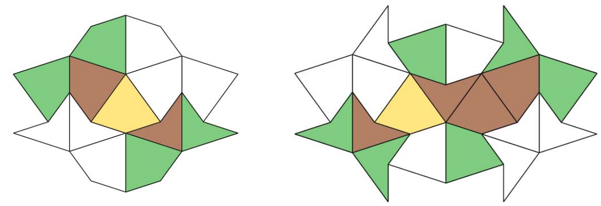


Fig. 8: A fully leafed Penrose tree with 19 vertices and 10 leaves. Note that this is a caterpillar. Tiles are colored with respect to their degrees.



(a) C_S (b) C_L

Fig. 9: Fully leafed Penrose caterpillars in bowtie patches: (a) (respectively (b)) Fully leafed caterpillar in the patch composed with the short (resp. long) bowtie and its adjacent tiles. Tiles are colored with respect to their degrees.

PROPOSITION 3 – *Let L be the leaf function of a graph G such that the maximal degree of a vertex in any induced subtree of G is at most 3. Then for any non-negative integer n, $L(n) \leq \lfloor n/2 \rfloor + 1$.*

Proof: In any graph, the inequality is true for $n \in [0, 1, 2]$. Lemma 1 then gives the inductive step, provided that induced subtrees of orders n and $n+1$ exist: if T_{n+1} is a fully leafed Penrose tree of order $n+1$, then by removing a leaf from T_{n+1} we obtain a tree T_n of order n which has at least $|T_{n+1}|$🍃-1 leaves, so that $L(n) \geq L(n+1) - 1$. The difference between $L(n)$ and $L(n+1)$ is at most 1 and these values are equal for at least half the values of n, hence the result for all $n \in [0..k]$, where k is the greatest integer for which an induced subtree of order k exists. Finally, for all $n > k$, since there is no subtree of order n we have $L(n) = -\infty \leq \lfloor n/2 \rfloor + 1$. ■

With Propositions 2 and 3, we have proved:

THEOREM 2 – *For any non-negative integer n,*

$$L_{kd}(n) \leq \lfloor n/2 \rfloor + 1.$$

Thanks to the algorithm presented in [9] and a particular Penrose tree (see Figure 8), we know that this upper bound is reached for all $n \leq 19$ (except for $n = 1$) and we suspect that it is not realized for $n \geq 20$ but we cannot prove it yet.

VI. PENROSE CATERPILLARS AND LOWER BOUND

As mentioned in Section III, arbitrarily long Conway worms cross Penrose tilings, and by local isomorphism there are infinitely many of them. Figure 5 shows how they correspond to parallel Ammann bars, so that we can use the properties of musical sequences: any subtree whose vertices are all tiles inside the worm is necessarily a path. It is then easy to build caterpillars whose derived paths are inside the worms.

This section gives an increasing family of Penrose trees with the maximum number of leaves among such caterpillars. In the best case scenario, we will have fully leafed Penrose trees and in the worst case scenario, we will have a lower bound for the leaf function, thus complementing the upper bound given in Theorem 2. The idea of caterpillars is suggested by the fact that fully leafed induced subtrees in the triangular and hexagonal lattices are caterpillars, or almost caterpillars, and the maximum degree of a vertex in a Penrose tree is 3 as in these lattices. Recall Figure 4, which shows the queen's kingdom around vertex C and the jack's kingdom around vertex P (colored tiles). We use the observations on kingdoms and bowties to show how tiles adjacent to a worm are forced.

LEMMA 2 – *Any long bowtie lies in a queen's kingdom.*

Proof: In Figure 4a notice that the dart in the center of the long bowtie forces two kites above it, so that the vertex configuration around point A is an ace. Moreover, since this dart is surrounded with two aces in the bowtie, the vertex configurations at points B and D are necessarily jacks, forcing the vertex neighborhood around C to be a queen (using the jack's kingdom). Hence the long bowtie is entirely inside a queen's kingdom. ■

LEMMA 3 – *Any short bowtie lies in a jack's kingdom.*

Proof: In Figure 4b the vertex neighborhood around P can only be a jack, then we can see how the short bowtie is inside the jack's kingdom. ■

Now let us first consider the patch composed with a short bowtie and its adjacent tiles. Figure 9a shows a fully leafed Penrose caterpillar C_S in such a patch: this caterpillar has 7 vertices including 4 leaves, and we know that $L_{kd}(7) \leq 4$ so it is a fully leafed Penrose tree in the whole tiling. In this patch, this is the biggest fully leafed Penrose tree and it is unique up to isometry. Furthermore, it is easily extendable on both sides (left and right) so that the derived path of C_S stays in the worm. In the patch composed with the long bowtie and its adjacent tiles, as shown in Figure 9b, the biggest fully leafed Penrose tree that can be extended on both sides following the worm has 11 vertices and 6 leaves. It is also unique up to isometry and it is a caterpillar which we call C_L.

