Coding Closed Geodesics on Modular Surfaces by Use of the Elias Type of Codes

Reginaldo Palazzo Jr., Marinaldo Felipe da Silva, and Henrique Lazari

Abstract— **In this paper we show a procedure for coding geodesics on modular surfaces by use of the Elias type of code for source coding. This procedure implies that the arithmetic codes associated with each primitive hyperbolic matrix can be viewed as a generalization of the Elias codes. The main result of this paper establishes the procedures to be followed in order to identify the arithmetic code and the axis of the geodesic when only the probability associated with the geodesic is given. Several examples are also considered.**

I. Introduction

The traditional source coding techniques such as Huffman, Elias and Lempel-Ziv, [3], are used to compact sequences of symbols at the source output. Each one of these sequence of symbols can be viewed as an orbit $\{\phi_n(x)\}_{n\in\mathbb{Z}}$ of a discrete-time dynamical system whose evolution of a point x occurs at $t = n \in \mathbb{Z}$. On the other hand, continuous-time dynamical systems whose orbits are parameterized by the set of real numbers, are also important. In general, such systems are modelled by differential equations whose solutions are orbits. An example of a continuous-time dynamical system is the geodesic flow on surfaces with constant negative curvature.

In this context, the aim of this paper is to show how the traditional process of coding sequence of symbols (orbits) with a probability δ (real number in [0,1)) using the Elias code is identified with the process of coding simple closed geodesic (closed curve without self-crossing) based on its repelling fixed point β (real number in [0,1)) in the hyperbolic plane by use of the arithmetic code. As a consequence, the latter procedure can be viewed as a generalization of the former. Although we do not know of any previous engineering work in this context, there are, however, important contributions in the mathematical literature involving coding of geodesics. For instance, in [8], Morse proposes coding geodesics with respect to a given Dirichlet region (fundamental region) of a Fuchsian group G. In [7], it is proposed the use of continued fractions for coding geodesics on modular surfaces. Related works appear in [1], [2], [4] and [9]. In [6], it is proposed an algorithm for the construction of arithmetic codes by use of the Gauss reduction theory.

This paper is organized as follows. In Section II we state the problem and the procedure employed in the identification of the process of coding simple closed geodesics and the source coding process. In Section III, we briefly review the concepts of dynamical systems and the Elias code for source coding. In Section IV, we show the procedure for coding geodesics by use of the arithmetic codes. Such codes are derived from the assumption that the hyperbolic plane is tesselated by ideal triangles (fundamental region $\mathcal{D}$, see Fig. 2 c)). The tessellation of the hyperbolic plane by ideal triangles has the following rationality: any geodesic in the hyperbolic plane intersecting the *fundamental region* cuts adjacent sides of this region. The modular surface is identified by the quotient $\mathbb{H}^2/G$ of the hyperbolic plane $\mathbb{H}^2$ by the modular group $G = SL(2,\mathbb{Z})$, where S means special (determinant equal to 1) and L means linear, group of 2×2 matrices over $\mathbb{Z}$. Let $\mathcal{D}$ be a fundamental region for G and $\pi : \mathbb{H}^2 \longrightarrow \mathbb{H}^2/G$ be the projection. The restriction of π to $\mathcal{D}$ identifies the congruent points of $\mathcal{D}$ and take $\mathcal{D}/\mathcal{G}$ to an oriented surface with possibly marked points and cusps. In the case being considered it is topologically characterized as a three times punctured spherical surface. We present some examples of arithmetic codes derived from continued fractions. Finally, in Section V the conclusions are drawn.

II. Problem Statement

For simplicity, consider a source outputting symbols a_1 and a_2 with probabilities p_1 and p_2, respectively, such that they add to one. We consider the process of coding such a source by use of the Elias code. Each sequence at the source output can be viewed as an orbit of a dynamical system. Associated to each sequence is its probability of occurrence. Note, however, that the set of sequences and the set of probabilities is not a one-to-one mapping, since two distinct sequences may have the same probability. In order to have a (an almost) one-to-one mapping some kind of transformation has to be used. A suitable two steps type of transformation is used for the purpose of this paper. The first step is to associate a q-ary symbol to each source symbol. In our case, a binary symbol, that is, $a_1 \to 0$ e $a_2 \to 1$. Therefore, the original sequence is transformed to a sequence of 0's and 1's. The second step is to think of this sequence as representing a real number in the interval $[0,1)$. For that, we have to use a comma in front of the binary sequence. With this, and disregarding exceptional cases, to each binary sequence we have a real number in the interval $[0,1)$ representing the probability of the corresponding sequence.

