CONTINUED FRACTIONS AND THE ERGODIC THEOREM

MICHAEL BRADFORD WILLIAMS

1. Introduction

"What is your favorite number?" Popular answers to this questions might include

$$0, 1, 2, e, \pi, \gamma, 17, 42, \phi, \text{etc.}$$

There are many criteria for why a number might be interesting or important. For example, it could satisfy a simple and elegant relation, it could have relevance to many different areas of mathematics (or beyond), or it could have properties that are fundamentally important. (Or it could satisfy all three!) We will see a number that is interesting primarily because of the utter *strangeness* of a property it possesses.

2. Continued Fractions

Put the previous question aside for the momen as we discuss objects of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_2}{a_3 + \cfrac{b_3}{a_4 + \dots}}}}$$

These are *continued fractions*. For us, $a_i, b_i \in \mathbb{R}$, although they could be from any set where it makes sense to divide elements, e.g., $\mathbb{C}, \mathbb{R}(x), \mathcal{O}^*_{\mathbb{R}}$, etc.

From a purely aesthetic standpoint, continued fractions are nice because many numbers have interesting expressions as continued fractions. Here are some examples:

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{\dots}}}}} \qquad \sqrt{3} = 1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{\dots}}}}}$$

$$\pi = \cfrac{4}{1 + \cfrac{1^2}{3 + \cfrac{2^2}{5 + \cfrac{3^2}{7 + \cfrac{4^2}{\dots}}}}} \qquad \phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\dots}}}}}$$

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{\dots}}}}}} \qquad \sqrt{\frac{\pi e}{2}} = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{2}{1 + \cfrac{3}{1 + \cfrac{4}{\dots}}}}}$$

Wow!

From now on, we will consider continued fractions where $b_i = 1$ for all i, and $a_i \in \mathbb{R}$. These continued fractions look like

$$r = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}}$$

and nice notation for this is

$$r = [a_0; a_1, a_2, \dots].$$

For example,

$$\sqrt{2} = [1; 2, 2, 2, \dots], \qquad \sqrt{3} = [1; 1, 2, 1, 2, \dots]$$
$$\phi = [1; 1, 1, 1, \dots], \qquad e = [2; 1, 2, 1, 1, 4, 1, 1, 6, \dots].$$

Question 2.1. Given a sequence $\{a_i\} \subset \mathbb{R}$, when is $[a_0; a_1, a_2, \dots]$ a real number?

Clearly, if the sequence $\{a_i\}$ is finite, then $[a_0; a_1, a_2, \dots, a_k] \in \mathbb{R}$. But if it is infinite, then there are convergence issues. Luckily, there is a rather nice theorem.

Theorem 2.2. *Let $\{a_i\}$ be a sequence of real numbers. Then*

$$[a_0; a_1, a_2, \dots] \in \mathbb{R} \quad \Longleftrightarrow \quad \sum_{i=1}^{\infty} a_i \text{ diverges.}$$

The proof is elementary, but long for the purposes of this note, so we won't give it here.

In all of our examples, the numbers involved were *positive integers.* By the previous theorem, all such continued fractions converge. We can also ask about the converse of this question.

Question 2.3. Can all real numbers be written as a continued fraction with positive integers?

The answer is yes! Here's how. Given $x \in (0, 1) \setminus \mathbb{Q}$, we want

$$x = [a_0; a_1, a_2, \dots], \quad a_i \in \mathbb{Z}^+.$$

Since $0 < x < 1$, we must have $a_0 = 0$. To find a_1, we have

$$x = \frac{1}{a_1 + \Box} \quad \Longrightarrow \quad \frac{1}{x} = a_1 + \Box,$$

and $\square$ is something less than one (it is a fraction with one in the numerator and denominator larger than one). This means we can round down:

$$a_1 = \left\lfloor \frac{1}{x} \right\rfloor .$$

Next,

$$\begin{aligned} x = \frac{1}{a_1 + \frac{1}{a_2 + \square}} &\implies \frac{1}{x} = a_1 + \frac{1}{a_2 + \square}, \\ &\implies \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor = \frac{1}{a_2 + \square}, \\ &\implies \frac{1}{\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor} = a_2 + \square. \\ &\implies a_2 = \left\lfloor \frac{1}{\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor} \right\rfloor . \end{aligned}$$

