INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE

Continued Fraction Algorithms, Functional Operators, and Structure Constants

Philippe FLAJOLET, Brigitte VALLÉE

N° 2931
Juillet 1996

PROGRAMME 2

Rapport de recherche

1996

ISSN 0249-6399

Continued Fraction Algorithms, Functional Operators, and Structure Constants[1]

Philippe Flajolet and Brigitte Vallée

Abstract. *Continued fractions lie at the heart of a number of classical algorithms like Euclid's greatest common divisor algorithm or the lattice reduction algorithm of Gauss that constitutes a 2-dimensional generalization. This paper surveys the main properties of functional operators, —transfer operators— due to Ruelle and Mayer (also following Lévy, Kuzmin, Wirsing, Hensley, and others) that describe precisely the dynamics of the continued fraction transformation. Spectral characteristics of transfer operators are shown to have many consequences, like the normal law for logarithms of continuants associated to the basic continued fraction algorithm, where better convergence terms are obtained, and a purely analytic estimation of the average number of steps of the Euclidean algorithm. Transfer operators also lead to a complete analysis of the "Hakmem" algorithm for comparing two rational numbers via partial continued fraction expansions, and of the "digital tree" algorithm for completely sorting n real numbers by means of their continued fraction representations. Thus, a small number of "structure constants" appear to govern the behaviour of a variety of continued fraction based algorithms.*

Algorithmes de fractions continues, opérateurs fonctionnels et constantes de structure

Résumé. Les fractions continues sont au cœur de nombreux algorithmes, comme l'algorithme d'Euclide ou l'algorithme de réduction des réseaux de Gauss qui en constitue une généralisation en dimension 2. Nous passons en revue les principales propriétés d'opérateurs fonctionnels —opérateurs de transfer— dus à Ruelle et Mayer (faisant suite à Lévy, Kuzmin, Wirsing, Hensley et d'autres), qui décrivent précisément la dynamique de la transformation des fractions continues. Les propriétés spectrales de ces opérateurs ont de nombreuses conséquences, comme: la loi normale des logarithmes de continuants où sont obtenus de meilleurs termes d'erreur, ainsi qu'une estimation purement analytique du nombre moyen d'étapes de l'algorithme d'Euclide. Les opérateurs de transfert conduisent de surcroît à une analyse complète de l'algorithme "Hakmem" de comparaison de rationnels et de l'algorithme de tri digital appliqué aux fractions continues. Il apparaît ainsi qu'un petit nombre de constantes de structure gouvernent le comportement d'une variété d'algorithmes fondés sur les fractions continues

[1]**Text of an invited lecture at the *Seventh International Conference on Fibonacci Numbers and their Applications*, Graz, July 1996.**

Continued Fraction Algorithms, Functional Operators, and Structure Constants

Philippe FLAJOLET
INRIA-Rocquencourt, F-78153 Le Chesnay (France)
[Philippe.Flajolet@inria.fr]
, Brigitte VALLÉE
GREYC, Université de Caen, F-14032 Caen (France)
[Brigitte.Vallee@info.unicaen.fr]

Abstract

Continued fractions lie at the heart of a number of classical algorithms like Euclid's greatest common divisor algorithm or the lattice reduction algorithm of Gauss that constitutes a 2-dimensional generalization. This paper surveys the main properties of functional operators, —transfer operators— due to Ruelle and Mayer (also following Lévy, Kuzmin, Wirsing, Hensley, and others) that describe precisely the dynamics of the continued fraction transformation. Spectral characteristics of transfer operators are shown to have many consequences, like the normal law for logarithms of continuants associated to the basic continued fraction algorithm, where better convergence terms are obtained, and a purely analytic estimation of the average number of steps of the Euclidean algorithm. Transfer operators also lead to a complete analysis of the "Hakmem" algorithm for comparing two rational numbers via partial continued fraction expansions, and of the "digital tree" algorithm for completely sorting n real numbers by means of their continued fraction representations. Thus, a small number of "structure constants" appear to govern the behaviour of a variety of continued fraction based algorithms.

1 Algorithms and operators

Perhaps the simplest of all algorithmic schemes is the following: *"Starting with an initial object of some domain $\mathcal{D}$, repeatedly apply a transformation $x \mapsto T(x)$ until a certain halting condition is met."* In other words, one computes a sequence of iterates

$$(S) \qquad x_0, \quad x_1 = T(x_0), \quad x_2 = T(x_1), \quad \ldots \quad , \quad x_n = T(x_{n-1}), \quad \ldots \, ,$$

until some (possibly vacuous) condition H is satisfied.

Two special cases of importance are the *binary expansion* transformation and the *continued fraction* transformation that are defined by

$$T_{BE}(x) = \{2x\}, \qquad T_{CF}(x) = \{\frac{1}{x}\}, \tag{1}$$

where $\{u\} = u \bmod 1 = u - \lfloor u \rfloor$ is the fractional part of u and $\mathcal{D} = [0, 1[$. These two transformations give rise to the binary expansion and to the continued fraction representation of x_0,

$$x_0 = \sum_{j=1}^{\infty} m_j 2^{-j}, \qquad x_0 = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots}}}, \tag{2}$$

where the "digits" $m_j = m_j(x_0)$ are obtained by either $m_j = \lfloor 2x_{j-1} \rfloor$ or $m_j = \lfloor 1/x_{j-1} \rfloor$.

The transformations of interest in this context are usually far from injective but at least have a finite or denumerable set of inverse branches. Keeping track of which branch of the inverse function corresponds to each application of T provides a *symbolic coding* of x_0, or equivalently a generalized *number representation system*. Here,

$$T_{BE}^{-1}(x) = \left\{ \frac{x+m}{2} \right\}_{m=0,1}, \qquad T_{CF}^{-1}(x) = \left\{ \frac{1}{m+x} \right\}_{m=1}^{\infty}, \tag{3}$$

and the rules for computing the "digits" associated to T_{BE} and T_{CF} are precisely of this nature. The common process is described by a canonical numbering of the branches of T^{-1}, setting $T^{-1}(x) = \{h_m(x)\}_{m \in M}$; the coding of x_0 is then $(m_1, m_2, m_3, \ldots)$ corresponding to the diagram

$$x_0 \quad \underset{h_{m_1}}{\overset{T}{\rightleftarrows}} \quad x_1 \quad \underset{h_{m_2}}{\overset{T}{\rightleftarrows}} \quad x_2 \quad \underset{h_{m_3}}{\overset{T}{\rightleftarrows}} \quad \cdots. \tag{4}$$

Loosely speaking, such a coding system works, because $T(x)$ is an expanding map, ($|T'(x)| > 1$), so that the inverse branches are contracting, ($|h_j'(x)| < 1$). The reader is referred to the book [2] for an outstanding exposition of the general theory.

One introduces the so-called *fundamental intervals* defined as

$$I_{m_1, m_2, \ldots, m_k} = h_{m_1} \circ h_{m_2} \circ \cdots \circ h_{m_k}([0, 1]),$$

where the parameter k is called the depth. Here, the inverse branches are all monotonic, either decreasing (T_{CF}) or increasing (T_{BE}), so that the interior

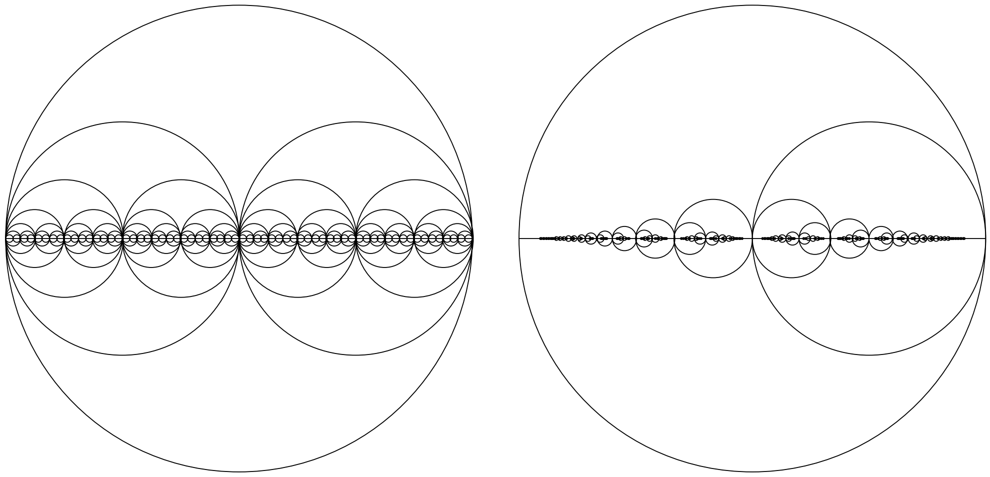

Figure 1: The fundamental circles associated with the binary expansion transformation (left) and the continued fraction transformation (right).

of $I_{m_1,m_2,\ldots,m_k}$ contains all the irrational numbers whose representation starts with $m_1, m_2, \ldots, m_k$. There are important boundary cases, the dyadic rationals for T_{BE} or the rationals for T_{CF}, where the expansion terminates. These sets are of measure zero, but play an important rôle in analysing the dynamics of T.

The geometry of fundamental intervals is well illustrated by a two-dimensional diagram where an interval J is represented by the circle of diameter J; the circles built over fundamental intervals are called *fundamental circles*. Such circle diagrams exhibit a fractal-like aspect, see Fig. 1. Apart from their esthetic aspect, they relate directly to algorithms as the areas of the fundamental circles give basic informations on the 2-dimensional algorithms to be considered later.

Analysis of algorithms. The purpose of analysis of algorithms is to characterize the cost of an algorithm under a well-defined probabilistic model that describes the initial distribution of its input x_0. The problematics of analysis of algorithms is well illustrated by Knuth's monumental series, *The Art of Computer Programming*, especially [18] that includes an interesting perspective on continued fractions and the Euclidean algorithm.

Consider for instance the binary expansion algorithm applied to an initial value x_0 that is uniformly distributed over $[0, 1]$ and halt it after n iterations of T_{BE}. Assume that we have to pay a toll of 1 each time a 1-bit is encountered. Each x_j is now uniformly distributed so that the "dynamics" of the algorithm is especially simple. Moreover, each bit m_j has independently of the other equal probability of being 0 or 1, so that the total cost C_n is a binomial (or Bernoulli) distributed random variable,

$$\Pr\{C_n = r\} = \frac{1}{2^n}\binom{n}{r}. \tag{5}$$

The mean of the cost is $\mu_n = n/2$, the standard deviation is $\sigma_n = \sqrt{n}/2$, and the famous Gauss-Laplace-De Moivre theorem even gives us a limit distribution,

$$\lim_{n\to\infty} \Pr\left\{\frac{C_n - \mu_n}{\sigma_n}\right\} \to \Phi(x), \qquad \text{with} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt, \tag{6}$$

the distribution function of a standard Gaussian variable. Our purpose in this paper is precisely to analyse characteristics of continued fraction based algorithms and detect laws that in a way generalize properties of the binary expansion algorithm.

