On coalgebra of real numbers

D. Pavlović

Kestrel Institute
Palo Alto, USA

V. Pratt

Dept. of Computer Science
Stanford University
Stanford, USA

Abstract

We define the continuum up to order isomorphism (and hence homeomorphism) as the final coalgebra of the functor $X \cdot \omega$, ordinal product with ω. This makes an attractive analogy with the definition of the ordinal ω itself as the initial algebra of the functor $1; X$, prepend unity, with both definitions made in the category of posets. The variants $1; (X \cdot \omega)$, $X^o \cdot \omega$, and $1; (X^o \cdot \omega)$ yield respectively Cantor space (surplus rationals), Baire space (no rationals), and again the continuum as their final coalgebras.

1 Introduction

Coinduction has only relatively recently been recognized as a genuine logical principle [2]. Before that, it was introduced and used mostly in the semantics of concurrency [13]. It has by now been presented from many different angles: [1,8,12,16–18], to name just a few contributors.

Why would so foundational a principle wait for the late 20th century to be discovered? In [14,16] the idea was put forward that coinduction is new only by name, while it had actually been around for a long time, concealed within the infinitistic methods of *mathematical analysis*. Roughly,

$$\frac{\text{induction}}{\text{arithmetic}} \approx \frac{\text{coinduction}}{\text{analysis}}$$

The infinitary constructions in elementary calculus are coinductive, just like the infinitary constructions in elementary arithmetic are inductive. However, all evidence presented in [16] was built upon a datatype of real numbers, assumed as given. In the present note, we describe several ways to derive this datatype from scratch, as a final coalgebra.

This paper will be published in
Electronic Notes in Theoretical Computer Science, Volume 19
URL: `www.elsevier.nl/locate/entcs`

In the simplest and the most abstract form, the idea is to view an infinite object x as a stream, i.e. an infinite list $[x_0, x_1, x_2 \ldots]$ of elements from some set Σ. Construed as a rudimentary process, this stream reduces to an *action*, and a *resumption*, i.e. the "head"

$$h(x) = x_0$$

and the "tail", another stream

$$t(x) = x' \quad = \quad [x_1, x_2, x_3 \ldots]$$

But the pair

$$< h, t > : \Sigma^\omega \to \Sigma \times \Sigma^\omega$$

is, of course, the final coalgebra structure for the functor $\Sigma \times (-) : \mathsf{Set} \to \mathsf{Set}$.

Although this simple picture, and the related coalgebraic ideas, are nowadays probably not far from becoming a standard part of the basic toolkit for designing datatypes, and perhaps even systems in general [8,18], it may nevertheless come as a surprise that the idea to implement real numbers along these lines can be dated as far back as 1971 — and to the writings of the first "hackers", of all places! In the famous HAKMEM report, R.W. Gosper and R. Schroeppel took up analyzing (among some 200 other computational themes) real arithmetic in terms of continued fractions. The main result appears to be Gosper's derivation [6, 101B] of a general algorithmic scheme for implementing arithmetic operations and methods of successive approximation.

He begins as follows:

> Let x be a continued fraction
>
> $$p_0 + \cfrac{q_0}{p_1 + \cfrac{q_1}{p_2 + \cdots}} = p_0 + \frac{q_0}{x'}$$
>
> where x' is again a continued fraction and the p's and q's are integers.[...]
>
> Instead of a list of p's and q's, let x be a computer *subroutine* which produces its next p and q each time it is called. Thus on its first usage, x will "output" p_0 and q_0 and, in effect, change itself into x'.

Real numbers are thus presented as streams of pairs of integers. The mentioned coalgebraic structure, although never spelled out, is then employed in the subsequent constructions, as well as in the discussion touching upon some of the still very active themes, such as guarded induction [14,15], or the role of redundancy in representation.

It may seem ironic that what we now consider to be a very general computational method had an early brief appearance, even on the background, among the primordial "hacks". But this is probably just an instance of the irony of language.

In any case, although HAKMEM was never published, it has remained available throughout the intervening years, it was widely read, and sometimes even cited. But while the algorithms derived in it have been acknowledged as the source of inspiration for some of the most interesting modern approaches to exact real arithmetic [5,10,20], the underlying coalgebraic idea seems to have gone unnoticed.

While hoping to point to this conceptual link, we must add that the constructions on the following pages should not be taken as a rational reconstruction of the Hakmem view of reals. In fact, they were obtained while we were trying to work out an effective underpinning for our wider calculus-by-coinduction effort, initiated in [14,16]. The coalgebras presented here are just one detail in that plan. However, the realization that a concrete coalgebra of reals was never worked out, although its effects were in use for quite a while, made us try to present it here.

