Asymptotics of some functions arising in number theory and analysis of algorithms via computation and Mellin transforms

Richard P. Brent
Mathematical Sciences Institute, ANU
CARMA, University of Newcastle

incorporating joint work with
Michael Coons, Donald Knuth,
Brigitte Vallée and Wadim Zudilin

Presented at Workshop on Mathematics and Computation, Newcastle, 20 June 2015

Abstract

We consider the asymptotic behaviour of some interesting functions that arise naturally in

- analysis of algorithms (analysis of the average behaviour of the binary Euclidean algorithm) and in
- number theory (proving algebraic independence results using Mahler's method).

The asymptotic behaviour of these functions was first explored via computation and later explained via Mellin transforms.

Summary

The talk has two parts:

Summary

The talk has two parts:

- Part I Analysis of the binary Euclidean algorithm
(with contributions by Don Knuth and Brigitte Vallée)

Summary

The talk has two parts:

- Part I Analysis of the binary Euclidean algorithm
  (with contributions by Don Knuth and Brigitte Vallée)
- Part II Asymptotics of a Mahler function
  (with contributions by Michael Coons and Wadim Zudilin)

Summary

The talk has two parts:

- Part I Analysis of the binary Euclidean algorithm (with contributions by Don Knuth and Brigitte Vallée)
- Part II Asymptotics of a Mahler function (with contributions by Michael Coons and Wadim Zudilin)

At first sight the two parts seem unrelated, but by considering Mellin transforms we'll see that they are very similar.

Part I — Analysis of the binary Euclidean algorithm

The binary Euclidean algorithm is a variant of the classical Euclidean algorithm for finding greatest common divisors.

Part I — Analysis of the binary Euclidean algorithm

The binary Euclidean algorithm is a variant of the classical Euclidean algorithm for finding greatest common divisors.

It avoids divisions and multiplications, except by powers of two, so is potentially faster than the classical algorithm on a binary machine.

Part I — Analysis of the binary Euclidean algorithm

The binary Euclidean algorithm is a variant of the classical Euclidean algorithm for finding greatest common divisors.

It avoids divisions and multiplications, except by powers of two, so is potentially faster than the classical algorithm on a binary machine.

I will describe the binary algorithm and consider its average case behaviour. In particular, I will discuss some conjectures which were verified computationally in the 1970s and recently proved by Ian Morris (2014), extending earlier work by Gérard Maze (2005) and by Brigitte Vallée in the 1990s.

Part I — Analysis of the binary Euclidean algorithm

The binary Euclidean algorithm is a variant of the classical Euclidean algorithm for finding greatest common divisors.

It avoids divisions and multiplications, except by powers of two, so is potentially faster than the classical algorithm on a binary machine.

I will describe the binary algorithm and consider its average case behaviour. In particular, I will discuss some conjectures which were verified computationally in the 1970s and recently proved by Ian Morris (2014), extending earlier work by Gérard Maze (2005) and by Brigitte Vallée in the 1990s.

Analogous results for the classical algorithm were conjectured by Gauss (1800), and eventually proved by Kuz'min (1928), Lévy (1929) and Wirsing (1974).

Notation

$\lg(x)$ denotes $\log_2(x)$.

Notation

$\lg(x)$ denotes $\log_2(x)$.

$\mathrm{Val}_2(u)$ denotes the dyadic valuation of the positive integer u, i.e. the greatest integer j such that $2^j \mid u$.

The binary Euclidean algorithm

The idea of the *binary* Euclidean algorithm is to avoid the "division" operation $r \leftarrow u \bmod v$ of the classical algorithm, but retain $O(\log N)$ worst (and average) case for inputs $u, v \leq N$.

The binary Euclidean algorithm

The idea of the *binary* Euclidean algorithm is to avoid the "division" operation $r \leftarrow u \bmod v$ of the classical algorithm, but retain $O(\log N)$ worst (and average) case for inputs $u, v \le N$.

We assume that the algorithm is implemented on a binary computer so division by a power of two is easy. In particular, we assume that the "shift right until odd" operation

$$u \leftarrow u/2^{\mathrm{Val}_2(u)}$$

or equivalently

while even(u) do $u \leftarrow u/2$

can be performed in constant time, although time $O(\mathrm{Val}_2(u))$ would be sufficient.

Definition of the algorithm

It is easy to take account of the largest power of two dividing the inputs, so for simplicity we assume that u and v are *odd* positive integers.

Following is a simplified version of the algorithm given in Knuth, *The Art of Computer Programming*, §4.5.2.

Algorithm B

B1. $t \leftarrow |u - v|$;
if $t = 0$ return u;

B2. $t \leftarrow t/2^{\mathrm{Val}_2(t)}$;

B3. if $u \geq v$ then $u \leftarrow t$ else $v \leftarrow t$;
go to B1.

History

The binary Euclidean algorithm is often attributed to Silver and Terzian (unpublished, 1962) and Stein (1967). However, it seems to go back almost as far as the classical Euclidean algorithm. Knuth (§4.5.2) quotes a translation of a first-century AD Chinese text *Chiu Chang Suan Shu* on how to reduce a fraction to lowest terms:

History

The binary Euclidean algorithm is often attributed to Silver and Terzian (unpublished, 1962) and Stein (1967). However, it seems to go back almost as far as the classical Euclidean algorithm. Knuth (§4.5.2) quotes a translation of a first-century AD Chinese text *Chiu Chang Suan Shu* on how to reduce a fraction to lowest terms:

> *If halving is possible, take half.*
>
> *Otherwise write down the denominator and the numerator, and subtract the smaller from the greater.*
>
> *Repeat until both numbers are equal.*
>
> *Simplify with this common value.*

History

The binary Euclidean algorithm is often attributed to Silver and Terzian (unpublished, 1962) and Stein (1967). However, it seems to go back almost as far as the classical Euclidean algorithm. Knuth (§4.5.2) quotes a translation of a first-century AD Chinese text *Chiu Chang Suan Shu* on how to reduce a fraction to lowest terms:

> *If halving is possible, take half.*
>
> *Otherwise write down the denominator and the numerator, and subtract the smaller from the greater.*
>
> *Repeat until both numbers are equal.*
>
> *Simplify with this common value.*

This looks very much like Algorithm B !

The worst case

At step B1, u and v are odd, so $t = |u - v|$ is even. Thus, step B2 always reduces t by at least a factor of two. Using this fact, it is easy to show that step B3 is executed at most

$$\lfloor \lg(u + v) \rfloor$$

times. Thus, if $N = \max(u, v)$, step B3 is executed at most $\lg(N) + O(1)$ times.

The worst case

At step B1, u and v are odd, so $t = |u - v|$ is even. Thus, step B2 always reduces t by at least a factor of two. Using this fact, it is easy to show that step B3 is executed at most

$$\lfloor \lg(u + v) \rfloor$$

times. Thus, if $N = \max(u, v)$, step B3 is executed at most $\lg(N) + O(1)$ times.

If step B2 is replaced by single-bit shifts

$$\text{while even}(t) \text{ do } t \leftarrow t/2$$

the overall worst case time is still $O(\log N)$.

The worst case

At step B1, u and v are odd, so $t = |u - v|$ is even. Thus, step B2 always reduces t by at least a factor of two. Using this fact, it is easy to show that step B3 is executed at most

$$\lfloor \lg(u + v) \rfloor$$

times. Thus, if $N = \max(u, v)$, step B3 is executed at most $\lg(N) + O(1)$ times.

If step B2 is replaced by single-bit shifts

$$\text{while even}(t) \text{ do } t \leftarrow t/2$$

the overall worst case time is still $O(\log N)$.

Hint for proof: consider $\lg(uv)$.

Numerical example: $\gcd(123, 456)$

Binary

$$
\begin{aligned}
(123, 456) &: \ 456 \to 456/2^3 = 57 \\
(123, 57) &: \ 123 - 57 = 66 \to 66/2 = 33 \\
(57, 33) &: \ 57 - 33 = 24 \to 24/2^3 = 3 \\
(33, 3) &: \ 33 - 3 = 30 \to 30/2 = 15 \\
(15, 3) &: \ 15 - 3 = 12 \to 12/2^2 = 3 \\
(3, 3) &: \ 3 - 3 = 0 \implies \gcd = 3
\end{aligned}
$$

Numerical example: $\gcd(123, 456)$

Binary

$$\begin{aligned}
(123, 456) &: \ 456 \to 456/2^3 = 57 \\
(123, 57) &: \ 123 - 57 = 66 \to 66/2 = 33 \\
(57, 33) &: \ 57 - 33 = 24 \to 24/2^3 = 3 \\
(33, 3) &: \ 33 - 3 = 30 \to 30/2 = 15 \\
(15, 3) &: \ 15 - 3 = 12 \to 12/2^2 = 3 \\
(3, 3) &: \ 3 - 3 = 0 \implies \gcd = 3
\end{aligned}$$

Classical

$$\begin{aligned}
(123, 456) &: \ 456 \mod 123 = 87 \\
(123, 87) &: \ 123 \mod 87 = 36 \\
(87, 36) &: \ 87 \mod 36 = 15 \\
(36, 15) &: \ 36 \mod 15 = 6 \\
(15, 6) &: \ 15 \mod 6 = 3 \\
(6, 3) &: \ 6 \mod 3 = 0 \implies \gcd = 3
\end{aligned}$$

A continuous model

To analyse the expected behaviour of Algorithm B, we can follow what Gauss did for the classical algorithm, and construct a continuous model. This was first done in my 1976 paper, and made rigorous by Vallée (1998), Maze (2005) & Morris (2014).

