2008

A Fundamental Region on Two Copies of the Hyperbolic Plane

Alyssa Jesse Soenksen

A FUNDAMENTAL REGION ON TWO COPIES OF THE HYPERBOLIC PLANE

A Thesis

Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors with Distinction

Alyssa Jesse Soenksen

University of Northern Iowa

May 2008

This Study By: Alyssa Soenksen
Entitled: A Fundamental Region on Two Copies of the Hyperbolic Plane

has been approved as meeting the thesis requirements for the Designation University Honors with Distinction.

5/5/08
Date

Theron Hitchman, Honors Thesis Advisor, Mathematics

5/9/08
Date

Jessica Moon, Director, University Honors Program

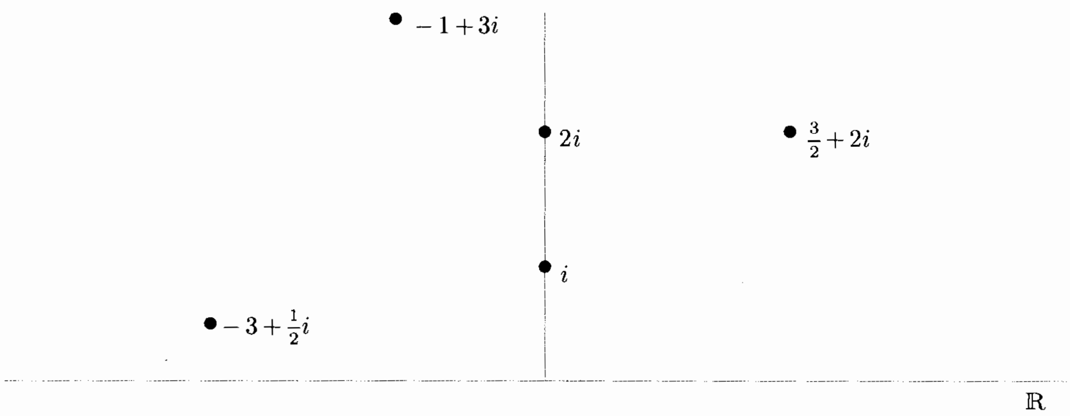
Hyperbolic geometry is a beautiful, non-Euclidean space that hosts spectacular patterns and infinite designs. To learn about this space, this paper will focus on how linear fractional transformations act on this space and the patterns that reveal themselves. To expand on this, I will construct a fundamental domain under these mappings and explore the coding of closed geodesics on the fundamental domain, which requires an understanding of continued fractions. My research will then be applied to two copies of the hyperbolic plane. My goal is to understand fundamental regions in this space and eventually the geodesics.

This paper is intended for a second- or third-year mathematics student who has completed multivariable calculus and a semester of modern algebra. Some proofs or examples may require some understanding of real analysis.

Hyperbolic Geometry

The following information is based largely on the work of Svetlana Katok, from both her work published in *Fuchsian Groups* [4] and a lecture series at MASS REU [3].

Two-dimensional hyperbolic geometry can be represented by many different models. For the purpose of this paper, we will only consider the upper half-plane model, or the *Lobachevski* plane, where $\mathfrak{H} = \{x \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$.



*Example of a few points contained in $\mathfrak{H}$

Notice that this model is bounded below by the real axis, $\mathbb{R} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) = 0\}$, but does not contain it; this is known as the *boundary at infinity*.

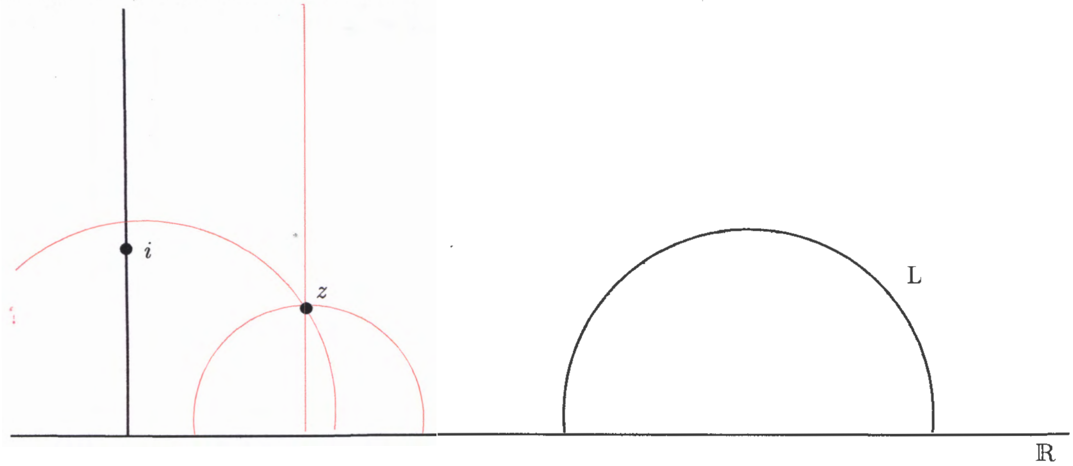
Definition 1. *In the presence of a metric, a* geodesic *is the shortest path between two points*

Consider the concept "the shortest distance between two points is a straight line." Geodesics are exactly these "straight lines." In the Euclidean plane, this metric is $\mathrm{ds} = \sqrt{\mathrm{dx}^2 + \mathrm{dy}^2}$.

In the upper-half plane model,

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

is defined as the *hyperbolic metric*. Under this metric, geodesics are represented by rays orthogonal to the real axis and by semi-circles with diameters on the real axis. Any two points are joined by a unique geodesic, and the distance between these points is measured along this line. This geometry is a Non-Euclidean geometry in that given a line L, and a point z not on L, there exists more than one line that passes through z but does not intersect L.



*Hyperbolic geometry is a non-Euclidean space. This is an image of a typical point z and L and geodesics that pass through z but do not intersect with L.

Let I = [0,1] and $\gamma = \{x(t) + i\, y(t) \epsilon \mathfrak{H} \,|\, t \epsilon I\}$, where $\gamma : \mathrm{I} \rightarrow \mathfrak{H}$ is a piecewise differentiable path. Then the hyperbolic length h(γ) is defined by

$$h(\gamma) = \int_0^1 \frac{\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}}{y(t)} dt = \int_0^1 \frac{|\frac{dz}{dt}|}{y(t)} dt$$

and the distance between two points $z,\ w \,\epsilon\, \mathfrak{H}$, or *hyperbolic distance*, is $\rho(z, w) = \inf h(\gamma)$, where the *infimum* is greatest lower bound over the set of paths.

Example 2. Let $z_1,\ z_2 \epsilon \mathfrak{H}$ such that $z_1 = ia$, $z_2 = ib$ and b>a. Let x(t)=0 and y(t) = a+(b-a)t, $t\epsilon$I=[0, 1] be the path from z_1 to z_2 along the y-axis. Then $\gamma{:}\, I \rightarrow \mathfrak{H}$ where $\gamma(t) = x(t) + iy(t)$ and

$$h(\gamma) = \int_0^1 \frac{\sqrt{0^2 + (b-a)^2}}{a + (b-a)\,t} dt = \int_0^1 \frac{b-a}{a + (b-a)\,t} dt = \int_0^1 \frac{1}{u} du, \; (u = a + (b-a)t)$$

$$= \ln|a + (b-a)t| \,|_0^1 = \ln|b| - \ln|a| = \ln\frac{b}{a}$$

Notice that $\gamma = x(t) + iy(t)$ is the path from z_1 to z_2 that follows the y-axis, and is obviously the shortest path between the points.

Now, consider the special linear group of degree 2 over ℝ,

$$\mathrm{SL}(2,\mathbb{R}) = \{M = [\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}], |\, a, b, c, d \epsilon \mathbb{R} \text{ and } \det(M) = a\,d - b\,c = 1\}$$

and the set of *Mobius transformations* of ℂ onto itself of the form

$$\{z \rightarrow \frac{a\,z + b}{c\,z + d} \,|\, a, b, c, d \epsilon \mathbb{R}, a\,d - b\,c = 1\}$$

which we will denote as PSL(2,$\mathbb{R}$). Notice that for $g \epsilon$PSL(2,$\mathbb{R}$), g and $-g$ provide the same transformation and therefore is the set SL(2,$\mathbb{R}$) up to the identification of g and $-g$. Notice that by letting $c=0, d=1$, we can see that PSL(2,$\mathbb{Z}$) contains all transformations of the form $z \rightarrow az+b$, and also for $a=0, b=-1, c=1$, and $d=0$, PSL(2,$\mathbb{Z}$) contains the transformation $z \rightarrow -\frac{1}{z}$.

Theorem 3. *PSL(2,$\mathbb{R}$) acts on $\mathfrak{H}$ by homeomorphisms (continuous and bijective mappings with a continuous inverse).*

Proof. Let T ϵ PSL(2, $\mathbb{R}$) and w=T(z)=$\frac{az+b}{cz+d}$. Then we can write

$$w = \frac{(az+b)(c\bar{z}+d)}{(cz+d)(c\bar{z}+d)} = \frac{(az+b)(c\bar{z}+d)}{|cz+d|^2} = \frac{ac|z|^2+adz+bc\bar{z}+bd}{|cz+d|^2}.$$

From this,

$$\mathrm{Im}(w) = \frac{w-\bar{w}}{2i} = \frac{(ac|z|^2+adz+bc\bar{z}+bd)-(ac|z|^2+ad\bar{z}+bcz+bd)}{2i\,|cz+d|^2}$$

$$= \frac{(ad-bc)z-(ad-bc)\bar{z}}{2i|cz+d|^2} = \frac{z-\bar{z}}{2i|cz+d|^2} = \frac{\mathrm{Im}(z)}{|cz+d|^2}$$

Since Im(z)>0, for $z\epsilon\mathfrak{H}$, then for each w such that Im(w)>0 we can find z such that T(z)=w and therefore T maps $\mathfrak{H}$ onto itself. Further, PSL(2, $\mathbb{R}$) are well-defined mappings in $\mathfrak{H}$.

