The Brun Gcd Algorithm in high dimensions is almost always subtractive

Valérie Berthé

IRIF, UMR CNRS 7089, Université Paris Diderot, France

Loïck Lhote

GREYC, UMR CNRS 6072, ENSICAEN & Université de Caen Normandie, France

Brigitte Vallée

GREYC, UMR CNRS 6072, Université de Caen Normandie, France

Abstract

We introduce and study an algorithm which computes the gcd of $d+1$ entries. This is a natural extension of the usual Euclid algorithm, and coincides with it for $d = 1$; it performs Euclidean divisions, between the largest entry and the second largest entry, and then re-orderings. This is the discrete version of a multidimensional continued fraction algorithm due to Brun, in d dimensions. We perform the average-case analysis of this algorithm, and prove that the mean number of steps is linear with respect to the size of the entry. The method is based on the study of the underlying Brun dynamical system, and relies on dynamical analysis. All the constants that arise in the analysis are mathematical constants which are defined via the Brun dynamical system: for instance, the main constant of the analysis involves the entropy of the system, and the other constants of the analysis are related to fine properties of the invariant measure of the system. We are then led to the asymptotic behaviour of the Brun dynamical system, when dimension d becomes large, which is of independent interest. We show that the asymptotic Brun dynamical system almost always involves quotients that are equal to 1, and is then almost always subtractive. This explains the ineffficiency of the Brun gcd algorithm, particularly when it is compared in high dimensions to another extension of the Euclid algorithm, proposed by Knuth, and already analyzed by the authors.

Email addresses: `berthe@irif.fr` (Valérie Berthé), `loick.lhote@ensicaen.fr` (Loïck Lhote), `brigitte.vallee@unicaen.fr` (Brigitte Vallée).

[1] Thanks to the ANR `DynA3S` (ANR-13-BS02-0003) and `AleaEnAmSud` AmSud-STIC Projects.

1. Introduction.

General Context. We study a multiple gcd algorithm that is a natural extension of the usual Euclid algorithm for $(d+1)$ integers, and coincides with it for $d = 1$. This is a discrete version of a multidimensional continued fraction algorithm introduced by Brun in [9]. This algorithm is itself based on a dynamical system. The Brun continued fraction algorithm and its associated dynamical system admit various descriptions, that can be found for instance in [1, 27, 8]. It appears under various names; it is closely related to the Podsypanin modified Jacobi–Perron algorithm [25], it is also called the d-dimensional Gauss transformation in [15] or the ordered Jacobi–Perron algorithm in [14]. For more precisions on the Brun algorithm, see the book [28].

This dynamical system belongs to the class of multidimensional unimodular continued fraction algorithms, described in the book [28]. The algorithms of this class produce simultaneous diophantine approximations of a real vector. Then, the literature, for instance in [21, 8, 29], mainly focuses on the convergence of these approximations, closely related to the Lyapunov exponents of the underlying dynamical systems. These algorithms have also important applications to discrete geometry [4, 18], where they are used in small dimensions.

With each unimodular multidimensional continued fraction algorithm, a gcd algorithm may be of course associated, such as underlined in [31]. However, the probabilistic analysis of such a class of gcd algorithms has not been yet considered. This is our general project, and we begin by the `BrunGcd` Algorithm, as its associated dynamical system is based on one of the most "natural" extensions of the Euclid dynamical system [2] and shares many important properties with this last system. In particular, the Brun invariant density admits a "semi-explicit" form which will be very useful in the analysis. The `BrunGcd` algorithm was already studied in [22], in relation with efficient exponentiation with precomputation proposed by de Rooij [10], but only in the worst-case.

Main results. We perform the probabilistic analysis of the `BrunGcd` algorithm. We first focus on the total number of steps and prove that it is on average linear in the size of the input. The dominant constant equals $(d+1)/\mathcal{E}_d$ where $\mathcal{E}_d$ is the entropy of the Brun dynamical system. This entropy is not precisely studied in the literature, but there exists a conjecture due to Hardscastle and Khanin [14] that states that $\mathcal{E}_d$ is $\Theta(1)$ for $d \to \infty$. However, we show that $\mathcal{E}_d$ satisfies $\mathcal{E}_d \sim \log d$ for $d \to \infty$ (see Theorem 2), so that the dominant constant for the number of steps grows as $d/\log d$.

We then compare the `BrunGcd` with another multiple gcd algorithm, the `PlainGcd` algorithm, described in Knuth's book [20], which deals with the classical one-dimensional Euclid dynamical system. The authors have already analyzed this algorithm in [6], and prove that the mean number of steps is also linear in the size of the output [3]. However, the dominant constant is independent of the dimension d and equals $2/\mathcal{E}_1$, where $\mathcal{E}_1$ is the entropy of the usual Euclid dynamical system (in dimension $d = 1$). We conclude that

[2] Another natural choice could have been the Jacobi–Perron algorithm, but its invariant density is not known at all; we will return to this choice in the conclusion.

[3] This is not exactly the same notion of size: in [6], the input size is the sum of the sizes of the input components, whereas, in the present paper, the input size is the maximum of the sizes of the input components.

the PlainGcd algorithm is much more efficient than the BrunGcd algorithm, in particular for large dimensions d.

We finally explain the inefficiency of the BrunGcd algorithm when d is large: almost all the divisions deal with a quotient equal to 1. Then the main operation performed is not a division but... a plain subtraction. This is reinforced by the comparison with the subtractive version of the algorithm, whose number of steps is proven for $d > 1$ to be also of linear complexity, moreover with a small multiplicative constant. This exhibits a strong difference with the classical Gcd algorithm (case $d = 1$).

Methods. We use here the methods of *dynamical analysis* such as developed in [3, 23, 30]: a gcd algorithm is viewed as a dynamical system, with each iterative step being a linear fractional transformation. Costs of interest are then described with Dirichlet generating functions that are algebraically related to transfer operators of the system. The main analytical property of these series, namely the existence of a dominant pole, is itself closely related to the existence of a dominant eigenvalue, together with a spectral gap, for the corresponding transfer operators. The asymptotic extraction of coefficients is then achieved by means of Tauberian theorems.

Plan of the paper. Section 2 introduces the algorithm and states the main results in Theorems 1 and 2. Then, in Section 3, we study the underlying Brun dynamical system in d dimensions, which involves the whole family of Brun dynamical systems in smaller dimension. We provide a description of its rational trajectories. The following sections perform the dynamical analysis of the algorithm, with its two steps, the combinatorial step (Section 4) and the analytic step (Sections 5 and 6). The analytic step is based on spectral properties of transfer operators and their quasi-inverses that are described in Section 5, whereas Section 6 gathers all required elements to conclude the proof of Theorem 1. Section 7 is devoted to the behaviour of the Brun dynamical system in high dimensions, and leads to the proof of Theorem 2. The paper ends with a conclusion and open problems.

This paper is an extended version of the short conference paper [5]. In this short paper, most of the results of the two Sections 5 and 7 are just mentioned, and not proven. We provide here a complete proof of these results. Moreover, we also present new results, as for instance the study of the "remaining phases" (that is, the phases that intervene when the number of entries decreases), performed in Sections 7.4 and 7.5.

Acknowledgements. We wish to thank Jeffrey Shallit for mentioning to us the article [22]. We also wish to thank all the members of the project DynA3S for many discussions about the class of multidimensional continued fractions algorithms, and specially Anne Broise-Alamichel for exchanges on the Brun Algorithm.

2. The Brun Gcd Algorithm.

We describe the BrunGcd algorithm and state the main complexity results (in the worst-case and in the average-case). We describe the main constants of the analysis in terms of characteristics of the underlying dynamical system. We compare the present algorithm with the PlainGcd algorithm introduced by Knuth.

2.1. General description.

The algorithm `BrunGcd`(d) computes the gcd of $(d+1)$ positive integers. It deals with the input set $\Omega_{(d)}$ which gathers the *ordered* $(d+1)$-uples $\boldsymbol{u}$ formed with *positive* and *distinct* integer numbers

$$\Omega_{(d)} := \{\boldsymbol{u} = (u_0, u_1, \ldots, u_d) \mid u_0 > u_1 > u_2 > \ldots > u_d > 0\}\,.$$

During the execution of the algorithm, some components "disappear" and the algorithm deals with the *disjoint union*

$$\Gamma_{(d)} = \bigoplus_{\ell=0}^{d-1} \Omega_{(d-\ell)}\,.$$

The algorithm `BrunGcd`(d) performs a sequence of steps, and each step deals with the pair (u_0, u_1) (that contains the two largest entries of $\boldsymbol{u}$) and the list $\texttt{End}\,\boldsymbol{u}$ which gathers all the components of $\boldsymbol{u}$ except u_0; it divides the first component u_0 by the second component u_1, and creates a remainder v_0 [4]

$$v_0 := u_0 - mu_1, \quad m := \left[\frac{u_0}{u_1}\right]\,.$$

Then, the procedure $\texttt{InsDis}\,(v_0, \texttt{End}\,\boldsymbol{u})$ inserts $v_0 \geq 0$ at a suitable position inside the list $\texttt{End}\,\boldsymbol{u}$, so that the result remains an ordered uple of distinct positive values: the second component u_1 becomes the largest one, and there are three possible cases for the insertion (or not) of v_0:

- (G) (Generic case) if v_0 is not present in the list $\texttt{End}\,\boldsymbol{u}$, this is a usual insertion;
- (Z) (Zero case) if $v_0 = 0$, we *do not* insert v_0;
- (E) (Equality case) if $v_0 \neq 0$ is already present in the list $\texttt{End}\,\boldsymbol{u}$ at position i, we *do not* insert v_0.

In each of the cases (Z) or (E), we do not insert v_0, but we memorize the potential insertion position (in case (E), we would have inserted v_0 "in front of" u_i). Finally, each *step* of the algorithm `BrunGcd`(d) is described by the map $U_{(d)} : \Gamma_{(d)} \to \Gamma_{(d)}$ which associates with $\boldsymbol{u}$

$$U_{(d)}(\boldsymbol{u}) = \texttt{InsDis}\,(u_0 \bmod u_1, \texttt{End}\,\boldsymbol{u})\,. \tag{1}$$

The algorithm `BrunGcd`(d), described in Figure 1, decomposes into d *phases*, labelled from $\ell = 0$ to $\ell = d-1$. During each phase, a component is "lost", and the ℓ-th phase, denoted by $\texttt{BrunGcd}_{(d,\ell)}$, transforms an element of $\Omega_{(d-\ell)}$ into an element of $\Omega_{(d-\ell-1)}$. The phase ends as soon as it meets case (Z), or [5] case (E), where it looses a component. The algorithm stops at the end of the $(d-1)$-th phase with an element of $\Omega_{(0)}$ which equals the gcd.

In the analysis of the `BrunGcd` algorithm, as usual, we gather the inputs with respect to their length. Here, we choose as the length of the input $\boldsymbol{u} := (u_0, u_1, \ldots u_d)$ the length of its maximal component, namely u_0. The size of the entry is then at most $d \log u_0$. We then introduce the set of inputs

$$\Omega_{(d,N)} := \{\boldsymbol{u} \in \Omega_{(d)} \mid u_0 \leq N\}\,, \tag{2}$$

[4] The subtractive version of the `BrunGcd` algorithm is defined in the same way with a subtraction $u_1 - u_0$ being performed instead of a division.

[5] The $(d-1)$-th phase always ends with the case (Z).

`BrunGcd`(d)
Input : $\boldsymbol{u} \in \Omega_{(d)}$
Ouput: $\boldsymbol{u} \in \mathbb{N}$
For $\ell = 0$ to $d-1$ **do** $\boldsymbol{u} := \texttt{BrunGcd}_{(d,\ell)}(\boldsymbol{u})$;

`BrunGcd`$_{(d,\ell)}$
Input : $\boldsymbol{u} \in \Omega_{(d-\ell)}$
Repeat $\boldsymbol{u} := U_{(d)}(\boldsymbol{u})$ **until** $u \in \Omega_{(d-\ell-1)}$.

Fig. 1. The `BrunGcd` algorithm.

and we study the maximum number of steps of the algorithm on $\Omega_{(d,N)}$ (in the worst-case analysis) or the mean number of steps of the algorithm when $\Omega_{(d,N)}$ is endowed with the uniform probability (in the average-case analysis).

2.2. Worst-case behaviour of the Brun Gcd Algorithm.

As shown in [22], the worst-case of the `BrunGcd` algorithm arises when the quotients are the smallest possible (all equal to 1, except the last one, equal to 2), and the insertion positions the largest possible. Then, the best worst-case bound involves an algebraic number τ_d which extends to general dimensions the inverse of the Golden ratio. The real τ_d will also intervene in the study of the associated dynamical system (see for instance Eq. (33) and (34)) and is central in a conjecture that is related to the average-case behaviour of the algorithm (see Section 7.4).

Proposition 1. For any integer d, and any integer N, the maximum number $Q_{(d,N)}$ of steps of the `BrunGcd` Algorithm on the set $\Omega_{(d,N)}$ defined in (2) satisfies

$$Q_{(d,N)} \sim \frac{1}{|\log \tau_d|} \log N \qquad (N \to \infty)\,, \tag{3}$$

and involves the smallest real root $\tau_d \in]0,1[$ of the polynomial $z^{d+1}+z-1$. When $d \to \infty$, the reals τ_d and $|\log \tau_d|$ admit the estimates

$$\tau_d^{d+1} = 1 - \tau_d = \frac{1}{d}\left(\log d - \log\log d + o(1)\right), \qquad \frac{1}{|\log \tau_d|} \sim \frac{d+1}{\log d}, \quad (d \to \infty)\,. \tag{4}$$

Proof. The proof is mainly based on the precise description of the worst-case configuration, which is provided in [22], and is now summarized. This worst-case configuration arises when all the quotients are equal to 1, and the insertion positions the largest possible. It corresponds to the sequence $\boldsymbol{x}_{(d)} = (\boldsymbol{x}_{(d),i})_{i \geq -d}$ defined by the following linear recurrence:

$$x_{(d),i} = 1 \quad (\text{for } -d \leq i \leq 0), \qquad x_{(d),i} = x_{(d),i-1} + x_{(d),i-1-d}, \quad (\text{for } i \geq 1)\,.$$

It is indeed proven in [22] that the `BrunGcd` algorithm performs exactly n steps on the input $(x_{(d),n}, x_{(d),n-1}, \ldots, x_{(d),n-d}) \in \Omega_{(d)}$. This entails that the maximum number $Q_{(d,N)}$ of steps performed by the `BrunGcd` Algorithm on $\Omega_{(d,N)}$ satisfies

$$Q_{(d,N)} := \max\{n \in \mathbb{N} \mid x_{(d),n} \leq N\}.$$

As the sequence $\boldsymbol{x}_{(d)}$ is defined by a linear recurrence, the generating function of the sequence $(\boldsymbol{x}_{(d),n})_{n\geq 0}$ is a rational fraction $F_{(d)}(z)$ that satisfies

$$F_{(d)}(z) := \sum_{n\geq 0} x_{(d),n} z^n = \frac{z^{d+1}-1}{z-1} \cdot \frac{1}{1-z-z^{d+1}}\,.$$

The polynomial $1-z-z^{d+1}$ has a unique real root in $]0,1[$: it is simple, denoted by τ_d, and any other root y satisfies $|y| > \tau_d$. Then, the function $F_{(d)}$ has a dominant pole at $z=\tau_d$, and the coefficient $x_{(d),n}$ of z^n in $F_{(d)}(z)$ is of exponential growth as $n\to\infty$, namely

$$x_{(d),n} \sim \chi_d \cdot \tau_d^{-n} \qquad (n\to\infty)\,. \tag{5}$$

For d fixed, this entails the estimate given in (3). Now, the asymptotics of τ_d (for $d\to\infty$) is obtained by letting $\tau_d := 1-\varepsilon_d$ in the equation $\tau_d^{d+1}+\tau_d = 1$. This gives the estimates for τ_d and $\log\tau_d$ given in (4). Replacing the estimate of τ_d in the residue of $F_{(d)}$ at $z=\tau_d$ gives an asymptotic estimate for χ_d, namely

$$\chi_d \sim \left(\frac{1}{\log d}\right) \tau_d^{-d} \qquad (d\to\infty)\,. \tag{6}$$

□

2.3. Probabilistic behaviour: main results

We now describe the precise probabilistic behaviour of the algorithm `BrunGcd`(d) on the set $\Omega_{(d,N)}$ defined in (2).

Theorem 1. When the algorithm `BrunGcd` acts on the set $\Omega_{(d,N)}$ endowed with the uniform distribution, the following holds when d is fixed and N tends to ∞:

(*a*) The total number L_d of steps and the number M_d of steps performed during the first phase satisfy

$$\mathbb{E}_N[L_d] \sim \mathbb{E}_N[M_d] \sim \frac{d+1}{\mathcal{E}_d} \cdot \log N, \qquad (N\to\infty)\,,$$

and involve the entropy $\mathcal{E}_d$ of the underlying Brun dynamical system.

(*b*) The total number R_d of steps performed during the remainder of the execution (after the first phase) has a mean value that is asymptotic to a constant ρ_d. For each $\ell > 1$, the number of steps $R_{d,\ell}$ performed during the ℓ-th phase of the execution has a mean value that is asymptotic to a constant $1+\rho_{d,\ell}$ (with $\rho_{d,\ell} > 0$).

(*c*) Let O_d be the number of quotients equal to 1 during the first phase. The ratio between the means $\mathbb{E}_N[O_d]$ and $\mathbb{E}_N[M_d]$ is asymptotic to a constant $\theta_d < 1$

$$\frac{\mathbb{E}_N[O_d]}{\mathbb{E}_N[M_d]} \to \theta_d, \qquad (N\to\infty)\,.$$

(*d*) Let Σ_d be the number of steps of the subtractive version of `BrunGcd` during the first phase. For $d>1$, the ratio between $\mathbb{E}_N[\Sigma_d]$ and $\mathbb{E}_N[M_d]$ is asymptotic to a constant σ_d,

$$\frac{\mathbb{E}_N[\Sigma_d]}{\mathbb{E}_N[M_d]} \to \sigma_d, \qquad (N\to\infty)\,,$$

whereas for $d=1$ (case of the Euclid Algorithm), this ratio is of order $\Theta(\log N)$.

