A Master Theorem for Discrete Divide and Conquer Recurrences*

Michael Drmota[†] Wojciech Szpankowski[‡]

Dedicated to Philippe Flajolet 1948-2011

Abstract

Divide-and-conquer recurrences are one of the most studied equations in computer science. Yet, *discrete* versions of these recurrences, namely

$$T(n) = a_n + \sum_{j=1}^{m} b_j T\left(\lfloor p_j n + \delta_j \rfloor\right) + \sum_{j=1}^{m} b'_j T\left(\lceil p_j n + \delta'_j \rceil\right)$$

for some known sequence a_n and given b_j, b'_j, p_j and δ_j, δ'_j, present some challenges. The discrete nature of this recurrence (represented by the floor function) introduces certain oscillations not captured by the traditional Master Theorem, for example due to Akra and Bazzi who primary studied the continuous version of the recurrence. We apply powerful techniques such as Dirichlet series, Mellin-Perron formula, and (extended) Tauberian theorems of Wiener-Ikehara to provide a complete and precise solution to this basic computer science recurrence. We illustrate applicability of our results on several examples including a popular and fast arithmetic coding algorithm due to Boncelet for which we estimate its average redundancy and prove the Central Limit Theorem for the phrase length. To the best of our knowledge, *discrete* divide and conquer recurrences were not studied in this generality and such detail; in particular, this allows us to compare the redundancy of Boncelet's algorithm to the (asymptotically) optimal Tunstall scheme.

Key Words: Divide-and-conquer recurrence, mergesort, Karatsuba algorithm, Strassen algorithm, Boncelet's data compression algorithm, Dirichlet series, Mellin-Perron formula, Tauberian theorem.

*The work of these authors was supported in part by the Austrian Science Foundation FWF Grant No. S9604, by NSF Science and Technology Center Grant CCF-0939370, NSF Grants DMS-0800568, and CCF-0830140, and the MNSW grant N206 369739.

[†]Inst. Discrete Mathematics and Geometry, TU Wien, A-1040 Wien, Austria, michael.drmota@tuwien.ac.at.

[‡]Department of Computer Science, Purdue University, West Lafayette, IN 47907-2066 U.S.A., spa@cs.purdue.edu

1 Introduction

Divide and conquer is a very popular strategy to design algorithms. It splits the input into several smaller subproblems, solving each subproblem separately, and then knitting together to solve the original problem. Typical examples include heapsort, mergesort, discrete Fourier transform, queues, sorting networks, compression algorithms, and so forth [7, 22, 26, 29, 30]. While it is relatively easy to determine the general growth order for the algorithm complexity, a precise asymptotic analysis is often appreciably more subtle. Our goal is to present such an analysis for *discrete* divide and conquer recurrences.

The complexity of a divide and conquer algorithm is well described by its divide and conquer recurrence. We assume that the problem is split into $m \geq 2$ subproblems. It is natural to assume that there is a cost associated with combining subproblems together to find the solution. We denote such a cost by a_n, where n is the size of the original problem. In addition, each subproblem may contribute in a different way to the final solution; we represent this by coefficients b_j and $\overline{b}_j$ for $1 \leq j \leq m$. Finally, we postulate that the original input n is divided into subproblems of size $\lfloor h_j(n) \rfloor$ and $\lceil \overline{h}_j(n) \rceil$, $1 \leq j \leq m$, where $h_j(x)$ and $\overline{h}_j(x)$ are functions that satisfy $h_j(x) \sim \overline{h}_j(x) \sim p_j x$ for $x \to \infty$ and for some $0 < p_j < 1$. We aim at presenting precise asymptotic solutions of *discrete* divide and conquer recurrences of the following form [7]

$$T(n) = a_n + \sum_{j=1}^{m} b_j T\left(\lfloor h_j(x) \rfloor\right) + \sum_{j=1}^{m} \overline{b}_j T\left(\left\lceil \overline{h}_j(x) \right\rceil\right) \qquad (n \geq 2). \tag{1}$$

