Ducci Sequences

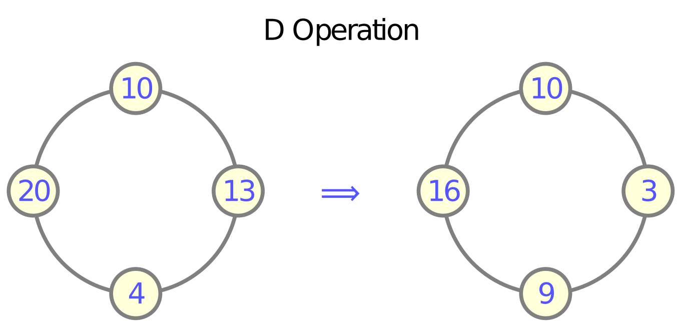
2023/06/05 v0.6

1 Classical Ducci

Here is a classical if slightly frivolous example of an apparently simple iteration that has lots of interesting properties.

> Place 4 integers on a circle. Compute the absolute values of all differences of adjacent pairs of numbers. Write these values between the corresponding numbers and erase the numbers themselves.
>
> Repeat. What happens?

This puzzle was introduced by Enrico Ducci (1864-1940) in the 1930s, and was mostly forgotten till Honsberger's great booklet "Ingenuity in Mathematics" appeared in 1970. As a first step towards understanding what is going on, note that we can ignore integers, after one step we are dealing only with natural numbers. Here is a picture of the operation.



1.1 Ducci's Theorem

Less informally, we want to iterate the following function on $\mathbb{N}^4$:

$$D(x_1, x_2, x_3, x_4) = (|x_1 - x_4|, |x_2 - x_1|, |x_3 - x_2|, |x_4 - x_3|)$$

Is there anything one can say about repeated application of this function? What is $D^t(x_1, x_2, x_3, x_4)$? Here are two small examples, we iterate D on "random" initial conditions.

0	10	13	4	20
1	10	3	9	16
2	6	7	6	7
3	1	1	1	1
4	0	0	0	0

0	94	68	11	85
1	9	26	57	74
2	65	17	31	17
3	48	48	14	14
4	34	0	34	0
5	34	34	34	34
6	0	0	0	0

After a few steps we reach the zero vector, which is a fixed point of the operation. Just two experiments are obviously not enough to form a conjecture, but more attempts always seem to produce the same result.

Before we try to establish the conjecture, let's clean up notation a bit. We write $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ for a vector of natural numbers, and $\mathbf{0} = (0,0,0,0)$ for the zero vector. $z\,\boldsymbol{x}$ means: multiply all the components of $\boldsymbol{x}$ by z, and similarly for division $\boldsymbol{x}/z$. Naturally, $\gcd \boldsymbol{x} = \gcd(x_1, x_2, x_3, x_4)$.

The first observation is that all orbits end in fixed point $\mathbf{0}$.

Theorem 1.1 (Ducci) *All orbits under D end in the fixed point $\mathbf{0}$.*

So we claim that all orbits look like so:

$$\boldsymbol{a}, D(\boldsymbol{a}), D^2(\boldsymbol{a}), \ldots, D^{t-1}(\boldsymbol{a}), D^t(\boldsymbol{a}) = \mathbf{0}$$

where t is sufficiently large. How would we go about proving this?

Perhaps the most natural approach is to show that the vectors "shrink" under repeated application of D, so that ultimately we must hit $\mathbf{0}$. To make this idea more precise, call the maximum element of a vector its *weight*. From the definition of D, weights cannot increase under an application of D. Unfortunately, the weight may or may not decrease during a single application of D; just think about $(1,0,0,0)$. A little fumbling shows that weight can be preserved only if the vector looks like $(0,a,b,c)$ where $a \geq b, c$, up to rotation and reflection. By this we mean that, for our purposes, any argument (x_1, x_2, x_3, x_4) can be replaced by (x_2, x_3, x_4, x_1) or (x_4, x_3, x_2, x_1). Now we can do a straightforward but exceedingly boring case analysis. Go ahead and try it.