A continuous model

To analyse the expected behaviour of Algorithm B, we can follow what Gauss did for the classical algorithm, and construct a continuous model. This was first done in my 1976 paper, and made rigorous by Vallée (1998), Maze (2005) & Morris (2014).

Assume that the initial inputs (u_0, v_0) to Algorithm B are uniformly and independently distributed in $(0, N)$, apart from the restriction that they are odd. Let (u_n, v_n) be the value of (u, v) after n iterations of step B3.

A continuous model

To analyse the expected behaviour of Algorithm B, we can follow what Gauss did for the classical algorithm, and construct a continuous model. This was first done in my 1976 paper, and made rigorous by Vallée (1998), Maze (2005) & Morris (2014).

Assume that the initial inputs (u_0, v_0) to Algorithm B are uniformly and independently distributed in $(0, N)$, apart from the restriction that they are odd. Let (u_n, v_n) be the value of (u, v) after n iterations of step B3.

Let

$$x_n = \frac{\min(u_n, v_n)}{\max(u_n, v_n)} ,$$

and let $F_n(x)$ be the probability distribution function of x_n (in the limit as $N \to \infty$). Thus $F_0(x) = x$ for $x \in [0, 1]$.

Plausible assumption

We make the plausible assumption that $\mathrm{Val}_2(t)$ takes the value k with probability 2^{-k} at step B2.

Plausible assumption

We make the plausible assumption that $\mathrm{Val}_2(t)$ takes the value k with probability 2^{-k} at step B2.

It is a plausible approximation because $\mathrm{Val}_2(t)$ at step B2 depends on the least significant bits of u and v, whereas the comparison at step B3 depends on the most significant bits, so one would expect the steps to be (almost) independent.

Plausible assumption

We make the plausible assumption that $\mathrm{Val}_2(t)$ takes the value k with probability 2^{-k} at step B2.

It is a plausible approximation because $\mathrm{Val}_2(t)$ at step B2 depends on the least significant bits of u and v, whereas the comparison at step B3 depends on the most significant bits, so one would expect the steps to be (almost) independent.

A rigorous justification has recently been given by Ian Morris, who shows that the assumption is correct in the limit as $N \to \infty$.

The recurrence for F_n

Consider the effect of steps B2 and B3. We can assume that $u > v$ so $t = u - v$. If $\mathrm{Val}_2(t) = k$ then $X = v/u$ is transformed to

$$X' = \min\left(\frac{u-v}{2^k v}, \frac{2^k v}{u-v}\right) = \min\left(\frac{1-X}{2^k X}, \frac{2^k X}{1-X}\right) .$$

The recurrence for F_n

Consider the effect of steps B2 and B3. We can assume that $u > v$ so $t = u - v$. If $\mathrm{Val}_2(t) = k$ then $X = v/u$ is transformed to

$$X' = \min\left(\frac{u-v}{2^k v}, \frac{2^k v}{u-v}\right) = \min\left(\frac{1-X}{2^k X}, \frac{2^k X}{1-X}\right) .$$

It follows that $X' < x$ iff

$$X < \frac{1}{1 + 2^k/x} \text{ or } X > \frac{1}{1 + 2^k x} .$$

The recurrence for F_n

Consider the effect of steps B2 and B3. We can assume that $u > v$ so $t = u - v$. If $\mathrm{Val}_2(t) = k$ then $X = v/u$ is transformed to

$$X' = \min\left(\frac{u-v}{2^k v}, \frac{2^k v}{u-v}\right) = \min\left(\frac{1-X}{2^k X}, \frac{2^k X}{1-X}\right) .$$

It follows that $X' < x$ iff

$$X < \frac{1}{1+2^k/x} \text{ or } X > \frac{1}{1+2^k x} .$$

Thus, the recurrence for $\widetilde{F}_n(x) = 1 - F_n(x)$ is

$$\widetilde{F}_{n+1}(x) = \sum_{k\geq 1} 2^{-k}\left(\widetilde{F}_n\left(\frac{1}{1+2^k/x}\right) - \widetilde{F}_n\left(\frac{1}{1+2^k x}\right)\right)$$

and $\widetilde{F}_0(x) = 1 - x$ for $x \in [0, 1]$.

The recurrence for f_n

Differentiating the recurrence for $\widetilde{F}_n$ we obtain a recurrence for the probability density $f_n(x) = F'_n(x) = -\widetilde{F}'_n(x)$:

$$\begin{aligned} f_{n+1}(x) &= \sum_{k\geq 1} \left(\frac{1}{x+2^k}\right)^2 f_n\left(\frac{x}{x+2^k}\right) \\ &+ \sum_{k\geq 1} \left(\frac{1}{1+2^k x}\right)^2 f_n\left(\frac{1}{1+2^k x}\right) . \end{aligned}$$

The recurrence for f_n

Differentiating the recurrence for $\widetilde{F}_n$ we obtain a recurrence for the probability density $f_n(x) = F_n'(x) = -\widetilde{F}_n'(x)$:

$$
\begin{aligned}
f_{n+1}(x) &= \sum_{k\geq 1} \left(\frac{1}{x+2^k}\right)^2 f_n\left(\frac{x}{x+2^k}\right) \\
&+ \sum_{k\geq 1} \left(\frac{1}{1+2^k x}\right)^2 f_n\left(\frac{1}{1+2^k x}\right) .
\end{aligned}
$$

This recurrence seems nicer than the one for $\widetilde{F}_n$ since the "weights" $(x+2^k)^{-2}$ and $(1+2^k x)^{-2}$ are positive.

The recurrence for f_n

Differentiating the recurrence for $\widetilde{F}_n$ we obtain a recurrence for the probability density $f_n(x) = F'_n(x) = -\widetilde{F}'_n(x)$:

$$\begin{aligned} f_{n+1}(x) &= \sum_{k\geq 1}\left(\frac{1}{x+2^k}\right)^2 f_n\left(\frac{x}{x+2^k}\right) \\ &+ \sum_{k\geq 1}\left(\frac{1}{1+2^k x}\right)^2 f_n\left(\frac{1}{1+2^k x}\right) . \end{aligned}$$

This recurrence seems nicer than the one for $\widetilde{F}_n$ since the "weights" $(x+2^k)^{-2}$ and $(1+2^k x)^{-2}$ are positive. On the other hand, $f_n(x)$ is unbounded on $(0, 1)$ (for $n \geq 1$), whereas $\widetilde{F}_n(x)$ is bounded on $[0, 1]$.

Conjectures (now proved)

In my 1976 paper I gave numerical and analytic evidence that $F_n(x)$ converges to a limiting distribution $F(x)$ as $n \to \infty$, and that $f_n(x)$ converges to the corresponding probability density $f(x) = F'(x)$.

Conjectures (now proved)

In my 1976 paper I gave numerical and analytic evidence that $F_n(x)$ converges to a limiting distribution $F(x)$ as $n \to \infty$, and that $f_n(x)$ converges to the corresponding probability density $f(x) = F'(x)$.

Assuming the existence of F, it is shown in my 1976 paper that the expected number of iterations of Algorithm B is $\sim K \lg N$ as $N \to \infty$, where $K = 0.705\ldots$ is a constant defined by

$$K = \ln 2 / E_\infty \ ,$$

and

$$E_\infty = \ln 2 + \int_0^1 \left(\sum_{k=2}^{\infty} \left(\frac{1 - 2^{-k}}{1 + (2^k - 1)x} \right) - \frac{1}{2(1+x)} \right) F(x)\, dx \ .$$

Conjectures (now proved)

In my 1976 paper I gave numerical and analytic evidence that $F_n(x)$ converges to a limiting distribution $F(x)$ as $n \to \infty$, and that $f_n(x)$ converges to the corresponding probability density $f(x) = F'(x)$.

