Rose-Hulman Undergraduate Mathematics Journal

Volume 5
Issue 1

Article 4

Generalized Cantor Expansions

Joseph Galante
University of Rochester, jg006k@mail.rochester.edu

Generalized Cantor Expansions

Joseph Galante
University of Rochester

1 Introduction

In this paper, we will examine the various types of representations for the real and natural numbers. The simplest and most familiar is base 10, which is used in everyday life. A less common way to represent a number is the so called Cantor expansion. Often presented as exercises in discrete math and computer science courses [8.2, 8.5], this system uses factorials rather than exponentials as the basis for the representation. It can be shown that the expansion is unique for every natural number. However, if one views factorials as special type of product, then it becomes natural to ask what happens if one uses other types of products as bases. It can be shown that there are an uncountable number of representations for the natural numbers. Additionally this paper will show that it is possible to extend the concept of mixed radix base systems to the real numbers. A striking conclusion is that when the proper base is used, all rational numbers in that base have terminating expansions. In such base systems it is then possible to tell whether a number is rational or irrational just by looking at its digits. Such a method could prove useful in providing easy irrationality proofs of mathematical constants.

2 Motivation

2.1 Definition The Cantor expansion of the natural number n is

$$n = a_k * k! + a_{(k-1)} * (k-1)! + \ldots + a_2 * 2! + a_1 * 1!$$

where all the a_i (digits) satisfy $0 \le a_i \le i$

This definition is standard and found in several sources [see 8.2, 8.5]. The identity

$$n! = 1 + \sum_{i=0}^{n-1} i(i!) \qquad \text{(Equation 2.1.1)}$$

is crucial for the Cantor Expansion since it allows "carries" to occur. When adding numbers, Equation 2.2.1 provides a meaning to advance to the next term in the base.

2.2 Example

23 = 3*3!+2*2!+1*1!
24 = 23 + 1 = 3*3!+2*2!+1*1! + 1 = 4!

It is reasonably easy to show via induction that all natural numbers have a unique Cantor expansion. Rather than prove uniqueness at this time, we will offer a generalization of this concept, and then show that the regular Cantor expansion is a special case.

3 Generalization

Knowing that identity 2.1.1 is the key, it is conceivable that a more general identity will yield a more general result. Realizing that the original expansion relies on factorials and their properties, a generalization should therefore rely on defining a new and more general product. Rather than multiplying together the sequence of numbers 1,2,3,4…, we will consider functions which multiply positive integers in a given list.

3.1 Lemma Let $S = \{1, x_1, x_2, \ldots |$ where x_i is a natural number strictly greater than one$\}$. (Note that the index for the first element will be n=0.) If $p(n)$ is the n^{th} element of S (note $p(0)=1$) and $P(n) = \prod_{i=0}^{n} p(i)$, then our generalized identity becomes

$$P(n+1) = 1 + \sum_{i=0}^{n} (p(i+1) - 1)P(i) \qquad \text{(Equation 3.1.1)}$$

An inductive proof of Lemma 3.1 is presented in section 7.1. Using the generalized identity we can extend the concept of a Cantor expansion.

3.2 Definition The generalized Cantor expansion (GCE) of the natural number n with respect to the ordered sequence S is

$$\begin{aligned} &n = a_k * P(k) + a_{(k-1)} * P(k-1) + \ldots . + a_1 * P(1) + a_0 * P(0) \qquad \text{(Equation 3.2.1)} \\ &0 \le a_i < p(i+1), \quad 0 \le i \le k \end{aligned}$$

where S, p, and P are as defined above in Lemma 3.1. The sequence S is referred to as the base set or base sequence. It is important to note that the order of the elements in S matters, and the a_k's are called the digits of n in its GCE representation.

Notation varies, but in this paper we will use the format $(a_k \ldots a_0)_S$ where S is the base set or sometimes just $(a_k \ldots a_0)$ when the base set is implied. Other notations use matrix format to combine the digits and base sets [see 8.13]. Lastly, as an example, we often measure time itself using a limited form of a mixed radix system, which has a base set corresponding to S={1,60,60,12,...} and has the common format of "hh:mm:ss."

3.3 Theorem Given a base set S, any natural numbers can be written in generalized Cantor expansion.

Proof The proof is by induction. It is easy to see that Equation 3.1.1 holds for n=0 and 1. That is, $0=0*P(0)$ and $1=1*P(0)$ since $P(0)=1$ and $p(1)>1$.

Now assume the first n natural numbers can be written in GCE form. We need to show that (n+1) can be written this way as well. So we know then that:
$n = a_k * P(k) + a_{(k-1)} * P(k-1) + ... + a_1 * P(1) + a_0 * P(0)$ for $0 \le a_i < p(i+1), \quad 0 \le i \le k$
where a_k is the first nonzero digit of n, so that $n \ge P(k)$. (If the first term was zero, we could consider a smaller number of terms and re-label subscripts accordingly.) We break the inductive step up into two cases. See 3.5 for concrete examples of the cases.

Case I: There exists a place i strictly less than k, such that the i^{th} digit a_i is strictly less than $p(i+1)$-1. This case will cover the addition of one to a number n without any arithmetic carries into the kth place. We want the number n to have some digit before the kth place which when one is added to it, will not produce a carry.

