Theoretical Computer Science 548 (2014) 25–53

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Asymptotic expansions for linear homogeneous divide-and-conquer recurrences: Algebraic and analytic approaches collated

Philippe Dumas

SpecFun Team, Inria-Saclay, France

ARTICLE INFO

Article history:
Received 8 April 2013
Received in revised form 2 June 2014
Accepted 18 June 2014
Available online 8 July 2014
Communicated by H. Prodinger

Keywords:
Divide-and-conquer recurrence
Radix-rational sequence
Spectral radius
Dilation equation
Cascade algorithm
Dirichlet series
Fourier series

ABSTRACT

Among all sequences that satisfy a divide-and-conquer recurrence, those which are rational with respect to a numeration system are certainly the most basic and the most essential. Nevertheless, until recently this specific class of sequences has not been systematically studied from the asymptotic standpoint. We recall how a mechanical process designed by the author permits to compute their asymptotic expansions. The process is based on linear algebra, and involves computing Jordan normal forms, joint spectral radii, and solving dilation equations. The main contribution of the present article is the comparison between our algebraic method and the classical analytic number theory approach. Moreover, we develop new ways to compute the Fourier series of the periodic functions involved in the expansion. The article comes with an extended bibliography.

1. Introduction

This article deals with the asymptotic study of a special type of divide-and-conquer sequences, namely *radix-rational sequences*, which are rational with respect to a numeration system. In most cases where the numeration system is the binary system, such sequences satisfy a recurrence which essentially links the values u_{2n} and u_{2n+1} to the value of u_n by a linear homogeneous equation. In the first part of the article (Sections 2 to 5), a new approach based on linear algebra is set out; in the second part (Sections 6 and 7) this algebraic method is compared with the more classical approach relying on analytic number theory.

1.1. Ubiquity

Radix-rational sequences emerge in various subfields of mathematics and computer science. The first example that comes to mind [1,2] is the binary sum-of-digits function, that is the number s_n of 1's in the binary expansion of an integer n [3]. This sequence satisfies $s_0 = 0$, $s_{2n} = s_n$, $s_{2n+1} = s_n + 1$. Although one may think that such a sequence is of limited interest, it actually occurs in many problems like the study of the maximum of a determinant of an $n \times n$ matrix with entries ± 1 [4] or in a merging process occurring in graph theory [5]. This basic example has been greatly generalized to the number of occurrences of some pattern in the binary code [6], or in the Gray code: Flajolet and Ramshaw [7] study the average case of

E-mail address: Philippe.Dumas@inria.fr.

http://dx.doi.org/10.1016/j.tcs.2014.06.036
0304-3975/

Batcher's odd–even merge by using the sum-of-digits function of the Gray code. Among the sequences directly related to a numeration system, the Thue–Morse sequence which writes $u_n = (-1)^{s_n}$ is certainly the one which has caused the greatest number of publications [8]. There exist variants involving some subsequences [9,10] or some binary patterns [6] other than the simple pattern 1: the Rudin–Shapiro sequence associated with the pattern 11 satisfies $u_{2n} = u_{4n+1} = u_n$, $u_{4n+3} = -u_{2n+1}$, $u_0 = u_1 = 1$, and was initially designed to minimize the L^∞ norm of a sequence of trigonometric polynomials with coefficients ± 1 [11,12]. The study of the complexity of algorithms is another source of radix-rational sequences. The cost c_n of computing the n-th power of a matrix by binary powering satisfies $c_0 = 0$, $c_{2n} = c_n + 1$, $c_{2n+1} = c_n + 2$. The idea of binary powering has been re-employed by Morain and Olivos [13] in the context of computations on an elliptic curve where the subtraction has the same cost than addition. Supowit and Reingold [14] have used the divide-and-conquer strategy to give heuristics for the problem of Euclidean matching and this has led them to a 4-rational sequence in the worst case. Number theory is another domain which provides examples like the number of odd binomial coefficients in the n-th row of Pascal's triangle [15], or the number of integers bounded by n, which are representable as sums of three perfect squares [16].

1.2. Context

Sequences related to a numeration system have a long history, but as far as asymptotic behavior is concerned, the first studies go back to the middle of the 20th century. The studies in question deal with one example at a time and use elementary methods. In our opinion, the most noteworthy article is the Brillhart, Erdős, and Morton's work [17], for it contains the seeds of almost all ideas about the topic: dilation equations, asymptotic expansions with periodic coefficients, Hölder continuous and non-differentiable functions, bounds for the periodic functions, Fourier series, Dirichlet series, residues, Mellin transforms. It can be and needs to be read and re-read many times in order to extract all its very substance.

A more systematic study has begun in the late seventies, with two independent methods. One is related to the combinatorics of words [18–21], and is summarized in [22, Sec. 4.1 and Sec. 6.1]. The other is based on analytic number theory [7, 23–25]. Its most evolved version is [26], even if it is essentially limited to positive sequences, due to the fact that it mixes several radices.

Until recently, the idea of radix-rational sequence was not emphasized, but the rationality property sometimes rises to the surface [27, Corollaire, p. 13-07], [28, Satz 3 and Satz 4]. This is not surprising, since it was defined only twenty years ago [29]. The sequences are viewed only as satisfying divide-and-conquer recurrences without this wording being necessarily employed. Nevertheless, a linear representation appears in [30], where a new approach based on Fourier transforms and probability theory is used (this last point seems to limit the use of the method to nonnegative linear representations). Besides, dilation equations arise in [31–33] and are systematically used in [34]. But these works limit themselves to scalar-valued functions. By contrast, [35] uses matrices, eigenvalues and dilation equations, but the study is very specific to the Thue–Morse sequence.

Stolarsky [15] provides a rich (although a bit outdated) bibliography on digital sequences, that is on sequences based on a numeration system. More recent bibliographies may be found in [3] and [36], or in the very recent [37]. Moreover, Stolarsky [15, p. 719] notes: *Whatever its mathematical virtues, the literature on sums of digital sums reflects a lack of communication between researchers*, a sentence which is still topical. This was a motivation for us to provide the reader with an extended bibliography, with the hope that this field of research can get organized better and that the basic results about this topic will be known to all interested people.

1.3. Outline

The rough outline of the article is as follows. Section 2 is devoted to the precise definition of radix rational sequences. In Sections 3 and 4, we provide the reader with an algorithmic method of easy use in most cases to determine the asymptotic behavior of radix-rational sequences. The used tools all come from linear algebra and this is the distinctive feature of the method. The method has already been exposed in [38], but for the sake of completeness a self-contained exposition is given. This results in Theorem 1, which states that a radix-rational sequence admits an asymptotic expansion with variable coefficients in the scale $N^\alpha \binom{\log N}{m}$, or equivalently $N^\alpha \log^m N$. The so-called *dichopile algorithm*, briefly recalled in Example 1 below, will be a thread in this paper. The study of this algorithm shows that our theory applies to settings with a practical relevance. The asymptotic behavior of its complexity is given by

$$f_N \underset{N\to+\infty}{=} \frac{1}{2} N \log_2 N + N\varphi(\log_2 N) + O\left(N^\varepsilon\right),$$

where $\varphi(t)$ is a 1-periodic function and $0 < \varepsilon < 1$. The function $\varphi(t)$ is not explicit, but it is completely defined by a system of functional equations and can be exactly computed for a dense set of real numbers t, namely the useful values $\log_2 N \bmod 1$. We present in Section 5 some remarks and open issues about the method and we summarize it into an algorithm. This ends the first part of the article. Next in Section 6, we compare our algebraic approach and the analytic number theory approach, popularized notably by Philippe Flajolet [23,24]. We pursue the comparison in Section 7, where we elaborate on the computation of the Fourier coefficients of the involved periodic functions. We review the classical method and we design two new methods, more or less sophisticated, whose merit is their generality.

2. Radix-rational sequences

Radix-rational sequences are a mere generalization of classical rational sequences, that is sequences which satisfy a linear homogeneous recurrence with constant coefficients. In the same manner they satisfy a linear homogeneous recurrence with constant coefficients, but where the shift $n \mapsto n+1$ is replaced by the pair of scaling transformations $n \mapsto 2n$ and $n \mapsto 2n+1$, in case the radix is 2. It is the merit of Allouche and Shallit [29,39] to have put all various scattered examples in a common framework, that leads to the idea of a linear representation with matrices. For the examples given in the introduction the matrix dimension d is usually small, say from 1 to 4, but there exist examples where d is much larger, like in the work of Cassaigne [40] for which $d = 30$. For such intricate cases a general method of study is necessary and we will elaborate one. But, it is first necessary to formalize the notion of a radix-rational sequence.

2.1. Forward direction

Let S be a rational formal series over a finite alphabet $\mathcal{X}$. For the sake of simplicity, we are assuming here and in all the sequel that the coefficients of the formal series are complex numbers, even if it is possible to consider in this first section the more general framework of a commutative field or even of a commutative semi-ring. If $\mathcal{X}$ is the set of digits for the numeration system with radix B, the rational series defines a sequence s_n in the following manner: for each non-negative integer n, we consider the word $w = n_{\ell-1} \cdots n_1 n_0$ which is the radix B expansion of n, that is

$$n = (n_{\ell-1} \cdots n_1 n_0)_B = n_0 + n_1 B + \cdots + n_{\ell-1} B^{\ell-1}, \tag{1}$$

and the term s_n of the sequence is the value (S, w) of the rational series S over the word w.

More concretely, if a linear representation of the rational series S is at our disposal, we can compute the value (S, w) for a word $w = w_1 \cdots w_\ell$ by the formula

$$(S, w) = \boldsymbol{\ell} \boldsymbol{A}_{w_1} \boldsymbol{A}_{w_2} \cdots \boldsymbol{A}_{w_\ell} \boldsymbol{c} = \boldsymbol{\ell} \boldsymbol{A}_w \boldsymbol{c}. \tag{2}$$

Here, the triple $\boldsymbol{\ell}$, $(\boldsymbol{A}_b)_{0 \le b < B}$, $\boldsymbol{c}$ is a linear representation of S. This means that $\boldsymbol{\ell}$ is a row vector, the matrices $\boldsymbol{A}_b$ are square matrices, and $\boldsymbol{c}$ is a column vector, all with a coherent size d, which is the dimension of the representation. Formula (2) translates immediately into

$$s_n = \boldsymbol{\ell} \boldsymbol{A}_{n_{\ell-1}} \cdots \boldsymbol{A}_{n_1} \boldsymbol{A}_{n_0} \boldsymbol{c} \tag{3}$$

and gives a way to compute the successive values of the sequence s associated with the rational series S.

The rational character of the formal series S has the following meaning [41,42]: there exists a vector space of finite dimension which contains the series S and which is left stable by the trimming operators T_r, with $0 \le r < B$, defined by (the digit r is viewed as a letter)

$$T_r S = \sum_{w = vr} (S, w) v.$$

The operator T_r extracts the part of the formal series associated with the words which end with r and trims the letter r. These operators translate immediately, via the numeration system, into the multisection operators $T_r s_n = s_{Bn+r}$. The correspondence is clearer if we use the ordinary generating function of the sequence s_n, which parallels the formal series S,

$$T_r \sum_{n=0}^{+\infty} s_n z^n = \sum_{n=0}^{+\infty} s_{Bn+r} z^n. \tag{4}$$

With this viewpoint, the linear representation of the formal series S is interpreted as follows for the sequence s_n: we have a finite-dimensional vector space $\mathcal{V}$ of sequences, equipped with a basis $(v^1, v^2, \ldots, v^d)$ (or more generally a generating system), which is left stable by the multisection operators and contains the sequence s. The column vector $\boldsymbol{c}$ gives the coordinates of the sequence s with respect to the basis $(v^i)_{1 \le i \le d}$; the matrix $\boldsymbol{A}_b$ is the expression of the multisection operator T_b; and the row vector $\boldsymbol{\ell}$ appears to be the linear form $u \mapsto u_0$, which evaluates the sequences for $n = 0$. As an example, with $B = 2$, $n = 5 = (101)_2$, the column vectors $\boldsymbol{c}$, $\boldsymbol{A}_1 \boldsymbol{c}$, $\boldsymbol{A}_0 \boldsymbol{A}_1 \boldsymbol{c}$, and $\boldsymbol{A}_1 \boldsymbol{A}_0 \boldsymbol{A}_1 \boldsymbol{c}$ are the expressions of the sequences s_n, s_{2n+1}, s_{4n+1}, and s_{8n+5} respectively. Taking the value for $n = 0$ of the last sequence provides the value s_5 and this explains the formula $s_5 = \boldsymbol{\ell} \boldsymbol{A}_1 \boldsymbol{A}_0 \boldsymbol{A}_1 \boldsymbol{c}$. A more meaningful wording for the linear representation could be: $\boldsymbol{\ell}$ is the initial values row-vector, $\boldsymbol{A}$ is the action, $\boldsymbol{c}$ is the coordinates vector.

We are led to the following definition, but with a slight change with respect to [29,39]. We prefer the attributive adjective *rational* to *regular*, because *rational* is understood by both computer scientists and mathematicians, while *regular* has a precise meaning for the former but remains hazy for the latter.

Definition 1. A (complex) sequence is said to be *rational* with respect to a numeration system, or more briefly *radix-rational*, or more precisely B-*rational* if its orbit under the action of the B-multisection operators lies in a finite dimensional vector space.

2.2. Reverse direction

Let us assume now that we have a sequence $s = (s_n)$ which is B-rational. We can build a linear representation of this sequence in the following way. If the sequence is the null sequence there is nothing to do, and we consider that the associated dimension d is 0. Otherwise we put $v^1 = s$ as the first vector of the basis we want to build. Next we consider successively the sequences $(s_{Bn}), (s_{Bn+1}), \ldots, (s_{Bn+B-1}), (s_{B^2n}), (s_{B^2n+1})$, and so on. Each of these sequences is either independent of the previous ones and we add it to the basis, say as v^k, or it can be expressed as a linear combination of the previous sequences that are already in the basis. The process necessarily stops since all these sequences live in a finite dimensional space. When it stops, we have a free family $(v^1, v^2, \ldots, v^d)$ which is a basis of a vector space $\mathcal{V}$. This vector space is left stable by the multisection operators and contains the sequence s. Moreover, we have a linear representation of the sequence. First the sequence s is the first vector of the basis, and this gives the vector column $\boldsymbol{c} = \boldsymbol{e}_1$, say. Next, we have noted in passing the dependence encountered in the process, between each of the sequence v^j and its images by the multisection operators $T_0 v^j, T_1 v^j, \ldots, T_{B-1} v^j$, and this gives us the matrices $\boldsymbol{A}_0, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_{B-1}$. And last the values $v_0^1, \ldots, v_0^d$ at 0 give the row vector $\boldsymbol{\ell}$.

By this process, we have associated with the B-rational sequence a rational formal series over the alphabet of the radix-B numeration system. Moreover, we see that the classical rational sequences, like the Fibonacci sequence, that is the sequences whose ordinary generating function is a rational function (for which 0 is not a pole), appear to be 1-rational because it is well known that such a sequence writes $s_n = \boldsymbol{\ell} A^n \boldsymbol{c}$ for some square matrix A. Hence, the notion of B-rational sequences generalizes the classical one of rational sequences.

2.3. Sensitivity to the leftmost zeroes

The above construction is not entirely satisfactory, for if a rational series determines a radix rational sequence, the converse is not true. The reason is that the sequence does not use the words which begin with some zeroes. (The B-ary word associated with the number 0 is the empty word.) A first idea to fill this gap is to extend the formal series by the value 0 for words that begin with zero. A more preferable method is to follow the process of the previous paragraph. It provides sequences v^j which satisfy obviously $(T_0 v^j)_0 = v_0^j$, that is, in one formula, $\boldsymbol{\ell}\boldsymbol{A}_0 = \boldsymbol{\ell}$. This property determines completely the formal series if its values for the B-ary expansions of the integers are known.

Definition 2. A linear representation $\boldsymbol{\ell}$, $(\boldsymbol{A}_b)_{0 \le b < B}$, $\boldsymbol{c}$ of a B-rational sequence is said to be insensitive to the leftmost zeroes, or simply zero-insensitive, if it satisfies $\boldsymbol{\ell}\boldsymbol{A}_0 = \boldsymbol{\ell}$.

This definition implies two properties. First, a radix-rational sequence always admits a linear representation which is insensitive to the leftmost zeroes, which is proved above. Second, because all the minimal linear representations of a rational formal series are isomorphic, the same property is satisfied for the radix-rational sequence if we use only zero-insensitive linear representations.

Example 1 *(Dichopile algorithm).* The *dichopile algorithm* is an attempt to find a balance between space and time in the random generation of words from a regular language [43,44]. To generate uniformly a word of length n the previously known methods use $O(n)$ in space and $O(n)$ in time, while the dichopile algorithm uses only $O(\log n)$ in space (a big saving) but $O(n \log n)$ in time (a small loss), as we will see. The time complexity is given by the sequence f_n defined as follows [43, p. 45],

$$\begin{aligned} f_n &= n + g_n, \\ g_n &= f_{\lfloor n/2 \rfloor - 1} + g_{\lceil n/2 \rceil}, \quad \text{for } n \ge 2, \end{aligned} \tag{5}$$

with $f_1 = 1$, $g_1 = 0$. We add $f_0 = 0$, $g_0 = 0$. Both sequences f_n and g_n are rational with respect to the radix $B = 2$.

As it will soon become clear, we are not interested in a linear representation for the sequence f_n but for the sequence of backward differences $u_n = \nabla f_n = f_n - f_{n-1}$ (with $f_{-1} = 0$), which is 2-rational too according to Lemma 1 below. A linear representation of dimension 6 is the following

$$\boldsymbol{\ell} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T,$$

$$\boldsymbol{A}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{A}_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6}$$

It uses the generating family $u_n = \nabla f_n = 1 + \nabla g_n$, ∇f_{n-1}, ∇g_{n+1}, ∇g_n, 1, and $1 - \delta_{0,n}$ (here $\delta_{x,y}$ is the usual Kronecker symbol). As a consequence it is not a reduced linear representation, because the first sequence is the sum of the fourth and the fifth. Nevertheless it is perfectly acceptable and, moreover, it is insensitive to the leftmost zeroes.

From a theoretical standpoint, we always can assume that we use a zero-insensitive linear representation.

2.4. Backward differences

As it is pointed out in Example 1, we do not use a linear representation of the sequence s_n we want to estimate asymptotically, but a linear representation of the sequence $u_n = \nabla s_n = s_n - s_{n-1}$ of its backward differences. The next assertion shows that the latter is radix-rational if the former is radix-rational.

Lemma 1. *If a sequence s_n is B-rational, then the sequence of its backward differences is B-rational too. Conversely, if a sequence u_n is B-rational, then the sequence of its partial sums s_n is B-rational too.*

Proof. We could give a proof based on linear representations as in [38, Lemma 1], but this is useless because we will not employ the change from one representation to the other. It is simpler to consider the ordinary generating functions

$$u(x) = \sum_{n=0}^{+\infty} u_n x^n, \qquad s(x) = \sum_{n=0}^{+\infty} s_n x^n,$$

which are related by

$$u(x) = (1-x)s(x), \qquad s(x) = \frac{1}{1-x} u(x).$$

Allouche and Shallit have proved that the product of two B-rational generating functions is B-rational, hence the title *The* ring *of k-regular sequences* of their article [29]. Moreover $1 - x$ is B-rational as a polynomial, and $1/(1 - x)$ is B-rational as a rational function whose poles are all roots of unity [29, Th. 3.3]. Accordingly, if $s(x)$ is B-rational, so is $u(x)$, and vice versa. □

3. The linear machine

3.1. Vector recurrences

Our aim is to describe a computational method for deriving an asymptotic expansion of a radix rational sequence s_n. It takes as input a linear representation of the sequence of backward differences $u_n = \nabla s_n = s_n - s_{n-1}$ (with $s_{-1} = 0$). Its output is an asymptotic expansion for the sequence s_n, which writes as a finite linear combination of terms of the form

$$n^\alpha \binom{\log_B n}{m} \psi(\log_B n). \tag{7}$$

Here α is a real number, B is the radix of the numeration system, m is a nonnegative integer, and $\psi(t)$ is an oscillatory function often 1-periodic in practice.

