REAL ANALYSIS
TOPIC 33A - SIMPLE CONTINUED FRACTIONS (DRAFT)

PAUL L. BAILEY

Abstract. We seek an example of a perfect set consisting only of irrationals. To this end, we would like to understand and verify the claim that “the set of infinite continued fractions using only ones and twos” is such a set.

1. Continued Fractions

1.1. Basis Definitions.

Definition 1. A *continued fraction* is an expression of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \ddots}}},$$

where $a_0 \in \mathbb{Z}$, and (a_n) and (b_n) are sequences of positive real numbers, for $n \geq 1$.

A continued fraction is *simple* if $b_n = 1$ for all n. We are only interested in simple continued fractions.

We admit the possibility of finite sequences. The *length* of the sequence (a_n) is the largest $N \in \mathbb{N}$ such that a_N exists. Sequences of length zero correspond to integers.

A continued fraction is *finite* (or *terminating*) if (a_n) is a finite sequence. Otherwise, it is *infinite* (or *nonterminating*).

A continued fraction represents a real number. This is clear for finite continued fractions; for example,

$$2 + \cfrac{1}{3 + \cfrac{1}{5 + \cfrac{1}{7}}} = 2 + \cfrac{1}{3 + \cfrac{7}{36}} = 2 + \frac{36}{115} = \frac{266}{115}.$$

We see that a finite continued fraction will always represents a rational number $x \in \mathbb{Q}$ such that $a_0 \leq x < a_0 + 1$. In turns out that infinite continued fractions always converge to some real number x, in a manner that we will make precise as we proceed.

Date: March 25, 2019.

We write the number x as given by this sequence with a bracket, followed by an integer, followed by a semicolon, followed by a sequence of integers separated by commas:

$$x = [a_0; a_1, a_2, a_3, \dots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}.$$

A finite sequence may be readily evaluated; for example,

$$[2; 1, 2, 3] = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{3}}} = 2.7.$$

Finite continued fractions may be written in two ways:

- $[a_0] = [a_0 - 1; 1]$;
- $[a_0; a_1, a_2, \dots, a_{n-1}, a_n] = [a_0; a_1, a_2, \dots, a_{n-1}, a_n - 1, 1]$.

For example,

$$[2; 1, 2, 3] = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{3}}} = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{1}}}} = [2; 1, 2, 2, 1].$$

We consider the form on the left to be the standard form; thus, we assume that finite continued fractions of length at least one do not end in 1.

1.2. **The Continued Fraction of a Real Number.** Every real number has a representation as a continued fraction.

Definition 2. Let $x \in \mathbb{R}$. The *floor* of x is

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z} \mid n \leq x\}.$$

Then $x - \lfloor x \rfloor \in [0, 1)$, so its reciprocal has a positive floor. Inductively define a sequence (x_n) by setting

$$x_0 = x \quad \text{and} \quad x_{n+1} = \frac{1}{x_n - \lfloor x_n \rfloor}.$$

The *continued fraction representation* of x is a sequence $[a_0; a_1, a_2, \dots]$, defined by $a_n = \lfloor x_n \rfloor$.

If x_n is an integer, then the sequence is finite and a_{n+1} does not exist.

Note that $a_0 < 0$ if and only if $x < 0$, and for $n \geq 1$, we have $a_n > 0$.

Let's find the continued fraction decomposition of $\frac{497}{95}$.

$$\frac{497}{95} = 5 + \frac{22}{95}$$
$$\frac{95}{22} = 4 + \frac{7}{22}$$
$$\frac{22}{7} = 3 + \frac{1}{7}$$

So,

$$\frac{497}{95} = [5;4,3,7] = 5 + \cfrac{1}{4 + \cfrac{1}{3 + \cfrac{1}{7}}}.$$

It should be noted that these are the digits obtained by the Euclidean Algorithm, which makes repeated use of the Division Algorithm to find the greatest divisor of two numbers.

$$\begin{aligned} 497 &= 5(95) + 22 \\ 95 &= 4(22) + 7 \\ 22 &= 3(7) + 1 \end{aligned}$$

Thus $\gcd(497, 95) = 1$. Since every rational number may be written as a fraction whose numerator and denominator are relatively prime, every rational number which is not an integer has a continued fraction expression which terminates in a fraction of the form $\frac{1}{n}$, where $n \in \mathbb{N}$ and $n \geq 2$.

