

FEATURES OF A HIGH SCHOOL OLYMPIAD PROBLEM

LAWRENCE SMOLINSKY

ABSTRACT. This paper is a supplement to a talk for mathematics teachers given at the 2016 LSU Mathematics Contest for High School Students. The paper covers more details and aspects than could be covered in the talk.

We start with an interesting problem from the 2009 Iberoamerican Math Olympiad concerning a particular sequence. We include a solution to the problem, but also relate it to several areas of mathematics. This problem demonstrates the countability of the rational numbers with a direct one-to-one correspondence. The problem also shows the one-to-one correspondence of finite continued fractions and rational numbers. The subsequence of odd indexed terms was constructed by Johannes Kepler and is discussed. We also show that the indices have an extension to the 2-adic integers giving a one-to-one correspondence between the positive real numbers and the 2-adic integers. Everything but the extension to the 2-adic integers is known to have appeared elsewhere.

1. THE PROBLEM

Let us start with problem five from the 2009 Iberoamerican Math Olympiad.

Problem 1.1. *The sequence $\{a_n\}_{n=1}^{\infty}$ satisfies $a_1 = 1$ and for $n \geq 1$,*

$$a_{2n} = a_n + 1; \quad a_{2n+1} = \frac{1}{a_{2n}}.$$

Prove that every positive rational number occurs in the sequence exactly once.

Let's write out the beginning of the sequence:

$$1, 2, \frac{1}{2}, 3, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, 4, \frac{1}{4}, \frac{4}{3}, \frac{3}{4}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{5}, 5, \frac{1}{5}, \frac{5}{4}, \frac{4}{5}, \frac{7}{3}, \frac{3}{7}, \frac{7}{4}, \frac{4}{7}, \frac{7}{2}, \frac{2}{7}, \frac{7}{5}, \frac{5}{7}, \frac{8}{3}, \frac{3}{8}, \dots$$

2. COUNTABLE AND UNCOUNTABLE.

For a finite set we can determine how large it is by counting the number of elements. One can view counting the elements in a finite set as putting them in one-to-one correspondence with a finite counting set. For example, $\{\clubsuit, \diamond, \spadesuit, \heartsuit\} \leftrightarrow \{1, 2, 3, 4\}$, and we can say the two sets are size or cardinality 4. We can also compare infinite sets using one-to-one correspondences. The measure of the size of a set in this sense is called cardinality, and two sets have the same cardinality if they can be put into one-to-one correspondence.

Date: March 1, 2016.

infinite and an infinite set (like $[0, 1]$) that cannot be put into one-to-one correspondence with \mathbb{N} is called uncountable.

What about the rational numbers?

3. ARE THE RATIONAL NUMBERS COUNTABLE OR UNCOUNTABLE?

One property the rational numbers possess is that between any two real numbers there are infinitely many rational numbers! It is called being dense in the real line and the real numbers also possess this property. When faced with the question of whether the rational numbers are countable like \mathbb{N} or uncountable like \mathbb{R} , many people intuit that the rational numbers are uncountable because they are dense like the real numbers. This is false.

The rational numbers are countable. In fact, if a student successfully answers Problem (1.1), then they will have proved that there is a one-to-one correspondence between the positive rational numbers and \mathbb{N} . The correspondence is simply:

$$n \leftrightarrow a_n.$$

It is now easy to construct a one-to-one correspondence between the rational numbers and \mathbb{N} . Let $s_1 = 0$, $s_{2n} = a_n$, $s_{2n+1} = -a_n$ and the correspondence is:

$$n \leftrightarrow s_n.$$

4. PROPERTIES OF THE SEQUENCE

Some properties of the sequence that will be required in subsequent sections are discussed. First recall the recursion relations from Problem (1.1):

$$a_1 = 1, a_{2n} = a_n + 1, \text{ and } a_{2n+1} = \frac{1}{a_{2n}}.$$

From the second equality we note that $a_{2^m k} = a_k + m$ and in particular $a_{2^m} = m + 1$.

Next we observe that the sequence in Problem (1.1) has the following properties:

Properties 4.1.

