INRIA
INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE

An Average-case Analysis of the Gaussian Algorithm for Lattice Reduction

Hervé Daudé , Philippe Flajolet , Brigitte Vallée

N ° 2798
Février 1996

PROGRAMME 2

Rapport de recherche

1996

ISSN 0249-6399

An Average-case Analysis of the Gaussian Algorithm for Lattice Reduction

Abstract. *The Gaussian algorithm for lattice reduction in dimension 2 is analysed under its standard version. It is found that, when applied to random inputs in a continuous model, the complexity is constant on average, the probability distribution decays geometrically, and the dynamics is characterized by a conditional invariant measure. The proofs make use of connections between lattice reduction, continued fractions, continuants, and functional operators. Analysis in the discrete model and detailed numerical data are also presented.*

Une analyse en moyenne de l'algorithme de Gauss de réduction des réseaux

Résumé. L'algorithme de réduction des réseaux en dimension 2 qui est dû à Gauss est analysé sous sa forme dite standard. Il est établi ici que, sous un modèle continu, sa complexité est constante en moyenne et que la distribution de probabilités associée decroît géométriquement tandis que la dynamique est caractérisée par une densité conditionnelle invariante. Les preuves font appel aux relations entre réduction des réseaux, fractions continues, continuants, et opérateurs fonctionnels. Une analyse du modèle discret complétée de données numériques est aussi présentée.

AN AVERAGE-CASE ANALYSIS OF THE GAUSSIAN ALGORITHM FOR LATTICE REDUCTION

Hervé Daudé[1], Philippe Flajolet[2], and Brigitte Vallée[3]

[1] LATP, URA 225, Département de Mathématiques, CMI Université de Provence, 39 rue F.Joliot-Curie F-13453 Marseille Cedex 13 (France) [daude@gyptis.univ-mrs.fr].

[2] Algorithms Project, INRIA-Rocquencourt, F-78153 Le Chesnay (France), [Philippe.Flajolet@inria.fr].

[3] GREYC, URA 1526, Département d'Informatique, Université de Caen, F-14032 Caen (France), [Brigitte.Vallee@info.unicaen.fr].

Abstract. The Gaussian algorithm for lattice reduction in dimension 2 is analysed under its standard version. It is found that, when applied to random inputs in a continuous model, the complexity is constant on average, the probability distribution decays geometrically, and the dynamics is characterized by a conditional invariant measure. The proofs make use of connections between lattice reduction, continued fractions, continuants, and functional operators. Analysis in the discrete model and detailed numerical data are also presented.

1 Introduction

The lattice reduction problem consists in finding a short basis of a lattice of Euclidean space given an initially skew basis. This reduction problem is well–known to be central to many areas of approximation and optimization with deep consequences in computational number theory, cryptography, and symbolic computation.

In dimension $d = 1$, lattice reduction may be viewed as a mere avatar of the Euclidean GCD algorithm and of continued fraction expansions. Lattice reduction *per se* really started with Gauss who gave an algorithm that solves the problem exactly using what resembles a lifting of the Euclidean algorithm to 2–dimensional lattices. In recent times, an important discovery was made by Lenstra, Lenstra and Lovász in 1982 [3, 14, 30]; their algorithm, called the LLL algorithm, is able to find reduced bases in all dimensions $d \geq 3$. The LLL algorithm itself proceeds by stages based on the Gaussian algorithm as the main reduction step.

The Euclidean algorithm and the continued fraction algorithm are by now reasonably well understood as regards complexity questions. Knuth's book [11] provides a detailed account till 1981. From results of Wirsing, Babenko, and Hensley, the following facts are known. The worst case complexity of the Euclidean algorithm is $\mathcal{O}(\log N)$ when applied to integers at most N (Lamé and Dupré); the average case on random inputs is also logarithmic (Dixon, Heilbronn); the distribution of the number of iterations obeys in the asymptotic limit a normal law with a variance that is logarithmic (a recent result of Hensley).

There are some deep connections between these properties and an invariant measure for the continued fraction transformation whose existence was first conjectured by Gauss and proved in this century by Lévy and Kuzmin. Most of these results are obtained by means of functional operators related to continued fractions and continuants of which extensive use will be made here. We refer in particular to the works of Wirsing [29], Babenko [1], Mayer [16, 17, 18, 19], and Hensley [9].

This paper provides a detailed analysis of the Gaussian algorithm, both in the average case and in probability. Like its one–dimensional counterpart, the algorithm is known to be of worst–case logarithmic complexity, a result due to Lagarias [13], with best possible bounds being provided by Vallée [25] and Kaib-Schnorr [10]. The probabilistic behaviour of the Gaussian algorithm turns out to be appreciably different however. The main results of the paper are as follows.

- The average–case complexity of the Gaussian algorithm (measured in the number of iterations performed) is asymptotically constant, and thus essentially independent of the size of the input vectors.
- The distribution of the number of iterations is closely approximated by a geometric law.

The paper also describes the evolution of data during the execution of the algorithm. One begins with an initial density of data inside some domain: What is the density inside the domain after k iterations of the algorithm? We establish the following result:

- The dynamics of the algorithm is governed by a (conditional) limit measure that constitutes the analogue of the limit measure first observed by Gauss for continued fractions.

In this paper, we mostly focus on the analysis of what we call the "standard" version of the Gaussian reduction algorithm, which generalizes the standard Euclidean algorithm. Precise characterizations of the behaviour of the algorithm are given here. In particular the geometric rate of decrease of the distribution of costs and the limit measure are expressed simply in terms of spectral properties of Ruelle-Mayer operators that generalize the Perron-Frobenius operator classically associated with Euclid's algorithm.

Our analytic results are naturally expressed as multiple infinite sums involving the continuants of continued fraction theory. As such sums tend to be rather slowly convergent, some attention is also paid to obtaining precise estimates by means of suitable convergence acceleration techniques. For instance, we establish that the average case complexity of the algorithm is asymptotic to the constant $\mu = 1.35113\,15744\ldots$ that can be expressed with a number of remarkable quantities like $\zeta(3)$ and the tetralogarithm of argument $1/2$.

On average, the Gaussian algorithm is thus of complexity $\mathcal{O}(1)$, which is of an order different from the worst-case. The case of dimension $d = 2$ therefore departs significantly from its 1–dimensional analogue, and it would be of interest to determine to which extent such a phenomenon propagates to higher dimensions. Our analytic knowledge of the LLL algorithm in higher dimensions is of course less advanced, but Daudé and Vallée [4] already succeeded in proving that the LLL algorithm, when applied to d–dimensional lattices, has an average–case complexity that is bounded from above by a constant K_d, where $K_d = \mathcal{O}(d^2 \log d)$. The present work thus fits as a component of a more global enterprise whose aim is to understand theoretically why the LLL algorithm performs in practice much better than worst–case bounds predict, and to quantify precisely the probabilistic behaviour of lattice reduction in higher dimensions.

An extended abstract of the present paper appears in [4].

2 Lattice reduction in dimension 2

Lattices and bases. This paper addresses specifically the reduction of 2–dimensional lattices. A *lattice* of rank 2 in the complex plane $\mathbb{C}$ is the set $\mathcal{L}$ of elements of $\mathbb{C}$ ("vectors") defined by

$$\mathcal{L} = \mathbb{Z}u \oplus \mathbb{Z}v = \{\lambda u + \mu v \mid \lambda, \mu \in \mathbb{Z}\},$$

where (u, v), called a *basis*, is a pair of $\mathbb{R}$–linearly independent elements of $\mathbb{C}$. A lattice is generated by infinitely many bases that are related to each other by integer matrices of determinant ± 1.

Amongst all the bases of a lattice $\mathcal{L}$, some that are called reduced enjoy the property of being formed with "short" vectors. In dimension 2, the best reduced bases are minimal bases that satisfy optimality properties: define u to be a first minimum of a lattice $\mathcal{L}$ if it is a nonzero vector of $\mathcal{L}$ that has smallest Euclidean norm; a second minimum v is any vector amongst the shortest vectors of the lattice that are linearly independent

Fig. 1. A lattice and two of its bases represented by the parallelogram they span. The first basis is skew, the second one is minimal (reduced).

---

of u. Then a basis is *minimal* if it comprises a first and a second miminum. Without loss of generality, one can always assume a minimal basis to be acute, since one of (u, v) and $(u, -v)$ is certainly acute.

A slightly weaker notion will play an important rôle in this paper. A basis is said to be *quasi–minimal* if the triangle that it defines $\langle u, v, u-v\rangle$ contains two mimima of the lattice. Then there is one amongst the six following pairs

$$(u, v), (v, u), (u, u-v), (v, v-u), (u-v, u), (v-u, v) \tag{1}$$

that defines a minimal basis of the lattice. Again, we may restrict attention to quasi–minimal bases that are acute.

The following result gives characterizations of acute minimal bases and acute quasi-minimal bases.

Proposition 1. *Let (u, v) be an acute basis. Then the following two conditions (a) and (b) are equivalent:*

(a) (u, v) is minimal;

(b) (u, v) satisfies the two simultaneous inequalities:

$$(M_1): \ |\frac{u}{v}| \le 1 \qquad \text{and} \qquad (M_2): \ 0 \le \Re(\frac{v}{u}) \le \frac{1}{2}.$$

Let (u, v) be an acute basis. Then the three following conditions $(c), (d), (e)$ are equivalent:

(c) (u, v) is quasi-minimal;

(d) the triangle that (u, v) defines is acute (it has three acute angles);

(e) (u, v) satisfies the two simultaneous inequalities:

$$(Q_1): \ 0 \le \Re(\frac{u}{v}) \le 1 \qquad \text{and} \qquad (Q_2): \ 0 \le \Re(\frac{v}{u}) \le 1.$$

Proof. $(a) \Rightarrow (b)$. It is clear that a minimal basis satisfies (M_1), and if acute also (M_2) since otherwise the vector $v-u$ would be shorter than v.

$(b) \Rightarrow (a)$. Let w be an arbitrary nonzero vector of $\mathcal{L}$, $w = \lambda u + \mu v$. Three cases are to be distinguished depending upon $\mu = 0$, $|\mu| = 1$, and $|\mu| \ge 2$.

Case $\mu = 0$. One has $|w| = |\lambda|\,|u|$ with $\lambda \ne 0$, so that $|w| \ge |u|$.

Case $|\mu| = 1$: Since (M_1) holds, the quantity $|\Re(v/u)|$ is minimal amongst all the $|\Re(w/u)|$ when w lies on the two straight lines corresponding to $\mu = \pm 1$, and thus $|v|$ itself is minimal amongst all the $|w|$ when $w = \lambda u + \mu v$ lies on the two straight lines corresponding to $|\mu| = 1$.

Case $|\mu| \ge 2$. From (M_1) and (M_2), the angle formed by (u, v) is in absolute value between $\pi/3$ and $\pi/2$. Thus, the orthogonal projection $p(v)$ of v on u satisfies $|p(v)| \ge \sqrt{3}/2|v|$. Therefore, for $|\mu| \ge 2$, one has

$$|w| = |\lambda u + \mu v| \ge 2|p(v)| \ge \sqrt{3}|v| > |v|.$$

Finally, (u, v) is a minimal basis of the lattice, and this completes the proof of $(b) \Rightarrow (a)$.

It is also clear that (d) and (e) are equivalent. Moreover, (c) implies (d): The two smaller sides of the triangle form a minimal basis, and satisfy (b), and thus (e), and finally (d).

$(d) \Rightarrow (c)$: The two vectors u and v formed with the two smaller sides of an acute triangle satisfy (b). Then, since (b) implies (a), they form a minimal basis. ■

The Gaussian reduction scheme. In general, a lattice reduction algorithm takes as input an arbitrary basis and produces as output a basis that is reduced. In dimension 2, the stronger notions of minimality and quasi-minimality give rise to two closely related algorithms:

- the standard Gaussian algorithm SGA produces an acute quasi–minimal basis;
- the centered Gaussian algorithm CGA produces an acute minimal basis.

What is common to these two algorithms is an iterative structure aimed at satisfying simultaneously the conditions of Proposition 1. The conditions (M_1) and (Q_1) are simply satisfied by exchanges between vectors. The conditions (M_2) and (Q_2) are met by integer translations of one of the following types:

- for SGA, $v := v - mu$ with $m = \lfloor \Re(v/u) \rfloor$;
- for CGA, $v := \varepsilon(v - mu)$ with $m = \lfloor \Re(v/u) \rceil$ and $\varepsilon = \mathrm{sign}(\Re(v/u) - m)$. (Here, $\lfloor x \rceil$ represents the integer nearest to the real x.)

The complex framework. Many structural characteristics of lattices and bases are invariant under linear transformations —similarity transformations in geometric terms— of the form $S_\lambda \ : \ z \mapsto \lambda z$ with $\lambda \in \mathbb{C} \setminus \{0\}$. An instance is the characterization of minimal and quasiminimal bases given in Proposition 1 that only depends on the ratio $z = v/u$. It is thus natural to consider lattices and bases taken up to equivalence under similarity. For such similarity invariant properties, it is sufficient to restrict attention to lattices generated by a basis of the form $(1, z)$.

