The Negs and Regs of Continued Fractions

Alex Eustis

Arthur T. Benjamin, Advisor

Sanjai Gupta, Reader

December 12, 2005

Harvey Mudd College

Department of Mathematics

Abstract

There are two main aims of this thesis. The first is to further develop and demonstrate applications of the combinatorial interpretation of continued fractions introduced in Benjamin and Quinn [2003]. The second is to investigate the theory of *negative* continued fractions, a relatively unresearched topic. That is, discuss the ways in which they are similar to and different from the regular class, describe how to convert between the two forms, and show that the central theorems concerning regular continued fractions also apply to the negative ones.

Contents

Chapter 1

Introduction to Continued Fractions

Continued fractions have been well-studied and have a rich history. They first appeared in the 16th century, although Brezinski [1991] argues that related concepts can be traced back to antiquity. Here we reproduce the basic definitions and theorems. Most texts always use numerators of 1, so our definitions are somewhat more general, but essentially the same.

1.1 Notation and Basic Theory

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

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{\ddots + \frac{a_n}{b_n}}}}. \tag{1.1}$$

Usually a_i and b_i are integers, although they can be real numbers, complex numbers, polynomials, etc. An *infinite continued fraction* is the limit of a sequence of finite continued fractions.

Definition 2. *A* regular *continued fraction is one in which*

- $b_k = 1$ *for all* k
- a_0 *is an integer*
- $a_1, a_2, \ldots$ *are integers* ≥ 1.

Regular continued fractions are the type encountered most often in the literature. For regular continued fractions we use the notation

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \frac{1}{a_n}}} = [a_0, a_1, \ldots, a_n]. \tag{1.2}$$

The coefficients a_k are known as the *partial quotients* of the continued fraction.

Perhaps the most basic fact about continued fractions is that they serve as rational approximations to real numbers, in the same way that decimal expansions do. For decimal expansions, this is summed up in the following propositions:

- If $d_1, d_2, \ldots$ are integers in the range $0 \leq d_i \leq 9$, then the series

  $$\sum_{i=1}^{\infty} \frac{d_i}{10^i}$$

  converges to some $x \in [0, 1]$.

- For all $x \in [0, 1]$, there exist $d_1, d_2, \ldots$, $0 \leq d_i \leq 9$, such that

  $$\sum_{i=1}^{\infty} \frac{d_i}{10^i} = x.$$

  The representation is unique for irrational x, and for rational x there are at most two such representations.

For regular continued fractions, it is well known that analogous theorems hold.

Theorem 1. *Let $a_0, a_1, \ldots$ satisfy the criteria for a regular continued fraction. Then the sequence*

$$\{c_n\} = [a_0, a_1, \ldots, a_n]$$

converges to some $x \in \mathbb{R}$. The rational numbers c_k are known as the convergents.

Theorem 2. *If x is an irrational number, then there exist unique $a_0, a_1, \ldots$, satisfying the criteria for a regular continued fraction, such that*

$$\lim_{n \to \infty} [a_0, a_1, \ldots, a_n] = x.$$

If x is rational, then

$$x = [a_0, \ldots, a_n]$$

for some a_i, with $a_n \geq 2$ (or $n = 0$). We also have

$$x = [a_0, \ldots, a_n - 1, 1],$$

and these are the only two representations for rational x.

For proofs of these theorems, consult Rockett and Szüsz [1992] or any other standard text on continued fractions.

Although the next two theorems are typically used for the purpose of proving the previous two, they are of interest to us in their own right. This first theorem shows how to compute $[a_0, \ldots, a_n]$ "forwards," rather than backwards as order of operations would suggest.

Theorem 3. *Define A_k, B_k recursively as follows:*

$$\begin{aligned} A_{-1} &= 1, & B_{-1} &= 0 \\ A_0 &= a_0, & B_0 &= 1 \end{aligned}$$

$$\left.\begin{aligned} A_k &= a_k A_{k-1} + A_{k-2} \\ B_k &= a_k B_{k-1} + B_{k-2} \end{aligned}\right. \quad (k \geq 1).$$

Then

$$[a_0, a_1, \ldots, a_k] = \frac{A_k}{B_k}$$

in lowest terms.

Theorem 4. *If A_k and B_k are defined as above, then*

$$A_k B_{k-1} - A_{k-1} B_k = (-1)^{k-1}.$$

Equivalently,

$$\frac{A_k}{B_k} - \frac{A_{k-1}}{B_{k-1}} = \frac{(-1)^{k-1}}{B_k B_{k-1}}, \tag{1.3}$$

which shows that $\{\frac{A_k}{B_k}\}$ converges at least exponentially (given that B_k increases exponentially, which can be seen from the previous theorem). It also shows that the convergents alternately approximate the limit from above and below.

1.2 Negative Continued fractions

In this paper we investigate an alternative definition, which is similar to the regular continued fraction in many ways.

Definition 3. *A* negative *continued fraction is one in which*

- $b_k = -1$ *for all* k
- a_0 *is an integer*
- $a_1, a_2, \ldots$ *are integers* ≥ 2.

For negative continued fractions we use the notation

$$a_0 - \cfrac{1}{a_1 - \cfrac{1}{\ddots - \frac{1}{a_n}}} = [a_0, a_1, \ldots, a_n]_-. \tag{1.4}$$

All four theorems from the previous section have direct analogues for negative continued fractions. We shall prove these analogues as we proceed.

