

# Drawing Closed Geodesics

Nate Collins

Jeff Ginn

Sarah Salko

August 9, 2003

## Abstract

This paper focuses on minus continued fractions, geodesics, and their relation to hyperbolic functions. Throughout this paper, we use definitions and notation based on those of Svetlana Katok. See [1] for any clarification.

## Acknowledgements

This research was done during Summer 2003 at Pennsylvania State University under the guidance of Dr. Misha Guysinsky and sponsorship of the National Science Foundation Research Experience for Undergraduates grant (NSF DMS-\*\*\*\*\*).

## 1 Minus Continued Fractions

Minus Continued Fractions (MCFs) are used to write real and rational numbers in the form of a fraction. The standard form of a minus continued fraction is:

$$\alpha = (a_0, a_1, \dots, a_n) = a_0 - \frac{1}{a_1 - \frac{1}{a_2 - \frac{1}{\ddots - \frac{1}{a_n}}}}$$

Using this process, most numbers reach a period, or a repeated sequence of values in the fraction's expansion.

After looking at MCF, we can now define quadratic irrationalities, which are one of the main operations used to find geodesics in our program.

**Definition.** A real number is a **quadratic irrationality** if it is a real root of the equation  $ax^2+bx+c$  where  $a, b, c \in \mathbb{Z}, c \neq 0$ , and the discriminant  $D = b^2 - 4ac$  is positive and not a perfect square.

Furthermore, a quadratic irrationality can be generalized in the form

$$\alpha = \frac{m_0 + \sqrt{D}}{l_0}$$

## 2 Hyperbolic Geometry

The hyperbolic plane  $H$  is the upper-half plane  $\{(x, y) \in \mathbb{C} \mid y \geq 0\}$  with the metric

$$ds = \frac{\sqrt{dx^2+dy^2}}{y}$$

By using this hyperbolic plane we can now look at geodesics in detail.

**Definition. Geodesics** are shown visually in  $\mathbb{R}^2$  as straight lines and semi-circles orthogonal to the real line. More importantly, they are defined as the shortest curve or length-minimizing path with respect to a certain metric, which in this case is  $H$ .

A geodesic has two endpoints, and with two unique points a geodesic can be drawn in  $\mathbb{R}^2$ .

The most important tie between the hyperbolic plane and geodesics comes from the source of the hyperbolic metric. This metric gives the length of curves in  $H$  in the same way a Euclidean metric does. Given some differentiable curve  $\gamma : (a, b) \subset \mathbb{R} \rightarrow H$ , its length is :

$$h(\gamma) = \int_a^b \frac{\sqrt{\frac{dx}{dt}^2 + \frac{dy}{dt}^2}}{y(t)} dt$$

with respect to the hyperbolic metric.

This formula is very helpful when trying to determine the endpoints of a geodesics.

Now it is important to define the group  $SL(2, \mathbb{R})$  as the set of real  $2 \times 2$  matrices with determinant one under matrix multiplication.  $PSL(2, \mathbb{R})$  contains the group  $SL(2, \mathbb{R})$  and also the positive and negative  $2 \times 2$  identity matrices.

**Definition.** A **Mobius transformation**, which is also known as a linear fractional transformation, is a quotient of two linear expressions that are a bijection on the extended complex plane. These transformations are isometries that preserve angles and are defined as

$$T(z) = \frac{az+b}{cz+d}$$

where  $a, b, c, d$  are complex numbers.

**Definition.** An **isometry** is found in  $H$  and acts on itself if it preserves the hyperbolic distance between two metric spaces.

Moving away from the basic definitions of hyperbolic geometry, we can not take about specific transformation, such as types of isometries. There are two types of isometries: *orientation – preserving* and *orientation – reversing*. They are defined by the signs of their determinants and the absolute value of their trace. *Orientation – preserving isometries* are found in the group  $PSL(2, \mathbb{R})$ , while *orientation – reversing isometries* are Mobius transformations with determinant equal to negative one. There are three types of orientation preserving isometries:

$$\begin{aligned} |t| > 2 & \text{ hyperbolic} \\ |t| < 2 & \text{ elliptic} \\ |t| = 2 & \text{ parabolic} \end{aligned}$$