Functional operators. For obvious reasons, in the analysis of algorithms, one has also to consider input distributions that may be rather nonuniform. Assume for instance that we run the binary expansion algorithm starting from an x_0 that is distributed like a random variable X with density $w(x)$,

$$\Pr\{x < X \le x + dx\} = w(x)\,dx.$$

Then, one iteration of T_{BE} produces a new random variable, $Y = T(X)$, with density

$$w^*(x) \equiv \frac{1}{dx}\Pr\{x < Y \le x + dx\} = \frac{1}{2}w(\frac{x}{2}) + \frac{1}{2}w(\frac{x+1}{2}),$$

as shown by a simple computation that tracks the possible origins of Y. This suggests to introduce an operator

$$\mathcal{G}_{BE}[f](x) := \frac{1}{2}f(\frac{x}{2}) + \frac{1}{2}f(\frac{x+1}{2}), \tag{7}$$

so that the distribution of x_k is precisely described by the density function

$$\mathcal{G}_{BE}^{k}[w](x) = \frac{1}{2^k}\sum_{i=0}^{2^k-1} w(\frac{x+i}{2^k}). \tag{8}$$

Thus, provided w is smooth enough, for instance $\mathcal{C}^1$, we have

$$\mathcal{G}_{BE}^{k}[w](x) = 1 + O(\frac{1}{2^k}), \tag{9}$$

as the sum in (8) is a Riemann sum of step 2^{-k}.

The distribution of the kth iterate x_k therefore approaches the uniform distribution exponentially fast as the algorithm proceeds. In particular, digits are asymptotically nearly equidistributed and a limit law like (6) persists (by exponential mixing, see [4, Sec.27]), although closed-form expressions like (5) are no longer available. Phrased differently, Eq. (9) expresses that the iterates of $\mathcal{G}_{BE}$ converge to 1, an eigenfunction corresponding to the eigenvalue 1, at a rate of $\frac{1}{2}$ that appears to be the subdominant (second) eigenvalue of $\mathcal{G}_{BE}$[1].

This simple example suggests a direct relationship between the dynamics of T, the computational complexity properties of algorithms based on the iteration of T, and spectral properties of an operator closely related to the way T transforms probability distributions. The basic ingredient, well-developed in dynamical systems theory, is the class of *transfer operators* [2, 30] whose general form is

$$\mathcal{G}_s[f](x) = \sum_m |h'_m(x)|^{s/2} f(h_m(x)), \tag{10}$$

(the sum is extended to all branches h_m of T^{-1}) so that the density function of x_k is $\mathcal{G}_2^k[w](x)$. The dynamics of the basic algorithm is described by $s = 2$ but insertion of the parameter s proves highly useful, as we shall see in the sequel.

For continued fractions, the operator $\mathcal{G}_s$ assumes the form

$$\mathcal{G}_s[f](x) = \sum_{m=1}^{\infty} \frac{1}{(m+x)^s} f(\frac{1}{m+x}), \tag{11}$$

and is called the *Ruelle-Mayer operator*, see [25, 30]. For $\Re(s) > 1$ and s near the real axis, this operator has a unique dominant eigenvalue $\lambda(s)$, so that $|\lambda(s)| > |\mu(s)|$ for $\mu(s)$ a subdominant eigenvalue. (See Section 3 for details).

It was realized by Gauss as early as 1800 that the operator $\mathcal{G}_2$ admits the density function

$$\psi_2(x) = \frac{1}{\log 2} \frac{1}{1+x} \tag{12}$$

as eigenfunction corresponding to the eigenvalue $\lambda(2) = 1$, a fact easily verified by a straightforward calculation. Equivalently, $\psi_2(x)$ is invariant under the continued fraction transformation T_{CF}. It took however more than a century to prove that the iterates $G_2^k[w]$ converge geometrically fast to the Gauss distribution $\psi_2(x)$,

$$\mathcal{G}_2^k[w](x) = \psi_2(x) + O(|\mu(2)|^k), \qquad \text{where} \quad \mu(2) \approx -0.30366\ldots, \tag{13}$$

for a wide class of distributions $w(x)$ that includes the uniform distribution. There $\mu(2)$, known as the Gauss-Kuzmin-Wirsing constant, is the subdominant

[1] The complete spectrum of $\mathcal{G}$ (acting on functions analytic in a disc that properly contains the interval $[0,1]$) is the geometric progression $\{2^{-r}\}_{r=0}^{\infty}$, with the rth Bernoulli polynomial $B_r(x)$ an eigenfunction corresponding to the eigenvalue 2^{-r}.

eigenvalue of $\mathcal{G}_2$. Nowadays, we may regard the convergence expressed by (13) but without its error term as a consequence of the ergodic theorem. We refer globally to the books by Knuth [18], by Khinchin [16], and by Rockett and Szűsz [29] for a detailed exposition of the elementary metric theory of continued fractions.

Properties of coefficients in continued fraction expansions are naturally strongly dependent on properties of $\mathcal{G}_2$ and $\psi_2(x)$. For instance, if x is uniformly distributed over $[0,1]$, then its kth continued fraction "digit" $m_k(x)$ has value ℓ whenever $T^{k-1}(x) \in [\frac{1}{\ell}, \frac{1}{\ell+1}[$, and by the law of Gauss,

$$\lim_{k\to+\infty} \Pr\{m_k(x) = \ell\} = \int_{1/(\ell+1)}^{1/\ell} \psi_2(t)\, dt = \log_2\left(1 + \frac{1}{\ell(\ell+2)}\right), \qquad (14)$$

with a speed of convergence that is again $O(|\mu(2)|^k)$. In the asymptotic limit, the continued fraction quotients are thus distributed like a random variable whose probability distribution is given by (14).

In the remainder of this paper, we shall examine related properties, like the fine distribution of continuants, and derive a Gaussian law that bears a superficial resemblance to (6). Such properties entertain in turn close ties with the distribution of the lengths of fundamental intervals and with the behaviour of the Euclidean algorithm. The problems are however more difficult than for binary expansions as continued fraction digits are inherently (though somewhat "weakly") correlated. The proof techniques there make a heavy use of spectral properties of the operators $\mathcal{G}_s$ when s lies in a complex neighbourhood of 2. In particular Lévy's constant, $\lambda'(2) = -\pi^2/(12\log 2)$, that relates to the entropy of fundamental intervals, intervenes for dominant asymptotics both in the law of continuants and in the number of steps of the Euclidean algorithm.

Two-dimensional algorithms. One of our goals is to analyse two-dimensional versions of the basic algorithmic scheme based on iteration of a function T. For binary expansions and continued fractions, T is piecewise monotonic, either increasing like T_{BE} or decreasing like T_{CF}. Then we may compare a pair of real numbers (x_0, y_0) by running "in parallel" two iterations of T that start with the two numbers x_0 and y_0, and halt as soon as a discrepancy is detected. (Special rational cases are easily tested and subjected to an appropriate treatment.)

The two real inputs cause *at least* k iterations of the algorithm if they both belong to a *fundamental square* of the form

$$Q_{m_1,\ldots,m_k} = I_{m_1,\ldots,m_k} \times I_{m_1,\ldots,m_k}.$$

These fundamental squares are analogous to the circle representations associated to fundamental intervals of Fig. 1. Their structure in the case of T_{CF} is illustrated in Fig. 4, where black and white are used to depict squares of odd and even depth respectively.

Assume that (x_0, y_0) is uniformly distributed over the unit square. Once more, the binary expansion algorithm is easy to cope with. In that case, the number L of iterations of the algorithm has mean value equal to 2 and a distribution that is geometric with parameter $1/2$, that is to say $\Pr\{L = r\} = 2^{-r}$. In the case of an arbitrary base b, the mean becomes $b/(b-1)$ and the law is a geometric one with parameter $1/b$.

The continued fraction version of the algorithm seems to have been first formulated for comparing two rational numbers in the celebrated HAKMEM memo [3]. In this context no multiprecision operation is required as the successive rationals $T^k(x_0), T^k(y_0)$ are of diminishing sizes. Interest in the algorithm was recently rekindled as it may be used to find the sign of 2×2 integer determinants, a recurrent problem of some importance in computational geometry that is equivalent to the comparison of rational numbers. The geometry of these problems is furthermore isomorphic to the one that arises in the analysis of the Gaussian algorithm for lattice reduction [5, 33]. Thus, multidimensional generalizations of the continued fraction algorithm are of interest in rather diverse situations.

For continued fractions, it is now the operator $\mathcal{G}_s$ taken at $s = 4$ that plays a rôle. As we shall see, the comparison of the continued fraction expansions of two reals necessitates computing L partial quotients, where L has a mean value $\mathrm{Ex}\{L\} \approx 1.35113$, an interesting constant that is expressible in terms of $\zeta(3)$ and the tetralogarithm $Li_2(1/2)$. Furthermore, the distribution of L decays exponentially and

$$\Pr\{L \geq r\} \underset{r \to +\infty}{\sim} C\,\lambda(4)^r, \qquad \text{where} \quad \lambda(4) \approx 0.19945$$

is the dominant eigenvalue of the operator $\mathcal{G}_4$. Under this measure, continued fractions behave roughly like base 5 representations, since $\lambda(4)^{-1} \approx 5$.

The algorithm generalizes to a test based on "digital trees" for deciding distinctness of n real numbers represented by continued fractions (this also provides a way to sort). The analysis provides interesting indications on the collective behaviour of n "random" continued fractions. In this context the quantities $-2\lambda'(2)$ and $\lambda(4)$ play a rôle analogous to the the entropy and the probability of coincident digits for other number representation systems.

Plan of the paper. We first make explicit the basic algebraic and analytic properties of transfer operators associated with continued fractions in Section 2. Next, in Section 3, we establish the Gaussian law of the logarithms of continuants (Philipp's theorem) and derive improved error terms for convergence to the asymptotic limit. In Section 4, we show that the basic average-case analysis of Euclid's algorithm has a simple formulation in terms of transfer operators, which leads to a simple proof of the Heilbronn-Dixon theorem in the style of

Constant	*Related properties*	*Ref.*
$\lambda(2) = 1$	Limit law of Gauss for continued fractions	Eq.(13)
$\mu(2) \approx -0.30366$ (Wirsing's constant)	Speed of convergence in the law of Gauss and in moments of continuants	Eq.(13) Th.1
$\lambda'(2) = \frac{-\pi^2}{12\log 2}$ (Lévy's constant)	Entropy of fundamental intervals	Eq.(31)
	Mean value of $\frac{1}{k}\log Q_k$	Th.1
	Mean number of steps of Euclid's algorithm	Th.2
	Size and path length of CF trees	Th.4
$\lambda''(2)$ (Hensley's constant)	Variance in the law of continuants and in Euclid's algorithm	Th.1 [13]
$\lambda(4) \approx 0.19945$ (Vallée's constant)	Rate of coincidence of CF digits	Th.3
	Height of CF trees	Th.5

Figure 2: A synopsis of occurrences of the major structure constants.

classical analytic number theory. Sections 5 and 6 are devoted to the analysis of algorithms for comparing 2 real numbers and sorting n real numbers, given their continued fraction representations.