In that case, the property for a basis to be minimal and acute corresponds to the fact that z belongs to the so–called *fundamental domain*

$$\mathcal{F} = \left\{ z \ \mid \ |z| \geq 1 \ \text{ and } \ 0 \leq \Re(z) \leq \frac{1}{2} \right\}. \tag{2}$$

Such a domain is familiar from the theory of modular forms [23] or the reduction theory of quadratic forms [22]. Similarly, a quasi-minimal and acute basis is determined by the fact that z belongs to the strip

$$\mathcal{B} = \{ \, z \mid 0 \leq \Re(z) \leq 1 \, \}, \tag{3}$$

without being in the disk $\mathcal{D}$ of diameter $[0, 1]$,

$$\mathcal{D} = \{ \, z \mid \Re(\frac{1}{z}) \geq 1 \}.$$

It should also be observed that exchange operations or translations introduced above only depend on the complex ratio $z = v/u$. Thus the execution traces of the Gaussian algorithms are invariant under similarity. This makes it possible to give a formulation of the Gaussian algorithms CGA and SGA entirely in terms of complex numbers. Let $(u_0, v_0), \ldots, (u_k, v_k)$ be a sequence of bases constructed by one of the Gaussian algorithms. We associate to it the sequence $(1, z_0), \ldots, (1, z_k)$ where $z_j = v_j / u_j$. The geometric transformation effected by each step of the algorithm consists of an exchange $(u, v) \mapsto (v, u)$, a translation $v \mapsto v - m\,u$, and a possible sign change $v \mapsto -v$. In the complex framework, this corresponds to an inversion-symmetry $S \ : \ z \mapsto 1/z$, followed by a translation $z \mapsto T^{-m} z$ with $T(z) = z + 1$, and by a possible sign change $z \mapsto Jz$ where $J(z) = -z$.

Algorithm SGA
Input. A complex z that belongs to disk $\mathcal{D}$ of diameter $[0,1]$
While $(z \in \mathcal{D})$ **do** $z := U(z)$;
Output. A complex z that belongs to domain $\mathcal{B} \setminus \mathcal{D}$ where $\mathcal{B} = \{0 \le \Re(z) \le 1\}$.

Fig. 2. The Standard Gaussian Algorithm [SGA].

In this context, the standard Gaussian algorithm brings z into the already defined strip

$$\mathcal{B} = \{ z \mid 0 \le \Re(z) \le 1 \},$$

while the complete Gaussian algorithm brings z into

$$\widetilde{\mathcal{B}} = \{ z \mid 0 \le \Re(z) \le \frac{1}{2} \},$$

at the expense of a possible sign change.

The standard algorithm. The next sections are devoted to the analysis of the standard Gaussian algorithm, SGA, that we specify now in full. The algorithm SGA produces a quasi-minimal basis whose transformation into a completely minimal basis is trivial (as we will see later), so that its analysis does model the core of the reduction process. At the same time, no sign-change is involved, a feature that gives a much regular structure to the algorithm: in many ways, SGA is to CGA what standard continued fractions are to centered continued fractions, and the close connection with standard continued fractions justifies our terminology.

The algorithm SGA is directed towards bringing z inside the strip $\mathcal{B}$ defined in (3), $\mathcal{B} = \{0 \le \Re(z) \le 1\}$. In order to do so, it suffices to consider a transformation U formed with an inversion-symmetry S and a translation T^{-m} aimed at bringing z into $\mathcal{B}$. It is readily realized that this is achieved by the transformation

$$U(z) = \frac{1}{z} - \lfloor \Re(\frac{1}{z}) \rfloor,$$

with $\lfloor u \rfloor$ the integer part of u. This transformation U is an extension to the complex domain of the operation defining standard continued fraction expansions, for which $U(x) = \frac{1}{x} - \lfloor \frac{1}{x} \rfloor$.

In the rest of the paper, we assume that the Gaussian algorithm is applied to complex numbers z such that $\Im(z) \neq 0$, which corresponds to nondegenerate lattices. One also operates with bases that are acute, so that z belongs to the half-plane $\Re(z) \ge 0$. For reasons already explained, the reduction algorithm takes as input complex numbers from the disk $\mathcal{D}$ of diameter $[0,1]$. The transformation U is then iterated till exit from that disk. The corresponding specification is given in Fig. 2.

For this algorithm, upon exit from the main iteration loop, it is no longer true that that the basis $(1,z)$ is minimal. It is only quasi–minimal and a minimal basis results from applying one amongst 6 possible permutations on the sides of the triangle generated by $(1,z)$, which transform the basis $(1,z)$ into a minimal basis $(1,z')$, where z' belongs to the set corresponding to (1)

$$\{z, \frac{1}{z}, 1-z, \frac{z-1}{z}, \frac{1}{1-z}, \frac{z}{z-1}\}$$

This corresponds to the action of the cross-ratio group and to the fact that the domain $\mathcal{B} \setminus \mathcal{D}$ is the union of six simple transforms of the fundamental domain $\mathcal{F}$ defined in (2), namely

$$\mathcal{B} \setminus \mathcal{D} = \mathcal{F} \cup S\mathcal{F} \cup TJ\mathcal{F} \cup TJS\mathcal{F} \cup STJ\mathcal{F} \cup TJSTJ\mathcal{F}.$$

These domains are represented in Fig. 3.

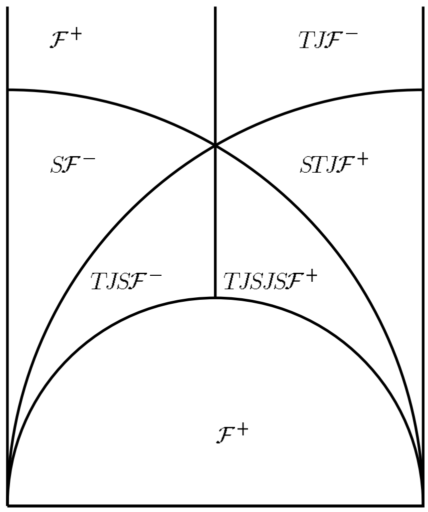

Fig. 3. The upper part of domain $\mathcal{B} \setminus \mathcal{D}$. Here $\mathcal{F}^+$ and $\mathcal{F}^-$ are resp. the upper part and lower parts of $\mathcal{F}$

Thus simply adding a trivial test produces an algorithm whose output is a minimal basis $(1, z)$. In addition, the analysis of the algorithm obtained in this way is then only a trivial variant of the analysis of the core algorithm SGA.

The centered algorithm. The standard algorithm when completed by the final phase just described produces a minimal basis. An alternative, corresponding to a path often taken in the literature, consists in using a centered division algorithm. The corresponding algorithm, CGA can be subjected to an analysis similar to the one exposed here, though more technical. In [4], we briefly point out some of the principles upon which the analysis can be based. It is found there that the average number of iterations equals 1.08922. This is less than the corresponding quantity for the standard algorithm (1.35113), but each iteration is computationally more complicated since a centered remainder routine is needed.

Probabilistic models. The question addressed here is the estimation of the number L of iterations performed by the standard algorithm. The model considered is in essence equivalent to applying the reduction algorithm to random bases, where similar bases are identified.

The *continuous model* is defined by the fact that the inputs are taken uniformly over the definition domain $\mathcal{D}$. The eventual goal is to analyse the behaviour of the algorithm under a *discrete model* where inputs are members of $\mathbb{Q}(i)$ of the form

$$\mathbb{Q}^{\langle N \rangle} = \{\frac{a}{N} + i\frac{b}{N} \mid b \neq 0\},$$

suitably restricted to disk $\mathcal{D}$. The random variable $L^{\langle N \rangle}$ then depends on N. However, as N gets large, it converges, both in moments and distribution, to its continuous counterparts, a fact to be proved in Section 5.

Thus, the results to be enounced later for the continuous model —the average number of iterations is constant and the probability distribution admits exponential tails— carry over to the discrete model. In other words, the behaviour of lattice reduction in dimension 2 is essentially insensitive to the size of the input vectors. This is a notable difference with the one–dimensional case of Euclid's algorithm.

3 Continued fractions and lattice reduction

The Gaussian algorithm is closely related to the linear fractional transformations (also called homographies) that are associated to continued fractions, and thus also to the classical continuant polynomials. The probability distribution, the average cost and the dynamic densities can be expressed as a function of continuants. In this way, a first average-case analysis of the Gaussian algorithm can be given.

The fundamental disks. The algorithm *SGA* produces a sequence $z_0, z_1, \ldots, z_k$ of transforms of $z_0 \in \mathcal{D}$ obtained by iterating the transformation U. As we saw, each step corresponds to a particular transformation

$$z_{j+1} = -m_{j+1} + \frac{1}{z_j} \qquad \text{or} \qquad z_j = \frac{1}{m_{j+1} + z_{j+1}}. \tag{4}$$

While z_j is in $\mathcal{D}$, $1/z_j$ satisfies $\Re(1/z_j) > 1$, so that we have the condition $m_j \geq 1$. Thus, from (4), there results that an execution of the Gaussian algorithm on input z_0 translates into a complex continued fraction expansion

$$z_0 = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots \atop m_k + z_k}}},$$

where the expansion is stopped as soon as z_k lies in $\mathcal{B} \setminus \mathcal{D}$. The number of iterations, L, then assumes the value $k \geq 1$. All the m_j are at least 1.

This leads us to introducing the set $\mathcal{L}_k$ of linear fractional transformations of *depth* k (for $k \geq 1$) defined as the collection of all $h(z)$ of the form

$$h_m(z) = h_{m_1, m_2, \ldots, m_k}(z) = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots \atop m_k + z}}}, \tag{5}$$

where the $m_j \in \mathbb{N} = \{1, 2, \ldots\}$. In the sequel, we denote by $|h|$ the depth of homography h: for an element h of $\mathcal{L}_k$, we have $|h| = k$.

The *event* $[L \geq k+1]$ coincides with the set of complex z such that all the $U^j(z)$, for $j = 0, \ldots, k$, lie in $\mathcal{D}$. As soon as an iterate $U^i(z_0)$ belongs to $\mathcal{B} \setminus \mathcal{D}$, the same property holds for all the further iterates $U^j(z_0)$ for $j \geq i$. Thus, defining $\mathcal{D}_k := U^{(-k)}(\mathcal{D})$ with $\mathcal{D}_0 = \mathcal{D}$, we have $[L \geq k+1] \equiv \mathcal{D}_k$. These domains form an infinite descending chain, $\mathcal{D}_0 \supset \mathcal{D}_1 \supset \mathcal{D}_2 \supset \cdots$. and each $\mathcal{D}_k$ is the disjoint union —up to boundary sets of measure 0— of transforms of $\mathcal{D}$ by the transformations of $\mathcal{L}_k$ of (5),

$$\mathcal{D}_k := [L \geq k+1] = \bigcup_{|h|=k} h(\mathcal{D}).$$

The disk $h(\mathcal{D})$ is the disk of diameter $h([0,1])$. Within the theory of continued fractions, the transform $h([0,1])$ of interval $[0,1]$ by an homography h of depth k is known as a fundamental interval of rank k. A rendering of disks $h(\mathcal{D})$, also called the fundamental disks, is given in Figure 4.

Since fundamental intervals of rank k are disjoint up to boundary sets of measure 0, fundamental disks of rank k are also disjoint in the same sense, and all these considerations imply that, under the uniform probabilistic model of use, the probability ϖ_k that the algorithm performs at least $k+1$ iterations is

$$\varpi_k = \frac{||\mathcal{D}_k||}{||\mathcal{D}||} = \frac{4}{\pi} \sum_{|h|=k} ||h(\mathcal{D})|| = \sum_{|h|=k} |h(0) - h(1)|^2, \tag{6}$$

where $||A||$ denotes the area of a domain $\mathcal{A}$ of the plane.

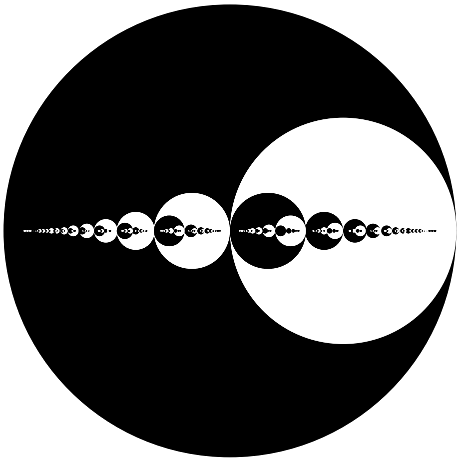

Fig. 4. The domains $\mathcal{D}_0 \setminus \mathcal{D}_1$, $\mathcal{D}_1 \setminus \mathcal{D}_2$, $\mathcal{D}_2 \setminus \mathcal{D}_3$, $\mathcal{D}_3 \setminus \mathcal{D}_4$, $\mathcal{D}_4 \setminus \mathcal{D}_5$ represented alternatively in black and white. (The largest disk is $\mathcal{D}_0 \equiv \mathcal{D}$ which is the disk of diameter $[0,1]$.)

Continuants. Homographies of $\mathcal{L}_k$ are naturally associated with continued fractions of depth k themselves expressible in terms of *continuants*, [11, p. 340] [21]. The continuant polynomials are defined by

$$Q_k(x_1, x_2, \ldots, x_k) = x_k Q_{k-1}(x_1, \ldots, x_{k-1}) + Q_{k-2}(x_1, \ldots, x_{k-2}), \tag{7}$$

with $Q_0 = 1$, $Q_1(x_1) = x_1$. As is well–known the continuant polynomial $Q_k(x_1, \ldots, x_k)$ is also the sum of all monomials that obtain by crossing out pairs $x_i x_{i+1}$ of consecutive variables in the product $x_1 x_2 \cdots x_k$. Continuants thus satisfy the symmetry property $Q_k(x_1, \ldots, x_k) = Q_k(x_k, \ldots, x_1)$

Classically, a function $h_m \in \mathcal{L}_k$ with $m = (m_1, \ldots, m_k)$ admits the expression

$$h_m(z) = \frac{P_k + zP_{k-1}}{Q_k + zQ_{k-1}},$$

where the four coefficients $Q_k(h), Q_{k-1}(h), P_k(h), P_{k-1}(h)$ can be expressed as elements of family Q_k

$$\begin{array}{ll} Q_k(h) = Q_k(m_1, \ldots, m_k), & Q_{k-1}(h) = Q_{k-1}(m_1, \ldots, m_{k-1}), \\ P_k(h) = Q_{k-1}(m_2, \cdots, m_k), & P_{k-1}(h) = Q_{k-2}(m_2, \ldots, m_{k-1}). \end{array} \tag{8}$$

For an homography h that is associated to $(m_1, \ldots, m_k)$, denote by $\widehat{h}$ the homography associated to $(m_k, \ldots, m_1)$. One has

$$Q_k(\widehat{h}) = Q_k(h), \quad Q_{k-1}(\widehat{h}) = P_k(h). \tag{9}$$

Note also the determinant identity

$$Q_k P_{k-1} - Q_{k-1} P_k = (-1)^k. \tag{10}$$

The diameter of a fundamental disk $h(\mathcal{D})$ is the interval $h([0,1])$. For $h \in \mathcal{L}_k$, the length ρ_h of this diameter can be solely expressed from (10) with continuants Q_k and Q_{k-1}

$$\rho_h = |h(0) - h(1)| = \frac{1}{Q_k(Q_k + Q_{k-1})}. \tag{11}$$

This simple fact has two main consequences, both for the worst-case analysis, and for the average case-analysis of the *SGA* algorithm.

