

On the Comparison of Numbers:

An excursion into non-classical corners of computational mathematics

Philippe Flajolet, INRIA, Rocquencourt

<http://algo.inria.fr/flajolet>

Based on joint work with Brigitte Vallée, Caen

A very simple problem

Let X, Y be uniformly distributed over $[0, 1]$.

Cost of deciding $X \lesseqgtr Y$?

A very simple solution

If X, Y are given in **some base**, e.g., $B = 2$

- scan digits;
- stop when **discrepant digits**/bits are found.

Expected cost (number of comparisons) is $E[C] = 2$.

Probability distribution is **geometric**: $P(C = k) = 2^{-k}$.

THE END!?

What if X and Y are given as fractions?

$$X = \frac{a}{b}, \quad Y = \frac{c}{d} \quad ?$$

E.g. Compare:

$$\frac{36}{113} \begin{matrix} \leq \\ > \end{matrix} \frac{113}{355}; \quad \frac{97.8}{307.1} \begin{matrix} \leq \\ > \end{matrix} \frac{307.1}{964.9} \quad ?$$

Still mathematically trivial $\text{sign}(ad - bc) \dots$

- Requires double precision
- Unsafe in fixed precision (floats)

Efficiency in symbolic manipulation systems and number-theory packages: $O(N^2)$ boolean complexity.

Robustness: Problems with floats.

Independent solutions:

- HAKMEM (Gosper 1972)
- Knuth's Metafont (late 1970s);
- computational geometry \equiv orientation (Boissonnat, et al.)

Introduction.

HAKMEM 1972

Item 101A (Gosper): 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.

$$\alpha = \frac{113}{36} = 3 + \frac{1}{7 + \frac{1}{\mathbf{5}}}, \quad \beta = \frac{355}{113} = 3 + \frac{1}{7 + \frac{1}{\mathbf{16}}},$$

= two "lazy" parallel executions of the continued fraction algorithm.

Simulations (10^6): HAKMEM $X \lesseqgtr Y$ for $X, Y \in [0, 1]$, uniform.

	Basic CF-sign	Centered CF-sign
$P(L = 1)$	0.710050	0.918003
$P(L = 2)$	0.241275	0.075710
$P(L = 3)$	0.038339	0.000422
$P(L = 4)$	0.008424	0.000035
$P(L = 5)$	0.001608	0.000005
$E(L)$	1.351612	1.088791

Decay roughly geometrically with k : 5^{-k} for BCF; 13^{-k} for CCF.

The expectations seem to be constants:

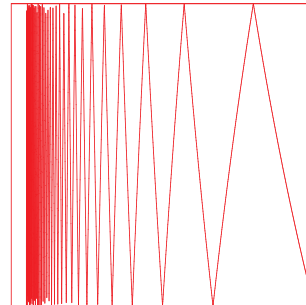
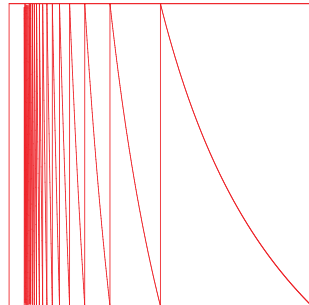
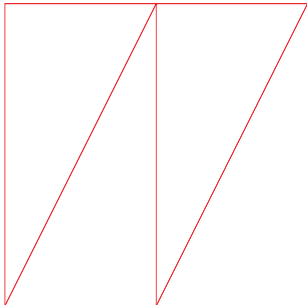
Anticipate $E(L) = 1.352 \pm 0.001$ and $E(\tilde{L}) = 1.089 \pm 0.001$.

1 Expanding maps & continued fractions

A piecewise monotonic function $U(x)$ (**shift**) from some interval \mathcal{I} that is **expanding**: $|U'(x)| > 1$.

Branches are indexed by $\mathcal{M} =$ **digits**.

