THE CONTINUED FRACTION SYSTEM (AND RELATED SYSTEMS)[1]
LaTeX'd by Kyle Kneisl[2]

1. Basic Definitions

Each $x \in \mathbb{R}$ has a continued fraction representation given by

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}} = [a_0; a_1, a_2, a_3, \dots] \in \mathbb{N}^{\mathbb{N}},$$

where $a_0 \in \mathbb{Z}$ and $a_i \in \mathbb{N}$ for all $i \geq 1$.

We now define some operations on $\mathbb{R}$ that will be frequently used in our discussion.

Definition 1 ($\lfloor\cdot\rfloor$ and $\langle\cdot\rangle$)**.** We define the *floor* function $\lfloor\cdot\rfloor : \mathbb{R} \longrightarrow \mathbb{Z}$ to be $x \mapsto z$, where z is the greatest integer less than or equal to x. The *fractional part* of x is given by the function $\langle\cdot\rangle : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $x \mapsto x - \lfloor x \rfloor$.

As $\langle x\rangle = [0; a_1, a_2, \dots]$, we will focus on $x \in [0, 1)$. That is to say, we will be taking $a_0 = 0$ in what follows and suppress the writing of $a_0 = 0$ so that by $[a_1, a_2, \dots]$ we mean $[0; a_1, a_2, \dots]$.

2. The Continued Fraction Algorithm

Define $T : (0, 1) \to (0, 1)$ by $Tx = \left\langle \frac{1}{x} \right\rangle$. T is called the *Gauss map.* The following shift property for T will be crucial in our analysis.

Proposition 2. $T[a_1, a_2, a_3, \dots] = [a_2, a_3, \dots]$. *That is, the Gauss map* $T : (0, 1) \to (0, 1)$ *corresponds to the shift on* $\mathbb{N}^{\mathbb{N}}$, $\sigma[a_1, a_2, \dots] = [a_2, a_3, \dots]$.

[1]These notes are the result of work done in an ergodic theory class presented by Karl Petersen during the Spring 2000 semester at the University of North Carolina. The first part follows Billingsley's book, while the later parts were based on papers of Jeff Lagarias and Caroline Series as well as the dissertation of Haijing Ma.

[2]I would like to thank Dr. Petersen and Léonard Choup of the Mathematics Department as well as Daniel Spitzner of the Statistics Department for allowing me to refer to their notes in addition to my own.

Proof.

$$\begin{aligned}
T[a_1, a_2, a_3, \dots] &= T\left[\cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}}}\right] \\
&= \left\langle a_1 + \cfrac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}} \right\rangle \\
&= \cfrac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}} \\
&= [a_2, a_3, \dots].
\end{aligned}$$

□

Put $x_0 = x \in \mathbb{R}$ and write:

$$\begin{aligned}
a_0 &= \lfloor x_0 \rfloor, & x_1 &= \langle x_0 \rangle, \\
a_1 &= \left\lfloor \frac{1}{x_1} \right\rfloor, & x_2 &= \left\langle \frac{1}{x_1} \right\rangle = Tx_1, \\
a_2 &= \left\lfloor \frac{1}{x_2} \right\rfloor, & x_3 &= \left\langle \frac{1}{x_2} \right\rangle = Tx_2, \\
&\dots & &\dots
\end{aligned}$$

This gives us a coding of $(0,1)$ by the shift space on a countable alphabet. Under this coding, the intervals $\left(\frac{1}{n+1}, \frac{1}{n}\right)$ correspond to the time-0 partition in $X = \mathbb{N}^{\mathbb{N}}$ when $x \in [0,1)$ is coded by $[a_1, a_2, a_3, \dots] \in \mathbb{N}^{\mathbb{N}}$: $\left(\frac{1}{n+1}, \frac{1}{n}\right) = \{x \in (0,1) : a_1(x) = n\}$.

Example. x=$\sqrt{2}/2 \in [0,1)$ has coding $[1, 2, 2, 2, 2, \dots]$.

3. A Recurrence Relation

Put

$$\frac{p_n}{q_n} = [a_1, a_2, a_3, \dots, a_n].$$

In slightly misleading terminology, $\frac{p_n}{q_n} = [a_1, a_2, a_3, \dots, a_n]$ is called the n'th *convergent* and a_n is called the n'th *partial quotient.* Setting

$\frac{p_{-1}}{q_{-1}} = \frac{1}{0}$ and $\frac{p_0}{q_0} = \frac{0}{1}$, we obtain

$$\frac{p_{n+1}}{q_{n+1}} = \frac{p_{n-1} + a_{n+1}p_n}{q_{n-1} + a_{n+1}q_n}, \qquad \text{for all } n \geq 0,$$

which we may write in the form

$$\begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix},$$

that is

$$\begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_{n+1} \end{pmatrix}.$$

Now, suppose that $x = [a_1, a_2, \ldots]$. Then it can be shown that

$$x - \frac{p_n}{q_n} = (-1)^n \frac{1}{q_n q_{n+1}} \theta_n$$

for some $\theta_n \in [0,1]$, and

$$\left| x - \frac{p_n}{q_n} \right| \leq \frac{1}{q_n q_{n+1}} < \frac{1}{{q_n}^2}$$

for all n.

4. The Gauss Measure on $[0,1)$

It turns out that T preserves the *Gauss measure* given in the following definition.

Definition 3 (Gauss measure)**.** The Gauss measure G on $[0,1)$ is given by

$$\mathrm{G}(\Omega) = \frac{1}{\log 2} \int_\Omega \frac{1}{1+x}\, dx$$

for all measurable $\Omega \subseteq [0,1]$.

It is clear that this is a measure and that G has the same null sets as Lebesgue measure, so that a statement holding almost everywhere with respect to Gauss measure also holds m-.a.e., where m is Lebesgue measure. It is also easy to see that

$$\left(\frac{1}{2\log 2}\right) m(A) \leq \mathrm{G}(A) \leq \left(\frac{1}{\log 2}\right) m(A)$$

for all measurable sets A.

Theorem 4. *The non-invertible map $T : (0,1) \to (0,1)$ preserves Gauss measure.*

Proof. Fix $t > 0$. We show that $\mathrm{G}T^{-1}[0,t] = \mathrm{G}[0,t]$, and this yields the result. It is easy to verify that

$$T^{-1}[0,t] = \{x : 0 \le Tx \le t\} = \bigcup_{n=1}^{\infty} \left[\frac{1}{n+t}, \frac{1}{n}\right],$$

whence

$$\begin{aligned}
GT^{-1}[0,t] &= \sum_{n=1}^{\infty} \mathrm{G}\left[\frac{1}{n+t}, \frac{1}{n}\right] \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \int_{\frac{1}{n+t}}^{\frac{1}{n}} \frac{dx}{1+x} \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \left[\log\left(1+\frac{1}{n}\right) - \log\left(1+\frac{1}{n+t}\right)\right] \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \log\left(\frac{n+1}{n} \cdot \frac{n+t}{n+t+1}\right) \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \left[\log\left(1+\frac{t}{n}\right) - \log\left(1+\frac{t}{n+1}\right)\right] \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \int_{t/(n+1)}^{t/n} \frac{dx}{1+x} \\
&= \frac{1}{\log 2} \int_0^t \frac{dx}{1+x} \\
&= \mathrm{G}[0,t].
\end{aligned}$$

□

In a letter to Laplace, Gauss stated essentially that the iterates of Lebesgue measure under T converge to G in that $m(T^{-n}I) \longrightarrow \mathrm{G}(I)$ for every interval I. Kuzmin, Lévy, and Wirsing showed that that the convergence is exponentially fast:

$$\left|m(T^{-n}I) - \mathrm{G}(I)\right| \le c\gamma^n \qquad \text{for all } I \text{ and } n,$$

with $\gamma \simeq 0.303663$. Better understanding of this phenomenon may involve the Riemann Hypothesis.