Worst-case analysis. We digress and show that the *SGA* algorithm always terminates for complex numbers z that are not real. The corresponding bounds intervene in our later analysis of the discrete model.

Proposition 2. *The domain $\mathcal{D}_k := U^{(-k)}\mathcal{D}$ is a subset of the horizontal strip*

$$\mathcal{I}_k = \{z \mid |\Im z| \leq \frac{1}{\phi^{2k}}\},$$

where $\phi = (1+\sqrt{5})/2$ is the golden ratio. On a nonreal input z ($\Im z \neq 0$), the number $L(z)$ of iterations of algorithm SGA satisfies

$$L(z) \;\leq\; \lceil \frac{1}{2} \log_\phi \frac{1}{|\Im z|} \rceil. \tag{12}$$

Proof. The smallest continuant Q_k is obtained, from (7), when all the m_j are equal to 1. Then, Q_k reduces to the $(k+1)$st Fibonacci number ϕ_{k+1}, defined by $\phi_0 = 0, \phi_1 = 1$ and $\phi_{k+1} = \phi_k + \phi_{k-1}$. If ϕ is the golden ratio, $\phi = (1+\sqrt{5})/2$, the $(k+1)$st Fibonacci number ϕ_{k+1} satisfies $\phi_{k+1} \geq \phi^{k-1}$. From (11), we deduce that the radius of the fundamental disk $h(\mathcal{D})$ satisfies, for all h of depth k,

$$\frac{\rho_h}{2} \leq \frac{1}{\phi^{2k}},$$

which shows the first assertion. If now the complex number z satisfies $|\Im z| \geq \phi^{-2k}$, then z cannot belong to $\mathcal{D}_k$, and $L(z)$ is at most equal to k, and we obtain (12). ■

Probabilistic analysis. The considerations above permit us to express the probability distribution of the Gaussian algorithm in terms of continuants.

Theorem 3. *The probability ϖ_k that algorithm SGA performs more than k iterations on a random input $z \in \mathcal{D}$ is expressible as*

$$\varpi_k := \Pr[L \geq k+1] = \sum_{|h|=k} \frac{1}{Q_k^2 (Q_k + Q_{k-1})^2},$$

where $Q_k = Q_k(m_1, \ldots, m_k)$, $Q_{k-1} = Q_{k-1}(m_1, \ldots, m_{k-1})$, and the sum is over all integers $m_j \geq 1$.

Proof. This statement results from combining the expression of diameter of $h(\mathcal{D})$ given in (11) with the form (6) already found for the probability distribution. ■

The following table displays the probability distribution of *SGA* computed by Theorem 3 and the numerical methods of Section 6 against the result of 10^8 simulations of the algorithm.

k	$\Pr[L \geq k+1]$	Simulations
1	0.28986	0.28984361
2	0.04848	0.04847104
3	0.01027	0.01027170
4	0.00200	0.00200478
5	0.00040	0.00040299
6	0.00008	0.00008031
7	0.00002	0.00001569
Expectation:	1.35113	1.351094

Average-case analysis. The mean number of iterations of the lattice reduction algorithm of Gauss admits a form that no longer involves continuants. We will see later in Section 6 that it is related in fact to a number of remarkable constants including the tetralogarithm of argument 1/2 and $\zeta(3)$.

Theorem 4. *The mean number of iterations $\varpi := E[L]$ of algorithm SGA applied to a random $z \in \mathcal{D}$ is given by the double sums,*

$$\varpi = \frac{5}{4} + \frac{180}{\pi^4} \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c < 2d} \frac{1}{c^2} = \frac{3}{4} + \frac{180}{\pi^4} \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c \leq 2d} \frac{1}{c^2} \tag{13}$$

Proof. By a standard transformation, the expected cost is

$$\varpi = \sum_{k \geq 1} \Pr[L \geq k] = 1 + \sum_{k \geq 1} \varpi_k,$$

and the last sum represents a sum over all possible values of continuants. The formula (13) results from simple number-theoretic arguments: each rational number c/d of the interval $]0,1[$ admits two continued fraction forms, the proper one (that finishes with $m_k \geq 2$) and the improper one (that finishes with $m_{k+1} = 1$. Thus, c/d can be represented in two different ways as a ratio of continuants P_k/Q_k. Accordingly, any integer pair (c,d) satisfying $\gcd(c,d) = 1$, $d \geq 2$, and $0 < c < d$ can be written in two different ways as a pair (P_k, Q_k). or equivalently as a pair (Q_{k-1}, Q_k), given the general properties (9) of continuants. In this manner, taking into account boundary cases, namely numbers 0 and 1, we find

$$\varpi = \frac{5}{4} + 2 \sum_{\substack{d \geq 2,\ 0 < c < d \\ \gcd(c,d)=1}} \frac{1}{d^2 (c+d)^2}.$$

The general term in the last sum is homogeneous of degree 4, so that the gcd condition is eliminated provided one divides the sum by $\zeta(4) = \pi^4/90$,

$$\varpi = \frac{5}{4} + \frac{2}{\zeta(4)} \beta \quad \text{with} \quad \beta := \sum_{d \geq 2} \frac{1}{d^2} \sum_{1 \leq c < d} \frac{1}{(c+d)^2} = \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c < 2d} \frac{1}{c^2}. \tag{14}$$

Now, adding the terms that correspond to $d = 1$ and $c = d$, we obtain

$$\varpi = \frac{3}{4} + \frac{2}{\zeta(4)} \beta' \quad \text{with} \quad \beta' := \sum_{d \geq 1} \frac{1}{d^2} \sum_{1 \leq c \leq d} \frac{1}{(c+d)^2} = \sum_{d \geq 1} \frac{1}{d^2} \sum_{d < c \leq 2d} \frac{1}{c^2}. \tag{15}$$

The constants β and β' are linked by the relation $\beta' - \beta = \zeta(4)/4$. ■

Dynamic analysis. We have already mentioned the importance of the invariant measure of Gauss in the Euclidean algorithm. No such invariant measure can exist here as the reduction algorithm terminates with probability 1. However, a rôle quite similar to the invariant measure of Gauss is played by a function that describes the limit distribution of successive transforms of the input as the reduction algorithms proceeds.

Initially, the input distribution is uniform inside the disk $\mathcal{D}$, so that to z_0 is associated the constant density function over $\mathcal{D}$. Assume now that the algorithm performs at least $k+1$ iterations. Then the kth iterate z_k is an element of $\mathcal{D}$. A natural question is to determine its distribution inside $\mathcal{D}$. The density of the k-th iterate z_k at point z is proportional to

$$\lim_{r \to 0} \frac{1}{\pi r^2} \Pr\left[z_k \in D(z,r)\right] = \lim_{r \to 0} \frac{1}{\pi r^2} \Pr\left[U^{(-k)}(D(z,r))\right]$$

where $D(z,r)$ is the disk of center z and radius r. The set $U^{(-k)}(D(z,r))$ is the disjoint union of all the disks $h(D(z,r))$ for h in $\mathcal{L}_k$. Thus the density of the k-th iterate z_k is also proportional to

$$\widetilde{F}_k(z) := \lim_{r\to 0} \frac{1}{\pi r^2} \sum_{|h|=k} ||h(D(z,r))||. \tag{16}$$

The proportionality factor is taken so as to ensure that the integral of the density over $\mathcal{D}$ equals 1, and the legitimate definition of the *conditional* density function inside $\mathcal{D}$ after k iterations of the algorithm is

$$F_k(z) = \frac{4}{\pi \varpi_k} \widetilde{F}_k(z), \tag{17}$$

since

$$\frac{\pi}{4}\varpi_k = ||U^{(-k)}\mathcal{D}|| = \iint_{\mathcal{D}} \widetilde{F}_k(z)dxdy.$$

We shall call F_k the *dynamic density* (of order k) of the algorithm.

Theorem 5. *The dynamic density F_k is given by*

$$F_k(z) = \frac{4}{\pi\varpi_k} \sum_{|h|=k} \frac{1}{|Q_{k-1}z + Q_k|^4}, \qquad \textit{where} \qquad \varpi_k = \Pr[L \geq k+1].$$

Proof. The jacobian of the linear transformation h is equal to $|h'(z)|^2$, so that

$$||h(D(z,r)|| \sim \pi r^2 |h'(z)|^2,$$

when r tends to 0. From the determinant equality, we have

$$|h'(z)|^2 = \frac{1}{|Q_{k-1}z + Q_k|^4},$$

and from summing over all h of depth k, with (16),

$$\widetilde{F}_k(z) = \sum_{|h|=k} \frac{1}{|Q_{k-1}z + Q_k|^4}. \tag{18}$$

We come back to F_k by (17). ■

Proposition 6. *The following functional relation holds between the dynamic densities F_k and F_{k-1},*

$$\frac{\varpi_k}{\varpi_{k-1}} F_k(z) = \sum_{m\geq 1} \frac{1}{|m+z|^4} F_{k-1}\left(\frac{1}{m+z}\right).$$

Proof. Using (18) and (7), we isolate the last component $m = m_k$ in the k-uple $m = (m_1, m_2 \ldots, m_k)$ associated to h, and get

$$\widetilde{F}_k(z) = \sum_{m\geq 1} \sum_{|h|=k-1} \frac{1}{|Q_{k-1}z + mQ_{k-1} + Q_{k-2}|^4}$$

$$= \sum_{m\geq 1} \frac{1}{|m+z|^4} \sum_{|h|=k-1} \frac{1}{|Q_{k-1} + \frac{Q_{k-2}}{m+z}|^4} = \sum_{m\geq 1} \frac{1}{|m+z|^4} \widetilde{F}_{k-1}\left(\frac{1}{m+z}\right).$$

We come back to F_k and F_{k-1} with (17). ■

Thus, assuming that F_k admits a limit F_∞ and that ratio ϖ_{k+1}/ϖ_k converges to some constant λ, the quantities λ and F_∞ satisfy

$$\lambda F_\infty(z) = \sum_{m\geq 1} \frac{1}{|m+z|^4} F_\infty(\frac{1}{m+z}).$$

So, the limit objects λ and F_∞ must be an eigenvalue and a corresponding eigenvector of the operator $\mathcal{H}$ defined by the right hand side of the previous identity, namely

$$\mathcal{H}[f](z) = \sum_{m\geq 1} \frac{1}{|m+z|^4} f(\frac{1}{m+z}).$$

In particular, the restriction of F_∞ to the real axis must be an eigenvalue of the operator $\mathcal{G}$ such that

$$\mathcal{G}[f](x) = \sum_{m\geq 1} \frac{1}{(m+x)^4} f(\frac{1}{m+x}).$$

This sharply motivates the introduction of the operator $\mathcal{G}$ in the next section, where we shall also establish the assumptions regarding F_k and ϖ_k.

4 The $\mathcal{G}$ operator

The complete analysis of the probability distribution and of the dynamics of the Gaussian algorithm depends on the introduction of an operator $\mathcal{G}_s$ formally defined by

$$\mathcal{G}_s[f](t) = \sum_{m\geq 1} \frac{1}{(m+t)^s} f(\frac{1}{m+t}), \tag{19}$$

and more specifically on the instance $s=4$ that we simply denote by $\mathcal{G} := \mathcal{G}_4$. (Note that continued fractions and the Euclidean algorithm correspond to the case $s=2$.) In fact, Proposition 6 involves a modified operator

$$\mathcal{H}_s[f](t) = \sum_{m\geq 1} \frac{1}{|m+t|^s} f(\frac{1}{m+t}), \tag{20}$$

and more specifically the instance $s=4$. However, in the case when the initial density F_0 is uniform, Proposition 7 will show that it is sufficient to study operator $\mathcal{G}$, which can be viewed as the holomorphic version of $\mathcal{H} := \mathcal{H}_4$. The general case when the initial density F_0 is no longer uniform requires properties of the whole family $\mathcal{H}_s$ and is treated in [27].

Such operators $\mathcal{G}_s$ are well-known under the name of Ruelle-Mayer operators. They permit to generate continuants, and at the same time their spectral properties, related to the Perron–Frobenius theory, have nice consequences for the average analysis of the Gaussian algorithm.

