A note on the computation of the fraction of smallest denominator in between two irreducible fractions

Isabelle Sivignon

Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, France
CNRS, GIPSA-Lab, F-38000 Grenoble, France

Abstract

Given two irreducible fractions f and g, with $f < g$, we characterize the fraction h such that $f < h < g$ and the denominator of h is as small as possible. An output-sensitive algorithm of time complexity $\mathcal{O}(d)$, where d is the depth of h is derived from this characterization.

Keywords: simplest fraction, continued fractions, algorithm

1. Introduction

Given two irreducible fractions defining an interval, we investigate the problem of finding the fraction of smallest denominator in this interval. Keeping integers as small as possible can be crucial for exact arithmetic computations. Indeed, even if exact arithmetic libraries [1] provide efficient tools to handle exact computations on arbitrarily large integers, dealing with big integers remains computationally expensive. Thus, when a parameter value has to be chosen in a given interval, it may be interesting to choose the value involving integers as small as possible to enable fast computations. This problem also has a nice geometric interpretation and arises for instance in the computation of the integer hull of a polygon [2].

In this note, we consider two irreducible proper fractions f and g such that $0 \leq f < g \leq 1$. We give a characterization of the irreducible proper fraction h such that $f < h < g$ and the denominator of h is as small as possible. Finally, we provide an output-sensitive algorithm to compute h from the continued fraction decompositions of f and g.

2. Problem statement and theorem

A *simple continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

The numbers a_i are called *partial quotients*. In the following, we restrict ourselves to the case where all a_i are positive integers. Any rational number f can be expressed as a simple continued fraction with a finite number of partial quotients. We denote $f = \frac{p}{q} = [a_0; a_1, a_2 \dots a_n]$, $p, q \in \mathbb{N}$. The m-th *principal convergent* of f is equal to $f_m = \frac{p_m}{q_m} = [a_0; a_1, a_2 \dots a_m]$. The sequence of fractions f_{2m}, $m \in 0 \dots \lfloor \frac{n}{2} \rfloor$, is increasing while the sequence of fractions f_{2m+1}, $m \in 0 \dots \lfloor \frac{n}{2} \rfloor$ is decreasing. Moreover, for any $a_i > 1$, the *intermediate*

Email address: isabelle.sivignon@gipsa-lab.grenoble-inp.fr (Isabelle Sivignon)

convergents are defined as $f_{i,k} = \frac{p_{ik}}{q_{ik}} = [a_0; a_1 \ldots a_{i-1}, k]$ with $1 \leq k < a_i$. When $k = a_i$, we have $f_{i,a_i} = f_i$. In the following, convergent stands for either a principal or an intermediate convergent.

For a given fraction f, we define two ordered sets of convergents as follows (see also [3, Chap. 32, Parag. 15]): $\Gamma_{even}(f) = \{f_{2m,k}, k \in 1 \ldots a_{2m}, m \in 0 \ldots \lfloor \frac{n}{2} \rfloor\}$, $\Gamma_{odd}(f) = \{f_{2m+1,k}, k \in 1 \ldots a_{2m+1}, m \in 0 \ldots \lfloor \frac{n-1}{2} \rfloor\}$. If $a_{2m} > 1$, we have $f_{2m,k-1} < f_{2m,k}$ for any $m \in 0 \ldots \lfloor \frac{n}{2} \rfloor$ and any $k \in 2 \ldots a_{2m}$, and if $n \geq 2$, $f_{2m,a_{2m}} < f_{2m+2,1}$ for any $m \in 0 \ldots \lfloor \frac{n}{2} \rfloor - 1$. Similarly, if $a_{2m+1} > 1$ we have $f_{2m+1,k-1} > f_{2m+1,k}$ for any $m \in 0 \ldots \lfloor \frac{n-1}{2} \rfloor$ and any $k \in 2 \ldots a_{2m}$, and if $n \geq 3$, $f_{2m+1,a_{2m}} > f_{2m+3,1}$ for any $m \in 0 \ldots \lfloor \frac{n-1}{2} \rfloor - 1$.

For the sake of clarity, we rename the elements of $\Gamma_{even}(f)$ and $\Gamma_{odd}(f)$ as follows: $\Gamma_{even}(f) = \{\gamma_0 = [a_0], \gamma_2 \ldots \gamma_{2i} \ldots\}$ and $\Gamma_{odd}(f) = \{\ldots \gamma_{2i+1} \ldots \gamma_3, \gamma_1 = [a_0; a_1]\}$.