In the rest of the article, we will constantly use the following notations. Let us stress upon one point: for what concerns the theory, the assumption that the linear representation is insensitive to the leftmost zeroes does not restrict the generality.

Notations. Row vectors and column vectors are denoted by bold lowercase letters and matrices are denoted by bold uppercase letters.

Throughout, $\boldsymbol{\ell}$, $(\boldsymbol{A}_b)_{0\le b<B}$, $\boldsymbol{c}$ is a zero-insensitive linear representation for $u_n = \nabla s_n$, which is d-dimensional. The matrix $\mathbf{I}_d$ is the identity matrix of size d. The matrix $\boldsymbol{Q}$ is the sum of the square matrices in the representation,

$$\boldsymbol{Q} = \sum_{0\le b<B} \boldsymbol{A}_b. \tag{8}$$

The integer K is the integer part of the logarithm to base B of N, and t is its fractional part,

$$K = \lfloor \log_B N \rfloor, \qquad t = \log_B(N) - K = \{\log_B N\}. \tag{9}$$

At last $(0.w)_B$ is the real number belonging to $[0, 1)$, whose B-ary expansion reads $0.w$, and $(w)_B$ is the integer whose B-ary expansion reads w.

The main idea is to link the partial sum of u_n up to N, that is s_N, with the sum of the rational series S behind u_n for some words of a given length.

Lemma 2. *Both sums*

$$s_N = \sum_{n=0}^{N} u_n, \quad \text{and} \quad S_K(x) = \sum_{\substack{|w|=K \\ (0.w)_B \le x}} (S, w), \tag{10}$$

are related by

$$s_N = S_{K+1}\left(B^{t-1}\right). \tag{11}$$

Proof. Formula (11) comes from cutting the interval $[0, N]$ by the powers of B. The integers in the interval $[B^k, B^{k+1})$ have a B-ary expansion which is a length $k+1$ word. The sum over all length $k+1$ words expresses with the matrix $\boldsymbol{Q}$, as follows

$$\sum_{|w|=k+1} \boldsymbol{\ell} \boldsymbol{A}_w \boldsymbol{c} = \sum_{|w|=k+1} \boldsymbol{\ell} \boldsymbol{A}_{w_1} \cdots \boldsymbol{A}_{w_{k+1}} \boldsymbol{c} = \boldsymbol{\ell}(\boldsymbol{A}_0 + \cdots + \boldsymbol{A}_{B-1})^{k+1} \boldsymbol{c} = \boldsymbol{\ell} \boldsymbol{Q}^{k+1} \boldsymbol{c}.$$

This leads to the formula

$$s_N = \boldsymbol{\ell}(\mathbf{I}_d - \boldsymbol{A}_0) \sum_{k=0}^{K} \boldsymbol{Q}^k \boldsymbol{c} + S_{K+1}\left(B^{t-1}\right). \tag{12}$$

More precisely, the B-ary expansions of the integers in $[B^k, B^{k+1})$ do not begin with a 0, hence we have to subtract $\boldsymbol{\ell} \boldsymbol{A}_0 \boldsymbol{Q}^k \boldsymbol{c}$ from $\boldsymbol{\ell} \boldsymbol{Q}^{k+1} \boldsymbol{c}$. This explains the first term in (12). The second term $S_{K+1}(B^{t-1})$ corresponds to the interval $[B^K, N]$. With the notations (9), the integer N writes $N = B^{K+t} = B^{K+1} B^{t-1}$ and for a length $K+1$ word w the inequality $(w)_B \le N$ is equivalent to $(0.w)_B \le B^{t-1}$, hence the occurrence of the term $S_{K+1}(B^{t-1})$. At this point, we take advantage of our assumption that we consider only zero-insensitive linear representations and Formula (12) simplifies into (11). $\square$

It will turn out that, for our purposes, it is not sufficient to consider the scalar-valued sum $S_K(x)$, and we reinforce the notations with a vector-valued sum.

Notations. The sum $S_K(x)$ writes $\boldsymbol{\ell} \boldsymbol{s}_K(x)$, where $\boldsymbol{s}_K(x)$ is the vector-valued sum associated with $\boldsymbol{c}$

$$\boldsymbol{s}_K(x) = \sum_{\substack{|w|=K \\ (0.w)_B \le x}} \boldsymbol{A}_w \boldsymbol{c}. \tag{13}$$

It is understood in the notation above that $\boldsymbol{s}_K(x)$ depends on the vector $\boldsymbol{c}$. According to our needs, we will change the vector $\boldsymbol{c}$ into another vector. We are interested by the behavior of $\boldsymbol{s}_K(x)$ for large values of K. Our starting point is the next formula.

Lemma 3. *The sum $\boldsymbol{s}_K(x)$ satisfies the recurrence*

$$\boldsymbol{s}_{K+1}(x) = \sum_{b_1 < x_1} \boldsymbol{A}_{b_1} \boldsymbol{Q}^K \boldsymbol{c} + \boldsymbol{A}_{x_1} \boldsymbol{s}_K(Bx - x_1), \tag{14}$$

valid for each $x = (0.x_1 x_2 \ldots)_B$ in $[0, 1)$.

Proof. The numbers from $[0, x]$ with a B-mantissa of length $K+1$ are sorted according to the prefixes of their B-ary expansions. This emphasizes the intervals $[0, (0.x_1)_B)$, $[(0.x_1)_B, (0.x_1 x_2)_B), \ldots [(0.x_1 x_2 \ldots x_K)_B, (0.x_1 x_2 \ldots x_{K+1})_B]$, and leads to the decomposition

$$\boldsymbol{s}_{K+1}(x) = \sum_{b_1 < x_1} \boldsymbol{A}_{b_1} \boldsymbol{Q}^K \boldsymbol{c} + \sum_{b_2 < x_2} \boldsymbol{A}_{x_1} \boldsymbol{A}_{b_2} \boldsymbol{Q}^{K-1} \boldsymbol{c} + \cdots + \sum_{b_{K+1} \le x_{K+1}} \boldsymbol{A}_{x_1} \boldsymbol{A}_{x_2} \cdots \boldsymbol{A}_{b_{K+1}} \boldsymbol{c}.$$

The above recursive Formula (14) is a mere consequence of this decomposition. $\square$

We will decompose the column vector $\boldsymbol{c}$ of the linear representation on a basis of generalized eigenvectors for the matrix $\boldsymbol{Q}$. For such vectors Formula (14) will reveal an associated functional equation, known as a dilation equation. But this will be possible only for sufficiently large eigenvalues, and we will start by explaining what 'sufficiently large' means in the context.

3.2. Joint spectral radius

The computation of a rational series uses products of square matrices of an arbitrary length. We need to evaluate the size of such products. We consider all products $\boldsymbol{A}_w$ for words w of a given length T and their norms, for some subordinate norm. Then the joint spectral radius of the set of matrices $\boldsymbol{A}_b$, $0 \le b < B$, is defined as [45,46]

$$\rho_* = \lim_{T \to +\infty} \max_{|w|=T} \|\boldsymbol{A}_w\|^{1/T}. \tag{15}$$

It is independent of the used subordinate norm. Moreover, the maximum occurring on the right-hand side of (15) is a nonincreasing function of T. More information can be found in [47] and the papers which come with it.

We apply immediately this definition to the study of the error term in the asymptotic expansion.

Lemma 4. *Let $\boldsymbol{v}$ be a generalized eigenvector of $\boldsymbol{Q}$ associated with the eigenvalue $\rho\omega$, $\rho \ge 0$, $|\omega| = 1$, which satisfies $\rho \le \rho_*$. Then for every $r > \rho_*$ and for every subordinate norm, the sequence $\boldsymbol{s}_K$ associated with $\boldsymbol{v}$,*

$$\boldsymbol{s}_K(x) = \sum_{\substack{|w|=K \\ (0.w)_B \le x}} \boldsymbol{A}_w \boldsymbol{v},$$

satisfies $\|\boldsymbol{s}_K(x)\| = O(r^K)$ uniformly with respect to x.

Proof. First, let us consider simply the case of an eigenvector and assume $\|\boldsymbol{A}_b\| \le r$ for $0 \le b < B$. With Formula (14) of Lemma 3, we readily obtain

$$\big\|\boldsymbol{s}_{K+1}(x)\big\| \le \rho^K \sum_{b_1 < x_1} \|\boldsymbol{A}_{b_1}\| \|\boldsymbol{v}\| + \|\boldsymbol{A}_{x_1}\| \big\|\boldsymbol{s}_K(Bx - x_1)\big\| \tag{16}$$

and, using the supremum norm,

$$\|\boldsymbol{s}_{K+1}\|_\infty \le \rho^K B r \|\boldsymbol{v}\| + r\|\boldsymbol{s}_K\|_\infty. \tag{17}$$

With $\rho \le \rho_* < r$, this recursive inequality leads to $\|\boldsymbol{s}_K\|_\infty = O(r^K)$.

Under the same assumption $\|\boldsymbol{A}_b\| \le r$ for $0 \le b < B$, if $\boldsymbol{v}$ is a generalized eigenvector for $\boldsymbol{Q}$, it satisfies $\boldsymbol{Q}^K \boldsymbol{v} = O(\rho^K K^{\nu-1})$ for some positive integer ν. Eq. (17) becomes $\|\boldsymbol{s}_{K+1}\|_\infty \le C\rho^K K^{\nu-1} Br\|\boldsymbol{v}\| + r\|\boldsymbol{s}_K\|_\infty$ for some constant C. The same conclusion follows.

In the general case, there exists a positive integer T such that all words of length T satisfy $\|\boldsymbol{A}_w\| \le r^T$. We change the basis of numeration B into $B' = B^T$. Equivalently, we consider the linear representation $\boldsymbol{\ell}$, $(\boldsymbol{A}_w)_{|w|=T}$, $\boldsymbol{c}$. The previous case applies and we obtain $\|\boldsymbol{s}_{KT}\|_\infty = O(r^{KT})$. Since there is only a finite number of integers s between 0 and T, we can extend the result into $\|\boldsymbol{s}_{KT+s}\|_\infty = O(r^{KT+s})$ for $0 \le s < T$ by using Formula (16). This provides us with the expected formula. □

Example 2 *(Dichopile algorithm, continued).* Here, we are using the maximum absolute column sum $\|\ \|_1$. Because $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ have nonnegative coefficients, computing $\|\boldsymbol{A}_w\|_1$ amounts to computing the row vector $\boldsymbol{u}\boldsymbol{A}_w$ and to taking the largest component $\|\boldsymbol{u}\boldsymbol{A}_w\|_\infty$, where $\boldsymbol{u}$ is the row vector with all components equal to 1. It turns out that all these products $\boldsymbol{u}\boldsymbol{A}_w$ write

$$\boldsymbol{v} = \begin{pmatrix} a & b & c & d & 1 & 1 \end{pmatrix}$$

because this is true for the empty word, and this writing is preserved by the product by $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$,

$$\boldsymbol{v}\boldsymbol{A}_0 = \begin{pmatrix} 1+b & 1+d & c & b & 1 & 1 \end{pmatrix}, \qquad \boldsymbol{v}\boldsymbol{A}_1 = \begin{pmatrix} 1+c & 1+b & a & c & 1 & 1 \end{pmatrix}$$

The formulæ above show that all the components for a word of length T are bounded by $T+1$. Moreover if $\boldsymbol{u}\boldsymbol{A}_1^T$ (here T is an exponent) writes $[a, b, c, d, 1, 1]$ where b is larger than all the other components, then $\boldsymbol{u}\boldsymbol{A}_1^{T+1}$ writes $[1+c, 1+b, a, c, 1, 1]$ and $1+b$ is the largest component, so that $\|\boldsymbol{u}\boldsymbol{A}_1^T\|_\infty = T+1$. We thus obtain

$$\max_{|w|=T} \|\boldsymbol{A}_w\|_1 = T+1.$$

As a consequence we conclude that $\rho_* = 1$, and that the error term will be $O(r^K)$ for every $r > 1$.

3.3. Dilation equations

Assume for a while that we take for $\boldsymbol{c}$ an eigenvector $\boldsymbol{v}$ of $\boldsymbol{Q}$ associated with the eigenvalue $\rho\omega$, with $\rho > 0$ and $|\omega| = 1$. With $\boldsymbol{f}_K(x) = \boldsymbol{s}_K(x)/(\rho\omega)^K$, Formula (14) of Lemma 3 becomes

$$\rho\omega\boldsymbol{f}_{K+1}(x) = \sum_{b_1<x_1} \boldsymbol{A}_{b_1}\boldsymbol{v} + \boldsymbol{A}_{x_1}\boldsymbol{f}_K(Bx - x_1).$$

If the sequence $\boldsymbol{f}_K(x)$ has a limit $\boldsymbol{f}(x)$, the previous equation gives in the limit

$$\rho\omega\boldsymbol{f}(x) = \sum_{b_1<x_1} \boldsymbol{A}_{b_1}\boldsymbol{v} + \boldsymbol{A}_{x_1}\boldsymbol{f}(Bx - x_1).$$

According to Definition (13), we must have $\boldsymbol{f}(1) = \boldsymbol{v}$, because of $\boldsymbol{s}_K(1) = \sum_{|w|=K} \boldsymbol{A}_w\boldsymbol{v} = \boldsymbol{Q}^K\boldsymbol{v} = (\rho\omega)^K\boldsymbol{v}$.

More generally, if $\boldsymbol{Q}$ admits a generalized eigenvector associated with the eigenvalue $\rho\omega$, there exists a free family $\mathcal{V} = (\boldsymbol{v}^{(j)})_{0\le j<\nu}$ of vectors in $\mathbb{C}^d$, such that the subspace generated by $\mathcal{V}$ is left stable by $\boldsymbol{Q}$ and the matrix induced with respect to the basis $\mathcal{V}$ of this subspace is the usual Jordan cell with size ν

$$\boldsymbol{J} = \boldsymbol{J}_{\nu,\rho\omega} = \begin{pmatrix} \rho\omega & 1 & & & \\ & \rho\omega & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \rho\omega \end{pmatrix}.$$

The same idea as in the case of an eigenvector leads us to consider the functional equation

$$\boldsymbol{F}(x)\boldsymbol{J} = \sum_{b_1<x_1} \boldsymbol{A}_{b_1}\boldsymbol{V} + \boldsymbol{A}_{x_1}\boldsymbol{F}(Bx - x_1). \tag{18}$$

This time $\boldsymbol{V}$ and $\boldsymbol{F}(x)$ are matrix-valued, precisely their values are in $\mathbb{C}^{[1,d]\times[0,\nu)}$. The matrix $\boldsymbol{V}$ is made from the column vectors $\boldsymbol{v}^{(0)}, \dots, \boldsymbol{v}^{(\nu-1)}$ in that order. Likewise the matrix $\boldsymbol{F}$ is made from the vector-valued functions $\boldsymbol{f}^{(0)}(x), \dots, \boldsymbol{f}^{(\nu-1)}(x)$ in that order. The three matrices $\boldsymbol{Q}$, $\boldsymbol{J}$ and $\boldsymbol{V}$ are related by

$$\boldsymbol{Q}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{J}. \tag{19}$$

Anew we must have $\boldsymbol{F}(1) = \boldsymbol{V}$. As the functions $\boldsymbol{s}_K(x)$ are defined on the segment $[0, 1]$, we are searching for a solution of Eq. (18) on $[0, 1]$. But we extend the function $\boldsymbol{F}(x)$ by

$$\boldsymbol{F}(x) = \boldsymbol{0} \quad \text{for } x \le 0, \qquad \boldsymbol{F}(x) = \boldsymbol{V} \quad \text{for } x \ge 1. \tag{20}$$

With this choice, Eq. (18) rewrites as a homogeneous equation

$$\boldsymbol{F}(x)\boldsymbol{J} = \sum_{0\le b<B} \boldsymbol{A}_b\boldsymbol{F}(Bx - b). \tag{21}$$

This equation is a *dilation equation* or *two-scale difference equation*. Such equations are common in the definition of wavelets [48] and refinement schemes [49], and have been studied in detail [50, Th. 2.2], [51,52]. They also occur in the study of toy-examples of dynamical systems [53,54].

The dilation equation may admit solutions, or not. Anyway, we can use the algorithm known as the *cascade algorithm* [55, §6.5, p. 206], which enters into the family of *interpolatory schemes* [56]. Let us assume for a while that the binary system is used. Then the boundary conditions (20) and the dilation equation (21) completely determine a function $\boldsymbol{F}$ at dyadic numbers in $[0, 1]$ by the cascade algorithm. First, with $\boldsymbol{F}(0)$ and $\boldsymbol{F}(1)$ as input, the cascade algorithm computes $\boldsymbol{F}(1/2)$, next $\boldsymbol{F}(1/4)$ and $\boldsymbol{F}(3/4)$, and in the end all the values $\boldsymbol{F}(i/2^K)$ with $0 \le i \le 2^K$ up to some depth K. All pictures in this article are computed this way. The algorithm will appear as an argument for the uniqueness of the solution in the proof of the next lemma and we will use it in the next example to best determine the solution of the involved dilation equation.

Lemma 5. *For an eigenvalue of $\boldsymbol{Q}$ with modulus larger than the joint spectral radius of the linear representation, Eq. (21) with boundary conditions (20) has a unique continuous solution.*

Proof. The existence of a continuous solution results from the Banach fixed-point theorem [48, §4]. More concretely, there is a power of the fixed-point operator

$$\mathcal{L}\boldsymbol{\Phi}(x) = \sum_{0\le b<B} \boldsymbol{A}_b\boldsymbol{\Phi}(Bx - b)\boldsymbol{J}^{-1}, \tag{22}$$

that is a contraction mapping. In order that such an assertion makes sense, we have to choose a norm on the space $\mathcal{C} = \mathcal{C}(\mathbb{R}, \mathbb{C}^{[1,d]\times[0,\nu)})$ of continuous functions from $\mathbb{R}$ into the space of rectangular matrices $\mathbb{C}^{[1,d]\times[0,\nu)}$.