Ruelle-Mayer operators and continuants. There is a close relationship between the iterates of $\mathcal{G}_s$ and continuants. Operator $\mathcal{G}_s$ can also be expressed as a sum on all the homographies of depth 1

$$\mathcal{G}_s[f](t) = \sum_{|h|=1} \widetilde{h}(t)^s f\circ h(t),$$

where $\widetilde{h}(t)$ is the holomorphic function that coincides with $\sqrt{|h'(t)|}$ on the interval $\mathcal{J}$. The iterate of order k of $\mathcal{G}_s$ uses all the homographies of depth k

$$\mathcal{G}_s^k[f](t) = \sum_{|h|=k} \widetilde{h}(t)^s f\circ h(t), \tag{21}$$

and generates the continuants of depth k in the following sense:

$$\mathcal{G}_s^k[f](t) = \sum_{|h|=k} \frac{1}{(Q_{k-1}t + Q_k)^s} f(\frac{P_{k-1}t + P_k}{Q_{k-1}t + Q_k});$$

in particular, using the symmetry relation (9) between Q_k and P_{k-1},

$$\mathcal{G}_s^k[f](0) = \sum_{|h|=k} \frac{1}{Q_k^s} f(\frac{P_k}{Q_k}) = \sum_{|h|=k} \frac{1}{Q_k^s} f(\frac{Q_{k-1}}{Q_k}). \tag{22}$$

The probability distribution and the dynamic densities involved in Theorem 3 and Theorem 5 precisely admit such expressions.

Proposition 7. *The probability distribution ϖ_k and the dynamic density F_k can be expressed as a function of iterates of order k of $\mathcal{G}$:*

$$\varpi_k := \Pr[L \geq k+1] = \mathcal{G}^k[u(t)](0) \qquad \text{where} \qquad u(t) = \frac{1}{(1+t)^2}$$

$$F_k(z) = \frac{1}{\varpi_k}\mathcal{G}^k[v_z(t)](0) \qquad \text{where} \qquad v_z(t) = \frac{1}{(1+tz)^2(1+t\bar{z})^2}.$$

The expectation ϖ admits the following expression

$$\varpi = (I - \mathcal{G})^{-1}[u](0).$$

Proof. Compare expressions involved in Theorem 3 and Theorem 5 with (22).

Analytic properties of the $\mathcal{G}_s$ operators have been investigated in detail by Mayer and we globally refer to [19] and references therein. Ruelle-Mayer operators enjoy three main properties: In suitable Banach spaces, they are nuclear of order 0; in convenient Hilbert spaces, they are isomorphic to integral operators, with Bessel functions as kernels, and they are diagonalizable; furthermore, for s real, their spectrum is real, and they verify a Perron-Frobenius property, which proves the existence of dominant spectral objects.

Let $\mathcal{J}$ denote a bounded open interval that contains strictly the segment $[0,1]$ and $\mathcal{V}$ the open disk with diameter $\mathcal{J}$. We require that $\mathcal{V}$ is strictly mapped inside itself by all the homographies of depth 1, *i.e.*,

$$h(\overline{\mathcal{V}}) \subset \mathcal{V} \text{ for } |h| = 1.$$

For all s with $\Re(s) > 1$, the operator $\mathcal{G}_s$ acts on the space $A_\infty(\mathcal{V})$ of functions f that are holomorphic in $\mathcal{V}$ and continuous on the closure $\overline{\mathcal{V}}$ of $\mathcal{V}$. The set $A_\infty(\mathcal{V})$ endowed with the sup–norm is a Banach space. Note that functions u and ν_z involved in Proposition 7 are elements of such a space.

Remark. Usually, one chooses for $\mathcal{J}$ and $\mathcal{V}$ the open interval and the open disk of center 1 and radius $3/2$, respectively. This choice is however not intrinsic, and all the results of this section can be easily adapted to other configurations of $\mathcal{J}$ and $\mathcal{V}$ provided that $\mathcal{J}$ is bounded, contains strictly the segment $[0,1]$, $\mathcal{V}$ is the disk of diameter $\mathcal{J}$, and $\mathcal{V}$ is strictly mapped inside itself by all the homographies of depth 1. A typical configuration is based on an interval $\mathcal{J}$ of the form $]-\delta, 1+3\delta[$ with $0 < \delta \leq 1/2$.

Nuclearity. Let B be a Banach space and $B^\star$ its dual space. An operator $\mathcal{M} : B \to B$ is *nuclear* of order 0 if it admits a representation

$$\mathcal{M}[f] = \sum_{i \in I} \mu_i \; e_i^\star(f) \; e_i \quad \text{for all } f \in B,$$

with $e_i \in B$, $e_i^\star \in B^\star$ such that $||e_i|| = ||e_i^\star|| = 1$ and the μ_i are p-summable for all $p > 0$ (*i.e.*, $\sum |\mu_i|^p < +\infty$).

Such operators have been introduced and studied by Grothendieck [7] [8]. They are compact and they have a discrete spectrum. Moreover, most of matrix algebra can be extended to them; in particular, one can define the trace of such an operator, and also the analogue of the characteristic polynomial known as the Fredholm determinant.

Some spaces have the nice property that every bounded map is nuclear of order 0. A typical example of such a space is the space $A_\infty(\mathcal{V})$. Since each operator $\mathcal{G}_s$ can be expressed as a convergent sum of bounded operators, the operator $\mathcal{G}_s$ is itself bounded, and thus nuclear of order 0.

Spectral decomposition. Ruelle-Mayer operators enjoy stronger properties when they operate on restricted spaces of functions holomorphic in half-planes, which can be endowed with a Hilbert space structure, and are called Hardy spaces. This point of view has been adopted first by Babenko, and further generalized by Mayer, whom we follow here.

Let us consider the half-plane $\mathcal{B}_\delta = \{z \mid \Re(z) > \delta\}$ and the Hilbert space H_s of holomorphic functions f on $\mathcal{B}_{-1/2}$, bounded on in each of the half-planes $\mathcal{B}_\delta$ $(\delta > -1/2)$ which admit an integral representation

$$f(z) = \int_0^\infty e^{-wz}\ \varphi(w)\ w^{(s-1)/2} \frac{dw}{e^w - 1}$$

where φ is square integrable with respect to measure $dm(w)$ with density $1/(e^w - 1)$.

The norm in H_s is defined as

$$|f|_{<s>} = \int_0^\infty |\varphi(w)|^2 \frac{dw}{e^w - 1},$$

and satisfies

$$|f(z)| \le r(z,s)\ |f|_{<s>}, \tag{23}$$

with a constant $r(z,s)$ that depends on s and $z \in \mathcal{B}_{-1/2}$. On the space H_s, the operator $\mathcal{G}_s$ can be expressed as

$$\mathcal{G}_s[f](z) = \int_0^\infty e^{-wz}\ \mathcal{K}_s[\varphi](w)\ w^{(s-1)/2} \frac{dw}{e^w - 1}$$

where $\mathcal{K}_s$ is an integral operator that involves the Bessel function J_{s-1} of index $s-1$

$$J_{s-1}(u) = \sum_{k=0}^\infty (\frac{u}{2})^{2k+s-1}\ \frac{(-1)^k}{k!\ \Gamma(k+s)}$$

under the form

$$\mathcal{K}_s[\varphi](w) = \int_0^\infty J_{s-1}(2\sqrt{vw})\ \varphi(v)\ \frac{dv}{e^v - 1}$$

In the space H_s, $\mathcal{G}_s$ is thus isomorphic to an integral operator, whose kernel is the Bessel function of index $s-1$. This representation shows that $\mathcal{G}_s$ is diagonalizable on H_s and it gives the spectral decomposition of the iterates $\mathcal{G}_s^k$ of $\mathcal{G}_s$ on the space H_s

$$\mathcal{G}_s^k[f] = \sum_i \lambda_{i,s}^k\ f_{i,s}^\star[f]\ f_{i,s}, \tag{24}$$

for all $k \ge 0$ and f in H_s. Here, the $\lambda_{i,s}$ are the eigenvalues of $\mathcal{G}_s$ taken in order of decreasing moduli, the functions $f_{i,s}$ are the associated eigenvectors, and $f_{i,s}^\star$ is the dual basis of $f_{i,s}$.

Since the function $\mathcal{G}_s[f]$ belongs to H_s as soon as f belongs to $A_\infty(\mathcal{V})$, the two spectra, the spectrum of $\mathcal{G}_s$ and the spectrum of its restriction to H_s, are the same. Moreover, for real s, the spectrum of $\mathcal{G}_s$ is real. Using the relations (23) and (24), one derives the spectral decomposition of the iterates $\mathcal{G}_s^k[f](t)$ of $\mathcal{G}_s$

$$\mathcal{G}_s^k[f](t) = \sum_i \lambda_{i,s}^k\ f_{i,s}^\star[f]\ f_{i,s}(t), \tag{25}$$

for all $k \ge 1$, f in $A_\infty(\mathcal{V})$ and $t \in \mathcal{V}$. Here, the $\lambda_{i,s}$ are the eigenvalues of $\mathcal{G}_s$ taken in order of decreasing moduli, the functions $f_{i,s}$ are the associated eigenvectors, and $f_{i,s}^\star$ is the dual basis of $f_{i,s}$ in H_s that one extends by

$$f_{i,s}^\star[f] := \frac{1}{\lambda_{i,s}}\ f_{i,s}^\star[\mathcal{G}_s[f]]$$

for f in to $A_\infty(\mathcal{V})$.

Positivity. Operators called u_0–positive have been introduced by Krasnoselsky [12]; they give a generalization of positive operators for finite dimensional spaces, and possess dominant spectral properties.

A set K in a real space B is called a proper cone if (i) for all $\rho > 0$ and all $f \in K$, $\rho f \in K$, (ii) $K \cap -K = \{0\}$. A proper cone is called reproducing if $B = K - K$, *i.e.*, every element f in B is equal to the difference of elements of K.

Let K be a proper, reproducing cone, with a non-empty interior $\operatorname{Int} K$. A linear operator $\mathcal{M} : B \to B$ is positive (with respect to cone K) if $\mathcal{M}(K)$ is a subset of K. Let u_0 be an element of $\operatorname{Int} K$; a positive operator $\mathcal{M}$ is u_0–positive with respect to cone K if there exists for every $f \neq 0$ in K a number p and reals $\alpha, \beta > 0$ such that

$$\beta\, u_0 \leq \mathcal{M}^p[f] \leq \alpha\, u_0,$$

where the order $\leq$ is related to K : $f \leq g$ if and only if $g - f \in K$.

Krasnoselsky [12] showed that a compact and u_0–positive operator satisfies a Perron-Frobenius property: it has a unique eigenvector $g \in \operatorname{Int} K$ and the associated eigenvalue λ is simple, positive, and in absolute value strictly larger than all other eigenvalues.

For real s $(s > 1)$, operator $\mathcal{G}_s$ leaves invariant the space $R_\infty(\mathcal{V})$ of elements of $A_\infty(\mathcal{V})$ which are real on $\mathcal{J}$; Space $R_\infty(\mathcal{V})$ is a real Banach space and the set K of functions f of $R_\infty(\mathcal{V})$ whose restriction to $\mathcal{J}$ is positive is a proper, reproducing cone K, with a non-empty interior $\operatorname{Int} K$. Function $u_0 = 1$ belongs to $\operatorname{Int} K$ and the restriction of $\mathcal{G}_s$ to $R_\infty(\mathcal{V})$ is u_0–positive with respect to cone K [19]. Krasnoselsky's theorem can be applied, and, since the spectrum of $\mathcal{G}_s$ is real, the two spectra —the spectrum of $\mathcal{G}_s$ and the spectrum of the restriction of $\mathcal{G}_s$ to $R_\infty(\mathcal{V})$— are the same.

Dominant spectral properties. From the previous facts, there results that the operator $\mathcal{G}_s$ has a Perron-Frobenius property.

Theorem 8. *[Mayer] For real $s > 1$, the operator $\mathcal{G}_s : A_\infty(\mathcal{V}) \to A_\infty(\mathcal{V})$ has a positive dominant eigenvalue $\lambda_s := \lambda_{1,s}$ which is simple and strictly larger than the other eigenvalues in absolute value. The corresponding eigenfunction f_s is strictly positive on $\mathcal{J}$, and is normalized by $f_s(0) = 1$. The adjoint operator $\mathcal{G}_s^\star : A_\infty^\star(\mathcal{V}) \to A_\infty^\star(\mathcal{V})$ has a positive eigenfunctional $f_s^\star$ with eigenvalue λ_s such that $f_s^\star(f) > 0$ if $f > 0$ on $\mathcal{J}$.*

If $\mathcal{P}_s$ denotes the projection on the dominant eigensubspace, $\mathcal{P}_s = f_s^\star \otimes f_s$, and μ_s denotes a subdominant eigenvalue of $\mathcal{G}_s$, then for every element f of $A_\infty(\mathcal{V})$ that is strictly positive on $\mathcal{J}$, for every $t \in \mathcal{J}$, one has

$$\mathcal{G}_s^k[f](t) = \lambda_s^k \mathcal{P}_s[f](t) \left[1 + \mathcal{O}(|\frac{\mu_s}{\lambda_s}|^k)\right], \quad \text{for } k \to \infty \tag{26}$$

where the implied constant in $\mathcal{O}$ error term depends on f, s and t.

For $s = 2$, the operator $\mathcal{G}_s$ has dominant spectral properties that are very well known: the dominant eigenvalue λ_2 is equal to 1, the subdominant eigenvalue, which is also simple, is equal to $\mu_2 \approx -0.303663$, the famous Gauss-Kuzmin-Wirsing constant; the dominant eigenvector f_2, which corresponds to the limit-density of the continued fraction algorithm, and the dominant eigenvector $f_2^\star$ of the adjoint operator are both explicit and equal respectively to

$$f_2(z) = \frac{1}{1+z} \quad \text{and} \quad f_2^\star[f] = \frac{1}{\log 2} \int_0^1 f(x)dx.$$

For other values of parameter s, the operators $\mathcal{G}_s$ have been less studied and to the best of our knowledge, they have never been used for any $s \neq 2$. In fact, for $s \neq 2$, the dominant spectral objects of operators $\mathcal{G}_s$ do not seem to be related to any known special functions. We will give later (in Section 6) some numerical estimates of these objects in the case $s = 4$.