The main constants which appear in the analysis, namely the constants $\mathcal{E}_d, \theta_d, \sigma_d, \rho_d$ are *mathematical* constants, more precisely *dynamical* constants, as they are defined via the dynamical system underlying the `BrunGcd` algorithm. This system will be precisely described in Section 3. It is defined on the simplex

$$\mathcal{J}_{(d)} = \{\boldsymbol{x} = (x_1, \ldots, x_d \mid 1 \geq x_1 \geq \ldots \geq x_d \geq 0\}\,,$$

and admits an invariant density Ψ_d on $\mathcal{J}_{(d)}$, described in Eq. (37) of Section 5.5. The main constants which appear in Theorem 1 are expressed with this invariant density, as it is now stated:

Theorem 2. Consider the measure ν_d associated with the invariant density Ψ_d defined on the simplex $\mathcal{J}_{(d)}$, and the function $\mu_d : [0,1] \to [0,1]$ which associates with y the measure $\nu_d(y\mathcal{J}_{(d)})$ of the homothetic simplex $y\mathcal{J}_{(d)}$. The following holds:

(*a*) The constants of the analysis admit expressions which involve the function μ_d

$$\mathcal{E}_d = (d+1)\int_0^1 \mu_d(y)\frac{dy}{y}, \qquad \theta_d = 1 - \mu_d\left(\frac{1}{2}\right), \qquad \sigma_d = \sum_{m\geq 1} \mu_d\left(\frac{1}{m}\right).$$

(*b*) Consider the two functions $y(d) := (\log d)/d$ and $\epsilon(d) = A(\log\log d)^{-1}$ for some[6] constant $A > 0$. When $d \to \infty$, the following asymptotic estimates hold for the function $\mu_d : y \mapsto \mu_d(y)$, and exhibit two different regimes

$$\begin{cases} \mu_d(y) \in \left[\left(\dfrac{\log dy}{\log d}\right)^{d-1/2+d\,\epsilon(d)}, \left(\dfrac{\log dy}{\log d}\right)^{d-1/2-d\,\epsilon(d)}\right] & \text{for } y \in [y(d), 1]\,, \\ \mu_d(y) \leq (\log d)^{1/2}\left(\dfrac{dy}{\log d}\right)^d & \text{for } y \in [0, y(d)]\,. \end{cases}$$

(*c*) The following asymptotic estimates hold for the main constants of the analysis:

$$\mathcal{E}_d = \log d\left[1 + O(\log\log d)^{-1}\right],\ \ \theta_d = 1 - (1/2)^{\Theta(1)d\log d},\ \ 1 \leq \sigma_d \leq 2 + O(\log\log d)^{-1}\,.$$

The case of the constants $\rho_{d,\ell}$ which intervene in Thm 1(*b*) is different. These constants admit a (complicate) mathematical expression from which we did not succeed to derive an exact asymptotics. However, we present a conjecture that is based on some plausible heuristics and seems to be validated by experimental results (see Section 7.4).

Conjecture. Denote by τ_k the algebraic number which is the unique root < 1 of the equation $x^{k+1} + x - 1 = 0$ and appears in the worst-case analysis of the algorithm in k dimensions. Then, there are two absolute constants $a_- > 0$ and a_+ for which the constant $\rho_{d,\ell}$ of Theorem 1(*b*) satisfies for any $d > 1$, $\ell \in [1, d]$,

$$\rho_{d,\ell} = A(d,\ell)\cdot\frac{\tau_{d-\ell}^{\ell}}{1-\tau_{d-\ell}^{\ell}} \qquad \text{with} \quad a_- \leq A(d,\ell) \leq a_+\,.$$

[6] The constant A can be explicitly computed.

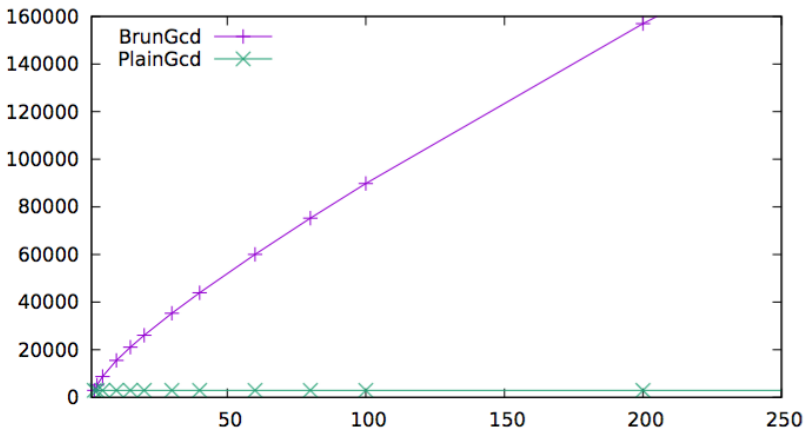

Fig. 2. Comparison between the mean number of steps performed by two algorithms, the `BrunGcd` algorithm, and the `PlainGcd` algorithm, as a function of the dimension d. The size N is fixed.

2.4. Comments on the main results.

We now comment and explain our main results.

In Thm 1(a), the total number of steps L_d is proven to remain on average linear in the size $\log N$. Moreover, (a) and (b) exhibit a strong difference between the first phase (where most of the computation is done), and the remainder of the execution (where the total number of steps R_d is on average asymptotically constant). The `PlainGcd` algorithm exhibits exactly the same phenomena, as it is shown in [6].

Figure 2 compares the number of steps of the `BrunGcd` and the `PlainGcd` algorithms, as a function of the dimension d, when the binary size is fixed to $\log_2 N = 5 \cdot 10^3$. The complexity of the `BrunGcd` algorithm appears to be sublinear with respect to d (as proven in Thm 1(a) and Thm 2(c)) whereas the complexity of the `PlainGcd` algorithm appears to be independent of d (as proven in [6] in the average case).

The two dominant constants, the ratio $(d+1)/\mathcal{E}_d$ which involves the entropy and intervenes in Thm 1(a), and Thm 2(c), on one hand, and the ratio $1/|\log \tau_d|$ which arises in the worst-case (see Proposition 1), on the other hand, both behave as $d/\log d$ for $d \to \infty$. This exhibits the same behaviour for the algorithm in the average-case and in the worst-case. As the worst-case is reached when the quotients are all equal to 1 (as proven in [22]), this seems to show that the `BrunGcd` Algorithm deals with quotients which are very often equal to 1.

This is indeed the case, as described in Thm 1(c) and Thm 2(c), and also illustrated in Figure 3, that exhibits the proportion of quotients equal to 1 during the first phase as a function of the dimension d. This proportion tends quickly to 1: when $d = 16$, more than 99% of the Euclidean divisions are in fact subtractions and for $d = 50$, the proportion is 99.99%. This is in adequation with the estimate given in Thm 2(c), where we show that the speed is "quasi-exponential" as a function of d, namely of order $O(2^{-\Theta(1)d/\log d})$.

The result given in Thm 1(d) exhibits a strong difference between the Euclid Algorithm (case $d = 1$) and the `BrunGcd` algorithm for $d \geq 2$. In the Euclid algorithm, the mean

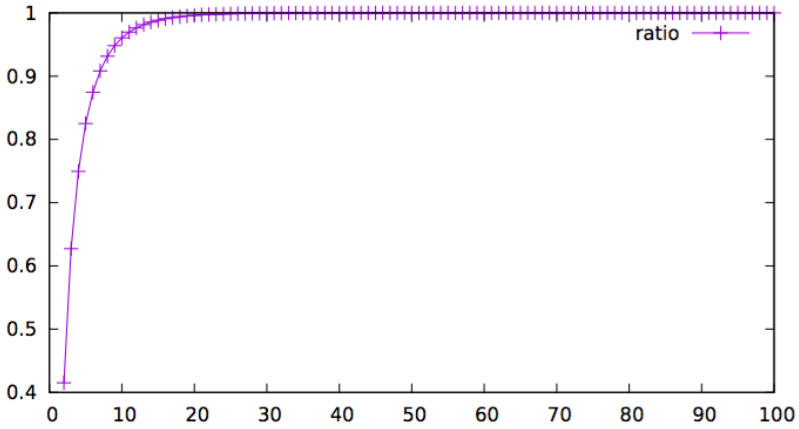

Fig. 3. Proportion of quotients equal to 1, during the first phase, as a function of the dimension d. The size N is fixed.

value of the quotient is infinite, whereas, in the present case, this mean value σ_d is finite. Moreover, Thm 2(c) provides an asymptotic estimate for the mean number of iterations of the subtractive algorithm (for $d \to \infty$), which shows that it is (asymptotically) sublinear with respect to d.

Moreover, the results given in Thm 2(b) and Thm 2(c) are of independent interest, as they describe the asymptotic behaviour of the `Brun` dynamical system (as $d \to \infty$), which is not often studied. In particular, the asymptotic estimates for the probability μ_d or the entropy $\mathcal{E}_d$ are new, and the estimate we provided for $\mathcal{E}_d$ contradicts the conjecture made by Hardcastle and Khanin in [14].

3. The underlying dynamical system.

We first describe a continuous extension of the algorithm, relate it with the Brun dynamical system, and provide an exact characterization of the trajectories that are related to the execution of the algorithm.

3.1. Continuous extension of the Brun Gcd algorithm.

We now extend the `BrunGcd`(d) algorithm into a continuous process. We use the projection π defined on $\Gamma_{(d)} \setminus \{\mathbf{0}\}$ as

$$\pi(\boldsymbol{u}) = \frac{1}{u_0}\mathtt{End}\,\boldsymbol{u}\,,$$

and the closure of the image $\pi(\Gamma_d \setminus \{\mathbf{0}\})$ is exactly the disjoint union $\mathcal{I}_{(d)} := \bigoplus_{\ell=0}^{d-1} \mathcal{J}_{(d-\ell)}$, where the simplex $\mathcal{J}_{(k)}$, is defined for $k \geq 1$ as

$$\mathcal{J}_{(k)} := \{\boldsymbol{x} = (x_1, \ldots, x_k) \mid 1 \geq x_1 \geq \ldots \geq x_k \geq 0\}\,. \tag{7}$$

The map $V_{(d)} : \mathcal{I}_{(d)} \to \mathcal{I}_{(d)}$ is now defined on each $\mathcal{J}_{(d-\ell)}$ by $V_{(d)}(\mathbf{0}^{\boldsymbol{d-\ell}}) = \mathbf{0}^{\boldsymbol{d-\ell}}$, and

$$V_{(d)}(\boldsymbol{x}) = \mathtt{InsDis}\left(\left\{\frac{1}{x_1}\right\}, \frac{1}{x_1}\mathtt{End}\,\boldsymbol{x}\right) \quad \text{for } \boldsymbol{x} \neq \mathbf{0}^{\boldsymbol{d-\ell}}\,,$$

where $\texttt{InsDis}(y_0, \boldsymbol{y})$ is now extended to $\mathcal{I}_{(d)}$. The map $V_{(d)}$ provides the extension of the conjugate with the projection π of the map $U_{(d)}$ defined in (1) and used in the $\texttt{BrunGcd}(d)$ algorithm. Indeed, the equality $V_{(d)} \circ \pi(\boldsymbol{u}) = \pi \circ U_{(d)}(\boldsymbol{u})$ leads to the definition of $V_{(d)}$ on the set $\pi(\Gamma_d) \setminus \{\mathbf{0}\})$ which is further extended to $\mathcal{I}_{(d)}$ "by continuity".

The dynamical system $(V_d, \mathcal{I}_{(d)})$ coincides with the usual ordered version of Brun dynamical system, and is precisely described in [15] for instance. It is defined on the simplex $\mathcal{J}_{(d)}$ with the transformation $T_{(d)}$ defined by $T_{(d)}(\mathbf{0^d}) = \mathbf{0^d}$, and

$$T_{(d)}(\boldsymbol{x}) = \texttt{Ins}\left(\left\{\frac{1}{x_1}\right\}, \frac{1}{x_1}\texttt{End}\,\boldsymbol{x}\right) \quad \text{for } \boldsymbol{x} \neq \mathbf{0^d}\,, \tag{8}$$

where now the map $\texttt{Ins}(y_0, \boldsymbol{y})$ is the usual insertion "in front of": it performs as `InsDis`, *without* removing zeroes and equal components, and the cases (Z) and (E) may be gathered into a unique case (ZE). There are now only two cases:

(G) if y_0 is not present in the list $\boldsymbol{y}$, this is an usual insertion;

(ZE) if y_0 is already present in the list $\boldsymbol{y}$, we insert y_0 *in front of the block* of components equal to y_0.

It is clear that $V_{(d)}(\boldsymbol{x}) = T_{(d)}(\boldsymbol{x})$ except in case (ZE). Then, the two dynamical systems coincide "almost everywhere" and the trajectories which do not meet case (ZE) will be the same for the two systems. But we are mainly interested in rational trajectories, and the rational trajectories may meet case (ZE) and differ in the two systems. For instance, the two trajectories of the input $\pi(4,3,2,1) = (3/4, 2/4, 1/4)$ are (the inserted component is in bold):

$$\begin{aligned}
&\text{for } V_{(d)}\text{:}\quad (3/4, 2/4, 1/4) \to (2/3, 1/3) \to (1/2, 1/3) \to (1/3) \to (0)\\
&\text{for } T_{(d)}\text{:}\quad (3/4, 2/4, 1/4) \to (2/3, \mathbf{1/3},\ 1/3) \to (\mathbf{1/2}, 1/2, 1/2) \to (1, 1, \mathbf{0}) \to (1, \mathbf{0}, 0) \to (\mathbf{0}, 0, 0)\,.
\end{aligned}$$

However, we will prove in Section 3.3 that a rational trajectory under $V_{(d)}$ –which exactly describes an execution of the $\texttt{BrunGcd}(d)$ algorithm– mainly decomposes into rational trajectories under $T_{(d-\ell)}$. Indeed, except possibly at the end of each phase, it uses `Ins` and not `InsDis`. This is why the `BrunGcd` algorithm will use, except at the end of each phase, the Brun dynamical system, which we now describe.

3.2. The Brun dynamical system.

The pair $(T_{(d)}, \mathcal{J}_{(d)})$ defines a dynamical system denoted as $\mathcal{D}_{(d)}$. For any $\boldsymbol{x} \in \mathcal{J}_{(d)}$, the map $T_{(d)}$ defined in (8) uses *digit* (m, j) formed with a *quotient* $m(\boldsymbol{x}) \geq 1$ and an *insertion index* $j(\boldsymbol{x}) \in [1..d]$. The set of digits is thus $\mathcal{A}_{(d)} := \mathbb{N}^* \times [1..d]$. We associate with (m, j) the subset

$$\mathcal{K}_{(d,m,j)} := \{\boldsymbol{x} \in \mathcal{J}_{(d)} \mid m(\boldsymbol{x}) = m, \quad j(\boldsymbol{x}) = j\}\,.$$

When (m, j) varies in $\mathcal{A}_{(d)}$, the subsets $\mathcal{K}_{(d,m,j)}$ form a topological partition of $\mathcal{J}_{(d)}$, and the restriction $T_{(d,m,j)}$ of $T_{(d)}$ to $\mathcal{K}_{(d,m,j)}$ is a bijection from $\mathcal{K}_{(d,m,j)}$ onto $\mathcal{J}_{(d)}$, written as

$$T_{(d,m,j)}(x_1, x_2, \ldots, x_d) = \left(\frac{x_2}{x_1}, \ldots, \frac{x_{j-1}}{x_1}, \frac{1}{x_1} - m, \frac{x_{j+1}}{x_1}, \ldots, \frac{x_d}{x_1}\right).$$

Its inverse is a bijection from $\mathcal{J}_{(d)}$ onto $\mathcal{K}_{(d,m,j)}$ written as

$$h_{(d,m,j)}(y_1,\dots,y_d) = \left(\frac{1}{m+y_j}, \frac{y_1}{m+y_j}, \dots, \frac{y_{j-1}}{m+y_j}, \frac{y_{j+1}}{m+y_j}, \dots, \frac{y_d}{m+y_j}\right) . \quad (9)$$

Any inverse branch of the map $T_{(d)}$ is called an *elementary inverse branch* (or an *inverse branch of depth one*) and the set of the elementary inverse branches is then

$$\mathcal{H}_{(d)} := \left\{ h_{(d,m,j)} \mid (m,j) \in \mathcal{A}_{(d)} \right\} , \quad (10)$$

whereas the inverse branches of the map $T^k_{(d)}$ are said to be of *depth* k and belong to the set $\mathcal{H}^k_{(d)}$. The set

$$\mathcal{H}^\star_{(d)} := \bigoplus_{k\geq 0} \mathcal{H}^k_{(d)} \quad (11)$$

describes all the truncated finite trajectories of the system $\mathcal{D}_{(d)}$.

3.3. Return to the Gcd Algorithm.

On an input $\boldsymbol{u} \in \Omega_{(d)}$, the execution of the $\texttt{BrunGcd}(d)$ algorithm is described by the trajectory of $\boldsymbol{u}$ under the map $U_{(d)}$ which ends at $\gcd(\boldsymbol{u})$. It proves useful to consider one more step, and now, the trajectory of $\boldsymbol{u}$ under the map $U_{(d)}$ ends at 0. It gives rise to the rational trajectory of the vector $\boldsymbol{x} := \pi(\boldsymbol{u})$ under $V_{(d)}$, that also ends at 0. This trajectory uses at each step a branch of the map $V_{(d)}$. We wish to precisely describe the set $\mathcal{B}_{(d)}$ of compositions of inverse branches of the map $V_{(d)}$ which are possibly used by such particular "stopping" trajectories. Then the equality $\pi(\boldsymbol{u}) = h(0)$ (for $h \in \mathcal{B}_{(d)}$) gives rise to a bijection between $\pi(\Omega_{(d)})$ and $\mathcal{B}_{(d)}$. Moreover, there is also a bijection between $\pi(\Omega_{(d)})$ and the set of coprime inputs

$$\underline{\Omega}_{(d)} := \{u \in \Omega_{(d)} \mid \gcd(\boldsymbol{u}) = 1\} , \quad (12)$$

and thus a bijection between $\mathcal{B}_{(d)}$ and $\underline{\Omega}_{(d)}$.

We now describe $\mathcal{B}_{(d)}$, and first focus, for each $\ell \in [0..d-1]$, on the set $\mathcal{P}_{(d-\ell)}$ used by the part of the trajectory associated with the ℓ-th phase, when the iterates of $\boldsymbol{u}$ under $U_{(d)}$ belong to $\Omega_{(d-\ell)}$. Then, the iterates of $\boldsymbol{x}$ under $V_{(d)}$ belong to the simplex $\mathcal{J}_{(d-\ell)}$. We study in a separate way

(a) the steps of the phase of index ℓ, which are not the last one, which constitute what we will call the *strict phase* of index ℓ;
(b) and the last step of the phase of index ℓ.