5. Ergodicity of T with Respect to Gauss Measure

Definition 5 (The relation $\asymp$)**.** For expressions u and v, $u \asymp v$ means that there are positive constants c_1 and c_2 such that $c_1 v \le u \le c_2 v$.

Definition 6. For $[a_1, a_2, \ldots, a_n]$ define the interval

$I_n[a_1, a_2, \ldots, a_n] = \{[a_1, a_2, \ldots, a_n, b_{n+1}, b_{n+2}, \ldots] : b_{n+i} \in \mathbb{N} \text{ for all } i > 0\}.$

Note that the intervals $I_n[a_1, \ldots, a_n]$ have lengths that shrink uniformly to 0, since $m(I_n) \leq c/2^n$.

Theorem 7. *T is ergodic with respect to the Gauss measure* G.

Proof. The key fact that we will use is that $\mathrm{G}(T^{-n}A|I_n) \asymp G(A)$ for all $A \in \mathcal{B}$. That is,

$$\frac{\mathrm{G}(T^{-n}A \cap I_n)}{G(I_n)} \asymp G(A),$$

or

$$\mathrm{G}(T^{-n}A \cap I_n) \asymp G(A)G(I_n),$$

whence

$$\mathrm{G}(T^{-n}A \cap B) \asymp G(A)G(B)$$

for all $B \in \mathcal{B}$.

Let us fix n and $a_1, \ldots, a_n$. Then,

$$x \in I_n \text{ if and only if } x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{\cdots}{1 + \cfrac{1}{a_n + \theta_n(x)}}}}$$

with $0 \leq \theta_n(x) \leq 1$. Now take $A = [u, v]$ where $0 \leq u \leq v \leq 1$. Then

$$m(I_n \cap T^{-n}[u, v]) = \theta_n^{-1}(v) - \theta_n^{-1}(u).$$

We now claim that

$$\theta_n^{-1}(t) = \frac{1}{a_1 + \cdots + \frac{1}{a_n + t}} = \frac{p_n + tp_{n-1}}{q_n + tq_{n-1}}.$$

To establish the claim, we check it first for $n = 1$:

$$\frac{p_1 + tp_0}{q_1 + tq_0} = \frac{1 + t \cdot 0}{a_1 + t \cdot 1},$$

since $p_{-1} = 1$, $p_0 = 0$, $q_{-1} = 0$, $q_0 = 1$ and $p_{n+1} = a_n p_n + p_{n-1}$ with $q_{n+1} = a_n q_n + q_{n-1}$.

Now, let us consider the $n+1$ case, assuming the induction hypothesis. We have

$$\frac{1}{a_1 + \frac{1}{\cdots + a_n + \frac{1}{a_{n+1} + t}}} = \frac{1}{a_1 + \frac{1}{\cdots + a_n + s}},$$

where $s = \frac{1}{a_{n+1}+t}$. By induction, the above is equal to

$$\frac{p_n + sp_{n-1}}{q_n + sq_{n-1}} = \frac{p_n + \left(\frac{1}{a_{n+1}+t}\right) p_{n-1}}{q_n + \left(\frac{1}{a_{n+1}+t}\right) q_{n-1}} = \frac{a_{n+1}p_n + p_{n-1} + tp_n}{a_{n+1}q_n + q_{n-1} + tq_n} = \frac{p_{n+1} + tp_n}{q_{n+1} + tq_n}.$$

This establishes the claim.

Next, we compute

$$\begin{aligned}
\frac{m\left(T^{-n}[u,v] \cap I_n\right)}{m(I_n)} &= \frac{\theta_n^{-1}(v) - \theta_n^{-1}(u)}{\theta_n^{-1}(1) - \theta_n^{-1}(0)} \\
&= \frac{\frac{p_n+vp_{n-1}}{q_n+vq_{n-1}} - \frac{p_n+up_{n-1}}{q_n+uq_{n-1}}}{\frac{p_n+p_{n-1}}{q_n+q_{n-1}} - \frac{p_n}{q_n}} \\
&= \dots \\
&= (v-u)\frac{q_n(q_n+q_{n-1})}{(q_n+vq_{n-1})(q_n+uq_{n-1})} \\
&= m[u,v]\frac{q_n(q_n+q_{n-1})}{(q_n+vq_{n-1})(q_n+uq_{n-1})}.
\end{aligned}$$

But $\dfrac{q_n(q_n+q_{n-1})}{(q_n+vq_{n-1})(q_n+uq_{n-1})} \asymp 1$, and so for every $A \in \mathcal{B}$ we have

$$\frac{m(T^{-n}A \cap I_n)}{m(I_n)} \asymp m(A).$$

Because m and G are comparable measures, we have now established that

$$\frac{\mathrm{G}(T^{-n}A \cap I_n)}{\mathrm{G}(I_n)} \asymp \mathrm{G}(A)$$

for every $A \in \mathcal{B}$.

Now suppose $T^{-1}A = A$ and $\mathrm{G}(A) > 0$. We have that

$$\frac{\mathrm{G}(T^{-n}A \cap I_n)}{\mathrm{G}(A)} \asymp \mathrm{G}(I_n)$$

for every n. Since the I_n generate, we have

$$\frac{\mathrm{G}(A \cap B)}{\mathrm{G}(A)} \asymp \mathrm{G}(B)$$

for every $B \in \mathcal{B}$. Now take $B = A^c$. Then $\mathrm{G}(B) \asymp 0$, so $\mathrm{G}(B) = 0$ and A is thus a set of full measure with respect to Gauss measure. This establishes G-ergodicity of T. □

6. f-Expansions

Recall that we have established that the Gauss map $Tx = \frac{1}{x} \bmod 1$

1. preserves the Gauss measure $d\mathrm{G} = \dfrac{dx}{(\log 2)(1+x)}$ on $[0,1]$, and
2. is ergodic with respect to this measure.

One can now generalize the continued fraction map. Take a piecewise monotonic function

$$f : [0,1] \to [0,1].$$

Let $\{A_n\}$ be the fixed partition of $[0,1]$ into intervals of monotonicity of f. Now define the function

$$a_k(x) = j \text{ if and only if } T^k x \in j\text{'th cell of the partition } \{A_n\}.$$

Often one can find an invariant measure for f that is absolutely continuous with repsect to m and analyze the resulting dynamical system.

7. Average Size of the Entries in a Typical Continued Fraction

We consider now estimates of the growth rates of a_n (and q_n in the next section) for typical x. First, note that

$$\begin{aligned}
\mathrm{G}\{a_1 = n\} &= \mathrm{G}\left[\frac{1}{n+1}, \frac{1}{n}\right] \\
&= \frac{1}{\log 2}\int_{\frac{1}{n+1}}^{\frac{1}{n}} \frac{dx}{1+x} \\
&= \frac{1}{\log 2}\log\left(\frac{(n+1)^2}{n(n+2)}\right) \\
&\asymp m\{a_1 = n\} \\
&= m\left[\frac{1}{n+1}, \frac{1}{n}\right] \\
&= \frac{1}{n} - \frac{1}{n+1} \\
&\asymp \frac{1}{n^2},
\end{aligned}$$

so

$$\int_0^1 a_1(x)\,d\mathrm{G}(x) = \sum_{n=1}^{\infty} n\mathrm{G}\{a_1 = n\} \asymp \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Thus, a_1 is finite almost everywhere but not integrable with respect to either the Gauss measure or Lebesgue measure. Hence, from a slight converse of the Ergodic Theorem, we obtain

$$\frac{1}{n}\left(a_1(x)+a_2(x)+\cdots+a_n(x)\right)=\frac{1}{n}\sum_{k=0}^{n-1} a_1\left(T^k x\right) \to \infty \text{ a.e.}\,.$$

In particular, for each k, $\{x : \ |a_n(x)| \leq k \text{ for all } n\}$ is a null set. Typical points x have many large $a_n(x)$'s.