We find the continued fraction expansion of $\sqrt{2}$. First note that if $x = 2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ddots}}}$, then $x = 2 + \frac{1}{x}$, so $x^2 - 2x - 1 = 0$, whence $x = \frac{2 \pm \sqrt{4+4}}{2} = 1 + \sqrt{2}$, since x is positive. Thus $\sqrt{2} = x - 1$, so

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ddots}}} = [1; 2, 2, 2, \ldots] = [1; \overline{2}].$$

Proposition 1. *Let $x \in \mathbb{R}$. Then x is irrational if and only if its continued fraction is infinite.*

Proof. It is easy to see that a finite continued fraction can be resolved into a fraction of the form $\frac{n}{m}$, where $m, n \in \mathbb{Z}$. We wish to see that a rational number always has a finite continued fraction. We write the fraction in the form $x = \frac{n}{m}$, where $\gcd(m, n) = 1$. By the division algorithm, write $n = mq + r$, where $0 \leq r < m$. Now $x = q + \frac{1}{m/r}$. We repeat this process with $\frac{m}{r}$. Since $r < m$, this process will eventually terminate with final remainder 1. Thus the continued fraction expression for $\frac{n}{m}$ is finite. $\square$

Proposition 2. *Let $x = [a_0; a_1, a_2, \ldots, a_n]$ with $a_0 \geq 1$. Then $x^{-1} = [0; a_0, a_1, a_2, \ldots, a_n]$.*

2. Convergents

The continued fraction expression for an irrational number may be viewed as a sequence of rational numbers, as follows.

Let x be an irrational number, so x has an infinite continued fraction expression, $x = [a_0; a_1, a_2, \dots]$. Let $c_n = [a_0; a_1, \dots, a_n]$; we call c_n the n^{th} *convergent* of x. We would like to know that the sequence (c_n) does indeed converge to x.

Each of the convergents is a rational number; thus let $\dfrac{p_n}{q_n} = c_n$, where $\gcd(p_n, q_n) = 1$. There is an inductive relationship between the p_n's, q_n's, and a_n's.

We maintain the following notation for the rest of this section. Keep in mind that the a_i's, p_i's, q_i's are positive integers (except a_0 maybe be any integer), and that the c_i's are rational numbers of the same sign as x.

Proposition 3. *For $i \geq 2$, we have*

$$p_i = a_i p_{i-1} + p_{i-2};$$
$$q_i = a_i q_{i-1} + q_{i-2}.$$

with initial values

$$p_0 = a_0, p_1 = a_0 a_1 + 1, q_0 = 1, q_1 = a_1.$$

Proof. We can verify the initial values directly. By induction, we assume that we have $p_i = a_i p_{i-1} + p_{i-2}$ and $q_i = a_i q_{i-1} + q_{i-2}$; we wish to show that $p_{i+1} = a_{i+1} p_i + p_{i-1}$ and $q_{i+1} = a_{i+1} q_i + q_{i-1}$.

The i^{th} convergent is

$$c_i = \frac{p_i}{q_i} = [a_0; a_1, \dots, a_{i-1}, a_i] = \frac{a_i p_{i-1} + p_{i-2}}{a_i q_{i-1} + q_{i-2}}.$$

The next continued fraction is formed by replacing a_i with $a_i + \dfrac{1}{a_{i+1}}$, so we plug this into the expressions we have for p_i and q_i:

$$\begin{aligned}
c_{i+1} &= \frac{p_{i+1}}{q_{i+1}} && \text{in lowest form} \\
&= [a_0; a_1, \dots, a_{i-1}, a_i + \frac{1}{a_i + 1}] && \text{replacing } a_i \text{ with } a_i + \frac{1}{a_{i+1}} \text{ in } c_i \\
&= \frac{\left(a_i + \frac{1}{a_{i+1}}\right) p_{i-1} + p_{i-2}}{\left(a_i + \frac{1}{a_{i+1}}\right) q_{i-1} + q_{i-2}} && \text{plugging this replacement into } p_i/q_i \\
&= \frac{a_i p_{i-1} + p_{i-2} + \frac{p_{i-1}}{a_{i+1}}}{a_i q_{i-1} + q_{i-2} + \frac{q_{i-1}}{a_{i+1}}} && \text{expanding} \\
&= \frac{p_i + \frac{p_{i-1}}{a_{i+1}}}{q_i + \frac{q_{i-1}}{a_{i+1}}} && \text{applying the inductive hypothesis} \\
&= \frac{a_{i+1} p_i + p_{i-1}}{a_{i+1} q_i + q_{i-1}} && \text{multiplying top and bottom by } a_{i+1}.
\end{aligned}$$

Thus $p_{i+1} = a_{i+1}p_i + p_{i-1}$ and $q_{i+1} = a_{i+1}q_i + q_{i-1}$. □

By convention, set $p_{-1} = 1$, $q_{-1} = 0$, $p_{-2} = 0$, and $q_{-2} = 1$; then the formulas in Proposition 3 hold for all $i \geq 0$.

Proposition 4. *The sequence (q_n) is increasing; that is, for $i \geq 1$, we have*

$$q_i > q_{i-1}.$$

Proof. We know that a_k and q_k are positive integers, for all $k \in \mathbb{N}$. Thus

$$q_i = a_i q_{i-1} + q_{i-2} > a_i q_{i-1} \geq q_{i-1}.$$

□

Proposition 5. *For $i \geq 1$, we have*

$$p_i q_{i-1} - p_{i-1} q_i = (-1)^{i-1}.$$

Proof. We verify the base case $i = 1$, using the initial values

$$p_0 = a_0, p_1 = a_0 a_1 + 1, q_0 = 1, q_1 = a_1.$$

Then $p_i q_{i-1} - p_{i-1} q_i = p_1 q_0 - p_0 q_1 = (a_0 a_1 + 1) \cdot 1 - a_0 \cdot a_1 = 1 = (-1)^{(1-1)}$.

By induction, we assume that $p_{i-1} q_{i-2} - p_{i-2} q_{i-1} = (-1)^{i-2}$, and attempt to show that $p_i q_{i-1} - p_{i-1} q_i = (-1)^{i-1}$.

By Proposition 3,

$$p_i = a_i p_{i-1} + p_{i-2} \text{ and } q_i = a_i q_{i-1} + q_{i-2}.$$

Solving each of these equations for a_i, then equating the a_i's, gives

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

Cross multiply to get

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

Rearrange this to obtain

$$p_i q_{i-1} - q_i p_{i-1} = -(p_{i-1} q_{i-2} - p_{i-2} q_{i-1}).$$

By our inductive hypothesis, the left hand side is $-(-1)^{i-2} = (-1)^{i-1}$, which completes the proof. □

Proposition 6. *For $i \geq 1$, we have*

$$c_i - c_{i-1} = \frac{(-1)^{i-1}}{q_{i-1} q_i}.$$

Proof. Divide the equation which is the result of Proposition 5 by $q_{i-1} q_i$. □

Consider the sequence of convergents (c_n); in light of the fact that the sequence (q_n) is increasing, the previous proposition shows that the terms of this sequence are getting closer together. We can say more.

Proposition 7. *For $i \geq 2$, we have*

$$c_i - c_{i-2} = \frac{a_i (-1)^i}{q_{i-2} q_i}.$$

Proof. Exercise. □

Call a convergent c_n *even* if it n is even, and otherwise call it *odd*.

Proposition 8. *The sequence of convergents* (c_n) *satisfies:*

- **(a)** *The sequence even convergents is increasing, and the sequence of odd convergents is decreasing.*
- **(b)** *Every even convergent is less than every odd convergent.*
- **(c)** *Every convergent lies between the two preceding convergents.*

Proof. First use Proposition 7.If i is even, $c_i - c_{i-2} = \dfrac{a_i}{q_{i-2}q_i}$ is positive, so the even convergents are increasing. On the other hand, if i is odd, $c_i - c_{i-2} = \dfrac{-a_i}{q_{i-2}q_i}$ is negative, so the odd convergents are decreasing. The proves **(a)**

Next use Proposition 6.