Generally, we are dealing with the hyperbolic case. A hyperbolic transformation has two fixed points  $u$  and  $w$ , which are found by solving the following formula:  $cz^2 + (d - a)z - b = 0$ . Now we can define the fixed points as

$$u, w = \frac{a-d \pm \sqrt{a+d^2-4}}{2c}$$

These fixed points  $u$  and  $w$  are known to be either attracting or repelling fixed points. These classifications can be obtained in the following way. Taking the general fixed point  $v_i$  as an eigenvector  $(x_i, y_i)$ , it can then be manipulated into the form of an eigenvalue  $\frac{x_i}{y_i} = v_i = T(w)$ . Then, the derivative of this function  $T(w)$  determines whether the fixed points are attracting or repelling, and are found by applying the formula

$$T'(v_i) = \frac{1}{(cv_i+d)^2}$$

where  $v_i = u, w$ .

**Definition.** Let  $\gamma$  be a modular group, which contain all Mobius transformations of the form  $\frac{az+b}{cz+d}$ , and  $\gamma$  is a subgroup of all the isometries in  $H$ . Then, there is a subset  $F$  of  $H$ , which is a **fundamental region** of  $\gamma$  if

- 1)  $F$  is a closed region in  $H$  bounded by a finite number of geodesics
- 2) no two distinct points in  $F$  are equivalent in  $\gamma$  and
- 3)  $z$  is in  $H$ , then there is a  $z'$  in  $F$  such that  $z' = z$  in  $\gamma$ .

At this point, we now return to all Mobius transformations, which will now be expressed as a series of matrices

$$T^{n_1} S T^{n_2} S \dots T^{n_k} S$$

$T^{n_k}$  can be identified as both  $T$  and  $T^{-1}$ . These functions  $T, T^{-1}, S$  can be written as both matrices and transformations with the subsequent forms:

$$\begin{aligned}
T &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = z + 1 \\
T^{-1} &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = z - 1 \\
S &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -1/z
\end{aligned}$$

The group  $PSL(2, \mathbb{Z})$  is generated by all three of these functions, and can be composed as a series of transformations, which will help map closed geodesics.

Looking at a basic picture of the fundamental region, the left vertical side is identified with the right vertical side through transformations  $T(z)$ . Additionally, the two arcs of the unit circle are interchanged by the transformation of order two  $S(z)$ .

A geodesic is closed when it is in one-to-one correspondence with classes of hyperbolic elements in  $\gamma$ , meaning that the geodesic will return to its initial point while the direction remains the same when the tangent line to the geodesic is taken. So conclusively, there is a string of these generators,  $T, T^{-1}, S$  that will return to the geodesic to its point of origin.

### 3 Coding Closed Geodesics

Here we introduce two finite sequences of integers defined up to cyclic permutations used to code closed geodesics.

**Definition.** Given a hyperbolic matrix  $A$  that fixes the axis of a closed geodesic  $C$ , the **arithmetic code** of  $C$  is the period of the minus continued fraction expansion of the attracting fixed point  $w$ . Suppose  $w = (n_0, \dots, n_k, \overline{n_{k+1}, \dots, n_{k+m}})$  then the arithmetic code is denoted

$$(A) = (n_{k+1}, n_{k+2}, \dots, n_{k+m})$$

**Definition.** Given a hyperbolic element  $\gamma$  that fixes a closed geodesic  $C$  the **geometric code** of  $C$  is denoted

$$[\gamma] = [n_0, n_1, \dots, n_m]$$

where  $\gamma = T^{n_1} S T^{n_2} S \dots T^{n_k} S, n_i \in \mathbb{Z}$ .

A convenient way to think of the geometric code is to associate the transformation  $T$  with the right vertical side of the fundamental region for  $SL(2, \mathbb{Z})$ ,  $T^{-1}$  with the left, and  $S$  to the two arcs. The geometric code can then be found by adding a positive integer equal to the number of times a segment of the closed geodesic exits through the right vertical side before exiting through one of the arcs, and a negative integer for the number of times a segment exits through the left vertical side before exiting through one of the arcs.

### 4 Reduction Theory

A reduction theory is defined on hyperbolic matrices.

