Persistent Patterns in Integer Discrete Circles

André Hoarau and Thierry Monteil

CNRS - Université Montpellier 2 http://www.lirmm.fr/~hoarau http://www.lirmm.fr/~monteil

Abstract. We study patterns that appear in discrete circles with integer center and radius. As the radius goes to infinity, the patterns get closer to digital straight segments: the notion of tangent words (described in Monteil DGCI 2011) allows to grasp their shape. Unexpectedly, some tangent convex words do not appear infinitely often due to deep arithmetical reasons related to an underlying Pell-Fermat equation. The aim of this paper is to provide a complete characterization of the patterns that appear in integer discrete circles for infinitely many radii.

Keywords: discrete circle, asymptotics, digital straight segment, tangent word, Pell-Fermat equation

1 Introduction

Freeman digitization schemes allow to associate a set of pixels to a planar object and a sequence of adjacent pixels to a planar curve in very natural ways; the Freeman code associates a finite word to a sequence of adjacent pixels [6]. The case of straight segments is well understood, and the associated words are known to be the balanced words [8], [11]. We are interested here in the words appearing in the Freeman code of another classical geometrical object: the discrete circles [3], [1], [7], [5].

In 1979, Zenon Kulpa [7] noticed that some "spikes" appear on the diagonal for arbitrary big radii in the Grid Intersect Quantization digitization of integer circles. For big radii, such spikes look unnatural since, as the curvature of circles of big radii goes to zero, we expect the digitization of the big circle to look locally like digital straight segments. Unfortunately, Kulpa could not go further since his remark is based on a visual description about the angle between three consecutive pixels.

In this paper, we shall use another digitization scheme, namely the Square Box Quantization (SBQ) described in [6] (see also Section 2.1). As we shall see, the same phenomenon appears in this case (Fig. 1). It turns out that many finite patterns not corresponding to digital straight segments do appear in integer circles for arbitrary large radii. For example, the patterns of Fig. 2 could be considered as the "next spikes", (b) is obtained by stretching (a), (c) is obtained by shearing (a).

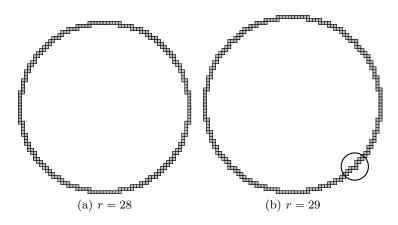


Fig. 1. The spike coded by 0011 appears in the digitization of big integer circles

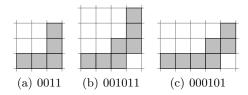


Fig. 2. The most elementary spikes

A word is said to be *persistent* if it appears in the Freeman code of integer discrete circles for infinitely many radii. Their complete description is the aim of this paper.

As we can consider the dual point of view where the circle radius does not grow but the grid mesh vanishes, we know that persistent words are tangent convex words, which were introduced in [9]. Experimentations lead us to the observation of an unexpected behaviour: while most tangent convex words seem to be persistent, some of them are not (Fig. 3).

As we shall see, the difference between both examples of Fig. 3 relies on a deep arithmetical reason. We can define a rational slope p/q for non-balanced tangent convex words, and say that it is *Pythagorean* if $p^2 + q^2$ is a square. Knowing whether a tangent convex word is persistent or not is related to a system of Pell-Fermat inequalities whose main parameter depends on the slope of the word. For example, the word (b) is not persistent since its slope 3/4 is Pythagorean $(3^2 + 4^2 = 5^2)$. We will prove the following

Theorem 1. A word is persistent if, and only if, it is tangent convex with a non-Pythagorean slope. In particular, balanced words are persistent.

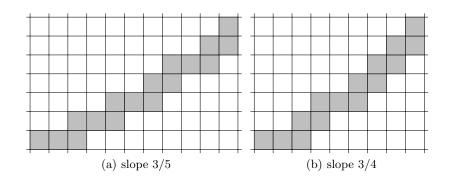


Fig. 3. Two tangent convex words: (a) is persistent, (b) is not

2 Preliminaries and tools

2.1 Framework

Discrete geometry introduces various schemes to associate a set of pixels to a plane continuous object. When the object is a curve, the pixels are ordered and the Freeman code associates to such a sequence a word over the alphabet $\mathbb{Z}/4\mathbb{Z}$ or $\mathbb{Z}/8\mathbb{Z}$ depending on whether two consecutive pixels have to share an edge or only a vertex. For example, the Freeman codes of digital straight segments are known to be precisely the *balanced* words using two consecutive letters (see e.g. [8], [11]), no matter the code or the digitization scheme. The case of discrete circles is wilder, there are various ways to define a discrete circle: the set of pixels can be described in an algorithmic way like in [3] or [13], it can also be described as the set of solutions of some analytic equation like in [1] or [5].