Now we estimate the average exponential growth rate of $a_n(x)$. Let

$$\gamma=\prod_{k=1}^{\infty}\left[1+\frac{1}{k(k+2)}\right]^{\log k/\log 2}.$$

Proposition 8. *$[a_1(x)a_2(x)\dots a_n(x)]^{\frac{1}{n}} \longrightarrow \gamma$ as $n\to\infty$ for almost every x, i.e.* $\frac{1}{n}\log[a_1(x)a_2(x)\dots a_n(x)] \to \sum_{k=1}^{\infty}\frac{\log k}{\log 2}\log\left[1+\frac{1}{k(k+2)}\right]$ *Lebesgue almost everywhere.*

Proof. Note that $\int \log a_1(x)\,d\mathrm{G}=\sum \log n\cdot \mathrm{G}\{a_1=n\} \asymp \sum \log n\cdot \frac{1}{n^2} < \infty$, so that $\log a_1 \in L^1$. By the Ergodic Theorem,

$$\begin{aligned}\frac{1}{m}\sum_{k=0}^{m-1}\log a_1(T^k x) \to \int \log a_1(x)\,d\mathrm{G}(x) &= \sum_{n=1}^{\infty}\log n\cdot \mathrm{G}\{a_1=n\}\\ &=\sum_{n=1}^{\infty}\left(\frac{\log n}{\log 2}\cdot\log\frac{(n+1)^2}{n(n+2)}\right)\end{aligned}$$

almost everywhere as $m\to\infty$. ☐

8. Average Rate of Growth of the Denominators $q_n = q_n(x)$

Recall that $[a_1,a_2,\dots,a_n]=\dfrac{p_n}{q_n}$.

Proposition 9. $\frac{1}{n}\log q_n(x) \longrightarrow \frac{\pi^2}{12\log 2}$ *a.e.; that is,*

$$q_n \sim e^{n\pi^2/(12\log 2)}$$

almost everywhere as $n\to\infty$.

Proof. We begin with the following calculations:

$$
\begin{aligned}
-\frac{1}{n}\sum_{k=0}^{n-1}\log(T^k x) \to -\int_0^1 \log x \, d\mathrm{G}(x) &= -\frac{1}{\log 2}\int_0^1 \frac{\log x}{1+x}\, dx \\
&= -\frac{1}{\log 2}\cdot\left[\log x \log(1+x) - \int \frac{\log(1+x)}{x}\, dx\right]_0^1 \\
&= \frac{1}{\log 2}\int_0^1 \frac{\log(1+x)}{x}\, dx \\
&= \frac{1}{\log 2}\int_0^1 \frac{1}{x}\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots\right) dx \\
&= \frac{1}{\log 2}\int_0^1 \left(1 - \frac{x}{2} + \frac{x^2}{3} - \frac{x^3}{4} + \dots\right) dx \\
&= \frac{1}{\log 2}\left[x - \frac{x^2}{4} + \frac{x^3}{9} - \frac{x^4}{16} + \dots\right]_0^1 \\
&= \frac{1}{\log 2}\left(1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \dots\right) \\
&= \frac{1}{\log 2}\left(\frac{\pi^2}{12}\right),
\end{aligned}
$$

where we have used the fact that $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \frac{\pi^2}{6}$. We claim that

$$
-\frac{1}{n}\sum_{k=0}^{n-1}\log(T^k x) - \frac{1}{n}\log q_n(x) \to 0
$$

a.e. as $n \to \infty$. The relationship

$$
p_{n+1}(x) = q_n(Tx)
$$

for all $n \geq 0$ and $x \in [0,1]$ can be easily verified using a "semi-strong induction" in the sense that the $n-1$ and $n-2$ cases are assumed in the induction hypothesis to establish the n case. To establish the claim, we proceed as follows:

$$
q_n(x) = \frac{q_n(x)}{p_n(x)} \cdot \frac{q_{n-1}(Tx)}{p_{n-1}(Tx)} \cdots \frac{q_1(T^{n-1}x)}{p_1(T^{n-1}x)} \cdot q_0(T^n(x)),
$$

whence

$$
\log q_n(x) = \sum_{j=1}^{n} \log \frac{q_j(T^{n-j}x)}{p_j(T^{n-j}x)}.
$$

But,

$$\left|-\log(T^{n-j}x) - \log\frac{q_j(T^{n-j}x)}{p_j(T^{n-j}x)}\right| = \left|\log\frac{p_j(T^{n-j}x)}{q_j(T^{n-j}x)} - \log(T^{n-j}x)\right|$$

$$\asymp m(I_j) \le \frac{c}{2^j}.$$

This is because when $x = [a_1, a_2, a_3, a_4, \ldots]$, we have

$$T^{n-j}x = [a_{n-j}, a_{n-j+1}, \ldots, a_n, \ldots] \in I_j[a_{n-j}, \ldots, a_{n-1}].$$

As $m(I_j) \asymp \frac{1}{2^j}$, we have

$$\left|\frac{p_j(T^{n-j}x)}{q_j(T^{n-j}x)} - T^{n-j}x\right| \le m(I_j).$$

We now show that $\left|x - \frac{p_n(x)}{q_n(x)}\right|$ and $\left|\frac{x}{p_n(x)/q_n(x)} - 1\right|$ are both small, whence $|\log(T^jx) - \log[a_j(x), \ldots, a_n(x)]| = r_j(x)$ is summable in j, uniformly in x. To that end, if $0 \le u \le 1$ and $\left|\frac{v}{u} - 1\right| < \varepsilon$, then applying the Mean Value Theorem yields

$$|\log u - \log v| \le \frac{1}{u}(v-u) = \left|1 - \frac{v}{u}\right| < \varepsilon.$$

Hence, if

$$x = \frac{1}{a_1 + \cdots + \frac{1}{a_n + t}} = \frac{p_n + tp_{n-1}}{q_n + tq_{n-1}} \quad \text{for } 0 \le t \le 1,$$

then we have

$$x - \frac{p_n}{q_n} = \frac{p_n + tp_{n-1}}{q_n + tq_{n-1}} - \frac{p_n}{q_n} = \cdots = t\,\frac{p_{n-1}q_n - p_nq_{n-1}}{q_n(q_n + tq_{n-1})}.$$

Since $p_{n-1}q_n - p_nq_{n-1} = (-1)^n$, the above expression becomes

$$x - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n\left(\frac{1}{t}q_n + q_{n-1}\right)}.$$

Noting that $a_{n+1} \le 1/t < a_{n+1} + 1$, we obtain

$$\frac{1}{q_n((a_{n+1}+1)q_n + q_{n-1})} \le \left|x - \frac{p_n}{q_n}\right| \le \frac{1}{q_n(a_{n+1}q_n + q_{n-1})}.$$

Using the recurrence relations in the denominator of the above expression yields

$$\frac{1}{q_n(q_{n+1} + q_n)} \le \left|x - \frac{p_n}{q_n}\right| \le \frac{1}{q_nq_{n+1}}.$$

Finally, the facts

$$p_n \geq q^{n-2/2} \quad , \quad q_n \geq q^{n-1/2}$$

for every $n \geq 2$ establish that

$$\left| \frac{x}{p_n(x)/q_n(x)} - 1 \right| \asymp \frac{1}{q^n}.$$

□

An important consequence is the determination of the speed with which the continued fraction convergents approximate their target point.

