Analysis of Regular Sequences: Summatory Functions and Divide-and-Conquer Recurrences

Clemens Heuberger[1]
Department of Mathematics, University of Klagenfurt, Austria

Daniel Krenn
Fachbereich Mathematik, Paris Lodron University of Salzburg, Austria

Tobias Lechner
Department of Mathematics, University of Klagenfurt, Austria

Abstract

In the asymptotic analysis of regular sequences as defined by Allouche and Shallit, it is usually advisable to study their summatory function because the original sequence has a too fluctuating behaviour. It might be that the process of taking the summatory function has to be repeated if the sequence is fluctuating too much. In this paper we show that for all regular sequences except for some degenerate cases, repeating this process finitely many times leads to a "nice" asymptotic expansion containing periodic fluctuations whose Fourier coefficients can be computed using the results on the asymptotics of the summatory function of regular sequences by the first two authors of this paper.

In a recent paper, Hwang, Janson, and Tsai perform a thorough investigation of divide-and-conquer recurrences. These can be seen as 2-regular sequences. By considering them as the summatory function of their forward difference, the results on the asymptotics of the summatory function of regular sequences become applicable. We thoroughly investigate the case of a polynomial toll function.

2012 ACM Subject Classification Mathematics of computing → Discrete mathematics

Keywords and phrases Regular sequence, Divide-and-Conquer Recurrence, Summatory Function, Asymptotic Analysis

Supplementary Material *Software (Code for Example 9)*: `https://arxiv.org/src/2403.06589/anc`

Funding *Clemens Heuberger*: This research was funded in part by the Austrian Science Fund (FWF) [10.55776/DOC78]. For open access purposes, the author has applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

1 Introduction

1.1 Overview

The aim of [7] is the study of the asymptotic behaviour of the summatory functions of regular sequences [1]—in simplest terms, a sequence x is called q-regular for some integer $q \geq 2$ if there are square matrices $A_0, \ldots, A_{q-1}$, a row vector u and a column vector w such that for all integers $n \geq 0$,

$$x(n) = uA_{n_0} \ldots A_{n_{\ell-1}} w \tag{1}$$

where $(n_{\ell-1}, \ldots, n_0)$ is the $q$ary expansion of n; an alternative definition will be given in Definition 1. Regular sequences have been introduced by Allouche and Shallit [1]; a plethora of examples have also been given in the same publication. We highlight two prototypical

[1] Corresponding author

examples at this point: the binary sum of digits function and the worst case number of comparisons in merge sort.

The main result of [7] is that the summatory function $N \mapsto \sum_{0 \le n < N} x(n)$ of a q-regular sequence x has an asymptotic expansion

$$\sum_{0 \le n < N} x(n) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > R}} N^{\log_q \lambda} \sum_{0 \le k < m_C(\lambda)} \frac{(\log N)^k}{k!} \, \Phi_{\lambda k}(\log_q N) + O\bigl(N^{\log_q R} (\log N)^{\kappa}\bigr) \tag{2}$$

as $N \to \infty$, where the $\Phi_{\lambda k}$ are suitable 1-periodic continuous functions and $\sigma(C)$, m_C, R, κ are a set, a function, and two quantities, respectively, depending on the regular sequence and which will be explained in Theorem 3 below. An algorithm is given to compute the Fourier coefficients of the periodic functions. The main question is whether there are $\lambda \in \sigma(C)$ with $|\lambda| > R$. In this case, we say that we established a *good* asymptotic expansion for the summatory function of the regular sequence. Otherwise, (2) reduces to an error term. Note that discussing the question of whether the periodic fluctuations vanish is beyond the scope of this paper.

Studying the summatory function was motivated by the fact that in several well-known examples of regular sequences, the sequences themselves are fluctuating too much so that it is impossible to establish a good asymptotic expansion for the regular sequence itself. For instance, for the binary sum of digits function s_2, we have $s_2(2^k - 1) = k$ and $s_2(2^k) = 1$ for all integers $k \ge 0$, so the most precise asymptotic expansion for $s_2(n)$ is $s_2(n) = O(\log n)$ for $n \to \infty$. However, the summatory function might admit a good asymptotic expansion: For the summatory function of the binary sum of digits function, we have $\sum_{0 \le n < N} s_2(n) = \frac{1}{2} N \log_2 N + N\Phi(\log_2 N)$ for some 1-periodic continuous function Φ as $N \to \infty$; see Delange [3]. In this particular example, there is no error term; in general, an error term is to be expected.

However, *a priori*, it is not clear whether the summatory function of a regular sequence will be smooth enough so that a good asymptotic expansion can be established. In fact, it is known [1, Theorems 2.6 and 2.5] that the forward difference $n \mapsto x(n+1) - x(n)$ of a regular sequence x is again regular. This means that the summatory function of the forward difference of the binary sum of digits function equals the binary sum of digits function and no good asymptotic expansion can be obtained. Thus, as we are able to go forth (summatory function) and back (forward difference), the question arises whether for every regular sequence, there is a non-negative integer k such that its k-fold summatory function admits a good asymptotic expansion. In this paper, we prove that this is the case for all regular sequences except for some degenerate cases (Theorem 5).

For other regular sequences, the sequence itself might admit a good asymptotic expansion. One example are sequences associated with divide-and-conquer schemes [10, 11], for example, the worst case analysis of the number of comparisons in the merge sort algorithm. These are closely related to the so-called "master theorems"; see the discussion in [11]. These sequences are easily seen [11, Equation (2.1)] to be regular sequences (as long as the toll function is regular). While Hwang, Janson, and Tsai [11] provide a direct proof for the asymptotic behaviour and give plenty of examples, the question is whether these results can also be obtained by using the results in [7]. In the present paper, we see such a sequence as the summatory function of its forward difference, and we show that for polynomial toll functions, we get a good asymptotic expansion in the vast majority of cases. The result is formulated in Theorems 7 and 8. In contrast to [11], we are not constrained to cases where the toll function is asymptotically smaller than the sequence and Fourier coefficients can be computed using the results of [7].

