Arnoux's coding of the geodesic flow on the modular surface

Thomas A. Schmidt

Oregon State University

25 October 2023
OSET

Arnoux's coding of the geodesic flow on the modular surface

Thomas A. Schmidt

Oregon State University

25 October 2023
OSET

The paper

- From MathSciNet

  Arnoux, Pierre: *Le codage du flot géodésique sur la surface modulaire*. (French. English, French summary) [Coding of the geodesic flow on the modular surface] Enseign. Math. (2) 40 (1994), no. 1-2, 29–48.

The paper

- From MathSciNet

  Arnoux, Pierre: *Le codage du flot géodésique sur la surface modulaire*. (French. English, French summary) [Coding of the geodesic flow on the modular surface] Enseign. Math. (2) 40 (1994), no. 1-2, 29–48.

- GOALS of paper: In an elementary manner, give explicit coordinates for the unit tangent bundle of the modular surface and thereby derive explicit expressions for the geodesic flow. Relate this to continued fractions.

The paper

- From MathSciNet

  Arnoux, Pierre: *Le codage du flot géodésique sur la surface modulaire*. (French. English, French summary) [Coding of the geodesic flow on the modular surface] Enseign. Math. (2) 40 (1994), no. 1-2, 29–48.

- GOALS of paper: In an elementary manner, give explicit coordinates for the unit tangent bundle of the modular surface and thereby derive explicit expressions for the geodesic flow. Relate this to continued fractions.

- Earlier work: Ford, Artin, Adler-Flatto, Series. For more on coding, see various works of S. Katok.

Möbius action induced identifications

- The modular surface is $\mathcal{M} := SL_2(\mathbb{Z})\backslash\mathbb{H}$ where $\mathbb{H} = \{x + iy \mid y > 0\}$ with $ds = (dx + dy)/y$.

Möbius action induced identifications

- The modular surface is $\mathcal{M} := SL_2(\mathbb{Z})\backslash\mathbb{H}$ where $\mathbb{H} = \{x + iy \mid y > 0\}$ with $ds = (dx + dy)/y$.
- Möbius action allows $\mathbb{H} \longleftrightarrow SL_2(\mathbb{R})/SO(2)$,

Möbius action induced identifications

- The modular surface is $\mathcal{M} := SL_2(\mathbb{Z})\backslash\mathbb{H}$ where $\mathbb{H} = \{x + iy \mid y > 0\}$ with $ds = (dx + dy)/y$.
- Möbius action allows $\mathbb{H} \longleftrightarrow SL_2(\mathbb{R})/SO(2)$,
- $T^1(\mathbb{H}) \longleftrightarrow PSL_2(\mathbb{R})$,

Möbius action induced identifications

- The modular surface is $\mathcal{M} := SL_2(\mathbb{Z})\backslash\mathbb{H}$ where $\mathbb{H} = \{x + iy \mid y > 0\}$ with $ds = (dx + dy)/y$.
- Möbius action allows $\mathbb{H} \longleftrightarrow SL_2(\mathbb{R})/SO(2)$,
- $T^1(\mathbb{H}) \longleftrightarrow PSL_2(\mathbb{R})$,
- and, $T^1(\mathcal{M}) \longleftrightarrow SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$.

Lattices of covolume one and $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$

We can also identify $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ with 'area one' lattices:

- A lattice $\Gamma \subset \mathbb{R}^2$ uniquely corresponds to the set of its positively oriented bases.

Lattices of covolume one and $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$

We can also identify $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ with 'area one' lattices:

- A lattice $\Gamma \subset \mathbb{R}^2$ uniquely corresponds to the set of its positively oriented bases.
- A positively oriented basis of a lattice uniquely corresponds to the rows of a 2×2 real matrix of positive determinant.