Again $\square < 1$, so we rounded down. We can repeat this procedure to find that

$$\begin{aligned} a_0(x) &= 0 \\ a_1(x) &= \left\lfloor \frac{1}{x} \right\rfloor \\ a_2(x) &= \left\lfloor \frac{1}{\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor} \right\rfloor \\ a_3(x) &= \left\lfloor \frac{1}{\frac{1}{\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor} - \left\lfloor \frac{1}{\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor} \right\rfloor} \right\rfloor \end{aligned}$$

What's the pattern? Let $T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$, then we can write

$$\begin{aligned} a_2(x) &= a_1(T(x)) \\ a_3(x) &= a_2(T(x)) = a_1(T^2(x)) \\ &\vdots \\ a_{k+1}(x) &= a_1(T^k(x)) \end{aligned}$$

where

$$T^k(x) = \underbrace{T(T(\dots(T(x))\dots))}_{k \text{ times}} .$$

This gives an algorithm for computing the continued fraction expression of the decimal part of any irrational number, and the expression is unique. If it is rational, this procedure would terminate when we get a zero in a denominator. In this case, the expression is not quite unique –there are two choices:

$$[a_0; a_1, \dots, a_{n-1}, a_n] = [a_0; a_1, \dots, a_{n-1}, a_n - 1, 1].$$

It is convention to go with the first, shorter form.

There are several reasons why continued fraction are useful for representing real numbers. First let's consider deficiencies of the decimal system:

- The base 10 is arbitrary, and if we change base, so do the digits in the representations of numbers:
$$\left(\frac{1}{4}\right)_{10} = 0.25, \left(\frac{1}{7}\right)_{10} = 0.\overline{142857}, \quad \left(\frac{1}{4}\right)_{7} = 0.\overline{15}, \left(\frac{1}{7}\right)_{10} = 0.1.$$
- Many rationals do not have finite decimal representations, e,g., $1/3 = 0.\overline{3}$.
- There can be ambiguity in notation, e.g., $0.\overline{9} = 1$.

The advantages of continued fraction representations are many. Here are a few.

- Continued fractions are *base neutral.* The whole numbers $\{a_i\}$ don't depend on the base, regardless of what base is used to represent them.
- Rational numbers always have finite continued fraction representations.
- Continued fractions make it easier to see structure/patterns in irrational (and even transcendental) numbers.
- There is no ambiguity in notation: numbers have unique continued fraction representations.

3. Detour into Measure Theory

Suppose that we have some coins (which we assume have size proportional to their value) lined up in a row on a table. We want to find the total dollar amount. In other words, we have a function and we want to compute its integral.

Riemann integration computes this value as follows. Go from left to right, multiply each coin by its value, and sum up all the values:
$$V = \sum_{\text{coins } c} \text{value}(c) \cdot 1.$$
This seems rather laborious, but there is another way.

Sort coins according to denomination, multiply the value of each denomination by the # of coins in that denomination. Sum over the denominations:
$$V = \sum_{\text{denominations } D} \text{value}(D) \cdot \#\{c \in D\}.$$

This is the basic idea of the Lebesque integral. This type of integration requies a way to determine the "size" of sets –in other words, a way to count (or measure) the number of coins per denomination. For us, a *measure* μ is a function
$$\mu : \{\text{subsets of } \mathbb{R}\} \longrightarrow [0, \infty]$$
that satisfies some nice properties, such as

- $U \subseteq V \Rightarrow \mu(U) \leq \mu(V)$
- $U = \coprod U_i \Rightarrow \mu(U) = \sum \mu(U_i)$
- $\mu(U) = \mu(U + h) \quad \forall h \in \mathbb{R}$.

An example of a measure is the *Lebesgue measure* on $\mathbb{R}$, define on intervals by
$$m([a, b]) = b - a.$$

How do we integrate? Let (X, μ) be a set with a measure. Start with a piecewise-constant function
$$s(x) = \sum c_i \chi_{U_i}(x),$$

and say that

$$\int_X s\,d\mu = \sum c_i \mu(U_i).$$

This is exactly what we did with the coins. If f is a non-negative "measurable" function on X, its integral is

$$\int_X f\,d\mu = \sup_{\substack{0\le s\le f \\ s \text{ simple}}} \int_X s\,d\mu.$$

This can be easily extended to functions with positive and negative values.

The Lebesgue integral agrees with the Riemann integral for all functions that are Riemann integrable (e.g., bounded, "almost everywhere" continuous functions). There do exist functions that are Lebesque integrable, but not Riemann integrable, functions that are not integrable in either sense.