An analogous statement of the source coding problem is the following: Given δ, a real number in the interval $[0,1]$ (probability associated with a sequence of symbols), and

Departamento de Telemática, FEEC-UNICAMP. This work was supported by FAPESP, CNPq, CAPES. email: palazzo@dt.fee.unicamp.br

Departamento de Matemática, Universidade Federal de Rondonia. email: felipe@unir.br

Departamento de Matemática, UNESP, Rio Claro. email: hlazari@rc.unesp.br

knowing the *apriori* probabilities of each symbol, let us say, p_1 and p_2, How the sequence can be identified? How the sequence can be encoded? The Elias procedure of chain partitioning intervals proportional to the apriori probabilities is a clever alternative to determine the sequence at the source output and to encode it. This procedure uses a tree, where to each branch the corresponding apriori probabilities are associated with the objective of matching a path probability to δ. When this occurs we have determined the sequence at the source output.

To understand the connection between the arithmetic and Elias codes we consider the concept of factorization of reduced matrices in $SL(2,\mathbb{Z})$, that is, hyperbolic matrices having attracting, α, and repelling, β, fixed points such that $\alpha > 1$ and $0 < \beta < 1$. One of the important results, [7], consists in considering G as a Fuchsian group and $\gamma_1, \gamma_2 \in G$ as hyperbolic elements having a common fixed point. Thus, the second fixed points also coincide, and consequently, have the same axis, and both are powers of a primitive matrix having the same axis.

The characterization of conjugate hyperbolic matrices in $SL(2,\mathbb{Z})$ with the same trace occurs by its attracting and repelling fixed points having periods in its continued fractions expansion that are cyclic permutations of one another. Hence, we have two other invariants of a closed geodesic, also defined from a cycle of permutation: the periods of the continued fractions expansion of its attracting and repelling fixed points. The first invariant is the *arithmetic code* of A denoted by (A). The second is the arithmetic code associated with the matrix (A^{-1}) that corresponds to the same geodesic associated with the matrix A with reversed orientation. Hence, it is convenient to have a result and a systematic procedure guaranteeing the existence of a set of hyperbolic matrices in $SL(2,\mathbb{Z})$ with a given trace t such that each matrix is always reduced by a finite number of conjugations. Under these conditions we may identify each hyperbolic matrix in $SL(2,\mathbb{Z})$ with its repelling fixed point, β. Since β takes values on the interval $[0,1)$, it follows that we may interpret β as the probability of occurrence of the corresponding geodesic.

The importance of this result is related to the fact that there exists a one-to-one correspondence between the set of reduced hyperbolic matrices (A-cycle) with the set of probabilities of each geodesic.

Analogously to the statement of the source coding problem we have the following coding of geodesics problem: Given a value of β, $\beta \in [0,1]$, How the corresponding geodesic can be identified? How the associated geodesic can be encoded?

The answer to these questions comes from the use of continued fractions with respect to the repelling fixed point. From Gauss reduction theorem we know that a hyperbolic matrix in $SL(2,\mathbb{Z})$ with a given trace may be reduced by a finite number of conjugations. On the other hand, the necessary and sufficient conditions for a hyperbolic matrix $A \in SL(2,\mathbb{Z})$ to be totally $\mathcal{D}$-reduced as a function of its arithmetic code, [7], is to show the equivalence of the following statements: 1) A is totally $\mathcal{D}$-reduced; and, 2) all the geodesic segments (geodesic components) in $\mathcal{D}$ corresponding to the conjugation class of A are clockwise oriented. Therefore, to identify all the reduced matrices as a function of its arithmetic codes, [7], is equivalent to: $(A) = (n_1, n_2, ..., n_m)$ is totally $\mathcal{D}$-reduced if and only if $\frac{1}{n_i} + \frac{1}{n_{i+1}} \leq \frac{1}{2}$ for every $i \ (mod\, m)$.

As a consequence of the previous facts, we have the following result.

Proposition II.1: Let β_A be the repelling fixed point, identified as the probability of occurrence of the geodesic associated with the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where A is a totally $\mathcal{D}$-reduced primitive hyperbolic matrix (geodesic intersecting the fundamental region $\mathcal{D}$) with positive trace ($|tr\{A\}| > 2$). Let $n_1, n_2, ..., n_m \geq 2$ be integer numbers such that $\frac{1}{\beta_A} = (\overline{n_m, n_{m-1}, ..., n_1})$, where the overbar denotes the period of the sequence in the minus continued fractions expansion. Then, the attracting fixed point α_A is given by $\alpha_A = (\overline{n_1, n_2, ..., n_m})$, and the matrix A may be represented as $A = T^{n_1} S T^{n_2} S ... T^{n_m} S$, where $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ are the transformations identifying the sides of the fundamental region $\mathcal{D}$. The associated arithmetic code is $(A) = (n_1, n_2, ..., n_m)$.

From Proposition II.1, we have identified the geodesic, or equivalently, its axis associated with the hyperbolic matrix A and the coding is realized by the arithmetic code.

III. Preliminaries

A. Review of Symbolic Dynamics

Symbolic dynamics, [5], by its relevance, occupies a great deal of the theory of dynamical systems. Due to its importance, we briefly review some concepts which are relevant to the purpose of this paper.