Corollary 10.

$$\frac{1}{n} \log \left| x - \frac{p_n(x)}{q_n(x)} \right| \longrightarrow -\frac{\pi^2}{6 \log 2}$$

almost everywhere.

Proof. Note that

$$-\log \left| x - \frac{p_n}{q_n} \right| \geq \log q_n + \log(q_{n+1} + q_n)$$

and

$$-\log \left| x - \frac{p_n}{q_n} \right| \leq \log q_n + \log q_{n+1}.$$

But $\log(q_{n+1} + q_n) \in (\log(2q_n), \log(2q_{n+1}))$, and the result follows. □

Corollary 11. *The entropy of the Gauss map*

$$Tx = \frac{1}{x} \mod 1$$

on $[0,1]$ *with the Gauss measure is* $\dfrac{\pi^2}{6 \log 2}$.

Proof. The basic intervals $I_{a_1 \ldots a_n} = \{x : \; a_1(x) = a_1, \ldots, a_n(x) = a_n\}$ generate the σ-algebra of Lebesgue measurable sets. The countable time-0 partition $\alpha = \{I_1, I_2, \ldots\}$ (with $I_n = \{a_n = n\}$) forms a countable generator for T:

$$I_{a_1 \ldots a_n} = I_{a_1} \cap T^{-1} I_{a_2} \cap \cdots \cap T^{-n+1} I_{a_n}.$$

Therefore, the entropy with respect to the Gauss measure is $h_{\mathrm{G}}(T) = h_{\mathrm{G}}(\alpha, T)$, the entropy with respect to the time-0 partition α.

By the Shannon-McMillan-Breiman Theorem,

$$h_{\mathrm{G}}(\alpha, T) = \lim_{n \to \infty} \frac{1}{n} \log(\alpha_n(x))$$

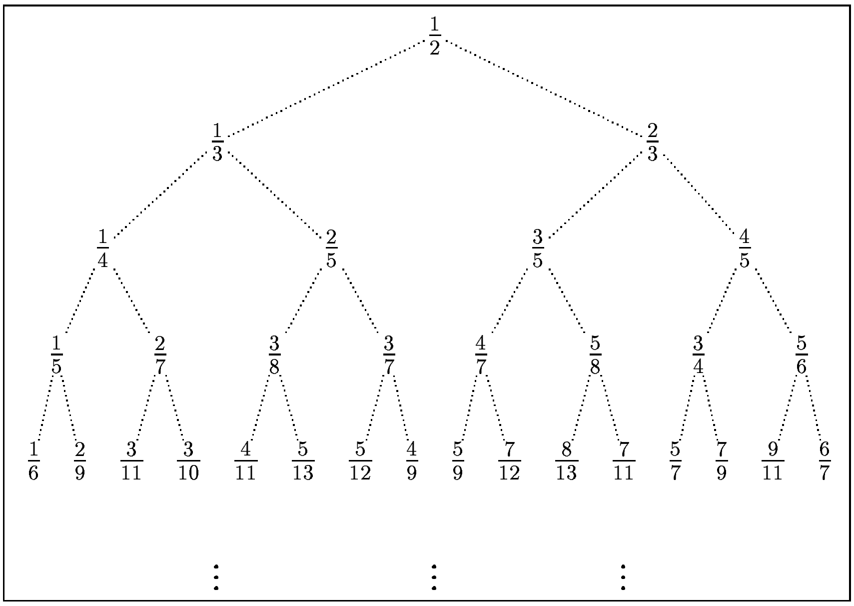

FIGURE 1. The top of the Farey tree.

almost everywhere, where $\alpha_n(x) = I_{a_1(x)\dots a_n(x)}$ is the cell of the partition $\alpha_n = \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha$ to which x belongs. From the above,

$$m(I_n) = \left| \frac{p_n + p_{n-1}}{q_n + q_{n-1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n(q_n + q_{n-1})},$$

so that

$$-\frac{1}{n} \log m(\alpha_n(x)) \longrightarrow \frac{\pi^2}{6 \log 2} \quad \text{a.e.}$$

Finally, since $G \asymp m$, it follows that

$$-\frac{1}{n} \log \mathrm{G}(\alpha_n(x)) \longrightarrow \frac{\pi^2}{6 \log 2} \quad \text{a.e.}$$

□

9. THE FAREY TREE

John Farey [1] was a surveyor, geologist and amateur astronomer, among other things. He occasionally found the time to write articles in the *Philosophical Magazine*. John Farey discovered the so-called *Farey tree* and Farey series in 1816 during Napoleonic times. These ideas were in reality first suggested by C. Haros circa 1802. However, Cauchy was apparently not aware of the work of Haros and attributed the tree to Farey, and this mistake has been repeated ever since.

The Farey tree can be constructed as follows, reminiscent of the construction of Pascal's Triangle.

1. Every column contains precisely one entry.
2. The only entry in the first row is $\frac{1}{2}$.
3. Each subsequent row defines twice as many columnar entries as the previous row.
4. No row is defined until the preceding row is defined. That is, the n'th row depends only upon the columns defined by the entries in rows 1 through $n-1$.
5. Each entry defining a column has numerator (resp. denominator) the sum of the numerators (resp. denominators) of the fractions which define the column to the immediate left and immediate right of the present entry.
6. If the present entry is in the leftmost column (and hence there is no fraction to its left), the entry is determined by summing the numerators (resp. denominators) of the fraction to its immediate right and the fraction $\frac{0}{1}$.
7. If the present entry is in the rightmost column (and hence there is no fraction to its right), the entry is determined by summing the numerators (resp. denominators) of the fraction to its immediate left and the fraction $\frac{1}{1}$.
8. All entries must be rendered in lowest terms.

Remarkably, every rational number in $(0,1)$ appears exactly one time in this array. Every real number in $(0,1)$ corresponds to a unique finite or infinite path on the Farey tree.

Example. In Figure 1, the entry $\frac{7}{12}$ was computed via $\frac{7}{12} = \frac{4+3}{7+5}$, since the entry to the immediate left was $\frac{4}{7}$ and the entry to the immediate right was $\frac{3}{5}$. Note that one of the summands was not in the row immediately previous to $\frac{7}{12}$.

Consider now two fractions $\frac{p}{q}$ and $\frac{r}{s}$ in lowest terms.

Definition 12 (Farey Neighbors and Farey Intervals)**.** We say that $\frac{p}{q}$ and $\frac{r}{s}$ are *Farey neighbors* if $ps - qr = \pm 1$. Two Farey neighbors define a *Farey interval.* Each Farey interval is "labeled" uniquely according to the *mediant* (child) of the neighbors, given by $\frac{p+r}{q+s}$.

Example. $\frac{1}{4}$ and $\frac{2}{7}$ are Farey neighbors whose mediant is $\frac{3}{11}$.

Let us now recall our earlier matrix product for continued fractions. We have $p_n = a_n p_{n-1} + p_{n-2}$ and $q_n = a_n q_{n-1} + q_{n-2}$ with $p_{-1} = q_0 = 1$

and $q_{-1} = p_0 = 0$. This gives rise to the matrix formula

$$\begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & a_1 \end{pmatrix}}_{BA^{a_1-1}} \cdots \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix}}_{BA^{a_n-1}},$$

where

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

That is, we may write any continued fraction representation in the following way:

$$B \underbrace{AA\ldots A}_{a_1-1 \text{ times}} B \underbrace{AA\ldots A}_{a_2-1 \text{ times}} B \underbrace{AA\ldots A}_{a_3-1 \text{ times}} \ldots$$

The results of the intermediate matrix products which arise this way correspond to the *intermediate convergents* or *Farey fractions* as will soon become clear.

