The μ-inverse for the HOL Light reals

Freek Wiedijk

Abstract

We present an alternative definition of the multiplicative inverse for the real numbers as formalized in John Harrison's HOL Light system.

1 Defining the real numbers

The two common ways to define the real numbers from the rational numbers are by means of Cauchy sequences or by means of Dedekind cuts. Both methods for defining the real numbers have many variations. (For instance, lazy streams of digits are a way to code Cauchy sequences, and Conway's way to represent numbers [1] are a variation on the theme of Dedekind cuts.)

2 Nearly-multiplicative functions

In [2] John Harrison defined the non-negative real numbers $\mathbb{R}_{\geq 0}$ as equivalence classes of certain sequences (a_n) called *nearly-multiplicative functions.* These sequences are sequences of *natural numbers*,[1] such that the sequence of rationals

$$\left(\frac{a_n}{n}\right)$$

is a Cauchy sequence with a convergence rate proportional to $1/n$. Specifically it satisfies, for some $a \in \mathbb{R}_{\geq 0}$ and $B \in \mathbb{R}_{>0}$ (which both depend on the sequence):

$$\left|\frac{a_n}{n} - a\right| \leq \frac{B}{n}$$

or, equivalently:

$$|a_n - an| \leq B$$

So (a_n) is a nearly-multiplicative function corresponding to the real number a if a_n is, up to a bounded 'error', equal to an. If one draws the graph of such a sequence, it stays within a straight 'band' of finite width.

There are two different ways to characterize these sequences without mentioning rational or real numbers, so just in terms of natural number arithmetic. Those are the requirements of *near-multiplicativity*:

$$\exists B \in \mathbb{N}. \forall m, n \in \mathbb{N}.\ |na_m - ma_n| \leq B\,(m + n)$$

[1] So the reals are defined without first defining the rational numbers or even the integers.

(which is the Cauchy-sequence property

$$\exists B \in \mathbb{N}.\, \forall m, n \in \mathbb{N}.\ \left|\frac{a_m}{m} - \frac{a_n}{n}\right| \leq B\Big(\frac{1}{m} + \frac{1}{n}\Big)$$

if one divides both sides by mn), and of *near-additivity*:

$$\exists B' \in \mathbb{N}.\, \forall m, n \in \mathbb{N}.\ |a_{m+n} - (a_m + a_n)| \leq B'$$

These two requirements are the same, as is shown in [2].

3 Equivalence and embedding of the naturals

The *equivalence* of two nearly-multiplicative sequences is defined by:

$$(a_n) \sim (b_n) \quad \Leftrightarrow \quad \exists C.\ |a_n - b_n| \leq C$$

Two nearly-multiplicative sequences represent the same non-negative real number in $\mathbb{R}_{\geq 0}$ if and only if they are equivalent.

For each natural number $k \in \mathbb{N}$ there exists a nearly-multiplicative sequence k^* representing it in $\mathbb{R}_{\geq 0}$, which is defined by

$$(k^*)_n = kn$$

In particular:

$$(0^*)_n = 0$$

and:

$$(1^*)_n = n$$

4 Arithmethical operations

The addition on the non-negative real numbers in terms of nearly-multiplicative sequences is defined by:

$$(a + b)_n = a_n + b_n$$

For multiplication there are two choices for a 'natural' definition:

- Either one defines it by:

  $$(ab)_n = a_n b_n \,\mathtt{DIV}\, n$$

  (where `DIV` is division in the natural numbers.)

- Alternatively, one can define it by:

  $$(ab)_n = a_{b_n}$$

  This works, because a_{b_n} is close to $a(b_n)$ which is close to $a(bn)$ which is $(ab)n$. With this definition proving that multiplication is commutative (which is done in [2]) becomes non-trivial.

Analogously there are two ways to define multiplicative inverse for nearly-multiplicative functions.

- Either one defines:
$$(a^{-1})_n = n^2 \;\mathtt{DIV}\; a_n$$

- Alternatively, one can define:
$$(a^{-1})_n = \mu m.\, a_m \geq n$$
(The notation $\mu m.\, P(m)$ means 'the least $m \in \mathbb{N}$ for which $P(m)$ holds.') So $(a^{-1})_n$ is the first position where the graph of a_m 'crosses' the line of constant height n. This definition has the property that:
$$(1^*)^{-1} = 1^*$$
(which is a reason to have in this definition $\geq$ instead of $>$.)

Both these definitions need to specify what happens when the sequence represents the real number 0.

- If the inverse is defined in the first way, the inverse of a sequence $(a_n) \sim 0^*$ won't be a nearly-multiplicative sequence (it will grow at least quadratically.) In [2] the result of the division for that case is *defined* to be the nearly-multiplicative sequence 0^* representing 0 in $\mathbb{R}_{\geq 0}$.

