On using bounded continued fractions to represent reals

Stephen Lucas[*]

(received 8 August 2003; revised 27 February 2004)

Abstract

The continued fraction representation of an arbitrary real will have partial quotients that exceed any specified upper limit, making their representation on a computer difficult. However, any real can be represented as the sum of two continued fractions with an upper bound on their partial quotients. Here, algorithms are developed to actually find these bounded continued fractions for arbitrary reals and rationals. Whereas the real form is not competitive with the standard representation of reals on a computer, the rational representation shows promise. We also conjecture that there are an infinite number of representations available for any real.

[*]School of Mathematics and Statistics, University of South Australia, Mawson Lakes, South Australia 5095, Australia. mailto:stephen.lucas@unisa.edu.au

Contents

1 Introduction

A simple continued fraction for a rational is

$$\frac{p}{q} = b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cfrac{\vdots}{b_{n-1} + \cfrac{1}{b_n}}}} \quad \equiv b_0 + \frac{1}{b_1 +} \frac{1}{b_2 +} \cdots + \frac{1}{b_n}$$

$$\equiv [b_0; b_1, b_2, \ldots, b_n],$$

where b_0 is an integer, and the b_i's for $i > 0$ are positive integers. The b_i's are called *partial quotients*. A straightforward procedure based upon the Euclidean algorithm can be used to find the simple continued fraction as

follows. Given x, set $x_0 = x$, then $b_i = \lfloor x_i \rfloor$ and $x_{i+1} = 1/(x_i - b_i)$ for $i = 0, 1, \ldots, n$. If x is irrational, the iteration continues indefinitely.

The *convergents* of x are the fractions $p_k/q_k = [b_0; b_1, \ldots, b_k]$ which approximate x. If k is even, $0 < x - p_k/q_k < 1/q_k^2$, else if k is odd, $-1/q_k^2 < x - p_k/q_k < 0$. Convergents can be calculated using

$$\frac{p_{-1}}{q_{-1}} = \frac{1}{0}, \quad \frac{p_0}{q_0} = \frac{b_0}{1} \quad \text{and} \quad \frac{p_i}{q_i} = \frac{b_i p_{i-1} + p_{i-2}}{b_i q_{i-1} + q_{i-2}} \quad \text{for} \quad i = 1, 2, \ldots .$$

Terminating continued fractions $[b_0; b_1, b_2, \ldots, b_n]$ can also be written as

$$[b_0; b_1, b_2, \ldots, b_n - 1, 1],$$

or as the infinite expansion

$$[b_0, b_1, b_2, \ldots, b_n, \infty, b_{n+2}, b_{n+3}, \ldots]$$

for arbitrary $b_{n+2}, b_{n+3}, \ldots$. Examples include

$$\begin{aligned}
27/31 &= [0; 1, 6, 1, 3], \\
\pi &= [3; 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, \ldots], \\
\sqrt{48} &= [6; 1, 12, 1, 12, \ldots], \\
\ln(2) &= [0; 1, 2, 3, 1, 6, 3, 1, 1, 2, 1, 1, 1, 1, 3, 10, 1, 1, 1, 2, 1, 1, 1, \ldots].
\end{aligned}$$

These, and many other properties of continued fractions can be found in texts like Hardy and Wright [7] or Olds [10].