Definition 13 (Farey fractions)**.** Continuing with $p_{-1} = q_0 = 1$ and $q_{-1} = p_0 = 0$, the sets $\{\frac{jp_{n-1} + p_{n-2}}{jq_{n-1} + q_{n-2}}\}$ for $1 \leq j < a_n$ are the n'th *Farey fractions* for a real number in the interval.

Example (Ordinary continued fraction (OCF) convergents). Let

$$x = [2, 3, 2, 4, \ldots].$$

Recall the recurrence $p_n = a_n p_{n-1} + p_{n-2}$ and $q_n = a_n q_{n-1} + q_{n-2}$ with $p_{-1} = q_0 = 1$ and $q_{-1} = p_0 = 0$. Now we have

$$n = 1 \qquad \frac{1}{2} = \frac{a_1 p_0 + p_{-1}}{a_1 q_0 + q_{-1}}$$

$$n = 2 \qquad \frac{1}{2 + \frac{1}{3}} = \frac{3}{7} = \frac{3 \cdot 1 + 0}{3 \cdot 2 + 1}$$

$$n = 3 \qquad \frac{1}{2 + \frac{1}{3 + \frac{1}{2}}} = \frac{7}{16} = \frac{2 \cdot 3 + 1}{2 \cdot 7 + 2}$$

$$n = 4 \qquad \frac{1}{2 + \frac{1}{3 + \frac{1}{2 + \frac{1}{4}}}} = \frac{31}{71} = \frac{4 \cdot 7 + 3}{4 \cdot 16 + 7}.$$

Thus, the first few OCF convergents are $\frac{1}{2}, \frac{3}{7}, \frac{7}{16}, \frac{31}{71}$.

Example (Farey fraction convergents). As before, let

$$x = [2, 3, 2, 4, \ldots].$$

We then have

$$\begin{aligned}
n=1, j=1: &\quad \frac{1\cdot 0+1}{1\cdot 1+0}=1\\
n=2, j=1: &\quad \frac{1\cdot 1+0}{1\cdot 2+1}=\frac{1}{3}\\
n=2, j=2: &\quad \frac{2\cdot 1+0}{2\cdot 2+1}=\frac{2}{5}\\
n=3, j=1: &\quad \frac{1\cdot 3+1}{1\cdot 7+2}=\frac{4}{9}\\
n=4, j=1: &\quad \frac{1\cdot 7+3}{1\cdot 16+7}=\frac{10}{23}\\
n=4, j=2: &\quad \frac{2\cdot 7+3}{2\cdot 16+7}=\frac{17}{39}\\
n=4, j=3: &\quad \frac{3\cdot 7+3}{3\cdot 16+7}=\frac{24}{55}.
\end{aligned}$$

Thus, the first few Farey fractions for x are:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}, \frac{7}{16}, \frac{10}{23}, \frac{17}{39}, \frac{24}{55}, \frac{31}{71},$$

where these fractions are those computed from the above list, followed by the OCF approximant for each n. Successive entries are Farey neighbors, as can be seen by descending to each of these entries in the Farey tree (see Figure 1).

The n'th OCF, p_n/q_n, is the *best rational approximation* to x whose denominator is no greater than q_n. The n'th Farey approximant s_n/t_n is the *best* one-sided *rational approximation* to x among the set of all rational numbers s/t with $t \leq t_n$. This means that there are no s/t closer to, and on the same side of, x than s_n/t_n with $t \leq t_n$ (although there might be an OCF convergent that is closer to x on the other side). When we arrive at $j = a_n$, we hit a new OCF convergent, on the current side of x, closer than the previous OCF convergent. These approximations can be visualized in terms of integer lattice points in the plane that are closer, in terms of the vertical distance, to the line of irrational slope x through the origin than others nearer the origin: see Figure 2.

Sinai [8] provides an explanation of OCF and Farey convergents in terms of rotations of the circle through an angle $2\pi x$, or addition by x mod 1 (for $x = [a_1, a_2, \ldots]$ irrational). Let I_0 denote the arc from 1 to $e^{2\pi i x}$ on the unit circle, or alternatively the subinterval $[0, x]$ of $[0, 1]$. Since $a_1 = \lfloor 1/x \rfloor$, we have $a_1 x \leq 1 < (a_1 + 1)x$, so that we can lay off

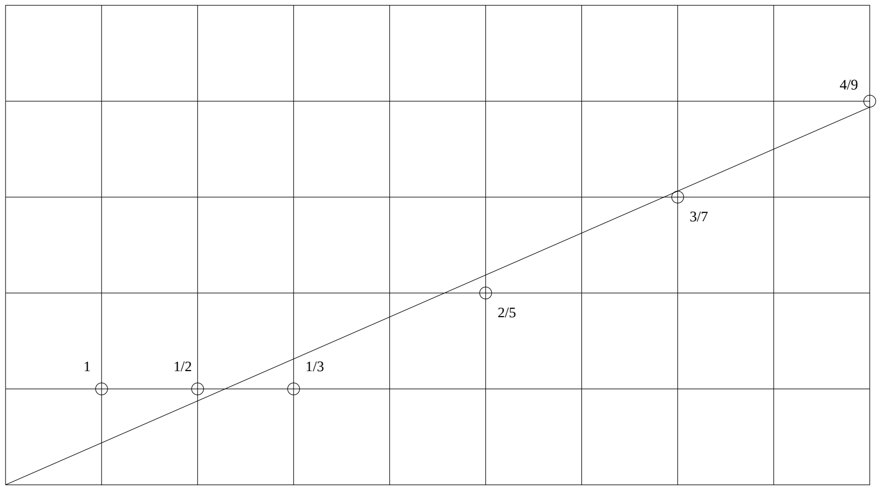

FIGURE 2. Visualization of OCF convergents.

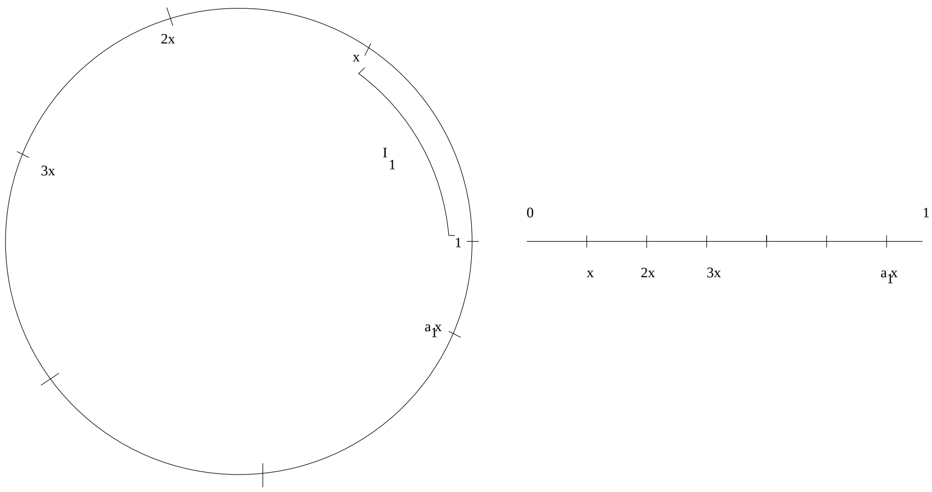

FIGURE 3. The laying off of arcs.

exactly a_1 arcs (intervals) of length $2\pi x$ (or x) before slopping over the initial point 1 (end): see Figure 3.

