Discrete Mathematics and Theoretical Computer Science DMTCS vol. **9**, 2007, 247–272

Asymptotic behaviour of a non-commutative rational series with a nonnegative linear representation

Philippe Dumas[1] and Helger Lipmaa[2] and Johan Wallén[3]

[1]*Algorithms Project, INRIA Rocquencourt, 78153 Le Chesnay Cedex, France*
`Philippe.Dumas@inria.fr`
[2]*Cybernetica AS and University of Tartu, Estonia*
[3]*Laboratory for Theoretical Computer Science, Helsinki University of Technology, P.O. Box 5400, FI-02015 TKK, Espoo, Finland*

received May 10, 2005, *revised September 10, 2007*, *accepted October 1st, 2007*.

We analyse the asymptotic behaviour in the mean of a non-commutative rational series, which originates from differential cryptanalysis, using elementary tools from analysis and linear algebra, and more sophisticated tools from analytic number theory. We show that a probability distribution function describes the asymptotic behaviour of the rational series according to the length of words. As a result, the non-classical rational sequence, obtained by interpreting this rational series via the octal numeration system, admits an oscillating asymptotic behaviour for its first-order summation function. The distribution function and the periodic function are differentiable almost everywhere and not differentiable on an everywhere dense set. We compute the moments of the distribution function using a functional equation, which brings to light a self-similarity phenomenon, and we derive a Fourier representation of the periodic function using a Dirichlet series with vector coefficients. The method is applicable to a wide class of sequences rational with respect to a numeration system essentially under the condition that they admit a linear representation with nonnegative coefficients.

Keywords: rational series, rational sequence, numeration system, self-similarity, asymptotics

1 Introduction

Sequences that are rational with respect to a numeration system occur in essentially two domains. The first concerns automatic sequences—that is, sequences generated by finite machines that result from interpreting integers as words via a numeration system. While their definitions are frequently very simple, such as a paper-folding sequence defined as the sequence of ridges and valleys obtained by unfolding a sheet of paper that has been folded in half again and again, the behaviour of these sequences may be really difficult to study and the answers, if they exist, use sophisticated or elementary but intricate arguments Allouche and Shallit (2003).

The second domain where these sequences arise is the study of the complexity of algorithms that use a divide-and-conquer strategy. In this context, the sequences have real or integer values, because they provide the cost of an algorithm for a given input size. For divide-and-conquer recurrences, sequences that are rational with respect to a numeration system are as fundamental as classical rational sequences for classical linear recurrences with constant coefficients. This class of non-classical rational sequences extends the class of automatic sequences, since automatic sequences are non-classical rational sequences that take only a finite number of values. Computer scientists often neglect their fine asymptotic behaviour and content themselves with a rough estimation (Cormen et al., 1990, Master theorem). In fact these sequences hide functions of a fractal nature, periodic in a logarithmic scale.

We are dealing with a particular problem that serves as a guide for developing the theory of the asymptotic behaviour of these non-classical rational sequences. Our motivation for developing this example in detail is that the method should be applicable to a wide class of non-classical rational sequences. Indeed, the use initiated by Philippe Flajolet of the Mellin-Perron formula, a tool from analytic number theory, in the study of the complexity of divide-and-conquer algorithms, digital sums or sequences related to binary numeration system has produced many papers, but no general result which allows to deduce systematically the asymptotic behaviour of the sequence under consideration from the recurrence it satisfies. (See (Flajolet and Golin, 1994), (Flajolet et al., 1994) as seminal works, or even (Flajolet, 1985), (Flajolet and Martin, 1985) for the idea of a periodic function as a consequence of a set of regularly spaced complex poles.)

We are considering first a rational series from formal language theory and then a rational sequence. A rational series associates a number to each word over a given alphabet. It can be defined by a linear representation—that is, a set of matrices—and the size of the matrices is called the dimension of the representation. The key condition for our study is the fact that the rational series admits a linear representation whose coefficients are all nonnegative. An additional condition on the Jordan reduced form of a certain matrix provides the behaviour of the first-order means. This last point provides a new argument. In previous studies, the qualitative behaviour of the first-order means is always obtained by some *ad hoc* argument, specific to each example.

As usual in the study of the fine behaviour of digital sums or sequences of this type, a periodic function of a fractal nature arises. The word "fractal" is generally used in a vague meaning in this context, but we show that a self-similar function occurs in the problem. The periodic function in our example is Hölder of order $2/3$, and $2/3$ is the best possible exponent, even locally. Moreover, the subset of points where this function is differentiable is an everywhere dense subset, and also the set where this function is non-differentiable is an everywhere dense subset of the real line. It seems that this mixed property appears for the first time in this domain.

The paper is organised as follows. Section 2 shows the cryptographic origin of the problem, and introduces the rational series adp, which is the subject of our study. Section 3 provides an overview of the method we follow to put light on the asymptotic behaviour of adp. We show in Section 4 how to obtain the qualitative asymptotic behaviour of the first-order Cesàro means of adp. This gives rise to a probability distribution function whose regularity is discussed. It turns out to be Hölder but non-differentiable on an everywhere dense set. Moreover in Section 5, we prove that the distribution function is the sum of the components of a vector function which satisfies a self-similarity property. This property permits us to compute exactly and efficiently values of the distribution function on an everywhere dense set and to compute exactly its moments. In Section 6, we translate the result into the asymptotic behaviour of a sequence sbs (for side-by-side) deduced from adp in a very simple manner. We obtain a periodic function of bounded variation and the rest of the paper is devoted to the explicit computation of its Fourier series. Section 7 is a preliminary one, where we provide the necessary tools of analytic number theory. Section 8 first provides an explicit asymptotic expansion for the second-order Cesàro means of sbs, and next for the first-order means of sbs.

Eventually the difference between our example and previous works (Flajolet and Golin, 1994), (Flajolet et al., 1994) is the high dimension of the linear representation, namely 8 in the first part (Sections 2–5) which is the most original, and 17 in the second part (Sections 6–8) which follows the same approach as the cited papers. This feature prevents us to use the smart but tricky arguments usually employed in the domain. Showing that rational series and rational sequences may be dealt in a systematic manner is indeed of interest.

2 Origin of the problem and rational series adp

In this paper we aim to present a detailed asymptotic analysis of a rational series adp. Recall (Sakarovitch, 2003) that a non-commutative formal power series S over the monoid of words from an alphabet $\mathcal{A}$ with real coefficients is a map which associates to each word w a real number $S(w)$, called the coefficient of w in S. Classically the series is written as a formal sum

$$S = \sum_{w \in \mathcal{A}^*} S(w) \cdot w,$$

indexed by the words. It is said to be a formal series in the non-commutative indeterminates which are the letters from $\mathcal{A}$.

Definition 1 *A non-commutative formal series is a rational series if and only if there is a square matrix A_x of size $d \times d$ for each letter x in the alphabet, a row vector L and a column vector C such that for each word $w = w_1 \cdots w_\ell$, the coefficient of w in the series is $S(w) = LA_{w_1} \cdots A_{w_\ell} C$. The family L, $(A_x)_x$, C is called a linear representation of dimension d of the rational series S.*

The series adp is motivated by differential cryptanalysis. Differential cryptanalysis (Biham and Shamir, 1991) is one of the most widely used methods for analysing symmetric ciphers. This method studies how differences propagate through functions. Let G, H be finite Abelian groups, and let $f\colon G \to H$ be a function. For each $\alpha \in G$, let D_α denote the difference operator $(D_\alpha f)(x) = f(x+\alpha) - f(x)$. Differential cryptanalysis is especially concerned with the probability that $(D_\alpha f)(x)$ has a given value $\beta \in H$. The mapping $(\alpha, \beta) \mapsto \Pr_{x \in G}[(D_\alpha f)(x) = \beta]$ is called the differential probability of f. We will consider the group $G_N = \{0, 1, \ldots, 2^N - 1\}$ with the usual addition modulo 2^N as the group operation. We identify G_N and the set $\mathbb{Z}_2^N$ of N-tuples of bits using the natural correspondence that maps $x_{N-1}2^{N-1} + \cdots + x_1 2 + x_0 \in G_N$ onto $(x_{N-1}, \ldots, x_1, x_0) \in \mathbb{Z}_2^N$. In this way the usual componentwise addition $\oplus$ in $\mathbb{Z}_2^N$ (or bitwise exclusive-or) carries over to a function $G_N \times G_N \to G_N$. We call the differential probability of this mapping the *additive differential probability* of exclusive-or and denote it by adp,

$$\mathrm{adp}(\alpha, \beta, \gamma) = \Pr_{x,y}[((x+\alpha) \oplus (y+\beta)) - (x \oplus y) = \gamma].$$

We will consider adp as a function of octal words by encoding the tuple (α, β, γ) as the octal word $w = w_{N-1} \cdots w_0$, where $w_i = \alpha_i 4 + \beta_i 2 + \gamma_i$. This defines adp as a function from the octal words of length N to the interval $[0, 1]$. As N varies in the set of nonnegative integers, we obtain a map over the octal words with values in $[0, 1]$, that is a formal series

$$\mathrm{adp} = \sum_{w \in \{0,1,\ldots,7\}^*} \mathrm{adp}(w) \cdot w.$$

The key result which permits us to investigate the asymptotic behaviour of adp is the following (Lipmaa et al., 2004, Theorem 1).

Lemma 1 *The formal series* adp *is a rational series over the alphabet of the octal numeration. It admits the* 8*-dimensional linear representation* L, A_k, $0 \leq k < 8$, C, *where*

$$L = \begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 & 1\end{pmatrix}, \qquad C = \begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix}^T,$$

$$A_0 = \frac{1}{4}\begin{pmatrix}4 & 0 & 0 & 1 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix},$$

and $A_k = T_k A_0 T_k$ *for* $0 < k < 8$, *with* T_k *the permutation matrix associated to the translation* $i \mapsto i \oplus k$.

3 Overview of the behaviour of adp in the mean

We will give a detailed asymptotic average-case analysis of the formal series adp. The analysis proceeds as follows.

We first view the family $(\mathrm{adp}(w)/4^N)_{|w|=N}$ as a probability distribution μ_N on the real segment $[0, 1]$ by interpreting the word w as the real number whose octal expansion is $(0.w)_8$. For each N, we have an associated distribution function $F_N(x)$. Using the linear representation for adp, we prove a limit theorem stating that the sequence of distribution functions converges uniformly to a distribution function $F(x)$. It turns out that it is Hölder of order $2/3$, and that this exponent is the best possible. As a by-product, we show that $F(x)$ is differentiable on an everywhere dense subset, and also non-differentiable on an everywhere dense subset (Section 4). Moreover $F(x)$ writes $L\mathbf{F}(x)$ where the vector function $\mathbf{F}(x)$ satisfies a self-similarity property, which explains the fractal character of $F(x)$. This permits us to compute the moments of the distribution function $F(x)$ (Section 5).