Lattices of covolume one and $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$

We can also identify $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ with 'area one' lattices:

- A lattice $\Gamma \subset \mathbb{R}^2$ uniquely corresponds to the set of its positively oriented bases.
- A positively oriented basis of a lattice uniquely corresponds to the rows of a 2×2 real matrix of positive determinant.
- A lattice of covolume one has its positively oriented bases forming elements of $SL_2(\mathbb{R})$. Oriented changes of basis for such a Γ are effectuated exactly by left multiplication by elements of $SL_2(\mathbb{Z})$.

Theorem 1, Canonical bases

Theorem (Arnoux's coset representatives) The set $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ is in 1-1 correspondence with the set of matrices $A = \begin{pmatrix} a & c \\ -b & d \end{pmatrix}$ such that $(a, b, c, d) \in \Omega_0 \cup \Omega_1 \cup \Omega_2$ where

$$\Omega_0 = \{(a, b, c, d) \in \mathbb{R}^4 \mid ad + bc = 1,\ 0 < b < 1 \leq a,\ 0 \leq c < d\},$$
$$\Omega_1 = \{(a, b, c, d) \in \mathbb{R}^4 \mid ad + bc = 1,\ 0 \leq a < 1 \leq b,\ 0 \leq d < c\},$$
$$\Omega_2 = \{(a, b, 0, d) \in \mathbb{R}^4 \mid ad = 1,\ 0 < b < a < 1\}.$$

Geodesic flow acts (locally) as $(a, b, c, d) \mapsto (ae^{t/2}, be^{t/2}, ce^{-t/2}, de^{-t/2})$.

Theorem 2: Zippered rectangles give (interiors of) Ω_0, Ω_1

Theorem (Arnoux's construction) To each lattice Γ, meeting the coordinate axes only at the origin, the *algorithm below* uniquely associates a positively oriented basis and a fundamental domain that is the union of two rectangles whose bases are aligned and such that one rectangle has width less than 1 and the other greater than one, with the wider taller than the narrower rectangle.

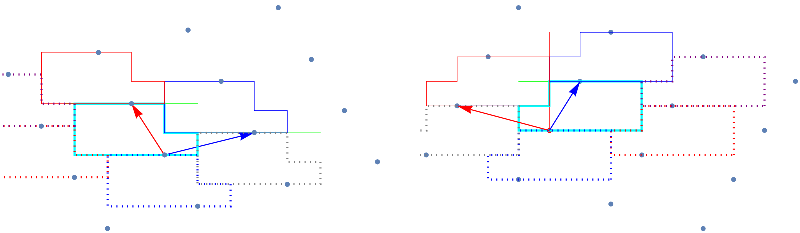


Figure: Cases: Ω_0 and Ω_1.

Theorem 3, Gauss map as factor

Theorem (Arnoux's extension) Let $\Sigma_0 \subset \Omega_0$ be the subset where $a = 1$, and $\Sigma_1 \subset \Omega_1$ be the subset where $b = 1$. Then
(1) the geodesic flow maps the 2D boundary face Σ_0 to a boundary face of Ω_0 that can naturally be identified with Σ_1;
(2) the geodesic flow maps Σ_1 to a face that can in turn be identified with Σ_0;
(3) and, there are real coordinates so that these maps can be given by

$$\Sigma_\epsilon \to \Sigma_{1-\epsilon}$$
$$(x, y) \mapsto (1/x - \lfloor 1/x \rfloor, x - x^2 y).$$

Lebesgue measure is invariant here. This gives a double cover of the natural extension of the standard Gauss map of regular continued fractions.