Unfortunately, it is difficult to store the continued fraction of an arbitrary real efficiently on a computer, because its partial quotients can be arbitrarily large. Gauss originally conjectured, and Kusmin proved, that a partial quotient of almost every arbitrary real takes the value n with probability $\log_2(1+1/n(n+2))$. For example, about 41.5% of partial quotients are ones, 17% are twos, and 0.00015 % are 1000s. There is no obvious way to choose a maximum integer size for partial quotients. The aim of this paper is to develop efficient methods of representing arbitrary reals as continued fractions that can be easily stored on a computer.
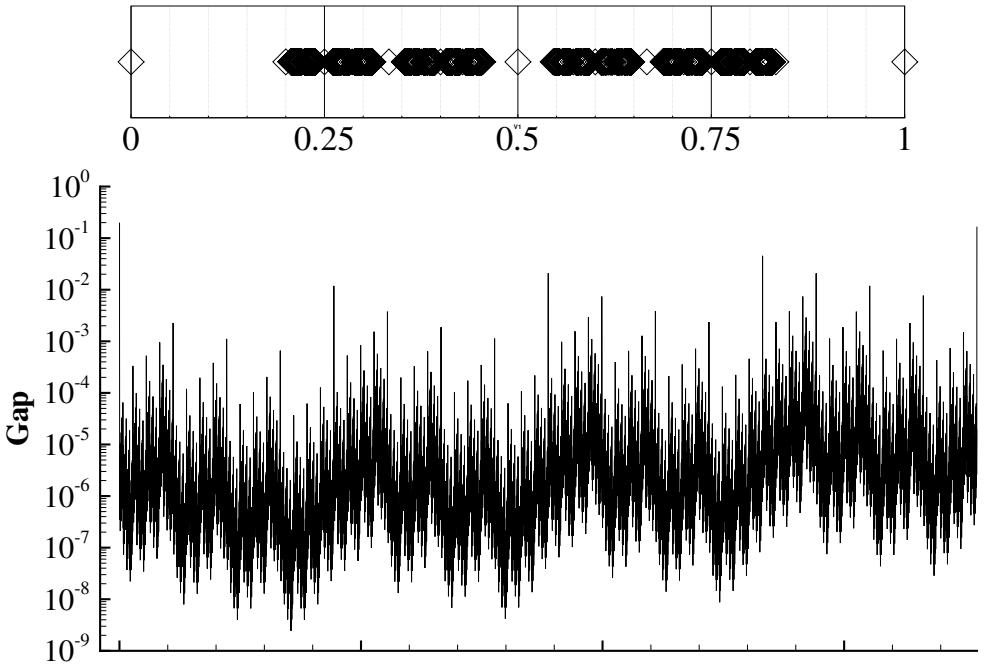
FIGURE 1: Distribution of elements of $F(4)$ with $b_0 = 0$ with up to eight partial quotients, and the gaps between these elements.

2 Bounded representations

A real x is called *badly approximable* if there exists an integer n such that for every rational p/q, $|x - p/q| > 1/nq^2$. Such numbers form a countably infinite set of measure zero. Let us define the set of bounded continued fractions $F(m)$ as the set $\{[b_0; b_1, b_2, \ldots] : b_0 \in \mathbb{Z},\ b_i \in \mathbb{N} \text{ and } 1 \leq b_i \leq m \text{ for } i \geq 1\}$. Then, for every $x \in F(m)$ and $p/q \in \mathbb{Q}$, $|x - p/q| > 1/(m+2)q^2$. Thus, all bounded continued fractions are badly approximable irrationals. For example, in Figure 1, the top graph plots all elements of $F(4)$ with $b_0 = 0$ and up to eight partial quotients — there are 17750 of them. The lower graph

shows the gaps between elements, and the sawtooth nature of these gaps is typical of the distribution of a Cantor set. The set $F(4)$ certainly cannot represent all reals.

In 1947, Hall [6] showed that $\mathbb{R} = F(4) + F(4)$, or that every real can be represented as the sum of two elements of $F(4)$. The proof is based on Cantor set theory, and is only valid for nonterminating continued fractions. If $b_0 = 0$, then $F(4) + F(4)$ includes the entire interval $(\sqrt{2} - 1, 4\sqrt{4} - 4)$. Later, Diviš [5] and Cusick [4] independently showed that $F(3) + F(3) \neq \mathbb{R}$ and $3F(2) \neq \mathbb{R}$, while $3F(3) = \mathbb{R}$ and $4F(2) = \mathbb{R}$. Hlavka [8] showed that $F(4) + F(3) = \mathbb{R}$, $F(4) + F(2) + F(2) = \mathbb{R}$, $F(3) + F(3) + F(2) = \mathbb{R}$, $F(7) + F(2) = \mathbb{R}$ but $F(4) + F(2) \neq \mathbb{R}$ and $F(3) + F(2) + F(2) \neq \mathbb{R}$. In a more recent sequence of papers, Astels [1, 2, 3] essentially completed this line of research and showed that $F(5) \pm F(2) = \mathbb{R}$, $F(3) \pm F(4) = \mathbb{R}$, $F(3) - F(3) = \mathbb{R}$, $F(3) \pm F(2) \pm F(2) = \mathbb{R}$, and $F(B_0) + F(B_0) = \mathbb{R}$, where B_0 is the set of positive odd integers. A good survey article in this area that predates Astel's is work is Shallit [11].

