While a

Loop invariants is the topic of the problem [1] in this note.

[1] Problem 4 on page 9 from A. Engel. *Problem-Solving Strategies*. Problem Books in Mathematics. Springer New York, 2013. ISBN 9781475789546. URL https://books.google.com/books?id=aUofswEACAAJ

Problem

We start with the state (a, b) where a, b are positive integers. To this initial state we apply the following algorithm:

```
while a > 0:
    if a < b:
        (a,b) = (2a, b - a)
    else:
        (a,b) = (a - b, 2b)
```

For which starting positions does the algorithm stop? In how many steps does it stop, if it stops? What can you tell about periods and tails?

We start with $a > 0$ and $b > 0$. We adopt the following notation: a_i, b_i are the values after $i \in \mathbb{N}_{\geq 0}$ times through the loop. Before the first time through the loop $a_0 = a$, $b_0 = b$. Let $n = a + b$.

Let's collect some invariants. We will prove all of them by induction on $i \in \mathbb{N}_{\geq 0}$.

Invariant 1.1.

$$\forall i \geq 0 : a_i + b_i = n$$

Proof. Base case $a_0 + b_0 = a + b = n$ holds by definition of n and (a_0, b_0). Assume $a_i + b_i = n$. For $a_{i+1} + b_{i+1}$ we have two cases:

Case $a_i < b_i$: Here we have $a_{i+1} = 2a_i$ and $b_{i+1} = b_i - a_i$. So

$$a_{i+1} + b_{i+1} = 2a_i + b_i - a_i = a_i + b_i = n$$

Case $a_i \geq b_i$: In this case we have $a_{i+1} = a_i - b_i$ and $b_{i+1} = 2b_i$. It follows

$$a_{i+1} + b_{i+1} = a_i - b_i + 2b_i = a_i + b_i = n$$

□

Invariant 1.2.

$$\forall i \geq 0 : b_i > 0$$

Proof. This follows almost immediately from definitions [2]. □

[2] Base case $b_0 = b > 0$ holds by definition of b. Assume $b_i > 0$. Again we have two cases. If $a_i < b_i$ then $b_{i+1} = b_i - a_i > 0$. If $a_i \geq b_i$ then $b_{i+1} = 2b_i > 0$.

Invariant 1.3.

$$\forall i \geq 0 : a_i \geq 0$$

Proof. This also follows from definitions [3]. □

[3] Base case $a_0 = a > 0$ holds by definition of a. Assume $a_i \geq 0$. Again we have two cases. If $a_i < b_i$ then $a_{i+1} = 2a_i \geq 0$. If $a_i \geq b_i$ then $a_{i+1} = a_i - b_i \geq 0$.

Invariant 1.4.

$$\forall i \geq 0 : a_i \equiv 2^i a \mod n$$

Proof. Base case $a_0 = a = 2^0 a$ trivially holds. Assume $a_i \equiv 2^i a \mod n$. For a_{i+1} we have two cases:

Case $a_i < b_i$: Here we have $a_{i+1} = 2a_i$. So

$$\begin{aligned} a_{i+1} &= 2a_i \\ &\equiv 2 \cdot 2^i a \mod n \\ &\equiv 2^{i+1} a \mod n \end{aligned}$$

Case $a_i \geq b_i$: In this case we have $a_{i+1} = a_i - b_i$. It follows

$$\begin{aligned} a_{i+1} &= a_i - b_i \\ &\equiv a_i + n - b_i \mod n \\ &\equiv a_i + a_i + b_i - b_i \mod n \\ &\equiv 2a_i \mod n \\ &\equiv 2 \cdot 2^i a \mod n \\ &\equiv 2^{i+1} a \mod n \end{aligned}$$

□

We will use these 4 invariants ($a_i \geq 0$, $b_i > 0$, $a_i + b_i = n$ and $a_i \equiv 2^i a \mod n$) to determine for which initial values a and b the loop terminates. To do so we consider $\frac{a}{n}$. Because $0 < a < n$ we know that $\frac{a}{n} \in (0, 1)$. We look at the expansion of $\frac{a}{n}$ in base 2.

Theorem 1.1. *If the expansion of $\frac{a}{n}$ is finite with k digits $d_i \in \{0, 1\}$*

$$\frac{a}{n} = \sum_{i=1}^{k} d_i 2^{-i}$$

then $a_k = 0$ and the loop terminates after k steps.