- (1) $a_1 = 1$
- (2) $a_{2k} = a_k + 1 > 1$
- (3) $a_{2k+1} = \frac{1}{a_{2k}} < 1$ for $k > 0$

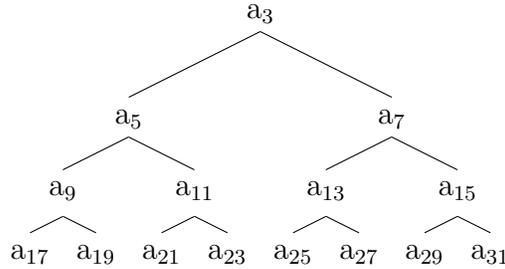
The even indexed sequence elements are the positive rational numbers greater than 1 and the odd indexed rational numbers are the positive rational numbers less than or equal to 1. In particular

$$n \leftrightarrow a_{2n+1}$$

gives a one-to-one correspondence between the rational numbers in $(0, 1)$ and \mathbb{N} , or in other words, the rational numbers in $(0, 1)$ as the subsequence of odd indices starting at 3: a_3, a_5, a_7, \dots .

5. A SEQUENCE OF JOHANNES KEPLER

The subsequence a_3, a_5, a_7, \dots , which enumerates the rational numbers strictly between 0 and 1, was essentially given by Johannes Kepler in *Harmonices Mundi, Book III*. A translation is by Aiton, Duncan and Field (see [1]) and is excerpted at the website [2]. While Kepler did not describe a sequence, he described the elements in an ordered manner by giving a tree structure. His tree structure relates to the subsequence we defined as follows:



Kepler's rules to propagate his tree are

$$\begin{array}{c}
 \frac{x}{y} \\
 / \quad \backslash \\
 \frac{x}{x+y} \quad \frac{y}{x+y}
 \end{array}$$

or in terms of the sequence the propagation is

$$\begin{array}{c}
 a_{2n+1} \\
 / \quad \backslash \\
 a_{2(2n)+1} \quad a_{2(2n+1)+1}
 \end{array}$$

It is a short exercise with the two recursion relations in Problem (1.1) to show that if $a_{2n+1} = \frac{x}{y}$, then $a_{2(2n)+1} = \frac{1}{\frac{1}{a_{2n+1}} + 1} = \frac{x}{x+y}$ and $a_{2(2n+1)+1} = \frac{1}{a_{2n+1} + 1} = \frac{y}{x+y}$.

Another interesting feature of Kepler's tree is that the subsequence obtained by moving down the right-hand edge of the tree are all ratios of successive Fibonacci numbers.

6. A SOLUTION TO PROBLEM (1.1).

The solution presented is very similar to Alexander Remorov's solution [3]. We first show that the sequence includes all positive rational numbers. Use induction on k with the induction hypothesis: if $\frac{p}{q}$ with $p, q \in \mathbb{N}$ has $p + q \leq k$, then $\frac{p}{q}$ occurs in the sequence.

The induction hypothesis holds for $k = 2$ since $a_1 = \frac{1}{1} = 1$. Now suppose $\frac{r}{s}$ is a rational number with $r + s = k + 1$, and the induction hypothesis holds for k . Note that $r \neq s$ by Properties (4.1) 2 and 3. Consider the case $r > s$. By the induction hypothesis, $\frac{r-s}{s} = a_m$ for some m . Then $\frac{r}{s} = 1 + \frac{r-s}{s} = a_{2m}$. Next, consider the case $s > r$. By the first case, there is an m with $a_m = r/s$. Therefore, by the recursion relation in Problem (1.1), $a_{m+1} = s/r$.

We next show that there are no repetitions in the sequence so that the sequence has distinct terms. Suppose the smallest index representing a number that occurs more than once is m

and $a_m = a_n$. Note that $m > 1$ by Properties (4.1) 2 and 3. Also note that by Properties (4.1) 2 and 3, m and n are the same parity. If m and n are even, then $a_{m/2} = a_{n/2}$, by the recursion in Problem (1.1). This contradicts m being minimum. Similarly, if m and n are odd, then $a_{m-1} = a_{n-1}$. Again, it follows by the recursion in Problem (1.1) that contradicts m being minimum.