A **coding** of a real x : $m(x), m(U(x)), m(U^2(x)), \dots$



procedure $(\mathcal{I}, \mathcal{U}, m)$ -**expansion** $(x : \mathcal{I})$

for $k := 1$ **to** $+\infty$ **do**

$m_k := m(x); x := \mathcal{U}(x);$

Set \mathcal{H}_1 of **branches** of $\mathcal{U}^{(-1)}$. Run backwards:

$$x_0 = h_{\mathbf{m}}(x_k) \quad \text{where} \quad h_{\mathbf{m}}(y) = h_{m_1} \circ h_{m_2} \circ \cdots \circ h_{m_k}(y).$$

The set \mathcal{H} of all compositions $h_{\mathbf{m}}$ is set of *inverse branches*.

Equivalence principle.

- the stochastic behaviour of a numbering system
- the dynamics of $\mathcal{U}(x)$ on the interval \mathcal{I}
- the dynamics of the semigroup of contractions \mathcal{H} .

Basic continued fraction expansions.

$$I = [0, 1], \quad U(x) = \left\{ \frac{1}{x} \right\}, \quad m(x) = \left[\frac{1}{x} \right].$$

Inverse branches of depth 1: $\mathcal{H}_1 := \{h(z) = \frac{1}{m+z} \mid m \geq 1\}$.

$$x_0 = h_{\mathbf{m}}(x_k) \quad \text{where} \quad h(y) = \frac{1}{m_1 + \frac{1}{m_2 + \frac{1}{\ddots + \frac{1}{m_k + y}}}}.$$

Centered continued fraction expansions.

$$\tilde{\mathcal{H}}_1 := \left\{ h(z) = \frac{1}{m+z} \mid m \geq 2 \right\} \cup \left\{ h(z) = \frac{1}{m-z} \mid m \geq 3 \right\}.$$

Sign algorithms.

Take two random points; follow their trajectories under the shift \mathcal{U} . Stop on discrepancy.

procedure $(\mathcal{I}, \mathcal{U}, m)$ -**sign**($x, x' : \mathcal{I}$)

for $k := 1$ **to** $+\infty$ **do**

$m_k := m(x); m'_k := m(x'); x := \mathcal{U}(x); x' = \mathcal{U}(x');$

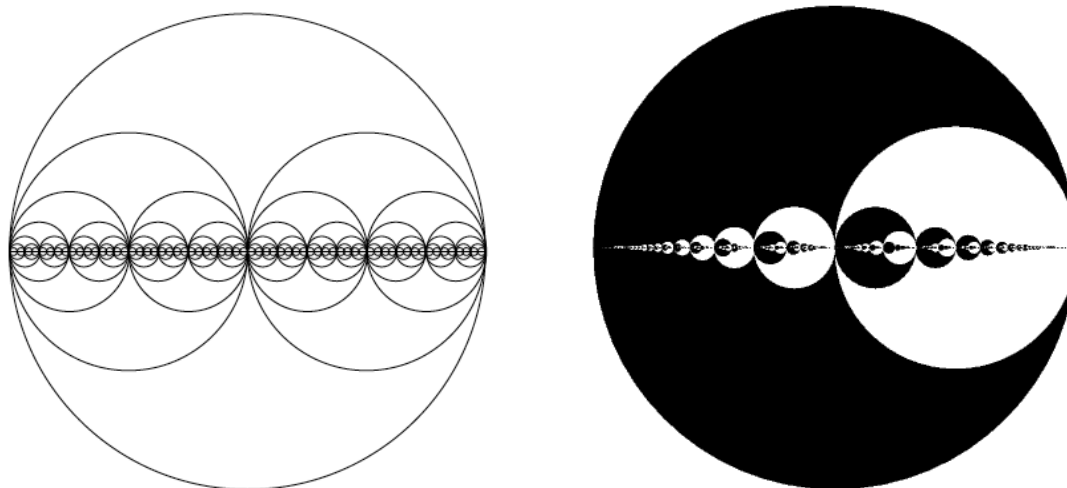
if $m_k \neq m'_k$ **then exit.**

2 Fundamental intervals & signs

Uniform probability model (Lebesgue) on \mathcal{I} .