A popular approach to solve this recurrence is to relax it to a *continuous* version of the following form (hereafter we assume $\overline{b}_j = 0$ for simplicity)

$$T(x) = a(x) + \sum_{j=1}^{m} b_j T(h_j(x)), \qquad x > 1, \tag{2}$$

(where $h_j(x) \sim p_j x$ with $0 < p_j < 1$) and solve it using a Master Theorem as for example in [7, 26]. This is usually quite powerful and provides order of the growth for $T(x)$. The most general solution of (2) is due to Akra and Bazzi [2] who proved (under certain regularity assumptions, namely that $a'(x)$ is of polynomial growth and that $h_j(x) - p_j x = O(x/(\log x)2))$

$$T(x) = \Theta\left(x^{s_0}\left(1 + \int_1^x \frac{a(u)}{u^{s_0+1}} du\right)\right),$$

where s_0 is a unique real root of

$$\sum_j b_j p_j^s = 1.$$

Actually this also leads directly to

$$T(n) = \Theta\left(n^{s_0}\left(1 + \sum_{j=1}^{n} \frac{a_j}{j^{s_0+1}}\right)\right)$$

in the discrete version provided that $a_{n+1} - a_n$ is at most of polynomial growth.

For more precise results of the continuous version one can apply Mellin transform techniques [14, 15, 30]. Indeed, let

$$t(s) = \int_0^{\infty} T(x) x^{s-1} dx$$

be the Mellin transform of $T(x)$. Then using standard properties of the Mellin transform applied to the (slightly simplified) divide and conquer recurrence $T(x) = a(x) + \sum_{j=1}^{m} b_j T(p_j x)$ we arrive at

$$t(s) = \frac{a(s) + g(s)}{1 - \sum_{j=1}^{m} b_j p_i^{-s}},$$

where $a(s)$ is the Mellin transform of $a(x)$, and $g(s)$ is an additional function due to the initial conditions. Suppose that $a(s)$ and $g(s)$ are analytic for $\Re(s) \geq -s_0$, where $-s_0$ is the root of $1 = \sum_j b_j p_i^{-s}$. Then we recover the asymptotics of $T(x)$ showing that

$$T(x) \sim C x^{s_0} \quad \text{or} \quad T(x) \sim \Psi(\log x) x^{s_0}$$

where C is a constant, and $\Psi(x)$ is a discontinuous periodic function when the logarithms $\log p_j$ are rationally related (i.e., $\log p_i$ are integer multiples of a real number as in Definition 1)

Discrete versions of the divide and conquer recurrence, given by (1) are more subtle and require a different approach. We first apply Dirichlet series (closely related to the Mellin transform) that better captures the discrete nature of the recurrence, and then use Tauberian theorems (and also the Mellin-Perron formula) to obtain asymptotics for $T(n)$. The corresponding result, a precise *Discrete Master Theorem*, is stated in Theorem 1 of the next section. As in the continuous case the solution depends crucially on the relation between $\log p_1, \ldots, \log p_m$; when $\log p_1, \ldots, \log p_m$ are rationally related the final solution will exhibit some oscillation that disappears when $\log p_1, \ldots, \log p_m$ are irrationally related. This phenomenon was already observed for other discrete recurrences [9, 12].

In the nutshell, in Theorem 1 for nondecreasing sequences a_n of the form $a_n = C n^{\sigma} (\log n)^{\alpha}$ with $C > 0$, and irrationally related $\log p_1, \ldots, \log p_m$ we prove

$$T(n) = \begin{cases} C_1 + o(1) & \text{if } \alpha = \sigma = 0 \text{ and } s_0 < 0, \\ C_2 \log n + C_2' + o(1) & \text{if } \alpha = \sigma = s_0 = 0, \\ C_3 (\log n)^{\alpha+1} \cdot (1 + o(1)) & \text{if } \sigma = s_0 = 0, \\ C_4 n^{s_0} \cdot (1 + o(1)) & \text{if } \sigma < s_0 \text{ and } s_0 > 0, \\ C_5 n^{s_0} (\log n)^{\alpha+1} \cdot (1 + o(1)) & \text{if } \sigma = s_0 > 0 \text{ and } \alpha \neq -1, \\ C_5 n^{s_0} \log\log n \cdot (1 + o(1)) & \text{if } \sigma = s_0 > 0 \text{ and } \alpha = -1, \\ C_6 (\log n)^{\alpha} (1 + o(1)) & \text{if } \sigma = 0 \text{ and } s_0 < 0, \\ C_7 n^{\sigma_a} (\log n)^{\alpha} \cdot (1 + o(1)) & \text{if } \sigma > s_0 \text{ and } \sigma > 0, \end{cases} \tag{3}$$

where the explicitly computable constants $C_1, C_2, C_3, C_4, C_5, C_6, C_7$ are positive and C_2' is real. When $\log p_1, \ldots, \log p_m$ are rationally related, $T(n)$ behaves as in the irrationally related case with the following two exceptions:

$$\begin{array}{ll} C_2 \log n + \Psi_2(\log n) + o(1) & \text{if } \alpha = \sigma = s_0 = 0, \\ \Psi_4(\log n)\, n^{s_0} \cdot (1 + o(1)) & \text{if } \sigma < s_0 \text{ and } s_0 > 0, \end{array} \tag{4}$$

where C_2 is positive and $\Psi_2(t), \Psi_4(t)$ are periodic functions with period L (with usually countably many discontinuities).