Proof. From

$$\frac{a}{n} = \sum_{i=1}^{k} d_i 2^{-i}$$

we get by multiplying both sides with $2^k n$:

$$2^k a = \sum_{i=1}^{k} n d_i 2^{k-i} \equiv 0 \mod n$$

Together with invariant 1.4 we get

$$a_k \equiv 2^k a \equiv 0 \mod n$$

and because $a_k \geq 0$, $b_k > 0$, $a_k + b_k = n$ we know that $0 \leq a_k < n$, so it must be that $a_k = 0$ and the loop terminates after at most k steps. To show that the loop terminates after exactly k steps, we need to show that $a_j > 0$ for $0 \leq j < k$. We will do this by finding a contradiction. Assume there exists a $j < k$ such that $a_j = 0$. Then it also holds that $2^j a \equiv 0 \mod n$.

From

$$\frac{a}{n} = \sum_{i=1}^{k} d_i 2^{-i}$$

we get by multiplying both sides with $2^j n$:

$$2^j a = \sum_{i=1}^{k} n d_i 2^{j-i} = \sum_{i=1}^{j} n d_i 2^{j-i} + \sum_{i=j+1}^{k} n d_i 2^{j-i} \equiv 0 \mod n$$

$2^j a \equiv 0 \mod n$, so $2^j a = nq$ for some $q \in \mathbb{Z}$. Then

$$q = \sum_{i=1}^{j} d_i 2^{j-i} + \sum_{i=j+1}^{k} d_i 2^{j-i}$$

We have $q \in \mathbb{Z}$, $\sum_{i=1}^{j} d_i 2^{j-i} \in \mathbb{Z}$, but $\sum_{i=j+1}^{k} d_i 2^{j-i} \notin \mathbb{Z}$, because $d_i \in \{0, 1\}$. This is a contradiction.

□

We arrived at a neat result: if the binary expansion of $\frac{a}{a+b}$ is finite with k digits, then the loop terminates after k steps.

What can we say if the expansion is not finite but instead has a repeating pattern with a prefix and a period (the only other option [4]) ? For starters, we can use a contradiction similar to the earlier one to prove that the loop does not terminate. Consider the infinite binary expansion:

[4] That is because $\frac{a}{a+b} \in \mathbb{Q}$. See below for why.

$$\frac{a}{n} = \sum_{i=1}^{\infty} d_i 2^{-i}$$

Assume there is a k for which $a_k = 0$. Then by multiplying the expansion with $2^k n$ we get:

$$2^k a = \sum_{i=1}^{k} n d_i 2^{k-i} + \sum_{i=k+1}^{\infty} n d_i 2^{k-i} \equiv 0 \mod n$$

So for some $q \in \mathbb{Z}$ such that $2^k a = nq$ we have

$$q = \sum_{i=1}^{k} d_i 2^{k-i} + \sum_{i=k+1}^{\infty} d_i 2^{k-i}$$

The left side and the first sum on the right both belong to $\mathbb{Z}$ but the second sum does not, which is a contradiction. This means, that $\forall k : a_k > 0$ and the loop does not terminate.

At this point we will do a small digression and prove some theorems about decimal expansion.

Theorem 1.2. *Given an integer $p > 1$, the series*

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

with $d_i \in \{0, 1, \ldots, p-1\}$ converges to a value $x \in [0, 1]$.

Proof.
$$\sum_{i=1}^{n} \frac{d_i}{p^i} \leq \sum_{i=1}^{n} \frac{p-1}{p^i} \xrightarrow[n \to \infty]{} 1$$

so the series is bounded and will converge. □

Theorem 1.3. *For every $x \in [0, 1]$ there exists a decimal expansion with base $p > 1$ such that*

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

with $d_i \in \{0, 1, \ldots, p-1\}$.