In general, from (26), the dominant eigenvalue λ_s can be obtained as

$$\lambda_s = \lim_k \, (\mathcal{G}_s^k[f](x))^{1/k} \tag{27}$$

for any analytic function f that is strictly positive on $\mathcal{J}$ and any x of $\mathcal{J}$. Choosing $f = 1$ and $x = 0$ and using (22), one gets

$$\lambda_s = \lim_k \, (\sum_{Q_k} \frac{1}{Q_k^s})^{1/k}.$$

The smallest continuant Q_k is the $(k+1)$st Fibonacci number ϕ_k that we already used in Proposition 2. It satisfies $\lim_k \phi_k^{1/k} = \phi$ where ϕ is the golden ratio $(1+\sqrt{5})/2$. This shows that,

$$\text{for} \quad u \geq 0, \;\; \lambda_{s+u} \leq \frac{1}{\phi^u} \, \lambda_s, \tag{28}$$

so that the map $s \mapsto \lambda_s$ defines a strictly decreasing function of s.

In the sequel, we are going to use dominant spectral properties of $\mathcal{G}_s$ (for $s = 4$) to derive an asymptotic behaviour for the probability distribution and for the dynamic densities of the Gaussian algorithm.

Asymptotic behaviour of probability distribution. The distribution of the number of iterations of Gaussian algorithm is closely approximated by a geometrical law, whose ratio is equal to the dominant eigenvalue λ_4 of $\mathcal{G}$.

Theorem 9. *There exist real numbers c_i such that, for all $k \geq 1$,*

$$\varpi_k := \Pr[L \geq k+1] = \sum_i c_i \; \lambda_{i,4}^k,$$

where the $\lambda_{i,4}$ are the eigenvalues of $\mathcal{G}$ ordered with respect to decreasing moduli. In particular, one has asymptotically:

$$\varpi_k = c\,\lambda_4^k \left[1 + \mathcal{O}(|\frac{\mu_4}{\lambda_4}|^k)\right].$$

Here, $\lambda_4 := \lambda_{1,4}$ is the dominant eigenvalue of $\mathcal{G}$, $\mu_4 := \lambda_{2,4}$ is a subdominant eigenvalue of $\mathcal{G}$. Numerically, one has:

$$\lambda_4 \approx 0.1994, \;\; |\mu_4| \leq 0.082, \;\; c \approx 1.3.$$

This theorem is in accordance with observation of the numerical data following Theorem 3, as the probabilities decay roughly like $(1/5)^n$. We come back later in Section 6 to the numerical estimates of the two dominant eigenvalues of $\mathcal{G}$.

Proof. We use Proposition 7, Theorem 8, and the spectral decomposition given in (25) associated with u defined in Proposition 7. The real c_i are defined by $c_i = f_{i,4}^\star[u]$ and Theorem 8 proves that the real $c := c_1$ is strictly positive. ■

Limit density. The limit density is an eigenfunction of the nonanalytic operator $\mathcal{H}$ and it is expressible in terms of the dominant eigenfunction f_4 of the analytic operator $\mathcal{G}$.

Theorem 10. *The dynamic density $F_k(z)$ converges geometrically to a limit density F_∞:*

$$F_\infty(x+iy) = \alpha \int_{-1}^{1} (1-w^2) \; f_4(x+iyw)\,dw.$$

There, f_4 is the dominant eigenfunction of $\mathcal{G}$ and α the normalization constant determined by $\iint F_\infty \,dx\,dy = 1$. The limit density F_∞ is an eigenfunction of the operator $\mathcal{H}$ and, on the real axis, it is proportional to the dominant eigenfunction f_4 of $\mathcal{G}$.

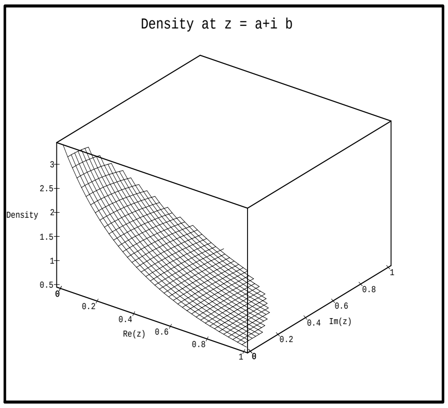

Fig. 5. The limit density F_∞.

Proof. From Proposition 7 and Theorem 8, it is clear that the limit density F_∞ exists and is proportional to the function $z \mapsto f_4^\star[\nu_z]$, since

$$F_\infty(z) = \lim_{k\to\infty} \frac{4}{\pi\varpi_k}\mathcal{G}^k[\nu_z](0) = \frac{4}{c\pi} f_4^\star[\nu_z].$$

Moreover, for real x, the function $F_\infty(x)$ is an eigenfunction of $\mathcal{G}$ relative to eigenvalue λ_4, and it is thus proportional to f_4. Thus, by the equality above and for x real, the three quantities $f_4^\star[\nu_x]$, $f_4(x)$ and $F_\infty(x)$ are proportional; then F_∞ extends the function f_4 in the same way as ν_z extends ν_x outside the real axis. For $z = x + iy \in \mathcal{D}$ and $t \in \mathcal{V}$, we have

$$\nu_z(t) = \frac{1}{(1+xt)^4}\frac{1}{[1+(\frac{yt}{1+xt})^2]^2} = \sum_k (-1)^k (k+1)\ y^{2k}\frac{t^{2k}}{(1+xt)^{4+2k}}.$$

Using

$$\frac{t^{2k}}{(1+xt)^{4+2k}} = \frac{3!}{(2k+3)!}\frac{d^{2k}}{dx^{2k}}(\frac{1}{1+xt})^4,$$

we obtain

$$\nu_z(t) = \sum_k \frac{(-1)^k}{(2k)!}\ \frac{3y^{2k}}{(2k+1)(2k+3)}\frac{d^{2k}}{dx^{2k}}\nu_x(t).$$

Since x is here a parameter, the linear form $f_4^\star$ commutes with derivation with respect to x, and, since $f_4^\star[\nu_x]$ is proportional to $f_4(x)$, we deduce that $f_4^\star(\nu_z)$ is proportional to

$$\sum_k \frac{(-1)^k}{(2k)!}\ \frac{3y^{2k}}{(2k+1)(2k+3)}\frac{d^{2k}}{dx^{2k}}f_4(x).$$

Finally, with α as a normalization constant,

$$F_\infty(z) = \alpha\int_{-1}^{1}(1-w^2)\ f_4(x+iyw)dw$$

that can also be written

$$F_\infty(z) = \alpha\int_\gamma \frac{(w-z)(w-\bar{z})}{(z-\bar{z})^2}f_4(w)dw$$

for any simple path that links z to $\bar{z}$. With the last form, one can easily check that F_∞ is an eigenfunction of operator $\mathcal{H}$. ■

5 The discrete model

The analysis of the standard algorithm under the discrete model where inputs are taken from the discrete set

$$\mathbb{Q}^{\langle N\rangle} = \{\frac{a}{N} + i\frac{b}{N} \mid b \neq 0\} \cap \mathcal{D}, \tag{29}$$

derives from the previous analyses thanks to "Gauss's principle". This principle relates the number of integer points in a domain to the area of the domain: a disk of radius ρ contains $\pi N^2\rho^2 + \mathcal{O}(\rho N + 1)$ lattice points of $\mathbb{Q}^{\langle N\rangle}$. From worst–case bounds of Proposition 2, we know also that the reduction algorithm performs a number of iterations that is $\mathcal{O}(\log N)$.

Theorem 11. *Let $L^{\langle N\rangle}$ be the number of iterations of the standard Gaussian algorithm applied to random inputs from $\mathbb{Q}^{\langle N\rangle} \cap \mathcal{D}$. The random variable $L^{\langle N\rangle}$ admits the following upper-bound (which corresponds to worst-case analysis)*

$$L^{\langle N\rangle}(z) \leq k_N \ \textit{with} \ \ k_N := \lceil \frac{1}{2}\log_\phi N \rceil.$$

The random variable $L^{\langle N\rangle}$ converges in moments and in distribution to the random variable L associated with the continuous model. In particular, the distribution $\varpi_k^{\langle N\rangle}$ of $L^{\langle N\rangle}$ converges uniformly to the distribution ϖ_k of L

$$\varpi_k^{\langle N\rangle} - \varpi_k = \mathcal{O}(\frac{1}{N}), \quad \textit{with} \ \ \varpi_k^{\langle N\rangle} = 0 \ \textit{for} \ \ k > k_N; \tag{30}$$

the mean value $\varpi^{\langle N\rangle} := E[L^{\langle N\rangle}]$ satisfies

$$\varpi^{\langle N\rangle} - \varpi = \mathcal{O}(\frac{\log N}{N}),$$

and the ℓ-th moment $M_\ell^{\langle N\rangle}$ of $L^{\langle N\rangle}$ converges to the ℓ-th moment M_ℓ of L

$$M_\ell^{\langle N\rangle} - M_\ell = \mathcal{O}(\frac{(\log N)^\ell}{N})$$

uniformly in ℓ and N provided that $\ell! = \mathcal{O}(N^{\gamma-1})$ where γ is a real > 1 defined as a function of the dominant eigenvalue λ_4 of operator $\mathcal{G}$,

$$\gamma = \frac{|\log \lambda_4|}{2\log\phi} \approx 1.675.$$

Proof. From Proposition, the integer k_N is an upperbound on the worst-case complexity in the sense that

$$L^{\langle N\rangle}(z) \leq k_N \ \ \text{for} \quad z \in \mathbb{Q}^{\langle N\rangle}, \ \ \varpi_k^{\langle N\rangle} = 0 \ \text{for} \ \ k > k_N.$$

We denote by $P^{\langle N\rangle}(h)$ the number of points of $\mathbb{Q}^{\langle N\rangle}$ that are contained inside the fundamental disk $h(\mathcal{D})$. The probability distribution $\varpi_k^{\langle N\rangle}$ admits the following expression

$$\varpi_k^{\langle N\rangle} = \frac{1}{P^{\langle N\rangle}(I)} \sum_{|h|=k} P^{\langle N\rangle}(h).$$

We remark that $P^{\langle N\rangle}(h) = 0$ as soon as the diameter ρ_h of disk $h(\mathcal{D})$ is less than $2/N$. Otherwise, by "Gauss's principle", a fundamental disk $h(\mathcal{D})$ of diameter ρ_h contains a number of points equal to

$$P^{\langle N\rangle}(h) = \frac{\pi}{4}N^2[\rho_h^2 + \mathcal{O}(\frac{\rho_h}{N})].$$

This is valid in particular for $\mathcal{D}$ itself, and

$$P^{\langle N\rangle}(I) = \frac{\pi}{4}N^2[1+\mathcal{O}(\frac{1}{N})].$$

Since the fundamental intervals $h([0,1])$ of rank k form a partition of $[0,1]$, from (11) and (6), we get

$$\sum_{|h|=k} \rho_h = 1, \quad \sum_{|h|=k} \rho_h^2 = \varpi_k,$$

and thus

$$\sum_{|h|=k, \rho_h \geq \frac{2}{N}} \rho_h^2 = \varpi_k + \mathcal{O}(\frac{1}{N}).$$

We then deduce the first two assertions of the theorem.

The ℓ-th moment $M_\ell^{\langle N\rangle}$ of $L^{\langle N\rangle}$,

$$M_\ell^{\langle N\rangle} := \sum_{k\geq 0}(k+1)^\ell \,\Pr[L^{\langle N\rangle} = k+1] = \sum_{k<k_N} (k+1)^\ell \ (\varpi_k^{\langle N\rangle} - \varpi_{k-1}^{\langle N\rangle}),$$

is to be compared with the ℓ-th moment M_ℓ of L,

$$M_\ell := \sum_{k\geq 0}(k+1)^\ell \ (\varpi_k - \varpi_{k-1}).$$

The difference between $M_\ell^{\langle N\rangle}$ et M_ℓ is composed of two terms

$$A := \sum_{k>k_N} (k+1)^\ell \ (\varpi_k - \varpi_{k-1}), \qquad B := \sum_{k\leq k_N} (k+1)^\ell \ (\delta_k - \delta_{k-1}),$$

where $\delta_k := \varpi_k^{\langle N\rangle} - \varpi_k$. With Abel's transformation, and (30), we deduce that

$$B = k_N^\ell \ \mathcal{O}(\frac{1}{N}).$$

The sum A can be written with Theorem 9 as a function of dominant eigenvalue λ_4 of operator $\mathcal{G}$

$$A = \mathcal{O}\ (\sum_{k\geq k_N} k^\ell \ \lambda_4^k) = \mathcal{O}\ (\int_{k_N}^{\infty} x^\ell \lambda_4^x dx).$$

A direct computation gives

$$A = \mathcal{O}\ \left(\ell! \ \lambda_4^{k_N} \ k_N^\ell\right) = \mathcal{O}(\frac{\ell! \ k_N^\ell}{N^\gamma}) \text{ with } \gamma = \frac{|\log\lambda_4|}{2\log\phi}.$$

The relation (28) proves that γ is greater than 1, so that, provided that $\ell! = \mathcal{O}(N^{\gamma-1})$, one obtains

$$A = k_N^\ell \ \mathcal{O}(\frac{1}{N}). \qquad \blacksquare$$

6 Numerical estimates

We have already mentioned numerical estimates for the mean and the probability distribution of the Gaussian algorithm, including the mean value of the cost $\varpi \approx 1.35113$ and the estimate of the dominant eigenvalue $\lambda_4 \approx 0.1994$. Most of the expressions obtained so far involve slowly converging sums. The purpose of this section is to give indications on series transformations that permit to evaluate the mean (Theorem 12), the probability distribution (Theorem 13) and traces (Theorem 15) to great accuracy. We also detail ways in which exact bounds can be proved on λ_4 and μ_4 using trace formulae and truncation of operators (Theorem 17).