Observe now that

$$\begin{aligned}\frac{l(I_1)}{l(I_0)} &= \frac{1 - a_1 x}{x} \\ &= \frac{1 - \frac{a_1}{a_1 + [a_2, a_3, \ldots]}}{\frac{1}{a_1 + [a_2, a_3, \ldots]}} \\ &= [a_2, a_3, \ldots] = Tx \ .\end{aligned}$$

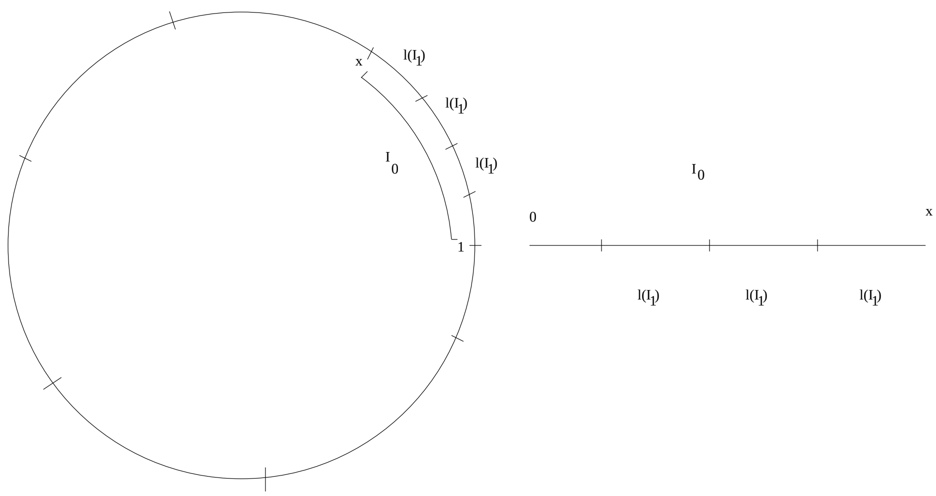

FIGURE 4. There are a_2 subintervals of length $l(I_1)$ inside I_0.

Consequently, if $kl(I_1) \leq l(I_0) < (k+1)l(I_1)$, then $k \leq 1/Tx < k+1$, so as above $k = \lfloor 1/Tx \rfloor = a_2$. Thus (in a kind of renormalization) when we begin at the right endpoint x of I_0, we can lay off exactly a_2 intervals of length $l(I_1)$ before slopping over the other end of I_0: see Figure 4.

The approximation $0 \approx x - a_2 l(I_1) = x - a_2(1 - a_1 x)$ is equivalent to the OCF approximation $x \approx a_2/(1 + a_1 a_2) = 1/(a_1 + 1/a_2)$. The extra division points which approximate 0, before switching sides, correspond to the Farey approximations to x.

10. Dynamics of Farey and Continued Fraction Expansions

Recall that we start with $p_{-1} = q_0 = 1$, $q_{-1} = p_0 = 0$, and

$$\begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (BA^{a_1-1}BA^{a_2-1}\dots BA^{a_{n-1}-1}),$$

where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. By $\frac{p_n}{q_n}$ we mean the partial fraction approximant $[a_1, \dots, a_n]$. If we write $x = [a_1, a_2, \dots]$, recall that the Gauss map T is given by $Tx = [a_2, a_3, \dots] = \frac{1}{x} \bmod 1$, which has finite ergodic invariant measure $d\mathrm{G} = \frac{1}{\log 2} \cdot \frac{dx}{1+x}$.

Label the edges of the Farey tree leaving each vertex by A and B as shown in Figure 5, with B on the edge toward the vertex with larger

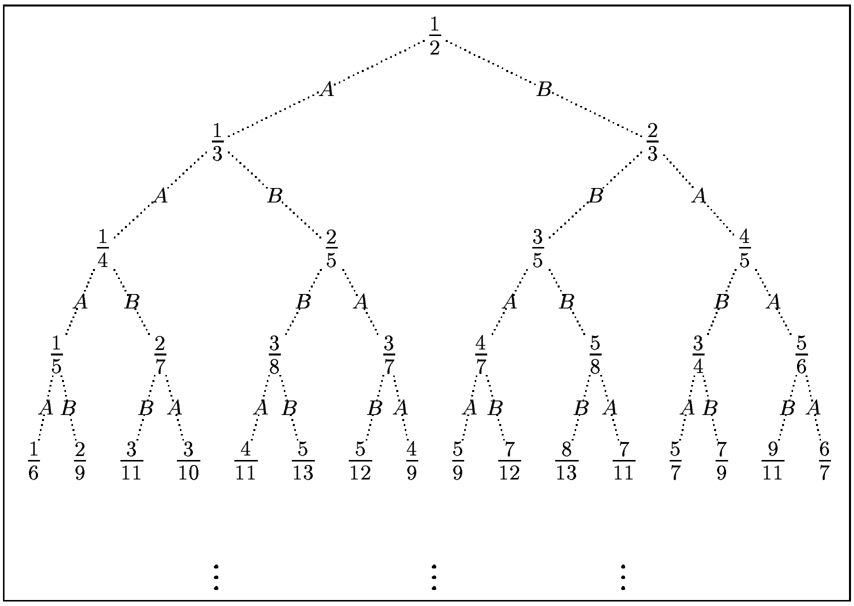

FIGURE 5. The top of the Farey tree, with A,B labels.

denominator. The partial products of the above matrix product, with the initial B suppressed, determine paths on the Farey tree. We agree that for rational x, the path will terminate with A^{∞}. As we descend on a path in the Farey tree (Figure 1), each decision we make in our descent will correspond to application of A or B. See Figure 5. For example, the path to $\dfrac{7}{16}$ and beyond begins with the label $ABAAB\ldots$. The computed matrix entries give the Farey—and occasionally OCF—approximations to the point $x = [a_1, a_2, \ldots]$.

The Gauss map T shifts the continued fraction expansion. What is the significance of shifting along the A, B labeling of the *entire* Farey path?

T corresponds to the shift to just after the next B. The shift on $\{A, B\}^{\mathbb{N}}$, called the *Farey shift*, corresponds to a certain map U : $[0, 1] \longrightarrow [0, 1]$.

We define U as follows:

$$U(x) = \begin{cases} \dfrac{x}{1-x}, & \text{if } 0 \leq x \leq \frac{1}{2} \\[2ex] \dfrac{1-x}{x}, & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

The Farey tree sets up the correspondence

$$\{A, B\}^{\mathbb{N}}, \sigma \longleftrightarrow [0, 1], U$$

via

$$A^{a_1-1}BA^{a_2-1}B\cdots \longleftrightarrow x=[a_1,a_2,\dots].$$

So if $x\in[\frac{1}{2},1]$ then $a_1=1$ and $x\sim BA^{a_2-1}B\dots$, i.e., for $x\in[\frac{1}{2},1]$ the first Farey label of x is B. So, the Gauss map T —our "shift past the next B"—corresponds to the *first return map* U^* to $[\frac{1}{2},1]$ of U. Gauss measure dG determines the *infinite* U-invariant measure dx/x: we view $([0,1],U)$ as a tower over $([\frac{1}{2},1],U^*)\approx([0,1],T)$.

So U^*, which corresponds to the shift to the next B in labeled Farey expansions, is analytically conjugate via

$$\phi(x)=\frac{1-x}{x}=U\Big|_{[\frac{1}{2},1]}$$

to T=Gauss map on $[0,1]\simeq$ shift past the next B in the labeled Farey tree path.