The analytic properties of transfer operators associated to continued fractions have been superbly worked out by Mayer in a series of papers [24, 25, 22, 23] that provide the technical background for the functional analysis aspects of the our paper. We also took inspiration from perturbative properties developed by Hensley [13] in his deep results relative to Euclid's algorithm.

Througout this paper, we encounter recurrently a few basic "structure constants" arising from transfer operators —$\lambda(2), \mu(2), \lambda'(2), \lambda''(2), \lambda(4)$, most notably. Figure 2 summarizes their basic connections to continued fraction. (See [9] for an attractive discussion of a wide class of constants.)

2 Continuants and operators

The basic continued fraction transformation and its digit-extraction function will be denoted simply by

$$T(x) = \left\{ \frac{1}{x} \right\}, \qquad m(x) = \left\lfloor \frac{1}{x} \right\rfloor.$$

The *continued fraction algorithm* then reads

Algorithm $CF(x)$ **for** $k := 1$ **while** $x \neq 0$ **do** $\{\ m_k := m(x);\ x := T(x)\ \}$.

The transfer operators are directly related to the iteration of T and their analytic properties are the key to asymptotic properties of continued fraction algorithms.

The algebra of transfer operators. To a real number x, the algorithm associates the sequence of iterates $x_0 = x, x_1, x_2, \ldots, x_k, \ldots$, and provided the kth iterate exists,

$$x_0 = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots \cfrac{}{m_k + x_k}}}}.$$

where by construction $m_j \geq 1$. The relation between x_0 and x_k is described by a *linear fractional transformation* (LFT), also called homography,

$$h_{\mathbf{m}}(z) = h_{m_1, m_2, \ldots, m_k}(z) = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots \cfrac{}{m_k + z}}}}.$$

The quantity k is called the *depth* of h and denoted by $|h|$; the collection of all LFTs of depth k gives all the inverse branches of T^k.

As is well-known [27], an LFT is expressible in terms of *continuant* polynomials,

$$h_{\mathbf{m}}(z) = \frac{P_k + zP_{k-1}}{Q_k + zQ_{k-1}}, \tag{15}$$

where

$$\begin{array}{rclrcl} Q_k & = & Q_k(m_1, \ldots, m_k), & Q_{k-1} & = & Q_{k-1}(m_1, \ldots, m_{k-1}), \\ P_k & = & Q_{k-1}(m_2, \cdots, m_k), & P_{k-1} & = & Q_{k-2}(m_2, \ldots, m_{k-1}). \end{array}$$

The continuant polynomials are determined by the recurrence

$$Q_k(m_1, m_2, \ldots, m_k) = m_k Q_{k-1}(m_1, \ldots, m_{k-1}) + Q_{k-2}(m_1, \ldots, m_k), \tag{16}$$

with $Q_0 = 1$, $Q_1(m_1) = m_1$. It is also well-known that the continuant polynomial $Q_k(\mathbf{m})$ is obtained by taking the sum of monomials obtained from $m_1 m_2 \cdots m_k$ by crossing out in all possible ways pairs of adjacent variables $m_i m_{i+1}$. Hence, continuants satisfy the symmetry property,

$$Q_k(m_1, \ldots, m_k) = Q_k(m_k, \ldots, m_1), \tag{17}$$

as well as the determinant identity,

$$Q_k P_{k-1} - Q_{k-1} P_k = (-1)^k. \tag{18}$$

The transfer operators of Ruelle-Mayer are defined by

$$\mathcal{G}_s[f](z) = \sum_{m \geq 1} \frac{1}{(m+z)^s} \, f(\frac{1}{m+z}),$$

for s a real number satisfying $\Re(s) > 1$. There the sum may be seen also as taken over all LFTs of depth 1. By the chain rule, the kth iterate of $\mathcal{G}_s$ is also

$$\mathcal{G}_s^k[f](z) = \sum_{|h|=k} \left((-1)^k h'(z)\right)^{s/2} \; f(h(z)),$$

with a sum now ranging over all LFTs of depth k. The determinant identity induces a simplification in the last sum since

$$h'_{\mathbf{m}}(z) = \frac{(-1)^k}{(Q_{k-1}z + Q_k)^2}.$$

This is summarized by the following proposition.

Proposition 1 *The iterate of order k of the operator $\mathcal{G}_s$ is expressible in terms of continuants of order k:*

$$\mathcal{G}_s^k[f](z) = \sum_{m_1 \ldots m_k} \frac{1}{(Q_{k-1}z + Q_k)^s} \; f(\frac{P_{k-1}z + P_k}{Q_{k-1}z + Q_k}). \tag{19}$$

In particular,

$$\mathcal{G}_s^k[f](0) = \sum_{m_1 \ldots m_k} \frac{1}{Q_k^s} \; f(\frac{P_k}{Q_k}) \;\; = \sum_{m_1 \ldots m_k} \frac{1}{Q_k^s} \; f(\frac{Q_{k-1}}{Q_k}) \tag{20}$$

An important consequence of the determinant identity is that the length of the fundamental interval relative to an LFT $h = h_{\mathbf{m}}$ is

$$u_h = |h(1) - h(0)| = \frac{1}{Q_k(Q_k + Q_{k-1})}, \tag{21}$$

with $Q_k = Q_k(m_1, \ldots, m_k)$ and $Q_{k-1} = Q_{k-1}(m_1, \ldots, m_{k-1})$ the continuants associated to h. Thus, from Prop. 1, many quantities related to these intervals are expressible in terms of transfer operators: see the law of continuants and Eq. (29–30), Euclid's algorithm and Lemma 1, the sign algorithm and Eq. (37), as well as Eq. (49) for the sorting algorithm.

The analysis of transfer operators. Asymptotic properties of continued fraction algorithms are closely related to *spectral properties* of transfer operators. The situation in a way parallels that of Markov chains where a dominant eigenvector gives the stationary distribution of the chain while the subdominant eigenvalue is an indicator of the speed of convergence to the stationary limit.

In order to develop such spectral properties, one first needs to make precise the functional spaces on which the $\mathcal{G}_s$ operators are applied. We let $\mathcal{J}$ and $\mathcal{V}$ denote respectively the interval and the open disk of centre 1 and diameter 5/4.

The disc and the interval are mapped strictly within themselves by an arbitrary LFT of depth 1, and hence by any LFT of arbitrary depth. The $\mathcal{G}_s$ operators for $\Re(s) > 1$ are then taken to act on the space $A_\infty(\mathcal{V})$ formed with all functions f that are holomorphic in the disc $\mathcal{V}$ and are continuous on the closed disc $\bar{\mathcal{V}}$. Endowed with the sup-norm,

$$||f|| = \sup\ \{|f(z)|,\ z \in \mathcal{V}\},$$

$A_\infty(\mathcal{V})$ is a Banach space. The starting point is the following easy result. There, $\zeta(s) = \sum_{n\geq 1} n^{-s}$ is the Riemann zeta function.

Proposition 2 *For $\Re(s) > 1$, the operator $\mathcal{G}_s$ is bounded,*

$$||\mathcal{G}_s f|| \leq \zeta(\sigma)\, ||f||,$$

and compact. Its spectrum is discrete, with only an accumulation point at 0.

Proof. Boundedness results from the triangular inequality. In addition, the operator is compact by classical properties of bounded sequences of analytic functions. Hence, its spectrum is discrete by a basic theorem on compact operators over Banach spaces [20]. (See Mayer's papers [25, 23] for details.) □

In this paper, we make essential use of the spectral radius $\rho(s)$ (the maximum modulus of eigenvalues) of the operator $\mathcal{G}_s$. The properties on which our analysis rests are the following:

(P_1) *Positivity.* When $s = \sigma$ is real, the operator $\mathcal{G}_s$ satisfies a strong positivity property related to the Perron-Frobenius theory [19, 25, 34]. There is a unique dominant eigenvalue (of largest modulus) $\lambda(\sigma)$ that is positive and has multiplicity 1. Thus $\lambda(\sigma)$ equals the spectral radius $\rho(\sigma)$. In particular, this makes it possible to separate $\lambda(\sigma)$ and a subdominant eigenvalue $\mu(\sigma)$ (one of second largest modulus): $\lambda(\sigma) > |\mu(\sigma)|$.

(P_2) *Perturbation.* By the classical theory of analytic perturbation [15], for s in a sufficiently small neighbourhood of any point σ of the real axis, unicity of the dominant eigenvalue $\lambda(s)$ is preserved, so that the separation of eigenvalues, $|\lambda(s)| > |\mu(s)|$, survives. In addition, the basic spectral quantities depend analytically on s. These results have been developed by Faivre [8] and Hensley [13].

(P_3) *Maximum property.* In any half-plane $\Re(s) \geq \sigma > 1$, the spectral radius $\rho(s)$ attains its unique maximum at $s = \sigma$ where it equals $\lambda(\sigma)$.

(P_4) *Special properties.* The dominant spectral objects at $s = 2$ are known. The dominant eigenvalue is $\lambda(2) = 1$, with corresponding eigenfunction the Gauss function $\psi_2(z) = (1+z)^{-1}$. Wirsing [34] has determined the subdominant eigenvalue $\mu(2) \approx -0.30366$.

We now detail properties P_1, P_2, P_3, P_4 on which this paper is built.

Positivity (P_1). By the Perron-Frobenius theory of positive operators [19], for *positive real* $s > 1$, the operator $\mathcal{G}_s$ has a unique dominant eigenvalue $\lambda(s)$ that is positive. The corresponding eigenfunction $\psi_s(x)$ is strictly positive on the real interval $\mathcal{J}$. Thus, $\mathcal{G}_s$ decomposes as

$$\mathcal{G}_s = \lambda(s)\mathcal{P}_s + \mathcal{N}_s,$$

where $\mathcal{P}_s$ is the projection over the eigenspace determined by ψ_s and $\mathcal{N}_s$ has a spectral radius strictly smaller than $\lambda(s)$. In other words, one has

$$\mathcal{G}_s[f](x) = \lambda(s)\, e_s[f]\, \psi_s(x) + \mathcal{N}_s[f](x), \tag{22}$$

for some continuous linear form $e_s[f]$.

If f is a function of $A_\infty(\mathcal{V})$, there results

$$\mathcal{G}_s^k[f](x) = \lambda(s)^k e_s[f]\psi_s(x) + \mathcal{N}_s^k[f](x), \tag{23}$$

for any $k \geq 1$ and any $z \in \mathcal{V}$, the spectral radius of $\mathcal{N}_s^k$ being $|\mu(s)|^k$.

Perturbation (P_2). The dependency of $\mathcal{G}_s$ with respect to s is an analytic (holomorphic) one. Therefore, by the standard theory of analytic perturbation, the relation (23) persists for s in sufficiently small neighbourhoods of any real point $\sigma > 1$. In summary:

Proposition 3 *For any positive $\sigma > 1$, there exist two positive reals r_0 and M such that, for all s satisfying $|s - \sigma| \leq r_0$, one has*

$$\mathcal{G}_s^k[f](x) = \lambda(s)^k e_s[f]\psi_s(x) + O(M^k\, ||f||), \qquad M < |\lambda(s)|, \tag{24}$$

with $\lambda(s), \psi_s(x), e_s[.]$ that depend analytically on s.