Definition 1. A fraction f is said to be *less complex* (resp. *more complex*) than a fraction g if and only if the denominator of f is strictly lower (resp. strictly greater) than the denominator of g.

Before stating our main theorem, let us recall some classical definitions. The *Farey sequence of order* n is the ascending sequence of irreducible fractions between 0 and 1 whose denominator do not exceed n [4]. The *mediant fraction* of two given fractions $\frac{p}{q}$ and $\frac{p'}{q'}$ is defined as $\frac{p+p'}{q+q'}$. Given three successive terms $\frac{p}{q}$, $\frac{p'}{q'}$ and $\frac{p''}{q''}$ of a Farey sequence, $\frac{p'}{q'}$ is the mediant of $\frac{p}{q}$ and $\frac{p''}{q''}$ and we have $p'q - pq' = 1$ (similarly, $p''q' - p'q'' = 1$).

Theorem 1. *Let f and g be two irreducible fractions such that $0 \leq f < g \leq 1$. Let $\mathcal{H} = \{\frac{p}{q} \mid f < \frac{p}{q} < g\}$, and $h_{\min} = \arg\min_{\frac{p}{q} \in \mathcal{H}}\{q\}$. Then,*

(i) the fraction $h_{\min}$ is well (uniquely) defined;

(ii) if f and g are successive terms of a Farey sequence, $h_{\min}$ is equal to the mediant of f and g;

(iii) otherwise, if g is a convergent of f, $g = \gamma_i \in \Gamma_{odd}(f)$, then $h_{\min} = \gamma_{i+2}$;

(iv) otherwise, if f is a convergent of g, $f = \gamma_i \in \Gamma_{even}(g)$, then $h_{\min} = \gamma_{i+2}$;

(v) otherwise, $h_{\min}$ is equal to the unique fraction in $\mathcal{H}$ that is a convergent of both f and g.

The proof uses the following lemmas.

Lemma 1. *([3, Chap. 32, Parag. 15]) For any fraction f, "it is impossible between any consecutive pair of either" $\Gamma_{even}(f)$ or $\Gamma_{odd}(f)$ "to insert a fraction which shall be less complex than the more complex of the two".*

Lemma 2. *Let $\frac{p}{q}$ and $\frac{p'}{q}$ be two irreducible fractions with $0 < \frac{p}{q} < \frac{p'}{q} < 1$, $q \neq 1$. There exists a fraction $\frac{a}{b}$ such that $b < q$ and $\frac{p}{q} < \frac{a}{b} < \frac{p'}{q}$.*

Proof. Suppose that there is no fraction of denominator strictly lower than q between $\frac{p}{q}$ and $\frac{p'}{q}$. Then $\frac{p}{q}$ and $\frac{p'}{q}$ are two successive terms of the Farey sequence of order q. Therefore, $p'q - qp = 1$, and rewriting p' as $p + k$ with $k \in \mathbb{N}, k \geq 1$, we get $qk = 1$, which leads to a contradiction.

Lemma 3. *For any fraction f, and for all $i, j \in \mathbb{N}$, if γ_i and γ_j are two elements of $\Gamma_{even}(f)$ or $\Gamma_{odd}(f)$, then γ_i is more complex than γ_j if and only if $i > j$.*

Proof. The proof is straightforward from the definition of principal and intermediate convergents.

Proof. (of Theorem 1)

(i) Suppose that there exist two fractions $h = \frac{a}{b}$ and $h' = \frac{a'}{b}$ such that $b = \min_{\frac{p}{q} \in \mathcal{H}}\{q\}$ and $h, h' \in \mathcal{H}$. From the Theorem hypothesis, $0 < h < 1$, and thus $b > 1$. Then Lemma 2 directly leads to a contradiction. Moreover, a similar argument shows that the denominator of $h_{\min}$ is different from both the denominators of f and g.

(ii) In this case, as stated in [3, Chap. 32, Parag. 12] for instance, any fraction lying between f and g is more complex than f and g. Consequently, h is the fraction such that f, h and g are successive terms of a Farey sequence of order strictly greater than $\max\{f_q, g_q\}$, and by definition, h is the mediant of f and g.