The dynamical viewpoint leads to results about the growth rates of Farey as well as OCF denominators. Recall that

$$\frac{a_1+a_2+\cdots+a_n}{n}\longrightarrow\infty\qquad \text{a.e. } d\text{G}$$

and that

$$\frac{\log q_n}{n}\longrightarrow\frac{\pi^2}{12\log 2}\qquad \text{a.e. } d\text{G}.$$

A more precise result due to Khintchine and Lévy is

$$\frac{a_1+a_2+\cdots+a_n}{n\log n}\longrightarrow\frac{1}{\log 2}\qquad \text{in measure.}$$

Diamond and Vaaler showed that if the largest a_k among $\{a_1,\dots,a_n\}$ is set to 0, then

$$\frac{a_1+a_2+\cdots+a_n}{n\log n}\longrightarrow\frac{1}{\log 2}\qquad a.e.\, d\text{G}.$$

If the n'th Farey approximant of x is s_n/t_n, then it is a result of Lagarias that

$$\frac{\log t_n}{n}\sim\frac{\pi^2}{12\log n}\ \text{in measure.}$$

Since $\Omega=\{A,B\}^{\mathbb{N}}$ has many natural finite shift-invariant measures, these will carry over to U-invariant measures on the interval $[0,1]$. For example, consider the Bernoulli measure $B(\frac{1}{2},\frac{1}{2})$ on Ω. This carries over to the *Minkowski measure* $d?$, where $?:[0,1]\longrightarrow[0,1]$ is the cumulative distribution function $?(x)=d?[0,x]$. Minkowski showed that ? maps the rational and quadratic irrationals in $[0,1]$ in a one-to-one order-preserving fashion onto the rationals in $[0,1]$. Ma studied growth rates of OCF, Farey, and f-expansion digits and denominators for points typical with respect to d? and the images of other measures on Ω.

Example (due to Ma). Start with $B(p, 1-p)$ on Ω, carry over to μ_p on $[0,1]$. Then

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \longrightarrow \frac{1}{1-p} \quad \text{a.e. } d\mu_p,$$

$$\frac{\log q_n}{n} \longrightarrow \lambda_p = -\int_0^1 \log x \, d\mu_p,$$

and (for a known constant c)

$$\frac{\log t_n}{n} \longrightarrow \frac{\lambda_p}{c} \quad \text{a.e. } d\mu_p.$$

Similar results hold for Markov measures.

11. Continued fractions and geodesic flows

One might guess a connection of continued fraction dynamics with the modular group, geodesic flow in hyperbolic space, etc., from the following observations:

1. Two points $x, y \in [0,1]$ have a *common tail* for their continued fraction expansions—i.e., there exist $m, n \geq 1$ such that $a_{m+j}(x) = a_{n+j}(y)$ for all $j \geq 0$—if and only if there exist four integers a, b, c, d, with $ad - bc = \pm 1$ and $x = \dfrac{ay+b}{cy+d}$. (This is equivalent to the existence of m and n with $T^{m-1}x = T^{n-1}y$.)
2. Suppose $w, w' \in \mathbb{C}$ generate a *lattice*

$$\mathcal{L} = \{mw + nw' : m, n \in \mathbb{Z}\},$$

i.e. there do not exist finite limit points of $\mathcal{L}$. Then, (θ, θ') is another pair that generates $\mathcal{L}$ if and only if there exist four integers a, b, c, d, with $ad - bc = \pm 1$ and

$$\begin{pmatrix} \theta \\ \theta' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w \\ w' \end{pmatrix}.$$

Example. $(\theta, \theta') = (w, 2w + w')$ (see Figure 6).

Definition 14 (Elliptic functions)**.** Meromorphic functions f which have two periods w and w' ($f(z+w) = f(z+w') = f(z)$ for every z in $\mathbb{C}$) are called *elliptic functions.*

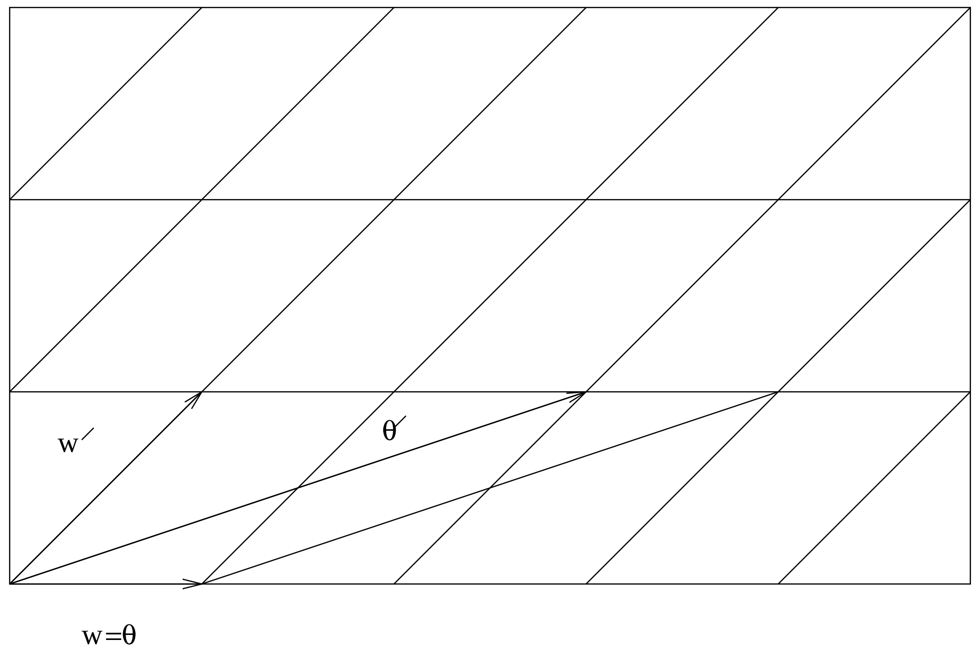

FIGURE 6. Two generators for one lattice

Definition 15 (Modular group)**.** The *modular group* Γ is the group of all linear fractional transformations $\mathbb{C} \longrightarrow \mathbb{C}$ which can be represented as $\dfrac{az+b}{cz+d}$ with a, b, c, $d \in \mathbb{Z}$, and $ad - bc = 1$.

$SL(2,\mathbb{Z})$ is the group of 2×2 matrices with integer entries and determinant 1. Note that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \simeq \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{as transformations,}$$

so that the modular group is ismorphic to $SL(2,\mathbb{Z})/\{\pm I\}$.

Definition 16 (Modular function)**.** A meromorphic function f on the upper half plane $\mathbb{H}$ which is invariant under the modular group—that is, $f(\gamma z) = f(z)$ for every $\gamma \in \mathrm{SL}(2,\mathbb{Z})$—is called a *modular function.*

Theorem 17. *The modular functions are the rational functions of a* single *modular function, namely, the Dedekind-Klein modular function J, which maps a fundamental region $\Omega = \Omega_1 \cup \Omega_2$ where*

$$\Omega_1 = \{z \in \mathbb{C} \,:\, Im\ z > 0, -\frac{1}{2} \leq Re\ z \leq 0, |z| \geq 1\}$$

and

$$\Omega_2 = \{z \in \mathbb{C} \,:\, Im\ z > 0, 0 \leq Re\ z < \frac{1}{2}, |z| > 1\}$$

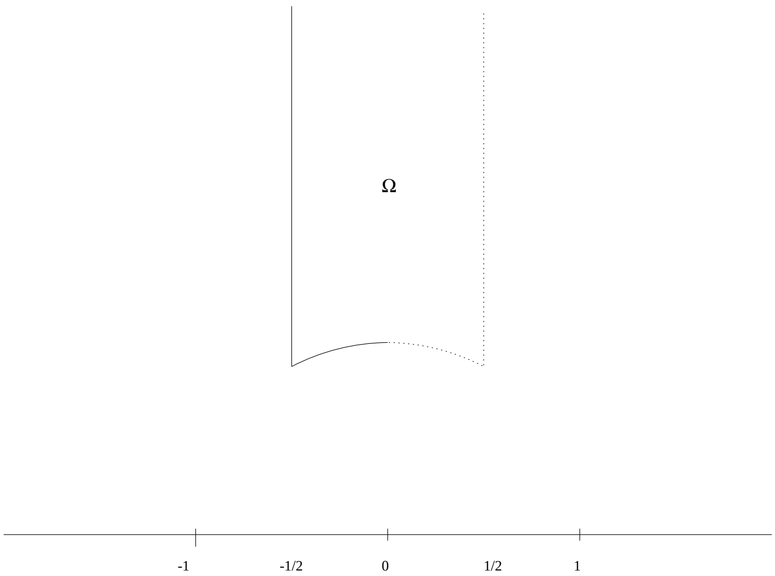

FIGURE 7. The Fundamental region of Theorem 17

one-to-one onto $\mathbb{C}$. (See [6], p. 390ff).

