

# CRITICAL CURVES OF A PIECEWISE LINEAR MAP

JOHN A G ROBERTS, ASAKI SAITO, AND FRANCO VIVALDI

ABSTRACT. We study the parameter space of a family of planar maps, which are linear on each of the right and left half-planes. We consider the set of parameters for which every orbit recurs to the boundary between half-planes. These parameters consist of algebraic curves, determined by the symbolic dynamics of the itinerary that connects boundary points. We study the algebraic and geometrical properties of these curves, in relation with such a symbolic dynamics.

## 1. INTRODUCTION

Piecewise linear (affine) maps are maps defined on a partitioned phase space where a different linear (affine) map acts on each region of the partition. Often this can be achieved to maintain continuity of the map. Dissipative versions arise naturally in engineering and physical models and the nature of the attractors in them and the possible bifurcations have received considerable attention [4].

Less well-studied is the conservative case; in particular, regular orbits in piecewise linear symplectic maps are not well understood. A much studied two-parameter family on the plane is [3, 8, 9, 14–16]

$$(1) \quad F(x, y) = \begin{cases} (ax - y, x) & x > 0 \vee (x = 0 \wedge y \leq 0) \\ (bx - y, x) & \text{otherwise} \end{cases}$$

where  $a$  and  $b$  are real parameters. (Our definition of  $F$  on the line  $x = 0$  is a variant of that found in the literature.) The map  $F$  can lay claim to being the normal form for a piecewise linear map acting on the partition into the left and right half-planes [9].

The map  $F$  sends rays through the origin into themselves, and the ray dynamics is a smooth circle map  $f$ . The purpose of this work is to study the symbolic dynamics of  $f$  with respect to the above binary partition of the plane, extending the works [14, 15] on the system (1), and complementing the works [9, 24–26] on mode-locking and bifurcations in piecewise-linear maps. This work is the first part of a planned broader study [23].

Each orbit of  $F$  corresponds to developing a product or *word* in the matrices  $A = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$  and  $B = \begin{pmatrix} b & -1 \\ 1 & 0 \end{pmatrix}$ , following the symbol sequence of the dynamics between the left and right-half planes. (We use roman fonts for the real parameters  $a, b$  and italic fonts for the symbols  $a, b$  of the corresponding half-planes; the latter appear as letters in words or indeterminates in algebraic expressions.) Relevant observables are the frequencies of the two type of matrices in the developing word, which are orbit-dependent. In this sense,  $F$

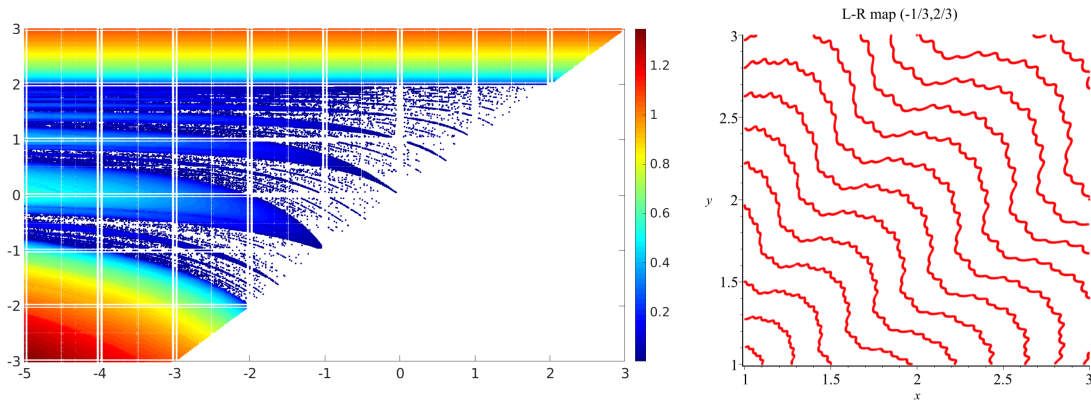


FIGURE 1. Left: the parameter space of the map (1), with mode-locking regions colour-coded by the escape rate (only the range  $a \leq b$  is shown). Right: some orbits for the parameter pair  $(a, b) = (-1/3, 2/3)$ , seemingly conjugate to an irrational rotation with symbolic complexity  $\mathcal{K}(n) = 2n$ .

generalises a model in condensed matter physics, the discrete Schrödinger equation on a one-dimensional lattice (see, e.g., [22] and references therein). Here one propagates the solution along the lattice by developing a matrix word in the aforementioned matrices  $A$  and  $B$  (with  $a$  and  $b$  being physically relevant parameters) according to a two-letter substitution rule which prescribes the asymptotic frequencies of  $A$  and  $B$  in the word. The case of quasiperiodic words (e.g., the Fibonacci sequence) has received much attention.

Matrix words of the type described also arise naturally in the classical study of continued fraction expansions in number theory where  $a$  and  $b$  are taken to be integers [19]. Nevertheless, the theory of *continuant polynomials* which is used to study continued fractions can equally well be applied to study the dynamics of  $F$  with  $a, b \in \mathbb{R}$  [2, 10, 19].

Our knowledge of parameter space of  $F$  is limited to four countable families of algebraic curves  $C(a, b) = 0$ , where  $C$  is a polynomial with integer coefficients. The first two families are known explicitly, while the other two are constructed using certain finite symbolic codes, whose general form is still unavailable. They are:

- i) An infinite family of line segments (together with their image under parameter exchange<sup>1</sup>), where the rotation number of  $f$  is known [3, 14].
- ii) An infinite family of hyperbolae, invariant under parameter exchange, where the map  $f$  is of finite order and the rotation number is constant [3, 14].
- iii) A family of algebraic curves defined by a dynamical condition, for which the map  $F$  is known to support invariant curves with rational or irrational rotation number, consisting of finitely many arcs of conic sections glued together [15].

<sup>1</sup>Exchanging  $a$  and  $b$  in (1) leads to a conjugate system, so it suffice to study the case  $a \leq b$ .

- iv) A family of algebraic curves defined by a dynamical condition, bounding the mode-locking regions [9, 24–26].

A central question regards the existence and properties of quasi-periodic orbits whose closure are topological circles, special cases of which occur for parameters of type i) and iii). General results are scarce. M. Herman’s work in the broader setting of the Froeschlé group [11, Theorem VIII.5.1] implies that any map  $F$  having an irrational rotation number  $\theta = \theta(a, b)$  with bounded partial quotients in its continued fraction expansion is topologically conjugate to a rotation of the plane, and hence has invariant circles. This set of parameter values has zero two-dimensional Lebesgue measure. The results of [15] on curves of type iii) are complementary, as they regard one-dimensional sets of parameters (see below). Numerical experiments suggest ubiquity of parameter pairs for which all orbits are dense on non-smooth curves, as in figure 1, right.

The present work is devoted to the development of a theory of the curves of type iii), which we call the **critical curves** in parameter space. They are defined by the condition that an initial boundary ray (the positive or negative ordinate semi-axis in  $\mathbb{R}^2$ ) be sent by the map  $F$  to the opposite boundary ray in a prescribed number of iterations. If the parameters  $a, b$  belong to critical curve, and if the rotation number of the corresponding circle map  $f$  is irrational, then the phase space of  $F$  foliates into piecewise smooth invariant curves, consisting of arcs of conic sections joined together [15, theorem 2.2]<sup>2</sup>.

We now summarise the contents and main results of this paper. In section 2 we provide some background on the symbolic dynamics of rotations and on the circle map  $f$ . We then give some preliminary results for the dynamics of  $F$  on a half-plane and its relation to the parameter. In the following section we introduce the main object of study, the critical curves, which correspond to the occurrence of Sturmian-type words in the symbolic dynamics (see [21, section 6]). Using the theory of continuants, we determine some general algebraic properties of these curves (propositions 5,7, section 3.1). We link the symbolic dynamics to the time-reversal symmetry of the map  $F$  (theorem 6), and show that critical curves have several disjoint branches, whose number is determined by the code. We establish that only one branch —indeed, only a part of it— is relevant to the dynamics of the map  $F$  (theorems 8 and 10), meaning that the code that defines the curve is the symbolic dynamics of a segment of an actual orbit of the map  $F$ , with initial condition of the appropriate boundary ray. Subsequently (section 3.2) we consider functions defined over critical curves, and derive formulae for the Poisson brackets of two curves (theorem 9), for later use.

In section 4 we consider intersections of critical curves, which lead to periodic orbits. We introduce the concept of intersection sequence, with which we classify the intersections of a curve with curves of lower degree generated by sub-words (theorem 11). We construct an associated geometrical object, called the polygonal of the curve, by means of which we formulate sufficient conditions for the transversality of these intersections (theorem 14). We also formulate conditions under which intersections delimit the portion of a curve which has dynamical significance.

---

<sup>2</sup>The authors of [15] do not introduce explicitly the notion of critical curve.

In section 5 we consider the first generation of curves of type iii). These are obtained by concatenating the symbolic words of two curves of type i), and then allowing for repetitions of the concatenated word (theorem 18).

Identities of Chebyshev polynomials are collected in an appendix.

## 2. BACKGROUND

**2.1. Rotational words.** For references for this section, see, e.g., [1, 21].

The **rotational words** are the symbolic dynamics of rotations with respect to a two-element partition of the circle. A rotational orbit has the form  $x_t = \{\theta t + x_0\}$  for some  $\theta, x_0 \in [0, 1)$ , where  $\{\cdot\}$  denotes the fractional part. Without loss of generality, we choose the partition  $I_a = [0, \rho)$  and  $I_b = [\rho, 1)$ , the subscript denoting the symbol associated to each interval. Thus a rotational word is determined by a triple  $(\theta, \rho, x_0)$ . Rotations are invertible, so all rotational words can be extended to the left; the extension is unique, apart from some notable special cases (see below).

If we denote by  $\mathcal{L}_n$  the collection of all sub-words (or *factors*) of length  $n$  in an infinite word  $W$ , then the *complexity function* of  $W$  is  $\mathcal{K}(n) = |\mathcal{L}_n|$ . An infinite word is eventually periodic if and only if  $\mathcal{K}(n) \leq n$  for some  $n$ , so we require  $\mathcal{K}(n) \geq n + 1$ . For rotational words,  $\mathcal{K}(n)$  is (eventually) an affine function of  $n$ . The precise form of  $\mathcal{K}$  depends on whether or not the boundary points 0 and  $\rho$  of the partition are on the same doubly-infinite orbit. If they are, we let  $\ell$  be the transit time from one to the other; if they aren't, we let  $\ell = \infty$ . Then, irrespective of the initial condition  $x_0$ , the rotational word has complexity [1, theorem 10]

$$(2) \quad \mathcal{K}(n) = \begin{cases} 2n & n \leq \ell \\ n + \ell & n > \ell. \end{cases}$$

The case  $\ell = 1$  are the much studied *Sturmian* words [21, section 6]. If  $1 < \ell < \infty$ , we call the word *quasi-Sturmian* (of length  $\ell$ ), and *free* if  $\ell = \infty$ .

All irrational rotations are minimal (and indeed uniquely ergodic), and this implies that for fixed irrational  $\theta$  and arbitrary  $\rho$ , each factor of any rotational word occurs infinitely often and with bounded time between successive occurrences, hence with a well-defined frequency; the latter may be computed as the length of an interval [1]. The symbolic language [the sequence  $(\mathcal{L}_n)$ ] does not depend on the initial condition  $x_0$  [21, p. 105], and for our purpose the parameter space for rotations will be the unit square  $[0, 1)^2$ . The parameters  $\theta$  and  $\rho$  will be referred to as the **rotational parameters**.

If the points 0 and  $\rho$  are on the same orbit, separated by  $\ell$  iterations ( $\ell$  could be negative), then  $\ell\theta \equiv \rho \pmod{1}$ , that is, the quasi-Sturmian configurations for a given  $\ell$  correspond to a finite collection of segments in rotational parameter space.

**2.2. The circle map.** Let  $f: [0, 1) \rightarrow [0, 1)$  be the circle map associated to (1), namely

$$(3) \quad f(e^{2\pi ix}) = \frac{F(e^{2\pi ix})}{\|F(e^{2\pi ix})\|} \quad e^{2\pi ix} \cong (\cos(2\pi x), \sin(2\pi x)) \in \mathbb{R}^2.$$

This is an orientation-preserving homeomorphism whose derivative is continuous and of bounded variation [14, theorem 3.1]. Hence  $f$  has a well-defined rotation number  $\theta = \theta(a, b)$ . The latter is a continuous function of the parameters, and its range is the interval  $[0, 1/2]$  [14, theorems 2.1 and 3.3]. From Denjoy theorem [12, p. 401], if  $\theta$  is irrational then  $f$  is topologically conjugate to a rotation by  $\theta$ , which in turn determines the symbolic language  $\mathcal{L}$ , irrespective of the initial conditions.

We divide the domain of the circle map  $f$  into two intervals  $I_a$  and  $I_b$ , and associate to an orbit a symbolic word  $W$  in the letters  $a$  and  $b$ , determined by the two branches of the map  $F$  in (1). We denote by  $W_n$  the prefix of  $W$  of length  $n$ , and by  $|W_n|_w$  the number of times the finite word (factor)  $w$  appears in  $W_n$ . The rotation number  $\theta(a, b)$  and the density  $\rho(a, b)$  are defined as follows:

$$(4) \quad \theta(a, b) = \lim_{n \rightarrow \infty} \frac{1}{n} |W_n(a, b)|_{ab}, \quad \rho(a, b) = \lim_{n \rightarrow \infty} \frac{1}{2n} (|W_n^+(a, b)|_a + |W_n^-(a, b)|_a),$$

where  $W$  is arbitrary and  $W^\pm$  is the word of the orbit with initial condition  $(0, \pm 1)$ .

The above limits exist because they are frequencies of factors of rotational words. Since every appearance of the factor  $ab$  corresponds to one loop around the origin, our definition of rotation number is equivalent to the standard one for circle homeomorphisms [12, chapter 11]; in particular, the rotation number is independent from the initial condition. This also holds for the density, as long as the rotation number is irrational, from minimality of irrational rotations (section 2.1). If the rotation number is rational, then the density may depend on the initial condition; in our definition (4) the union of the orbits corresponding to  $W^\pm$  is invariant under time-reversal symmetry [because  $F(0, \mp 1) = (\pm 1, 0)$ ]. This ensures that  $\rho(a, a) = 1/2$  for all real  $a$ , so that  $\rho$  is continuous on the main diagonal in parameter space.

We consider the level sets of the rotation number:

$$(5) \quad \Theta(x) = \{(a, b) \in \mathbb{R}^2 \mid \theta(a, b) = x\} \quad 0 \leq x \leq 1/2.$$

If  $x = p/q$  is rational, then  $\Theta(p/q)$  will be called a *resonance*.<sup>3</sup> A point  $(a, b)$  of a resonance is a *pinch-point* if the set  $\Theta(p/q) \setminus \{(a, b)\}$  is locally disconnected near  $(a, b)$ . Note that the density  $\rho$  is not necessarily continuous within a resonance, because the same rotation number may be associated to more than one code. For instance, the periodic codes  $aabb$  and  $abbb$  have the same rotation number  $1/4$ , but distinct densities  $1/2$  and  $1/4$ .

We now remove from the parameter space  $\mathbb{R}^2$  of the circle map  $f$  the interior of the resonances  $\Theta(0)$  and  $\Theta(1/2)$ , given by

$$(6) \quad \Theta(1/2) = \{(a, b) \in \mathbb{R}^2 \mid ab \leq 4, a, b < 0\}, \quad \Theta(0) = \{(a, b) \in \mathbb{R}^2 \mid \max(a, b) \geq 2\}.$$

We obtain an infinite strip  $\mathcal{A}$ , the region lying between the boundaries of  $\Theta(0)$  and  $\Theta(1/2)$ . (If we identify the points  $(-\infty, \zeta)$  and  $(\zeta, -\infty)$ , for  $0 \leq \zeta \leq 2$ , then  $\mathcal{A}$  becomes a compact set—a topological annulus.)

A detailed investigation of parameter space outside resonances requires replacing original parameters  $a, b$  with the rotational parameters  $\theta, \rho$ , which are suited for an arithmetical analysis of critical curves. This is the subject of a forthcoming investigation [23].

<sup>3</sup>We prefer this term to the customary terms *tongue* or *sausage*.

**2.3. Basic dynamical properties.** The area-preserving map  $F$  of (1) has a common form of a **reversible** map, i.e., a map conjugate to its inverse via an involution, specifically

$$(7) \quad F^{-1} = R \circ F \circ R, \quad R : (x, y) \mapsto (y, x).$$

Reversible maps are well-studied [17] and an equivalent definition is that they can be written as the composition of two involutions (e.g.,  $F$  is the composition of  $F \circ R$  and  $R$ ). An orbit of  $F$  is called *symmetric* if it is  $R$ -invariant and *asymmetric* otherwise, in which case it forms an asymmetric pair with its  $R$ -image. A symmetric orbit must contain one or two points from the *symmetry lines*

$$Fix(R) \cup Fix(F \circ R),$$

and is periodic if and only if it contains two points, where

$$(8) \quad Fix(R) := \{(x, x) : x \in \mathbb{R}\}, \quad Fix(F \circ R) := \left\{ \left( \frac{a}{2} y, y \right) : y > 0 \right\} \cup \left\{ \left( \frac{b}{2} y, y \right) : y \leq 0 \right\}.$$

As also realised in [14], a special role in the dynamics of  $F$  is played by the positive and negative ordinate semi-axes and we define:

$$(9) \quad L^+ := \{(0, y) : y \geq 0\}, \quad L^- := \{(0, -y) : y \geq 0\}.$$

We use the terminology  $L^+$ -orbit of  $F$  for an orbit that contains  $L^+$  and likewise an  $L^-$ -orbit. An orbit that contains neither  $L^+$  or  $L^-$  will be called a non- $L^\pm$  orbit. Because of the scale invariance of  $F$ , it suffices to study the orbit of  $(0, 1)$  to find an  $L^+$  orbit and the orbit of  $(0, -1)$  to find an  $L^-$  orbit. As a result, we will often identify  $L^+$ , respectively  $L^-$ , with  $(0, 1)$ , respectively  $(0, -1)$ . More generally, the scale invariance of  $F$  allows us to talk interchangeably about an orbit of points in the plane and the associated orbit of rays where each point is embedded into its position vector from the origin.