A real number α is said to be *polynomial time computable* if there exists an integer r such that an approximation of α to accuracy 10^{-d} can be computed in time $\mathcal{O}(d^r)$. We let P denote the class of such numbers. A major problem is to find which of the constants of this paper are polynomial time computable. Effective numerical procedures usually go along with proofs of membership in P.

Expected cost. The expected number of iterations ϖ of the Gaussian algorithm admits an expression that involves the remarkable constants $\zeta(3)$ and $Li_4(\frac{1}{2})$, where $Li_4(z)$ is the *tetralogarithm*,

$$Li_4(z) = \frac{z}{1^4} + \frac{z^2}{2^4} + \frac{z^3}{3^4} + \cdots = \sum_{n=1}^{\infty} \frac{z^n}{n^4}.$$

Given that the zeta function is easily computable [6], this proves that ϖ is a member of P while providing a fast computation scheme.

Theorem 12. *The mean number of iterations of the Gaussian algorithm SGA can be expressed with sums of generalized harmonic numbers*

$$\varpi = \frac{17}{4} + \frac{360}{\pi^4} \sum_{d=1}^{\infty} \frac{(-1)^d}{d^2} \sum_{c=1}^{d} \frac{1}{c^2} \tag{31}$$

$$\varpi = \frac{15}{2} - \frac{720}{\pi^4} \sum_{d=1}^{\infty} \frac{(-1)^d}{d^3} \sum_{c=1}^{d} \frac{1}{c}, \tag{32}$$

or with $\zeta(3)$ and the tetralogarithm $Li_4(\frac{1}{2})$,

$$\varpi = -\frac{60}{\pi^4} \left[24 Li_4(\frac{1}{2}) + 21\zeta(3)\log 2 + (\log 2)^4 \right] + \frac{60}{\pi^2} (\log 2)^2 + 17. \tag{33}$$

Thus, ϖ lies in the class P of polynomial time computable constants and

$$\varpi = 1.35113\,15744\,91659\,00179\,38680\,05256\,46466\,84404\,78970\,85087 \pm 10^{-50}.$$

Proof. Our computations are based on [5, 20, 24], see [2] for a vivid introduction to the subject. The starting point is Equations (14) and (15) from the proof of Theorem 4. Our aim is to transform the constants β and β' defined there,

$$\beta := \sum_{d \ge 2} \frac{1}{d^2} \sum_{1 \le c < d} \frac{1}{(c+d)^2} = \sum_{d \ge 1} \frac{1}{d^2} \sum_{d < c < 2d} \frac{1}{c^2},$$

$$\beta' := \sum_{d \ge 1} \frac{1}{d^2} \sum_{1 \le c \le d} \frac{1}{(c+d)^2} = \sum_{d \ge 1} \frac{1}{d^2} \sum_{d < c \le 2d} \frac{1}{c^2},$$

that are linked by the elementary relation $\beta' - \beta = \zeta(4)/4$. These constants are also conveniently expressed in terms of generalized harmonic numbers

$$H_n^{(r)} := \sum_{j=1}^{n} \frac{1}{j^r},$$

and the related sums

$$S(r,s) := \sum_{n=1}^{\infty} \frac{1}{n^s} H_n^{(r)} \qquad A(r,s) := \sum_{n=1}^{\infty} \frac{(-1)^n}{n^s} H_n^{(r)}, \qquad P(r,s) := \sum_{n=1}^{\infty} \frac{(-1)^n}{n^s} H_{2n}^{(r)}.$$

The following elementary relations hold:

$$S(r,s) + A(r,s) = \sum_{n=1}^{\infty} \frac{(1+(-1)^n)}{n^s} H_n^{(r)} = \sum_{n=1}^{\infty} \frac{2}{2^s} \frac{H_{2n}^{(r)}}{n^s} = \frac{1}{2^{s-1}} P(r,s).$$

With the pairs $(r,s) = (2,2)$ and $(r,s) = (1,3)$, one gets

$$P(2,2) = 2[S(2,2) + A(2,2)], \quad P(1,3) = 4[S(1,3) + A(1,3)]. \tag{34}$$

Proof of (31). We first use the pair $(2,2)$

$$\beta' = \sum_{d \geq 1} \frac{1}{d^2} [H_{2d}^{(2)} - H_d^{(2)}] = P(2,2) - S(2,2) = 2A(2,2) + S(2,2).$$

Furthermore, squaring $\zeta(2)$ and folding by symmetry yields

$$S(2,2) = \frac{1}{2}\zeta^2(2) + \frac{1}{2}\zeta(4) = \frac{7}{4}\zeta(4), \quad \beta' = \frac{7}{4}\zeta(4) + 2A(2,2),$$

which is equivalent to the second expression for ϖ.

Proof of (32). We now use the pair $(1,3)$. We link $S(2,2)$ with $S(1,3)$, $P(2,2)$ with $P(1,3)$, hence also $A(2,2)$ with $A(1,3)$, by means of an old trick of Euler and Nielsen that is based on partial fraction decomposition [20]. We take

$$2\left[\frac{1}{c^3(d-c)} - \frac{1}{c^3 d}\right] = \frac{1}{c^2 d^2} + \frac{1}{c^2(d-c)^2} - \frac{1}{d^2(d-c)^2}. \tag{35}$$

$$2\left[\frac{1}{c^3 d} - \frac{1}{c^3(c+d)}\right] = \frac{1}{c^2 d^2} + \frac{1}{c^2(c+d)^2} - \frac{1}{d^2(c+d)^2}. \tag{36}$$

and propose to sum over all pairs (c,d) with $d \geq 2$ and $1 \leq c < d$. From (35), we obtain

$$2S(1,3) = [\zeta(2)^2 - S(2,2)] + \zeta(2)^2 + [S(2,2) - \zeta(4)] = \frac{5}{2}\zeta(4)$$

and, from (36),

$$2[P(1,3) - S(1,3)] = [\zeta(2)^2 - S(2,2)] + [\zeta(2)^2 - P(2,2)] - \beta.$$

These equations relate $A(1,3)$ to $A(2,2)$, so that β becomes expressible in terms of $A(1,3)$:

$$A(1,3) = \frac{13}{32}\zeta(4) - \frac{1}{2}A(2,2), \quad \beta = \frac{25}{8}\zeta(4) - 4A(1,3)$$

Proof of (33). Sitaramachandrarao [24] notes that $A(1,3)$ is related to a quantity already considered by Ramanujan and can be expressed with the tetralogarithm of argument $1/2$ and $\zeta(3)$. De Doelder [5, p. 128] provides the explicit integral manipulations needed to reduce $A(1,3)$ to $\zeta(3)$ and $Li_4(1/2)$, which starts from the integral representation

$$A(1,3) - \zeta(4) = -\frac{1}{2}\int_0^1 \log t \log^2(1+t)\, \frac{dt}{t}.$$

and gives

$$A(1,3) = -\frac{7}{64}\zeta(4) + \frac{7}{8}\zeta(3)\log(2) - \frac{\pi^2}{24}(\log 2)^2 + \frac{1}{24}(\log 2)^4 + Li_4(\frac{1}{2}).$$

Since the expression (32) of ϖ involves $A(1,3)$, this eventually relates our constant to the tetralogarithm. Finally, the numerical form is easily obtained (in polynomial time) since the tetralogarithm series converges geometrically. ■

Probability distribution. The probability distribution of the Gaussian algorithm can be expressed in terms of complicated series involving the zeta function, the resulting expressions being especially useful for small values of k.

Theorem 13. *The probability distribution of the number of iterations of the Gaussian algorithm has initial values:* $\varpi_0 := \Pr[L \geq 1] = 1$, *and*

$$\begin{aligned}
\varpi_1 &:= \Pr[L \geq 2] = \frac{\pi^2}{3} - 3, \\
\varpi_2 &:= \Pr[L \geq 3] = -5 + \frac{2\pi^2}{3} - 2\zeta(3) + 2\sum_{n=0}^{\infty}(-1)^n(n+1)\zeta(n+4)(\zeta(n+2)-1).
\end{aligned}$$

In general, each $\varpi_k = \Pr\{L \geq k+1\}$ *lies in the class* P *of polynomial time computable constants.*

The proof is based on the fact that summations of suitable analytic functions at integer points can be computed in polynomial time via representations in terms of zeta function series. The following proposition summarizes the general process.

Proposition 14. *Let* $F(z_1, z_2, \ldots, z_k)$ *be a rational function of* $\mathbb{Q}(z_1, \ldots, z_k)$ *such that*

$$F(z_1, z_2, \ldots, z_k) = z_1^2 z_2^2 \cdots z_k^2 f(z_1, \ldots, z_k),$$

where $f(z_1, \ldots, z_k)$ *is analytic at all points of the product domain* Δ *defined by* $0 \leq \Re(z_i) \leq 2$, $\Im(z_i) = 0$. *Then the constant*

$$\varphi = \sum_{n_1, \ldots, n_k = 1}^{\infty} F(\frac{1}{n_1}, \frac{1}{n_2}, \ldots \frac{1}{n_k}) \tag{37}$$

is polynomial time computable.

Proof. Univariate case. We first treat the case of one complex variable ($k = 1$) and set $z = z_1$. A discussion of the basic technique is given for instance in Vardi's entertaining book [28]. The sum defining φ converges as $F(z) = \mathcal{O}(z^2)$ for z near 0. Let M be an integer such that $F(z)$ is analytic for $|z| < \frac{1}{M}$. The summation domain in (37) is split according to

$$\varphi = \sum_{n=1}^{M} F(\frac{1}{n}) + \sum_{n=M+1}^{\infty} F(\frac{1}{n}),$$

and we let $\varphi_{\leq}$ and $\varphi_{>}$ denote the two partial sums.

The sum $\varphi_{\leq}$ requires a finite number of evaluations of F at rational points, and thus it lies in P. The sum $\varphi_{>}$ is evaluated by means of the Taylor expansion of F at the origin:

$$F(z) = \sum_{r=2}^{\infty} f_r z^r.$$

Define the truncated zeta function as

$$\zeta_M(s) = \sum_{n=M+1}^{\infty} \frac{1}{n^s} = \zeta(s) - \sum_{n=1}^{M} \frac{1}{n^s}.$$

Then using the Taylor expansion of F and interchanging summations in the definition of $\varphi_{>}$, one gets

$$\varphi_{>} = \sum_{n=M+1}^{\infty} \sum_{r=2}^{\infty} f_r(\frac{1}{n^r}) = \sum_{r=2}^{\infty} f_r \zeta_M(r).$$

The given series exhibits geometric convergence since $\zeta_M(r) = \mathcal{O}((M+1)^{-r})$ and it is in P since the value of $\zeta(r)$ can be computed uniformly in time polynomial in r and the number d of digits required, as effects from standard algorithms based on Euler–Maclaurin summation [6, 28].

Multivariate case. By compacity of the analyticity domain of F, there exists a fixed real number ε such that F is analytic in the polydisc $|\zeta_i - z_i| < \varepsilon$ for each $(z_1, \ldots, z_k) \in \Delta$. In particular, the Taylor expansion of F has radius of convergence $\geq \varepsilon$ at each point of Δ. We then choose an integer cut point M such that $\frac{1}{M} < \varepsilon$.

The summation domain of (37) subdivides as

$$\mathbb{N}^k = \Big([1\,..\,M] \cup [M+1\,..\,+\infty]\Big)^k,$$

which induces a splitting of the sum giving φ into 2^k subregions. The sum taken over $[1\,..\,M]^k$ is a finite one that requires only a fixed number of rational function evaluations, and thus it is in P. The k-fold infinite sum is transformed by a process similar to the univariate case into

$$\sum_{r_1,\ldots,r_k} f_{r_1,\ldots,r_k} \zeta_M(r_1) \cdots \zeta_M(r_k), \tag{38}$$

which still exhibits geometric convergence and thus lies in P, as the Taylor coefficients of F are polynomial time computable. The remaining $2^k - 2$ sums are lower dimensional sums that are computable by the same process as (38) with some variables instantiated to values of the form $\frac{1}{n_j}$. Geometric convergence is ensured by the choice of M dictated by the compacity argument given above.

The proof also entails a representation of φ in the form

$$\varphi = c^{(0)} + \sum_{r_1 > M} c^{(1)}_{r_1} \zeta_M(r_1) + \cdots + \sum_{r_1,\ldots,r_k > M} c_{r_1,\ldots,r_k} \zeta_M(r_1) \cdots \zeta_M(r_k),$$

where the c–coefficients are rational numbers that are easily computable. ■

A version of Proposition 14 is clearly valid for larger classes of analytic functions provided they are both computable and expandable in polynomial time. Theorem 13 directly results from the proposition.

Proof of Theorem 13. The two cases of ϖ_1 and ϖ_2 give the idea of the general strategy. First, from the definition of continuants Q_0, Q_1, we have

$$\varpi_1 = \sum_{m_1 \geq 1} \frac{1}{m_1^2 (m_1+1)^2},$$

and, from the partial fraction decomposition of $m_1^{-2}(m_1+1)^{-2}$, we get $\varpi_1 = 2\zeta(2) - 3$.

Similarly, when $k = 2$, we have

$$\varpi_2 = \sum_{m_1, m_2 \geq 1} \frac{1}{(m_1 m_2 + 1)^2 (m_1 m_2 + m_1 + 1)^2}.$$

The partial fraction expansion of the general term (taken with respect to m_2)

$$\frac{1}{m_1^2\,(m_2 m_1 + 1)^2} + \frac{1}{m_1^2\,(m_2 m_1 + m_1 + 1)^2} - \frac{2}{m_1^3\,(m_2 m_1 + 1)} + \frac{2}{m_1^3\,(m_2 m_1 + m_1 + 1)}$$

can be subjected to summation: the contributions of the last two terms telescope; the first 2 terms, upon factoring out $(m_1 m_2)^{-2}$ and $(m_1(m_1 m_2 + 1))^{-2}$ and using the Taylor series of $(1+y)^{-2}$, can then be expanded in terms of zeta functions.