Using this idea we see that the digits of n satisfy for some i,

$$n > a_i * P(i) + ... + a_0 * P(0) = y$$

The number y will always be strictly less than n since n will always have an additional nonzero term which y does not, namely the term a_k*$P(k)$.

When adding one, the strict inequality, becomes only the inequality

$$n \ge a_i * P(i) + ... + a_0 * P(0) + 1$$

By the inductive hypothesis, we see that there exists a valid generalized Cantor expansion for y since $n \ge y$. Thus we can rewrite y+1 as

$$y + 1 = a_i * P(i) + ... + a_0 * P(0) + 1 = (a_i') * P(i) + ... + (a_0') * P(0)$$

Our initial assumption about a_i will tell us that it will not grow larger than $p(i+1)$-1 from a carry and so the rewrite will not require using another place.

It now follows that

$$n + 1 = a_k * P(k) + ... + (a_{i+1})P(i+1) + (a_i * P(i) + ... + a_0 * P(0) + 1)$$
$$n + 1 = a_k * P(k) + ... + (a_{i+1})P(i+1) + (a_i') * P(i) + ... + (a_0') * P(0)$$

which is a valid generalized Cantor expansion.

Case II: Digits in all places except the k^{th} place are equal to $p(i+1)$-1.
This case covers a series of carries that terminates at the greatest digit place of the number, or possibly advances to the next place.

In this case $n = a_k * P(k) + (p(k) - 1) * P(k-1) + ... + (p(1) - 1) * P(0)$

Therefore,

$$
\begin{aligned}
n+1 &= a_k * P(k) + (p(k)-1) * P(k-1) + ... + (p(1)-1) * P(0) + 1 \\
&= a_k * P(k) + P(k) \quad \text{from Lemma 3.1} \\
&= (a_k + 1) * P(k) + 0 * P(k-1) + ... + 0 * P(0)
\end{aligned}
$$

which is a valid generalized Cantor expansion if $a_k+1< p(k+1)$

Otherwise, if $a_k+1=p(k+1)$ then

$$n = (p(k+1)-1) * P(k) + (p(k)-1) * P(k-1) + ... + (p(1)-1) * P(0)$$

and

$$n+1 = (p(k+1)-1) * P(k) + (p(k)-1) * P(k-1) + ... + (p(1)-1) * P(0) + 1$$

$$n+1 = 1 * P(k+1) + 0 * P(k) + ... + 0 * P(0) \quad \text{from Lemma 3.1}$$

which is a valid generalized Cantor expansion.

Therefore by the principle of induction, we can conclude that all natural numbers can be expressed in a generalized Cantor expansion. ■

Now that we know every natural number has a generalized Cantor expansion, the question of uniqueness arises.

3.4 Theorem The generalized Cantor expansion of a natural number is unique.

The proof of this theorem uses induction and is complicated slightly by needing several different cases. The proof is found in appendix section 7.3 for the curious reader.

3.5 Examples

3.5.1 – Base 10, S={1,10,10,10…}

Example of Case 1
Let n=32378 in regular base 10, and then n+1 = 32378 + 1 = 32379. The addition of the number one does not produce a carry which changes the digit in the leftmost place. Thus we can think of n as 30000+2378, and n+1 as 30000+2378+1. In the proof, we used the inductive hypothesis to argue that 2378 is strictly less than n, and so 2378+1 is less than or equal to n, and so has a valid expansion, in this case 2379. Then 30000+2379=32379 = n+1 is also valid as an expansion. Our other initial assumption in this case was that 2378 was a number such that 2378+1 would not yield 10000.

Example of Case 2
Let n=39999 and then n+1 = 39999 +1 = 40000. This case uses carries so that when one is added to the lowest (rightmost) digit, it effects other digits.

Example of Case 2
Let n=99999 and then n+1 = 99999 +1 = 100000. This case uses carries so that when one is added to the lowest (rightmost) digit, it affects all the other digits and results in n+1 requiring a 6 digit representation.

3.5.2 – Mixed Radix, S=set of squares of natural numbers={1,4,9,16,…}
$p(0)=1$, $p(1)=4$, $p(2)=9$
$P(0)=1$, $P(1)=4*1$, $P(2) = 9*4*1$

Example of Case 1
Convert 137 into general Cantor expansion for the given S
$137 = 3*P(2) + 7*P(1)+1*P(0) = 3*(9*4*1) + 7*(4*1) + 1*(1)$
$137 \Rightarrow (3\ 7\ 1)_S$

$137+1=138= 3*P(2) + (7*P(1) + 1*P(0)+1)= 3*(9*4*1) + 7*(4*1) + 2*(1)$
$138 \Rightarrow (3\ 7\ 2)_S$

Example of Case 2
Convert 71 into general Cantor expansion for the given S
$71 = 1*P(2) + 8*P(1)+3*P(0) = 1*(9*4*1) + 8*(4*1) + 3*(1)$
$71 \Rightarrow (1\ 8\ 3)_S$

$71 + 1 = 72 = 2*P(2) + 0*P(1) + 0*P(0) = 2*(9*4*1)$
$72 \Rightarrow (2\ 0\ 0)_S$

Case 2 used:
Convert 575 into general Cantor expansion for the given S
$575 = 15*P(2) + 8*P(1)+3*P(0) = 15*(9*4*1) + 8*(4*1) + 3*(1)$
$575 \Rightarrow (15\ 8\ 3)_S$

$575 + 1 = 576 = 1*P(3) + 0*P(2) + 0*P(1) + 0*P(0) = 1*(16*9*4*1)$
$576 \Rightarrow (1\ 0\ 0\ 0)_S$

These examples additionally illuminate the fact that some representations using GCE need not cover every possible combination of digits used. We note that if we start counting upwards using the base set in example 3.5.2, our representations leap from $(1\ 8\ 3)_S$ to $(2\ 0\ 0)_S$. Thus the range of digits from $(1\ 8\ 4)_S$ to $(1\ 9\ 9)_S$ is off limits since it breaks the rules of our representation. If we choose to break the rules, then we lose uniqueness of representations.

