AN ERGODIC ANALYSIS OF CONTINUED FRACTIONS

ADAM QUINN JAFFE

These are my notes on applications of ergodic theory to the continued fraction expansion of a real number, which was the final project of a ten-week directed reading program at Stanford. I was advised by Jimmy He, and we mainly read Einsiedler and Ward's *Ergodic Theory with a view towards Number Theory.* The full program covered far more material than what is written, and this document only covers the topics that I presented at our final colloquium.

1. Introduction

Ergodic theory focuses on studying the long-term bevaior of dynamical systems, and was first motivated by statistical mechanics. An illustrative and historical example to keep in mind is the following:

Example 1.1. Suppose that a large number of particles are allowed to move freely inside a box, and let their time-discrete evolution be described by a function T. Now let f be any function whose domain is the box. For any particle x, we may study its trajectory $\{T^i(x)\}_{i=1}^{\infty}$ in the sense of the time *time average*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(T^i(x))$$

or in the sense of the *space average*

$$\frac{1}{\mathrm{Vol(Box)}}\int_{\mathrm{Box}} f(x)dx.$$

A motivating question of ergodic theory asks when these two quantities agree.

Modern research in ergodic theory is abstracted to a more mathematical setting. This development has allowed it to find applications to susprising fields like number theory, combinatorics, and ramsey theory.

In this talk we will introduce the basic theory of ergodic theory and highlight a classical result. Then we will explore applications of this result to number theory, specifically we will prove that the set of all badly approximable numbers has Lebesuge measure zero.

As ergodic theory is closely related to probability theory, measure theory, and functional analysis, there is a bit of background required to understand it in full detail. However, to make this talk more amenable to a general audience, I may be informal with some of these technical details.

2. Ergodic Theory

Let us define the basic objects of our study.

Definition 2.1. Let $(X, \mathcal{F}, \mu)$ be a probability space and let $T : X \to X$ be a map satisfying $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{F}$. Then we call the tuple $(X, \mathcal{F}, \mu, T)$ a *measure-preserving system.*

Examples of measure preserving systems are abundant.

Example 2.2. Let $X = \{0,1\}^{\mathbb{N}}$ be the set of all infinite binary strings, endowed with the product measure $\mu = \prod_{i=1}^{\infty} \tilde{\mu}$ for $\tilde{\mu}(\{1\}) = \tilde{\mu}(\{0\}) = \frac{1}{2}$. Also consider the shift map $\sigma : X \to X$ defined by $\sigma(x_1, x_2, \dots) = (x_2, x_3, \dots)$. Then $(X, \mathcal{B}(X), \mu, \sigma)$ is a measure-preserving system.

Example 2.3. Consider $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ with the usual Lebesgue (probability) measure λ. For any $\rho \in \mathbb{R}$ define the map $T_\rho(x) = x + \rho$. Then $(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1), \lambda, T_\rho)$ is a measure-preserving system.

The next definition identifies a specific class of measure-preserving systems which will be of great interest to us.

Definition 2.4. A measure preserving system $(X, \mathcal{F}, \mu, T)$ is called *ergodic* if any $A \in \mathcal{F}$ with $T^{-1}(A) = A$ has either $\mu(A) = 0$ or $\mu(A) = 1$.

In this language, we see that Example 2.2 is ergodic, and that Example 2.3 is ergodic if and only if ρ is irrational.

Finally, we can state a classical and powerful result:

Theorem 2.5 (Birkhoff)**.** Suppose that $(X, \mathcal{F}, \mu, T)$ is a measure preserving system. Then for any function $f \in L^1(X, \mathcal{F}, \mu)$, there exists a set $A \in \mathcal{F}$ with $\mu(A) = 0$ such that any $x \in X \setminus A$ gives

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)).$$

Moreover, if $(X, \mathcal{F}, \mu, T)$ is ergodic, then the limit is exactly

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_X f d\mu.$$

That is, the convergent does not depend on x.

In the next section we will review the theory of continued fractions so that we may later apply Birkhoff's theorem.

3. Continued Fractions

Given an irrational α, one might like to ask: Which rational numbers are the "best" approximation to α? It turns out that continued fractions are a convenient way to make this precise, as we will later see.

But first, what are continued fractions? Informally, this theory will allow us to write irrational numbers as a limit of rational numbers in the following canonical way:

$$\alpha = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \dots}}.$$

Naturally, we may truncate this to a finite number of terms will give a rational number which approximates α. This expansion consists of finitely many terms if and only if α is rational, and the theory is generally uninteresting in this case. We therefore focus our attention to the case of α irrational. We also lose no generality in restricting attention to the case of $\mathbb{S}^1$, or equivalently, $a_0 = 0$.

Making this precise is the following:

Definition 3.1. Define the map $T(x) = \frac{1}{x} \mod 1$. Then for any $\alpha \in \mathbb{S}^1 \setminus \mathbb{Q}$, we may generate the *continued fraction coefficients* $\{a_n\}_{n=1}^{\infty}$ according to

$$a_n = \left\lfloor \frac{1}{T^{n-1}(x)} \right\rfloor.$$

Then we generate the *continued fraction convergents* $\{p_n/q_n\}_{n=1}^{\infty}$ according to

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \dots \frac{1}{a_n}}.$$

A famous result of Dirichlet is the following:

Theorem 3.2 (Dirichlet)**.** Let $\{p_n/q_n\}_{n=1}^{\infty}$ be the continued fraction convergents of $\alpha \in \mathbb{S}^1 \setminus \mathbb{Q}$. Then we have the bound:

$$\left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2 \sqrt{5}}$$

The question naturally arises if Dirichlet's bound can be improved. It turns out that it cannot; it is sharp, for example, for $\alpha = \frac{1+\sqrt{5}}{2}$. Identifying the numbers for which Theorem 3.2 is sharp leads us to the following definition:

Definition 3.3. An irrational number $\alpha \in \mathbb{S}^1 \setminus \mathbb{Q}$ is called *badly approximable* if there exists some $\varepsilon \in (0, 1/\sqrt{5})$ such that

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{\varepsilon}{q^2}$$

holds for all rationals p/q.