A very important idea in measure theory is that if a set has *measure zero*, i.e.,

$$\mu(U) = 0,$$

then it is "too small to matter." For example, any countable set in $\mathbb{R}$ has Lebesgue measure zero, but there are also some uncountable ones, such as the Cantor set. These sets are too small to matter because you can change a function on a set of measure 0, and the integral doesn't see the difference!

For example, the function

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 1 & \text{if } x \notin \mathbb{Q} \cap [0,1] \end{cases}$$

is famously not integrable in the Riemann sense. But, we see that this function differs from the constant function $f(x) \equiv 1$ only on a set of measure 0. Therefore

$$\int_{[0,1]} f(x)\,dm = \int_{[0,1]} dm = 1 - 0 = 1.$$

So, the Lebesgue integral is more general, and we can integrate functions over weird domains –but not too weird, there do exists non-measureable sets.

4. The Ergodic Theorem

Let (X,μ) be a subset of $\mathbb{R}$ with a measure. A function $\tau : X \to X$ is *ergodic* if

- τ preserves measure: $\mu(\tau^{-1}(U)) = \tau(U)$;
- τ is "invariant": if U and $\tau^{-1}(U)$ differ by a set of measure 0, then either $\mu(U) = 0$ or $\mu(X \setminus U) = 0$.

This says that there is no stretching or compressing. Note that τ may not be continuous. The word "ergodic" comes from the Greek words for "work" and "path."

For example, the circle $S^1 = \mathbb{R}/\mathbb{Z}$ has a measure induced by m on $\mathbb{R}$. Let $\alpha \in \mathbb{R}$ be irrational. The function

$$\tau(x) = x + \alpha \pmod 1$$

is ergodic.

A main result concerning such functions is the following.

Theorem 4.1 (The Ergodic Theorem). *Let (X, μ) be such that $\mu(X) = 1$ and let $\tau : X \to X$ be ergodic. If $f : X \to \mathbb{R}$ is integrable, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)) \overset{a.e.}{=} \int_X f\, d\mu,$$

that is, almost everywhere (except perhaps on a set of measure 0).

Thinking of τ as some operation repeated over time, the theorem says "the time average of f equals the space average of f." Ergodic theory is the study of dynamical systems (e.g., a space with a function that is repeatedly applied) with properties like the one in the theorem. It was developed in part by Boltzmann to address problems in statistical mechanics.

5. Back to Continued Fractions

Let

$$X = (0, 1), \quad \mu(U) = \frac{1}{\log 2} \int_U \frac{dm}{1 + x},$$

then μ is a measure on X, $\mu(X) = 1$, and μ has the same sets of measure 0 as m.

Let $T : X \to X$ be given by

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

as before. One can show that this function is ergodic with respect to μ. (This is difficult.) Since any $x \in (0, 1)$ can be written as a unique (essentially) continued fraction $x = [0; a_1, a_2, \dots]$ with $a_i \in \mathbb{Z}$, let

$$f(x) = \log a_1(x).$$

Now let's apply the Ergodic theorem!

$$\frac{1}{\log 2} \int_{(0,1)} \frac{\log a_1(x)}{1 + x}\, dm \overset{a.e.}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log a_1(T^k(x))$$

$$\overset{a.e.}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log a_{k+1}(x).$$

Now exponentiate:

$$\lim_{n\to\infty} \big(a_1(x) \dots a_n(x)\big)^{1/n}(x) \overset{a.e.}{=} \exp\left(\frac{1}{\log 2} \int_{(0,1)} \frac{\log\lfloor \frac{1}{x} \rfloor}{1 + x}\, dm \right).$$

The left side is exp of a definite integral, so it is a number!

Conclusion: Take virtually any real number $x \in (0, 1)$ and write it (uniquely) as a continued fraction, $x = [0; a_1, a_2, \dots]$. Then

$$\lim_{n\to\infty} \big(a_1(x) \cdots a_n(x)\big)^{1/n} \overset{a.e.}{=} K.$$

The left sides is a limit of geometric means, which depends on x, and the right side is a number, independent of x! This number is called *Khinchin's Constant*, and

$$K \cong 2.68454\dots.$$

Here are some facts about Khinchin's Constant:

- Almost every number has this property, but some numbers that *do not* have this property include rationals, roots of quadratics with rational coefficients, ϕ, and e.
- No number has been *proven* to actually have this property!
- There is numerical evidence that π, γ, and K itself do have this property.

"What is your favorite number now?"