Let M be a metric space, the set of possible values that the dynamical system can take on. Let us assume the laws governing such a system are time-invariant. The time dependence of the system provides a mapping $\phi : M \to M$, usually continuous. Thus, if $\phi_0(x) = x$ describes the system at time $t = 0$ sec, then $\phi_1(x)$ describes the system at time $t = 1$ sec, $\phi_2(x)$ at time $t = 2$ sec, and so on. Hence, in order to study the behavior of the system it suffices to study the behavior of the sequence $\phi_0(x), \phi_1(x), \phi_2(x), ...$ These considerations lead to the following notion.

A *dynamical system* (M, ϕ) consists of a compact metric space M and a continuous mapping $\phi : M \to M$. If ϕ is a homeomorphism, then (M, ϕ) is said to be an *invertible dynamical system*. Usually, (M, ϕ) is denoted by ϕ in order to emphasize the *dynamics*.

Let (M, ϕ) be a dynamical system. The *orbit* of a point $x \in M$ is the set $\{\phi_n(x)\}_{n \in \mathbb{Z}}$ when ϕ is invertible and, $\{\phi_n(x)\}_{n \geq 0}$, otherwise. A *periodic point* is a point $x \in M$ such that $\phi_n(x) = x$ for some $n > 0$. An orbit is called a *periodic orbit* if x is a periodic point.

A realistic model of a physical system uses time continuity to model the evolution. The mathematical model for this is called a *continuous flow* consisting of a family

$\{\phi_t\}_{t\in\mathbb{R}} : X \rightarrow X$ of homeomorphisms ϕ_t of a compact metric space over itself, such that: 1) $\phi_t(x)$ is jointly continuous in t and x; and, 2) $\phi_s \circ \phi_t = \phi_{s+t}$ for every $s, t \in \mathbb{R}$. The *orbit* of a point $x \in X$ over a continuous flow $\{\phi_t\}$ is the set $\{\phi_t(x) : t \in \mathbb{R}\}$. The solution to a system of differential equations may be viewed as a continuous flow. A class of continuous flow of great interest in symbolic dynamics is the *class of geodesic flow.*

Given a surface S, there are classical notions such as the tangent vector, curvature and geodesic. *Geodesics* are curves that minimize distance. Let S be a compact surface with constant negative curvature. For each point x on S and for each vector u tangent to S in x, there exists a unique geodesic $\gamma_{u,x}$ passing by x in the direction of u. Geodesics are normally parameterized by $\mathbb{R}$, where time is the parameter in consideration. We denote by $\gamma_{u,x}(t)$ the position of the geodesic $\gamma_{u,x}$ at time t. The *geodesic flow* is a continuous flow $\{\phi_t\}_{t\in\mathbb{R}} : X \rightarrow X$ defined by $\phi_t(x, u) = (y, v)$, where $y = \gamma_{u,x}(t)$, and v is the unitary vector tangent to $\gamma_{u,x}$ in y,

In the case of discrete-time dynamical systems, the idea of the symbolic dynamics is to associate sequences of symbols with orbits of continuous flows and, to deduce the properties of these orbits by the properties of the sequences.

Two geodesic flows $\{\phi_t\}_{t\in\mathbb{R}} : X \rightarrow X$ and $\{\psi_t\}_{t\in\mathbb{R}} : Y \rightarrow Y$ are *equivalents* if the mapping $\pi : X \rightarrow Y$ takes orbits of $\{\phi_t\}$ into orbits of $\{\psi_t\}$ then π is a homeomorphism that preserves orientation.

Consider $r \in \mathbb{R}$ being written as $r = n + s$, where $n \in \mathbb{Z}$ and $s \in [0, 1)$. Thus, $r \equiv s \;(mod 1)$ implies that s is the *fractional part* of the real number r. With this notation, the mapping $\phi_r : \mathbb{R} \rightarrow \mathbb{R}$ defined by $\phi_r(x) = x+r \;(mod 1)$ is a homeomorphism. Hence, the pair $(\mathbb{R},\phi_r)$ is an invertible dynamical system. This leads to the Elias technique for source coding, [3], and it can be viewed as a geodesic flow associated with a sequence of symbols (representing a real number) to be encoded as the fractional part of this real number. Therefore, characterized as an arithmetic code since a type of continued fractions expansion will be used.

B. *Tree codes of the Elias type*

As it is well known, the entropy of a source is given by $-\sum_j p_j \log p_j$, where p_j denotes the probability of occurrence of the j-th symbol of the source. On the other hand, the average length of the codewords of a Huffman code is given by $\sum_j p_j l_j$, where l_j denotes the codeword length associated with the j-th symbol of the source. In order to encode at a rate equal to the entropy of the source, the codewords have to be chosen such that the length satisfies $l_j = -\log p_j$. Since the codeword length has to be an integer number, then it is not possible to satisfy such a condition but for a few exceptional cases. An alternative is to use tree codes.