The argument by cases is really unpleasant, it's just the kind of proof any sane person would like to avoid. Here is a much better and much less obvious attack, based on extracting common factors from the argument vector whenever possible. If the vector elements are coprime; nothing changes; otherwise, we reduce the weight of the vector.

We need to keep track of the common factors we pull out. To this end, define $E : \mathbb{N}^4 \times \mathbb{N} \to \mathbb{N}^4 \times \mathbb{N}$ by $E(\boldsymbol{x}, c) = (D(\boldsymbol{x})/d, c\,d)$, where $d = \gcd D(\boldsymbol{x})$, if $D(\boldsymbol{x}) \neq \mathbf{0}$. Otherwise, set $E(\boldsymbol{x}, c) =$

$(\mathbf{0}, c)$. An easy induction shows that

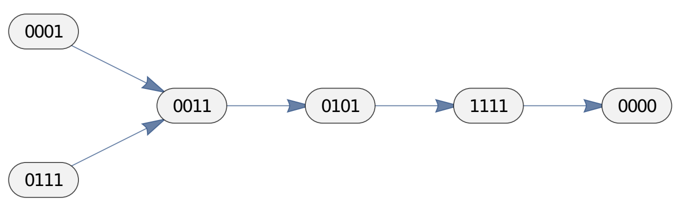
$$E^t(\boldsymbol{x}, 1) = (D^t(\boldsymbol{x})/d, d) \qquad \text{where } d = \gcd D^t(\boldsymbol{x})$$

as long as $D^t(\boldsymbol{x}) \neq \mathbf{0}$. But then D converges to $\mathbf{0}$ iff E converges to some $(\mathbf{0}, d)$. Here is an example, with both operations side by side.

t	$D^t(\boldsymbol{x})$	$E^t(\boldsymbol{x}, 1)$	d
0	$(4, 7, 13, 24)$	$(4, 7, 13, 24)$	1
1	$(20, 3, 6, 11)$	$(20, 3, 6, 11)$	1
2	$(9, 17, 3, 5)$	$(9, 17, 3, 5)$	1
3	$(4, 8, 14, 2)$	$(2, 4, 7, 1)$	2
4	$(2, 4, 6, 12)$	$(1, 2, 3, 6)$	2
5	$(10, 2, 2, 6)$	$(5, 1, 1, 3)$	2
6	$(4, 8, 0, 4)$	$(1, 2, 0, 1)$	4
7	$(0, 4, 8, 4)$	$(0, 1, 2, 1)$	4
8	$(4, 4, 4, 4)$	$(1, 1, 1, 1)$	4
9	$(0, 0, 0, 0)$	$(0, 0, 0, 0)$	4

Here the common factors are always 1 or 2, but that is just coincidence.

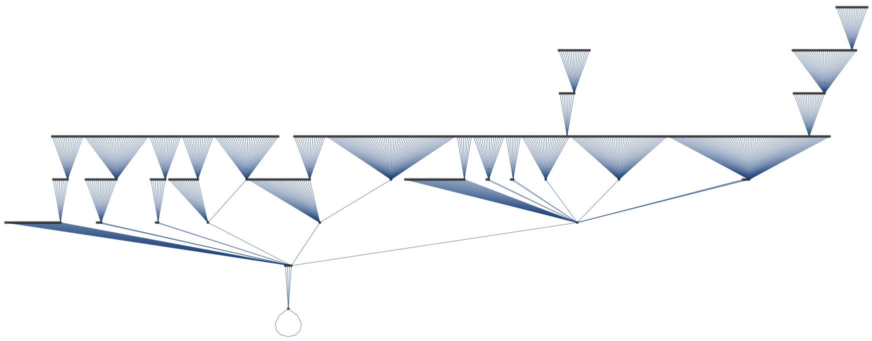
OK, so we have a slightly better way of thinking about repeated applications of D, but how do we know that E converges? Focus on the case where the common factor is 2, all the elements are even. If we ignore the actual numerical values and consider only the parity of the numbers, as before up to rotations and reflections, a single application of D has the effect expressed in the following diagram:



Make sure you understand why the diagram is correct. But then we can pull out a factor of 2 after at most 4 intermediate steps, done.