This theory, then, allows us to represent a real as the sum of two bounded continued fractions. In terms of efficiency, Hall's original result is particularly appealing; only two bits are required to represent each partial quotient. Figure 2 shows elements of $F(4) + F(4)$ with $b_0 = 0$ and up to eight partial quotients — 35938 in total. The gaps between elements get smaller as the number of partial quotients increase.

Unfortunately, all the results presented in the papers referenced here are existence proofs. No methods to actually construct the bounded continued fractions for an arbitrary real are described. In the next section, we derive such an algorithm.
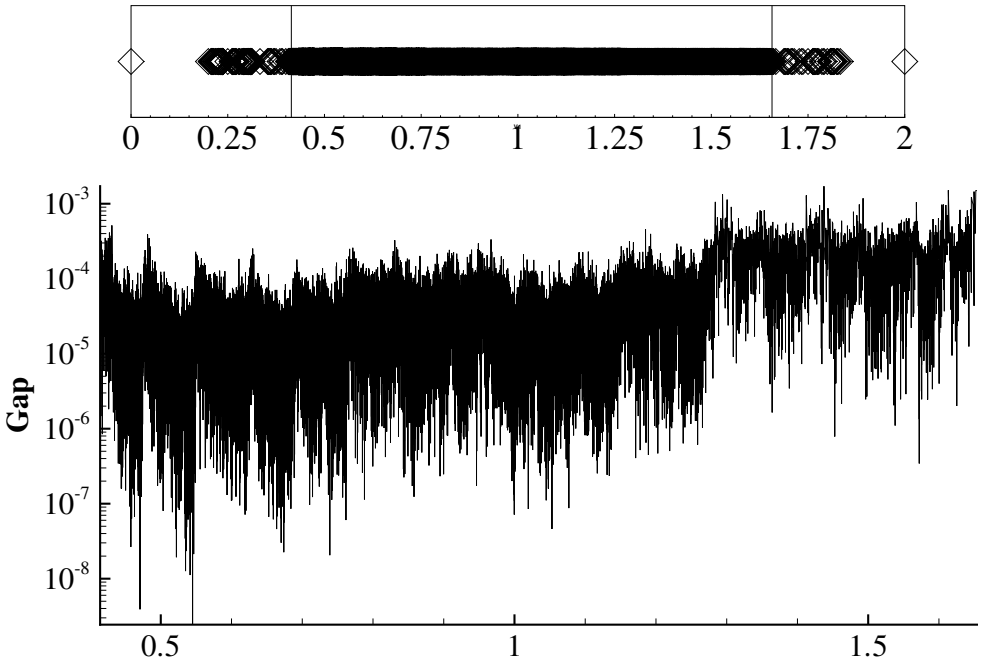
FIGURE 2: Distribution of elements of $F(4) + F(4)$ with $b_0 = 0$ with up to eight partial quotients, and the gaps between these elements.

3 An algorithm for reals

Hall [6] proved that any real x can be represented as $x_1 + x_2$, where $x_1, x_2 \in F(4)$. If $\{p_i/q_i\}$, $\{r_i/s_i\}$ are convergents for x_1, x_2 respectively, then if i is even (a similar result holds if i is odd), $1/(6q_i^2) < x_1 - p_i/q_i < 1/(q_i^2)$ and $1/(6s_i^2) < x_2 - r_i/s_i < 1/(s_i^2)$. These can be added to give

$$\frac{1}{6}\left(\frac{1}{q_i^2} + \frac{1}{s_i^2}\right) < x - \left(\frac{p_i}{q_i} + \frac{r_i}{s_i}\right) < \frac{1}{q_i^2} + \frac{1}{s_i^2}. \tag{1}$$