Second, we introduce the sequence sbs obtained by putting side-by-side the values of $\mathrm{adp}(w)$ according to their length and rank in the lexicographic ordering of the octal words w. The limit theorem translates to a formula

$$\sum_{n<\nu} \mathrm{sbs}(n) = \nu^{2/3} G_{2/3}(\log_8 \nu) + O(\nu^{1/3}), \tag{1}$$

where $G_{2/3}$ is a 1-periodic function, which inherits the properties of F. In particular, the function is of bounded variation. As a consequence, its Fourier series converges uniformly towards $G_{2/3}$ (Section 6).

Third, the linear representation of adp gives rise to a 17-dimensional linear representation of the 8-rational sequence sbs. We consider the seventeen sequences associated with the linear representation for $\mathrm{sbs}(n)$ by taking each canonical basis vector of $\mathbb{Q}^{17}$ as the column vector. Let U_n be the row vector of these sequences, and let $U(s)$ be its Dirichlet series. Each sequence is bounded and the Dirichlet series have abscissa of convergence that is not greater than 1. The function $U(s)$ is analytic for $\sigma > 1$ and satisfies the functional equation

$$U(s)(I_{17} - 8^{-s}Q') = \nabla U(s), \tag{2}$$

where Q' is the sum of the square matrices in the linear representation for sbs and the function $\nabla U(s)$ is analytic for $\sigma > 0$. This formula provides a meromorphic extension of $U(s)$ to $\sigma > 0$. The rightmost possible singularities have $\sigma = 2/3$. A change of coordinates using a Jordan form $J = P^{-1}Q'P$ transforms U_n and $U(s)$ into the sequence V_n and its Dirichlet series $V(s)$. This implies that $2/3$ is indeed a singularity for the first component $v^1(s)$ of $V(s)$. The singularities of the other components have $\sigma \le 1/3$. Finally, the order of $U(s)$, *i.e.* the exponent in the order of growth at $\pm i\infty$ as a function of σ, is at most $1 - \sigma$ for $0 < \sigma < 1$ and 0 for $\sigma > 1$ (Section 7).

Fourth, we apply a Mellin-Perron formula to get an integral expression for the sums of the second-order of sbs. The integral is evaluated using the residue theorem by pushing the vertical line of integration to the left. This gives the asymptotic expansion

$$\sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} \mathrm{sbs}(k) \underset{\nu\to\infty}{=} \nu^{5/3} H_{5/3}(\log_8 \nu) + \nu^{4/3} H_{4/3}(\log_8 \nu) + O(\nu^{1+\epsilon}),$$

where $H_{5/3}$ and $H_{4/3}$ are 1-periodic continuous functions and $0 < \epsilon < 1/3$. A pseudo-Tauberian argument combined with (1) gives the Fourier series

$$G_{2/3}(\lambda) = \frac{1}{\ln 8} \sum_{k\in\mathbb{Z}} \frac{\nabla v^1(2/3 + k\chi)}{2/3 + k\chi} e^{2\pi i k\lambda},$$

where $\nabla v^1(s)$ is an analytic function and $\chi = 2\pi i/\ln 8$. The Fourier series converges uniformly towards $G_{2/3}$ (Section 8).

4 Limit theorem

Our point of departure is the following result. The section is devoted to its proof and immediate consequences.

Theorem 1 *There exists a distribution function $F(x)$ such that the summation function of $\mathrm{adp}(w)$ for words of length N satisfies*

$$\sum_{\substack{|w|=N \\ (w)_8 < 8^N x}} \mathrm{adp}(w) \underset{N\to\infty}{=} 4^N \cdot F(x) + O(2^N)$$

uniformly for $x \in [0, 1)$. Moreover, the function $F(x)$ is Hölder of order $2/3$ and $2/3$ is the best exponent for $F(x)$.

4.1 Convergence

Let L, $A_0, \ldots, A_7$, and C be the linear representation of adp, and let $Q = A_0 + A_1 + \cdots + A_7$. The matrix Q contains almost all the knowledge we need in order to study the asymptotic behaviour of adp. It has a dominant eigenvalue 4, which is simple. More precisely, the matrix Q is diagonalizable with eigenvalues 4 (simple), 2 (triple), 1 (quadruple). The vector $V = (E_1 + \cdots + E_8)/8$, where (E_j) is the canonical basis, is an eigenvector for the eigenvalue 4. With this choice, we have $C = V + V_2 +$

V_1 where V_2 and V_1 are some eigenvectors associated to the eigenvalues 2 and 1 respectively. As a consequence, $Q^n C = 4^n V + 2^n V_2 + V_1$ for each nonnegative integer n.

For each integer $N > 0$, we define a probability distribution μ_N on the real segment $[0,1]$ by its distribution function

$$F_N(x) = 4^{-N} \sum_{\substack{|w|=N \\ (w)_8 < 8^N x}} \mathrm{adp}(w) = 4^{-N} \sum_{\substack{|w|=N \\ (w)_8 < 8^N x}} L A_w C,$$

where $(w)_8 = w_{|w|-1} 8^{|w|-1} + w_{|w|-2} 8^{|w|-2} + \cdots + w_0$ is the octal integer represented by w and $A_w = A_{w_{|w|-1}} \cdots A_{w_0}$. This is indeed a probability distribution, since

$$\mu_N([0,1]) = \frac{1}{4^N} \sum_{|w|=N} \mathrm{adp}(w) = \frac{1}{4^N} L Q^N C = 1, \tag{3}$$

as can be easily checked.

We need a result which is no more than a mere remark, but which is basic in our study. The rational series adp admits the linear representation L, A_0, …, A_7, C. Varying the vector C, we obtain a vector space of rational series which is nothing else than the vector space generated by adp under the action of the monoid of octal words.

Lemma 2 *Let $\mathcal{V}$ be a bounded set of column vectors. Then all rational series with a linear representation which has the same matrices L, A_0, …, A_7 as* adp *but an arbitrary column vector from $\mathcal{V}$ admit a common upper bound.*

Proof: We observe that the maximum absolute column sum norm of each square matrix in the linear representation of adp is 1. □

We are now in position to prove the convergence of the sequence (F_N).

Lemma 3 *The sequence of distribution functions (F_N) converges uniformly to a distribution function F with a speed of convergence of order 2^{-N}.*

Proof: We will prove that $(F_N(x))$ is a Cauchy sequence. Let M and N two integers with $0 \le M \le N$. For $x \in [0,1)$, let k be the integer part of $8^M x$. The nonnegative character of adp provides us with the inequalities

$$4^{-N} \sum_{\substack{|w|=N \\ (w)_8 \le 8^{N-M} k}} \mathrm{adp}(w) \le F_N(x) \le 4^{-N} \sum_{\substack{|w|=N \\ (w)_8 \le 8^{N-M}(k+1)}} \mathrm{adp}(w).$$

Substituting $w'w''$ for w with w' a word of length M and w'' a word of length $N-M$, the inequality $(w)_8 \le 8^{N-M} k$ becomes $(w')_8 8^{N-M} + (w'')_8 \le 8^{N-M} k$. This is certainly the case if $(w')_8 \le k-1$. Similarly the inequality $(w)_8 \le 8^{N-M}(k+1)$ becomes $(w')_8 8^{N-M} + (w'')_8 \le 8^{N-M}(k+1)$. It implies $(w')_8 \le k+1$. From this remark and anew the nonnegative character of adp follow the inequalities

$$4^{-N} \sum_{\substack{|w'|=M \\ |w''|=N-M \\ (w')_8 \le k-1}} \mathrm{adp}(w'w'') \le F_N(x) \le 4^{-N} \sum_{\substack{|w'|=M \\ |w''|=N-M \\ (w')_8 \le k+1}} \mathrm{adp}(w'w'').$$

By making $\mathrm{adp}(w'w'') = LA_{w'}A_{w''}C$ explicit we obtain

$$4^{-N} \sum_{\substack{|w'|=M \\ (w')_8 \leq k-1}} LA_{w'}Q^{N-M}C \leq F_N(x) \leq 4^{-N} \sum_{\substack{|w'|=M \\ (w')_8 \leq k+1}} LA_{w'}Q^{N-M}C.$$

Besides,

$$F_M(x) = 4^{-M} \sum_{\substack{|w'|=M \\ (w')_8 < 8^M x}} LA_{w'}C.$$

Then the difference $F_N(x) - F_M(x)$ is bounded by the absolute value of

$$4^{-N} \sum_{\substack{|w'|=M \\ (w')_8 \leq k}} LA_{w'}Q^{N-M}C - 4^{-M} \sum_{\substack{|w'|=M \\ (w')_8 \leq k}} LA_{w'}C = 4^{-N} \sum_{\substack{|w'|=M \\ (w')_8 \leq k}} LA_{w'}(Q^{N-M}C - 4^{N-M}C)$$

plus or minus a term $4^{-N}LA_{w'}Q^{N-M}C$, corresponding to $(w')_8 = \ell$ with $\ell = k-1$ or $k+1$, and a term $4^{-M}LA_{w'}C$, corresponding to $(w')_8 = k$, which are of order 4^{-M} and negligeable in the problem, as we will see. We will prove that we have a bound of order 2^{-M}.

The difference $Q^{N-M}C - 4^{N-M}C$ writes $2^{N-M}(1-2^{N-M})V_2 + (1-4^{N-M})V_1$. We will deal with the term associated to V_2. The other associated to V_1 may be dealt with in the same manner. (It turns out to be of order 4^{-M}.) We then come up with $S_M(x) = L\mathbf{S}_M(x)$ where $\mathbf{S}_M(x)$ is defined by

$$\mathbf{S}_M(x) = \sum_{\substack{|w'|=M \\ (w')_8 < 8^M x}} A_{w'}V_2.$$

The sum $S_M(x)$ is similar to the sum which defines $F(x)$. We anticipate on the idea of Theorem 3. Let x' be the most significant figure of $8^M x$. The above sum writes

$$\mathbf{S}_M(x) = \sum_{r<x'} A_r Q^{M-1}V_2 + \sum_{\substack{|w''|=M-1 \\ (w'')_8 < 8^{M-1}(8x-x')}} A_{x'}A_{w''}V_2 = 2^{M-1}\sum_{r<x'} A_r V_2 + A_{x'}\mathbf{S}_{M-1}(8x-x').$$

Using the 1-norm for vectors and its induced norm for matrices, that is the maximum absolute column sum norm, we obtain immediately $\|\mathbf{S}_M(x)\| \leq 2^{M-1}\,\|Q\|\,\|V_2\| + \|\mathbf{S}_{M-1}(8x-x')\|$ (because the maximum absolute column sum norm of each matrix A_r, $0 \leq r \leq 7$ is 1) and for the supremum norm $\|\mathbf{S}_M\|_\infty \leq 2^{M-1}\,\|Q\|\,\|V_2\| + \|\mathbf{S}_{M-1}\|_\infty$. By induction, we conclude $\|\mathbf{S}_M(x)\| = O(2^M)$ with a constant in the big oh which is independant of x. The same result is valid for $S_M(x)$ and we get

$$4^{-N} \sum_{\substack{|w'|=M \\ (w')_8 \leq k}} 2^{N-M}(1-2^{N-M})LA_{w'}V_2 = 4^{-N}2^{N-M}(1-2^{N-M})O(2^M) = O(2^{-M}).$$

Eventually we obtain

$$F_N(x) - F_M(x) = O(2^{-M})$$

uniformly with respect to x. As a consequence $(F_N(x))$ is a Cauchy sequence which converges uniformly towards a function $F(x)$. Taking the limit over N we obtain $F(x) - F_M(x) = O(2^{-M})$ uniformly with respect to x. The limit function $F(x)$ is obviously a distribution function (because of $LV = 1$) and μ stands for the associated measure. □

4.2 Hölder condition

Recall that a function f is Hölder of order $\alpha \in (0, 1]$ if it satisfies $|f(x) - f(y)| \leq C|x - y|^\alpha$ for some constant C, and all x, y.