When $e_\sigma[f] \neq 0$, the relation (24) leads to approximations by quantities of the form $\lambda(s)^k$, where the relative error terms are exponentially small.

Maximum property (P_3). From the properties above in conjunction with Prop. 1, additional features of the dominant eigenvalue $\lambda(s)$ can be derived for s in or near the real axis. First $\lambda(s)$ is alternatively defined by

$$\lambda(s) = \lim_k \left(\sum_{m_1,\ldots,m_k} \frac{1}{Q_k^s} \right)^{1/k},$$

since the sum on the right is $\mathcal{G}_s^k[1](0)$ to which (23) applies. The smallest continuant Q_k being equal to the $(k+1)$st Fibonacci number, one has

$$\lambda(s+u) \leq \frac{1}{\phi^u}\, \lambda(s), \quad \phi = \frac{1+\sqrt{5}}{2}, \tag{25}$$

so that $\lambda(s)$ strictly decreases along the real axis, $s > 1$. In addition, Faivre [8] and others have proved, by a judicious use of the triangular inequality, that for any $\sigma > 1$, one has $\rho(\sigma + it) < \rho(\sigma)$ for $t \neq 0$.

Special properties (P_4). At $s = 2$, the dominant spectral quantities are explicit:

$$\lambda(2) = 1, \qquad \psi_2(x) = \frac{1}{\log 2}\frac{1}{1+x}, \qquad e_2[f] = \int_0^1 f(x)\,dx. \tag{26}$$

The fact that $\mathcal{G}_s\psi_2 = \psi_2$ constitutes Gauss's observation of the invariance of the density $\psi_2(x)$ under the continued fraction transformation. The corresponding projector e_2 of (22) is also explicit. This results from the property of $\mathcal{G}_2$ to preserve probability distributions in conjunction with Prop. 3: for $w(x) > 0$ when x is real,

$$\int_0^1 \mathcal{G}_2^k[w](t)\,dt = \int_0^1 w(t)\,dt = e_2[w]\int_0^1 \psi_2(t)\,dt + O(\mu(2)^k),$$

from which $e_2[w]$ results.

Wirsing [34] first observed that the deflated operator $-\mathcal{N}_2 = -\mathcal{G}_2 + e_2\psi_2$ is itself positive and used the Perron-Frobenius theory to prove that its dominant eigenvalue $-\mu(2) \approx 0.30366$ is real and simple. This constant is known as the Gauss-Kuzmin-Wirsing constant.

Most objects relative to spectral properties of $\mathcal{G}_s$ do not seem to relate to classical functions and constants. A notable exception is the dominant properties of $\mathcal{G}_s$ in Eq. (26) as well as the related identity

$$\lambda'(2) = -\frac{\pi^2}{12\log 2}, \tag{27}$$

that defines Lévy's constant. The proof of (27) starts from the relation $\mathcal{G}_s[\psi_s] = \lambda(s)\psi_s$, which upon differentiation with respect to s yields at $s = 2$,

$$(I - \mathcal{G}_2)\psi_2' = \mathcal{G}_2'\psi_2 - \lambda'(2)\psi_2, \tag{28}$$

with $\psi_s' = \frac{d}{ds}\psi_s$, etc. In (28), the l.h.s. has a projection that equals 0, so that applying e_2, which means integrating between 0 and 1, one determines $\lambda'(2)$.

$$\int_0^1 \mathcal{G}_2'[\psi_2](t)\,dt = \lambda'(2)\int_0^1 \psi_2(t)\,dt.$$

Note 1. In fact, more is known about the complete spectrum. By Grothendieck's theory of nuclear spaces and operators, the generalized traces corresponding to the set $\{\lambda_j(s)\}$ of eigenvalues and defined by $\sum_j \lambda_j(s)^\varepsilon$ exist for all $\varepsilon > 0$. In addition, Babenko [1] and Mayer [25, 22] have shown the existence of a hidden Hilbert space structure (via Bessel functions and Hankel transforms) for s real where the operator is selfadjoint. Thus, the eigenvalues $\lambda_j(s)$ are all real when $s > 1$. These facts, together with the associated trace formulæ have consequences regarding the numerical determination of the spectrum of $\mathcal{G}_s$, see [5, 22]. They also entail in a few cases full expansions or at least improved error terms (compare for instance Thm. 5 and [5, 33]).

3 The law of continuants

This section is devoted to the analysis of the denominator of the kth convergent of the continued fraction representation of a real number x. For an irrational number, the Q_k grow at least exponentially. The rate of growth together with its possible deviations from the mean value are basic to the understanding of the continued fraction process.

By a slight abuse of notations, we set

$$Q_k(x) := Q_k(m_1(x), m_2(x), \ldots, m_k(x)).$$

Fig. 3 gives the values of $\frac{1}{k}\log Q_k(x)$ for $x \in \{\phi, \log 2, 2^{1/3}, \pi, \gamma, e^{\pi\sqrt{163}}, e\}$ and for values of k till 1000. The constants $\phi = (1+\sqrt{5})/2$ and $e = \exp(1)$ are clearly anomalous, but they are known to have an explicit continued fraction representation,

$$\phi = 1 + /1, 1, 1, 1, 1, 1, \ldots/, \qquad e = 2 + /1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8 \ldots/,$$

so that

$$\frac{1}{k}\log Q_k(\phi) \to \log\phi \approx 0.481, \qquad \frac{1}{k}\log Q_k(e) \to \infty.$$

The other five numbers exhibit a certain tendency of $\frac{1}{k}\log Q_k$ to converge within an interval like $[1.15, 1.22]$ —with a mean for $k = 1,000$ that equals 1.18614— while the theory to be developed below precisely predicts convergence towards

$$-\lambda'(2) = \frac{\pi^2}{12\log 2} \approx 1.18656$$

with overwhelming probability.

The quantity $Q_k(x)$ is defined almost everywhere; it is also a step function that is constant on any fundamental interval of depth k. Here we regard it as a random variable (also written Q_k) and examine characteristics of its distribution when x is uniformly distributed over the interval $[0, 1[$.

Philipp [28] recognized that the distribution of $\log Q_k$ obeys in the limit a Gaussian law. Here we show how this result follows rather directly from the operator approach. In passing, we derive optimal error terms of an order of $O(1/\sqrt{k})$, while Philipp obtained $O(k^{-1/5})$, later improved by Mischyavichyus [26] to $O(\log k/\sqrt{k})$.

$x\backslash k$	10	20	50	100	200	500	1000
ϕ	0.4007	0.4409	0.4651	0.4731	0.4771	0.4796	0.4804
$\log 2$	0.7407	0.7490	0.9448	0.9675	1.0896	1.1428	1.1571
$2^{1/3}$	0.8368	1.0396	1.2295	1.2728	1.2378	1.1770	1.1672
π	1.2807	1.1139	1.2764	1.1772	1.1851	1.1718	1.1774
γ	0.8567	1.0136	1.2229	1.0747	1.1703	1.1689	1.2055
$e^{\pi\sqrt{163}}$	3.4317	2.2099	1.7377	1.4099	1.3047	1.2440	1.2235
e	0.6284	0.8078	1.0914	1.3219	1.5372	1.8381	2.0691

Figure 3: The quantity $\frac{1}{k}\log Q_k(x)$ for various values of x (to be compared to $\pi^2/(12\log 2) \approx 1.186$). The cases of ϕ and e exemplify the two extreme cases.

The quantity $Q_k(x)$ needs first to be related to operators. The value of this continuant is constant and equal to $Q_k(m_1, \ldots, m_k)$ over the interval $h_{\mathbf{m}}([0,1[)$ whose length is $Q_k^{-1}(Q_k + Q_{k-1})^{-1}$, by (21). In particular, the moment generating function of $\log Q_k(x)$,

$$M_k(s) := \mathrm{Ex}\{\exp(s\,\log Q_k(x))\} = \mathrm{Ex}\{Q_k(x)^s\}$$

satisfies

$$M_k(s) = \sum_{m_1 \ldots m_k} Q_k^s \frac{1}{Q_k(Q_k + Q_{k-1})} = \sum_{m_1 \ldots m_k} \frac{1}{Q_k^{2-s}} \frac{1}{1 + \frac{Q_{k-1}}{Q_k}}.$$

By Prop. 1, this is in turn expressible in terms of the transfer operators:

$$M_k(x) = \mathcal{G}_{2-s}^k \left[\frac{1}{1+u}\right](0) \tag{29}$$

Now, by the perturbation properties (Prop. 3) of the transfer operators, there exists a sufficiently small complex neighbourhood of $s = 2$ where

$$M_k(s) = \exp\left(k \log \lambda(2-s) + V(s)\right) \cdot \left(1 + O(\alpha^k)\right), \tag{30}$$

with

$$V(s) = \log\left(e_{2-s}[\frac{1}{1+u}]\,\psi_{2-s}(0)\right)$$

that is analytic near $s = 2$, and α any number satisfying $|\mu(2)| < \alpha < 1$, with $\mu(2)$ the Gauss-Kuzmin-Wirsing constant.

Equation (30) means that $M_k(s)$ behaves nearly like a large power (kth power) of the fixed function $\lambda(s)$. The *central limit theorem* of probability theory asserts that such large powers —in the "pure" power case $V = \alpha = 0$ at least— induce Gaussian laws in the asymptotic limit. There are two differences here: one is the analytic factor $e^{V(s)}$; the other corresponds to the error term $O(\alpha^k)$ which is negligible in the scale of the problem.

The extension of the central limit theorem to "quasi-powers" of the form (30) has been done in a general setting by Hwang [14]. Hwang's technology is based on the Berry-Essen inequality.

Theorem (Hwang's quasipower theorem) *Let Z_k be a sequence of random variables whose moment generating functions admit the asymptotic estimate*

$$M_k(s) := Ex\{\exp(s\,Z_k)\} = \exp(kU(s) + V(s))\left(1 + O(\frac{1}{W_k})\right), \qquad W_k \to \infty,$$

the error term being uniform for s in a disc $|s| \le s_0$ for some $s_0 > 0$. Assume that $U(s)$ and $V(s)$ are analytic for $|s| \le s_0$ and $U(s)$ satisfies the "second moment condition" $U''(0) \neq 0$. Then, the distribution of Z_k is asymptotically Gaussian:

$$\Pr\{\frac{Z_k - kU'(0)}{\sqrt{kU''(0)}} < t\} = \Phi(t) + O(\frac{1}{S(k)}) \quad \textit{where} \quad \Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-w^2/2}\,dw$$

uniformly for all x in R, as k tends to ∞, with $S(k) = \min(\sqrt{k}, W_k)$.

