NOTES ON CONTINUED FRACTIONS

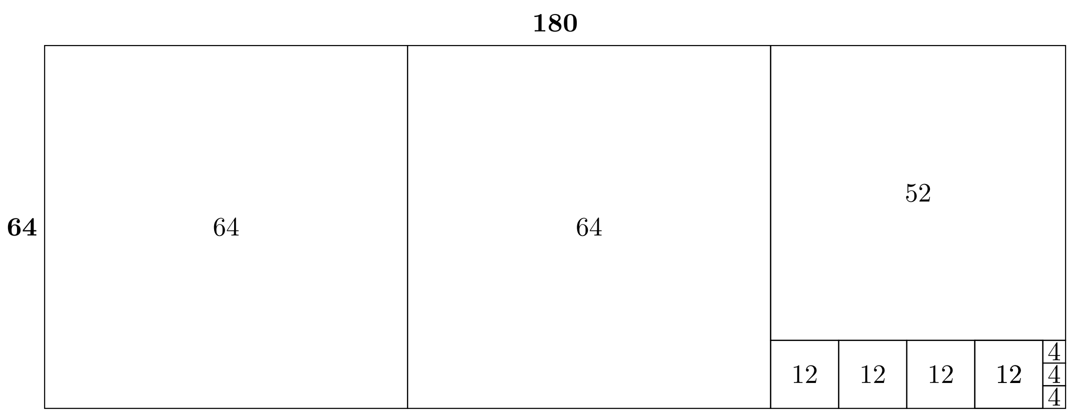
ERIC MARTIN

1. Paving a rectangle by squares, Euclid's algorithm for computing the greatest common divisor, and finite continued fractions

Euclid's algorithm determines that $\gcd(180, 64) = 4$ by performing the computations displayed in red in the following:

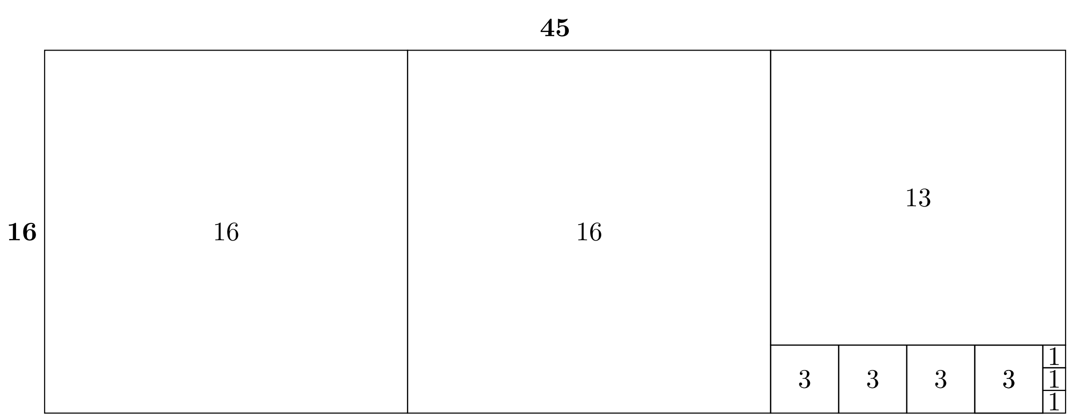
```
180     =     180 // 64 * 64 + 180 % 64     =     2 * 64 + 52
64      =      64 // 52 * 52 + 64  % 52     =     1 * 52 + 12
52      =      52 // 12 * 12 + 52  % 12     =     4 * 12 +  4
12      =      12 //  4 *  4 + 12  %  4     =     3 * 12 +  0
```

It corresponds to finding out that 4 is the size of the largest square thanks to which it is possible to pave a rectangle of size 180 by 64, based on the following geometric construction:



So when the gcd is 1, the paving of the rectangle can only be achieved with squares of size 1 by 1:

```
45     =     45  // 16 * 16 + 45 % 16     =     2 * 16 + 13
16     =     16  // 13 * 13 + 16 % 13     =     1 * 13 +  3
13     =     13  //  3 *  3 + 13 %  3     =     4 *  3 +  1
 3     =      3  //  1 *  3   + 3 %  1    =     3 *  1 +  0
```



Date: Session 1, 2017.