Lemma 4 *The limiting distribution function F is Hölder of order* $2/3$.

Proof: Let x, y be arbitrary such that $0 \leq x < y \leq 1$. The numbers x and y satisfy $8^{-(\nu+1)} \leq y - x < 8^{-\nu}$ for some well defined integer ν. We can find a semi-open interval $[u, v)$ such that its ends are octal numbers of the form $u = (0.u_1 \cdots u_\nu)_8$, $v = u + k8^{-\nu}$, where k is 1 or 2, and the semi-open interval $(x, y]$ is included in $[u, v)$. The interval $[u, v)$ may overhang outside of the interval $[0, 1]$, but we consider that distribution functions are extended by 0 on the left of 0 and by 1 on the right of 1. The number $u + k8^{-\nu}$ has an octal expansion of the form $u + k8^{-\nu} = (u'_0.u'_1 \cdots u'_\nu)_8$. For all $N \geq \nu$, the term $\mu_N\,[u, v)$ is $\mu_N\,[u, u + 8^{-\nu})$ or $\mu_N\,[u, u + 8^{-\nu}) + \mu_N\,[u + 8^{-\nu}, u + 2.8^{-\nu})$, depending on whether $k = 1$ or $k = 2$. Numbers from an interval like $[u, u + 8^{-\nu})$ have octal expansions $(0.u_1 \cdots u_\nu w_{\nu+1} \cdots)_8$ with figures $w_{\nu+i}$ arbitrary. As a consequence $\mu_N\,[u, v)$ admits an upper bound which is one term or the sum of two terms of the form

$$4^{-N} \sum_{\substack{|w|=N \\ u \leq (0.w_1 \cdots w_N)_8 < u + 8^{-\nu}}} \operatorname{adp}(w) = 4^{-N} \sum_{w_{\nu+1}, \cdots, w_N} LA_{u_1} \cdots A_{u_\nu} A_{w_{\nu+1}} \cdots A_{w_N} C,$$

that is, with $Q = A_0 + A_1 + \cdots + A_7$,

$$4^{-N} \sum_{\substack{|w|=N \\ u \leq (0.w_1 \cdots w_N)_8 < u + 8^{-\nu}}} \operatorname{adp}(w) = 4^{-N} LA_{u_1} \cdots A_{u_\nu} Q^{N-\nu} C.$$

The matrix Q appears in Lemma 3 and the formula $Q^{N-\nu}C = 4^{N-\nu}V + 2^{N-\nu}V_2 + V_1$ shows that

$$4^{-N} LA_{u_1} \cdots A_{u_\nu} Q^{N-\nu} C = 4^{-\nu} LA_{u_1} \cdots A_{u_\nu} V + \frac{1}{2^{N+\nu}} LA_{u_1} \cdots A_{u_\nu} V_2 + \frac{1}{4^N} LA_{u_1} \cdots A_{u_\nu} V_1$$

remains bounded when N goes through the integers. Therefore we obtain

$$\forall N \geq \nu, \quad \mu_N\,[u, v) \leq K 4^{-\nu},$$

for some constant K independent of N, ν, u, v, x, and y. But the hypothesis $8^{-(\nu+1)} \leq y - x < 8^{-\nu}$ gives $\mu_N\,[u, v) \leq 4K(y-x)^{2/3}$ and the fact that the positive limit measure μ is a non-decreasing function of sets, we have

$$0 \leq F(y) - F(x) = \mu\,(x, y] \leq \mu\,[u, v) \leq 4K(y - x)^{2/3}.$$

We conclude that the distribution function $F(x)$ is Hölder of order $2/3$. □

The Hölderian character of a function like $F(x)$ has been studied in (Dumont and Thomas, 1989), but in a slightly different context of unusual numeration system related to substitutions of words, and the employed argument not seems to be the same.

4.3 Best exponent

A natural question arises: is $2/3$ the best (that is, the largest) exponent that can be used in the Hölder condition for $F(x)$? We will show that there is no subinterval in which $F(x)$ satisfies a Hölder condition with an exponent larger than $2/3$. Towards this end, we will exhibit an inequality of the form

$$\liminf_{y\to x^+} \frac{F(y)-F(x)}{(y-x)^{2/3}} \geq K > 0.$$

Such an inequality cannot be valid for every x. We thus restrict ourselves to rational and even octal numbers x, and this is sufficient for our purpose.

Let x be an octal number and y be a real number satisfying $0 \leq x < y < 1$. As in the previous subsection, we consider octal expansions $x = (0.x_1 \cdots x_M)_8$, $y = (0.y_1 \cdots y_M y_{M+1} \cdots)_8$. Without loss of generality, we may assume that $8^{-(\nu+1)} \leq y - x < 8^{-\nu}$ and $\nu \geq M+1$. Once again, the nonnegative character of $\operatorname{adp}(w)$ gives for $N > \nu$ the inequality

$$\sum_{\substack{|w|=N\\ x\leq(0.w)_8<y}} \operatorname{adp}(w) \geq \sum_{\substack{|w|=N\\ x\leq(0.w)_8<x+8^{-(\nu+1)}}} \operatorname{adp}(w) = LA_{x_1}\cdots A_{x_M}A_0^{\nu+1-M}Q^{N-\nu-1}C,$$

where Q is still $A_0+\cdots+A_7$. For the same reason as above, we obtain $Q^{N-\nu-1}C = 4^{N-\nu-1}V + 2^{N-\nu-1}V_2 + V_1$. The term $LA_{x_1}\cdots A_{x_M}A_0^{\nu+1-M}(2^{N-\nu-1}V_2+V_1)$ is $O(2^{N-\nu})$ because all the rational series in the vector space generated from $\operatorname{adp}(w)$ under the action of the monoid of octal words are uniformly bounded for a column vector in a bounded set. Next the matrix A_0 is diagonalizable with a dominant eigenvalue 1 with eigenvector $V^0 = 1/4\,E_1$. The others eigenvalues are $1/4$ and 0. As a consequence V writes $V = V^0 + V^0_{1/4} + V^0_0$ where $V^0_{1/4}$ and V^0_0 are eigenvectors for the eigenvalues $1/4$ and 0 respectively. From these facts follows the equality

$$4^{N-\nu-1}LA_{x_1}\cdots A_{x_M}A_0^{\nu+1-M}V = 4^{N-\nu-1}LA_{x_1}\cdots A_{x_M}V^0 + 4^{N-\nu}O\left(4^{-\nu+M}\right) + 4^{N-\nu}O\left(2^{N-\nu}\right).$$

Hence, dividing by 4^N, we obtain

$$F_N(y) - F_N(x) \geq 4^{-\nu}\left(\frac{1}{4}LA_{x_1}\cdots A_{x_M}V^0 + O\left(4^{-\nu+M}\right) + O(2^{-N-\nu})\right).$$

(The consideration of $\left[x, x+8^{-(\nu+1)}\right)$ in place of $\left(x, x+8^{-(\nu+1)}\right]$ inserts a small error, which is not greater than $2/4^N$ because $\operatorname{adp}(w)$ is bounded by 1, and this will cause no problem.) In the limit this provides us with the formula

$$F(y) - F(x) \geq K4^{-\nu}$$

for some absolute constant (depending only on x) and next with

$$\liminf_{y\to x^+} \frac{F(y)-F(x)}{(y-x)^{2/3}} \geq K \geq \frac{1}{4}LA_{x_1}\cdots A_{x_M}V^0. \tag{4}$$

We are not far from the announced inequality.

The lower bound K is of interest only if it is positive. Hence we consider the rational sequence which admits the linear representation L, $A_0, \ldots, A_7$, and V^0, and we look for its support $\mathcal{L}$—that is, the rational

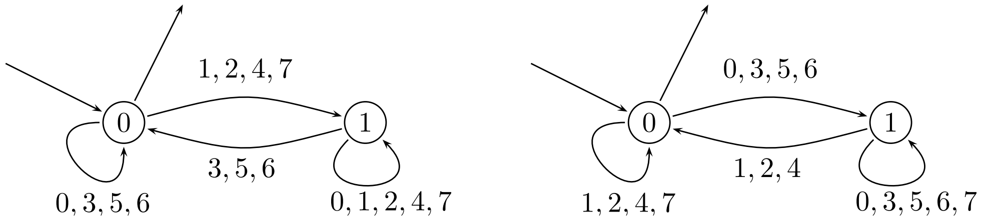

Fig. 1: The automaton on the left recognises the language $\mathcal{L}$, which is the support of the series adp.

language of words which give a nonzero value by this series. We let $S(w)$ denote the associated formal series. A word $w = w_0 w_1 \cdots w_{N-1}$ gives a positive value for the series S if and only if there is a positive term in the sum

$$S(w) = \sum_{i_0, i_1, \cdots, i_N} L_{i_0} A_{w_0, i_0, i_1} \cdots A_{w_{N-1}, i_{N-1}, i_N} V^0_{i_N}.$$

This means that we can construct the following nondeterministic automaton for recognising the words w such that $S(w) \neq 0$. The states are the indices of the matrices in the representation, say the integers from 1 to 8, augmented by an initial state $init$. The transitions are labelled by the figures $0, 1, \ldots, 7$ of the octal numeration, and the empty word ε. There is a transition labelled ε from the initial state $init$ to state i if and only if the coefficient L_i is positive. There is a transition labelled r from state i to state j if and only if the coefficient $A_{r,i,j}$ is positive. There are no other transitions. A state i is accepting if and only if the coefficient V^0_i is positive. The initial state is not accepting.

This automaton has nine states. If we turn this automaton into a minimal deterministic automaton (Sakarovitch, 2003), we obtain the automaton with only two states in Figure 1 on the left hand side. If the word corresponding to the octal number $x = (0.x_1 \cdots x_M)$ is accepted by the automaton, the lower bound (4) shows that the function $F(x)$ cannot satisfy a Hölder condition with a larger exponent than $2/3$ in a right neighbourhood of x. Moreover, these octal numbers are dense in $[0, 1)$, since every word is a prefix of a word accepted by the automaton (the graph is strongly connected).

Lemma 5 *The subset of numbers for which $2/3$ is locally the best exponent in the Hölder condition for $F(x)$ is everywhere dense.*

Additionally, we can deal with the left limit of the quotient $(F(x) - F(y))/(x - y)^{2/3}$ by reversing the roles of x and y. The only differences compared to the right limit is that we have to replace the matrix A_0 by the matrix A_7. The vector V^0 becomes $V^7 7 = 1/4\, E_8$, a fact which is not surprising in view of the definition of matrices A_k. We arrive at the same conclusion, using the automaton in Figure 1 on the right hand side. The notable point is that some octal numbers have a local Hölder exponent $2/3$ and some others have a greater local exponent.