Proof. We divide the interval $[0, 1]$ into p intervals $[\frac{i}{p}, \frac{i+1}{p}]$ with $0 \leq i < p$. Since $[0, 1] = \bigcup_{i=0}^{p-1} [\frac{i}{p}, \frac{i+1}{p}]$ we know there exists at least one index i with $x \in [\frac{i}{p}, \frac{i+1}{p}]$. We set $d_1 = i$ and subdivide $[\frac{i}{p}, \frac{i+1}{p}]$ into p segments $[\frac{i}{p}, \frac{i+1}{p}] = \bigcup_{j=0}^{p-1} [\frac{d_1}{p} + \frac{j}{p^2}, \frac{d_1}{p} + \frac{j+1}{p^2}]$. x is in one of these subintervals and we set d_2 to be the index of that subinterval and continue in this manner recursively defining all d_i. Because of the nested interval property with monotone decreasing length this converges to x.

Another way to prove it is like this:

The case where $x = 0$ is trivial (just set all $d_i = 0$).

For $x > 0$ we have:

The set $N_1 = \{k \in \mathbb{N}_0 : \frac{k}{p} < x\}$ is a set of non-negative integers strictly bounded above by p, so it has a largest element and we set $d_1 = \max(N_1)$. Then $x \leq \frac{d_1+1}{p}$ (otherwise $d_1 + 1 \in N_1$ and d_1 wouldn't be the largest element of N_1). We therefore have

$$\frac{d_1}{p} < x \leq \frac{d_1 + 1}{p}$$

We continue and look at $N_2 = \{k \in \mathbb{N}_0 : \frac{d_1}{p} + \frac{k}{p^2} < x\}$. Again the set N_2 is strictly bounded above by p and we set $d_2 = \max(N_2)$. Again we have:

$$\frac{d_1}{p} + \frac{d_2}{p^2} < x \leq \frac{d_1}{p} + \frac{d_2 + 1}{p^2}$$

Having defined $d_1, d_2, \ldots d_{n-1}$ we can recursively define $d_n = \max(N_n)$ with

$$N_n = \{k \in \mathbb{N}_0 : \sum_{i=1}^{n-1} \frac{d_i}{p^i} + \frac{k}{p^n} < x\}$$

Again $p \notin N_n$, so the definition is valid and the following inequalities hold:

$$\sum_{i=1}^{n} \frac{d_i}{p^i} < x \leq \sum_{i=1}^{n-1} \frac{d_i}{p^i} + \frac{d_n + 1}{p^n}$$

We define $u_n = \sum_{i=1}^{n} \frac{d_i}{p^i}$, $v_n = \sum_{i=1}^{n-1} \frac{d_i}{p^i} + \frac{d_n+1}{p^n}$ and $w_n = \frac{d_{n+1}+1}{p^{n+1}}$. u_n is monotone increasing and bounded above, so it converges. For v_n we have

$$\begin{aligned}
&v_n \geq v_{n+1} \\
&\Leftrightarrow \sum_{i=1}^{n-1} \frac{d_i}{p^i} + \frac{d_n + 1}{p^n} \geq \sum_{i=1}^{n} \frac{d_i}{p^i} + \frac{d_{n+1} + 1}{p^{n+1}} \\
&\Leftrightarrow \frac{d_n + 1}{p^n} \geq \frac{d_n}{p^n} + \frac{d_{n+1} + 1}{p^{n+1}} \\
&\Leftrightarrow \frac{1}{p^n} \geq \frac{d_{n+1} + 1}{p^{n+1}} \\
&\Leftrightarrow p \geq d_{n+1} + 1
\end{aligned}$$

which holds by definition of d_{n+1}. So v_n is monotone decreasing and bounded below, therefore it converges too. w_n converges to zero and $v_n = u_{n-1} + w_n$ therefore

$$lim_{n\to\infty}u_n = lim_{n\to\infty}v_n = x$$

□

Theorem 1.4. *Given is base $p > 1$ and*

$$x = \sum_{i=1}^{n} \frac{d_i}{p^i}$$

with $d_i \in \{0, 1, \dots, p-1\}$ and $d_n \neq 0$. Then there are two base p expansions of x.

