Introducing the reals as a totally ordered additive group without using the rationals

de Bruijn, N.G.

Published: 01/01/1975

Citation for published version (APA):
Introducing the reals as a totally ordered additive group without using the rationals

by

N.G. de Bruijn
Introducing the reals as a totally ordered additive group without using the rationals by N.G. de Bruijn.

1. The purpose of this note is to define the reals as infinite decimal fractions (we use the word decimal also for the case where the base is not ten but any positive integer b), and to derive its additive structure and its order without reference to rational numbers. Actually our treatment can be made entirely combinatorial. The arithmetic operations we use in order to define operations on decimal fractions, are restricted to a limited range (especially if b = 2), and can be replaced by combinatorial operations.

Our point of departure is the set \( \mathbb{Z} \) of all integers. It can be characterized as a non-empty totally ordered set without maximal or minimal element, with the property that every non-empty subset which is bounded above (below) has a maximal (minimal) element. We use \( z + 1 \) and \( z - 1 \) in order to indicate successor and predecessor.

The author intends to write another version of this paper where the rôle of the addition algorithm is taken over by the subtraction algorithm. (This is slightly simpler: the study of inequality between decimal fractions is related to the subtraction algorithm). That version will have a more extensive introduction, and will say something on the prospects of multiplication.
2. Notation. \( \mathbb{Z} \) is the set of all integers.

- \( b \) is a fixed integer \( >1 \).

- If \( P \) and \( Q \) are sets then \( Q^P \) is the set of all mappings of \( P \) into \( Q \).

\( \Sigma \) is the set of all \( f \in \{0,1,\ldots,b-1\}^\mathbb{Z} \) with the following property: for every \( z \in \mathbb{Z} \) there exists an \( x \in \mathbb{Z} \) with \( x > z \) and \( f(x) < b-1 \).

- If \( f, g \in \mathbb{Z}^\mathbb{Z} \) then \( f + g \) is defined by \( (f+g)(z) = f(z) + g(z) \) (\( z \in \mathbb{Z} \)). This \( f + g \) is not the same thing as \( \text{sum}(f, g) \) (to be introduced in Section 4).

- If \( f \in \mathbb{Z}^\mathbb{Z} \) then \( (f) \mod b \) is the function \( h \) that satisfies, for all \( z \in \mathbb{Z} \),
  \[
  h(z) \in \mathbb{Z}, \quad h(z) - f(z) \equiv 0 \pmod{b}, \quad 0 \leq h(z) < b.
  \]

- If \( \lambda \) is any rational number, then \( \lfloor \lambda \rfloor \) denotes the largest integer \( \leq \lambda \).

3. Carry. If \( k \in \{0,1,\ldots,2b-2\}^\mathbb{Z} \) we define an element \( p \) of \( \{0,1\}^\mathbb{Z} \) which is to be called the carry of \( k \), denoted as \( p = \text{carr}(k) \). For every \( z \in \mathbb{Z} \) we define \( p(z) \) as follows: \( p(z) = 0 \) if there exists an \( x \in \mathbb{Z} \), with \( x > z \), \( k(x) < b-1 \), and such that \( k(u) \leq b-1 \) for all \( u \) with \( u \in \mathbb{Z} \), \( z < u < x \). In all other cases we put \( p(z) = 1 \).

**Theorem 3.1.** If \( k \in \{0,1,\ldots,2b-2\}^\mathbb{Z} \), \( p = \text{carr}(k) \) then we have, if \( z \in \mathbb{Z} \),

\[
  p(z-1) = \lfloor (k(z) + p(z))/b \rfloor.
\]

**Proof.** From the definition of \( \text{carr} \) we infer: if \( k(z) < b-1 \) then \( p(z-1) = 0 \), if \( k(z) > b-1 \) then \( p(z-1) = 1 \), if \( k(z) = b-1 \) then \( p(z-1) = p(z) \).

4. Sum. If \( f, g \in \{0,1,\ldots,b-1\}^\mathbb{Z} \) we have \( f+g \in \{0,1,\ldots,2b-2\}^\mathbb{Z} \), and we define \( \text{sum}(f, g) \) by

\[
  \text{sum}(f, g) = (f+g + \text{carr}(f+g)) \mod b.
\]

(4.1)
Theorem 4.1. If $f, g \in \{0,1,\ldots,b-1\}^\mathbb{Z}$ then $\text{sum}(f,g) \in \{0,1,\ldots,b-1\}^\mathbb{Z}$ and $\text{sum}(f,g) = \text{sum}(g,f)$.

Theorem 4.2. If $f, g \in \Sigma$ then $\text{sum}(f,g) \in \Sigma$.