4.4 Regularity result for the limit function

Because the limit function $F(x)$ is nondecreasing, it is differentiable almost everywhere (Billingsley, 1995, Theorem 3.12). Nevertheless, the function $F(x)$ cannot be differentiable in the everywhere dense set where $2/3$ is locally the best exponent in the Hölder condition.

Theorem 2 *The limit function $F(x)$ is differentiable almost everywhere. The set of points where $F(x)$ is not differentiable is everywhere dense.*

5 Self-similarity

5.1 Functional equation

For each vector E_i of the canonical basis we may consider the rational series which associates to an octal word w the value $f_i(w) = E_i^T A_w C$. Following exactly the same way as for adp we get

$$4^{-N} \sum_{\substack{|w|=N \\ (w)_8 \le 8^N x}} f_i(w) \underset{N\to+\infty}{=} F_i(x) + o(1)$$

for some continuous function $F_i(x)$. Because $L = E_1^T + \cdots + E_8^T$, we have $F(x) = F_1(x) + \cdots + F_8(x)$.

For $0 \le r < 8$, $r/8 \le x < (r+1)/8$ and $1 \le i \le 8$, and considering the leftmost letter of each word w, we may write ($A_{r,i,j}$ is the component i, j of matrix A_r)

$$\begin{aligned}
4^{N+1} F_i(x) + o(4^N) &= \sum_{\substack{|w|=N+1 \\ (w)_8 \le 8^N \cdot 8x}} E_i^T A_w C \\
&= \sum_{0\le k<r} \sum_{|w|=N} E_i^T A_k A_w C + \sum_{\substack{|w|=N \\ 8^N r + (w)_8 \le 8^N \cdot 8x}} E_i^T A_r A_w C \\
&= \sum_{0\le k<r} \sum_{j=1}^{8} E_i^T A_k Q^N C + \sum_{\substack{|w|=N \\ (w)_8 \le 8^N (8x-r)}} \sum_{j=1}^{8} A_{r,i,j} E_j^T A_w C \\
&= \sum_{0\le k<r} \sum_{j=1}^{8} E_i^T A_k (4^N V + o(4^N)) + \sum_{j=1}^{8} A_{r,i,j} (4^N F_j(8x-r) + o(4^N)).
\end{aligned}$$

Dividing by 4^N and taking the limit we obtain

$$4F_i(x) = \sum_{0\le k<r} E_i^T A_k V + \sum_{j=1}^{8} A_{r,i,j} F_j(8x-r).$$

Introducing the column vector $\mathbf{F}(x)$ whose components are $F_1(x)$, …, $F_8(x)$, this result writes

$$4\mathbf{F}(x) = \sum_{0\le k<r} A_k V + A_r \mathbf{F}(8x-r).$$

This equation leads us to consider and solve the following problem.

Lemma 6 *The problem*

- $\Phi(x)$ *is a continuous function from the interval* $[0,1]$ *into the space of* 8*-dimensional vectors.*

- $\Phi(0) = 0$, $\Phi(1) = V$.
- *For every octal figure* r *and for* x *in* $[r/8, (r+1)/8)$,

$$\Phi(x) = \frac{1}{4} \sum_{0 \leq k < r} A_k V + \frac{1}{4} A_r \Phi(8x - r),$$

where A_0, …, A_7 *are the square matrices of the linear representation of* adp, 4 *is the dominant eigenvalue of* $Q = A_0 + \cdots + A_7$, *and* V *is an associated eigenvector,*

has a unique solution.

Proof: The space of continuous functions from $[0, 1]$ into the space $\mathbb{R}^8$ is equipped with the norm of the maximum $\|\Phi\|_\infty = \max_x \sum_i |\Phi_i(x)|$. It is known that this space is complete. The subspace of such continuous functions Φ which satisfy $\Phi(0) = 0$ and $\Phi(1) = V$ is closed and therefore complete. The equation of the problem appears as a fixed point equation $\Phi = \mathcal{L}\Phi$ in this space, where the operator $\mathcal{L}$ is defined by $\mathcal{L}\Phi = \Psi$ with

$$\Psi(x) = \frac{1}{4} \sum_{0 \leq k < r} A_k V + \frac{1}{4} A_r \Phi(8x - r)$$

for $0 \leq r < 8$ and $r/8 \leq x < (r + 1)/8$. It is sufficient to see that the subspace under consideration is left invariant by $\mathcal{L}$ and that $\mathcal{L}$ is a contraction in order to prove the lemma.

We have to show that Ψ is continuous as Φ is continuous and satisfies $\Psi(0) = 0$, $\Psi(1) = V$. Thanks to the piecewise definition of Ψ this amounts to consider the left and right behaviour of Ψ at the points $r/8$ for $0 \leq r \leq 8$. The definition of Ψ and the continuity of Φ give immediately

$$\Psi(0) = \frac{1}{4} A_0 \Phi(0) = 0, \qquad \Psi(\frac{r}{8} + 0) = \Psi(\frac{r}{8}) = \frac{1}{4} \sum_{0 \leq k < r} A_r V,$$

$$\Psi(\frac{r}{8} - 0) = \frac{1}{4} \sum_{0 \leq k < r-1} A_r V + \frac{1}{4} A_r \Phi(1) = \frac{1}{4} \sum_{0 \leq k < r} A_r V,$$

$$\Psi(1 - 0) = \frac{1}{4} \sum_{0 \leq k < 7} A_r V + \frac{1}{4} A_7 V = \frac{1}{4} QV = V, \qquad \Psi(1) = V,$$

and the constraints are satisfied.

Let us assume that we have two functions Φ_1 and Φ_2 in the subspace under consideration and let Ψ_1 and Ψ_2 be their images by $\mathcal{L}$. From $\Psi_1(x) - \Psi_2(x) = (1/4)A_r(\Phi_1(8x - r) - \Phi_2(8x - r))$ follows the inequality $\|\Psi_1 - \Psi_2\| \leq (1/4) \|\Phi_1 - \Phi_2\|$ because the maximum absolute column sum norm of each matrix A_r is 1. As a consequence $\mathcal{L}$ is a contraction and this ends the proof. □

Putting together the computation on the F_i's and the previous lemma, we obtain the following representation of the limiting distribution function $F(x)$.

Theorem 3 *The distribution function* $F(x)$ *writes* $F(x) = L\mathbf{F}(x)$ *where* $\mathbf{F}(x)$ *is the unique continuous function from* $[0, 1]$ *into the space* $\mathbb{R}^8$ *which satisfies*

- $\mathbf{F}(0) = 0$, $\mathbf{F}(1) = V$,

- *for* $0 \leq r < 8$ *and* $x \in [r/8, (r+1)/8)$, $\mathbf{F}(x) = \dfrac{1}{4} \sum_{0 \leq k < r} A_k V + \dfrac{1}{4} A_r \mathbf{F}(8x - r)$,

where $A_0, \ldots, A_7$ *are the square matrices of the linear representation of* adp, 4 *is the dominant eigenvalue of* $Q = A_0 + \cdots + A_7$, *and* $V = (E_1 + \cdots + E_8)/8$ *is the associated eigenvector.*

Billingsley (1995, eq. (7.30), p. 104; eq. (31.17), p. 409) gives such examples of functional equation, but for one dimensional problem. The theorem is reminiscent from the iterated function systems defined by Hutchinson (and popularized by Barnsley). Indeed Hutchinson (1981, Section 3.5) introduces the idea of a parametrized curve like $\mathbf{F}(x)$ defined by a system comparable to the one in Lemma 6.

5.2 Moments

It is possible to show that the distribution function F has a characteristic function $\phi(t) = \int e^{itx} dF(x)$ which writes $\phi(t) = L\Phi(t)C$ where $\Phi(t)$ is a square matrix defined as a convergent infinite product, namely $\Phi(t) = \Sigma(t/8)\Sigma(t/8^2) \cdots \Sigma(t/8^N) \cdots$ with $\Sigma(t) = (1/4) \sum_{0 \leq r < 8} e^{itr} A_r$. The proof is the same as for the functional equation of F, using the rightmost letter of each word (Grabner et al., 2005). In principle the knowledge of $\phi(t)$ allows to compute all the moments of the distribution function F.

We will follows a clumsier though more effective way to compute these moments. Let us denote the moments

$$\mu_\ell = \int_0^1 x^\ell \, dF(x)$$

and

$$m_\ell = \int_0^1 x^{\ell-1} F(x) \, dx, \qquad M_\ell = \int_0^1 x^{\ell-1} \mathbf{F}(x) \, dx, \qquad M_{\ell,r} = \int_{r/8}^{(r+1)/8} x^{\ell-1} \mathbf{F}(x) \, dx$$

for $\ell \geq 1$ and $0 \leq r < 8$. The moment μ_ℓ is related to m_ℓ by the formula $\mu_\ell = 1 - \ell m_\ell$, and m_ℓ may be computed from the vector M_ℓ by $m_\ell = LM_\ell$. Hence we are trying to determine the M_ℓ's. Let us begin with M_1. Using the functional equation of $\mathbf{F}(x)$ and the change of variable $y = 8x - r$, we have

$$\begin{aligned} 4M_{1,r} &= \int_{r/8}^{(r+1)/8} \left(\sum_{0 \leq k < r} A_k V + A_r \mathbf{F}(8x - r) \right) dx \\ &= \frac{1}{8} \sum_{0 \leq k < r} A_k V + A_r \int_0^1 \mathbf{F}(y) \frac{dy}{8} = \frac{1}{8} \sum_{0 \leq k < r} A_k V + \frac{1}{8} A_r M_{1,r}. \end{aligned}$$

Adding the equalities for $r = 0, \ldots, 7$, we see that M_1 satisfies

$$(32\, \mathrm{I}_8 - Q) M_1 = \sum_{0 \leq r < 8} \sum_{0 \leq k < r} A_k V = \sum_{0 \leq r < 7} (7 - r) A_r V.$$

The number 32 is not an eigenvalue of Q and the equation has a unique solution

$$M_1 = \frac{1}{120} \begin{pmatrix} 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 \end{pmatrix}^T.$$

We obtain $m_1 = LM_1 = 1/2$ and $\mu_1 = 1/2$.