It is interesting to consider the cardinality of the set of all possible generalized Cantor expansions.

3.6 Theorem There are an uncountable number of base sets S which can be used to make generalized Cantor expansions.

Proof

Each expansion has a unique base set S which characterizes the expansion.

$$S = \{1, x_1, x_2, \ldots \mid where\ x_i > 1, x_i \text{ a natural number}\}$$

We can construct another base set *S'* with a smaller range of terms such that

$$S' = \{1, x_1 \bmod 10, x_2 \bmod 10, \ldots \mid \text{where } x_i \in S\}$$

There is a one to one correspondence between the elements of *S'* and the digits of a real number y = $1.y_1\, y_2\, y_3 \ldots$ (standard notion of base 10 representation for real numbers is used here).

Since every value of every variable will be reached if we consider the set of all *S*'s, then every real number on the continuous interval [1,2) will be reached via its decimal expansion. Thus the cardinality of the set of all *S'* is the same as that of the real numbers, and thus is uncountable. The set of all *S*, which is a larger set, is also then uncountable.

Therefore there are an uncountable number of base sets *S* which can be used to make generalized Cantor expansions. ■

Now with powerful new facts about generalized Cantor expansions, we can examine how specific number systems fit into this definition. For example, we will show how the Cantor expansion and also our regular base 10 number system fit into the picture.

3.7 Factorials and the Cantor Expansion

To see how the original Cantor expansion is a special case of the GCE, let $S = \{1,2,3,4,5\ldots.\}$. It easily follows that $p(i) = i+1$ and $P(i) = (i+1)!$
Identity 2.1.1 becomes

$$1+\sum_{i=0}^{n-1}(p(i+1)-1)*P(i) = 1+\sum_{i=0}^{n-1}(i+1)(i+1)! = 1+\sum_{i=1}^{n} i*(i!) = 1+\sum_{i=0}^{n} i*(i!) = (n+1)!$$

which, is actually the original identity shifted up by one iteration. Then

$$\begin{aligned} n &= a_k*P(k) + \ldots a_1*P(1) + a_0*P(0) \\ &= a_k*(k+1)! + \ldots a_1*2! + a_0*1! \end{aligned}$$

Where $0 \le a_i \le i+1$. The more natural notation in this example would be to let a_i be associated with $i!$ Therefore shifting indices yields

$$a_{k+1}*(k+1)! + \ldots a_1*2! + a_1*1!$$

Where $0 \le a_i \le i$.

Proof of the factorial identity (Equation 2.2.1) is an exercise in [8.2].

3.8 Base-b

In general, base-b numbers can be represented using $S = \{1,b,b,b,\ldots\}$ for $b>1$, b a natural number. It follows that $p(0)=1$, $p(\text{i}) = b$ for $i>0$, and $P(i) = b^i$.

This reduces equation 3.1.1 to

$$P(n+1) = 1 + \sum_{i=0}^{n} (b-1) * b^i = b^{(n+1)} \qquad \text{(Equation 3.8.1)}$$

Additionally the coefficients will be between $0 \le a_i \le (b - 1)$ and

$$n = a_k * b^k + a_{(k-1)} * b^{(k-1)} + \ldots + a_0$$

which is the standard definition of a number in base-b notation

(Proof of equation 3.8.1 is an example in [8.2])

3.9 Other Interesting Examples of Mixed Radix

There are several other interesting cases to consider for the base set S.

Letting $S = \{1,2,3,5,7,11,13 \ldots \text{(primes of increasing order)} \ldots\}$
The consecutive products become what are known as primorials. (See [8.11] for an overview of the properties of primorials.) By picking this base set, we can write numbers as sums of primorials.

17 = 2*(3*2*1) + 2*(2*1) + 1*(1) => $(2\ 2\ 1)_S$
42 = 1*(5*3*2*1) + 2*(3*2*1)+0*(2*1)+0*(1) => $(1\ 2\ 0\ 0)_S$

We can also go the other direction and note that certain prime numbers have nice sums attached to them.

$(1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1)_S$ =>
1*(31*29*23*19*17*13*11*7*5*3*2*1) + 1*1=200560490131 which is prime

$(9\ 8\ 7\ 6\ 5\ 4\ 3\ 2\ 1)_S$ => 9*(19*17*13*11*7*5*3*2*1) +
8*(17*13*11*7*5*3*2*1)+7*(13*11*7*5*3*2*1)+6*(11*7*5*3*2*1)+5*(7*5*3*2*1)+
4*(5*3*2*1)+3*(3*2*1)+2*(2*1)+1*(1) = 91606553 which is prime

4 The Leap to The Reals

Having considered several cases with the natural numbers, it becomes logical to question whether Generalized Cantor Expansions can be extended to the real numbers.