The Elias code is a tree code of variable length used to compact source symbols. As an example, consider a binary source whose alphabet is $\{a_1, a_2\}$ with probabilities $\{0.6, 0.4\}$, respectively. The entropy of the source is 0.97 bits. At time zero, the source emits a very long sequence of output symbols. This semi-infinite sequence at the source output may be represented as a sequence of bits. For instance, ,011010111010......, where the symbols a_1 and a_2 are represented by 0 and 1, respectively. Since we have placed a comma in front of the zero, we let this sequence to be considered as an infinite binary expansion of a real number. Hence, we may say that this expansion represents a real number δ in the interval $[0, 1)$.

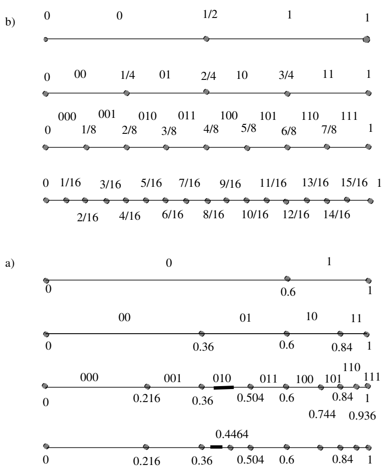

Fig. 1. Elias coding; a) source output; b) source encoder output

Taking as reference Fig. 1, divide the interval $[0, 1)$ in two sub-intervals $[0, 0.6)$ and $[0.6, 1)$. Use the first symbol of the source if it is 0, to specify that δ is in the first sub-interval, if it is 1, to specify that δ is in the second sub-interval. Again, divide each sub-interval in the same proportion, that is, in sub-intervals $[0, 0.36)$ and $[0.36, 0.6)$ or in $[0.6, 0.84)$ and $[0.84, 1)$. Use the second symbol of the source to specify one of these sub-intervals. Each source output is used in the divisions that follows.

This procedure of partitioning intervals in sub-intervals proportional to the corresponding a priori probabilities of the source symbols is similar to the procedure of partitioning intervals in sub-intervals by use of the Farey sequences. Whereas the first tessellates the interval $[0, 1)$, and by translations tessellates $\mathbb{R}$, the second tessellates the semi-infinite strip $[0, 1)$ of the hyperbolic plane, and by translations tessellates the whole hyperbolic plane.

In this way, the specified interval for any symbol sequence of the source has length equal to the probability of that sequence. Note that the intervals decrease as the length of the sequence increases. In the limit when the length of the sequence goes to infinite, exactly a point is specified.

The codeword, on the other hand, is a binary representation of δ. Consider the case shown in Fig. 1 b). As soon as the encoder has a sufficient number of source symbols, it is in the condition of determining if δ is in the interval $[0, 0.5)$ or in the interval $[0.5, 1)$, and from this the first bit of the codeword may be sent. Similarly, the second bit of the codeword may be sent from the knowledge of the interval in which δ is in, that is, $[0, 0.25)$ or $[0.25, 0.5)$ or in one of the sub-intervals $[0.5, 0.75)$ or $[0.75, 1)$. Hence, the coding procedure follows.

The decoder starts recovering the source symbols only after receiving some bits of the codeword. For instance, if the binary sequence begins with 011..., then the real number δ must be between 0.375 and 0.5; observe that in the third row of Fig. 1 a) the interval $[0.375, 0.5]$ is in the interval $[0.36, 0.504]$, therefore, the first two source symbols are $a_1 a_2$. If the binary sequence begins with 0110...., then the real number δ must be between 0.375 and 0.4375, see Fig. 1 b); then the first three source symbols are $a_1 a_2 a_1$, for the interval $[0.375, 0.4375]$ is in the interval $[0.36, 0.4464]$ in the fourth row of Fig. 1 a).

IV. Coding Simple Closed Geodesics

In this section we consider coding simple closed geodesics associate with the class of hyperbolic matrices in $SL(2,\mathbb{Z})$ by use of the *arithmetic code*. The arithmetic code is derived from the concept of continued fractions, makes use of the Gauss reduction theory, and it is specific for the modular group $SL(2,\mathbb{Z})$.

The coding of geodesics in the hyperbolic plane is associated with a primitive geodesic (geodesic which intersects the fundamental region $\mathcal{D}$) of a given set of conjugate reduced matrices of a matrix $A \in SL(2,\mathbb{Z})$, called an A-cycle. The arithmetic codes are associated with an "arithmetic" (continued fractions) of the fixed points of a fractional linear transformation (Mbius transformations). To this transformation is associated a hyperbolic matrix of the modular group $SL(2,\mathbb{Z})$.