An orbit of  $F$  is typically made from patching together **orbit segments** that occupy the domain of  $M_a = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$  in the right half plane with orbit segments that switch to occupy the domain of  $M_b$  in the left half plane, before repeating this alternating behaviour. For reference, we denote these domains:

$$(10) \quad \text{dom}(M_a) := \left\{ (x, y) \in \mathbb{R}^2 : -\frac{\pi}{2} \leq \arctan_2(y, x) < \frac{\pi}{2} \right\}, \quad \text{dom}(M_b) := \mathbb{R}^2 \setminus \text{dom}(M_a).$$

In proposition 1 below, we study the dynamics within a single orbit segment in the right half plane (for ray counting in that result, note that  $L^- \in \text{dom}(M_a)$  but  $L^+ \in \text{dom}(M_b)$ ).

Define the eigenvalues of  $M_a$  by

$$(11) \quad \lambda = \lambda_\pm = \frac{a}{2} \pm \sqrt{\left(\frac{a}{2}\right)^2 - 1}, \quad \lambda_+ \lambda_- = 1.$$

When  $|a| \leq 2$ , the elliptic case, we shall make use of the quantity

$$(12) \quad \zeta_j := 2 \cos(\pi/j), \quad j = 1, 2, \dots$$

When  $|a| > 2$ , the hyperbolic case, the real eigenvectors associated to  $\lambda_\pm$  are denoted  $V_\pm$  and have slope  $\lambda_\mp$ . From (69) in the appendix, the powers of  $M_a$  can be expressed in terms of polynomials  $U_n(a)$  in  $a$  of degree  $n - 1$ , with  $U_n(a) = \bar{U}_{n-1}(a/2)$ , where  $\bar{U}_n$  is the  $n$ th Chebyshev polynomial of the second kind. The polynomials  $U_n$  satisfy

$$(13) \quad U_{-1}(a) = -1, \quad U_0(a) = 0, \quad U_{n+1}(a) = aU_n(a) - U_{n-1}(a), \quad n \geq 0.$$

From (69), the forward images of  $(1, 0) = M_a(0, -1)$  by  $M_a^j$  for  $j \geq 0$ , as long as they remain in  $\text{dom}(M_a)$ , are the rays

$$(14) \quad (U_{j+1}(a), U_j(a)) \in \text{dom}(M_a), \quad \text{with slope } m_j := U_j(a)/U_{j+1}(a).$$

Likewise, by reversibility, the images of  $R(1, 0) = (0, 1)$  by  $M_a^{-j}$  are the rays

$$(15) \quad \mathbf{r}_j(a) := (U_j(a), U_{j+1}(a)) \in \text{dom}(M_a), \quad \text{with slope } 1/m_j.$$

The Möbius transformation that relates the slope  $m$  of an initial ray in  $\text{dom}(M_a)$  to the slope  $\mu_a(m)$  of its image is:

$$(16) \quad m \mapsto \mu_a(m) = \frac{1}{a - m}, \quad m \in \mathbb{R} \cup \infty.$$

It inherits a reversing symmetry from  $R$ , given by

$$(17) \quad m \mapsto \frac{1}{m}.$$

We also note that order preservation of the induced circle map corresponding to  $M_a$  (or  $M_b$ ) means that if one ray is obtained from another by an anti-clockwise rotation then the same is true of their images. We use the anti-clockwise direction to order rays in the orbit segment from ‘first’ to ‘last’.

**Proposition 1.** *Consider the dynamics of  $M_a = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$  in  $\text{dom}(M_a)$  of (10).*

- i) *For all  $a > 2$ , all rays between  $L^-$  and the contracting eigenvector ray  $V_- \cap \text{dom}(M_a)$  with slope  $\lambda_+ > 1$  converge onto the expanding eigenvector ray  $V_+ \cap \text{dom}(M_a)$  with slope  $\lambda_- < 1$  —the points on these rays escape to infinity along this direction; the rays between  $V_- \cap \text{dom}(M_a)$  and  $L^+$  eventually escape to  $\text{dom}(M_b)$ .*
- ii) *For any  $\kappa \geq 2$ , we have  $a = \zeta_\kappa$  of (12) if and only if there is a segment of a symmetric orbit that goes from  $L^-$  to  $L^+$  with  $\kappa$  rays in  $\text{dom}(M_a)$  and  $M_a^\kappa = -Id$ .*
- iii) *When  $a \in (\zeta_{\kappa-1}, \zeta_\kappa)$ ,  $\kappa \geq 2$ , the possible orbit segments of  $M_a$  in  $\text{dom}(M_a)$  comprise a sequence of rays rotating anti-clockwise from the fourth quadrant into the first quadrant and eventually escaping to the left half-plane. The possibilities are:*
  - a) *a  $L^+$ -orbit segment with  $\kappa-1$  rays and first ray  $(U_{\kappa-1}(a), U_\kappa(a))$  in the interior of the fourth quadrant;*
  - b) *a  $L^-$ -orbit segment with  $\kappa$  rays comprising  $L^- = (0, -1)$  and  $(1, 0)$  and the  $R$ -image of the  $L^+ = (0, 1)$ -orbit segment contained within the first quadrant.*
  - c) *a non- $L^\pm$  orbit segment with  $\kappa - 1$  rays if the first ray has slope  $m$  with  $U_\kappa(a)/U_{\kappa-1}(a) < m < 0$  and  $\kappa$  rays if the first ray has slope  $m$  with  $m < U_\kappa(a)/U_{\kappa-1}(a) < 0$ .*
- iv) *The rays  $\mathbf{r}_j(a)$  of (15) that delineate the sectors in item iii a) rotate anti-clockwise as  $a$  increases.*

**PROOF.** Firstly, we make a general comment about orbit segments. Since  $(0, -1)$  maps anti-clockwise to  $(1, 0)$ , independent of  $a$ , any orbit segment in  $\text{dom}(M_a)$  not starting parallel to  $(0, -1)$  must have a single ray, its first, in the interior of the fourth quadrant (which is the image under  $M_b$  of the last ray in the preceding orbit segment). This first ray with positive  $x$ -coordinate has an image with positive  $y$ -coordinate since  $y' = x$  in  $F$ . On the other hand, the image of  $(0, 1)$  by  $M_a^{-1}$  is  $(1, a)$  and the last ray of any orbit segment of  $M_a$  that escapes to the left half plane must be the single ray of the orbit segment located in the wedge sector determined by  $(1, a)$  and  $L^+$ .

We prove i). When  $a > 2$ ,  $U_{j+1}(a) > U_j(a) \geq j$  for  $j \geq 1$  so (14) and order preservation implies the forward orbit of  $(1, 0)$ , and hence that of any initial ray in the fourth quadrant, is confined thereafter to the first quadrant. Analysis of (16) confirms that the fixed point at  $m_- = \lambda_-$  of (11) is attracting on  $[-\infty, \lambda_+)$ . The fixed point at  $m_+ = \lambda_+$  is repelling and the interval  $(\lambda_+, \infty)$  eventually is mapped to  $m < 0$  (which in this case corresponds to rays moving to the second quadrant). Of course, the fixed points  $m_{\pm} = \lambda_{\pm}$  correspond to the eigenvectors of  $M_a$ .

We prove ii). For  $M_a^{-j}(0, 1) = (U_j(a), U_{j+1}(a))$ , requiring  $U_j(a) > 0$  for  $1 \leq j < \kappa$  and  $U_{\kappa}(a) = 0$  produces  $\kappa - 1$  rays clockwise in the first quadrant, the last one parallel to  $(1, 0)$ . Since the pre-image of  $(1, 0)$  is  $(0, -1)$ , we obtain  $\kappa$  rays in total. For  $\kappa = 2$ , the condition on  $a$  for this orbit is that  $U_2(a) = a = 0$ . For  $\kappa > 2$ , the condition  $U_j(a) > 0$  for  $1 \leq j < \kappa$  and  $U_{\kappa}(a) = 0$  finds the right-most root of  $U_{\kappa}$ , namely  $a = \zeta_{\kappa}$ . This follows from the well known roots of the Chebyshev polynomials  $\bar{U}_n$ , as does the fact that  $U_{\kappa-1}(\zeta_{\kappa}) = 1$ . Thus the  $(\kappa - 1)$ th ray is exactly  $(1, 0)$ , not just parallel to it. Substituting  $U_{\kappa-1}(\zeta_{\kappa}) = 1$ ,  $U_{\kappa}(\zeta_{\kappa}) = 0$  and  $U_{\kappa+1}(\zeta_{\kappa}) = -1$  into (69) gives  $M_a^{\kappa} = -Id$ . Since  $(1, 0) = F(0, -1)$  and  $(0, 1)$  are mapped to one another by  $R$ , the finite orbit segment from  $L^-$  to  $L^+$  is part of a symmetric orbit.

iii) For an  $L^+$ -orbit segment of  $\kappa - 1$  rays, we require from (15) that  $U_j(a) > 0$  for  $1 \leq j < \kappa$  and  $U_{\kappa}(a) < 0$ , whence  $a \in (\zeta_{\kappa-1}, \zeta_{\kappa})$ , giving part a). These  $\kappa - 1$  rays, together with  $L^+$ , delineate  $\kappa$  sectors in  $\text{dom}(M_a)$ . Taking any ray internal to the first sector bounded by  $L^-$  and  $(U_{\kappa-1}(a), U_{\kappa}(a))$  and iterating gives an orbit segment of  $\kappa$  rays in  $\text{dom}(M_a)$ , with one ray in each sector by order preservation. The first ray has slope  $m : m < U_{\kappa}(a)/U_{\kappa-1}(a) < 0$ . Taking any ray internal to the second sector starting at  $(U_{\kappa-1}(a), U_{\kappa}(a))$  that is also in the fourth quadrant, with slope  $m : U_{\kappa}(a)/U_{\kappa-1}(a) < m < 0$ , gives an orbit segment of  $\kappa - 1$  rays. This proves part c). A  $L^-$  orbit segment of  $\kappa$  rays begins with  $L^-$  and then a ray parallel to  $(1, 0) = R(0, 1)$ . From (14) and (15), the forward orbit of  $(1, 0)$  contained in the first quadrant is the  $R$ -image of the backwards orbit of  $(0, 1)$  contained in the first quadrant, which from b) comprises  $\kappa - 1$  rays including now  $(0, 1)$  itself.

We prove iv). The inverse of (16) is  $m \mapsto \mu_a^{-1}(m) = a - \frac{1}{m}$  for  $m \in \mathbb{R} \cup \infty$ , and it gives the slope  $m(F^{-1}z) = \mu_a^{-1}(m)$  in terms of the slope  $m = m(z)$  of a ray  $z \in \mathbb{R}^2$ . We apply this to the rays  $\mathbf{r}_j(a)$  of (15) that delineate the sectors in proposition 1 iii) a). Defining  $r_j(a) := m(\mathbf{r}_j(a))$ , we have

$$r_{j+1}(a) = a - \frac{1}{r_j(a)} \implies r'_{j+1} = 1 + \frac{r'_j}{r_j^2},$$

where the prime denotes differentiation with respect to  $a$ . Since  $r_0 = \infty, r'_0 = 0$  and  $r_1 = a, r'_1 = 1$ , then by induction, we see  $r'_j \geq 1$  for  $j \geq 1$ . Hence as  $a$  increases,  $\mathbf{r}_j(a)$  moves anti-clockwise.  $\square$

There is the obvious version of the above for  $M_b$ .

This first dynamical analysis allows some preliminary bounds on rotation number and density in parameter space.



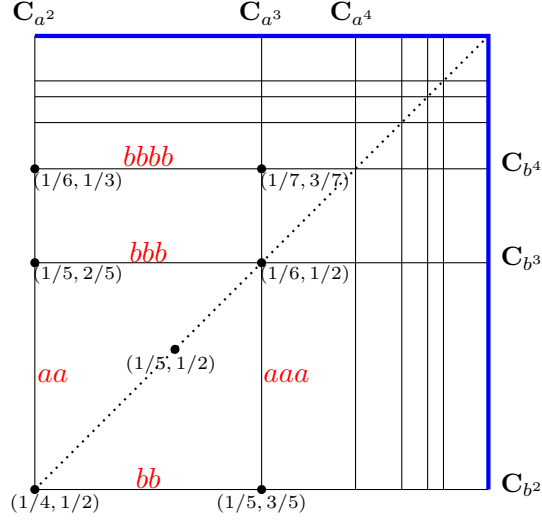


FIGURE 2. Partition of the first quadrant  $[0, 2]^2$  of the parameter space  $\mathcal{A}$  into rotational domains, determined by the critical curves of proposition 3. The boundary of  $\Theta(0)$  is represented in blue, and positive and negative critical words in red. The rational pairs near highlighted points represent the value of rotational parameters at those points.

**Proposition 2.** *Let  $\kappa \geq 2$ . If  $a \in (\zeta_{\kappa-1}, \zeta_\kappa]$ , then the number of rays in each orbit segment in  $\text{dom}(M_a)$  is  $\kappa - 1$  or  $\kappa$ . If  $b \in (\zeta_{\ell-1}, \zeta_\ell]$ ,  $\ell \geq 2$ , the analogous result holds in  $\text{dom}(M_b)$  and in addition:*

$$(i) \theta(a, b) \in \left[ \frac{1}{\kappa + \ell}, \frac{1}{\kappa + \ell - 2} \right]; \quad (ii) \rho(a, b) \in \left[ \frac{\kappa - 1}{\kappa + \ell - 1}, \frac{\kappa}{\kappa + \ell - 1} \right].$$

PROOF. The statement on the number of rays versus parameter value follows directly from the list of possibilities in proposition 1. If we use  $N_a^i = N_a^i(a, b)$ ,  $i \in \mathbb{Z}$ , to label the bi-infinite sequence of the number of rays in  $\text{dom}(M_a)$  indexed by revolution  $i$ , and similarly for  $N_b^i$ , then

$$\begin{aligned} \theta(a, b) &= \lim_{k \rightarrow \infty} \frac{2k + 1}{\sum_{i=-k}^{i=k} N_a^i + N_b^i}, \\ \rho(a, b) &= \lim_{k \rightarrow \infty} \frac{\sum_{i=-k}^{i=k} N_a^i}{\sum_{i=-k}^{i=k} N_a^i + N_b^i} \\ &= \lim_{k \rightarrow \infty} \left( 1 + \frac{\sum_{i=-k}^{i=k} N_b^i}{\sum_{i=-k}^{i=k} N_a^i} \right)^{-1}. \end{aligned}$$

Using the bounds  $(2k + 1)(\kappa - 1) \leq \sum_{i=-k}^{i=k} N_a^i \leq (2k + 1)\kappa$  etc., gives the results. Also we observe

$$\rho(a, b) = K(a) \theta(a, b),$$

where  $K(a) := \lim_{k \rightarrow \infty} \frac{\sum_{i=-k}^{i=k} N_a^i}{2k+1}$  and  $\kappa - 1 \leq K \leq \kappa$ .  $\square$

## 3. CRITICAL CURVES

A **boundary parameter** is a pair  $(a, b) \in \mathcal{A}$  for which there is an F-orbit (a **boundary orbit**) that contains two (not necessarily distinct) boundary rays, namely the positive and negative ordinate semi-axes  $L^+$  and  $L^-$  of (9). Such a dynamical condition is accompanied by a finite **boundary word**  $w$ , which encodes the itinerary between two boundary rays; this part of the boundary orbit will be called a **boundary segment**. Because the two components of the map F coincide on boundary rays, changing any letter that corresponds to a boundary ray has no effect on the orbit. (This always applies to the first letter of a boundary word.) The resulting ambiguity defines an equivalence relation on words, whereby we write  $w \sim w'$  to indicate that  $w'$  is obtained from  $w$  by any change in the symbols allocated to a boundary ray. A boundary word which is equivalent, but not equal, to the symbolic dynamics of the map F is said to be **improper**. These words appear at the intersections of curves (section 4).

The **rank**  $\iota$  of a boundary word  $w$  is given by  $\iota = 1 + |w|_{ab} + |w|_{ba}$  [cf. definition (4)]. If  $w$  encodes the symbolic dynamics of F, then  $\iota$  is the number of half-loops that occur in the transit between the initial and final boundary rays. A boundary word begins and ends with the same symbol in the odd-rank cases ( $L^+ \mapsto L^-$  or  $L^- \mapsto L^+$ ) and with distinct symbols in the even-rank case ( $L^+ \mapsto L^+$  or  $L^- \mapsto L^-$ ). The **sign** of a boundary word is positive if the word begins with the symbol  $a$  (the orbit starts from  $L^-$ ), and negative otherwise.

If the rank of  $w$  is even, then the critical curve is the parametric locus of a periodic orbit having (at least) one point on the partition boundary. These are the periodic  $\beta$ -orbits in [9, section 3.1], in which case the length and rank of the boundary word are, respectively, the denominator and twice the numerator of the rotation number. If an even rank curve supports both positive and negative boundary segments, then we call it an **axis** (of a resonance).

If the rank of a word  $w$  is odd, then the set of parameters for which  $w$  is a boundary word will be called a **critical curve**, examples of which are given in [15, examples 4.1–3] (see proposition 3 below). The associated boundary orbit of F will be called a **critical orbit** in this case.