4.1 The New Identity

Let S be a sequence of natural numbers where the first element is one and all other elements are greater than one. (Note that the index for the first element will be n=0.)

Define $p(n)$ = the n^{th} element of S (note p(0)=1) and $P(n) = 1/[\prod_{i=0}^{n} p(i)]$.

Our generalized identity then is

$$1 = \sum_{i=0}^{n} ((p(i)-1) * P(i)) + P(n) \qquad \text{(Equation 4.1.1)}$$

An inductive proof of 4.1.1 can be found in section 7.2. With 4.1.1, which in some respects is similar to 3.1.1 for natural numbers, we can create a new definition.

4.2 Definition

A number x, $0 \le x < 1$ can be represented in a Fractional Generalized Cantor Expansion (FGCE) with respect to the base set S if and only if

$$x = c_1 * P(1) + c_2 * P(2) + \ldots = \sum_{i=1}^{\infty} (c_i * P(i)) \qquad \text{(Equation 4.2.1)}$$

where $0 \le c_i < p(i)$, with p, P, and S as defined in 4.1. The sequence S is referred to as the base set or base sequence. It is important to note that the order of the elements in S matters, and the c_i's are called the digits of n in its FGCE representation.

We can write in short hand $x = (\,.\,c_1\, c_2\, c_3 \ldots)_S$

It would be nice if all FGCE's converge so that our definition is well defined, but first we must know some of the properties of the function P.

4.3 Lemma *P* converges to zero as *n* approaches infinity.

Proof

Since $x_i \ge 2$ for all $i \ge 1$, then

$$0 < P(n) = \frac{1}{x_1 * x_2 * \ldots * x_n} \le \frac{1}{2*2*\ldots*2} = (1/2)^n \quad \text{for all } n$$

As n approaches infinity, $(1/2)^n$ approaches zero, and thus P converges to zero by the squeeze theorem. ■

With this nice property of P, we can continue.

4.4 Theorem For a given base set *S*, all FGCE series are convergent.

Proof

If we use $c_i=p(i)-1$ for each i , then Equation 4.1.1 becomes

$$1=\sum_{i=1}^{n}(\,(p(i)-1)*P(i)\,)\;+\;P(n)$$

and it follows that

$$1>\sum_{i=1}^{n}(\,(p(i)-1)*P(i)\,)\geq 0$$

Thus the largest FGCE is bounded for any n. We then have

$$0\leq(1-\sum_{i=1}^{n}(\,(p(i)-1)*P(i)\,)\,)=\;P(n)$$

which converges by Lemma 4.3. Thus the sum converges as well.

We then note that

$$0\leq\sum_{i=1}^{n}(ci*P(i)))\leq\sum_{i=1}^{n}(\,(p(i)-1)*P(i))$$

since all coefficients c_i are satisfy $0\leq c_i\leq p(i)-1$, and we have convergence of the smaller sum by the comparison test.

Therefore all FGCE series converge. ■

4.5 Definition A terminating FGCE of length n is an FGCE that contains only a finite number of nonzero terms such that all the nonzero terms occur before the n+1 term, for some nonnegative integer n.

Example In base 10, 0.1742 would have a terminating FGCE of length 4, since all the nonzero terms occur before the 5th place. (The initial 0. does not count as a place.)

Example In base 10, 1/3=0.3333333… would not have a terminating FGCE.

4.6 Lemma Terminating FGCE's of length n divide the interval [0,1) up into increments of $P(n)$ for a given n and given base set S.

Proof

In this proof, we will be using both the GCE and FGCE, so we will denote the GCE base set as S' and the GCE P and p functions as P' and p'.

Let $S=\{1, x_1, x_2, \ldots, x_n\}$ (We do not care about terms after x_n)
Let $S'=\{1, x_n, x_{(n-1)}, \ldots, x_1\}$.
The reason for the strange indexing becomes apparent later, but note that $P(n)=P'(n)$.

Let $l = m/P(n)$, where $0 \leq m \leq P(n)-1$. We can see that as m varies between 0 and $P(n)-1$, that the l's divide up [0,1) into increments of length $P(n)$.

We now want to write m as a GCE

$$m = c_n + c_{n-1}(x_n) + c_{n-2}(x_n * x_{n-1}) + \ldots + c_2(x_n * \ldots * x_4 * x_3) + c_1(x_n * \ldots * x_3 * x_2)$$

With $0 \leq c_i < p'(i+1)$ (Note that we are counting down from n with our c_i's so from definition 3.2 $a_k = c_{n-k}$) Also we omit the term c_0 since it is zero, as $m < P'(n) = x_1 * x_2 * \ldots * x_n$

So $l = m/P(n) = m/(x_1 * x_2 * \ldots * x_n)$

After some shuffling of terms, we get

$$l = \frac{c_1\, x_2 * x_3 * \ldots * x_n}{x_1 * x_2 * \ldots * x_n} + \frac{c_2\, x_3 * x_4 * \ldots * x_n}{x_1 * x_2 * \ldots * x_n} + \ldots + \frac{c_n}{x_1 * x_2 * \ldots * x_n}$$