The same computation provides us with an equation for M_ℓ, namely

$$(8^\ell \times 4\,\mathrm{I}_8 - Q)M_\ell = \frac{1}{\ell} \sum_{0 \le r < 8} \left((r+1)^\ell - r^\ell\right) \sum_{0 \le k < r} A_k V + \sum_{j=1}^{\ell-1} \binom{\ell-1}{j-1} \sum_{0 \le r < 8} r^{\ell-j} A_r M_j.$$

The previous equation has a unique solution because $8^\ell \times 4$ is not an eigenvalue of Q and this equation is a recursion which permits us to successively compute the vectors M_ℓ. Thus we get the exact values of the moments

$$\mu_1 = \frac{1}{2}, \quad \mu_2 = \frac{1}{3}, \quad \mu_3 = \frac{1}{4}, \quad \mu_4 = \frac{1}{5}, \quad \mu_5 = \frac{1}{6},$$
$$\mu_6 = \frac{383289787}{2683033605}, \quad \mu_7 = \frac{127762673}{1022108040}, \quad \mu_8 = \frac{276816430187784766 1255549}{249138337012784135418 32325},$$
$$\mu_9 = \frac{498264047820839145275 8817}{498276674025568270836 64650}, \mu_{10} = \frac{6463661131650900677276012464206762702709565 77}{7110310784656332672513849866851231905086977 875}.$$

It is worth noting the difference $\delta_\ell = \mu_\ell - 1/(\ell+1)$ with the moments of the uniform distribution are the following

$$\delta_1 = 0, \quad \delta_2 = 0, \quad \delta_3 = 0, \quad \delta_4 = 0, \quad \delta_5 = 0,$$
$$\delta_6 = -0.0000002713346559, \quad \delta_7 = -0.0000008140039677, \quad \delta_8 = -0.000001583164964,$$
$$\delta_9 = -0.000002533974675, \quad \delta_{10} = -0.000003625207098.$$

5.3 Numerical computation

A way to see how the function $F(x)$ looks like is to choose a sufficiently large integer N and to compute all the numbers $\mathrm{adp}(w)$ for w of length N. This approach needs to compute the 8^N vectors LA_w and demands about 8^N multiplications matrix by vector. With this method we obtain values of $F(x)$ with an error of the order of 2^{-N}.

Another way is the use of the functional equation satisfied by $\mathbf{F}(x)$. We know the vectors $\mathbf{F}(0)$ and $\mathbf{F}(1)$. We precompute all the vectors $(1/4)A_kV$ and their partial sums at a cost $O(1)$. Assume we have computed all the vectors $\mathbf{F}(k/8^{N-1})$ with $0 \le k \le 8^{N-1}$, then the functional equation gives the values $\mathbf{F}((8k+r)/8^N)$ with $1 \le r < 8$ for a cost of about 8^N multiplications matrix by vector. Hence with a global cost of about 8^N we get all the 8^N values $\mathbf{F}(k/8^N)$, $0 \le k < 8^N$. The difference with the previous method is that the obtained values of $F(x)$ are exact.

Lemma 6 and its proof shed another light on the second method. We start with the parametrized curve xV, whose ends are 0 and V, and we iterate the operator $\mathcal{L}$ associated with the fixed point equation satisfied by $\mathbf{F}(x)$. This produces a sequence of continuous piecewise linear function $L\mathbf{F}_k(x)$ which converges towards the function $L\mathbf{F}(x) = F(x)$.

Note that we are able to compute values not only for octal numbers but for rational numbers too because their expansion is ultimately periodic. The basic case is the case where the expansion of x is purely periodic with a period T and $\mathbf{F}(x)$ is the only solution of $\mathbf{F}(x) = \mathcal{L}^T\mathbf{F}(x)$. In the general case, we have $\mathbf{F}(x) = \mathcal{L}^P\mathbf{G}(x)$ and $\mathbf{G}(x) = \mathcal{L}^T\mathbf{G}(x)$ if T is the period of the expansion and P is the length of the non periodic prefix part.

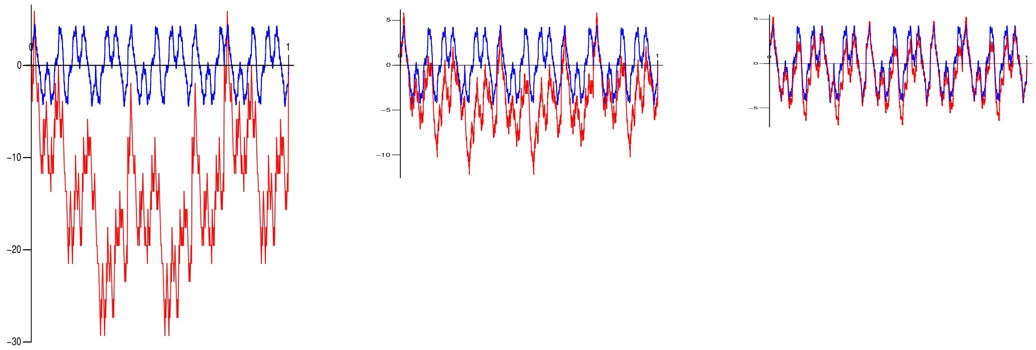

Fig. 2: The difference $F(x) - x$ for the limit function $F(x)$ and its approximations of rank 3, 4, 5 computed from the partial sums of adp. The vertical unit is 10^{-3}.

For the problem under study the graph of $F(x)$ is not interesting because $F(x)$ is very close to the identity. The cause is the origin of the problem: the cipher produces a distribution function which looks like the uniform distribution. Hence it is more interesting to consider the difference $F(x) - x$. In Figure 2, we have drawn the graphic (blue curve) of the function $\Delta(x) = F(x) - x$ by computing the 8^5 values $\Delta(k/8^5)$ for $0 \leq k < 8^5$. The three graphics in red correspond to the partial sums of adp for words of length 3, 4, and 5 from left to right. The vertical unit is 10^{-3}. The speed of convergence proves to be of order 2^{-N} as expected.

6 From rational series to rational sequences

6.1 Rational sequence sbs

Let $\mathrm{sbs}(n)$ (for *side-by-side*) be the sequence obtained by putting side-by-side the values of $\mathrm{adp}(w)$ according to the length and rank in the lexicographic order of the octal words w (that is, according to the radix order). In the correspondence between adp and $\mathrm{sbs}(n)$, the words of length N correspond to the integer interval from $(8^N - 1)/7$ to $(8^{N+1} - 1)/7 - 1$, since $\sum_{n=0}^{N-1} 8^n = (8^N - 1)/7$.

The sequence $\mathrm{sbs}(n)$ does not exist alone, but is part of a family of sequences linked by a linear representation, like in the case of rational series. A sequence s_n is called k-*rational* or rational with respect to the radix $k \geq 1$ (Allouche and Shallit, 2003) if and only if there exists a $1 \times d$ row matrix L, k square matrices A_i, $i = 0, \ldots, k-1$ of size $d \times d$, and a $d \times 1$ column matrix C such that if we write n in base k as $n = (n_\ell \cdots n_0)_k$, the value of s_n is given by

$$s_n = L A_{n_\ell} \cdots A_{n_0} C.$$

By convention, $s_0 = LC$. In this definition, it is assumed that the expansion of the integer n satisfies $n_\ell \neq 0$ because generally speaking we do not have $LA_0 = L$, but this property will be satisfied in the example under consideration. The family L, (A_i), C is called a *linear representation* of dimension d of the sequence. The following lemma shows that $\mathrm{sbs}(n)$ is a rational sequence with respect to the octal numeration.

Lemma 7 *Let S be a rational series over the alphabet $\{0,\dots,k-1\}$. The sequence s_n obtained by putting side-by-side the values of $S(w)$ according to the length and rank in the lexicographic order of words is a k-rational sequence. Moreover, if L, $(A_i)_{i=0}^{k-1}$, C is a linear representation of S with dimension d, the following is a linear representation of s_n with dimension $2d+1$: $L' = \begin{pmatrix}1 & L & 0 & \cdots & 0\end{pmatrix}$, $C' = \begin{pmatrix}0 & C & 0 & \cdots & 0\end{pmatrix}^T$,*

$$A'_0 = \begin{pmatrix}1 & L & 0\\ 0 & 0 & 0\\ 0 & A_{k-1} & A_{k-2}\end{pmatrix},\ A'_1 = \begin{pmatrix}0 & 0 & L\\ 0 & A_0 & 0\\ 0 & 0 & A_{k-1}\end{pmatrix},$$

$$\text{and} \qquad A'_r = \begin{pmatrix}0 & 0 & 0\\ 0 & A_{r-1} & A_{r-2}\\ 0 & 0 & 0\end{pmatrix} \qquad \text{for } 1 < r < k.$$

Proof: The rational character of the rational formal series S means that there exists a finite dimensional vector space $\mathcal{V}$ containing S, which is left invariant by the action of the free monoid of words defined by $r.S(w) = S(wr)$. The linear representation comes from the existence of a finite family of formal series $(f^i)_{1\le i\le d}$ generating $\mathcal{V}$, and the matrices A_r express the action (the superscripts in (f^i) are not exponents but indices). We have a similar property for a rational sequence. To prove the lemma, it suffices to verify that the family $(e^0, u^1, \dots, u^d, v^1, \dots, v^d)$ defined below is a generating family of a vector space $\mathcal{W}$ left invariant by the action $r.u_n = u_{kn+r}$ of the free monoid on the space of sequences. The sequence e^0 is defined by $e^0_0 = 1$ and $e^0_n = 0$ for $n \ne 0$. The sequences u^1, …, u^d are those which correspond to the formal series f^1, …, f^d via the numeration with radix k, and the sequences v^1, …, v^d are variations of the sequences u^i defined by $v^i_0 = 0$ and $v^i_n = u^i_{n-1}$ for $n > 0$. The result follows from a simple algebraic manipulation. □

Applying Lemma 7 to the linear representation of adp gives a linear representation L', $A'_0, \dots, A'_7$, C' of dimension 17 of the 8-rational sequence $\mathrm{sbs}(n)$.

By the previous lemma and Lemma 2, we obtain immediately the following property, which results from the fact that the rational sequence $\mathrm{sbs}(n)$ and the rational series $\mathrm{adp}(w)$ take the same values.

Lemma 8 *All sequences in the vector space generated by $\mathrm{sbs}(n)$ under the action of the monoid of octal words are bounded.*

6.2 Periodic function

For all real $x \in [0,1)$, we thus have

$$\sum_{n < \frac{8^N-1}{7} + 8^N x} \mathrm{sbs}(n) = \sum_{k=0}^{N-1} \sum_{|w|=k} \mathrm{adp}(w) + \sum_{\substack{|w|=N\\ (w)_8 < 8^N x}} \mathrm{adp}(w)$$

$$= \sum_{k=0}^{N-1} 4^k + \sum_{\substack{|w|=N\\ (w)_8 < 8^N x}} \mathrm{adp}(w) \underset{N\to\infty}{=} 4^N \left(\frac{1}{3} + F(x) + O(2^{-N})\right).$$

Let $\nu = (8^N - 1)/7 + 8^N x$. Since

$$\nu^{2/3} = \frac{4^N}{7^{2/3}} \left(1 + 7x - \frac{1}{8^N}\right)^{2/3} \underset{N\to\infty}{=} \frac{4^N}{7^{2/3}} (1+7x)^{2/3} + O(2^{-N}),$$

we obtain

$$\frac{1}{\nu^{2/3}} \sum_{n<\nu} \mathrm{sbs}(n) \underset{N\to\infty}{=} \Gamma(x) + O\left(\frac{1}{\nu^{1/3}}\right) \qquad \text{with} \qquad \Gamma(x) = 7^{2/3} \frac{\frac{1}{3} + F(x)}{(1+7x)^{2/3}}.$$

We find easily $\lim_{x\to 1^-} \Gamma(x) = \lim_{x\to 0^+} \Gamma(x)$ and $\Gamma(x)$ extends to the real line as a continuous 1-periodic function which is Hölder with exponent $2/3$, according to the formula

$$\Gamma(x+\delta) \underset{\delta\to 0}{=} 7^{2/3} \frac{\frac{1}{3} + F(x) + O(\delta^{2/3})}{(1+7x)^{2/3} + O(\delta)} = \Gamma(x) + O(\delta^{2/3})$$

and the constant implied in the big oh is independant of x (because the constant implied in the Hölder property for $F(x)$ is independant of x and $x \mapsto (1+7x)^{2/3}$ is Lipschitz on $[0,1]$).