The model we are considering is the Square Box Quantization (SBQ) of integer circles [6]. The SBQ of a curve is the set of pixels that it intersects. An integer circle is a circle with integer radius r and centered at (0, 0). It is denoted by C(r).

Note that an integer circle cannot meet any pixel vertex since such points have coordinates of the form (p + 1/2, q + 1/2) for some $(p, q) \in \mathbb{Z}^2$ and $(p + 1/2)^2 + (q+1/2)^2$ cannot be an integer. In such a non-ambiguous case, the Square Box Quantization is also known as the *supercover* [4] or the *standard model* [2].

The Freeman code associated to the Square Box Quantization of a curve γ also corresponds to what dynamicists call the *cutting sequence* $c(\gamma)$: it is obtained by reading the letters associated to the edges (see figure on the right) of the pixels that γ intersects along the time. Such a word defines a *pattern*, that is, a finite set of pixels defined up to an integer translation.



Hence, to see if a word u belongs to the cutting sequence of an integer circle of radius r, we have to ensure that, for some integer translation, each edge of the associated translated pattern corresponding to each letter is crossed by the

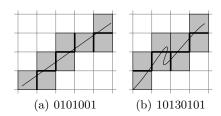


Fig. 4. Examples of cutting sequences, and their associated pattern (in gray)

circle. This is done by checking that, for each such edge, the distance from the center to one of the vertices of the edge is less than r, and the distance from the center to the other vertex is greater than r. Those vertices play a central role and are called *control points* (see Fig. 9).

Since the center of an integer circle is also the center of a pixel, the global symmetries allow us to restrict our study to the seventh octant, where the Freeman code only uses the letters 0 and 1, and the slope is between 0 and 1.

2.2 Tangent convex words

Tangent convex words were described in [9] Section 4.5. We recall their structure in a self-contained way and explain the geometric conditions that ensure their existence as a persistent word.

Let \mathcal{C}^+ denote the set of *convex curves*, that is, the set of smooth regular curves with positive curvature. When γ belongs to \mathcal{C}^+ , we denote by $T(\gamma)$ the set of words that appear in the coding of γ with grids of arbitrary small meshes:

$$T(\gamma) = \bigcap_{\epsilon > 0} \bigcup_{mesh(G) \leqslant \epsilon} F(\gamma, G) \tag{1}$$

where $F(\gamma, G)$ denotes the set of factors appearing in the Freeman code of γ by the grid G. A word u is *tangent* to a curve γ if $u \in T(\gamma)$. The *tangent convex words* are all the words which are tangent to some convex curve: $T(\mathcal{C}^+) = \bigcup_{\gamma \in \mathcal{C}^+} T(\gamma)$. A non-balanced tangent convex word is called a *spike*.

The relationship with integer circles is the following: up to renormalization, we can consider that we code a single unit circle centered at (0, 0) with grids of meshes 1/n $(n \in \mathbb{N}^*)$. Since the unit circle (parametrized counterclockwise) belongs to \mathcal{C}^+ , we deduce that the persistent words of integer circles form a subset of the set of tangent convex words $T(\mathcal{C}^+)$. This allow us to restrict our search among this set, which is well structured and has small complexity: the number of elements of $T(\mathcal{C}^+)$ of length n is equivalent to $n^3/6$ [10].

Let $\gamma = (\gamma_x, \gamma_y) : [0, 1] \to \mathbb{R}^2$ be a convex curve. With almost no loss of generality and since it corresponds to the seventh octant assumption done in

Subsection 2.1, we assume that γ is going in the North and East direction, that is, $\forall t \in [0,1], \ \gamma'_x(t) > 0$ and $\gamma'_y(t) > 0$. The coding of such curves only uses the letters 0 and 1 (Fig. 5). For any $t \in [0,1]$, we denote the slope of γ at t by $\rho(\gamma'(t)) = \frac{\gamma'_y(t)}{\gamma'_x(t)}$.