The remaining paper is structured as follows. In Section 1.2, we recall the definition and the relevant results on regular sequences. This is followed in Section 1.3 by the statement of our new result on the k-fold summatory function. In Section 1.4, we present the state of the art for divide-and-conquer sequences and state our version the result in Section 1.5. An explicit example is discussed in Section 1.6. Sections 2 and 3 are devoted to the proofs of our theorems.

1.2 Regular Sequences: Definition and State of the Art

We recall the definition[2] of a *regular sequence*; see Allouche and Shallit [1, 2] for characterisations, properties, and an abundance of examples.

▶ **Definition 1.** *Let $q \geq 2$ be an integer. A sequence $x \in \mathbb{C}^{\mathbb{N}_0}$ is said to be q-*regular[3] *if there are a non-negative integer D, a family $A = (A_r)_{0 \leq r < q}$ of $D \times D$ matrices over $\mathbb{C}$, a vector $u \in \mathbb{C}^{1\times D}$ and a vector-valued sequence $v \in (\mathbb{C}^{D\times 1})^{\mathbb{N}_0}$ such that for all $n \in \mathbb{N}_0$, we have*

$$x(n) = uv(n),$$

and such that for all $0 \leq r < q$ and all $n \in \mathbb{N}_0$, we have

$$v(qn + r) = A_r v(n). \tag{3}$$

We call $(u, A, v(0))$ a linear representation *of x and v the* right vector-valued sequence *associated with this linear representation.*

Note that (1) easily follows from (3) by induction; the other direction is contained in [1, Lemma 4.1].

In [7] asymptotic properties were studied. To formulate an abbreviated version of its main result, we first need to recall the notion of the joint spectral radius of a set of square matrices as bounds on matrix products are relevant in view of the representation (1). We fix a vector norm $\|\cdot\|$ on $\mathbb{C}^D$ and consider its induced matrix norm.

▶ **Definition 2.** *Let D be a positive integer and $\mathcal{G}$ be a finite set of $D \times D$ matrices over $\mathbb{C}$.*

1. *The joint spectral radius of $\mathcal{G}$ is defined as*

$$\rho(\mathcal{G}) \coloneqq \lim_{k\to\infty} \sup\{\|G_1 \dots G_k\|^{1/k} \mid G_1, \dots, G_k \in \mathcal{G}\}.$$

2. *We say that $\mathcal{G}$ has the* simple growth property *if*

$$\|G_1 \dots G_k\| = O(\rho(\mathcal{G})^k)$$

holds for all $G_1, \dots, G_k \in \mathcal{G}$ and $k \to \infty$.

For a family $G = (G_i)_{i\in I}$ of $D \times D$ matrices, we set $\rho(G) \coloneqq \rho(\{G_i \mid i \in I\})$ and we say that G has the simple growth property if $\{G_i \mid i \in I\}$ has the simple growth property.

---

[2] Strictly speaking, this is an algorithmic characterisation of a regular sequence which is equivalent to the definition given by Allouche and Shallit [1], who first introduced this concept: they define a sequence x to be q-regular if the kernel

$$\left\{x \circ (n \mapsto q^j n + r) \,\middle|\, j,\, r \in \mathbb{N}_0 \text{ with } 0 \leq r < q^j\right\}$$

is contained in a finitely generated module.

[3] In the standard literature, the basis is frequently denoted by k instead of our q here.

We note that the joint spectral radius and the simple growth property are independent of the chosen norm; cf. [8, Remark 4.2].

For a square matrix M, let $\sigma(M)$ denote the set of eigenvalues of M and by $m_M(\lambda)$ the size of the largest Jordan block of M associated with some $\lambda \in \mathbb{C}$. In particular, we have $m_M(\lambda) = 0$ if $\lambda \notin \sigma(M)$. Finally, we let $\{z\} := z - \lfloor z \rfloor$ denote the fractional part of a real number z. We use Iverson's convention: For a statement S, we set $[\![S]\!] = 1$ if S is true and 0 otherwise; see also Graham, Knuth and Patashnik [6, p. 24].

▶ **Theorem 3** ([7, Theorem A], [4, 5])**.** *Let x be a q-regular sequence with linear representation (u, A, w), and set*

$$B_r := \sum_{0 \le s < r} A_s, \qquad\qquad C := \sum_{0 \le s < q} A_s$$

for $0 \le r < q$.

We choose $R > 0$ as follows: If A has the simple growth property, then we set $R = \rho(A)$. Otherwise, we choose $R > \rho(A)$ such that there is no eigenvalue $\lambda \in \sigma(C)$ with $\rho(A) < |\lambda| \le R$.

Then we have

$$\sum_{0 \le n < N} x(n) = \sum_{\substack{\lambda \in \sigma(C) \\ |\lambda| > R}} N^{\log_q \lambda} \sum_{0 \le k < m_C(\lambda)} \frac{(\log N)^k}{k!} \, \Phi_{\lambda k}(\{\log_q N\}) + O\bigl(N^{\log_q R} (\log N)^{\max\{m_C(\lambda) \colon |\lambda| = R\}}\bigr) \tag{4}$$

as $N \to \infty$, where $\Phi_{\lambda k}$ are suitable 1-periodic functions. If there are no eigenvalues $\lambda \in \sigma(C)$ with $|\lambda| \le R$, the O-term can be omitted.

For $|\lambda| > R$ and $0 \le k < m_C(\lambda)$, the function $\Phi_{\lambda k}$ is Hölder continuous with any exponent smaller than $\log_q(|\lambda|/R)$.

Note that [7] also contains results on how to compute the Fourier coefficients of the periodic fluctuations $\Phi_{\lambda k}$.

1.3 Summatory Functions of Regular Sequences

As announced in Section 1.1, within this paper, we show that for all regular sequences except for some degenerate cases, there is a non-negative integer k such that the k-fold summatory function admits a good asymptotic expansion. In order to formulate our result, we first fix a notation for summatory functions.

▶ **Definition 4.** *For a sequence $x \colon \mathbb{N}_0 \to \mathbb{C}^D$ (for some positive integer D), define the sequence $\Sigma x \colon \mathbb{N}_0 \to \mathbb{C}^D$ by*

$$(\Sigma x)(N) = \sum_{0 \le n < N} x(n).$$

We use the convention that Σ binds more strongly than evaluation, i.e., we write $\Sigma x(N)$ instead of $(\Sigma x)(N)$.