Theorem 4, verifying Lévy's constant

- **Theorem (Lévy 1936)** For almost every real $x \in [0,1]$ the series of its regular continued fraction denominators $(q_n)_{n\geq 0}$ satisfies
$$\lim_{n\to\infty} \frac{\log q_n}{n} = \frac{\pi^2}{12\log 2}.$$

Theorem 4, verifying Lévy's constant

- **Theorem (Lévy 1936)** For almost every real $x \in [0,1]$ the series of its regular continued fraction denominators $(q_n)_{n\geq 0}$ satisfies

$$\lim_{n\to\infty} \frac{\log q_n}{n} = \frac{\pi^2}{12\log 2}.$$

- Arnoux easily shows that

$$\frac{\pi^2}{12\log 2} = \frac{\mathrm{vol}\; T^1(\mathcal{M})}{2\mathrm{Area}(\Sigma_0 \cup \Sigma_1)}.$$

Theorem 4, verifying Lévy's constant

- **Theorem (Lévy 1936)** For almost every real $x \in [0,1]$ the series of its regular continued fraction denominators $(q_n)_{n\geq 0}$ satisfies

$$\lim_{n\to\infty} \frac{\log q_n}{n} = \frac{\pi^2}{12\log 2}.$$

- Arnoux easily shows that

$$\frac{\pi^2}{12\log 2} = \frac{\mathrm{vol}\, T^1(\mathcal{M})}{2\mathrm{Area}(\Sigma_0 \cup \Sigma_1)}.$$

- Let τ_n be the n-th return time of (thus, the flow value bringing) $(x,0) \in \Sigma_0$ to $\Sigma_0 \cup \Sigma_1$. He shows that q_n/n and $\tau_n/2n$ limit to the same value and argues by Birkhoff sums using the ergodicity of the geodesic flow to get the result.

Algorithm

- Choose $x \in \Gamma$.

Algorithm

- Choose $x \in \Gamma$.
- For each $y \in \Gamma$ with $y \neq x$, let H_y be the horizontal open interval segment of length 2 centered at y. (Green in figure.)

Algorithm

- Choose $x \in \Gamma$.
- For each $y \in \Gamma$ with $y \neq x$, let H_y be the horizontal open interval segment of length 2 centered at y. (Green in figure.)
- The upwards vertical ray emanating from x meets a unique first H_y. (Red in figure.) **Assume this intersection is to the left of y.** This is the case of Ω_1.

Figure: Construction begins

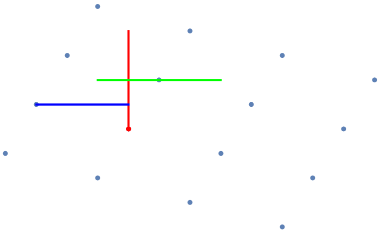


Figure: Beginning of construction, case of intersection left of y.

Algorithm, cont'd

- Let a be the (horizontal) distance from the intersection to y and c be the (vertical) distance from it to x. Thus, $y = x + a + ic$ using complex notation.

Algorithm, cont'd

- Let a be the (horizontal) distance from the intersection to y and c be the (vertical) distance from it to x. Thus, $y = x + a + ic$ using complex notation.

- Let V_x be the open vertical segment from x up to H_y. Of those $z \in \Gamma$ such that H_z extends to intersect V_x, there is a unique z minimizing the distance from z to V_x. (See blue in figure.) Choose this z, label that (horizontal) distance b and label the (vertical) distance from x up to the point of intersection as d.

Algorithm, cont'd

- Let a be the (horizontal) distance from the intersection to y and c be the (vertical) distance from it to x. Thus, $y = x + a + ic$ using complex notation.
- Let V_x be the open vertical segment from x up to H_y. Of those $z \in \Gamma$ such that H_z extends to intersect V_x, there is a unique z minimizing the distance from z to V_x. (See blue in figure.) Choose this z, label that (horizontal) distance b and label the (vertical) distance from x up to the point of intersection as d.
- Note that $\boxed{0 < a < 1 \leq b; \quad 0 < d < c}$.

The reversed ‘L’

- Form the bottom of our reversed ‘L’ as the union of the horizontal segments: of length a to the left of x and of length b to the right.

The reversed ‘L’

- Form the bottom of our reversed ‘L’ as the union of the horizontal segments: of length a to the left of x and of length b to the right.
- Above the left portion of the bottom, erect the rectangle of height d.