Assuming the existence of F, it is shown in my 1976 paper that the expected number of iterations of Algorithm B is $\sim K \lg N$ as $N \to \infty$, where $K = 0.705\ldots$ is a constant defined by

$$K = \ln 2 / E_\infty \; ,$$

and

$$E_\infty = \ln 2 + \int_0^1 \left(\sum_{k=2}^{\infty} \left(\frac{1 - 2^{-k}}{1 + (2^k - 1)x} \right) - \frac{1}{2(1+x)} \right) F(x) \, dx \; .$$

These conjectures are now theorems, thanks to Morris (2014).

Simplifications

We can simplify the expression for K to obtain

$$K = 2/b \; ,$$

where

$$b = 2 - \int_0^1 \lg(1-x) f(x)\, dx \; .$$

Using integration by parts we obtain an equivalent expression

$$b = 2 + \frac{1}{\ln 2} \int_0^1 \frac{1 - F(x)}{1 - x}\, dx \; .$$

A discrepancy

In my 1976 paper I claimed that, for all $n \ge 0$ and $x \in (0,1]$,

$$F_n(x) = \alpha_n(x)\,\lg(x) + \beta_n(x)\ , \qquad (1)$$

where $\alpha_n(x)$ and $\beta_n(x)$ are analytic and regular in the disk $|x| < 1$. From (1) we can derive recurrence relations for the functions $\alpha_n(x)$ and $\beta_n(x)$, e.g.

$$2\alpha_{n+1}(2x) - \alpha_{n+1}(x) = \alpha_n\left(\frac{x}{1+x}\right) - 3f_n(1)x,$$

and similarly for $\beta_n(x)$.

A discrepancy

In my 1976 paper I claimed that, for all $n \geq 0$ and $x \in (0, 1]$,

$$F_n(x) = \alpha_n(x)\,\lg(x) + \beta_n(x)\;, \qquad (1)$$

where $\alpha_n(x)$ and $\beta_n(x)$ are analytic and regular in the disk $|x| < 1$. From (1) we can derive recurrence relations for the functions $\alpha_n(x)$ and $\beta_n(x)$, e.g.

$$2\alpha_{n+1}(2x) - \alpha_{n+1}(x) = \alpha_n\left(\frac{x}{1+x}\right) - 3f_n(1)x,$$

and similarly for $\beta_n(x)$. Using this method, Knuth (1997) found

$$K = 0.70597\ 12461\ 019\underline{45}\cdots$$

A discrepancy

In my 1976 paper I claimed that, for all $n \geq 0$ and $x \in (0, 1]$,

$$F_n(x) = \alpha_n(x)\,\lg(x) + \beta_n(x)\ , \qquad (1)$$

where $\alpha_n(x)$ and $\beta_n(x)$ are analytic and regular in the disk $|x| < 1$. From (1) we can derive recurrence relations for the functions $\alpha_n(x)$ and $\beta_n(x)$, e.g.

$$2\alpha_{n+1}(2x) - \alpha_{n+1}(x) = \alpha_n\left(\frac{x}{1+x}\right) - 3f_n(1)x,$$

and similarly for $\beta_n(x)$. Using this method, Knuth (1997) found

$$K = 0.70597\ 12461\ 019\underline{45}\cdots$$

but using a direct discretisation method I found

$$K = 0.70597\ 12461\ 019\underline{16}\cdots$$

A discrepancy

In my 1976 paper I claimed that, for all $n \geq 0$ and $x \in (0, 1]$,

$$F_n(x) = \alpha_n(x)\,\lg(x) + \beta_n(x)\ , \qquad (1)$$

where $\alpha_n(x)$ and $\beta_n(x)$ are analytic and regular in the disk $|x| < 1$. From (1) we can derive recurrence relations for the functions $\alpha_n(x)$ and $\beta_n(x)$, e.g.

$$2\alpha_{n+1}(2x) - \alpha_{n+1}(x) = \alpha_n\left(\frac{x}{1+x}\right) - 3f_n(1)x,$$

and similarly for $\beta_n(x)$. Using this method, Knuth (1997) found

$$K = 0.70597\ 12461\ 019\underline{45}\cdots$$

but using a direct discretisation method I found

$$K = 0.70597\ 12461\ 019\underline{16}\cdots$$

Why the discrepancy?

Some detective work

After a flurry of emails we (Brent and Knuth) tracked down the error. It was found independently, and at the same time, by Flajolet and Vallée, who were in email contact with us.

Some detective work

After a flurry of emails we (Brent and Knuth) tracked down the error. It was found independently, and at the same time, by Flajolet and Vallée, who were in email contact with us.
(Knuth was in a hurry to finalise the third edition of volume 2 of *The Art of Computer Programming*.)

Some detective work

After a flurry of emails we (Brent and Knuth) tracked down the error. It was found independently, and at the same time, by Flajolet and Vallée, who were in email contact with us. (Knuth was in a hurry to finalise the third edition of volume 2 of *The Art of Computer Programming*.)

We found that eqn. (1): $F_n(x) = \alpha_n(x)\,\mathrm{lg}(x) + \beta_n(x)$ is incorrect for $n \geq 1$. A small oscillatory term, not expressible in this form with $\alpha_n(x), \beta_n(x)$ regular in the disk $|x| < 1$, is missing!

Some detective work

After a flurry of emails we (Brent and Knuth) tracked down the error. It was found independently, and at the same time, by Flajolet and Vallée, who were in email contact with us. (Knuth was in a hurry to finalise the third edition of volume 2 of *The Art of Computer Programming*.)

We found that eqn. (1): $F_n(x) = \alpha_n(x)\,\mathrm{lg}(x) + \beta_n(x)$ is incorrect for $n \geq 1$. A small oscillatory term, not expressible in this form with $\alpha_n(x), \beta_n(x)$ regular in the disk $|x| < 1$, is missing!

To explain this, we need to consider Mellin transforms.

Mellin transforms

The *Mellin transform* of a function $g(x)$ is defined by

$$g^*(s) = \int_0^\infty g(x) x^{s-1} dx \; .$$

Mellin transforms

The *Mellin transform* of a function $g(x)$ is defined by

$$g^*(s) = \int_0^\infty g(x)x^{s-1}dx\ .$$

It is easy to see that, if

$$h(x) = \sum_{k\ge1} 2^{-k} g(2^k x)\ ,$$

then the Mellin transform of h is

$$h^*(s) = \sum_{k\ge1} 2^{-k(s+1)} g^*(s) = \frac{g^*(s)}{2^{s+1}-1}\ .$$

Mellin inversion

Under suitable conditions we can apply the Mellin inversion formula to obtain

$$h(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} h^*(s) x^{-s} ds \ ,$$

where c is a real constant lying in a certain interval.

Mellin inversion

Under suitable conditions we can apply the Mellin inversion formula to obtain

$$h(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} h^*(s)x^{-s}ds\ ,$$

where c is a real constant lying in a certain interval.
Applying these results to $g(x) = 1/(1+x)$, whose Mellin transform is $g^*(s) = \pi/\sin\pi s$ for $0 < \Re s < 1$, we find

$$h(x) = \sum_{k\geq 1}\frac{2^{-k}}{1+2^k x}$$

as a sum of residues of

$$\left(\frac{\pi}{\sin\pi s}\right)\frac{x^{-s}}{2^{s+1}-1}$$

in the left half-plane $\Re s \leq 0$.

Application of Mellin inversion

This gives

$$h(x) = xP(\lg x) + x\lg x + 1 + \frac{x}{2} - \frac{2}{1}x^2 + \frac{4}{3}x^3 - \frac{8}{7}x^4 + \cdots ,$$

where

$$P(t) = \frac{2\pi}{\ln 2}\sum_{n=1}^{\infty}\frac{\sin 2n\pi t}{\sinh(2n\pi^2/\ln 2)}$$

comes from the poles of $1/(2^{s+1}-1)$ at

$$s = -1 \pm \frac{2\pi i n}{\ln 2},\ \ n \in \{1, 2, 3, \ldots\}.$$

The "wobbles" caused by $P(t)$

Because the residues at the non-real poles are tiny, thanks to the $\sinh$ term in the denominator, $P(t)$ is a very small periodic function:

$$|P(t)| < 7.8 \times 10^{-12}$$

for real t.

The "wobbles" caused by $P(t)$

Because the residues at the non-real poles are tiny, thanks to the $\sinh$ term in the denominator, $P(t)$ is a very small periodic function:

$$|P(t)| < 7.8 \times 10^{-12}$$

for real t.

Thus, numerical computations performed using single-precision (36-bit) floating-point arithmetic did not reveal the error incurred by omitting the term involving $P(t)$.