Proof. Assume that $f$ and $g$ are in $\Sigma$ but $s$ is not, where $s = \text{sum}(f,g)$. So there is some $w \in \mathbb{Z}$ such that $s(z) = b-1$ for all $z > w$. We put $f+g = k$, $\text{carr}(k) = p$, whence $s = (k+p) \mod b$. If $z > w$ we have either $k(z) + p(z) = b-1$ or $k(z) + p(z) = 2b-1$. Since $k(z) < 2b-1$, $p(z) \leq 1$, we have just three possibilities if $z > w$:

(i) $k(z) = b-1$, $p(z) = 0$, $p(z-1) = 0$;
(ii) $k(z) = b-2$, $p(z) = 1$, $p(z-1) = 0$;
(iii) $k(z) = 2b-2$, $p(z) = 1$, $p(z-1) = 1$.

(the values of $p(z-1)$ follow from Theorem 3.1). We cannot have case (i) all the time, for if $k(z) = b-1$ for all $z > w$ we have $p(w) = 1$. If $z$ is in case (ii) or (iii), then its successor $z+1$ is in case (iii). So there is some $v > w$ such that we are in case (iii) for all $z > v$. But $k(z) = 2b-2$ implies $f(z) = g(z) = b-1$. Since $f \in \Sigma$ we have a contradiction.

5. Associativity of sum.

Theorem 5.1. If $f, g, h \in \Sigma$ then we have

$$\text{sum}(\text{sum}(f,g),h) = \text{sum}(f,\text{sum}(g,h)).$$

Proof. We put $p = \text{carr}(f+g)$, $s = \text{sum}(f,g)$, $q = \text{carr}(s+h)$, $r = p+q$, $t = \text{sum}(s,h)$.

It follows that

$$t = (f+g+h+r) \mod b.$$  (5.1)
With some fixed \( z \in \mathbb{Z} \) we put
\[
\lambda = (f(z) + g(z) + p(z))/b,
\]
and we obtain
\[
s(x)/b = \lambda - \lfloor \lambda \rfloor.
\]
From Theorem 3.1 we get
\[
p(z-1) = \lfloor \lambda \rfloor, \quad q(z-1) = \left\lfloor \lambda - \lfloor \lambda \rfloor + \frac{h(z) + q(z)}{b} \right\rfloor,
\]
whence
\[
r(z-1) = \left\lfloor \left( (f(z) + g(z) + h(z) + r(z))/b \right) \right\rfloor. \tag{5.2}
\]

We next interchange the roles of \( f \) and \( h \), and we write \( p^* = \text{carr}(h+g), \) \( s^* = \text{sum}(h,g), \) \( q^* = \text{carr}(s^*+h), \) \( r^* = p^*q^*, \) \( t^* = \text{sum}(s^*,f) \). As a counterpart to (5.2) we get
\[
r^*(z-1) = \left\lfloor \left( (f(z) + g(z) + h(z) + r^*(z))/b \right) \right\rfloor. \tag{5.3}
\]

Since \( p,q,p^*,q^* \) have values 0 and 1 only, we have \( |r(z) - r^*(z)| \leq 2 \) for all \( z \), whence, by (5.2) and (5.3), \( |r(z-1) - r^*(z-1)| \leq 1 \) for all \( z \).

Therefore \( |r(z) - r^*(z)| \leq 1 \) for all \( z \).

Let us assume that there is a \( z \in \mathbb{Z} \) with \( r(z-1) < r^*(z-1) \). Comparing (5.2) and (5.3), and using \( |r^*(z) - r(z)| \leq 1 \), we derive that \( r^*(z) - r(z) = 1 \), and next that \( f(z) + g(z) + h(z) + r(z) \equiv b-1 \pmod{b} \) (for otherwise (5.2) and (5.3) would give \( r(z-1) = r^*(z-1) \)). Thus \( r(z) < r^*(z) \) and \( t(z) = b-1 \).

Repeating the argument we get \( r(y) < r^*(y) \) and \( t(y) = b-1 \) for all \( y \geq z \).

This contradicts \( t \in \mathbb{E} \) (cf. Theorem 4.2). The conclusion is that \( r(z) < r^*(z) \) can hold for no \( z \in \mathbb{Z} \). Similarly we can show that \( r^*(z) < r(z) \) for no \( z \in \mathbb{Z} \). Hence \( r = r^* \), and from (5.1) (plus its counterpart for \( t^*,r^* \)), we get \( t = t^* \). This proves the theorem.
6. **Zero element.** By $f_0$ we denote the element of $\Sigma$ with $f_0(z) = 0$ for all $z \in \mathbb{Z}$.