After simplification of the fractions, we get

$$l = \frac{c_1}{x_1} + \frac{c_2}{x_1 * x_2} + \ldots + \frac{c_n}{x_1 * x_2 * \ldots * x_n}$$

So,

$$l = \sum_{i=1}^{n} (c_i * P(i)) = c_1 * P(1) + \ldots + c_n * P(n)$$

At this point, we notice that the number l has been put into a terminating FGCE of length n since the coefficients satisfy $0 \leq c_i < p'(i+1) = x_i = p(i)$. So the l's, which are rational numbers represented by terminating FGCE's, divide up the interval [0,1) into increments of length $P(n)$. ■

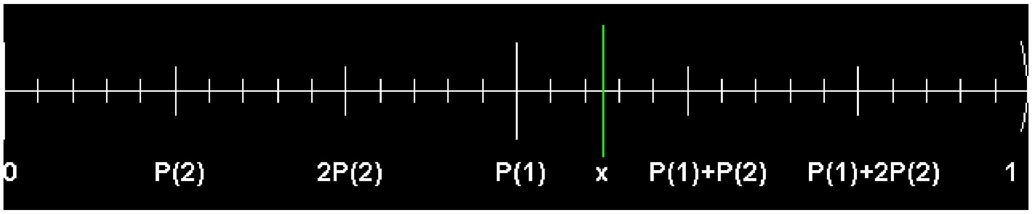
Next we show that each real number has a FGCE associated with it. The following theorem extends the concept of a Generalized Cantor Expansion to the real numbers in the unit interval [0,1). Once the numbers in [0,1) have FGCEs then it is relatively easy to extend the concept to all real numbers.

4.7 Theorem For a given base set S, each real number $0 \leq x < 1$ has a FGCE

Proof

This figure nicely illustrates the process which we will be employing.

Figure 4.7.1 - P's dividing up [0,1) with an x in between ($S = \{1,2,3,5\ldots\}$ shown)



We notice that for a given x, and a fixed n, we can pick c_i's so that

$$c_1*P(1)+c_2*P(2)+\ldots+c_n*P(n) \leq x \leq c_1*P(1)+c_2*P(2)+\ldots+(c_n+1)*P(n) \quad \text{(Eqn. 4.7.2)}$$

We pick the c_i's (digits) so that x's location in [0,1) can be specified within an interval of length $P(n)$. The coefficients determine the start and end points of this interval. Additionally, we want as tight a bound on x as possible with our choices. To do this, it is easiest to think of the process of picking the digits as an algorithm where you pick the largest c_1 available under the FGCE definition to satisfy $c_1*P(1) \leq x \leq (c_1+1)*P(1)$, then pick the largest c_2 available to satisfy $c_1*P(1)+c_2*P(2)\leq x \leq c_1*P(1)+(c_2+1)*P(2)$, and so on, each time keeping the c_i's from the previous step. You can keep doing this procedure indefinitely to obtain as many digits of x as desired.

(As an example, in figure 4.7.1 with S={1,2,3,5…},
for n=1, c_1=1 since $1*P(1) \leq x \leq (1+1)*P(1)$,
for n=2, c_1=1 and c_2=0 since $1*P(1)+0*P(2) \leq x \leq 1*P(1)+(0+1)*P(1)$,
for n=3, c_1=1, c_2=0, and c_3=2 ,
since $1*P(1)+0*P(2)+2*P(3)\leq x \leq 1*P(1)+0*P(1)+(2+1)*P(3)$.
In base 10, this procedure would be equivalent to saying $0.5\leq x<1$, $0.50\leq x\leq 0.66$, $0.566\leq x\leq 0.600$,…)

Our goal is to have the c_i's (digits) we just picked form a convergent series which goes to x as we add more digits. Rewriting equation 4.7.2, we see that our choice of c_i's for given n's actually results in partial sums of an FGCE. So for a given S, and an x in the interval [0,1), we have picked an n and c_i's (digits) that satisfies, for x a real number:

$$\sum_{i=1}^{n}(c_i * P(i)) \leq x \leq \sum_{i=1}^{n}(c_i * P(i)) + P(n)$$

Then,

$$0 \leq x - \sum_{i=1}^{n}(c_i * P(i)) \leq P(n)$$

As n approaches infinity, $x - \sum_{i=1}^{n}(c_i * P(i))$ is squeezed in between zero and $P(n)$, which is known to converge to zero from lemma 4.3. So, $x - \sum_{i=1}^{n}(c_i * P(i))$ approaches zero as n

approaches infinity, which is to say that $\sum_{i=1}^{n}(c_i * P(i))$ approaches x as n approaches infinity. Thus for the given x, there is an FGCE series which converges to x. ■

Unlike the natural number expansions, we cannot show that FGCE's for a given x and S are unique. Since FGCE's for real numbers are infinite series, they may have different coefficients, provided that they converge to x. All that is needed is that the expansions agree enough so that the difference between the expansion and x is arbitrarily small. For example in decimal, 1=0.9999999999999… This is the only type of non-uniqueness found using the real numbers.

As a useful convention, we can define the <u>Proper Fractional Generalized Cantor Expansion</u> of a real number x as the terminating expansion which is exactly equal to x. The proper FCGE may not exist for some numbers, since the number may have an infinite number of digits in the expansion to begin with, regardless of the base set S. For example in decimal (S={1,10,10,10…}), the number π would not have a proper FGCE since there is no finite sum of fractions which will equal π exactly.