Let $\{\lambda\} = \lambda - \lfloor \lambda \rfloor$ denote the fractional part of λ. Since $0 \le x < 1$, we have $1 \le 1 + 7x < 8$ and $\{\log_8(7\nu + 1)\} = \log_8(1 + 7x)$. Besides, let x' be such that $\{\log_8(7\nu)\} = \log_8(1 + 7x')$. The number x' is well given because the map $z \mapsto \log_8(1 + 7z)$ defines a bijection from $[0, 1)$ onto itself. For $8^{-N}/7 \le x < 1$ we have $\lfloor \log_8(7\nu) \rfloor = N$ and $x - x' = 8^{-N}/7$. For $0 \le x < 8^{-N}/7$ we have $\lfloor \log_8(7\nu) \rfloor = N - 1$ and $8^{-N}/7 < 1 + x - x' \le 8 \cdot 8^{-N}/7$. Thanks to the periodic character of $\Gamma(x)$ we obtain, in every case, $\Gamma(x') = \Gamma(x + \delta)$ with $|\delta| \le 8 \cdot 8^{-N}/7$, hence $\delta = O(1/\nu)$ uniformly with respect to x. Because $\Gamma(x)$ is Hölder with exponent $2/3$, this gives $\Gamma(x') = \Gamma(x) + O(1/\nu^{2/3})$ and the asymptotic formula for the running sum rewrites

$$\sum_{n<\nu} \mathrm{sbs}(n) \underset{\nu\to\infty}{=} \nu^{2/3}\Gamma(x') + O(\nu^{1/3}).$$

We conclude by expressing x' as a function of $\log_8 \nu$ that

$$\sum_{n<\nu} \mathrm{sbs}(n) \underset{\nu\to\infty}{=} \nu^{2/3} G_{2/3}(\log_8 \nu) + O(\nu^{1/3}),$$

where $G_{2/3}(\lambda)$ is a 1-periodic function, namely

$$G_{2/3}(\lambda) = 7^{2/3} \frac{\frac{1}{3} + F\left(\frac{8^{\{\lambda + \log_8 7\}} - 1}{7}\right)}{4^{\{\lambda + \log_8 7\}}}.$$

As a remark, this definition provides us with the formula $G_{2/3}(-\log_8 7) = 7^{2/3}/3$, since $F(0) = 0$, and the bounds $7^{2/3}/12 \le G_{2/3}(\lambda) \le 4 \cdot 7^{2/3}/3$ for all λ, since F takes values in $[0, 1]$.

We note that the big oh is uniform with respect to x or ν in all the previous formulæ. This point arises from the uniform convergence in Theorem1 and from the previous computations. This justifies the uniform character of the convergence in the following theorem.

Theorem 4 *There exists a strictly positive* 1*-periodic function* $G_{2/3}$ *such that*

$$\sum_{n<\nu} \mathrm{sbs}(n) \underset{\nu\to\infty}{=} \nu^{2/3} G_{2/3}(\log_8 \nu) + O(\nu^{1/3})$$

uniformly with respect to ν.

The occurrence of the exponent $2/3$ may be obtained by elementary means. Indeed, equation (3), and the nonnegativity of $\mathrm{adp}(w)$ give

$$\frac{4^N-1}{3} \le \sum_{n<\nu} \mathrm{sbs}(n) < \frac{4^{N+1}-1}{3}$$

for $(8^N-1)/7 \le \nu < (8^{N+1}-1)/7$, and the relation $\sum_{n<\nu} \mathrm{sbs}(n) \asymp \nu^{2/3}$ follows.

The properties of $F(x)$, the periodicity of $G_{2/3}(\lambda)$, and the equality

$$\lim_{\lambda'\to\lambda^+} G_{2/3}(\lambda') = \lim_{\lambda'\to\lambda^-} G_{2/3}(\lambda')$$

for $\lambda = -\log_8 7$ lead to the following result about the regularity of $G_{2/3}(\lambda)$.

Corollary 1 *The function* $G_{2/3}(\lambda)$ *is Hölder of order* $2/3$ *and* $2/3$ *is locally the best Hölder exponent on an everywhere dense subset of the real line. Moreover it is differentiable almost everywhere and non-differentiable on an everywhere dense subset.*

The problem of non-differentiability was addressed in (Tenenbaum, 1997), but the argument relies on the assumption that the sequence under consideration u_n satisfies $\inf_n |u_n| > 0$. This argument does not apply here because $\mathrm{sbs}(n)$ has the value 0 infinitely often. Moreover, the argument gives only the non-differentiability at every point. It seems that our example is the first where the function is differentiable on an everywhere dense subset, and also not differentiable on an everywhere dense subset.

6.3 Bounded variation and convergence of the Fourier series

Because $G_{2/3}(\lambda)$ is a continuous function, we may consider its Fourier series, but at this point we do not know if this Fourier series is convergent. Let λ be a real number and ν_N the integral part of $8^{N+\lambda}$. The formula

$$G_{2/3}(\lambda) \underset{N\to+\infty}{=} \frac{1}{4^\lambda} 4^{-N} \sum_{0\le n<\nu_N} \mathrm{sbs}(n) + o(1)$$

provides the decomposition $\log_8 G_{2/3}(\lambda) = -2\lambda/3 + H(\lambda)$ with

$$H(\lambda) = \lim_{N\to+\infty} \log_8 4^{-N} \sum_{0\le n<\nu_N} \mathrm{sbs}(n).$$

For a given N, the last sum is a nondecreasing function of λ, hence $H(\lambda)$ is a nondecreasing function of λ. Since the function $\log_8 G_{2/3}(\lambda)$ is the difference of two nondecreasing bounded functions, it is of bounded variation in every finite interval. It follows that $G_{2/3}(\lambda)$ is of bounded variation.

There is another way to obtain this conclusion. The function $G_{2/3}(\lambda)$ appears to be a composition of functions which are piecewise indefinitely differentiable, at the exception of $F(x)$ which is Hölder of order $2/3$. By repeated applications of the mean value theorem, we see that $G_{2/3}(\lambda)$ is Hölder of order $2/3$ too. Hence it is of bounded variation.

Lemma 9 *The function $G_{2/3}(\lambda)$ is Hölder of order $2/3$ and of bounded variation.*

This permits us to apply the Jordan-Dirichlet theorem, taking into account the continuity of $G_{2/3}(\lambda)$.

Theorem 5 *The Fourier series of $G_{2/3}(\lambda)$ converges towards $G_{2/3}(\lambda)$, and the convergence is uniform.*

According to (Zygmund, 1977, Thm. 10.9, p. 64) (and more specifically Remark (a), p. 120), the Fourier series of $G_{2/3}(\lambda)$ converges with a kth remainder of the order of $k^{-2/3}$. Hence the convergence is uniform but very slow. This is to be compared with the convergence of the sequence (F_N) with an error of order 2^{-N}.

7 Dirichlet series

The precise study of the Fourier expansion of $G_{2/3}(\lambda)$ relies on the use of Dirichlet series associated to $\mathrm{sbs}(n)$. Recall that the Dirichlet series $f(s)$ of a sequence (a_n) is defined by $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ for $s \in \mathbb{C}$. Following tradition, we will write $s = \sigma + it$ where $\sigma, t \in \mathbb{R}$. Each Dirichlet series has an abscissa of convergence σ_c such that $f(s)$ converges for $\sigma > \sigma_c$ and diverges for $\sigma < \sigma_c$. If $f(s)$ diverges or converges for all s, $\sigma_c = \pm\infty$. The following result (Hardy and Riesz, 1964) can be used for computing the abscissa of convergence.

Lemma 10 *If the abscissa of convergence of the Dirichlet series for a sequence (a_n) is positive, it is given by*

$$\sigma_c = \limsup_{n\to\infty} \frac{\log|\sum_{k=1}^{n} a_k|}{\log n}.$$

Combined with Theorem 4, this shows that the abscissa of convergence σ_c of the Dirichlet series for $\mathrm{sbs}(n)$ is $\sigma_c = 2/3$.

We consider the seventeen rational sequences $u_n^1, u_n^2, \dots, u_n^{17}$ that arise from the linear representation L', $A'_0, \dots, A'_7$ and each vector from the canonical basis of $\mathbb{Q}^{17}$. Note that $u_n^2 = \mathrm{sbs}(n)$. Let $u^i(s)$ be the Dirichlet series of u_n^i, and let $U(s)$ be the row vector of the series $u^i(s)$.

According to Lemma 8, all sequences $u_n^1, u_n^2, \dots, u_n^{17}$ are bounded. As a consequence each Dirichlet series $u^i(s)$ has an abscissa of convergence ≤ 1. Thus, $U(s)$ has abscissa of convergence $\sigma_c \leq 1$.

7.1 Meromorphicity

Our next goal is to show that the analytic function $U(s)$, defined in the half-plane $\sigma > \sigma_c$, admits an meromorphic extension. Let $U_n = (u_n^1, \dots, u_n^{17})$. Following (Allouche and Cohen, 1985), we write

$$\begin{aligned} U(s) &= \sum_{n=1}^{\infty} \frac{U_n}{n^s} = \sum_{n=1}^{\infty} \frac{U_{8n}}{(8n)^s} + \sum_{r=1}^{7} \left(\frac{U_r}{r^s} + \sum_{n=1}^{\infty} \frac{U_{8n+r}}{(8n+r)^s} \right) \\ &= \sum_{n=1}^{\infty} \frac{U_n A'_0}{(8n)^s} + \sum_{r=1}^{7} \left(\frac{U_r}{r^s} + \sum_{n=1}^{\infty} \frac{U_n A'_r}{(8n+r)^s} \right) = \sum_{r=1}^{7} \frac{U_r}{r^s} + \sum_{r=0}^{7} \sum_{n=1}^{\infty} \frac{U_n A'_r}{(8n+r)^s}. \end{aligned}$$

Denote $Q' = \sum_{r=0}^{7} A'_r$ and

$$\nabla U(s) = \sum_{r=1}^{7} \frac{U_r}{r^s} + \sum_{r=1}^{7} \sum_{n=1}^{\infty} U_n A'_r \left(\frac{1}{(8n+r)^s} - \frac{1}{(8n)^s} \right). \tag{5}$$

Thus we have $U(s)(\mathrm{I}_{17} - 8^{-s}Q') = \nabla U(s)$. Let Δ_h be the difference operator $\Delta_h u(n) = u(n+h) - u(n)$. Then we can write $\nabla U(s)$ as

$$\nabla U(s) = \sum_{r=1}^{7} \frac{U_r}{r^s} + \frac{1}{8^s} \sum_{r=1}^{7} \sum_{n=1}^{\infty} U_n A'_r \Delta_{r/8} \frac{1}{n^s}. \tag{6}$$

Since $\Delta_{r/8} n^{-s} = (n + r/8)^{-s} - n^{-s} = -s \int_n^{n+r/8} u^{-(s+1)}\, du$,

$$\left| \Delta_{r/8} \frac{1}{n^s} \right| \le |s| \int_n^{n+r/8} \frac{du}{u^{\sigma+1}} = \frac{|s|}{\sigma} \left| \Delta_{r/8} \frac{1}{n^\sigma} \right|. \tag{7}$$

For fixed s, we have $\Delta_{r/8}\, n^{-\sigma} \sim_{n\to\infty} \sigma n^{-(\sigma+1)}$. Since the components of U_n are bounded sequences, it follows that

$$\frac{1}{8^s} U_n A'_r \Delta_{r/8} \frac{1}{n^s} \underset{n\to\infty}{=} O\left(\frac{1}{n^{\sigma+1}}\right).$$

We conclude that the series (5) converges absolutely in the half-plane $\sigma > 0$. Thus, the function $\nabla U(s)$ is analytic in this half-plane.