2 Lists and streams

Definition 2.1 *Let* Σ *be a set. A* Σ-stream algebra *is a set* A *together with an isomorphism*

$$A \underset{c}{\overset{<h,t>}{\rightleftarrows}} \Sigma \times A \quad (\cong)$$

In other words,

$$\begin{aligned} h(a :: x) &= a \\ t(a :: x) &= x \\ h(x) :: t(x) &= x \end{aligned}$$

where $a :: \beta$ *is the infix notation for* $c(a, x)$.

Definition 2.2 *A* Σ-list algebra *is a set* B *with a distinguished element* $[]$*, and an isomorphism*

$$B^\bullet \underset{c}{\overset{<h,t>}{\rightleftarrows}} \Sigma \times B \quad (\cong)$$

where $B = \{[]\} + B^\bullet$.

A list algebra structure on B thus corresponds to an isomorphism $B \cong 1 + \Sigma \times B$

Examples. The basic example of a list algebra is, of course, the set Σ^* of lists from Σ: the operations are clearly $h =$ head, $t =$ tail and $c =$ cons. The distinguished element $[]$ is the empty list.

The basic example of a stream algebra is the set Σ^ω of streams, or infinite lists from Σ. The set Σ^∞ of lists and streams together, i.e. of finite and infinite lists, is again a list algebra.

The set $\mathbb{A}$ of analytic functions (say) at 0 also forms a stream algebra, with the structure

$$\begin{aligned} h(f) &= f(0) \\ t(f) &= f' \\ a :: f(x) &= a + \int_0^x f(t)\,dt \end{aligned}$$

The Taylor expansion induces an isomorphism of $\mathbb{A}$ with the subalgebra of $\mathbb{R}^\omega$, consisting of the streams of Taylor coefficients. A method for implementing parts of basic differential calculus in terms of stream algebra has been outlined in [16].

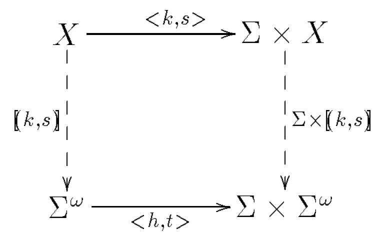
Coalgebras. List and stream algebras are usually derived from initial algebras and final coalgebras.

By definition, given a functor $F : \mathcal{C} \to \mathcal{C}$, an F-coalgebra is simply an arrow $A \stackrel{a}{\to} FA$. A coalgebra homomorphism from $A \stackrel{a}{\to} FA$ to $B \stackrel{b}{\to} FB$ is an arrow $g : A \to B$ such that $Fg \circ a = b \circ g$.

As mentioned in the introduction,

$$\Sigma^\omega \xrightarrow{<h,t>} \Sigma \times \Sigma^\omega$$

is the final coalgebra for the functor $\Sigma \times (-) : \mathsf{Set} \to \mathsf{Set}$. Its finality means that every coalgebra $X \stackrel{<k,s>}{\longrightarrow} \Sigma \times X$ induces a unique coalgebra homomorphism $[\![k,s]\!]$.



The image of $x \in X$ is the stream $[\![k,s]\!](x) = [x_0, x_1, \ldots]$, where $x_i = ks^i(x)$.

Dually, an algebra for $F : \mathcal{C} \to \mathcal{C}$ is an arrow $FA \stackrel{a}{\to} A$. The algebra homomorphisms and the notion of initiality can be obtained by reversing the arrows of the corresponding coalgebra statements. An initial algebra for the functor $1 + \Sigma \times (-) : \mathsf{Set} \to \mathsf{Set}$ is

$$1 + \Sigma \times \Sigma^* \xrightarrow{[[],\mathrm{cons}]} \Sigma^*,$$

induced by the list algebra structure. The final coalgebra for the same functor is the set Σ^∞ of finite and infinite lists, with the structure map derived from list algebra again.

By the Lambek lemma [11], the structure map of every initial algebra, and every final coalgebra, must be an isomorphism. Therefore, the initial algebras and the final coalgebras for the functor $1 + \Sigma \times (-) : \mathsf{Set} \to \mathsf{Set}$ always satisfy the list algebra equations. The initial algebra for the functor $\Sigma \times (-) : \mathsf{Set} \to \mathsf{Set}$ is empty, but the final coalgebras yield stream algebras. [1]

[1] One might ask why we are calling them algebras, rather than coalgebras then. Someone else might reply that they can be made into an algebraic theory; on the other hand, it needs to be mentioned that algebras for an algebraic theory are not the same thing as algebras

Reals as streams. The goal of the present paper is to derive real numbers as list and stream algebras. There are infinitely many irredundant presentations of reals as lists or streams of positive integers, some of them convenient for one purpose, some for another. We spell out three crucial examples, and explain their relations. This should suffice for extracting other examples, although no surprises are to be expected there.