According to the assumption of the lemma, we have $\rho/\rho_* > 1$ where ρ is the modulus of the eigenvalue under consideration. So that we can choose a positive real number $\varepsilon > 0$ satisfying $\rho/\rho_* > (1+\varepsilon)^2$. With the above notations, the inverse matrix of the Jordan cell $\boldsymbol{J}$ has a spectral radius $1/\rho$. Then there exists a subordinate norm $\|\ \|$ satisfying $1/\rho \le \|\boldsymbol{J}^{-1}\| \le (1+\varepsilon)/\rho$. To this norm we associate the supremum norm $\|\ \|_\infty$ on the space $\mathcal{C}$. Moreover, by the definition of the joint spectral radius, there exists an integer T such that all matrices $\boldsymbol{A}_w$ associated with a word w of length T have their norm bounded by $\rho_*^T(1+\varepsilon)^T$. The T-th power of the operator $\mathcal{L}$ writes

$$\mathcal{L}^T\boldsymbol{\Phi}(x) = \sum_{|w|=T} \boldsymbol{A}_w \boldsymbol{\Phi}\big(B^T x - (w)_B\big)\boldsymbol{J}^{-T}. \tag{23}$$

Recall that $(w)_B$ is the integer whose B-ary expansion is equal to w. Eq. (23) is nothing but Eq. (21) for radix B^T in place of B. If $\boldsymbol{\Phi}_1(x)$ and $\boldsymbol{\Phi}_2(x)$ are two elements of $\mathcal{C}$, whose images by $\mathcal{L}^T$ are $\boldsymbol{\Psi}_1(x)$ and $\boldsymbol{\Psi}_2(x)$, then their respective differences $\boldsymbol{\Delta}(x) = \boldsymbol{\Phi}_2(x) - \boldsymbol{\Phi}_1(x)$ and $\mathcal{L}^T\boldsymbol{\Delta}(x) = \boldsymbol{\Psi}_2(x) - \boldsymbol{\Psi}_1(x)$ are related by

$$\mathcal{L}^T\boldsymbol{\Delta}(x) = \boldsymbol{A}_{x_1\ldots x_T}\boldsymbol{\Delta}\big((0.x_{T+1}\ldots)_B\big)\boldsymbol{J}^{-T},$$

where $0.x_1\ldots x_T x_{T+1}\ldots$ is the B-ary expansion of x. Consequently, the supremum norm of $\mathcal{L}^T\boldsymbol{\Delta}$ is bounded by

$$\big\|\mathcal{L}^T\boldsymbol{\Delta}\big\|_\infty \le \rho_*^T(1+\varepsilon)^T\|\boldsymbol{\Delta}\|_\infty\frac{(1+\varepsilon)^T}{\rho^T}.$$

Using the inequality $\rho/\rho_* > (1+\varepsilon)^2$, we see that $\mathcal{L}^T$ is a contraction operator on the space $\mathcal{C} = \mathcal{C}(\mathbb{R}, \mathbb{C}^{[1,d]\times[0,\nu)})$.

The continuous functions $\boldsymbol{\Phi}(x)$ in $\mathcal{C}$ that satisfy the boundary conditions $\boldsymbol{\Phi}(0) = \boldsymbol{0}$ and $\boldsymbol{\Phi}(1) = \boldsymbol{V}$ are the elements of a closed subspace $\mathcal{B}$ of $\mathcal{C}$. Since $\mathcal{C}$ is known to be complete, $\mathcal{B}$ is itself complete. To apply the fixed-point theorem, it remains to show that the subspace $\mathcal{B}$ is left stable by the operator $\mathcal{L}^T$. As a matter of fact, it is left stable by the operator $\mathcal{L}$. According to Formula (22), the image $\mathcal{L}\boldsymbol{\Phi}(x)$ is continuous if $\boldsymbol{\Phi}(x)$ is continuous. Moreover, again with (22), we have $\mathcal{L}\boldsymbol{\Phi}(0) = \sum_{0\le b<B}\boldsymbol{A}_b \times \boldsymbol{0} \times \boldsymbol{J}^{-1} = \boldsymbol{0}$ and $\mathcal{L}\boldsymbol{\Phi}(1) = \sum_{0\le b<B}\boldsymbol{A}_b \times \boldsymbol{V} \times \boldsymbol{J}^{-1} = \boldsymbol{Q}\,\boldsymbol{V}\,\boldsymbol{J}^{-1} = \boldsymbol{V}$, according to (19). We obtain a partial conclusion: Eq. (23) has a unique solution in $\mathcal{B}$, that we denote by $\boldsymbol{F}_0$.

Let us consider the sequence of functions in $\mathcal{B}$ defined by $\boldsymbol{F}_k = \mathcal{L}^k\boldsymbol{F}_0$. By definition of $\boldsymbol{F}_0$ we have $\boldsymbol{F}_0 = \mathcal{L}^T\boldsymbol{F}_0$, hence the sequence is T-periodic: for every k, we have $\boldsymbol{F}_k = \mathcal{L}^T\boldsymbol{F}_k$. But the fixed-point theorem says that the fixed-point equation $\boldsymbol{\Phi} = \mathcal{L}^T\boldsymbol{\Phi}$ has a unique solution in $\mathcal{B}$, which is $\boldsymbol{F}_0$. So the sequence is constant. In particular, we have $\boldsymbol{F}_0 = \boldsymbol{F}_1$, that is $\boldsymbol{F}_0 = \mathcal{L}\boldsymbol{F}_0$. This proves the existence of a solution for Eq. (21) with boundary conditions (20). Besides, the solution is continuous and the cascade algorithm determines the solution on all the numbers j/B^K, with $0 \le K$, $0 \le j \le B^K$, that form a dense set of real numbers in $[0,1]$, hence the uniqueness of the solution. □

Let us recall that a function f from the real line into a normed space is Hölder continuous with exponent α if it satisfies $\|f(y) - f(x)\| \le c|y-x|^\alpha$ for some constant c. The following result is common in the study of dilation equations and a proof can be found in [50, p. 1039].

Lemma 6. *For an eigenvalue of $\boldsymbol{Q}$ with modulus $\rho > \rho_*$, the solution of Eq. (21) is Hölder continuous with exponent $\log_B(\rho/r)$ for every $r > \rho_*$.*

Example 3 *(Dichopile algorithm, continued).* With the previous linear representation for the sequence $u_n = \nabla f_n$, the matrix

$$\boldsymbol{Q} = \boldsymbol{A}_0 + \boldsymbol{A}_1 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 2 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

admits the Jordan normal form $\boldsymbol{\Lambda} = \boldsymbol{P}^{-1}\boldsymbol{Q}\,\boldsymbol{P}$, with (the coefficient $1/12$ in $\boldsymbol{P}$ is useless at this stage, but permits us to later use linear combinations with integers coefficients)

$$\boldsymbol{\Lambda} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{P} = \frac{1}{12}\begin{pmatrix} 0 & 2 & 0 & 6 & -2 & -6 \\ 0 & 4 & 0 & -6 & 2 & 0 \\ 0 & 4 & 0 & 6 & 2 & 0 \\ 0 & 2 & 0 & -6 & -2 & 6 \\ 12 & -16 & 6 & 15 & 1 & 0 \\ 0 & 10 & -6 & -15 & -1 & 6 \end{pmatrix}. \tag{24}$$

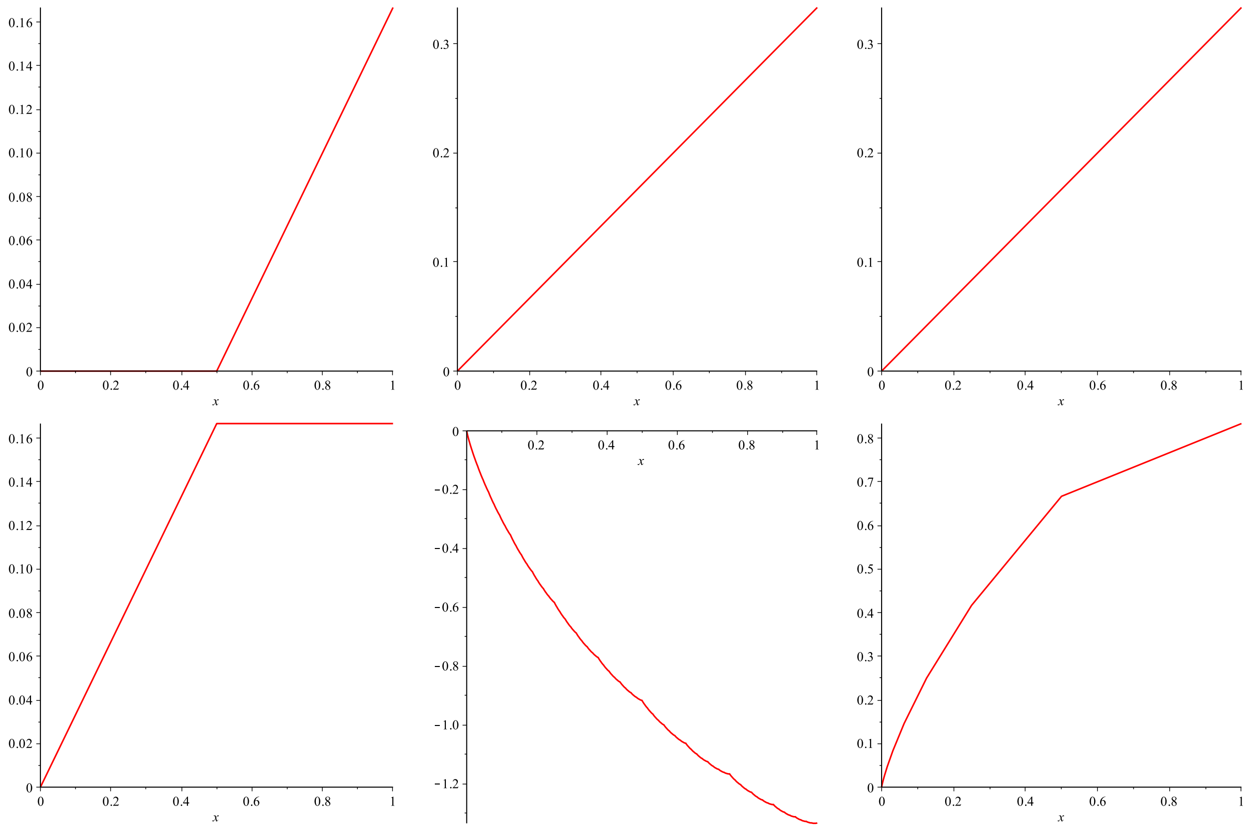

Fig. 1. The six components (from left to right, and from top to bottom) of the vector-valued function $\boldsymbol{g}(x) = \boldsymbol{f}^{(1)}(x)$ for the dilation equation associated with the dichopile algorithm.

Let us consider only the 2×2 Jordan block at the upper left corner of $\boldsymbol{\Lambda}$. We have to find two vector-valued functions $\boldsymbol{f}^{(0)}(x)$ and $\boldsymbol{f}^{(1)}(x)$, which we rename $\boldsymbol{f}(x)$ and $\boldsymbol{g}(x)$. The equation for $\boldsymbol{g}(x)$ writes

$$\begin{aligned}
g_1(x) &= \frac{1}{2} g_3(2x-1) - \frac{1}{2} f_1(x), \\
g_2(x) &= \frac{1}{2} g_1(2x) + \frac{1}{2} g_4(2x) + \frac{1}{2} g_2(2x-1) - \frac{1}{2} f_2(x), \\
g_3(x) &= \frac{1}{2} g_3(2x) + \frac{1}{2} g_1(2x-1) + \frac{1}{2} g_4(2x-1) - \frac{1}{2} f_3(x), \\
g_4(x) &= \frac{1}{2} g_2(2x) - \frac{1}{2} f_4(x), \\
g_5(x) &= \frac{1}{2} g_5(2x) + \frac{1}{2} g_1(2x-1) + \frac{1}{2} g_5(2x-1) + \frac{1}{2} g_6(2x-1) - \frac{1}{2} f_5(x), \\
g_6(x) &= \frac{1}{2} g_1(2x) + \frac{1}{2} g_2(2x) + \frac{1}{2} g_6(2x) + \frac{1}{2} g_2(2x-1) - \frac{1}{2} f_6(x).
\end{aligned}$$

It is non-homogeneous and needs the knowledge of $\boldsymbol{f}(x)$ to be solved. Besides, the equation for $\boldsymbol{f}(x)$ is the homogeneous version of the previous equation for $\boldsymbol{g}(x)$, obtained by suppressing the terms $f_i(x)$ and by replacing $g_i(x)$ by $f_i(x)$. To solve the equations we use the cascade algorithm. First, for $\boldsymbol{f}(x)$, we draw the graphs of the functions $f_i(x)$ and we obtain $f_5(x) = x$ and $f_i(x) = 0$ for the other indices i. (Obviously these expressions are valid only for x in $[0, 1]$.) Second, for $\boldsymbol{g}(x)$, we draw the pictures of Fig. 1, and we deduce explicit expressions for g_1, g_2, g_3, g_4, like

$$g_1(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1/2, \\ (x - 1/2)/3 & \text{if } 1/2 \le x \le 1. \end{cases}$$

Such empirical results are not a proof, but the system has a unique continuous solution and a check by mere substitution establishes the formulæ. Hence at this stage we know $\boldsymbol{f}(x)$ and the first four components of $\boldsymbol{g}(x)$. Next we find elementarily

$$g_6(x) = \begin{cases} x/3 + 1/2 & \text{if } 1/2 \le x \le 1, \\ (k+2)x/3 + 1/(3 \times 2^k) & \text{if } 1/2^{k+1} \le x \le 1/2^k \text{ with } k \ge 1. \end{cases}$$

However, we do not have an explicit expression for $g_5(x)$. We only know that it is well and completely defined by the dilation equation

$$g_5(x) = \frac{1}{2} g_5(2x) + \frac{1}{2} g_5(2x - 1) + h(x) \quad \text{for } 0 \le x \le 1,$$

where $h(x)$ is a known, but slightly complicated, piecewise affine function, with the boundary conditions $g_5(0) = 0$, $g_5(1) = -4/3$. Moreover, it is Hölder continuous with exponent $\log_2(2/r)$ for every $r > \rho_* = 1$, that is with exponent ε for every $\varepsilon < 1$.

4. Asymptotic expansion

4.1. Asymptotic expansion for the rational series

We proceed to a Jordan reduction of the matrix $\boldsymbol{Q}$, which is, let us call it to mind, the sum of the square matrices of the linear representation for the sequence u_n. Next we expand the column vector $\boldsymbol{c}$ over the Jordan basis, so that we have to consider the sums $\boldsymbol{s}_K(x)$ relating to each vector of the Jordan basis. That is, using the same notations as in the previous section, we have to consider

$$\boldsymbol{s}_K(x) = \sum_{\substack{|w|=K \\ (0.w)_B \le x}} \boldsymbol{A}_w \boldsymbol{v}^{(\nu-1)}. \tag{25}$$

When x goes from 0 to 1, the sum goes from $\boldsymbol{A}_0^K \boldsymbol{v}^{(\nu-1)}$ for $x = 0$ to the vector $\boldsymbol{Q}^K \boldsymbol{v}^{(\nu-1)}$ for $x = 1$. This last quantity evaluates to

$$\begin{aligned} \boldsymbol{Q}^K \boldsymbol{v}^{(\nu-1)} &= \binom{K}{\nu-1} (\rho\omega)^{K-\nu+1} \boldsymbol{v}^{(0)} + \binom{K}{\nu-2} (\rho\omega)^{K-\nu+2} \boldsymbol{v}^{(1)} + \cdots \\ &\quad + \binom{K}{1} (\rho\omega)^{K-1} \boldsymbol{v}^{(\nu-2)} + (\rho\omega)^K \boldsymbol{v}^{(\nu-1)}, \end{aligned} \tag{26}$$

with the help of the equality $\boldsymbol{Q}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{J}$ indicated in (19) (as in the previous section, matrix $\boldsymbol{V}$ has size $[1, d] \times [0, \nu)$ and its columns are $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(\nu-1)}$). It is not too astonishing that the intermediate values may be described by the following expression

$$\begin{aligned} \boldsymbol{a}_K(x) &= \binom{K}{\nu-1} (\rho\omega)^{K-\nu+1} \boldsymbol{f}^{(0)}(x) + \binom{K}{\nu-2} (\rho\omega)^{K-\nu+2} \boldsymbol{f}^{(1)}(x) + \cdots \\ &\quad + \binom{K}{1} (\rho\omega)^{K-1} \boldsymbol{f}^{(\nu-2)}(x) + (\rho\omega)^K \boldsymbol{f}^{(\nu-1)}(x), \end{aligned} \tag{27}$$

at least asymptotically, since the functions $\boldsymbol{f}^{(k)}(x)$ are obtained by a refinement scheme whose inputs are both values at the endpoints of the interval. The reader may notice that, in some sense, we are filling the gap of the usual studies of divide-and-conquer recurrences where only the powers of 2 are taken into account and the intermediate indices are neglected.

Lemma 7. *For a Jordan vector $\boldsymbol{v}^{(\nu-1)}$ associated with an eigenvalue $\rho\omega$, whose modulus ρ is larger than the joint spectral radius, that is $\rho > \rho_*$, the expression $\boldsymbol{a}_K(x)$ above is an asymptotic expansion for the associated sum $\boldsymbol{s}_K(x)$ defined in (25). The error term is $O(r^K)$ for every r between ρ and ρ_*, and it is uniform with respect to x.*

Proof. By substitution of $\boldsymbol{s}_K(x) = \boldsymbol{a}_K(x) + \boldsymbol{\delta}_K(x)$ into the basic recurrence Formula (14), we find $\boldsymbol{\delta}_{K+1}(x) = \boldsymbol{A}_{x_1} \boldsymbol{\delta}_K(Bx - x_1)$. More precisely, we use the recursion (14) of Lemma 3, the dilation equation (18), and the fundamental recursion of binomial coefficients. With the assumption $\rho > r > \rho_*$, we choose a length T such that all matrices $\boldsymbol{A}_w$ associated with a length T word w have a norm smaller than r. We readily obtain $\|\boldsymbol{\delta}_{KT}\|_\infty = O(r^{KT})$. Next the recursion (14) shows that we have $\|\boldsymbol{\delta}_{KT+s}\|_\infty = O(r^{KT+s})$ for $0 \le s < T$, hence the result. □

Example 4 *(Dichopile algorithm, continued).* Lemma 7 permits us to find an asymptotic expansion for $S_K(x)$ as follows. We expand the column vector $\boldsymbol{c}$ of the linear representation (6) onto the Jordan basis. We find

$$\boldsymbol{c} = \boldsymbol{v}_2^{(0)} + \boldsymbol{v}_2^{(1)} + \boldsymbol{v}_1^{(1)} - 2\boldsymbol{v}_{-1}^{(0)} - \boldsymbol{v}_0^{(0)},$$

with natural notations: $\boldsymbol{v}_\lambda^{(j)}$ is a generalized eigenvector for the eigenvalue λ satisfying $\boldsymbol{Q}\,\boldsymbol{v}_\lambda^{(j)} = \lambda\boldsymbol{v}_\lambda^{(j)} + \boldsymbol{v}_\lambda^{(j-1)}$. According to Lemma 4, we have to take into account only the 2×2 Jordan block at the upper left corner of $\boldsymbol{\Lambda}$ of Formula (24), and consequently only $\boldsymbol{v}_2^{(0)} + \boldsymbol{v}_2^{(1)}$.

We apply Lemma 7 and specifically Formula (27) to obtain the contribution of each involved generalized eigenvector. With the notations of Example 3, the contribution of $\boldsymbol{v}_2^{(0)}$ to the sum $\boldsymbol{s}_K(x)$ relating to $\boldsymbol{c}$ is $2^K\boldsymbol{f}^{(0)}(x) = 2^K\boldsymbol{f}(x)$. Similarly, the contribution of $\boldsymbol{v}_2^{(1)}$ is $2^{K-1}K\boldsymbol{f}^{(0)}(x) + 2^K\boldsymbol{f}^{(1)}(x) = 2^{K-1}K\boldsymbol{f}(x) + 2^K\boldsymbol{g}(x)$. So that we obtain

$$\boldsymbol{s}_K(x) \underset{K\to+\infty}{=} 2^{K-1}(K+2)\boldsymbol{f}(x) + 2^K\boldsymbol{g}(x) + O\left(r^K\right)$$

for $1 < r < 2$. Henceforth the sum $S_K(x) = \boldsymbol{\ell}\boldsymbol{s}_K(x)$, defined in Formula (10), satisfies

$$S_K(x) \underset{K\to+\infty}{=} 2^{K-1}(K+2)x + 2^K g_5(x) + O\left(r^K\right). \tag{28}$$

4.2. Asymptotic expansion for the radix-rational sequence

Eq. (11) in Lemma 2, which relates s_N and S_{K+1}, gives us a way to obtain the asymptotic expansion of the sequence s_N.