Consider first (a). In this case, such a step does not loose a component, and only involves a step of type (G). Then, the trajectory uses branches of the map $T_{(d-\ell)}$, and each step uses possibly any inverse branch in the set $\mathcal{H}_{(d-\ell)}$ defined in (10). Hence, the set of the inverse branches used during the strict ℓ-phase is the set $\mathcal{H}^\star_{(d-\ell)}$.

Consider (b). Now, a component is lost since the insertion is not done, and there are two possible cases, namely Case (Z) and Case (E).

Case (Z). The quotient $1/x_1$ is equal to an integer $m \geq 2$, the position of potential insertion is $j = d - \ell$. The equality $V_{(d)}(\boldsymbol{x}) = m \cdot \texttt{End}\,\boldsymbol{x}$ holds and the inverse branch associates with the $(d-\ell-1)$-uple $\boldsymbol{y}$ the $(d-\ell)$-uple

$$z_{(d-\ell,m)}(\boldsymbol{y}) = \frac{1}{m}(1, \boldsymbol{y}) . \quad (13)$$

The set $\mathcal{Z}_{(d-\ell)}$ of inverse branches used in case (Z) is thus

$$\mathcal{Z}_{(d-\ell)} := \{z_{(d-\ell,m)} \mid m \geq 2\}\,.$$

Case (E). An equality of the form $(1/x_1) - m = x_i/x_1$ holds with a quotient $m \geq 1$ and a potential insertion position [7] $j = i - 1 < d - \ell$. Then, Case (E) cannot occur for $\ell = d-1$. The equality $V_{(d)}(\boldsymbol{x}) = (\mathtt{End}\,\boldsymbol{x})/(1-x_i)$ holds and the inverse branch associates with the $(d-\ell-1)$-uple $\boldsymbol{y}$ the $(d-\ell)$-uple

$$s_{(d-\ell,m,j)}(\boldsymbol{y}) = \frac{1}{m+y_j}\,(1,\boldsymbol{y})\,. \tag{14}$$

The set of inverse branches $\mathcal{S}_{(d-\ell)}$ used in case (E) is thus

$$\mathcal{S}_{(d-\ell)} := \emptyset \quad (\text{for } \ell = d-1), \qquad \mathcal{S}_{(d-\ell)} := \{s_{(d-\ell,m,j)} \mid m \geq 1,\, j < d-\ell\} \quad (\text{for } \ell < d-1).$$

In summary, the set of possible inverse branches used during the last step of the ℓ-th phase is

$$\mathcal{F}_{(d-\ell)} := \mathcal{S}_{(d-\ell)} \cup \mathcal{Z}_{(d-\ell)}. \tag{15}$$

Finally, we have proven the following (recall that $\mathcal{H}^{\star}_{(d-\ell)}$ is defined in (11)):

Proposition 2. The $\mathtt{BrunGcd}(d)$ algorithm builds a bijection between the set $\underline{\Omega}_{(d)}$ of coprime inputs of $\Omega_{(d)}$ and the set $\mathcal{B}_{(d)}$ of inverse branches possibly used by the rational trajectories of the shift $V_{(d)}$. This set is written as

$$\mathcal{B}_{(d)} := \mathcal{P}_{(d)} \circ \mathcal{P}_{(d-1)} \circ \ldots \circ \mathcal{P}_{(1)} = \mathcal{P}_{(d)} \circ \mathcal{B}_{(d-1)}$$

and involves the sets $\mathcal{P}_{(d-\ell)}$ of inverse branches used by the $\mathtt{BrunGcd}_{(d,\ell)}$, characterized as

$$\mathcal{P}_{(d-\ell)} := \mathcal{H}^{\star}_{(d-\ell)} \circ \mathcal{F}_{(d-\ell)}.$$

3.4. Why InsDis rather than Ins?

We have decided to deal with the set $\Gamma_{(d)}$. We are then led to use `InsDis` to stay inside $\Gamma_{(d)}$. But there is a natural question: why not dealing with the set $\Gamma_{(=,d)}$ with possible blocks of equal non-zero components? Then, we could use `Ins` and stay inside $\Gamma_{(=,d)}$. This defines another algorithm, the $\mathtt{BrunGcd}_{=}(d)$ algorithm, whose continuous extension directly leads to the Brun dynamical system. Even though the whole path seems more natural, we have to memorize the position of each block of equal components, and this leads to a quite involved analogue set $\mathcal{B}_{(=,d)}$ that describes the rational trajectories of the $\mathtt{BrunGcd}_{=}(d)$ algorithm.

4. Dynamical analysis (I).

We now begin the analysis of the algorithm, and introduce the main objects: the class of costs of interest, the (Dirichlet) generating functions, the generating operators. The main result of this section (Theorem 3) relates generating functions and generating operators.

[7] We recall that we would have inserted "in front of".

4.1. Additive costs.

We consider here costs that are said to be additive. One begins with a nonnegative *elementary* cost c defined on each inverse branch in $\mathcal{B}_{(d)}$ of depth one[8]. Such a cost is then extended in an additive way on $\mathcal{B}_{(d)}$, namely

$$c(h_1 \circ h_2 \circ \cdots \circ h_p) := \sum_{i=1}^{p} c(h_i).$$

Now, a cost C defined on $\Omega_{(d)}$ is said to be *additive* if it is associated with such a cost c, and satisfies

$$C(\boldsymbol{u}) := c(h) \qquad \text{when } \pi(\boldsymbol{u}) = h(0).$$

Then $C(\boldsymbol{u})$ equals the total cost on the trajectory of $\pi(\boldsymbol{u})$ and satisfies $C(\boldsymbol{u}) = C(\lambda \boldsymbol{u})$ for any integer $\lambda \neq 0$.

There are three main additive costs of interest here.

(*i*) The total number $C_{d,\ell}$ of steps during the strict ℓ-th phase, associated with the characteristic function c of the set $\mathcal{H}_{(d-\ell)}$; with the family $C_{d,\ell}$, we return to the number of steps mentioned in Theorem 1, with the relations

$$L_d = M_d + R_d, \qquad M_d = 1 + C_{d,0}, \qquad R_d = \sum_{\ell=1}^{d-1} R_{d,\ell} \qquad R_{d,\ell} = 1 + C_{d,\ell}\,. \quad (16)$$

(*ii*) The number O_d of steps of the strict first phase with a quotient equal to 1, associated with the characteristic function c of the subset

$$\{h_{(d,m,j)} \in \mathcal{H}_{(d)} \mid m = 1\}\ .$$

(*iii*) The number Σ_d of steps of the subtractive algorithm during the strict first phase, associated with the cost c which associates with an inverse branch h of the first phase its quotient m.

The associated elementary costs deal with a specific strict phase : the strict ℓ-phase for $C_{d,\ell}$ and the strict first phase of index $\ell = 0$ for M_d, O_d, R_d. The cost C is said to *be concentrated* on this (strict) phase: more precisely, the cost $C_{d,\ell}$ is concentrated on the strict phase of index ℓ, and the costs M_d, O_d on the strict first phase of index 0.

4.2. Dirichlet generating functions.

The (basic) Dirichlet generating function $S_{(d)}(s)$, of the input set $\Omega_{(d)}$ relative to the length $\|\boldsymbol{u}\| := u_0$, is defined as

$$S_{(d)}(s) := \sum_{\boldsymbol{u} \in \Omega_d} \frac{1}{||\boldsymbol{u}||^s}\,. \quad (17)$$

In the same vein, with a cost $C : \Omega_{(d)} \to \mathbb{N}$, we associate the cumulative generating function of the cost

$$\widehat{S}_{(d,C)}(s) := \sum_{\boldsymbol{u} \in \Omega_d} \frac{C(\boldsymbol{u})}{||\boldsymbol{u}||^s} = \sum_{n \geq 1} n^{-s} \sum_{\|\boldsymbol{u}\| = n} C(\boldsymbol{u})\,. \quad (18)$$

When $\Omega_{(d,N)}$ is endowed with the uniform distribution, the mean $\mathbb{E}_N[C]$ of cost C on $\Omega_{(d,N)}$ is expressed with a quotient which involves the coefficients of the two previous

[8] Conditions required on the growth of the cost c in the general case are highlighted in Section 5.7.

generating functions, as

$$\mathbb{E}_N[C] = \frac{1}{\Phi_N\left[S_{(d)}\right]}\Phi_N\left[\widehat{S}_{(d,C)}\right], \tag{19}$$

where $\Phi_N[f]$ is defined as

$$\Phi_N[f] = \sum_{n\leq N} a_n \quad \text{when} \quad f(s) = \sum_{n\geq 1}\frac{a_n}{n^s}. \tag{20}$$

Generally speaking, singularity analysis relates the asymptotics of the coefficients of a generating function to the nature and the position of its dominant singularity (as we will see in Section 6.2). Then, we need an alternative expression of the series which highlights the singularities. Such an expression exists for the series $S_{(d)}(s)$, namely

$$S_{(d)}(s) = \frac{1}{d!}\sum_{n\geq d+1}\frac{1}{n^s}(n-1)(n-2)\dots(n-d).$$

As its "dominant" series is $(1/d!)\zeta(s-d)$, the series $S_{(d)}(s)$ has a dominant simple pole at $s = d+1$ with residue $1/d!$.

However, such a "dominant" behaviour is not known for the other Dirichlet series. The remaining of the paper is devoted to this task and is based on the dynamical analysis approach where generating functions are expressed with generating transfer operators. For this aim, we introduce an intermediate tool, the Dirichlet (bivariate) generating function,

$$S_{(d,C)}(s,w) := \sum_{\boldsymbol{u}\in\Omega_d}\frac{w^{C(\boldsymbol{u})}}{||\boldsymbol{u}||^s} \tag{21}$$

whose derivative is related to the cumulative generating function. One has

$$\widehat{S}_{(d,C)} = \Delta[S_{(d,C)}]$$

and involves the functional Δ defined as

$$\Delta[A](s) := \frac{\partial}{\partial w}A(s,w)\Big|_{w=1}. \tag{22}$$

As we wish to use the bijection described in Proposition 2, we also introduce the three underlined series that are the exact counterparts of series defined in (17), (18), (21), now with respect to the set $\underline{\Omega}_{(d)}$ of coprime inputs. As C is an additive cost, the non-underlined series are related to their underlined counterparts as

$$\frac{S_{(d)}(s)}{\underline{S}_{(d)}(s)} = \frac{S_{(d,C)}(s,w)}{\underline{S}_{(d,C)}(s,w)} = \frac{\widehat{S}_{(d,C)}(s)}{\underline{\widehat{S}}_{(d,C)}(s)} = \zeta(s). \tag{23}$$

4.3. Generating operators.

Let $\mathcal{G}$ be a set of inverse branches; we say that h is a *factor* of $\mathcal{G}$ if h is a factor of an element of $\mathcal{G}$ (with respect to the composition), and we denote this situation as $h \propto \mathcal{G}$. For instance in the same vein, a set $\mathcal{H}$ is said to be a factor of $\mathcal{G}$ if each element of $\mathcal{H}$ is a factor of $\mathcal{G}$, and this will be denoted as $\mathcal{H} \propto \mathcal{G}$. We will consider in the sequel factors of $\mathcal{B}_{(d)}$. The following easy result is central in this work.

Lemma 1. Any inverse branch $h \propto \mathcal{B}_{(d)}$ (defined in Prop. 2) is a linear fractional transformation (LFT), and any factor of $h \propto \mathcal{P}_{(d-\ell)}$ is written as

$$h = \frac{1}{D[h]}\Big(N_1[h], N_2[h], \ldots, N_{d-\ell}[h]\Big)\,,$$

where the denominator $D[h]$ and the numerators $N_i[h]$ are co-prime affine functions. When $h \in \mathcal{H}_{(d-\ell)}$, the determinant $J[h]$ of its Jacobian matrix is related to the denominator $D[h]$ via the equality

$$|J[h]| = |D[h]|^{-(d-\ell+1)}\ . \tag{24}$$

Proof. The proof (by recurrence on the depth of h) begins with the expression of the branches of depth 1 given in (9),(13),(14). □

Transfer operators are central tools for studying probabilistic properties of trajectories in dynamical systems; see e.g. [3, 23, 30]. Here, we adapt these tools to our needs, strongly use the fact that inverse branches are LFT, and also consider an additive cost c.

It proves useful to deal first with *component operators*: with a linear fractional transformation h and a cost c, we associate the (plain) operator and its cumulative version, i.e.,

$$\mathbf{G}_{s,\langle h\rangle}[f](\boldsymbol{x}) := |D[h](\boldsymbol{x})|^{-s}\, f\circ h(\boldsymbol{x}), \qquad \widehat{\mathbf{G}}_{s,\langle h\rangle}[f](\boldsymbol{x}) := c(h)|D[h](\boldsymbol{x})|^{-s}\, f\circ h(\boldsymbol{x})\,,$$

and also its (bivariate) weighted version,

$$\mathbf{G}_{s,w,[c],\langle h\rangle}[f](\boldsymbol{x}) := |D[h](\boldsymbol{x})|^{-s}\, w^{c(h)} f\circ h(\boldsymbol{x})\,.$$

Now, we associate with a subset $\mathcal{G} \propto \mathcal{B}_{(d)}$ the *generating operator* of the subset $\mathcal{G}$, under its three versions, the *plain*, the *cumulative*, and the *weighted versions*, namely

$$\mathbf{G}_{s,\langle\mathcal{G}\rangle} = \sum_{h\in\mathcal{G}} \mathbf{G}_{s,\langle h\rangle},\ \widehat{\mathbf{G}}_{s,[c],\langle\mathcal{G}\rangle} = \sum_{h\in\mathcal{G}} c(h)\,\mathbf{G}_{s,\langle h\rangle},\ \mathbf{G}_{s,w,[c],\langle\mathcal{G}\rangle} := \sum_{h\in\mathcal{G}} \mathbf{G}_{s,w,[c],\langle h\rangle}\,. \tag{25}$$

As for series, the first two operators are related to the third one via the equalities

$$\mathbf{G}_{s,\langle\mathcal{G}\rangle} = \mathbf{G}_{s,1,[c],\langle\mathcal{G}\rangle}, \qquad \widehat{\mathbf{G}}_{s,[c],\langle\mathcal{G}\rangle} = \Delta[w\mapsto \mathbf{G}_{s,w,[c],\langle\mathcal{G}\rangle}]\,.$$

4.4. Dynamical dictionary.

Then, there exists a "dynamical" dictionary on the generating operators that is similar to the analytic combinatorics dictionary, described for instance in the book of Flajolet and Sedgewick [12]. First, for disjoint factors $\mathcal{G}_1$ and $\mathcal{G}_2$ of $\mathcal{B}_{(d)}$, one has

$$\mathbf{G}_{s,w,[c],\langle\mathcal{G}_1+\mathcal{G}_2\rangle} = \mathbf{G}_{s,w,[c],\langle\mathcal{G}_1\rangle} + \mathbf{G}_{s,w,[c],\langle\mathcal{G}_2\rangle}\ .$$

If now $\mathcal{G}_1$ and $\mathcal{G}_2$ may be composed and satisfy $(\mathcal{G}_1\circ\mathcal{G}_2) \propto \mathcal{B}_{(d)}$, multiplicative properties of the denominator, together with additive properties of the cost, entail the equality

$$\mathbf{G}_{s,w,[c],\langle\mathcal{G}_1\circ\mathcal{G}_2\rangle} = \mathbf{G}_{s,w,[c],\langle\mathcal{G}_2\rangle}\circ\mathbf{G}_{s,w,[c],\langle\mathcal{G}_1\rangle}\ .$$

In particular, when $\mathcal{G}^\star \propto \mathcal{B}_{(d)}$, the generating operator of $\mathcal{G}^\star$ is the quasi-inverse

$$\mathbf{G}_{s,w,[c],\langle\mathcal{G}^\star\rangle} = \big(\mathbf{I} - \mathbf{G}_{s,w,[c],\langle\mathcal{G}\rangle}\big)^{-1}\ .$$

This occurs in our context for each strict phase of index ℓ, with the set $\mathcal{H}_{(d-\ell)} \propto \mathcal{B}_{(d)}$ of inverse branches of the Brun dynamical system $\mathcal{D}_{(d-\ell)} = (T_{(d-\ell)}, \mathcal{J}_{(d-\ell)})$, and the

generating operator of the strict phase of index ℓ is

$$\left(\mathbf{I} - \mathbf{G}_{s,w,[c],\langle\mathcal{H}_{(d-\ell)}\rangle}\right)^{-1} .$$

Moreover, the generating operator $\mathbf{G}_{s,w,[c],\langle\mathcal{H}_{(d-\ell)}\rangle}$ is closely related to the (usual) weighted transfer operator $\mathbf{H}_{s,w,[c],(d-\ell)}$ of the Brun dynamical system $\mathcal{D}_{(d-\ell)}$ defined by

$$\mathbf{H}_{s,w,[c],(d-\ell)}[f](\boldsymbol{x}) := \sum_{h\in\mathcal{H}_{(d-\ell)}} |J[h](\boldsymbol{x})|^s \, w^{c(h)} f\circ h(\boldsymbol{x}) \,, \tag{26}$$

and, with the scale change $s \mapsto s/(d-\ell+1)$ together with (24), the generating operator of the strict phase of index ℓ is finally

$$(I - \mathbf{H}_{s/(d-\ell+1),w,[c],(d-\ell)})^{-1} .$$

4.5. Generating operators used in the `BrunGcd`(d) *algorithm.*

With Proposition 2, this yields the following characterization of the generating operator of the set $\mathcal{B}_{(d)}$ used by the `BrunGcd`(d) algorithm.

Proposition 3. Let c be an additive cost on the set $\mathcal{B}_{(d)}$.