Application to $F_1(x)$

Using the results obtained by Mellin transforms, we find that

$$F_1(x) = 1 + h(1/x) - h(x) = -xP(\lg x) + \alpha_1(x)\lg(x) + \beta_1(x),$$

where

$$\alpha_1(x) = -x,$$

$$\beta_1(x) = \frac{x(5x-1)}{6(1+x)} + \frac{3}{2}\sum_{j=2}^{\infty} \frac{(-2x)^j}{(2^{j-1}-1)(2^{j+1}-1)}.$$

Application to $F_1(x)$

Using the results obtained by Mellin transforms, we find that

$$F_1(x) = 1 + h(1/x) - h(x) = -xP(\lg x) + \alpha_1(x)\lg(x) + \beta_1(x),$$

where

$$\alpha_1(x) = -x,$$

$$\beta_1(x) = \frac{x(5x-1)}{6(1+x)} + \frac{3}{2}\sum_{j=2}^{\infty}\frac{(-2x)^j}{(2^{j-1}-1)(2^{j+1}-1)}.$$

Observe that $\alpha_1(x)$ and $\beta_1(x)$ are regular in $|x| < 1$, but the term $xP(\lg x)$ is not. Thus, the assumption underlying Knuth's evaluation of K is incorrect, although the numerical effect is small since the function $P(t)$ is so small.

A conjecture of Vallée

Let $\lambda = f(1)$, where $f = F'$ is the limiting probability density as above. Brigitte Vallée (1997/8) conjectured that

$$\frac{\lambda}{b} = \frac{2\ln 2}{\pi^2} \, ,$$

or equivalently that

$$K\lambda = \frac{4\ln 2}{\pi^2} \, . \tag{2}$$

Numerical results

Using an improvement of the discretisation method of my 1976 paper, and the equivalent of more than fifty decimal places working precision, I computed the limiting probability density f, then K, $\lambda = f(1)$, and $K\lambda$. The results were

$$\begin{aligned} K &= 0.7059712461\ 0191639152\ 9314135852\ 8817666677 \\ \lambda &= 0.3979226811\ 8831664407\ 6707161142\ 6549823098 \\ K\lambda &= 0.2809219710\ 9073150563\ 5754397987\ 9880385315 \end{aligned}$$

These are believed to be correctly rounded values.

Numerical results

Using an improvement of the discretisation method of my 1976 paper, and the equivalent of more than fifty decimal places working precision, I computed the limiting probability density f, then K, $\lambda = f(1)$, and $K\lambda$. The results were

$$\begin{aligned} K &= 0.7059712461\ 0191639152\ 9314135852\ 8817666677 \\ \lambda &= 0.3979226811\ 8831664407\ 6707161142\ 6549823098 \\ K\lambda &= 0.2809219710\ 9073150563\ 5754397987\ 9880385315 \end{aligned}$$

These are believed to be correctly rounded values.

Vallée's conjecture (2) is that

$$K\lambda = 4\ln 2/\pi^2 \ .$$

The computed value of $K\lambda$ agrees with $4\ln 2/\pi^2$ to the 40 decimals shown (in fact to 44 decimals).

Proofs

Vallée proved her conjecture under the assumption of a spectral condition.

Proofs

Vallée proved her conjecture under the assumption of a spectral condition.

The conjecture was proved rigorously by Ian Morris (2014) without assuming any spectral condition.

Proofs

Vallée proved her conjecture under the assumption of a spectral condition.

The conjecture was proved rigorously by Ian Morris (2014) without assuming any spectral condition.

The proofs depend on some rather sophisticated functional analysis, e.g. the theory of Hardy spaces and Ruelle operators, and are too long to give here – if you are interested, see the original papers by Vallée, Maze and Morris.

Part II – Asymptotics of a Mahler function

One of the first significant contributions of Mahler is an approach, now called "Mahler's method", yielding transcendence and algebraic independence results for the values at algebraic points of a large family of power series satisfying functional equations of a certain type. In the seminal paper [9][1] Mahler established that the Fredholm series $f(z) = \sum_{k\geq 0} z^{2^k}$*, which satisfies* $f(z^2) = f(z) - z$*, takes transcendental values at any nonzero algebraic point in the open unit disk.*

[1] K. Mahler, *Math. Ann.* 101 (1929), 342–366.

Part II – Asymptotics of a Mahler function

> *One of the first significant contributions of Mahler is an approach, now called "Mahler's method", yielding transcendence and algebraic independence results for the values at algebraic points of a large family of power series satisfying functional equations of a certain type. In the seminal paper [9][1] Mahler established that the Fredholm series $f(z) = \sum_{k \ge 0} z^{2^k}$, which satisfies $f(z^2) = f(z) - z$, takes transcendental values at any nonzero algebraic point in the open unit disk.*
>
> *J. Borwein, Y. Bugeaud and M. Coons*
> *The legacy of Kurt Mahler*
> *AustMS Gazette, March 2014, pg. 16.*

[1] K. Mahler, *Math. Ann.* 101 (1929), 342–366.

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}).$$

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}).$$

$F(z)$ is related to the Stern sequence.

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4 F(z^{16}).$$

$F(z)$ is related to the Stern sequence.
NB: F here is unrelated to the F of Part I.

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}).$$

$F(z)$ is related to the Stern sequence.
NB: F here is unrelated to the F of Part I.

We consider the asymptotic behaviour of $F(z)$ as $z \to 1^-$. This has applications to algebraic independence results.

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4 F(z^{16}).$$

$F(z)$ is related to the Stern sequence.
NB: F here is unrelated to the F of Part I.

We consider the asymptotic behaviour of $F(z)$ as $z \to 1^-$. This has applications to algebraic independence results.

Specifically, BCZ (2015) proved that the functions $F(z), F(z^4), F'(z)$, and $F'(z^4)$ are algebraically independent over $\mathbb{C}(z)$;

The function $F(z)$

Dilcher and Stolarsky [Acta Arithmetica, 2009] introduced a Mahler function $F(z) = 1 + z + \cdots$ satisfying the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4 F(z^{16}).$$

$F(z)$ is related to the Stern sequence.
NB: F here is unrelated to the F of Part I.

We consider the asymptotic behaviour of $F(z)$ as $z \to 1^-$. This has applications to algebraic independence results.

Specifically, BCZ (2015) proved that the functions $F(z), F(z^4), F'(z)$, and $F'(z^4)$ are algebraically independent over $\mathbb{C}(z)$; it follows (thanks to a result of Kumiko Nishioka) that $F(\alpha), F(\alpha^4), F'(\alpha)$, and $F'(\alpha^4)$ are independent over $\mathbb{Q}$ for any nonzero algebraic number α in the unit disk.

The Stern sequence

Stern's diatomic sequence (or Stern-Brocot sequence) is defined by

$$\begin{aligned} a_0 &= 0, \\ a_1 &= 1, \\ a_{2n} &= a_n \text{ for } n > 0, \\ a_{2n+1} &= a_n + a_{n+1} \text{ for } n > 0. \end{aligned}$$

The Stern sequence

Stern's diatomic sequence (or Stern-Brocot sequence) is defined by

$$\begin{aligned} a_0 &= 0, \\ a_1 &= 1, \\ a_{2n} &= a_n \text{ for } n > 0, \\ a_{2n+1} &= a_n + a_{n+1} \text{ for } n > 0. \end{aligned}$$

This sequence has many interesting properties (see the OEIS entry A002487). For example, a_n/a_{n+1} runs through all the reduced nonnegative rationals exactly once.

Some properties of $F(z)$

Dilcher and Stolarsky (2009) defined $F(z)$ using a polynomial analogue of the Stern sequence, and deduced the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}). \tag{3}$$

Some properties of $F(z)$

Dilcher and Stolarsky (2009) defined $F(z)$ using a polynomial analogue of the Stern sequence, and deduced the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}). \tag{3}$$

However, for our purposes it is simpler to *define* $F(z)$ by the recurrence (3) and the auxiliary condition $F(z) = 1 + O(z)$ as $z \to 0$.

Some properties of $F(z)$

Dilcher and Stolarsky (2009) defined $F(z)$ using a polynomial analogue of the Stern sequence, and deduced the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}). \qquad (3)$$

However, for our purposes it is simpler to *define* $F(z)$ by the recurrence (3) and the auxiliary condition $F(z) = 1 + O(z)$ as $z \to 0$.

Using Mahler's method, Adamczewski (2010) proved that $F(\alpha)$ is transcendental for every algebraic α, $0 < |\alpha| < 1$.

Some properties of $F(z)$

Dilcher and Stolarsky (2009) defined $F(z)$ using a polynomial analogue of the Stern sequence, and deduced the recurrence

$$F(z) = (1 + z + z^2)F(z^4) - z^4F(z^{16}). \quad (3)$$

However, for our purposes it is simpler to *define* $F(z)$ by the recurrence (3) and the auxiliary condition $F(z) = 1 + O(z)$ as $z \to 0$.