We are now able to formulate our result.

▶ **Theorem 5.** *Let x be a q-regular sequence with linear representation (u, A, w) and set $C := \sum_{0 \le r < q} A_r$. Assume that $C \ne 0$.*

Then there is a non-negative integer k such that $\Sigma^k x$ admits a good asymptotic expansion.

This theorem is proved in Section 2.

1.4 Divide-and-Conquer Sequences: Definition and State of the Art

Hwang, Janson, and Tsai [11] study sequences x with

$$x(n) = \alpha x\Big(\Big\lfloor \frac{n}{2} \Big\rfloor\Big) + \beta x\Big(\Big\lceil \frac{n}{2} \Big\rceil\Big) + g(n) \tag{5}$$

for $n \geq 2$, where α and β are two given positive constants, g is a given function, called the toll function, and $x(1)$ is given.

The simplest version of their result is summarised in the following theorem; more general (weaker assumptions on g) versions are also available.

▶ **Theorem 6** ([11, Corollary 2.14])**.** *Let x be a sequence satisfying (5). Assume that there is an $\varepsilon > 0$ such that $g(n) = O(n^{\log_2(\alpha+\beta)-\varepsilon})$. Then*

$$x(n) = n^{\log_2(\alpha+\beta)}\Phi(\{\log_2 n\}) + O(n^{\log_2(\alpha+\beta)-\varepsilon})$$

for $n \to \infty$ where Φ is a continuous, 1-periodic function.

1.5 Divide-and-Conquer Sequences: Polynomial Toll Function

For divide-and-conquer sequences g, a sequence satisfying the recurrence (5) can be seen as a regular sequence: It is not hard to see that we have

$$\begin{aligned} x(2n) &= (\alpha+\beta)x(n) + g(2n), \\ x(2n+1) &= \alpha x(n) + \beta x(n+1) + g(2n+1) \end{aligned} \tag{6}$$

for $n \geq 1$. Thus x is a 2-recursive sequence in the sense of [8] and therefore 2-regular by [8, Corollary D]. Alternatively, a linear representation for x can also be constructed directly from (6): the associated right vector-valued sequence consists of $n \mapsto x(n)$, $n \mapsto x(n+1)$ and the right vector-valued sequence associated to a linear representation of g; the matrices of the linear representation can then easily be reconstructed from (6) and the linear representation of g. The fact that (6) holds only for $n \geq 1$ (instead of $n \geq 0$) can be fixed; see [1, Proof of Lemma 4.1] or [8, Theorem B].

As announced in Section 1.1, our goal is to see what can be said about the asymptotics of $x(n)$ for $n \to \infty$ using Theorem 3. While the method works for arbitrary regular toll functions; see Remark 14 (although good asymptotic expansions cannot be guaranteed in all cases), we formulate our main result for polynomial toll functions; first versions are contained in the master's thesis [13] of the third author.

▶ **Theorem 7.** *Let $g(n) = \sum_{i=0}^{k} c_i n^i$ be a polynomial of degree $k \geq 1$, x be a sequence satisfying (5). Then the asymptotic behaviour of $x(n)$ for $n \to \infty$ can be described as follows, where Φ and Ψ are 1-periodic continuous functions.*

- ***Case 1a.*** *If $\alpha+\beta > 2^k$ and $2^k > \max\{\alpha,\beta\}$, then*

  $$x(n) = n^{\log_2(\alpha+\beta)}\Phi(\{\log_2 n\}) + n^k\Psi(\{\log_2 n\}) + O(n^{\log_2 \max\{\alpha,\beta\}}).$$

- ***Case 1b.*** *If $\alpha+\beta > 2^k$ and $\max\{\alpha,\beta\} \geq 2^k$, then*

  $$x(n) = n^{\log_2(\alpha+\beta)}\Phi(\{\log_2 n\}) + O(n^{\log_2 \max\{\alpha,\beta\}}(\log n)^{[\![\max\{\alpha,\beta\}=2^k]\!]}).$$

- ***Case 2.*** *If $\alpha+\beta = 2^k$, then*

  $$x(n) = n^k(\log n)\Phi(\{\log_2 n\}) + n^k\Psi(\{\log_2 n\}) + O(n^{\log_2 \max\{\alpha,\beta\}+[\![\alpha=\beta]\!]\varepsilon})$$

  for any $\varepsilon > 0$.

- ***Case 3.*** *If $2^k > \alpha + \beta > 2^{k-1}$, then*

$$x(n) = n^k \Phi(\{\log_2 n\}) + n^{\log_2(\alpha+\beta)} \Psi(\{\log_2 n\}) \\ + O(n^{\log_2 \max\{\alpha,\beta,2^{k-1}\} + [\![\max\{\alpha,\beta\}=2^{k-1}]\!]\varepsilon} (\log n)^{[\![\max\{\alpha,\beta\}<2^{k-1}]\!]})$$

for any $\varepsilon > 0$.

- ***Case 4.*** *If $2^{k-1} \geq \alpha + \beta$, then*

$$x(n) = n^k \Phi(\{\log_2 n\}) + O(n^{k-1} (\log n)^E),$$

where

$$E \coloneqq 1 + [\![\alpha + \beta = 2^{k-1}]\!]([\![k \geq 2 \text{ and } c_{k-1} \neq 0]\!] + [\![k = 1 \text{ and } d_0 + d_1 \neq 0]\!])$$

with

$$d_0 \coloneqq (1 - \beta)x(1) - g(1) + g(0), \qquad d_1 \coloneqq g(1) - (1 - \beta)x(1).$$

This theorem is proved in Section 3.

The case of a constant toll function is somewhat simpler.