(*a*) The generating operator $\mathbf{B}_{s,w,[c],(d)}$ of the set $\mathcal{B}_{(d)}$ used by the `BrunGcd`(d) algorithm decomposes as

$$\mathbf{B}_{s,w,[c],(d)} = \mathbf{P}_{s,w,[c],(1)} \circ \mathbf{P}_{s,w,[c],(2)} \circ \cdots \circ \mathbf{P}_{s,w,[c],(d)} \,,$$

and involves the generating operator $\mathbf{P}_{s,w,[c],(k)}$ of the set $\mathcal{P}_{(k)}$ used by the `BrunGcd`(d) algorithm during its phase of index $d-k$, for $1 \le k \le d$. The following recurrence relation holds:

$$\mathbf{B}_{s,w,[c],(d)} = \mathbf{B}_{s,w,[c],(d-1)} \circ \mathbf{P}_{s,w,[c],(d)}.$$

(*b*) For any $k \in [1..d]$, the operator $\mathbf{P}_{s,w,[c],(k)}$, which is relative to the phase of index $d-k$, involves the quasi-inverse of the generating operator $\mathbf{G}_{s,w,[c],(k)}$ of the set $\mathcal{H}_{(k)}$ and the generating operator $\mathbf{F}_{s,w,[c],(k)}$ associated with the final set $\mathcal{F}_{(k)}$ of the phase of index $d-k$, and is equal to

$$\mathbf{P}_{s,w,[c],(k)} = \mathbf{F}_{s,w,[c],(k)} \circ (I - \mathbf{G}_{s,w,[c],(k)})^{-1} .$$

(*c*) For any $k \in [1..d]$, the generating operator $\mathbf{G}_{s,w,[c],(k)}$ of the set $\mathcal{H}_{(k)}$ is closely related to the (weighted) transfer operator $\mathbf{H}_{s,w,[c],(k)}$ of the dynamical system $\mathcal{D}_{(k)}$ defined in (26) via the change of scale $t_k : s \mapsto s/(k+1)$, namely

$$\mathbf{G}_{s,w,[c],(k)} = \mathbf{H}_{t_k(s),w,[c],(k)} = \mathbf{H}_{s/(k+1),w,[c],(k)} \,. \tag{27}$$

4.6. Relating generating functions to generating operators.

The following result is one of the key ingredients of the paper. It is typical in dynamical analysis as it relates generating functions and generating operators.

Theorem 3. Consider the `BrunGcd`(d) algorithm acting on the set $\Omega_{(d)}$, together with an additive cost C, related to some elementary cost c. Then, the three following relations hold between

(*i*) the three generating functions $S_{(d)}, \widehat{S}_{(d,C)}, S_{(d,C)}$, defined in (17),(18),(21),

(*ii*) the three analogous generating operators of the set $\mathcal{B}_{(d)}$, i.e., the plain, the cumulative and the weighted one, defined in (25) and in Proposition 3, and denoted respectively as $\mathbf{B}_{s,(d)}, \widehat{\mathbf{B}}_{s,[c],(d)}, \mathbf{B}_{s,w,[c],(d)}$:

$$S_{(d)}(s) = \zeta(s) \cdot \mathbf{B}_{s,(d)}[1](0)\,,$$
$$\widehat{S}_{(d,C)}(s) = \zeta(s) \cdot \widehat{\mathbf{B}}_{s,[c],(d)}[1](0)\,,$$
$$S_{(d,C)}(s,w) = \zeta(s) \cdot \mathbf{B}_{s,w,[c],(d)}[1](0)\,.$$

Proof. The proof is indeed quite short. Let $h \in \mathcal{B}_{(d)}$. The equality

$$\pi(\boldsymbol{u}) = \left(\frac{u_1}{u_0}, \frac{u_2}{u_0}, \ldots, \frac{u_d}{u_0}\right) = h(0)$$

together with $\gcd(\boldsymbol{u}) = 1$ proves that the denominator of the LFT h satisfies $|D[h](0)| = u_0$. Moreover, as C is an additive cost associated with cost c, the equality $C(\boldsymbol{u}) = c(h)$ holds. Now, the bijection between $\underline{\Omega}_{(d)}$ and the set $\mathcal{B}_{(d)}$, together with Eq. (23), entail the relations

$$\begin{aligned} S_{(d,C)}(s,w) = \zeta(s) \cdot \underline{S}_{(d,C)}(s) &= \zeta(s) \cdot \sum_{\boldsymbol{u} \in \underline{\Omega}_{(d)}} \frac{w^{C(\boldsymbol{u})}}{u_0^s} \\ &= \zeta(s) \cdot \sum_{h \in \mathcal{B}_{(d)}} w^{c(h)}\, |D[h](0)|^{-s} = \zeta(s) \cdot \mathbf{B}_{s,w,[c],(d)}[1](0)\,. \end{aligned}$$

□

4.7. Derivatives of quasi-inverses.

When the cost C is a cost concentrated on the strict phase of index d–k, the functional Δ defined in (22) is applied to the quasi-inverse $(I - \mathbf{G}_{s,w,[c],(k)})^{-1}$ and produces what we call a "double quasi-inverse", formed with two (plain) quasi-inverses and a middle cumulative operator associated with the cost c. Now, via (27), we return to the usual weighted transfer operator $\mathbf{H}_{t,w,[c],(k)}$, defined in (26), with the scale-change $t = s/(k{+}1)$, and obtain

$$\Delta[w \mapsto (I - \mathbf{G}_{s,w,[c],(k)})^{-1}] = (I - \mathbf{H}_{t,(k)})^{-1} \circ\ \widehat{\mathbf{H}}_{t,[c],(k)} \circ\ (I - \mathbf{H}_{t,(k)})^{-1}\,, \tag{28}$$

which involves the (usual) plain operator and its cumulative version, i.e.,

$$\mathbf{H}_{t,(k)}[f] := \sum_{h \in \mathcal{H}_{(k)}} |J[h]|^t \cdot f \circ h, \qquad \widehat{\mathbf{H}}_{t,[c],(k)}[f] := \sum_{h \in \mathcal{H}_{(k)}} c(h) \cdot |J[h]|^t \cdot f \circ h\,. \tag{29}$$

We now focus on the quasi-inverses of the transfer operators (simple or double) which will play a central role in the analysis, and define

$$\mathbf{Q}_{t,(k)} := (I - \mathbf{H}_{t,(k)})^{-1}, \qquad \mathbf{Q}^{[2]}_{t,[c],(k)} := (I - \mathbf{H}_{t,(k)})^{-1} \circ\ \widehat{\mathbf{H}}_{t,[c],(k)} \circ\ (I - \mathbf{H}_{t,(k)})^{-1}\,. \tag{30}$$

According to Proposition 3(a) and (b), the plain Dirichlet series $S_{(d)}(s)$, associated with the plain operator $\mathbf{B}_{s,(d)}$, involves one quasi-inverse for each strict phase, whereas the cumulative Dirichlet series relative to a cost C concentrated on the strict phase of index $d{-}k$ involves a double quasi-inverse for the phase of index $d{-}k$, and only a plain quasi-inverse for each other phase. This is now precisely stated:

Proposition 4. Associate with $k \in [1..d]$ the scale change $t_k : s \mapsto s/(k+1)$. Then the following holds:

(*a*) The plain Dirichlet series $S_{(d)}(s)$ involves a unique quasi-inverse $\mathbf{Q}_{t_k(s),(k)} = (I - \mathbf{H}_{t,(k)})^{-1}$ for each phase of index $d-k$ with $k \in [1..d]$.

(*b*) If C is a cost concentrated on the strict phase of index $d-k$ (with $k \in [1..d]$), the cumulative series $\widehat{S}_{(d,C)}(s)$ of cost C involves one (plain) quasi-inverse $\mathbf{Q}_{t_j(s),(j)}$ for each phase of index $d-j$ with $j \neq k$ and a "double" quasi-inverse $\mathbf{Q}^{[2]}_{t_k(s),[c],(k)}$ for the phase of index $d-k$.

5. Study of quasi-inverses.

This ends the first part of our analysis, of a combinatorial nature. We begin the second part of our analysis, of an analytic nature. We then have now to study the Dirichlet series, from an analytic point of view, and discover their singularities. With Proposition 4, their singularities are brought by quasi-inverses. This is why this section is devoted to the study of the transfer operator

$$\mathbf{H}_{t,(k)} := \sum_{h \in \mathcal{H}_{(k)}} |J[h]|^t \, f \circ h = \sum_{h \in \mathcal{H}_{(k)}} |D[h]|^{-t(k+1)} \; f \circ h$$

and its quasi-inverse $(I - \mathbf{H}_{t,(k)})^{-1}$. We will apply this study to the Dirichlet series in the following Section 6.

5.1. Functional space.

For $k = 1$, the space $C^1([0,1])$ of continuously differentiable functions on the unit interval proves to be "well-adapted" to the analysis of the transfer operator of the Euclid system, as the "upper" part of its spectrum is formed with isolated eigenvalues. This fact was proven and used for instance in [2, 3, 7]. Here, we deal with the space $C^1(\mathcal{J}_{(k)})$ which gathers the continuously differentiable functions on the simplex $\mathcal{J}_{(k)}$. For such a function f, we denote by $\boldsymbol{D}f(\boldsymbol{x})$ the differential of the function f at $\boldsymbol{x}$. We consider any fixed norm on the space $\mathbb{R}^k$, denoted as $||\cdot||_{(k)}$, and we define the two following norms on the space $C^1(\mathcal{J}_{(k)})$, namely the norms $||\cdot||_0$ and $||\cdot||_1$ defined as

$$||f||_0 := \sup_{\boldsymbol{x} \in \mathcal{J}_{(k)}} |f(\boldsymbol{x})|, \qquad ||f||_1 = ||f||_0 + \sup_{\boldsymbol{x} \in \mathcal{J}_{(k)}} ||\boldsymbol{D}f(\boldsymbol{x})||_{(k)} \, .$$

Then, the operator $\mathbf{H}_{t,(k)}$ is well-defined for $\Re t > 1/(k+1)$ as Eq. (24) relates its norm $||\mathbf{H}_{t,(k)}]]_0$ with the Riemann Zeta function $\zeta(t/(k+1))$. It acts on the Banach space $\big(C^1(\mathcal{J}_{(k)}), \|\cdot\|_1\big)$, it is bounded, and the map $t \mapsto \mathbf{H}_{t,(k)}$ is analytic.

We now describe the main properties of the operator $\mathbf{H}_{t,(k)}$, and will show how they generalize the properties of the one-dimensional case $d = 1$.

5.2. Contraction and distorsion properties

The *contraction ratio* δ_k of the dynamical system $\mathcal{D}_{(k)}$, defined as

$$\delta_k := \limsup_{n \to \infty} \sup_{\substack{h \in \mathcal{H}^n_{(k)} \\ \boldsymbol{x} \in \mathcal{J}_{(k)}}} ||\boldsymbol{D}h(\boldsymbol{x})||^{1/n}_{(k)} \, , \tag{31}$$

involves the differential $\boldsymbol{D}h(\boldsymbol{x})$ and its norm derived from the norm $||\cdot||_{(k)}$. This contraction ratio is considered in [8] where the strict inequality $\delta_k < 1$ is proved: the Brun dynamical system $\mathcal{D}_{(k)}$ is thus *contracting*. Moreover, as the inverse branches are linear fractional transformations with positive coefficients, the *distortion* property holds: there exists $L_k > 0$ which relates the differential of the Jacobian to the Jacobian itself,

$$||\boldsymbol{D}J[h](\boldsymbol{x})||_{(k)} \leq L_k\,|J[h](\boldsymbol{x})|, \qquad \forall h \in \mathcal{H}^{\star}_{(k)},\, \forall \boldsymbol{x} \in \mathcal{J}_{(k)}\,.$$

These properties entail an inequality of Lasota-Yorke type for the operator $\mathbf{H}_{t,(k)}$, namely,

$$||\mathbf{H}^n_{t,(k)}[f]||_1 \leq K \cdot \big(|t| \cdot \underline{\delta}_k^n \cdot ||f||_1 + ||f||_0\big) \cdot ||\mathbf{H}^n{}_{\Re t,(k)}||_0 \tag{32}$$

for some constant $K > 0$ (related to the distortion constant), and a real $\underline{\delta}_k \in]\delta_k, 1[$. With theorems due to Hennion [16], and Ionescu Tulcea and Marinescu [17], such a Lasota-Yorke inequality entails the *quasi-compacity* of the operator $\mathbf{H}_{t,(k)}$ for real values of parameter t. We recall this notion in the next section (see also the book [2]).

5.3. Quasi-compacity

The spectrum of an operator $\mathbf{G}$ is the set of the complex numbers λ for which $\lambda I - \mathbf{G}$ is not invertible. The spectral radius $R(\mathbf{G})$ of an operator $\mathbf{G}$ is defined as

$$R(\mathbf{G}) = \sup\{|\lambda|, |\ \lambda \in \mathtt{Sp}(\mathbf{G})\}\,.$$

An operator is said *quasi-compact* if the "upper" part of its spectrum is formed with isolated eigenvalues of finite multiplicity: there exists a radius $r < R(\mathbf{G})$ for which any $\lambda \in \mathtt{Sp}(\mathbf{G})$ with $|\lambda| \geq r$ is an isolated eigenvalue with finite multiplicity. Inequality (32) proves the existence of such a radius r for any real t: the inequality $r \leq \underline{\delta}_k R(\mathbf{H}_{t,(k)})$ holds and entails the quasi-compacity for the operator $\mathbf{H}_{t,(k)}$.

Proposition 5. The following holds:

(*i*) For a real $t > 1/(k+1)$, the transfer operator $\mathbf{H}_{t,(k)}$ is quasi-compact. This property extends when t belongs to a complex neighborhood of the real line $t > 1/(k+1)$.

(*ii*) Consider the algebraic number τ_k associated with the worst-case behaviour of the `BrunGcd` Algorithm. The equality holds

$$\tau_k^{k+1} = \lim_{n\to\infty} \sup_{h\in\mathcal{H}^n_{(k)}} |J[h](\boldsymbol{x})|^{1/n} \qquad \text{for any } \boldsymbol{x} \in \mathcal{J}_{(k)}\,. \tag{33}$$

(*iii*) The spectral radius $R(\mathbf{H}_{t,(k)})$ is strictly decreasing as a function of t, and the inequality holds

$$R(\mathbf{H}_{t+u,(k)}) \leq \tau_k^{(k+1)u}\, R(\mathbf{H}_{t,(k)}) \qquad \text{for } u \geq 0,\ t > 1/(k+1)\,. \tag{34}$$

(*iv*) An *aperiodicity* property holds: For any complex number t, with $\Re t > 1/(k+1)$, the spectral radius satisfies the strict inequality

$$R(\mathbf{H}_{t,(k)}) < R(\mathbf{H}_{\Re t,(k)}) \qquad \text{for any non-real number } t\,.$$

Remark. There is a relation between the two constants δ_k and τ_k: the contraction ratio δ_k is associated with the dominant eigenvalue of a linear mapping (namely, the differential $\boldsymbol{D}h(\boldsymbol{x})$) whereas the constant τ_k^{k+1} is associated with the absolute value of its determinant. Then the relation $\tau_k^{k+1} \leq \delta_k^k$ holds.

Proof. (i): It uses analytic perturbation (see the book [19]).
(ii): As we saw in (5), the algebraic number τ_k associated with the worst-case bound is defined as

$$\tau_k^{-1} = \lim_{n\to\infty} \inf_{h\in\mathcal{H}^n_{(k)}} |D[h](\mathbf{0})|^{1/n}$$

and Eq.(24), together with the distortion property, lead to the result.
(iii): We begin with the inequality that uses the estimates given in (5) and (6)

$$\mathbf{H}^n_{t+u,(k)}[1](\mathbf{0}) = \sum_{h\in\mathcal{H}^n_{(k)}} |J[h]|^{t+u}(\mathbf{0})$$

$$\leq \Big(\sup_{h\in\mathcal{H}^n_{(k)}} |J[h]|^u(\mathbf{0})\Big)\cdot\Big(\sum_{h\in\mathcal{H}^n_{(k)}} |J[h]|^t(\mathbf{0})\Big) \leq \chi_k^{-(k+1)u}\cdot\tau_k^{n(k+1)u}\cdot\mathbf{H}^n_{t,(k)}[1](\mathbf{0})\,,$$

and we use the Radius Spectral Theorem.
The proof of (iv) is of a different flavour. It mainly uses the fact that inverse branches are linear fractional transformations. It is similar to the proof in case $k=1$ (the Euclid algorithm) which is given in [30]. □

5.4. Spectral decomposition

For real t, the operator $\mathbf{H}_{t,(k)}$ is quasi-compact. Moreover, the dynamical system is *ergodic*, as proven e.g. in [1]. This entails the following: the transfer operator $\mathbf{H}_{t,(k)}$ admits a unique simple dominant eigenvalue $\lambda_{(k)}(t)$, separated from the remainder of the spectrum by a spectral gap. Moreover, there exists a spectral decomposition

$$\mathbf{H}_{t,(k)} = \lambda_{(k)}(t)\mathbf{A}_{t,(k)} + \mathbf{K}_{t,(k)}\,,$$

that involves the dominant eigenvalue $\lambda_{(k)}(t)$, the projection $\mathbf{A}_{t,(k)}$ on the dominant eigenspace, and a "remainder" operator $\mathbf{K}_{s,(k)}$ whose spectral radius is strictly smaller than $\lambda_{(k)}(t)$. Then, the relation $\mathbf{A}_{t,(k)}\circ\mathbf{K}_{t,(k)} = \mathbf{K}_{t,(k)}\circ\mathbf{A}_{t,(k)} = 0$ leads to the spectral decomposition for the quasi-inverse

$$\mathbf{Q}_{t,(k)} = (I-\mathbf{H}_{t,(k)})^{-1} = \frac{\lambda_{(k)}(t)}{1-\lambda_{(k)}(t)}\mathbf{A}_{t,(k)} + (\mathbf{I}-\mathbf{K}_{t,(k)})^{-1}\,, \tag{35}$$

where the norm of the "remainder quasi-inverse" $(\mathbf{I}-\mathbf{K}_{t,(k)})^{-1}$ satisfies

$$||(\mathbf{I}-\mathbf{K}_{t,(k)})^{-1}||_1 \leq \frac{1}{1-\lambda_{(k)}(t)}\,. \tag{36}$$

5.5. Dominant spectral objects at $t=1$.

For $t=1$, the transfer operator is the density transformer, and 1 is an eigenvalue of maximum modulus, the unique simple dominant eigenvalue of maximum modulus, moreover isolated from the remainder of the spectrum by a spectral gap. When $t=1$, the dominant eigenfunction is explicitly exhibited in [1], and involves the set $\mathfrak{S}_k$ of permutations on $[1..k]$,

$$\psi_k(\boldsymbol{x}) = \sum_{\sigma\in\mathfrak{S}_k}\prod_{i=1}^{k}\frac{1}{1+x_{\sigma(1)}+x_{\sigma(2)}+\ldots+x_{\sigma(i)}}\,. \tag{37}$$

However, as an explicit expression for the integral

$$\kappa_k := \int_{\mathcal{J}_{(d)}} \psi_k(\boldsymbol{x})\, d\boldsymbol{x} \tag{38}$$

is not known, except for small values of k, the density Ψ_k associated with this eigenfunction is itself not completely explicitly known. At $t = 1$, the projection $\mathbf{A}_{1,(k)}$ involves the dominant eigenvalue ψ_k defined in (37), the integral I_k on the simplex $\mathcal{J}_{(k)}$ and the constant $\kappa_k = I_k[\psi_k]$. Indeed the projection $\mathbf{A}_{1,(k)}$ is defined as

$$\mathbf{A}_{1,(k)}[f](\boldsymbol{x}) = I_k[f] \cdot \Psi_k(\boldsymbol{x})\,, \tag{39}$$

$$\text{where} \qquad I_k[f] := \int_{\mathcal{J}_{(k)}} f(\boldsymbol{u}) d\boldsymbol{u}, \qquad \kappa_k = I_k[\psi_k], \qquad \Psi_k(\boldsymbol{x}) := \frac{\psi_k(x)}{\kappa_k}\,.$$

Remark that Eq. (34) relates the dominant eigenvalue $\lambda_{(k)}$ and the algebraic number τ_k which intervenes in the worst case, via the inequality

$$\lambda_{(k)}(1+u) \le \tau_k^{(k+1)u} \qquad \text{for any real } u \ge 0. \tag{40}$$

5.6. Quasi-inverses.

With perturbation arguments [19], there exists a complex neighborhood of the real axis for which the decomposition holds. With the equality $\lambda_{(k)}(1) = 1$, this proves that the quasi-inverse has a pole at $t = 1$, with a residue which involves in particular the entropy $\mathcal{E}_k = -\lambda'_{(k)}(1)$ and the projection $\mathbf{A}_{t,(k)}$, whose expression was given in (39). Moreover, with Assertions (ii) and (iii) of Proposition 5, the map $t \mapsto \mathbf{Q}_{t,(k)}$ is analytic on the punctured half-plane $\{\Re t > 1,\ t \neq 1\}$.