To show one-to-one, let $z, w \epsilon \mathfrak{H}$ such that $z \neq w$. By way of contradiction, assume T(z) = T(w), then

$$\frac{az+b}{cz+d} = \frac{aw+b}{cw+d}.$$

Using algebra,

$$adz - adw - bcz + bcw = 0 \Rightarrow (ad-bc)(z-w) = 0$$

Since ad-bc=1, then z=w. This is a contradiction and T is one-to-one and therefore bijective.

If w=T(z), then we can solve $z = -\frac{b-dw}{a-cw}$, so since a,c cannot both be zero and b,d cannot both be zero, it follows that T and T^{-1} are continuous and all conditions are satisfied. □

An *isometry* is a homeomorphism that preserves distance. Proving that elements of PSL(2,$\mathbb{R}$) are isometric mappings of $\mathfrak{H}$ will be crucial in finding geodesics and eventually measuring distances.

Theorem 4. *The mappings defined by PSL(2, $\mathbb{R}$) are isometries of $\mathfrak{H}$.*

Proof. Consider I = [0,1] and $\gamma: I \rightarrow \mathfrak{H}$ as a piecewise differentiable path given by $z(t) = x(t) + iy(t)$. Let TϵPSL(2, $\mathbb{R}$) and $w(t) = T(z(t)) = u(t) + iv(t)$. It is necessary to show that h(γ)=h(T(γ)).

Notice, from w = $\frac{az+b}{cz+d}$

$$\frac{\mathrm{dw}}{\mathrm{dz}} = \frac{a(cz+d)-c(az+b)}{(cz+d)^2} = \frac{ad-bc}{(cz+d)^2} = \frac{1}{(cz+d)^2}$$

From Theorem 1, Im(w) = $\frac{\mathrm{Im}(z)}{|cz+d|^2}$. This implies $v = \frac{y}{|cz+d|^2}$ and $|\frac{\mathrm{dw}}{\mathrm{dz}}| = |\frac{v}{y}|$. Therefore

$$h(T(\gamma)) = \int_0^1 \frac{|\frac{\mathrm{dw}}{\mathrm{dt}}|}{v(t)}\mathrm{dt} = \int_0^1 \frac{|\frac{\mathrm{dw}}{\mathrm{dz}}\frac{\mathrm{dz}}{\mathrm{dt}}|}{v(t)}\mathrm{dt} = \int_0^1 \frac{|\frac{v(t)}{y(t)}\frac{\mathrm{dz}}{\mathrm{dt}}|}{v(t)}\mathrm{dt} = \int_0^1 \frac{|\frac{\mathrm{dz}}{\mathrm{dt}}|}{y(t)}\mathrm{dt} = h(\gamma)$$

And $\mathrm{PSL}(2,\mathbb{R}) \subset \mathrm{Isom}(\mathfrak{H})$. □

Proposition 5. *Given L, a straight line or semicircle orthogonal to the real axis, then for carefully chosen $\alpha, \beta \in \mathbb{R}$, $T = -(z-\alpha)^{-1} + \beta$ maps L to the imaginary axis.*

First, it's important to clarify that $T \in \mathrm{PSL}(2,\mathbb{R})$.

$$-\frac{1}{z-\alpha} + \beta = \frac{-1+\beta(z-\alpha)}{z-\alpha} = \frac{\beta z - (1+\alpha\beta)}{z-\alpha} = \begin{pmatrix} \beta & -(1+\alpha\beta) \\ 1 & -\alpha \end{pmatrix} z$$

Notice that $\det(T) = -\alpha\beta + (1+\alpha\beta) = 1$. And therefore $T \in \mathrm{PSL}(2,\mathbb{R})$.

Case 1: Let L be a straight line orthogonal to the real axis at the point $\alpha \in \mathbb{R}$. Then $z = \alpha + iy(t)$ and

$$T(z) = -(\alpha + iy(t) - \alpha)^{-1} + \beta = -\frac{1}{iy(t)} + \beta = \frac{i}{y(t)} + \beta$$

so for $\beta = 0$, T maps L to the imaginary axis.

Case 2: Let C be a semicircle orthogonal to the real axis with α an endpoint. Let $h \in \mathbb{R}$ be the center of the circle and r be the radius. Without loss of generality, let $h - r = \alpha$. Consider $z = r\cos t + h + i r \sin t$, $t \in (0,\pi)$, then

$$T(z) = \frac{-1}{r\cos t + h + ir\sin t - (h-r)} + \beta = \frac{-1}{(r + r\cos t) + ir\sin t} + \beta$$

$$= \frac{-r - r\cos t + ir\sin t}{(r+r\cos t)^2 + (r\sin t)^2} + \beta = \frac{-r - r\cos t + ir\sin t}{2r^2 + 2r^2\cos t} + \beta = \frac{1}{2r}\left[-1 + i\frac{\sin t}{1+\cos t}\right] + \beta$$

so for $\beta = \frac{1}{2r}$, T maps C to the imaginary axis.

The following theorem shows how geodesics appear in the upper-half plane. Since the distance between two points is measured along a unique geodesic, the results will help develop a formula for computing $\rho(z,w)$ where $z, w \in \mathfrak{H}$.

Theorem 6. *Geodesics in $\mathfrak{H}$ are semicircles and straight lines orthogonal to the real axis.*

Proof. Let $z_1, z_2 \in \mathfrak{H}$ such that $z_1 = ia$ and $z_2 = ib$, $b > a$. Let $\gamma : I \to \mathfrak{H}$ be a piecewise differentiable path joining z_1 and z_2, with $\gamma(t) = (x(t), y(t))$, then

$$h(\gamma) = \int_0^1 \frac{\sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}}{y(t)}\,dt \geq \int_0^1 \frac{|\frac{dy}{dt}|}{y(t)}\,dt \geq \int_0^1 \frac{\frac{dy}{dt}}{y(t)}\,dt = \int_a^b \frac{dy}{y} = \ln\frac{b}{a}$$

which, from Example 1, is the hyperbolic length of the segment of the y-axis joining ia and ib. This shows that the geodesic joining z_1 and z_2 is the imaginary axis joining them.

For arbitrary $z_1, z_2 \in \mathfrak{H}$, let L be the unique Euclidean circle or straight line orthogonal to the real axis that passes through these points. From Proposition 5, there exists a mapping in $\mathrm{PSL}(2,\mathbb{R})$ that sends L to the imaginary axis, which implies that L is exactly the geodesic joining z_1 and z_2. □

Theorem 7. *For $z, w \in \mathfrak{H}$*

$$\rho(z,w) = \ln\frac{|z-\bar{w}| + |z-w|}{|z-\bar{w}| - |z-w|}$$

Proof. Let L be the unique geodesic joining z and w. From Proposition 5, there exists a TϵPSL(2,$\mathbb{R}$)that maps L to the imaginary axis and by Theorem 4, $\rho(z, w)$ is invariant under T. Now all that is left to check is when $z = i$a, $w = i$b (a<b).

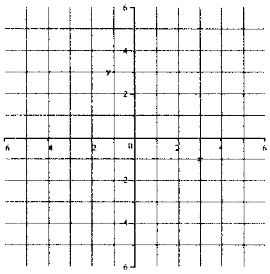
$$\ln\frac{|ia+ib|+|ia-ib|}{|ia+ib|-|ia-ib|} = \ln\frac{(a+b)+(b-a)}{(a+b)-(b-a)} = \ln\frac{2b}{2a} = \ln\frac{b}{a} = \rho(z,w)$$

□

A Fundamental Region in $\mathfrak{H}$

Put simply, a fundamental region is the smallest region which, based on symmetry, represents the whole object. This region is important because we are able to understand properties about the whole object by only considering the fundamental region.

First, let's begin with a simple, familiar example. Consider the Euclidean plane $\mathbb{R}^2$ and the transformations A and B where $A(x, y) = (x+1, y)$ and $B(x, y) = (x, y+1)$. A, B generate the lattice $\mathbb{Z}^2$ on the plane. The fundamental domain D is the set $[0, 1]^2$, which will cover the entire plane with no overlapping of interior points by infinite translations under $A, A^{-1}, B,$ and B^{-1}.



*The transformations of A and B over the plane.

A quotient space can be thought of as the surface that results from "gluing" together certain points. By gluing together parallel edges of D, you can see that the quotient space $\mathbb{R}^2/\mathbb{Z}^2$ is the torus.

Definition 8. *Let Γ be a subgroup of the group Isom($\mathfrak{H}$). We call a subset $F \subset \mathfrak{H}$ a* fundamental region *for Γ if it satisfies the following conditions:*

1. *F is a closed region in $\mathfrak{H}$ bounded by a finite number of geodesics.*
2. *The images T(F) for T$\epsilon\Gamma$ cover $\mathfrak{H}$.*
3. *For $T_1 \neq T_2$, the images $T_1(F)$ and $T_2(F)$ have no interior points in common.*

The proof for the following theorem is omitted, since it requires significant topological understanding. If the reader chooses to investigate the proof, it can be found in Katok's MASS REU lecture [3].