The probability that $x \in \mathcal{I}$ belongs to the **fundamental interval** determined by the branch h is, with $|\mathcal{I}| = 1$,

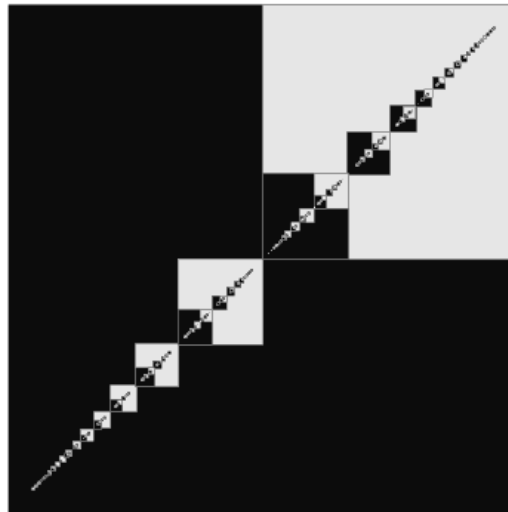
$$u_h = |h(0) - h(1)|.$$



The square $\mathcal{I} \times \mathcal{I}$.

$$\text{Event } [L \geq k + 1] = \bigcup_{|h|=k} h(\mathcal{I}) \times h(\mathcal{I}). \quad (1)$$

$$P([L \geq k + 1]) = \sum_{|h|=k} u_h^2 = \sum_{|h|=k} |h(0) - h(1)|^2. \quad (2)$$



The **average number of iterations** is obtained by summation:

$$E[L] = \sum_{\mathbf{h}} u_{\mathbf{h}}^2 = \sum_{\mathbf{h}} |\mathbf{h}(0) - \mathbf{h}(1)|^2. \quad (3)$$

More generally, we define the “**moment sums**”

$$\rho^{(\ell)} := \sum_{\mathbf{h}} u_{\mathbf{h}}^{\ell} = \sum_{\mathbf{h}} |\mathbf{h}(0) - \mathbf{h}(1)|^{\ell}$$

which are central in the **n-sorting** algorithm (\leadsto later).

Continued fraction sign-algorithms.

Theorem 1. *Expected costs of BCF/CCF-sign are **lattice sums**,*

$$E(L) = \rho^{(2)}, \quad E(\tilde{L}) = \tilde{\rho}^{(2)},$$

where, the “moments” of index ℓ are

$$\begin{aligned} \rho^{(\ell)} &= 1 + \frac{1}{2^\ell} + \frac{2}{\zeta(2\ell)} \sum_{(d,c) \in \mathcal{C}(1,2)} \frac{1}{c^\ell d^\ell} \\ \tilde{\rho}^{(\ell)} &= \frac{2^\ell}{\zeta(2\ell)} \sum_{(d,c) \in \mathcal{C}(\phi, \phi^2)} \frac{1}{c^\ell d^\ell}, \quad (\phi = (1 + \sqrt{5})/2), \end{aligned}$$

w/ *lattice sum*: $\mathcal{C}(\beta, \gamma) := \{(m, n) \in \mathbf{N}^2 \mid m\beta < n < m\gamma\}$.

Proof (elementary!):

— The set \mathcal{H} of all possible LFT's used by the BCF–algorithm is

$$\mathcal{H} := \left\{ \frac{az + b}{cz + d} \mid (a, b, c, d) \in \mathbf{N}^4, |ad - bc| = 1, c \leq d, a \leq c, b \leq d \right\}.$$

= unimodular h + relatively prime coeffs. + inequalities.

— The set $\tilde{\mathcal{H}}$ of all possible LFT's used by the CCF–algorithm is

$$\tilde{\mathcal{H}} := \left\{ \frac{az + b}{cz + d} \mid (b, d) \in \mathbf{N}^2, (a, c) \in \mathbf{Z}^2, ac \geq 0, |ad - bc| = 1, \frac{-d}{\phi^2} < c < \frac{d}{\phi}, |a| \leq \frac{|c|}{2}, b \leq \frac{d}{2} \right\}.$$

3 The sorting algorithm

Given $X := \{x_1, x_2, \dots, x_n\}$ with $x_j \in \mathcal{I}$. Build *digital tree*, $\text{trie}(X)$:

(R₁) If $X = \{x\}$ then $\text{trie}(X)$ is single *leaf node* \boxed{x} .