A. Arithmetic Codes

The group $PSL(2,\mathbb{R}) = SL(2,\mathbb{R})/\{\pm I\}$ acts on the upper-half plane $\mathbb{H}^2 = \{z \in \mathbb{C} : Im\, z > 0\}$ by *Möbius transformations*: $\gamma : z \to \frac{az+b}{cz+d}$, which are oriented-preserving isometries in $\mathbb{H}^2$ and with a hyperbolic metric $ds^2 = \frac{dx^2+dy^2}{y^2}$. This action extends to the border of $\mathbb{H}^2$, $\mathbb{R} \cup \{\infty\}$. The fixed points of $\gamma \in PSL(2,\mathbb{R})$ are the solutions of the following equation: $z = \gamma(z) = \frac{az+b}{cz+d}$. If the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, associated with the transformation is hyperbolic, $|tr\{A\}| > 2$, then γ has two fixed points in $\mathbb{R} \cup \{\infty\}$, namely, the roots of a quadratic polynomial $cz^2 + (d-a)z - b = 0$, whose discriminant Δ is equal to $\Delta = (a+d)^2 - 4 > 0$.

A fixed point, denoted by α, is called *attracting*, that is, $A'(\alpha) = \frac{d}{dz}\gamma(z)_{|z=\alpha} = \frac{1}{(c\alpha+d)^2} < 1$, and the other fixed point, denoted by β, is called *repelling*, $A'(\beta) = \frac{d}{dz}\gamma(z)_{|z=\beta} = \frac{1}{(c\beta+d)^2} > 1$.

A geodesic in $\mathbb{H}^2$ from β to α, called *axis* of γ, is invariant when we choose matrix $-A$ rather than matrix A. If γ belongs to a Fuchsian group G, that is, a discrete subgroup of $PSL(2,\mathbb{R})$, its axis becomes a *closed geodesic* in the quotient space $\mathbb{H}^2/G$, or equivalently, on the modular surface identified by the fundamental region $\mathcal{D}$.

In this point, we make use of the concept of continued fractions in the construction of codes to be used for coding geodesics on modular surfaces by use of the Gauss reduction theory, [7]. Such codes are sequences of integer numbers $(n_1, n_2, ..., n_m)$, with $n_i \geq 2$, defined as a permutation cycle. They are called *arithmetic code* of the conjugation class of A, denoted by (A).

Assume now we have a set with an equivalence relation. In general terms, the *reduction theory* is an algorithm with the objective of determining the canonical representatives in each class of equivalence. Such representatives are called *reduced elements*. Each equivalence class contains a canonical set (finite and non-empty) of reduced elements forming a cycle. Following the reduction algorithm, [7], we may pass from a given element in an equivalence class to a reduced element in a finite number of steps. Applying the algorithm to this reduced element, another reduced element will be generated, and so on. Thus, all the reduced elements in a cycle are obtained.

In order to clarify these concepts, consider the reduction algorithm for co-compact Fuchsian groups, [6]. The elements whose axis intersect a given fundamental region $\mathcal{D}$ are the reduced elements. The cycle of its G-conjugate reduced elements are all the reduced elements with the same Morse code, and the intersection of its geodesics with the region $\mathcal{D}$ has all the closed geodesics associate with this particular conjugation class.

A hyperbolic matrix in $SL(2,\mathbb{Z})$ is called *reduced* if its attracting and repelling fixed points α and β, respectively, satisfy $\alpha > 1$, and $0 < \beta < 1$.

The set of all conjugated reduced matrices of a given matrix $A \in SL(2,\mathbb{Z})$ is called *A-cycle*. On the other hand, we know that for any matrix $A \in SL(2,\mathbb{Z})$ the *A-cycle* consisting of all the matrices B such that the arithmetic code associate with B is equal to the arithmetic code associate with A, that is, $(B) = (A)$, is finite and non-empty.

Let $\mathcal{D}$ be a fundamental region in $\mathbb{H}^2$. A matrix $A \in SL(2,\mathbb{Z})$ is called *$\mathcal{D}$-reduced* if A is reduced and its axis intersects $\mathcal{D}$. A hyperbolic matrix in $SL(2,\mathbb{Z})$ is called *totally $\mathcal{D}$-reduced* if all the matrices in the A-cycle are $\mathcal{D}$-reduced. The direction of the axis of a hyperbolic transformation is not invariant by conjugation, that is, some geodesics in $\mathcal{D}$ may be clockwise or counter-clockwise oriented.

B. Minus continued fractions

Let $n_0, n_1, n_2, ...$ be a sequence of integer numbers satisfying $n_i \geq 2$, for $i \geq 0$. Let us denote by $(n_0, n_1, ..., n_s)$ the finite minus continued fractions

$$(n_0, n_1, ..., n_s) = n_0 - \cfrac{1}{n_1 - \cfrac{1}{n_2 - \cfrac{1}{\ddots \, \cfrac{1}{n_s}}}},$$

and by $(n_0, n_1, n_2, ...)$ the limit $\lim_{s\to\infty}(n_0, n_1, ..., n_s)$. Conversely, (by the uniqueness of the limit), every real number α has a unique expansion in continued fractions $\alpha = \alpha_0 = (n_0, n_1, n_2, ...)$, with $n_i \in \mathbb{Z}$ and $n_1, n_2, ... \geq 2$, by the following procedure:

- $n_0 = [[\alpha_0]] + 1$ and, inductively;

• $n_i = [[\alpha_i]] + 1$, where $\alpha_{i+1} = \frac{1}{n_i - \alpha_i}$ and $[[x]]$ denotes the integer part of x.