We see that the orbits of proposition 1 ii) are critical orbits of rank 1 with critical curve  $a = \zeta_\kappa$  and boundary word  $a^\kappa$  and this is true independently of the value of  $b$  and the dynamics in  $\text{dom}(M_b)$ . This is also the family i) mentioned in the introduction, which will be used in section 5. The following result extends example 3.2 in [14] (cf. also propositions 1 and 2 of the previous section):

**Proposition 3.** *For  $\kappa = 2, 3, \dots$ , the set*

$$(18) \quad \mathbf{C}_{a^\kappa} := \{(\zeta_\kappa, b) : b \in \mathbb{R}\},$$

with  $\zeta_\kappa$  of (12) has rotational parameters

$$(19) \quad \rho(\zeta_\kappa, b) = \kappa\theta(\zeta_\kappa, b), \quad \theta(\zeta_\kappa, b) = \begin{cases} 0 & b > 2 \\ \frac{\arccos(b/2)}{\kappa \arccos(b/2) + \pi} & |b| \leq 2 \\ \frac{1}{\kappa + 1} & b < -2. \end{cases}$$

Exchanging  $a$  and  $b$  yields the twin sets  $\mathbf{C}_{b\ell} := \{(a, \zeta_\ell) : a \in \mathbb{R}\}$ , with the corresponding formula for the rotation number and  $\rho(a, \zeta_\ell) = 1 - \rho(\zeta_\ell, a)$ .

PROOF. The rotation number in (19) for  $|b| \leq 2$  follows from example 3.2 and theorem 2.1 in [14]. The value of  $\theta$  for  $b > 2$  follows from (the  $b$ -part of) proposition 1 and the proof of proposition 2. If  $b \leq -2$ , then by (19) we have  $\theta(\zeta_\kappa, -2) = 1/(\kappa + 1)$ , and by [14, theorem 2.1],  $\theta(a, b)$  is nonincreasing in  $b$ , so that  $\theta(\zeta_\kappa, b) \geq 1/(\kappa + 1)$ . However, (ii) of the same theorem implies  $0 \leq \theta(\zeta_\kappa, b) \leq 1/(\kappa + 1)$  for any  $b$ , and thus we have  $\theta(\zeta_\kappa, b) = 1/(\kappa + 1)$  for  $b \leq -2$ .

The density in (19) follows from  $K(a)$  defined in the proof of proposition 2, noting from proposition 1 ii) that for  $a = \zeta_\kappa$ , we have  $N_a^i = \kappa$  for all  $i$ , whence  $K(\zeta_\kappa) = \kappa$ . Exchanging  $a$  and  $b$  yields the same analysis; the obvious modification for the density results from the congruence  $\rho \equiv \kappa\theta \pmod{1}$  with negative  $\kappa$ .  $\square$

Note that in the interval  $|b| \leq 2$  ( $|a| \leq 2$ ) these components do not intersect the interior of any resonance, by virtue of the fact that the rotation number is nowhere constant. The points  $(\zeta_\kappa, \pm 2)$  of  $\mathbf{C}_{a^\kappa}$  lie on the boundary of the resonances with rotation number 0 and  $1/(\kappa + 1)$ .

The critical curves are our main object of study, and the closure of the set of boundary parameters which belong to some critical curve will be called the *critical set*  $\Xi \subset \mathcal{A}$ . The simplest components of the critical set are those described in proposition 3. Indeed, for  $\kappa \geq 2$ , the positive word  $a^\kappa = \underbrace{a \cdots a}_\kappa$  is critical of rank one, with components  $\mathbf{C}_{a^\kappa}$ .

Likewise,  $b^\ell$  is negative of rank one, with component  $\mathbf{C}_{b^\ell}$ .

For later use, we collect the following which are known or easily verified:

**Lemma 4.** *For the map  $F = F_{ab}$  of (1) and its reversing symmetry  $R$  of (7), we have:*

- i) *From reversibility, the  $R$ -image of the forward (backward) orbit of  $(0, y)$  is the backward (forward) orbit of  $R(0, y) = (y, 0) = F(0, -y)$ , i.e.,*

$$F^{-k+1}(0, -y) = R F^k(0, y), \quad k \in \mathbb{Z}, \quad y \in \mathbb{R}^+.$$

- ii) *A critical orbit of odd rank exists if and only if  $(0, 1)$  and  $R(0, 1) = (1, 0) = F(0, -1)$  are in the same orbit. Since the point  $(0, 1)$  and its image under  $R$  are in the same orbit, the critical orbit is symmetric.*
- iii) *With  $T : (x, y) \mapsto (-x, -y)$ , we have*

$$F_{ab} = T \circ F_{ba} \circ T^{-1}.$$

In particular, it suffices to study critical orbits where  $(0, -1)$  goes to  $(0, 1)$  in forward time as those from  $(0, 1)$  to  $(0, -1)$  for  $F_{ab}$  map to the former for  $F_{ba}$ .

- iv) There can be no critical orbit (of rank  $> 1$ ) from  $(0, -1)$  to  $(0, 1)$  if  $a > 2$  ( $b > 2$ ).
- v) We have  $F(0, \pm 1) = (\mp 1, 0)$  together with  $F(1, 0) = (a, 1)$ ,  $F(-1, 0) = (-b, -1)$  and  $F^{-1}(0, 1) = (1, a)$ ,  $F^{-1}(0, -1) = (-1, -b)$ . Additionally:
  - (a) each orbit segment in  $\text{dom}(M_a)$  ( $\text{dom}(M_b)$ ) only has at most one point in the interior of the fourth (second) quadrant, with its image in the first or second (third or fourth) quadrant;
  - (b) if  $a, b \in (-2, 2)$ , and if an orbit segment has a point  $P$  in the interior of the first (third) quadrant, then some image  $F^j P$  with  $j \geq 1$  is in the second (fourth) quadrant and some pre-image  $F^{-k} P$  with  $k \geq 1$  is in the fourth (second) quadrant.

PROOF. Property i) is a standard result for reversible maps [17]. Property ii) is from [14, theorem 3.4]. Property iii) is from [14, equation (2.4)]. Property iv) follows from Proposition 1 i) and its corresponding version for  $M_b$ . Property v) follows from  $F$  being order preserving and  $F : (x, y) \mapsto (x', y')$  having  $y' = x$  and  $F^{-1} : (x', y') \mapsto (x, y)$  having  $x = y'$ .  $\square$

**3.1. Algebraic properties.** Requiring that an orbit contain two boundary rays leads to algebraic curves over  $\mathbb{Q}$ . We introduce the necessary formalism, and establish some properties of these curves.

Let

$$(20) \quad M_x = \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}$$

where  $x$  is an indeterminate and consider the finite word

$$w = w_0 w_1 \cdots w_{t-1},$$

where we regard, for now, each letter of the word  $w$  as an indeterminate (in the case of  $F$ , we will later specialise to  $w_i \in \{a, b\}$ ). We form a product of matrices of type (20) as follows:

$$(21) \quad M_{\emptyset} = \text{Id}, \quad M_{[w]} = M_{w_{t-1}} M_{w_{t-2}} \cdots M_{w_0},$$

where  $\emptyset$  denotes the empty word.

The entries of  $M_{[w]}$  are polynomials in the indeterminates  $w_0, w_1, \dots, w_{t-1}$ . Define the recursive sequence of polynomials  $Q_t = Q_t(w_1, \dots, w_{t-1}) \in \mathbb{Z}[w_0, w_1, \dots, w_{t-1}]$  by the three-term recurrence:

$$(22) \quad Q_{-1} = -1 \quad Q_0 = 0 \quad Q_{t+1} = w_t Q_t - Q_{t-1} \quad t \geq 0.$$

Hence

$$\begin{aligned} Q_1 &= 1, \\ Q_2(w_1) &= w_1, \\ Q_3(w_1, w_2) &= w_1 w_2 - 1, \\ Q_4(w_1, w_2, w_3) &= w_1 w_2 w_3 - w_1 - w_3, \\ Q_5(w_1, w_2, w_3, w_4) &= w_1 w_2 w_3 w_4 - w_1 w_4 - w_3 w_4 - w_1 w_2 + 1. \end{aligned}$$

Alternatively, we can use the tridiagonal determinant representation:

$$(23) \quad Q_t(w_1, \dots, w_{t-1}) := \begin{vmatrix} w_1 & 1 & 0 & \cdots & 0 \\ 1 & w_2 & 1 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 1 & w_{t-2} & 1 \\ 0 & \cdots & 0 & 1 & w_{t-1} \end{vmatrix}, \quad t \geq 2,$$

with the recurrence relation (22) following by expanding the determinant along the last row. Note we can also run the recurrence (22) backwards using indeterminates  $w_{-i}$  to find:

$$(24) \quad Q_{-t} = -Q_t(w_i \mapsto w_{-i}) \quad t \geq 0.$$

With  $Q_t$  defined, it follows by induction:

$$(25) \quad M_{[w_1 \cdots w_{t-1}]} = \begin{pmatrix} Q_t(w_1, \dots, w_{t-1}) & -Q_{t-1}(w_2, \dots, w_{t-1}) \\ Q_{t-1}(w_1, \dots, w_{t-2}) & -Q_{t-2}(w_2, \dots, w_{t-2}) \end{pmatrix},$$

and

$$(26) \quad M_{[w_0 \cdots w_{t-1}]} = M_{[w_1 \cdots w_{t-1}]} M_{[w_0]} = \begin{pmatrix} w_0 Q_t - Q_{t-1} & -Q_t \\ w_0 Q_{t-1} - Q_{t-2} & -Q_{t-1} \end{pmatrix} \quad t \geq 1.$$

The above results follow from the theory of *continuants* (or *continuant polynomials*), a name given to the tridiagonal determinant (23). Continuants and their properties were studied by Euler [10, Section 6.7] in connection with generalised continued fractions involving arbitrary real (or complex) numbers instead of integers, see proposition 5 i) below. The associated three-term recurrence for Euler's original continuant is for the polynomial sequence  $q_t$  satisfying  $q_{-1} = -1$ ,  $q_0 = 0$  and  $q_{t+1} = w_t q_t + q_{t-1}$ . The polynomial  $Q_t$  of (22) belongs to a class of generalisations of the classic continuant, variously called signed continuant polynomials or generalised Chebyshev polynomials, which have arisen in the study of cluster algebras and frieze patterns [2, 19]. The relationship between  $Q_t$  and  $q_t$  is  $Q_t(w_1, \dots, w_{t-1}) = (-i)^{t-1} q_t(i w_1, \dots, i w_{t-1})$  with  $i = \sqrt{-1}$ .

It follows from Euler that  $Q_t$  can be generated as the sum of the product  $w_1 w_2 \dots w_{t-1}$ , its leading term, together with all possible ways of writing the leading term again but striking out adjacent product pairs  $w_i w_{i+1}$  and replacing such a pair with  $-1$  (e.g., the four non-leading terms of  $Q_5$  above are obtained by striking out, respectively,  $w_2 w_3$ ,  $w_1 w_2$ ,  $w_3 w_4$  and the two pairs  $w_1 w_2$  and  $w_3 w_4$ ). Note that from (22) and (13),

$$(27) \quad Q_{n+1}(x, x, \dots, x) = U_{n+1}(x) = \bar{U}_n(x/2),$$

where  $\bar{U}_n$  are the Chebyshev polynomials of the second kind —see the Appendix.

We shall need the following properties of continuants, found in references [10, Chapter 6.7] [19, Section 5.1], [2, Section 2.2]: [2, 10, 19],

**Proposition 5** (Continuant Polynomials). *For all  $t > 1$ , indeterminates  $w_1, \dots, w_{t-1} \in \mathbb{R}$  and integers  $k, l \geq -1$ , we have*

- i) 
$$\frac{Q_t(w_1, \dots, w_{t-1})}{Q_{t-1}(w_2, \dots, w_{t-1})} = w_1 - \frac{1}{w_2 - \frac{1}{\dots - \frac{1}{w_{t-1}}}}$$
- ii)  $Q_t(w_1, \dots, w_{t-2}, w_{t-1}) = Q_t(w_{t-1}, w_{t-2}, \dots, w_1)$
- iii)  $Q_{k+l+1}(w_1, \dots, w_{k+l}) = Q_{k+1}(w_1, \dots, w_k) Q_{l+1}(w_{k+1}, \dots, w_{k+l}) - Q_k(w_1, \dots, w_{k-1}) Q_l(w_{k+2}, \dots, w_{k+l}).$

We now specialise to the case of odd rank, for which the structure of words and curves is constrained by the *reversibility* of the map. The **reduced word**  $\underline{w}$  of  $w$  is defined as  $\underline{w} = w_1 \cdots w_{n-1}$  (again reversibility dictates considering to drop the first letter  $w_0$  – see below). We say that  $\underline{w}$  is a *palindrome* if

$$w_i = w_{n-i} \quad i = 1, \dots, n-1.$$

For F, the structure of a (palindromic) reduced boundary word of odd rank can also be encoded by the (palindromic) integer exponent sequence  $(i_k)$  of odd length  $2m-1$ , using  $i_k \in \mathbb{Z}^+$ ,  $1 \leq k \leq m$ , that we call its **block sequence**, e.g.,

$$(28) \quad \underline{w} = a^{i_1} b^{i_2} a^{i_3} \dots b^{i_{m-1}} a^{i_m} b^{i_{m-1}} \dots a^{i_3} b^{i_2} a^{i_1}.$$

In this form, the rank is now obvious, being  $2m-1$ , whereas

$$(29) \quad i_m + 2 \sum_{k=1}^{m-1} i_k = n-1.$$

We take the convention that  $i_k$ ,  $k$  odd, counts powers of  $a$  if the associated word is positive as in (28) (recall this means the word begins with  $a$ ) and powers of  $b$  if the word is negative. We see for a positive palindromic word that the middle block comprises powers of  $a$  when  $m$  is odd and powers of  $b$  when  $m$  is even. We recall from proposition 2 that necessarily:

$$(30) \quad a \in (\zeta_{\kappa-1}, \zeta_{\kappa}) \implies i_k \in \{\kappa-1, \kappa\}, \quad b \in (\zeta_{\ell-1}, \zeta_{\ell}) \implies i_k \in \{\ell-1, \ell\}.$$

The significance of palindromic words to boundary curves of the map F is established by the following result (it suffices to prove the case of a positive word noting ii) and iii) of lemma 4).

**Theorem 6.** *Consider a non-periodic critical orbit of F that contains  $(0, -1)$  and  $(0, 1)$  and the associated positive boundary word  $w$  of length  $n$  that encodes the itinerary between them. We have:*

- i) *The reduced boundary word of rank  $2m-1$  is a palindrome (28) of  $2m-1$  blocks (equivalently its block sequence is a palindromic  $(2m-1)$ -tuple of positive integers)*
- ii) *The integer  $n$  is odd and  $i_m$  is even if and only if the boundary segment intersects  $\text{Fix}(\mathbf{R})$  once in  $\text{dom}(\mathbf{M}_a)$  ( $\text{dom}(\mathbf{M}_b)$ ), whence  $m$  is odd (even).*
- iii) *The integer  $n$  is even and  $i_m$  is odd if and only if the boundary segment intersects  $\text{Fix}(\mathbf{F} \circ \mathbf{R})$  once. The intersection is in  $\text{dom}(\mathbf{M}_a)$  ( $\text{dom}(\mathbf{M}_b)$ ) if  $i_m = 1$  and  $m$  is even (odd) or if  $i_m > 1$  and  $m$  is odd (even).*

PROOF. i). From reversibility, for arbitrary  $z \in \mathbb{R}^2$ , we have  $z' = \mathbf{F}(z) \iff \mathbf{R}z = \mathbf{F}(\mathbf{R}z')$ . We claim the symbol sequence of the F-orbit leaving the point  $z \neq (0, \pm 1)$ , developing to the right, is the same as the symbol sequence for the F-orbit arriving to  $\mathbf{R}z$ , developing to the left. To see this, note the first symbol for the former is determined by  $z$

and the first symbol for the latter is determined by  $Rz'$  and we claim that these symbols are the same. From proposition 4 v), we have that if  $z$  is in the first or fourth quadrant, hence encoded with symbol  $a$  for its forward image, then  $z'$  is in the first or second quadrant, whence  $Rz'$  is in the first or fourth quadrant, and  $Rz'$  is encoded similarly to  $z$ . A similar result is true for  $z$  in the second or third quadrant and the symbol  $b$ . The argument is then iterated, next with  $z'$  and  $RF(z')$ . If we take a critical orbit with  $z = (1, 0) = F(0, -1)$ , we see the finite orbit connecting  $z$  and  $Rz = (0, 1)$  must have a palindromic symbol sequence, i.e., the reduced word is a palindrome as in (28).

We prove ii). We know a critical orbit is symmetric from lemma 4. In general [17], for any  $G$  that acts as a reversor of a reversible map  $L$ , so  $L^{-1} = G^{-1} \circ L \circ G$ , an orbit is  $G$ -invariant if and only if  $Gz = L^{k_z}z$  with  $k_z \in \mathbb{Z}$  for each point  $z \in \mathbb{R}^2$  of the orbit. The latter is true if and only if it is true for one point. Then either:

- (a)  $GL^{j_z}z = L^{j_z}z$  if  $k_z = 2j_z$ , i.e.,  $L^{j_z}z$  is fixed by  $G$ ; or
- (b)  $LG L^{j_z+1}z = L^{j_z+1}z$  if  $k_z = 2j_z + 1$ , i.e.,  $L^{j_z+1}z$  is fixed by  $LG$ . If a symmetric orbit is not periodic, it contains one point fixed by  $G$  or one point fixed by  $LG$ .

We are interested in  $L = F$  and  $G = R$  and we can take the point  $z = (1, 0)$  and  $k_z = n - 1$ . If  $n$  is odd so  $k_z$  is even, we have case (a). Hence the forward orbit of the midway point  $F^{k_z/2}(1, 0) \in \text{Fix}(R)$  is the reflection by  $R$  of its backwards orbit and the slope of the rays in the forward and backwards orbit are reciprocals of each other from (17). The involution  $R$  preserves the first and third quadrants and interchanges the second and the fourth. If  $m > 1$  is even (odd), the middle block in (28) is built from the letter  $b$  (the letter  $a$ ) and the number of points  $i_m$  of this orbit segment in  $\text{dom}(M_b)$  ( $\text{dom}(M_a)$ ) must be even. It comprises  $F^{k_z/2}(1, 0)$  on the line  $y = x$  and pairs of its forward and backward iterates in the interior of the third (first) quadrant, plus the necessary additional point in the second (fourth) quadrant guaranteed from lemma 4. If  $m = 1$ , we have only a single letter in the word and then  $i_m$  in the reduced word is even (and the power is odd in the full word, cf. proposition 1). Hence  $i_m$  is even as claimed, which can also be seen from (29).