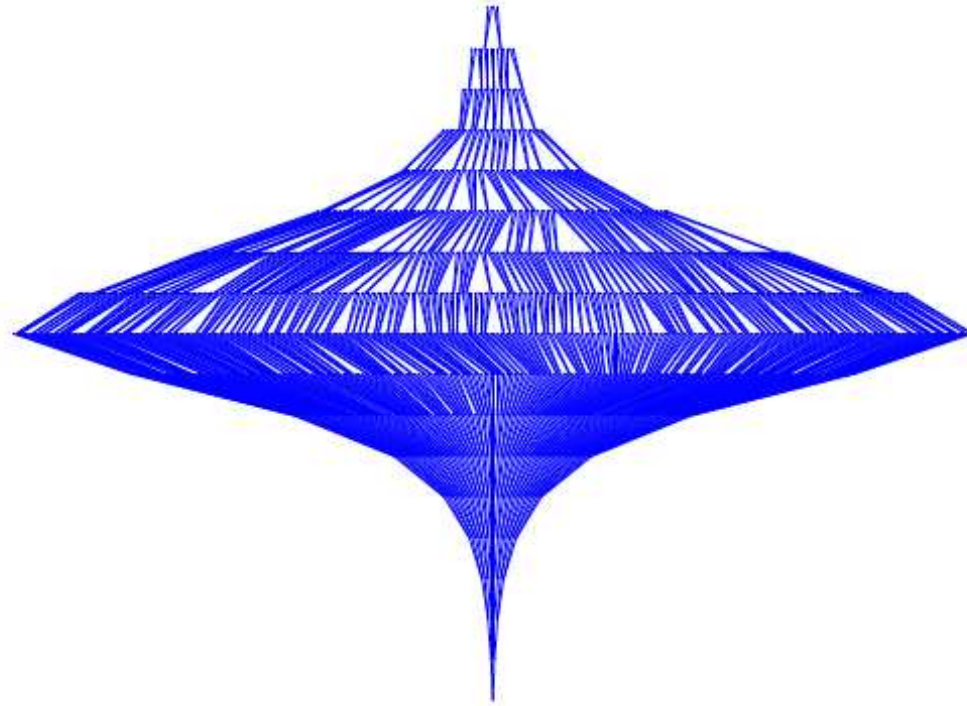
(R₂) If $\text{card}(X) \geq 2$, then $\text{trie}(X)$ is *built recursively*

$$\text{trie}(X) = \langle o, \text{trie}(X_1), \text{trie}(X_2), \dots, \text{trie}(X_r) \rangle,$$

where $X_i = \{U(x) \mid m(x) = b_i, x \in X\}$.

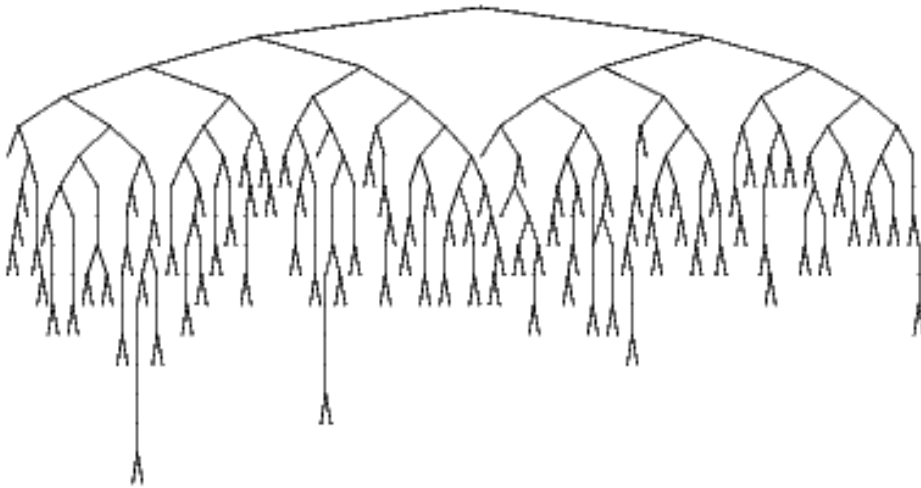
 Follow n random trajectories and stop each when it “deviates”.