This establishes a one-to-one correspondence between the set of real numbers α and the set of infinite sequences $(n_0, n_1, n_2, ...)$ with $n_i \in \mathbb{Z}$, and $n_1, n_2, ... \geq 2$.

This correspondence satisfies the following properties:

P1 - α is rational if and only if from a given n_i, the remaining numbers are equal to 2;

P2 - α is a quadratic irrationality, that is, a root of a quadratic polynomial with coefficients in $\mathbb{Z}$ if and only if its continued fractions expansion is eventually periodic, that is, from a certain point on the sequence repeats, $\alpha = (n_0, n_1, ...n_k, \overline{n_{k+1}, n_{k+2}, ..., n_{k+m}})$, where the overbar in $n_{k+1}, n_{k+2}, ..., n_{k+m}$ means that these numbers repeat periodically, and its minimum period is m;

P3 - α has a pure periodic expansion in continued fractions if and only if $\alpha > 1$, and $0 < \beta < 1$, where β is the conjugate root of α, that is, it is the other root of the same quadratic polynomial from which α is a root;

P4 - If $\alpha = (\overline{n_1, n_2, ..., n_m})$, then $1/\beta = (\overline{n_m, n_{m-1}, ..., n_1})$.

Property P3 is very important, for it provides an equivalent definition of a reduced matrix, that is, a matrix is a reduced matrix if and only if its attracting fixed point has an expansion in continued fractions which is pure and periodic. Katok [7] shows that the period of an expansion in continued fractions is a complete system of $SL(2,\mathbb{Z})$-invariants.

Example IV.1: Consider the matrices $A = \begin{pmatrix} 14 & -3 \\ 5 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 14 & -5 \\ 3 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 16 & -7 \\ 7 & -3 \end{pmatrix}$, $D = \begin{pmatrix} 15 & -8 \\ 2 & -1 \end{pmatrix}$. $A, B, C, D \in SL(2,\mathbb{Z})$ since its elements are integer numbers and the corresponding determinants are equal to 1. The matrices A, B, C, and D are hyperbolic since the corresponding traces are greater than 2.

The solutions of the equation $\gamma(z) = \frac{az+b}{cz+d} = z$, for each one of the given matrices are the corresponding values of its fixed points, $\alpha_A = 2.7845$, $\beta_A = .2154$; $\alpha_B = 4.640872$, $\beta_B = .359128$; $\alpha_C = 2.2746594$, $\beta_C = .4396262$; $\alpha_D = 7.46$, $\beta_D = .54$. Representing the corresponding attracting fixed points as continued fractions we have $\alpha_A = (\overline{3,5})$; $\alpha_B = (\overline{5,3})$; $\alpha_C = (\overline{3,2,2,3})$; and $\alpha_D = (\overline{8,2})$. Consequently, the corresponding arithmetic codes are $(A) = (3,5)$; $(B) = (5,3)$; $(C) = (3,2,2,3)$; and $(D) = (8,2)$.

Example IV.2: Let us determine $\frac{1}{\beta_A} = (\overline{5,3})$ by use of the continued fractions expansion.

• Let $\beta_0 = \frac{1}{\beta_A} \backsimeq 4.640872$, then:
• $n_0 = [[\beta_0]] + 1 = [[4.640872]] + 1 = 5$;
• $n_1 = [[\beta_1]] + 1$, with $\beta_1 = \frac{1}{n_0 - \beta_0} = 2.78452325 \Longrightarrow n_1 = [[2.78452325]] + 1 = 3$;
• $n_2 = [[\beta_2]] + 1$, with $\beta_2 = \frac{1}{n_1 - \beta_1} = 4.680872 \Longrightarrow n_2 = [[4.680872]] + 1 = 5$;
• $n_3 = [[\beta_3]] + 1$, with $\beta_3 = \frac{1}{n_2 - \beta_2} = 2.78452325 \Longrightarrow n_3 = [[2.78452325]] + 1 = 3$, and so on.

As a consequence, $\frac{1}{\beta_A} = (5,3,5,3,5,3,\cdots) = (\overline{5,3})$.