Theorem 1. *Let s_N be a radix-rational sequence, where the sequence of backward differences is defined by a linear representation $\boldsymbol{\ell}$, $(\boldsymbol{A}_b)_{0\le b<B}$, $\boldsymbol{c}$, insensitive to the leftmost zeroes. Then the sequence s_N admits an asymptotic expansion which is a sum of terms*

$$N^{\log_B\rho}\binom{\log_B N}{m} \times e^{i\vartheta\log_B N} \times \varphi(\log_B N), \tag{29}$$

indexed by the eigenvalues $\rho e^{i\vartheta}$ of $\boldsymbol{Q} = \boldsymbol{A}_0 + \cdots + \boldsymbol{A}_{B-1}$ whose modulus ρ is larger than the joint spectral radius ρ_* of the linear representation, with m a nonnegative integer and $\varphi(t)$ a 1-periodic function. The error term is $O(N^{\log_B r})$ for every $r > \rho_*$. Moreover the 1-periodic function $\varphi(t)$ is Hölder continuous with exponent $\log_B(\rho/r)$ for every r between ρ and ρ_*.

It should be noted that the assumption of insensitivity is only a convenience that simplifies the computation. See [38] for a general version.

Proof. The error term is $O(r^{\log_B N}) = O(N^{\log_B r})$ for every $r > \rho_*$ according to Lemma 7. A mere substitution in $S_{K+1}(x)$ (not in $S_K(x)$, however) translates the result about $S_K(x)$ into the expected result about s_N. First we write $\omega = e^{i\vartheta}$ in (27) with ϑ a real number. Next, for each term $\binom{K}{\ell}(\rho\omega)^{K-\ell}\boldsymbol{f}(x)$, we change K into $K+1$ and x into $B^{\{t\}-1}$. Moreover we rewrite $K = \lfloor\log_B N\rfloor$ as $\log_B N - \{t\}$ (implicitly $t = \log_B N$) and we obtain the regular part of an asymptotic expansion for s_N as a sum of terms

$$\binom{\log_B N + 1 - \{t\}}{\ell}\left(\rho e^{i\vartheta}\right)^{\log_B N - (\ell - 1 + \{t\})}\boldsymbol{\ell}\boldsymbol{f}\left(B^{\{t\}-1}\right).$$

Using the Chu–Vandermonde identity and the equality $\rho^{\log_B N} = N^{\log_B\rho}$, this rewrites

$$\begin{aligned}
&N^{\log_B\rho}e^{i\vartheta\log_B N}\rho^{-\ell+1-\{t\}}e^{-i\vartheta(\ell-1+\{t\})}\boldsymbol{\ell}\boldsymbol{f}\left(B^{\{t\}-1}\right) \times \sum_{m=0}^{\ell}\binom{\log_B N}{m}\binom{1-\{t\}}{\ell-m} \\
&\quad = \sum_{m=0}^{\ell} N^{\log_B\rho}\binom{\log_B N}{m} \times e^{i\vartheta\log_B N} \times \rho^{-\ell+1-\{t\}}e^{-i\vartheta(\ell-1+\{t\})}\boldsymbol{\ell}\boldsymbol{f}\left(B^{\{t\}-1}\right)\binom{1-\{t\}}{\ell-m}.
\end{aligned} \tag{30}$$

It turns out that the expansion is a linear combination of the more elementary terms (29).

The proof is complete, except the point about the Hölderian character of $\varphi(t)$. This property comes from Lemma 5, but the use of the fractional part $\{t\}$ needs a complement: we have to verify the continuity at integers. The point needs that we stand back a little from the computation and think that we are practicing linear algebra. The expression (27) for the term $\boldsymbol{a}_K(x)$ is not just the result of chance. It is the last column of the product $\boldsymbol{F}(x)\boldsymbol{J}^K$. The change from the rational series to the rational sequence makes us to consider $\boldsymbol{F}(B^{\{t\}-1})\boldsymbol{J}^{K+1} = \boldsymbol{\Phi}(t)\boldsymbol{J}^t$ with the 1-periodic part $\boldsymbol{\Phi}(t) = \boldsymbol{F}(B^{\{t\}-1})\boldsymbol{J}^{1-\{t\}}$. We compute the limits at the integers from above and from below $\boldsymbol{\Phi}(1^-) = \boldsymbol{F}(1) = \boldsymbol{V}$ and $\boldsymbol{\Phi}(0^+) = \boldsymbol{F}(1/B)\boldsymbol{J} = \boldsymbol{A}_0\boldsymbol{V}\boldsymbol{J}^{-1}\boldsymbol{J} = \boldsymbol{A}_0 V$. Using the insensitivity of the linear representation to the leftmost zeroes, we obtain $\boldsymbol{\ell}\boldsymbol{\Phi}(1^-) = \boldsymbol{\ell}\boldsymbol{\Phi}(0^+)$ and the expected continuity of $\varphi(t)$ which is the last component of the row vector $\boldsymbol{\ell}\boldsymbol{\Phi}(t)$. $\square$

Example 5 *(Dichopile algorithm, the end).* Applying the substitution explained above, we obtain

$$f_N \underset{N\to+\infty}{=} \frac{1}{2}N\log_2 N + N\varphi(\log_2 N) + O\left(N^{\varepsilon}\right), \tag{31}$$

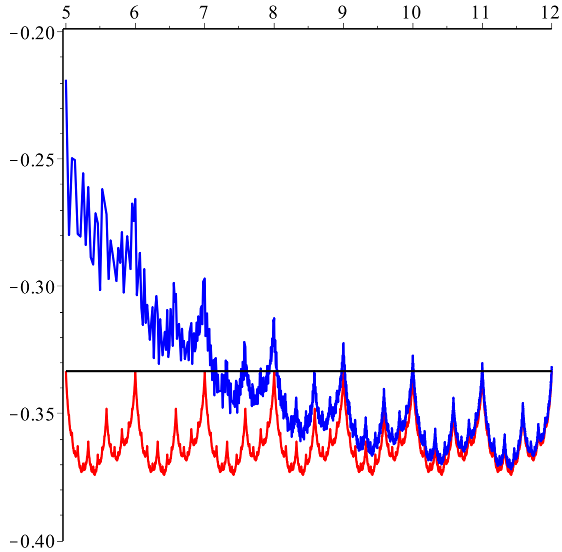

Fig. 2. Theory is closely akin to practice: for the dichopile algorithm, the normalized sequence $N^{-1}(f_N - N\log_2(N)/2)$ approaches the 1-periodic function $\varphi(t)$ as N goes towards infinity. Abscissæ are in logarithmic scale for N. The maximum value of the periodic function is $-1/3$.

for every ε satisfying $0<\varepsilon<1$ and with a 1-periodic function

$$\varphi(t)=\frac{3-\{t\}}{2}+2^{1-\{t\}}g_5\big(2^{\{t\}-1}\big), \tag{32}$$

which is Hölder with exponent ε for $0<\varepsilon<1$. Let us detail the change from one asymptotic expansion to the other. First, using Eq. (28), we write $S_{K+1}(x)=2^K(K+3)x+2^{K+1}g_5(x)+O(r^K)$ for $1<r<2$. Next we replace K by $\log_2 N-\{t\}$ and x by $2^{\{t\}-1}$ to obtain

$$\begin{aligned} f_N \underset{N\to+\infty}{=} & 2^{\log_2 N-\{t\}}(\log_2 N-\{t\}+3)2^{\{t\}-1}+2^{\log_2 N-\{t\}+1}g_5\big(2^{\{t\}-1}\big)+O\big(N^{\log_2 r}\big) \\ = & \frac{1}{2}2^{\log_2 N}\log_2 N+\frac{1}{2}2^{\log_2 N}\big(3-\{t\}\big)+2^{\log_2 N}2^{1-\{t\}}g_5\big(2^{\{t\}-1}\big)+O\big(N^{\log_2 r}\big) \end{aligned}$$

and the writing above appears. Fig. 2 shows the comparison between the sequence $N^{-1}(f_N-N\log_2(N)/2)$ and the periodic function $\varphi(t)$.

5. Comments on the algebraic approach

5.1. Lazy approach

Perhaps the reader finds that all this algebraic machinery is too complicated if he wants only an asymptotic equivalent for the sequence under consideration. In that case, it suffices to simplify the process of computation.

Example 6 *(Dichopile algorithm, the true end).* Let us assume that we are only interested in an equivalent for the cost f_N of the dichopile algorithm. We compute the spectral radius of the matrix $\boldsymbol{Q}$ and we find it is equal to 2. Using the maximum absolute column sum norm, we find numerically for the length $T=10$

$$\max_{|w|=10}\|\boldsymbol{A}_w\|_1^{1/10}\simeq 1.271<2.$$

(As a matter of fact, the example is so simple that we find by hand the value $\sqrt{3}\simeq 1.732$ for the length 2 and this is sufficient, but we try to be a little more generic.) We proceed as in Example 4, but we retain only the contribution of the vector $\boldsymbol{v}_2^{(1)}$, and moreover we take only the first term in (27), that is $\binom{K}{\nu-1}(\rho\omega)^{K-\nu+1}\boldsymbol{f}^{(0)}(x)$. As it is often the case for the cost of an algorithm, the function $\boldsymbol{f}^{(0)}(x)=\boldsymbol{f}(x)$ is explicit (Example 3) and we obtain

$$S_K(x)\underset{K\to+\infty}{\sim}2^{K-1}Kx, \qquad \text{henceforth} \qquad f_N\underset{N\to+\infty}{\sim}\frac{1}{2}N\log_2 N.$$

5.2. Improvement of the error term

It is slightly irritating that in using Theorem 1, we have always to write *the error term is* $O(N^{\log_B r})$ *for every* $r > \rho_*$. It would be simpler if we could write *the error term is* $O(N^{\log_B \rho_*})$. This is not true in full generality, but there is a circumstance that guarantees this property. In the definition of the joint spectral radius (15), it can happen that there exists a subordinate norm and a length T such that the equality

$$\rho_* = \max_{|w|=T} \|\boldsymbol{A}_w\|^{1/T} \tag{33}$$

holds. In such a case, it is said that *the set of matrices has the finiteness property* [57]. We will also say that the linear representation has the finiteness property.

The coordinate vector $\boldsymbol{c}$ of the representation decomposes into the Jordan basis used to reduce the matrix $\boldsymbol{Q}$. The eigenvalues associated with the generalized eigenvectors which occur in this decomposition can be sorted into two sets: first the set $\Lambda_>$ of eigenvalues of $\boldsymbol{Q}$ larger than the joint spectral radius, next its complementary part $\Lambda_\leq$. The first set $\Lambda_>$ provides the regular part of the asymptotic expansion. The other set $\Lambda_\leq$ provides the error term.

Lemma 8. *If the linear representation has the finiteness property, the error term in the expansion announced by Theorem 1 writes*

- $O(N^{\log_B \rho_*})$ *if* $\Lambda_\leq$ *does not contain the joint spectral radius* ρ_*,
- $O(N^{\log_B \rho_*} \log_B^m N)$ *if* ρ_* *is a member of* $\Lambda_\leq$ *and m is the maximal size of the Jordan cells associated with* ρ_* *and involved in the decomposition of the coordinate vector* $\boldsymbol{c}$ *over the Jordan basis.*

Proof. We deal with every eigenvalue not larger than ρ_*. We use the same notations as in Lemma 4. In the first case $\rho < \rho_*$, the proof of this proposition works with ρ_* in place of r. In the second case, let us assume for a while that we have $T = 1$ in (33). With the recursive formula (14) and $\boldsymbol{Q}^K \boldsymbol{v}^{\nu-1} = O(\rho_*^K K^{\nu-1})$, we obtain $\|\boldsymbol{s}_{K+1}\|_\infty \leq c\rho_*^K K^{\nu-1} + \rho_* \|\boldsymbol{s}_K\|_\infty$ for some positive constant c. This recurrence solves into $\|\boldsymbol{s}_K\|_\infty = O(\rho_*^K K^\nu)$. For the general case $T \geq 1$, we use the subsequences $\boldsymbol{s}_{KT+s}(x)$ with $0 \leq s < T$. To have a bound at our disposal we employ m that is the maximal ν encountered in the decomposition of $\boldsymbol{c}$. Substituting $\log_B N$ for K, we arrive at the announced result. □

The next example enables us to see, with our eyes, that the order of the error term predicted by the above proposition is tight.

Example 7 *(Triangular tiling).* This example is slightly different from the other examples in the article because first we do not consider a sequence but a rational series, second the linear representation is not insensitive to the leftmost zeroes. This last point is of no importance here because it was only useful in the change from a rational series to a radix-rational sequence. For a real ϑ, we consider the rotation matrix

$$\boldsymbol{R}_\vartheta = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}.$$

In what follows we use $\boldsymbol{v}_0 = \boldsymbol{e}_1$ the first vector of the canonical basis and $\boldsymbol{v}_\vartheta = \boldsymbol{R}_\vartheta \boldsymbol{v}_0$. The example is based on the linear representation, with radix 2,

$$\boldsymbol{\ell} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \boldsymbol{A}_0 = \boldsymbol{R}_{-\vartheta}, \qquad \boldsymbol{A}_1 = \boldsymbol{R}_\vartheta, \qquad \boldsymbol{c} = \boldsymbol{v}_0.$$

Because we use orthogonal matrices, the joint spectral radius is $\rho_* = 1$. The matrix $\boldsymbol{Q}$ is the diagonal matrix $2\cos\vartheta\, \mathbf{I}_2$. We particularize the case to $\vartheta = \pi/3$, so that the only eigenvalue is $\rho = \rho_*$ with $\nu = 1$, and $m = 1$.

We consider the words $u = 11$, $v = 01$, and the rational numbers x_{2K+2} and y_{2K} whose binary expansions are $(0.uv^K)_2$ and $(0.v^K)_2$ respectively. According to the functional equation (14) satisfied by $\boldsymbol{s}_K(x)$, we have

$$\boldsymbol{s}_{2K+2}(x_{2K+2}) = \boldsymbol{R}_{-\pi/3}\boldsymbol{v}_0 + \boldsymbol{R}_{\pi/3}\big(\boldsymbol{R}_{-\pi/3}\boldsymbol{v}_0 + \boldsymbol{R}_{\pi/3}\boldsymbol{s}_{2K}(y_{2K})\big) = \boldsymbol{v}_{-\pi/3} + \boldsymbol{v}_0 + \boldsymbol{R}_{2\pi/3}\boldsymbol{s}_{2K}(y_{2K}),$$
$$\boldsymbol{s}_{2K}(y_{2K}) = \boldsymbol{R}_{-\pi/3}\big(\boldsymbol{R}_{-\pi/3}\boldsymbol{v}_0 + \boldsymbol{R}_{\pi/3}\boldsymbol{s}_{2K-2}(y_{2K-2})\big) = \boldsymbol{v}_{-2\pi/3} + \boldsymbol{s}_{2K-2}(y_{2K-2}).$$

With $\boldsymbol{s}_0(0) = \boldsymbol{v}_0$, we obtain by induction $\boldsymbol{s}_{2K}(y_{2K}) = K\boldsymbol{v}_{-2\pi/3} + \boldsymbol{v}_0$ and

$$\boldsymbol{s}_{2K+2}(x_{2K+2}) = (K+1)\boldsymbol{v}_0 + \boldsymbol{v}_{-\pi/3} + \boldsymbol{v}_{2\pi/3} = (K+1)\boldsymbol{v}_0.$$

Hence the estimate $O(K^m \rho_*^K)$ provided by Lemma 8 is satisfactory. The orbit of the parameterized curve $\boldsymbol{s}_K(x)$ is illustrated in Fig. 3.

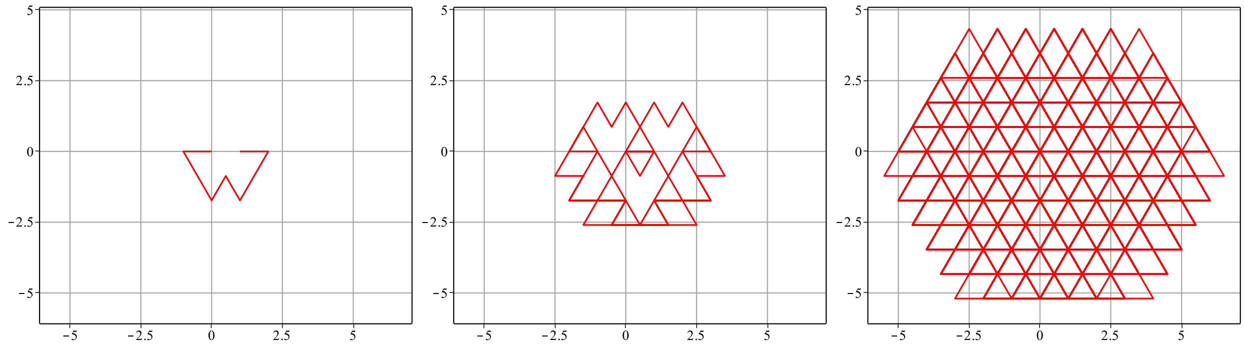

Fig. 3. The image of the parameterized curve $\boldsymbol{s}_K(x)$ from Example 7 grows linearly with K. Here we see $\boldsymbol{s}_3$, $\boldsymbol{s}_6$, and $\boldsymbol{s}_{12}$.

5.3. Algorithm

To summarize the obtained results and clarify the process of computation we write it as an algorithm (Algorithm 1) named `LRtoAE` for Linear Representation to Asymptotic Expansion. To tell the truth, it is a *pseudo*-algorithm, for there exist two essential difficulties: the computation of the joint spectral radius [58] and the resolution of the dilation equations. In Line 3, we perhaps are obliged to content ourselves with a real number r slightly larger than ρ_* as in Example 6. In that case the question about the finiteness property (Line 4) is no longer meaningful. In the same way, it is likely that in Line 14 we cannot solve explicitly the dilation equation. This is not an obstacle that prevents us from writing the expansion, that provides us with the qualitative behavior of the sequence.

5.4. Related topics

Certainly improvements are possible. For example, in our theorem we consider a Hölder exponent. It must be understood that this exponent is a global lower bound, which means it is valid uniformly in the whole interval of reference. A deeper approach is presented in [50, Th. 4.2, p. 1054] where the idea of a local Hölder exponent is described and studied in the framework of dilation equations.

In [59], Tenenbaum shows that some periodic functions are nowhere differentiable. It would be a misunderstanding of Tenenbaum's article to think that this is the general case. As an example, we consider below a probability distribution function and, according to the Lebesgue differentiation theorem for monotone functions, it is almost everywhere differentiable. Actually, Tenenbaum assumes that the Hölder exponent is an upper bound and not a lower bound, in mathematical notations $\Omega(|h|^\alpha)$ and not $O(|h|^\alpha)$. A good bibliography about nowhere differentiable functions and their history can be found in [60,61].

Example 8 *(Biased coin distribution function).* Billingsley [62, Ex. 31.1, p. 407] studied the random variable $X = \sum_{n\geq 0} X_n/2^n$ where X_n is the result of a coin tossing with probabilities p_0 and p_1 for $X_n = 0$ and $X_n = 1$ respectively. This defines a rational series with dimension 1, radix 2 and a linear representation $\boldsymbol{\ell} = (1)$, $\boldsymbol{A}_0 = (p_0)$, $\boldsymbol{A}_1 = (p_1)$, $\boldsymbol{c} = (1)$ with $0 < p_0, p_1 < 1$, $p_0 + p_1 = 1$. From the standpoint of number theory, the associated sequence is a completely 2-multiplicative function [36, Def. 8.1.5]. We have $\boldsymbol{Q} = (1)$, $\rho = 1$, $\boldsymbol{v} = (1)$, $\rho_* = \max(p_0, p_1)$. The distribution function F is the solution of the dilation equation (18) and it is Hölder with exponent $\log_2(1/\rho_*)$. We illustrate the example with $p_0 = 1/5$, $p_1 = 4/5$ and the exponent is $\log_2(4/5) \simeq 0.322$, which provides us with Fig. 4, already sketched out in [63, p. 268–269] or [64], where the distribution function is quoted as Lebesgue's singular function. A similar picture appears also in [54, Fig. 6] about baker-type maps and in [65], which is a gentle introduction to the idea of box dimension.