The reversed ‘L’

- Form the bottom of our reversed ‘L’ as the union of the horizontal segments: of length a to the left of x and of length b to the right.
- Above the left portion of the bottom, erect the rectangle of height d.
- Above the right portion of the bottom, erect the rectangle of height c.

The reversed ‘L’

- Form the bottom of our reversed ‘L’ as the union of the horizontal segments: of length a to the left of x and of length b to the right.
- Above the left portion of the bottom, erect the rectangle of height d.
- Above the right portion of the bottom, erect the rectangle of height c.
- Form the matrix $A = \begin{pmatrix} a & c \\ -b & d \end{pmatrix}$.

The fundamental domain $\mathcal{F}$; basis: rows of A

Let $\mathcal{F}$ be the reversed ‘L’ so formed. One shows that $\mathcal{F}$ is a fundamental domain for Γ and also the rows of A are a positively oriented basis of Γ. (See figure, next slide.) The theorem thus holds in this case.

Figure: tiling by images of $\mathcal{F}$; identifications give torus

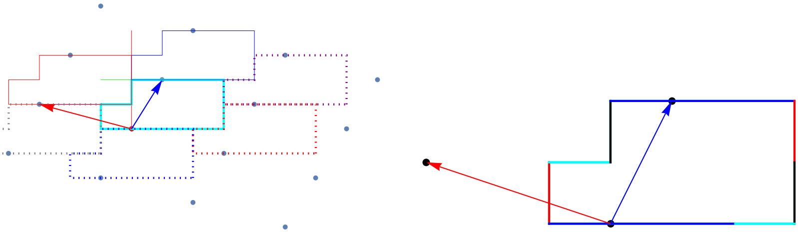
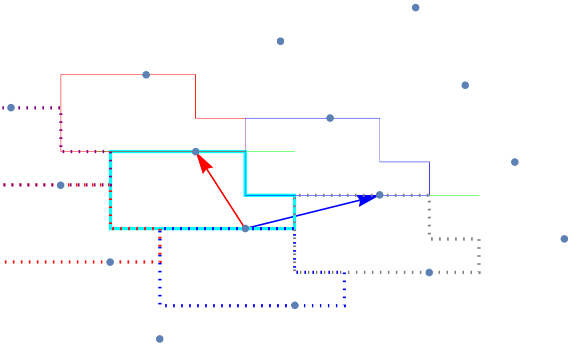


Figure: Case of intersection left of y. Left: $\mathcal{F}$ and its translates. Right: the fundamental domain identifies to a torus. Note that it is the suspension of an IET. Recall $A = \begin{pmatrix} a & c \\ -b & d \end{pmatrix}$.

Ω_0, case of 'L' shape

When the first intersection above is to the **right** of y, we label differently so as to still find an oriented basis corresponding to $A = \begin{pmatrix} a & c \\ -b & d \end{pmatrix}$. Now, $\boxed{0 < b < 1 \leq a; \quad 0 < c < d}$.



Accounting for *all* Γ, 1

- Try to apply the algorithm to a general Γ. It could be Γ includes a 'short' vertical vector and we find that y lies directly above x. If the algorithm succeeds other than this, we choose to include this in the case of Ω_1. That is, we include the possibility of $a = 0$ in the definition of Ω_1.

Accounting for *all* Γ, 1

- Try to apply the algorithm to a general Γ. It could be Γ includes a 'short' vertical vector and we find that y lies directly above x. If the algorithm succeeds other than this, we choose to include this in the case of Ω_1. That is, we include the possibility of $a = 0$ in the definition of Ω_1.