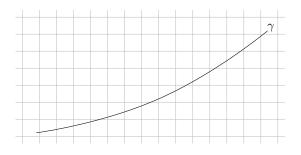


Fig. 5. u = 000010001001001001001

Let u be a tangent word of γ : there exists a sequence (G_n) of grids whose meshes go to 0 and such that $\forall n \in \mathbb{N}, u \in F(\gamma, G_n)$. In particular, for any integer n, there exist two sequences $(t_n^1), (t_n^2)$ in [0, 1] such that u is the Freeman code of $\gamma|_{[t_n^1, t_n^2]}$ with respect to the grid G_n . Up to taking a subsequence (the segment [0, 1] is compact), we can assume that (t_n^1) and (t_n^2) both converge to some $t \in [0, 1]$: we say that $\rho(\gamma'(t))$ is a slope of u (it may not be unique since u can be tangent to various curves).

Slope different from 1: desubstitution. Assume that $\rho(\gamma'(t)) < 1$. This implies that the word 11 is not a factor of u. We can use this information to construct a word $\delta_0(u)$ from u such that $|\delta_0(u)| < |u|$ and $\delta_0(u) \in T(\delta_0(\gamma))$, where $\delta_0(\gamma)$ is another convex curve and T is defined by equation 1. For each n, we add diagonal edges to the grid G_n which we label by the letter 5. Hence, we can associate a new cutting sequence u' by inserting the letter 5 between any two consecutive 0 (Fig. 6).

Then, we remove the vertical lines of the grid, and associate the corresponding word u'' obtained from u' by removing the letter 0 (Fig. 7).

We then renormalize the parallelogram grid back to G_n by applying the shear matrix

$$M_0 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1}.$$

The word obtained by replacing the letter 5 by 0 in u'' is denoted by $\delta_0(u)$ and is a factor of $F(\delta_0(\gamma), G_n)$, where $\delta_0(\gamma) = M_0 \circ \gamma$ is the curve obtained by applying the matrix M_0 to γ (Fig. 8).

The matrix M_0 corresponds to a linear bijective bi-uniformly continuous map, it therefore preserves the regularity of the curve. Moreover, since it has positive

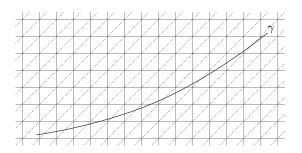


Fig. 6. u' = 0505050105050105010501010501

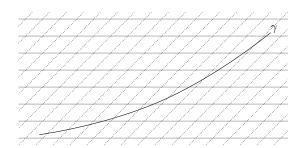


Fig. 7. u'' = 55515515151151

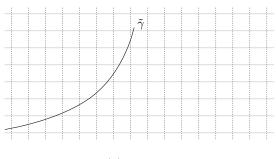


Fig. 8. $\delta_0(u) = 00010010101101$

determinant, the sign of the curvature of the curve is preserved: $\delta_0(\gamma)$ belongs to \mathcal{C}^+ . All the operations are reversible, hence u is in $T(\gamma)$ if, and only if, $\delta_0(u)$ is in $T(\delta_0(\gamma))$.

In the same way, if $\rho(\gamma'(t)) > 1$ then 00 is not a factor of u, we insert a 5 between two consecutive 1, and renormalize to obtain a word $\delta_1(u)$, and the matrix used for the renormalization is

$$M_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1}.$$

Slope equal to 1: diagonal words. If $\rho(\gamma'(t)) = 1$, we say that the word u is *diagonal*. For large n, the curve γ cannot intersect more than one integer diagonal (a line of slope 1 passing through a corner of G_n). As explained in [9], the diagonal word can oscillate a lot for smooth curves, but in the case of a convex curve, the oscillation is very limited: for n large enough, the curve γ intersects the integer diagonal at most twice between t_n^1 and t_n^2 (Fig. 9).

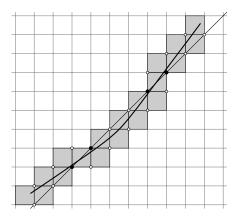


Fig. 9. A diagonal word and its principal (black) and secondary (white) control points

When the number of such intersections is less than 2, we get a balanced word. When this number is 2, the first intersection corresponds to an occurrence of 00 in u, and the second corresponds to an occurrence of 11 (in particular the slope of u is unique and must be equal to 1). The remaining of the diagonal word u of $T(\mathcal{C}^+)$ is made by an alternation of 0 and 1, hence it is completely determined by the three lengths: before 00, between 00 and 11, after 11.