**Theorem 6.1.** For all $f \in \Sigma$ we have $\text{sum}(f, f_0) = f$.

**Proof.** Follows from the definition of sum. Note that $\text{carr}(f) = f_0$.

7. **Additive inverse.** If $f \in \Sigma$ we define $\text{min}(f)$ by

$$
\text{min}(f) = (g + \text{carr}(g)) \mod b,
$$

where $g$ is defined by

$$
g(z) = b^{-1} - f(z) \quad (z \in \mathbb{Z}).
$$

**Theorem 7.1.** If $f \in \Sigma$ then $\text{min}(f) \in \Sigma$ and $\text{sum}(f, \text{min}(f)) = f_0$.

**Proof.** We define $g$ by (7.2), and further $h = \text{carr}(g)$, $k = \text{min}(f)$, $r = f + k$, $s = \text{carr}(r)$. We distinguish three cases.

(i) $f = f_0$. Now $g(z) = b^{-1}$ for all $z$, and $h(z) = 1$ for all $z$. Hence $\text{min}(f) = f_0$, and $\text{sum}(f_0, \text{min}(f_0)) = f_0$ by Theorem 6.1.

(ii) There is an $x \in \mathbb{Z}$ with $f(x) \neq 0$ and such that $f(z) = 0$ for all $z > x$. Now $g(x) \neq b^{-1}$, and $g(z) = b^{-1}$ for all $z > x$. Next we observe $h(z) = 1$ ($z \geq x$), $h(z) = 0$ ($z < x$), and therefore $k(z) = b^{-1} - f(z)$ ($z < x$), $k(x) = b - f(x)$, $k(z) = 0$ ($z > x$). This shows $k \in \Sigma$. Next we note that $r(z) = b^{-1}, b$ or $0$ according to $z < x$, $z = x$, $z > x$, and $s(z) = 1, 0, 0$ according to $z < x$, $z = x$, $z > x$. It follows that $\text{sum}(f, k) = (r + s) \mod b = f_0$.

(iii) For all $y \in \mathbb{Z}$ there is an $x > y$ with $f(x) \neq 0$. Obviously $h = f_0$, whence $k = g$. For all $y$ we have an $x > y$ with $k(x) = b^{-1} - f(x) \neq b^{-1}$, whence $k \in \Sigma$. Finally $r(x) = b^{-1}$, $s(x) = 1$ for all $x \in \mathbb{Z}$, whence $\text{sum}(f, k) = (r + s) \mod b = 0$. 

8. The additive group. It follows that at once from Theorems 4.1, 4.2, 5.1, 6.1, 7.1 that \((E, \text{sum})\) is an abelian group with neutral element \(f_0\). In particular, if \(f \in E\), \(g \in E\) then the unique solution \(h\) of \(\text{sum}(g, h) = f\) equals \(\text{sum}(f, \text{min}(g))\). A consequence is that \(\text{min}(\text{min}(g)) = g\) for all \(g \in E\).

9. The reals. An \(f \in E\) is called positive if \(f \neq 0\) and there is an \(x \in \mathbb{Z}\) with \(f(y) = 0\) for all \(y < x\). We say that \(f \in E\) is negative if there is an \(x \in \mathbb{Z}\) such that \(f(y) = b-1\) for all \(y < x\). An \(f \in E\) is called real if \(f\) is positive or negative or zero (i.e. \(f_0\)). These three alternatives exclude each other.

The set of all real elements of \(E\) is denoted by \(\mathbb{R}\).

**Theorem 9.1.** If \(f\) is positive then \(\text{min}(f)\) is negative. If \(f\) is negative then \(\text{min}(f)\) is positive.

**Proof.** (i) Let \(f\) be positive. Then \(x\) exists such that \(f(x) \neq 0\) and such that \(f(y) = 0\) for all \(y < x\). Now with \(g\) of (7.2) we have \(g(y) = b-1\) for all \(y < x\), and \(\text{carr}(g)\) has just zeros to the left of \(x\) (note that \(g(x) < b-1\)). Hence \(f(y) = b-1\) \((y < x)\).

(ii) Let \(f\) be negative. Since \(f \in E\) we do not have \(f(z) = b-1\) for all \(z \in \mathbb{Z}\). Let \(x\) be the least integer with \(f(x) \neq b-1\). Now \(f(y) = b-1\) for all \(y < x\), whence both \(g\) and \(\text{carr}(g)\) have just zeros to the left of \(x\), and \(\text{min}(f)\) appears to be positive.

**Theorem 9.2.** If \(f\) and \(g\) are real then \(\text{sum}(f, g)\) is real. If \(f\) and \(g\) are positive then \(\text{sum}(f, g)\) is positive. If \(f\) and \(g\) are negative then \(\text{sum}(f, g)\) is negative.