The binary trie



is another saga . . .

$$\begin{aligned}
\phi - 1 &= /1, 1, 1, 1, 1, 1, \dots./ \\
\gamma &= /1, 1, 2, 1, 2, 1, \dots./ \\
\exp(1) - 2 &= /1, 2, 1, 1, 4, 1, \dots./ \\
\log 2 &= /1, 2, 3, 1, 6, 3, \dots./ \\
\{\exp(\pi\sqrt{163})\} &= /1, \mathbf{1333462407511}, 1, 8, 1, 1, \dots./ \\
2^{1/3} - 1 &= /3, 1, 5, 1, 1, 4, \dots./ \\
\pi - 3 &= /7, 15, 1, 292, 1, 1, \dots./
\end{aligned}$$



Typical shape??

Theorem 2. *The expectation of the number of digit inspections*

$$P(n) = n \sum_{\ell=2}^n (-1)^{\ell} \binom{n-1}{\ell-1} \rho^{\langle \ell \rangle},$$

where $\rho^{\langle \ell \rangle}$ is a moment sum. [Similarly for CCF.]

= **finite differences** of moment sums!

$$P(2) = 2 \rho^{(2)}, \quad P(3) = 6 \rho^{(2)} - 3 \rho^{(3)}, \quad P(4) = 12 \rho^{(2)} - 12 \rho^{(3)} + 4 \rho^{(4)}.$$

Proof: (i) **Poissonize** and get independence;

(ii) **de-Poissonize** “algebraically”.

Cf. Clément-FI-Vallée, *Algorithmica* 2001.

4 Multiple zeta values [Euler-Zagier-Borwein²]

With $H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$:

$$\sum_{n=1}^{\infty} \frac{(H_{n-1})^4}{n^3} = \frac{185}{8} \zeta(7) - \frac{43}{2} \zeta(3)\zeta(4) + 5\zeta(2)\zeta(5)$$

$$\zeta^+(s) \equiv \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \zeta^-(s) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^s} = (2^{1-s} - 1)\zeta(s),$$

$$\zeta^{-+}(s, t) = \sum_{n=1}^{\infty} \sum_{q=1}^{n-1} \frac{(-1)^n}{n^s q^t} = \sum_{n=1}^{\infty} (-1)^n \frac{H_{n-1}^{(t)}}{n^s}.$$

A complete evaluation means a reduction to a polynomial form in zeta values $\zeta(2), \zeta(3), \dots, \bar{\zeta}(1) = -\log 2$, and possibly other constants.

——— (Zagier 1994), (Borwein*), (Fl-Salvy, *Exp. Math.* 1998) ...

$$\text{Li}_m(z) = \frac{z}{1^m} + \frac{z^2}{2^m} + \frac{z^3}{3^m} + \cdots = \sum_{n=1}^{\infty} \frac{z^n}{n^m}. \quad (4)$$

Theorem 3. *The mean number $\rho^{\langle 2 \rangle}$ of comparisons in BCF–Sign can be expressed in terms of double zeta values as*

$$\rho^{\langle 2 \rangle} = \frac{17}{4} + \frac{360}{\pi^4} \sum_{d=1}^{\infty} \frac{(-1)^d}{d^2} \sum_{c=1}^d \frac{1}{c^2} = \frac{15}{2} - \frac{720}{\pi^4} \sum_{d=1}^{\infty} \frac{(-1)^d}{d^3} \sum_{c=1}^d \frac{1}{c},$$

or with $\zeta(3)$ and the tetralogarithm $\text{Li}_4(\frac{1}{2})$,

$$\rho^{\langle 2 \rangle} = -\frac{60}{\pi^4} \left(24\text{Li}_4\left(\frac{1}{2}\right) - \pi^2(\log 2)^2 + 21\zeta(3)\log 2 + (\log 2)^4 \right) + 17. \quad (5)$$

Thus, $\rho^{\langle 2 \rangle}$ is in class **P** = polynomial time computable:

$$\rho^{\langle 2 \rangle} = 1.35113\ 15744\ 91659\ 00179\ 38680\ 05256\ 46466\ 84404\ 78970\ 85087 \pm 10^{-50}.$$

Moment sums for BCF

♠ Original sum are **very slowly convergent**.

♡ Following Zeilberger, define $f(z)$ to be *holonomic* if its satisfies a *linear differential equation* with coefficients in $\mathbb{Q}(z)$.