The grid points on the integer diagonal surrounding the two intersections are called *principal control points* because they contain all the information concerning the localization of the convex curve coded by u for n large enough. The other control points are said to be *secondary* (the satisfaction of those control points — that is, the satisfaction of the equations derived from those control points — is guaranteed by the fact that the curve is convex or by the fact that, when n is large enough, the curve becomes very close to the integer diagonal).

The minimal non-balanced diagonal words of $T(\mathcal{C}^+)$ are of the form $u = 00(10)^{\ell}11$ for some integer ℓ which is called the *width* of the spike u. Any non-balanced diagonal word of $T(\mathcal{C}^+)$ can be extended in a single way to the left and to the right, by adding alternately a 0 or a 1. Diagonal words of $T(\mathcal{C}^+)$ are the words recognized by the non-deterministic automaton depicted in Fig. 10, where all states are considered as initial and final.

The two transitions without return correspond to the two possible intersections with the integer diagonal, i.e. to the occurrences of 00 and 11. Both

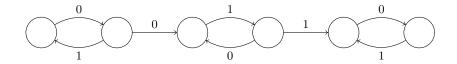


Fig. 10. Automaton recognizing diagonal words of $T(\mathcal{C}^+)$

transitions are used in recognizing u if, and only if, u is not balanced. The width ℓ corresponds to the number of central loops made during the recognition of u.

Iteration. Those considerations allow us to recognize the words of $T(\mathcal{C}^+)$: given a word u, we apply δ_0 (resp. δ_1) to u as long as 11 (resp. 00) does not appear in u. The length of the considered words is strictly decreasing, hence after finitely many steps, we reach one of the following two cases:

- 1. the word is empty. Hence, the initial word u is balanced [8], therefore it belongs to $T(\mathcal{C}^+)$.
- 2. 00 and 11 both appear in the word: it suffices to check that it is recognized by the automaton of Fig. 10 to conclude that the initial word u is in $T(\mathcal{C}^+)$.

This provides a desubstitution algorithm which recognizes the words of $T(\mathcal{C}^+)$.

Note that when a word $u \in T(\mathcal{C}^+)$ is not balanced, then the iterated desubstitution leads to a diagonal word where both 00 and 11 appear: the slope of uis then unique, and independent of any supporting curve γ , time t or sequence of grids (G_n) . It can be computed using the continued fraction associated to the alternation of δ_0 and δ_1 leading to the diagonal word whose slope equals 1. For balanced words however, the set of possible slopes is an interval of nonempty interior.

Conversely, we can generate the words of $T(\mathcal{C}^+)$ by going backward in the algorithm: we construct a diagonal word by describing a path in the automaton of Fig. 10. Then we apply one of the following operations, as much as we want:

- 1. add a 0 to the right of each 1,
- 2. add a 0 to the left of each 1,
- 3. add a 1 to the right of each 0,
- 4. add a 1 to the left of each 0.

The words of $T(\mathcal{C}^+)$ are the factors of the words we obtain that way. It is important to notice that the primary control points can be followed during this construction by applying to the initial primary control points the matrix M_0^{-1} in the first two cases and the matrix M_1^{-1} in the last two cases. Hence, we know how to localize a curve around a word of $T(\mathcal{C}^+)$ of any slope.

2.3 Generalized Pell-Fermat equations

A generalized Pell-Fermat equation is a Diophantine equation of the form $x^2 - Dy^2 = K$, where D is a positive integer parameter, K is an integer parameter, and x and y are the integer unknowns. When D is not a square, and when the equation admits at least one solution, then it admits infinitely many solutions [12]. The set of solutions is a finite union of geometric progressions.

3 Main result

Theorem 1. A word is persistent if, and only if, it is tangent convex with a non-Pythagorean slope. In particular, balanced words are persistent.

We shall first deal with the spikes, which are somehow more rigid (they have a single slope and can be studied through a few principal control points), and then we will extend the obtained results to balanced words.

3.1 Spikes crossing the octants: slopes 0 and 1

Due to the lack of space, the description of persistent spikes for the extremal slopes 0 (no spike of the form $0^k 30^\ell 10^m$ is persistent) and 1 (all spikes of the form $(01)^k 00(10)^\ell 11(01)^m$ are persistent) is postponed, we will concentrate on the general case of slopes in the open interval (0, 1) which contains all the ideas. We can still notice that the behaviour of those two particular cases is coherent with this general description, since 0/1 is Pythagorean and 1/1 is not.

3.2 Spikes with Pythagorean slopes in (0,1)

Lemma 1. A spike with a Pythagorean slope in (0,1) is not persistent.