Proof. The first expansion is $x = \sum_{i=1}^{\infty} \frac{d_i}{p^i}$ with $d_i = 0$ for $i > n$. For the second expansion we define the following series:

$$y = \sum_{i=1}^{n-1} \frac{d_i}{p^i} + \frac{d_n - 1}{p^n} + \sum_{i=n+1}^{\infty} \frac{p-1}{p^i}$$

and prove that $y = x$. Then the two expansions are $0.d_1d_2\dots d_n00000\dots$ and $0.d_1d_2\dots(d_n-1)(p-1)(p-1)(p-1)\dots$

To prove that $y = x$ we look at

$$\begin{aligned}
\sum_{i=n+1}^{\infty} \frac{p-1}{p^i} &= \frac{p-1}{p^n} \sum_{i=1}^{\infty} \frac{1}{p^i} \\
&= \frac{p-1}{p^n} (\sum_{i=0}^{\infty} \frac{1}{p^i} - 1) \\
&= \frac{p-1}{p^n} (\frac{p}{p-1} - 1) \\
&= \frac{p-1}{p^n} \frac{1}{p-1} \\
&= \frac{1}{p^n}
\end{aligned}$$

So y becomes

$$y = x - \frac{1}{p^n} + \frac{1}{p^n} = x$$

□

Theorem 1.5. *If we disallow series with infinitely repeated $(p-1)$ tail, any $x \in [0, 1]$ has a unique decimal expansion in base p.*

Proof. Assume two decimal expansions where both agree until index $k-1$ and index k is the first index where they differ.

$$x = \sum_{i=1}^{k-1} \frac{d_i}{p^i} + \frac{e_k}{p^k} + \sum_{i=k+1}^{\infty} \frac{e_i}{p^i}$$
$$y = \sum_{i=1}^{k-1} \frac{d_i}{p^i} + \frac{f_k}{p^k} + \sum_{i=k+1}^{\infty} \frac{f_i}{p^i}$$

Without loss of generality assume $e_k < f_k$.
We have

$$\begin{aligned} y - x &= \sum_{i=1}^{k-1} \frac{d_i}{p^i} + \frac{f_k}{p^k} + \sum_{i=k+1}^{\infty} \frac{f_i}{p^i} - \sum_{i=1}^{k-1} \frac{d_i}{p^i} - \frac{e_k}{p^k} - \sum_{i=k+1}^{\infty} \frac{e_i}{p^i} \\ &= \frac{f_k - e_k}{p^k} + \sum_{i=k+1}^{\infty} \frac{f_i}{p^i} - \sum_{i=k+1}^{\infty} \frac{e_i}{p^i} \\ &= \frac{f_k - e_k}{p^k} + \frac{1}{p^k}(\sum_{i=1}^{\infty} \frac{f_{k+i}}{p^i} - \sum_{i=1}^{\infty} \frac{e_{k+i}}{p^i}) \end{aligned}$$

We denote $u = \sum_{i=1}^{\infty} \frac{f_{k+i}}{p^i}$ and $v = \sum_{i=1}^{\infty} \frac{f_{k+i}}{p^i}$. Since we disallowed repeated $(p-1)$ tail, we know that $0 \leq u < 1$ and $0 \leq v < 1$, so $-1 < u - v < 1$. It follows that

$$0 \leq \frac{f_k - e_k - 1}{p^k} < y - x < \frac{f_k - e_k + 1}{p^k}$$

and $x \neq y$. □

Theorem 1.6. $x \in [0,1] \cap \mathbb{Q}$ *if and only if its decimal expansion in base* $p > 1$ *is either finite or has a prefix (of length zero or more) and an infinitely repeating non-zero length pattern tail.*

Proof.
($\Rightarrow$):
$x \in [0,1] \cap \mathbb{Q}$, so there exist $m, n \in \mathbb{N}$ with $m < n$ and $x = \frac{m}{n}$. We basically do the long division and present an expansion that will have a repeating tail (if it isn't finite). Let $k \in \mathbb{N}$ be the smallest integer such that $mp^k \geq n$ and we do division:

$$mp^k = nq + r$$

with $0 \leq r < n$. Because k is the smallest integer with $mp^k \geq n$ we have $np > mp^k$ (otherwise $k-1$ would be a smaller integer satisfying the same). That means $np > nq + r$ and thus $p > \frac{np-r}{n} > q$. This gives us $k-1$ zeros and the first non-zero digit in the expansion, namely q:

$$\begin{aligned}\frac{m}{n} &= \frac{1}{p^k}\frac{mp^k}{n} \\ &= \frac{1}{p^k}\frac{nq+r}{n} \\ &= \frac{q}{p^k} + \frac{r}{n}\end{aligned}$$

We repeat this process with $\frac{r}{n}$. There are only n possible remainders, so if it doesn't end with a remainder of zero it must eventually get a previously seen remainder and so the expansion will repeat itself. This creates an expansion with an infinitely repeating non-zero length pattern tail. Since it isn't finite, we can disallow repeating $(p-1)$ and from the expansion uniqueness theorem we have proved the $(\Rightarrow)$ direction.