Using now for each phase of index $d-k$ the scale change $t_k(s) = s/(k+1)$, we describe the behaviour of the quasi inverse $s \mapsto \mathbf{Q}_{t_k(s),(k)}$.

Proposition 6. Consider the phase of index $d-k$ with $k \in [1..d]$ and the scale change $s \mapsto s/(k+1)$. Then, the quasi-inverse $s \mapsto \mathbf{Q}_{t_k(s),(k)}$ relative to the phase of index $d-k$ is analytic on the punctured half-plane $\{\Re s \ge k+1, s \neq k+1\}$ and admits a simple pole at $s = k+1$, with a residue $\mathbf{T}_{(k)}$ defined as

$$\mathbf{T}_{(k)}[g](\boldsymbol{x}) = \frac{k+1}{\mathcal{E}_k} \cdot I_k[g] \cdot \Psi_k(x) \tag{41}$$

which involves the entropy $\mathcal{E}_k$, the integral I_k on the simplex $\mathcal{J}_{(k)}$ and the invariant density Ψ_k defined in (39).

5.7. Double quasi-inverses.

As we saw in Section 4.7, the analysis of the cumulative Dirichlet series $\widehat{S}_C(s)$ associated with an elementary cost c that is concentrated on the phase of index $d-k$ deals with the "double quasi-inverse"

$$\mathbf{Q}^{[2]}_{t,[c],(k)} = (I - \mathbf{H}_{t,(k)})^{-1} \circ \widehat{\mathbf{H}}_{t,[c],(k)} \circ (I - \mathbf{H}_{t,(k)})^{-1} \quad t = t_k(s) := s/(k+1)\,. \tag{42}$$

Each quasi-inverse "brings" a simple pole at $t = 1$ (i.e., $s = k+1$), and the double quasi-inverse is singular at $s = k+1$, with at least a pole of order 2.

We first focus on the middle operator. After an easy change of variables, we remark that, for any density Ψ, the integral

$$\int_{\mathcal{J}_{(k)}} \widehat{\mathbf{H}}_{1,[c],(k)}[\Psi](\boldsymbol{x})d\boldsymbol{x} = \sum_{h\in\mathcal{H}_{(k)}} c(h) \int_{h(\mathcal{J}_{(k)})} \Psi(\boldsymbol{y})\, d\boldsymbol{y}$$

coincides with the mean value of the elementary cost c with respect to the measure ν of density Ψ. We adopt the notation $\underline{\mathbb{E}}_{(k)}[\cdot]$ for the mean value, when it is relative to the invariant measure ν_k of density Ψ_k. Then the equality holds

$$\underline{\mathbb{E}}_{(k)}[c] = \int_{\mathcal{J}_{(k)}} \widehat{\mathbf{H}}_{1,[c],(k)}[\Psi_k](\boldsymbol{x})d\boldsymbol{x} = I_k\left[\widehat{\mathbf{H}}_{1,[c],(k)}\left[\Psi_k\right]\right] .$$

We now assume that the mean value $\underline{\mathbb{E}}_{(k)}[c]$ is finite and return to the study of the double quasi-inverse which admits in this case a pole of order 2 at $s = k+1$; moreover, the "dominant" constant is obtained via the spectral decomposition of each quasi-inverse at $s = k+1$, which involves the projection given in (39), and

$$\frac{(k+1)^2}{|\lambda'_{(k)}(1)|^2}\, \mathbf{A}_{1,(k)} \circ \widehat{\mathbf{H}}_{1,[c],(k)} \circ \mathbf{A}_{1,(k)}[g](\boldsymbol{x})$$

$$= \left(\frac{k+1}{\mathcal{E}_k}\right)^2 I_k[g] \cdot I_k\left[\widehat{\mathbf{H}}_{1,[c],(k)}\left[\Psi_k\right]\right] \cdot \Psi_k(\boldsymbol{x}) .$$

We have thus shown:

Proposition 7. Consider a cost C concentrated on the phase of index $d-k$ with $k \in [1..d]$, and associated with an elementary cost c whose mean value $\underline{\mathbb{E}}_{(k)}[c]$ with respect to the measure ν_k is finite. Consider the scale change $t = s/(k+1)$. Then, the double quasi-inverse relative to the phase of index $d-k$ and defined in (42) is analytic on the punctured half-plane $\{\Re s \geq k+1,\ s \neq k+1\}$ and admits a pole of order 2 at $s = k+1$, with a dominant operator $\mathbf{U}_{[c],(k)}$ defined as

$$\mathbf{U}_{[c],(k)}[g](\boldsymbol{x}) = \left(\frac{k+1}{\mathcal{E}_k}\right)^2 \cdot I_k[g] \cdot \underline{\mathbb{E}}_{(k)}[c] \cdot \Psi_k(\boldsymbol{x}) , \tag{43}$$

which involves the entropy $\mathcal{E}_k$, the integral I_k on the simplex $\mathcal{J}_{(k)}$, the invariant density Ψ_k and the mean value $\underline{\mathbb{E}}_{(k)}[c]$.

Remark. The mean value $\underline{\mathbb{E}}_{(k)}[c]$ of cost c occurs in the study of weighted real trajectories by applying Birkhoff's ergodic theorem. The occurrence of the same mean value $\underline{\mathbb{E}}_{(k)}[c]$ in the dominant operator (43), and thus in the study of weighted rational trajectories, is one of the key features in dynamical analysis. It leads to the following statement : "Rational trajectories behave in the average-case exactly as the truncated real trajectories behave generically".

5.8. Mean values of costs associated with digits.

Consider the simplex $\mathcal{J}_{(d)}$ and the variable m defined on $\mathcal{J}_{(d)}$ which associates with $\boldsymbol{x} \in \mathcal{J}_{(d)}$ the quotient $m(\boldsymbol{x})$ computed by the `BrunGcd` algorithm on the input $\boldsymbol{x}$. For $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$, the quotient $m(\boldsymbol{x})$ is equal to $\lfloor 1/x_1 \rfloor$, and the event $[m(\boldsymbol{x}) \geq m]$

coincides with the event $[x_1 \leq 1/m]$. As $\boldsymbol{x} \in \mathcal{J}_{(d)}$, all the components x_i satisfy $x_i \leq (1/m)$ too, and the equality between the two subsets $\{\boldsymbol{x} \mid m(\boldsymbol{x}) \geq m\}$ and $= (1/m)\,\mathcal{J}_{(d)}$ holds. Then, the study of the variable m leads to consider homothetic images $y\mathcal{J}_{(d)}$ of the simplex $\mathcal{J}_{(d)}$. Moreover, the previous section highlights the role that the measure ν_d associated with the invariant density Ψ_d plays in our analyses of the `BrunGcd` algorithm. This is why the function μ_d which associates with $y \in [0,1]$ the measure $\nu_d\left(y\mathcal{J}_{(d)}\right)$ will play a central role in the distribution of quotients, in particular in Section 7 below. The next result will be important in our analysis.

Proposition 8. The following equalities hold for $d \geq 1$

$$\mathbb{P}_{(d)}[m = 1] = \mu_d(1/2), \quad \mathbb{E}_{(d)}[m] := \sum_{m \geq 1} \mu_d\,(1/m)\ . \tag{44}$$

For $d \geq 2$, the mean value $\mathbb{E}_{(d)}[m]$ is finite, and it is infinite for $d = 1$.

Proof. The first assertion is clear. The second assertion is due to the fact that $\mu_d(y)$ is $\Theta(y^d)$ for $y \to 0$ and any d (with hidden constants in the Θ which depend on d). Then the sum of the series which appears in the expression of $\mathbb{E}_{(d)}[m]$ is finite for $d \geq 2$ and infinite for $d = 1$ (the case of the Euclid algorithm). □

6. Dynamical analysis (II).

We now return to the number of steps that the `BrunGcd` algorithm performs during its various phases, and to the proof of Theorem 1.

6.1. Analytic properties of Dirichlet series.

With Theorem 3 and Proposition 4 which relate generating functions and generating operators, Eq. (16) which recalls the additive costs of interest, Propositions 6 and 7 which describe the behaviour of quasi-inverses and double quasi-inverses near their singularities, we describe the analytic properties of the Dirichlet series of interest.

Proposition 9. Consider the `BrunGcd` algorithm when acting on $\Omega_{(d)}$, for a dimension $d \geq 2$, together with the following costs:

- the total number L_d of steps,
- the number M_d of steps in the first phase (of index $\ell = 0$),
- the number O_d of quotients equal to 1 during the first phase,
- the number Σ_d of subtractive steps during the first phase,
- the number $R_{d,\ell}$ of steps during the phase of index $\ell > 0$.

with the six Dirichlet generating functions of interest, namely, the plain series and the five cumulative series relative with the number of steps $L_d, M_d, O_d, \Sigma_d, R_{d,\ell}$. The following holds:

(*a*) The six series are analytic in the punctured half-plane $\{\Re s \geq d+1, s \neq d+1\}$.

(*b*) The plain series and the cumulative series relative to $R_{d,\ell}$ with $\ell > 0$ admit a simple pole at $s = d+1$. Their residues are denoted as T_d and $\rho_{d,\ell} \cdot T_d$.

(c) The cumulative series relative to L_d, M_d, O_d, Σ_d admit a pole of order 2 at $s = d+1$, and their dominant constants $\mathtt{Dom}$ are expressed with the residue T_d and three other constants a_d, θ_d, σ_d. One has:

$$\mathtt{Dom}[L_d] = \mathtt{Dom}[M_d] = a_d\, T_d\,, \qquad \mathtt{Dom}[O_d] = a_d\, \theta_d\, T_d, \quad \mathtt{Dom}[\Sigma_d] = a_d\, \sigma_d\, T_d\,.$$

(d) The constants a_d, θ_d and σ_d involve the entropy $\mathcal{E}_d$, the function $\mu_d(y)$, defined in Section 5.8, and the expectations of Proposition 8,

$$a_d = \frac{d+1}{\mathcal{E}_d}, \quad \theta_d = \mathbb{P}_{(d)}[m=1] = 1 - \mu_d\,(1/2)\,, \quad \sigma_d = \mathbb{E}_{(d)}[m] = \sum_{m\geq 1} \mu_d\,(1/m)\,. \tag{45}$$

Proof. The generating operator associated with the complete execution of the `BrunGcd` algorithm is the operator $\mathbf{B}_{s,(d)}$ described in Theorem 3. We consider now two parts of the execution: the strict first phase (namely, the first phase, except its last step) and the remaining of the execution. The generating operator relative to the strict first phase is the quasi-inverse $\mathbf{Q}_{t_d(s),(d)} = (I - \mathbf{H}_{s/(d+1),(k)})^{-1}$ whereas the operator relative to the remaining of the execution is $\mathbf{D}_{s,(d)} := \mathbf{B}_{s,(d-1)} \circ \mathbf{F}_{s,(d)}$. This remaining operator contains one quasi-inverse for each phase. For $k \in [1..d-1]$, the quasi-inverse of the phase of index $d-k$ is $\mathbf{Q}_{t_k(s),(k)}$. Note also, that, according to Section 4.1, it is sufficient to consider costs concentrated either on the first phase, or on another phase.

Plain series. With Proposition 6, the operator $\mathbf{Q}_{t_d(s),(d)}$ of the strict first phase has a dominant pole at $s = d+1$, whereas the remaining operator $\mathbf{D}_{s,(d)}$ remains regular at $s = d+1$. Then the "total" operator $\mathbf{B}_{s,(d)}$ has a pole with a residue equal to $\mathbf{D}_{d+1,(d)} \circ \mathbf{T}_{(d)}$ which involves the operator $\mathbf{T}_{(d)}$ described in (41). According to Theorem 3, the generating function $S_{(d)}(s)$ has thus a simple pole at $s = d+1$ with a residue

$$T_d = \zeta(d+1) \cdot \mathbf{D}_{d+1,(d)} \circ \mathbf{T}_{(d)}[1](0) = b_d \cdot \left(\frac{d+1}{\mathcal{E}_d}\right) \quad \text{with} \quad b_d = \zeta(d+1)\mathbf{D}_{d+1,(d)}[\Psi_d](0)\,.$$

Cumulative generating function relative to a cost concentrated on the first phase. Consider a cost C concentrated on the strict first phase, associated with an elementary cost c. Then the cumulative generating operator relative to this cost C is

$$\mathbf{D}_{s,(d)} \circ \mathbf{Q}^{[2]}_{t_d(s),[c],(d)}\,.$$

With Proposition 7, it has a pole of order 2 at $s = d+1$, with a dominant operator equal to $\mathbf{D}_{d+1,(d)} \circ \mathbf{U}_{[c],(d)}$ which involves the operator $\mathbf{U}_{[c],(d)}$ described in (43). With Theorem 3, the cumulative generating series $S_{(d,C)}(s)$ has a pole at $s = d+1$ of order 2 with a dominant constant

$$\mathtt{Dom}[C] = \zeta(d+1) \cdot \mathbf{D}_{d+1,(d)} \circ \mathbf{U}_{[c],d}[1](0) = b_d \cdot \left(\frac{d+1}{\mathcal{E}_d}\right)^2 \cdot \underline{\mathbb{E}}_{(d)}[c]\,.$$

The mean values $\underline{\mathbb{E}}_{(d)}[c]$ of costs c associated with M_d, O_d and Σ_d are expressed in terms of μ_d in Proposition 8 and proven there to be finite.

Cumulative generating function relative to a cost concentrated on another phase. Consider now a cost C concentrated on a remaining phase (not the first one) of index $\ell = d-k$

with $k < d$. Then, the cumulative generating operator relative to this cost is almost the same as the plain generating operator: there is only a double quasi-inverse $\mathbf{Q}^{[2]}_{t_k(s),[c],(k)}$ that replaces the (simple) quasi-inverse $\mathbf{Q}_{t_k(s),(k)}$. This incursion of a new quasi-inverse does *not* change the nature of the dominant singularity, which remains a simple pole at $s = d+1$. However, the residue of the generating function changes and we now denote it by $T_d\, \rho_{d,\ell}$. Note that the insertion of the quasi-inverse "in the middle" creates difficulties for the asymptotic estimates of the constant $\rho_{d,\ell}$. We will return to this question in Section 7.4 $\square$

6.2. Extraction of coefficients and final step in the proof of Theorem 1.

It remains to relate the asymptotics of the coefficients of the Dirichlet series and their dominant singularities, and this is done via the Delange Tauberian Theorem.

Theorem 4 (Delange)**.** For $\sigma > 0$, consider a Dirichlet series $S(s)$ with non-negative coefficients which converges for $\Re s > \sigma$. Assume that the following holds:

(*i*) $S(s)$ is analytic on $\Re s = \sigma, s \neq \sigma$;

(*ii*) near σ, $S(s)$ satisfies $S(s) = A(s)/(s-\sigma)^{\gamma+1}$ where A is analytic in σ, $A(\sigma) \neq 0$ and $\gamma \geq 0$.

Then as $N \to \infty$, the following asymptotics holds

$$\Phi_N[S] \sim \frac{A(\sigma)}{\sigma\Gamma(\gamma+1)} N^\sigma (\log N)^\gamma, \qquad \text{for } \Phi_N \text{ defined in (20).}$$

With Proposition 9, all the Dirichlet series of interest satisfy the hypotheses of the Delange Tauberian Theorem with $\sigma = d+1$, the series of (b) with $\gamma = 0$ and the series of (c) with $\gamma = 1$. With Eq. (19), the mean value of any cost C of interest equals a ratio whose numerator and denominator can be estimated with the Delange Tauberian Theorem. This concludes the proof of Theorem 1, and the main constants involved are constants $\rho_{d,\ell}, a_d, \theta_d, \sigma_d$.

We do not know much about the asymptotics of $\rho_{d,\ell}$ (we return to this point in Section 7.4), but we focus in the next section on the study of the other constants, namely a_d (closely related to the entropy $\mathcal{E}_d$), θ_d, σ_d, and their asymptotics for $d \to \infty$.

7. Behaviour of the Brun dynamical system in high dimensions.

We now study the dependency of the main constants of the analysis κ_d, $\mathcal{E}_d$, θ_d, σ_d with respect to the dimension d and prove Theorem 2. Even if these constants occur in the analysis of the `BrunGcd` algorithm, they are only defined via the Brun dynamical system in d dimensions. The two main constants, namely κ_d and $\mathcal{E}_d$, are considered in 7.1 with a proof of Theorem 2(a). Asymptotic estimates for the function μ_d with a proof of Theorem 2(b) are handled in Section 7.2, whereas Section 7.3 is devoted to the proof of Theorem 2(c). Lastly, in Section 7.4, we state a conjecture on the constants $\rho_{d,\ell}$ that is based on some natural heuristics and validated by some preliminary experiments.