Using Mahler's method, Adamczewski (2010) proved that $F(\alpha)$ is transcendental for every algebraic α, $0 < |\alpha| < 1$.

Independently, Michael Coons (2010) proved that $F(z)$ is a transcendental function, along with results on transcendence at algebraic arguments.

The auxiliary function $\mu(z)$

We are interested in the behaviour of $F(z)$ for $z \in [0, 1)$, and in particular the asymptotic behaviour of $F(z)$ as $z \to 1^-$.

The auxiliary function $\mu(z)$

We are interested in the behaviour of $F(z)$ for $z \in [0, 1)$, and in particular the asymptotic behaviour of $F(z)$ as $z \to 1^-$.

It is useful to define an auxiliary function $\mu : [0, 1) \mapsto \mathbb{R}$ by

$$\mu(z) = \frac{F(z)}{F(z^4)}\,. \tag{4}$$

The auxiliary function $\mu(z)$

We are interested in the behaviour of $F(z)$ for $z \in [0,1)$, and in particular the asymptotic behaviour of $F(z)$ as $z \to 1^-$.

It is useful to define an auxiliary function $\mu : [0,1) \mapsto \mathbb{R}$ by

$$\mu(z) = \frac{F(z)}{F(z^4)} \,. \qquad (4)$$

From the recurrence for $F(z)$ and (4), $\mu(z)$ satisfies the recurrence

$$\mu(z) = 1 + z + z^2 - \frac{z^4}{\mu(z^4)} \,. \qquad (5)$$

The auxiliary function $\mu(z)$

We are interested in the behaviour of $F(z)$ for $z \in [0, 1)$, and in particular the asymptotic behaviour of $F(z)$ as $z \to 1^-$.

It is useful to define an auxiliary function $\mu : [0, 1) \mapsto \mathbb{R}$ by

$$\mu(z) = \frac{F(z)}{F(z^4)} \,. \tag{4}$$

From the recurrence for $F(z)$ and (4), $\mu(z)$ satisfies the recurrence

$$\mu(z) = 1 + z + z^2 - \frac{z^4}{\mu(z^4)} \,. \tag{5}$$

Our strategy is to analyse the asymptotic behaviour of $\mu(z)$ and then deduce the corresponding behaviour of $F(z)$.

$\mu(z)$ as a continued fraction

Observe that $\mu(z)$ may be written as a continued fraction

$$\begin{aligned}\mu(z) &= (1+z+z^2) - z^4/\mu(z^4)\\ &= (1+z+z^2) - \frac{z^4}{(1+z^4+z^{2\cdot 4}) - z^{4^2}/\mu(z^{4^2})} = \cdots\end{aligned}$$

$\mu(z)$ as a continued fraction

Observe that $\mu(z)$ may be written as a continued fraction

$$\begin{aligned} \mu(z) &= (1 + z + z^2) - z^4/\mu(z^4) \\ &= (1 + z + z^2) - \frac{z^4}{(1 + z^4 + z^{2\cdot 4}) - z^{4^2}/\mu(z^{4^2})} = \cdots \end{aligned}$$

Since $\mu(z) = F(z)/F(z^4)$, we have an explicit expression for $F(z)$ as an infinite product:

$$F(z) = \prod_{k=0}^{\infty} \mu\left(z^{4^k}\right). \qquad (6)$$

$\mu(z)$ as a continued fraction

Observe that $\mu(z)$ may be written as a continued fraction

$$\begin{aligned}\mu(z) &= (1 + z + z^2) - z^4/\mu(z^4)\\ &= (1 + z + z^2) - \frac{z^4}{(1 + z^4 + z^{2\cdot 4}) - z^{4^2}/\mu(z^{4^2})} = \cdots\end{aligned}$$

Since $\mu(z) = F(z)/F(z^4)$, we have an explicit expression for $F(z)$ as an infinite product:

$$F(z) = \prod_{k=0}^{\infty} \mu\left(z^{4^k}\right). \tag{6}$$

In this sense we have an explicit solution for $F(z)$ as an infinite product of continued fractions.

Some properties of $F(z)$ as an analytic function

Lemma

The Maclaurin series

$$F(z) = \sum_{n=0}^{\infty} f_n z^n$$

has coefficients $f_n \in \{0, 1\}$. *Also,* $F(z)$ *is strictly monotonic increasing and unbounded for* $z \in [0, 1)$, *and can not be analytically continued past the unit circle.*

Some properties of $F(z)$ as an analytic function

Lemma

The Maclaurin series

$$F(z) = \sum_{n=0}^{\infty} f_n z^n$$

has coefficients $f_n \in \{0, 1\}$*. Also,* $F(z)$ *is strictly monotonic increasing and unbounded for* $z \in [0, 1)$*, and can not be analytically continued past the unit circle.*

From the functional equation for $F(z)$ it follows that $F(z)$ has a singularity at $z = \exp(2\pi i/2^k)$ for all non-negative integers k.

Some properties of $F(z)$ as an analytic function

Lemma

The Maclaurin series

$$F(z) = \sum_{n=0}^{\infty} f_n z^n$$

has coefficients $f_n \in \{0, 1\}$*. Also,* $F(z)$ *is strictly monotonic increasing and unbounded for* $z \in [0, 1)$*, and can not be analytically continued past the unit circle.*

From the functional equation for $F(z)$ it follows that $F(z)$ has a singularity at $z = \exp(2\pi i/2^k)$ for all non-negative integers k. Thus, there is a dense set of singularities on the unit circle, which is a natural boundary.

Properties of $\mu(z)$

Lemma

If $\mu_1 := \lim_{x\to 1-} \mu(x)$ *and* $\mu_1' := \lim_{x\to 1-} \mu'(x)$*, then*

$$\mu_1 = \frac{3+\sqrt{5}}{2} = \rho^2 \approx 2.618 \tag{7}$$

and

$$\mu_1' = \frac{21+8\sqrt{5}}{11} \approx 3.535. \tag{8}$$

Properties of $\mu(z)$

Lemma

If $\mu_1 := \lim_{x\to 1-} \mu(x)$ *and* $\mu_1' := \lim_{x\to 1-} \mu'(x)$*, then*

$$\mu_1 = \frac{3+\sqrt{5}}{2} = \rho^2 \approx 2.618 \tag{7}$$

and

$$\mu_1' = \frac{21+8\sqrt{5}}{11} \approx 3.535. \tag{8}$$

Sketch of proof.

Let $Q(x)$ be the larger root of $Q(x) = 1 + x + x^2 - x^4/Q(x)$.
Show that $\mu(x) < Q(x)$ for all $x \in (0, 1)$.

Properties of $\mu(z)$

Lemma

If $\mu_1 := \lim_{x \to 1-} \mu(x)$ *and* $\mu_1' := \lim_{x \to 1-} \mu'(x)$, *then*

$$\mu_1 = \frac{3 + \sqrt{5}}{2} = \rho^2 \approx 2.618 \tag{7}$$

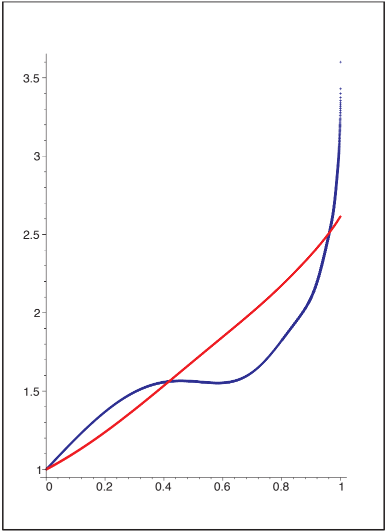
and

$$\mu_1' = \frac{21 + 8\sqrt{5}}{11} \approx 3.535. \tag{8}$$

Sketch of proof.

Let $Q(x)$ be the larger root of $Q(x) = 1 + x + x^2 - x^4/Q(x)$. Show that $\mu(x) < Q(x)$ for all $x \in (0, 1)$. Hint – use induction on $x = x_0^{4^{-n}}$, where x_0 is sufficiently small. □

$\mu(x)$ and $\mu'(x)$ for $x \in [0, 1)$



What can we say about $\mu''(x)$?

It appears from the graph of $\mu'(x)$ that $\mu''(x)$ is unbounded as $x \to 1^-$, and this is indeed true. We have the following result, where the constant $2\lg(\rho)$ is best possible.[2]

[2] As in Part I, We write $\lg(x)$ for $\log_2(x)$.

What can we say about $\mu''(x)$?