Proof. Let u be a spike of slope p/q, such that 0 < p/q < 1 and $p^2 + q^2 = s^2$ for some integer s. We want to prove that u is not persistent. Assume by contradiction that for an increasing sequence (r_n) of positive integers, u appears in the cutting sequence of the circle $C_n = C(r_n)$.

The position of the principal control points a_n, b_n, c_n, d_n of u relative to C_n is depicted in Fig. 11 (it is obtained from Fig. 9 by applying shear matrices).

Since C_n crosses both segments $[a_n, b_n]$ and $[c_n, d_n]$, the line of equation y = -(q/p)x crosses the segment $[a_n, d_n]$, so we have a rough localization of the pattern associated to u. On this line, the points of the form (pi, -qi) where i is an integer are at distance $\sqrt{p^2 + q^2}$ from each other. Hence, for each n, there is a point of coordinate $(pi_n, -qi_n)$ in the rectangle whose base is $[a_n, d_n]$ and whose height has length $\sqrt{p^2 + q^2}$ (depicted in gray on the figure).

Let (x_n, y_n) be the vector from $(pi_n, -qi_n)$ to the point a_n . Since $(pi_n, -qi_n)$ has integer coordinates and stays in a bounded rectangle around the pattern associated to u, the vector (x_n, y_n) takes only finitely many values. Up to taking a subsequence, we can assume that it is constant. Since $(pi_n, -qi_n)$ is the center

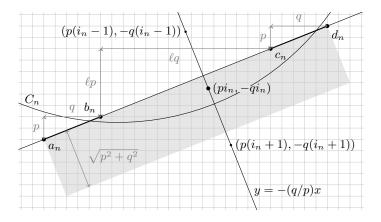


Fig. 11. Situation around the pattern associated to u for C_n

of a pixel and a_n is a corner of a pixel, there exists integers X and Y such that, for any n, $(x_n, y_n) = (X + 1/2, Y + 1/2)$.

Since the point a_n is outside the circle C_n , we have

$$r_n^2 \le (pi_n + X + 1/2)^2 + (-qi_n + Y + 1/2)^2$$

which can be rewritten as $R_n^2 - I_n^2 \leq K_a$, where $R_n = 2sr_n$, $I_n = 2s^2i_n + (2X+1)p - (2Y+1)q$ and $K_a = ((2X+1)p - (2Y+1)q)^2 + (X+1/2)^2 + (Y+1/2)^2$.

The same argument with the point b_n inside the circle C_n leads to an inequality of the form $K_b \leq R_n^2 - I_n^2$, where K_b is another constant (R_n and I_n are the same thanks to a nice cancellation).

Now, both sequences (R_n) and (I_n) are nonnegative integer sequences tending to infinity. Since the difference between their squares is bounded, it eventually equals zero. Hence, for n large enough, $R_n = I_n$. Since (p, q, s) is a primitive Pythagorean triple, exactly one of p or q is odd, hence I_n is odd. Unfortunately, R_n is even, a contradiction.

3.3 Spikes with non-Pythagorean slopes in (0,1)

Lemma 2. A spike with a non-Pythagorean slope in (0,1) is persistent.

Proof. Let u be a spike of slope p/q, such that 0 < p/q < 1 and $p^2 + q^2$ is not a square. Let ℓ be its width. We want to prove that u is persistent.

We start by finding an integer circle C which is well located with respect to the principal control points of u, that is, we require that C crosses both segments [a, b] and [c, d] (we do not care about the secondary control points, hence u need not be a factor of the cutting sequence of C but this will become true for larger circles).

Let c_1 denote the center of the pixel located down-left of the control point b (see Fig. 12), and let C_1 denote the circle centered at c_1 with radius 1. For any

positive integer i, let c_i denote the integer point $c_1 + (i - 1, 0)$, and let C_i denote the circle with center c_i and radius i.

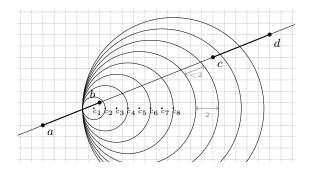


Fig. 12. Construction of an integer circle satisfying the principal control points of u

By construction, each C_i is an integer circle that crosses the segment [a, b]. There is another intersection between C_i and the line defined by (a, d): from i to i+1, this intersection goes forward in the direction from a to d and the distance between two intersections is a positive constant strictly less than 2. Since the length of the segment [c, d] is $\sqrt{p^2 + q^2} > \sqrt{5} > 2$, there is an integer i such that C_i crosses [c, d]. Let $c = c_i$ and $C = C_i$. The integer circle C satisfies the control points of u.