All the obtained representations of reals turn out to be based on final coalgebras for the functors $\mathbb{N} \times (-) : \mathsf{Set} \to \mathsf{Set}$ and $1 + \mathbb{N} \times (-) : \mathsf{Set} \to \mathsf{Set}$ respectively, where $\mathbb{N}$ is the set of natural numbers. The finality of the coalgebras of reals accounts for their *coinductive* nature, just like the initiality of the algebra $1 + \mathbb{N} \xrightarrow{[z,s]} \mathbb{N}$ of natural numbers is well known to account for the induction on them. The presented constructions can thus be viewed as a further piece of evidence for the idea that the coinduction plays in mathematical analysis a role similar to that of the induction in arithmetic.

3 Coalgebras on $[0, 1)$

While analytic functions decompose by the Taylor expansion into streams of real numbers, real numbers will be analysed as streams of natural numbers. [2]

It is of course a matter of taste "which" set of natural number to use in the following presentation. The more complicated among our examples appear a bit simpler with $\mathbb{N} = \{1, 2, 3, \dots\}$.

3.1 Monotone dyadics

The simplest stream algebra, say, on the interval $[0, 1)$, can be derived directly from the usual binary notation. It is well-known that disallowing the infinite tails of 1 makes this notation irredundant: each element $a \in [0, 1)$ can be written in exactly one way, e.g. $a = 0.0111010011111000\dots$. We assume that these streams are always infinite, padded with zeros if necessary.

By partitioning the stream by the occurrences of 0s, and recording the length of the substrings $1 \cdots 110$, each legal binary stream induces a unique $\mathbb{N}$-stream, and vice versa. For example $011101001111000\dots$ corresponds to $[1, 4, 2, 1, 5, 1, 1, \dots]$, by

$$\underbrace{0}_{1}\underbrace{1110}_{4}\underbrace{10}_{2}\underbrace{0}_{1}\underbrace{11110}_{5}\underbrace{0}_{1}\underbrace{0}_{1}\cdots$$

Of course, one could start $\mathbb{N}$ from 0 as well, and not count the 0 into the length of $1 \cdots 110$, as to get $[0, 3, 1, 0, 4, 0, 0, \dots]$ instead, the choice made for section 5. Here it is natural to read the 0's as commas and the blocks of 1's

for a functor. Trying to avoid overloading algebra seems hopeless.

[2] As an initial algebra for the functor $1 + (-) : \mathsf{Set} \to \mathsf{Set}$, natural numbers can, of course, be viewed as finite lists of a single symbol. This is their *unary* notation.

as tally notation. The above choice, no zeros, is made with an eye on later examples in sections 3 and 4.

This idea, which can be found in Hausdorff [7, §10], yields stream algebra on $[0,1)$

$$\begin{aligned} h(x) &= \mu n.\ 1 - \frac{1}{2^n} > x \qquad (= 1 - \lceil \log(1-x) \rceil) \\ t(x) &= 2 - 2^{h(x)}(1-x) \\ n :: x &= \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} + \frac{x}{2^n} \qquad \left(= 1 - \frac{2-x}{2^n}\right) \end{aligned} \tag{1}$$

where $\mu n.\ \Phi(n)$ denotes the least n satisfying Φ. The head $h_m(x)$ is thus the length of the string $1 \cdots 110$ leading the binary presentation of x. The tail $t(x)$ is the real number corresponding to the binary stream obtained when this string is deleted.

The induced coalgebra $[0,1) \xrightarrow{<h,t>} \mathbb{N} \times [0,1)$ is final, i.e. isomorphic with the familiar coalgebra $\mathbb{N}^\omega$ of streams. The isomorphism $[0,1) \cong \mathbb{N}^\omega$ can be understood as follows.

In order to get a stream $[a_0, a_1, a_2, \ldots]$ corresponding to a number $a \in [0,1)$,

- keep adding $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} \ldots$ until adding for the first time $\frac{1}{2^n}$ overshoots a; set $a_0 = n$;
- to the sum $\frac{1}{2} + \cdots + \frac{1}{2^{n-1}}$ (which is thus $\leq a$) further add $\frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \frac{1}{2^{n+3}} + \ldots$ until adding $\frac{1}{2^m}$ overshoots a; set $a_1 = m - n$; and so on.

In terms of the described stream algebra structure, a_i, of course, is just $ht^i(a)$.

The other way around, the isomorphism will assign to a stream $[a_0, a_1, \ldots] \in \mathbb{N}^\omega$ the number

$$a = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{\overline{a}_{i-1}} \sum_{k=1}^{a_i - 1} \left(\frac{1}{2}\right)^k \tag{2}$$

where $\overline{a}_{-1} = 0$ and $\overline{a}_i = \sum_{k=0}^{i} a_k$. For instance, $[1,4,2,1,\ldots]$ will thus go to

$$\begin{aligned} &\overbrace{\frac{1}{2}}^{1}\Big(\overbrace{\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16}}^{4}\Big(\overbrace{\frac{1}{2} + \frac{1}{4}}^{2}\big(\ldots\big)\Big)\Big) \\ &= \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^7}\big(\ldots\big) \end{aligned}$$

i.e. to the binary number $0.0111010\ldots$.