Under these strong analyticity conditions, the mean and variance of Z_k are obtained by differentiation of the asymptotic form of the moment (or the cumulant) generating functions:

$$E\{Z_k\} = kU'(0) + V'(0) + O(\frac{1}{W_k}), \quad var\{Z_k\} = kU''(0) + V''(0) + O(\frac{1}{W_k}).$$

The theorem is applicable to the relation (30), with $U(s) = \log\lambda(2-s)$ that is an analytic function near $s = 0$ and Hensley [13] has established the second moment condition $U''(0) \neq 0$. Thus, we can state:

Theorem 1 *The distribution of the random variable $\log Q_k(x)$ is asymptotically Gaussian,*

$$\Pr\left[\frac{\log Q_k(x) - Ak}{\sqrt{Bk}} < t\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-w^2/2}\,dw + \quad O(\frac{1}{\sqrt{k}}),$$

uniformly for $x \in R$, as $k \to +\infty$. The constants A and B are

$$A = -[\log\lambda(s)]'_{s=2} = -\lambda'(2) \quad \textit{and} \quad B = [\log\lambda(s)]''_{s=2} = \lambda''(2) - \lambda'(2)^2.$$

More precisely,

$$Ex\{\log Q_k(x)\} = Ak + C + O(\alpha^k) \quad \textit{and} \quad var\{\log Q_k\} = Bk + D + O(\alpha^k),$$

where α is any real that is larger than the Wirsing constant $\alpha > |\mu(2)| \approx 0.30366$.

Detailed proofs and generalizations are given in [33]. The speed of convergence, the second order asymptotics of the variance, and the exponentially small error terms improve on earlier known results. In addition, exponential tails result from this approach.

Note 1. This result also determines the *entropy* H of the partition induced by the fundamental intervals that is defined by

$$H := \lim_{k\to\infty} \frac{-1}{k} \sum_{|h|=k} u_h \log u_h = \lim_{k\to\infty} \frac{1}{k} \mathrm{Ex}\{\log\left(Q_k(Q_k + Q_{k-1})\right)\}.$$

The expectation is $2\,\mathrm{Ex}\{\log Q_k\} + O(1)$, so that

$$H = -2\lambda'(2) = \frac{\pi^2}{6\log 2}. \tag{31}$$

4 The analysis of Euclid's algorithm

Heilbronn [11] and Dixon [6] near 1970 have determined the order of growth of the expected number of steps of the Euclidean gcd algorithm. The problem is equivalent to determining the expectation of the depth of the continued fraction associated to a random number p/q with $1 \le p < q < N$; the bound N is made to tend to infinity. Recently, Hensley [13] has succeeded in proving the deep result that the depth is in the asymptotic limit distributed like a Gaussian variable. Here, we establish connections between Euclid's algorithm and transfer operators, show how to rederive a weak form of the Heilbronn-Dixon theorem by operator methods followed by Tauberian arguments, and explain how some of Hensley's results are related to a natural analytic conjecture.

Algebra. We consider here the two sets

$$\begin{array}{lcl} \omega_n & := & \{p \mid 1 \le p \le n,\ \gcd(p,n) = 1\} \\ \Omega_N & := & \{(p,q) \mid 1 \le p \le q \le N,\ \gcd(p,q) = 1\}, \end{array}$$

and the two random variables x_n and X_N that represent the depth of continued fractions associated to rationals of ω_n and Ω_N. The essential quantity is the number $\nu_n^{\langle k\rangle}$ of fractions $p/n \in \omega_n$ with depth exactly k. The cardinality of ω_n is $\phi(n)$, the Euler totient function, so that

$$\sum_{k\ge 1} \nu_n^{\langle k\rangle} = \phi(n), \qquad \Pr\{x_n = k\} = \frac{\nu_n^{\langle k\rangle}}{\phi(n)}, \qquad \mathrm{Ex}\{x_n\} = \frac{1}{\phi(n)} \sum_k k\nu_n^{\langle k\rangle}, \tag{32}$$

with similar expressions for X_N that involve an additional summation over n.

For analysing the arithmetic quantities $\nu_n^{\langle k\rangle}$, we make use of Dirichlet series and introduce

$$A_k(s) := \sum_{n\geq 1} \frac{\nu_n^{\langle k\rangle}}{n^s}, \qquad S(s,u) := \sum_{k\geq 1} u^k A_k(s) = \sum_{n\geq 1} \frac{1}{n^s} \sum_{k\geq 1} u^k \ \nu_n^{\langle k\rangle}. \qquad (33)$$

The function $S(s,u)$ is both a Dirichlet series in s and a power series in u. The algebraic connection between Euclid's algorithm and operators is then summarized by the following lemma.

Lemma 1 *For all $k \geq 1$, one has*

$$A_k(s) = \mathcal{G}_s^{k-1}\,[\zeta_s]\,(0), \qquad S(s,u) = u\ (I - u\,\mathcal{G}_s)^{-1}\ [\zeta_s]\ (0),$$

where ζ_s is a Hurwitz zeta function,

$$\zeta_s(v) := \sum_{m\geq 2} \frac{1}{(m+v)^s}.$$

Proof. The quantity $\nu_n^{\langle k\rangle}$ represents exactly the number of times the integer n appears as a continuant of order k associated to a *proper* continued fraction expansion. Thus, $\nu_n^{\langle k\rangle}$ is the number of k-tuples $(m_1, m_2, \ldots, m_k)$ that satisfy the two conditions,

$$n = Q_k(m_1, m_2, \ldots, m_k), \quad m_k > 1.$$

As a consequence of this observation, one has

$$A_k(s) := \sum_{n\geq 1} \frac{\nu_n^{\langle k\rangle}}{n^s} = \sum_{\substack{m_1\ldots m_k \\ m_k>1}} \frac{1}{Q_k^s} \ = \sum_{m_1\ldots m_k} \frac{1}{Q_k^s} - \sum_{\substack{m_1\ldots, m_k \\ m_k=1}} \frac{1}{Q_k^s}.$$

Now, for $m_k = 1$, we have $Q_k = Q_{k-1} + Q_{k-2}$, so that the second sum in the expression of $A_k(s)$ rewrites as

$$\sum_{m_1\ldots m_{k-1}} \frac{1}{(Q_{k-1}+Q_{k-2})^s} \ = \sum_{m_1\ldots m_{k-1}} \frac{1}{Q_{k-1}^s(1+\frac{Q_{k-2}}{Q_{k-1}})^s}.$$

From Prop. 1, $A_k(s)$ is expressible as a kth iterate of the $\mathcal{G}_s$ operator,

$$A_k(s) = \mathcal{G}_s^k[1](0) - \mathcal{G}_s^{k-1}[\frac{1}{(1+v)^s}](0) = \mathcal{G}_s^{k-1}\,[\zeta_s]\,(0),$$

where ζ_s is the Hurwitz zeta function. The expression for the bivariate generating function $S(s,u)$ then results directly from the form of $A_k(s)$. $\square$

Analysis. The quantity $S(s,1)$ is by Lemma 1 a Dirichlet series of $\phi(n)$:

$$S(s,1) = (I - \mathcal{G}_s)^{-1}\,[\zeta_s]\,(0) = \frac{\zeta(s-1)}{\zeta(s)} - 1. \tag{34}$$

The expectations of x_n and X_N are related to the derivative of $S(s,u)$ with respect to u taken at $u=1$, so that we set

$$T(s) := \left.\frac{\partial S(s,u)}{\partial u}\right|_{u=1}.$$

With $T_n = [n^{-s}]\,T(s)$, the coefficient of n^{-s} in $T(s)$, we have

$$\mathrm{Ex}\{X_N\} = \frac{\sum_{n\le N} T_n}{\sum_{n\le N}\phi(n)},$$

By Lemma 1, $T(s)$ satisfies

$$T(s) = (I - \mathcal{G}_s)^{-2}\,[\zeta_s]\,(0). \tag{35}$$

On the punctured halfplane $\mathcal{H} := \{s \mid \Re(s) \ge 2, s \ne 2\}$, the spectral radius of $\mathcal{G}_s$ is strictly less than 1. The operator $I - \mathcal{G}_s$ is thus invertible, and it depends analytically on s. Therefore, the function $T(s)$ is analytic on $\mathcal{H}$.

We now examine $T(s)$ in a sufficiently small neighbourhood V of $\sigma = 2$. There, Prop. 3 applies and we have

$$\mathcal{G}_s^k[f](0) = d(s)\lambda(s)^k \; + \mathcal{N}_s^k[f](0),$$

with $d(s) = \psi_s(0)\,e_s[\zeta_s]$ that is analytic in V. In $V \cap \mathcal{P}$, the spectral radius of $\mathcal{G}_s$ is strictly less than 1, while in the whole of V the spectral radius of $\mathcal{N}_s$ is strictly less than some $M < 1$, by Prop. 3. We thus have, for $s \in V \cap \mathcal{P}$,

$$\begin{aligned} S(s,1) &= \frac{d(s)}{1-\lambda(s)} + R_1(s) \\ T(s) &= \frac{d(s)}{(1-\lambda(s))^2} + R_2(s), \end{aligned} \tag{36}$$

where

$$R_1(s) = (I - \mathcal{N}_s)^{-1}[\zeta_s](0), \qquad R_2(s) = (I - \mathcal{N}_s)^{-2}[\zeta_s](0)$$

are analytic in V. Since $d(s), \lambda(s)$ are analytic at $s = 2$, the relations (36) provide the analytic continuation of $S(s,1), T(s)$ around $s = 2$ and show that these two functions have a pole there (simple and a double respectively).

This is a typical situation where analytic information on a Dirichlet series can be transferred to asymptotic information on its coefficients. Here the transfer is effected by Delange's Tauberian theorem [32]:

Theorem (Delange's Tauberian theorem) *Let $F(s)$ be a Dirichlet series with nonnegative coefficients a_n such that $F(s)$ converges for $\Re(s) > \sigma > 0$. Assume that $F(s)$ is analytic on $\Re(s) = \sigma$, $\neq \sigma$ and that for some $w \geq 0$,*

$$F(s) = \frac{g(s)}{(s-\sigma)^{w+1}} + h(s),$$

where g, h are analytic at σ, with $g(\sigma) \neq 0$. Then, as $N \to \infty$,

$$\sum_{n \leq N} a_n = \frac{g(\sigma)}{\sigma \Gamma(w+1)} \; N^\sigma \log^w N \quad [1 + \varepsilon(N)] \qquad \varepsilon(N) \to 0.$$

By (34), the Tauberian theorem applies to $T(s)$ for which

$$T(s) \sim \frac{d(2)}{\lambda'(2)^2} \frac{1}{(s-2)^2} \qquad (s \to 2).$$

The quantity $\lambda'(2)$ is known from Section 2, and $d(2) = 1/(2 \log 2)$ results from the explicit form of the projector e_2 in (26).