Now, for any integer radius r, we try to find a solution by locating the pattern associated to u in such a way that the point c coincides with some (pi, -qi) where i is an integer. Similarly to the case of Pythagorean slope, the fact that the circle of radius r is well located with respect to the control points is equivalent to the existence of some integer i such that

$$\max(K_b, K_c) \le 4(p^2 + q^2)r^2 - I^2 \le \min(K_a, K_d)$$

where $I = 2(p^2 + q^2)i + (2X + 1)p - (2Y + 1)q$ and K_a , K_b , K_c , K_d are some constants.

Hence, the radii for which the circles are well located with respect to the principal control points of u and such that the point c corresponds to some (pi, -qi) satisfy a Pell-Fermat inequation, whose set of solutions is the union of the solutions of the family of generalized Pell-Fermat equations: $4(p^2 + q^2)r^2 - I^2 = K$ for K in the set $[\max(K_b, K_c), \min(K_a, K_d)] \cap \mathbb{Z}$.

By construction of C, the pair $(r_0, 0)$ is a particular solution of one of those equations. Hence, since $D = 4(p^2 + q^2)$ is not a square, Subsection 2.3 ensures that that equation has infinitely many solutions: there are infinitely many radii r such that the circle C(r) is well located with respect to the principal control points of u. When such an r is large enough, the secondary control points become automatically satisfied as well, hence u appears in the cutting sequence of C(r).

Therefore, u is persistent.

3.4 Balanced persistent words

Lemma 3. Every balanced word is persistent.

Proof. Let u be a balanced word. As already noticed, the set of slopes associated to u contains a open interval I, hence, by density, there exists a non-Pythagorean rational number p/q in I (it suffice to ensure that p and q are odd). Then, the word u can be extended to a spike of slope p/q, which is persistent. Hence, u is persistent, since it is the factor of a persistent word.

Proof of Theorem 1. Since persistent words are tangent convex, since the tangent convex words are either spikes or balanced, we just proved the Theorem 1.

Decision algorithm. As explained in [9], deciding whether a word belongs to $T(\mathcal{C}^+)$ can be done in linear time. The computation of the slope can be done along the process described in Section 2.2, "Iteration", hence deciding whether a given word is persistent can be done in linear time.

References

- Eric Andres. Discrete circles, rings and spheres. Computers & Graphics 18, 5:695– 706, 1994.
- Eric Andres. Defining discrete objects for polygonalization: The standard model. In DGCI, pages 313–325, 2002.
- Jack Bresenham. A linear algorithm for incremental digital display of circular arcs. Commun. ACM, 20(2):100–106, 1977.
- Daniel Cohen-Or and Arie Kaufman. Fundamentals of surface voxelization. Graphical Models and Image Processing, 57(6):453 – 461, 1995.
- Christophe Fiorio, Jean-Luc Toutant, and Damien Jamet. Discrete Circles: an arithmetical approach with non-constant thickness. In Vision Geometry XIV, IS&T/SPIE 18th Annual Symposium Electronic Imaging, pages 60660C01– 60660C12. SPIE, January 2006.
- Herbert Freeman. Computer processing of line-drawing images. Computing Surveys, 6:57–97, March 1974.
- Zenon Kulpa. On the properties of discrete circles, rings, and disks. Computer Graphics and Image Processing, 10(4):348–365, 1979.
- 8. M. Lothaire. Algebraic combinatorics on words, volume 90 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 2002.
- Thierry Monteil. Another definition for digital tangents. In Discrete geometry for computer imagery, volume 6607 of Lecture Notes in Comput. Sci., pages 95–103. Springer, Heidelberg, 2011.
- Thierry Monteil. The complexity of tangent words. In Petr Ambroz, Stepan Holub, and Zuzana Masáková, editors, WORDS, volume 63 of EPTCS, pages 152–157, 2011.
- 11. N. Pytheas Fogg. Substitutions in dynamics, arithmetics and combinatorics, volume 1794 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2002.
- 12. John P. Robertson. Solving the generalized Pell equation $x^2 Dy^2 = N$. Preprint available at http://www.jpr2718.org/pell.pdf, 2004.
- Laure Tougne. Circle digitization and cellular automata. In Proceedings of the 6th International Workshop on Discrete Geometry for Computer Imagery, DCGA '96, pages 283–294, London, UK, 1996. Springer-Verlag.