3.2 Alternating dyadics

A slightly different procedure of approximating $a \in [0,1)$ is to

- keep adding $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} \ldots$ and stop only *after* the sum reaches or overshoots a; but stop *immediately* after this, so that the difference between the sum and a remains less than the last summand added, say $\frac{1}{2^n}$. Set $a_0 = n$

- Since the tail $\frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \frac{1}{2^{n+3}} \ldots$ adds up to $\frac{1}{2^n}$, subtracting its elements one by one from the previous sum, must eventually, say after subtracting m elements of this tail, lead below a. Set $a_1 = m$.
- Then start adding again the further elements of the tail, until we get above a, and so on.

This is the alternating binary approximation. The structure induced on $[0, 1)$ is

$$\begin{aligned} h(x) &= \mu n.\ 1 - \frac{1}{2^n} \geq x \qquad (= -\lfloor \log(1-x) \rfloor) \\ t(x) &= 2^{h(x)}(1-x) - 1 \\ n :: x &= \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} + \frac{1-x}{2^n} \qquad \left(= 1 - \frac{1+x}{2^n}\right) \end{aligned} \tag{3}$$

Note however that $h(0) = 0$ falls out of $\mathbb{N}$. But without 0 on the left hand side, the above structure yields an isomorphism $(0, 1) \overset{\cong}{\to} \mathbb{N} \times [0, 1)$. In fact, here we have a *list* algebra $[0, 1) \xrightarrow{<h,t>} 1 + \mathbb{N} \times [0, 1)$, where h and t can be viewed as undefined on 0, or assumed to send it to $1 = \{< 0, 0 >\}$.

In any case, $[0, 1)$ still appears a final coalgebra, albeit for a different functor. It is isomorphic with the "canonical" final coalgebra

$$\mathbb{N}^\infty \xrightarrow{\text{<head,tail>}} 1 + \mathbb{N} \times \mathbb{N}^\infty$$

consisting of the finite and infinite lists. The element of 1 plays the role of the head and the tail of the empty list.

The isomorphism $[0, 1) \cong \mathbb{N}^\infty$ takes $a \in [0, 1)$ to the list $[a_0, a_1, \ldots]$, where $a_i = ht^i(a)$. This list terminates after n entries if $t^n(a) = 0$, so that $ht^n(a)$ falls out of $\mathbb{N}$.

The other way around, given a list $[a_0, a_1, \ldots]$, the corresponding number will be

$$a = \sum_{i=0}^{\infty} \frac{(-1)^i}{2^{\overline{a}_{i-1}}} \sum_{k=1}^{a_i} \left(\frac{1}{2}\right)^k \tag{4}$$

again with $\overline{a}_{-1} = 0$ and $\overline{a}_i = \sum_{i=0}^{i} x_i$. This time, the stream $[1, 4, 2, \ldots]$ will correspond to

$$\begin{aligned} &\frac{1}{2} - \frac{1}{2}\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16}\right) + \frac{1}{32}\left(\frac{1}{2} + \frac{1}{4}\right) \cdots \\ &= \underbrace{\frac{1}{2}}_{1} - \underbrace{\frac{1}{2^2} - \frac{1}{2^3} - \frac{1}{2^4} - \frac{1}{2^5}}_{4} + \underbrace{\frac{1}{2^6} + \frac{1}{2^7}}_{2} \cdots \end{aligned}$$

The real interval $[0, 1)$ is thus identified not only with the streams of positive integers, but also, in a different way, with their lists.

3.3 Continued fractions

A regular continued fraction[3] expansion of $a \in [0,1)$ is in the form

$$a = \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}}}$$

with

$$a_0 = \left\lfloor \frac{1}{a} \right\rfloor$$
$$a_{i+1} = \left\lfloor \frac{1}{\frac{1}{a_{i-1}} - a_i} \right\rfloor$$

Here $\lfloor x \rfloor$ denotes the greatest integer below x, and $a_{-1} = a$. This yields another bijection from the real interval $[0,1)$ to the lists of positive integers $\mathbb{N}^\infty$, finite or infinite, with the empty list corresponding to 0.