7.1. The main constants of the analysis: integral expressions for κ_d and $\mathcal{E}_d$.

We already explained in Section 5.8 the role played by the simplex $y\mathcal{J}_{(d)}$, and its measure $\mu_d(y) := \nu_d\left((y\mathcal{J}_{(d)}\right)$ with respect to the measure ν_d associated with the invariant density Ψ_d. We then introduce the function

$$\kappa_d(y) := \int_{y\mathcal{J}_{(d)}} \psi_d(\boldsymbol{x}) d\boldsymbol{x} \tag{46}$$

which involves the eigenfunction ψ_d (whose expression is given in (37)). This integral extends the integral κ_d defined in (38), in the sense that $\kappa_d = \kappa_d(1)$. As the equality $\mu_d(y) = \kappa_d(y)/\kappa_d$ holds, the study of $y \mapsto \kappa_d(y)$ is the main step in the study of the function μ_d.

We first provide one-dimensional integral expressions for the distribution constant $\kappa_d(y)$ and the entropy $\mathcal{E}_d$. The expression for $\kappa_d(y)$ is obtained via the Laplace transform and the expression of the entropy is obtained via the Rohlin formula.

Proposition 10. The constant $\kappa_d(y)$ and the entropy $\mathcal{E}_d$ are expressed as one-dimensional integrals which involve the functions α and β defined as

$$\alpha(u) := \int_0^1 e^{-tu} dt = \frac{1-e^{-u}}{u}, \qquad \beta(u) := \int_0^u \alpha(v) dv\,, \tag{47}$$

and the following holds

$$\kappa_d(y) = \frac{1}{d!} \int_0^\infty e^{-u} \beta(uy)^d du, \qquad \mathcal{E}_d = (d+1) \int_0^1 \frac{1}{y} \frac{\kappa_d(y)}{\kappa_d} dy\,.$$

Proof.
Expression for $\kappa_d(y)$. As the function ψ_d defined in (37) involves the set of permutations, we remark that $\kappa_d(y)$ defined in (46) admits an alternative form

$$\kappa_d(y) = \int_{[0,y]^d} \varphi_d(\boldsymbol{x}) d\boldsymbol{x}, \qquad \text{with} \quad \varphi_d(\boldsymbol{x}) = \frac{1}{1+x_1} \cdots \frac{1}{1+x_1+x_2+\cdots+x_d}\,. \tag{48}$$

For $x \geq 0$, the equality

$$\frac{1}{1+x} = \int_0^{+\infty} e^{-tx} e^{-t} dt$$

gives rise to the equality

$$\varphi_d(x) = \int_{[0,+\infty]^d} H(\boldsymbol{t}, \boldsymbol{x}) d\boldsymbol{t}\,, \qquad \text{where} \quad H(\boldsymbol{t}, \boldsymbol{x}) := e^{-(t_1+t_2+\ldots t_d)}\, G(\boldsymbol{t}, \boldsymbol{x})$$

involves the function $G(\boldsymbol{t}, \boldsymbol{x}) = e^{-t_1 x_1} \ldots e^{-t_d(x_1+x_2+\ldots+x_d)} = e^{-x_1(t_1+t_2+\ldots t_d)} \ldots e^{-x_d t_d}$.
Under this last form, $G(\boldsymbol{t}, \boldsymbol{x})$ is a product of factors, each of them only involving the variable x_i, and the following d-dimensional integral is a product of one-dimensional integrals, namely

$$\int_{[0,y]^d} G(\boldsymbol{t}, \boldsymbol{x}) d\boldsymbol{x} = y^d A(y, \boldsymbol{t})\,,$$

where the function A is written in terms of the function α defined in (47) as

$$A(y, \boldsymbol{t}) = \alpha(yt_d) \cdot \alpha(y(t_d + t_{d-1})) \cdot \ldots \cdot \alpha(y(t_1 + \ldots t_d))\,.$$

Returning to the integral $\kappa_d(y)$, and after interverting integrations, we obtain

$$\kappa_d(y) = \int_{[0,+\infty]^d} e^{-(t_1+t_2+\dots t_d)}\, d\boldsymbol{t} \left(\int_{[0,y]^d} G(\boldsymbol{t},\boldsymbol{x}) d\boldsymbol{x} \right) = y^d \int_{[0,+\infty]^d} e^{-(t_1+\dots+t_d)}\, A(y,\boldsymbol{t}) d\boldsymbol{t}\,.$$

Letting $u_d = t_d$, $u_{d-1} = t_{d-1} + t_d, \dots, u_1 = t_1 + \dots + t_d$, and dealing with the simplex $\mathcal{S}_u := \{0 \le u_d \le u_{d-1} \le u_2 \le u\}$, we obtain

$$\frac{\kappa_d(y)}{y^d} = \int_0^\infty e^{-u} \alpha(yu) du \left(\int_{\mathcal{S}_u} \left[\prod_{i=2}^d \alpha(yu_i) \right] du_2\, du_3\, du_d \right).$$

The second integral involves a function that is symmetric with respect to the variables u_i for $i \in [2,d]$ on the simplex $\mathcal{S}_u$. Then the equality holds

$$\int_{\mathcal{S}_u} \left[\prod_{i=2}^d \alpha(yu_i) \right] du_2\, du_3\, du_d = \frac{1}{(d-1)!} \left(\int_{[0,u]} \alpha(yv) dv \right)^{d-1} = \frac{1}{(d-1)!} \left[\frac{\beta(yu)}{y} \right]^{d-1}$$

and involves the function $\beta(u) = \int_0^u \alpha(v) dv$. Finally, we obtain

$$d!\, \kappa_d(y) = \int_0^\infty e^{-u} d\, y\, \alpha(yu)\, \beta(yu)^{d-1}\, du$$

and an integration by parts leads to the final expression for $\kappa_d(y)$.

Expression for $\mathcal{E}_d$. The Rohlin formula relates the entropy of a dynamical system to the integral of the Jacobian of the shift T with respect to the invariant measure, and

$$\mathcal{E}_d = \frac{1}{\kappa_d} \int_{\mathcal{J}_d} |\log |J[T](\boldsymbol{x})||\ \psi_d(\boldsymbol{x})\, d\boldsymbol{x} = \frac{d+1}{\kappa_d} \int_{\mathcal{J}_d} |\log x_1|\, \psi_d(\boldsymbol{x})\, d\boldsymbol{x}\,.$$

The function

$$L(y) := \int_{\substack{\boldsymbol{x}\in\mathcal{J}_d \\ x_1 = y}} \psi_d(\boldsymbol{x}) d\boldsymbol{x}$$

coincides with the derivative of the map $y \mapsto \kappa_d(y)$. Then, with an integration by part, we obtain

$$\mathcal{E}_d := \frac{d+1}{\kappa_d} \int_0^1 |\log y|\, L(y) dy = \frac{d+1}{\kappa_d} \int_0^1 \frac{\kappa_d(y)}{y} dy\,.$$

Together with (45), this end the proof of Theorem 2(*a*). □

7.2. Asymptotic estimates for the function μ_d. Proof of Theorem 2(b).

We already mentioned that the integral $\kappa_d(y)$ is not exactly computable, except for small values of d, and Eq. (48) leads to easy bounds

$$\frac{1}{d!} y^d \le \kappa_d(y) \le y^d\,, \tag{49}$$

which are not of great use when $d \to \infty$. This is why we will conduct a finer asymptotic study, mainly based on Laplace estimates.

Proposition 11. Consider the two functions $y(d) := (\log d)/d$ and $\epsilon(d) = A(\log\log d)^{-1}$, for some constant A. The following asymptotic estimates hold for the function $y \mapsto \mu_d(y)$, when $d \to \infty$, and exhibit two different regimes

$$\begin{cases} \mu_d(y) \in \left[\left(\dfrac{\log dy}{\log d} \right)^{d-1/2+d\epsilon(d)} , \left(\dfrac{\log dy}{\log d} \right)^{d-1/2-d\epsilon(d)} \right] & \text{for } y \in [y(d), 1]\,, \\ \mu_d(y) \leq (\log d)^{1/2} \left(\dfrac{dy}{\log d} \right)^d & \text{for } y \in [0, y(d)]\,. \end{cases}$$

Proof. The integral which defines $\kappa_d(y)$ involves the function

$$\exp[h_{d,y}(u)], \qquad \text{with} \quad h_{d,y}(u) = d \log \beta(uy) - u\,. \tag{50}$$

As we will see, the function $h_{d,y}$ admits a maximum, attained at a point $u_{d,y}$ where its second derivative is not zero. The Laplace method provides asymptotic estimates for the integral, and relates the asymptotics of the integral to the maximum value $h_{d,y}(u_{d,y})$. We now recall these Laplace estimates, described for instance in [12] page 758.

Consider the integral

$$I_d = \int_0^\infty \exp\left[F_d(v)\right]\, dv$$

which involves a function F_d of class $\mathcal{C}^\infty$. Assume moreover the following[9]

(a) *it admits a unique maximum on $[0, \infty[$, attained at a point denoted as $v = v_d$,*
(b) *its second derivative F_d'' is not zero at v_d,*
(c) *the quotient $\gamma_d := \left| F_d'''(v_d)/(F_d''(v_d))^{3/2} \right|$ tends to zero.*

Then, the following estimate holds:

$$I_d \sim \sqrt{2\pi} \cdot \left[|F_d''(v_d)|^{-1/2} \cdot \exp\left[F_d(v_d)\right] \right] \Big(1 + O(\gamma_d) \Big)\,.$$

We first check that the previous properties hold for the function $h_{d,y}$ described in (50). As the derivative of the function $h_{d,y}$ is expressed with the derivative $\gamma := \alpha/\beta$ of the function $\log \beta$, which involves the functions α and β defined in (47), we first study the properties of γ and its derivatives.

Properties of γ. The main properties of γ are easily deduced from properties of α and β. For $u \to \infty$, the maps α and β admit the estimates

$$\alpha(u) \sim \frac{1}{u} \qquad \beta(u) \sim \log u \qquad \gamma(u) \sim \frac{1}{u \log u}\,.$$

Moreover, the function $\gamma : [0, +\infty[\to \mathbb{R}$ is decreasing from ∞ to 0, and, for $x > 0$, there exists a unique real $u = u(x)$ for which $\gamma(u(x)) = 1/x$. Furthermore, there exists a function $P(x)$ for which

$$u(x) = \frac{x}{P(x)} \qquad \text{with} \quad P(x) \sim \log x \quad (x \to \infty)\,.$$

We will use in the sequel the following properties of the map $x \mapsto P(x)$, valid for $x \to \infty$:

[9] The classical framework deals with the case when F_d is written as $F_d = dF$, and this will not be the case here.

(*i*) $P(x) = \log x + O(\log\log x)$,
(*ii*) The maps $x \mapsto P(x)$ and $x \mapsto x/P(x)$ are increasing.

The first two derivatives of γ satisfy

$$\begin{cases} |\gamma'(u)| \sim u^{-2}(\log u)^{-1} \quad (u \to \infty), & |\gamma'(u(x))| \sim |\log x| \cdot x^{-2} \quad (x \to \infty)\,. \\ |\gamma''(u)| \sim 2u^{-3}(\log u)^{-1} \quad (u \to \infty), & |\gamma''(u(x))| \sim 2\,|\log x|^2 \cdot x^{-3} \quad (x \to \infty)\,. \end{cases}$$

Properties of $h_{d,y}$. The equality $h'_{d,y}(u) = d \cdot y \cdot \gamma(uy) - 1$ holds; then, the function $h_{d,y}$ admits a maximum at $u_{d,y} = (1/y)\,u(dy)$, Property (a) holds, and the maximum value $h_{d,y}(u_{d,y})$ is written as $dH(dy)$ with

$$H(x) = -\frac{1}{P(x)} + \log\alpha(u(x)) + \log x = -\frac{1}{P(x)} + \log(1 - e^{-u(x)}) + \log P(x)\,.$$

We restrict ourselves to the case when $x \to \infty$. In this case, the term $\log(1 - e^{-u(x)})$ is exponentially small and completely negligible, and we "forget" it. We then study the simplified form of H, namely

$$\widehat{H}(x) = \log P(x) - \frac{1}{P(x)}\,. \tag{51}$$

The two terms relative to the second and the third derivatives then satisfy

$$\begin{cases} h''_{d,y}(u_{d,y}) = dy^2\gamma'(u(dy)), & |h''_{d,y}(u_{d,y})| \sim |\log dy| \cdot d^{-1}\,. \\ h''_{d,y}(u_{d,y}) = dy^3\gamma''(u(dy)), & |h'''_{d,y}(u_{d,y})| \sim 2\,|\log dy|^2 \cdot d^{-2}\,. \end{cases}$$

Finally, Properties (b) and (c) hold with a sequence γ_d of order $d^{-1/2}\,(\log dy)^{1/2}$.

Estimates for $\mu_d(y)$ for any y. For $d \to \infty$, with (51), the inequality $\widehat{H}(d) > \log\log d - 1$ holds. The Laplace method applies to κ_d itself, and this entails, together with the Stirling estimate, the lower bound for κ_d,

$$\kappa_d > \sqrt{2\pi}\left(\frac{d}{\log d}\right)^{1/2}\frac{1}{d!}\left(\frac{\log d}{e}\right)^d \quad\Longrightarrow\quad \kappa_d > \left(\frac{1}{\log d}\right)^{1/2}\left(\frac{\log d}{d}\right)^d\,.$$

With the trivial bound $\kappa_d(y) \le y^d$ described in (49) which holds for any $y \in [0,1]$ and is interesting when y is small, this provides an upper bound for $\mu_d(y)$, for any $y \le 1$,

$$\mu_d(y) \le (\log d)^{1/2}\ \left(\frac{yd}{\log d}\right)^d\,.$$

Of course, this bound is not very useful as soon as $y \ge y(d)$, and we now study this case.

Estimates for $\mu_d(y)$ for $y \ge y(d)$. With (51), the equality holds

$$\widehat{H}(dy) - \widehat{H}(d) = \log\left(\frac{P(dy)}{P(d)}\right) + \frac{1}{P(d)} - \frac{1}{P(dy)}\,.$$

Letting $v = P(dy)/P(d)$, and using the estimate $v \sim \log dy/\log d$ (for $d \to \infty$), one has

$$\widehat{H}(dy) - \widehat{H}(d) = (\log v)\,[1 - \eta(d,v)]\,, \qquad \text{with} \quad \eta(d,v) = \frac{1}{P(d)}\frac{(1-v)}{v\log v}$$

When y belongs to $[y(d),1]$, then the estimate $v=\Omega(\log\log d/\log d)$ holds, and entails the estimate $|\eta(d,v)|=o(\log\log d)^{-1}$. Now, using more precise estimates of $P(x)$, namely $P(x)=\log x\,(1+O(\log\log x/\log x))$, one has

$$\log P(dy)=\log\log dy+O\left(\frac{\log\log dy}{\log dy}\right)=\log\log dy\left[1+O\left(\frac{1}{\log dy}\right)\right]$$

and finally

$$\log\frac{P(dy)}{P(d)}=\log\frac{\log dy}{\log d}\left[1+O\left(\frac{1}{\log\log d}\right)\right].$$

Then, with $\epsilon(d)=O(\log\log d)^{-1}$, one has finally

$$\widehat{H}(dy)-\widehat{H}(d)\in\left[\log\frac{\log dy}{\log d}\,(1+\epsilon(d))\,,\ \ \log\frac{\log dy}{\log d}\,(1-\epsilon(d))\right].$$

Now, the ratio of the second derivatives satisfies

$$\left|\frac{h_d''(u_d)}{h_{d,y}''(u_{d,y})}\right|^{1/2}\sim\left(\frac{\log d}{\log dy}\right)^{1/2},$$

and the estimate for γ_d on the segment $[y(d),1]$ is of order $O(\log\log d)^{-1}$. One obtains the first estimate of the proposition.

Now if y belongs to some interval $[y_0(d),1]$ where $y_0(d)$ is $d^{\xi(d)}$ with $\xi(d)\to 0$, then the estimate becomes more precise: one has indeed in this case

$$\log\frac{\log dy}{\log d}=\log\left(1+\frac{\log y}{\log d}\right)=\frac{\log y}{\log d}\,\left[1+O(|\xi(d)|)\right],$$

and

$$\log\frac{P(dy)}{P(d)}=\frac{\log y}{\log d}\left[1+O\left(|\xi(d)|+\frac{1}{\log\log d}\right)\right]. \tag{52}$$

This is in particular the case when y is constant, with $\xi(d)=0$. □

7.3. Final estimates for the constants of the analysis. Proof of Theorem 2(c).

We finally provide the asymptotic estimates stated in Theorem 2 for the entropy $\mathcal{E}_d$, together with the constants θ_d and σ_d. The estimate for the entropy $\mathcal{E}_d$ contradicts a conjecture of Hardcastle and Khanin [14] which states that the entropy is asymptotically constant.

Proposition 12. The asymptotic estimates hold for the main constants of the analysis,

$$\mathcal{E}_d=\log d\left[1+O(\log\log d)^{-1}\right],\ \ \theta_d=1-(1/2)^{\Theta(1)d\log d},\ \ 1\leq\sigma_d\leq 2+O(\log\log d)^{-1}.$$

Proof. (*a*) We begin with the integral expression of $\mathcal{E}_d$ in terms of the function μ_d, and we "insert" there the previous estimates. We first consider the part brought by the integral of the function $y\mapsto(1/y)\mu_d(y)$ over the interval $[y(d),1]$. We deal with some exponent c which will be equal later to $d(1-1/2\pm\epsilon(d))$, and we perform the change of variables $u:=\log dy/\log d$. Then, the equalities hold, with $a(d)=(\log\log d)/(\log d)$,

$$\int_{y(d)}^1\left(\frac{\log dy}{\log d}\right)^c\frac{dy}{y}=\log d\int_{a(d)}^1 u^c du=\frac{\log d}{c+1}\left[1-a(d)^{c+1}\right]\sim\frac{\log d}{c+1}. \tag{53}$$

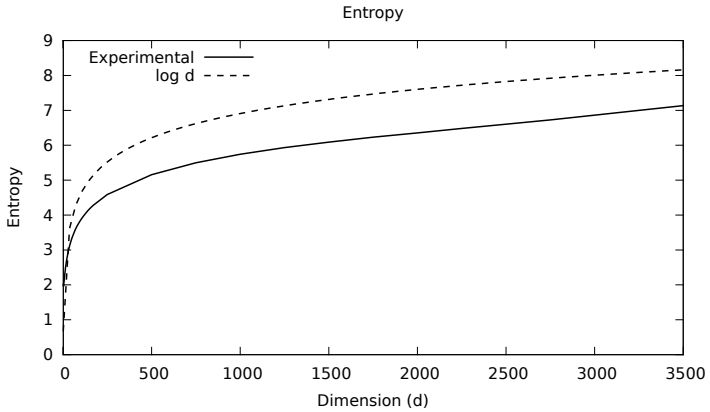

Fig. 4. Comparison between the mathematical curve $d \mapsto \mathcal{E}_d$ and the value of the entropy provided by the mean number of steps of the `BrunGcd` algorithm.