7.2 Poles

The poles of $U(s)$ come from the term $I_{17} - 8^{-s}Q'$. The eigenvalues of Q' are 4, 2, 1, 1/4 and 0, with multiplicities 1, 3, 6, 3 and 4, respectively. Let $J = P^{-1}Q'P$ be the Jordan form of Q', where J is the quasi-diagonal matrix

$$J = \operatorname{diag}\begin{pmatrix} 4 & 2 & 2 & 2 & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & 1 & 1 & 1 & 1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We make a change of coordinates to get a new sequence V_n with $U_n = V_n P^{-1}$ and $U(s) = V(s)P^{-1}$. More precisely, if L' $A'_0, \ldots, A'_7$, C' is the linear representation of U_n, we get the linear representation $L'' = L'P$, $C'' = P^{-1}C'$, $A''_0 = P^{-1}A'_0 P$, ..., $A''_7 = P^{-1}A'_7 P$. Applying this change of coordinates to (5) gives $V(s)(I_{17} - 8^{-s}J) = \nabla V(s)$, where

$$\nabla V(s) = \sum_{r=1}^{7} \frac{V_r}{r^s} + \sum_{r=1}^{7} \sum_{n=1}^{\infty} V_n A''_r \left(\frac{1}{(8n+r)^s} - \frac{1}{(8n)^s} \right).$$

The function $\nabla V(s)$ is analytic for $\sigma > 0$.

The equation $V(s)(I_{17} - 8^{-s}J) = \nabla V(s)$ gives a system of equations of the form $v^j(s)(1 - J_{jj} \cdot 8^{-s}) = \nabla v^j(s)$, $j = 1, \ldots, 17$, at the exception of $j = 6$, where the equation is $(v^5(s) + v^6(s))(1 - 8^{-s}) = \nabla v^6(s)$.

From these equations we see, with $\chi = 2\pi i/\ln 8$, that

- The function $v^1(s)$ is meromorphic in the half plane $\sigma > 0$, with possible poles as $2/3 + k\chi$, $k \in \mathbb{Z}$.
- The functions $v^2(s), \ldots, v^4(s)$ are meromorphic in the half plane $\sigma > 0$, with possible poles as $1/3 + k\chi$, $k \in \mathbb{Z}$.

- The functions $v^5, \dots, v^{17}(s)$ are analytic in the half plane $\sigma > 0$.

Recall that the Dirichlet series $u^2(s)$ of $\mathrm{sbs}(n)$ has abscissa of convergence $2/3$, and $U(s) = V(s)P^{-1}$. Hence, $u^2(s)$ extends to a meromorphic function in $\sigma > 0$. Since $\mathrm{sbs}(n)$ is nonnegative, $2/3$ is a singularity of $u^2(s)$. If $2/3$ would not be a pole of $v^1(s)$, the argument above would show that $u^2(s)$ is analytic in $\sigma > 1/3$—a contradiction. Thus, $2/3$ indeed is a pole for $v^1(s)$. For the other Dirichlet series $u^j(s)$, we do not know exactly their abscissa of convergence, but since the Dirichlet series have nonnegative coefficients, and the rightmost possible singularity is at $2/3$, we know that all the Dirichlet series have abscissa of convergence not greater that $2/3$.

7.3 Order of growth

Recall that the order of growth $\mu_g(\sigma)$ of a function $g(s)$ along the vertical line of abscissa σ is

$$\mu_g(\sigma) = \inf\left\{\lambda \mid g(\sigma + it) \underset{|t|\to\infty}{=} O(|t|^\lambda)\right\}.$$

Since the Dirichlet series defining $u^j(s)$ have nonnegative coefficients and abscissæof convergence $\leq 2/3$, their order of growth is $\mu_{u^j}(\sigma) = 0$ for $\sigma > 2/3$.

From (6) and (7), we obtain the inequality $\|\nabla U(s)\| \leq C_0 + C_1|s|8^{-\sigma}\zeta(\sigma + 1)$ for some constants C_0 and C_1. We see that $\mu_{\nabla u^j}(\sigma) \leq 1$ for $0 < \sigma < 1$. Moreover it is evident that the series of (5) is bounded for every $\sigma > 1$, because the sequence U_n is bounded. According to Lindelöf's theorem (Hardy and Riesz, 1964, Theorem 14), $\mu(\sigma)$ is a convex function, and thus

$$\mu_{\nabla u^j}(\sigma) \leq 1 - \sigma \quad \text{for } 0 < \sigma < 1 \text{ and } \quad \mu_{\nabla u^j}(\sigma) = 0 \quad \text{for } \sigma > 1.$$

Since the function $s \mapsto \mathrm{I}_{17} - 8^{-s}Q'$ is periodic with respect to t, this is valid also for the functions u^j. Since the functions $v^j(s)$ are linear combinations of the $u^j(s)$, the same result holds for all the $v^j(s)$.

8 Fourier series

8.1 Mellin-Perron formula

We now apply the following Mellin-Perron formula (Tenenbaum, 1995). Let (c) denote the vertical line $\sigma = c$.

Lemma 11 (Mellin-Perron formula) *Let $f(s) = \sum_{k=1}^{\infty} f_k k^{-s}$ be the Dirichlet series of the sequence (f_k). Let the line (c), with $c > 0$, lie inside the half-plane of absolute convergence of $f(s)$. With these hypotheses, the sums of first-order and second-order of the sequence f_n are given by the formulæ*

$$\sum_{1\leq k<\nu} f_k + \frac{1}{2}f_\nu = \frac{1}{2\pi i}\int_{(c)} f(s)\nu^s \frac{ds}{s}, \tag{8.1}$$

$$\frac{1}{\nu}\sum_{1\leq k\leq n<\nu} f_k = \frac{1}{\nu}\sum_{1\leq k<n\leq\nu} f_k = \sum_{1\leq k<\nu} f_k\left(1 - \frac{k}{\nu}\right) = \frac{1}{2\pi i}\int_{(c)} f(s)\nu^s \frac{ds}{s(s+1)}. \tag{8.2}$$

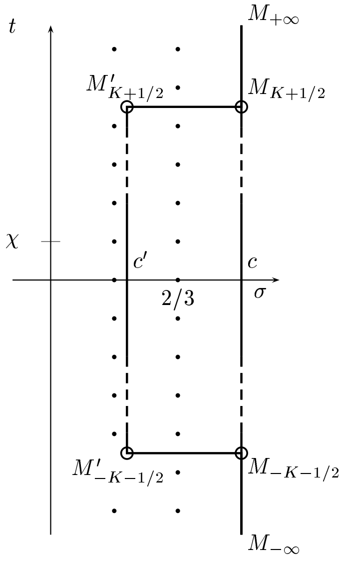

Fig. 3: Poles of $U(s)$.

In the first case, the integral is a principal value integral.

The first idea which comes in mind is to apply the first formula to the Dirichlet series $V(s)$. Precisely, we want to use a vertical line (c) with $c > 2/3$, and for a positive integer K we want to replace (Figure 3) the vertical segment $[M_{-K-1/2}, M_{K+1/2}]$ by the polygonal line $M_{-K-1/2}$, $M'_{-K-1/2}$, $M'_{K+1/2}$, $M_{K+1/2}$. The horizontal segment $[M'_{K+1/2}, M_{K+1/2}]$ goes through the middle of the poles $2/3+K\chi$ and $2/3+(K+1)\chi$ (with $\chi = 2\pi i/\ln 8$), and the abscissa c' of $M'_{K+1/2}$ is between $1/3$ and $2/3$. The figure is symmetric with respect to the σ-axis. Collecting the residues of the poles which are inside the rectangle, we obtain a trigonometric polynomial multiplied by $\nu^{2/3}$, and we are not far from $\nu^{2/3}G_{2/3}(\log_8 \nu)$.

However we cannot let K go to infinity, because we know only that $U(s)/s$ is $O(|t|^{-c'})$ if c' is the abscissa of the new vertical line, and c' is between $1/3$ and $2/3$. This does not ensure the absolute convergence of the integral on the line (c'). Hence we are led to use the second formula.

We apply the formula for the second-order sums to the row vector $V(s)$, and push the line of integration to the left, taking the residues of the function into account. Because of the factor $s(s+1)$, we may consider both lines of poles at abscissæ $2/3$ and $1/3$. Hence we introduce a new vertical line (ϵ) with $0 < \epsilon < 1/3$. For $v^1(s)$, we get

$$\frac{1}{\nu}\sum_{1\le n\le\nu}\sum_{k=1}^{n-1} v_k^1 = \sum_{k\in\mathbb{Z}} \operatorname*{Res}_{s=\frac{2}{3}+k\chi} \frac{\nabla v^1(s)\nu^s}{(1-4\cdot 8^{-s})s(s+1)} + \frac{1}{2\pi i}\int_{(\epsilon)} \frac{\nabla v^1(s)\nu^s}{1-4\cdot 8^{-s}}\frac{ds}{s(s+1)}.$$

For $j = 2, 3, 4$, we get

$$\frac{1}{\nu}\sum_{1\le n\le\nu}\sum_{k=1}^{n-1} v_k^j = \sum_{k\in\mathbb{Z}} \operatorname*{Res}_{s=\frac{1}{3}+k\chi} \frac{\nabla v^j(s)\nu^s}{(1-2\cdot 8^{-s})s(s+1)} + \frac{1}{2\pi i}\int_{(\epsilon)} \frac{\nabla v^j(s)\nu^s}{1-2\cdot 8^{-s}}\frac{ds}{s(s+1)}.$$

For $j = 5, \dots, 17$, at the exception of $j = 6$, we get

$$\frac{1}{\nu} \sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} v_k^j = \frac{1}{2\pi i} \int_{(\epsilon)} \frac{\nabla v^j(s) \nu^s}{1 - J_{jj} \cdot 8^{-s}} \frac{ds}{s(s+1)}.$$

For $j = 6$, we get

$$\frac{1}{\nu} \sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} v_k^6 = \frac{1}{2\pi i} \int_{(\epsilon)} \frac{(\nabla v^6(s) - \nabla v^5(s)) \nu^s}{1 - 8^{-s}} \frac{ds}{s(s+1)}.$$

All the integrals above can be bounded as

$$\left| \frac{1}{2\pi i} \int_{(\epsilon)} \frac{\nabla v^j(s) \nu^s}{1 - J_{jj} \cdot 8^{-s}} \frac{ds}{s(s+1)} \right| \underset{\nu \to \infty}{=} O(\nu^\epsilon).$$

By computing the residues, we obtain

$$\frac{1}{\nu} \sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} v_k^1 \underset{\nu \to \infty}{=} \frac{\nu^{2/3}}{\ln 8} \sum_{k \in \mathbb{Z}} \frac{\nabla v^1(2/3 + k\chi)}{(2/3 + k\chi)(5/3 + k\chi)} \exp(2\pi i k \log_8 \nu) + O(\nu^\epsilon).$$

For $j = 2, 3, 4$, we obtain

$$\frac{1}{\nu} \sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} v_k^j \underset{\nu \to \infty}{=} \frac{\nu^{1/3}}{\ln 8} \sum_{k \in \mathbb{Z}} \frac{\nabla v^j(1/3 + k\chi)}{(1/3 + k\chi)(4/3 + k\chi)} \exp(2\pi i k \log_8 \nu) + O(\nu^\epsilon).$$

Finally, for $j = 5, \dots, 17$, we have $\frac{1}{\nu} \sum_{1 \le n < \nu} \sum_{k=1}^{n-1} v_k^j =_{\nu \to \infty} O(\nu^\epsilon)$.