Such bijections correspond to the list algebra structures on $[0,1)$. Writing $\lfloor x \rfloor$ in terms of μn, the above expansions yield

$$\begin{aligned} h(x) &= \mu n.\ \frac{1}{n+1} < x \\ t(x) &= \frac{1}{x} - h(x) \\ n :: x &= \frac{1}{n+x} \end{aligned} \tag{5}$$

The list $[a_0, a_1, a_2, \ldots]$, corresponding to a given $a \in [0,1)$ can, of course, be computed as $[h(a), ht(a), ht^2(a), \ldots]$, but the following procedure, essentially from [6], shows better what is going on:

- test $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ until $\frac{1}{n+1}$ falls below a, then set $a_0 = n$; $\frac{p_0}{q_0} = \frac{1}{a_0}$ is the first *convergent* of a;
- then go up: $\frac{1}{a_0+1}, \frac{2}{2a_0+1}, \frac{3}{3a_0+1}$, until $\frac{(m+1)p_0}{(m+1)q_0+1}$ comes above a; set $a_1 = m$; the second convergent of a is $\frac{p_1}{q_1} = \frac{a_1}{a_1 a_0 + 1}$ (which is always a reduced fraction);
- down again: $\frac{p_0+p_1}{q_0+q_1}, \frac{p_0+2p_1}{q_0+2q_1}, \ldots$, until $\frac{p_0+(\ell+1)p_1}{q_0+(\ell+1)q_1} < a$; set $a_2 = \ell$, and $\frac{p_2}{q_2} = \frac{p_0+a_2 p_1}{q_0+a_2 q_1}$;
- then up again, always adding $\frac{p_{i-1}+p_i}{q_{i-1}+q_i}, \frac{p_{i-1}+2p_i}{q_{i-1}+2q_i}, \ldots$, until the next convergent $\frac{p_{i+1}}{q_{i+1}} = \frac{p_{i-1}+a_{i+1}p_i}{q_{i-1}+a_{i+1}q_i}$ is reached ...

The convergents thus alternate, with $\frac{p_{2i}}{q_{2i}} \geq a$ and $\frac{p_{2i+1}}{q_{2i+1}} \leq a$. Approximating a can be understood as finding the convergent $\frac{p_{2i+2}}{q_{2i+2}}$ yet closer to a, in the form $\frac{p_{2i}+a_{2i+2}p_{2i+1}}{q_{2i}+a_{2i+2}q_{2i+1}}$, thus belonging to the interval $\left(\frac{p_{2i+1}}{q_{2i+1}}, \frac{p_{2i}}{q_{2i}}\right]$. The number a is now

[3] General continued fractions allow more general numerators.

in $\left(\frac{p_{2i+1}}{q_{2i+1}}, \frac{p_{2i+2}}{q_{2i+2}}\right]$, and the next convergent $\frac{p_{2i+3}}{q_{2i+3}}$ is sought there.

Reversing this process, the number $a \in [0,1)$ corresponding to a given list $[a_0, a_1, \ldots]$ along the isomorphism $[0,1) \cong \mathbb{N}^\infty$ is obtained as

$$a = \lim_{i \to \infty} \frac{p_i}{q_i}$$

where

$$\begin{aligned} p_{-1} &= 0 & q_{-1} &= 1 \\ p_0 &= 1 & q_0 &= a_0 \\ p_{i+1} &= p_{i-1} + a_{i+1} p_i & q_{i+1} &= q_{i-1} + a_{i+1} q_i \end{aligned}$$

3.4 Comparisons

Each of the described structures is based on a suitable decomposition of the interval $[0,1)$ on countably many subintervals homeomorphic to it, which are then mapped by $< h, t >$ to its countably many copies in $\mathbb{N} \times [0,1)$. In the alternating dyadic coalgebra, this mapping is arranged as follows:

$$\begin{array}{ccccccccccc} [0,1) & = & \{0\} & + & (0,1/2] & + & (1/2,3/4] & + & (3/4,7/8] & + \ldots \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\ 1+\mathbb{N}\times[0,1) & = & \{<0,0>\} & + & \{1\}\times[0,1) & + & \{2\}\times[0,1) & + & \{3\}\times[0,1) & + \ldots \end{array} \tag{6}$$

The monotone dyadic coalgebra differs only by the fact that the subintervals are in the form $[0,1/2), [1/2,3/4)$ and so on, so that 1 is not needed.

But the continued fraction coalgebra looks different:

$$\begin{array}{ccccccccc} [0,1) & = & \{0\} & + \cdots + & \left[\frac{1}{n+1}, \frac{1}{n}\right) & + \cdots + & [1/3,1/2) & + & [1/2,1) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 1+\mathbb{N}\times[0,1) & = & \{\infty\} & + \cdots + & \{n\}\times[0,1) & + \cdots + & \{2\}\times[0,1) & + & \{1\}\times[0,1) \end{array} \tag{7}$$

The transformation on $[0,1)$, induced as the isomorphism of these two final coalgebra structures is continuous, but rather cumbersome to describe. It maps $(0,1/2]$ to $[1/2,1)$, $(1/2,3/4]$ to $[1/3,1/2)$ and so on.