▶ **Theorem 8.** *Let $g(n) = c_0$ be a constant toll function and let x be a sequence satisfying* (5)*. Let d_0 and d_1 be defined as in Theorem 7. Then the asymptotic behaviour of $x(n)$ for $n \to \infty$ can be described as follows, where Φ is a 1-periodic continuous function.*

- ***Case 1.*** *If $d_0 = d_1 = 0$, then*

$$x(n) = n^{\log_2(\alpha+\beta)} \Phi(\{\log_2 n\}).$$

- ***Case 2a.*** *If $d_0 \neq 0$ or $d_1 \neq 0$, and $\alpha + \beta > 1$, then*

$$x(n) = n^{\log_2(\alpha+\beta)} \Phi(\{\log_2 n\}) \\ + O(n^{\log_2 \max\{\alpha,\beta,1\} + [\![\max\{\alpha,\beta\}=1]\!]\varepsilon} (\log n)^{[\![\max\{\alpha,\beta\}<1]\!]})$$

for any $\varepsilon > 0$.

- ***Case 2b.*** *If $d_0 \neq 0$ or $d_1 \neq 0$, and $\alpha + \beta \leq 1$, then*

$$x(n) = O((\log n)^{[\![\alpha+\beta=1 \text{ and } d_0+d_1\neq 0]\!]}).$$

This theorem is also proved in Section 3.

1.6 Example

We conclude this introductory section with one example to illustrate the results.

▶ **Example 9.** Consider the divide-and-conquer algorithm for finding the minimum and the maximum of a list of n elements. The number $x(n)$ of comparisons needed satisfies (5) for $n \geq 3$ with $a = b = 1$ and $g(n) = 2$ for $n \geq 3$ and with $x(1) = 0$ and $x(2) = 1$; cf. [10, Example 3.2].

By Theorem 8 (and [8, Theorem B] to deal with the fact that the divide-and-conquer recurrence is only valid for $n \geq 3$ instead of $n \geq 2$; see Section 4 for details), we get

$$x(n) = n\Phi(\{\log_2 n\}) + O(n^\varepsilon)$$

for some 1-periodic continuous function Φ and any $\varepsilon > 0$. The Fourier coefficients of Φ can be computed; cf. Figure 1.

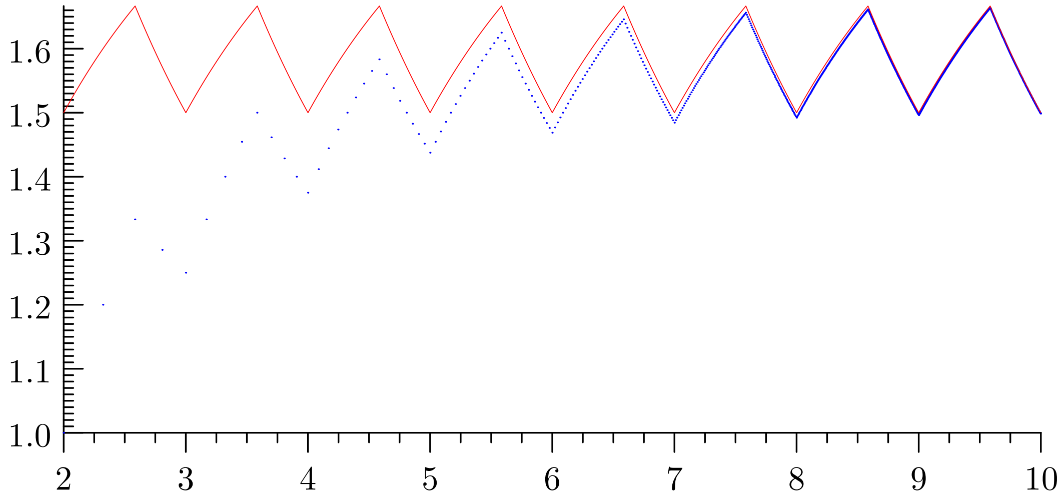

Figure 1 Comparison of the 1-periodic function Φ in the asymptotic expansion determined using Theorem 8 with the empirical values of the sequence.

2 Summatory Functions: Proof of Theorem 5

Before proving Theorem 5, we collect two lemmata on the linear representations of summatory functions and k-fold summatory functions.

The following lemma is implicitly shown in [7, Lemma 12.2], however, it is crucial for our purposes, so we provide a precise formulation and will prove it for self-containedness.

▶ **Lemma 10.** *Let x be a q-regular sequence with linear representation (u, A, w) and associated right vector-valued sequence v. Set*

$$B_r := \sum_{0\le s<r} A_s \text{ for } 0 \le r < q \text{ and } C := \sum_{0\le s<q} A_s. \tag{7}$$

Then we have

$$\Sigma v(qN + r) = C\,\Sigma v(N) + B_r v(N) \tag{8}$$

for all $N \geq 0$ and $0 \leq r < q$.

Additionally, Σx is regular with linear representation $(\widetilde{u}, \widetilde{A}, \widetilde{w})$ with

$$\begin{aligned} \widetilde{u} &:= (u, 0),\\ \widetilde{A}_r &:= \begin{pmatrix} C & B_r \\ 0 & A_r \end{pmatrix} \text{ for } 0 \le r < q,\\ \widetilde{w} &:= \begin{pmatrix} 0 \\ w \end{pmatrix}; \end{aligned}$$

the associated right vector-valued sequence is $\binom{\Sigma v}{v}$.

Proof. By definition, we have $\Sigma v(0) = 0$. Let $N \geq 0$ and $0 \leq r < q$. Then

$$\Sigma v(qN + r) = \sum_{0\le n<qN} v(n) + \sum_{qN\le n<qN+r} v(n).$$

Replacing n by $qm+s$ for $m \in \mathbb{Z}$ and $0 \le s < q$ in the first sum and replacing n by $qN+s$ for $0 \le s < r$ in the second sum yields

$$\Sigma v(qN+r) = \sum_{0\le m<N} \sum_{0\le s<q} v(qm+s) + \sum_{0\le s<r} v(qN+s).$$

Using the linear representation yields

$$\Sigma v(qN+r) = \sum_{0\le m<N} \sum_{0\le s<q} A_s v(m) + \sum_{0\le s<r} A_s v(N) = C\,\Sigma v(N) + B_r v(N).$$

In other words, we have shown (8).