It appears from the graph of $\mu'(x)$ that $\mu''(x)$ is unbounded as $x \to 1^-$, and this is indeed true. We have the following result, where the constant $2\lg(\rho)$ is best possible.[2]

Lemma

Let $\alpha \le 2\lg(\rho) \approx 1.388$. Then, for $t \in (0,1)$ we have

$$\mu''(e^{-t}) = O(t^{\alpha-2}) \tag{9}$$

[2] As in Part I, We write $\lg(x)$ for $\log_2(x)$.

What can we say about $\mu''(x)$?

It appears from the graph of $\mu'(x)$ that $\mu''(x)$ is unbounded as $x \to 1^-$, and this is indeed true. We have the following result, where the constant $2\lg(\rho)$ is best possible.[2]

Lemma

Let $\alpha \le 2\lg(\rho) \approx 1.388$. Then, for $t \in (0,1)$ we have

$$\mu''(e^{-t}) = O(t^{\alpha-2}) \tag{9}$$

and

$$\mu(e^{-t}) = \mu_1 - t\mu_1' + O(t^{\alpha}). \tag{10}$$

[2] As in Part I, We write $\lg(x)$ for $\log_2(x)$.

Why the exponent $\alpha \approx 1.388$?

Differentiating the recurrence for $\mu(z)$ twice, we obtain

$$\mu''(e^{-t}) = A(t) + B(t)\mu''(e^{-4t}),$$

where $A(t)$ is bounded, and

$$B(t) = 16e^{-10t}/\mu(e^{-4t})^2 = 16/\mu_1^2 + O(t).$$

Why the exponent $\alpha \approx 1.388$?

Differentiating the recurrence for $\mu(z)$ twice, we obtain

$$\mu''(e^{-t}) = A(t) + B(t)\mu''(e^{-4t}),$$

where $A(t)$ is bounded, and

$$B(t) = 16e^{-10t}/\mu(e^{-4t})^2 = 16/\mu_1^2 + O(t).$$

The exponent α is chosen so that $16/\mu_1^2 \le 4^{2-\alpha}$, since this inequality is necessary (and sufficient) for the inductive proof to go through.

Why the exponent $\alpha \approx 1.388$?

Differentiating the recurrence for $\mu(z)$ twice, we obtain

$$\mu''(e^{-t}) = A(t) + B(t)\mu''(e^{-4t}),$$

where $A(t)$ is bounded, and

$$B(t) = 16e^{-10t}/\mu(e^{-4t})^2 = 16/\mu_1^2 + O(t).$$

The exponent α is chosen so that $16/\mu_1^2 \le 4^{2-\alpha}$, since this inequality is necessary (and sufficient) for the inductive proof to go through.

Since $\mu_1 = \rho^2$, we have to choose $\alpha \le 2\lg(\rho) \approx 1.388$.

Mellin transforms

Our strategy is to deduce the asymptotic behaviour of $\mu(z)$ and $F(z)$ as $z \to 1^-$ from certain Mellin transforms.

Mellin transforms

Our strategy is to deduce the asymptotic behaviour of $\mu(z)$ and $F(z)$ as $z \to 1^-$ from certain Mellin transforms.

Specifically, define

$$\mathcal{F}(s) := \int_0^\infty \ln(F(e^{-t}))\, t^{s-1}\, \mathrm{d}t$$

and

$$\mathcal{M}(s) := \int_0^\infty \ln(\mu(e^{-t}))\, t^{s-1}\, \mathrm{d}t.$$

Mellin transforms

Our strategy is to deduce the asymptotic behaviour of $\mu(z)$ and $F(z)$ as $z \to 1^-$ from certain Mellin transforms.

Specifically, define

$$\mathcal{F}(s) := \int_0^\infty \ln(F(e^{-t}))\, t^{s-1}\, \mathrm{d}t$$

and

$$\mathcal{M}(s) := \int_0^\infty \ln(\mu(e^{-t}))\, t^{s-1}\, \mathrm{d}t.$$

The integrals converge in the half-plane $\Re(s) > 0$.

Mellin transforms

Our strategy is to deduce the asymptotic behaviour of $\mu(z)$ and $F(z)$ as $z \to 1^-$ from certain Mellin transforms.

Specifically, define

$$\mathcal{F}(s) := \int_0^\infty \ln(F(e^{-t}))\, t^{s-1}\, \mathrm{d}t$$

and

$$\mathcal{M}(s) := \int_0^\infty \ln(\mu(e^{-t}))\, t^{s-1}\, \mathrm{d}t.$$

The integrals converge in the half-plane $\Re(s) > 0$. For $\Re(s) \le 0$ we define $\mathcal{F}(s)$ and $\mathcal{M}(s)$ by analytic continuation (if possible).

Properties of the Mellin transforms

Since

$$\ln \mu(e^{-t}) = \ln F(e^{-t}) - \ln F(e^{-4t}),$$

we see that

$$\mathcal{M}(s) = (1 - 4^{-s})\mathcal{F}(s).$$

Properties of the Mellin transforms

Since

$$\ln \mu(e^{-t}) = \ln F(e^{-t}) - \ln F(e^{-4t}),$$

we see that

$$\mathcal{M}(s) = (1 - 4^{-s})\mathcal{F}(s).$$

We can deduce the behaviour of $\ln F(e^{-t})$ for small positive t from knowledge of the singularities of $\mathcal{F}(s)$.

Properties of the Mellin transforms

Since

$$\ln \mu(e^{-t}) = \ln F(e^{-t}) - \ln F(e^{-4t}),$$

we see that

$$\mathcal{M}(s) = (1 - 4^{-s})\mathcal{F}(s).$$

We can deduce the behaviour of $\ln F(e^{-t})$ for small positive t from knowledge of the singularities of $\mathcal{F}(s)$.

Since $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$, it is sufficient to determine the singularities of $\mathcal{M}(s)$ and (easy) those of $(1 - 4^{-s})^{-1}$.

Properties of the Mellin transforms

Since

$$\ln \mu(e^{-t}) = \ln F(e^{-t}) - \ln F(e^{-4t}),$$

we see that

$$\mathcal{M}(s) = (1 - 4^{-s})\mathcal{F}(s).$$

We can deduce the behaviour of $\ln F(e^{-t})$ for small positive t from knowledge of the singularities of $\mathcal{F}(s)$.

Since $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$, it is sufficient to determine the singularities of $\mathcal{M}(s)$ and (easy) those of $(1 - 4^{-s})^{-1}$.

First we use the Lemmas above to extend the domain of definition of $\mathcal{M}(s)$ into the left half-plane.

Analytic continuation of $\mathcal{M}(s)$

Define

$$\widetilde{\mu}(t) := \ln(\mu(e^{-t})) - \ln(\mu_1)e^{-\lambda t},$$

where

$$\lambda := \frac{\mu_1'}{\mu_1 \ln \mu_1} \approx 1.403\,.$$

Analytic continuation of $\mathcal{M}(s)$

Define

$$\widetilde{\mu}(t) := \ln(\mu(e^{-t})) - \ln(\mu_1)e^{-\lambda t},$$

where

$$\lambda := \frac{\mu_1'}{\mu_1 \ln \mu_1} \approx 1.403\,.$$

Since $\lambda \geq 1$, $\widetilde{\mu}(t) = O(e^{-t})$ as $t \to +\infty$.

Analytic continuation of $\mathcal{M}(s)$

Define

$$\widetilde{\mu}(t) := \ln(\mu(e^{-t})) - \ln(\mu_1)e^{-\lambda t},$$

where

$$\lambda := \frac{\mu_1'}{\mu_1 \ln \mu_1} \approx 1.403\,.$$

Since $\lambda \geq 1$, $\widetilde{\mu}(t) = O(e^{-t})$ as $t \to +\infty$.
Also, from the Lemmas above, as $t \to 0^+$ we have

$$\widetilde{\mu}(t) = (\lambda \ln \mu_1 - \mu_1'/\mu_1)t + O(t^\alpha).$$

Analytic continuation of $\mathcal{M}(s)$

Define

$$\widetilde{\mu}(t) := \ln(\mu(e^{-t})) - \ln(\mu_1)e^{-\lambda t},$$

where

$$\lambda := \frac{\mu_1'}{\mu_1 \ln \mu_1} \approx 1.403\,.$$

Since $\lambda \geq 1$, $\widetilde{\mu}(t) = O(e^{-t})$ as $t \to +\infty$.
Also, from the Lemmas above, as $t \to 0^+$ we have

$$\widetilde{\mu}(t) = (\lambda \ln \mu_1 - \mu_1'/\mu_1)t + O(t^\alpha).$$

Our choice of λ makes the coefficient of t vanish, so $\widetilde{\mu}(t) = O(t^\alpha)$.

Analytic continuation of $\mathcal{M}(s)$

Let

$$\widetilde{\mathcal{M}}(s) := \int_0^\infty \widetilde{\mu}(t) t^{s-1} \, \mathrm{d}t.$$

Since $\widetilde{\mu}(t) = O(t^\alpha)$, the integral converges for $\Re(s) > -\alpha$.