1.2 Counting Steps

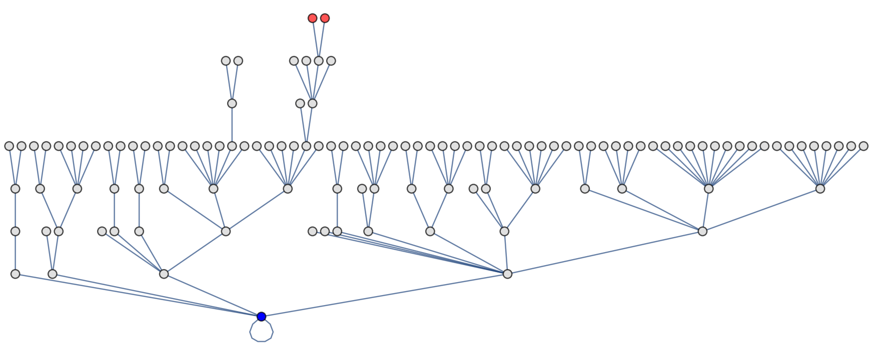
The last proof shows that, no matter which starting vector $\boldsymbol{a}$ we choose, after a sufficient number of applications of D, we wind up in the fixed point $\mathbf{0}$. That immediately brings up the question of haw many steps it takes to get there. Let's call this number the stopping time for $\boldsymbol{a}$, in symbols $\sigma(\boldsymbol{a})$. A good way of phrasing the stopping time question is to think about the Ducci tree, a graph whose vertices are all vectors $\boldsymbol{x} \in \mathbb{N}^4$ and whose edges look like $\boldsymbol{x} \to D(\boldsymbol{x})$. The stopping time is then none other than the distance of $\boldsymbol{a}$ to $\mathbf{0}$ in the tree. It would be nice to have a picture of the tree, at least for vectors of limited weight. Unfortunately, the number of nodes of weight at most n is $(n+1)^4$, so things get out of hand very quickly. Here is the picture for weight at most 4.



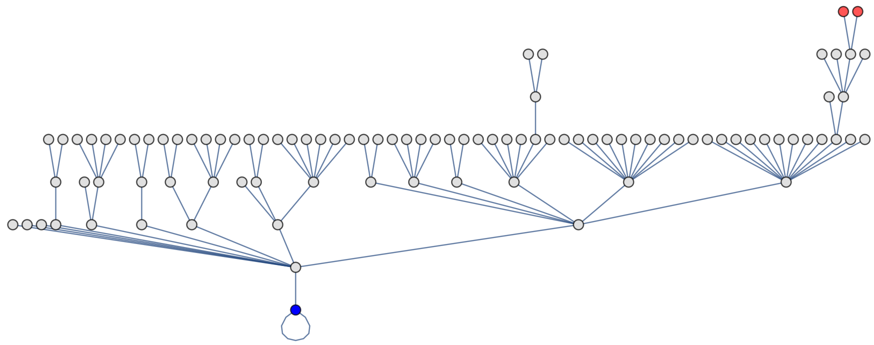
Pretty, but not too terribly useful. We need to filter out useless information. One way to do this to return to our casual comment that "rotations and reflections" don't matter. What we mean by this more precisely is the following: write L for a cyclic rotation to the left, R for a cyclic rotation to the right and rev for reversing a vector. Thus rev is an involution (applying rev twice changes nothing) and, more interestingly, $L(R(x)) = R(L(x)) = x$ and $L(\mathsf{rev}(x)) = \mathsf{rev}(R(x))$. Then one can verify that