4.8 Example – Proper FGCE

Let S={1,2,2,2,2…} which corresponds to the binary number system
Then we have 1/2= (**.** 1,0,0,0….)$_S$ or 1/2=(**.** 0,1,1,1,1…)$_S$
In this case (**.** 1,0,0,0….)$_S$ is the proper expansion since after one term, the expansion when evaluated, will give back the exact number it is meant to represent.

In regular decimal notation,
1.0000… is a proper FGCE for one, as opposed to 0.9999999…

Armed with the definition of a proper FGCE, we can notice that these expansions produce some very interesting properties with rational numbers. Here are some examples of expansions of both rational and irrational numbers to help guide us.

4.9 Example - FGCE Expansions

Consider the base set S = {1,2,3,4,5…}

e-2=(1 1 1 1..)$_S$ which is expected since $e=\sum_{k=0}^{\infty}1/k!$

Log(2) = (1 1 0 3 1 0 3 6 2 5 4 6 11..)$_S$ (pattern not apparent)

24/25 = (1 2 3 0 1 1 3 1 8 0 0 0 0..)$_S$

1/7 = (0 0 3 2 0 6 0 0 0 0..)$_S$

In base-10, different fractions have different periods. If the fraction contain only powers of 2 and 5 in the denominator, it will terminate. Otherwise it will repeat and its digits will form cycles. (See [8.4] for more about the decimal cycles of fractions.) It is also known that in other bases, different fractions terminate than those of base 10. For example in base 3, 1/3 has a terminating expansion, whereas 1/2 will have a repeating expansion. In fact in base 3, all fractions which have only powers of 3 in the denominator will have terminating expansions. It is interesting to look at Fractional Generalized Cantor Expansions which could have all powers of all primes in the base set. Then all fractions will have terminating expansions. The following theorem will show this.

4.10 Theorem Given a base set S and a rational number $x=a/b$ with $0 \leq x < 1$, x has a proper FGCE if and only if there exists an i such that b divides $x_1*x_2*...*x_i$.

Proof

We take $x=a/b$ with a and b in reduced form, so that a and b are relatively prime.

(Forward Direction)
Assume that x has a proper FGCE.

Then

$$x=\frac{a}{b}=\frac{c_1}{x_1}+\frac{c_2}{x_1*x_2}+\ldots+\frac{c_i}{x_1*x_2*...*x_i}$$

By simply putting all the fractions in common denominator form and adding them, we get

$$x=\frac{a}{b}=\frac{t}{x_1*x_2*...*x_i}<1$$

If $t / (x_1*x_2*...*x_i)$ is in reduced form, then $a=t$ and $b=(x_1*x_2*...*x_i)$, and so b divides $x_1*x_2*...*x_i$, since b divides itself.

If $t / (x_1*x_2*...*x_i)$ is not in reduced form, then both t and $(x_1*x_2*...*x_i)$ have some common factor $c > 1$. If we divide both t and $(x_1*x_2*...*x_i)$ by this common factor, then (t/c) and $((x_1*x_2*...*x_i)/c)$ become relatively prime. But then this would mean that $a=t/c$ and $b=((x_1*x_2*...*x_i)/c)$.

$b=((x_1*x_2*...*x_i)/c)$ and so $b*c=(x_1*x_2*...*x_i)$ which implies b divides $(x_1*x_2*...*x_i)$

In either situation b divides $x_1*x_2*...*x_i$.

(Reverse Direction)

Assume there exists an i such that b divides $x_1*x_2*...*x_i$

Then $b*d = x_1*x_2*...*x_i$ for some d.

So,

$$\frac{a}{b} = \frac{a*d}{b*d} = \frac{a*d}{x_1*x_2*...*x_i} < 1$$

By lemma 4.5, a FGCE divides the unit interval up into fractions with denominators of $P(i) = x_1*x_2*...*x_i$ for a given i, and a/b is a fraction of such a form. Therefore a/b has a finite, terminating expansion of length i, and thus it has proper FGCE. ∎

A necessary condition for S to give a specific fraction a terminating expansion is that it contains the factors of the denominator of x. Thus an appropriate choice of S to make all rational numbers have terminating expansions would be an S in which all possible factors are included. Once we have selected such a base set, theorem 4.10 holds for all rational numbers.

This is an interesting result since it implies that there could be base set in which all rational numbers between zero and one have a terminating expansion. We present such a base set now. Consider $S = \{1,2,3,4,5\ldots\}$. In this base set, there is always an i such that b divides $x_1*x_2*...*x_i$. This is so since all natural numbers are contained in the set and we can pick the i such that $x_i = i = b$. Then b divides $x_1*x_2*...*x_i$, since b divides $x_1*x_2*...*b$. And so this base has the remarkable property of making the separation of rational and irrational numbers apparent by their numerical representations. If we have a number whose irrationality we are not sure of, we can expand it out in with our special base set and if it terminates it must be rational. Conversely if we can prove that its representation will require an infinite number of nonzero digits with this base set, then it must be irrational.