The blue part in both previous sets of equations is the same, and the pictures illustrate that

$$\frac{180}{64} = \frac{45}{16} = 2 + \cfrac{1}{1 + \frac{1}{4+\frac{1}{3}}}$$

The pictures illustrate that more generally, any rational number can be written as:

$$a_0 + 1/(a_1 + 1/(a_2 + \cdots + 1/a_k \overbrace{)\dots)}^{k}$$

where $a_0 \in \mathbb{Z}$, $k \in \mathbb{N}$, and $a_1, \dots, a_k \in \mathbb{N} \setminus \{0\}$ with $a_k \neq 1$, which is the general form of a finite continued fraction, that it is convenient to denote by $[a_0, a_1, a_2, \dots, a_k]$. Note that we could allow a finite continued fraction to end in 1 because for all $b \in \mathbb{N} \setminus \{0, 1\}$, $b = b - 1 + \frac{1}{1}$; that would make $[a_0, a_1, a_2, \dots, a_k - 1, 1]$ an alternative representation to $[a_0, a_1, a_2, \dots, a_k]$.

2. Computation of a finite continued fraction

More generally, given $k \in \mathbb{N} \setminus \{0\}$ and $r_1, \dots, r_k \in \mathbb{R}$ with $r_2, \dots, r_k$ at least equal to 1, let $[r_1, \dots, r_k]$ be defined as r_1 if $k = 1$, and as $r_1 + \frac{1}{[r_1, \dots, r_k]}$ if $k > 0$. For all $i \in \{-1, \dots, k\}$:

- let p_i be equal to 0 if $i = -1$, to 1 if $i = 0$, and to $r_k p_{k-1} + p_{k-2}$ if $k \geq 1$;
- let q_i be equal to 1 if $i = -1$, to 0 if $i = 0$, and to $r_k q_{k-1} + q_{k-2}$ if $k \geq 1$.

A trivial proof by induction shows that for all nonzero $j \leq k$, $q_k > 0$. Then:

$$[r_1, \dots, r_k] = \frac{p_k}{q_k} \tag{1}$$

which provides an effective method for computing $[r_1, \dots, r_k]$.

Towards proving (1), first define for all $j \leq k$ the matrix M_j as

$$\begin{bmatrix} p_j & q_j \\ p_{j-1} & q_{j-1} \end{bmatrix}$$

It is immediately verified by induction that for all nonzero $j \leq k$,

$$M_j = \begin{bmatrix} r_j & 1 \\ 1 & 0 \end{bmatrix} M_{j-1},$$

from which it follows that for all nonzero $j \leq k$,

$$\begin{bmatrix} p_j & q_j \\ p_{j-1} & q_{j-1} \end{bmatrix} = \begin{bmatrix} r_j & 1 \\ 1 & 0 \end{bmatrix} \dots \begin{bmatrix} r_1 & 1 \\ 1 & 0 \end{bmatrix} \tag{2}$$

As the transpose of the product of two matrixes A and B is the product of B by A, we have that for all nonzero $j \leq k$,

$$\begin{bmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{bmatrix} = \begin{bmatrix} r_1 & 1 \\ 1 & 0 \end{bmatrix} \dots \begin{bmatrix} r_j & 1 \\ 1 & 0 \end{bmatrix}$$

which implies that for all nonzero $j \leq k$,

$$\begin{bmatrix} p_j \\ q_j \end{bmatrix} = \begin{bmatrix} r_1 & 1 \\ 1 & 0 \end{bmatrix} \dots \begin{bmatrix} r_j & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{3}$$

Now proof of (1) is by induction on the length of finite continued fractions and application of (3). It is trivial that if $k = 1$ then (1) holds. Assume that $k > 1$. Denoting $[r_2, \dots, r_k]$ by $\frac{a}{b}$, we have by induction and (3) that

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r_2 & 1 \\ 1 & 0 \end{bmatrix} \dots \begin{bmatrix} r_k & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Set

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} r_2 & 1 \\ 1 & 0 \end{bmatrix} \dots \begin{bmatrix} r_k & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then $[r_1, \dots, r_k] = r_1 + \frac{1}{[r_2, \dots, r_k]} = r_1 + \frac{b}{a} = \frac{ar_1 + b}{a}$. It then follows from (3) again that

$$\begin{bmatrix} p_j \\ q_j \end{bmatrix} = \begin{bmatrix} r_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} ar_1 + b & cr_1 + d \\ a & c \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} ar_1 + b \\ a \end{bmatrix}$$

hence $[r_1, \dots, r_k] = \frac{p_1}{q_j}$, which completes the proof of (1).