Theorem 9. *If p$\epsilon\mathfrak{H}$ is not fixed by any element of $\Gamma - \{\mathrm{Id}\}$, then $D_p(\Gamma)$ is a connected fundamental region for Γ, where*

$$D_p(\Gamma) = \{z \,\epsilon\, \mathfrak{H} \,|\, \rho(z, p) \leq \rho(T(z), p) \text{ for all } T\epsilon\Gamma\}$$

The region $D_p(\Gamma)$ is read as the *Dirichlet region for Γ centered at p*, and can be thought of as the set of points z which are closer than T(z) in the hyperbolic metric to p, for all T$\epsilon\Gamma$.

Theorem 10. *The group PSL(2,$\mathbb{Z}$)is generated by three elements, T(z) $= z+1$, $T^{-1} = z-1$ and S(z)$= -\frac{1}{z}$.*

Proof. First, notice that $T(z)$ corresponds to $T=\begin{pmatrix}1&1\\0&1\end{pmatrix}$, $T^{-1}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$, and $S=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ Let $g(z)=\frac{az+b}{cz+d}$ be a transformation in PSL(2,$\mathbb{Z}$). We will show that it can be represented as the composition of a finite number of transformations T, T^{-1}, and S (we don't need to consider S^{-1} since S and S^{-1} are the same translation). Since $ad-bc=1$, the integers a, c are relatively prime or one of them is equal to zero. If $a=0$, either $b=-1, c=1$, or vice versa. In the first case, $g(z)=\frac{-1}{z+d} \Rightarrow T^{-d}S\circ g=I_2$ and hence $g=S\circ T^d$. In the second case, $g(z)=\frac{1}{-z+d}\Rightarrow T^dS\circ g=I_2$, so $g=S\circ T^{-d}$. If $c=0$, then $g(z)=z+b$ or $g(z)=z-b$ and then clearly $g=T^b$ or $g=T^{-b}$, respectively.

Now assume that $a,\, c\neq 0$. The algorithm of factorization of the matrix corresponding to g is essentially the Euclidean algorithm for finding the greatest common divisor of $|a|$ and $|c|$, which in this case is equal to 1. Assume $c>0$; if $|a|>c$, then we can write $|a|=qc+r$, where q, r are positive integers and $r<c$. If $a>0$, then apply T^{-q} to g to obtain

$$T^{-q}\circ g(z)=\frac{rz+b'}{cz+d}$$

And then by applying S,

$$S\circ T^{-q}\circ g(z)=\frac{-cz-d}{rz+b'}$$

Similarly, if $a<0$, then

$$S\circ T^{q}\circ g(z)=\frac{-cz-d}{rz-b''}$$

In either case, we get $\frac{a_1z+b_1}{c_1z+d_1}$ with $|a_1|\geq|c_1|$ and $|a_1|<|a|$. After finitely many steps, we arrive at $\frac{a_nz+b_n}{c_nz+d_n}$ with $a_n=\pm 1$ and $c_n=0$, which is already solved above. If $|a|<|c|$, then first apply S to reduce to the case already considered.

□

To construct a fundamental region for PSL(2,$\mathbb{Z}$) on $\mathfrak{H}$, we will let $p=2i$. To show that p is not fixed for any element in Γ except for I_2, let $A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ such that $A(2i)=2i$. Then,

$$\frac{a(2i)+b}{c(2i)+d}=2i\Rightarrow 2ai+b=-4c+2di\Rightarrow b+4c+2(a-d)i=0$$

Therefore, $b+4c=0$ and $a=d$. So, $A=\begin{pmatrix}a&-4c\\c&a\end{pmatrix}$ and $a^2+4c^2=1$. Since $a, c\epsilon\mathbb{Z}$, the only solution is $a=1$ and $c=0$. This means that the identity matrix is the only element that fixes $2i$.

Now, to cover all of $\Gamma=$ PSL(2,$\mathbb{Z}$), we need to only consider the generators of the group and their inverses, namely T, T^{-1}, and S as defined in the previous theorem. To find the Dirichlet region, we need to find $\{z\epsilon\mathfrak{H}\,|\,\rho(z,p)\leq\rho(T(z),p) \text{ for all } T\epsilon\Gamma\}$. First consider $T(z)=z+1$. Then, we need to solve

$$\ln\frac{|z+2i|+|z-2i|}{|z+2i|-|z-2i|}\leq\ln\frac{|z+1+2i|+|z+1-2i|}{|z+1+2i|-|z+1-2i|}\Rightarrow|z+2i|\,|z+1-2i|=|z-2i|\,|z+1+2i|$$

Let $z=x+iy$. Then

$$(x^2+(y+2)^2)\,((x+1)^2+(y-2)^2)\leq(x^2+(y-2)^2)\,((x+1)^2+(y+2)^2)$$

$$(x^2+2x+1)(y^2+4y+4)+x^2(y^2-4y+4)\leq(x^2+2x+1)(y^2-4y+4)+x^2(y^2+4y+4)$$

$$8xy+4y\leq-8xy-4y\Rightarrow 2xy\leq-y\Rightarrow x\geq-\frac{1}{2}$$

For when $T(z)=z-1$, we can use similar algebra to find the set $x \leq \frac{1}{2}$. Now consider $S(z)=-\frac{1}{z}$.

$$\ln\frac{|z+2i|+|z-2i|}{|z+2i|-|z-2i|} \leq \ln\frac{|-\frac{1}{z}+2i|+|-\frac{1}{z}-2i|}{|-\frac{1}{z}+2i|-|-\frac{1}{z}-2i|} \Rightarrow |z+2i||-\frac{1}{z}-2i| = |z-2i||-\frac{1}{z}+2i|$$

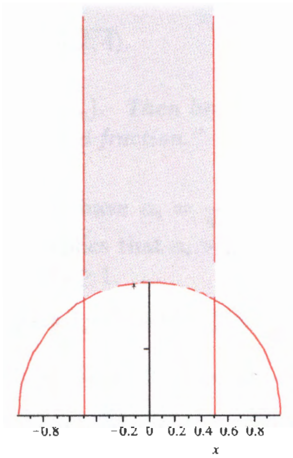
With $z=x+iy$, $-\frac{1}{z}=-\frac{1}{x+iy}\frac{x-iy}{x-iy}=-\frac{x-iy}{x^2+y^2}$. And so,

$$(x^2+(y+2)^2)\frac{x^2+(y+2x^2+2y^2)^2}{(x^2+y^2)^2} \leq (x^2+(y-2)^2)\frac{x^2+(2x^2+2y^2-y)^2}{(x^2+y^2)^2}$$

which reduces to

$$x^2+y^2 \geq 1$$

The intersection of these three sets is the shaded area below



Therefore, the fundamental region for $\mathfrak{H}$ under PSL(2,$\mathbb{Z}$) is

$$F=\{z\epsilon\mathfrak{H} \mid \text{Re}(z) \leq |\frac{1}{2}| \text{ and } |z| \geq 1\}$$

Consider the point $z=i$ which divides the circular boundary into two equal sections. The quotient space is constructed by "gluing" together these two sections, as well as gluing the two vertical boundaries together. This space becomes almost a sphere, with a cusp at infinity. In addition, notice that at the points $z=i$ and $z=\pm\frac{\sqrt{3}}{2}+\frac{1}{2}i$, the surface "buckles" to a cone.

Continued Fractions

Our discussion on continued fractions follows largely from Katok's MASS REU publication [3] as well as a publication by Caroline Series [7].

A "minus" continued fraction is expressed as

$$a_0-\cfrac{1}{a_1-\cfrac{1}{a_2-\cfrac{1}{a_3-\frac{1}{\cdots}}}}$$

where each a_i is a positive integer. To conserve space, we will write $\alpha=(a_0,a_1,a_2,a_3,\ldots)$

Let α be a given real number. Let $a_0 = [\alpha] + 1$, where $[\cdot]$ denotes the largest integer smaller than or equal to a number. Then, let $\alpha_1 = \frac{1}{a_0 - \alpha}$ and continue in the fashion such that

$$a_n = [\alpha_n] + 1, \alpha_{n+1} = \frac{1}{a_n - \alpha_n}$$

Example 11. Let $\alpha = 3 + \sqrt{2}$. Find the first five terms of the "minus continued fraction" of α. Since $\alpha \approx 4.41421$, then $a_0 = [3 + \sqrt{2}] + 1 = 5$, $\alpha_1 = \frac{1}{5 - (3 + \sqrt{2})} \approx 1.70711, a_1 = 2, \alpha_2 = \frac{1}{2 - \frac{1}{5 - (3 + \sqrt{2})}} \approx 3.41421$ Continuing in this fashion, we are able to fill out the following table:

n	α_n	a_n
0	4.41421	5
1	1.70711	2
2	3.41421	4
3	1.70711	2
4	3.41421	4

Table 1.

So α can be written as $(5, 2, 4, 2, 4, \ldots)$ or $(5, \overline{2, 4})$.

Theorem 12. *Let $r_n = (a_0, a_1, \ldots, a_{n-1}, a_n)$. Then $\lim_{n \to \infty} r_n = \alpha$, i.e., any real number α can be represented as an infinite "minus continued fraction."*

Proof. First, notice that for $i \geq 1$, we have $\alpha_i = \frac{1}{a_{i-1} - \alpha_{i-1}}$. If α_{i-1} is not an integer, then $0 < a_{i-1} - \alpha_{i-1} < 1$ and $\alpha_i > 1$, which implies that $a_i > 2$. If α_{i-1} is an integer, then $a_{i-1} - \alpha_{i-1} = 1$ and hence $a_i = 2$. Therefore $a_i \geq 2$ for all $i \geq 1$.