Another interesting property of the base set $S = \{1,2,3,4,5\ldots\}$ is its expansion of e-2=$(1\ 1\ 1\ 1..)_S$. Using this set, irrational numbers may exhibit patterns in their infinite length expansions similar to the way some rational numbers do in bases without mixed radixes (ex. in base 10, 1/9 = 0.1111…). The argument in Theorem 3.6 can be used again on the base sets for FGCE's, so we know there are an uncountable number of base sets to choose from. It would be interesting for someone to examine these types of patterns further and look for other patterns in other base sets which meet the requirements for Theorem 4.10. Are there an infinite number of these types of base sets and do they all produce patterns with irrational numbers? I leave it to the curious reader to explore this further.

Finally, we realize that every real number consists of an integer and "decimal" part. The GCE allows all the integer parts to be represented, and the FGCE allows all "decimal" parts to be represented. Thus the expansion of a real number can be thought of as a list of coefficients for the GCE and FGCE with respect to two base sets. We can write for a real number

$$x=(a_n \ldots a_2\, a_1 \,.\, c_1\, c_2\, c_3 \ldots.)_{S,\,S'}$$

with the S and numbers to the left of the decimal point representing a GCE and the S' and numbers to the right of the decimal point representing a FGCE. Since there are an uncountable number of valid choices for both base sets, then there are an uncountable number of ways to represent the real numbers.

4.11 Example

If $S = S' = \{1,2,3,5,7\ldots$primes of increasing order$\}$, then
$\pi^{\pi} = 36 + 0.46215960\ldots$
$= (1\ 1\ 0\ 0\ .\ 0\ 2\ 3\ 6\ 0\ 7\ 11\ 1\ 21\ 4\ 28\ \ldots.)_{S,\,S'}$ (no apparent pattern).

5 Conclusion

Having taken a brief excursion into the world of numerical representations, we recall that some representation methods are common day, such as decimal, others are more exotic, such as the Cantor base expansion, and still others have many practical applications, such as binary. By extending our definition of base to a mixed radix system, the Generalized Cantor Expansion gives us an uncountable number of systems to represent the natural and real numbers. Additionally we have seen that by picking our representations carefully we can have special types of bases where is it easy to identify rational and irrational numbers by their expansions and that these bases may have other rich properties as well.

But the question now becomes "how general can we make our representations?" Donald Knuth [8.9] said that "π is 10 in base π". Perhaps the generalized representations presented here can be extended further to include rational and real numbers into the base sets. There are instances of using the golden ratio as a base [8.14] as well as negative numbers [see 8.15]. All of these representations could be very useful in digit related number theory. Now that we have shown that there are bases in which all rational numbers have terminating expansions, we can extend previously discrete and finite methods to include rational numbers. For example, the concepts of additive persistence and digital roots normally use base 10 numbers, however, now since it has been shown there are an uncountable number of bases to pick from, those concepts can be expanded [see 8.8, 8.10]. Such may be true of much in the field of digit related number theory.

6 Thanks

Special thanks to Ken McMurdy for constructive feedback, comments, and support that helped make this paper possible. Thanks to Will Jennings and George Bruhn for proof reading this paper and providing constructive feedback. Thanks to the referees for providing many helpful suggestions.

7 Appendix

This appendix contains proofs of lemmas and theorems left out of the main body. The proofs are of an uninteresting character, but are presented to be thorough.

7.1 – Proof of Lemma 3.1

We proceed by induction on n.

Basis: n=0
$p(1)=p(1)$
$p(1)*1=p(1)-1+1$
$P(1) = 1 + (p(1)-1)*1$ since $P(1)=1*p(1)$
$P(0+1) = 1 + (p(0+1)-1)*P(0)$ since $P(0)=p(0)=1$
Thus the basis case is established

We assume that equation 3.1.1 works for the first n natural numbers, then we need to show it works for $n+1$ as well.

$P((n+1)+1) = P(n+2) = p(n+2)*P(n+1)$ (from definition of $P(n)$)

$$= (p(n+2)-1)*P(n+1) + P(n+1)$$
$$= (p(n+2)-1)*P(n+1) + \left(1+\sum_{i=0}^{n}(p(i+1)-1)P(i)\right) \quad \text{(By the Inductive Hypothesis)}$$

$$P(n+2) = 1+\sum_{i=0}^{n+1}(p(i+1)-1)P(i)$$

Therefore by the principle of induction,

$$P(n+1) = 1+\sum_{i=0}^{n}(p(i+1)-1)P(i)$$ Holds for all natural numbers n. ■

7.2 – Proof of Lemma 4.1

We proceed by induction on n.

Basis:
n=0
$1=0*1+1=(1-1)*1+1=(p(0)-1)*P(0) + P(0)$
n=1
$1=p(1)/p(1)$
$1=(p(1)-1)/p(1) + 1/p(1)$
and since $P(1) =1/(p(1)*p(0))=1/(p(1)*1) = 1/p(1)$ and $(p(0)-1)*P(1) =0$

$1 = ((p(0)-1)*P(0) \;+\; (p(1)-1)*P(1)) + P(1)$
Thus the basis case has been established.