7. CONTINUED FRACTION REPRESENTATION OF ELEMENTS AND BASE 2 REPRESENTATION OF INDICES

We can read the sequence entry for a_n from the representation of n , the index of the sequence entry. (A similar treatment is given in [5]). First write n in base two as

$$n = 2^{m_0} + 2^{m_1} + 2^{m_2} + \dots + 2^{m_k}$$

with m_0 a nonnegative integer, the other m_i 's positive integers, and $m_i < m_{i+1}$. Next define $n_0 = m_0$ and $n_i = m_i - m_{i-1}$ for $i = 1, \dots, k$. We may then write

$$(4) \quad n = 2^{n_0} + 2^{n_0+n_1} + 2^{n_0+n_1+n_2} + 2^{n_0+n_1+n_2+n_3} + \dots + 2^{n_0+n_1+n_2+n_3+\dots+n_k},$$

which we may express as

$$(5) \quad n = 2^{n_0}(1 + 2^{n_1}(1 + 2^{n_2}(1 + 2^{n_3}(\dots 2^{n_{k-1}}(1 + 2^{n_k})\dots))))$$

where the intergers n_i satisfy $n_0 \geq 0$ and $n_i > 0$ for $i = 1, \dots, k$.

To unravel the value of a_n , we can use the expression for n in Eq (5) and the recursion relations in Problem (1.1). Start from the inside and work out:

$$\begin{aligned} a_{1+2^{n_k}} &= \frac{1}{a_{2^{n_k}}} = \frac{1}{1 + n_k} \\ a_{1+2^{n_{k-1}}(1+2^{n_k})} &= \frac{1}{n_{k-1} + \frac{1}{1+n_k}} \\ &\vdots \\ a_{2^{n_0}(1+2^{n_1}(1+2^{n_2}(1+2^{n_3}(\dots 2^{n_{k-1}}(1+2^{n_k})\dots))))} &= n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_{k-1} + \frac{1}{1+n_k}}}} \end{aligned}$$

This final expression for a_n as a long compound fraction

$$(6) \quad a_n = n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_{k-1} + \frac{1}{1+n_k}}}}$$

is called a continued fraction.

The n_i can be obtained from counting the number of zeros between the base two digits of n as can be seen in Eq (4). Therefore, from the base two representation of the index n ,

one can read off the sequence value a_n by counting the number of zeros between ones or the decimal point. The instructions are as follows:

Instructions 7.1.

- (1) n_0 is the number of zeros between the decimal point and the first one (with perhaps n_0 being 0).
- (2) n_m-1 is the number of zeros between the m -th one and the next one for $0 < m < k$.
- (3) n_k-2 is the number of zeros between the $k-1$ -st one and the last (k -th) one.

Note the last nested denominator of the partial fraction is $1 + n_k$. For example, if $600 = 1001011000.0$, so

$$a_{600} = 3 + \frac{1}{1 + \frac{1}{2 + \frac{1}{4}}}$$

Furthermore for any $n > 600$ that has a base two representation agreeing with the start of 600, i.e., $\dots 1011000$, a_n has the same beginning as a continued fraction as a_{600} (up through $n_{k-1} = 2$),

$$a_{600} = 3 + \frac{1}{1 + \frac{1}{2 + \frac{1}{\ddots}}}$$

If you combine the observation that a_n has the form given in Eq (6) with Problem (1.1), then you see that each positive rational number has a unique representation as a continued fraction of the form given in Eq (6), i.e., with for some whole number k , $n_0 \geq 0$, and $n_i > 0$ for $i = 1, \dots, k$ ¹. It follows as a simple exercise that

Theorem 7.2. *For every rational number q there are unique choices of z an integer, k a whole number, and if $k > 0$, natural numbers n_i for $i = 1, \dots, k$ so that q may be uniquely represented as*

$$z + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{\ddots + \frac{1}{n_{k-1} + \frac{1}{1+n_k}}}}}$$

Finally, note that if you start with a positive rational number q and write down its continued fraction representation, then you have a recipe for finding the place of q in the sequence, i.e., for finding the n with $q = a_n$. The recipe for expressing q as a continued fraction follows from an algorithm that is thought to go back to the Pythagoreans and is called the Euclidean Algorithm as it appeared in Euclid's *Elements*.

¹In some presentations, n_k is allowed to be zero and then there may be two representations

8. EXTENSION TO POSITIVE REAL CONTINUED FRACTIONS AND A SEQUENCE INDEXED BY THE 2-ADIC INTEGERS.

We move beyond Problem (1.1) in this section to an extension of the sequence. We still write of this extension as a sequence even though it is more properly a function.