We now define two sequences of integers $\{p_n\}$ and $\{q_n\}, n \geq -2$, inductively:

$$p_{-2} = 0,\ p_{-1} = 1;\ p_i = a_i p_{i-1} - p_{i-2} \text{ for } i \geq 0$$

$$q_{-2} = -1, q_{-1} = 0;\ q_i = a_i q_{i-1} - q_{i-2} \text{ for } i \geq 0$$

We will prove that $r_n = \frac{p_n}{q_n}$. By induction, we will find

$$1 = q_0 < q_1 < q_2 < \ldots < q_n < \ldots$$

and so $\lim_{n \to \infty} q_n = \infty$. We have $q_1 = a_1 q_0 - q_{-1} = a_1 \geq 2 > q_0$ as the base of induction. Now assume that $1 = q_0 < q_1 < \ldots < q_{n-1}$. Then

$$q_n = a_n q_{n-1} - q_{n-2} > a_n q_{n-1} - q_{n-1} \geq 2q_{n-1} - q_{n-1} = q_{n-1}$$

which supports our claim. Now, suppose that

$$(a_0, a_1, a_2, \ldots, a_{n-1}, x) = a_0 - \cfrac{1}{a_1 - \cfrac{1}{a_2 - \cfrac{1}{\ddots - \frac{1}{x}}}}$$

Then for any $x \geq 1$,

$$(a_0, a_1, \ldots, a_{n-1}, x) = \frac{x p_{n-1} - p_{n-2}}{x q_{n-1} - q_{n-2}},$$

which follows by induction from the definition of $\{p_n\}$ and $\{q_n\}$.

Also by the definition of $\{p_n\}$ and $\{q_n\}$ we see that

$$p_{i-1} q_i - p_i q_{i-1} = p_{i-2} q_{i-1} - p_{i-1} q_{i-2} = \ldots = p_{-2} q_{-1} - p_{-1} q_{-2} = 1$$

Then with $x = a_n$, we obtain $r_n = \frac{a_n p_{n-1} - p_{n-2}}{a_n q_{n-1} - q_{n-2}} = \frac{p_n}{q_n}$.

Notice that $r_n - r_{n+1} = \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} = \frac{1}{q_n q_{n+1}} > 0$, and so $\{r_n\}$ is monotone-decreasing. Also, $\{r_n\}$ is bounded below by $a_0 - 1$, and hence has a limit, which is a real number. Let $\alpha = (a_0, a_1, \ldots, a_{n-1}, \alpha_n)$ where $\alpha_n = \frac{1}{a_{n-1} - \alpha_{n-1}}$. Then,

$$\frac{p_{n-1}}{q_{n-1}} - \alpha = \frac{\alpha_n p_{n-1} - p_{n-2}}{\alpha_n q_{n-1} - q_{n-2}} = \frac{1}{q_{n-1}\alpha_n q_{n-1} - q_{n-2}}, \leq \frac{1}{q_{n-1}} \to 0$$

Thus $\lim_{n\to\infty} r_n = \alpha$.

□

Theorem 13. *A real number α is rational if and only if there exists a positive integer n such that for all $k \geq n$, $\alpha_k = 2$.*

First, it is useful to construct a minus continued fraction expansion that is only a repeated 2. That is

$$x = 2 - \cfrac{1}{2 - \cfrac{1}{2 - \cfrac{1}{2 - \ldots}}}$$

Notice, then, that

$$2 - x = \frac{1}{x} \Rightarrow x^2 - 2x + 1 = (x-1)^2 = 0$$

Therefore, the minus continued fraction $(\bar{2}) = 1$, and any expansion with a tail of 2's can be written as a rational number.

Proof. First, since $(2, 2, \ldots) = 1$ then for $\alpha = n \epsilon \mathbb{Z}$, we have $n = (n+1, 2, 2, \ldots)$. Now assume that $\alpha \epsilon \mathbb{Q}$. Then $\alpha = c_0/d_0$, with $c_0, d_0 \epsilon \mathbb{Z}$ and the "tails" are also rational numbers written in the least terms,

$$\alpha_1 = (a_1, a_2, \ldots) = \frac{c_1}{d_1}$$

$$\alpha_2 = (a_2, a_3, \ldots) = \frac{c_2}{d_2}$$

$$\alpha_n = (a_n, a_{n+1}, \ldots) = \frac{c_n}{d_n}$$

Since $a_n \geq 2$ for $n \geq 1$, we have $c_n/d_n > 1$. We will show that $d_0 > d_1 > \ldots$, and therefore for some $N \epsilon \mathbb{Z}$ we have $d_N = 1$, which implies that $\alpha_N \epsilon \mathbb{Z}$ and $a_i = 2$ for $i > N$. First, notice that

$$\frac{c_0}{d_0} = a_0 - \frac{1}{\frac{c_1}{d_1}} = a_0 - \frac{d_1}{c_1}, \text{ so } \frac{c_0}{d_0} + \frac{d_1}{c_1} = a_0.$$

Similarly,

$$\frac{c_n}{d_n} + \frac{d_{n+1}}{c_{n+1}} = a_n \text{ or } c_n c_{n+1} + d_n d_{n+1} = a_n d_n c_{n+1}$$

Since c_{n+1} divides the other two terms of the above formula, we can assert that $c_{n+1} | d_n d_{n+1}$, but $(c_{n+1}, d_{n+1}) = 1$ and therefore $c_{n+1} | d_n$ and $c_{n+1} < d_n$. Using the fact that $c_{n+1} > d_{n+1}$, we know that $d_{n+1} < d_n$.

Conversely, as discussed above, if α has a tail of 2's, we have $\alpha = (a_0, a_1, \ldots, 1)$. Hence α is a rational number. □

Example 14. Let $\alpha = \frac{7}{5}$. Then $a_0 = 2$ and $\alpha_1 = \frac{1}{2 - \frac{7}{5}} = \frac{5}{3}$, $a_1 = 2$, $\alpha_2 = \frac{1}{2 - \frac{5}{3}} = 3$, and so

n	α_n	a_n
0	1.4	2
1	1.67	2
2	3	4
3	1	2
4	1	2

Table 2.

So in the minus continued fraction for α, $a_i = 2$ for $i \geq 3$ and $\alpha = \frac{7}{5} = (2, 2, 4, \bar{2})$.

Definition 15. *A real number is defined as a* quadratic irrationality *if it is a real root of the quadratic equation $ax^2 + bx + c$, with coefficients $a, b, c \in \mathbb{Z}$, $c \neq 0$ and the discriminant $D = b^2 - 4ac$ is positive and not a perfect square.*

Consider the minus continued fraction for $\alpha = 3 + \sqrt{2}$ from Example 11, which we found to be $\alpha = (5, \overline{2, 4})$. To find the quadratic polynomial, consider

$$x = 5 - \frac{1}{2 - \frac{1}{4 - \frac{1}{\cdots}}} \Rightarrow 5 - x = \frac{1}{2 - \frac{1}{4 - (5 - x)}} = \frac{1}{2 - \frac{1}{x - 1}} = \frac{x - 1}{2x - 3}$$

$$(2x - 3)(5 - x) = x - 1 \Rightarrow x^2 - 6x + 7 = 0$$

Since $(3 + \sqrt{2})^2 - 6(3 + \sqrt{2}) + 7 = 0$, then we know our observation that α infinitely repeats is true and, also, α is a quadratic irrationality. Explicitly, $\alpha = 3 + \sqrt{2} = (5, \overline{2, 4})$

Theorem 16. *A real number α is a quadratic irrationality if and only if its minus continued fraction expansion is eventually periodic with the periodic part being anything but a repeated 2.*

Proof. Let α be a quadratic irrationality; we will need to show that the minus continued fraction expansion of α is eventually periodic, with the periodic part not a repeated 2. Notice that α can be written such that

$$\alpha = a_0 = \frac{m_0 + \sqrt{D}}{l_0}$$

where $m_0, l_0, D \in \mathbb{Z}, l_0 \neq 0$ and $D > 0$ is not a square. Also, let $\alpha = (a_0, a_1, a_2, \ldots, a_n, \ldots)$. Let α_n be the "tail" of the minus continued fraction for α, where $\alpha_n = (a_n, a_{n+1}, \ldots)$ and then $\alpha = (a_0, a_1, \ldots, a_{n-1}, \alpha_n)$.