We assume equation 4.1.1 is true for *n* and we need to show it is true for (*n*+1).

$$1 = \sum_{i=0}^{n+1} ((p(i)-1) * P(i)) \; + \; P(n+1)$$

$$1 = \sum_{i=0}^{n} ((p(i)-1) * P(i)) \; + \; (p(n+1)-1)P(n+1) + P(n+1)$$

$$1 = \sum_{i=0}^{n} ((p(i)-1) * P(i)) \; + \; p(n+1)P(n+1)$$

$$1 = \sum_{i=0}^{n} ((p(i)-1) * P(i)) \; + \; p(n+1) * (P(n)/p(n+1))$$

$$1 = \sum_{i=0}^{n} ((p(i)-1) * P(i)) \; + \; P(n)$$

which by the inductive hypothesis is assumed to be true.

Therefore the identity is true for all $n \geq 1$. ■

7.3 – Proof of Theorem 3.4

We proceed by induction on *n*, the number of digits.

As a basis case, we first note that all single digit GCE's have a unique expansion. So if $n_1=a_0$ and $n_1=b_0$ are less than $P(1)$, then $a_0=b_0$. In a sense we consider single digits as the building blocks of all representations.

Now we will consider $n \geq P(1)$, and suppose that all numbers with *k*-1 digit representations have unique representations. We need to show that all *k* digit numbers have unique representations.

Suppose *n* has two *k*-digit representations.
$n = a_k * P(k) + a_{(k-1)} * P(k-1) + ... + a_1 * P(1) + a_0 * P(0)$ and
$n = (a_k)' * P(k) + (a_{(k-1)})' * P(k-1) + ... + (a_1)' * P(1) + (a_0)' * P(0)$

Let $m = a_k * P(k)$ and $r = a_{(k-1)} * P(k-1) + ... + a_1 * P(1) + a_0 * P(0)$
Let $m' = (a_k)' * P(k)$ and $r' = (a_{(k-1)})' * P(k-1) + ... + (a_1)' * P(1) + (a_0)' * P(0)$

So $n=m+r$ and $n=m'+r'$. Then $n-n = 0 = (m-m')+(r-r')$

Then $\text{abs}(m-m')=c*P(k)$ and $\text{abs}(r-r') < P(k)$

Case I: $m-m'>0$ and $r-r'>0$
Then $(m-m')+(r-r') > 0$, which is a contradiction.

Case II: $m-m'>0$ and $r-r' < 0$
Then $(m-m')+(r-r') \geq 1*P(k) + r-r' > 1*P(k) - P(k) = 0$
But $(m-m')+(r-r') > 0$ is a contradiction.

Case III: $m-m'<0$ and $r-r' > 0$
Then $(m-m')+(r-r') \leq -1*P(k) + r-r' < -1*P(k) + P(k) = 0$
But $(m-m')+(r-r') < 0$ is a contradiction.

Case IV: $m-m'<0$ and $r-r'<0$
Then $(m-m')+(r-r') < 0$, which is a contradiction.

So the only possible choice is then for $m-m'=0$. So then $m=m'$ and $r=r'$

Since $r=r'$ and r has k-1 digit representation, then r has a unique k-1 digit representation. Thus we conclude $a_0=a_0'\ldots a_{k-1} = a_{k-1}'$. Since $m=m'$ then $a_k=a_k'$. Thus n has a unique generalized Cantor expansion.

Therefore the generalized Cantor expansion of a natural number is unique. ■

8 References

8.1 Friedberg, Insel, Spence. Linear Algebra. Third Edition. Prentice Hall. New Jersey ISBN 0-13-233859-9

8.2 Johnsonbaugh, Richard. Discrete Mathematics. Second Edition. Macmillan Publishing Company. New York. ISBN 0-02-359690-2

8.3 Lady, E.L. "The Cantor Expansion of a Number." http://www.math.hawaii.edu/~lee/discrete/Cantor.pdf

8.4 Ore, Oystein. Number Theory and Its History. Dover Publications. Inc. New York. ISBN 0-486-65620-9

8.5 Rosen, Kenneth H. Discrete Mathematics. Fourth Edition. McGraw Hill. ISBN 0-07-289905-0

8.6 Sidorov, Nikita and P. Glendinning. Unique representations of real numbers in non-integer bases *Math. Res. Letters* **8** (2001), 535-543.

8.7 Spivak, Michael. Calculus. Third Edition. Publish or Perish Inc. ISBN 0-914098-89-6

8.8 Weisstein, Eric. MathWorld. "Additive Persistence"
http://mathworld.wolfram.com/AdditivePersistence.html

8.9 Weisstein, Eric. MathWorld. "Base." http://mathworld.wolfram.com/Base.html
Note: The reference to Knuth was taken off of this site, which in turn got the reference from Knuth's Book:
Knuth, D. E. "Positional Number Systems." §4.1 in *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms, 3rd ed.* Reading, MA: Addison-Wesley, pp. 195-213, 1998

8.10 Weisstein, Eric. MathWorld. "Digital Root"
http://mathworld.wolfram.com/DigitalRoot.html

8.11 Weisstein, Eric. MathWorld. "Primorial"
http://mathworld.wolfram.com/Primorial.html

8.12 Wikipedia. "Dedekind Cut" http://www.wikipedia.org/wiki/Dedekind_cut

8.13 Wikipedia. "Mixed Radix" http://en2.wikipedia.org/wiki/Mixed_radix

8.14 Wikipedia "Golden Mean Base" http://en2.wikipedia.org/wiki/Golden_mean_base

8.15 Wikipedia "Numeral Systems" http://en.wikipedia.org/wiki/Numeral_system