- This was an innocuous choice, as the $(b = 0)$-boundary of Ω_0 as a subset of $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$ is identified with the $(a = 0)$-boundary of Ω_1:

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

Accounting for *all* Γ, 2

- The algorithm in the restricted setting remains unchanged if we include x in V_x. Now a general Γ could have a 'short' horizontal vector so x and z are horizontally aligned. We hence include the possibility of $c = 0$ in the Ω_0 and $d = 0$ in the Ω_1.

Accounting for *all* Γ, 2

- The algorithm in the restricted setting remains unchanged if we include x in V_x. Now a general Γ could have a 'short' horizontal vector so x and z are horizontally aligned. We hence include the possibility of $c = 0$ in the Ω_0 and $d = 0$ in the Ω_1.
- Similarly, if Γ has a 'very short' horizontal vector, it could be that there are two choices for y. Choose the option whose intersection point is to the right; thus, $0 < b < 1$. Let the horizontal distance between those two choices be a. Then we find a basis given by $(a, 0), (-b, d)$. The case of $a \geq 1$ is already in Ω_0. Otherwise, we have $0 < b < a < 1$, $c = 0$ and define Ω_2 to address exactly this case.

Accounting for *all* Γ, 2

- The algorithm in the restricted setting remains unchanged if we include x in V_x. Now a general Γ could have a 'short' horizontal vector so x and z are horizontally aligned. We hence include the possibility of $c = 0$ in the Ω_0 and $d = 0$ in the Ω_1.
- Similarly, if Γ has a 'very short' horizontal vector, it could be that there are two choices for y. Choose the option whose intersection point is to the right; thus, $0 < b < 1$. Let the horizontal distance between those two choices be a. Then we find a basis given by $(a, 0), (-b, d)$. The case of $a \geq 1$ is already in Ω_0. Otherwise, we have $0 < b < a < 1,\ c = 0$ and define Ω_2 to address exactly this case.
- With this, all covolume one Γ are accounted for with the union

$$\Omega_0 = \{(a, b, c, d) \in \mathbb{R}^4 \mid ad + bc = 1,\ 0 < b < 1 \leq a,\ 0 \leq c < d\},$$

$$\Omega_1 = \{(a, b, c, d) \in \mathbb{R}^4 \mid ad + bc = 1,\ 0 \leq a < 1 \leq b,\ 0 \leq d < c\},$$

$$\Omega_2 = \{(a, b, 0, d) \in \mathbb{R}^4 \mid ad = 1,\ 0 < b < a < 1\}.$$

Geodesic flow

- Under the identification of $T^1\mathcal{M}$ with $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$, the geodesic flow is effectuated by right multiplication by diagonal matrices of the form $g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$.

Geodesic flow

- Under the identification of $T^1\mathcal{M}$ with $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$, the geodesic flow is effectuated by right multiplication by diagonal matrices of the form $g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$.
- This flow acts on the various $\mathcal{F}$, stretching horizontally while contracting vertically. Similarly for the image of our positively oriented bases.

Geodesic flow

- Under the identification of $T^1\mathcal{M}$ with $SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})$, the geodesic flow is effectuated by right multiplication by diagonal matrices of the form $g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$.
- This flow acts on the various $\mathcal{F}$, stretching horizontally while contracting vertically. Similarly for the image of our positively oriented bases.
- In particular, $(a, b, c, d) \mapsto (ae^{t/2}, be^{t/2}, ce^{-t/2}, de^{-t/2})$.

Entrance and exit; 3-D

- With mild abuse,
  $\Omega_0 = \{(a, b, c) \in \mathbb{R}^3 \mid 0 < b < 1 \leq a,\ 0 \leq c < 1/(a+b)\}$.
  The flow enters on the $a = 1$-face of boundary and exits on $b = 1$-face of boundary. (See next figure, then come back.)

Entrance and exit; 3-D

- With mild abuse,
  $\Omega_0 = \{(a, b, c) \in \mathbb{R}^3 \mid 0 < b < 1 \leq a,\ 0 \leq c < 1/(a+b)\}$.
  The flow enters on the $a = 1$-face of boundary and exits on $b = 1$-face of boundary. (See next figure, then come back.)

