Continued Fractions from Euclid till Present

Ilan Vardi
IHES, Bures sur Yvette

October 19, 1998

[summary by Philippe Flajolet]

Continued fractions have fascinated mankind for centuries if not millennia. The timeless construction of a rectangle obeying the "divine proportion" (the term is in fact from the Renaissance) and the "self-similarity" properties that go along with it are nothing but geometric counterparts of the continued fraction expansion of the golden ratio,

$$\phi \equiv \frac{1+\sqrt{5}}{2} = \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\cdots}}}.$$

Geometry was developed in India from the rules for the construction of altars. The *Sulva Sūtra* (a part of the *Kalpa Sūtra* hypothesized to have been written around 800 BC) provides a rule[1] for doubling an area that corresponds to the near-equality:

$$\sqrt{2} \doteq 1 + \frac{1}{3} + \frac{1}{3\times 4} - \frac{1}{3\times 4\times 34} \qquad \text{(correct to } 2\,10^{-6}\text{)}. \tag{1}$$

Exclusively for these seminar proceedings, we propose the original observation that the third and fourth partial sums in (1), namely $\frac{17}{12}$ and $\frac{577}{408}$, are respectively the fourth and eighth convergents to $\sqrt{2}$.

Accordingly, in the classical Greek world, there is evidence of knowledge of the continued fraction for $\sqrt{2}$ which appears in the works of Theon of Smyrna (discussed in Fowler's reconstruction [6] and in [17]) and possibly of Plato in *Theaetetus*, see [3]. As every student knows, Euclid's algorithm is a continued fraction expansion algorithm in disguise, and Archimedes' Cattle Problem (*circa* 250 BC) most probably presupposes on the part of its author some amount of understanding of quadratic irrationals, Pell's equation, and continued fractions; see [17] for a discussion.

The continued fraction convergent $\pi \approx \frac{355}{113}$ was known to Tsu Ch'ung Chi born in Fan-yang, China in 430 AD. More recently, the Swiss mathematician Lambert proved the 2,000 year conjecture (it already appears in Aristotle) that π is irrational, this thanks to the continued fraction expansion of the tangent function,

$$\tan z = \cfrac{z}{1-\cfrac{z^2}{3-\cfrac{z^2}{5-\cdots}}},$$

[1] "Increase the measure by its third part, and this third part by its own fourth, less the thirty-fourth part of that fourth". See vol. I of Dutt's book [4, p. 272] for context including otherwise rational approximations to $\sqrt{\pi/4}$.

and Apéry in 1979 gave in "a proof that Euler missed" [12] nonstandard expansions like

$$\zeta(3) \equiv \sum_{n=1}^{\infty} \frac{1}{n^3} = 1 + \cfrac{1}{2\cdot 2 + \cfrac{1^3}{1 + \cfrac{1^3}{2\cdot 6 + \cfrac{2^3}{1 + \cfrac{2^3}{2\cdot 10 + \cfrac{3^3}{1+\cdots}}}}}},$$

from which the irrationality of $\zeta(3)$ eventually derives.

Earlier, Euler had estimated that

$$\sum_{n=0}^{\infty} (-1)^n n! = 0.596\,347\,362\ldots$$

by means of a formal continued fraction expansion of the series $\sum_n n!(-z)^n$. A famous memoir of Stieltjes contains many other series identities like

$$\int_0^{\infty} \tanh u\, e^{-zu}\, du = \cfrac{1}{z^2 + \cfrac{1\cdot 2}{1 + \cfrac{2\cdot 3}{z^2 + \cfrac{3\cdot 4}{1 + \cfrac{4\cdot 5}{z^2+\cdots}}}}},$$

which, a century later, enabled the author of this summary to provide the first analyses in a dynamic context of some basic data structures of computer science. The fascination continues and in 1985, Gosper found 17 million continued fraction digits of π, which corresponds to about 18 million decimal digits. Such quests are by the way not totally senseless: for instance, from similar data, we can know for sure that, should Euler's constant γ be rational, then its numerator and denominator would have at least 242,080 decimal digits (Papanikolaou, 1997)!

1. **Arithmetica**

This brilliant and erudite talk (see especially [14, 15]) discusses the many facets of standard (or "regular") continued fraction expansions of real numbers, which will be written occasionally as

$$x = [a_0, a_1, \ldots] \quad \text{if} \quad x = \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cdots}}.$$

One refers to the a_j as continued fraction digits or partial quotients (thinking of Euclid's algorithm).