We will prove by induction that α_n is a quadratic irrationality of the form $\frac{m_n + \sqrt{D}}{l_n}$ for $n = 0, 1, 2, \ldots$. The base for induction is already shown, so let $\alpha_n = \frac{m_n + \sqrt{D}}{l_n}$. Then

$$\alpha_n = a_n - \frac{1}{\alpha_{n+1}} \Rightarrow \frac{1}{\alpha_{n+1}} = a_n - \frac{m_n + \sqrt{D}}{l_n} = \frac{a_n l_n - m_n - \sqrt{D}}{l_n}$$

$$\alpha_{n+1} = \frac{l_n}{a_n l_n - m_n - \sqrt{D}} = \frac{a_n l_n - m_n + \sqrt{D}}{\left(\frac{(a_n l_n - m_n)^2 - D}{l_n}\right)}$$

and so $\alpha_{n+1} = \frac{m_{n+1} + \sqrt{D}}{l_{n+1}}$, which satisfies the recurrent equations:

$$m_{n+1} = a_n l_n - m_n, l_{n+1} = \frac{m_{n+1}^2 - D}{l_n}$$

Let $\overline{\alpha_n} = \frac{m_n - \sqrt{D}}{l_n}$, where $\overline{\alpha_n}$ is called "the conjugate of α_n." From Theorem 12,

$$\alpha_0 = (a_0, a_1, \ldots, a_{n-1}, \alpha_n) = \frac{\alpha_n p_{n-1} - p_{n-2}}{\alpha_n q_{n-1} - q_{n-2}}.$$

Taking the conjugate of both sides, we get

$$\overline{\alpha_0} = (a_0, a_1, \ldots, a_{n-1}, \overline{\alpha_n}) = \frac{\overline{\alpha_n} p_{n-1} - p_{n-2}}{\overline{\alpha_n} q_{n-1} - q_{n-2}}.$$

Solving for $\overline{\alpha_n}$, we obtain

$$\overline{\alpha_n} = \frac{\overline{\alpha_0} q_{n-2} - p_{n-2}}{\overline{\alpha_0} q_{n-1} - p_{n-1}} = \frac{q_{n-2}}{q_{n-1}} \left(\frac{\overline{\alpha_0} - \frac{p_{n-2}}{q_{n-2}}}{\overline{\alpha_0} - \frac{p_{n-1}}{q_{n-1}}} \right)$$

Since both $\frac{p_{n-1}}{q_{n-1}}$ and $\frac{p_{n-2}}{q_{n-2}}$ tend to $\alpha_0 \neq \overline{\alpha_0}$, then $\frac{\overline{\alpha_0} - \frac{p_{n-2}}{q_{n-2}}}{\overline{\alpha_0} - \frac{p_{n-1}}{q_{n-1}}}$ tends to 1. Since the sequence $\{q_n\}$ is increasing, we see that $\frac{q_{n-2}}{q_{n-1}} < 1$. From Theorem 13, the minus continued fraction of α has infinitely many terms greater than 2. More precisely, there exists a subsequence $\{n_k\}$ such that $a_{n_k - 1} \geq 3$, and then $a_{n_k - 1} - 1 \geq 2$ and $a_{n_k} \geq 2$, and we obtain

$$\begin{aligned} q_{n_k} &= a_{n_k} q_{n_k - 1} - q_{n_k - 2} = a_{n_k}(a_{n_k - 1} q_{n_k - 2} - q_{n_k - 3}) - q_{n_k - 2} \\ &= (a_{n_k} a_{n_k - 1} - 1) q_{n_k - 2} - a_{n_k} q_{n_k - 3} \geq (a_{n_k}(a_{n_k - 1} - 1) - 1) q_{n_k - 2} \\ &\geq 3 q_{n_k - 2} \end{aligned}$$

Thus

$$\frac{q_{n_k - 2}}{q_{n_k}} = \frac{q_{n_k - 2}}{q_{n_k - 1}} \cdot \frac{q_{n_k - 1}}{q_{n_k}} \leq \frac{1}{3},$$

which implies that one of the above fractions is $\leq \frac{1}{\sqrt{3}} < 1$. Now we conclude that there exists an $N > 1$ such that $0 < \overline{\alpha_N} < 1$. Since $a_N \geq 2$, we have $\alpha_N > 1$.

Now by induction, we can show that for any $n \geq N$, $0 < \overline{\alpha_n} < 1$ and $\alpha_n > 1$. The basis was just proved. Let $0 < \overline{\alpha_k} < 1$ and $\alpha_k > 1$. Then

$$\alpha_{k+1} = \frac{1}{a_k - \alpha_k} > 1, \qquad 0 < \overline{\alpha_{k+1}} = \frac{1}{a_k - \overline{\alpha_k}} < 1,$$

the last inequality holds since $a_k \geq 2$ and $\overline{\alpha_k} < 1$. Thus we can conclude that there exists an $N > 1$ such that for all $n > N$, we have $0 < \overline{\alpha_n} < 1$. Since $a_n \geq 2$, we have $\alpha_n > 1$.

It follows that for $n \geq N$

$$0 < \frac{m_n - \sqrt{D}}{l_n} < 1, \qquad \frac{m_n + \sqrt{D}}{l_n} > 1.$$

Since $\alpha_n - \overline{\alpha_n} = \frac{2\sqrt{D}}{l_n} > 0$, we conclude that $l_n > 0$, and therefore the previous statement implies that $|m_n - l_n| < \sqrt{D}$, and hence we can take only finitely many values for a given D. We have $D - (m_n - l_n)^2 > 0$, and this expression also can take only finitely many values. From the recurrence equations that solve α_{n+1} from α_n, we can write

$$D - (m_n - l_n)^2 = D - m_n^2 - l_n^2 + 2m_n l_n = -l_n l_{n-1} - l_n^2 + 2m_n l_n = l_n(-l_{n-1} - l_n + 2m_n)$$

Thus, $l_n | (D - (m_n - l_n)^2)$ and l_n and m_n can each take only finitely many values. Therefore, for some $j \neq k$, $\alpha_j = \alpha_k$. In other words, the "tails" of α coincide or, more precisely, $a_j = [\alpha_j] + 1 = [\alpha_k] + 1 = a_k$, and so on. This shows that the minus continued fraction expansion of α is eventually periodic. Then, the periodic part cannot be a repeated 2 since in this case α would be rational by Theorem 13.

Now assume that α have an eventually periodic continued fraction expansion. Then, it is a root of a quadratic equation with integer coefficients. Since the periodic part of its minus continued fraction expansion is anything but a repeated 2, α is irrational from Theorem 13, so the second root is irrational as well and α is a quadratic irrationality. □

Closed Geodesics

The information in this section was largely deduced from the works of Katok and Ilie Ugarcovici, [5] and [6].

A transformation $A \epsilon \mathrm{PSL}(2, \mathbb{Z})$ is said to be *hyperbolic* if A moves along a circular path. A hyperbolic transformation has two fixed points in $\mathbb{R} \cup \{\infty\}$, one attracting and one repelling, which we will call u and w, respectively. The geodesic in $\mathfrak{H}$ connecting u and w is called the *axis* of A and is denoted $C(A)$. Notice that if $z \epsilon C(A)$ then $A^n(z) \epsilon C(A)$ and $A^n(z) \to w$ as $n \to \infty$.

For this section, we will consider the modular surface $M = \mathrm{PSL}(2, \mathbb{Z}) \backslash \mathfrak{H}$, with respect to the fundamental domain $F = \{z \epsilon \mathfrak{H} \mid \mathrm{Re}(z) \leq |\frac{1}{2}| \text{ and } |z| \geq 1\}$.

Theorem 17. *Closed geodesics on the modular surface M are in a one-to-one correspondence with classes of hyperbolic elements in* $\mathrm{PSL}(2, \mathbb{Z})$.

This theorem tells us that given a closed geodesic on F, we can find a finite string of generators of $\mathrm{PSL}(2, \mathbb{Z})$ such that after their successive application, we return to the original geodesic in $\mathfrak{H}$. The elements in PSL(2,$\mathbb{Z}$) that correspond to closed geodesics are those that fix the points u, $w \epsilon \mathbb{R} \cup \{\infty\}$.

Example 18. Let $A = \begin{pmatrix} 15 & -8 \\ 2 & -1 \end{pmatrix} \epsilon \mathrm{PSL}(2, \mathbb{Z})$. The geodesic that corresponds to A is exactly the one with endpoints u and w such that $A(u) = u$ and $A(w) = w$. Let $z = x + y\,i$. Then $A(z) = \frac{15z - 8}{2z - 1} = \frac{15(x + y\,i) - 8}{2(x + y\,i) - 1} = x + y\,i$ exactly when $x = 4 \pm 2\sqrt{3}$, $y = 0$.

Therefore, the geodesic that corresponds to $A = \begin{pmatrix} 15 & -8 \\ 2 & -1 \end{pmatrix}$ has endpoints $u = 4 - 2\sqrt{3}$ and $w = 4 + 2\sqrt{3}$. Further, this is the semicircle with a diameter on the real line, centered at $x = 4$ and radius $2\sqrt{3}$.

Definition 19. *An oriented geodesic in $\mathfrak{H}$ with a repelling point u and an attracting point w, for u,w irrationals, is called* reduced *if* $0 < u < 1$ and $w > 1$.

Reduction Algorithm: Let γ be an arbitrary geodesic on $\mathfrak{H}$ with end points u and w. Consider the following sequence of real pairs $\{(u_k, w_k)\}$, where $(u_0, w_0) = (u.w)$ and

$$u_{k+1} = ST^{-a_k} \ldots ST^{-a_1} ST^{-a_0} u, \qquad w_{k+1} = ST^{-a_k} \ldots ST^{-a_1} ST^{-a_0} w$$

where $(a_0, a_1, \ldots, a_k, \ldots)$ is the minus continued fraction expansion for w.

The *arithmetic code* of a reduced geodesic γ with endpoints (u, w) is $(\gamma) = (\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots)$, which is created by juxtaposing the minus continued fraction expansions of $1/u = (a_{-1}, a_{-2}, \ldots)$ and $w = (a_0, a_1, \ldots)$.