Chapter 2

Combinatorial Interpretation of Continued Fractions

2.1 Square-and-domino tilings

First, we define the main combinatorial object in this paper, which is a *square-and-domino tiling*. For complete proofs of the ideas in this section, as well as a more thorough discussion, see Proofs That Count[Reference me???].

Definition 4. *Let n be a nonnegative integer. Consider a linear board of length n ("n-board" for short). At our disposal we have square tiles of length 1 and domino tiles of length 2. A square-and-domino tiling is any arrangement of tiles that covers the board completely with no overlapping tiles. Let $\mathcal{F}_n$ denote the set of all square-and-domino tilings.*

Theorem 5. *Let $F_1 = 1, F_2 = 1, F_2 = 2, F_3 = 3, F_4 = 5, \ldots$ be the Fibonacci sequence. Then*

$$|\mathcal{F}_n| = F_{n+1}.$$

This theorem is proved by showing that $|\mathcal{F}_n| = |\mathcal{F}_{n-1}| + |\mathcal{F}_{n-2}|$, which follows by conditioning on the whether the last tile is a square or a domino. We are not much concerned with Fibonacci numbers in this paper, but it explains the notation $\mathcal{F}_n$.

2.1.1 Weighted tilings

Next, we introduce the idea of a weighted tiling and some associated notation. Let $a_0, \ldots, a_n$ and $b_1, \ldots, b_n$ be any real numbers, and consider an

$n+1$-board with cells labeled 0 to n. For a given tiling $t \in \mathcal{F}_{n+1}$, we assign a weight to each tile as follows: a square in cell i has weight a_i and a domino covering cells $i-1$ and i has weight b_i.

Definition 5. *Given a tiling $t \in \mathcal{F}_{n+1}$, its* weight, *denoted $w(t)$, is the product* (not *the sum) of the weights of its tiles. We also define the* weighted sum *of the tilings in $\mathcal{F}_{n+1}$ and notate it as follows:*

$$|0:n| = \sum_{t \in \mathcal{F}_{n+1}} w(t) \tag{2.1}$$

Whenever we use the notation $|0:n|$, the weights $a_0, \ldots, a_n, b_1, \ldots, b_n$ are implied from context. For integers $0 \leq i \leq j \leq n$, we may also write $|i:j|$ to mean the weighted sum of the tilings of the sub-board starting at cell i and ending at cell j.

It is also consistent and convenient to define $|i:i-1| = 1$, if we think of this quantity as the number of ways to tile a zero-length board.

The simplest nontrivial example is when $a_i = b_i = 1$. Then each tiling has unit weight, and so $|0:n| = F_{n+2}$ by theorem 5.

2.1.2 A few identities

The following identities apply to weighted tilings in general. Since we may let $a_i = b_i = 1$ as above, they have direct analogues for Fibonacci numbers, and the proofs are essentially the same.

The first identity could be used to recursively compute $|0:n|$ starting from $|0:-1| = 1$ and $|0:0| = a_0$. It also generalizes theorem 3, as we shall see.

Identity 1.

$$|0:n| = a_j|0:n-1| + b_j|0:n-2|$$

Proof. There are two types of tilings: those that end in a square and those that end in a domino. Consider the sum of the weights of all tilings ending in a square. These weights all have a common factor of a_n which distributes out of the sum, leaving us with the weights of all the tilings of the $0:n-1$ sub-board. This gives the $a_n|0:n-1|$ term. Of course, the same argument shows that the weighted sum of all tilings ending in a domino is $b_n|0:n-2|$. □

The same identity from the other end of the tiling reads:

$$|0:n| = a_0|1:n| + b_1|2:n|. \tag{2.2}$$

The next identity is known as the "tail-swapping" identity. It is also the generalization of theorem 4.

Identity 2.

$$|0:n||1:n-1| = |0:n-1||1:n| + (-1)^{n+1}\prod_{i=1}^{n} b_i.$$

Proof. This one really needs a diagram badly, but I'll get around to it eventually. □

2.1.3 The weighted sum as a determinant

We can also view this weighted sum as the determinant of a certain matrix, as per the following theorem.

Theorem 6. *Let* $a_0, \ldots, a_n, b_0, \ldots, b_n$ *be real numbers, let* P *be the* $(n+1) \times (n+1)$ *matrix*

$$P = \begin{bmatrix} a_0 & 1 & 0 & \cdots & 0 \\ -b_1 & a_1 & 1 & & \vdots \\ 0 & -b_2 & a_2 & & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & 0 & -b_n & a_n \end{bmatrix}.$$

Then

$$|0:n| = \det(P)$$

Several authors (for instance, see Clarke et al. [1999]) use this determinant instead of continued fractions. The above theorem shows that the two approaches are equivalent.

2.2 Continued fractions

2.2.1 General Case

The following theorem gives the connection between continued fractions and combinatorics. More specifically, it shows how to express a continued fraction in terms of the weighted sums defined above.