Analytic continuation of $\mathcal{M}(s)$

Let

$$\widetilde{\mathcal{M}}(s) := \int_0^\infty \widetilde{\mu}(t) t^{s-1}\,\mathrm{d}t.$$

Since $\widetilde{\mu}(t) = O(t^\alpha)$, the integral converges for $\Re(s) > -\alpha$. Now

$$\mathcal{M}(s) = \widetilde{\mathcal{M}}(s) + \ln(\mu_1)\lambda^{-s}\Gamma(s)$$

gives the analytic continuation of $\mathcal{M}(s)$ into the half-plane

$$\mathcal{H} := \{s \in \mathbb{C} : \Re(s) > -2\lg(\rho)\}.$$

Analytic continuation of $\mathcal{M}(s)$

Let

$$\widetilde{\mathcal{M}}(s) := \int_0^\infty \widetilde{\mu}(t) t^{s-1}\,\mathrm{d}t.$$

Since $\widetilde{\mu}(t) = O(t^\alpha)$, the integral converges for $\Re(s) > -\alpha$. Now

$$\mathcal{M}(s) = \widetilde{\mathcal{M}}(s) + \ln(\mu_1)\lambda^{-s}\Gamma(s)$$

gives the analytic continuation of $\mathcal{M}(s)$ into the half-plane

$$\mathcal{H} := \{s \in \mathbb{C} : \Re(s) > -2\lg(\rho)\}.$$

In $\mathcal{H}$, the only singularities of $\mathcal{M}(s)$ occur at the singularities of $\Gamma(s)$, i.e. at $s \in \{0, -1\}$.

Singularities of $\mathcal{F}(s)$ in $\mathcal{H}$

The Mellin transform $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ has three types of singularities in $\mathcal{H}$.

Singularities of $\mathcal{F}(s)$ in $\mathcal{H}$

The Mellin transform $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ has three types of singularities in $\mathcal{H}$.

(a) A double pole at $s = 0$, since $\Gamma(s)$ has a pole there, and the denominator $1 - 4^{-s}$ vanishes at $s = 0$.

Singularities of $\mathcal{F}(s)$ in $\mathcal{H}$

The Mellin transform $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ has three types of singularities in $\mathcal{H}$.

(a) A double pole at $s = 0$, since $\Gamma(s)$ has a pole there, and the denominator $1 - 4^{-s}$ vanishes at $s = 0$.

(b) Poles at $s = ik\pi/\ln(2)$ for $k \in \mathbb{Z}\backslash\{0\}$, since the denominator $1 - 4^{-s}$ vanishes at these points.

Singularities of $\mathcal{F}(s)$ in $\mathcal{H}$

The Mellin transform $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ has three types of singularities in $\mathcal{H}$.

(a) A double pole at $s = 0$, since $\Gamma(s)$ has a pole there, and the denominator $1 - 4^{-s}$ vanishes at $s = 0$.

(b) Poles at $s = ik\pi/\ln(2)$ for $k \in \mathbb{Z}\backslash\{0\}$, since the denominator $1 - 4^{-s}$ vanishes at these points.

(c) A pole at $s = -1$, since $\Gamma(s)$ has a pole there.

Asymptotics of $\ln F(e^{-t})$

Theorem

For arbitrary $\varepsilon > 0$ and small positive t,

$$\ln F(e^{-t}) = -\lg(\rho)\ln(t) + c_0 + \sum_{k=1}^{\infty} a_k(t) + c_1 t + O(t^{2\lg(\rho)-\varepsilon}),$$

Asymptotics of $\ln F(e^{-t})$

Theorem

For arbitrary $\varepsilon > 0$ and small positive t,

$$\ln F(e^{-t}) = -\lg(\rho)\ln(t) + c_0 + \sum_{k=1}^{\infty} a_k(t) + c_1 t + O(t^{2\lg(\rho)-\varepsilon}),$$

where $c_0 \approx 0.1216$ and $c_1 \approx 0.4501$ are constants

Asymptotics of $\ln F(e^{-t})$

Theorem

For arbitrary $\varepsilon > 0$ and small positive t,

$$\ln F(e^{-t}) = -\lg(\rho)\ln(t) + c_0 + \sum_{k=1}^{\infty} a_k(t) + c_1 t + O(t^{2\lg(\rho)-\varepsilon}),$$

where $c_0 \approx 0.1216$ and $c_1 \approx 0.4501$ are constants, and

$$a_k(t) = \frac{1}{\ln 2}\,\Re\left(\mathcal{M}\left(\frac{ik\pi}{\ln 2}\right)\exp(-ik\pi\lg(t))\right).$$

Asymptotics of $\ln F(e^{-t})$

Theorem

For arbitrary $\varepsilon > 0$ and small positive t,

$$\ln F(e^{-t}) = -\lg(\rho)\ln(t) + c_0 + \sum_{k=1}^{\infty} a_k(t) + c_1 t + O(t^{2\lg(\rho)-\varepsilon}),$$

where $c_0 \approx 0.1216$ and $c_1 \approx 0.4501$ are constants, and

$$a_k(t) = \frac{1}{\ln 2}\Re\left(\mathcal{M}\left(\frac{ik\pi}{\ln 2}\right)\exp(-ik\pi\lg(t))\right).$$

Note. It is easy to see that $a_k(4t) = a_k(t)$, so the $a_k(t)$ are periodic in $\log(t)$.

The oscillatory terms $a_k(t)$

We can write

$$a_k(t) = A_k \cos(k\pi \lg(t)) + B_k \sin(k\pi \lg(t)).$$

The oscillatory terms $a_k(t)$

We can write

$$a_k(t) = A_k \cos(k\pi \lg(t)) + B_k \sin(k\pi \lg(t)).$$

Define

$$C_k := \sqrt{A_k^2 + B_k^2} = \max_{t>0} |a_k(t)| = \frac{|\mathcal{M}(ik\pi / \ln 2)|}{\ln 2} .$$

The oscillatory terms $a_k(t)$

We can write

$$a_k(t) = A_k \cos(k\pi \lg(t)) + B_k \sin(k\pi \lg(t)).$$

Define

$$C_k := \sqrt{A_k^2 + B_k^2} = \max_{t>0} |a_k(t)| = \frac{|\mathcal{M}(ik\pi / \ln 2)|}{\ln 2}\,.$$

Numerically, we find
$C_1 \approx 2.1 \times 10^{-3}$, $C_2 \approx 2.2 \times 10^{-6}$, $C_3 \approx 2.8 \times 10^{-9}$,
$C_4 \approx 3.3 \times 10^{-12}$, ...

The oscillatory terms $a_k(t)$

We can write

$$a_k(t) = A_k \cos(k\pi \lg(t)) + B_k \sin(k\pi \lg(t)).$$

Define

$$C_k := \sqrt{A_k^2 + B_k^2} = \max_{t>0} |a_k(t)| = \frac{|\mathcal{M}(ik\pi/\ln 2)|}{\ln 2} .$$

Numerically, we find
$C_1 \approx 2.1 \times 10^{-3}$, $C_2 \approx 2.2 \times 10^{-6}$, $C_3 \approx 2.8 \times 10^{-9}$,
$C_4 \approx 3.3 \times 10^{-12}$, ...

The constants C_k appear to decrease exponentially fast as $k \to \infty$.

Sketch proof of the theorem

Consider the singularity of type (a).

Sketch proof of the theorem

Consider the singularity of type (a).

Define $L(s) := \mathcal{M}(s)/\Gamma(s)$.

Sketch proof of the theorem

Consider the singularity of type (a).

Define $L(s) := \mathcal{M}(s)/\Gamma(s)$. Then

$$L(0) = 2\ln\rho, \quad L'(0) = \widetilde{\mathcal{M}}(0) - 2\ln(\lambda)\ln(\rho) \approx 0.06.$$

Sketch proof of the theorem

Consider the singularity of type (a).

Define $L(s) := \mathcal{M}(s)/\Gamma(s)$. Then

$$L(0) = 2\ln\rho, \;\; L'(0) = \widetilde{\mathcal{M}}(0) - 2\ln(\lambda)\ln(\rho) \approx 0.06.$$

Near the double pole at $s = 0$,

$$\mathcal{F}(s) = \frac{L(0)}{2\ln 2}\, s^{-2} + c_0\, s^{-1} + O(1),$$

where

$$c_0 = \frac{(\ln 2 - \gamma)L(0) + L'(0)}{2\ln 2}\,.$$

Sketch proof of the theorem

Consider the singularity of type (a).