Theorem 2 (Heilbronn-Dixon) *The mean number of steps of the Euclidean algorithm applied to the set Ω_N satisfies*

$$Ex\{X_N\} = \frac{-1}{\lambda'(2)} \log N \, (1 + o(1)) \sim \frac{12 \log 2}{\pi^2} \log N.$$

Note 1. Eq. (36) leads to a curious observation that connects spectral properties of $\mathcal{G}_s$ to the Riemann hypothesis. Let β be a non trivial zero of the zeta function, $0 < \Re(\beta) < 1$. Then, the right hand side of (34) has a simple pole at $s = 2$ and poles at points β. It is otherwise known that the operator $\mathcal{G}_s$ admits a meromorphic continuation to the whole of the complex plane. Therefore, Eq. (34) implies that *the points where the (continued) operator $\mathcal{G}_s$ admits* 1 *as an eigenvalue include all the nontrivial zeros of the Riemann zeta function.* Actually, a great deal more is known, see Efrat's paper [7].

Note 2. A similar reasoning applies to the bivariate generating function $S(s, e^t)$ and shows that the moment generating function of X_N satisfies

$$\mathrm{Ex}\{\exp(tX_N)\} = B(t) N^{\xi(t)-2}(1 + \varepsilon_t(N)).$$

There $\xi(t)$ is defined for t near 0 as the root near 2 of $\lambda(\xi(t)) = e^{-t}$, and $B(t)$, which is defined in terms of projections, is an analytic function of t. The Tauberian technology only shows that, *pointwise* for each t, one has $\varepsilon_t(N) \to 0$. We conjecture that in fact the convergence of the error terms to 0 is *uniform*. Assuming this conjecture, Hwang's theorem is applicable and this would imply a particularly simple proof of Hensley's theorem [13]:

The number of iterations of the Euclidean gcd algorithm, represented by the variable X_N, obeys asymptotically a Gaussian law.

A possible path to a completely analytic proof of Hensley's theorem would be to show that $\lambda(s)$ does not come "too close" to 1 "too often", so that $S(s,u)$ has a pole-free region extending that of $S(s,1)$. Hensley has also determined the variance of X_N to be of the form

$$\chi \log N\,(1+o(1)), \qquad \text{with} \qquad \chi = -\frac{\lambda''(2)-\lambda'(2)^2}{\lambda'(2)^3}.$$

If true, our conjecture would imply further connections between $\mathcal{G}$ operators and Porter's constant [17], as well as the fact that the variance is of the form $\chi \log N + \widehat{\chi} + o(1)$, for some Porter-like constant $\widehat{\chi}$.

Note 3. The Tauberian approach also provides density estimates for rational numbers in Ω_N whose continued fraction quotients all lie in a fixed set Q, under the asymptotic form $C(Q)\cdot N^{\xi(Q)}$. The exponent ξ there is known to be related both to spectral properties of transfer operators and to Hausdorff dimension [12].

5 The sign algorithm

It is easy to compare two real numbers x and y, or equivalently determine the sign of $(x-y)$, by a modification of the continued fraction algorithm. The principle of the method goes back to Gosper around 1972 who described it in the celebrated "Hakmem" memorandum of 1972 (item 101 of [3]):

> "Numerical comparison of continued fractions is slightly harder than in decimal, but much easier than with rationals – just invert the decision as to which is larger whenever the first discrepant terms are even-numbered. Contrast this with the problem of comparing the rationals 113/36 and 355/113."

The skeleton of the algorithm is thus:

Algorithm $CF_2(x,y)$ **for** $k := 1$ **while** $m(x) = m(y)$ **do**
{ $x := T(x)$; $y := T(y)$; }
if $k \bmod 2 = 1$ **then** return(sign($m(x)$-$m(y)$)) **else** return(sign($m(y)$-$m(x)$)).

This algorithm works as described whenever x and y are distinct irrational and it is readily supplemented by additional tests in order to correctly deal with all cases, including rational entries where $m(x)$ and/or $m(y)$ may be 0.

Analysis of CF_2 ***over the reals.*** We consider a pair (x,y) of reals uniformly distributed over the unit square $\mathcal{Q}=[0,1]^2$, and let L be the random variable

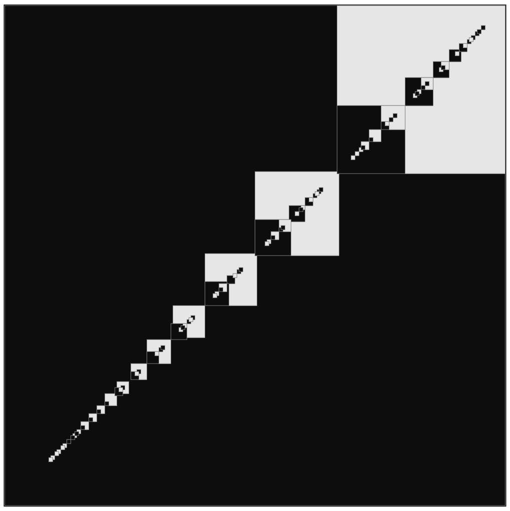

Figure 4: The fundamental domains $\{L = k\} \cap [0, \frac{1}{2}]^2$ of the sign algorithm CF_2.

that represents the number of iterations of the algorithm (the value of k upon termination of the algorithm). Then, the event $\{L \geq k+1\}$ is formed of all pairs (x, y) whose continued fraction representations coincide till depth k. In that case, x and y belong to a common fundamental interval of depth k. Thus, up to sets of measure 0, one has

$$\{L \geq k+1\} = \bigcup_{|h|=k} h(\mathcal{I}) \times h(\mathcal{I}),$$

and, with u_h the length of the fundamental interval of h,

$$\rho_k := \Pr\{L \geq k+1\} = \sum_{|h|=k} (u_h)^2 = \sum_{m_1,\ldots,m_k} \frac{1}{Q_k^2(Q_k + Q_{k-1})^2}, \tag{37}$$

by (15). The quantity ρ_k therefore represents the probability that two random real numbers have continued fraction representations that coincide till depth k.

By Prop. 1, the probability distribution $\{\rho_k\}$ is then expressible in terms of the iterates of $\mathcal{G}_4$:

$$\rho_k = \mathcal{G}_4^k \left[\frac{1}{(1+x)^2}\right](0).$$

The spectral properties summarized by Prop. 3 now apply and, as a consequence, the ρ_k decrease geometrically:

$$\rho_k = C \cdot \lambda(4)^k (1 + O(\alpha^k)), \tag{38}$$

for some real $C > 0$ and some constant α with $0 < \alpha < 1$ (one may take any $\alpha > |\mu(4)|/\lambda(4)$).

The expectation of L satisfies

$$\mathrm{Ex}\{L\} \equiv \sum_{k \geq 0} \rho_k = (I - \mathcal{G}_4)^{-1} \left[\frac{1}{(1+x)^2} \right] (0). \tag{39}$$

An elementary argument now shows that the operator $\mathcal{G}_4$ in the expression of the expectation can be dispensed with. Each rational number of $[0, 1]$ admits two continued fraction representations: a proper one where the last quotient satisfies $m_k \geq 2$ and an improper one where it satisfies $m_{k+1} = 1$. Thus, each rational number c/d, with $\gcd(c, d) = 1$, $d \geq 2$ and $0 < c < d$, admits two representations as (P_k/Q_k) or equivalently as (Q_{k-1}/Q_k) given the symmetry property (17) of continuants. Taking care of the special cases 0 and 1, one gets

$$\mathrm{Ex}\{L\} = \frac{5}{4} + 2 \sum_{c,d} \frac{1}{c^2 (c+d)^2}, \tag{40}$$

where the sum is over all pairs (c, d) with $\gcd(c, d) = 1$, $d \geq 2$ and $0 < c < d$. The gcd condition is then easily eliminated by multiplying the sum by $\zeta(4) = \sum_{n \geq 1} n^{-4} = \pi^4/90$. In summary, we have proved:

Theorem 3 *Under the uniform distribution over the unit square, the number of iterations of the continued fraction based sign algorithm CF_2 has a probability distribution that satisfies*

$$\rho_k := \Pr\{L \geq k+1\} = C \cdot \lambda(4)^k (1 + O(\alpha^k)),$$

for some real $C > 0$ and some constant α with $0 < \alpha < 1$, with $\lambda(4) \approx 0.1994$ the dominant eigenvalue of the $\mathcal{G}_4$ operator. The expected number of iterations is

$$Ex\{L\} = \frac{3}{4} + \frac{180}{\pi^4} \sum_{d=1}^{\infty} \sum_{c=d+1}^{2d} \frac{1}{c^2 d^2} \approx 1.35113$$

Numerical procedures for estimating $\lambda(4)$ have been described in [5]. There, it is also shown, following Sitaramachandrarao [31], that the expectation is expressible in terms of $\zeta(3)$ and the tetralogarithm $\mathrm{Li}_4(1/2) = \sum_{n \geq 1} 2^{-n} n^{-4}$.

The CF_2 algorithm applied to real numbers thus has features roughly comparable to the binary expansion algorithm discussed in Section 1. The quantity $\lambda(4) = \lim \rho_k^{1/k}$ dictates the asymptotic rate at which digits in two continued fraction expansions coincide.

Note 1. The analysis generalizes directly to *nonuniform* densities over the unit square that are proportional to $|x - y|^r$ for some real parameter $r > -1$. The case $r = 0$ is the uniform model; the cases $-1 < r < 0$ correspond to giving a heavier weight to the "difficult" input configurations of the algorithm. Statements similar to Theorem 3 then hold true but with the operator $\mathcal{G}_{4+2r}$ replacing $\mathcal{G}_4$. This has been studied systematically by Vallée [33] who observed that, as r tends to -1, the behaviour of the algorithm resembles that of the 1–dimensional continued fraction algorithm. Vallée introduced in fact a family of 2–dimensional operators that unifies all these analyses and leads to a powerful generalization of Theorems 1,3 when the initial distribution of x is nonuniform.

Note 2. The geometry of fundamental domains for the sign algorithm is isomorphic to the geometry of the 2-dimensional lattice reduction for which Gauss had given an optimal algorithm. Thus, the analysis conducted here parallels that of our paper [5] where the alternative expression for $\mathrm{Ex}\{L\}$ and the precise numerical evaluation of constants are discussed in detail.

Note 3. As is clear from the HAKMEM quotation, the CF_2 algorithm was conceived initially as an economical way of comparing two rationals. In this context, its interest stems from the fact that the sizes of the rational numbers x and y steadily decrease in the course of the algorithm, so that no multiprecision operation is required. In a way, algorithm CF_2 then corresponds to running in parallel two "lazy" versions of the gcd algorithm, one on (a, c), the other on (b, d), if $x = a/c$ and $y = b/d$. Execution is halted as soon as a discrepancy of quotients is detected. The same algorithm may be used to estimate the sign of 2×2-determinants with integer entries and again no multiprecision manipulation is required, a feature of particular interest in computational geometry.

It is easily proved that the CF_2 applied to random rationals with denominators in the range $[N, 2N]$ has, up to error terms that are $O(1/N)$, a behaviour described by the continuous model and Theorem 3. Thus, although the worst-case cost of CF_2 under this model is

$$\frac{\log N}{\log \phi} + O(1), \qquad \phi = \frac{1 + \sqrt{5}}{2},$$

the average case is constant and the probability of executing a large number of steps decays exponentially —for instance ρ_5 is of the order of 10^{-3}, ρ_{10} of the order of 10^{-7}.