Every complete surface of constant negative curvature is isometric to a quotient of the upper half plane $\mathbb{H}$, with its hyperbolic metric $ds^2 = \dfrac{dx^2 + dy^2}{y^2}$, by a discrete subgroup of isometries. The quotient of $\mathbb{H}$ by the modular group Γ is called the *modular surface*; it can be pictured as a torus with one point removed, dragged to ∞. Geodesics in $\mathbb{H}$ are semicircles orthogonal to the real axis and vertical lines. The connection between the Gauss map (shift on continued fraction expansions) and the geodesic flow on $\mathbb{H}/\Gamma$ has been examined by E. Artin and, more recently, C. Series.

Series constructs the *Farey tessellation* $\mathfrak{F}$ of $\mathbb{H}$: the geodesic triangle τ with vertices at 0, 1, and ∞ and all of its images under Γ. See Figure 8. The real endpoints of the semicircles are exactly the pairs of Farey neighbors. The basic triangle τ is formed from the fundamental region Ω for J as follows. The half of Ω in the left half plane is translated one unit to the right and adjoined to the right half, obtaining a region R. Letting

$$M = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \in \Gamma,$$

we obtain

$$\tau = R \cup MR \cup M^2 R.$$

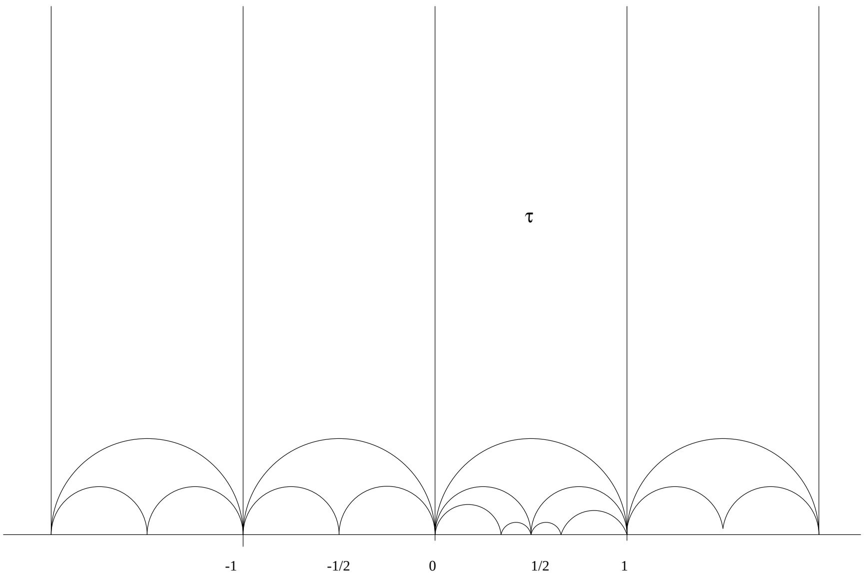

FIGURE 8. The geodesic triangle τ and its images

Given $x \in [0,1]$, choose $x' < 0$ and let γ denote the oriented geodesic in $\mathbb{H}$ from x to x'. The edges of the triangles in $\mathfrak{F}$ divide γ into segments. Label the segments by L or R depending on whether there are one or two vertices of the triangle being traversed to the left of the geodesic. The labeling so obtained looks like

$$\dots R^{n_{-1}}.L^{n_0}R^{n_1}\dots,$$

the . denoting the crossing of the imaginary axis. Then $x = [n_0; n_1, n_2, \dots]$!! Note that $n_0 = 0$ if $x \in [0,1]$ and that varying the x' only changes the left half of this L,R coding of γ.

Let S denote the orbit under Γ of the set of unit tangent vectors with base points on the imaginary axis whose codings change from L to R or vice versa at that point. Then Series shows the following:

Theorem 18. *The Gauss map (shift on continued fraction expansions) is almost topologically conjugate (up to a countable set) to the Poincaré map of the geodesic flow on* $\mathbb{H}/\Gamma$ *for the Poincaré section* S.

The geometrical interpretation allows Series to display Gauss measure on $[0,1]$ as the natural dynamical measure coming from the invariant hyperbolic measure for the geodesic flow on $\mathbb{H}$. Ergodicity and other dynamical properties of the Gauss map T are also consequences.

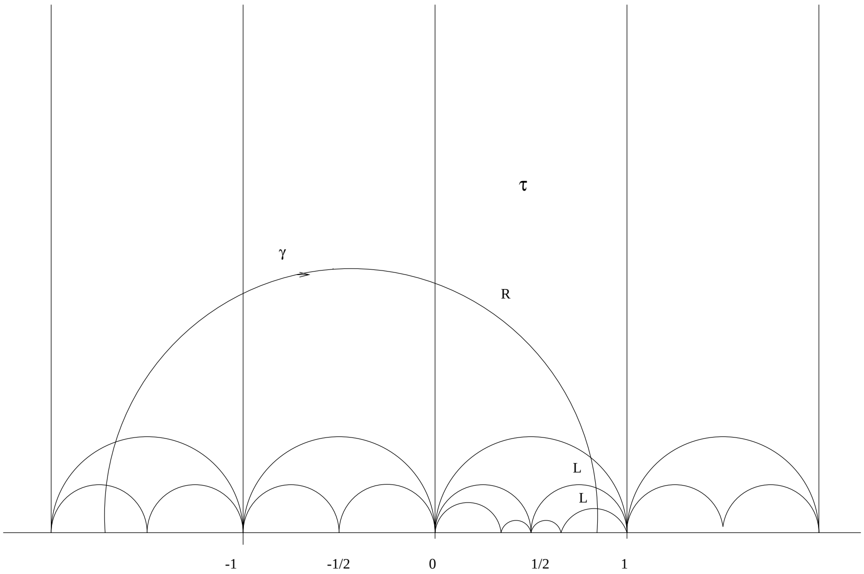

FIGURE 9. A labeled geodesic

REFERENCES

[1] Beiler, A., *Recreations in the Theory of Numbers*, Dover Publications, New York, 1966.
[2] Billingsley, P., *Ergodic Theory and Information*, J. Wiley & Sons, New York, 1965.
[3] Lagarias, J., The Farey shift, preprint 1991.
[4] ———, Number theory and dynamical systems, Proc. Symp. Appl. Math. **46**(1992), 35–72.
[5] Ma, H., *Growth Rates of Digits and Denominators for a Class of f-Expansions*, Ph.D. Dissertation, University of North Carolina at Chapel Hill, 1995.
[6] Saks, S. and Zygmund, A., *Analytic Functions*, Elsevier, Amsterdam, 1971.
[7] Series, C., The modular surface and continued fractions, J. London Math. Soc. **31**(1985), 69–80.
[8] Sinai, Ya. G., *Topics in Ergodic Theory*, Princeton Univ. Press, 1994.

June 1, 2000.