It should be remarked that the order of magnitude of $T(n)$ can be checked easily by the Akra-Bazzi theorem [2]. If we just know an upper bound for a_n which is of the form $a_n = O(n^{\sigma} (\log n)^{\alpha})$ – even if a_n is not necessarily increasing – the Akra-Bazzi theorem provides an upper bound for $T(n)$ which is of form stated in (4). Furthermore, if we know a_n only approximately (e.g., if $a_n = n^{\sigma} (\log n)^{\alpha} + O((n_1^{\sigma} (\log n)_1^{\alpha})$ with $\sigma_1 < \sigma$ and if $\sigma = \sigma_1$ but $\alpha_1 < \alpha$), then our results still holds approximately.

As a featured application of our results and techniques developed for solving the general discrete divide and conquer recurrence, we shall present a comprehensive analysis of a data compression algorithm due to Boncelet [4], where we need even more precise results than stated in Theorem 1. Boncelet's algorithm is a variable-to-fixed data compression scheme. One of the best variable-to-fixed scheme belongs to Tunstall [31]; another variation is due to Khodak [20]. Boncelet's algorithm is based on the divide and conquer strategy, and therefore is very fast and easy to implement. The question arises how it compares to the (asymptotically) optimal Tunstall algorithm. In Theorem 2 and Corollary 1 we provide an answer by first computing the redundancy of the Boncelet scheme (i.e., the excess of code length over the optimal code length) and compare it to the redundancy of the Tunstall code. In this case precise asymptotics of the Boncelet recurrence are crucial. We also prove in Theorem 3 that the phrase length of the Boncelet's algorithm obeys the central limit law, as for the Tunstall algorithm [10].

The literature on continuous divide and conquer recurrence is very extensive. We mention here [2, 7, 6]. Unfortunately, the discrete version of the recurrence has received much less attention, especially with respect to precise asymptotics. Flajolet and Golin [13] and Cheung et al. [5] use similar techniques to ours, however, their recurrence is a simpler one with $p_1 = p_2 = 1/2$. Erdős et al. [11] apply renewal theory and Hwang [18] (cf. also [19]) analytic techniques when dealing with similar recurrences. However, the approach presented in this paper is generalized and somewhat simplified by using a combination of methods such as Tauberian theorems and Mellin-Perron techniques. To the best of our knowledge, there is no comprehensive analysis of the discrete divide and conquer recurrences and therefore there is no precise redundancy analysis for the Boncelet's algorithm.

The paper is organized as follows. In the next two sections we present our main results regarding the discrete divide and conquer recurrence, and the Boncelet's algorithm. All proofs are delayed till the last section. In the Appendix A we discuss analytic continuations properties of certain Dirichlet series, and in the Appendix B we present the Wiener-Ikehara Tauberian theorem and several extensions.

2 Main Results

In this section we present our main results, including an asymptotic solution to a general discrete divide and conquer recurrence, and its application to an arithmetic coding algorithm due to Boncelet [4].

2.1 Divide and Conquer Recurrence

For $m \geq 2$, let $b_1, \ldots, b_m$ and $\overline{b}_1, \ldots, \overline{b}_m$ be positive real numbers and $h_j(x)$ and $\overline{h}_j(x)$ non-decreasing positive functions with $h_j(x) = p_j x + O(x^{1-\delta})$ and $\overline{h}_j(x) = p_j x + O(x^{1-\delta})$ for some positive numbers $p_j < 1$ and some $\delta > 0$ (for $1 \leq j \leq m$). We consider a (general) divide and conquer recurrence: given $T(0) \leq T(1)$ for $n \geq 2$ we set