As we have $x = uv$, we also have

$$\Sigma x = u\,\Sigma v = u\,\Sigma v + 0 = \widetilde{u}\begin{pmatrix}\Sigma v\\ v\end{pmatrix}.$$

We conclude that Σx has the given linear representation and associated right vector-valued sequence. ◀

▶ Remark 11. In [7, Lemma 12.2], a very similar result appears in Equation (12.1) there. The difference between Lemma 10 and that equation is an additional summand $(I-A_0)[\![qN+r>0]\!]$.

The reason for the additional summand is that in general, $f(0) = A_0 f(0)$ (with the notations there) does not hold, cf. also [9] for a discussion of this condition.

Iterating the results in Lemma 10 leads to the following lemma.

▶ **Lemma 12.** *Let x be a q-regular sequence with linear representation (u, A, w), $k \ge 1$, and use the notations from (7). Then $\Sigma^k x$ is q-regular with linear representation $(\widetilde{u}, \widetilde{A}, \widetilde{w})$ where $\widetilde{A}_r$ is a block upper triangular matrix with diagonal blocks $q^{k-1}C$, $q^{k-2}C$, …, qC, C, A_r for $0 \le r < q$ and $\widetilde{u}$ and $\widetilde{w}$ are vectors; the associated right vector-valued sequence is $(\Sigma^k v, \Sigma^{k-1} v, \dots, \Sigma v, v)^\top$.*

Proof. We claim that for $m \ge 1$, there are matrices $M_{m,0}$, …, $M_{m,m-1}$ such that

$$\Sigma^m v(qN+r) = q^{m-1} C\,\Sigma^m v(N) + \sum_{0\le j<m} M_{m,j}\,\Sigma^j v(N) \tag{9}$$

holds for all $N \ge 0$ and $0 \le r < q$.

We show (9) by induction on m. For $m = 1$, this is (8). To show (9) for m replaced by $m+1$, we use (8) for v replaced by the regular sequence with associated right vector-valued sequence $\binom{\Sigma v}{v}$ studied in Lemma 10 and the linear representation given there. We obtain

$$\Sigma^m \begin{pmatrix}\Sigma v\\ v\end{pmatrix}(qN+r) = q^{m-1}\begin{pmatrix} qC & \sum_{0\le r<q} B_r\\ 0 & C\end{pmatrix} \Sigma^m \begin{pmatrix}\Sigma v\\ v\end{pmatrix}(N) + \sum_{0\le j<m} \widetilde{M}_{m,j}\,\Sigma^j \begin{pmatrix}\Sigma v\\ v\end{pmatrix}(N)$$

for suitable matrices $\widetilde{M}_{m,j}$ for $0 \le j < m$. Considering the first block row of this equation and collecting terms by powers of Σ leads to (9) with m replaced by $m+1$.

Using (9) for $1 \le m \le k$ yields the linear representation as described in the lemma. ◀

Proof of Theorem 5. Let ρ be the joint spectral radius of A and r the spectral radius (largest absolute value of an eigenvalue) of C.

For some fixed k which will be chosen appropriately later, Lemma 12 yields a linear representation of $\Sigma^k x$ with the properties given there. Let $\widetilde{C} := \sum_{0 \le r < q} \widetilde{A}_r$.

The k-fold summatory function $\Sigma^k x$ admits a good asymptotic expansion if the spectral radius of $\widetilde{C}$ is larger than the joint spectral radius of $\widetilde{A}$. So we compute both.

By Lemma 12, $\widetilde{C}$ is a block upper triangular matrix with diagonal blocks $q^k C$, …, qC, C. So the spectral radius of $\widetilde{C}$ is $q^k r$.

The joint spectral radius of a family of block upper triangular matrices is the maximum of the joint spectral radii of the diagonal blocks; see [12, Proposition 1.5]. This implies that the joint spectral radius of $\widetilde{A}$ is $\max\{q^{k-1}r, q^{k-2}r, \dots, qr, r, \rho\} = \max\{q^{k-1}r, \rho\}$.

It is clear that $q^k r > q^{k-1} r$ so the only condition which needs to be satisfied is $q^k r > \rho$. Such a k exists because $r > 0$ (as $C \neq 0$). ◀

3 Divide-and-Conquer Recurrences: Proof of Theorems 7 and 8

For the proof of Theorems 7 and 8, we will first consider a general regular toll function g and summarise our findings in the general case in Remark 14. Afterwards, we will specialise to a polynomial toll function.

As announced in Section 1.1, we write x as the summatory function of the forward difference of x, i.e.,

$$x(N) = \sum_{0 \le n < N} (x(n+1) - x(n)) + x(0)$$

for $N \ge 0$. Note that strictly speaking, $x(0)$ is not defined in (5). However, we may assume that $g(1)$ and $g(0)$ are somehow defined: they are not used in (5), but we can extend the definition of g if $g(0)$ and $g(1)$ should be undefined.

In order to use Theorem 3, we need a linear representation of the forward difference of x. A first step is the following lemma.

▶ **Lemma 13.** *Let x be a sequence satisfying* (5) *for some toll function g and set $x(0) := 0$ and $h(n) := x(n+1) - x(n)$ for $n \ge 0$. Then*

$$\begin{aligned} h(2n) &= \beta h(n) + g(2n+1) - g(2n) + d_0 \delta_0(n), \\ h(2n+1) &= \alpha h(n) + g(2n+2) - g(2n+1) + d_1 \delta_0(n) \end{aligned} \tag{10}$$

for $n \ge 0$ with d_0, d_1 as in Theorem 7 and $\delta_0(n) := [\![n = 0]\!]$ for $n \ge 0$.

Proof. We can rewrite (6) as

$$\begin{aligned} x(2n) &= (\alpha + \beta) x(n) + g(2n) - g(0)\delta_0(n), \\ x(2n+1) &= \alpha x(n) + \beta x(n+1) + g(2n+1) + ((1-\beta)x(1) - g(1))\delta_0(n) \end{aligned} \tag{11}$$

for $n \ge 0$.