Theorem 7. *Let* $a_0, \ldots, a_n, b_1, \ldots, b_n$ *be real numbers. Then*

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{\ddots + \cfrac{b_n}{a_n}}} = \frac{|0:n|}{|1:n|}.$$

Proof. The most straightforward way to compute the LHS which requires no background knowledge of continued fractions is to just use the order of operations suggested by the way the expression is written. That is, starting at $k = n$ and working backwards towards $k = 0$, we compute the quantity underneath the first k division bars. For $k = n$ this quantity is just a_n, which we shall write as $\frac{a_n}{1}$ for reasons that will soon become apparent. Now, suppose that we have determined the quantity under the first k division bars to be equal to $\frac{p_k}{q_k}$. (Note: this notation is not meant to suggest that the quantity must be rational; only that we have found real numbers p_k and q_k such that the above holds.) In order to "move up" one division bar we do the following computation:

$$a_{k-1} + \frac{b_k}{p_k/q_k} = \frac{a_{k-1}p_k + b_k q_k}{p_k}. \tag{2.3}$$

To continue this process, we define the sequences p_k, q_k in terms of the a_k, b_k as follows:

$$\begin{array}{l|l} p_n = a_n & q_n = 1 \\ p_{k-1} = a_{k-1}p_k + b_k q_k & q_{k-1} = p_k \quad (1 \leq k \leq n) \end{array}$$

so that we have

$$\frac{p_k}{q_k} = a_k + \frac{b_{k+1}}{\ddots + \frac{b_n}{a_n}}. \tag{2.4}$$

Now, we want to prove by induction that

$$p_k = |k : n| \text{ and } q_k = |k+1 : n|. \tag{2.5}$$

(Both statements must be proven at once). The base case $k = n$ is clear. Now, supposing the statements are true for some k, we have

$$q_{k-1} = p_k = |k : n|$$

as desired, and

$$p_{k-1} = a_{k-1}p_k + b_k q_k = a_{k-1}|k : n| + b_k|k+1 : n| = |k-1 : n|$$

by identity 1.

This shows that, in particular,

$$a_0 + \frac{b_1}{a_1 + \frac{b_2}{\ddots + \frac{b_n}{a_n}}} = \frac{p_0}{q_0} = \frac{|0 : n|}{|1 : n|}.$$

□

2.2.2 Integer values

So far, we have only assumed that $a_0, \ldots, a_n, b_1, \ldots, b_n$ are real numbers. In fact, there is nothing to stop us from considering them to complex numbers or in fact elements of any field. However, for our purposes they will usually be integer values. When a_k, b_k are nonnegative integers, we can give an interpretation for $|0 : n|$ which is "even more" combinatorial, because it is the size of a set rather than just a sum of numbers. The objects in the set are still square-and-domino tilings, but we do away with the weights and instead allow squares and dominoes to be stacked.

Definition 6. *Let $a_0, \ldots, a_n$ and $b_1, \ldots, b_n$ be nonnegative integers. We define a* tiling of an $n+1$-board with height conditions *as follows. Again we must cover all cells 0 to n, but on a given cell i, instead of allowing only a single square tile, we allow a stack of squares of height 1 to a_i. Over cells $i-1, i$, we also allow dominoes to be stacked in a the same way up to height b_i.*

Of course, we don't allow squares to be stacked on top of dominoes or any other such combinations. An equivalent conception of a tiling with height conditions is a colored tiling, where a square on cell i can be any of a_i colors, and a domino covering cells $i-1, i$ can be any of b_i colors.

Theorem 8. *For nonnegative integers $a_0, \ldots, a_n$ and $b_1, \ldots, b_n$, the number of tilings of an $n+1$-board with these height conditions is equal to $|0 : n|$.*

Proof. For any tiling $t \in \mathcal{F}_{n+1}$, $w(t)$ gives the number of different ways to form a height conditioned tiling whose square and domino stacks occur in the same places as the squares and dominoes (respectively) in t. It follows that $\sum_{t \in \mathcal{F}_{n+1}} w(t)$ gives the total number of height conditioned tilings overall. □

2.2.3 Negative Dominoes

Since we are dealing with negative continued fractions, we also want to know how to generalize theorem 8 to interpret $|0 : n|$ as the size of a set even when $b_i < 0$. Let us first define the tilings that will make up this set.

Definition 7. *Let $a_0, \ldots, a_n$ be nonnegative integers, and let $b_1, \ldots, b_n$ be integers satisfying $b_k > -a_k$. Then we define a* mixed tiling of an $(n+1)$-board with height conditions a_k, b_k *as follows. In cell k we allow stacks of squares up to height a_k. If $b_k \geq 0$ then over the cells $k-1, k$ we also allow stacks of dominoes up to height b_k. If $b_k < 0$, then instead of allowing dominoes we impose a* directionality requirement *on cells $k-1, k$. This requirement is defined as follows:*

1. *A stack of squares on cell* $k-1$ *of maximum height (that is, height* a_{k-1}*) counts as* $\rightarrow$.

2. *A stack of squares up to height* $|b_k|$ *on cell* k *counts as* $\leftarrow$.

3. *The directionality requirement prohibits* $\rightarrow\leftarrow$ *from appearing consecutively on cells* $k-1, k$.

Note that the requirement $b_k > -a_k$ guarantees that no maximum height stack of squares can satisfy 2; that is, no cell can be $\rightarrow$ and $\leftarrow$ at the same time.

These mixed tilings are of a more general class than the unmixed tilings we defined earlier, because in the case $b_k \geq 0$ for all k, the mixed tilings are precisely the same as the unmixed ones. We now prove that theorem 8 generalizes fully to these new tilings.

Theorem 9. *If* $a_0, \ldots, a_n, b_1, \ldots, b_n$ *are integers satisfying* $a_k \geq 0$, $b_k > -a_k$, *the number of mixed tilings of an* $n+1$*-board with these height conditions is* $|0:n|$.