The best approximant properties of continued fractions are well-known [7]. Indeed, any convergent p/q (obtained by a finite truncation) of the continued fraction expansion of α approximates α better than any other fraction with a smaller denominator and better even than many with a larger denominator (compare π with $\frac{16}{5}, \frac{314}{100}$ against $\frac{22}{7}$).

A fruitful consequence of the best approximation property is the following: let D be a non-square and let p/q be a "good" approximant of $\sqrt{D}$. Then, $p^2 - q^2 D$ is small. In particular, as discovered by Lagrange, the continued fraction expansion of $\sqrt{D}$ hides a solution to Pell's equation,

$p^2 - Dq^2 = 1$. The solutions are essentially powers of a minimal solution that corresponds to a so-called fundamental unit of $\mathbb{Q}(\sqrt{D})$. Archimedes' Cattle Problem can be solved in this perspective, the full solution involving numbers with 206,545 digits! In an entertaining note [17], Vardi even shows that the solution can be enounced with only a few symbols: the total number of cattle is

$$\left\lceil \frac{25194541}{184119152}(109931986732829734979866232821433543901088049 + 50549485234315033074477819735540408986340\sqrt{4729494})^{4658} \right\rceil .$$

Continued fraction algorithms and the Euclidean algorithm are central to many other areas of number theory. For instance, Rademacher observed that the Jacobi symbol $\left(\frac{d}{c}\right)$ (that tells us in essence whether d is a square modulo c) is expressible as

$$\left(\frac{d}{c}\right) = (-1)^{(3-d-(d^{-1} \bmod c)+c\sum(-1)^i a_i)/4}, \tag{2}$$

where $d/c = [0, a_1, \ldots, a_r]$ with r even. In fact, the exponent in (2) is closely related to the Dedekind sum classically defined as

$$s(d,c) := \sum_{h=1}^{c-1} ((hd/c))((h/c)), \tag{3}$$

where the symbol $((x))$ signifies 0 if x is an integer and $x - \lfloor x \rfloor - \frac{1}{2}$ otherwise (see also the last section, especially formula (5)). Zagier discovered that the continued fraction expansion "by excess",

$$\frac{d}{c} = b_0 - \cfrac{1}{b_1 - \cfrac{1}{b_2 - \cfrac{1}{b_3 - \cdots}}}$$

gives rise to an identity related to (2) (and to (5) below),

$$s(d,c) = \frac{1}{12}\left(-6 + \frac{c + (c^{-1} \bmod d)}{d} - \sum_i (b_i - 3)\right). \tag{4}$$

In fine, properties of the type (2) and (4) result from simple matrix algebra, the point being that linear fractional transformations that arise from continued fractions are generated by $S(z) = z + 1$, and $T(z) = -1/z$, and are representable as 2×2 integer matrices of determinant 1, i.e., the group $SL(2, \mathbb{Z})$, while there is exactly one additive character on this group.

Amongst other interesting connections, we find Cornacchia's deterministic algorithm (1908) that makes it possible to write a prime as a sum of two squares directly from the Euclidean algorithm. (Legendre and Hermite had already introduced continued fractions for this problem.) Continued fraction algorithms are discussed in the celebrated HAKMEM ("Hackers' Memorandum") document [2], a text well worth reading.

2. **Statistica**

Examination of large slices of digits in continued fraction expansions of real constants (like $\pi, \gamma, \zeta(3)$ and many others, but unlike ϕ or e) reveals that the digits $1, 2, 3, \ldots$ tend to occur with a frequency that shows little deviation from the values 41%, 17%, 9%, ..., respectively. Such observations belong to the "metric" or "statistical" theory of continued fractions. They are essentially related to properties of the continued fraction transformation $T(x) = \{1/x\}$ (with $\{\cdot\}$ the fractional part function) under iteration.

Gauss first observed around 1800 that the probability density

$$\psi(x) = \frac{1}{\log 2} \frac{1}{1+x},$$

is invariant under the transformation T, as a simple computation based on partial fractions shows. The corresponding measure is called Gauss' measure. It was then natural to conjecture (and so did Gauss) that iteration of T on random data that obey a probability density, for instance a uniform $(0,1)$ density, produces distributions that converge to the invariant ψ.