Let c_i be an even convergent. Since the odd convergents decrease, we have $c_{i+1} < c_{i-1}$. The sign of $c_i - c_{i-1}$ is $(-1)^{i-1} = -1$, so $c_i < c_{i-1}$. However, the sign of $c_{i+1} - c_i$ is $(-1)^i$, so $c_{i+1} > c_i$. These inequalities are reversed if i is odd. Together, we have

$$c_i < c_{i+1} < c_{i-1} \text{ if } i \text{ is even} \qquad \text{and} \qquad c_{i-1} < c_{i+1} < c_i \text{ if } i \text{ is odd }.$$

This proves **(c)**.

Finally, pick two convergents c_i and c_j, where i is even and j is odd. Suppose $i < j$. Then $i \geq j-1$, and $j-1$ is even. Then $c_i \leq c_{j-1}$ since the even convergents are increasing, but $c_{j-1} < c_j$ be the previous paragraph. Thus $c_i \leq c_{j-1} < c_j$. On the other hand, if $i > j$, we have $i \geq j+1$, and $c_i < c_{j+1} < c_j$. This proves **(b)**. □

Proposition 9. *The sequence of convergents* (c_n) *converges.*

Proof. The sequence of even terms is increasing and is bounded above by every odd term, so it converges, say to L_1. The sequence of odd terms is decreasing and is bounded below by every even term, so it also converges, say to L_2. Clearly $L_1 \leq L_2$. Let $\epsilon = L_2 - L_1$, and suppose that $\epsilon > 0$. Then there exists N_1 such that all $n \geq N_1$ and n even implies $|L_1 - c_n| < \epsilon/4$, and there exists N_2 such that all $n \geq N_1$ and n odd implies $|L_2 - c_n| < \epsilon/4$. For n bigger than N_1 and N_2, $|c_n - c_{n-1}| > \dfrac{\epsilon}{2}$.

Let N_3 be so large that $n \geq N_3$ implies $q_n > \dfrac{2}{\epsilon}$. Let $N = \max\{N_1, N_2, N_3\}$.

Let $n > N$. Then $|c_n - c_{n-1}| = \dfrac{1}{q_{n-1}q_n} < \dfrac{1}{q_n} < \dfrac{\epsilon}{2}$. This contradiction proves the proposition. □

3. Convergents Converge to Home

Consider two finite convergents which differ only in the last position, say

$$x = [a_0; a_1, \ldots, a_n, r] \text{ and } y = [a_0; a_1, \ldots, a_n, s].$$

We wish to allow that r and s be any positive real numbers. It is relatively clear that $x = y$ if and only if $r = s$. What are the conditions under which $x < y$ or $x > y$?

First, let us drop the requirement that the a_i be integers. Instead, we let $\alpha_0, \dots, \alpha_n \in \mathbb{R}$ be positive real numbers, and set

$$[\alpha_0, \alpha_1, \dots, \alpha_{n-1}, \alpha_n] = \alpha_0 + \cfrac{1}{\alpha_1 + \cfrac{1}{\ddots + \cfrac{1}{\alpha_{n-1} + \cfrac{1}{\alpha_n}}}}.$$

Note that

$$[\alpha_0, \alpha_1, \dots, \alpha_{n-2}, \alpha_{n-1}, \alpha_n] = [\alpha_0, \alpha_1, \dots, \alpha_{n-2}, \alpha_{n-1} + \frac{1}{\alpha_n}].$$

Also, it is clear that

$$[\alpha_0, \alpha_1, \dots, \alpha_k, \dots, \alpha_n] = [\alpha_0, \alpha_1, \dots, \alpha_{k-1}, [\alpha_k, \dots, \alpha_n]].$$

We may use this last observation to perform induction.