The fact that each ϖ_k for $k \geq 3$ belongs to P results from Proposition 14. For instance, when $k = 3$, we have

$$\varpi_3 = \sum_{n_1,n_2,n_3=1}^{\infty} F(\frac{1}{n_1}, \frac{1}{n_2}, \frac{1}{n_3})$$

with

$$F(z_1, z_2, z_3) = \frac{z_1^4 z_2^4 z_3^4}{(1 + z_1 z_2 + z_2 z_3)^2 (1 + z_3 + z_1 z_2 + z_2 z_3 + z_1 z_2 z_3)^2},$$

which satisfies the conditions of Proposition 14. ∎

The following values have been determined in this way to great accuracy:

$$\begin{aligned}
\varpi_1 &= 0.28986\,81336\,96452\,87294 \\
\varpi_2 &= 0.04848\,08014\,49463\,63270 \\
\varpi_3 &= 0.01027\,81647\,79066\,59643.
\end{aligned}$$

Notice however that the computational complexity has an exponent that increases with k.

Trace formulæ The traces of powers of $\mathcal{G}$ contain informations on the eigenvalues since they are nothing but power sums of the eigenvalues. A computational process very similar to the one employed for determining the ϖ_k applies here. First, we state a trace formula originally due to Babenko and Meyer.

Theorem 15. *(i) The trace of the operator $\mathcal{G}_s^k$ satisfies*

$$Tr\, \mathcal{G}_s^k \equiv \sum_i \lambda_{i,s}^k = \sum_{|h|=k} \frac{\tau(h)^{-s}}{1 - (-1)^k \tau(h)^{-2}}, \text{ with } \tau(h) = \frac{(Q_k + P_{k-1}) + \sqrt{(Q_k + P_{k-1})^2 - 4(-1)^k}}{2}. \tag{39}$$

(ii) For $s = 4$, each $Tr\,\mathcal{G}^k$ is in the class P of constants computable in polynomial time. For instance, $Tr\,\mathcal{G}$ admits the explicit form

$$Tr\,\mathcal{G} = \frac{7}{2} - \frac{2}{\sqrt{5}} - \frac{7}{\sqrt{2}} + \frac{1}{2} \sum_{n=2}^{\infty} (-1)^n \frac{n-1}{n+1} \binom{2n}{n} [\zeta(2n) - 1 - \frac{1}{2^{2n}}].$$

Proof (Sketch). We follow [19, p. 189]. From (21), the operator $\mathcal{G}_s^k$ satisfies

$$\mathcal{G}_s^k[f](z) = \sum_{|h|=k} \widetilde{h}(z)^s f \circ h(z),$$

where $\widetilde{h}(z)$ is the holomorphic function that coincides with $\sqrt{|h'(z)|}$ on the interval $\mathcal{J}$. The operator $\mathcal{G}_s^k$ is thus an infinite sum of operators each of the general form

$$\mathcal{L}_{\alpha,\beta}[f](z) = \alpha(z)\ f \circ \beta(z),$$

with α and β belonging to $\mathcal{A}_\infty(\mathcal{V})$. Here, the functions β map the domain $\mathcal{V}$ strictly inside itself; then, they have exactly one fixed point z_0 in $\mathcal{V}$ for which one has $|\beta'(z_0)| < 1$ and $\alpha(z_0) \neq 0$. Thus, the spectrum of $\mathcal{L}_{\alpha,\beta}$ is exactly a geometric progression of the form

$$\{\alpha(z_0)\beta'(z_0)^n\}_{n=0}^{\infty}, \tag{40}$$

as is shown by a direct construction. Then, the corresponding trace is

$$Tr\,\mathcal{L}_{\alpha,\beta} = \frac{\alpha(z_0)}{1 - \beta'(z_0)}.$$

The quantity $Tr\,\mathcal{G}_s^k$ is the sum of the $Tr\,\mathcal{L}_{\alpha,\beta}$ where the component operators $\mathcal{L}_{\alpha,\beta}$ are taken with $\beta = h$ and $\alpha = \widetilde{h}^s$, as follows from (extended) additivity of the trace. A simple computation of the fixed point of a homography h then yields the result stated in (i).

Next, we have

$$P_0 + Q_1 = m_1, \;\; P_1 + Q_2 = m_1 m_2 + 2, \;\; P_2 + Q_3 = m_1 m_2 m_3 + m_1 + m_2 + m_3.$$

The special formula for the trace of $\mathcal{G}_4$ results from the identity

$$Tr\,\mathcal{G}_4 = \sum_{m\geq 1} \frac{\tau_m^{-4}}{1+\tau_m^{-2}} \text{ where } \tau_m := \tau(h) \text{ for } h(z) = \frac{1}{m+z}.$$

After expanding the algebraic function that involves

$$\tau_m = \frac{1}{2}\left(m + \sqrt{m^2+4}\right),$$

in inverse powers of m, we perform resummation in terms of zeta functions like in the proof of Theorem 13. The general case results from a simple extension of Proposition 14 to algebraic functions (the rational forms in $\tau(h)$ that are subjected to summation). ■

In this way, the trace $Tr\,\mathcal{G}$ is found to be

$$Tr\,\mathcal{G} = 0.14446\,23962\,46160\,81588\,24990\,90525\,48320\,38136 \pm 10^{-40},$$

and one determines

$$Tr\,\mathcal{G}^2 = 0.04647\,18256\,42727\,93983\,52797\,53170 \pm 10^{-30},$$

a value that is especially precious for numerical checks (see below).

Eigenvalues and test functions. The last numerical task is to estimate the dominant and subdominant eigenvalues, λ_4 and μ_4 that characterize the rate of the geometric decay of the probabilities ϖ_k.

The general approach to determining the dominant eigenvalue uses a method already employed by Wirsing [29] in the analysis of the Euclidean algorithm. It is based on the positivity of the operator $\mathcal{G}_s$ for real $s > 1$, valid in particular for $s = 4$, and is summarized by the following proposition:

Proposition 16. *Let f be a function that is analytic and strictly positive on the interval $[0,1]$. Assume that there exist two constants α, β with $0 < \alpha < \beta$ such that*

$$\alpha < \frac{\mathcal{G}[f](x)}{f(x)} < \beta, \tag{41}$$

for all $x \in [0,1]$. Then the dominant eigenvalue λ_4 of $\mathcal{G}$ satisfies

$$\alpha \leq \lambda_4 \leq \beta.$$

Proof. If f satisfies the assumptions, it satisfies similar assumptions in an open rectangle of the form $\mathcal{R} :=]-\delta, 1+3\delta[\times]-\rho, +\rho[$ with $\rho > 0$ and $\delta > 0$ sufficiently small. If $\widetilde{\mathcal{V}}$ is the disk of diameter $]-\delta, 1+3\delta[$, there exists an integer k_0 such that

$$\bigcup_{|h|=k} h(\widetilde{\mathcal{V}}) \subset \mathcal{R}, \quad \text{for all } \; k \geq k_0.$$

Then, the iterate $g := \mathcal{G}^{k_0}(f)$ is an element of $\mathcal{A}_\infty(\widetilde{\mathcal{V}})$, and, from the remark in the beginning of Section 4, we can adapt all the properties of Section 4 to such a space. When iterating the relation (41), we have

$$\alpha^k < \frac{\mathcal{G}^k[g](x)}{g(x)} < \beta^k \;\text{ or }\; \alpha < \left(\frac{\mathcal{G}^k[g](x)}{g(x)}\right)^{1/k} < \beta.$$

On the other hand, using (27), we deduce that λ_4 has to be contained in the interval $[\alpha, \beta]$. ■

Thus, *proving* effective bounds on the dominant eigenvalue reduces to *finding* (by whatever means!) suitable "test" functions that satisfy the inequality (41), with $\beta - \alpha$ sufficiently small.

A first class of bounds is obtained by adapting Wirsing's approach [29] and introducing a specific set of test functions $\psi_a(t)$ whose transforms have a simple manageable form,

$$\begin{cases} \psi_a(t) = \dfrac{1}{(1+at)(1+(a+1)t)(1+(a+2)t)(1+(a+3)t)} \\ \phi_a(t) = \mathcal{G}[\psi_a](t) = \dfrac{1}{3}\,\dfrac{1}{(t+a+1)(t+a+2)(t+a+3)}. \end{cases} \tag{42}$$

Choosing $a = \frac{487}{1000}$, we discover that the ratio $\phi_a(t)/\psi_a(t)$ always lies in the interval $]0.170, 0.205[$ for $t \in [0,1]$. (This fact can be checked by purely algebraic computations if desired, since it is only relative to exactly known rational functions.) We therefore obtain a first proven bound of

$$\frac{170}{1000} < \lambda_4 < \frac{205}{1000}.$$

Eigenvalues and truncations. In order to obtain more refined estimates on λ_4, we need to find test functions that are expected to be "good" approximations to the dominant eigenfunctions, while still being computable efficiently. The idea is to approximate the effect of $\mathcal{G}$ on analytic functions represented by their Taylor expansion at some point a of the interval $[0,1]$.

We thus examine the transforms

$$h_j(x) = \mathcal{G}[(x-a)^j](x) = \sum_{n=1}^{\infty} \frac{1}{(n+x)^4}\left(\frac{1}{n+x} - a\right)^j.$$

By the binomial theorem, we find

$$h_j(x) = \sum_{\ell=0}^{j} \binom{j}{\ell} (-a)^{j-\ell}\, \zeta(4+\ell, x+1),$$

where $\zeta(s,w) = \sum_{n=0}^{\infty}(n+w)^{-s}$ is the Hurwitz zeta function. The Hurwitz zeta function is itself easily expanded at any point a:

$$\begin{aligned} \zeta(s, x+1) &= \zeta(s, a+1) - \binom{s}{1}\zeta(s+1, a+1)\,(x-a) + \binom{s+1}{2}\zeta(s+2, a+1)\,(x-a)^2 - \cdots, \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{s+i-1}{i}\zeta(s+i, 1+a)\,(x-a)^i. \end{aligned}$$

Thus, viewed as acting on series expansions at $x = a$, the operator $\mathcal{G}$ is expressed by an infinite matrix $M = (M_{i,j})$, where $M_{i,j}$ is the coefficient of $(x-a)^i$ in $\mathcal{G}[(x-a)^j]$, that is

$$M_{i,j} = (-1)^i \sum_{\ell=0}^{j} \binom{j}{\ell}\binom{i+\ell+3}{i}(-a)^{j-\ell}\zeta(4+\ell+i, a+1).$$

For instance, the matrix at $a = 0$ assumes the simple form

$$M_{i,j} = (-1)^i\binom{i+j+3}{i}\zeta(i+j+4).$$

In the sequel, guided by numerical experiments, we adopt the value $a = 1/2$ that gives faster convergence; in that case, the quantity $\zeta(s, \frac{3}{2})$ even reduces to the classical Riemann zeta function

$$\zeta(s, \frac{3}{2}) = 2^s \left[\frac{1}{3^s} + \frac{1}{5^s} + \dots\right] = 2^s \left[-1 + \zeta(s)(1 + 2^{-s})\right].$$

For each m, the finite matrix $T^{[m]}$ is defined as the square submatrix obtained from $(M_{i,j})$ by restricting the indices to $0 \le i, j < m$, and accordingly it may be viewed as operating on polynomials of degree less than m: $T^{[m]}[f]$ is the truncated Taylor series at a of the transform by $\mathcal{G}$ of the truncated Taylor series of f at a. For that reason, we refer to the $T^{[m]}$ as truncated operators.

The $T^{[m]}$ provide a sequence of approximations to operator $\mathcal{G}$ so that their eigenvalues and eigenfunctions may be expected to provide reasonably good approximations to the corresponding spectral characteristics of $\mathcal{G}$. We let $\lambda_4^{[m]}$ denote the dominant eigenvalue of $T^{[m]}$ and $p^{[m]}$ be the corresponding eigenfunction which is a polynomial of degree at most equal to $m - 1$.

Using computer algebra, we have determined the eigenfunctions of the $T^{[m]}$ for many values of $m \le 100$ and numerical accuracy up to 150 digits. (The heaviest such computation required about 10^{12} machine cycles, that is to say at the moment a few hours of computing time). The dominant eigenfunctions considered are all positive on the interval $[0, 1]$ and thus are good candidates as test functions along the lines of Proposition 16. In that case, the transforms $\mathcal{G}[p^{[m]}]$ are computed numerically by mean of well-known (and efficient) algorithms available for the zeta function. We discover that the ratios

$$\rho^{[m]}(x) = \frac{\mathcal{G}[p^{[m]}](x)}{p^{[m]}(x)}$$

have an amplitude of variation,

$$\xi^{[m]} = \sup_{x\in[0,1]} \rho^{[m]}(x) - \inf_{x\in[0,1]} \rho^{[m]}(x),$$

that is very small. For instance, a rationalized form of $p^{[4]}$ is

$$p^{[4]}(x) \approx 1 - \frac{87}{50}(x - \frac{1}{2}) + \frac{391}{200}(x - \frac{1}{2})^2 - \frac{1687}{1000}(x - \frac{1}{2})^3.$$

It is seen that $\rho^{[4]}(x)$ varies smoothly between 0.194 and 0.226 over the interval, so that $\xi^{[4]} \approx 0.032$ and this provides the bounds

$$\frac{194}{1000} \le \lambda_4 \le \frac{226}{1000}.$$

Moving to higher values of m, we discover that the variation $\xi^{[32]}$ is less than $3 \cdot 10^{-10}$ with a pronounced plateau over $[0.1, 0.8]$ suggesting the more precise value

$$\lambda_4 = 0.19945\,88183\,43767 \pm 10^{-15}, \tag{43}$$

which is then fully confirmed by the computation at $m = 64$.