Note that $\nabla v^1(2/3 + it) =_{|t| \to \infty} O(|t|^{1/3})$ and $\nabla v^j(1/3 + it) =_{|t| \to \infty} O(|t|^{2/3})$ for $j \ge 2$. Thus, the series above converge absolutely. It follows that the trigonometric series define 1-periodic continuous functions. Since the sequence $\mathrm{sbs}(n)$ is a linear combination of the sequences $v^1, \dots, v^{17}$, we obtain the following result.

Theorem 6 *For all* $0 < \epsilon < 1/3$*,*

$$\sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} \mathrm{sbs}(k) \underset{\nu \to \infty}{=} \nu^{5/3} H_{5/3}(\log_8 \nu) + \nu^{4/3} H_{4/3}(\log_8 \nu) + O(\nu^{1+\epsilon}),$$

where $H_{5/3}(\lambda)$ *and* $H_{4/3}(\lambda)$ *are* 1*-periodic continuous functions.*

8.2 From double to simple sums

By Theorem 4

$$\sum_{1 \le n < \nu} \mathrm{sbs}(n) \underset{\nu \to \infty}{=} \nu^{2/3} G_{2/3}(\log_8 \nu) + O(\nu^{1/3}),$$

where $G_{2/3}(\lambda)$ is a 1-periodic continuous function. We will use the following pseudo-Tauberian result (Flajolet et al., 1994, Proposition 6.4) to derive a Fourier series expansion for $G_{2/3}(\lambda)$.

Lemma 12 *Let $f(\lambda)$ be a 1-periodic continuous function and let τ be a complex number with positive real part. Then there exists a 1-periodic continuously differentiable function $g(\lambda) = g(f, \tau\,;\,\lambda)$, depending on $f(\lambda)$, τ, and λ, such that*

$$\frac{1}{\nu^{\tau+1}} \sum_{1 \le n < \nu} n^\tau f(\log_8 n) \underset{\nu \to \infty}{=} g(f, \tau\,;\, \log_8 \nu) + o(1).$$

Moreover, the function $g(\lambda) = g(f, \tau\,;\,\lambda)$ satisfies

$$\int_0^1 g(\lambda)\, d\lambda = \frac{1}{\tau + 1} \int_0^1 f(\lambda)\, d\lambda$$

and

$$g\left(f(\lambda) e^{-2\pi i \lambda}, \tau + \frac{2\pi i}{\ln 8}\,;\, \lambda \right) = g(f(\lambda), \tau\,;\, \lambda)\, e^{-2\pi i \lambda}.$$

Lemma 12 (with $\tau = 2/3$) implies that there exists a 1-periodic and continuously differentiable function $G_{5/3}(\lambda)$ such that

$$\sum_{1 \le n \le \nu} \sum_{k=1}^{n-1} \mathrm{sbs}(k) \underset{\nu \to \infty}{=} \nu^{5/3} G_{5/3}(\log_8 \nu) + o(\nu^{5/3}).$$

The uniqueness of asymptotic expansion with variable coefficients (Bourbaki, 1976, Chapter V) shows that $G_{5/3}(\lambda) = H_{5/3}(\lambda)$. Let $c_k(F)$ denote the Fourier coefficients of the periodic function $F(\lambda)$. By Lemma 12, we get with $\chi = 2\pi i / \ln 8$

$$\begin{aligned} c_k(H_{5/3}) &= \int_0^1 H_{5/3}(\lambda) e^{-2\pi i k \lambda}\, d\lambda = \int_0^1 g(G_{2/3}(\lambda) e^{-2\pi i k \lambda}, 2/3 + \chi k\,;\, \lambda)\, d\lambda \\ &= \frac{1}{2/3 + \chi k + 1} \int_0^1 G_{2/3}(\lambda) e^{-2\pi i k \lambda}\, d\lambda = \frac{1}{5/3 + \chi k} c_k(G_{2/3}). \end{aligned}$$

This shows that the Fourier coefficients of $G_{2/3}$ are given by

$$c_k(G_{2/3}) = \left(\frac{5}{3} + k\chi \right) c_k(H_{5/3}) = \frac{1}{\ln 8} \frac{\nabla v^1(2/3 + k\chi)}{2/3 + k\chi}.$$

Theorem 7 *The function $G_{2/3}(\lambda)$ admits the Fourier series*

$$G_{2/3}(\lambda) = \frac{1}{\ln 8} \sum_{k \in \mathbb{Z}} \frac{\nabla v^1(2/3 + k\chi)}{2/3 + k\chi} e^{2\pi i k \lambda},$$

where $\chi = 2\pi i / \ln 8$.

8.3 More natural approach

Now we know that the residues of $v^1(s)$ on the rightmost line of poles give the Fourier coefficients of $G_{2/3}(\lambda)$. This permits us to take a slightly different approach to our computation. Let γ_K be the polygonal line $M_{-\infty}$, $M_{-K-1/2}$, $M'_{-K-1/2}$, $M'_{K+1/2}$, $M_{K+1/2}$, $M_{+\infty}$ of Figure 3. Cauchy's theorem gives, with $\lambda = \log_8 \nu$,

$$\frac{1}{2\pi i}\int_{(c)} v^1(s)\nu^s \frac{ds}{s} - \frac{1}{2\pi i}\int_{\gamma_K} v^1(s)\nu^s \frac{ds}{s} = \nu^{2/3} \sum_{k=-K}^{K} c_k(G_{2/3}) e^{2\pi i k\lambda}.$$

The integrals over the horizontal segments go to zero when K tends towards infinity, because the function inside them is $O(1/K^{c'})$. The partial sum of the Fourier series converges towards $G_{2/3}(\lambda)$. As a consequence, the integral over the vertical segment $[M'_{-K-1/2}, M'_{K+1/2}]$ is convergent. In the limit, we obtain

$$\frac{1}{2\pi i}\int_{(c)} v^1(s)\nu^s \frac{ds}{s} - \frac{1}{2\pi i}\int_{(c')} v^1(s)\nu^s \frac{ds}{s} = \nu^{2/3} G_{2/3}(\lambda).$$

In other words, the idea we exposed at the beginning of this section gives the correct result. The use of the second-order sum was only a technical mean in order to justify the idea.

References

Jean-Paul Allouche and Henri Cohen. Dirichlet series and curious infinite products. *Bull. London Math. Soc.*, 17(6):531–538, 1985. ISSN 0024-6093.

Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences*. Cambridge University Press, Cambridge, 2003. ISBN 0-521-82332-3. Theory, applications, generalizations.

Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. *J. Cryptology*, 4(1): 3–72, 1991. ISSN 0933-2790.

Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. ISBN 0-471-00710-2.

Nicolas Bourbaki. *Éléments de mathématique: Fonctions d'une variable réelle; Théorie élémentaire*. Hermann, 1976. Translated into *Functions of a real variable: Elementary theory. Elements of Mathematics.* Springer-Verlag, Berlin, 2004. xiv+338 pp. ISBN: 3-540-65340-6.

Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to algorithms*. The MIT Electrical Engineering and Computer Science Series. MIT Press, Cambridge, MA, 1990. ISBN 0-262-03141-8.

Jean-Marie Dumont and Alain Thomas. Systèmes de numération et fonctions fractales relatifs aux substitutions. *Theoret. Comput. Sci.*, 65(2):153–169, 1989. ISSN 0304-3975.

Philippe Flajolet. Approximate counting: a detailed analysis. *BIT*, 25(1):113–134, 1985. ISSN 0006-3835.

Philippe Flajolet and Mordecai Golin. Mellin transforms and asymptotics. The mergesort recurrence. *Acta Inform.*, 31(7):673–696, 1994. ISSN 0001-5903.

Philippe Flajolet and G. Nigel Martin. Probabilistic counting algorithms for data base applications. *J. Comput. System Sci.*, 31(2):182–209, 1985. ISSN 0022-0000. Special issue: Twenty-fourth annual symposium on the foundations of computer science (Tucson, Ariz., 1983).

Philippe Flajolet, Peter Grabner, Peter Kirschenhofer, Helmut Prodinger, and Robert F. Tichy. Mellin transforms and asymptotics: digital sums. *Theoret. Comput. Sci.*, 123(2):291–314, 1994. ISSN 0304-3975.

Peter J. Grabner, Clemens Heuberger, and Helmut Prodinger. Counting optimal joint digit expansions. *Integers*, 5(3):A9, 19 pp. (electronic), 2005. ISSN 1553-1732.

G. H. Hardy and M. Riesz. *The general theory of Dirichlet's series*. Stechert-Hafner, Inc., New York, 1964. Cambridge Tracts in Mathematics and Mathematical Physics, No. 18. Also available from `http://library5.library.cornell.edu/math.html`.

John E. Hutchinson. Fractals and Self Similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

Helger Lipmaa, Johan Wallén, and Philippe Dumas. On the additive differential probability of exclusive-or. In Bimal Roy and Willi Meier, editors, *Fast Software Encryption 2004*, volume 3017 of *Lecture Notes in Computer Science*, pages 317–331. Springer-Verlag, 2004.

Jacques Sakarovitch. *Éléments de théorie des automates*. Vuibert, 2003. ISBN 2-7117-4807-3. Translated into *Elements of Automata Theory*, Cambridge University Press, 2005, ISBN 0-521-84425-8.

Gérald Tenenbaum. *Introduction à la théorie analytique et probabiliste des nombres*, volume 1 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, second edition, 1995. ISBN 2-85629-032-9. Translated into *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995, ISBN 0-521-41261-7.

Gérald Tenenbaum. Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires. In *The mathematics of Paul Erdős, I*, volume 13 of *Algorithms Combin.*, pages 117–128. Springer, Berlin, 1997.

A. Zygmund. *Trigonometric series. Vol. I, II*. Cambridge University Press, Cambridge, 1977. ISBN 0-521-07477-0. Reprinting of the 1968 version of the second edition with Volumes I and II bound together.