A constant that is the value $f(z_0)$ of a holonomic function at an algebraic point z_0 is called a *holonomic constant*.

Such holonomic constants are **polynomial-time computable (P)**.

♡ Usual series-acceleration (Cohen-Villegas-Zagier; Pari, Maple) works just fine in practice!

BCF-Sorting: Get $P(n)$ to 16D for $n = 0, \dots, 200$ via $\rho^{(\ell)}$ to 200D.

What about Centred Continued Fractions (CCF)?

$$\tilde{\rho}^{(2)} = \frac{2^2}{\zeta(4)} \sum_{c\varphi < d < c\varphi^2} \frac{1}{c^2 d^2}, \quad (\varphi = (1 + \sqrt{5})/2)$$

Quiz:

— Can you compute to 100D : $\sum_{n \geq 1} \frac{1}{[n\phi]^2}$?

(= 1.29106 03681 14387 48950 47876 ...)

— Do you believe this? (Borwein² *AMM* 1992)

$$\sum_{n=1}^{\infty} \frac{\lfloor ne^{\pi\sqrt{163}} \rfloor}{2^n} \doteq 1280640 \quad \text{with error} < 10^{500,000,000}.$$

Theorem 4. *The expectation of the cost of the CCF-sign algorithm is in \mathbf{P} :*

$$\tilde{\rho}^{(2)} = 1.08922\ 14740\ 95380\ \dots$$

Two ingredients:

Lemma 1 (Mahler, B2, FV). *Let $\theta < 1$ be an irrational number with convergents $\{p_n/q_n\}_{n=0}^\infty$. The generating function of the lattice cone $\mathcal{C}(0, \theta)$ is given by*

$$\sum_{(m,n) \in \mathcal{C}(0,\theta)} x^m y^n = \sum_{k=0}^{\infty} (-1)^k \frac{x^{q_k+q_{k+1}} y^{p_k+p_{k+1}}}{(1-x^{q_k} y^{p_k})(1-x^{q_{k+1}} y^{p_{k+1}})}.$$

Proof: Cut a lattice by a \mathbb{Q} -slanted line; do inclusion-exclusion.

+ **Mellin trick:** for seq. (a_n) and $A(z) := \sum_n z^{a_n}$, we have

$$\sum_n \frac{1}{(a_n)^s} = \frac{1}{\Gamma(s)} \int_0^\infty A(e^{-t}) t^{s-1} dt = \frac{1}{\Gamma(s)} \int_0^\infty A(w) \left(\log \frac{1}{w} \right)^{s-1} \frac{dt}{t}.$$

5 Transfer operators and sign algorithms

If X is a random variable with density $f(x)$, then $U(X)$ has density

$$g(x) = \mathcal{G}[f](x), \quad \text{where} \quad \mathcal{G}[f](x) := \sum_{h \in U^{(-1)}} |h'(x)| f \circ h(x).$$

This is the **density transformer** aka **PF-operator**:

The Ruelle **transfer operator** is the generalization:

$$\mathcal{G}_s[f](x) := \sum_{h \in U^{(-1)}} |h'(x)|^s f \circ h(x).$$

Get the dynamics of U from analysis of \mathcal{G}_s ?

For general sign/sorting algorithms, Vallée introduces **secant generalization**...

An aside: Euclid's algorithm (1-dim.)

$$\mathcal{G}_s[f](x) := \sum_{m \geq 0} \frac{1}{(m+x)^{2s}} f\left(\frac{1}{m+x}\right).$$

Theorem (Hensley–Baladi–Vallée) *Euclid's algorithm is Gaussian!!*



Basic continued fractions (BCF) & sign algorithm

The transfer operator is

$$\mathcal{G}_s[f](x) := \sum_{m \geq 0} \frac{1}{(m+x)^{2s}} f\left(\frac{1}{m+x}\right).$$

- Gauß conjectured in 1800 the stationary distribution of CFs as $\gamma(x) \propto \frac{1}{1+x}$. Certainly $\gamma(x)$ is eigenvalue of \mathcal{G}_1 , but more was needed (Lévy, Kuzmin, Wirsing).
- Around 1975, Babenko proposed numerical estimates of subdominant eigenvalues of \mathcal{G}_1 to be found in (Knuth 1981)
- In 1994+, FV needed dominant eigenvalues of \mathcal{G}_2 for sign algorithm and on the way suspected some spurious values; extensive tests led to conjecture **first 37 eigenvalues of \mathcal{G}_1 to 25D with “almost certainty”**.
- In 2005, **Loïc Lhote** (Caen) could provide certified algorithms for Wirsing’s constant $\lambda_2(1)$, Vallée’s constant $\lambda_1(2)$, and Hensley’s constant $\lambda''(1)$.