Theorem 20. *The above algorithm produces in finitely many steps a reduced geodesic PSL(2,$\mathbb{Z}$)-equivalent to γ.*

Proof. Let γ be an arbitrary geodesic on $\mathfrak{H}$ with irrational end points u and w where $w = (a_0, a_1, a_2, \ldots)$. Using the reduction algorithm, let $u_0 = u$ and $w_0 = w$ and

$$u_{k+1} = ST^{-a_k} \ldots ST^{-a_1} ST^{-a_0} u, \quad w_{k+1} = ST^{-a_k} \ldots ST^{-a_1} ST^{-a_0} w$$

Since w is irrational, $w_{k+1} = (a_{k+1}, a_{k+2}, \ldots) > 1$. From Theorem 12,

$$u = T^{n_0} S T^{n_1} S \ldots T^{n_k} S(u_{k+1}) = \frac{p_k u_{k+1} - p_{k-1}}{q_k u_{k+1} - q_{k-1}}, \quad w = T^{n_0} S T^{n_1} S \ldots T^{n_k} S(w_{k+1}) = \frac{p_k w_{k+1} - p_{k-1}}{q_k w_{k+1} - q_{k-1}}$$

and hence,

$$u_{k+1} = \frac{q_{k-1} u - p_{k-1}}{q_k u - p_k} = \frac{q_{k-1}}{q_k} + \frac{1}{q_k^2 (\frac{p_k}{q_k} - u)} = \frac{q_{k-1}}{q_k} + \varepsilon_k$$

where $\varepsilon_k \to 0$. As shown in Theorem 12, $\frac{p_k}{q_k} \to w$. From this, we can assume that $\left|\frac{p_k}{q_k} - u\right| > \frac{1}{2}|w - u|$ for a large enough k and therefore

$$|\varepsilon_k| = \frac{1}{q_k^2 \left|\frac{p_k}{q_k} - u\right|} < \frac{2}{q_k^2 \, |w - u|}$$

Since $1 < q_0 < q_1 < \ldots$, then for a large enough k, $|\varepsilon_k| < \frac{1}{q_k}$ and

$$0 < \frac{q_{k-1}}{q_k} - \frac{1}{q_k} < u_{k+1} = \frac{q_{k-1}}{q_k} + \varepsilon_k < \frac{q_{k-1}}{q_k} + \frac{1}{q_k} \leq 1.$$

Therefore, there exists a positive integer l such that $0 < u_{l+1} < 1$. The geodesic with points u_{l+1} and w_{l+1} is reduced and PSL(2,$\mathbb{Z}$) equivalent to γ. □

It is important to note that any further application of the reduction algorithm to a reduced geodesic yields other reduced geodesics whose codes are left shifts of the code of the initial reduced geodesic.

Example 21. Let γ be a geodesic in $\mathfrak{H}$ from $u = -5 + 3\sqrt{7}$ and $w = 3 + \sqrt{2}$. We've already shown that $w = (5, \overline{2, 4})$. Using the given algorithm, we find

$$u_1 = ST^{-5}u = \frac{10 + 3\sqrt{2}}{82}, \qquad w_1 = ST^{-5}w = \frac{2 + \sqrt{2}}{2}$$

After just one step, we have $0 < u_1 < 1$ and $w_1 > 1$. Therefore, the geodesic with endpoints at $(\frac{10 + 3\sqrt{2}}{82}, \frac{2 + \sqrt{2}}{2})$ is reduced and PSL(2, $\mathbb{Z}$)-equivalent to γ. Since $1/u_1 = 10 - 3\sqrt{2} = (6, 5, 2, \ldots)$ and $w_1 = (2, 4, 2, \ldots)$, the arithmetic code of $\gamma = (\ldots, 2, 5, 6, 2, 4, 2 \ldots)$.

Notice that the definition of a *reduced geodesic* requires irrational endpoints. This is because the orbit of infinity $= 1/0$ under PSL(2,$\mathbb{Z}$) is the rational numbers. To clarify, consider $T \epsilon$PSL(2, $\mathbb{Z}$) such that $T = (\begin{smallmatrix} a & b \\ c & d \end{smallmatrix})$. Then $T(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}) = \frac{a}{c}$. So for any geodesic with endpoints rational, the oriented geodesic on M goes to the cusp. Notice that this implies that if an arithmetic code as a tail of 2's in either direction, the corresponding geodesic goes to the cusp. Therefore, for a geodesic to be closed, it must have irrational endpoints.

Let $T_x M$ be the unit tangent space of M at the point x (i.e. the vector space that contains all unit tangent vectors at x) and let SM be the unit tangent bundle of M, or the union of each $T_x M$ for all $x \epsilon M$. A cross section for the geodesic flow is a subset of the unit tangent bundle SM in which each geodesic visits infinitely. Our construction of the cross section for the geodesic flow on M, denoted C_G, will be so that successive returns of a geodesic γ to C_G correspond to left-shifts in the arithmetic code of γ.

Construction of the Cross-Section C_G: Let P be the set of all tangent vectors with base points on the circular boundary of F and pointing inward, such that the corresponding geodesic on $\mathfrak{H}$ is reduced, and let Q be the set of all tangent vectors with base points on the right vertical boundary of F and pointing inward, such that if γ is the corresponding geodesic, then $TS(\gamma)$ is reduced. If $C_G = P \cup Q$, then $C_G \subset SM$ and is a cross-section of SM.

Theorem 22. *C_G is a cross-section of the geodesic flow on M.*

Proof. Let γ be an oriented geodesic on M that is presented as a bi-infinite sequence of PSL(2,$\mathbb{Z}$)-equivalent geodesic segments on $F \subset \mathfrak{H}$. Any segment extended to a geodesic on $\mathfrak{H}$ can be reduced according to Theorem 20. Let $\pi: S\mathfrak{H} \rightarrow SM$ be the projection of the unit tangent bundles. Then there exists a reduced geodesic γ' on $\mathfrak{H}$ such that $\pi(\gamma') = \gamma$. Since γ' must intersect the unit semicircle $|z| = 1$, then either $\gamma' \cap F$ or $ST^{-1}(\gamma') \cap F$ is one of the segments of γ on F. In either case, γ intersects C_G at least once. Denote this intersection by $\mathbf{x}_0 \epsilon C_G$; we will follow the geodesic from this starting point.

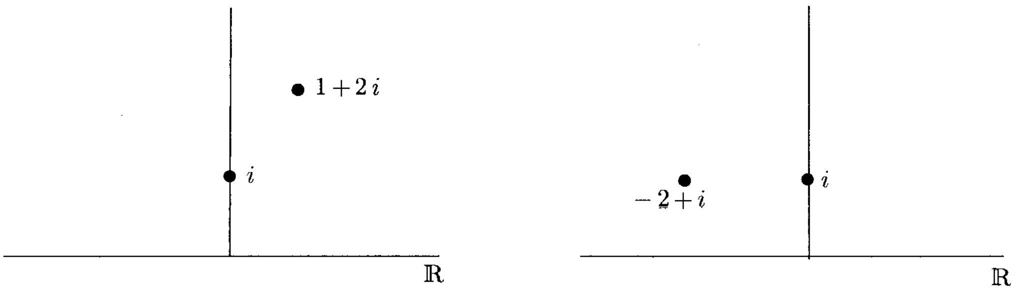
If $\mathbf{x}_o \epsilon P$ then the corresponding geodesic γ' on $\mathfrak{H}$ is reduced. To show that γ' intersects C_G again, notice that γ' intersects with the left side of the semi-circle $|z - \lceil w \rceil| = 1$ and $\lceil w \rceil = a_0$. Let $\mathbf{x}_1'$ be the unit tangent vector at this intersection. Then $ST^{-a_0}(\mathbf{x}_1')$ moves $\mathbf{x}_1'$ to $|z| = 1$ and $\mathbf{x}_1 = \pi(\mathbf{x}_1') \epsilon C_G$ and the first intersection of γ with C_G after $\mathbf{x}_0$ is at $\mathbf{x}_1$.

If $\mathbf{x}_o \epsilon Q$, then γ' on $\mathfrak{H}$ corresponding to $TS(\mathbf{x}_o)$ is reduced and the same argument follows. □

Every oriented geodesic γ on M can be represented as a bi-infinite sequence of segments σ_i between successive returns to C_G. To each segment σ_i, we associate the corresponding reduced geodesic γ_i on $\mathfrak{H}$ and thus obtain a sequence of reduced geodesics $\{\gamma_i\}_{i=-\infty}^{\infty}$ representing γ. If the arithmetic code $(\gamma_i) = (\ldots, a_{-2}, a_{-1}, a_0, a_1, \ldots)$, then $\gamma_{i+1} = ST^{-a_0}(\gamma_i)$, and the coding sequence is shifted one symbol to the left. Thus all reduced geodesics γ_i in the sequence produce the same, up to a shift, bi-infinite coding sequence (or arithmetic code). The left shift of the sequence corresponds to the return of the geodesic to the cross-section C_G.

Fundamental Region of $\mathfrak{H} \times \mathfrak{H}$

To take this problem to the next level, we will attempt the construction of a fundamental region on two copies of the upper-half plane $\mathfrak{H} \times \mathfrak{H}$. If $z \epsilon \mathfrak{H} \times \mathfrak{H}$, then $z = (z_1, z_2)$, where z_1 and z_2 are in separate copies.