The part brought by the integral of the function $y \mapsto (1/y)\mu_d(y)$ over the interval $[0, y(d)]$ is at most

$$(\log d)^{1/2} \left(\frac{d}{\log d}\right)^d \int_0^{y(d)} y^{d-1} dy = (\log d)^{1/2} \left(\frac{d}{\log d}\right)^d \frac{y(d)^d}{d} \leq \frac{1}{d} (\log d)^{1/2} \tag{54}$$

With (53) applied with $c = d(1 - 1/2 \pm \epsilon(d))$ together with (54), we obtain the estimate for the entropy.

(b) When y is fixed, with (52), the estimate becomes $\mu_d(y) = y^{\Theta(1)\cdot(d/\log d)}$.

(c) We now study the constant σ_d and begin with the series of general term $\mu_d(1/m)$ which is compared to the corresponding integral via the inequalities

$$1 \leq \sum_{m \geq 1} \mu_d \left(\frac{1}{m}\right) \leq 1 + I \quad \text{with} \quad I := \int_1^\infty \mu_d(1/x)\, dx = \int_0^1 \mu_d(y) \frac{dy}{y^2}$$

We now estimate the integral I, separately dealing with the interval $[y(d), 1]$ where we use the bound $1/y \leq d/(\log d)$ together with the computation of (53), and we obtain

$$\int_{y(d)}^1 \mu_d(y) \frac{dy}{y^2} = 1 + O(\log\log d)^{-1},$$

whereas on the interval $[0, y(d)]$ the integral is at most

$$(\log d)^{1/2} \left(\frac{d}{\log d}\right)^d \int_0^{y(d)} y^{d-2} dy \leq (\log d)^{1/2} \left(\frac{d}{\log d}\right)^d y(d)^{d-1} \frac{1}{d-1} \leq \frac{1}{(\log d)^{1/2}}.$$

□

Figure 4 compares the theoretical curve $\mathcal{E}_d = \log d$ with experimental values of $\mathcal{E}_d$ which are obtained via the number of steps during the first phase.

7.4. A conjecture on the constants $\rho_{d,\ell}$.

We consider the strict ℓ-th phase in d dimensions, with $\ell \geq 1$ and we wish to study the cost $C_{d,\ell} = R_{d,\ell} - 1$ equal to the number of iterations of the `BrunGcd` algorithm(d) during this *strict* ℓ-th phase. The transfer operator of the remaining execution of the algorithm after the strict first phase $\ell = 0$ is the operator $\mathbf{D}_{s,(d)}$ defined in Proposition 9, and, as we are interested on the strict ℓ-th phase, we decompose the operator $\mathbf{D}_{s,(d)}$ as

$$\mathbf{V}_s \circ (I - \mathbf{H}_{t_k(s),(k)})^{-1} \circ \mathbf{W}_s, \qquad t_k(s) = s/(k+1)\,, \qquad (\text{with } \ell := d-k)\,,$$

whereas the cumulative transfer operator relative to the cost $C_{d,\ell}$ is

$$\mathbf{V}_s \circ \mathbf{H}_{t_k(s),(k)} \circ (I - \mathbf{H}_{t_k(s),(k)})^{-2} \circ \mathbf{W}_s\,.$$

Here, the operator $\mathbf{V}_s$ describes the execution of the algorithm after the strict ℓ-th phase, and the operator $\mathbf{W}_s$ describes the algorithm after the strict first-phase, until the beginning of the ℓ-th phase. Moreover, with Proposition 9, we are interested in the value of these two operators at $s = (d+1)$, and this leads to the study of the quasi-inverse

$$(I - \mathbf{H}_{t,(k)})^{-1}, \qquad \text{with } t = t_k(d+1) := \tfrac{d+1}{k+1}\,. \tag{55}$$

Proposition 9 proves that the mean value $\mathbb{E}_N[C_{d,\ell}]$ is asymptotic to a quotient, namely

$$\mathbb{E}_N[C_{d,\ell}] \sim \rho_{d,\ell} \quad (N \to \infty), \qquad \text{with} \quad \rho_{d,\ell} = \frac{\mathbf{V}[\psi](0)}{\mathbf{V}[\varphi](0)}\,, \tag{56}$$

which involves the operator $\mathbf{V} = \mathbf{V}_{d+1}$, together with the functions ψ and φ, themselves defined as

$$\begin{aligned} \varphi &:= (I - \mathbf{H}_{t,(k)})^{-1}[\phi], \quad \phi := \mathbf{W} \circ \mathbf{T}_{(d)}[1] \\ \psi &:= \mathbf{H}_{t,(k)} \circ (I - \mathbf{H}_{t,(k)})^{-1}[\varphi] = \mathbf{H}_{t,(k)} \circ (I - \mathbf{H}_{t,(k)})^{-2}[\phi]\,. \end{aligned} \tag{57}$$

Moreover, the real t is defined in (55), we let $\mathbf{W} := \mathbf{W}_{d+1}$, and the residue operator $\mathbf{T}_{(d)}$ is defined in Eq. (41) of Proposition 6.

We now study the ratio in (56), with three main steps.

First step. We relate this ratio to the spectral decomposition of the operator $\mathbf{H}_{t,(k)}$. With this spectral decomposition described in (35), we obtain

$$(I - \mathbf{H}_{t,(k)})^{-1}[G](\boldsymbol{x}) = \frac{\lambda_{(k)}(t)}{1 - \lambda_{(k)}(t)} \quad \left[\Psi_{t,(k)}](\boldsymbol{x}) \cdot I_{t,(k)}[G] + \mathbf{Y}^{(1)}_{t,(k)}[G](\boldsymbol{x})\right]\,,$$

$$\mathbf{H}_{t,(k)}(I - \mathbf{H}_{t,(k)})^{-2}[G](\boldsymbol{x}) = \frac{\lambda^2_{(k)}(t)}{(1 - \lambda_{(k)}(t))^2} \left[\Psi_{t,(k)}](\boldsymbol{x}) \cdot I_{t,(k)}[G] + \mathbf{Y}^{(2)}_{t,(k)}[G](\boldsymbol{x})\right]\,.$$

Here, the operators which intervene in the remainder term are defined from the operator $\mathbf{K}_{t,(k)}$ of (36) as

$$\mathbf{Y}^{(1)}_{t,(k)} = \mathbf{X}_{t,(k)}, \quad \mathbf{Y}^{(2)}_{t,(k)} = \mathbf{K}_{t,(k)}\mathbf{X}_{t,(k)}, \quad \text{with} \quad \mathbf{X}_{t,(k)} := \frac{1 - \lambda_{(k)}(t)}{\lambda_{(k)}(t)} \cdot (I - \mathbf{K}_{t,(k)})^{-1}\,.$$

Applied to $G = \phi$, this gives the decompositions

$$\begin{cases} \varphi(\boldsymbol{x}) &= \dfrac{\lambda_{(k)}(t)}{1 - \lambda_{(k)}(t)} \quad \left[\Psi_{t,(k)}(\boldsymbol{x}) \cdot I_{t,(k)}[\phi] + \mathbf{Y}^{(1)}_{t,(k)}[\phi](\boldsymbol{x})\right], \\ \psi(\boldsymbol{x}) &= \dfrac{\lambda^2_{(k)}(t)}{(1 - \lambda_{(k)}(t))^2} \quad \left[\Psi_{t,(k)}(\boldsymbol{x}) \cdot I_{t,(k)}[\phi] + \mathbf{Y}^{(2)}_{t,(k)}[\phi](\boldsymbol{x})\right]. \end{cases}$$

The images by the operator $\mathbf{V}$ decompose as

$$\begin{cases} \mathbf{V}[\varphi](0) &= \dfrac{\lambda_{(k)}(t)}{1-\lambda_{(k)}(t)} \qquad \Big[\mathbf{V}[\Psi_{t,(k)}](0)\cdot I_{t,(k)}[\phi] + \mathbf{V}\circ \mathbf{Y}^{(1)}_{t,(k)}[\phi](0)\Big]\,, \\ \mathbf{V}[\psi](0) &= \dfrac{\lambda^2_{(k)}(t)}{(1-\lambda_{(k)}(t))^2} \quad \Big[\mathbf{V}[\Psi_{t,(k)}](0)\cdot I_{t,(k)}[\phi] + \mathbf{V}\circ \mathbf{Y}^{(2)}_{t,(k)}[\phi](0)\Big]\,. \end{cases}$$

Then the ratio of the two functions which appears in (56) decomposes as

$$\frac{\mathbf{V}[\psi](0)}{\mathbf{V}[\varphi](0)} = \left(\frac{1+\boldsymbol{\epsilon}_2(t,k,\phi)}{1+\boldsymbol{\epsilon}_1(t,k,\phi)}\right)\cdot\left(\frac{\lambda_{(k)}(t)}{1-\lambda_{(k)}(t)}\right)\,, \tag{58}$$

and the first factor involves itself two positive ratios,

$$\boldsymbol{\epsilon}_i(t,k,\phi) = \frac{\mathbf{V}\circ \mathbf{Y}^{(i)}_{t,(k)}[\phi](0)}{\mathbf{V}[\Psi_{t,(k)}](0)\cdot I_{t,(k)}[\phi]}\,. \tag{59}$$

We now return to the study of the (strict) ℓ-th phase of the algorithm in d dimensions. With $\ell := d\text{–}k$, the triple (t,k,ϕ) depends on the pair (d,ℓ) with $k = d\text{–}\ell$, $t = (d{+}1)/(k{+}1)$ and ϕ itself also depends on the pair (d,ℓ), as the decomposition in (57) depends on this pair. The ratio of (58) is written as

$$\rho_{d,\ell} = B(d,\ell)\,\left(\frac{\lambda(d,\ell)}{1-\lambda(d,\ell)}\right)\,; \tag{60}$$

the second factor involves the eigenvalue

$$\lambda(d,\ell) := \lambda_{(d-\ell)}(\tfrac{d+1}{d-\ell+1}) = \lambda_{(d-\ell)}(1+\tfrac{\ell}{d-\ell+1})\,,$$

whereas the first factor

$$B(d,\ell) := \frac{1+\boldsymbol{\epsilon}_2(t,k,\phi)}{1+\boldsymbol{\epsilon}_1(t,k,\phi)}\,, \qquad \text{with } k = d-\ell,\quad t = \tfrac{d+1}{k+1} \tag{61}$$

is expressed itself with two ratios $\boldsymbol{\epsilon}_i$ defined in (59).

Second step. We first consider the second factor, and we "replace" the eigenvalue by the algebraic number of the worst case, which is indeed reached when all the quotients are equal to 1 (and we know that this almost always happens in high dimensions). We indeed use Eq. (40) which relates here the eigenvalue $\lambda(d,\ell)$ and the algebraic number of the worst case $\tau_{d-\ell}$, and provides the sharp bounds

$$\exp[-\mathcal{E}\tfrac{\ell}{d-\ell+1}] \le \lambda(d,\ell) \le \tau^\ell_{d-\ell}, \qquad \frac{\lambda(d,\ell)}{1-\lambda(d,\ell)} \le \frac{\tau^\ell_{d-\ell}}{1-\tau^\ell_{d-\ell}}\,. \tag{62}$$

We thus let

$$\rho_{d,\ell} = A(d,\ell)\cdot\widehat{\rho}_{d,\ell} \qquad \text{with} \qquad \widehat{\rho}_{d,\ell} = \frac{\tau^\ell_{d-\ell}}{1-\tau^\ell_{d-\ell}}\,, \tag{63}$$

where the constants $A(d,\ell)$ are smaller than the initial constants $B(d,\ell)$.

We now study the ratio $\widehat{\rho}_{d,\ell}$. Using the asymptotic estimate of τ_k given in (4), we first exhibit three different regimes for the function $\ell \mapsto \tau^\ell_{d-\ell}$ according to whether ℓ be "small", "moderate" or "large".

(*i*) When $\ell = o(d/\log d)$, one has $\tau_{d-\ell}^{\ell} \to 1$

$$\tau_{d-\ell}^{\ell} := \exp(\ell \log \tau_{d-\ell}) \sim \exp\left(-\frac{\ell}{d}\log d\right) \qquad \text{with} \quad \frac{\ell}{d}\log d \to 0\,.$$

(*ii*) When $\ell = \Theta(d)$, one lets $\ell := ad$, one has $\tau_{d-\ell}^{\ell} \to 0$ and

$$\tau_{d-\ell}^{\ell} := \exp(\ell \log \tau_{d-\ell}) \sim -K(a)\left(\frac{\log d}{d}\right)^{a/(1-a)} \quad \text{with} \quad K(a) = (1-a)^{-a/(1-a)}\,.$$

(*iii*) When $d - \ell$ is bounded by some constant K, one has $\tau_{d-\ell}^{\ell} \le \tau_K^{\ell}$.

The following result thus provides *proven* results for the function $\widehat{\rho}_{d,\ell}$.

Proposition 13. For a fixed dimension d, there are (at least) three regimes for the function $\ell \mapsto \widehat{\rho}_{d,\ell}$ defined in (63):

$$\begin{cases} \text{For } \ell \text{ small}, \quad \ell = o(d/\log d): & \widehat{\rho}_{d,\ell} \sim \left(\dfrac{d}{\log d}\right)\dfrac{1}{\ell}, \\ \text{For } \ell \text{ moderate}, \quad \ell = ad: & \widehat{\rho}_{d,\ell} \sim K(a)\left(\dfrac{\log d}{d}\right)^{a/(1-a)}, \\ \text{For } \ell \text{ large}, \quad \ell > d - K: & \widehat{\rho}_{d,\ell} \le \tau_K^{\ell}\,. \end{cases}$$

Remark. For a fixed dimension d, the three regimes for the function $\ell \mapsto \widehat{\rho}_{d,\ell}$ can be informally described as follows: for small values of the phase index ℓ, the function decreases as $(1/\ell)$; for moderate values of ℓ, the function is polynomially decreasing (with an exponent which depends on the position of ℓ in the range); finally, in the large range, the function has an exponential decrease.

Third step. We now state our conjecture.

Conjecture. *Denote by τ_k the algebraic number which is the unique root < 1 of the equation $x^{k+1} + x - 1 = 0$ and appears in the worst-case analysis of the algorithm in k dimensions. Then, there are two absolute constants $a_- > 0$ and a_+ for which the constant $\rho_{d,\ell}$ which appears in Theorem 1(b) satisfies for any $d > 1$, $\ell \in [1, d]$,*

$$\rho_{d,\ell} = A(d,\ell) \cdot \frac{\tau_{d-\ell}^{\ell}}{1 - \tau_{d-\ell}^{\ell}} \qquad \textit{with} \quad a_- \le A(d,\ell) \le a_+\,. \tag{64}$$

Fourth Step. We now (try to) study the family $A(d,\ell)$. We recall that $A(d,\ell)$ satisfies $A(d,\ell) \le B(d,\ell)$, and $B(d,\ell)$ is itself expressed in (60) and (61) with ratios $\boldsymbol{\epsilon}_i$. We have then to estimate each ratio defined in (59), and we provide estimates of values $\mathbf{V}[F](0)$ which intervene in such ratios.

Lemma 2. For any $d \ge 2$, any $k \in [2..d-1]$, any F of class $\mathcal{C}^1$ on $\mathcal{J}_{(k)}$, the following estimate of $\mathbf{V}[F](0)$ involves the integral $I_{(k)}[F]$ of F on $\mathcal{J}_{(k)}$,

$$\mathbf{V}[F](0) = \frac{\zeta(d+1-k)}{\zeta(k)}\, I_{(k)}[F] + I_{(k)}[1]\, Z(d,k)\; O\left(|F||_1\right)$$

where $Z(d,k)=\zeta(d)$ for $k\geq 3$ and $Z(d,2)=\zeta'(2)$. Moreover, the constant in the O-term is uniform, and does not depend on d,k,F.

Proof. By Lemma 1, the operator $\mathbf{V}$ transforms a (general) function F defined on $\mathcal{J}_{(k)}$ into the function

$$\sum_{h\in\mathcal{M}} |D[h]|^{-(d+1)}\, F\circ h$$

defined on $\mathcal{J}_{(1)}$. Here, the set $\mathcal{M}$ gathers all the inverse branches which intervene in the execution of the algorithm after the strict ℓ-th phase. Then, the equality holds

$$\Big\{h(0)\mid h\in\mathcal{M}\Big\}=\bigcup_{q\geq 1} A_q,\ \text{ with }\ A_q:=\left\{\left(\frac{p_1}{q},\dots,\frac{p_k}{q}\right)\in\mathcal{J}_{(k)}\mid \gcd(p_1,p_2,\dots,p_k,q)=1\right\},$$

and entails another expression for $\mathbf{V}[F](0)$, i.e.,

$$\mathbf{V}[F](0):=\sum_{h\in\mathcal{M}}|D[h](0)|^{-(d+1)}F\circ h(0)=\sum_{q\geq 1}\frac{1}{q^{d+1-k}}\left(\frac{1}{q^k}\sum_{\boldsymbol{x}\in A_q}F(\boldsymbol{x})\right). \tag{65}$$

Here, the last factor is the "coprime" Riemann sum (with denominator q) associated with the integral $I_{(k)}[F]$ of the function F on the simplex $\mathcal{J}_{(k)}$. It is denoted as $\widehat{R}_q[F,\mathcal{J}_{(k)}]$ and studied for instance in [26]. Applied to our context, this proves that the "coprime" Riemann sum satisfies

$$\widehat{R}_q[F,\mathcal{J}_{(k)}]=\frac{1}{\zeta(k)}\,I_{(k)}[F]+\left(\frac{1}{q^{k-1}}\right)\,I_{(k)}[1]\,O\left(||F||_1\right),\qquad \text{for } k\geq 3,$$

$$\widehat{R}_q[F,\mathcal{J}_{(2)}]=\frac{1}{\zeta(2)}\,I_{(2)}[F]+\left(\frac{\log q}{q}\right)\,I_{(2)}[F]\,O\left(||F||_1\right).$$

With Eq. (65), this leads to the final estimate of $\mathbf{V}[F](0)$. □

When replacing the general function F of the Lemma by our functions of interest (which themselves depend on d and k), the Lemma provides an estimate of each $\boldsymbol{\epsilon}_i$ of Eq. (59), which is probably useful. But, our knowledge of the function ϕ is not precise enough, and we do not know how to use it for proving that each $\boldsymbol{\epsilon}_i$ is bounded. We thus do not succeed in proving that the family $A(d,\ell)$ is bounded.