We prove iii). Now consider the case (b) above, i.e.,  $n$  is even so  $k_z$  is odd and  $q = F^{j_z+1}(1, 0) \in \text{Fix}(FR)$ . For  $y > 0$ , the involution  $FR : x' = -x + ay, y' = y$  is a horizontal reflection about the point  $\frac{a}{2}y$ ; for  $y < 0$ , it is a horizontal reflection about the point  $\frac{b}{2}y$ . We have  $F^j q = F^j (FR q) = (FR)(F^{-j} q)$ , so the forward and backward iterates of  $q$  under  $F$  at corresponding times in the upper half-plane and in the lower half-plane must be pairs under these reflections. As an example, suppose  $q \in \text{dom}(M_a)$  in the first quadrant. This means the single ray in the second quadrant when  $F^r q$  first enters it (cf. lemma 4) must force the iterates of  $q$  in the first quadrant to have  $r - 1$  rays to the left of  $x = \frac{a}{2}y$  in forwards time,  $r \geq 1$ , and  $r$  rays to the right in backward time. Counting  $q$  itself and the single ray  $F^{-(r+1)} q$  in the fourth quadrant gives  $2r + 1$  for the number  $i_m$  of the letter  $a$  in the middle block, whence  $i_m$  is odd. Similar reasoning applies for the other possibilities.  $\square$

This result leads us to study continuant polynomials for palindromic words (interestingly, palindromic continuants were used in Smith's 1855 proof of the Fermat two-square theorem [6]). For generality, we again let  $w = w_0 \cdots w_{n-1}$  be a word in  $n$  letters (not just

two letters  $a$  and  $b$ ). Let

$$(31) \quad \mathbf{k} = \lfloor n/2 \rfloor.$$

and build the particular polynomials in  $\mathbb{Z}[w_0, \dots, w_{n-1}]$  from (22):

$$(32) \quad C_w := \begin{cases} Q_{\mathbf{k}+1} - Q_{\mathbf{k}} & n \text{ odd} \\ Q_{\mathbf{k}+1} - Q_{\mathbf{k}-1} & n \text{ even} \end{cases} \quad \tilde{C}_w := \begin{cases} Q_{\mathbf{k}+1} + Q_{\mathbf{k}} & n \text{ odd} \\ Q_{\mathbf{k}} & n \text{ even.} \end{cases}$$

For completeness, if the rank of  $w$  is *even*, we let  $C_w = Q_n$  and  $\tilde{C}_w = 1$ .

The following result collects some algebraic properties of the  $Q_t$  polynomials for palindromic words in any alphabet.

**Proposition 7.** *If  $w_1 \cdots w_{n-1}$  is a palindrome (in any alphabet), then the following holds for the polynomials  $Q_t$  of (22) with  $t = 0, \dots, n$  (noting also (24) for (ii)):*

$$(33) \quad i) \quad Q_n = C_w \tilde{C}_w$$

$$(34) \quad ii) \quad Q_{n-t} - Q_t = C_w \times \begin{cases} Q_{t-\mathbf{k}+1}(w_{\mathbf{k}} = -1, w_{\mathbf{k}+1}, \dots, w_{t-1}) & n \text{ odd} \\ -Q_{t-\mathbf{k}}(w_{\mathbf{k}+1}, \dots, w_{t-1}) & n \text{ even.} \end{cases}$$

PROOF. i) From proposition 5 ii) and iii) with  $l = t - 1$ ,  $k = n - t$  and the palindromic word, we have for  $t = 1, \dots, n - 1$ :

$$(35) \quad \begin{aligned} & Q_n(w_1, \dots, w_{n-1}) = \\ & Q_{n-t+1}(w_1, \dots, w_{n-t}) Q_t(w_1, \dots, w_{t-1}) - Q_{n-t}(w_1, \dots, w_{n-t-1}) Q_{t-1}(w_1, \dots, w_{t-2}). \end{aligned}$$

For odd  $n = 2\mathbf{k} + 1$ , we take  $t = \mathbf{k} + 1$ , to obtain  $Q_n = Q_{\mathbf{k}+1}^2 - Q_{\mathbf{k}}^2 = C_w \tilde{C}_w$ . For even  $n = 2\mathbf{k}$ , we take  $t = \mathbf{k} = n - t$ , giving  $Q_n = Q_{\mathbf{k}+1} Q_{\mathbf{k}} - Q_{\mathbf{k}} Q_{\mathbf{k}-1} = C_w \tilde{C}_w$ .

ii) We prove the result by induction, in the first instance for  $t \geq \mathbf{k}$ . Consider the polynomial  $A_t = Q_{n-t} - Q_t$ . If  $n = 2\mathbf{k} + 1$ , from (32) we find:

$$A_{\mathbf{k}} = Q_{\mathbf{k}+1} - Q_{\mathbf{k}} = C_w = C_w Q_1 \quad A_{\mathbf{k}+1} = Q_{\mathbf{k}} - Q_{\mathbf{k}+1} = -C_w = C_w Q_2(w_{\mathbf{k}} = -1).$$

If  $n = 2\mathbf{k}$ , then

$$A_{\mathbf{k}} = Q_{\mathbf{k}} - Q_{\mathbf{k}} = 0 = C_w Q_0 \quad A_{\mathbf{k}+1} = Q_{\mathbf{k}-1} - Q_{\mathbf{k}+1} = -C_w = C_w (-Q_1).$$

The above data serves as the base case for induction. Assume that (34) holds for all  $i$  in the range  $\mathbf{k} \leq i \leq t$ , for some  $t \geq \mathbf{k} + 1$ . Then

$$\begin{aligned} A_{t+1} &= Q_{n-(t+1)} - Q_{t+1} = w_{n-t} Q_{n-t} - Q_{n-(t-1)} - w_t Q_t + Q_{t-1} \\ &= w_t (Q_{n-t} - Q_t) - (Q_{n-(t-1)} - Q_{t-1}) = C_w (w_t S_t - S_{t-1}) = C_w S_{t+1}, \end{aligned}$$

where  $S_t$  denotes the particular solutions in (34) for the respective cases of  $n$  even and  $n$  odd, which obviously satisfy  $S_{t+1} = w_t S_t - S_{t-1}$ . This completes the induction for the range  $t \geq \mathbf{k}$ . To see that (34) holds also for  $0 \leq t < \mathbf{k}$ , note that both sides become their negatives under  $t \mapsto n - t$ , using (24) on the right hand side.  $\square$

We now apply the above results to the dynamics of  $F$ . Let  $w = w_0 \cdots w_{n-1}$  be a boundary word – necessarily  $n \geq 2$  for a boundary word since one matrix of the form (20)



cannot map a boundary ray to a boundary ray. Let  $x_{-1} = -\text{sign}(w), x_0 = 0, \dots, x_n$  be the corresponding orbit segment. From the equation

$$\begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix} = M_{[w_0 \dots w_{t-1}]} \begin{pmatrix} 0 \\ -\text{sign}(w) \end{pmatrix}$$

and (26), we obtain

$$x_t = \text{sign}(w) Q_t(w_1, \dots, w_{t-1}), \quad t \geq 1.$$

An equation for the boundary curve with word  $w$  is obtained by requiring that  $x_n = 0$ :

$$(36) \quad Q_n(a, b) = 0.$$

This equation stores redundant information. It expresses the fact that the image of one boundary ray under the matrix product  $M_{[w]}$  is another boundary ray, and such an action may be realised without any reference to the symbolic dynamics of the map  $F$ . As a result, the curve

$$(37) \quad \mathcal{Q}_n = \{(a, b) \in \mathbb{R}^2 : Q_n(a, b) = 0\}$$

has several branches, as we shall see below. A parameter pair  $c = (a, b)$  for which  $w$  is equivalent<sup>4</sup> to the symbolic word of a boundary segment of  $F$  will be called a **legal point** of the curve, and a **legal branch** of the curve is one containing legal points. If a point is not legal, then it may happen that the rank of the word does not correspond to the number of half-turns performed by the orbit segment. We call the latter the **orbital rank** of the point  $c$ .

**Theorem 8.** *Let  $\mathcal{Q}_n$  be the curve (37), with reduced word  $\underline{w} = w_1 \dots w_{n-1}$ ,  $n \geq 2$ . Then*

- i)  $\mathcal{Q}_n$  has  $n - 1$  disjoint branches, of which precisely one is legal. On the legal branch, rank and orbital rank coincide, and vice-versa.
- ii) If the rank of  $w$  is greater than 1, then each branch is represented by a decreasing function  $b = b(a)$ .
- iii)  $\mathcal{Q}_n$  has  $n - 1$  asymptotes,  $|\underline{w}|_b$  of which horizontal and  $|\underline{w}|_a$  vertical, including multiplicities.

**PROOF.** We prove i). From (27), we see that for  $a = b$  we have  $Q_n(a, a) = \bar{U}_{n-1}(a/2)$ . Since the latter Chebyshev polynomial has  $n - 1$  distinct real roots, the curve (37) intersects the line  $a = b$  in  $n - 1$  distinct points, which are

$$(38) \quad a = b = \zeta_{n,j} = 2 \cos\left(\frac{\pi j}{n}\right), \quad j = 1, \dots, n - 1.$$

[The number  $\zeta_{n,1}$  above corresponds to  $\zeta_n$  in (12).] Over that line, we have  $M_{[w]} = M_a^n$ , and as  $a$  decreases from 2 to  $-2$ , the rotation number of  $M_a$  increases monotonically from 0 to  $1/2$ . At the value  $a = \zeta_{n,j}$  the image of a boundary ray will reach a boundary ray after  $j$  half-turns, and therefore the orbit of  $M_a$  will have the correct rank precisely for  $j = \iota$ .

Now prolong each branch starting from the corresponding point (38). Since the prolongation preserves the initial and final rays, as well as the orbital rank, distinct branches cannot intersect. Thus the boundary curve has a unique legal branch, namely that where rank and orbital rank coincide, which is the branch containing the point  $a = b = \zeta_{|w|, \iota(w)}$ . The proof of i) is complete.

<sup>4</sup>In the sense mentioned at the beginning of section 3.

We prove ii). Let  $c = (a, b)$  be a (finite) point on  $\mathcal{Q}_n$ . We consider the  $M_{[w]}(c)$ -orbit of the appropriate ray  $L_1 = (\pm 1, 0)$ , the sign agreeing with that of  $w$  (the notation  $M_{[w]}(c)$  refers to putting the specified parameter values  $c = (a, b)$  into the matrix entries of (25)). We don't require  $c$  to be legal, so the word  $w$  may be unrelated to the symbolic trajectory of the orbit of  $F$ . Since the rank of  $w$  is greater than 1, the orbit segment of  $w$  will have at least one non-boundary ray acted upon by each matrix  $M_a$  and  $M_b$ , where  $M_x$  is given in (20).

One verifies that for any ray  $L$  and parameters  $\xi$  and  $\epsilon$ , the ray  $M_{\xi+\epsilon}L$  is obtained from  $M_\xi L$  by rotating clockwise if  $\epsilon > 0$ , and anticlockwise if  $\epsilon < 0$ , unless  $L = L^\pm$ , in which case the two rays coincide<sup>5</sup>. Considering that the circle map  $f$  is orientation-preserving [14, theorem 3.1], by repeating the above argument it follows that for any  $\epsilon > 0$ , both  $M_{[w]}(c + (\epsilon, 0))L_1$  and  $M_{[w]}(c + (0, \epsilon))L_1$  are obtained from the vertical boundary ray  $M_{[w]}(c)L_1$  via a clockwise rotation.

From the above argument, it follows that the partial derivatives  $\partial Q_n/\partial a$  and  $\partial Q_n/\partial b$  are non-zero and agree in sign<sup>6</sup>, that is, the tangent at any point of the curve has negative slope. (This also shows that  $\mathcal{Q}_n$  has no isolated points.) We have proved ii).

We prove iii). Let  $n = |w|$ ,  $n_a = |w|_a$ ,  $n_b = |w|_b$ , and

$$(39) \quad Q_n(a, b) = a^{n_a} b^{n_b} + \sum_{k=0}^{n-2} \sum_{i=0}^k c_{i, k-i} a^i b^{k-i}, \quad \text{for some } c_{i,j} \in \mathbb{Z}.$$

We rewrite the above as follows

$$(40) \quad \begin{aligned} \frac{1}{a^{n_a}} Q_n(a, b) &= b^{n_b} + \sum_{j=0}^{n_b-1} c_{n_a, j} b^j + O\left(\frac{1}{a}\right) & |a| \rightarrow \infty \\ \frac{1}{b^{n_b}} Q_n(a, b) &= a^{n_a} + \sum_{j=0}^{n_a-1} c_{j, n_b} a^j + O\left(\frac{1}{b}\right) & |b| \rightarrow \infty. \end{aligned}$$

As  $|a|$  or  $|b|$  tends to infinity, the corresponding polynomial on the RHS of (40) must vanish, each root giving the equation of an asymptote. This gives  $n_a + n_b = n - 1$  asymptotes, counting multiplicities. Since the curve  $\mathcal{Q}_n$  has order  $n - 1$ , by Bézout's theorem, it cannot have more than  $n - 1$  points on the line at infinity, so there are no other asymptotes.  $\square$

Let us return to equation (36). If the rank of  $w$  is even, then  $Q_n$  may be (and typically is) irreducible, as in the case  $w = a^2 b^2$  for which  $Q_4 = ab^2 - a - b$ . Thus the equation (36) is the minimal description of a boundary curve of even rank, in general.

For odd rank, we consider the factorisation (33), and replace (37) by

$$(41) \quad \mathcal{C}_w = \{(a, b) \in \mathbb{R}^2 : C_w(a, b) = 0\},$$

where  $C_w$  is defined in (32). This is justified as follows. From reversibility we have  $x_t = x_{n-t}$  for all  $t$ . If  $n$  is odd, then the point  $z_{\mathbf{k}+1} = (x_{\mathbf{k}+1}, x_{\mathbf{k}}) = \pm(Q_{\mathbf{k}+1}, Q_{\mathbf{k}})$  lies on the symmetry line. Therefore the polynomial  $Q_{\mathbf{k}+1} - Q_{\mathbf{k}} = C_w$  must vanish on the legal branch of the curve. In general, there is no further factorisation, as shown by the

<sup>5</sup>This is theorem 3.2 (i) of [14].

<sup>6</sup>They are both positive if the rank is odd, and negative if the rank is even.

example  $w = a^3$ , for which the polynomial  $C_w = a - 1$  is irreducible. If  $n$  is even, then the points  $z_{\mathbf{k}+1}$  and  $z_{\mathbf{k}}$  are placed symmetrically with respect to the symmetry line, and hence  $Q_{\mathbf{k}+1} = Q_{\mathbf{k}-1}$ , that is,  $C_w$  vanishes. For  $w = a^4$  we have  $C_w = a^2 - 2$ , which is irreducible. From theorem 8, we see that  $\mathbf{k} = \lfloor n/2 \rfloor$  branches of the curve  $\mathcal{Q}_n$  belong to the curve  $C_w$ , while the remaining  $n - 1 - \mathbf{k}$  branches belong to  $\tilde{C}_w = 0$ . The former comprises all parameters corresponding to paths of odd rank, while the latter those of even rank. Here the term rank refers to the *orbital rank*, namely the number of half-turns around the origin, which, as already noted, may not be related to the number of factors  $ab$  and  $ba$  in the word.

The existence of non-legal branches cannot be avoided by considering only irreducible curves. For instance, the legal branch for the word  $w = a^\kappa$  is the line  $a = \zeta_\kappa$  [cf. (18)], and  $\zeta_\kappa$  is a root of the irreducible polynomial  $\Psi_{2\kappa}(a)$  [cf. (67)]. This polynomial has degree  $\phi(2\kappa)/2$ , where  $\phi$  is Euler's function [20, p 37]. For  $\kappa \geq 4$  such a degree is greater than one, corresponding to as many branches; so there is a non-legal branch.

**3.2. Congruences.** In this section we consider functions defined on a critical curve  $\mathcal{C}_w$ , with word  $w = w_0 \cdots w_{n-1}$ . Unless indicated otherwise, the results of this section will apply to the more general setting of reduced palindromic words, as in proposition 7. To lighten up the notation, we omit explicit reference to  $w$  and write  $\mathcal{C}$  for  $\mathcal{C}_w$  etc.

Given the polynomial  $C(a, b)$  of a curve [see (32)], we consider the polynomial ideal  $\langle C \rangle = C(a, b)\mathbb{Z}[a, b]$  of all the multiples of  $C$  in  $\mathbb{Z}[a, b]$ . (For background, see, e.g., [7].) The quotient ring  $\mathbb{Z}[a, b]/\langle C \rangle$  of residue classes modulo  $\langle C \rangle$ , namely the sets of the form  $P + \langle C \rangle$  for  $P \in \mathbb{Z}[a, b]$ , represents the polynomial functions  $P : \mathcal{C} \rightarrow \mathbb{R}$ . We write  $P \equiv Q \pmod{C}$  to mean that  $P - Q \in \langle C \rangle$ , in which case  $P$  and  $Q$  represent the same function on  $\mathcal{C}$ .