**Definition.** A hyperbolic matrix  $A$  with attracting and repelling fixed points  $w, u$ , respectively, is **reduced** if

$$w > 1, 0 < u < 1$$

The following lemma states this condition is equivalent to the minus continued fractions expansion of  $w$  being purely periodic. For proof see [2].

**Lemma.** Let  $\alpha \in \mathbb{R}$ .  $\alpha$  has a purely periodic minus continued fraction expansion iff  $\alpha > 1$ , and  $0 < \alpha' < 1$  where  $\alpha'$  is the other root of the same quadratic polynomial as  $\alpha$  is.

**Reduction Algorithm** There is a finite number of reduced matrices in  $SL(2, \mathbb{Z})$  with a given trace  $t$ ,  $|t| > 2$ . Any hyperbolic matrix in  $SL(2, \mathbb{Z})$  with a trace  $t$  can be reduced by a finite number of standard conjugations. Applied to a reduced matrix  $A$ , this conjugation gives another reduced one. Any reduced matrix conjugate to  $A$  is obtained from  $A$  by a number of standard conjugations. Thereby the set of reduced matrices is decomposed into disjoint cycles of conjugate matrices [2].

The complete proof may be found in [2]. We present an outline of the proof below.

**Idea of Proof** Let  $k = a - d - 2c$ .  $|k| < \sqrt{D}$  where  $D = (a + b)^2 - 4 = t^2 - 4$  where  $t$  is the trace of a hyperbolic matrix  $A$ . This implies  $k$  may only assume a finite number of values. To establish finiteness express  $a, b, c, d$  in terms of  $k, D, t$  where  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Given a hyperbolic matrix  $A$  with fixed point  $w_B = (n_0, \dots, n_k, \overline{n_{k+1}, \dots, n_{k+m}})$  we obtain a matrix  $B$  with fixed point  $w_B = (\overline{n_{k+1}, \dots, n_{k+m}})$  by conjugation of  $A$  by the matrix  $P = ST^{-n_k} ST^{-n_{k-1}} \dots ST^{n_0}$ . From the above lemma  $B$  is reduced.

## 5 Algorithms

Of the many algorithms involved in finding the closure of a hyperbolic geodesic two, if any, are of interest.

**Definition** Let  $\mathcal{Q} = \{\frac{a \pm \sqrt{b}}{c} \mid \in \mathbb{R} \setminus \mathbb{Q}, a, b, c \in \mathbb{Z}\}$  denote the set of quadratic irrationalities introduced earlier.

It is obvious that  $T^n(\mathcal{Q}) \subset \mathcal{Q}$ , and  $n\mathcal{Q} \subset \mathcal{Q}$ . The calculation

$$\frac{c}{a \pm \sqrt{b}} = \frac{c(a \mp \sqrt{b})}{(a \pm \sqrt{b})(a \mp \sqrt{b})} = \frac{ca \mp \frac{|c|}{c} \sqrt{c^2 b}}{a^2 - b} \in \mathcal{Q}$$

shows that  $\frac{1}{\mathcal{Q}} \subset \mathcal{Q}$  and thus  $S(\mathcal{Q}) \subset \mathcal{Q}$ .

To compute the minus continued fraction expansion of some  $q \in \mathcal{Q}$ , the numbers  $q_0 = q, q_{i+1} = \frac{1}{\lceil q_i \rceil - q_i}$  must be calculated exactly. It remains only to show that the ceiling of a  $q \in \mathcal{Q}$ , then all  $q_i$  can be computed exactly.

**Claim**  $\forall q = \frac{a + \sqrt{b}}{c} \in \mathcal{Q}, \exists n \in \mathbb{Z}$  s.t.  $(n - 1)^2 < b < n^2$ , and  $\lceil q \rceil = \lceil \frac{a+n}{c} \rceil$ . Well, its obvious that  $x = \frac{a+n-1}{c} < q = \frac{a+\sqrt{b}}{c} < \frac{a+n}{c} = y$  and so  $y - q < y - x = \frac{1}{c}$ . So suppose by contradiction that  $\lceil y \rceil > \lceil q \rceil = k$  then  $q < \frac{kc}{c} < \frac{kc+d}{c} = y, d \geq \in \mathbb{N}$ , and thus  $y - q > \frac{d}{c} \geq \frac{1}{c}$ .