$$D(x) = R(D(L(x)))$$
$$D(x) = R(\mathsf{rev}(D(\mathsf{rev}(x))))$$

Hence $\sigma(x) = \sigma(L(x)) = \sigma(R(x)) = \sigma(\mathsf{rev}(x))$. If we are only interested in stopping times, we can can lump together all vectors that can be obtained from each other using rotations and reflections. This cuts down on the number of nodes in the Ducci tree and can helps in producing images. For example, there are 625 vectors of weight at most 4, but, if we don't distinguish between variants, there are only 120 to deal with. The following picture shows the improved Ducci tree up to weight 4.



Much better. The depth of the tree is 7, as witnessed by vectors $(0, 2, 3, 4)$ and $(0, 1, 2, 4)$. Can we do even better than this? Recall from the first section that we can pull out common factors without changing the stopping time, and we might as well combine that simplification with our rotation/reflection method to further reduce the number of nodes. This time, we wind up with just 95 nodes.



A minor improvement, the root now has just one child where there used to be 4. Still, not overwhelming. If we want to consider larger weights, we will have to confine ourselves to individual experiments. For instance, here are the minimal (in the lexicographic sense) vectors of weight at most 10 that have stopping times from 0 to 8:

$$(0,0,0,0),(1,1,1,1),(0,1,0,1),(0,0,1,1),(0,0,0,1),(0,1,2,3),(0,0,1,3),(0,1,2,4),(0,1,4,9)$$

Note that weight 9 actually suffices to get to 8 steps.

One might feel firmly convinced that stopping times must become arbitrarily large for some vectors. Alas, it is not so easy to produce a vector that takes, say, 1000 steps to get to $\mathbf{0}$. Try it, simply choosing large numbers in the vector won't work: they tend to cancel each other out very quickly.

Lemma 1.1 (Webb 1982) *There is no bound on the number of steps needed to reach* $\mathbf{0}$.

The idea is to consider a family $(\boldsymbol{a}_n)$ of special vectors with the hope that $D(\boldsymbol{a}_n) = \boldsymbol{a}_{n-1}$. This will not quite work out, but we will get close enough. To construct initial conditions leading to long transients, consider "tribonacci numbers"

$$T_n = T_{n-1} + T_{n-2} + T_{n-3} \qquad T_0 = 0 \qquad T_1 = T_2 = 1$$

These are modeled after the famous Fibonacci numbers $F_n = F_{n-1} + F_{n-2}$, $F_0 = 0$, $F_1 = 1$. The first few tribonacci values are

$$0, 1, 1, 2, 4, 7, 13, 24, 44, 81, 149, 274, 504, 927, 1705, 3136, \ldots$$

Now consider the vectors $\boldsymbol{a}_n = (T_{n-3}, T_{n-2}, T_{n-1}, T_n)$. A little bit of work shows

$$D^3(\boldsymbol{a}_n) = 2 \cdot L(\boldsymbol{a}_{n-2})$$

where L is again the cyclic shift to the left. This means $\sigma(\boldsymbol{a}_n) = \sigma(\boldsymbol{a}_{n-2}) + 3$ and thus $\sigma(\boldsymbol{a}_{2n}) = 3(n+1)$. Looking back at the D versus E table above, you may notice that our initial vector is none other than $\boldsymbol{a}_7$. As predicted, in row 3 we find $\boldsymbol{a}_5$ rotated, at least in the E column.

2 Generalized Ducci

There are two obvious ways we could try to generalize Ducci sequences:

- Replace the natural numbers by, say, the reals.
- Change the length of the vectors.

And we could consider combinations of both, but life is short.