Thus, given approximations $[b_0; b_1, b_2, \ldots, b_{n-1}]$ and $[c_0; c_1, c_2, \ldots, c_{n-1}]$ for x_1 and x_2, we can calculate $[b_0; b_1, b_2, \ldots, b_{n-1}, i] + [c_0; c_1, c_2, \ldots, c_{n-1}, j]$ for $i, j = 1, 2, 3, 4$ (all sixteen of them), and choose cases that satisfy the appropriate inequalities. Hall's result guarantees that there will be at least one solution. The question then reduces to which solution to choose if there are more than one. Two possibilities are keeping all solutions in a tree structure, or choosing the best at any level, with the possibility of backtracking if there are no solutions on the current branch. In this context, the level of the tree equals the number of terms in each continued fraction.

3.1 Building the entire tree

As an example, let us consider $x = 0.90196078431373$. Building the tree of valid continued fraction approximation is straightforward. As an example, given the current continued fractions $[b_0; b_1] = [0; 1]$ and $[c_0; c_1] = [0; 3]$, the set of valid next partial quotients are $(b_2, c_2) = (1, 1)$, $(1, 2)$, $(1, 3)$, $(1, 4)$. Also, no continued fractions can be added to the tree below $[0; 1, 1] + [0; 4, 1]$, since none of the sixteen possibilities satisfy the required inequalities.

In fact, the number of valid continued fractions at each level of the tree grows quite quickly. For $x = 0.90196078431373$, Table 1 shows the number of valid $F(4) + F(4)$ representations at each level (where the level is the number

TABLE 1: Given $x = 0.90196078431373$, number of valid $F(4) + F(4)$ representations, the best at each level, and the errors.

Level	Number	Best Combination		Error	Power
2	8	[02]	+ [02]	9.8e-2	2.7e-2
3	23	[024]	+ [024]	1.3e-2	3.8e-4
4	71	[0241]	+ [0242]	2.6e-3	1.1e-4
5	197	[02413]	+ [02424]	1.4e-4	1.5e-6
6	577	[024134]	+ [02424]	2.1e-5	4.1e-7
7	1436	[0241323]	+ [0242324]	3.3e-6	5.8e-9
8	4998	[02423223]	+ [02413243]	2.2e-7	1.6e-9
9	13618	[024232234]	+ [024132434]	4.1e-8	2.3e-11
10	38652	[0242324213]	+ [0241323343]	3.8e-9	6.3e-12

of partial quotients for x_1 and x_2), the best representations, and their errors. The dramatic increase in the number of representations was unexpected, but occurred for every other real considered. The numerical evidence suggests the conjecture that the number of $F(4) + F(4)$ representations of an arbitrary real is in fact infinite. Returning to the original work of Hall [6], a proof suggests itself involving overlaps in the sums of intermediate intervals in the construction of Cantor sets, but is yet to be established.

3.2   Efficient representation of reals

The original motivation for this line of research was the efficient representation of reals on a computer. Hall's original paper showed that bounded continued fractions can represent reals, and the bound on the partial quotients means that a small number of bits is required at each level. Now that an algorithm has been developed to construct such a bounded continued fraction representation, we evaluate its efficiency.

As well as the errors for the best $F(4) + F(4)$ representation, Table 1 includes a column of errors labelled "Power" that list the errors if we use the standard binary form to represent x. To ensure that both the continued fraction and standard representations of x use the same number of bits, and we are making a fair comparison, the number of bits used in the standard representation is $4(l-1)$, where l is the level — increasing the length of the bounded continued fractions by one (adding two integers in the range one to four) increases the number of bits required by four. We see that the standard representation is superior to that using bounded continued fractions. Unfortunately, this result is typical, and shows that the bounded continued fraction representation is less efficient.

3.3   A greedy algorithm

To conclude this section, we consider the second suggested algorithm, where we only add to the tree of possible continued fractions on the node with lowest error. If there is no valid continued fraction below the best node, backtrack and continue down the next best. This will substantially reduce the computational effort, but at the expense of not searching the entire tree for the best representation. Returning to our $x = 0.90196078431373$ example once more, we find that the greedy algorithm finds the best representation up to the sixth level, but is less accurate at later levels. The best solution at the eighth level is also a dead end, and we need to backtrack to the second best approximation at the sixth level before continuing.