Hence the  $q_i$  can be calculated exactly. Now the integers  $a_i$  mentioned above are simply  $\lceil q_i \rceil$ , and  $q_i$  determines all  $q_{i+j}, j \geq 1 \in \mathbb{N}$ , so if  $q_m = q_n$  then  $q_{m+j} = q_{n+j} \forall j$  and so  $\alpha$ , the MCF of  $q$  is periodic with period  $(\lceil q_m \rceil, \dots, \lceil q_n \rceil)$ .

**Definition** Let  $\mathcal{F}$  be the set  $\{a = x + iy \in \mathbb{C} \mid \frac{-1}{2} \leq x \leq \frac{1}{2}, |a| \geq 1\}$ .  $\mathcal{F}$  is a *fundamental region* for  $H$  i.e. given any  $h \in H$  there is one and only one point in  $\partial \mathcal{F}$ , the interior of  $\mathcal{F}$ , from the orbit of  $h$  under  $PSL(2, \mathbb{Z})$ .

**Definition** Let  $\mathcal{C}(q_1, q_2)$  denote the unique geodesic in  $H$  determined by  $q_1 < q_2 \in \mathcal{Q}$ , namely  $\mathcal{C}(q_1, q_2) = \{re^{i\theta} + c \mid r = \frac{q_2 - q_1}{2}, c = q_1 + r, 0 \leq \theta \leq \pi\}$ .

When attempting to plot geodesics  $\mathcal{C}(q_1, q_2)$  on  $\mathcal{F}$  and  $\mathcal{C}(q_1, q_2) \cap \mathcal{F} = \emptyset$  the question of finding some  $\gamma \in PSL(2, \mathbb{Z})$  s.t.  $\mathcal{C}(\gamma(q_1), \gamma(q_2)) \cap \mathcal{F} \neq \emptyset$  arises.

There are 3 mutually exclusive possibilities:

1.  $\mathcal{C}(T^n(q_1), T^n(q_2)) \cap \mathcal{F} \neq \emptyset$  for some  $n \in \mathbb{Z}$ .
2.  $\nexists n$  as in 1) but  $\exists k \in \mathbb{Z}$  s.t.  $q_1 < k < q_2$ .
3.  $\exists m \in \mathbb{Z}$  and  $0 < \delta_1 < \delta_2 < 1 \in \mathcal{Q}$  s.t.  $q_1 = m + \delta_1$  and  $q_2 = m + \delta_2$ .

The description of 1) gives  $\gamma$ . For 2) consider  $\mathcal{C}(T^{-k}(q_1), T^{-k}(q_2))$ , and let  $\Lambda = (0, \lambda)$  be the point on  $\mathcal{C}(T^{-k}(q_1), T^{-k}(q_2))$  with  $x = 0$ . Since 1) is satisfied for no  $n$  we have  $\lambda < 1$  and so  $|S(\Lambda)| = \frac{1}{\lambda} > 1$ . Hence  $\mathcal{C}(ST^{-k}(q_1), ST^{-k}(q_2))$  intersects  $\mathcal{F}$ . For 3) either  $\delta_1 < 1/2$  or  $1 - \delta_2 < 1/2$ . Suppose  $\delta_1 < 1/2$ , the other case is similar, and consider  $\mathcal{C}(ST^{-m}(q_1), ST^{-m}(q_2))$ .  $|\mathcal{C}(ST^{-m}(q_1), ST^{-m}(q_2))| = |\frac{-1}{\delta_1} - \frac{-1}{\delta_2}| = \frac{|\delta_2 - \delta_1|}{|\delta_1 \delta_2|} > \frac{|\delta_2 - \delta_1|}{|1/2| \cdot 1} = 2(q_2 - q_1)$  and so after a finite number of steps we arrive at case 1) or case 2).

## References

- [1] S. Katok. Continued Fractions, Hyperbolic Geometry, and Quadratic Forms. MASS REU Lectures. Summer 2001. pp. 121-160.
- [2] S. Katok. Coding of Closed Geodesics After Gauss and Morse.