Define $L(s) := \mathcal{M}(s)/\Gamma(s)$. Then

$$L(0) = 2\ln\rho, \quad L'(0) = \widetilde{\mathcal{M}}(0) - 2\ln(\lambda)\ln(\rho) \approx 0.06.$$

Near the double pole at $s = 0$,

$$\mathcal{F}(s) = \frac{L(0)}{2\ln 2}\, s^{-2} + c_0\, s^{-1} + O(1),$$

where

$$c_0 = \frac{(\ln 2 - \gamma)L(0) + L'(0)}{2\ln 2}\,.$$

Standard arguments applied to the inverse Mellin transform now give the first two terms $(-\lg(\rho)\ln(t) + c_0)$.

Sketch proof continued

Now consider the singularities of type (b).

Sketch proof continued

Now consider the singularities of type (b).

These are simple poles at $s = ik\pi/\ln 2$ for $k \in \mathbb{Z}\backslash\{0\}$.

Sketch proof continued

Now consider the singularities of type (b).

These are simple poles at $s = ik\pi/\ln 2$ for $k \in \mathbb{Z}\backslash\{0\}$.

From the pole at $ik\pi/\ln 2$ we get a term

$$T_k(t) := \frac{1}{\ln 4}\mathcal{M}\left(\frac{ik\pi}{\ln 2}\right)\exp(-ik\pi\lg(t)).$$

Sketch proof continued

Now consider the singularities of type (b).

These are simple poles at $s = ik\pi/\ln 2$ for $k \in \mathbb{Z}\backslash\{0\}$.

From the pole at $ik\pi/\ln 2$ we get a term

$$T_k(t) := \frac{1}{\ln 4}\mathcal{M}\left(\frac{ik\pi}{\ln 2}\right)\exp(-ik\pi\lg(t)).$$

Combining the terms $T_k(t)$ and $T_{-k}(t)$ for $k \geq 1$, the imaginary parts cancel and we are left with the oscillatory term $a_k(t)$.

Sketch proof continued

Now consider the singularity of type (c).

Sketch proof continued

Now consider the singularity of type (c).

At $s = -1$, $\mathcal{F}(s)$ has a pole with residue

$$c_1 = \frac{\lambda \ln \mu_1}{3} = \frac{\mu_1'}{3\mu_1} = \frac{23 + 3\sqrt{5}}{66}\,.$$

Sketch proof continued

Now consider the singularity of type (c).

At $s = -1$, $\mathcal{F}(s)$ has a pole with residue

$$c_1 = \frac{\lambda \ln \mu_1}{3} = \frac{\mu_1'}{3\mu_1} = \frac{23 + 3\sqrt{5}}{66}\,.$$

This accounts for the term $c_1 t$.

Sketch proof continued

Now consider the singularity of type (c).

At $s = -1$, $\mathcal{F}(s)$ has a pole with residue

$$c_1 = \frac{\lambda \ln \mu_1}{3} = \frac{\mu_1'}{3\mu_1} = \frac{23 + 3\sqrt{5}}{66}\,.$$

This accounts for the term $c_1 t$.

Finally, the error term $O(t^{2\lg(\rho)-\varepsilon})$ allows for the fact that we have only considered the singularities of $\mathcal{F}(s)$ in $\mathcal{H}$.

Sketch proof continued

Now consider the singularity of type (c).

At $s = -1$, $\mathcal{F}(s)$ has a pole with residue

$$c_1 = \frac{\lambda \ln \mu_1}{3} = \frac{\mu_1'}{3\mu_1} = \frac{23 + 3\sqrt{5}}{66}\,.$$

This accounts for the term $c_1 t$.

Finally, the error term $O(t^{2\lg(\rho)-\varepsilon})$ allows for the fact that we have only considered the singularities of $\mathcal{F}(s)$ in $\mathcal{H}$.

There could be (in fact are) other singularities in the half-plane

$$\{s \in \mathbb{C} : \Re(s) \le -2\lg(\rho)\}.$$

A corollary

All that we actually need for the applications is the following.

A corollary

All that we actually need for the applications is the following.

Corollary

For $z \in [0, 1)$,

$$F(z) = \frac{C(z)}{(1 - z)^{\lg \rho}} \,,$$

where $C(z)$ is a positive oscillatory term, bounded away from zero and infinity.

A corollary

All that we actually need for the applications is the following.

Corollary

For $z \in [0, 1)$,

$$F(z) = \frac{C(z)}{(1-z)^{\lg \rho}} ,$$

where $C(z)$ is a positive oscillatory term, bounded away from zero and infinity.

Remark

We do not need the full machinery of Mellin transforms to deduce the Corollary. Instead we could use the quantitiative version of Perron's theorem due to Coffman (1964).

A conjecture

We conjecture that $\mathcal{M}(s)$ and $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ have poles at $s = -2\lg(\rho) + ik\pi/\ln(2)$ for $k \in \mathbb{Z}$.

A conjecture

We conjecture that $\mathcal{M}(s)$ and $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ have poles at $s = -2\lg(\rho) + ik\pi/\ln(2)$ for $k \in \mathbb{Z}$.

This would account for numerical evidence that the error $e_1(t)$ in the linear approximation to $\mu(e^{-t})$ is of order $t^{2\lg(\rho)}$ but $e_1(t)/t^{2\lg(\rho)}$ does not tend to a limit as $t \to 0^+$; instead it has small oscillations that are periodic in $\lg t$.

A conjecture

We conjecture that $\mathcal{M}(s)$ and $\mathcal{F}(s) = (1 - 4^{-s})^{-1}\mathcal{M}(s)$ have poles at $s = -2\lg(\rho) + ik\pi/\ln(2)$ for $k \in \mathbb{Z}$.

This would account for numerical evidence that the error $e_1(t)$ in the linear approximation to $\mu(e^{-t})$ is of order $t^{2\lg(\rho)}$ but $e_1(t)/t^{2\lg(\rho)}$ does not tend to a limit as $t \to 0^+$; instead it has small oscillations that are periodic in $\lg t$.

k	$t = 2^{-k}$	$\mu(e^{-t})$	$e_1(t)$	$e_1(t)/t^{2\lg\rho}$
20	9.5367e-7	2.6180306	1.1708e-8	2.6790
21	4.7684e-7	2.6180323	4.4999e-9	2.6958
22	2.3842e-7	2.6180331	1.7079e-9	2.6787
23	1.1921e-7	2.6180336	6.5648e-10	2.6956
24	5.9605e-8	2.6180338	2.4917e-10	2.6786

Approximation of $\mu(e^{-t})$ for $t = 2^{-k}$, $20 \le k \le 24$, $e_1(t) = \mu(e^{-t}) - (\mu_1 - t\mu_1')$.

References

B. Adamczewski, Non-converging continued fractions related to the Stern diatomic sequence, *Acta Arith.* **142** (2010), 67–78.

R. P. Brent, Analysis of the binary Euclidean algorithm, *New Directions and Recent Results in Algorithms and Complexity*, Academic Press, New York, 1976, 321–355.

R. P. Brent, Twenty years' analysis of the binary Euclidean algorithm, *Millennial Perspectives in Computer Science* . . . , Palgrave, 2000, 41–53.

R. P. Brent, M. Coons and W. Zudilin, Algebraic independence of Mahler functions via radial asymptotics, *IMRN*, to appear.

P. Bundschuh and K. Väänänen, Transcendence results on the generating functions of the characteristic functions of certain self-generating sets, I, *Acta Arith.* **162** (2014), 273–288.

C. V. Coffman, Asymptotic behavior of solutions of ordinary difference equations, *Trans. AMS* **110** (1964), 22–51.

M. Coons, The transcendence of series related to Stern's diatomic sequence, *Int. J. Number Theory* **6** (2010), 211–217.

K. Dilcher and K. Stolarsky, Stern polynomials and double-limit continued fractions, *Acta Arith.* **140** (2009), 119–134.

P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, CUP, 2009. Appendix B.7: *Mellin transforms*.

D. E. Knuth, *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms* (3rd ed.), Addison-Wesley, 1997.

G. Maze, Existence of a limiting distribution for the binary GCD algorithm, *J. Discrete Algorithms* **5** (2007), 176–186.

I. D. Morris, *A rigorous version of R. P. Brent's model for the binary Euclidean algorithm*, arXiv:1409.0729, 2 Sept. 2014.

M. A. Stern, Über eine zahlentheoretische Funktion, *J. Reine Angew. Math.*, **55** (1858), 193–220. See also OEIS A002487.

B. Vallée, Dynamics of the binary Euclidean algorithm: functional analysis and operators, *Algorithmica* **22** (1998), 660–685.

D. Zagier, The Mellin transform and other useful analytic techniques, in E. Zeidler, *Quantum Field Theory I ...*, Springer-Verlag, 2006, 305–323.