- If the inverse is defined in the second way, we will get the expression $\mu m.\, \bot$ (because for sufficiently large n, no a_m will be greater than n.) If we define:
$$\mu m.\, \bot = 0$$
to make the μ operator total, then this will lead to a total multiplicative inverse, and we will automatically (so not by definition) get:
$$(0^*)^{-1} = 0^*$$

5 The properties of the inverse

In [2] the inverse is defined according to the first approach (although multiplication is defined according to the second!), and the second way to define the multiplicative inverse is not even mentioned. We will now show how to prove the basic properties of the inverse if it is defined according to the second approach. This theory has been formalized using the HOL Light system.[2]

We first present the HOL Light definitions of the μ operator and of the μ-inverse:

[2]This formalization is on the Web at `<http://www.cs.kun.nl/~freek/notes/muinv.ml>`

```
parse_as_binder "mu";;
let mu = new_definition
  `(mu) P = (@n. P n /\ (!m. m < n ==> ~P m) \/
                (n = 0) /\ (!m. ~P m))`;;
let muinv = new_definition
  `muinv(x) = afn(\n. mu m. fn x m >= n)`;;
```

When in the rest of this note we write $^{-1}$ we refer to this `muinv` function. The propositions that follow show that it is well defined and that it behaves like a multiplicative inverse. The propositions are labelled with their name in the HOL Light formalization.

In the following, $a = (a_n)$ is a nearly-multiplicative function.

Proposition 1 (= `MUINV_TO_0`) *If $a \sim 0^*$, then there exists some N such that $(a^{-1})_n = 0$ for all $n \geq N$.*

Proof Because $a \sim 0^*$, there is a B with $a_n \leq B$ for all $n \in \mathbb{N}$. Take $N = B+1$, then for $n \geq N$ clearly $a_m \geq n$ can't happen and so is equivalent to $\perp$ and so $(a^{-1})_n = \mu m.\, a_m \geq n = \mu m.\, \perp = 0$. □

Proposition 2 (= `MUINV_EQ_0`) *If $a \sim 0^*$, then $a^{-1} \sim 0^*$.*

Proof Take N as in the previous proposition. $(a^{-1})_n$ is bounded on $n < N$ (because it only takes finitely many values there) and is zero for $n \geq N$, so it is bounded everywhere, and so $a^{-1} \sim 0^*$. □

This is about all there is to say about the case that $a \sim 0^*$. So in the rest of this note, suppose that $a \not\sim 0^*$. This means that a_n can get arbitrarily big, so the μ in the definition of $^{-1}$ will never be degenerated.

Proposition 3 (= `NADD_CLOSE`) *For some B we have that:*

$$a_0 \leq B$$
$$a_n \leq a_{n-1} + B$$

So this says that the a_n can't jump up too far.

Proof Directly from the property of near additivity. □

Proposition 4 (= `MUINV_CLOSE'`) *For some B:*

$$n \leq a_{(a^{-1})_n} \leq n + B$$

Proof Take B as in the previous proposition. Let $m = (a^{-1})_n$, and suppose that $m > 0$. Then because m is the smallest number with $n \geq a_m$, we have:

$$a_{m-1} < n \leq a_m$$

From this and the inequalities of the previous proposition the required inequalities follow.

The case $m = 0$ is similar. □

Proposition 5 (= `MUINV_LINV`) *The function a^{-1} is the multiplicative inverse of a, i.e.:*

$$a^{-1} \cdot a \sim a \cdot a^{-1} \sim 1^*$$

Proof Directly from the previous proposition (because multiplication is function composition and $\sim$ is bounded difference.) $\square$

Proposition 6 (= `MUINV_UBOUND_ALL`) *The function $(a^{-1})_n$ is bounded by a linear function, so there exists a B such that:*

$$(a^{-1})_n \leq Bn$$

Proof It is enough to show that $(a^{-1})_n$ will be bounded for sufficiently large n. This follows using proposition 3 and the fact that a_n is bounded from below, which means that there is an A such that:

$$n \leq Aa_n$$

(It can be shown that $B = 2A$ is a sufficient estimate.) $\square$

Proposition 7 (= `MUINV_NADD`) *The function $(a^{-1})_n$ is a nearly-multiplicative function.*

Proof From proposition 4 follows that the difference between $|n(a^{-1})_m - m(a^{-1})_n|$ and $|a_{(a^{-1})_n}(a^{-1})_m - a_{(a^{-1})_m}(a^{-1})_n|$ is bounded by a multiple of $|(a^{-1})_m + (a^{-1})_n|$. Furthermore because a is a nearly-multiplicative function, $|a_{(a^{-1})_n}(a^{-1})_m - a_{(a^{-1})_m}(a^{-1})_n|$ is also bounded by a multiple of $|(a^{-1})_m + (a^{-1})_n|$. Together this gives that $|n(a^{-1})_m - m(a^{-1})_n|$ is bounded by a multiple of $|(a^{-1})_m + (a^{-1})_n|$. But then by the previous proposition it is bounded by a multiple of $|m + n|$. $\square$

6 Relevance

We presented an alternative way to define the multiplicative inverse for the real numbers in John Harrison's HOL Light system. Although this alternative definition is elegant and corresponds more closely to the multiplication of the system than the current definition of inverse, its theory is not substantially simpler. Also because once the key properties of inverse have been proved no use is made of its definition, it really doesn't matter which definition one uses.

References

[1] J.H. Conway. *On Numbers and Games.* Academic Press, New York, 1976.

[2] J.R. Harrison. *Theorem Proving with the Real Numbers.* Springer-Verlag, Berlin, Heidelberg, New York, 1998.