6 The sorting algorithm

In this section[2], we consider the problem of sorting n real numbers given their continued fraction representations. The algorithm proceeds by elementary com-

[2] We present here a preliminary report of results, some of which have been obtained jointly with Julien Clément.

parisons between continued fraction "digits" (*i.e.*, quotients) and it is a generalization of the sign algorithm of Section 5. Its underlying structure is a tree, called a *digital tree* or *trie* [17, 21]. Though motivated by algorithmic considerations, the analysis developed here has also a direct interpretation in terms of continued fractions: given n uniformly independently drawn random real numbers of $[0, 1]$, it answers probabilistic questions like:

— what is the "closest pair" of numbers, where distance is measured by the number of coincident continued fraction digits?

— how many digits in total must be determined in order to distinguish, hence sort, the n numbers?

— how many fundamental intervals are necessary to separate the n numbers?

These parameters have a natural formulation in terms of the associated digital tree structure. The structural constants $-2\lambda'(2)$ (the entropy) and $\lambda(4)$ (the rate of coincident digits) play an essential rôle.

Digital trees. Let $A \subseteq N$ be a set of elements called digits, and A^∞ the set of all infinite sequences built over A,

$$A^\infty = \{\alpha = (a_1, a_2, \ldots) \mid a_j \in A\}.$$

For $\alpha \in A^\infty$, the head and tail functions are defined by

$$\mathrm{head}(\alpha) = a_1, \qquad \mathrm{tail}(\alpha) = (a_2, a_3, \ldots).$$

To any finite set X of elements of A^∞, one associates a *digital tree*, $\mathrm{tree}(x)$, defined by the following recursive rules:

(R_1) If $X = \{\alpha\}$ has cardinality equal to 1, then $\mathrm{tree}(X)$ consists of a single *leaf node* that contains α.

(R_2) If X has cardinality at least 2, then $\mathrm{tree}(X)$ is an *internal node* represented generically by 'o' to which are attached r subtrees, where $r = \mathrm{card}\{\mathrm{head}(\alpha),\ \alpha \in X\}$ is the number of different head symbols in X. Let $b_1 < b_2 < \cdots < b_r$ be these different head symbols; $\mathrm{tree}(x)$ is defined by

$$\mathrm{tree}(X) = \langle o, \mathrm{tree}(X_1), \mathrm{tree}(X_2), \ldots, \mathrm{tree}(X_r)\rangle,$$

where

$$X_i = \{\mathrm{tail}(\alpha) \mid \mathrm{head}(\alpha) = b_i,\ \alpha \in X\}.$$

Such a tree structure underlies classical radix sorting methods. The tree can be built by following the recursive rules R_1, R_2 and a traversal of leaves gives rise to an algorithm that sorts any set X lexicographically.

$$
\begin{array}{lclcl}
\mathrm{a} & = & \phi - 1 & = & /\mathbf{1,\ 1,\ 1},\ \mathit{1,\ 1,\ 1,\ldots}/ \\
\mathrm{b} & = & \gamma & = & /\mathbf{1,\ 1,\ 2},\ \mathit{1,\ 2,\ 1,\ldots}/ \\
\mathrm{c} & = & \exp(1) - 2 & = & /\mathbf{1,\ 2,\ 1},\ \mathit{1,\ 4,\ 1,\ldots}/ \\
\mathrm{d} & = & \log 2 & = & /\mathbf{1,\ 2,\ 3},\ \mathit{1,\ 6,\ 3,\ldots}/ \\
\mathrm{e} & = & \{\exp(\pi\sqrt{163})\} & = & /\mathbf{1,\ 1333462407511},\ \mathit{1,\ 8,\ 1,\ 1,\ \ldots}/ \\
\mathrm{f} & = & 2^{1/3} - 1 & = & /\mathbf{3},\ \mathit{1,\ 5,\ 1,\ 1,\ ,\ 4,\ldots}/ \\
\mathrm{g} & = & \pi - 3 & = & /\mathbf{7},\ \mathit{15,\ 1,\ 292,\ 1,\ 1,\ldots}/
\end{array}
$$

Figure 5: The trie corresponding to the fractional parts of the 7 quantities of Fig. 2 has size 4, height 3, and path length 16. The digits in bold represent the set of 16 digits (the path length) needed to distinguish globally between the 7 numbers; the maximum number of digits needed to compare two elements is 3 (the height); the number of separation intervals is $4 + 7 = 11$.

The *level* of a node in a digital tree is the number of edges that connect it to the root. The *height* of the tree is the maximum level of any leaf, the *path length* of the tree is the sum of the levels of all leaves and the *size* of the tree is the number of its internal nodes.

From now on, we specialize the discussion to digital trees built on digit sequences of continued fraction representations of real numbers, so that $A = \{1, 2, \ldots\}$. Such trees are called CF–*trees*, and suitable traversals of a CF-tree sort X either lexicographically or according to the natural order over the reals. The height of tree(X) is then a measure of distance between the two closest elements of X (in terms of continued fraction expansions). The path length equals the total number of digits that need to be examined in order to distinguish all elements of X (or even sort X), and the size of the tree plus the cardinality of X is the number of fundamental intervals necessary to isolate all elements of X.

Rather than fixing the cardinality n of the set X (this model is called the Bernoulli model), we consider a variable number N of elements that obeys a Poisson law of parameter x,

$$\Pr\{N = k\} = e^{-x}\frac{x^k}{k!}.$$

This model is called the Poisson model of rate x. In this model, N is strongly concentrated near its mean x with a high probability so that the rate x plays a rôle much similar to size. It can be proved in fact that the analyses under the Bernoulli model of index n coincide, to first asymptotic order, with the corresponding ones under a Poisson model of rate $x = n$.

The interest of the Poisson model is that there is complete independence of what happens in disjoint subintervals of $[0, 1]$. In particular, the number of elements that fall into any interval of length q is itself distributed as a Poisson variable of rate qx.

Algebra. The strong independence property of the Poisson model gives access to the Poisson generating functions of basic parameters. The rôle of fundamental intervals is crucial and for a continued fraction LFT h, the length u_h of the associated interval is given by (21).

Consider the Poisson model of rate x. A random CF–tree has height $\leq k$ provided no fundamental interval of depth k contains more than 1 element. The probability of this event, is given by the independence property of the Poisson model,

$$\pi_k(s) = \prod_{|h|=k} (1 + xu_h)e^{-xu_h}, \tag{41}$$

which yields the expected height

$$D(x) = \sum_{k=0}^{\infty} [1 - \pi_k(x)]\,. \tag{42}$$

For size, we observe that an internal node in the tree corresponds to a fundamental interval occupied by at least two numbers. Similarly, path length is equal to the sum over all internal nodes of the number of elements that fall into the associated fundamental interval. The expectations of size and path length result, since

$$\left[1 - (1+\lambda)e^{-\lambda}\right], \qquad \lambda\left[1 - e^{-\lambda}\right], \tag{43}$$

are respectively the probability that a Poisson variable of rate λ has value ≥ 2 and the expectation of such a variable conditioned to be at least 2. In summary:

Lemma 2 *The expectations of height, size, and path length are, under the Poisson model of rate x,*

$$\begin{aligned} D(x) &= \sum_k \left[1 - \prod_{|h|=k} (1+xu_h)e^{-xu_h}\right] \\ S(x) &= \sum_h \left[1 - (1+xu_h)e^{-xu_h}\right] \\ P(x) &= \sum_h xu_h\left[1 - e^{-xu_h}\right], \end{aligned} \tag{44}$$

where the sums and products are indexed by linear fractional transformations h associated to continued fractions, and $u_h = |h(1) - h(0)|$.

Analysis of size and path length. We perform the asymptotic analysis of tree parameters by means of *Mellin transforms.* The Mellin transform of a function $g(x)$ is defined by

$$g^*(s) := \int_0^\infty g(x)x^{s-1}\,dx, \tag{45}$$

and the Mellin transform of e^{-x} is the Euler gamma function $\Gamma(s)$. The Mellin transform is known to map the asymptotic expansions of the original function $g(x)$ at 0 and $+\infty$ to the singularities of the transformed function $g^*(s)$. In addition, it transforms a so-called harmonic sum

$$G(x) = \sum_j c_j g(\omega_j x), \qquad \omega_j > 0, \tag{46}$$

into a factored form,

$$G^*(s) = \Omega(s)\cdot g^*(s), \qquad \Omega(s) = \sum_j c_j(\omega_j)^{-s}. \tag{47}$$

These two properties make it possible to analyse asymptotically a large number of sums that arise in combinatorial theory, see [10] for a detailed review of the method. In particular, the original analysis of digital tries built on standard binary representations has been obtained in this way by De Bruijn and Knuth [17, 21].

The general relations (45–47) apply to $S(x)$ and $P(x)$ whose Mellin transforms are found to be

$$S^*(s) = -\Lambda(-s)(s+1)\Gamma(s), \qquad P^*(s) = -\Lambda(-s)s\Gamma(s). \tag{48}$$

There, the key quantity $\Lambda(s)$ is the *Dirichlet series of fundamental intervals*,

$$\Lambda(s) = \sum_{|h|\geq 0} (u_h)^s = (I - \mathcal{G}_{2s})^{-1}\left[(1+x)^{-s}\right](0), \tag{49}$$

considered now for complex values of s. From (49), $\Lambda(s)$ has a single pole at $s = 1$.

By the principles of Mellin asymptotics, the precise location of singularities of $\Lambda(s)$ is needed. From the argument already used in (40) for the average-case analysis of the sign algorithm, one has

$$\Lambda(s) = 1 - \frac{1}{2^s} + \frac{2}{\zeta(s)}\sum_{q=1}^{\infty}\sum_{p=1}^{q}\frac{1}{q^s(p+q)^s}. \tag{50}$$

Euler-Maclaurin summation applies, resulting in

$$\zeta(2s)(\Lambda(s) - 1 + \frac{1}{2^s}) = 2\sum_{1<q}\frac{1}{q^{2s-1}}\left[\int_1^2\frac{dt}{t^s} + O(\frac{1}{q})\right],$$

for $\Re(s) > 1$. Thus $\Lambda(s)$ can be put under the form

$$\Lambda(s) = 2\frac{\zeta(2s-1)}{\zeta(2s)}\frac{2^{1-s}-1}{1-s} + \frac{R(s)}{\zeta(2s)}, \tag{51}$$

where $R(s)$ involves a series whose general term is $O(q^{-2s})$ that is analytic for $\Re(s) > \frac{1}{2}$. In particular the singular expansion of $\Lambda(s)$ at $s = 1$ reads

$$\Lambda(s) = \frac{6\log 2}{\pi^2}\frac{1}{s-1} + O(1) \qquad (s \to 1), \tag{52}$$

and the function is continuable to the left of $\Re(s) = 1$, till $\Re(s) = 1/2$ at least.