Thus equation (34) yields

$$(42) \quad Q_{n-t}(w_1, \dots, w_{n-t-1}) \equiv Q_t(w_1, \dots, w_{t-1}) \pmod{C}, \quad t = 0, \dots, n,$$

and considering that  $Q_n \equiv Q_0 = 0$  and  $Q_{n-1} \equiv Q_1 = 1$ , on the curve  $\mathcal{C}$  the matrix (26) takes the form

$$(43) \quad M_{[w]} \equiv \begin{pmatrix} -Q_{n-1}(w_2, \dots, w_{n-1}) & 0 \\ w_0 - Q_{n-2}(w_2, \dots, w_{n-2}) & -1 \end{pmatrix} \equiv \begin{pmatrix} -1 & 0 \\ w_0 - Q_{n-2} & -1 \end{pmatrix} \pmod{C}$$

where the second congruence follows from the fact that  $\det(M)$  has unit determinant (cf. [14, theorem 3.4]).

The variation of a function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  along the curve  $C$  is given by the Poisson brackets

$$\{f, C\} = \frac{\partial f}{\partial a} \frac{\partial C}{\partial b} - \frac{\partial f}{\partial b} \frac{\partial C}{\partial a}.$$

We consider the observables

$$(44) \quad \varphi_t = \arctan_2(Q_{t-1}, Q_t), \quad t = 0, \dots, n,$$

which represent the angle of the rays<sup>7</sup> in the orbit of the map  $F$ , that is, the points of the orbit segment of the circle map. This follows from the relation  $x_t = \text{sign}(w)Q_t$ , where  $(x_0, \dots, x_n)$  is the one-dimensional orbit segment associated to the curve.

We have

$$(45) \quad \{\varphi_t, C\} = \frac{Q_t^2}{Q_t^2 + Q_{t-1}^2} \{Q_{t-1}/Q_t, C\} = \frac{Q_t\{Q_{t-1}, C\} - Q_{t-1}\{Q_t, C\}}{Q_t^2 + Q_{t-1}^2}.$$

Thus  $\{\varphi_t, C\}$  has the form  $\Delta_t/\|z_t\|^2$ , where

$$(46) \quad \Delta_t = Q_t\{Q_{t-1}, C\} - Q_{t-1}\{Q_t, C\}, \quad t = 0, \dots, n,$$

and  $\|z_t\|^2 = Q_t^2 + Q_{t-1}^2$ . The latter has no real roots, because any common root of  $Q_t$  and  $Q_{t-1}$  would be common to all  $Q_i$ s, which is impossible since  $Q_1$  has no roots.

In preparation for the next statement, we consider the polynomials

$$(47) \quad \Xi_i = \begin{cases} Q_i^2 \frac{\partial C}{\partial a} & \text{if } w_i = b \\ -Q_i^2 \frac{\partial C}{\partial b} & \text{if } w_i = a, \end{cases} \quad i = 0, \dots, n-1$$

If  $\underline{w}$  is a palindrome, we have  $\Xi_{n-i} = \Xi_i$ , for  $i = 1, \dots, n-1$ .

We now establish formulae for  $\Delta_t$ .

**Theorem 9.** *Let  $\mathcal{C}$  be a boundary curve, and let  $\Delta$  and  $\Xi$  be as above. The following holds:*

$$\text{i) } \Delta_0 = 0, \quad \Delta_t = \sum_{i=0}^{t-1} \Xi_i, \quad t = 1, \dots, n, \quad \Delta_n \equiv 0 \pmod{\mathcal{C}}$$

If, in addition,  $\mathcal{C}$  is a critical curve, we have ( $\mathbf{k} = \lfloor n/2 \rfloor$ )

$$\text{ii) } \Delta_{n-t} \equiv -\Delta_{t+1} \pmod{\mathcal{C}}, \quad \Delta_t \equiv -\frac{1}{2} \sum_{i=t}^{n-t} \Xi_i \pmod{\mathcal{C}}, \quad t = 1, \dots, \mathbf{k}.$$

If  $n = 2\mathbf{k} + 1$ , then  $\Delta_{\mathbf{k}+1} = \{Q_{\mathbf{k}}, Q_{\mathbf{k}+1}\}C \equiv 0 \pmod{\mathcal{C}}$ .

$$\text{iii) } \sum_{t=1}^n \{\varphi_t, C\} \equiv 0 \pmod{\mathcal{C}}.$$

PROOF. i) Using linearity and Liebnitz rule for Poisson brackets, (46) becomes

$$\begin{aligned} \Delta_t &= (w_{t-1}Q_{t-1} - Q_{t-2})\{Q_{t-1}, C\} - Q_{t-1}\{w_{t-1}Q_{t-1} - Q_{t-2}, C\} \\ &= Q_{t-1}\{Q_{t-2}, C\} - Q_{t-2}\{Q_{t-1}, C\} - Q_{t-1}^2\{w_{t-1}, C\}. \end{aligned}$$

This gives

$$(48) \quad \Delta_t = \Delta_{t-1} + Q_{t-1}^2\{C, w_{t-1}\},$$

with

$$\{C, w_i\} = \begin{cases} \partial C / \partial a & \text{if } w_i = b \\ -\partial C / \partial b & \text{if } w_i = a. \end{cases}$$

<sup>7</sup> $\varphi_t$  lies in the interval  $(-\pi, \pi]$ .

From (46) we have  $\Delta_0 = \Delta_1 = 0$  and iterating the above recursion we find, for  $t > 0$

$$\Delta_t = \sum_{i=0}^{t-1} Q_i^2 \{C, w_i\} = \sum_{i=0}^{t-1} \Xi_i,$$

as desired. To establish the congruence, we let  $Q_n = C\tilde{C}$  [see (32) and following remark], and compute

$$\begin{aligned} \Delta_n &= C\tilde{C}\{Q_{n-1}, C\} - Q_{n-1}\{C\tilde{C}, C\} \\ &= C\tilde{C}\{Q_{n-1}, C\} - Q_{n-1}C\{\tilde{C}, C\} \equiv 0 \pmod{C}. \end{aligned}$$

We have proved i).

We prove ii). Using (42) and (46), we obtain

$$(49) \quad \Delta_{t+1} \equiv -\Delta_{n-t} \pmod{C}, \quad t = 0, \dots, n-1,$$

which, for  $n = 2\mathbf{k} + 1$ , yields  $\Delta_{\mathbf{k}+1} \equiv 0 \pmod{C}$ . More precisely,

$$\begin{aligned} \Delta_{\mathbf{k}+1} &= Q_{\mathbf{k}+1}\{Q_{\mathbf{k}}, Q_{\mathbf{k}+1} - Q_{\mathbf{k}}\} - Q_{\mathbf{k}}\{Q_{\mathbf{k}+1}, Q_{\mathbf{k}+1} - Q_{\mathbf{k}}\} \\ &= \{Q_{\mathbf{k}}, Q_{\mathbf{k}+1}\}(Q_{\mathbf{k}+1} - Q_{\mathbf{k}}) = \{Q_{\mathbf{k}}, Q_{\mathbf{k}+1}\}C. \end{aligned}$$

Iterating (48) backward and using i) we obtain the formula

$$(50) \quad \Delta_t \equiv -\sum_{i=t}^{n-1} \Xi_i \pmod{C}, \quad t = 1, \dots, n-1.$$

Since  $\Xi_{n-i} = \Xi_i$ , we find, for  $t = 1, \dots, \mathbf{k}$  the above sum becomes

$$\begin{aligned} \sum_{i=t}^{n-1} \Xi_i &= \sum_{i=t}^{n-t} \Xi_i + \sum_{i=n-t+1}^{n-1} \Xi_i \\ &\equiv \sum_{i=t}^{n-t} \Xi_i + \sum_{i=1}^{t-1} \Xi_i \pmod{C}. \end{aligned}$$

Since  $0 = Q_0 \equiv Q_n \pmod{C}$ , the range of the rightmost sum may be extended to include  $i = 0$ . Then, substituting the above expression in (50), and adding the latter to formula i), we obtain

$$\Delta_t \equiv -\frac{1}{2} \sum_{i=t}^{n-t} \Xi_i \pmod{C} \quad t = 1, \dots, \mathbf{k},$$

which completes the proof of ii).

We prove iii). If  $n$  is odd, then from ii) we obtain  $\Delta_{\mathbf{k}+1} \equiv 0 \pmod{C}$ . Keeping this in mind, we find

$$\begin{aligned} \sum_{t=1}^n \{\varphi_t, C\} &\equiv \sum_{t=1}^{\lfloor n/2 \rfloor} \left( \frac{\Delta_t}{Q_t^2 + Q_{t-1}^2} + \frac{\Delta_{n-t+1}}{Q_{n-t+1}^2 + Q_{n-t}^2} \right) \\ &\equiv \sum_{t=1}^{\lfloor n/2 \rfloor} \frac{\Delta_t + \Delta_{n-t+1}}{Q_t^2 + Q_{t-1}^2} \equiv 0 \pmod{C}. \end{aligned}$$

This establishes iii) and the proof of the theorem is complete.  $\square$

Some remarks on theorem 9 are in place. Part ii) says that  $\Delta_{\mathbf{k}+1}$  vanishes identically on  $\mathcal{C}$  for  $n = 2\mathbf{k} + 1$ . This is due to time-reversal symmetry: the point  $z_{\mathbf{k}+1}$  of the orbit segment never leaves the symmetry axis.

The statement iii) says that along an orbit segment of a critical curve, the sum of the angles is constant. To find the value of the constant, we represent the points of the orbit of  $F$  as complex numbers  $z_t = x_t + ix_{t-1}$ , with  $z_0 = -\text{sign}(w)i$ . We seek the value of

$$S = \sum_{t=1}^n \varphi_t = \sum_{t=1}^n \arg(z_t) \quad z_t = F^t(z_0),$$

for an arbitrary point on  $\mathcal{C}$ . We rewrite this sum as

$$(51) \quad S = \frac{1}{2} \sum_{t=0}^{n-1} [\arg(z_{t+1}) + \arg(z_{n-t})].$$

For any non-zero complex number  $z$  we have  $\arg(z) + \arg(R(z)) \equiv \frac{\pi}{2} \pmod{2\pi}$ . From reversibility, (51) becomes

$$S = \frac{1}{2} \sum_{t=0}^{n-1} \frac{\pi}{2} \pmod{2\pi} \equiv \frac{\pi n}{4} \pmod{\pi}.$$

In the above formula the modulus may be removed. The value of the sum can be shown to depend on the numbers of rays lying in the third quadrant, which is constant along  $\mathcal{C}$  as long as the code doesn't change.

In the next section we shall examine further geometrical consequences of theorem 9.

#### 4. INTERSECTIONS OF CURVES

A **double point** is a point of intersection of two distinct critical curves. Many geometrical properties of a critical curve  $\mathcal{C}$  are determined by its intersections with other critical curves. For instance,  $\mathcal{C}$  has a single legal branch [theorem 8 i)], but in general not all points of that branch are legal. It turns out that the legal part of the branch, which we call the **legal arc**, is delimited by certain double points.

**Theorem 10.** *For  $(a, b) \in (-2, 2) \times (-2, 2)$ , the legal points on the single legal branch of the critical curve  $\mathcal{C}_w$  of (41) are not isolated. They form legal arcs that are delimited by double points.*

**PROOF.** It suffices to consider the case of positive  $w$ . From theorem 8 and its proof, we can assume the existence of a parameter pair  $(a_0, b_0)$  that is a non-isolated point of  $\mathcal{C}_w$ , for which there is critical orbit, from  $(0, -1)$  to  $(1, 0)$ , with given positive word  $w$  whose reduced word is the palindrome (28). We seek to investigate whether there persists a critical orbit with the same word for parameters in an open neighbourhood of  $(a_0, b_0)$  on  $\mathcal{C}_w$ .

From (25), we have for  $z_1 = (1, 0)$  and  $t > 1$ ,  $z_t = (Q_t(w_1, \dots, w_{t-1}), Q_{t-1}(w_1, \dots, w_{t-2}))$ , depending on the first  $t - 1$  symbols of  $\underline{w}$ . The coordinates of  $z(t)$  are polynomial, hence continuous, functions of  $a$  and  $b$ . By assumption, when  $a = a_0 \in (\zeta_{\kappa-1}, \zeta_{\kappa})$ ,  $\kappa \geq 2$  and

$b = b_0 \in (\zeta_{\ell-1}, \zeta_\ell)$ ,  $\ell \geq 2$ , we have  $z_n = (0, 1)$ . From proposition 1, the domain  $\text{dom}(M_a)$  is divided into  $\kappa$  sectors by  $L^\pm$  and the  $\kappa - 1$  rays  $\mathbf{r}_j(a)$  of (15) that are functions of  $a$  and rotate anti-clockwise as  $a$  increases. Likewise the domain  $\text{dom}(M_b)$  is divided into  $\ell$  sectors by  $L^\pm$  and the analogous  $\ell - 1$  rays that depend on  $b$  only and rotate anti-clockwise as  $b$  increases. The existence of a critical orbit with positive word can be viewed as the occurrence of a non-empty intersection of the forward semi-infinite  $L^-$ -orbit with the backward semi-infinite  $L^+$  orbit. Under the assumption that  $t = n$  is the first visit to  $(0, 1)$ , we have that the forward orbit of  $(0, -1)$  avoids all sector boundaries in  $\text{dom}(M_a)$  and  $\text{dom}(M_b)$  until it coincides with the first ray  $(U_{\kappa-1}(a), U_\kappa(a))$  in the final visit to  $\text{dom}(M_a)$  before completing the critical segment (otherwise it would arrive at  $L^-$  or  $L^+$  earlier than claimed). Because the sector boundaries in  $\text{dom}(M_a)$  or  $\text{dom}(M_b)$  are all iterates of their first ray, a critical (boundary) orbit for a positive word exists if and only if the forward orbit of  $(0, -1)$  coincides at some point with the first ray of  $\text{dom}(M_a)$ . As soon as this happens, the forward orbit of  $(0, -1)$  and the backward orbit of  $(0, 1)$  coincide along a finite critical orbit segment. Necessarily, this condition on parameters  $(a, b)$  of coinciding with the first ray of  $\text{dom}(M_a)$  is equivalent to belonging to  $\mathcal{C}_w$ .

As the critical orbit segment is finite there are a finite number of possible first ray collisions that can occur as we vary  $a$  and  $b$ . If at  $(a_0, b_0)$ , the only collision is with the first ray of  $\text{dom}(M_a)$  in the final block, we can maintain this collision by staying on  $\mathcal{C}_w$  and, by continuity of the orbit rays in  $a$  and  $b$ , continue to avoid first ray collisions in other blocks of the word on some open arc containing  $(a_0, b_0)$ . This arc will be delimited by parameter values corresponding to other first ray collisions before the final block, i.e., double points.  $\square$

We defer the question of whether there is a unique legal arc on the legal branch to a later paper [23]. This has to do with whether the earlier first ray collisions are transverse. We claim each of the rays in the orbit  $F^j(0, -1)$ ,  $j \geq 1$ , move clockwise as  $a$  or  $b$  increases, i.e., the opposite direction to the first rays which move anti-clockwise. To see this, realise that the analysis in theorem 9 applies for arbitrary  $\mathcal{C}$ , not just the critical curve. Taking  $\mathcal{C}$  to be a horizontal or vertical line shows that  $\{\varphi_t, b\} = \frac{\partial \varphi_t}{\partial a}$  and  $\{\varphi_t, a\} = -\frac{\partial \varphi_t}{\partial b}$  are both negative. So moving to the right and down on the critical curve is necessary to maintain the final ray at  $(0, 1)$ .

Let  $c = (a, b) \in \mathcal{C}$  be such that for some  $t \neq 0, n$  we have  $Q_t(c) = 0$ . We collect all values of  $t$  for which  $Q_t(c) = 0$  to form the finite sequence

$$(52) \quad T = T_w(c) = (t_1, t_2, \dots) \quad t_j < t_{j+1},$$

called the **intersection sequence** of the curve at  $c$ . In preparation for the next statement, we denote by  $Q_{t,u}$  the  $t$ -th  $Q$ -polynomials (22) for the word  $u$ , and by  $\mathcal{Q}_{t,u}$  the corresponding curve. As before, we write  $Q_t$  for  $Q_{t,w}$ .

**Theorem 11.** *Let a critical curve  $\mathcal{C}_w$  have non-empty intersection sequence  $T(c) = (t_1, t_2, \dots)$  at a point  $c$ . Then  $c$  is a double point,  $|T|$  is even, and the orbit at  $c$  is periodic with minimal period  $t_2$ . If we let, for  $j = 1, \dots, |T|/2$ ,*

$$(53) \quad w = uvu', \quad \text{where } u = w_0 \cdots w_{t_j-1} \quad \text{and} \quad |u'| = |u|,$$

*then  $c$  lies at the intersection of  $\mathcal{C}_w$  and three curves, namely*

$$(54) \quad \mathcal{C}_{u,j} = \mathcal{Q}_{t_j,u}, \quad \mathcal{C}_{v,j} = \mathcal{Q}_{n-2t_j,v}, \quad \mathcal{C}_{uv,j} = \mathcal{Q}_{n-t_j,uv}.$$

The orbital rank of  $\mathcal{C}_{v,j}$  is odd, while those of  $\mathcal{C}_{u,j}$  and  $\mathcal{C}_{uv,j}$  have, respectively, the same and the opposite parity as  $j$ .

PROOF. Let  $t \in T$ . Then  $n-t \in T$ , from (42). However, we cannot have  $t = n-t = n/2$ , because then from  $Q_{n/2} = 0$  and  $C = Q_{n/2+1} - Q_{n/2-1} = 0$  we would have that all  $Q_i$ s vanish at  $c$ , but  $Q_1 = 1$  does not. So  $|T|$  is even. By symmetry ( $w$  is a palindrome), the  $t$ -th and the  $(n-t)$ -th rays are distinct boundary rays, so at  $c$  the boundary segment visits both rays twice, and is therefore periodic. Moreover one of  $\mathcal{Q}_t$  and  $\mathcal{Q}_{n-t}$  is a critical curve, and hence  $c$  is a double point. During one period the orbit must visit both boundary rays, so the period cannot be  $t_1$ , which is minimal. For the same reason, the rays visited at consecutive  $t_j$ s must be different. It follows that the minimal period of the orbit is  $t_2$ , and that the orbital rank of  $\mathcal{Q}_{t_j}$  has the same parity as  $j$ , while that of  $\mathcal{Q}_{n-t_j}$  has opposite parity.