4   An algorithm for rationals

The example we have been considering up to now is in fact a decimal approximation to $46/51$. All the theory on representing reals by bounded

continued fractions is predicated upon the continued fractions having an in-finite number of partial quotients, and so the previous algorithms will never terminate. However, it turns out that $46/51$ can be represented by *finite* bounded continued fractions, as for example $4/17 + 2/3 = [0; 4, 4] + [0; 1, 2]$ or $5/17 + 31/51 = [0; 3, 2, 2] + [0; 1, 1, 1, 1, 4, 2]$. While the previous algorithm could find the first form if a check for equality is included at each step, it would never find the second, since the number of partial quotients in each bounded continued fraction is different.

Let us define the set of finite bounded continued fractions $F'(m)$ as the set $\{[b_0; b_1, b_2, \ldots, b_n] : n \in \mathbb{N},\ b_0 \in \mathbb{Z},\ b_i \in \mathbb{N} \text{ and } 1 \leq b_i \leq m \text{ for } 1 \leq i \leq n\}$. Can every rational x/y $(x, y \in \mathbb{N})$ be represented as a sum of two elements from $F'(4)$?

An algorithm can be developed based upon the idea of representing x/y as a sum of fractions, then choosing those sums where each fraction is an element of $F'(4)$. We thus need a technique for stepping through every possible sum of fractions. A simple procedure is to let $x/y = p/q + r/s$ where $x, y, p, q, r, s \in \mathbb{N}$, then set $q = s = ny$ for $n = 1, 2, 3, \ldots$, and for each n, let $p = 1, 2, \ldots, \lfloor nx/2 \rfloor$ and $q = nx - p$ as long as $\gcd(n, p) = 1$. This final gcd condition is to ensure we do not repeat fractions. We then rewrite the fractions in lowest form, and check whether they are in $F'(4)$. The easiest way of checking is to simply find the continued fraction representation of each fraction, and check its partial quotients.

Figure 3 shows an example of applying this procedure to $13/12$, and also shows which sums are of the form $F'(4) + F'(4)$. We see that stepping through in this manner will go through all sums of fractions adding to p/q, and in some sense list them in increasing magnitude. In this case, the shortest representation is $13/12 = 1/3 + 3/4 = [0; 3] + [0; 1, 3]$.

Returning to $x = 46/51$, we can find all $F'(4) + F'(4)$ representations as n increases. Table 2 shows the cumulative number of valid representations, which increases in a roughly linear manner. This behavior is typical of every

```
13/12:
= 1/12 + 12/12 = 1/12 + 1/1
= 2/12 + 11/12 = 1/6 + 11/12
= 3/12 + 10/12 = 1/4 + 5/6 = [04] + [0141] (Len 6)
= 4/12 + 9/12 = 1/3 + 3/4 = [03] + [013] (Len 5)
= 5/12 + 8/12 = 5/12 + 2/3 = [0222] + [012] (Len 7)
= 6/12 + 7/12 = 1/2 + 7/12 = [02] + [01122] (Len 7)
= 1/24 + 25/24
= 3/24 + 23/24 = 1/8 + 23/24
= 5/24 + 21/24 = 5/24 + 7/8
= 7/24 + 19/24  = [0323] + [01314] (Len 9)
= 9/24 + 17/24 = 3/8 + 17/24 = [0212] + [01223] (Len 9)
= 11/24 + 15/24 = 11/24 + 5/8
= 13/24 + 13/24
= 1/36 + 38/36 = 1/36 + 19/18
= 2/36 + 37/36 = 1/18 + 37/36
= 4/36 + 35/36 = 1/9 + 35/36
= 5/36 + 34/36 = 5/36 + 17/18
= 7/36 + 32/36 = 7/36 + 8/9
= 8/36 + 31/36 = 2/9 + 31/36
= 10/36 + 29/36 = 5/18 + 29/36
= 11/36 + 28/36 = 11/36 + 7/9 = [03312] + [0132] (Len 9)
```

$\vdots$

FIGURE 3: Stepping through all sums of fractions adding to 13/12.