$$\begin{aligned} T(n) &= a_n + \sum_{j=1}^{m} b_j T\left(\lfloor h_j(x) \rfloor\right) + \sum_{j=1}^{m} \overline{b}_j T\left(\left\lceil \overline{h}_j(x) \right\rceil\right) \qquad (n \geq 2), \qquad (5) \\ &= a_n + \sum_{j=1}^{m} b_j T\left(\left\lfloor p_j x + O(x^{1-\delta}) \right\rfloor\right) + \sum_{j=1}^{m} \overline{b}_j T\left(\left\lceil p_j x + O(x^{1-\delta}) \right\rceil\right) \end{aligned}$$

where $(a_n)_{n\geq 2}$ is a known *non-negative* and *non-decreasing* sequence. We also assume that $h_j(2) < 2$ and $\overline{h}_j(2) \leq 1$ (for $1 \leq j \leq m$) so that the recurrence is well defined. It follows by induction that $T(n)$ is nondecreasing, too. In order to solve recurrence (5), we use Dirichlet series [3, 30]. In fact, in the proof presented in Section 4 we make use of the following Dirichlet series

$$\widetilde{T}(s) = \sum_{n=1}^{\infty} \frac{T(n+2) - T(n+1)}{n^s} \tag{6}$$

from which we can calculate $\sum_{i=1}^{n-2} T(i+2) - T(i+1) = T(n) - T(2)$.

For an asymptotic solution of recurrence (5), we will make some assumptions regarding the Dirichlet series of the known sequence a_n. We postulate that the abscissa of absolute convergence σ_a of the Dirichlet series

$$\widetilde{A}(s) = \sum_{n=1}^{\infty} \frac{a_{n+2} - a_{n+1}}{n^s} \tag{7}$$

is finite (or $-\infty$), hence $\widetilde{A}(s)$ represents an analytic function for $\Re(s) > \sigma_a$. For example, if we know that a_n is non-decreasing and

$$a_n = O(n^{\sigma}(\log n)^{\alpha})$$

for some real number σ and α, then $\widetilde{A}(s)$ converges (absolutely) for all s with $\Re(s) > \sigma$. In particular, we have $\sigma_a \leq \sigma$.

Analytically, these observations follow from the fact, proved in Section 4, that the Dirichlet series $\widetilde{T}(s)$ can be expressed as

$$\widetilde{T}(s) = \frac{\widetilde{A}(s) + B(s)}{1 - \sum_{j=1}^{m}(b_j + \overline{b}_j)\, p_j^s} \tag{8}$$

for some analytic function $B(s)$ and $\widetilde{A}(s)$ as in (7). For the asymptotic analysis, we appeal to the Tauberian theorem by Wiener-Ikehara and an analysis based on the Mellin-Perron formula (see Appendix B and Section 4.3). Both approaches rely on the singular behaviour of $\widetilde{T}(s)$. By the Mellin-Peron formula, we shall observe that

$$T(n) = T(2) + \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \widetilde{T}(s) \frac{(n-\frac{3}{2})^s}{s}\, ds. \tag{9}$$

Hence, the asymptotic behavior of $T(n)$ depends on the singular behaviour of $\widetilde{A}(s)$, $s = 0$, and roots of the denominator in (8), that is, roots of the *characteristic equation*

$$\sum_{j=1}^{m}(b_j + \overline{b}_j)\, p_j^s = 1. \tag{10}$$

We denote by s_0 the unique real solution of this equation.

A master theorem as presented in this paper has usually three (major) parts. In the first case, the asymptotics of $T(n)$ is driven by the recurrence and does not depend on a_n, in the second case, there is an *interaction* between the internal structure of the recurrence and the sequence a_n (resonance), and in the third case, the (asymptotic) behaviour of a_n dominates. Informally, the