With both patches side by side it is easy to see what happens when a short and a long bowtie follow each other in a worm: if for instance the short one is on the left, both darts in C_S coincide with the two ones on the left side of C_L, so that we lose two leaves but the resulting subgraph is a tree. Then

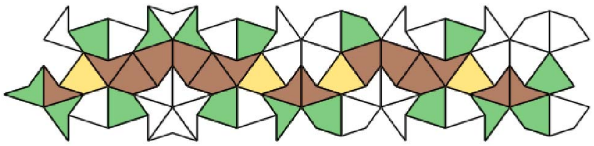


Fig. 10: A Penrose caterpillar along a worm, with 16 leaves and 35 vertices. Tiles are colored with respect to their degrees.

we have a caterpillar of order 16 with $4+6-2=8$ leaves, and the same happens when the short bowtie is placed to the right of the long one (except the tiles that coincide are kites). Note that this caterpillar is not fully leafed, so the function to be presented here is not the leaf function but only a lower bound for it. Since a worm cannot have two consecutive short bowties, the only remaining possible configuration is two long bowties in a row. In this case, one caterpillar is the reflection of the one in Figure 9b and the concatenation, in either way, gives a caterpillar with $2 \cdot 6 - 2 = 10$ leaves and 20 vertices. Figure 10 gives an example of a Penrose caterpillar constructed along a Conway worm.

Let C'_S and C'_L be the caterpillars obtained from C_S and C_L by deleting the leaves on the sides – the ones that eventually coincide with inner vertices during concatenation –, then C'_S has 2 leaves and 5 vertices, and C'_L has 4 leaves and 9 vertices. Let q_S (resp. q_L) be the ratio of number of leaves over order in C'_S (resp. C'_L), then $q_S = 2/5$ is lower than $q_L = 4/9$. So we consider caterpillars that are concatenations of C'_S and C'_L, and we would like them to lie on as many long bowties as possible. Figure 5 shows how a set of parallel Ammann bars determines a sequence of bowties in a worm: each bowtie (long or short) crosses two half L-intervals, and each long bowtie additionally crosses one S-interval between both L's. Thus each long bowtie corresponds to $L+S$ and each short bowtie to L. Hence L corresponds to C'_S which means 2 leaves in 5 tiles, and S corresponds to what is left of C'_L when tiles corresponding to L are removed, that is 2 leaves in 4 tiles.

As a result, we can now focus only on L and S intervals. Since each interval give 2 leaves, we just have to determine the number k of intervals in the caterpillar and how many of them are long. Before completing our main argument, we need a last proposition.

PROPOSITION 4 (PROPOSITION 10.6.10 OF [4]) – *If k is the number of intervals the caterpillar crosses, then the number x_k of L intervals can only take two values: x and $x+1$ such that*

$$\frac{x}{k} < \frac{1}{\varphi} < \frac{x+1}{k} .$$

Hence, we have $x_k = \lfloor k/\varphi \rfloor$ or $x_k = \lceil k/\varphi \rceil$. The correct value will actually be the one that best approximates $1/\varphi$, and when we need to choose, since we would not want to exceed the real value for $L_{kd}(n)$, we take the value that corresponds to less L intervals, that is $x_k = \lfloor k/\varphi \rfloor$.

If n is the order of the caterpillar and k is the number of intervals in it, since we prefer to take a smaller value for k and then add a few tiles to equal n, then we want the greatest k such that

$$5\ \frac{k}{\varphi} + 4\left(k - \frac{k}{\varphi}\right) \leq n$$

that is

$$\frac{k}{\varphi} + 4k \leq n$$

then

$$k\left(\frac{4\varphi+1}{\varphi}\right) \leq n$$

which yields

$$k \leq \frac{\varphi n}{4\varphi+1}.$$

So finally we have

$$k = \left\lfloor \frac{\varphi n}{4\varphi+1} \right\rfloor.$$

The number of leaves over the intervals is then $2k$ and since $4 < (4\varphi+1)/\varphi < 5$ there are at most 5 remaining tiles to arrange at one end of the caterpillar or the other. Let Δ be the number of remaining tiles. If $0 \leq \Delta \leq 2$ then we just have to add Δ leaves to the ends ; if $3 \leq \Delta \leq 4$ the best strategy is to add one tile (a leaf) to an end and 2 or 3 tiles on the other, giving respectively 1 or 2 leaves ; if $\Delta = 5$ then we add 3 tiles/2 leaves to an end and 2 tiles/1 leaf to the other. We have proved:

THEOREM 3 – *Let $\ell_{kd}(n)$ denote the number of leaves of a caterpillar of order n constructed as described above. Then ℓ_{kd} is a lower bound for the leaf function L_{kd} and we have*

$$\ell_{kd}(n) = 2k + \begin{cases} \Delta & \text{if } 0 \leq \Delta \leq 2 \\ \Delta - 1 & \text{if } 3 \leq \Delta \leq 4 \\ 3 & \text{if } \Delta = 5 \end{cases}$$

where $k = \lfloor \varphi n/(4\varphi+1) \rfloor$ and $\Delta = n - 4k - \lceil k/\varphi \rceil$.