C. Coding geodesics by use of ideal triangles

In this subsection we consider coding geodesics in the hyperbolic plane, however with two main differences with respect to the coding of geodesics as described in subsection 4.1. The first difference is that the upper-half plane $\mathbb{H}^2 = \{z \in \mathbb{C}; Imz > 0\}$ is tessellated by ideal triangles whose vertices are on the border of the Poincaré disc, more precisely, one vertex is at infinite and the other two are in $\mathbb{R} \cup \{\infty\}$, see Fig. 2 c), whereas in subsection 4.1, the upper-half plane is tessellated by triangles with one vertex at infinite and the other two in $\mathbb{H}^2$, see Fig. 2 a). The reason for this consideration is that every geodesic is $\mathcal{D}$-reduced. The second difference is related to the elements of the coding sequence. In subsection 4.1, the elements of the sequence are power of the transformations T and T^{-1} separated by S, in this subsection such elements are geodesic elements labelled by D (right) and E (left), also with a constraint that such segments do not coincide with *cusps*. It should be clear that, as it occurs with the classical techniques for source coding, there are constraints that must be considered related to the arithmetic code just presented, one example is that the repelling fixed point can not be a reducible rational number for it is associated with a parabolic transformation (cusp).

As in subsection 4.1, the objective is to establish a connection between geodesics on modular surfaces M and the continued fractions. This connection was observed by Artin [2] when applying continued fractions deduced the existence of dense geodesics on M.

We consider the upper half-plane $\mathbb{H}^2$ with a hyperbolic metric. Geodesics in $\mathbb{H}^2$ are semi-circles whose centers are in $\mathbb{R}$ or they are orthogonal to the x-axis. The modular group $SL(2,\mathbb{Z})$ acts on $\mathbb{H}^2$ by isometries, namely; by Möbius transformations. The upper-half plane $\mathbb{H}^2$ is projected onto the modular surface M by $\pi : \mathbb{H}^2 \rightarrow \mathbb{H}^2/SL(2,\mathbb{Z})$, called *projection function*. Geodesics on M are exactly the images of the geodesics of $\mathbb{H}^2$ by the transformation π.

The idea that a sequence in which a geodesic γ cuts certain fixed lines on M (or on its lifting to $\mathbb{H}^2$), was the object of several researches, [1], [4], and [9]. In general, the borders of the canonical tessellation of $\mathbb{H}^2$ are used by copies of the fundamental region $\mathcal{D} = \{z \in \mathbb{C}; |Rez| \leq \frac{1}{2}, |z| \geq 1\}$, with the objective of determining a relation between the cutting sequence γ and the expansions in continued fractions of the ending points of proper lifting of γ.

We consider substituting the canonical tessellation by the *Farey tessellation* $\mathbb{F}$. Such a tessellation is a tessellation of $\mathbb{H}^2$ by ideal triangles, where the set of vertices is precisely $\mathbb{Q} \cup \{\infty\}$.

An oriented geodesic in $\mathbb{H}^2$ is divided in segments when it transversely cuts the triangles that compose $\mathbb{F}$. When cutting such a triangle Δ, a segment s cuts two sides meeting at a vertex which is on the border $(\mathbb{R} \cup \{\infty\})$. These oriented segments are labelled according to the position of the vertex with respect to the geodesic. If the vertex is to the right, it is labelled D, otherwise, E. This labelling is invariant by the action of the group $SL(2,\mathbb{Z})$. Consequently, for any geodesic $\overline{\gamma}$ on M (topologically, the modular surface M is a sphere three times punctured, with singularities in the images of i, $\frac{1}{2}\left(1+i\sqrt{3}\right)$ and ∞, respectively), we may associate a cutting sequence of the form $E^{n_1}D^{n_2}...$, where $n_i \in \mathbb{N}$, and the positive ending point of γ over $\mathbb{R}$ is given by the plus continued fractions of γ_∞, that is, $\gamma_\infty = n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3+...}}$.

Hence, we use the same notation $(n_1, n_2, ...)$ for representing the plus continued fractions of γ_∞, and also $[[x]]$ for the *integer part* of x, $x > 0$.

C.1 Farey tessellation

The fundamental region $\mathcal{D} = \{z \in \mathbb{C} : |Re(z)| \leq 1/2,\ |z| \geq 1\}$ for $SL(2,\mathbb{Z})$ is divided in half by the imaginary axis (dashed line in Fig. 2 a)). By translation of the left-side (denoted by A) of the region, shown in Fig. 2 b), by the transformation $z \to z+1$ and placing the two parts together, a new fundamental region is obtained for $SL(2,\mathbb{Z})$, namely the four sided polygon as shown in Fig. 2 b).

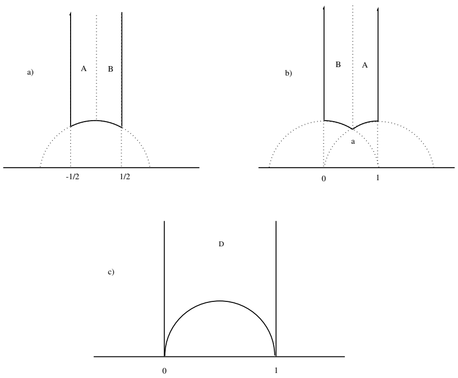

Fig. 2. Construction of the fundamental region of the ideal triangle

If S' denotes the matrix $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \in SL(2,\mathbb{Z})$ (an element of order 3), then the three images of each triangle (A and B) by I, S' and ${S'}^2$ leads to exactly the ideal triangle Δ whose vertices are 0, 1 and ∞, as may be observed in Fig. 2c. The images of Δ by $SL(2,\mathbb{Z})$ tessellates $\mathbb{H}^2$. Notice that S' results from $S' = S^{-1}T^{-1}$, where $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ are the translation and inversion matrices, respectively, and that S' fixes the point $a = \frac{1}{2}(1+\sqrt{3}i)$, see Fig. 2. This tessellation is called *Farey tessellation*, and is denoted by $\mathbb{F}$. Note that $\mathbb{F}$ may be considered as the images of the imaginary axis by $SL(2,\mathbb{Z})$.