We have already seen and proved that there is a one-to-one correspondence between the rational numbers and the continued fractions. These continued fractions had finitely many quotients terminating with $n_k + 1$. We will call these finite continued fractions. The notion of continued fractions may be extended to infinite expressions:

$$(7) \quad z + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{n_4 + \frac{1}{\ddots}}}}}$$

for $z \in \mathbb{Z}$ and $n_i \in \mathbb{N}$.

Theorem 8.1. *There is a one-to-one correspondence between the infinite continued fractions, i.e., expression of the form given in (7), and the irrational numbers. The continued fraction (7) is nonnegative if and only if $z \geq 0$.*

An infinite expression of the form (7) must be interpreted using the theory of limits just as infinite decimals must be interpreted using limits. One may terminate the infinite continued fraction (7) at a finite level, say at n_m , to get the rational number c_{m+1} . The c_m 's are called *convergents* of the continued fraction. The infinite continued fraction (7) is viewed as a limit of its sequence of convergents:

$$(8) \quad z, z + \frac{1}{n_1}, z + \frac{1}{n_1 + \frac{1}{n_2}}, z + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3}}}, z + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{n_4}}}} \cdots$$

The limit $\lim_{m \rightarrow \infty} c_m$ always exists and is an irrational number, r . The odd subsequence of convergents $\{c_{2n-1}\}_{n=1}^{\infty}$ monotonically increases to r , and the even subsequence of convergents $\{c_{2n}\}_{n=1}^{\infty}$ monotonically decreases to r . The details are beyond this talk but reader may find them in many sources such as [4].

It is interesting to note that continued fraction representation of a number may be more easily accessed or understood than a decimal representation. For example, while both $\sqrt{2}$ and e have infinite decimal representations that are not completely known, the continued fraction representations are known. Square roots all have periodic continued fraction representations and

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\ddots}}}}$$

or $n_i = 2$ for $i = 1, 2, 3, \dots$. The representation of e was first published in 1744 by Euler:

$$e = 2 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{1 + \frac{1}{4 + \frac{1}{1 + \frac{1}{\ddots}}}}}}}}$$

or $n_0 = 2$ and for $i > 0$, $n_i = \begin{cases} 2k & \text{if } i = 3k - 1 \\ 1 & \text{otherwise} \end{cases}$. For a proof see [6].

We now turn to the second aspect needed to extend the sequence. The index set of the sequence used to obtain the rational numbers are the natural numbers. Representing the index n as a base two numeral allows us to directly produce continued fraction representation of a_n . We extend the index set to a number system called the 2-adic integers to produce a type of 2-adic sequence that yields the positive real numbers without repetition.

Infinite expressions like infinite continued fractions or infinite decimals are given meaning using the notion of limits. For example the infinite decimal $0.333333333 \dots$ is given meaning as the limit of the sequence $0.3, 0.33, 0.333, 0.3333$, etc. The later terms are close together because their difference is small in the usual absolute value or norm. The digits farther to the right of the decimal point are in smaller place values than those to the left in terms of the norm.

There are number systems introduced (in the late 1890's by Kurt Hensel) that extend the counting numbers using norms motivated by number theory. For each prime number p , there is a number system called the p -adics. These p -adic numbers turn out to be very important in mathematics. For example, Wiles's proof of Fermat's last theorem uses p -adic numbers. We consider only the 2-adics. Every integer can be written as a power of two times an odd number, $n = 2^k m$ with m odd. The 2-adic norm is $|n|_2 = |2^k m|_2 = 1/2^k$. If we right counting numbers in base two, then the digits farther to the left are in the smaller place value in terms of the 2-adic norm. The 2-adic integer numerals are base two numerals that extend infinitely to the left! A 2-adic integer is of the form $z = \sum 2^{m_i}$ for m_i an increasing finite or infinite sequence of counting numbers. An example of is $\dots 010101010 = \sum_{m=1}^{\infty} 2^{2m-1} = -2/3$, which is a 2-adic integer. (You can see it is $-2/3$ by dividing by 2, multiplying by 3, and adding 1 to get zero).