Proof. We say a domino is in *negative position* if its weight is negative; that is, if it covers cells $k-1, k$ where $b_k < 0$. Since squares always have non-negative weight, it is clear that

$$w(t) = (-1)^{d(t)}|w(t)|, \tag{2.6}$$

where $d(t)$ denotes the number of dominoes the tiling t has in negative position. Hence we may write

$$|0:n| = \sum_{t \in \mathcal{F}_{n+1}} (-1)^{d(t)}|w(t)|. \tag{2.7}$$

Now, let $\mathcal{U}$ denote the set of unmixed tilings with height conditions $a_0, \ldots, a_n, |b_0|, \ldots, |b_n|$. By theorem 8 and its proof, $|w(t)|$ is the number of number of tilings $s \in \mathcal{U}$ with square and domino stacks in the same places as the squares and dominoes of t. In particular, each of these corresponding tilings $s \in \mathcal{U}$ satisfies $d(s) = d(t)$. It follows that

$$|0:n| = \sum_{s \in \mathcal{U}} (-1)^{d(s)}. \tag{2.8}$$

To say this another way, if we let $\mathcal{E} \subseteq \mathcal{U}$ be the set of unmixed tilings s with $d(s)$ even, and let $\mathcal{O}$ be the set of tilings with $d(s)$ odd, then $|0:n| = |\mathcal{E}| - |\mathcal{O}|$.

Now, we can identify $\mathcal{M}$, the set of *mixed* tilings with height conditions $a_0, \ldots, a_n, b_1, \ldots, b_n$, as a subset of $\mathcal{E}$, since $\mathcal{M}$ consists of those tilings with *zero* domino stacks in negative position, and which also satisfy the directionality requirements.

We wish to show that $|\mathcal{E} - \mathcal{M}| = |\mathcal{O}|$, from which it will follow that $|0 : n| = |\mathcal{E}| - |\mathcal{O}| = |\mathcal{M}|$ as desired. We proceed by finding a one-to-one correspondence between $\mathcal{E} - \mathcal{M}$ and $\mathcal{O}$.

Define the word *breach* to mean either a stack of dominoes in negative position or a place where a directionality requirement fails. We may also define the *height* of a breach to mean the number of dominoes in the stack (if it is a stack of dominoes) or the number of squares in the stack corresponding to $\leftarrow$ (if it is $\rightarrow\leftarrow$).

By definition, every tiling in $\mathcal{E} - \mathcal{M}$ or in $\mathcal{O}$ must have at least one breach. The correspondence is given by finding the last breach that occurs and replacing it with one of the opposite kind but the same height.

[Diagram here]

It is easy to see that this correspondence maps tilings in $\mathcal{E} - \mathcal{M}$ to ones in $\mathcal{O}$, and vice versa. We can also see that the correspondence is its own inverse because it does not change the position of the last breach. Note that this is where we need to use the fact that no cell can be both $\rightarrow$ and $\leftarrow$; that is, $\leftrightarrow$ does not exist. Otherwise this correspondence could replace $\boxed{}\leftarrow$ with $\rightarrow\leftrightarrow\leftarrow$, which *would* change the position of the last breach.

Therefore we have proven that $|\mathcal{E} - \mathcal{M}| = |\mathcal{O}|$, which is what we wanted to show. $\square$

2.3 How to Convert between Regs and Negs

If $x \in \mathbb{R}$, then there exists a direct relationship between the partial quotients for the regular continued fraction of x and those of the negative continued fraction. In what follows we shall describe and prove this relationship.

2.3.1 Mixed tilings with $b_k = \pm 1$

Although the tilings in the previous section allowed $|b_k| > 1$, we now restrict ourselves to the case $b_k = \pm 1$. In this case, dominoes may no longer be stacked, and $\leftarrow$ is always a single unstacked square.

For this kind of mixed tiling we prefer a slightly different interpretation. Rather than think of $\leftarrow$ and $\rightarrow$ as stacks of squares, we instead consider them to be objects unto themselves with which we may tile the board.

Whenver $b_k = -1$, cell $k-1$ may contain $\rightarrow$ and cell k may contain $\leftarrow$, but not both at once. Under this interpretation, a_k no longer represents the maximum height of square stacks, but rather, it is the number of different objects allowed in cell k, including square stacks and arrows but excluding dominoes.

Also, we adopt another notation for $|0 : n|$ in which the values a_k and b_k are expressed explicitly. When $b_k = 1$ we use the symbol $\oplus$, and when $b_k = -1$ we use the symbol $\ominus$, so for example the notation

$$|a_0 \oplus a_1 \ominus a_2| \tag{2.9}$$

means the number of ways to tile a mixed board with conditions $a_0, a_1, a_2, b_1 = 1, b_2 = -1$.

We now present a theorem which will allows us to convert a small section of a board that uses dominoes ($\oplus$) into a section that uses arrows ($\ominus$) without changing the number of tilings.

Theorem 10. *For any appropriate values of* $a_0, \dots, a_n$*,*

$$\begin{gathered} |\dots a_{m-1} \oplus a_m \oplus a_{m+1} \dots| = \\ |\dots (a_{m-1}+1) \ominus \underbrace{2 \ominus 2 \ominus 2}_{a_m - 1 \text{ times}} \ominus (a_{m+1}+1) \dots|. \end{gathered} \tag{2.10}$$

In the above equation, the ellipses on the left are placeholders for all the values $a_0, a_1, \dots, a_{m-2}, b_1, \dots, b_{m-1}$, which may be any valid values at all, as long as they are the same values for the LHS and the RHS. Similar comments apply to the ellipses on the right.