2.1 Real Vectors

Switching the argument vectors to $\mathbb{R}^4$ means we have to deal with irrational or even transcendental numbers like $\sqrt{2}$ or π. One might suspect that the switch could easily lead to divergence; after all, these numbers are hugely more complicated than the naturals. Here is one attempt that fails spectacularly.

$$\begin{array}{c|llll}
0 & 0 & \sqrt{2} & \sqrt{7} & \pi \\
1 & \pi & \sqrt{2} & -\sqrt{2}+\sqrt{7} & -\sqrt{7}+\pi \\
2 & \sqrt{7} & -\sqrt{2}+\pi & 2\sqrt{2}-\sqrt{7} & -\sqrt{2}+2\sqrt{7}-\pi \\
3 & \sqrt{2}-\sqrt{7}+\pi & \sqrt{2}+\sqrt{7}-\pi & -3\sqrt{2}+\sqrt{7}+\pi & -3\sqrt{2}+3\sqrt{7}-\pi \\
4 & 4\sqrt{2}-4\sqrt{7}+2\pi & -2\sqrt{7}+2\pi & -4\sqrt{2}+2\pi & -2\sqrt{7}+2\pi \\
5 & 4\sqrt{2}-2\sqrt{7} & 4\sqrt{2}-2\sqrt{7} & 4\sqrt{2}-2\sqrt{7} & 4\sqrt{2}-2\sqrt{7} \\
6 & 0 & 0 & 0 & 0
\end{array}$$

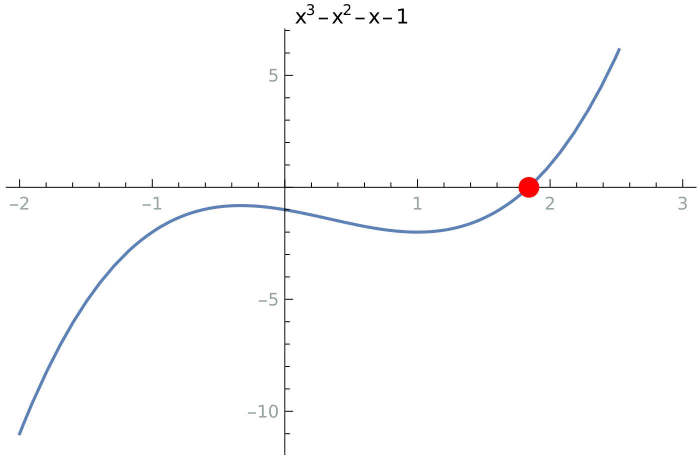
Note how throwing the transcendental number π into the mix does not help at all, never mind the irrational roots. After a few more experiments along these lines one might reluctantly conclude that convergence to $\mathbf{0}$ always occurs, no matter which starting vector in $\mathbb{R}^4$ we choose. Alas, that is false, though it is rather tricky to come up with a counterexample.

Theorem 2.1 (Lotan 1949) *There is a real vector $\boldsymbol{a} \in \mathbb{R}^4$ that produces a divergent orbit under D.*

The construction of this vector goes like this: let q be the real root of $x^3 - x^2 - x - 1 = 0$, so

$$q = \frac{1}{3}\left(1 + \sqrt[3]{19 - 3\sqrt{33}} + \sqrt[3]{19 + 3\sqrt{33}}\right) \approx 1.83929$$

If you worry about the existence of this root, here is a calming picture.



Now for the big surprise: the vector $(1, q, q^2, q^3)$ does not converge. E.g., after 5 steps we get a term

$$\frac{1}{9}\left(24 - 4\sqrt[3]{19+3\sqrt{33}} - 6\left(19+3\sqrt{33}\right)^{2/3} + \left(19-3\sqrt{33}\right)^{2/3} \cdot \left(-6+4\sqrt[3]{19+3\sqrt{33}}\right) + 4\sqrt[3]{19-3\sqrt{33}}\left(-1+\left(19+3\sqrt{33}\right)^{2/3}\right)\right)$$

The further we go, the worse these expressions become. Of course, that is not a proof of non-convergence.