We can extend the index set of the sequence in Problem (1.1) infinite 2-adic numerals by using part 1 and 2 of Instructions (7.1). So for example,

$$a_{-2/3} = a_{\dots 010101010} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\ddots}}}}}}}}$$

which happens to be $\sqrt{2}$. In general, following Instructions (7.1) means expressing a 2-adic integer as

$$(9) \quad n = \sum_{m=0}^{\infty} 2^{\sum_{i=0}^m n_i}$$

then the instructions yield

$$(10) \quad a_n = n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{n_4 + \frac{1}{n_5 + \ddots}}}}}$$

This sets up an one-to-one correspondence between the 2-adic integers and the positive real numbers by Theorems (7.2) and (8.1).

We close by noting that although this map is not continuous, there is a consistency result. Every 2-adic integer is a limit of counting numbers specified by the numeral representation since

$$\lim_{m \rightarrow \infty} \sum_{i=0}^m 2^{m_i} = \sum_{i=0}^{\infty} 2^{m_i}.$$

The consistency result is

Theorem 8.2. *Let $z = \sum_{k=0}^{\infty} 2^{\sum_{i=0}^k n_i}$ with $n_i \in \mathbb{Z}$, $n_0 \geq 0$, and $n_i > 0$ for $i > 0$. Then*

$$(11) \quad \lim_{m \rightarrow \infty} a_{\sum_{k=0}^m 2^{\sum_{i=0}^k n_i}} = n_0 + \frac{1}{n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{n_4 + \frac{1}{n_5 + \ddots}}}}} = a_z$$

Proof. To show Eq.(11), we need to refer to the theory of continued fractions. The continued fraction in Eq.(11) has a real limit r . The k -th convergent is obtained by retaining truncating the continued fraction after the k -th denominator, i.e., in Eq.(11) dropping, or setting to zero, n_i for $i > k$. The even convergents form an increasing sequence with limit a_z and the odd convergents form a decreasing sequence with limit a_z .

Now, $a_{\sum_{k=0}^m 2^{\sum_{i=0}^k n_i}}$ is between its $(m - 1)$ -st and $(m - 2)$ -nd convergent, but the $(m - 1)$ -st and $(m - 2)$ -nd convergents of $a_{\sum_{k=0}^m 2^{\sum_{i=0}^k n_i}}$ and a_z are the same. Eq.(11) follows since $a_{\sum_{k=0}^m 2^{\sum_{i=0}^k n_i}}$ is sandwiched between a monotonically increasing and decreasing sequences converging to a_z . □

Example 8.3. Consider right-hand edge of Kepler's tree in Section (5). This edge consists of ratios of successive Fibonacci numbers. In terms of the sequence from Problem (1.1), the edge is

$$a_3, a_7, a_{15}, \dots, a_{2^n - 1}, \dots$$

If we write the indices in base two, it is

$$(12) \quad a_{11}, a_{111}, a_{1111}, \dots, a_{\underbrace{1 \dots 1}_{n \text{ one's}}}, \dots$$

By Theorem (8.2),

$$\lim_{n \rightarrow \infty} a_{2^n - 1} = \lim_{n \rightarrow \infty} a_{1 \dots 1} = a_{\dots 111}$$

and

$$(13) \quad a_{\dots 111} = \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\ddots}}}}}$$

by the instructions. This continued fraction is the reciprocal of the golden ratio or $\frac{\sqrt{5} - 1}{2}$. The sequence of the right-hand edge of the tree, Eq. (12), is the sequence of convergents for the continued fraction (13).

Acknowledgements: The author is grateful to Andrew McDaniel for reviewing this manuscript.

REFERENCES

- [1] Johannes Kepler with E.J. Aiton, A.M. Duncan, and J.V. Field, trans., *The Harmony of the World* (Philadelphia, Pennsylvania: American Philosophical Society, 1997)
- [2] <http://search.cpan.org/~kryde/Math-PlanePath-122/lib/Math/PlanePath/FractionsTree.pm>
- [3] Remorov, A. Alexander Remorov's homepage (2015). *Math Competition: Math Competition Handouts: Olympiad Level: Sequences*. Downloaded from <http://www.mit.edu/~alexrem/Sequences.pdf>
- [4] Olds, C.D. *Continued Fractions*, Random House, New York 1963.
- [5] Czyz, J. and Self, W. The Rationals Are Countable-Euclid's Proof. *The College Mathematics Journal* 34(2003)367-369.
- [6] Cohn, H. A Short Proof of the Simple Continued Fraction Expansion of e. *The American Mathematical Monthly*, 113 (2006)57-62.

DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803, USA
E-mail address: smolinsk@math.lsu.edu