It is not difficult to see that the images of $\{0, 1, \infty\}$ by $SL(2,\mathbb{Z})$ are exactly the points $\mathbb{Q} \cup \{\infty\}$, and that two points $\frac{p}{q}$ and $\frac{p'}{q'}$ (irreducible rational numbers) are connected by a segment of $\mathbb{F}$ if and only if, $\begin{pmatrix} p & p' \\ q & q' \end{pmatrix} \in SL(2,\mathbb{Z})$.

A description of $\mathbb{F}$ with respect to the Farey sequences is as follows: the n-th Farey sequence $\mathbb{F}_n$, is the set of rational numbers $\frac{p}{q}$ with $|p|, |q| \leq n$ arranged in increasing order. As a consequence, we have

$$\begin{aligned}
\mathbb{F}_1 &= \{-\infty, -1, 0, 1, \infty\} \\
\mathbb{F}_2 &= \{-\infty, -2, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2, \infty\} \\
\mathbb{F}_3 &= \{-\infty, -3, -2, -\frac{3}{2} - 1, -\frac{2}{3}, -\frac{1}{2}, -\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1, \frac{3}{2}, 2, 3, \infty\}
\end{aligned}$$

and so on. Therefore, $\mathbb{F}$ may be obtained by the trace of the vertical line passing by the point 0 and connecting the adjacent points in each Farey sequence, as may be seen in Fig. 3, where we have illustrated only the positive elements of each Farey sequence.

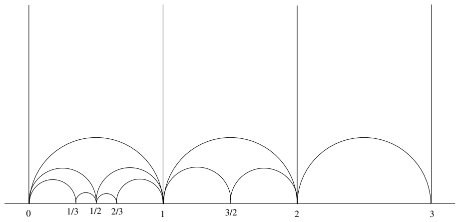

Fig. 3. Farey sequences $\mathbb{F}_1$, $\mathbb{F}_2$, and $\mathbb{F}_3$

Example IV.3: Consider the matrices A, B, C and D from Example 4.1. Hence, from the values of α_A, α_B, α_C and α_D obtained from β_A, β_B, β_C and β_D, the codes associate with the corresponding geodesics by use of the plus continued fractions expansion are

$$\begin{aligned}
\alpha_A = 2.7845 &\quad\longrightarrow\quad (2, 1, 3, 1, 1, 1, \cdots); \\
\alpha_B = 4.640872 &\quad\longrightarrow\quad (4, 1, 1, 1, 1, \cdots); \\
\alpha_C = 2.2746594 &\quad\longrightarrow\quad (2, 3, 1, 1, \cdots); \\
\alpha_D = 7.46 &\quad\longrightarrow\quad (7, 2, 3, \cdots)
\end{aligned}$$

V. Conclusions

In this paper we have shown the connection between coding geodesics on modular surfaces by use of arithmetic codes and Elias codes for source coding. The main results of this paper established the procedures in order to identify both the arithmetic code and the axis of the geodesic when only the probability associated with the geodesic is given.

References

[1] R. Adler, and L. Flato, "Cross section maps for geodesic flows," *Ergodic Theory and Dynamical Systems*, Progress in Mathematics 2, ed. A. Katok, Birkhäuser, Boston, 1980.

[2] E. Artin,"Ein Mechanisches System mit quasiergodischen Bahnen," *Collected Papers*, Addison Wesley, Reading, Mass., 1965, pp.499-501.

[3] T.M. Cover, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.

[4] G.D. Hedlund, "A metrically transitive group defined by the modular group," *Amer. J. Math.*, 57, (1935), pp.668-678.

[5] D. Lind, and M. Briam, *Symbolic Dynamics and Coding*, Cambridge University Press.

[6] S. Katok, "Reduction theory for fuchsian groups," *Mathematische Annalen*, 273, (1986), pp.461-470.

[7] S. Katok, "Coding of geodesics after Gauss and Morse," *Geometriae Dedicata*, 63, 1996, pp.123-145.

[8] M. Morse, *Symbolic Dynamics*, (unpublished), *Institute for Advanced Study Notes*, Princeton, (1966).

[9] C. Series, "Geometrical Markov coding of geodesic on surfaces of constant negative curvature," *Ergod. Th. & Dynam. Sys.*, 6, (1986), pp.601-625.