2.2 Different Dimensions

Next consider vectors in $\mathbb{N}^d$; we refer to d as the dimension of the Ducci sequence. So for dimension 4 we always have convergence to $\mathbf{0}$. Does your intuition tell you anything about what might happen with other dimensions? We claim that, regardless of dimension, every orbit must be ultimately periodic (though it need not necessarily reach the fixed point $\mathbf{0}$). To see why, recall our weight argument from section 1. Since weight cannot increase, there are only finitely many vectors in play, so, for any particular starting vector $\boldsymbol{a}$, we are really dealing with a finite domain.

The cheapest non-boring example is $d = 5$. We have broken the orbit into the transient and the periodic part. It takes 12 steps to reach the limit cycle.

0	1	2	3	4	5	6	7	8	9	10	11	12
4	1	1	6	4	3	2	2	2	1	1	1	1
6	2	1	0	6	2	1	1	1	1	0	1	0
10	4	2	1	1	5	3	2	1	0	1	1	0
1	9	5	3	2	1	4	1	1	0	0	1	0
3	2	7	2	1	1	0	4	3	2	2	2	1

The limit cycle has length 15 and looks rather unexpected.

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	0	1	0	1	1	0	0	1	0	0	0	1	1	1	1
0	1	1	0	0	1	0	0	0	1	1	1	1	0	1	0
0	0	1	0	0	0	1	1	1	1	0	1	0	1	1	0
0	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0
1	1	1	1	0	1	0	1	1	0	0	1	0	0	0	1

All vector components on the limit cycle are either 0 or 1. Could this be coincidence?

Lemma 2.1 *All Ducci sequences over the integers are ultimately periodic. If the limit cycle is not* $\mathbf{0}$*, then the vectors on the limit cycle are two-valued in the sense that* $x_i \in \{0, c\}$ *for some* c.

Proof.

Let's start with a general observation: $|a-b| \leq \max(a, b)$ with equality only in the case where at least one of a or b is 0. As already mentioned, the weight argument shows that the orbits must be ultimately periodic. Let $c > 0$ be the weight in the periodic part and $\boldsymbol{x}$ its starting vector. Suppose that there is some intermediate value b that appears in $\boldsymbol{x}$, say, $0 < b < c$. By rotation we may assume that $x_1 = b$. A simple induction shows that $D^t(\boldsymbol{x})_j < c$ for all $1 \leq j \leq t+1$. But then $D^{d-1}(\boldsymbol{x})$ has weight less than c, contradiction.

□

Recall the trick of pulling out common factors from above. Any vector with entries in $\{0, c\}$ turns into a binary vector when we pull out the factor c, and we can give a very simple description of the way D operates on binary vectors. In fact, on binary vectors, the Ducci operator degenerates into exclusive or:

$$D(\boldsymbol{x}) = \boldsymbol{x} \oplus R(\boldsymbol{x})$$

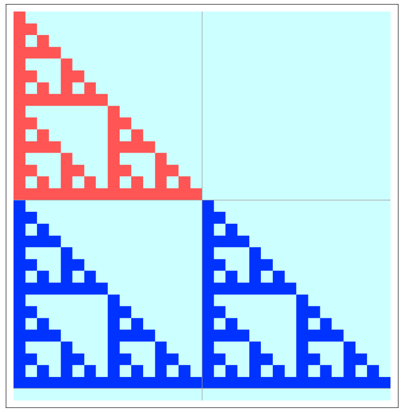
where R denotes the cyclic right-shift operation, and $\oplus$ is the xor operation, applied pointwise. For example, $D(1,1,1,0,1) = (1,1,1,0,1) \oplus (1,1,1,1,0) = (0,0,0,1,1)$.

It remains to determine when we reach $\mathbf{0}$ as opposed to some non-zero binary vector. More precisely, we would like to understand for which dimensions all orbits end in $\mathbf{0}$. This may sound difficult, but it is really a question about the behavior of one-dimensional cellular automata, and these devices are fairly well understood. Well, at least cellular automata based on xor are well understood.