We observe that, if required, such numerical computations can be completely validated by purely algebraic manipulations only; it suffices to use truncated Taylor expansions of the Hurwitz zeta function, then rational coefficient approximations, and finally exact computations with algebraic numbers (for determining the maximum and minimum of a rational functions with rational coefficients). We have not deemed this tedious verification necessary since the numerical evidence gathered is overwhelming. In addition, a further check is possible since the traces are known to great accuracy (see below).

Given the approximation (43) and the value of the trace, $Tr\,\mathcal{G}^2 = 0.04647$, we easily deduce a bound on the subdominant eigenvalue, since $|\mu_4| < (Tr\,\mathcal{G}^2 - \lambda_4^2)^{1/2} = 0.081780$.

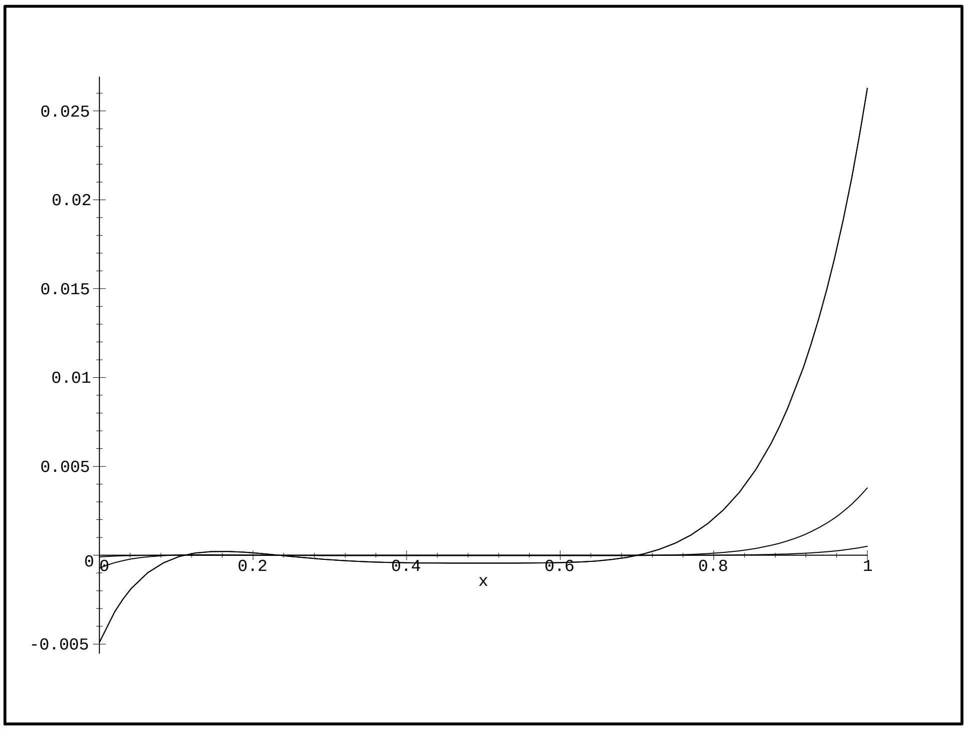

Fig. 6. Plot of the ratios $\rho^{[m]}(x) - \lambda_4 = \dfrac{\mathcal{G}[p^{[m]}](x)}{p^{[m]}(x)} - \lambda_4$ for $m = 4, 6, 8$.

Theorem 17. *The dominant and subdominant eigenvalues of $\mathcal{G}$ satisfy*

$$\lambda_4 = 0.19945\,88183\,43767 \pm 10^{-15}, \qquad |\mu_4| < 0.082.$$

If we allow ourselves to give up some rigour, it is possible to go much further. Examination of all the eigenvalues of the truncated operators $T^{[m]}$ reveals that most of them stabilize to a fixed set of definite values, with the occasional occurrence of some spurious values that eventually disappear as m increases. It is then natural to conjecture that the stable limit values do yield the complete spectrum of $\mathcal{G}$. Again a check via traces is possible and, in this way, we have obtained convincing (but not proven) values for the first 37 eigenvalues of operator $\mathcal{G}$ with an accuracy almost certainly better than 10^{-25}. Here is a listing of our first estimates:

$$\begin{aligned}
\lambda_{1,4} &\doteq +0.19945\,88183\,43767\,26019\,18456\\
\lambda_{2,4} &\doteq -0.07573\,95140\,84360\,60892\,78089\\
\lambda_{3,4} &\doteq +0.02856\,64037\,69818\,52783\,00174\\
\lambda_{4,4} &\doteq -0.01077\,74165\,76612\,69829\,31408\\
\lambda_{5,4} &\doteq +0.00407\,09406\,93426\,42144\,86407.
\end{aligned}$$

The 37 eigenvalues found are all simple and they alternate in sign: this confirms, two properties that have been conjectured but not established for all the $\mathcal{G}_s$ operators for real values $s > 1$. We also observe that the agreement between the sum of these values and $Tr\,\mathcal{G}$ is of the same order as the last eigenvalue found while the sum of squares agrees to the stated precision of $\pm 10^{-25}$ with our earlier computation of $Tr\,\mathcal{G}^2$.

Finally, examination of the ratios $r_j = \lambda_{j,4}/\lambda_{j+1,4}$ shows a remarkable stability, for instance

$$r_1 = -2.633,\ r_2 = -2.651,\ r_3 = -2.650,\ r_4 = -2.647,\ r_5 = -2.644,$$

and so on. The spectrum of $\mathcal{G}$ is thus very nearly a geometric progression of ratio -2.64. A simplified model that we now explain sheds some light on such regularities.

For large enough s, the operator $\mathcal{G}_s$ should be dominated by its first term,

$$\mathcal{C}_s[f](x) = \frac{1}{(1+x)^s} f(\frac{1}{1+x}),$$

since the other terms composing $\mathcal{G}_s$ are of an exponentially smaller order as a function of s. From considerations recalled above (40), the spectrum of $\mathcal{C}_s$ forms a geometric progression determined exactly from local properties of the operator near the fixed point of $(1+x)^{-1}$, namely $1/\phi$, where $\phi = (1+\sqrt{5})/2$ is the golden ratio. Here, the dominant eigenvalue of $\mathcal{C}_s$ is $\widehat{\lambda}_{1,s} = \phi^{-s}$ and the other eigenvalues are given by $\widehat{\lambda}_{j,s} = (-1)^{j-1}\phi^{-s-2j}$. This model yields $\widehat{\lambda}_{1,4} = \phi^{-4} = 0.14589$, and the ratio between successive eigenvalues is $-\phi^2 = -2.61803$. This is a fairly good approximation to what is observed for the eigenvalues of $\mathcal{G}_4$ and it provides yet another "plausible" confirmation of our numerical data.

The Gauss-Kuzmin-Wirsing constant. The numerical methods developed here for the $\mathcal{G}_4$ operator apply also to the $\mathcal{G}_2$ operator that is closely tied with continued fractions and the Euclidean algorithm. Thus, as a further test, we have applied them to the computation of the Gauss-Kuzmin-Wirsing constant that is, up to sign, the second eigenvalue $\lambda_{2,2}$ of the operator $\mathcal{G}_2$. We find to 30 digits of accuracy:

$$\lambda_{2,2} = -0.30366\,30028\,98732\,65859\,74481\,21901.$$

The next eigenvalues of the $\mathcal{G}_2$ operator determined in this way are

$$\begin{aligned}
\lambda_{3,2} &= +0.10088\,45092\,93104\,07530\\
\lambda_{4,2} &= -0.03549\,61590\,21659\,84540\\
\lambda_{5,2} &= +0.01284\,37903\,62440\,26481\\
\lambda_{6,2} &= -0.00471\,77775\,11571\,03107\\
\lambda_{7,2} &= +0.00174\,86751\,24305\,51191\\
\lambda_{8,2} &= -0.00065\,20208\,58320\,50290
\end{aligned}$$

In particular this computation indicates the presence of a spurious value amongst the ones attributed to Babenko and printed in [11, p. 350]. Scepticism concerning the values given in [11, p. 350] has been expressed repeatedly, for instance by Mayer, Knuth (private communication, 1994) and MacLeod [15]. Our computations that are well validated by trace formulæ provide an independent confirmation of the observations of these authors.

Conclusion. This paper has demonstrated that the lattice reduction algorithm in dimension 2 deriving from Gauss's scheme is in a strong sense a 2–dimensional lifting of the continued fraction algorithm. An essential rôle is played by functional analysis methods and dominant spectral properties of Ruelle-Mayer operators. These operators generalize the Perron-Frobenius operator developed for continued fractions and the Euclidean algorithm. The methods used lead to a complete analysis of the Gaussian algorithm in the average case, in probability, and in dynamic behaviour, under either a continuous or a discrete model.

The centered algorithm discussed in Section 2 is briefly examined in [4] and further studied in [26]. Also, it has been recently noticed [26] that the methods of this paper apply to the analysis of an old algorithm for comparing (multiprecision) rationals or equivalently for computing the sign of 2×2 determinants with integer entries. Such problems are of general interest in symbolic computation as well as in computational geometry. The algorithm compares two rationals by computing the corresponding continued fraction expansions on a call-by-need basis. The analysis of [26] is similar to the ones given here and in [4], except that fundamental squares take the place of fundamental disks.

In this paper, we have considered the case when the input density is uniform. A wide class of nonuniform input distributions can be treated by means of generalized Ruelle-Mayer operators. In this way a whole range of analyses encompassing continued fractions and 2-dimensional lattice reduction can be obtained [27].

Acknowledgements. The work of Philippe Flajolet has been supported 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. BAILEY, D. H., BORWEIN, J. M., AND GIRGENSOHN, R. Experimental evaluation of Euler sums. *Experimental Mathematics 3*, 1 (1994), 17–30.
3. COHEN, H. *A Course in Computational Algebraic Number Theory.* No. 138 in Graduate Texts in Mathematics. Springer–Verlag, 1993.
4. DAUDÉ, H., FLAJOLET, P., AND VALLÉE, B. An analysis of the Gaussian algorithm for lattice reduction. In *Algorithmic Number Theory Symposium* (1994), L. Adleman, Ed., no. 877 in Lecture Notes in Computer Science, pp. 144–158. Proceedings of *ANTS'94*.
5. DE DOELDER, P. J. On some series containing $\psi(x) - \psi(y)$ and $(\psi(x) - \psi(y))^2$ for certain values of x and y. *Journal of Computational and Applied Mathematics 37* (1991), 125–141.
6. EDWARDS, H. M. *Riemann's Zeta Function.* Academic Press, 1974.
7. GROTHENDIECK, A. *Produits tensoriels topologiques et espaces nucléaires.* No. 16 in Memoirs of the American Mathematical Society. A.M.S., Providence, 1955.
8. GROTHENDIECK, A. La théorie de Fredholm. *Bulletin de la Société Mathématique de France 84* (1956), 319–384.
9. HENSLEY, D. The number of steps in the Euclidean algorithm. *Journal of Number Theory 49*, 2 (1994), 142–182.
10. KAIB, M., AND SCHNORR, C. P. A sharp worst–case analysis of the Gaussian lattice basis reduction algorithm for any norm. Preprint, 1992. To appear in *J. of Algorithms.*
11. KNUTH, D. E. *The Art of Computer Programming*, 2nd ed., vol. 2: Seminumerical Algorithms. Addison-Wesley, 1981.
12. KRASNOSELSKY, M. *Positive solutions of operator equations.* P. Noordhoff, Groningen, 1964.
13. LAGARIAS, J. C. Worst–case complexity bounds for algorithms in the theory of integral quadratic forms. *Journal of Algorithms 1*, 2 (1980), 142–186.
14. LENSTRA, A. K., LENSTRA, H. W., AND LOVÁSZ, L. Factoring polynomials with rational coefficients. *Mathematische Annalen 261* (1982), 513–534.
15. MACLEOD, A. J. High-accuracy numerical values in the Gauss-Kuzmin continued fraction problem. *Computers and Mathematics with Applications 26*, 3 (1993), 37–44.
16. 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.
17. 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.
18. MAYER, D. H. On a ζ function related to the continued fraction transformation. *Bulletin de la Société Mathématique de France 104* (1976), 195–203.
19. MAYER, D. H. Continued fractions and related transformations. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, M. K. Tim Bedford and C. Series, Eds. Oxford University Press, 1991, pp. 175–222.
20. NIELSEN, N. *Die Gammafunktion.* Chelsea Publishing Company, New York, 1965. Reprinted from *Handbuch der Theorie der Gammafunktion* (1906) and *Theorie der Integrallogarithmus und verwandter Transzendenten* (1906).
21. ROCKETT, A., AND SZŰSZ, P. *Continued Fractions.* World Scientific, Singapore, 1992.
22. SCHARLAU, W., AND OPOLKA, H. *From Fermat to Minkowski, Lectures on the Theory of Numbers and its Historical Developments.* Undergraduate Texts in Mathematics. Springer-Verlag, 1984.
23. SERRE, J.-P. *A Course in Arithmetic.* Graduate Texts in Mathematics. Springer Verlag, 1973.
24. SITARAMACHANDRARAO, R. A formula of S. Ramanujan. *Journal of Number Theory 25* (1987), 1–19.
25. VALLÉE, B. Gauss' algorithm revisited. *Journal of Algorithms 12* (1991), 556–572.
26. VALLÉE, B. Évaluation du signe d'un déterminant 2×2. Rapport de Recherche de l'Université de Caen, Les Cahiers du GREYC, 1995.
27. 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 # 1995-4, 1995.
28. VARDI, I. *Computational Recreations in Mathematica.* Addison Wesley, 1991.
29. WIRSING, E. On the theorem of Gauss–Kusmin–Lévy and a Frobenius–type theorem for function spaces. *Acta Arithmetica 24* (1974), 507–528.
30. ZIPPEL, R. *Effective Polynomial Computations.* Kluwer Academic Publishers, Boston, 1993.