Example of $z \epsilon \mathfrak{H} \times \mathfrak{H}$ where $z = (1 + 2i, -2 + i)$

Let $\Gamma = \mathrm{PSL}(2, \mathbb{Z}[\sqrt{2}])$ be the *Mobius transformations* $z \rightarrow \frac{az+b}{cz+d}$, where $a, b, c, d \epsilon \mathbb{Z}[\sqrt{2}]$ and $ad - bc = 1$. The actions of Γ cover $\mathfrak{H} \times \mathfrak{H}$ under two embeddings. Let $A \epsilon \mathrm{PSL}(2, \mathbb{Z}[\sqrt{2}])$ and $(z_1, z_2) \epsilon \mathfrak{H} \times \mathfrak{H}$, then $A(z_1, z_2) = (A(z_1), \bar{A}(z_2))$. For example, let $A = \begin{pmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{pmatrix}$. Then $A(z_1, z_2) = (z_1 + \sqrt{2}, z_2 - \sqrt{2})$.

Recall that the Dirichlet region aound a point p is also the fundamental region for the space and is defined as

$$D_p(\Gamma) = \{z \epsilon \mathfrak{H} \,|\, \rho(z,p) \leq \rho(T(z),p) \text{ for all } T \epsilon \Gamma\}$$

where p is not fixed for any $T\epsilon\Gamma$. Since we are now in two copies of the upper-half plane, by Pythagorean's Theorem our distance formula becomes

$$\rho(z,w) = \sqrt{\left(\ln\frac{|z_1 - \overline{w_1}| + |z_1 - w_1|}{|z_1 - \overline{w_1}| - |z_1 - w_1|}\right)^2 + \left(\ln\frac{|z_2 - \overline{w_2}| + |z_2 - w_2|}{|z_2 - \overline{w_2}| - |z_2 - w_2|}\right)^2}$$

where $z = (z_1, z_2)$ and $w = (w_1, w_2)$.

Let $p = (2i, 2i)$, as it can be shown that p is fixed only by the identity transformation of PSL(2, $\mathbb{Z}[\sqrt{2}]$). We need to solve

$$\left(\ln\frac{|z_1+2i|+|z_1-2i|}{|z_1+2i|-|z_1-2i|}\right)^2 + \left(\ln\frac{|z_2+2i|+|z_2-2i|}{|z_2+2i|-|z_2-2i|}\right)^2 \leq \ln\left(\frac{|A(z_1)+2i|+|A(z_1)-2i|}{|A(z_1)+2i|-|A(z_1)-2i|}\right)^2 + \ln\left(\frac{|\bar{A}(z_2)+2i|+|\bar{A}(z_2)-2i|}{|\bar{A}(z_2)+2i|-|\bar{A}(z_2)-2i|}\right)^2$$

which reduces to

$$\left(\frac{|z_1+2i|+|z_1-2i|}{|z_1+2i|-|z_1-2i|}\right)\left(\frac{|z_2+2i|+|z_2-2i|}{|z_2+2i|-|z_2-2i|}\right) \leq \left(\frac{|A(z_1)+2i|+|A(z_1)-2i|}{|A(z_1)+2i|-|A(z_1)-2i|}\right)\left(\frac{|\bar{A}(z_2)+2i|+|\bar{A}(z_2)-2i|}{|\bar{A}(z_2)+2i|-|\bar{A}(z_2)-2i|}\right)$$

for each $A\epsilon\Gamma$.

The proper construction of the Dirichlet region requires having at least the generators of Γ, which are unknown. Instead, we look at the region constructed by some carefully chosen elements and hopefully reduce those elements to a set that contains the generators.

Constructing images in four-dimensions is not easy, especially with such a complicated equation. I attempted reducing this equation to something more manageable, but with little gain. Therefore, to help picture the region, I used the graphing software MAPLE. MAPLE graphs in Euclidean two- and three-space so I was required to reduce the dimensions of our space. To do this, I fixed the value of one of the parameters, plotted the resulting surface and found which side of the surface had the desired points that satisfied the inequality. Changing the value of the fixed parameter revealed information about the four-dimensional space.

Example 23. I have included this example to illustrate my method. Consider the torus constructed by the equation

$$(x^2 + y^2 + z^2 + 3)^2 = 16\,(x^2 + z^2).$$

A fixed value of z gives us a curve in the plane that corresponds to that z-value.

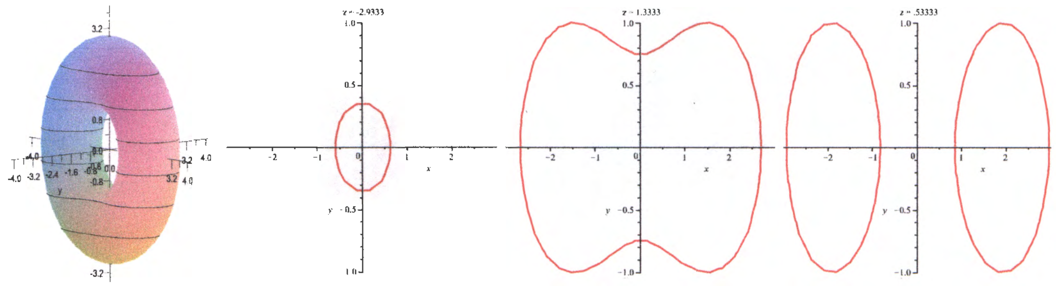
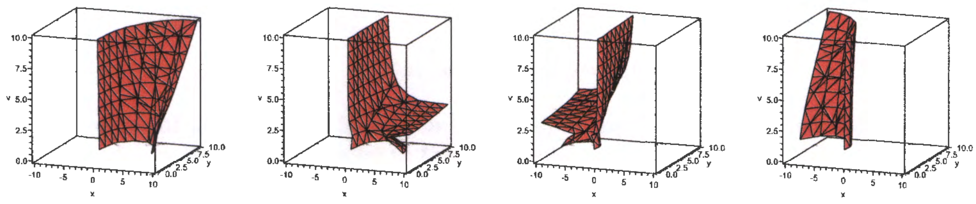


*Image 1 is torus given by the equation above. The other three images are the curves that results from choosing $z \approx 3, 4/3,$ and $1/2$.

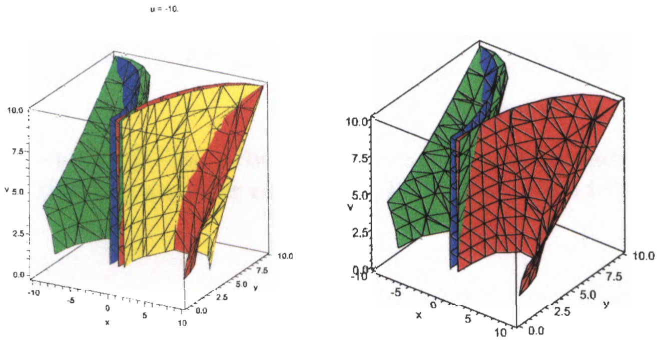
By considering curves together in 3-space, we can understand the surface that the curves represent. We will use a similar process to understand a four-dimensional region. First, we will fix one of the four parameters and then study the region as the value changes. For a translation $A \epsilon \mathrm{PSL}(2, \mathbb{Z}[\sqrt{2}])$, the boundary surface will be denoted as Π_A and the region as Σ_A. I should note that most of the information below is observation; an important area for future research will be defining the surfaces and curves that bound the regions.

To begin, we will first consider $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, where $T(z) = (z_1 + 1, z_2 + 1)$. Let $z_1 = x + yi$ and $z_2 = u + vi$. The following surfaces are Π_T for $u = -10, -\frac{5}{6}, 0, 10$.

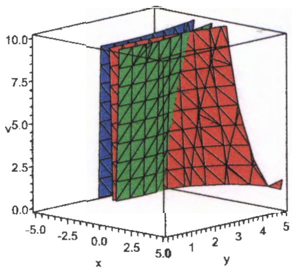


Further investigation revealed that $x = -\frac{1}{2}$ serves as a plane of symmetry, and Π_T converges to this plane as $u \to -\frac{1}{2}$. It is important to note that when $u < -\frac{1}{2}$, Σ_T is the back side of the space when considering the orientation as pictured above. For $u = -\frac{1}{2}$, Π_T is the plane $x = -\frac{1}{2}$ and Σ_T is every point such that $x \geq -\frac{1}{2}$, and for $u > -\frac{1}{2}$, Σ_T is in front of or above the surface.

Let $\bar{\Sigma} = \Sigma_T \cap \Sigma_{T^{-1}} \cap \Sigma_R \cap \Sigma_{R^{-1}}$, where $R = \left(\begin{smallmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{smallmatrix}\right)$ and $R^{-1} = \left(\begin{smallmatrix} 1 & -\sqrt{2} \\ 0 & 1 \end{smallmatrix}\right)$. In the below pictures, the red surface is Π_T, the yellow $\Pi_{T^{-1}}$, the blue Π_R and the green $\Pi_{R^{-1}}$. $\bar{\Sigma}$ is the intersection of the points between Π_T and $\Pi_{T^{-1}}$ and between Π_R and $\Pi_{R^{-1}}$.



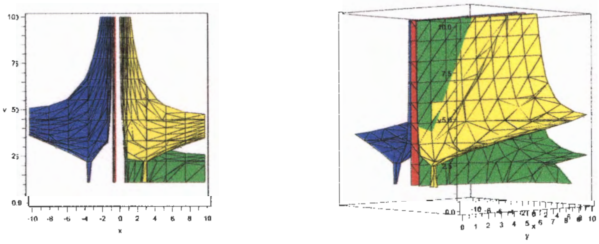
*The four surfaces for $u = -10$ are represented in the image on the left. A better view of $\bar{\Sigma}$ comes from removing $\Pi_{T^{-1}}$, which is pictured on the right. In this image, $\bar{\Sigma}$ is bounded entirely by Π_T (red) and $\Pi_{R^{-1}}$ (green).