Consider now the decomposition (53) for fixed  $j$ . Since both  $u$  and  $uv$  are prefixes of  $w$ , with  $|u| = t_j$  and  $|uv| = n-t_j$ , we find that  $\mathcal{C}_{u,j} = \mathcal{Q}_{t_j}$  and  $\mathcal{C}_{uv,j} = \mathcal{Q}_{n-t_j}$ . This establishes the parity of the ranks of  $\mathcal{C}_{u,j}$  and  $\mathcal{C}_{uv,j}$ . To show that at  $c$  the orbit on the curve  $\mathcal{C}_{v,j}$  is the same as the middle segment of the orbit on the boundary curve  $\mathcal{C}$ , we must verify that the polynomial  $Q_{n-2t_j,v}$  has the correct initial conditions prescribed by (22).

There are two cases. If  $j$  is odd, then  $\mathcal{C}_{u,j}$  is a critical curve. From (42) we then have  $Q_{1,u}(c) = Q_{t_j-1,u}(c) = 1$ , and hence  $(Q_{t_j}(c), Q_{t_j-1}(c)) = (0, 1) = (Q_{0,v}(c), -Q_{-1,v}(c))$ . Therefore, for  $t = t_j, \dots, n-t_j$  we have  $Q_t(c) = -Q_{t-t_j,v}(c)$ , and in particular  $Q_{n-2t_j,v}(c) = Q_{n-t_j}(c) = 0$ . Thus  $\mathcal{C}_{v,j}$  is a boundary curve of odd orbital rank.

If  $j$  is even, then  $\mathcal{C}_{u,j}$  is not a critical curve, and to compute  $Q_{t_j-1}(c) = Q_{-1,v}(c)$  we cannot use (42). However, since  $t_j$  is a (not necessarily minimal) period, for some  $0 < t < t_j$  the  $t$ -th point of the orbit must visit the end ray of  $\mathcal{C}$ . Then, by concatenating two critical curves [equivalently, by composing two matrices of type (43)], we conclude that  $Q_{-1,v} = \pm 1$ . The analysis proceeds as before, and we conclude again that  $\mathcal{C}_{v,j}$  is a boundary curve with odd orbital rank.  $\square$

Theorem 11 identifies a prominent set of double points associated with a critical curve  $\mathcal{C}_w$ . They occur at intersections with curves of lower rank, corresponding to factors of  $w$ . Several points are worth considering.

- i) Since  $Q_t(c) = Q_{n-t}(c) = 0$ , the symbols  $w_t$  and  $w_{n-t}$  may be changed independently without affecting the dynamics. As a result, at a legal double point the code  $w$  may be, and typically is, *improper* (see the beginning of section 3).
- ii) The  $Q$ -polynomials of odd rank may be replaced by the corresponding  $C$ -divisor, according to proposition 7 i), eliminating the branches with even orbital rank.
- iii) From the palindrome property of  $w$  and proposition 5 ii), we find that the words  $u$  and  $u'$  generate the same curve, as so do  $uv$  and  $vu'$ . So the sub-words  $u, v$ , and  $uv$  describe completely the decomposition (53).
- iv) At a double point  $c$  of a critical curve  $\mathcal{C}_w$  there are  $|T_w(c)|/2$  distinct decompositions, each involving intersections of boundary curves of lower rank. Thus the minimum number of decomposition is 1, while the maximum is  $(\iota - 1)/2$ , where  $\iota$  is the rank of  $w$ .



For illustration, consider the rank 7 palindrome let  $w = (a^3b^4)^3a^2$  [cf. theorem 18, ii), section 5]. At the double point  $c = (0, 2\cos(\pi/5))$ , we have  $T = (2, 7, 9, 14, 16, 21)$ , so the period is equal to 7. The symbols  $w_i$  are improper for  $i = 2, 9, 16$ , so the proper code is  $(a^2b^5)^3a^2$  [see theorem 18 ii) 1 below]. Theorem 11, applied to the proper code, gives  $|T|/2 = 3$  decompositions:

$$(55) \quad \begin{array}{ccccc} t & u & v & u' & \text{ranks} \\ \hline 2 & a^2 & (b^5a^2)^2b^5 & a^2 & 1,5,1 \\ 7 & a^2b^5 & a^2b^5a^2 & b^5a^2 & 2,3,2 \\ 9 & a^2b^5a^2 & b^5 & a^2b^5a^2 & 3,1,3 \end{array}$$

The proper code is a prefix of the periodic word  $(a^2b^5)^\infty$  of period 7.

At  $c = (1, 2\cos(3\pi/11))$  we have  $T(c) = (3, 20)$ , so the period is 20. The symbol  $w_{20} = w_{n-3}$  is improper. The proper code is  $(a^3b^4)^2a^3b^3a^3$ , which is a prefix of the periodic word  $((a^3b^4)^2a^3b^3)^\infty$  of period 20. We have  $|T|/2 = 1$  decomposition:

$$t = 3, \quad u = a^3, \quad v = (b^4a^3)^2b^3, \quad u' = a^3 \quad \text{ranks: } 1, 5, 1.$$

Next we provide a formula for the rotation number at a double point, and a partial converse of theorem 11.

**Lemma 12.** *Let  $C_w$  and  $C_{w'}$  be distinct critical curves of ranks  $\iota_w$  and  $\iota_{w'}$  (with  $\iota_{w'} \leq \iota_w$ , say), which intersect at a legal point  $c$ . Then*

$$(56) \quad \theta(c) = \begin{cases} \frac{\iota_w + \iota_{w'}}{2(|w| + |w'|)} & \text{if } \text{sign}(w) \neq \text{sign}(w') \\ \frac{\iota_w - \iota_{w'}}{2(|w| - |w'|)} & \text{if } \text{sign}(w) = \text{sign}(w'), \end{cases}$$

where  $w'' = w'$  if  $\iota_{w'} \neq \iota_w$ , and otherwise  $w''$  is any prefix of  $w'$  whose length is an odd-order element of the intersection sequence of  $w'$  at  $c$ .

PROOF. Let  $w$  and  $w'$  have opposite sign, with  $w$  positive (say). Then the orbit segment of  $w$  maps  $L^-$  to  $L^+$  in  $|w|$  iterates, and that of  $w'$  maps  $L^+$  to  $L^-$  in  $|w'|$  iterates. Thus at the double point  $c$  the (non necessarily minimal) period is  $|w| + |w'|$ . Let  $\iota_w$  and  $\iota_{w'}$  be the rank of the words  $w$  and  $w'$  at  $c$ . Since  $c$  is legal, the rotation number at  $c$  is equal to half the combined rank divided by the period, regardless of whether the codes are proper or improper, the ranks being unaffected by this property.

Let now  $w$  and  $w'$  have the same sign (positive, say), hence the same initial ray at  $c$ . We have two cases. If  $\iota_{w'} < \iota_w$ , then the word  $w'$  is equivalent to a prefix of  $w$ , so  $w \sim w'u'$  for some non-empty word  $u'$ , and  $L^+$  is periodic under  $u'$ , with (not necessarily minimal) period  $|u'| = |w| - |w'|$ , and even rank  $\iota_{w'} = \iota_w - \iota_{w'}$ . Letting  $w'' = w'$  and computing the rotation number from period and rank, as above, gives the desired formulae.

If  $\iota_{w'} = \iota_w$ , then since the two orbit segments have the same initial condition, we have  $|w| = |w'|$ . Because  $w \neq w'$ , the two words will differ at some boundary ray, that is, their common intersection sequence  $T$  at  $c$  is non-empty. In particular,  $T$  has some odd-order element  $t_j$ . It now suffices to let  $w''$  be the prefix of  $w'$  of length  $t_j$  and proceed as above.

□

The intersection of critical curves described in lemma 12 always leads to a decomposition  $w = uvu'$  of type (53). Indeed, in the equal sign case, letting  $n = |w|$  and  $n - t = |w''|$  we have  $Q_{n-t,w}(c) = 0$ , and, by symmetry,  $Q_{t,w}(c) = 0$ . Then both  $t$  and  $n - t$  belong to the intersection sequence  $T_w(c)$ . The required decomposition is obtain by letting  $u = w_0 \cdots w_{t'-1}$  where  $t' = \min(t, n - t)$ . In the unequal sign case we also obtain a decomposition of type (53), by considering the word  $ww'w$ .

To study intersections of curves, we introduce a sequence of vectors in  $\mathbb{Z}[a, b]^2$ :

$$(57) \quad \mathcal{G} = (\Gamma_0, \dots, \Gamma_{n-1})$$

where

$$(58) \quad \Gamma_t = \sum_{i=0}^t \gamma_i, \quad \gamma_t = \begin{cases} (Q_t^2, 0) & \text{if } w_t = a \\ (0, Q_t^2) & \text{if } w_t = b. \end{cases}$$

A key property of this sequence is derived from theorem 9. From i) we have  $\Delta_n \equiv 0 \pmod{C}$ , and using again i) and (47) we find

$$\sum_{i=1}^{n-1} \Xi_i = \frac{\partial C}{\partial a} \sum_{\substack{0 \leq i < n \\ w_i = b}} Q_i^2 - \frac{\partial C}{\partial b} \sum_{\substack{0 \leq i < n \\ w_i = a}} Q_i^2 \equiv 0 \pmod{C}.$$

The rightmost congruence expresses the vanishing of a determinant on  $\mathcal{C}$ , which establishes the following geometrical fact.

**Corollary 13.** *If  $\mathcal{C}$  is a boundary curve, then the vectors*

$$(59) \quad \nabla C = \left( \frac{\partial C}{\partial a}, \frac{\partial C}{\partial b} \right) \quad \text{and} \quad \Gamma_{n-1}$$

*are parallel at every point of the curve.*

Thus the normal to a boundary curve may be determined without computing derivatives. More precisely, there is a rational function  $\lambda \in \mathbb{Q}(a, b)$ , such that  $\nabla C = \lambda \Gamma_{n-1}$  on  $\mathcal{C}$ . Since both vectors in (59) are non-zero, the function  $\lambda$  is regular and non-zero on  $\mathcal{C}$ . From the corollary we also deduce at once that if the rank of  $\mathcal{C}$  is greater than one, then the partial derivatives of  $C$  have the same sign [cf. theorem 8 ii)].

For every parameter pair  $c = (a, b) \in \mathbb{R}^2$ , the sequence (57) defines a polygonal on the plane (still denoted by  $\mathcal{G}$ ), obtained by connecting the elements of the sequence by line segments (figure 3). If  $c \in \mathcal{C}_w$ , then  $\mathcal{G}(c)$  is called the **polygonal of  $\mathcal{C}_w$**  (or of  $w$ ). The terms ‘legal’ for points on curves, and ‘rank’ for words or curves, will also be used for polygonals.

By construction [see (57)] the vertices of  $\mathcal{G}$  correspond to code changes, that is, to the occurrence of the factor  $ab$  or  $ba$  in  $w$ . Therefore  $\mathcal{G}$  has  $\iota$  edges (line segments) and  $\iota - 1$  vertices, where  $\iota$  is the rank of  $w$ . If  $\mathcal{C}$  is a critical curve, then from (42) and the fact that the reduced word is a palindrome, we have  $\gamma_i \equiv \gamma_{n-i} \pmod{C}$ . As a consequence, the polygonal of a critical curve is symmetrical with respect to its barycentre.

The elements of  $\mathcal{G}$  are not necessarily distinct, and if  $\Gamma_t = \Gamma_{t-1}$  for some  $t$ , then we say that  $\Gamma_t$  is an **intersection point** of  $\mathcal{G}$ . If  $\Gamma_t$  is an intersection point of  $\mathcal{G}$ , then  $\gamma_t = (0, 0)$ , whence  $Q_t(c) = 0$ . Since  $1 = Q_1 \equiv Q_{n-1} \pmod{C}$ , the end-points of  $\mathcal{G}$  are not intersection

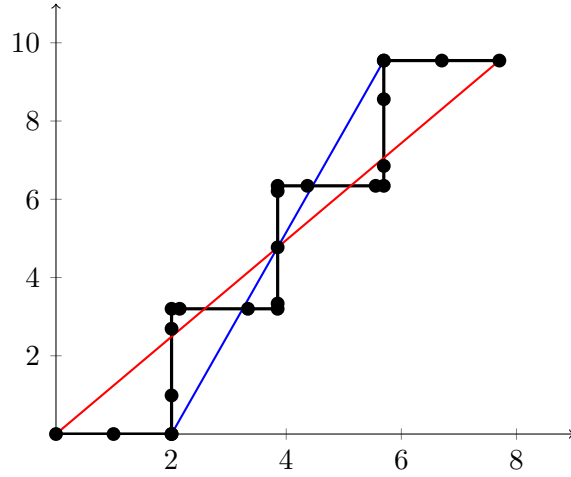


FIGURE 3. The odd-rank legal polygonal  $\mathcal{G}_w(c)$  for the word  $w = (a^3b^4)^3a^2$  at the intersection  $c = (1, 2\cos(3\pi/11))$  of the curve  $\mathcal{C}_w$  with  $\mathcal{C}_{a^3}$ . The added line segments are the medians of  $\mathcal{G}_w$  and of  $\mathcal{G}_v$  with  $v = b^4(a^3b^3)^2$ , which represent the normal to the curves  $\mathcal{C}_w$  and  $\mathcal{C}_v$  at  $c$ . These words and curves are described in theorems 11 and 18 ii).

points, so  $Q_t$  lies in the interior of the boundary segment, that is,  $c$  is a double point of  $\mathcal{C}$ , and  $t$  belongs to the intersection sequence  $T(c)$ , see (52). This argument may be reversed, to show that at a double point, the intersection points of the polygonal and the elements of the intersection sequence are in bi-unique correspondence.

Intersection points may occur at an arbitrary position on the polygonal. However, a legal intersection point necessarily corresponds to a code change, even if the code is improper. Thus the intersection points of a legal polygonal  $\mathcal{G}$  must occur at the vertices, and they occur in pairs, symmetrically placed with respect to the barycentre of  $\mathcal{G}$ . Thus all vertices of a legal polygonal are intersection points if and only if the intersection sequence is maximal:  $|T| = \iota - 1$ .

Let  $\mathcal{G}$  be a polygonal of odd rank. The segment joining its end-points will be called the **median** of  $\mathcal{G}$ , which bisects  $\mathcal{G}$ 's middle segment.  $\mathcal{G}$  is said to be **regular** if no intersection point of  $\mathcal{G}$  lies on the median. To formulate a sufficient condition for regularity, we consider the intersection sequence  $T_w(c) = (t_1, t_2, \dots)$  [cf. (52)] of a curve  $\mathcal{C}_w$  at a point  $c$ . We say that  $T_w(c)$  is **simple** if the the word  $w_0 \cdots w_{t_1-1}$  is equivalent to a word of rank 1 at  $c$ . An empty intersection sequence will also be considered simple.

**Theorem 14.** *An odd-rank legal polygonal with simple intersection sequence is regular.*

PROOF. First we show that the intersection points of an odd-rank polygonal either all lie on the median, or none of them does. Let  $T = (t_1, t_2, \dots)$  be the intersection sequence of  $\mathcal{G} = (\Gamma_0, \dots, \Gamma_{n-1})$ . The statement is trivially true if  $T$  is empty (e.g., for rank-one polygonals), since there are no intersection points. If  $T$  is non-empty, then the orbit at  $c$  is periodic with minimal period  $t_2$ , by theorem 11. By periodicity, we have  $\Gamma_{t_{2k}} = k\Gamma_{t_2}$ ,  $k = 1, \dots, |T|/2$ , that is, all even-rank intersection points lie on the segment joining the

origin  $\Gamma_0$  to the last intersection point  $\Gamma_{t_{|T|}}$ . By symmetry, all odd-rank intersection points lie on the segment joining  $\Gamma_{n-1}$  to the first odd-rank intersection point  $\Gamma_{t_1}$ . Therefore  $\Gamma_{t_1}$  lies on the median, if and only if all intersection points of  $\mathcal{G}$  lie on the median, as desired. (Note that this statement holds even if the polygonal is not legal.)

It now suffices to show that that if  $\mathcal{G}$  is legal and  $T$  is simple, then one vertex of  $\mathcal{G}$  does not belong to the median. For  $\iota > 1$  (the rank-1 case holds by definition)  $\mathcal{G}$  has at least two vertices, and all the intersection points (if any) are at the vertices. Since  $T$  is simple, the first vertex is an intersection point, which not on the median by construction. The proof of the theorem is complete.  $\square$

Next we show that regular polygonals bring about transversal intersections of curves.

**Lemma 15.** *With the notation of theorem 11, let the polygonal of a critical curve  $\mathcal{C}_w$  be regular at a legal double point  $c$ . Then the following holds:*

- i)  $\mathcal{C}_w$  intersect transversally the curves  $\mathcal{C}_{u,j}, \mathcal{C}_{v,j}, \mathcal{C}_{uv,j}$ , for  $j = 1, \dots, |T|/2$ .
- ii) All pairwise intersections of the families of curves
  - ii.1)  $\mathcal{C}_{u,j}, \mathcal{C}_{v,j}, \mathcal{C}_{uv,j}$ , for fixed  $j$ ;
  - ii.2)  $\mathcal{C}_{v,j}, \mathcal{C}_{v,j'}$ , for any  $j \neq j'$ .
 are transversal.
- iii) If  $|T| > 2$ , then the curves  $\mathcal{C}_{uv,j}, \mathcal{C}_{u,k}$  with  $j$  odd and  $k$  even are tangent.