Theorem 5. *The probability that the sign algorithm performs $k+1$ iterations satisfies*

$$P(L \geq k + 1) = C\lambda_1(2)^k + \text{exp. small},$$

where $\lambda_1(2) \doteq 0.19945\ 88183\ 43767\ \dots$

Proof. (i) For continued fraction need $\sum u_h^2$, where the quantity u_h is $|\mathbf{h}(0) - \mathbf{h}(1)| = \mathbf{h}'(1)^2$. Here can express everything with \mathcal{G}_2^k at depth k .

(ii) Need numerical access to $\text{Spec}(\mathcal{G}_2)$?

Ideas: (a) Use projection on space of polynomials (\rightsquigarrow below)

(b) Expect spectrum of \mathcal{G}_2 to be exponentially decreasing, simple, with sign alternations. Filter based on conjecture.

(c) Use a technique of “test functions” based on positivity to certify dominant eigenvalue.

(d) Use trace formulae of Mayer–Roepstorff for $\text{Trace}(\mathcal{G}_2^2)$ to check filtering; evaluation based on Fl-Vardi’s “zeta trick”.

Projection on polynomials

Fact: Transform of x^m is a Hurwitz zeta whose Taylor coefficients involve zeta values:

We wind up inverting matrices of **zeta values** \otimes **binomials**, e.g.

$$\begin{pmatrix} \zeta(4) & \zeta(5) & \zeta(6) & \zeta(7) \\ -4\zeta(5) & -5\zeta(6) & -6\zeta(7) & -7\zeta(8) \\ 10\zeta(6) & 15\zeta(7) & 21\zeta(8) & 28\zeta(9) \\ -20\zeta(7) & -35\zeta(8) & -56\zeta(9) & -84\zeta(10) \end{pmatrix}.$$

This passes all **consistency tests** and it works (Lhote 2005)!

6 Sorting and RH!?

The problem is **analytically well posed**

$$P(n) = n \sum_{\ell=1}^{n-1} (-1)^{\ell-1} \binom{n-1}{\ell} \rho^{(\ell+1)},$$

$$\rho^{(s)} = 2^{-s} - 2^{1-s} + (2^{s-1} - 1) \frac{\zeta(s)^2}{\zeta(2s)} + 2^s \frac{\zeta^{-+}(s)}{\zeta(2s)},$$

$$\zeta^{-+}(s) ::= \sum_{m < n}^{\infty} \frac{(-1)^n}{m^s n^s},$$

but there are **huge cancellations**.

Theorem 6. *The expected cost of sorting n uniform real numbers given by their basic continued fraction representations is*

$$P(n) = K_0 n \log n + K_1 n + Q(n) + K_2 + o(1),$$

where K_0 is Lévy's entropic constant, K_1 is Porter-like

$$K_0 = \frac{6 \log 2}{\pi^2}, \quad K_1 = 18 \frac{\gamma \log 2}{\pi^2} + 9 \frac{(\log 2)^2}{\pi^2} - 72 \frac{\log 2 \zeta'(2)}{\pi^4} - \frac{1}{2}.$$

The function $Q(u)$ is an oscillating function with mean value 0:

$$Q(n) = O(n^{\delta/2}),$$

where δ is any number such that

$$\delta > \sup \{ \Re(s) \mid \zeta(s) = 0 \}.$$

The Nörlund–Rice trick: If (f_k) admits an analytic lifting $\phi(s)$:

$$\sum_{k=1}^n \binom{n}{k} f_k = \frac{n!}{2i\pi} \int_L \phi(s) \frac{ds}{s(s-1)\cdots(s-n)}.$$

Singularities of $\phi(s)$ matter!

Here, we get [Riemann Hypothesis](#) into the game...
(but it is also exponentially offset).