**Proof.** Let \(f, g \in \mathbb{R}\). Put \(k = f+g\), \(p = \text{carr}(k)\), \(s = \text{sum}(f, g)\). Let \(x\) be such that \(f(y) = \gamma\) and \(g(y) = \delta\) for all \(y \leq x\), where both \(\gamma\) and \(\delta\) are either 0 or \(b-1\). If \(\gamma = \delta = b-1\) then \(p(y) = 1\) for \(y < x\), whence \(s(y) = b-1\) for \(y < x\), and \(s\) is negative. If \(\gamma = \delta = 0\) then \(p(y) = 0\) for \(y < x\), whence
s(y) = 0 for y < x, and s is positive or zero. If f and g are positive the case that s is zero is excluded: it would imply f = min(g), which contradicts Theorem 9.1. Finally if γ ≠ δ we have k(y) = b⁻¹ (y ≤ x) and Theorem 3.1 shows that p(y⁻¹) = p(y) for y ≤ x. It follows that either p(y) = 0, s(y) = b⁻¹ for all y ≤ x (i.e. s is negative) or p(y) = 1, s(y) = 0 for all y ≤ x (i.e. s is positive or zero).

Theorem 9.3. \((\mathbb{R}, \text{sum})\) is a subgroup of \((\Sigma, \text{sum})\).

Proof. Follows from Theorems 9.1 and 9.2.

10. Order. If f, g ∈ \(\mathbb{R}\) then we say that \(f > g\) if \(\text{sum}(f, \text{min}(g))\) is positive.

Theorem 10.1. If \(f, g, h ∈ \mathbb{R}\), we have

(i) Exactly one of the relations \(f > g\), \(f = g\), \(g > f\) holds.
(ii) If \(f > g\), \(g > h\) then \(f > h\).
(iii) If \(f > g\) then \(\text{sum}(f, h) > \text{sum}(g, h)\).
(iv) \(f\) is positive if and only if \(f > f₀\), and negative if and only if \(f₀ > f\).

Proof. (i) From the group properties we derive that if \(k = \text{sum}(f, \text{min}(g))\), \(k = \text{sum}(g, \text{min}(f))\), then \(k = \text{min}(k)\). Since \(k\) is either positive or zero or negative, application of Theorem 9.1 is sufficient.

(ii) This follows from the group properties if we use (Theorem 9.2) that two positive elements have a positive sum.

(iii) From the group properties we derive

\[ \text{sum}(\text{sum}(f, h), \text{min}(\text{sum}(g, h))) = \text{sum}(f, \text{min}(g)). \]

(iv) Note that \(\text{sum}(f, \text{min}(f₀)) = f\), \(\text{sum}(f₀, \text{min}(f)) = \text{min}(f)\), and apply Theorem 9.1.
Theorem 10.2. (i) If \( f \) is positive and \( g \) negative then \( f > g \). (ii) If \( f \) and \( g \) have the same sign (i.e. both positive or both negative), we have \( f > g \) if and only if there is an \( x \in \mathbb{Z} \) with \( f(z) > g(x) \) and such that \( f(y) = g(y) \) for all \( y < x \).

Proof. (i) Let \( f \) be positive and \( g \) negative. Then the \( h \) with \( f = \text{sum}(g, h) \) cannot be negative or zero since Theorem 9.1 would tell us that \( f \) is negative. Hence \( k \) is positive, whence \( f > g \).

(ii) Let \( f \) and \( g \) have the same sign. Let \( h \) be the element with \( f = \text{sum}(g, h) \). We put \( g + h = k \), \( \text{carry}(k) = p \), whence \((k+p) \mod b = f \).

Let \( x \) be the last integer such that \( f(x) \neq g(x) \). We first assume \( f(x) > g(x) \). Since \( g(x) + h(x) + p(x) \equiv f(x) \pmod{b} \), \( 0 \leq h(x) < b \), \( 0 \leq p(x) \leq 1 \) we infer \( 0 \leq h(x) + p(x) \leq b \), \( g(x) + h(x) + p(x) = f(x) \), whence, by Theorem 3.1, \( p(x-1) = 0 \). Since \( f(x-1) = g(x-1) \), \( 0 \leq h(x-1) + p(x-1) < b \), the same argument leads to \( p(x-2) = 0 \), and by induction we get \( p(y) = 0 \) for all \( y < x \). Since \( g(y) = f(y) \) for all \( y < x \) it follows that \( h(y) = 0 \) for all \( y < x \). So \( h \) is positive, whence \( f > g \).

In the case where \( g(x) > f(x) \) we interchange the roles of \( f \) and \( g \) and we get \( g > f \).