Final result. We now collect the four steps of our study, and obtain our final result:

Proposition 14. There are (at least) three regimes for the function $\ell\mapsto\rho_{d,\ell}$ which involve the family of constants $A(d,\ell)$ together with the constants $a\mapsto K(a)$ which appear in the "moderate" case of Proposition 13.

$$\begin{cases}\text{For } \ell \text{ small,}\quad \ell=o(d/\log d): & \rho_{d,\ell}\sim A(d,\ell)\left(\dfrac{d}{\log d}\right)\dfrac{1}{\ell},\\[2ex] \text{For } \ell \text{ moderate,}\quad \ell=ad: & \rho_{d,\ell}\sim A(d,ad)\,K(a)\left(\dfrac{\log d}{d}\right)^{a/(1-a)},\\[2ex] \text{For } \ell \text{ large,}\quad \ell>d-K: & \rho_{d,\ell}\leq A(d,d-\ell)\,\tau_K^{\ell}\,.\end{cases}$$

Under the validity of the conjecture, the family $A(d,\ell)$ is uniformy bounded for any (d,ℓ) with $d\geq 2$, and $\ell\in[2,d]$.

7.5. Experimental studies for the family of constants $\rho_{d,\ell}$.

We have performed an experimental study that is summarized in Figures 5 and 6. These experiments provide experimental values for $\rho_{d,\ell} := \mathbb{E}_N[C_{d,\ell}]$ which are denoted as $\widetilde{\rho}_{d,\ell}$. Comparing experimental values $\widetilde{\rho}_{d,\ell}$ with the theoretical values $\widehat{\rho}_{d,\ell}$ will give hints on the behaviour of the family $A(d,\ell)$ and thus on the validity of our conjecture.

Figure 5 exhibits three experimental curves which are the graphs of the function $\ell \mapsto 1+\widetilde{\rho}_{d,\ell}$. They are performed for dimensions $d \in \{1000, 2000, 4000\}$, and small phase indices $\ell \in [1..50]$. They are obtained with an input size $\log_2 N = 10^3$ and a number $M = 10^4$ of experiments. The conjecture predicts that the curves are hyperbols, and this appears to be actually the case.

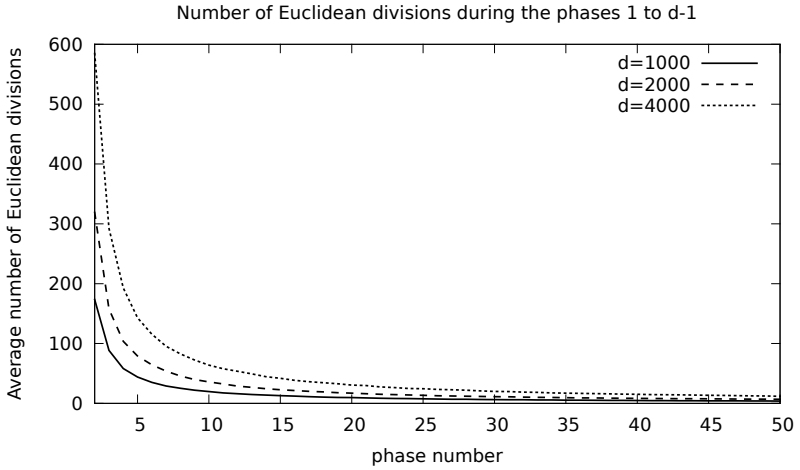

Fig. 5. Three experimental curves $\ell \mapsto \rho_{d,\ell}$ obtained with an input size $\log_2 N = 10^3$ for dimensions $d \in \{1000, 2000, 4000\}$, and small phase indices $\ell \in [1..50]$

The table of Figure 6 gathers experimental values $\widetilde{\rho}_d$ for $\rho_d := \mathbb{E}_N[R_d]$ and $\widetilde{\rho}_{d,\ell}$ for $\rho_{d,\ell} := \mathbb{E}_N[R_{d,\ell}] - 1$ for dimensions $d \in \{20, 50, 100, 250\}$ and two groups of phase indices ℓ: the columns on the left gather experiments performed for small values of ℓ, whereas the colums on the right gather experiments obtained for moderate values of ℓ. The values in parentheses are the theoretical values $\widehat{\rho}_{d,\ell}$ obtained in Proposition 13. The values in brackets are the ratios $\widetilde{A}(d,\ell) := \widetilde{\rho}_{d,\ell}/\widehat{\rho}_{d,\ell}$: they should be close to 1. Moreover, for small values of ℓ, the values in $\langle\cdot\rangle$ are the ratios $(1+\widetilde{\rho}_{d,\ell})/(1+\widetilde{\rho}_{d,1})$ between experimental values for successive phases: they have to be compared to the sequence $\ell \mapsto 1/\ell$.

We observe the following two experimental facts:

(*a*) ***Small values of phase indices*** ℓ**.** The conjecture predicts that, for small values of ℓ, the ratios $(1+\widetilde{\rho}_{d,\ell})/(1+\widetilde{\rho}_{d,1})$ will be close to $1/\ell$. There is a very good fit indeed between the values in $\langle\cdot\rangle$ of the fourth line and the sequence $\ell \mapsto 1/\ell$.
Always for small values of the index phase ℓ, the experimental ratios $\widetilde{A}(d,\ell) := \widetilde{\rho}_{d,\ell}/\widehat{\rho}_{d,\ell}$ (shown in brackets at the third line on the left) appear to be close enough to 1 (when d is large enough).

(*b*) ***Moderate values of phase indices*** ℓ**.** The experimental values $\widetilde{A}(d,\ell)$ for the ratio $A(d,\ell)$ (shown in brackets at the third line) seem to show the existence of an upper bound for the family $A(d,\ell)$, that seems at most equal to 2. However, the existence of a strictly positive lower bound a_- as stated in the conjecture appears to be more problematic: the function $d \mapsto \widetilde{A}(d,\ell)$ seems to be decreasing both on a fixed column (when ℓ/d is fixed and d is increasing) and on a fixed line too (when d is fixed and ℓ is increasing). This existence of a_- seems possible on the "lower moderate" range (for values $\ell \leq d/2$), but is not clear at all in the "higher moderate" range ....

	$\mathbb{E}_N[R_d]$	$\rho_{d,\ell} := \mathbb{E}_N[R_{d,\ell}] - 1$							
d \ ℓ		1	2	3	4	$d/5$	$2d/5$	$d/2$	$3d/5$
20	35.3	7.07	3.16	1.88	1.24	1.25	0.33	0.17	0.08
		(6.67)	(3.34)	(2.23)	(1.67)	(0.66)	(0.39)	(0.30)	(0.23)
		[1.05]	[0.94]	[0.84]	[0.74]	[1.89]	[0.84]	[0.56]	[0.34]
		⟨1.0⟩	⟨0.51⟩	⟨0.36⟩	⟨0.28⟩				
50	91.5	14.58	6.91	4.38	3.10	0.85	0.17	0.07	0.028
		(12.8)	(6.39)	(4.26)	(3.19)	(0.56)	(0.26)	(0.16)	(0.09)
		[1.13]	[1.08]	[1.02]	[0.97]	[1.51]	[0.65]	[0.43]	[0.31]
		⟨1.0⟩	⟨0.51⟩	⟨0.35⟩	⟨0.26⟩				
100	185.9	25.5	12.4	8.02	5.79	0.652	0.105	0.036	0.0109
		(21.7)	(10.9)	(7.24)	(5.43)	(0.49)	(0.181)	(0.092)	(0.039)
		[1.17]	[1.13]	[1.10]	[1.06]	[1.35]	[0.58]	[0.39]	[0.28]
		⟨1.0⟩	⟨0.506⟩	⟨0.340⟩	⟨0.256⟩				
250	470.83	53.79	26.71	17.46	13.02	0.46	0.049	0.0121	0.002
		(45.2)	(22.6)	(15.1)	(11.32)	(0.41)	(0.11)	(0.044)	(0.013)
		[1.19]	[1.18]	[1.15]	[1.15]	[1.12]	[0.44]	[0.27]	[0.15]
		⟨1.0⟩	⟨0.505⟩	⟨0.337⟩	⟨0.256⟩				

Fig. 6. This table gathers experimental values for the family of mean values, namely, $\widetilde{\rho}_d$ for $\rho_d := \mathbb{E}_N[R_d]$ and $\widetilde{\rho}_{d,\ell}$ for $\rho_{d,\ell} := \mathbb{E}_N[R_{d,\ell}] - 1$. These experiments are obtained with an input size $\log_2 N = 5 \cdot 10^3$ for dimensions $d \in \{20, 50, 100, 250\}$ and two groups of phase indices ℓ: the columns on the left gather experiments performed for small values of ℓ, whereas the colums on the right gather experiments obtained for moderate values of ℓ. The values in parentheses are the theoretical values $\widehat{\rho}_{d,\ell}$ obtained in Proposition 13. The values in brackets are the ratios $\widetilde{A}(d,\ell) := \widetilde{\rho}_{d,\ell}/\widehat{\rho}_{d,\ell}$: they should be close to 1. Moreover, for small values of ℓ, the values in $\langle\cdot\rangle$ are the ratios $(1+\widetilde{\rho}_{d,\ell})/(1+\widetilde{\rho}_{d,1})$ between experimental values for successive phases: they have to be compared to the sequence $\ell \mapsto 1/\ell$.

In conclusion, our conjecture seems to be true in the "small" range. In the "moderate" range, the family $A(d,\ell)$ appears to admit an upper bound, but the existence of a lower bound is not clear.

8. Conclusion and open problems.

We have precisely described the probabilistic behaviour of the `BrunGcd`(d) algorithm, on the set $\Omega_{(d,N)}$ which gathers the vectors in $d+1$ dimensions whose components are all at most equal to N. We have then studied its behaviour when d is fixed and $N \to \infty$, which is defined via a set of constants $a_d, \theta_d, \sigma_d, \rho_d$ which only depends on the ambient dimension. We have used fine properties of the invariant measure of the underlying dynamical system, which has not the same behaviour as in the one dimensional-case (the Euclid algorithm associated with the Gauss map). This proves in particular that the mean number of subtractive steps is finite, contrarily to the case $d = 1$. We have also elucidated the asymptotics of these constants when $d \to \infty$ (at least for the first three constants). All these constants play an important role in the description of the Brun dynamical system in d dimensions, and our study leads to a better understanding of the Brun dynamics when the dimension d tends to ∞, which had never been thoroughly studied, except for some aspects in [14].

The present study also leads us to a precise description of the Brun rational trajectories. As highlighted in Section 5.7, our dynamical methods prove that these finite trajectories behave in the average-case exactly as the truncated Brun real trajectories. We may use this knowledge in order to study simultaneous approximations associated with rational vectors, which corresponds to the real algorithmic situation.

However, we stress that we *do not* analyze the behaviour of the algorithm `BrunGcd` (d) on the set $\Omega_{(d,N)}$ when *both* the size N of the inputs and the ambient dimension d tend to ∞. The methods presented here do not allow such an analysis, as they only provide the dominant terms of the asymptotics, without *remainder* terms. For performing a more detailed analysis, which would lead to asymptotic estimates for $(d, N) \to \infty$, we need to know the behaviour of our quasi-inverses $\mathbf{Q}_{t,(d)}$ when t is on the left of the vertical line $\Re t = 1$. If there exists a *vertical strip* on the left of $\Re t = 1$ where the quasi-inverse is *of polynomial growth* when $|\Im t| \to \infty$ (what we call a VSPG), then is possible to "replace" the Delange Theorem by the Perron formula, which now provides remainder terms. In the case of the Euclid algorithm (case $d = 1$), the existence of a VSPG is based on deep results due to Dolgopyat [11] and was proven in [3]. The article [3] exhibits an asymptotic expansion for the mean number of steps

$$\mathbb{E}_N[M_1] = a_1 \log N + p_1 + O(N^{-\alpha}) \qquad (N \to \infty)\,.$$

Here, the exponent α is related to the width of the vertical strip, and the constant p_1 is a variant of the Porter constant: this is a computable constant, with an intricate expression due to Porter which involves many characteristics of the Gauss dynamical system. (See [13] for a precise discussion on variants of the Porter constant.)

If we dream a little bit, we could expect a VSPG in d dimensions associated with a width α_d; in this case, there would be an asymptotic expansion for the mean number of steps during the first phase, namely

$$\mathbb{E}_N[M_d] = a_d \log N + p_d + O_d(N^{-\alpha_d}) \qquad (N \to \infty)\,,$$

where the exponent α_d is the width of the VSPG in d dimensions, and the constant p_d is the extension of the Porter constant in d dimensions. In the same vein, we would also obtain a similar result for the mean number of steps during the other phases

$$\mathbb{E}_N[R_d] = \rho_d + O_d(N^{-\alpha_d}) \qquad (N \to \infty)\,.$$

Then, the mean value of the total number of steps would be

$$\mathbb{E}_N[L_d] = a_d \log N + (p_d + \rho_d) + O_d(N^{-\alpha_d}) \qquad (N \to \infty)\,.$$

Even in the case when such asymptotic estimates (for $N \to \infty$) could be proven for each dimension d, a complete analysis for $(d, N) \to \infty$ would depend on the asymptotics of the three constants a_d, ρ_d, p_d for $d \to \infty$. The first one is known, and we have stated a plausible conjecture about the second one. However, the asymptotic study of the Porter constant p_d seems very difficult to perform, and we have no idea, even about a possible conjecture.

The present article also proves that the `BrunGcd` algorithm is not efficient. This is probably the case for all the gcd algorithms which are based on multidimensional continued fraction algorithms in high ambient dimensions. For instance, the Gcd algorithm based on the Jacobi-Perron dynamical system has probably the same principal features as the `BrunGcd` algorithm, even if it is more complex to study: it is indeed not complete (although Markovian), and its invariant density is not known at all, whereas we strongly rely here on the "semi-explicit" form of the Brun invariant density to derive the behaviour of the Brun algorithm in high dimensions.

Finally, it would be interesting to compare with the extended gcd algorithm based on the LLL algorithm and described in [24]. The analysis of such an algorithm is surely intricate, as its underlying system is quite complex to deal with.

References

[1] P. Arnoux and A. Nogueira. Mesures de Gauss pour des algorithmes de fractions continues multidimensionnelles. *Ann. Sci. Ecole Norm. Sup.*, 26:645–664, 1993.

[2] V. Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[3] V. Baladi and B. Vallée. Euclidean algorithms are Gaussian. *J. Number Theory*, 110:331–386, 2006.

[4] V. Berthé, J. Bourdon, T. Jolivet, and A. Siegel. Generating discrete planes with substitutions. In *Combinatorics on words*, LNCS 8079 (2013), pages 58–70. Springer, 2013.

[5] V. Berthé, L. Lhote, and B. Vallée. Analysis of the brun gcd algorithm. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2016*, pages 87–94, 2016.

[6] V. Berthé, L. Lhote, and B. Vallée. Probabilistic analyses of the plain multiple gcd algorithm. *Journal of Symbolic Computation*, 74:425–474, 2016.

[7] A. Broise. Transformations dilatantes de l'intervalle et théorèmes limites. *Astérisque*, 238:1–109, 1996. Études spectrales d'opérateurs de transfert et applications.

[8] A. Broise-Alamichel and Y. Guivarc'h. Exposants caractéristiques de l'algorithme de Jacobi-Perron et de la transformation associée. *Annales de l'Institut Fourier*, 51(3):565–686, 2001.

[9] V. Brun. Algorithmes euclidiens pour trois et quatre nombres. In *13e congrès des mathématiciens scandinaves, Helsinki 1957*, pages 45–64. Mercators Tryckeri, Helsinki, 1958.

[10] P. de Rooij. Efficient exponentiation using precomputation and vector addition chains. In *Advances in Cryptology, EUROCRYPT '94 Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 389–399. Springer-Verlag, 1995.

[11] D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math.*, 147(2):357–390, 1998.

[12] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[13] P. Flajolet and B. Vallée. Continued fractions, comparison algorithms, and fine structure constants. *Proceedings of Canadian Mathematical Society*, 27:53–82, 2000.

[14] D. M. Hardcastle and K. Khanin. On almost everywhere strong convergence of multi-dimensional continued fraction algorithms. *Ergodic Theory Dynam. Systems*, 20(6):1711–1733, 2000.

[15] D. M. Hardcastle and K. Khanin. The d-dimensional Gauss transformation : strong convergence and lyapounov exponents. *Experiment. math.*, 11(1):119–129, 2002.

[16] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118(2):627–634, 1993.

[17] C. T. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)*, 52:140–147, 1950.

[18] D. Jamet, N. Lafrenière, and X. Provençal. Generation of digital planes using generalized continued-fractions algorithms. In *Discrete geometry for computer imagery*, volume 9647 of *Lecture Notes in Comput. Sci.*, pages 45–56. Springer, [Cham], 2016.

[19] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin-New York, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.

[20] D. E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, 3rd edition, 1998.

[21] J. Lagarias. The quality of the diophantine approximations found by the Jacobi-Perron algorithm, and related algorithms. *Mh. Math*, 115:299–328, 1993.

[22] C. Lam, J. Shallit, and S. Vanstone. Worst-case analysis of an algorithm for computing the greatest common divisor of n inputs. *in Proceedings of an International Conference on Coding Theory, Cryptography and Related Areas,*, 1998.

[23] L. Lhote and B. Vallée. Gaussian laws for the main parameters of the Euclid algorithms. *Algorithmica*, 50(4):497–554, 2008.

[24] B. S. Majewski and G. Havas. The complexity of greatest common divisor computations. In *Algorithmic number theory (Ithaca, NY, 1994)*, volume 877 of *Lecture Notes in Comput. Sci.*, pages 184–193. Springer, Berlin, 1994.

[25] E. V. Podsypanin. A generalization of the continued fraction algorithm that is related to the Viggo Brun algorithm. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 67:184–194, 227, 1977. Studies in number theory (LOMI), 4.

[26] P. Rotondo and B. Vallée. The recurrence function of a random sturmian word. In *Proceedings of ANALCO*, 2017.

[27] F. Schweiger. Ergodische Eigenschaften der Algorithmen von Brun und Selmer. *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II*, 191(8-9):325–329, 1982.

[28] F. Schweiger. *Multidimensional continued fractions*. Oxford University Press, 2000.

[29] F. Schweiger. A note on Lyapunov theory for Brun algorithm. In *Diophantine approximation*, volume 16 of *Dev. Math.*, pages 371–379. SpringerWienNewYork, Vienna, 2008.

[30] B. Vallée. Euclidean dynamics. *Discrete and Continuous Dynamical Systems*, 1(15):281–352, 2006.

[31] M. S. Waterman. A Jacobi algorithm and metric theory for greatest common divisors. *J. Math. Anal. Appl.*, 59(2):288–300, 1977.