The provided exponent is the best uniform exponent possible. The positive character of the representation permits us to show (see [66] for an example) that, assuming $p_0 \leq p_1$, at every dyadic point the best local Hölder exponent is $\log_2(1/p_0)$ on the right-hand side and $\log_2(1/\rho_*)$ on the left-hand side. Except in the case $p_0 = p_1 = 1/2$, this gives $\log_2(1/p_0) > 1$ and this explains the right-hand sided horizontal tangents in the pictures. See [64] for details about the derivative and a bibliography.

A recurrent question is about the bounds of the periodic functions. It is dealt with for example in [17,10] and more recently in [32,67].

An issue which does not seem to have been tackled yet concerns the symmetries of the solutions of dilations equations. It appears slightly in [61] with the symmetry $\tau(1-x) = \tau(x)$ of the Takagi function. We exemplify it on the Rudin–Shapiro sequence.

Algorithm 1: `LRtoAE`.

Input : A linear representation $\boldsymbol{\ell}$, $(\boldsymbol{A}_b)_{0\le b<B}$, $\boldsymbol{c}$ insensitive to the leftmost zeroes for the sequence of backward differences $u_n = \nabla s_n$.

Output: An asymptotic expansion of the sequence s_N with respect to the scale $N^\alpha \binom{\log_B N}{m}$, $\alpha \in \mathbb{R}$, $m \in \mathbb{N}_{\ge 0}$.

$\boldsymbol{Q} := \sum_{0\le b<B} \boldsymbol{A}_b$;
compute a Jordan basis $\mathcal{V}$ for the matrix $\boldsymbol{Q}$;
compute the joint spectral radius ρ_* of the linear representation;
if the linear representation has the finiteness property **then**
  $r := \rho_*$
else
  $r :=$ any number between ρ_* and the infimum of the modulus of eigenvalues of $\boldsymbol{Q}$ greater than ρ_*
end
expand the column vector $\boldsymbol{c}$ of the linear representation over the Jordan basis, as $\boldsymbol{c} = \sum_{\boldsymbol{v}\in\mathcal{V}} \gamma_{\boldsymbol{v}} \boldsymbol{v}$;
$\mathcal{V}_> :=$ the set of $\boldsymbol{v}$ in $\mathcal{V}$ such that $\gamma_{\boldsymbol{v}} \ne 0$ and the associated eigenvalue has a modulus $\rho > r$;
for each generalized eigenvector $\boldsymbol{v}$ in $\mathcal{V}_>$ **do**
  $\rho\omega :=$ the eigenvalue associated with $\boldsymbol{v}$;
  $\nu :=$ the size of the Jordan cell $\boldsymbol{J}$ associated with $\boldsymbol{v}$;
  compute the solution $\boldsymbol{F}(x)$ of the dilation system $\boldsymbol{F}(x)\boldsymbol{J} = \sum_b \boldsymbol{A}_b \boldsymbol{F}(Bx-b)$ with the boundary conditions $\boldsymbol{F}(x) = \boldsymbol{0}$ for $x \le 0$ and $\boldsymbol{F}(x) = \boldsymbol{V}$ for $x \ge 1$;
  write down the expansion, with $\boldsymbol{f}^{(-1)}(x) = 0$,

$$\boldsymbol{e}_{\boldsymbol{v},N} := \sum_{\ell=0}^{\nu-1} \binom{K+1}{\ell} (\rho\omega)^{K+1-\ell} \left[\rho\omega \boldsymbol{f}^{(\nu-\ell-1)}\left(B^{\{t\}-1}\right) + \boldsymbol{f}^{(\nu-\ell-2)}\left(B^{\{t\}-1}\right)\right]$$

  substitute $\log_B N - \{t\}$ for K in $\boldsymbol{e}_{\boldsymbol{v},N}$ (implicitly $t = \log_B N$);
  expand the binomial coefficients $\binom{\log_B N+1-\{t\}}{\ell}$ by the Chu–Vandermonde formula;
  collect $\boldsymbol{e}_{\boldsymbol{v},N}$ according to the binomial coefficients $\binom{\log_B N}{m}$
end
$\boldsymbol{e}_N := \sum_{\boldsymbol{v}\in\mathcal{V}_>} \gamma_{\boldsymbol{v}} \boldsymbol{e}_{\boldsymbol{v},N}$;
collect $\boldsymbol{e}_N$ according to the powers of N and next to the index m of the binomial coefficients to obtain

$$\boldsymbol{e}_N = \sum_{\rho,\vartheta,m} N^{\log_B \rho} \binom{\log_B N}{m} e^{i\vartheta \log_B N} \boldsymbol{\varphi}_{\rho,\vartheta,m}(\log_B N)$$

where $\boldsymbol{\varphi}_{\rho,\vartheta,m}(t)$ is 1-periodic and Hölder $\log_B(r'/\rho_*)$ for every $r' \in (\rho_*, \rho)$;
error term$_N := O(N^{\log_B r})$;
if the linear representation has the finiteness property **then**
  $\mathcal{V}_= :=$ the set of generalized eigenvectors $\boldsymbol{v}$ in $\mathcal{V}$ such that $\gamma_{\boldsymbol{v}} \ne 0$ and the associated eigens value $\rho\omega$ has a modulus $\rho = \rho_*$;
  $\nu_{\max} :=$ the maximal size of the Jordan cell for the vectors in $\mathcal{V}_=$;
  error term$_N := O(N^{\log_B \rho_*} \log_B^{\nu_{\max}} N)$
end
return

$$s_N \underset{N\to+\infty}{=} \boldsymbol{\ell}\boldsymbol{e}_N + \underset{N}{\text{error term}}$$

Example 9 *(Rudin–Shapiro sequence).* The Rudin–Shapiro sequence may be defined as $u_n = (-1)^{e_{2;11}(n)}$ where $e_{2;11}(n)$ is the number of (possibly overlapping) occurrences of the pattern 11 in the binary expansion of the integer n [68]. It is 2-rational: it admits the generating family (u_n, u_{2n+1}) and the reduced linear representation, insensitive to the leftmost zeroes,

$$\boldsymbol{\ell} = \begin{pmatrix} 1 & 1 \end{pmatrix}, \qquad \boldsymbol{A}_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \boldsymbol{A}_1 = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Theorem 1 gives the already known asymptotic expansion for the Rudin–Shapiro sequence [17]

$$\sum_{n\le N} u_n \underset{N\to+\infty}{=} \sqrt{N}\varphi(\log_4 N) + O(1)$$

where φ is the 1-periodic function defined by $\varphi(t) = 2^{1-\{t\}} f(4^{\{t\}-1})$ and f is defined through a dilation equation for radix 4. More precisely, we use the radix-4 linear representation $\boldsymbol{\ell}' = \boldsymbol{\ell}$, $\boldsymbol{A}'_0 = \boldsymbol{A}_0^2$, $\boldsymbol{A}'_1 = \boldsymbol{A}_0\boldsymbol{A}_1$, $\boldsymbol{A}'_2 = \boldsymbol{A}_1\boldsymbol{A}_0$, $\boldsymbol{A}'_3 = \boldsymbol{A}_1^2$, $\boldsymbol{c}' = \boldsymbol{c}$. Functions f and φ are illustrated in Fig. 5. Let us denote π_k the part of the graph of f which corresponds to the interval $[k/8, (k+1)/8]$ for $0 \le k < 8$. It is clear that the parts of odd index on the one hand, and the part of even index on the other hand, reproduce the same pattern (with a piece upside down). This can be proved elementarily, playing at the same time with the radix-2 representation and the radix-4 representation.

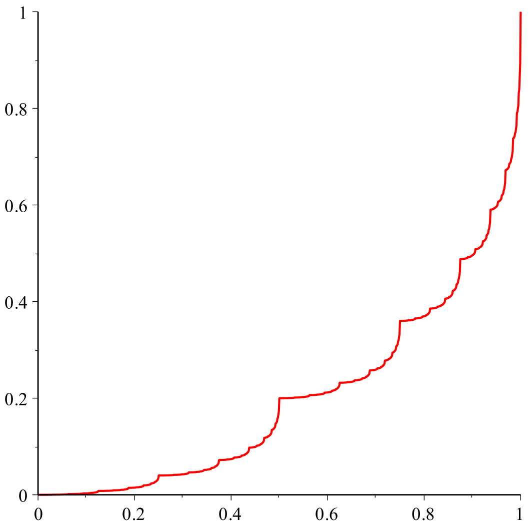

Fig. 4. Limit probability distributions which come from a Bernoulli process described in Example 8.

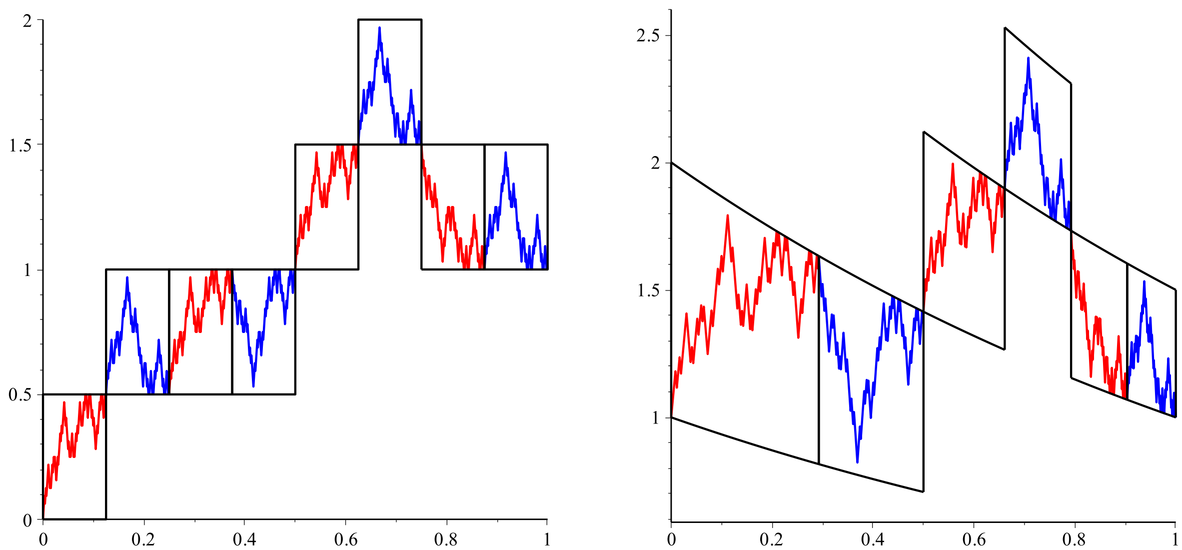

Fig. 5. The solution of the dilation equation (left-hand side) and the periodic function (right-hand side) associated with the Rudin–Shapiro sequence of Example 9 show some symmetries. They are obvious for the first one but quite buried for the second.

The symmetries of f are translated to φ, but they lose their obvious character. The pieces π_0 and π_1 disappear and the pieces π_k, $2 \le k < 8$ become the pieces π'_k associated with the intervals $[\log_4(k/2), \log_4((k+1)/2)]$. The links between the pieces become more intricate. For example the pieces π'_2 and π'_3 on the one side, and π'_6 and π'_7 on the other side, are linked by the formula $2^s\varphi(s) + 2^t\varphi(t) = 4$ under the condition that both numbers $s \in [0, 1/2]$ and $t \in [\log_4 3, 1]$ are related by $4^t - 4^s = 2$.

6. A matter of taste

Our point of interest is the comparison between the analytic number theory method and our linear algebra method. The analytic method is based on meromorphic functions, that are defined by Dirichlet series, and computation of some residues. A good and concise account of the method is given in [36, Sec. 8.2.3]. The underlying idea is geometrically obvious, but its application needs dexterity. The computations in [69] provides a good illustration of this fact. By contrast our approach is elementary but less intuitive. Nevertheless, it would be a mistake to oppose both approaches. Each of them has its own merit and eventually it is a matter of taste whether one chooses between analysis and algebra. But we think that, for the particular case of radix-rational sequences, the linear algebraic method is more appropriate because the issue is of a linear nature and the application of the method is easier in the context.

6.1. Multidimensional Dirichlet series

We first describe the running of the analytic number theory approach.

Example 10 *(Dichopile algorithm, another way-1).* The use of the analytic number theory approach begins with the Dirichlet series

$$u(s)=\sum_{n\geq 1}\frac{\nabla f_n}{n^s}$$

associated with the sequence $u_n=\nabla f_n$, or better with the family of Dirichlet series associated with each of the sequences in the basis used to define the linear representation (Example 1). All these Dirichlet series are gathered into a row-vector-valued Dirichlet series $\boldsymbol{u}(s)$, associated with the row-vector-valued sequence $\boldsymbol{u}_n$ whose components are the sequences of the basis. Consequently the component of interest is recovered by $u(s)=\boldsymbol{u}(s)\boldsymbol{c}$.

The first question is to evaluate the abscissa of absolute convergence of $\boldsymbol{u}(s)$. Usually this results from an *ad hoc* computation, often a bound obtained by elementary arguments in a former work [17], [24, Formulæ (6.7), (6.10)]. For us this results from the computation of the joint spectral radius. The value $\rho_*=1$ shows that all sequences in the basis are $O(n^\varepsilon)$ for every $\varepsilon>0$ and this asymptotic estimate is the best possible (among the comparisons with a power of n), so that the abscissa of absolute convergence is $\sigma_a=1$.

Next, we filter the integer n in the sum that defines $\boldsymbol{u}(s)$ according to its parity. This provides us with

$$\boldsymbol{u}(s)=\sum_{n=1}^{+\infty}\frac{\boldsymbol{u}_{2n}}{(2n)^s}+\boldsymbol{u}_1+\sum_{n=1}^{+\infty}\frac{\boldsymbol{u}_{2n+1}}{(2n+1)^s}=\frac{1}{2^s}\boldsymbol{u}(s)\boldsymbol{A}_0+\boldsymbol{u}_1+\frac{1}{2^s}\boldsymbol{u}(s)\boldsymbol{A}_1+\sum_{n=1}^{+\infty}\left(\frac{1}{(2n+1)^s}-\frac{1}{(2n)^s}\right)\boldsymbol{u}_n\boldsymbol{A}_1,$$

or in other words

$$\boldsymbol{u}(s)\left(\mathbf{I}_6-2^{-s}\boldsymbol{Q}\right)=\nabla\boldsymbol{u}(s)\qquad\text{with}\quad\nabla\boldsymbol{u}(s)=\boldsymbol{u}_1+\sum_{n=1}^{+\infty}\left(\frac{1}{(2n+1)^s}-\frac{1}{(2n)^s}\right)\boldsymbol{u}_n\boldsymbol{A}_1.\tag{34}$$

The last series converges absolutely for complex numbers s with a positive real part, because of the difference $(2n+1)^{-s}-(2n)^{-s}$, which is of order n^{-s-1}. Eq. (34) shows that $\boldsymbol{u}(s)$ extends on the half-plane $\sigma>0$ as a meromorphic function. More precisely the poles are the logarithms to base 2 of the eigenvalues of the matrix $\boldsymbol{Q}$. As a consequence, on the boundary of the half-plane of absolute convergence there is a vertical line of poles $1+\chi_k$ with $\chi_k=2k\pi i/\ln 2$ and k integer, associated with the dominant eigenvalue 2. There are no other poles in the vertical strip between 0 and 1.

The next stage is the use of the Mellin–Perron summation formula [70, Th. 13]

$$\sum_{1\leq k<N}\boldsymbol{u}_k+\frac{1}{2}\boldsymbol{u}_N=\frac{1}{2\pi i}\int_{(c)}\boldsymbol{u}(s)N^s\frac{ds}{s}.\tag{35}$$

In this formula the integral is taken along a vertical line (c) at an abscissa c larger than the abscissa of absolute convergence. We express the function $\boldsymbol{u}(s)$ using Formula (34) and to emphasize the poles at abscissa 1, we write $(\mathbf{I}_6-2^{-s}\boldsymbol{Q})^{-1}=\boldsymbol{R}(s)/(1-2^{1-s})^2$ with $\boldsymbol{R}(s)$ a matrix-valued meromorphic function, analytic on the right of 0. The function to be integrated has an expansion near $s=1+\chi_k$ of the form

$$\frac{\nabla\boldsymbol{u}(s)\boldsymbol{R}(s)}{(1-2^{1-s})^2}\frac{N^s}{s}\underset{s\to 1+\chi_k}{=}\boldsymbol{c}_{2,k}N^{1+\chi_k}\frac{1}{(s-1-\chi_k)^2}+\left(\boldsymbol{c}_{1,k}N^{1+\chi_k}\ln(N)+\boldsymbol{c}_{0,k}N^{1+\chi_k}\right)\frac{1}{s-1-\chi_k}+O(1).$$

Using Cauchy's residue theorem we change Formula (35) into

$$\sum_{1\leq k<N}\boldsymbol{u}_k+\frac{1}{2}\boldsymbol{u}_N=\sum_{k=-\infty}^{+\infty}\boldsymbol{c}_{1,k}N^{\chi_k}\times N\ln(N)+\sum_{k=-\infty}^{+\infty}\boldsymbol{c}_{0,k}N^{\chi_k}\times N+\frac{1}{2\pi i}\int_{(\varepsilon)}\boldsymbol{u}(s)N^s\frac{ds}{s},\tag{36}$$

where this time ε is between 0 and 1. The term $\boldsymbol{u}_N/2$ and the integral are of order N^ε. Overall N^{χ_k} is nothing but $\exp(2k\pi i\log_2 N)$ and some trigonometric series appear,

$$\sum_{1\leq k\leq N}\boldsymbol{u}_k=N\ln(N)\sum_{k=-\infty}^{+\infty}\boldsymbol{c}_{1,k}\exp(2k\pi i\log_2 N)+N\sum_{k=-\infty}^{+\infty}\boldsymbol{c}_{0,k}\exp(2k\pi i\log_2 N)+O\left(N^\varepsilon\right).\tag{37}$$

We have obtained an asymptotic expansion in the scale $N^\alpha\ln^\beta(N)$ with variable coefficients. The method is simple, concrete, obvious, natural: we see the poles, we see the line of integration, we push the line to the left, we catch the residues, and we have the asymptotic expansion of the partial sum.

6.2. Obstacles

Unfortunately the radiant sun that illumines this method was soon overshadowed by thick clouds.

Example 11 *(Dichopile algorithm, another way-2).* We review in turn each point that causes an issue in the previous example. First, we do not know if the points $1+\chi_k$ are really poles of $\boldsymbol{u}(s)$. In this example, the only point that is certainly a pole is the abscissa of absolute convergence $\sigma_a = 1$, according to Landau's theorem [70, Th. 10], because the matrices $\boldsymbol{\ell}$, $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ have nonnegative coefficients so that all the components of $\boldsymbol{u}_n$ are nonnegative. The same phenomenon occurs with the constant sequence of value 1, whose associated Dirichlet series is the Riemann zeta function $\zeta(s)$. In that case, Eq. (34) is the usual link between $\zeta(s)$ and the alternate Riemann zeta function. Because of the factor $1-2^{1-s}$ it seems that $\zeta(s)$ has a line of poles $1+\chi_k$, but we know that the only pole is 1. Here, comparing (37) and our expansion (31), which begins with a dominant term $N\log_2(N)/2$, we see that 1 may be a double pole but that the other points $1+\chi_k$ are certainly at most simple poles.