PROOF. We fix  $j$  in (53), and let  $k(j) = |T| - j + 1$ . Then  $j$  and  $k$  have opposite parity, while  $\Gamma_{t_j}$  and  $\Gamma_{t_k}$  are symmetric with respect to the centre of  $\mathcal{G}$ . We define three subsequences of  $\mathcal{G}$ :

$$(60) \quad \mathcal{G}_{u,j} = (\Gamma_0, \dots, \Gamma_{t_j-1}), \quad \mathcal{G}_{v,j} = (\Gamma_{t_j}, \dots, \Gamma_{t_k-1}), \quad \mathcal{G}_{uv,j} = (\Gamma_0, \dots, \Gamma_{t_k-1}).$$

We want to show that  $\mathcal{G}_{u,j}, \mathcal{G}_{v,j} - \Gamma_{t_j}$  and  $\mathcal{G}_{uv,j}$  are, respectively, the polygonals of the curves  $\mathcal{C}_{u,j}, \mathcal{C}_{v,j}$  and  $\mathcal{C}_{uv,j}$ ; by construction these polygonals are then embedded in  $\mathcal{G}$ .

For  $u$  and  $uv$  the result is immediate, since these words are prefixes of  $w$ . For the word  $v$ , we must verify that for  $i = -1, \dots, n-2t_j-1$  the sequences  $Q_{t_j+i}(c)$  and  $Q_{i,v}(c)$  agree in absolute value. Indeed the corresponding codes are the same, and the verification that the initial conditions  $(i = -1, 0)$  are the same has already been given in the proof of theorem 11. It follows that  $\mathcal{G}_{v,j}$  is an embedding of the polygonal of  $\mathcal{C}_{v,j}$  into  $\mathcal{G}$  by translation (see figure 3).

Next we deal with transversality. From corollary 13, the intersections of two curves is transversal if and only if the medians of the corresponding polygonals are not parallel; so we shift our attention to medians.

The medians of the polygonals  $\mathcal{G}_{u,j}, \mathcal{G}_{v,j}$  and  $\mathcal{G}_{uv,j}$  are given by

$$A_{u,j} = [\Gamma_0, \Gamma_{t_j-1}] \quad A_{v,j} = [\Gamma_{t_j}, \Gamma_{t_k-1}] \quad A_{uv,j} = [\Gamma_0, \Gamma_{t_k-1}].$$

By symmetry, the  $A_{v,j}$ s have a common mid-point. Since  $\mathcal{G}$  is regular, from the proof of theorem 14 we have that one end-point of  $A_{v,j}$  belongs to the segment joining  $\Gamma_0$  to  $\Gamma_{t_{|T|}}$ , while the other belongs to the segment joining  $\Gamma_{n-1}$  to  $\Gamma_{t_1}$ . For this reason, no two  $A_{v,j}$  can be parallel, which is ii.2).

Let  $A = [\Gamma_0, \Gamma_{n-1}]$  be the median of  $\mathcal{G}$ . Since  $\mathcal{G}$  is regular,  $A = [\Gamma_0, \Gamma_{n-1}]$  and  $A_{v,j}$  intersect transversally at their common mid-point. Thus  $A$  and  $A_{v,j}$  are the diagonals of a parallelogram which has one vertex at the origin  $\Gamma_0$ . The medians  $A_{u,j}$  and  $A_{uv,j}$  of the other polygonals connect the origin to the vertices of  $A_{v,j}$ , so no pair of medians can be parallel. This is ii.1).

In total, we obtain  $|T|/2$  distinct non-degenerate parallelograms, sharing one diagonal  $A$ , with their second diagonals  $A_{v,j}$  forming a pencil through the centre of the parallelogram. This suffices to establish that  $A$  is transversal to all medians  $A_{u,j}$ ,  $A_{v,j}$  and  $A_{uv,j}$ , which implies the transversal intersections of the corresponding curves. This is i).

Suppose that  $|T| > 2$ . Then the medians

$$A_{uv,1}, A_{u,2}, A_{uv,3}, A_{u,4}, \dots$$

have one end-point at  $\Gamma_0$ , while (as pointed out earlier) the other end-points are collinear. This implies the tangential intersection of the corresponding curves, which is iii).

The proof of the lemma is complete.  $\square$

**Lemma 16.** *Let  $\mathcal{C}$  be a critical curve, and let  $c \in \mathcal{C}$  be a legal point. The following statements are equivalent:*

- i) *For some  $t$  and  $s \neq t$ , we have  $\varphi_t(c) - \varphi_s(c) = 0$  [cf. (44)].*
- ii) *The intersection sequence  $T(c)$  is non-empty.*

PROOF. If  $\varphi_t(c) - \varphi_s(c) = 0$ , with  $t < s$ , say, then, from the continuity and invertibility of the circle map, we have that  $\varphi_{t+j}(c) - \varphi_{s+j}(c) = 0$ ,  $j = -t, \dots, n-s$ . Thus the orbit segment is part of a periodic orbit of period  $k := s-t$ , which contains both  $\varphi_0$  and  $\varphi_n$ , that is,  $c$  is a double point. The value  $j = -t$  corresponds to a collision of the  $k$ th ray with the initial boundary ray ( $z_k = z_0$ ), while  $j = n-s$  yields the same phenomenon for the end boundary ray ( $z_{n-k} = z_n$ ). Thus  $Q_k(c) = 0$ , which means that  $k \in T(c)$ , that is,  $T(c)$  is non-empty. Conversely, if  $T(c)$  is non-empty, then by theorem 11 we have  $\varphi_{t_2}(c) - \varphi_0(c) = 0$ .  $\square$

Condition i) expresses the collision of points of the orbit of the circle map at  $c$ . The transversality—or lack of it—of such a collision is expressed by the validity—or lack of it—of the inequality  $\varphi'_t(c) \neq \varphi'_s(c)$ , where  $\varphi' = \{\varphi, C\}$ . This condition is independent from the choice of  $t$  and  $s$  along the orbit. Indeed, suppose that  $\varphi'_t(c) - \varphi'_s(c) = 0$ . Then (45) and (46) give  $(Q_t^2 + Q_{t-1}^2)\Delta_s = (Q_s^2 + Q_{s-1}^2)\Delta_t$ . Using the recursion (48), and keeping in mind that  $w_s = w_t$  and that  $Q_s Q_{t-1} = Q_t Q_{s-1}$ , we obtain  $\Delta_{s\pm 1}(Q_t^2 + Q_{t-1}^2) = \Delta_{t\pm 1}(Q_s^2 + Q_{s-1}^2)$ . Since  $z_t = \lambda z_s$ , for some  $\lambda$ , then also  $z_{t\pm 1} = \lambda z_{s\pm 1}$ , from the local linearity of  $F^{\pm 1}$ . Thus  $\Delta_{s\pm 1}\|z_{t\pm 1}\|^2 = \Delta_{t\pm 1}\|z_{s\pm 1}\|^2$ , that is,  $\{\varphi_{s\pm 1}, C\} = \{\varphi_{t\pm 1}, C\}$ . Repeating this argument an appropriate number of times, we find that all intersections are non-transversal.

An **end-point** of a curve is a point on the boundary of the curve's legal arc.

**Theorem 17.** *At an end-point  $c$  of a critical curve  $\mathcal{C}$  the intersection sequence  $T(c)$  is non-empty. Conversely, if  $T(c)$  is non-empty and  $\mathcal{G}(c)$  is regular, then  $c$  is an end-point.*

PROOF. The initial and final rays of a critical curve are boundary rays, which, by definition, remain fixed along the curve. If  $c \in \mathcal{C}$  is an end-point of  $C_w$ , then at  $c$  some intermediate ray must become a boundary ray. Equivalently, the intersection sequence  $T_w(c)$  is non-empty.

Suppose that  $T(c)$  is non-empty. Then, according to theorem 11, the curve  $\mathcal{C}$  intersects the curves of lower rank given in (54). If  $\mathcal{G}(c)$  is regular, then these curves intersect  $\mathcal{C}$  transversally, from lemma 15 i). This in turn means that the corresponding rays intersect boundary rays transversally [that is,  $\{\varphi, C\} \neq 0$  at  $c$ , see (45)], so that the code becomes illegal at  $c$ .  $\square$

The regularity of  $\mathcal{G}$  is not necessary for  $c$  to be an end-point. Indeed using lemma 16, the end-points may be characterised in terms of collisions of points of the orbit of the circle map, and these collisions need not be transversal to render a code illegal.

## 5. FIRST-GENERATION CRITICAL CURVES

Proposition 3 established all critical curves of rank 1. In this section we construct two classes of boundary curves of higher rank. First, all rank-2 curves, obtained by concatenating two rank-1 words of opposite sign. Second, all the critical curves of the **first generation**. They are constructed by concatenating an arbitrary number of copies the same rank-2 word, and then extending the resulting word—to the right or to the left—with a suitable rank-1 word. First-generation curves are regrouped to form infinite pencils, incident to the same point, the **basis** of the pencil. For reason of brevity, we shall only deal with curves which lie in the first quadrant of the (a, b)-parameter space, leaving the general case to the sequel of this paper [23].

We begin by partitioning the first quadrant into rectangular **rotational domains**  $\mathcal{D}_{\kappa, \ell}$ , given by  $\zeta_\kappa \leq a \leq \zeta_{\kappa+1}$  and  $\zeta_\ell \leq b \leq \zeta_{\ell+1}$  (figure 2). We will show that the legal arc of a rank-2 word lies within a single rotational domain, while that of a first-generation critical curve occupies two neighbouring domains (see figure 4).

We establish some notation. The boundary words  $a^\kappa$  and  $b^\ell$  have rank 1 and opposite sign. Thus, from equation (18) and lemma 12, at the intersection of the corresponding curves we have the double point

$$(61) \quad \mathbf{c}_{\kappa, \ell} := \mathbf{C}_{a^\kappa} \cap \mathbf{C}_{b^\ell} = (\zeta_\kappa, \zeta_\ell) \quad \theta(\mathbf{c}_{\kappa, \ell}) = \frac{1}{\kappa + \ell}.$$

We shall make repeated use of the functions

$$(62) \quad \vartheta_{j, m}^\pm = \frac{m}{mj \pm 1}, \quad \zeta_{j, m}^\pm = 2 \cos \pi \vartheta_{j, m}^\pm.$$

In the rest of this paper we write  $\mathcal{C}_w$  to mean the legal branch of  $C_w$ .

We now state and prove the main result of this section.

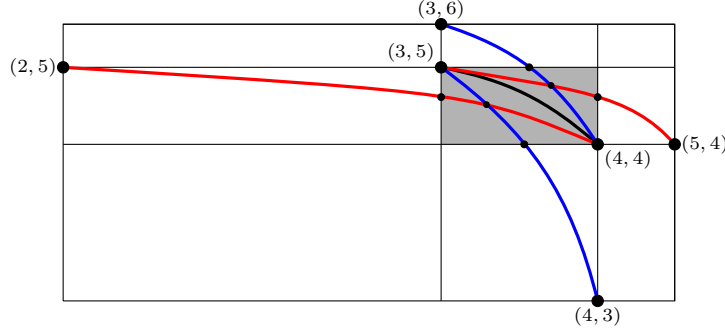


FIGURE 4. Illustration of theorem 18 for the domain  $\mathcal{D}_{3,4}$  (the grey rectangle). The black curve, of type i) in the theorem, is the axis of the domain. The blue and red curves are critical curves, of type ii) and iii), respectively and  $m = 1$ ; they intersect the boundary of the domain at the median rotation number of the adjacent vertices. The integer pairs represent the subscripts of the vertices  $\mathbf{c}_{\kappa,\ell}$ , which are joined by the curves.

### Theorem 18.

- i) For all  $\kappa, \ell \geq 2$  the words  $a^{\kappa+1}b^\ell$  and  $b^{\ell+1}a^\kappa$  give the same rank-2 curve with end-points  $\mathbf{c}_{\kappa,\ell+1}$  and  $\mathbf{c}_{\kappa+1,\ell}$ .
- ii) The following sequences of words  $w^{(m)}$ ,  $m = 1, 2, \dots$  define pencils of critical curves, with end-points  $\mathbf{e}$  (the basis) and  $\mathbf{e}_m$ , given by

	$w^{(m)}$	$\mathbf{e}$	$\mathbf{e}_m$	$\theta(\mathbf{e}_m)$	
1)	$(a^{\kappa+1}b^\ell)^m a^\kappa$	$\mathbf{c}_{\kappa,\ell+1}$	$(\zeta_{\kappa+1}, \zeta_{\ell,m}^-)$	$\vartheta_{\kappa+\ell+1,m}^-$	$\kappa \geq 2, \ell > 2$
2)	$a^{\kappa+1}(b^\ell a^\kappa)^m$	$\mathbf{c}_{\kappa+1,\ell-1}$	$(\zeta_\kappa, \zeta_{\ell,m}^+)$	$\vartheta_{\kappa+\ell,m}^+$	
3)	$(b^{\ell+1}a^\kappa)^m b^\ell$	$\mathbf{c}_{\kappa+1,\ell}$	$(\zeta_{\kappa,m}^-, \zeta_{\ell+1})$	$\vartheta_{\kappa+\ell+1,m}^-$	$\kappa > 2, \ell \geq 2,$
4)	$b^{\ell+1}(a^\kappa b^\ell)^m$	$\mathbf{c}_{\kappa-1,\ell+1}$	$(\zeta_{\kappa,m}^+, \zeta_\ell)$	$\vartheta_{\kappa+\ell,m}^+$	

while  $\theta(\mathbf{e})$  is given in (61).

PROOF. We denote by  $\Sigma_i$ ,  $i = 1, \dots, 4$  the  $i$ th quadrant in phase space.

We prove i). The reduced words of  $w = a^{\kappa+1}b^\ell$  and  $w' = b^{\ell+1}a^\kappa$  are mapped into one another by a reflection symmetry. From proposition 5 ii), it then follows that the corresponding curves are the same.

We consider the double points  $\mathbf{c}_{\kappa,\ell+1}$  and  $\mathbf{c}_{\kappa+1,\ell}$ , with proper positive codes  $(a^\kappa b^{\ell+1})^\infty$  and  $(a^{\kappa+1} b^\ell)^\infty$ , respectively. At the former point,  $w \sim a^\kappa b^{\ell+1}$ ; at the latter,  $w$  is proper. Likewise,  $w'$  is proper at  $\mathbf{c}_{\kappa,\ell+1}$  and  $w' \sim b^{\ell+1} a^\kappa$ . Thus both points are legal for both codes and from theorem 8 i), there is precisely one arc of  $\mathcal{C}_w$  connecting them. As we proceed along this arc from  $\mathbf{c}_{\kappa,\ell+1}$  to  $\mathbf{c}_{\kappa+1,\ell}$ , the  $\kappa$ -th ray  $(Q_\kappa, Q_{\kappa-1})$  of the positive orbit rotates clockwise into  $\Sigma_1$ , since  $\kappa > 1$ . The  $(\kappa+1)$ -st ray remains in  $\Sigma_2$ , because, by construction, there are no active branches of positive rank-1 curves in the interior of  $\mathcal{D}_{\kappa,\ell}$ .

Thus the legal arc of  $\mathcal{C}_w$  contains  $\mathcal{C}_w \cap \mathcal{D}_{\kappa,\ell}$ . The monotonicity of ray rotations with a prevents the prolongation of the legal arc outside  $\mathcal{D}_{\kappa,\ell}$ . An analogous argument shows that the legal arc of  $\mathcal{C}_{w'}$  is the same as that of  $\mathcal{C}_w$ . Statement i) is proved.

Next we turn to critical curves, making a preliminary remark. If a positive curve  $\mathcal{C}_w$  of rank greater than 1 intersects  $\mathcal{C}_{a^\kappa}$  at a legal point  $c$ , then  $c$  is necessarily an end-point of  $\mathcal{C}_w$ . Indeed, since the code at  $c$  is  $a^\kappa b \cdots$ , the word  $a^\kappa$  is equivalent to a prefix of  $w$ , and hence at  $c$  the intersection sequence of  $\mathcal{C}_w$  is simple. Then, by theorem 14 the polygonal of the curve is regular at  $c$ , and hence  $c$  is an end-point for  $\mathcal{C}_w$ , from theorem 17. The same holds for a negative curve intersecting  $\mathcal{C}_{b^\ell}$ .

We prove ii) part 1. Let  $w^{(m)} = (a^{\kappa+1}b^\ell)^m a^\kappa$ . At the double point  $\mathbf{c}_{\kappa,\ell+1} = (\zeta_\kappa, \zeta_{\ell+1})$ , the proper code is  $(a^\kappa b^{\ell+1})^\infty$ , and  $w^{(m)} \sim (a^\kappa b^{\ell+1})^m a^\kappa$ . Thus, for all  $m$ , the point  $\mathbf{c}_{\kappa,\ell+1}$  is a legal point of the curve  $\mathcal{C}_{w^{(m)}}$ , hence an end-point, from the above remark. We have shown that  $\mathbf{e} = \mathbf{c}_{\kappa,\ell+1}$ .

We now prolong  $\mathcal{C}_{w^{(m)}}$  inside the domain  $\mathcal{D}_{\kappa,\ell-1} \cup \mathcal{D}_{\kappa,\ell}$ . We write  $\mathbf{C}_{\kappa,\ell}$  for the axis  $\mathcal{C}_{a^{\kappa+1}b^\ell}$  and we let

$$(63) \quad L := M_{[w^{(m)}]}L^- = M_{[a^\kappa]}M_{[(a^{\kappa+1}b^\ell)^m]}L^-.$$

We fix  $a$  in the range  $\zeta_\kappa < a < \zeta_{\kappa+1}$ , and we proceed by induction on  $m$ .