(iii) Let $h \in \mathcal{H}, h \neq \gamma_{i+2}$. We show that h is more complex than γ_{i+2}. First, if $\gamma_{i+2} < h < g$, then Lemma 1 enables to conclude. If there exists j such that $h = \gamma_j$, then $j > i + 2$ since $h \in \mathcal{H}$, and we conclude using Lemma 3. Finally, if there exists j such that $\gamma_{j+2} < h < \gamma_j$, $j > i + 2$ then from Lemma 3, γ_j is more complex than γ_{i+2} and from Lemma 1, h is more complex than γ_j. So any fraction of $\mathcal{H}$ is more complex than γ_{i+2}. Moreover, since f and g are not successive terms of a Farey sequence, $\gamma_{i+2} \neq f$, which proves that $h_{\min} = \gamma_{i+2}$.

(iv) The proof is similar to item (iii).

(v) Since f and g are not convergents of one another and not successive terms of a Farey sequence, let γ_i and γ_{i+2} be the two convergents of f such that $\gamma_{i+2} < g < \gamma_i$. Similarly, let γ'_j and γ'_{j+2} be the two convergents of g such that $\gamma'_j < f < \gamma'_{j+2}$. Then, $h_{\min}$ is equal to γ_{i+2} or γ'_{j+2}. Indeed, there is no fraction less complex than γ_{i+2} between f and γ_{i+2} (by definition and Lemma 1), and if there was such a fraction between γ_{i+2} and g, then there would be a fraction less complex than γ_{i+2} between γ_{i+2} and γ_i, which is not possible from Lemma 1. A similar reasoning can be done with γ'_{j+2}. All in all, this implies that either $\gamma_{i+2} = \gamma'_{j+2}$ or $\gamma_{i+2} \neq \gamma'_{j+2}$ and they have the same denominator. Since this latter case is not possible (see (i)), $h_{\min} = \gamma_{i+2} = \gamma'_{j+2}$ which is the unique fraction in $\mathcal{H}$ that is a convergent for both f anf g.

The different cases of Theorem 1 have a nice interpretation using the Stern-Brocot tree [5, 6]. Cases (iii) and (iv) correspond to the case where f is an ancestor of g (or conversely) when f and g are not successive terms of a Farey sequence. Case (v) is the case where f and g have a unique common ancestor in the tree, which is the fraction of smallest denominator between them.

3. Algorithm and complexity

Algorithm 1 is an efficient algorithmic translation of Theorem 1[1]. It is based on the representation of f, g and $h_{\min}$ as continued fractions.

Let $f = [a_0; a_1, a_2 \dots a_n]$ and $g = [b_0; b_1, b_2 \dots b_m]$, with $a_n > 1$ and $b_m > 1$. We note that, when f and g are not successive terms of a Farey sequence, $h_{\min}$ is a convergent of f, g or both. In any case, this convergent can be defined by the knowledge of the common partial quotients of f and g.

In the algorithm, the partial quotients of f and g are computed on the fly using the Euclidean algorithm. We denote by $r_i(f)$ the difference between f and its i-th principal convergent: $r_i(f) = f - f_i$.

Theorem 2. *Given two proper irreducible fractions f and g such that $0 \leq f < g \leq 1$, Algorithm 1 computes the fraction h of smallest denominator such that $f < h < g$.*

PROOF. On line 1, the algorithm tests whether f and g are successive terms of a Farey sequence. If so, and following Theorem 1, h is set to the mediant of f and g.

Otherwise, the partial quotients of f and g are pairwise compared in the `while` loop of lines $7 - 9$. The algorithm exits the loop when either the i-th partial quotients of f and g differ or the i-th partial quotient is the last one for f or g. Indeed, the partial quotient a_i (resp. b_i) is the last partial quotient of f (resp. g) if and only if $r_i(f) = 0$ (resp. $r_i(g) = 0$). The tests between line 10 and line 33 cover all the possible exit cases. We review these cases here and prove case by case that the fraction h returned is the correct one.