Second, there is no reason (at this stage) for the trigonometric series above, which is the coefficient of N in (37), to be a convergent series and define a periodic function. Usually, an extra argument proves the occurrence of a periodic function (see for example [7, Lemma H], [10, Sec. 2] or [17, Sec. 2]). For us, Formula (32) defines a periodic function.

Third, the order of growth of $\boldsymbol{u}(s)$ along a vertical line does not permit the use of the Mellin–Perron formula. Eq. (34) shows, by the study of the right-hand side, that at infinity $\boldsymbol{u}(\sigma+it)$ does not grow faster than $|t|^0$ for $\sigma>1$ and than $|t|^1$ for $0<\sigma<1$. Hence $\boldsymbol{u}(\sigma+it)$ does not grow faster than $|t|^{1-\sigma}$ for $0<\sigma<1$. It is a consequence of the Lindelöf theorem [70, Sec. III.4], which asserts that the order of growth is a convex function with respect to σ. The bound is not small enough to guarantee the absolute convergence of the integral on the line (ε). However the Mellin–Perron formula, say of the second order,

$$\frac{1}{N}\sum_{1\le k\le n<N}\boldsymbol{u}_k=\sum_{1\le k<N}\boldsymbol{u}_k\left(1-\frac{k}{N}\right)=\frac{1}{2\pi i}\int_{(c)}\boldsymbol{u}(s)N^s\frac{ds}{s(s+1)}$$

can be used because if we change the line of integration (c) into (ε) the integrand becomes $O(|t|^{-1-\varepsilon})$. We obtain an expansion

$$\frac{1}{N}\sum_{1\le k\le n<N}u_k\underset{N\to+\infty}{=}N\log_2(N)\psi_1(\log_2 N)+N\psi_0(\log_2 N)+O\left(N^{\varepsilon}\right),$$

where $\psi_1(t)$ and $\psi_0(t)$ are 1-periodic functions defined as sums of convergent trigonometric series. But Proposition 6.4 of [24] provides us, by summation of the expansion (31), with

$$\frac{1}{N}\sum_{1\le k\le n<N}u_k\underset{N\to+\infty}{=}\frac{1}{4}N\log_2(N)-\frac{1}{8}N+N\psi(\log_2 N)+o(N). \tag{38}$$

The uniqueness of the asymptotic expansion shows that $\psi_1(t)=1/4$ and $\psi_0(t)=-1/8+\psi(t)$. Moreover Proposition 6.4 of [24] gives the link between the Fourier coefficients $c_k(\varphi)$ of the function $\varphi(t)$, which appears in the asymptotic expansion (31), and the Fourier coefficients $c_k(\psi)$ of the function $\psi(t)$ in (38),

$$c_k(\varphi)=(2+\chi_k)c_k(\psi).$$

This enables us to show that the Fourier series of $\varphi(t)$ is the product of the row vector $\boldsymbol{\ell}$ and the vector-valued trigonometric series which appears in (37) as the coefficient of N. In other words, to push the line of integration on the left and collect the residues provides us with the right Fourier series, even if this first seems to be a wrong process. Bernstein's theorem [71, Th. VI.3.1], quoted in [25, Prop. 6], guarantees the uniform convergence of the Fourier series for a Hölder continuous function with exponent $>1/2$. We know by our algebraic approach that this is the case for the dichopile algorithm. In this example, the Fourier series converges, but it is not the general case, and it may be necessary to consider Féjer sums [24, Th. 6.2].

Let us consider another example a little bit more complicated because two lines of poles play a part in the result.

Example 12 *(Rational series and cryptanalysis).* We consider a rational series that originates in differential cryptanalysis [66]. It uses the octal system of numeration, hence the alphabet $\{0,1,\dots,7\}$, and admits the 8-dimensional linear representation defined as follows. The initial values and the coordinates are

$$\boldsymbol{\ell}=\begin{pmatrix}1&1&1&1&1&1&1&1\end{pmatrix},\qquad \boldsymbol{c}=\begin{pmatrix}1&0&0&0&0&0&0&0\end{pmatrix}^T.$$

The first matrix of the action is

$$\boldsymbol{A}_0 = \frac{1}{4}\begin{pmatrix} 4 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The other ones, for $0 < b < 8$, are obtained from $\boldsymbol{A}_0$ as the conjugate $\boldsymbol{A}_b = \boldsymbol{P}_b \boldsymbol{A}_0 \boldsymbol{P}_b$, where $\boldsymbol{P}_b$ is the matrix of the involution $k \mapsto k \oplus b$ with $\oplus$ the bitwise xor, in formula $(\boldsymbol{A}_b)_{ij} = (\boldsymbol{A}_0)_{i\oplus b,\, j\oplus b}$. We consider the 8-rational sequence u_n associated with this representation and we want an asymptotic expansion of its partial sums s_N.

Let us begin with an elementary approach. The matrix

$$\boldsymbol{Q} = \sum_{0\le b<8} \boldsymbol{A}_b = \frac{1}{4}\begin{pmatrix} 7 & 2 & 2 & 1 & 2 & 1 & 1 & 0 \\ 2 & 7 & 1 & 2 & 1 & 2 & 0 & 1 \\ 2 & 1 & 7 & 2 & 1 & 0 & 2 & 1 \\ 1 & 2 & 2 & 7 & 0 & 1 & 1 & 2 \\ 2 & 1 & 1 & 0 & 7 & 2 & 2 & 1 \\ 1 & 2 & 0 & 1 & 2 & 7 & 1 & 2 \\ 1 & 0 & 2 & 1 & 2 & 1 & 7 & 2 \\ 0 & 1 & 1 & 2 & 1 & 2 & 2 & 7 \end{pmatrix}$$

reduces to a diagonal matrix $\boldsymbol{\Lambda} = \boldsymbol{P}^{-1}\boldsymbol{Q}\,\boldsymbol{P} = \operatorname{diag}(4, 2, 2, 2, 1, 1, 1, 1)$ with the change of basis matrix

$$\boldsymbol{P} = \frac{1}{24}\begin{pmatrix} 3 & 21 & 12 & 0 & 48 & 36 & 18 & 30 \\ 3 & -5 & -8 & -8 & -30 & -24 & -12 & -12 \\ 3 & 19 & 16 & 4 & -6 & 0 & 0 & -12 \\ 3 & -7 & -4 & -4 & -12 & -12 & -6 & -6 \\ 3 & 7 & 4 & 4 & -18 & -12 & 0 & -12 \\ 3 & -19 & -16 & -4 & 0 & 0 & -6 & -6 \\ 3 & 5 & 8 & 8 & -24 & -24 & -18 & -6 \\ 3 & -21 & -12 & 0 & 42 & 36 & 24 & 24 \end{pmatrix}.$$

Because all terms of the sequence u_n are nonnegative, the sum over the integers between two consecutive powers of 8, say 8^K and 8^{K+1}, is not larger than $\boldsymbol{\ell}\,\boldsymbol{Q}^{K+1}\boldsymbol{c}$, which is $O(4^K)$. Moreover if we denote by $\boldsymbol{p}^1, \boldsymbol{p}^2, \dots, \boldsymbol{p}^8$ the columns of $\boldsymbol{P}$, we have $\boldsymbol{c} = \boldsymbol{p}^1 + \boldsymbol{p}^2 - \boldsymbol{p}^3 + \boldsymbol{p}^5 - \boldsymbol{p}^6$. The vector $\boldsymbol{c}$ has a component over $\boldsymbol{p}^1$, so $\boldsymbol{Q}^K\boldsymbol{c}$ is of order 4^K and it can be verified that $\boldsymbol{\ell}(\boldsymbol{Q} - \boldsymbol{A}_0)\boldsymbol{Q}^K\boldsymbol{c}$ is of order 4^K too. We conclude

$$s_N \underset{N\to+\infty}{\asymp} N^{2/3}. \tag{39}$$

Let us go to the analytic approach. We consider the basis $(u_n^1, u_n^2, \dots, u_n^8)$ that lies behind the linear representation (u_n^1 is nothing but u_n according to the value of $\boldsymbol{c}$) and we see it as a sequence $\boldsymbol{u}_n$ of row vectors. To this one we associate the row-vector-valued Dirichlet series $\boldsymbol{u}(s)$. We change the linear representation using $\boldsymbol{P}$ into

$$\boldsymbol{\ell}' = \boldsymbol{\ell}\boldsymbol{P}, \qquad \boldsymbol{c}' = \boldsymbol{P}^{-1}\boldsymbol{c}, \qquad \boldsymbol{A}'_b = \boldsymbol{P}^{-1}\boldsymbol{A}_b\boldsymbol{P} \quad \text{for } 0 \le b < 8.$$

Accordingly the row-vector $\boldsymbol{u}_n$ and the Dirichlet series $\boldsymbol{u}(s)$ are changed into $\boldsymbol{v}_n = \boldsymbol{u}_n\boldsymbol{P}$ and $\boldsymbol{v}(s) = \boldsymbol{u}(s)\boldsymbol{P}$. The new Dirichlet series $\boldsymbol{v}(s)$ satisfies

$$\boldsymbol{v}(s)\big(\mathrm{I}_8 - 8^{-s}\boldsymbol{\Lambda}\big) = \nabla\boldsymbol{v}(s) \tag{40}$$

with

$$\nabla\boldsymbol{v}(s) = \sum_{b=1}^{7}\frac{\boldsymbol{v}_b}{b^s} + \sum_{b=1}^{7}\sum_{n=1}^{+\infty}\boldsymbol{v}_n\boldsymbol{A}'_b\left(\frac{1}{(8n+b)^s} - \frac{1}{(8n)^s}\right). \tag{41}$$

Because all the coefficients in the matrices are nonnegative and the column sums are equal to 0 or 1, the absolute column sum norm of each of the eight matrices are all equal to 1 and the spectral radius ρ_* is not larger than 1. But all matrices admits 1 as an eigenvalue, henceforth $\rho_* = 1$. (It is noteworthy that the linear representation has the finiteness property of Section 5.2.) Therefore the sequence $\boldsymbol{u}_n$ is bounded as well as the sequence $\boldsymbol{v}_n$. Eq. (41) shows that $\nabla\boldsymbol{v}(s)$ is an analytic function in the half-plane $\sigma > 0$. Henceforth the singularities of $\boldsymbol{v}(s)$ in this half-plane come from the term $\mathrm{I}_8 - 8^{-s}\boldsymbol{\Lambda}$ and $\boldsymbol{v}(s)$ as well as $\boldsymbol{u}(s)$ are meromorphic in this half-plane.

More precisely Eq. (40) translates into scalar equations $v^j(s)(1-8^{-s}\lambda_j) = \nabla v^j(s)$, where λ_j is the jth eigenvalue in $\boldsymbol{\Lambda}$. With $\chi_k = 2k\pi i/\ln 8$ we see that

- $v^1(s)$ admits a line of poles $2/3+\chi_k$ at abscissa $2/3$,
- $v^2(s)$, $v^3(s)$, $v^4(s)$ admit a line of poles $1/3+\chi_k$ at abscissa $1/3$,
- $v^5(s)$, $v^6(s)$, $v^7(s)$, $v^8(s)$ are analytic in $\sigma > 0$.

As we already noticed, the poles under consideration are perhaps fictitious. The only certain point is that $2/3$ is a pole of $v^1(s)$ because the abscissa of convergence of $u(s)$ has value $\limsup_{N\to+\infty} \ln(s_N)/\ln(N)$ according to [70, p. 8], that is $\sigma_c = 2/3$ by the elementary approach above, and $u(s)$ has positive coefficients [70, p. 10].

For $0<\sigma<1$, the order of growth of $\boldsymbol{v}(s)$ is bounded by $1-\sigma$. We can apply the Mellin–Perron formula of order 2 and push the line of integration from an abscissa $c>1$ to an abscissa ε between 0 and $1/3$. We obtain for example the expansion

$$\frac{1}{N}\sum_{1\le m\le n<N} v_m^1 = \sum_{k\in\mathbb{Z}} \operatorname{Res}\left[\frac{\nabla v^1(s)N^s}{(1-4\cdot 8^{-s})}\frac{1}{s(s+1)}, s=\frac{2}{3}+\chi_k\right] + \frac{1}{2\pi i}\int_{(\varepsilon)} \frac{\nabla v^1(s)N^s}{(1-4\cdot 8^{-s})}\frac{ds}{s(s+1)}, \tag{42}$$

that is

$$\frac{1}{N}\sum_{1\le m\le n<N} v_m^1 \underset{N\to+\infty}{=} \frac{N^{2/3}}{\ln 8}\sum_{k\in\mathbb{Z}} \frac{\nabla v^1(2/3+\chi_k)}{(2/3+\chi_k)(5/3+\chi_k)}\exp(2\pi i k\log_8 N) + O\left(N^{\varepsilon}\right). \tag{43}$$

Similarly we find

$$\frac{1}{N}\sum_{1\le m\le n<N} v_m^2 \underset{N\to+\infty}{=} \frac{N^{1/3}}{\ln 8}\sum_{k\in\mathbb{Z}} \frac{\nabla v^2(1/3+\chi_k)}{(1/3+\chi_k)(4/3+\chi_k)}\exp(2\pi i k\log_8 N) + O\left(N^{\varepsilon}\right). \tag{44}$$

Combining Eqs. (43) and (44), we arrive at

$$\sum_{1\le m\le n<N} u_m \underset{N\to+\infty}{=} N^{5/3}\psi_{5/3}(\log_8 N) + N^{4/3}\psi_{4/3}(\log_8 N) + O\left(N^{1+\varepsilon}\right) \tag{45}$$

for $0<\varepsilon<1$, where $\psi_{5/3}(t)$ and $\psi_{4/3}(t)$ are 1-periodic continuous functions. At this point, we are blocked because we do not know if the sequence s_N has an expansion with oscillating coefficients. An external information is needed.

Let us turn towards the algebraic approach. A great proportion of the computations has already been made. Our theorem gives directly an expansion of s_N,

$$s_N \underset{N\to+\infty}{=} N^{2/3}\varphi_{2/3}(\log_8 N) + N^{1/3}\varphi_{1/3}(\log_N) + O(\log N) \tag{46}$$

and similarly for the summatory function $\boldsymbol{s}_N$ of the vector-valued sequence $\boldsymbol{u}_n$. The expression of the error term uses the finiteness property (Section 5.2). Both functions are not defined by their Fourier series but through dilation equations. Moreover they are Hölder continuous respectively with exponent $2/3$ and $1/3$.

Once we know the existence of the asymptotic expansion, we can use [24, Prop. 6.4] to link (45) and (46) together, more precisely $\psi_{5/3}(t)$ with $\varphi_{2/3}(t)$. This gives us the Fourier series of $\varphi_{2/3}(t)$, namely

$$\varphi_{2/3}(t) = \frac{1}{\ln 8}\sum_{k\in\mathbb{Z}} \frac{\nabla v^1(2/3+\chi_k)}{2/3+\chi_k}\exp(2\pi i k t). \tag{47}$$

If we want to concretely proceed to the computation, we have to note that the linear representation is not insensitive to the leftmost zeroes. Hence we have to use Formula (12) in place of Formula (11). But the additional term is simply $(4^{K+1}-3K-4)/6$ with the notations used throughout the paper. Let us denote by $\boldsymbol{f}(x)$, respectively $\boldsymbol{g}(x)$, the solution of the dilation equation that uses the vector $\boldsymbol{w}_4 = \boldsymbol{p}_1$, resp. $\boldsymbol{w}_2 = \boldsymbol{p}_2 - \boldsymbol{p}_3$, for the boundary condition on the righthand side. Because $\boldsymbol{\ell}$ has all its components equal to 1 the useful functions are

$$f(x) = \sum_{k=1}^{8} f_k(x), \qquad g(x) = \sum_{k=1}^{8} g_k(x),$$

and the periodic functions write

$$\varphi_{2/3}(t) = 4^{1-\{t\}}\left(\frac{1}{6} + f\left(8^{\{t\}-1}\right)\right), \qquad \varphi_{1/3}(t) = 2^{1-\{t\}}g\left(8^{\{t\}-1}\right).$$

Fig. 6 shows the comparison between the theoretical version and the empirical version for each of the functions $\varphi_{2/3}(t)$ and $\varphi_{1/3}(t)$. Fig. 7 compares the use of the cascade algorithm and of the Fourier series.

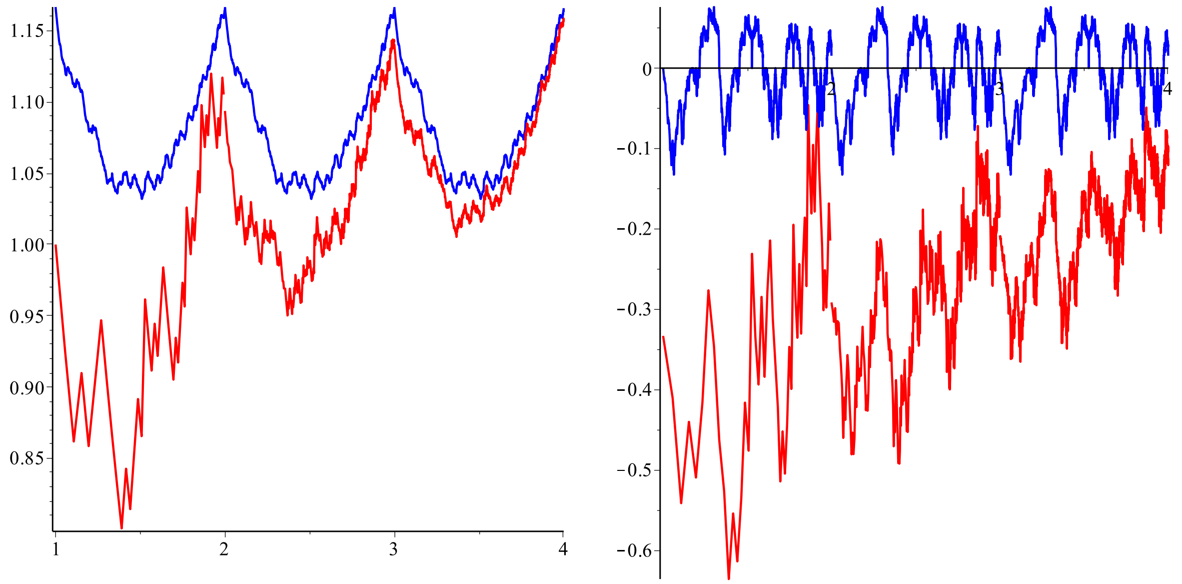

Fig. 6. From Example 12, on the left, the periodic function $\varphi_{2/3}(t)$ and the sequence $s_N/N^{2/3}$; on the right, the periodic function $\varphi_{1/3}(t)$ and the sequence $(s_N - N^{2/3}\varphi_{2/3}(\log_8 N))/N^{1/3}$.