Proof. We want to find a one-to-one correspondence between tilings of the first board and of the second. To do so we need a bit of terminology for board 2. The sequences of 2s is called the *arrow region* because the only objects allowed there are $\leftarrow$ and $\rightarrow$. The cell immediately to the left of the arrow region (with height condition $a_{m-1}+1$) is the *left border* and the cell to the right is the *right border*. We now provide a dictionary which we use to map the cells $m-1$, m, and $m+1$ of board 1 onto the the left border, arrow region, and right border (respectively) of board 2.

The only object allowed in cell $m-1$ of board 1 that is not allowed in the left border of board 2 is the left half of a domino, which we abbreviate $\sqsubset$. Also, the left border of board 2 can have $\rightarrow$, whereas cell $m-1$ of board 1 cannot. This naturally suggests that we use the rule

$$\sqsubset \quad \longmapsto \quad \rightarrow \tag{2.11}$$

for this cell. Any other object appearing in cell $m-1$ of board 1 may be replaced by the same object in board 2. Similarly,

$$\sqsupset \quad \longmapsto \quad \leftarrow \tag{2.12}$$

for cell $m+1$.

For cell m, rules 2.11 and 2.12 force us to use the following in order to satisfy directionality:

$$\sqsupset \quad \longmapsto \quad \underbrace{\rightarrow\rightarrow}_{a_m-1} \tag{2.13}$$

$$\sqsubset \quad \longmapsto \quad \underbrace{\leftarrow\leftarrow}_{a_m-1} \tag{2.14}$$

The remaining possibility is that a stack of squares appears in cell m. We observe that the number of ways to tile the arrow region is exactly a_m because we must begin with some number of $\leftarrow$ between 0 and $a_m - 1$ inclusive, and the remaining cells (if any) must be $\rightarrow$. This suggests that for a stack of h squares, $\boxed{\text{h}}$, we should use the rule

$$\boxed{\text{h}} \quad \longmapsto \quad \underbrace{\leftarrow\leftarrow}_{h-1}\underbrace{\rightarrow\rightarrow}_{a_m-h}. \tag{2.15}$$

This never breaks directionality because we already covered the cases in which $\rightarrow$ or $\leftarrow$ appear in the border squares.

Of course, the rest of the board is mapped to itself without any changes. Now that we have defined the correspondence between board 1 and board 2, it is really no more work to see that the correspondence is in fact a bijection. This is because if we start with a tiling of board 2, we can simply apply the above rules in reverse to determine a unique tiling of board 1. The only ambiguity is when the arrow region contains all arrows of the same kind; this might be mapped onto either half a domino or a minimal/maximal stack of squares. But, whether or not one *more* arrow appears in the appropriate border square distinguishes which case to apply. □

2.3.2 A conversion formula

What we investigate next is that when theorem 10 is used repeatedly, we can convert an entire board with $b_k = 1$ to one with $b_k = -1$. First, we note a simple modification that can be made to the end of a board.

Identity 3.

$$|a_n \ldots| = |1 \oplus (a_n - 1) \ldots|.$$

Proof. The correspondence between boards 1 and 2 is quite simple. If the first cell of board 1 is a maximal stack of squares, then replace it with a domino; otherwise leave it be and add a square to the beginning. □

Theorem 10 and the above identity give us a procedure for converting between regular and negative continued fractions.

For example, consider the regular continued fraction

$$6 + \cfrac{1}{2 + \frac{1}{3 + \frac{1}{3 + \frac{1}{2}}}}.$$

By theorem 7, this may be rewritten as

$$\frac{|6 \oplus 2 \oplus 3 \oplus 3 \oplus 2|}{|2 \oplus 3 \oplus 3 \oplus 2|}.$$

We can now use theorem 10 and identity 3 to convert these two reg-boards into similar neg-boards:

$$\begin{aligned} & |6 \oplus 2 \oplus 3 \oplus 3 \oplus 2| \\ = \; & |7 \ominus 2 \ominus 4 \oplus 3 \oplus 2| \\ = \; & |7 \ominus 2 \;\ominus 5 \ominus 2 \ominus 2 \ominus 3|, \end{aligned}$$

$$\begin{aligned} & |2 \oplus 3 \oplus 3 \oplus 2| \\ = \; & |1 \oplus 1 \oplus 3 \oplus 3 \oplus 2| \\ = \; & |2 \ominus 4 \oplus 3 \oplus 2| \\ = \; & |2 \ominus 5 \ominus 2 \ominus 2 \ominus 3| \end{aligned}$$

Thus we conclude that

$$6 + \cfrac{1}{2 + \frac{1}{3 + \frac{1}{3 + \frac{1}{2}}}} = \frac{|6 \oplus 2 \oplus 3 \oplus 3 \oplus 2|}{|2 \oplus 3 \oplus 3 \oplus 2|} = \frac{|7 \ominus 2 \;\ominus 5 \ominus 2 \ominus 2 \ominus 3|}{|2 \ominus 5 \ominus 2 \ominus 2 \ominus 3|} = 7 - \cfrac{1}{2 - \frac{1}{5 - \frac{1}{2 - \frac{1}{2 - \frac{1}{3}}}}}.$$

This same procedure can be made to work for any regular continued fraction. We prove this in the next few theorems.