[1] A python implementation of this algorithm is available on `www.gipsa-lab.fr/~isabelle.sivignon/Code/simplestFraction.py`

- The cases where $a_i = b_i$ and $r_i(f)$ or $r_i(g)$ is equal to zero are covered by lines $10-14$. Note that we cannot have $r_i(f) = r_i(g) = 0$ otherwise f would be equal to g. If $r_i(f) = 0$, then we have $f = [a_0; a_1 \ldots a_i]$ and $g = [a_0; a_1 \ldots a_i, b_{i+1}, \ldots b_m]$ with $m > i+1$ since f and g are not consecutive terms of a Farey sequence. Thus, by definition f is a principal convergent of g, and since $f < g$ by hypothesis, f is an even convergent of g. From Theorem 1 (iv), h is the next even convergent of g. We have $h = [a_0; a_1 \ldots a_i, b_{i+1}, 1] = [a_0; a_1 \ldots a_i, b_{i+1}+1]$. The case where $r_i(g) = 0$ is symmetrical and the proof is similar.

- If $a_i \neq b_i$ and neither $r_i(f)$ nor $r_i(g)$ are equal to zero (lines $16-17$), then we have $f = [a_0; a_1 \ldots a_{i-1}, a_i \ldots a_n]$ with $n > i$, $a_n > 1$, and $g = [a_0; a_1 \ldots a_{i-1}, b_i \ldots b_m]$ with $b_i \neq a_i$, $m > i$ and $b_m > 1$. f is not a convergent of g, nor conversely, and this case corresponds to case (v) of Theorem 1. The fraction $h = [a_0; a_1 \ldots a_{i-1}, \min(a_i, b_i) + 1]$ is a convergent of both f and g since $n, m > i$.

- If $a_i \neq b_i$ and $r_i(f) = r_i(g) = 0$ (lines $19-20$), then we have $f = [a_0; a_1, \ldots a_{i-1}, a_i]$ and $g = [a_0; a_1, \ldots a_{i-1}, b_i]$ with $b_i \neq a_i$. If $b_i < a_i$, g is an odd convergent of f, and f is an even convergent of g otherwise. In the first case, following Theorem 1 (iii), h is the next odd convergent of f, and $h = [a_0; a_1, \ldots a_{i-1}, b_i + 1]$. We have $h \neq f$ (i.e. $b_i + 1 \neq a_i$) since otherwise f and g would be successive convergents of f, and this case has been excluded by the test on line 1. The case $a_i < b_i$ is similar and proved using Theorem 1 (iv), leading to $h = [a_0; a_1, \ldots a_{i-1}, \min(a_i, b_i) + 1]$.

- If $a_i \neq b_i$ and $r_i(f) = 0$ and $r_i(g) \neq 0$ (lines $22-27$), then $f = [a_0; a_1, \ldots a_{i-1}, a_i]$ and $g = [a_0; a_1, \ldots a_{i-1}, b_i, \ldots b_m]$, with $m > i$ and $b_m > 1$. Three subcases arise, according to the respective values of a_i and b_i (keeping in mind that $a_i \neq b_i$):
  - if $a_i < b_i$ (line 23), then f is a convergent of g, and according to Theorem 1 (iv), $h = [a_0; a_1, \ldots a_{i-1}, a_i + 1]$;
  - if $a_i = b_i + 1$ (line 25), then $f = [a_0; a_1, \ldots a_{i-1}, b_i, 1]$ and f is an even convergent of g ($i+1$ is even). Using Theorem 1 (iv) again, h is the next even convergent of g. If $b_{i+1} > 1$, then $h = [a_0; a_1, \ldots a_{i-1}, b_i, 2]$. Otherwise, the next even convergent of g is equal to $[a_0; a_1, \ldots a_{i-1}, b_i, b_{i+1}, b_{i+2}, 1] = [a_0; a_1, \ldots a_{i-1}, b_i, b_{i+1}, b_{i+2} + 1]$.
  - otherwise, $a_i > b_i + 1$ (line 27), and f and g are not convergent of one another. Theorem 1 states that h is the unique fraction that is a convergent of both f and g. It is easy to see that the fraction $h = [a_0; a_1, \ldots b_i + 1] = [a_0; a_1, \ldots b_i, 1]$ is a convergent of both f (since $b_i + 1 < a_i$) and g (since $b_{i+1} \geq 2$).

- The last case, $a_i \neq b_i$ and $r_i(g) = 0$ and $r_i(f) \neq 0$ (lines $28-33$) is symmetrical to the previous one.