It is not hard to see that any decomposition of $[0,1)$ into countably many subintervals homeomorphic to it will induce a list or stream algebra, isomorphic to either $\mathbb{N}^\infty$ or $\mathbb{N}^\omega$. However, while "invisible" for coalgebra homomorphisms, the choice of the decomposition determines what is easy and what hard to do with each such representation.

Exploiting this difference is one of main tricks of *coinductive programming* [14,15]. E.g., while isomorphic as coalgebras, analytic and coanalytic functions [16] bear significant computational differences: the Laplace transform maps differential equations over analytic functions into algebraic equations over coanalytic functions; the inverse Laplace transform maps back the solutions of the latter into the solutions of the former. *So isomorphisms do make a difference!*

For instance, the alternating dyadics yield the best dyadic approximation of $a \in [0,1)$. If an n-th partial sum σ_n of (4) has the reduced form $\frac{p}{2^q}$, then

$$|a - \sigma_n| \le |a - \frac{r}{2^q}|$$

holds for all integers r. Indeed, by the construction of σ_n, q is $\overline{a}_n$ and $|a - \sigma_n| < \frac{1}{2^q}$.

On the other hand, it is well-known, and fairly clear from the described construction, that continuous fractions yield the best *rational* approximation, i.e.

$$|a - \frac{p_n}{q_n}| \le |a - \frac{r}{q_n}|$$

holds for any convergent $\frac{p_n}{q_n}$ of a, and any integer r.

4 Coalgebras on $[0, \infty)$

Transferred along the homeomorphism

$$\begin{aligned} [0,1) &\longleftrightarrow [0,\infty) \\ x &\longmapsto \frac{x}{1-x} \\ \frac{y}{1+y} &\longleftarrow\!\shortmid\; y \end{aligned}$$

the described coalgebras on $[0,1)$ can be given explicitly as follows.

We have first the monotone dyadics,

$$\begin{aligned} h(y) &= \mu n.\ 2^n - 1 > y \qquad (= 1 + \lfloor \log(1+y) \rfloor) \\ t(y) &= \frac{1+y}{2^{h(y)} - (1+y)} - 1 \\ n :: y &= 2^n \frac{1+y}{2+y} - 1 \end{aligned} \tag{8}$$

the alternating dyadics,

$$\begin{aligned} h(y) &= \mu n.\ 2^n - 1 \ge y \qquad (= \lceil \log(1+y) \rceil) \\ t(y) &= \frac{1}{2}\left(\frac{2^{h(y)-1}}{1 + y - 2^{h(y)-1}} - 1\right) \\ n :: y &= 2^n \frac{1+y}{1+2y} - 1 \end{aligned} \tag{9}$$

and the continued fractions,

$$
\begin{aligned}
h(y) &= \mu n.\ \frac{1}{n} < y \qquad \left(= 1 + \left\lfloor \frac{1}{y} \right\rfloor\right) \\
t(y) &= \frac{y}{y h(y) - 1} - 1 \\
n :: y &= \frac{1+y}{n(1+y)-1}.
\end{aligned} \tag{10}
$$

5 Continuum as Coalgebra, Formalized

We now formalize the above intuitions categorically. We start by fixing the ambient category to be Pos, posets. We know of no analogous technique that works in Set short of externally imposing the desired order and/or topology on the coinductively defined sets. In Pos we can obtain the desired structure without such *deus ex machina* intervention.

The significance of our definition of the reals is that it exposes the following appealing parallel between the continuum and the natural numbers. Whereas the ordinal ω has a natural presentation as the *initial algebra* of the functor $1; X$ (converse ordinal sum with 1, i.e. prepend unity to the given poset), the ordered continuum has just as natural a presentation as the *final coalgebra* of $X \cdot \omega$ (ordinal product with ω, i.e. arrange ω copies of the given poset in order).

To pass from $\mathbb{R}$ *qua* poset to $\mathbb{R}$ *qua* topological space, observe that the order type of any chain uniquely determines its topology when the latter is taken to be the order interval topology,[4] as is the usual case for the continuum. Hence our definition characterizes the continuum not only up to order isomorphism but as a corollary up to homeomorphism.

5.1 Ordinal Sum and Product

In the category Pos of posets, concatenation or *ordinal sum* $X; Y$ is definable as cardinal sum (coproduct or juxtaposition) $X + Y$ with its order augmented with $x \le y$ for all $x \in X, y \in Y$. Similarly, lexicographic or *ordinal product* $X \cdot Y$ is definable as ordinary or cardinal product $X \times Y$ with its order augmented with $(x, y) \le (x', y')$ for all x, x' in X and all $y < y'$ in Y (so Y supplies the "high order digit"). Equivalently, $X \cdot Y$ is the result of substituting one copy of X for each element of Y, with the resulting order being that of X within each copy, and that of Y between copies.