Theorem 2.2 *Every Ducci sequence of dimension* d *ends in fixed point* $\mathbf{0}$ *if, and only if,* d *is a power of 2.*

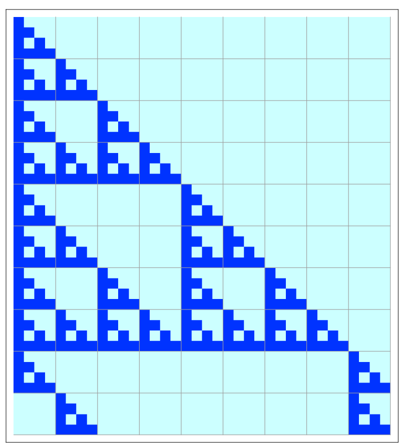
Sketch of proof.

We will outline an argument based on geometry, see the exercises for more details. First, we show that $d = 2^k$ implies that all vectors evolve to $\mathbf{0}$ after d steps. For example, a one-point binary seed for dimension $d = 32$ evolves like this.



The image shows a 33×32 matrix, the red triangle corresponds the time 0 through 15. There are two blue copies of this triangle during times 16 through 31; at time 32 we have reached $\mathbf{0}$. The picture starts with unit vector $\boldsymbol{e}_1$, but our game is invariant under rotations, so the same argument works for any unit vector. Moreover, D is additive on binary vectors in the sense that $D(\boldsymbol{x} \oplus \boldsymbol{y}) = D(\boldsymbol{x}) \oplus D(\boldsymbol{y})$. If follows that all vectors evolve to $\mathbf{0}$ at time d (at the latest).

For the opposite direction consider $d = 2^k \ell$ where $\ell > 1$ odd. Here is the case $d = 4 \times 9$.



The key observation is this: consider the 4×4 blocks indicated in the image and think of a block as 1 if it contains a triangle, 0 otherwise. The evolution of these blocks is the same as for single cells. Hence it suffices that the claim holds for odd d, see the exercises.

$\square$

Exercises

Exercise 2.1 Fill in all the missing pieces in the proof of Ducci's theorem.

Exercise 2.2 In section 1.2 we claimed that $\sigma(x) = \sigma(L(x)) = \sigma(R(x)) = \sigma(\mathsf{rev}(x))$. Give a proof using the identities about L, R and rev.

Exercise 2.3 Let $S_t = \{\, \boldsymbol{x} \mid \sigma(\boldsymbol{x}) = t \,\}$ be the vectors with stopping time t. Is $D(S_{t+1}) = S_t$?

Exercise 2.4 Why are Ducci sequences over the rationals not particularly interesting?

Exercise 2.5 Note that $D(\boldsymbol{x}) = D(\boldsymbol{x} + c)$ where the addition is meant to be componentwise. Can we identify vectors that differ only by an additive offset, the way we identified the results of reflections/rotations/factoring?

Exercise 2.6 Show that the tribonacci vectors really work as advertised.

Exercise 2.7 Explain why attempts similar to the one in section 2.1 to produce divergence with roots and transcendental numbers will fail.

Exercise 2.8 Try to give a proof of Lotan's theorem. A good first step would be to try to understand what q has to do with the tribonacci numbers in Webb's theorem.

Exercise 2.9 Show that D is additive on binary vectors in the sense that $D(\boldsymbol{x} \oplus \boldsymbol{y}) = D(\boldsymbol{x}) \oplus D(\boldsymbol{y})$ where $\oplus$ stands for bit-wise xor.

Exercise 2.10 Show that all orbits end in $\mathbf{0}$ when the dimension d is a power of 2.

Exercise 2.11 Show that there are orbits not ending in $\mathbf{0}$ when the dimension d is odd. Think about getting to $\mathbf{0}$. How does this prove our claim?

Exercise 2.12 Is there any computational shortcut to figure out how many steps a Ducci sequence will take to reach $\mathbf{0}$? I do not know the answer myself.