Lemma 4. *In a computing model where standard arithmetic operations are done in constant time, Algorithm 1 has a complexity of $\mathcal{O}(d)$ where d is the number of partial quotients of h.*

PROOF. The representation of f and g as continued fractions is computed on the fly, and only the partial quotients used to define the result fraction h are computed. The computation of the partial quotients a_i and b_i consists in one step of the Euclidean algorithm that uses standard arithmetic operations. Thus, the `while` loop runs in $\mathcal{O}(d)$ time, and all other operations from line 10 to line 33 are done in constant time. The computation of the fraction h from its representation as a continued fraction also runs in $\mathcal{O}(d)$ time since standard arithmetic operations are done in constant time.

References

[1] T. Granlund, the GMP development team, GNU MP: The GNU Multiple Precision Arithmetic Library, 5th Edition, `http://gmplib.org/` (2012).

[2] W. Harvey, Computing two-dimensional integer hulls, SIAM Journal on Computing 28 (6) (1999) 2285–2299.

[3] G. Chrystal, Algebra: An Elementary Text-book for the Higher Classes of Secondary Schools and for Colleges, no. vol. 2, Chelsea Publishing Company, 1964.

[4] G. H. Hardy, E. M. Wright, An Introduction to the Theory of Numbers, Oxford Mathematics, 2008.
[5] B. Hayes, On the teeth of wheels, Amer. Sci 88 (2000) 296–300.
[6] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics: A Foundation for Computer Science, 2nd Edition, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1994.

Algorithm 1: `FractionOfSmallestDenominatorInBetween(Proper fraction` f`, Proper fraction` g`)`

Input : $f = \frac{f_p}{f_q}$, $g = \frac{g_p}{g_q}$ such that $0 \leq f < g \leq 1$
Output: h the fraction of smallest denominator such that $f < h < g$

if $g_p f_q - g_q f_p = 1$ **then**
  $h \leftarrow \frac{f_p + g_p}{f_q + g_q}$
  return h
$i \leftarrow 0$
$a_i \leftarrow$ first partial quotient of f
$b_i \leftarrow$ first partial quotient of g
while $a_i = b_i$ *&* $r_i(f) \neq 0$ *&* $r_i(g) \neq 0$ **do**
  $i \leftarrow i + 1$
  compute a_i and b_i
if $a_i = b_i$ **then**
  if $r_i(f) = 0$ **then** f is an even principal convergent of g
    $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i, b_{i+1} + 1]$
  else $r_i(g) = 0$ and g is an odd principal convergent of f
    $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i, a_{i+1} + 1]$
else
  if $r_i(f) \neq 0$ *&* $r_i(g) \neq 0$ **then** case (v) of Theorem 1
    $h \leftarrow [a_0; a_1 \dots a_{i-1}, \min(a_i, b_i) + 1]$
  else
    if $r_i(f) = 0$ *&* $r_i(g) = 0$ **then** g is a convergent of f or conversely
      $h \leftarrow [a_0; a_1 \dots a_{i-1}, \min(a_i, b_i) + 1]$
    else
      if $r_i(f) = 0$ **then**
        if $a_i < b_i$ **then** $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i + 1]$ f is a convergent of g, i is even
        else
          if $a_i = b_i + 1$ **then** f is a convergent of g, i is odd
            if $b_{i+1} = 1$ **then** $h \leftarrow [a_0; a_1 \dots a_{i-1}, b_i, b_{i+1}, b_{i+2} + 1]$
            else $h \leftarrow [a_0; a_1 \dots a_{i-1}, b_i, 2]$
          else $a_i > b_i + 1$, case (v) of Theorem 1
            $h \leftarrow [a_0; a_1 \dots a_{i-1}, b_i + 1]$
      else $r_i(g) = 0$
        if $b_i < a_i$ **then** $h \leftarrow [a_0; a_1 \dots a_{i-1}, b_i + 1]$
        else
          if $b_i = a_i + 1$ **then**
            if $a_{i+1} = 1$ **then** $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i, a_{i+1}, a_{i+2} + 1]$
            else $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i, 2]$
          else $b_i > a_i + 1$
            $h \leftarrow [a_0; a_1 \dots a_{i-1}, a_i + 1]$
return h