TABLE 2: Given $x = 46/51$, cumulative number of valid $F'(4) + F'(4)$ representations as n increases.

n	10	20	30	40	50	60	70	80	90	100
Cumulative	24	53	85	113	135	163	196	233	265	297

example considered, and we conjecture that not only does $\mathbb{Q} = F'(4) + F'(4)$, but that the number of representations of each rational is infinite. Of course, the best case will be the one with the smallest number of partial quotients.

As it stands, this algorithm for representing rationals will be increasingly inefficient as the denominator becomes large. However, we will be able to dramatically cut down on the number of fractions to consider. By listing the elements of $F'(4)$, we discover that for a given integer p, $p/(p+q)$ is in $F'(4)$ when $q = k + pi$, $i = 0, 1, 2, 3$, $k \in \{1, 2, \ldots, p-1\}$, and $\gcd(p, p+k) = 1$. The possible values of k are limited due to the size of p. For example, with $p > 3$ and $\gcd(p, 3) = 1$,

$$\frac{p}{p+3} = \begin{cases} [0; 1, \lfloor p/3 \rfloor, 1], & \mod(p, 3) = 1, \\ [0; 1, \lfloor p/3 \rfloor, 1, 2], & \mod(p, 3) = 2, \end{cases}$$

so $p/(p+3) \in F'(4)$ only for $p < 15$. While it is difficult to generalise this result, and hence easily establish which k's to consider, this result will still reduce the number of fractions to check.

5   Conclusion

We have seen that bounded continued fractions can be used to represent both reals and rationals in what we conjecture are an infinite number of ways. Algorithms to actually construct these continued fractions have been presented, and while the real bounded continued fraction representation is not competitive compared to the standard computer representation of reals, the rational representation shows promise.

There are a variety extensions to consider on this work. The conjectures on the infinite number of representations need to be formally verified, and an efficient storage method for rationals developed more carefully. There are also a number of more efficient continued fraction representations worth

considering. Astels pointed out that $\mathbb{R} = F(4) \pm F(4)$, and so a more efficient representation could be formed at the cost of a single bit. Since $\mathbb{R} = F(4) + F(4)$, this means that $\mathbb{R} = F(n) + F(n)$ for $n \geq 4$. A different value of n may increase convergence rates — particularly for rationals. Finally, so-called *nearest integer continued fractions* allow for negative partial quotients, eliminating ones. They converge more quickly than simple continued fractions, and may be useful here.

References

[1] S. Astels, Sums of numbers with small partial quotients, *Proc. Am. Math. Soc.*, **130** (2002) 637–642. C925

[2] S. Astels, Sums of numbers with small partial quotients II, *J. Number Th.*, **91** (2001) 187–205. C925

[3] S. Astels, Cantor sets and numbers with restricted partial quotients, *Trans. Am. Math. Soc.*, **352** (2000) 133–170. C925

[4] T. W. Cusick, On M. Hall's continued fraction theorem, *Trans. Am. Math. Soc.*, **38** (1973) 253–254. C925

[5] B. Diviš, On the sums of continued fractions, *Acta. Arith.*, **22** (1973) 157–173. C925

[6] M. Hall Jr., On the sum and product of continued fractions, *Ann. Math.*, **48** (1947) 966–993. C925, C927, C928

[7] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers, 5th edition*, Oxford University Press, London, 1979. C923

[8] J. L. Hlavka, Results on sums of continued fractions, *Trans. Am. Math. Soc.*, **211** (1975) 123–134. C925

[9] D. E. Knuth, *The art of computer programming, volume 2: seminumerical algorithms (3rd edition)*, Addison–Wesley, Reading, Massachusetts, 1998.

[10] C. D. Olds, *Continued Fractions*, Random House, 1963. C923

[11] J. Shallit, Real numbers with bounded partial quotients: a survey, *Enseign. Math, (2)*, **38** (1992) 151–187. C925