Near $s = -1$, one has $\Gamma(s) = -1/(s+1) + O(1)$. The singular expansions of $S^*(s), P^*(s)$ obtained from (48) and (52) then yield the asymptotic expansions of $S(x), P(x)$ at infinity by standard Mellin technology [10].

Theorem 4 *The expectations of size and path length of CF–tree satisfy, under the Poisson model of rate x,*

$$S(x) = \frac{6\log 2}{\pi^2}x + o(x), \qquad P(x) = \frac{6\log 2}{\pi^2}x\log x + O(x).$$

Analysis of height. The analysis of height is based on estimates of the individual probabilities $\pi_k(x)$ of (41) followed by a Mellin analysis. First, taking logarithms, one gets

$$\log \pi_k(x) = \sum_{|h|=k}\big[-xu_h + \log(1+xu_h)\big]. \tag{53}$$

The largest interval of depth k corresponds to the worst case of Euclid's algorithm and has a length expressible in terms of Fibonacci numbers,

$$\max_{|h|=k}\{u_h\} = \frac{1}{F_{k+1}F_{k+2}} = O(\phi^{-2k}). \tag{54}$$

Thus, the condition

$$k \geq \frac{\log x}{2\log\phi}, \tag{55}$$

which is assumed from now on, guarantees, in the region (55), that

$$\log \pi_k(x) = -\frac{1}{2}x^2\sum_{|h|=k}u_h^2 + O\left(x^3\sum_{|h|=k}u_h^3\right). \tag{56}$$

On the other hand, the transfer operators $\mathcal{G}_s$ give access to the "moments" of fundamental intervals:

$$\begin{aligned}\sum_{|h|=k}u_h^r &= \sum\frac{1}{Q_k^r(Q_k+Q_{k-1})^r}\\ &= \mathcal{G}_{2r}^k\left[(1+x)^{-r}\right](0).\end{aligned} \tag{57}$$

By spectral properties in the style of Prop. 3, there is a constant C_r such that for all $\varepsilon > 0$, one has

$$\sum_{|h|=k} u_h^r = C_r \, \lambda(2r)^k + O((|\mu(2r)| + \varepsilon)^r). \tag{58}$$

Thus, under the sole condition (55), Eq. (56), (58) imply

$$\log \pi_k(x) = -\frac{c_2 x^2}{2} \lambda(4)^k + O(x^2(|\mu(4)| + \varepsilon)^k + x^3 \lambda(6)^k), \tag{59}$$

where the constants $\lambda(4), \mu(4), \lambda(6)$ are known to great accuracy from [5]:

$$\lambda(4) \approx 0.199458, \qquad \mu(4) \approx -0.075739, \qquad \lambda(6) \approx 0.063402.$$

For the expected height, it proves justified to use the approximation

$$\widetilde{\pi}_k(x) = \exp(-\frac{c_2}{2} x^2 \lambda(4)^k), \tag{60}$$

so that

$$D(x) = \widetilde{D}(x) + o(x), \qquad \widetilde{D}(x) = \sum_{k=0}^{\infty} \left[1 - \widetilde{\pi}_k(x)\right].$$

The Mellin analysis is similar to the case of digital trees built on binary expansions. The transform of the approximation $\widetilde{D}(x)$ in (60) is, for $-2 < \Re(s) < 0$,

$$\widetilde{D}^*(s) = -\frac{1}{2} \left(\frac{c_2}{2}\right)^{-s/2} \frac{\Gamma(s/2)}{1 - \lambda(4)^{-s}}.$$

The double pole at $s = 0$ induces a logarithmic term while the imaginary poles at $s = 2ik\pi / \log \lambda(4)$ contribute a periodic term.

Theorem 5 *The expected height of a CF-tree is, under the Poisson model of rate x,*

$$D(x) = \frac{2}{|\log \lambda(4)|} \log x + P(x) + o(1),$$

with $P(u)$ a continuous periodic function.

Note 1. The results obtained here are to be compared to the classical results on digital trees built on a finite set of digits $A = \{1, 2, \ldots, r\}$ where digit i has probability p_i. In that case, the expected size and path length under the Poisson model of rate x are approximated by

$$\frac{x}{H}, \qquad \frac{x \log x}{H}, \qquad \text{with} \qquad H = \sum_{i=1}^{r} p_i \log p_i$$

being the entropy function. The expected height is asymptotic to

$$\frac{2\log x}{\log K}, \qquad \text{with} \qquad K = \sum_{i=1}^{r} p_i^2$$

the probability that two digits coincide. These results parallel those of Theorems 4, 5, with $-2\lambda'(2)$ playing once more the rôle of entropy, and $\lambda(4)$ replacing the probability of coinciding digits, which is in accordance with results of the previous section regarding the sign algorithm CF_2.

Note 2. The fluctuations in the expected height are of a standard form with Fourier coefficients that are values of the gamma function at regularly spaced points of the imaginary axis. In contrast, the fluctuations for size and path length have an asymptotically smaller order that is dictated by the location of nontrivial zeros of the zeta function and hence related to the Riemann hypothesis.

Note 3. Similar estimates hold for the more natural Bernoulli model of index n, where the expectations s_n, p_n, d_n of size, path length, and height are found to satisfy:

$$s_n \sim S(n), \qquad p_n \sim P(n), \qquad d_n \sim D(n) \qquad (n \to +\infty).$$

For size and path length, this results from a Mellin analysis applied to exact forms of Bernoulli expectations by means of a modified Dirichlet series. For height, we relie on saddle point asymptotics, using an extension to complex values of x of the approximation of $D(x)$.

Acknowledgements. This work was supported in part by the Long Term Research Project ALCOM-IT (# 20244) of the European Union.

References

[1] BABENKO, K. I. On a problem of Gauss. *Soviet Mathematical Doklady 19*, 1 (1978), 136–140.

[2] BEDFORD, T., KEANE, M., AND SERIES, C., Eds. *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*. Oxford University Press, 1991.

[3] BEELER, M., GOSPER, R. W., AND SCHROEPPEL, R. HAKMEM. Memorandum 239, M.I.T., Artificial Intelligence Laboratory, Feb. 1972.

[4] BILLINGSLEY, P. *Probability and Measure*, 2nd ed. John Wiley & Sons, 1986.

[5] DAUDÉ, H., FLAJOLET, P., AND VALLÉE, B. An average-case analysis of the Gaussian algorithm for lattice reduction. To appear in *Combinatorics, Probability and Computing*, Oct. 1995. 30 pages.

[6] DIXON, J. D. The number of steps in the Euclidean algorithm. *Journal of Number Theory 2* (1970), 414–422.

[7] EFRAT, I. Dynamics of the continued fraction map and the spectral theory of $SL(2, Z)$. *Inventiones Mathematicae 114* (1993), 207–218.

[8] Faivre, C. Distribution of Lévy's constants for quadratic numbers. *Acta Arithmetica 61*, 1 (1992), 13–34.

[9] Finch, S. Favorite mathematical constants. Available under World Wide Web at `http: //www.mathsoft.com/asolve/constant/constant.html`, 1995.

[10] Flajolet, P., Gourdon, X., and Dumas, P. Mellin transforms and asymptotics : Harmonic sums. *Theoretical Computer Science 144*, 1–2 (June 1995), 3–58.

[11] Heilbronn, H. On the average length of a class of continued fractions. In *Number Theory and Analysis* (New York, 1969), P. Turan, Ed., Plenum Press, pp. 87–96.

[12] Hensley, D. The Hausdorff dimensions of some continued fraction Cantor sets. *Journal of Number Theory 33* (1989), 182–198.

[13] Hensley, D. The number of steps in the Euclidean algorithm. *Journal of Number Theory 49*, 2 (1994), 142–182.

[14] Hwang, H.-K. *Théorèmes limites pour les structures combinatoires et les fonctions arithmetiques.* PhD thesis, École Polytechnique, Dec. 1994.

[15] Kato, T. *Perturbation Theory for Linear Operators.* Springer-Verlag, 1980.

[16] Khinchin, A. I. *Continued Fractions.* University of Chicago Press, Chicago, 1964. A translation of the Russian original published in 1935.

[17] Knuth, D. E. *The Art of Computer Programming*, vol. 3: Sorting and Searching. Addison-Wesley, 1973.

[18] Knuth, D. E. *The Art of Computer Programming*, 2nd ed., vol. 2: Seminumerical Algorithms. Addison-Wesley, 1981.

[19] Krasnoselsky, M. *Positive solutions of operator equations.* P. Noordhoff, Groningen, 1964.

[20] Lorch, E. R. *Spectral Theory.* Oxford University Press, New York, 1962.

[21] Mahmoud, H. *Evolution of Random Search Trees.* John Wiley, New York, 1992.

[22] Mayer, D., and Roepstorff, G. On the relaxation time of Gauss's continued fraction map. I. The Hilbert space approach. *Journal of Statistical Physics 47*, 1/2 (Apr. 1987), 149–171.

[23] Mayer, D., and Roepstorff, G. On the relaxation time of Gauss's continued fraction map. II. The Banach space approach (transfer operator approach). *Journal of Statistical Physics 50*, 1/2 (Jan. 1988), 331–344.

[24] Mayer, D. H. On a ζ function related to the continued fraction transformation. *Bulletin de la Société Mathématique de France 104* (1976), 195–203.

[25] Mayer, D. H. Continued fractions and related transformations. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, T. Bedford, M. Keane, and C. Series, Eds. Oxford University Press, 1991, pp. 175–222.

[26] Mischyavichyus, G. A. Estimate of the remainder in the limit theorem for the denominators of continued fractions. *Litovskiĭ Matematičeskiĭ Sbornik 21*, 3 (1987), 63–74.

[27] Perron, O. *Die Lehre von der Kettenbrüchen*, vol. 1. Teubner, 1954.

[28] Philipp, W. Ein zentraler Grenzwertsatz mit Anwendungen auf die Zahlentheorie. *Zeitschrift für Wahrscheinlichkeitstheorie 8* (1967), 195–203.

[29] Rockett, A., and Szűsz, P. *Continued Fractions.* World Scientific, Singapore, 1992.

[30] Ruelle, D. *Dynamical Zeta Functions for Piecewise Monotone Maps of the Interval*, vol. 4 of *CRM Monograph Series.* American Mathematical Society, Providence, 1994.

[31] Sitaramachandrarao, R. A formula of S. Ramanujan. *Journal of Number Theory 25* (1987), 1–19.

[32] Tenenbaum, G. *Introduction à la théorie analytique des nombres*, vol. 13. Institut Élie Cartan, Nancy, France, 1990.

[33] Vallée, B. Opérateurs de Ruelle-Mayer généralisés et analyse des algorithmes d'Euclide et de Gauss. Rapport de Recherche de l'Université de Caen, Les Cahiers du GREYC # 4, 1995. To appear in *Acta Arithmetica.*

[34] Wirsing, E. On the theorem of Gauss–Kusmin–Lévy and a Frobenius–type theorem for function spaces. *Acta Arithmetica 24* (1974), 507–528.