Definition 8. *Let $a_0, \ldots, a_n$ be positive integers. Define the* neg-sequence for *$a_0, \ldots, a_n$ as*

$$\bar{a}_0, \ldots, \bar{a}_m = (a_0 + 1), \underbrace{2, \ldots, 2}_{a_1 - 1}, (a_2 + 2), \underbrace{2, \ldots, 2}_{a_3 - 1}, (a_4 + 2), \ldots.$$

Also, define the alternate neg-sequence for $a_0, \ldots, a_n$*as*

$$\bar{a}'_0, \ldots, \bar{a}'_m = a_0, (a_1+1), \underbrace{2, \ldots, 2}_{a_2-1}, (a_3+2), \underbrace{2, \ldots, 2}_{a_4-1}, \ldots.$$

These sequences end in either (a_n+1) *or* $\underbrace{2, \ldots, 2}_{a_n-1}$*, according to the alternating pattern. If* $a_0, a_1, \ldots$ *is an infinite sequence, we define its neg-sequences in the same way.*

Theorem 11. *Using the notation in the above definition,*

$$|a_0 \oplus \ldots \oplus a_n| = |\bar{a}_0 \ominus \bar{a}_1 \ominus \bar{a}_2 \ominus \ldots \ominus \bar{a}_m| = |\bar{a}'_0 \oplus \bar{a}'_1 \ominus \bar{a}'_2 \ominus \ldots \ominus \bar{a}'_m|.$$

Proof. The idea for this proof is demonstrated in the above examples. Beginning with the board $|a_0, \oplus \ldots \oplus a_n|$, make the following changes to it:

1. If necessary, apply identity 3 to the right edge so that there are an even number of $\oplus$ signs altogether (odd for the alternate neg sequence).
2. Starting with the first two $\oplus$ signs (second and third for the alternate), apply theorem 10 to this and each subsequent pair to convert them to $\ominus$.

□

We are now ready to prove the main result of this section, which shows that the neg-sequence is in fact the sequence of partial quotients for the negative continued fraction.

Theorem 12. *Let* $x \in \mathbb{R}^+$*, let* $a_0, a_1, \ldots$ *be the sequence of partial quotients for the regular continued fraction for* x*, and let* $\bar{a}_k, \bar{a}'_k$ *be the corresponding infinite neg-sequences. Then, for* n *odd,*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \frac{1}{a_n}}} = \bar{a}_0 - \cfrac{1}{\bar{a}_1 - \cfrac{1}{\ddots - \frac{1}{\bar{a}_m}}}$$

and for n *even,*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \frac{1}{a_n}}} = \bar{a}'_0 + \cfrac{1}{\bar{a}'_1 - \cfrac{1}{\ddots - \frac{1}{\bar{a}'_m}}}$$

Proof. Consider the odd case first. By theorem 7, it is sufficient to show

$$|a_0 \oplus \ldots \oplus a_n| = |\bar{a}_0 \ominus \ldots \ominus \bar{a}_m| \text{ and } |a_1 \oplus \ldots \oplus a_n| = |\bar{a}_1 \ominus \ldots \ominus \bar{a}_m|.$$

We note that the the neg-sequence for $a_0, \ldots, a_n$ coincides with the first m terms of the infinite neg-sequence $\bar{a}_0, \bar{a}_1, \ldots$. This is because the only place they could differ is in the last term ($(a_n + 1)$ versus $(a_n + 2)$), but n being odd circumvents this discrepancy. Therefore, the first equation follows directly from theorem 11 and the second follows from a very similar method: apply identity 3 to *both* ends of $|a_1 \oplus \ldots \oplus a_n|$ and then convert.

For the even case: again by theorem 7, it is sufficient to show

$$|a_0 \oplus \ldots \oplus a_n| = |\bar{a}'_0 \oplus \bar{a}'_1 \ominus \ldots \ominus \bar{a}'_m| \text{ and } |a_1 \oplus \ldots \oplus a_n| = |\bar{a}'_1 \ominus \ldots \ominus \bar{a}'_m|.$$

For the same reasons as before, the alternate neg-sequence for $a_0, \ldots, a_n$ corresponds with the first m terms of the infinite alternate neg-sequence $\bar{a}'_0, \bar{a}'_1, \ldots$. So, as before, the equations follow from theorem 11. □

Corrolary 13. *With the notation as above, $\bar{a}_0, \bar{a}_1, \ldots$ is the sequence of partial quotients for the negative continued fraction of x, and $-\bar{a}'_0, \bar{a}'_1, \bar{a}'_2, \ldots$ is the sequence of partial quotients for the negative continued fraction of $-x$.*

Proof. This follows by taking the limit as $n \to \infty$ in the previous theorem, and noting that

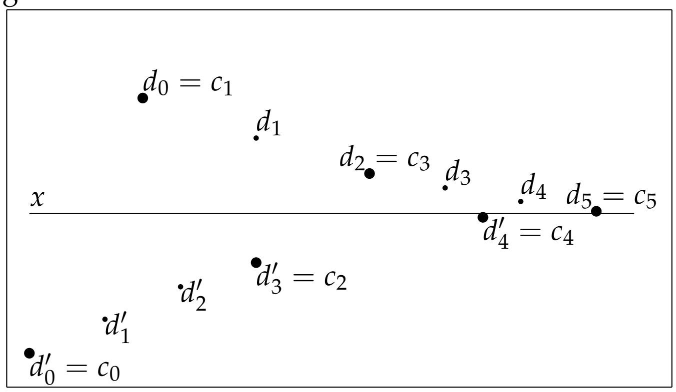
$$-x = -\left(\bar{a}'_0 + \cfrac{1}{\bar{a}'_1 - \cfrac{1}{\ddots}}\right) = -a_0 - \cfrac{1}{\bar{a}'_1 - \cfrac{1}{\ddots}}. \tag{2.16}$$