*Closer look at $\bar{\Sigma}$ for $u = -10$

For larger values of u, $\bar{\Sigma}$ is a similar region but shifts in the positive direction of the v-axis. The intersection of $\bar{\Sigma}$ with a plane parallel to the x, y-plane appears parabolic. As $u \to 0$, the curve of intersection moves in the direction of becoming parallel to the x, y-plane and eventually the surfaces no longer intersect. For $u < -\frac{\sqrt{2}}{2}$, the intersection of $\Pi_{R^{-1}}$ and Π_T defines the region and when $u > \frac{\sqrt{2}}{2}$, the intersection of Π_R and $\Pi_{T^{-1}}$ defines the region. The region is symmetrical around $u = 0$.

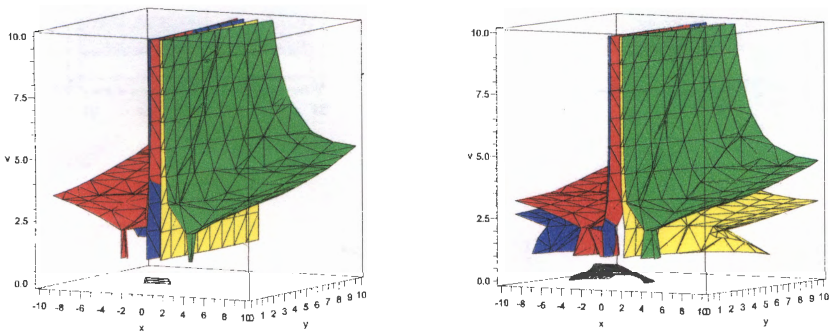
For the smaller values of u, $\bar{\Sigma}$ begins to change dramatically. The surfaces no longer intersect and $\bar{\Sigma}$ is most points in $\mathfrak{H} \times \mathfrak{H}$. This happens approximately when $|u| < \frac{1}{2}$.



*$\bar{\Sigma}$ is the region inbetween $\Pi_{R^{-1}}$ and Π_T for $u = -\frac{1}{2}$. Notice that $\Pi_{T^{-1}}$ (yellow) also bounds part of the region.

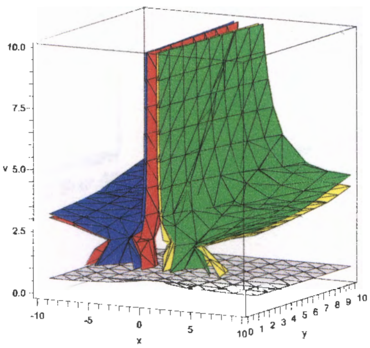
There are some properties of these surfaces that are worth discussing. Just as Π_T converges to the plane $x=-\frac{1}{2}$ as $u\to-\frac{1}{2}$, $\Pi_{T^{-1}}\to x=\frac{1}{2}$ for $u\to\frac{1}{2}$, $\Pi_R\to x=-\frac{\sqrt{2}}{2}$ as $u\to\frac{\sqrt{2}}{2}$, and $\Pi_{R^{-1}}\to x=\frac{\sqrt{2}}{2}$ as $u\to-\frac{\sqrt{2}}{2}$.

Now consider Π_S where $S=\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. For larger values of u, $\bar{\Sigma}$ is unaffected by Σ_S. However, for smaller values of u, $\Sigma_S\cap\bar{\Sigma}$ is a smaller region than $\bar{\Sigma}$. In the picture below, Π_S is gray.



*Surfaces for $u=\frac{1}{2}$ and $u=\frac{1}{4}$

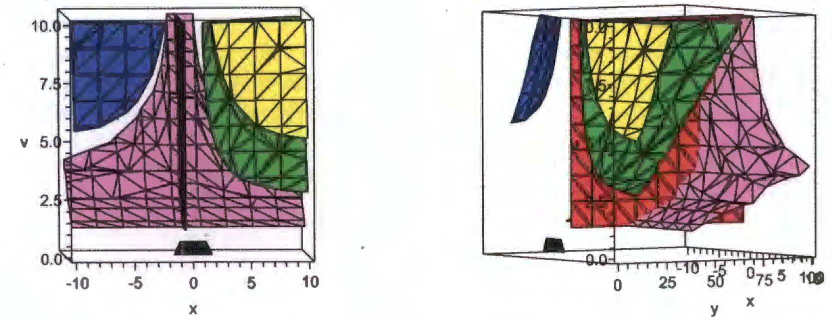
As $u\to 0$, Π_S flattens and is almost a level surface. $\bar{\Sigma}\cap\Sigma_S$ is the region above Π_S.



*Surfaces for $u=0$.

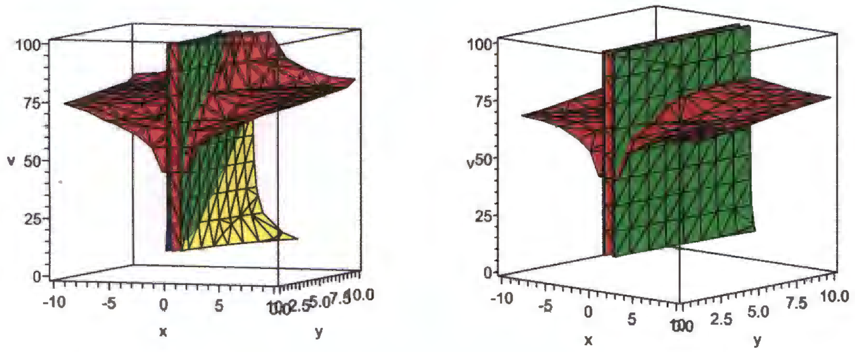
Let $Q=\left(\begin{smallmatrix}3-2\sqrt{2} & 0\\ 0 & 3+2\sqrt{2}\end{smallmatrix}\right)$. In the image below, Π_Q is light purple. Σ_Q is the region with smaller y-values. $\Sigma_Q\cap\bar{\Sigma}=\bar{\Sigma}$ except for approximately $|u|<\frac{1}{2}$, which is when $\bar{\Sigma}$ gets large very quickly.

In other words, Σ_Q bounds the region from above along the y-axis.



*Surfaces for $u=-\frac{1}{2}$.

The last translation in $\mathrm{PSL}(2,\mathbb{Z}[\sqrt{2}])$ we will consider is $Q^{-1}=\begin{pmatrix}3+2\sqrt{2} & 0\\ 0 & 3-2\sqrt{2}\end{pmatrix}$. $\Pi_{Q^{-1}}$ intersects with the region $\bar{\Sigma}$ for all u-values. For larger u-values, the cross section of $\Sigma_{Q^{-1}}\cap\bar{\Sigma}$ with a plane parallel to the x,y-plane is closer to a horseshoe than the filled parabola of $\bar{\Sigma}$. $\Pi_{Q^{-1}}$ is orange in the images below.



*The left image is the surfaces for $u=-10$ and the right is for $u=0$.

When $|u|<\frac{1}{2}$, $\Pi_{Q^{-1}}$ bounds $\bar{\Sigma}$ above along the v-axis.

In Conclusion

Examining $\Sigma=\Sigma_S\cap\Sigma_Q\cap\Sigma_{Q^{-1}}\cap\bar{\Sigma}$ for $u\epsilon(-\infty,\infty)$ provides an idea of the Dirichlet region of $\mathrm{PSL}(2,\mathbb{Z}[\sqrt{2}])$ on $\mathfrak{H}\times\mathfrak{H}$ centered at $z=(2i,2i)$ and the fundamental region of our group. The region is understandable for $|u|>1$, as the shape of the region does not change but its location in space does, i.e., it shifts in the positive direction of the v-axis. For $|u|<1$, Σ gets more interesting and less predictable. In these cross sections, Σ is bounded from above along the y-axis by Π_Q, below along the v-axis by Π_S and above along the v-axis by $\Pi_{Q^{-1}}$. The shape of the region changes to be almost a rectangular prism around $u = 0$.

Drawing a region in four-dimensions is not possible, but defining the surfaces and curves that bound the space is necessary to completely understand the desired region. Using this information, we would be able to consider the modular surface and the behavior of geodesics.

The slow processing time and high memory use of the MAPLE software reduced the amount of images I was able to create. Also, the program did not display parts of the surfaces with small y and v values, even as I zoomed in on the region. Therefore, I was only able to speculate the behavior of the region as $y \to 0$ and as $v \to 0$.

Bibliography

[1] Ford, Lester R. On a class of continued fractions. Edinburgh Mathematical Society, 1916. 38-45.

[2] Hilbert, David. *The Theory of Algebraic Number Fields.* Springer, 1998. 3-74.

[3] Katok, Svetlana. Continued fractions, hyperbolic geometry and quadratic forms. *MASS REU Lectures, Summer 2001.* American Mathematical Society, 2003. 121-159.

[4] Katok, Svetlana. *Fuchsian Groups.* University of Chicago Press, 1992. 1-22, 49-56.

[5] Katok, Svetlana; Ugarcovici, Ilie. Arithmetic Coding of Geodesics on the Modular Surface via Continued Fractions. Mathematics Subject Classification, 2000.

[6] Katok, Svetlana; Ugarcovici, Ilie. Symbolic Dynamics for the Modular Surface and Beyond. American Mathematical Society, 2006. 87-132.

[7] Series, Caroline. Non-Euclidean geometry, continued fractions and ergodic theory. Math Intelligencer 4, 1982. 24-31.

[8] Van der Geer, Gerard. *Hilbert Modular Surfaces.* Springer-Verlag, 1987. 1-11.