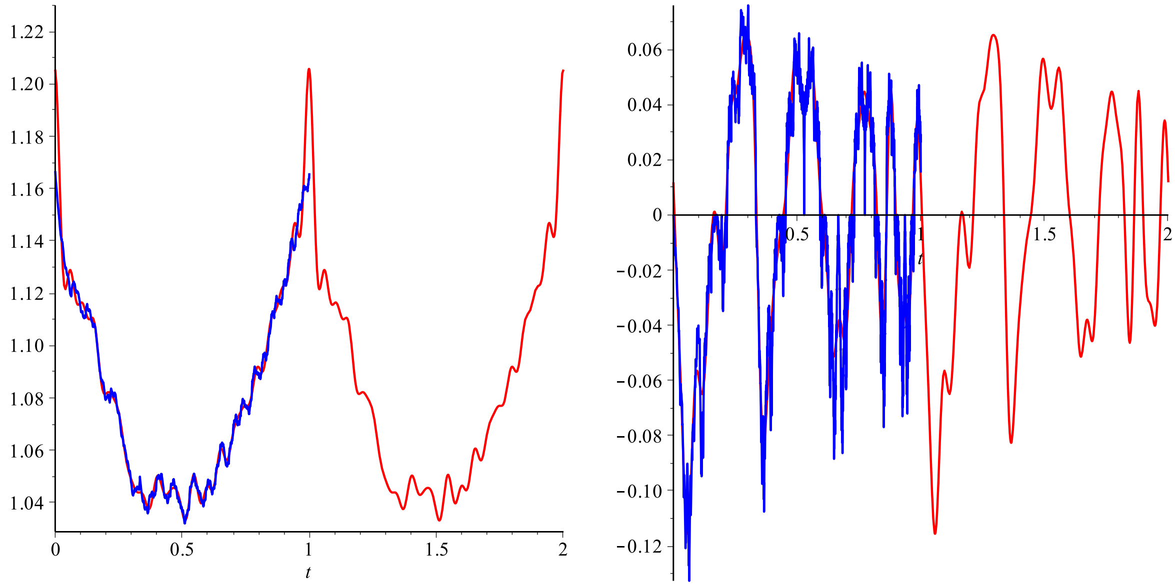

Fig. 7. For both functions $\varphi_{2/3}(t)$ (left) and $\varphi_{1/3}(t)$ (right) from Example 12, their graph computed by the cascade algorithm (between 0 and 1) and by the partial sum of their Fourier series with twenty harmonics (between 0 and 2).

As Examples 10, 11, and 12 have shown, the analytic number theory method and the linear algebraic method are not very far apart and the computations are very similar. They look different because number theorists hide the linear nature of the problem and deal only with scalar sequences.

Both methods have their weaknesses. For the analytic number theory method, there are three weak points: the order of growth at infinity of the Dirichlet series is difficult to evaluate, the apparent poles are perhaps not poles, and often an external information is needed about the existence of a periodic function. Besides, for the linear algebraic method, there are two weak points: the exact spectral radius is difficult to evaluate, the dilation equations are difficult to solve. But the problems that arise in the analytic method are far more sophisticated than in the linear method. An approximation of the spectral radius often suffices to deal with the asymptotic expansion and the solution of a dilation equation can always be gazed at with the help of the cascade algorithm.

The analytic number theory method has a wider scope. For example it is able do deal with a non-homogeneous equation where some logarithmic term appears on the right-hand side ($\log n$ is not a radix-rational sequence). Nevertheless the linear algebraic approach is the best tool to draw attention to the rational series of the formal language theory that are hidden behind the radix-rational sequences. The process that we have described comes as an extension of the simple idea to consider the behavior of the sequence at the integer powers of the radix. It may seem to be a trick, but it amounts to consider the matrix denoted Q all along the article, that is nothing but the equivalent for rational series of the density for formal languages. Our approach directly follows from the idea of rational series.

Moreover linear algebra is very little sophisticated and commonly taught, so that we can hope for example that *algorithmists* employ it in the study of complexity of algorithms, while they might be put off by the analytic number theory. As a matter of fact, linear algebra is well implemented in the computer algebra systems. The computation of a Jordan form, the evaluation of a spectral radius, or the implementation of the cascade algorithm are not an impediment and the algebraic approach can be automated, as we have shown with Algorithm 1.

7. Fourier coefficients

Having made the link between the 1-periodic function and its Fourier series, it is natural to compute its Fourier coefficients. At this point there are two possibilities. The first one is the fine case, where the Dirichlet series $\boldsymbol{u}(s)$ (notation of Example 10) is explicitly known, as in many examples dealt with in [24] or [25], which use the Riemann zeta function.

Example 13 *(Dichopile algorithm, another point of view).* One of the referees has kindly shown us that the dichopile enters into the fine case. We reproduce the proof (partially), with his/her agreement obviously. The sequence g_n satisfies the recursion $g_n = g_{\lfloor n/2\rfloor-1} + g_{\lceil n/2\rceil} + \lfloor n/2\rfloor - 1$, for $n \geq 2$. Hence its ordinary generating function $g(z)$ is a solution of the Mahlerian equation $g(z) = (1+z)(1+z^3)g(z^2)/z + z^4/((1+z)(1-z^2))$. This provides

$$g(z) = \frac{z}{(1-z)(1-z^3)} \sum_{j\geq 0} \frac{z^{3\cdot 2^j}(1-z^{3\cdot 2^j})}{1-z^{2^{j+1}}},$$

as it can be seen using $\bar{g}(z) = (1-z)(1-z^3)g(z)/z$. But the coefficients of this last function $\bar{g}_n = (g_{n+1} - g_n) - (g_{n-2} - g_{n-3})$ appear as differences of order 2 of g_n. This permits to apply the Mellin–Perron formula of order 2. The integral uses the Dirichlet series

$$\bar{g}^*(s) = \frac{(1-2^{1-s})\zeta(s) - 1 + 2^{-s} + 4^{-s}}{1-2^{-s}},$$

that depends on the Riemann zeta function. With the principles exposed in Example 10, the sequence $f_N = N + g_N$ admits the expansion

$$f_N \underset{N\to+\infty}{=} \frac{1}{2} N \log_2 N + N\varphi(\log_2 N) + O(\log N),$$

with a periodic function $\varphi(t)$ given by its Fourier expansion ($\chi_k = 2k\pi i/\ln 2$)

$$\varphi(t) = \frac{\log \pi}{6\log 2} + \frac{1}{12} - \frac{1}{2\log 2} + \frac{1}{3\log 2} \sum_{k\neq 0} \frac{1-\zeta(\chi_k)}{\chi_k(\chi_k+1)} e^{2k\pi it}. \tag{48}$$

The error term $O(\log N)$ is noteworthy, but is obtained by a specific argument.

The second possibility is the generic one, when the Dirichlet series has no properties other than the ones that come from the divide-and-conquer recurrence. This is the case we will focus on. We continue the computations with the dichopile algorithm, even if we know that it is not a generic case, but we are searching for generic methods.

7.1. Residues

We have at our disposal several methods to compute numerically the Fourier coefficients. The first method is the direct application of the analytic number theory approach. The Fourier coefficients are obtained through some residues.

Example 14 *(Dichopile algorithm, computation of the Fourier coefficients-1).* The basic formula is (34), that is $\boldsymbol{u}(s)(\mathbf{I}_6 - 2^{-s}\boldsymbol{Q}) = \nabla\boldsymbol{u}(s)$. The idea is to compute the right member for a pole $1+\chi_k$ of $\boldsymbol{u}(s)$ to obtain the residue at $1+\chi_k$. It is practically easier to use the change of basis of Example 3, that is to use a basis for which the matrix $\boldsymbol{Q}$ is in Jordan form. (We proceeded in that way in Example 12.) This emphasizes the components of $\boldsymbol{u}(s)$ which are really involved in the line of poles $1+\chi_k$.

However, the computation of $\nabla\boldsymbol{u}(s)$ is not so easy, because the convergence of the series (34) that defines $\nabla\boldsymbol{u}(s)$ is slow. Grabner and Hwang [25] deal with examples where they speed up the convergence by considering differences of the second order in place of the first order. Their goal is to ease the process of pushing the line described in Example 10. Overall, they use their so-called $1/2$-*balancing principle*. We will not insist on this point because it is well explained in [25]. Essentially it leads to consider series with a factor $1/4^m$ in place of series with a factor $1/2^m$, hence a clear gain for the convergence speed. It must be noticed that the description of the method with a factor $1/16^m$ is rather optimistic. It works only for very specific examples, and a factor $1/4^m$ is the general case.

7.2. Crude method

The second method is to return to the definition of the Fourier coefficients.

Example 15 *(Dichopile algorithm, computation of the Fourier coefficients-2).* By definition, the Fourier coefficients of $\varphi(t)$ are

$$c_k = \int_0^1 \varphi(t) e^{-2\pi ikt}\, dt.$$

The change of variable $t = 1 + \log_2 x$ with $1/2 \le x \le 1$ (that already appears in Delange's article [2, p. 37]) transforms this formula into

$$c_k = \frac{1}{\ln 2}\int_{1/2}^1 \varphi(\log_2 x) e^{-2k\pi i \log_2 x}\frac{dx}{x} = c_k = \frac{1}{\ln 2}\int_{1/2}^1 \left(1 - \frac{1}{2}\log_2 x + \frac{g_5(x)}{x}\right) e^{-2k\pi i \log_2 x}\frac{dx}{x}.$$

Distinguishing the case $k = 0$, we have

$$c_0 = \frac{5}{4} + \frac{1}{\ln 2}\int_{1/2}^1 \frac{g_5(x)}{x^2}\,dx, \qquad c_k = -\frac{i}{4k\pi} + \frac{1}{\ln 2}\int_{1/2}^1 \frac{g_5(x)}{x^2} e^{-2k\pi i\log_2 x}\,dx. \tag{49}$$

A crude approach is to compute these integrals by the trapezoidal rule, with the nodes $j/2^K$ for $2^{K-1} \le j \le 2^K$ and a given K. Obviously, this uses the cascade algorithm. As $g_5(x)$ is Hölder with exponent α for every $0 < \alpha < 1$, the error is of order $1/2^{K\alpha}$ for every $0 < \alpha < 1$, that is essentially $1/2^K$. This method has the advantage of simplicity, but we cannot speed up the computation à la Richardson because there is no asymptotic expansion of the error. It gives only a rough estimation. Below are the values obtained with $K = 10$ on the left-hand side and with $K = 12$ on the right-hand side. The computations have been made with 15 digits to avoid the rounding errors. The correct digits are written in bold.

$$\begin{array}{ll}
c_0 \simeq -0.\mathbf{362}6476334 & c_0 \simeq -0.\mathbf{3627}354935 \\
c_1 \simeq +0.\mathbf{00328}58226 + 0.\mathbf{0019}776043i & c_1 \simeq +0.\mathbf{003284}7914 + 0.\mathbf{0019}880642i \\
c_2 \simeq +0.\mathbf{00306}95432 - 0.\mathbf{00062}87349i & c_2 \simeq +0.\mathbf{0030691}945 - 0.\mathbf{00062}20152i \\
c_3 \simeq +0.\mathbf{00168}57421 + 0.\mathbf{0012}276124i & c_3 \simeq +0.\mathbf{0016854}542 + 0.\mathbf{00123}35442i \\
c_4 \simeq +0.\mathbf{00054}24079 - 0.\mathbf{0011}900529i & c_4 \simeq +0.\mathbf{0005422}948 - 0.\mathbf{0011}8836225i \\
c_5 \simeq +0.\mathbf{00113}28340 + 0.\mathbf{0005}285686i & c_5 \simeq +0.\mathbf{00113}25212 + 0.\mathbf{00053}50851i \\
c_6 \simeq +0.\mathbf{000616}9547 + 0.\mathbf{0003}775004i & c_6 \simeq +0.\mathbf{000616}6015 + 0.\mathbf{00038}47058i \\
c_7 \simeq +0.\mathbf{000678}9418 - 0.\mathbf{0004}749207i & c_7 \simeq +0.\mathbf{0006785}365 - 0.\mathbf{00046}64313i \\
c_8 \simeq -0.\mathbf{00031}21218 + 0.\mathbf{0002}826099i & c_8 \simeq -0.\mathbf{00031}18626 + 0.\mathbf{00029}12762i \\
c_9 \simeq +0.\mathbf{00033}87154 - 0.\mathbf{0004}014520i & c_9 \simeq +0.\mathbf{00033}84147 - 0.\mathbf{00039}14805i \\
c_{10} \simeq +0.\mathbf{00050}26412 + 0.\mathbf{000}0947029i & c_{10} \simeq +0.\mathbf{00050}21373 + 0.\mathbf{00010}51334i
\end{array}$$

This permits us to draw a picture (Fig. 8) of the truncated Fourier series against the periodic function.

7.3. Mellin transform

If we look scrupulously to the proof of Theorem 1 and particularly to Formula (30), we see that the occurrence of periodic functions in the expansion of the sequence comes from the term

$$\varphi(t) = \rho^{1-\{t\}} e^{i\vartheta(1-\{t\})} f\left(B^{\{t\}-1}\right)\binom{1-\{t\}}{p} \tag{50}$$

with $f(x) = \boldsymbol{\ell f}(x)$ for some nonnegative integer p. This integer p is $\ell - m$ in the Chu–Vandermonde type Formula (30).

Let us focus on the case $p = 0$. With the change of variable $t = 1 + \log_B x$, $1/B \le x \le 1$, the Fourier coefficients of the 1-periodic function $\varphi(t)$ are

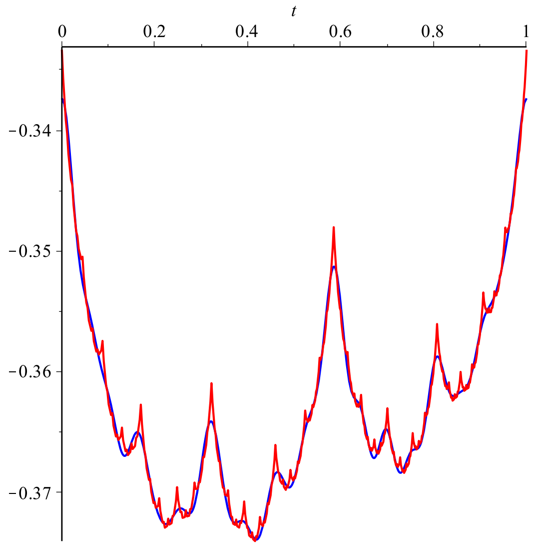

Fig. 8. The periodic function $\varphi(t)$ of the dichopile algorithm and its Fourier sum of order 10.

$$c_k = \int_0^1 \varphi(t) e^{-2\pi ikt}\, dt = \int_{1/B}^1 \rho^{\log_B x} e^{i\vartheta \log_B x} f(x) e^{-2\pi ik \log_B x} \frac{dx}{x \ln B},$$

that is

$$c_k = \frac{1}{\ln B} \int_{1/B}^1 x^{-1+\frac{\ln\rho + i\vartheta}{\ln B}} f(x) x^{-\chi_k}\, dx,$$

with $\chi_k = 2k\pi i/\ln B$. It turns out that these coefficients can be viewed as special values of a Mellin transform, namely values of

$$f^*(s) = \int_{1/B}^1 x^{s-1} f(x)\, dx$$

at points $s = -\chi_k + \log_B \lambda$ with $\lambda = \rho e^{i\vartheta}$ the eigenvalue of $\boldsymbol{Q}$ under consideration. According to Lemma 10, set out in Appendix A, we conversely know the value of the Mellin transform at positive integers,

$$f^*(\ell) = \sum_{r=1}^{B-1} m_{\ell,r}$$

where the $m_{\ell,r}$ are the partial moments of $f(x)$, defined by (A.1) in Appendix A.

Lemma 9. *The Mellin transform of the matrix-valued function $\boldsymbol{F}(x)$ defined by Lemma 5,*

$$\boldsymbol{F}^*(s) = \int_{1/B}^1 x^{s-1} \boldsymbol{F}(x)\, dx,$$

is an entire function, which can be expressed as the sum of a Newton series

$$\boldsymbol{F}^*(s) = \sum_{n=0}^{+\infty} \binom{s-1}{n} \Delta^n \boldsymbol{F}^*(1), \tag{51}$$

where Δ is the forward difference operator, acting on the variable s of $\boldsymbol{F}^*(s)$. A partial sum gives the value $\boldsymbol{F}^*(s)$ with an error of order the first neglected term.

Proof. It must be noticed that this function $\boldsymbol{F}^*(s)$ is analytic in the whole complex plane. Actually a Mellin transform is an integral from 0 to $+\infty$, and the behavior of the function at the endpoints of the interval determines the vertical strip in the complex plane where the Mellin transform is defined and analytic. Here the integrand vanishes in the neighborhood of 0 and of $+\infty$, so the transform is an entire function. We write the term x^{s-1} as $(1-(1-x))^{s-1}$ and we expand it by the binomial theorem. This gives

$$\boldsymbol{F}^*(s) = \sum_{n=0}^{+\infty} \binom{s-1}{n} \int_{1/B}^{1} \boldsymbol{F}(x)(x-1)^n\,dx.$$

Again with the binomial theorem, the last integral appears as the nth order difference of the sequence $\boldsymbol{F}^*(\ell)$ evaluated at $\ell = 1$. In other words we obtain Eq. (51). Moreover the integral expression of $\Delta^n \boldsymbol{F}^*(1)$ permits to show that it behaves like $(\boldsymbol{F}(1-1/B)+o(1))(1-1/B)^{n+1}/(n+1)$. As a consequence, the series converges at least as fast as a geometric series of ratio $1-1/B$. Hence the assertion about the error term holds. $\square$

The above result provides a third method to compute the Fourier coefficients. It may suggest that the calculation is particularly simple. This is not quite true, as the following example demonstrates.

Example 16 *(Dichopile algorithm, computation of the Fourier coefficients-3).* According to (49), we have to compute the integrals

$$e_k = \int_{1/2}^{1} g_5(x) x^{-\chi_k-2}\,dx = g_5^*(-1-\chi_k)$$

and $g_5^*(s)$ expresses as

$$g_5^*(s) = \sum_{n=0}^{+\infty} \binom{s-1}{n} \Delta^n g_5^*(1) = \sum_{n=0}^{+\infty} \binom{s-1}{n} \Delta^n m_{1,1}^{(1,5)},$$

where Δ acts on the first lower index ℓ of $m_{\ell,1}^{(1,5)}$. Let us point out that $m_{\ell,1}^{(1,5)}$ is the fifth component of the vector $\boldsymbol{m}_{\ell,1}^{(1)}$, which is itself the partial moment number 1 of order ℓ for the vector-valued function $\boldsymbol{f}^{(1)}(x) = \boldsymbol{g}(x)$. At this point, two problems arise. First the moments $m_{\ell,1}^{(1,5)}$ are of order $1/\ell$, while the differences $\Delta^n m_{1,1}^{(1,5)}$ are of order $2^n/n$ according to the proof of the above proposition. Hence there is a strong cancellation in the computation of differences $\Delta^n m_{1,1}^{(1,5)}$. To take this phenomenon into account, we do not compute these differences as floating point numbers but exactly as they are rational numbers. Second for $s = -1+\chi_k$ the modulus of $\binom{s-1}{n}/(2^n n)$ increases first up to a maximal value of order

$$10^{2.06|k|-0.65\ln|k|-1.08}$$

obtained for $n \simeq 5|k|$, and next decreases towards 0. Henceforth, to obtain the sum within an error $10^{-m}/2$ we have to sum the terms until they become smaller than $10^{-m}/2$ and we use mantissæ of length $2.06|k| - 0.65\ln|k| - 1 + m + 5$ (with 5 digits to control the rounding errors). Practically, we sum the series up to the nth term with $n \simeq (\pi/\ln(2))^2|k| + m\log_2 10 - 3/2\log_2|k| \simeq 20.5|k| + 3.3m - 2.2\ln|k|$. This analysis is not valid in the case $k = 0$, where the absolute value of the sequence is decreasing from the beginning.