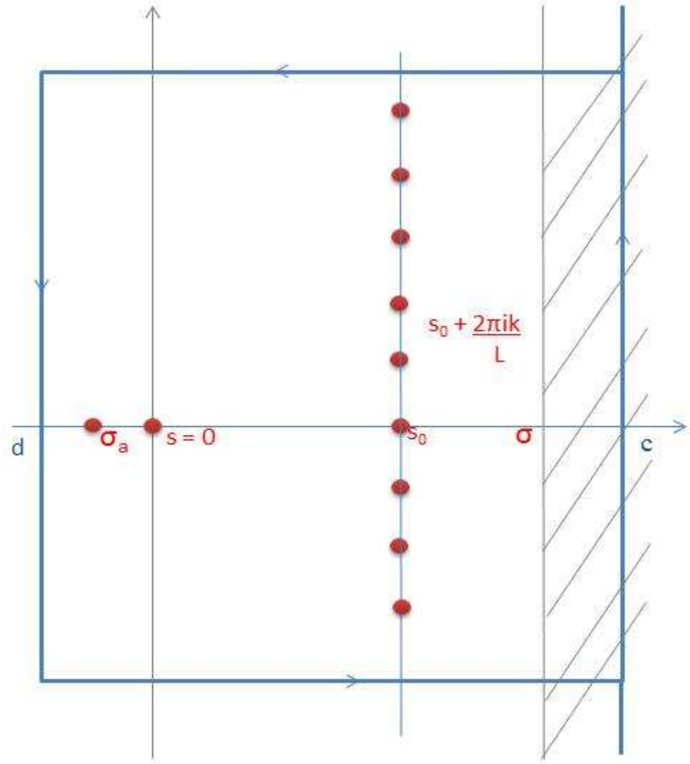

Figure 1: Illustration to the asymptotic analysis of the divide and conquer recurrence

first case corresponds to $s_0 > \sigma_a$, the second case to $s_0 = \sigma_a$, and the third one to $s_0 < \sigma_a$. This is illustrated in Figure 1 as an application of the residue analysis applied to the integral in (9). The interplay between the poles at $s = 0$, $s = \sigma_a$ and s_0 determines the asymptotic behavior.

We will handle these cases separately. Nevertheless, if $s_0 = \sigma_a$ or if $s_0 < \sigma_a$ we have to assume some regularity properties about the sequence a_n in order to cope with the asymptotics of $T(n)$. We assume that $\widetilde{A}(s)$ has a certain extension to a region that contains the line $\Re(s) = \sigma_a$ with a pole-like singularity at $s = \sigma_a$. To be more precise, we will assume that there exist functions $\tilde{F}(s)$, $g_0(s), \ldots, g_J(s)$ that are analytic in a region that contains the half plane $\Re(s) \ge \sigma_a$ such that

$$\widetilde{A}(s) = g_0(s)\frac{\left(\log \frac{1}{s-\sigma_a}\right)^{\beta_0}}{(s-\sigma_a)^{\alpha_0}} + \sum_{j=1}^{J} g_j(s)\frac{\left(\log \frac{1}{s-\sigma_a}\right)^{\beta_j}}{(s-\sigma_a)^{\alpha_j}} + \tilde{F}(s), \tag{11}$$

where $g_0(\sigma_a) \ne 0$, β_j are non-negative integers, α_0 is real, and $\alpha_1, \ldots, \alpha_J$ are complex numbers with $\Re(\alpha_j) < \alpha_0$ $(1 \le j \le J)$, and β_0 is non-negative if α_0 is contained in the set $\{0, -1, -2, \ldots\}$.

As demonstrated in Appendix A, this is certainly the case if a_n is a linear combination of sequences of the form

$$n^\sigma (\log n)^\alpha$$

(or related to such sequences with floor and ceiling functions). For example, if α is not a negative integer, then the corresponding Dirichlet series $\widetilde{A}(s)$ (in (7)) of the sequence $a_n = n^\sigma(\log n)^\alpha$ can be expressed as

$$\widetilde{A}(s) = \sigma\frac{\Gamma(\alpha+1)}{(s-\sigma)^{\alpha+1}} + \frac{\Gamma(\alpha+1)}{(s-\sigma)^{\alpha}} + \tilde{F}(s),$$

where $\tilde{F}(s)$ is analytic for $\Re(s) > \sigma - 1$, see Theorem 8 of Appendix A. Therefore,

$$\sigma_a = \sigma \quad \text{and} \quad \alpha_0 = \alpha + 1.$$

We will discuss several examples in Section 2.3.

If $s_0 = \sigma_a$ or if $s_0 > \sigma_a$, then the zeros of the equation (10) influence the analysis. It turns out we need to consider two different scenarios depending on a certain property of $p_1, \ldots, p_m$.

Definition 1. *We say that* $\log(1/p_1), \ldots, \log(1/p_m)$ *are rationally related if there exists a positive real number* L *such that* $\log(1/p_1), \ldots, \log(1/p_m)$ *are integer multiples of* L*, that is,* $\log(1/p_j) = n_j L$, $n_j \in \mathbb{Z}$, $(1 \le j \le m)$. *Without loss of generality we can assume that* L *is as large as possible which is equivalent to* $\gcd(n_1, \ldots, n_m) = 1$. *Similarly, we say that* $\log(1/p_1), \ldots, \log(1/p_m)$ *are irrationally related if they are not rationally related.*

Example. If $m = 1$, then we are always in the rationally related case. In the binary case $m = 2$, the numbers $\log(1/p_1)$, $\log(1/p_2)$ are rationally related if and only if the ratio $\log(1/p_1)/\log(1/p_2)$ is rational.