Let  $m = 1$ . We choose  $b$  so that  $(a, b)$  lies in the region bounded by  $\mathbf{C}_{\kappa,\ell}$  and  $\mathbf{C}_{\kappa,\ell-1}$ . If  $(a, b) \in \mathbf{C}_{\kappa,\ell}$ , then i) and (63) give  $L = M_a^\kappa L^- \in \Sigma_1$ , with the code  $w^{(1)}$ . If  $(a, b) \in \mathbf{C}_{\kappa,\ell-1}$ , then

$$(64) \quad M_{[a^{\kappa+1}b^{\ell-1}]}L^- = L^- \quad \text{and} \quad M_{[a^{\kappa+1}]}L^- \in \Sigma_2.$$

An infinitesimal increase in  $b$  changes the code  $a^{\kappa+1}b^{\ell-1}a^{\kappa+1}$  to  $w^{(1)}$ , while leaving the orbit segment unchanged. From (64) we conclude that just above  $\mathbf{C}_{\kappa,\ell-1}$  we have  $L \in \Sigma_2$ . Thus, from continuity and monotonicity, there must be a unique value of  $b$  for which  $L = L^+$ , which shows that  $w^{(1)}$  is a boundary word of rank 3, and that the curve  $\mathcal{C}_{w^{(1)}}$  has a legal branch between  $\mathbf{C}_{\kappa,\ell-1}$  and  $\mathbf{C}_{\kappa,\ell}$ .

Suppose now that for some  $m \geq 1$  the curve  $\mathcal{C}_{w^{(m)}}$  has a legal branch between  $\mathbf{C}_{\kappa,\ell-1}$  and  $\mathbf{C}_{\kappa,\ell}$ . We fix  $a$  and choose  $b$  so that  $(a, b)$  lies between  $\mathcal{C}_{w^{(m)}}$  and  $\mathbf{C}_{\kappa,\ell}$ . If  $(a, b) \in \mathbf{C}_{\kappa,\ell}$ , then i) and (63) give  $M_{[w^{(m+1)}]}L^- = M_{[a^\kappa]}L^- \in \Sigma_1$ . If  $(a, b) \in \mathcal{C}_{w^{(m)}}$  then, by the inductive hypothesis  $M_{[w^{(m)}]}L^- = L^+$ , while  $M_{[b^{\ell+1}a^\kappa]}L^+ \in \Sigma_2$ . An infinitesimal increase in  $b$  causes the code  $w^{(m)}b^{\ell+1}a^\kappa$  to become  $w^{(m+1)}$ , without changing the orbit segment; thus on  $\mathcal{C}_{w^{(m)}}$  we have  $M_{[w^{(m+1)}]}L^- \in \Sigma_2$ . It then follows that there is a unique value of  $b$  in the specified range for which  $M_{[w^{(m+1)}]}L^- = L^+$ , which completes the induction.

The point  $\mathbf{c}_m = \mathcal{C}_{w^{(m)}} \cap \mathcal{C}_{a^{\kappa+1}}$  has the form  $(\zeta_{\kappa+1}, b)$ , for some  $b = 2 \cos 2\pi\vartheta_b$  to be determined. Such a point is legal by continuity, hence is an end-point of  $\mathcal{C}_{w^{(m)}}$  (by the remark above), so that  $\mathbf{e}_m = \mathbf{c}_m$ . To determine  $b$  we first find  $\theta(\mathbf{c}_m)$ . The words  $w^{(m)}$  and  $a^{\kappa+1}$  have the same sign, and  $w^{(m)} - a^{\kappa+1} = b^\ell (a^{\kappa+1}b^\ell)^{m-1} a^\kappa$ ; lemma 12 then gives

$$(65) \quad \theta(\mathbf{c}_m) = \frac{m}{m(\kappa + \ell + 1) - 1} = \vartheta_{\kappa+\ell+1, m}^-.$$



The parameter  $\vartheta_b$  is related to the rotation number  $\theta(\zeta_{\kappa+1}, b)$  by equation (19). Solving the latter for  $\vartheta_b$  and using (65), we obtain  $b = \zeta_{\ell, m}^-$ , that is  $\mathbf{e}_m = (\zeta_{\kappa+1}, \zeta_{\ell, m}^-)$ , as desired.

We have shown that for any  $m \geq 1$  the curve  $\mathcal{C}_{w^{(m)}}$  has a legal arc in  $\mathcal{D}_{\kappa, \ell-1} \cup \mathcal{D}_{\kappa, \ell}$  with the stated end-points. The proof of ii) 1 is complete.

We prove ii) part 2. Let  $w^{(m)} = a^{\kappa+1}(b^\ell a^\kappa)^m$ . An argument analogous to that used in ii) 1 shows that  $\mathbf{c}_{\kappa+1, \ell-1}$  is a legal point of the curve  $\mathcal{C}_{w^{(m)}}$  for all  $m$ , so that  $\mathbf{e} = \mathbf{c}_{\kappa+1, \ell-1}$ .

To prolong  $\mathcal{C}_{w^{(m)}}$  inside the domain  $\mathcal{D}_{\kappa, \ell} \cup \mathcal{D}_{\kappa, \ell-1}$  we proceed by induction on  $m$ , keeping in mind that the base case is already established, since for  $m = 1$  the words for the two statements are the same. Thus assume that for some  $m \geq 1$  the curve  $\mathcal{C}_{w^{(m)}}$  has a legal branch between  $\mathbf{C}_{\kappa, \ell-1}$  and  $\mathbf{C}_{\kappa, \ell}$ . We fix  $a$  in the range  $\zeta_\kappa < a < \zeta_{\kappa+1}$  and  $b$  so that  $(a, b)$  lies between  $\mathbf{C}_{\kappa, \ell-1}$  and  $\mathcal{C}_{w^{(m)}}$ .

If  $(a, b) \in \mathbf{C}_{\kappa, \ell-1}$ , then from part i)  $M_{[(b^\ell a^\kappa)^{m+1}]}$  maps  $L^+$  to itself, and hence  $\Sigma_2$  to itself. The second equation in (64) then implies that  $M_{[w^{(m+1)}]}L^- \in \Sigma_2$ . If  $(a, b) \in \mathcal{C}_{w^{(m)}}$ , then, by the inductive hypothesis  $M_{[w^{(m)}]}L^- = L^+$ , while  $M_{[b^\ell a^\kappa]}L^+ \in \Sigma_1$ . Since  $w^{(m)}b^\ell a^\kappa = w^{(m+1)}$ , on  $\mathcal{C}_{w^{(m)}}$  we have  $M_{[w^{(m+1)}]}L^- \in \Sigma_1$ . It then follows that there is a unique value of  $b$  in the specified range for which  $M_{[w^{(m+1)}]}L^- = L^+$ , which completes the induction.

As above,  $\mathbf{e}_m$  is an end-point of the curve, and the formulae for  $\mathbf{e}_m$  and  $\theta(\mathbf{e}_m)$  are computed as the corresponding formulae in ii) 1. The proof of ii) 2 is complete.

The proof of ii) 3,4 is obtained from that of ii) 1,2 by merely exchanging parameters.  $\square$

Let us examine theorem 18 at the light of the material of section 4. We consider the curves of case ii) 1, of rank  $2m+1$  the other cases being analogous. At the left end-point  $\mathbf{e} = (\zeta_\kappa, \zeta_{\ell+1})$  we have  $w = (a^{\kappa+1}b^\ell)^m a^\kappa \sim (a^\kappa b^{\ell+1})^m a^\kappa$ . With reference to the decomposition (53), the  $Q$ -polynomials of the  $2m$  prefixes  $u = a^\kappa, a^\kappa b^{\ell+1}, a^\kappa b^{\ell+1} a^\kappa, \dots, (a^\kappa b^{\ell+1})^m$  vanish, leading to the following maximal intersection sequence

$$T(\mathbf{e}) = (\kappa, \kappa + \ell + 1, 2\kappa + \ell + 1, 2\kappa + 2(\ell + 1), \dots, m\kappa + m(\ell + 1)).$$

From theorem 14 and 17 we conclude that the polygonal  $\mathcal{G}(\mathbf{e})$  is regular and that all these intersections are transversal.

At the right-end point we find, from (62) and ii) 1:

$$\mathbf{e}_m = (\zeta_{\kappa+1}, \zeta_{\ell, m}^-), \quad \zeta_{\ell, m}^- = 2 \cos(\pi m / (m\ell - 1)).$$

We now show that

$$T(\mathbf{e}_m) = (t_1, t_2) = (\kappa + 1, m - \kappa - 1).$$

The factor  $a^{\kappa+1}$  produces a rotation by  $\pi$ , so  $t_1 = \kappa + 1$ . The factor  $b^\ell$  produces a rotation by an angle greater than  $\pi$ . For  $m > 1$ , the quantity  $mt / (m\ell - 1)$  is an integer for  $t = m\ell - 1 = (m - 1)\ell + \ell - 1$ , and no smaller positive  $t$ , as easily verified. It follows that  $T(\mathbf{e}_m)$  has no term  $t$  between  $t_1$  and  $t_2 = m - t_1$ , and that at  $\mathbf{e}_m$  we have  $w \sim (a^{\kappa+1}b^\ell)^{m-1} a^{\kappa+1} b^{\ell-1} a^{\kappa+1}$ . Thus the intersection sequence is maximal only for  $m = 1$ .

However, the polygonal  $\mathcal{G}(\mathbf{e}_m)$  is still regular for every  $m$ , because the first intersection point lies on the first vertex.

When combined with lemma 12, theorem 18 has the following immediate corollary, which gives the rotation number of the points of intersection of the odd-rank curves in parts ii) within the domain  $\mathcal{D}_{\kappa,\ell}$ . We extend the parameters to include the case  $m = 0$ , which corresponds to the words  $a^\kappa$  and  $b^\ell$ , respectively (see proposition 3). This is legitimate, since we only make use of the fact that all these curves are legal in  $\mathcal{D}_{\kappa,\ell}$ .

**Corollary 19.** *For any  $\kappa, \ell \geq 2$  and  $n, m \geq 0$ , the following holds*

$$(66) \quad \begin{array}{ccc} w & w' & \theta(\mathcal{C}_w \cap \mathcal{C}_{w'}) \\ \hline (a^{\kappa+1}b^\ell)^n a^\kappa & (b^{\ell+1}a^\kappa)^m b^\ell & \theta_{\kappa+\ell+1, n+m+1}^- \\ a^{\kappa+1}(b^\ell a^\kappa)^n & b^{\ell+1}(a^{\kappa+1}b^\ell)^m & \theta_{\kappa+\ell+1, n+m+1}^+ \end{array}$$

Note that the rotation number depends on the parameters through the sums  $\kappa + \ell$  and  $n + m$ .

#### ACKNOWLEDGEMENTS

JAGR and AS kindly thank Queen Mary, University of London, for their hospitality. JAGR and FV thank Valerie Berthé for stimulating discussions on rotational words. JAGR thanks Tim Siu for his help in creating the parameter space plot in Figure 1. This research was supported by the Australian Research Council and by JSPS KAKENHI Grant No. JP16KK0005.

#### APPENDIX: POLYNOMIAL IDENTITIES

For the purpose of factoring various polynomials appearing in our analysis, we introduce the sequence of polynomials

$$(67) \quad \Psi_1(X) = X - 2, \quad \Psi_2(X) = X + 2, \quad \Psi_n(x + x^{-1}) = C_n(x)x^{-\phi(n)/2} \quad n = 3, 4, \dots,$$

where  $C_n(X)$  is the  $n$ -th cyclotomic polynomial (that is, the roots of  $C_n$  are the primitive  $n$ -th roots of unity [18, section 2.4]), and  $\phi$  is Euler's function [20, p 37]. For  $n > 2$ ,  $\Psi_n$  is a monic polynomial in  $X = x + x^{-1}$ , of degree  $\phi(n)/2$ . Moreover,  $\Psi_n$  is irreducible for all  $n$ , and its roots are the distinct numbers  $2 \cos(2\pi k/n)$ , with  $k$  coprime to  $n$ . These properties of  $\Psi_n$  are established from the fact that the polynomial  $C_n$  has degree  $\phi(n)$ , is irreducible and reflexive<sup>8</sup>, together with the repeated use of the identity

$$(68) \quad x^k + x^{-k} = (x + x^{-1})(x^{k-1} + x^{1-k}) - (x^{k-2} + x^{2-k}) \quad k \in \mathbb{Z}.$$

<sup>8</sup>meaning that  $X^{\phi(n)}F_n(X^{-1}) = F_n(X)$

5.1. **Iterates.** Let now  $M_X$  be as in (20). A straightforward induction shows that the matrix  $M_X^n$  can be written as

$$(69) \quad M_X^n = \begin{pmatrix} U_{n+1} & -U_n \\ U_n & -U_{n-1} \end{pmatrix}, \quad n \geq 0,$$

where  $U_n$  satisfies the recursion relation

$$(70) \quad U_{-1}(X) = -1, \quad U_0(X) = 0, \quad U_{n+1}(X) = XU_n(X) - U_{n-1}(X), \quad n \geq 0.$$

We see that for  $n \geq 1$ ,  $U_n$  is a polynomial in  $X$  with integer coefficients and degree  $n - 1$ ; the term of degree  $k$  is nonzero if and only if  $k$  has the same parity as  $n - 1$ . The recursion (70) is a special case of the more general relation

$$U_n = U_k U_{n-k+1} - U_{k-1} U_{n-k} \quad n, k \in \mathbb{Z}$$

which is obtained from (69) and the identity  $M^n = M^{n-k}M^k$ . Using (70), one sees that for  $n > 0$ ,  $U_n(X) = \bar{U}_{n-1}(X/2)$ , where  $\bar{U}_n$  is the  $n$ th Chebyshev polynomial of the second kind, whence  $U_n(2 \cos \theta) = \frac{\sin n\theta}{\sin \theta}$ .

## REFERENCES

- [1] P Alessandri and V Berthé, Three distance theorems and combinatorics on words, *Enseignement Mathématique* **44** (1998) 103–132.
- [2] V Bazier-Matte, D Racicot-Desloges and T McMillan, Friezes and continuant polynomials with parameters, *Bol. Soc. Parana. Mat.* (3) **36** (2018) 57–81.
- [3] A F Beardon, S R Bullett, P J Rippon, Periodic orbits of difference equations, *Proc. Roy. Soc. Edinburgh* **125** (1995) 657–674.
- [4] M di Bernardo, C J Budd, A R Champneys and P Kowalczyk, *Piecewise-smooth Dynamical Systems. Theory and Applications*, Springer-Verlag, London (2008).
- [5] M Boshernitzan, A condition for minimal interval-exchange maps to be uniquely ergodic, *Duke Math. J.* **52** (1985) 723–752.
- [6] F W Clarke, W N Everitt, L L Littlejohn and S J R Vorster, H J S Smith and the Fermat two squares theorem, *Amer. Math. Monthly*, **106** (1999) 652–665.
- [7] D Cox, J Little, and D O’Shea, *Ideals, Varieties, and Algorithms*, Springer-Verlag, New York (1997).
- [8] R L Devaney A Piecewise linear model for the zones of instability of an area preserving map, *Physica D* **10** (1984) 387–393.
- [9] L B Garcia-Morato, E Freire Macias, E Ponce Nuñez and F Torres Perat, Bifurcation patterns in homogeneous area-preserving piecewise-linear map, *Qualitat. Theory of Dyn. Sys.* (2018).
- [10] R L Graham, D E Knuth and O Patashnik *Concrete Mathematics*, Addison Wesley, New York (1989).
- [11] M. B. Herman, Sur la conjugation différentiable des diff’eomorphismes du cercle à des rotations, *Inst. Hautes Études Sci. Publ. Math* **49** (1979) 5–33.
- [12] A Katok and B Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge (1995).
- [13] A. Ya. Khinchin, *Continued Fractions*, English translation: The University of Chicago Press, Chicago (1964).
- [14] J C Lagarias and E Rains, Dynamics of a family of piecewise-linear area-preserving plane maps I. Rational rotation numbers, *J. Difference Eqns. Appl.* **11** (2005) 1089–1108.
- [15] J C Lagarias and E Rains, Dynamics of a family of piecewise-linear area-preserving plane maps II. Invariant circles, *J. Difference Eqns. Appl.* **11** (2005) 1137–1163.
- [16] J C Lagarias and E Rains, Dynamics of a family of piecewise-linear area-preserving plane maps III. Cantor set spectra, *J. Difference Eqns. Appl.* **11** (2005) 1205–1224.
- [17] J S W Lamb and J A G Roberts, Time-reversal symmetry in dynamical systems: a survey, *Physica D* **112** (1998) 1–39.
- [18] P J McMarthy, *Algebraic Extensions of Fields*, Dover Publications, New York (1991).

- [19] S Morier-Genoud and V Ovsienko, Farey Boat: Continued Fractions and Triangulations, Modular Group and Polygon Dissections, *Jahresbericht der Deutschen Mathematiker-Vereinigung* **121** (2019) 91-136.
- [20] I Niven, *Irrational numbers*, The Mathematical Association of America, Washington DC (1956).
- [21] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Springer-Verlag, Berlin-Heidelberg (2002).
- [22] J A G Roberts, Escaping orbits in trace maps, *Physica A* **228** (1996) 295–325.
- [23] J A G Roberts, A Saito and F Vivaldi, Rotational parameters of a family of piecewise linear maps, in preparation.
- [24] D J Simpson, The structure of mode-locking regions of piecewise-linear continuous maps: I. Nearby mode-locking regions and shrinking points. *Nonlinearity* **30** (2017) 382–444.
- [25] D J Simpson, The structure of mode-locking regions of piecewise-linear continuous maps: II. Skew sawtooth maps. *Nonlinearity* **31** (2018) 1905–1939.
- [26] D J Simpson and J D Meiss, Shrinking point bifurcations of resonance tongues for piecewise-smooth, continuous maps, *Nonlinearity* **22** (2009) 1123–1144.

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NEW SOUTH WALES, SYDNEY, NSW 2052, AUSTRALIA

*Email address:* jag.roberts@unsw.edu.au

FUTURE UNIVERSITY HAKODATE, 116-2 KAMEDANAKANO-CHO, HAKODATE, HOKKAIDO 041-8655, JAPAN

*Email address:* saito@fun.ac.jp

SCHOOL OF MATHEMATICAL SCIENCES, QUEEN MARY, UNIVERSITY OF LONDON, LONDON E1 4NS, UK

*Email address:* f.vivaldi@maths.qmul.ac.uk