Then (10) follows from

$$\begin{aligned} h(2n) &= x(2n+1) - x(2n), \\ h(2n+1) &= x(2n+2) - x(2n+1) \end{aligned}$$

and inserting (11) into these equations. ◀

▶ Remark 14. Note that h occurs only as $h(n)$ on the right-hand side of of (10). Therefore, a right vector-valued sequence associated with h can be constructed by using $h(n)$ in its first component, then whatever is needed to express $g(2n+2)-g(2n+1)$ and $g(2n+1)-g(2n)$, followed by $\delta_0(n)$. The matrices A_0 and A_1 of the linear representation will thus be in block triangular form; the upper left diagonal elements of A_0 and A_1 being β and α, respectively.

Assume that the contributions of g to the linear representation are small in comparison to α and β. Then the joint spectral radius of the linear representation of h will be $\max\{\alpha, \beta\}$, whereas the matrix C as in Theorem 3 will have a dominant eigenvalue $\alpha+\beta$, which is larger than the joint spectral radius. Thus we will have a good asymptotic expansion in this case.

We now turn to the case of a polynomial toll function.

▶ **Lemma 15.** *Let x and h be as in Lemma 13 with a polynomial toll function $g(n) = \sum_{i=0}^{k} c_i n^i$ for some $k \ge 0$ and some constants c_0, …, c_k with $c_k \neq 0$.*

Set

$$b_{0j} := \sum_{i=j+1}^{k} \binom{i}{j} 2^j c_i, \qquad b_{1j} := \sum_{i=j+1}^{k} \binom{i}{j} (2^i - 2^j) c_i,$$

$$a_{0ij} := [\![j = i]\!] 2^i, \qquad a_{1ij} := \binom{i}{j} 2^j$$

for $0 \le i < k$ and $0 \le j < k$ and

$$b_r := (b_{r(k-1)}, \dots, b_{r0}), \qquad \widetilde{A}_r := (a_{rij})_{\substack{i=k-1,\dots,0,\\ j=k-1,\dots,0}}$$

$$\mu_r := \begin{cases} \beta & \text{if } r = 0, \\ \alpha & \text{if } r = 1, \end{cases} \qquad A_r := \begin{pmatrix} \mu_r & b_r & d_r \\ 0 & \widetilde{A}_r & 0 \\ 0 & 0 & [\![r = 0]\!] \end{pmatrix}$$

for $r \in \{0, 1\}$ and

$$u := (1, 0, \dots, 0) \in \mathbb{C}^{1\times(k+2)}, \qquad w := \begin{cases} (x(1), 0, \dots, 0, 1, 1)^\top \in \mathbb{C}^{(k+2)\times 1} & \text{if } k \ge 1, \\ (x(1), 1)^\top \in \mathbb{C}^{2\times 1} & \text{if } k = 0. \end{cases}$$

Then (u, A, w) is a linear representation for h.

If $d_0 = d_1 = 0$, then $(\widetilde{u}, \widetilde{A}, \widetilde{w})$ is also a linear representation for h where $\widetilde{u}$, $\widetilde{A}$, and $\widetilde{w}$ arise from from u, A, and w by removing the last column, the last row and column, and the last row, respectively.

We remark that A_r is an upper triangular matrix for $r \in \{0, 1\}$ because $\binom{i}{j} = 0$ for $j > i$ (and indices in $\widetilde{A}_0$ and $\widetilde{A}_1$ are decreasing).

Proof. As a right vector-valued sequence v, we choose

$$n \mapsto (h(n), n^{k-1}, \dots, 1, \delta_0(n))^\top.$$

We immediately check that $v(0) = w$ and that $uv(n) = h(n)$ holds for all $n \ge 0$.

Using the binomial theorem repeatedly, we get

$$\begin{aligned}
g(2n+1) - g(2n) &= \sum_{i=0}^{k} c_i \sum_{j=0}^{i-1} \binom{i}{j} 2^j n^j &&= \sum_{j=0}^{k-1} b_{0j} n^j, \\
g(2n+2) - g(2n+1) &= \sum_{i=0}^{k} c_i \sum_{j=0}^{i-1} \binom{i}{j} 2^j n^j (2^{i-j} - 1) &&= \sum_{j=0}^{k-1} b_{1j} n^j, \\
(2n)^i &= 2^i n^i &&= \sum_{j=0}^{k-1} a_{0ij} n^j \qquad \text{for } 0 \le i < k, \\
(2n+1)^i &= \sum_{j=0}^{i} \binom{i}{j} 2^j n^j &&= \sum_{j=0}^{k-1} a_{1ij} n^j \qquad \text{for } 0 \le i < k.
\end{aligned}$$

We verify that (10) translates into

$$v(2n+r) = A_r v(n)$$

for $r \in \{0,1\}$ and $n \ge 0$.

If $d_0 = d_1 = 0$, then A_0 and A_1 are block diagonal matrices. The lower right block is not taken into account when multiplying by u, so the lower right block can be omitted.

Thus the result follows. ◀

▶ **Lemma 16.** *Let x and h be as in Lemma 13 and g, (u, A, w) be as in Lemma 15. Assume that $k \ge 1$. Set $C = A_0 + A_1$.*

Then $\rho(A) = \max\{\alpha, \beta, 2^{k-1}\}$ and $\sigma(C) = \{\alpha + \beta, 2^k, 2^{k-1}, \dots, 2, 1\}$. If $\max\{\alpha, \beta\} \ne 2^{k-1}$, then A has the simple growth property.

Proof. By Lemma 15, A_0 is an upper triangular matrix with diagonal elements β, 2^{k-1}, …, 1, 1 and A_1 is an upper triangular matrix with diagonal elements α, 2^{k-1}, …, 1, 0.

The joint spectral radius of a set of upper triangular matrices is the maximum of the diagonal elements of the matrices; see [12, Proposition 1.5]. This implies that $\rho(A) = \max\{\alpha, \beta, 2^{k-1}\}$. By [8, Lemma 4.5], A has the simple growth property if the joint spectral radius of A occurs only once as a maximum of corresponding diagonal elements of A_0 and A_1. This is the case if $\max\{\alpha, \beta\} \ne 2^{k-1}$.