$(\Leftarrow)$:

This direction is easy. If it is a finite sum, then it is rational since all the parts are rational. If it is infinite repeating we can eliminate the non-repeating prefix since it is finite and rational and shift the rest. So we can concentrate on a repeating series with a period of length $k-1$:

$$\begin{aligned}x &= \sum_{i=0}^{\infty}\left(\frac{1}{p^{ki}}\sum_{j=1}^{k-1}\frac{d_j}{p^j}\right) \\ &= \left(\sum_{j=1}^{k-1}\frac{d_j}{p^j}\right)\sum_{i=0}^{\infty}\frac{1}{p^{ki}} \\ &= \left(\sum_{j=1}^{k-1}\frac{d_j}{p^j}\right)\left(1+\sum_{i=1}^{\infty}\frac{1}{p^{ki}}\right) \\ &= \left(\sum_{j=1}^{k-1}\frac{d_j}{p^j}\right)\left(1+\sum_{i=1}^{\infty}\left(\frac{1}{p^k}\right)^i\right) \\ &= \left(\sum_{j=1}^{k-1}\frac{d_j}{p^j}\right)\left(1+\frac{p^k}{p^k-1}\right)\end{aligned}$$

which is a rational expression. □

We return to our problem. We now know the expansion of $\frac{a}{a+b}$ is repeating a period if it doesn't terminate. We will show that the loop also repeats a period of the same length.

Theorem 1.7. *If $\frac{a}{n}$ has an expansion in base p which repeats a period of k digits infinitely, then*

$$ap^k \equiv a \mod n$$

Proof. We have $\frac{a}{n} = 0.\overline{d_1 d_2 d_3 \ldots d_k}$ which means

$$\begin{aligned}
\frac{a}{n} &= 0.\overline{d_1 d_2 d_3 \ldots d_k} \\
&= \sum_{i=1}^{k} \frac{d_i}{p^i} + \frac{1}{p^k}(\sum_{i=1}^{k} \frac{d_i}{p^i} + \frac{1}{p^k}(\sum_{i=1}^{k} \frac{d_i}{p^i} + \ldots \\
&= \sum_{i=1}^{k} \frac{d_i}{p^i} + \frac{1}{p^k}\frac{a}{n}
\end{aligned}$$

We multiply both sides by np^k and get

$$ap^k = \sum_{i=1}^{k} nd_i p^{k-i} + a$$

which proves the theorem. □

Theorem 1.8. *If $\frac{a}{n}$ has an expansion in base p which has a prefix and then repeats a period of k digits infinitely, then*

$$ap^k \equiv a \mod n$$

Proof. We have $\frac{a}{n} = 0.e_1 e_2 e_3 \ldots e_l \overline{d_1 d_2 d_3 \ldots d_k}$ which means

$$\begin{aligned}
\frac{a}{n} &= 0.e_1 e_2 e_3 \ldots e_l \overline{d_1 d_2 d_3 \ldots d_k} \\
&= \sum_{i=1}^{l} \frac{e_i}{p^i} + \frac{1}{p^l}(0.\overline{d_1 d_2 d_3 \ldots d_k})
\end{aligned}$$

This means

$$\frac{ap^l - \sum_{i=1}^{l} p^{l-i} e_i n}{n} = 0.\overline{d_1 d_2 d_3 \ldots d_k}$$

We can then apply the previous theorem to a new $a' := ap^l - \sum_{i=1}^{l} p^{l-i} e_i n$ and see that

$$a' p^k \equiv a' \mod n$$

But $a' \equiv ap^l \mod n$, so

$$ap^{k+l} \equiv ap^l \mod n$$

or $ap^k \equiv a \mod n$. □

We combine this last result with the invariant 1.4 to see that $a_{i+k} = a_i$ and the loop repeats values with period k.

Bibliography

A. Engel. *Problem-Solving Strategies*. Problem Books in Mathematics. Springer New York, 2013. ISBN 9781475789546. URL https://books.google.com/books?id=aUofswEACAAJ.