Nowadays, we know two approaches to such problems: one, more qualitative, is based on ergodic theory, while the other, more precise, is based on functional analysis (following Kuzmin, Lévy, Wirsing) and transfer operators (following Ruelle, Babenko, Mayer, Hensley, Vallée). Indeed, the ergodic theorem grants us that properties of trajectories of real numbers under x are governed (almost surely) by Gauss' measure, while functional analysis identifies ψ with the dominant eigenfunction of the transfer operator $\mathcal{G}_1$, where

$$\mathcal{G}_s[f](x) = \sum_{m=1}^{\infty} \frac{1}{(m+x)^{2s}} f(\frac{1}{m+x}).$$

Properties of Gauss' measure give access to properties of continued fraction digits. For instance, the limit frequency of digit k in the continued fraction expansions of a random real number is (with probability 1) equal to

$$\varpi_k = \int_{1/k}^{1/(k+1)} \psi(x)\, dx = \log_2 \left(1 + \frac{1}{k(k+2)}\right),$$

so that $\varpi_k = O(k^{-2})$ has a "soft" tail (and infinite expectation). From this type of argument, one can characterize expected properties of continued fractions of real numbers. One has, almost surely, for any function f that is not too wild,

$$\lim_{n\to+\infty} \frac{1}{n} \left(f(a_1(x)) + \cdots + f(a_n(x))\right) = K_f = \sum_{k=1}^{\infty} f(k) \log_2 \left(1 + \frac{1}{k(k+2)}\right).$$

For instance, the geometric and harmonic means of partial quotients satisfy (almost surely)

$$\sqrt[n]{a_1 \cdots a_n} \to e^{K_{\log}}, \qquad e^{K_{\log}} = \prod_{k=1}^{\infty} \left(1 + \frac{1}{k(k+2)}\right)^{\log_2 k}$$

$$\frac{n}{\frac{1}{a_1} + \cdots + \frac{1}{a_n}} \to \frac{1}{K_{\text{inv}}}, \qquad K_{\text{inv}} = \sum_{k=1}^{\infty} \frac{1}{k} \log_2 \left(1 + \frac{1}{k(k+2)}\right).$$

Such constants can be evaluated systematically by the "zeta-function trick" discussed in Vardi's book [13]. One finds the values

$$e^{K_{\log}} \doteq 2.68554, \qquad \frac{1}{K_{\text{inv}}} \doteq 1.74540,$$

where the first one is known as Khinchine's constant and the second one is the harmonic mean constant [1]. See Finch's beautiful site on Constants [5] for details and pointers to original sources. For instance, the harmonic mean of the 17 million continued fraction digits of π, as computed by Gosper, is 1.74594, a fact that that brings additional credibility to the conjecture that the continued fraction expansion of π is "normal" (that is, random-looking).

A similar type of question is: *How much information does a continued fraction digit convey?* This conduces to Lévy's constant that is related to the entropy of the continued fraction map. For x a number of the interval $(0,1)$, let us write

$$\frac{P_n(x)}{Q_n(x)} = [0, a_1(x), \ldots, a_n(x)].$$

Lévy proved that (almost surely)

$$\lim_{n\to\infty} \frac{1}{n} \log Q_n(x) \to \frac{\pi^2}{12 \log 2},$$

by making use of an approximation of the type

$$\log Q_n(x) \approx \sum_{k=1}^{n} \log T^k(x),$$

to which ergodic theory can be applied. Since the fundamental interval[2] of rank n that is determined by x has length about $1/Q_n(x)^2$, there results that very roughly,

$$Q_n(x)^{-2} \approx \left(e^{\pi^2/(6\log 2)}\right)^n \doteq 10^{-1.03064n}.$$

In other words, n continued fraction digits discriminate a number as well as about $1.03064n$ decimal digits. (This is again consistent with Gosper's data on the number π.) Refinements on Lévy's estimates are due to Philipp (who first proved that the distribution of $\log Q_n(x)$ is asymptotically Gaussian) and many others. Lévy's constant also appears (for good reasons, see Vallée's works, e.g., [11]) in the mean number of iterations of Euclid's algorithm which, for rational numbers p/q with $1 \le p < q \le n$, is

$$\frac{12 \log 2}{\pi^2} \log n + O(1).$$

This last estimate is due to Heilbronn and Dixon around 1969. A companion Gaussian limit distribution further holds, which is a "hard" result due to Hensley, 1994.

Arithmetic means of partial quotients take us to a different circle of ideas and the probabilistic phenomena at stake become now quite irregular. Define the sum-of-digits function,

$$S_n(x) = \sum_{k=1}^{n} a_k(x).$$

The point is that the mean value of a continued fraction digit satisfies

$$\sum_{k=1}^{\infty} k \log\left(1 + \frac{1}{k(k+2)}\right) = +\infty.$$

Thus, the arithmetic means S_n/n do not behave at all like their geometric or harmonic counterparts: exceptionally large values of partial quotients a_n and of sums S_n are to be encountered not too infrequently.

[2]The fundamental interval of rank n determined by x is defined by all the numbers whose representation starts with $[0, a_1(x), \ldots, a_n(x)]$.