For instance, the caterpillar in Figure 10 has $n = 35$ vertices and 16 leaves, and we can verify that $k = 7$ and $\Delta = 2$, which yields $\ell_{kd}(35) = 16$. By removing the 5 rightmost tiles, we would get a caterpillar with 30 vertices and 14 leaves, and we have $\ell_{kd}(30) = 14$. Finally, from Theorems 2 and 3 we have

COROLLARY 1 – *For all $n \in \mathbb{N}$,*

$$2\varphi n/(4\varphi+1) \leq L_{kd}(n) \leq \lfloor n/2 \rfloor + 1.$$

VII. CONCLUSION

The problem studied in this article is at the junction between graphs, tilings and words. It is quite intricate due to the non-periodicity of Penrose tilings and of the Fibonacci word. Still, even though we have not been able to find the leaf function for Penrose tilings yet, we managed to restrict the range of possible values. Since $2\varphi/(4\varphi+1) \simeq 0.433$, the gap between lower and upper bounds for the leaf function is relatively small. We have good hope that we can find L_{kd} or at least a finer lower bound thanks to a particular patch which has more leaves than a Penrose caterpillar of the same order constructed as described in section VI. If the leaf function for kites and darts tilings can be found, a similar procedure could be pursued in order to find the analogous function for Penrose tilings by rhombs, and possibly for other aperiodic tilings which share similar properties.

REFERENCES

[1] R. Penrose, "The Rôle of Aesthetics in Pure and Applied Mathematical Research," *Bulletin of the Institute of Mathematics and its Applications*, pp. 266–271, 1974, reprinted in [8].

[2] ——, "Pentaplexity," *Math. Intelligencer*, vol. 2, no. 1, pp. 32–37, 1978.

[3] ——, "Set of tiles for covering a surface," *United States Patent*, no. 4,133,152, 1979, jan. 9, 1979.

[4] B. Grünbaum and G. C. Shephard, *Tilings and Patterns*. Dover Publications, Inc., 2016, second edition. ISBN 978-0-486-46981-2.

[5] M. Gardner, *Penrose Tiles to Trapdoor Ciphers*, revised ed. The Mathematical Association of America, 1997, iSBN 0-883856-521-6.

[6] B. Grünbaum and G. C. Shephard, *Tilings and Patterns*. W. H. Freeman and Company, New York, 1987, iSBN 0-7167-1193-1.

[7] N. G. de Bruijn, ""Algebraic theory of Penrose's non-periodic tilings of the plane"," *Mathematics Proceedings*, vol. A84, pp. 39–66, 1981, reprinted in [8].

[8] P. J. Steinhardt and S. Ostlund, *The Physics of Quasicrystals*. World Scientific, 1987, collection of reprints. ISBN 9971-50-226-7.

[9] A. Blondin Massé, J. de Carufel, A. Goupil, M. Lapointe, É. Nadeau, and É. Vandomme, "Fully leafed induced subtrees," in *Combinatorial Algorithms*, C. Iliopoulos, H. W. Leong, and W.-K. Sung, Eds. Cham: Springer International Publishing, 2018, pp. 90–101.

[10] A. Blondin Massé, J. de Carufel, A. Goupil, and M. Samson, "Fully leafed tree-like polyominoes and polycubes," in *Combinatorial Algorithms*, L. Brankovic, J. Ryan, and W. F. Smyth, Eds. Cham: Springer International Publishing, 2018, pp. 206–218.

[11] A. Blondin Massé, J. de Carufel, and A. Goupil, "Saturated Fully Leafed Tree-Like Polyforms and Polycubes," *ArXiv e-prints*, Mar. 2018. [Online]. Available: https://arxiv.org/abs/1803.09181v1

[12] A. Blondin Massé, J. de Carufel, A. Goupil, M. Lapointe, É. Nadeau, and É. Vandomme, "Leaf realization problem, caterpillar graphs and prefix normal words," *ArXiv e-prints*, Dec. 2017. [Online]. Available: https://arxiv.org/abs/1712.01942

[13] P. Burcsi, G. Fici, Z. Lipták, F. Ruskey, and J. Sawada, "On prefix normal words and prefix normal forms," *Theoretical Computer Science*, vol. 659, pp. 1 – 13, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0304397516305680

[14] F. Harary and A. J. Schwenk, "The number of caterpillars," *Discrete Mathematics*, vol. 6, no. 4, pp. 359 – 365, 1973. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0012365X73900678

[15] J.-P. Allouche and J. O. Shallit, *Automatic sequences : theory, applications, generalizations*. Cambridge, UK: Cambridge University Press, 2003.

[16] M. W. Reinsch, "Lattice representations of Penrose tilings of the plane," *arXiv e-prints*, pp. math–ph/9 911 024, Nov 1999.