Proposition 10. *Let $\alpha_0, \alpha_1, \dots, \alpha_n, \alpha'_n \in \mathbb{R}$ be positive real numbers. Let $x = [\alpha_0, \alpha_1, \dots, \alpha_{n-1}, \alpha_n]$ and $x' = [\alpha_0, \alpha_1, \dots, \alpha_{n-1}, \alpha'_n]$. Suppose $\alpha_n < \alpha'_n$. Then*

- $x < x'$ *if n is even*
- $x > x'$ *if n is odd*

Proof. Note that

$$x = [\alpha_0, [\alpha_1, \alpha_2, \dots, \alpha_{n-1}, \alpha_n]] \text{ and } x' = [\alpha_0, [\alpha_1, \alpha_2, \dots, \alpha_{n-1}, \alpha'_n]].$$

Let $y = [\alpha_1, \alpha_2, \dots, \alpha_{n-1}, \alpha_n]$ and $y' = [\alpha_1, \alpha_2, \dots, \alpha_{n-1}, \alpha'_n]$.

By induction on the length of the sequence, $y < y'$ if $n-1$ is even, and $y > y'$ if $n-1$ is odd.

Suppose n is even. Then $n-1$ is odd, so $y > y'$, so $\frac{1}{y} < \frac{1}{y'}$. Thus, in this case, $x = \alpha_0 + \frac{1}{y} < \alpha_0 + \frac{1}{y'} = x'$.

Suppose n is odd. Then $n-1$ is even, so $y < y'$, so $\frac{1}{y} > \frac{1}{y'}$. Thus, in this case, $x = \alpha_0 + \frac{1}{y} > \alpha_0 + \frac{1}{y'} = x'$. □

Proposition 11. *Let c_i and c_j be convergents of x, so i is even and j is odd. Then*

$$c_i < x < c_j.$$

Proof. Recall the sequence (x_n) derived from x by setting $x_0 = x$ and $x_{n+1} = x_n - \lfloor x_n \rfloor$. Then $a_n = \lfloor x_n \rfloor$, so $x_n \geq a_n$.

Let $c_k = [a_0; a_1, \dots, a_{k-1}, a_k]$, and note that $x = [a_0; a_1, \dots, a_{k-1}, x_k]$. If k is even, $c_k < x$, and if k is odd, $x < c_k$. □

Proposition 12. *The sequence (c_n) of convergents of x converges to x.*

Proof. We already know the sequence converges, say to L. But x is an upper bound for the even terms, so $x \geq L$, and x is a lower bound for all the odd terms, so $x \leq L$. Thus $x = L$. □

4. A Perfect Set

Proposition 13. *Let $x \in \mathbb{R}$, and let $y = [a_0; a_1, a_2, \dots, a_{k-1}, a_k + 1] = [a_0; a_1, \dots, a_{k-1}, a_k, 1]$ be a rational number. Then y is a convergent of x if and only if x is strictly between z_1 and z_2, where*

$$z_1 = [a_0; a_1, \dots, a_{k-1}, a_k, 2], \text{ and}$$
$$z_2 = [a_0; a_1, \dots, a_{k-1}, a_k + 2].$$

Then $z_1 < x < z_2$ if k is even, and $z_2 < x < z_1$ if k is odd.

Proposition 14. *Let $y \in \mathbb{Q}$ be positive, and let*

$$U_y = \{x \in \mathbb{R} \mid y \text{ is a convergent of } x\}.$$

Then U_y is open.

Proposition 15. *Let*

$$P = \{x \in [0,1] \smallsetminus \mathbb{Q} \mid \text{ the continued fraction expression for } x \text{ has only } 1\text{'s and } 2\text{'s}\}.$$

Then P is perfect set of irrational numbers.

Proof. Why? □

Is the set P above homeomorphic to the Cantor set?

References

[1] Saxena and Shah, *Introduction to Real Variable Theory*, International Textbook Co., 1972
[2] Webpage: `https://mathcs.org/analysis/reals/topo/proofs/pfctuncb.html`
[3] Webpage: `https://en.wikipedia.org/wiki/Continued_fraction`
[4] Webpage: `https://www.cut-the-knot.org/proofs/SqContinuedFraction.shtml`

Department of Mathematics and CSci, BASIS Scottsdale
Email address: `paul.bailey@basised.com`