This part of the discussion aims at understanding the finesses of large fluctuations in continued fraction digits. First, a Borel-Cantelli argument shows that, given any function $\phi(n)$, then (with λ the Lebesgue measure)

$$\lambda \{x \in (0,1) \ : \ a_n(x) > \phi(n) \text{ (infinitely often)}\} = 1 \quad \Longleftrightarrow \quad \sum \frac{1}{\phi(n)} = +\infty.$$

Thus, we should expect $a_n > n \log n \log\log n$ infinitely often, but $a_n < n \log n (\log\log n)^2$ only at a finite number of places. In fact, Khinchine proved that

$$\lambda \left\{ x \ : \ \left| \frac{S_n(x)}{n \log_2 n} - 1 \right| > \epsilon \right\} \underset{n\to\infty}{\longrightarrow} 0,$$

for any ϵ. In other words, the arithmetic mean of n digits is "typically" close to $\log_2 n$.

Heinrich [8] obtained a more precise limiting distribution result: for any $x \in \mathbb{R}$ and with μ the Gauss measure:

$$\left| \mu \left\{ x \ : \ \frac{1}{n} \sum_{k=1}^{n} a_k(x) - \frac{\log n - \gamma}{\log 2} < t \right\} - G_{1,1}\left(t, \frac{\pi}{2\log 2}\right) \right| = O\left(\frac{(\log n)^2}{n}\right).$$

There γ denotes Euler's constant and $G_{1,1}(t,\lambda)$, is the stable distribution function whose characteristic function equals

$$\exp(-\lambda |u| (1 + 2i \log|u| \operatorname{sgn} u/\pi)).$$

Diamond and Vaaler further established, in a precise sense, that for almost all x's, there is at most one large partial quotient $a_n(x)$: one has

$$\left(S_n(x) - \max_{k\geq n} a_k(x) \right) \sim n \log_2 n,$$

with probability 1. Finally, the St. Petersburg game (play head-and-tails and double the stake till you win) is a well-known example of a sequence of i.i.d. random variables with infinite expectation and was considered as a paradox since no single 'fair' entry fee exists. In relation with the discussion of arithmetic means, Vardi's note [16] shows how the sequence of continued fraction digits of a random real number makes a reasonable choice of entry fees.

In the discrete world, these results indicate that the running time of the subtractive Euclidean algorithm (equal to the sum of the continued fraction digits) should be about $\log n \log\log n$, in the sense that if $S(p/q)$ is the running time, then $S(p/q)/(\log q \log\log q)$ should have a limiting distribution (consistent with Heinrich's result). This would imply that typically, the subtractive algorithm is more expensive than the standard algorithm by a factor of $\log\log n$. Knuth and Yao showed that, on average, the sum of all the partial quotients of all the regular continued fractions for m/n, with $1 \leq m \leq n$, is

$$\frac{6}{\pi^2} n (\log n)^2 + O(n \log n (\log\log n)^2),$$

where Lévy's constant pokes its nose again. (See Vallée's recent works [11] for the light shed by transfer operators on such phenomena.) This tells us that the *average* behavior of the subtractive algorithm is much different from its *typical* behaviour. Such results are not uncommon in number theory, e.g., in divisor problems; see [10].

3. Analytica

The last part of the talk is based on Vardi's paper [15]. The aim is to characterize the distribution of the *alternating sum-of-digits*

$$T_n(x) := \sum_{k=1}^{n} (-1)^k a_k(x),$$

for the continuous case, and of the discrete counterpart

$$T(d/c) = T_r(d/c), \quad \text{where} \quad d/c = [0, a_1, \ldots, a_r].$$

The mathematics go deeper and involve the hyperbolic Laplacian, Kloosterman sums, modular forms, and Eisenstein series. (See Sarnak's introductory lectures [9] for some related background.) The motivation comes largely from the need to understand the distribution of Dedekind sums defined in (3) that also satisfy a formula of Dean Hickerson (compare with Zagier's formula (4)), namely

$$s(d,c) = \frac{1}{12}\left(-3 + \frac{d + (d^{-1} \bmod c)}{c} - \sum_{i=1}^{r}(-1)^i a_i\right), \tag{5}$$

where the regular continued fraction expansion of $d/c = [0, a_1, \ldots, a_r]$ is normalized in such a way that r is even. This formula shows that Dedekind sums should really be thought of as alternating sums of regular continued fraction coefficients.