To see that many badly approximable numbers exist, we refer to a famous result of Lagrange:

Theorem 3.4 (Lagrange)**.** Quadratic irrationals are badly approximable. That is, if $\alpha \in \mathbb{S}^1 \setminus \mathbb{Q}$ is a solution to $ax^2 + bx + c = 0$ for $a, b, c \in \mathbb{Z}$, then α is badly approximable.

However, Theorem 3.4 only guarantees the existence of countably many badly approximable numbers. In fact, there are at least uncountably many of them! To see this, we will need the following characterization, which is not too hard to prove:

Lemma 3.5. An irrational number $\alpha \in \mathbb{S}^1 \setminus \mathbb{Q}$ is badly approximable if and only if its continued fraction coefficients $\{a_n\}_{n=1}^{\infty}$ form a bounded sequence.

Intuitively, Lemma 3.5 holds because the boundedness of $\{a_n\}_{n=1}^\infty$ implies that $\{p_n/q_n\}_{n=1}^\infty$ converges to α very slowly, and vice versa. But now it is clear that any bounded sequence of nonnegative integers gives rise to a badly approximable number, so there at least uncountably many of them. It is not known whether any algebraic numbers besides quadratic irrationals are badly approximable.

However, we can study more about the set of all badly approximable numbers than just its cardinality. We will see via ergodic theory in the next section that the set of badly approximable numbers forms a measure zero set with respect to the Lebesgue measure.

4. Applications

Recall from the previous section that we defined the map $T : X \to X$ for $X = \mathbb{S}^1 \setminus \mathbb{Q}$. Let us write $\mathcal{F} = \mathcal{B}(\mathbb{S}^1 \setminus \mathbb{Q})$ and λ for the Lebesgue measure on $\mathbb{S}^1$. Unfortunately, the system $(X, \mathcal{F}, \lambda, T)$ is not ergodic, nor is it measure-preserving. However, we can define a different measure which is better behaved with respect to the map T, which we call the *Gauss measure* μ on $\mathbb{S}^1$:

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{1}{1+x} dx.$$

We observe that λ and μ are *equivalent* in the sense that they agree on which sets in X have measure zero and measure one. In other words, they "smoothly" distort the space of X. So, since $(X, \mathcal{F}, \mu, T)$ is an ergodic measure-preserving system, we can study it as a surrogate for $(X, \mathcal{F}, \lambda, T)$. We omit the proof that $(X, \mathcal{F}, \mu, T)$ is measure-preserving and ergodic for brevity.

Now we state and prove our first interesting result:

Theorem 4.1. There exists a set $A \in \mathcal{F}$ with $\mu(A) = 0$ such that, for all $x \in X \setminus A$, the continued fraction coefficients $\{a_n\}_{n=1}^\infty$ of x satisfy:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} a_i = \infty.$$

Proof. Define the function $g : X \to \mathbb{R}$ by $g(x) = a$ for $x \in (\frac{1}{a+1}, \frac{1}{a})$, and observe that $g(T^i(x)) = a_{n+1}$. Then for each positive integer N, we define

$$g_N(x) = \begin{cases} g(x) & \text{if } g(x) \le N \\ 0 & \text{else} \end{cases}.$$

Note that $g_N(x) \le g(x)$ holds for all $x \in X$, and that we can compute:

$$\int g_N d\mu = \frac{1}{\log 2} \sum_{a=1}^{N} \int_{\frac{1}{a+1}}^{\frac{1}{a}} a dx = \frac{1}{\log 2} \sum_{a=1}^{N} \frac{1}{a+1}.$$

This proves $\int g_N d\mu < \infty$, $\int g d\mu = \infty$, and $\int g_N d\mu \to \infty$. Now that we have $g_N \in L^1(X, \mathcal{F}, \mu)$, we can use Theorem 2.5 to get the bound:

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) &\geq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} g_N(T^i(x))\\ &= \int g_N d\mu\end{aligned}$$

for any N. Then take $N \to \infty$ to get

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} a_i = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} g(T^i(x)) = \infty,$$

as desired.

□

Corollary 4.2. The set B of all badly approximable numbers in $\mathbb{S}^1$ has $\lambda(B) = 0$.

Proof. Take any $x \in B$ and denote its continued fraction coefficients by $\{a_n\}_{n=1}^{\infty}$. By Lemma 3.5, there is some $M > 0$ such that $a_n \leq M$ holds for all n, so $\frac{1}{n}\sum_{i=1}^{n} a_i \leq M$ holds for all n. Therefore, the conclusion of Theorem 4.1 fails, so we must have $x \in A$. Hence $B \subseteq A$. But $\mu(A) = 0$ implies $\mu(B) = 0$, and $\mu(B) = 0$ implies $\lambda(B) = 0$, which completes the proof. □