We also conclude that C is an upper triangular matrix with diagonal elements $\alpha+\beta$, 2^k, 2^{k-1}, …, 2, 1. Thus the assertion for $\sigma(C)$ follows. ◀

Proof of Theorem 7. To prove Theorem 7 using Theorem 3, we need to determine the eigenvalues of C which are greater or equal than the joint spectral radius of A_0 and A_1 (with the notations of Lemma 16) and the size of the largest Jordan block associated with any such eigenvalue.

As $\rho(A) = \max\{\alpha, \beta, 2^{k-1}\} \ge 2^{k-1} \ge 1$, it is clear that the only relevant eigenvalues of C are contained in the set $\{\alpha + \beta, 2^k, 2^{k-1}\}$.

The main case distinction of Theorem 7 concerns the order of $\alpha + \beta$, 2^k, and 2^{k-1}.

- **Case 1:** $\alpha + \beta > 2^k$. This implies that $\max\{\alpha, \beta\} > 2^{k-1}$ and therefore $\rho(A) = \max\{\alpha, \beta\}$ and A has the simple growth property. The eigenvalue $\alpha + \beta$ of C is larger than the joint spectral radius and is a simple eigenvalue. The eigenvalue 2^k is also a simple eigenvalue. If it is larger than the joint spectral radius, we are in Case 1a and have two asymptotic terms larger than the error term. If $2^k \le \max\{\alpha, \beta\}$, we are in Case 1b and have one asymptotic term larger than the error term; there is a logarithmic factor in the error term if and only if 2^k equals the joint spectral radius.

- **Case 2:** $\alpha+\beta = 2^k$. In this case, $2^k = \alpha+\beta$ has algebraic multiplicity 2 as an eigenvalue of C. We note that $C - 2^k I$ has the shape
$$C - 2^k I = \begin{pmatrix} 2^k - 2^k & \binom{k}{k-1}2^k c_k & b' \\ 0 & 2^k - 2^k & b'' \\ 0 & 0 & C' \end{pmatrix} = \begin{pmatrix} 0 & k2^k c_k & b' \\ 0 & 0 & b'' \\ 0 & 0 & C' \end{pmatrix}$$
for some vectors b', b'' and some upper triangular matrix C' with diagonal elements $2^{k-1} - 2^k$, $2^{k-2} - 2^k$, $\dots$, $2^0 - 2^k$. As $c_k \neq 0$ by assumption, the kernel of $C - 2^k I$ has dimension 1. We conclude that the size $m_C(2^k)$ of the largest Jordan block of C associated with the eigenvalue 2^k equals 2. So we have a logarithmic factor in the asymptotic main term.
We also note that $\max\{\alpha, \beta\} \geq 2^{k-1}$ holds in this case with equality if and only if $\alpha = \beta = 2^{k-1}$. So the joint spectral radius $\rho(A)$ equals $\max\{\alpha, \beta\}$ and A has the simple growth property unless $\alpha = \beta = 2^{k-1}$.
- **Case 3:** $2^k > \alpha + \beta > 2^{k-1}$. In this case, we have $\max\{\alpha, \beta\} > 2^{k-2}$ and C has two simple dominant eigenvalues 2^k and $\alpha+\beta$. We do not have additional information about the joint spectral radius. If $\max\{\alpha, \beta\} \neq 2^{k-1}$, then A has the simple growth property. If $\max\{\alpha, \beta\} < 2^{k-1}$, then the joint spectral radius of A is 2^{k-1}. As C has 2^{k-1} as an eigenvalue as well, there is a logarithmic factor in the error term in exactly this situation.
- **Case 4:** $2^{k-1} \geq \alpha + \beta$. In this case, we have $\max\{\alpha, \beta\} < 2^{k-1}$, so $\rho(A) = 2^{k-1}$ and A has the simple growth property. There is only one eigenvalue of C larger than this joint spectral radius, namely 2^k. We have to determine $m_C(2^{k-1})$ in order to find out the exponent of $\log n$ in the error term. The algebraic multiplicity of 2^{k-1} as an eigenvalue of C equals $1 + [\![\alpha + \beta = 2^{k-1}]\!]$. So if $\alpha + \beta < 2^{k-1}$, we have $m_C(2^{k-1}) = 1$ and a factor $\log n$ in the error term.
We now consider the case $\alpha + \beta = 2^{k-1}$ and $k \geq 2$. We note that $C - 2^{k-1} I$ has the shape
$$C - 2^{k-1} I = \begin{pmatrix} 2^{k-1} - 2^{k-1} & \binom{k}{k-1}2^k c_k & \binom{k-1}{k-2}2^{k-1}c_{k-1} + \binom{k}{k-2}2^k c_k & b' \\ 0 & 2^k - 2^{k-1} & \binom{k-1}{k-2}2^{k-2} & b'' \\ 0 & 0 & 2^{k-1} - 2^{k-1} & b''' \\ 0 & 0 & 0 & C' \end{pmatrix}$$
$$= \begin{pmatrix} 0 & k2^k c_k & (k-1)2^{k-1}c_{k-1} + k(k-1)2^{k-1}c_k & b' \\ 0 & 2^{k-1} & (k-1)2^{k-2} & b'' \\ 0 & 0 & 0 & b''' \\ 0 & 0 & 0 & C' \end{pmatrix}$$
for suitable vectors b', b'', b''' and a regular upper triangular matrix C'. Subtracting $2kc_k$ times the second row from the first row does not change the kernel, so we get
$$\ker(C - 2^{k-1} I) = \ker \begin{pmatrix} 0 & 0 & (k-1)2^{k-1}c_{k-1} & b' - 2kc_k b'' \\ 0 & 2^{k-1} & (k-1)2^{k-2} & b'' \\ 0 & 0 & 0 & b''' \\ 0 & 0 & 0 & C' \end{pmatrix}.$$
We conclude that $\dim\ker(C - 2^{k-1} I) = 1 + [\![c_{k-1} = 0]\!]$ and therefore $m_C(2^{k-1}) = 1 + [\![c_{k-1} \neq 0]\!]$.
Finally we turn to the case that $\alpha + \beta = 2^{k-1}$ and $k = 1$. Then we have
$$C - I = \begin{pmatrix} 0 & 2c_1 & d_0 + d_1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If $d_0 = d_1 = 0$, by the last statement of Lemma 15, the last row and column of $C - I$ are omitted and 1 has algebraic multiplicity 1 as an eigenvalue of C. We conclude that $\dim\ker C - I = 1 + [\![d_0 + d_1 = 0]\!] - [\![d_0 = 0]\!][\![d_1 = 0]\!]$ and therefore $m_C(2^{k-1}) = 1 + [\![d_0 + d_1 \neq 0]\!]$.