□

2.3.3 Convergents

As we know, given a real number x, the convergents $c_0, c_1, c_2, \ldots$ for its regular continued fraction alternately underestimate and overestimate x. On the other hand, the convergents $d_0, d_1, \ldots$ of its negative continued fraction always overestimate x. Furthermore, if we calculate the negative continued fraction for $-x$, its convergents $-d'_0, -d'_1, \ldots$ overestimate $-x$, which means that $d'_0, d'_1, \ldots$ underestimate x. The new information that the above theorem gives us is that, in fact, the negative convergents $d_0, d_1 \ldots$ include the odd regular convergents $c_1, c_3, \ldots$ as a subsequence, and the alternate negative convergents $d'_0, d'_1, \ldots$ include the even regular convergents $c_0, c_2, \ldots$

as a subsequence. This means that if we write the negative continued fractions for x and $-x$, we get *all* the convergents from the regular continued fraction, plus some additional convergents. The situation is pictured in the diagram below.



This naturally leads to the question: what are these extra convergents that we obtain from negative continued fractions? Our combinatorial interpretation can answer this question as well. First, an example.

Let $x = 2711/1731$. The first four convergents for the regular continued fraction of x are $1/1, 2/1, 3/2, 11/7$. The first four convergents for the negative continued fraction of x are $2/1, 5/3, 8/5, 11/7$. As expected, we eventually obtained $2/1$ and $11/7$, but we also obtained two fractions in between. The sequence of numerators is $2, 5, 8, 11$, which is an arithmetic progression. The sequence of denominators is $1, 3, 5, 7$, which is another arithmetic progression. Furthermore, these are the longest arithmetic progressions that can go from 2 to 11 and from 1 to 7 respectively, such that both progressions have the same number of terms.

This suggests that, given the reg-convergents $2/1$ and $11/7$, we could have predicted the neg-convergents that appear in between. By calculating $\gcd(11-2, 7-1) = 3$, we realize that the longest pair of arithmetic progressions of the above form has three steps, and so we would guess the correct sequence of convergents.

Indeed, we can prove that this works in general. First, we need to use a result from elementary number theory:

Lemma 1. *If* $a, b, c, d \in \mathbb{Z}$ *and* $ab - cd = \pm 1$, *then* $\gcd(a-c, b-d) = 1$.

Proof. Observe that

$$(a-c)(b+d) + (a+c)(b-d) = 2(ab-cd) = \pm 2,$$

so all we have to show is that $a - c$ and $b - d$ cannot both be even. If they were, then so would be $a + c$ and $b + d$, and so the LHS of the above equation would be divisible by 4 whereas the RHS would not. This contradiction proves the claim. Incidentally, this also proves that $\gcd(a + c, b + d) = 1$. □

Now, the theorem about the "in-between" convergents is as follows.

Theorem 14. *Let $c_{2n-1} = \frac{p_1}{q_1}, c_{2n+1} = \frac{p_2}{q_2}$ be two consecutive odd convergents of the regular continued fraction for x, and let $d_0, d_1, \ldots$ be the sequence of convergents of the negative continued fraction for x. From the previous theorem, $c_{2n-1} = d_m$ for some m, and $c_{2n+1} = d_{m+g}$.*

In addition,

$$g = gcd(p_2 - p_1, q_2 - q_1), \tag{2.17}$$

and the numerators and denominators of $d_m, d_{m+1}, \ldots, d_{m+g}$ each form an arithmetic progression.

The analogous result holds comparing the even convergents c_{2n}, c_{2n+2} to $\bar{d}_0, \bar{d}_1, \ldots$ (the neg-convergents to $-x$).

Proof. Let $a_0, a_1, \ldots$ be the sequence of partial quotients for the regular continued fraction for x. For $0 \leq j \leq g$, write $d_{m+j} = e_j / f_j$ in lowest terms. From the previous thoerems, we know how to express e_j and f_j in terms of the a_i as follows.

$$\begin{array}{lllllllll} e_j = & |(a_0+1) & \ominus & 2\ldots2 & \ominus\ldots\ominus & 2\ldots2 & \ominus & (a_{2n}+2)\ominus 2\ldots2| \\ f_j = & & & |\underbrace{2\ldots2}_{a_1-1} & \ominus\ldots\ominus & \underbrace{2\ldots2}_{a_{2n-1}-1} & \ominus & \underbrace{(a_{2n}+2)\ominus 2\ldots2}_{j}|. \end{array}$$

First, we show that $e_0, \ldots, e_g$ and $f_0, \ldots, f_g$ are arithmetic progressions. Using identity 1, we have, for $2 \leq j \leq g$,

$$\begin{aligned} e_j &= 2e_{j-1} - e_{j-2} \\ \Rightarrow e_j - e_{j-1} &= e_{j-1} - e_{j-2}. \end{aligned} \tag{2.18}$$

Thus, the difference between consecutive terms remains the same. The same argument holds for f_j.