Both operations are associative, not commutative, and preserve the posetal extremes of each of linearity (chainhood) and discreteness (sethood), as pointed out by Birkhoff [3,4], the inventor of the posetal unification of cardinal and ordinal arithmetic.

[4] The *order interval topology* on a chain (more generally any lattice) has for its open sets arbitrary unions of open intervals of the chain [19, §5.15].

Birkhoff's colleague Mac Lane at Harvard invented functors too late for any impact on Birkhoff's invention. The sum (coproduct) functor on Set is reflected into Pos to define the extension of $X;Y$ from objects to morphisms, readily seen to be a functor. One might suppose the same to be true of ordinal product, but functoriality fails for the second argument as may be illustrated with the quotient $f : 2 \to 1$ of the two-element chain $2 = \{0 < 1\}$ to the singleton poset $\{0\}$. Consider the morphism $1_2 \cdot f : (2 \cdot 2) \to (2 \cdot 1)$. Viewing $2 \cdot 2$ as $00 < 10 < 01 < 11$ and $2 \cdot 1$ as $00 < 10$, $2 \cdot f$ must take 00 and 01 to 00, and 10 and 11 to 10. But then $(2 \cdot f)(01) < (2 \cdot f)(10)$, violating monotonicity.

Ordinal product is however functorial in its first argument. To see this, let X vary while holding Y fixed. It suffices to show that for any monotone function $f : X \to X'$ (the agent of variation), the function $f \cdot 1_Y : X \cdot Y \to X' \cdot Y$ given by $(f \cdot 1_Y)(x, y) = (f(x), y)$ is monotone. That is, if $(x, y) \leq (x', y')$ then $(f(x), y) \leq (f(x'), y')$. But when $y \neq y'$ the order depends only on y and y', while when $y = y'$ it depends only on x and x', satisfying monotonicity in either case.

5.2 First Functor

Define $F_1 : \mathsf{Pos} \to \mathsf{Pos}$ as

$$F_1(X) = X \cdot \omega.$$

Theorem 5.1 *The final coalgebra of F_1 is order isomorphic to the unit real interval $[0, 1)$ (equivalently the nonnegative reals $\mathbb{R}^+$) standardly ordered.*

Proof. Since Pos has all limits including the empty limit 1, we may construct the final coalgebra of F_1 as the limit of $\ldots \to F_1(F_1(1)) \to F_1(1) \to 1$, where the map from $F^{i+1}(1)$ to $F^i(1)$ is $F^i(!)$ and $! : F(1) \to 1$ is the unique map to 1. Since limits in Pos are computed as for Set on the elements (unlike the case of colimits, coequalizers being the troublemaker) it follows that the final coalgebra of F_1 has as underlying set ω^ω, i.e. streams of natural numbers. The induced order on ω^ω can then be verified to be lexicographic, giving the monotone dyadics of section 3, the first representation we treated, which we saw there to be order isomorphic to the reals. □

Corollary 5.2 *ω^ω with the order interval topology induced from its lexicographic ordering is homeomorphic to $[0, 1)$ with its usual topology, since the latter is also the order interval topology.*

5.3 Second Functor

Define $F_2 : \mathsf{Pos} \to \mathsf{Pos}$ as

$$F_2(X) = 1; (X^o \cdot \omega).$$

Here X^o denotes the order dual of poset X while 1 denotes the singleton poset. The constituent operations of the definition all being functors, so is F_2. (Note that X^o as a functor does not reverse morphisms, only objects.) Since **Pos** is complete, F_2 has a final coalgebra.

Theorem 5.3 *The final coalgebra of F_2 is order isomorphic to* $[0, 1)$.

Proof. As for F_1 we start from the corresponding final coalgebra for **Set**. The only relevant difference is 1;, which in **Set** becomes 1+, for which the final coalgebra augments the set ω^ω of streams with the set ω^* of lists.

Passing now to **Pos**, the effect of X^o is to alternate the lexicographic order: the zeroth, second, fourth, ... numbers in the stream are ordered standardly while those in the odd positions are reverse ordered. (Interchanging even and odd here is immaterial, yielding an isomorphic final coalgebra.) This ordering corresponds to the alternating dyadics, the second representation of the continuum treated in section 3. □

Other compositions of familiar functors on **Pos** suggest themselves. Their final coalgebras will necessarily exist, **Pos** being complete as noted above, but they need not be order-isomorphic to the continuum. Two obvious functors intermediate between $F_1(X)$ and $F_2(X)$ are $\omega \cdot X^o$ and $1 + \omega \cdot X$. Up to homeomorphism, the final coalgebra of the former is $\mathbb{R} - \mathbb{Q}$, the irrationals, or Baire space. That of the latter is $\mathbb{R} + \mathbb{Q} + \{-\infty\}$, the rationals duplicated (as an ordered pair with nothing in between), or Cantor space. [5]