So, to summarise, $m_C(2^{k-1}) = 1 + E$ where E is defined in Theorem 7.

◀

Proof of Theorem 8. If $d_0 = d_1 = 0$, then Lemma 15 yields $A_0 = (\beta)$, $A_1 = (\alpha)$, and $C = (\alpha + \beta)$, so the joint spectral radius of A equals $\max\{\alpha, \beta\}$ which is strictly less than the unique eigenvalue $\alpha + \beta$ of C. As there is no eigenvalue of C less than the joint spectral radius of A, there is no error term. The result follows in this case.

From now on, we assume that $d_0 \neq 0$ or $d_1 \neq 0$. Lemma 15 yields

$$A_0 = \begin{pmatrix} \beta & d_0 \\ 0 & 1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} \alpha & d_1 \\ 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} \alpha + \beta & d_0 + d_1 \\ 0 & 1 \end{pmatrix}.$$

We see that the joint spectral radius of A is $\max\{\alpha, \beta, 1\}$ and that C has eigenvalues $\alpha + \beta$ and 1. It is now easy to deduce the assertions of the theorem. ◀

4 Details on Example 9

We proceed as outlined in Remark 14. Setting $x(0) = 0$ as usual, we have

$$x(n) = x(\lfloor n/2 \rfloor) + x(\lceil n/2 \rceil) + 2 - [\![n = 2]\!] - 2[\![n = 1]\!] - 2[\![n = 0]\!]$$

for $n \geq 0$. Equivalently, we have

$$\begin{aligned} x(2n) &= 2x(n) + 2 - [\![n = 1]\!] - 2[\![n = 0]\!], \\ x(2n + 1) &= x(n) + x(n + 1) + 2 - 2[\![n = 0]\!] \end{aligned}$$

for $n \geq 0$. Setting $h(n) \coloneqq x(n + 1) - x(n)$ leads to

$$\begin{aligned} h(2n) &= h(n) + [\![n = 1]\!], \\ h(2n + 1) &= h(n) + [\![n = 0]\!] \end{aligned}$$

for $n \geq 0$. This defines a 2-regular sequence with a linear representation (u, A, w) with associated right vector-valued sequence v defined by $v(n) = (h(n), \delta_1(n), \delta_0(n))$ with

$$\begin{aligned} A_0 &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & A_1 &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \\ u &= (1, 0, 0), & w &= (0, 0, 1)^\top. \end{aligned}$$

Here, δ_1 is defined by $\delta_1(n) \coloneqq [\![n = 1]\!]$ for $n \geq 0$.

We can now use SageMath[4] to compute Fourier coefficients and to produce Figure 1.

[4] The code for this example is available at `https://arxiv.org/src/2403.06589/anc`; it uses the code accompanying [7] which is available at at `https://gitlab.com/dakrenn/regular-sequence-fluctuations`.

References

1 Jean-Paul Allouche and Jeffrey Shallit. The ring of k-regular sequences. *Theoret. Comput. Sci.*, 98(2):163–197, 1992. doi:10.1016/0304-3975(92)90001-V.

2 Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences: Theory, applications, generalizations.* Cambridge University Press, Cambridge, 2003. doi:10.1017/CBO9780511546563.

3 Hubert Delange. Sur la fonction sommatoire de la fonction “somme des chiffres”. *Enseign. Math. (2)*, 21:31–47, 1975. doi:10.5169/seals-47328.

4 Philippe Dumas. Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences. *Linear Algebra Appl.*, 438(5):2107–2126, 2013. doi:10.1016/j.laa.2012.10.013.

5 Philippe Dumas. Asymptotic expansions for linear homogeneous divide-and-conquer recurrences: Algebraic and analytic approaches collated. *Theoret. Comput. Sci.*, 548:25–53, 2014. doi:10.1016/j.tcs.2014.06.036.

6 Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics. A foundation for computer science.* Addison-Wesley, second edition, 1994.

7 Clemens Heuberger and Daniel Krenn. Asymptotic analysis of regular sequences. *Algorithmica*, 82(3):429–508, 2020. doi:10.1007/s00453-019-00631-3.

8 Clemens Heuberger, Daniel Krenn, and Gabriel F. Lipnik. Asymptotic analysis of q-recursive sequences. *Algorithmica*, 84(9):2480–2532, 2022. doi:10.1007/s00453-022-00950-y.

9 Clemens Heuberger, Daniel Krenn, and Gabriel F. Lipnik. A note on the relation between recognisable series and regular sequences, and their minimal linear representations. *J. Symbolic Comput.*, 2024. doi:10.1016/j.jsc.2023.102295.

10 Hsien-Kuei Hwang, Svante Janson, and Tsung-Hsi Tsai. Exact and asymptotic solutions of a divide-and-conquer recurrence dividing at half: Theory and applications. *ACM Trans. Algorithms*, 13(4):Art. 47, 43 pp., October 2017. doi:10.1145/3127585.

11 Hsien-Kuei Hwang, Svante Janson, and Tsung-Hsi Tsai. Identities and periodic oscillations of divide-and-conquer recurrences splitting at half. *Adv. in Appl. Math.*, 155:Paper No. 102653, 53, 2024. doi:10.1016/j.aam.2023.102653.

12 Raphaël Jungers. *The joint spectral radius. Theory and applications*, volume 385 of *Lecture Notes in Control and Information Sciences.* Springer-Verlag, Berlin, 2009. doi:10.1007/978-3-540-95980-9.

13 Tobias Lechner. Application of theory for regular sequences on divide and conquer algorithms. Master’s thesis, University of Klagenfurt, 2024. urn:nbn:at:at-ubk:1-52067. URL: https://resolver.obvsg.at/urn:nbn:at:at-ubk:1-52067.