- Cut-and-stack of our rectangles (see figure) shows that exiting face is identified with the entering face, also given by $b = 1$, of $\Omega_1 = \{(a, b, d) \in \mathbb{R}^4 \mid 0 \leq a < 1 \leq b,\ 0 \leq d < 1/(a+b)\}$, as

$$\begin{pmatrix} 1 & \lfloor a \rfloor \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & c \\ -1 & d \end{pmatrix} = \begin{pmatrix} \{a\} & c + \lfloor a \rfloor d \\ -1 & d \end{pmatrix},$$

  since $0 \leq \{a\} = a - \lfloor a \rfloor < 1$ and
  $ad = (1 - c) \leq 1 < a/(1 + \{a\})$ (since the 'exit' has $a > 1$).

Flow figures

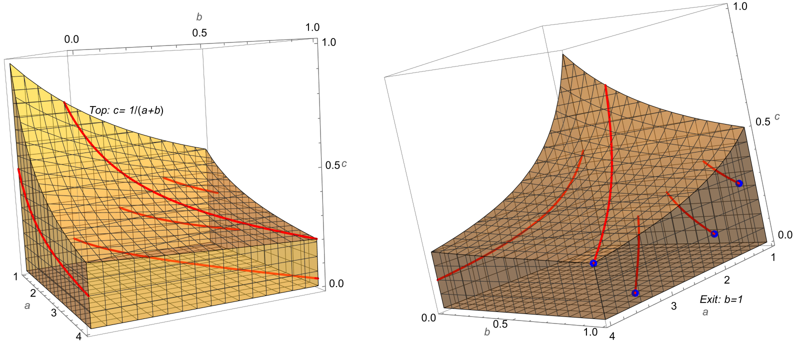


Figure: Two views of flow through Ω_0 from $a = 1$ to $b = 1$ faces.

Cut-and-stack figure

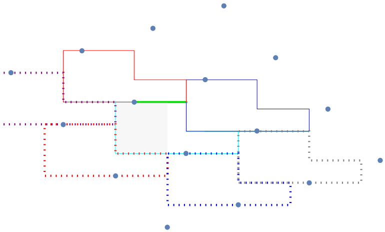


Figure: Here $b = 1$, and (for my ease) $\lfloor a \rfloor = 1$, thus one grey box is cut and then stacked above on the right. In particular, $(-b, d)$ is unchanged.

One 2D map

- The $a = 1$ face of Ω_0 can be parametrized by (b, c) with $0 < b < 1$, $0 < c < 1/(1 + b)$; rename as $(x, y)_0$.

One 2D map

- The $a = 1$ face of Ω_0 can be parametrized by (b, c) with $0 < b < 1$, $0 < c < 1/(1 + b)$; rename as $(x, y)_0$.
- The points flow to the $b = 1$-face, giving values $(a', b', c', d') = (1/b, 1, bc, b(1 - bc)\,)$. Using the identification above, this is the point $(a'', b'', c'', d'') = (1/b - \lfloor 1/b \rfloor, 1, *, b(1 - bc)\,)$.

One 2D map

- The $a = 1$ face of Ω_0 can be parametrized by (b, c) with $0 < b < 1$, $0 < c < 1/(1 + b)$; rename as $(x, y)_0$.
- The points flow to the $b = 1$-face, giving values $(a', b', c', d') = (1/b, 1, bc, b(1 - bc))$. Using the identification above, this is the point $(a'', b'', c'', d'') = (1/b - \lfloor 1/b \rfloor, 1, *, b(1 - bc))$.
- The $b = 1$-face of Ω_1 can be parametrized by (a, d) with $0 \leq a < 1$, $0 \leq d < 1/(a + 1)$. The map is then

$$(x, y)_0 \mapsto (1/x - \lfloor 1/x \rfloor, x - x^2 y)_1.$$

Second 2D map, Theorem 2 is proven.