Now, we have to show that $g = \gcd(p_2 - p_1, q_2 - q_1)$. Recall that $p_1 = e_0$, $p_2 = e_g$, $q_1 = f_0$, $q_2 = f_g$. Also, by the general formula for an arithmetic progression,

$$e_g = e_0 + (e_1 - e_0)g \qquad f_g = f_0 + (f_1 - f_0)g, \tag{2.19}$$

so

$$\gcd(p_2 - p_1, q_2 - q_1) = \gcd(g(e_1 - e_0), g(f_1 - f_0)) = g \gcd(e_1 - e_0, f_1 - f_0). \quad (2.20)$$

However, e_0/f_0 and e_1/f_1 are consecutive convergents of a negative continued fraction. So, by identity 2,

$$e_0 f_1 - e_1 f_0 = 1, \quad (2.21)$$

therefore $\gcd(e_1 - e_0, f_1 - f_0) = 1$ by our above lemma. This shows that $\gcd(p_2 - p_1, q_2 - q_1) = g$ as desired.

For even convergents, the proof is identical except for the fact that the beginnings of the above tilings look a little different. We leave the details to the reader. □

2.4 Other Identities

There are a number of identities involving continued fractions, and as an exercise in the usefulness of our combinatorial approach, we show how to prove several of these.

2.4.1 Van der Poorten's Ripple Lemma

Identity 4.

$$|a_0 \oplus \ldots \oplus a_n| = |a_0 - 1 \oplus 1 \ominus a_1 \oplus 1 \ominus a_2 \oplus 1 \ominus \ldots|.$$

Proof postponed for later. This identity appeared in van der Poorten [1998] in a somewhat different form.

2.4.2 An advanced identity

Identity 5. *If L_n denotes the nth Lucas number and F_n denotes the nth Fibonacci number, then*

$$\frac{F_{(t+1)m}}{F_{tm}} = L_m - \cfrac{(-1)^m}{L_m - \cfrac{(-1)^m}{\ddots - \frac{(-1)^m}{L_m}}}$$

Proof also postponed, but we can rewrite it as

$$F_{(n_1)m} = |0 : n - 1| F_m,$$

where $|0 : n-1|$ refers to a weighted tiling with $a_i = L_m$ and $b_i = (-1)^{m-1}$, then prove it combinatorially.

This identity appeared in Benjamin and Quinn [2003] as the last of the "open" problems (i.e. those still needing a combinatorial proof.)

Chapter 3

Future Work

3.1 Two-squares Theorems

Fermat proved the following well-known theorem:

Theorem 15. *Let p be a prime. Then*

$$p = x^2 + y^2$$

for some integers x, y if and only if $p \equiv 1 \mod 4$.

Several other versions of this theorem have been proven since then. The following appear in Nagell [1964]:

Theorem 16. *Let $p > 2$ be a prime.*

- *p can be written as $x^2 + y^2$ for some integers x, y if and only if $p \equiv 1 \mod 4$.*
- *p can be written as $x^2 + 3y^2$ if and only if $p \equiv 1 \mod 6$.*
- *p can be written as $x^2 + 2y^2$ if and only if $p \equiv 1$ or $3 \mod 8$*
- *p can be written as $x^2 + 7y^2$ if and only if $p \equiv 1$ or 9 or $11 \mod 14$.*
- *p can be written as $2x^2 + 3y^2$ if and only if $p \equiv 5$ or $11 \mod 24$*

Fermat's orginal two-squares theorem has a very neat proof, originally due to Smith (1826-1883) involving palindromic continued fractions Clarke et al. [1999]. The other versions of the theorem may have similar proofs involving palindromic continued fractions with $b_k \neq 1$, but I've yet to get all the details to work. More on this later.

3.2 Periodic Continued Fractions

May want to say a few words about how $a_0, a_1, \ldots$ is a periodic sequence of integers if and only if $[a_0, a_1, \ldots]$ is a quadratic number. Similar results undoubtedly hold for negative continued fractions.

3.3 Class Numbers

There is some mysterious connection between continued fractions and class numbers of quadratic fields, hinted at in Ireland and Rosen [1990] and zag. This could be a very neat culminating result if I could find out more about it.

Bibliography

Arthur T. Benjamin and Jennifer J. Quinn. *Proofs that really count*, volume 27 of *The Dolciani Mathematical Expositions*. Mathematical Association of America, Washington, DC, 2003. ISBN 0-88385-333-7. The art of combinatorial proof.

Claude Brezinski. *History of continued fractions and Padé approximants*, volume 12 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1991. ISBN 3-540-15286-5.

F. W. Clarke, W. N. Everitt, L. L. Littlejohn, and S. J. R. Vorster. H. J. S. Smith and the Fermat two squares theorem. *Amer. Math. Monthly*, 106(7): 652–665, 1999. ISSN 0002-9890.

Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990. ISBN 0-387-97329-X.

Trygve Nagell. *Introduction to number theory*. Second edition. Chelsea Publishing Co., New York, 1964.

Andrew M. Rockett and Peter Szüsz. *Continued fractions*. World Scientific Publishing Co. Inc., River Edge, NJ, 1992. ISBN 981-02-1047-7.

Alf van der Poorten. Formal power series and their continued fraction expansion. In *Algorithmic number theory (Portland, OR, 1998)*, volume 1423 of *Lecture Notes in Comput. Sci.*, pages 358–371. Springer, Berlin, 1998.