All things considered, the description of the periodic functions that appear in the asymptotic expansion can be made either by Fourier series or by dilation equations and a change of variable. Because of the Hölderian character of the periodic function, say with exponent α in $(0,1)$, the difference between the function and the partial Fourier sum of order n is of order $O(\ln(n)/n^\alpha)$ and the convergence is rather slow [71, Th. II.10.8], as it can be seen in Figs. 7 and 8. It is striking to see that in the numerous articles that study digital sequences, it is very rare that a drawing with a partial Fourier sum illustrates the convergence towards the periodic function, even if the Fourier coefficients have been computed. This is a good reason to consider that Fourier series are illusory while the dilation equations are a better tool. In the frequent case where the linear representation has rational components, a dilation equation permits us to compute the values of its solution exactly on a everywhere dense set. Only the change of variable introduces floating numbers for the values of the periodic function at points $\log_B N \bmod 1$.

Appendix A. Moments

For a function $f(x)$ continuous on the interval $[0,1]$, we define its moments and partial moments, respectively

$$m_\ell = \int_0^1 f(z) z^{\ell-1}\,dz, \qquad m_{\ell,r} = \int_{r/B}^{(r+1)/B} f(z) z^{\ell-1}\,dz, \tag{A.1}$$

where ℓ is a positive integer and $0 \le r < B$ with $B \ge 2$. Remarkably for the solution of a dilation equation of the type considered in Section 3.3, the moments and the partial moments can be computed exactly. Similar computations appear in [72, 73] (Cantor distribution) and [74, Eq. 3.1 or 4.2] (coefficients for wavelets) or [54, p. 5] (computation of a Riemann–Stieltjes integral).

Lemma 10. *With the notations of Section 3.3 and under the assumption that none of the numbers $B^\ell \rho$, $\ell \ge 1$, is an eigenvalue of $\boldsymbol{Q}$, the recurrences (with $\boldsymbol{W}_b = \boldsymbol{A}_b \boldsymbol{V}$ for $0 \le b < B$)*

$$B^\ell \boldsymbol{M}_\ell \boldsymbol{J} - \boldsymbol{Q}\,\boldsymbol{M}_\ell = \frac{1}{\ell} \sum_{b=0}^{B-2} \big(B^\ell - (b+1)^\ell\big) \boldsymbol{W}_b + \sum_{k=1}^{\ell-1} \binom{\ell-1}{k-1} \left(\sum_{r=0}^{B-1} r^{\ell-k} \boldsymbol{A}_b \right) \boldsymbol{M}_k, \tag{A.2}$$

$$\boldsymbol{M}_{\ell,r} \boldsymbol{J} = \sum_{b<r} \frac{(r+1)^\ell - r^\ell}{\ell B^\ell} \boldsymbol{W}_b + \frac{1}{B^\ell} \sum_{k=1}^{\ell} \binom{\ell-1}{k-1} r^{\ell-k} \boldsymbol{M}_k, \quad 0 \le r < B, \tag{A.3}$$

determine all the moments and partial moments of the matrix-valued function $\boldsymbol{F}(x)$ defined by Lemma 5.

Proof. Using the dilation equation (18) we find readily (A.3) by a mere change of variable and the binomial theorem.

It may look troublesome to have to solve Eq. (A.3), as it involves multiples of $\boldsymbol{M}_\ell$ on both the left-hand and the right-hand sides. However, we can emphasize the structure of the Jordan cell by writing it $\boldsymbol{J} = \rho\,\mathbf{I}_{[0,\nu)} + \boldsymbol{N}$ and the structure of the $d \times \nu$ matrix $\boldsymbol{M}_\ell$ by viewing it as the collection of its columns $\boldsymbol{m}_\ell^{(0)}, \boldsymbol{m}_\ell^{(1)}, \dots, \boldsymbol{m}_\ell^{(\nu-1)}$. We add $\boldsymbol{m}_\ell^{(-1)} = \mathbf{0}$ for convenience. With these notations (A.3) rewrites

$$\big(B^\ell \rho\,\mathbf{I}_d - \boldsymbol{Q}\big) \boldsymbol{m}_\ell^{(j)} = -B^\ell \boldsymbol{m}_\ell^{(j-1)} + \frac{1}{\ell} \sum_{b=0}^{B-2} \big(B^\ell - (b+1)^\ell\big) \boldsymbol{w}_b^j + \sum_{k=1}^{\ell-1} \binom{\ell-1}{k-1} \left(\sum_{r=0}^{B-1} r^{\ell-k} \boldsymbol{A}_b \right) \boldsymbol{m}_k^{(j)}, \quad 0 \le j < \nu, \tag{A.4}$$

because of the equality $\boldsymbol{M}_\ell \boldsymbol{N} = (\mathbf{0}, \boldsymbol{m}_\ell^{(0)}, \boldsymbol{m}_\ell^{(1)}, \dots, \boldsymbol{m}_\ell^{(\nu-2)})$. It is now clear that Eq. (A.4) enables us to compute successively $\boldsymbol{m}_1^{(0)}, \boldsymbol{m}_1^{(1)}, \dots, \boldsymbol{m}_1^{(\nu-1)}, \boldsymbol{m}_2^{(0)}, \dots, \boldsymbol{m}_2^{(\nu-1)}$ and more generally all the moments, if the numbers $B^\ell \rho$ are not eigenvalues of $\boldsymbol{Q}$. Last, Eq. (A.3) provides us with the partial moments because $\boldsymbol{J}$ is invertible. $\square$

Example 17 *(Dichopile algorithm, computation of the moments).* For the dichopile algorithm, Lemma 10 translates into the formulæ

$$\big(2^{\ell+1} \mathbf{I}_6 - \boldsymbol{Q}\big) \boldsymbol{m}_\ell^{(0)} = \frac{2^\ell - 1}{\ell} \boldsymbol{A}_0 \boldsymbol{v}_2^{(0)} + \boldsymbol{A}_1 \sum_{j=1}^{\ell-1} \binom{\ell-1}{j-1} \boldsymbol{m}_j^{(0)}, \tag{A.5}$$

$$\big(2^{\ell+1} \mathbf{I}_6 - \boldsymbol{Q}\big) \boldsymbol{m}_\ell^{(1)} = \frac{2^\ell - 1}{\ell} \boldsymbol{A}_0 \boldsymbol{v}_2^{(1)} + \boldsymbol{A}_1 \sum_{j=1}^{\ell-1} \binom{\ell-1}{j-1} \boldsymbol{m}_j^{(1)} - 2^\ell \boldsymbol{m}_\ell^{(0)}, \tag{A.6}$$

$$2\boldsymbol{m}_{\ell,1}^{(0)} = \frac{1 - 1/2^\ell}{\ell} \boldsymbol{A}_0 \boldsymbol{v}_2^{(0)} + \frac{1}{2^\ell} \boldsymbol{A}_1 \sum_{j=1}^{\ell} \binom{\ell-1}{j-1} \boldsymbol{m}_j^{(0)},$$

$$2\boldsymbol{m}_{\ell,0}^{(1)} = \frac{1 - 1/2^\ell}{\ell} \boldsymbol{A}_0 \boldsymbol{v}_2^{(1)} + \frac{1}{2^\ell} \boldsymbol{A}_1 \sum_{j=1}^{\ell} \binom{\ell-1}{j-1} \boldsymbol{m}_j^{(1)} - \boldsymbol{m}_{\ell,1}^{(0)}. \tag{A.7}$$

Since some components of $\boldsymbol{f}(x) = \boldsymbol{f}^{(0)}(x)$ and $\boldsymbol{g}(x) = \boldsymbol{f}^{(1)}(x)$ are known, we can verify the result of the computation in that particular cases.

References

[1] J.R. Trollope, An explicit expression for binary digital sums, Math. Mag. 41 (1968) 21–25.
[2] H. Delange, Sur la fonction sommatoire de la fonction “somme des chiffres”, Enseign. Math. (2) 21 (1) (1975) 31–47.
[3] M. Drmota, J. Gajdosik, The distribution of the sum-of-digits function, J. Théor. Nr. Bordx. 10 (1) (1998) 17–32.
[4] G.F. Clements, B. Lindström, A sequence of (±1)-determinants with large values, Proc. Amer. Math. Soc. 16 (1965) 548–550.
[5] M.D. McIlroy, The number of 1’s in binary integers: bounds and extremal properties, SIAM J. Comput. 3 (1974) 255–261.

[6] D.W. Boyd, J. Cook, P. Morton, On sequences of ± 1's defined by binary patterns, Dissertationes Math. (Rozprawy Mat.) 283 (1989) 64.
[7] P. Flajolet, L. Ramshaw, A note on Gray code and odd–even merge, SIAM J. Comput. 9 (1) (1980) 142–158.
[8] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet–Thue–Morse sequence, in: Sequences and Their Applications, Singapore, 1998, in: Springer Ser. Discrete Math. Theor. Comput. Sci., Springer, London, 1999, pp. 1–16.
[9] D.J. Newman, On the number of binary digits in a multiple of three, Proc. Amer. Math. Soc. 21 (1969) 719–721.
[10] J. Coquet, A summation formula related to the binary digits, Invent. Math. 73 (1) (1983) 107–115.
[11] W. Rudin, Some theorems on Fourier coefficients, Proc. Amer. Math. Soc. 10 (1959) 855–859.
[12] H.S. Shapiro, Extremal problems for polynomials and power series, Master's thesis, Massachusets Institute of Technology, 1951.
[13] F. Morain, J. Olivos, Speeding up the computations on an elliptic curve using addition-subtraction chains, RAIRO Inform. Théor. Appl. 24 (6) (1990) 531–543.
[14] K.J. Supowit, E.M. Reingold, Divide and conquer heuristics for minimum weighted Euclidean matching, SIAM J. Comput. 12 (1) (1983) 118–143.
[15] K.B. Stolarsky, Power and exponential sums of digital sums related to binomial coefficient parity, SIAM J. Appl. Math. 32 (4) (1977) 717–730.
[16] A.H. Osbaldestin, P. Shiu, A correlated digital sum problem associated with sums of three squares, Bull. Lond. Math. Soc. 21 (4) (1989) 369–374.
[17] J. Brillhart, P. Erdős, P. Morton, On sums of Rudin–Shapiro coefficients. II, Pacific J. Math. 107 (1) (1983) 39–69.
[18] J. Coquet, P. Van Den Bosch, A summation formula involving Fibonacci digits, J. Number Theory 22 (2) (1986) 139–146.
[19] J.-M. Dumont, A. Thomas, Systemes de numération et fonctions fractales relatifs aux substitutions, Theoret. Comput. Sci. 65 (2) (1989) 153–169.
[20] J.-M. Dumont, Summation formulae for substitutions on a finite alphabet, in: Number Theory and Physics, Les Houches, 1989, in: Springer Proc. Phys., vol. 47, Springer, Berlin, 1990, pp. 185–194.
[21] J.-M. Dumont, N. Sidorov, A. Thomas, Number of representations related to a linear recurrent basis, Acta Arith. 88 (4) (1999) 371–396.
[22] G. Barat, V. Berthé, P. Liardet, J. Thuswaldner, Dynamical directions in numeration, Ann. Inst. Fourier (Grenoble) 56 (7) (2006) 1987–2092.
[23] P. Flajolet, M. Golin, Mellin transforms and asymptotics. The mergesort recurrence, Acta Inform. 31 (7) (1994) 673–696.
[24] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, R.F. Tichy, Mellin transforms and asymptotics: digital sums, Theoret. Comput. Sci. 123 (2) (1994) 291–314.
[25] P.J. Grabner, H.-K. Hwang, Digital sums and divide-and-conquer recurrences: Fourier expansions and absolute convergence, Constr. Approx. 21 (2) (2005) 149–179.
[26] M. Drmota, W. Szpankowski, A master theorem for discrete divide and conquer recurrences, in: D. Randall (Ed.), Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, 2011, pp. 342–361.
[27] R. Béjian, H. Faure, Discrépance de la suite de Van der Corput, in: Séminaire Delange–Pisot–Poitou, 19e année: 1977/78, Théorie des nombres, Fasc. 1, Secrétariat Math., Paris, 1978, Exp. No. 13, 14.
[28] J. Brillhart, P. Morton, Über Summen von Rudin–Shapiroschen Koeffizienten, Illinois J. Math. 22 (1) (1978) 126–148.
[29] J.-P. Allouche, J. Shallit, The ring of k-regular sequences, Theoret. Comput. Sci. 98 (2) (1992) 163–197.
[30] P.J. Grabner, C. Heuberger, H. Prodinger, Counting optimal joint digit expansions, Integers 5 (3) (2005) #A09.
[31] L. Berg, M. Krüppel, A simple system of discrete two-scale difference equations, Z. Anal. Anwend. 19 (4) (2000) 999–1016.
[32] P.C. Allaart, K. Kawamura, Extreme values of some continuous nowhere differentiable functions, Math. Proc. Cambridge Philos. Soc. 140 (2) (2006) 269–295.
[33] M. Krüppel, De Rham's singular function, its partial derivatives with respect to the parameter and binary digital sums, Rostock. Math. Kolloq. 64 (2009) 57–74.
[34] R. Girgensohn, Digital sums and functional equations, Integers 12 (1) (2012) 141–160.
[35] S. Goldstein, K.A. Kelly, E.R. Speer, The fractal structure of rarefied sums of the Thue–Morse sequence, J. Number Theory 42 (1) (1992) 1–19.
[36] M. Drmota, P.J. Grabner, Analysis of digital functions and applications, in: Combinatorics, Automata and Number Theory, in: Encyclopedia Math. Appl., vol. 135, Cambridge Univ. Press, Cambridge, 2010, pp. 452–504.
[37] L.H.Y. Chen, H.-K. Hwang, V. Zacharovas, Distribution of the sum-of-digits function of random integers: a survey, arXiv:1212.6697.
[38] P. Dumas, Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences, Linear Algebra Appl. 438 (5) (2013) 2107–2126.
[39] J.-P. Allouche, J. Shallit, Automatic Sequences: Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.
[40] J. Cassaigne, Counting overlap-free binary words, in: STACS 93, Würzburg, 1993, in: Lecture Notes in Comput. Sci., vol. 665, Springer, Berlin, 1993, pp. 216–225.
[41] J. Berstel, C. Reutenauer, Rational Series and Their Languages, Monogr. Theoret. Comput. Sci. EATCS Ser., vol. 12, Springer-Verlag, Berlin, 1988.
[42] J. Sakarovitch, Elements of Automata Theory, Cambridge University Press, 2009.
[43] J. Oudinet, Approches combinatoires pour le test statistique à grande échelle, Ph.D. thesis, Université Paris-Sud XI, 2010.
[44] J. Oudinet, A. Denise, M.-C. Gaudel, A new dichotomic algorithm for the uniform random generation of words in regular languages, Theoret. Comput. Sci. 502 (2013) 165–176.
[45] G.-C. Rota, G. Strang, A note on the joint spectral radius, Ned. Akad. Wet. Proc. Ser. A 63 = Indag. Math. 22 (1960) 379–381.
[46] V.D. Blondel, The birth of the joint spectral radius: an interview with Gilbert Strang, Linear Algebra Appl. 428 (10) (2008) 2261–2264.
[47] V.D. Blondel, M. Karow, V.Y. Protasov, F.R. Wirth, Linear Algebra Appl. 428 (10) (2008) 2259–2404, Special Issue on the Joint Spectral Radius: Theory, Methods and Applications.
[48] I. Daubechies, J.C. Lagarias, Two-scale difference equations. I. Existence and global regularity of solutions, SIAM J. Math. Anal. 22 (5) (1991) 1388–1410.
[49] C.A. Micchelli, H. Prautzsch, Uniform refinement of curves, Linear Algebra Appl. 114/115 (1989) 841–870.
[50] I. Daubechies, J.C. Lagarias, Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals, SIAM J. Math. Anal. 23 (4) (1992) 1031–1079.
[51] C. Heil, Methods of solving dilation equations, in: Probabilistic and Stochastic Methods in Analysis, with Applications, Il Ciocco, 1991, in: NATO Adv. Stud. Inst. Ser., Ser. C, Math. Phys. Sci., vol. 372, Kluwer Acad. Publ., Dordrecht, 1992, pp. 15–45.
[52] O. Rioul, Simple regularity criteria for subdivision schemes, SIAM J. Math. Anal. 23 (6) (1992) 1544–1576.
[53] S. Tasaki, I. Antoniou, Z. Suchanecki, Deterministic diffusion, De Rham equation and fractal eigenvectors, Phys. Lett. A 179 (2) (1993) 97–102.
[54] S. Tasaki, T. Gilbert, J.R. Dorfman, An analytical construction of the SRB measures for baker-type maps, Chaos 8 (1998) 424–443, Chaos and irreversibility (Budapest, 1997).
[55] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 61, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
[56] N. Dyn, D. Levin, Subdivision schemes in geometric modelling, Acta Numer. 11 (2002) 73–144.
[57] R.M. Jungers, V.D. Blondel, On the finiteness property for rational matrices, Linear Algebra Appl. 428 (10) (2008) 2283–2295.
[58] J.N. Tsitsiklis, V.D. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate, Math. Control Signals Systems 10 (1) (1997) 31–40.
[59] G. Tenenbaum, Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires, in: The Mathematics of Paul Erdős, I, in: Algorithms Combin., vol. 13, Springer, Berlin, 1997, pp. 117–128.
[60] P.C. Allaart, K. Kawamura, The Takagi function: a survey, Real Anal. Exchange 37 (1) (2011) 1–54.

[61] J.C. Lagarias, The Takagi function and its properties, in: K. Matsumoto, S. Akiyama, K. Fukuyama, H. Nakadaand, H. Sugita, A. Tamagawa (Eds.), Functions and Number Theory and Their Probabilistic Aspects, RIMS Kokyuroku Bessatsu B34, 2012, pp. 153–189.

[62] P. Billingsley, Probability and Measure, third edn., Wiley Ser. Prob. Math. Stat., Prob. Math. Stat., John Wiley & Sons Inc., New York, 1995.

[63] Z. Lomnicki, S. Ulam, Sur la théorie de la mesure dans les espaces combinatoires et son application au calcul des probabilités: I. Variables indépendantes, Fund. Math. 23 (1934) 237–278.

[64] K. Kawamura, On the set of points where Lebesgue's singular function has the derivative zero, Proc. Japan Acad. Ser. A Math. Sci. 87 (9) (2011) 162–166.

[65] T. Bedford, Hölder exponents and box dimension for self-affine fractal functions, Constr. Approx. 5 (1989) 33–48.

[66] P. Dumas, H. Lipmaa, J. Wallén, Asymptotic behavior of a non-commutative rational series with a nonnegative linear representation, Discrete Math. Theor. Comput. Sci. 9 (1) (2007) 247–274.

[67] M. Krüppel, On the extrema and the improper derivatives of Takagi's continuous nowhere differentiable function, Rostock. Math. Kolloq. 62 (2007) 41–59.

[68] J. Brillhart, L. Carlitz, Note on the Shapiro polynomials, Proc. Amer. Math. Soc. 25 (1970) 114–118.

[69] H.-K. Hwang, Asymptotics of divide-and-conquer recurrences: Batcher's sorting algorithm and a minimum Euclidean matching heuristic, Algorithmica 22 (4) (1998) 529–546.

[70] G.H. Hardy, M. Riesz, The General Theory of Dirichlet's Series, Camb. Tracts Math. Math. Phys., vol. 18, Stechert-Hafner, Inc., New York, 1915.

[71] A. Zygmund, Trigonometric Series, Vol. I, II, third edn., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2002, with a foreword by Robert A. Fefferman.

[72] G.C. Evans, Calculation of moments for a Cantor–Vitali function, Amer. Math. Monthly 64 (8 part II) (1957) 22–27.

[73] F. Lad, W. Taylor, The moments of the Cantor distribution, Stat. Probab. Lett. 13 (1992) 307–310.

[74] W.C. Shann, C.C. Yen, Matrices and quadrature rules for wavelets, Taiwanese J. Math. 2 (4) (1998) 435–446.