3. Infinite continued fractions

Extend the notation of the previous section with $c_j = \frac{p_j}{q_j}$ for all strictly positive $j \leq k$. Then for all $j \in \{2, \ldots, k\}$, $c_j - c_{j-1}$ is equal to $\frac{p_j q_{j-1} - p_{j-1} q_j}{q_j q_{j-1}}$. Note that for all $j \leq k$, $p_j q_{j-1} - p_{j-1} q_j$ is the determinant of the matrix M_j, and it then follows from (3) that it is equal to $(-1)^r$. Hence for all strictly positive $j \leq k$,

$$c_j - c_{j-1} = \frac{(-1)^r}{q_j q_{j-1}}$$

Moreover, it is immediately verify by induction that $(q_j)_{2 \leq j \leq k}$ is a strictly increasing sequence. This shows that given $a_0 \in \mathbb{Z}$ and a sequence $(a_j)_{j \in \mathbb{N} \setminus \{0\}}$ of members of $\mathbb{N} \setminus \{0\}$, the sequence $([a_0, \ldots, a_j])_{j \in \mathbb{N}}$ converges; it is called an infinite continued fraction and it is denoted either as

$$a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

or as $[a_0, a_1, a_2, a_3 \ldots]$.

It follows from the previous observations that given an infinite continued fraction $[a_0, a_1, a_2, a_3 \ldots]$, $j \in \mathbb{N} \setminus \{0, 1\}$ and $n \in \mathbb{N}$, if $[a_0, \ldots, a_j]$ and $[a_0, \ldots, a_{j+1}]$ agree up to n digits after the decimal point, then $[a_0, \ldots, a_j]$ and $[a_0, a_1, a_2, a_3 \ldots]$ agree up to n digits after the decimal point. This allows one to compute exactly any approximation of $[a_0, a_1, a_2, a_3 \ldots]$.

4. Negating continued fractions

Obviously, for all $a \in \mathbb{Z}$, $-[a] = -a$ and $-[a, 2] = [-a-1, 2]$.

Given $a \in \mathbb{Z}$, $b \in \mathbb{N} \setminus \{0\}$, and $N \in \mathbb{N} \setminus \{0\} \cup \{\infty\}$,

$$-\left(a + \frac{1}{b + \frac{1}{N}}\right) = -a - 1 + 1 - \frac{1}{b + \frac{1}{N}} = -a - 1 + \frac{b + \frac{1}{N} - 1}{b + \frac{1}{N}} = -a - 1 + \frac{1}{\frac{b + \frac{1}{N}}{b + \frac{1}{N} - 1}} = -a - 1 + \frac{1}{1 + \frac{1}{b - 1 + \frac{1}{N}}}$$

Given $a \in \mathbb{Z}$, $b \in \mathbb{N} \setminus \{0\}$ and $N \in \mathbb{N} \setminus \{0\} \cup \{\infty\}$,

$$a + \frac{1}{0 + \frac{1}{b + \frac{1}{N}}} = a + b + \frac{1}{N}$$

It follows that, using $\ldots$ to denote the possibly missing terms of a finite or infinite continued fraction,

- for all $a \in \mathbb{Z}$ and $b \in \mathbb{N} \setminus \{0, 1, 2\}$, $-[a, b \ldots] = [-a-1, 1, b-1 \ldots]$;
- for all $a \in \mathbb{Z}$ and $c \in \mathbb{N} \setminus \{0\}$, $-[a, 2, c \ldots] = [-a-1, 1, 1, c \ldots]$;
- for all $a \in \mathbb{Z}$ and $c \in \mathbb{N} \setminus \{0\}$, $-[a, 1, c \ldots] == [-a-1, 1+c \ldots]$.

COMP9021 Principles of Programming