6 Future work

As explained already by Brouwer, canonical representatives make algebraic operations on reals undecidable. The problem is circumvented by redundant representations. Indeed, we originally obtained the alternating dyadics as a retract of a coalgebra with redundant representations of reals, developed following Conway's game theoretic constructions. Conway's description of the field structure readily lifts to this coalgebra. It seems likely that the structure can then be transferred to alternating dyadics along the retraction, but this has not yet been worked out in detail. The structure could then also be transferred to continued fractions along the isomorphism described in section 3.4. We are hoping to return to this theme in a forthcoming paper.

While not suitable for direct algebraic operations, the irredundant coalgebras described here will probably have a role in providing quick output, e.g. in particular applications of Newton or Runge-Kutta method. This follows from

[5] To see the correspondence between the recursive middle-third construction of Cantor space and the duplicate-rationals construction, also obtained as the order filters of $\mathbb{Q}$ in the course of Dedekind's construction, follow the removal of $(\frac{1}{3}, \frac{2}{3})$ by stretching the lower third $[0, \frac{1}{3}]$ to $[0, \frac{1}{2}]$ and dually for the upper third, thereby restoring everything while duplicating $\frac{1}{2}$. The recursion similarly duplicates the remaining dyadic rationals, a dense set in $[0, 1)$ and therefore sufficient for the claim. This argument points up the importance of defining the middle third to be open: taking it to be closed would instead produce Baire space, while taking it to be half-open would simply reproduce the continuum!

their "best approximation" properties, but also obviously requires detailed research.

We have been unable to answer the following question arising out of this work. The two functors $1;X$ and $X \cdot \omega$ involve respectively sum and product, albeit of the ordinal kind. This hints at some sort of duality between numbers and reals. One can talk vaguely of disconnected vs. connected, but does a more formal duality lurk in the shadows there?

References

[1] P. Aczel, *Non-Well-Founded Sets.* Lecture Notes 14 (CSLI 1988)

[2] J. Barwise and L.S. Moss, *Vicious Circles* (CSLI 1997)

[3] G. Birkhoff, An Extended Arithmetic, Duke Mathematical Journal, 3(1937), 311–316

[4] G. Birkhoff, Generalized Arithmetic, Duke Mathematical Journal, 9(1942), 283–302

[5] A. Edalat and P.J. Potts, A new representation for exact real numbers. *Electr. Notes in Theor. Comput. Sci.* 6(1997)

[6] R.W. Gosper, HAKMEM items 101A and 101B, Research report 239(1972), Artificial Intelligence Laboratory, MIT

[7] F. Hausdorff, *Grundzüge der Mengenlehre*, 3rd edition (Leipzig 1937)

[8] B. Jacobs, Coalgebraic specifications and models of deterministic hybrid systems. In: M. Wirsing and M.Nivat (eds.), *Proceedings of AMAST '96* Lecture Notes in Computer Science 1101 (Springer 1996) 520–535

[9] W.B. Jones and W.J. Thron, *Continued Fractions. Analytic Theory and Applications.* Encyclopedia of Mathematics and its Applications, Vol. 11 (Addison-Wesley 1980)

[10] P. Kornerup and D.W. Matula, finite precision lexicographic continued fraction number systems. *Proceedings of the Seventh Symposium on Computer Arithmetic* (IEEE 1985) 207–214

[11] J. Lambek, Subequalizers. *Canadian Math. Bull.* 13(1970) 337–349

[12] N.P. Mendler, P. Panangaden and R.L. Constable, Infinite Objects in Type Theory. *Proceedings of the First LICS*, (IEEE 1986) 249–255

[13] D.M.R. Park, *Concurrency and Automata on Infinite Sequences*, Lecture Notes in Computer Science 104 (Springer 1984)

[14] D. Pavlović, Guarded induction on final coalgebras. *Electr. Notes in Theor. Comput. Sci.* 11(1998) 1–18

[15] D. Pavlović, Towards semantics of guarded induction. *Submitted*

[16] D. Pavlović and M.H. Escardó, Calculus in coinductive form. *Proceedings of LICS '98* (IEEE 1998) 408–417

[17] H. Reichel, An approach to object semantics based on terminal co-algebras. *Math. Structures Comput. Sci.* 5(1995) 129–152

[18] J.J.M.M. Rutten, Universal coalgebra: a theory of systems. To appear in *Theoret. Comput. Sci.*

[19] E. Schecter, *Handbook of Analysis and Its Foundations* (Academic Press 1997)

[20] J. Vuillemin, Exact real computer arithmetic with continued fractions. *IEEE Trans. Comp.* 39(1990) 1087–1105