First, in the continuous case, the distribution of T_n is likely to resemble the difference between two random variables of the S_n type, themselves each distributed according to a $G_{1,1}$ stable law as discussed earlier (Heinrich's theorem). Now the difference of two $G_{1,1}$ is a Cauchy distribution, so that we are led to expect

$$\lim_{n\to\infty} \lambda\left\{x \ : \ T_n(x) < yn\right\} = \frac{1}{\pi}\int_{-\infty}^{y} \frac{\xi}{\xi^2 + x^2}\, dx, \tag{6}$$

with $\xi = \pi/(2\log 2)$.

In the discrete case, Vardi [15] has succeeded in proving that an analogue of (6) holds. In the language of Dedekind sums, the main result of [15] is then

$$\begin{aligned} \lim_{n\to\infty} \frac{\#\{0 < d < c < n, (d,c) = 1, s(d,c) < x\log c\}}{\#\{0 < d < c < n, (d,c) = 1\}} &= \frac{1}{\pi}\int_{-\infty}^{x} \frac{(2\pi)^{-1}}{(2\pi)^{-2} + s^2}\, ds \\ &= \frac{1}{\pi}\arctan(2\pi x) + \frac{1}{2}. \end{aligned} \tag{7}$$

(By Hickerson's formula (5), this gives the very same limiting distribution result for the alternating sum of continued fraction coefficients.) The proof uses the full machinery of modular forms and harmonic analysis on $SL(2,\mathbb{Z})$, see [9]. It would be of interest to find a simpler proof, as this would very likely have the advantage of providing the limiting distribution for the sum of continued fraction coefficients. The fact that there is a unique additive character for $SL(2,\mathbb{Z})$ makes it unlikely that this could be obtained using harmonic analysis.

Bibliography

[1] Bailey (David H.), Borwein (Jonathan M.), and Crandall (Richard E.). – On the Khintchine constant. *Mathematics of Computation*, vol. 66, n° 217, 1997, pp. 417–431.

[2] Beeler (M.), Gosper (R. W.), and Schroeppel (R.). – *HAKMEM*. – Memorandum n° 239, M.I.T., Artificial Intelligence Laboratory, February 1972. Available on the WorldWide Web at `http://www.inwap.com/pdp10/hbaker/hakmem/hakmem.html`.

[3] Caveing (J.). – *L'irrationalité, dans les mathématiques grecques jusqu'à Euclide.* – Presses Universitaires du Septentrion, Paris, 1998.
[4] Dutt (Romesh Cunder). – *A History of Civilisation of Ancient India Based on the Sanskrit Literature.* – Vishal Publishers, Delhi, 1972. A reprint of the original edition, 1888.
[5] Finch (Steven). – Favorite mathematical constants. – Available on the World Wide Web at the URL `http://www.mathsoft.com/asolve/constant/constant.html`, 1995.
[6] Fowler (D. H.). – *The Mathematics of Plato's Academy: A New Reconstruction.* – Clarendon Press, Oxford, 1987.
[7] Hardy (G. H.) and Wright (E. M.). – *An Introduction to the Theory of Numbers.* – Oxford University Press, 1979, fifth edition.
[8] Heinrich (Lothar). – Rates of convergence in stable limit theorems for sums of exponentially ψ-mixing random variables with an application to metric theory of continued fractions. *Mathematische Nachrichten*, vol. 131, 1987, pp. 149–165.
[9] Sarnak (Peter). – *Some Applications of Modular Forms.* – Cambridge University Press, 1990, *Cambridge Tracts in Mathematics*, vol. 99.
[10] Tenenbaum (Gérald). – *Introduction to analytic and probabilistic number theory.* – Cambridge University Press, Cambridge, 1995, xvi+448p. Translated from the second French edition (1995) by C. B. Thomas.
[11] Vallée (Brigitte). – A unifying framework for the analysis of a class of Euclidean algorithms. – Preprint, 2000. To appear in *Proceedings of LATIN'2000, Lecture Notes in Computer Science*, in press.
[12] Van der Poorten (Alfred). – A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$. *Mathematical Intelligencer*, vol. 1, 1979, pp. 195–203.
[13] Vardi (Ilan). – *Computational Recreations in Mathematica.* – Addison Wesley, 1991.
[14] Vardi (Ilan). – The distribution of Dedekind sums. – Preprint, October 1992. 20 pages.
[15] Vardi (Ilan). – Dedekind sums have a limiting distribution. *International Mathematics Research Notices*, n° 1, 1993, pp. 1–12.
[16] Vardi (Ilan). – The St. Petersburg game and continued fractions. *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique*, vol. 324, n° 8, 1997, pp. 913–918.
[17] Vardi (Ilan). – Archimedes' cattle problem. *The American Mathematical Monthly*, vol. 105, n° 4, 1998, pp. 305–319.