- The $b = 1$ face of Ω_1 flows to its $a = 1$ boundary face. (Except those points with $a = 0$.) This gives $(x, y)_1 = (a, d) - - > (a', b', c', d') = (\,1, a^{-1}, a(1 - ad), ad\,)$.

Second 2D map, Theorem 2 is proven.

- The $b = 1$ face of Ω_1 flows to its $a = 1$ boundary face. (Except those points with $a = 0$.) This gives $(x, y)_1 = (a, d) - - > (a', b', c', d') = (\,1, a^{-1}, a(1 - ad), ad\,)$.
- Identify to $a = 1$-face of Ω_0

$$\begin{pmatrix} 1 & 0 \\ \lfloor b' \rfloor & 1 \end{pmatrix} \begin{pmatrix} 1 & c' \\ -b' & d' \end{pmatrix} = \begin{pmatrix} 1 & c' \\ -\{b'\} & d' + \lfloor b' \rfloor c' \end{pmatrix},$$

Second 2D map, Theorem 2 is proven.

- The $b = 1$ face of Ω_1 flows to its $a = 1$ boundary face. (Except those points with $a = 0$.) This gives $(x, y)_1 = (a, d) - - > (a', b', c', d') = (\,1, a^{-1}, a(1 - ad), ad\,)$.

- Identify to $a = 1$-face of Ω_0

$$\begin{pmatrix} 1 & 0 \\ \lfloor b' \rfloor & 1 \end{pmatrix} \begin{pmatrix} 1 & c' \\ -b' & d' \end{pmatrix} = \begin{pmatrix} 1 & c' \\ -\{b'\} & d' + \lfloor b' \rfloor c' \end{pmatrix},$$

- Using the (new) (b, c), the map is then

$$(x, y)_1 \mapsto (\,1/x - \lfloor 1/x \rfloor, x - x^2 y\,)_0.$$

Theorem 3, Gauss map as factor. Repeated!

Theorem (Arnoux's extension) Let $\Sigma_0 \subset \Omega_0$ be the subset where $a = 1$, and $\Sigma_1 \subset \Omega_1$ be the subset where $b = 1$. Then
(1) the geodesic flow maps the 2D boundary face Σ_0 to a boundary face of Ω_0 that can naturally be identified with Σ_1;
(2) the geodesic flow maps Σ_1 to a face that can in turn be identified with Σ_0;
(3) and, there are real coordinates so that these maps can be given by

$$\Sigma_\epsilon \to \Sigma_{1-\epsilon}$$
$$(x, y) \mapsto (1/x - \lfloor 1/x \rfloor, x - x^2 y).$$

Lebesgue measure is invariant here. This gives a double cover of the natural extension of the standard Gauss map of regular continued fractions.

The natural extension.

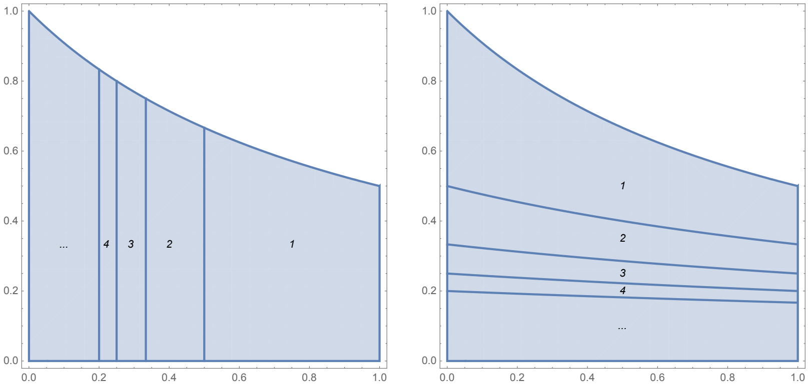


Figure: The natural extension map.

Thanks!