The following property of the roots of (10) is due to Schachinger [28] (cf. also [10, 16]).

Lemma 1. *Let* s_0 *be the unique real solution of equation (10). Then all other solutions* s' *of (10) satisfy* $\Re(s') \le s_0$.
(i) *If* $\log(1/p_1), \ldots, \log(1/p_m)$ *are irrationally related, then* s_0 *is the only solution of (10) on* $\Re(s) = s_0$.
(ii) *If* $\log(1/p_1), \ldots, \log(1/p_m)$ *are rationally related, then there are infinitely many solutions* s_k, $k \in \mathbb{Z}$, *with* $\Re(s_k) = s_0$ *which are given by*

$$s_k = s_0 + k\frac{2\pi i}{L} \qquad (k \in \mathbb{Z}),$$

where $L > 0$ *is the largest real number such that* $\log(1/p_j)$ *are all integer multiples of* L*. Furthermore, there exists* $\delta > 0$ *such that all remaining solutions of (10) satisfy* $\Re(s) \le s_0 - \delta$.

2.2 Discrete Master Theorem

We are now ready to formulate our main results regarding the asymptotic solutions of discrete divide and conquer recurrences. Note that the irrational case is easier to handle whereas the rational case needs additional assumptions on the Dirichlet series. Nevertheless these assumptions are usually easy to establish in practice.

As discussed, our Discrete Master Theorem shows that for sequences a_n of practical importance such as

$$a_n = n^{\sigma}(\log n)^{\alpha}$$

the solution $T(n)$ of the divide and conquer recurrence grows as

$$T(n) \sim C\, n^{\sigma'}(\log n)^{\alpha'}(\log\log n)^{\beta'} \tag{12}$$

(with $\sigma' = \max\{\sigma, s_0\}$) when $\log p_1, \ldots \log p_m$ are irrationally related. For rationally related $\log p_1, \ldots \log p_m$, it is either of the form (12) or (if $s_0 > \sigma$) there appears an oscillation in the form of

$$T(n) \sim \Psi(\log n)\, n^{s_0} \tag{13}$$

with a discontinuous periodic function $\Psi(x)$; see Figure 3 of Section 2.3.

More precisely, in Section 4 we prove the following result.

Theorem 1 (Discrete Master Theorem). *Let $T(n)$ be the divide and conquer recurrence defined in (5), where b_j and $\overline{b}_j$ are non-negative with $b_j+\overline{b}_j>0$, $h_j(x)$ and $\overline{h}_j(x)$ are increasing and non-negative functions with $h_j(2)<2$, $\overline{h}_j(2)\le 1$, and with $h_j(x)=p_jx+O(x^{1-\delta})$ and $\overline{h}_j(x)=p_jx+O(x^{1-\delta})$ for positive numbers $p_j<1$ and some $\delta>0$. Furthermore assume that the sequence $(a_n)_{n\ge 2}$ is non-negative and non-decreasing. Let σ_a denote the abscissa of absolute convergence of the Dirichlet series $\widetilde{A}(s)$ and s_0 the real root of (10). If $\sigma_a\ge s_0\ge 0$ assume further that $\widetilde{A}(s)$ has a representation of the form (11), where $\tilde{F}(s)$, $g_0(s),\ldots,g_J(s)$ are analytic in a region that contains the half plane $\Re(s)\ge\sigma_a$, $g_0(\sigma_a)\ne 0$, α_0 is real and $\Re(\alpha_j)<\alpha_0$ $(1\le j\le J)$, β_j are non-negative integers such that $\beta_0>0$ if α_0 is not contained in the set $\{0,-1,-2,\ldots\}$.*

(i) If $\log(1/p_1),\ldots,\log(1/p_m)$ are irrationally related and if α_0 is not contained in the set $\{0,-1,-2,\ldots\}$, then as $n\to\infty$

$$T(n)=\begin{cases} C_1+o(1) & \text{if } \sigma_a<0 \text{ and } s_0<0,\\ C_2\log n+C_2'+o(1) & \text{if } \sigma_a<s_0 \text{ and } s_0=0,\\ C_3(\log n)^{\alpha_0+1}(\log\log n)^{\beta_0}\cdot(1+o(1)) & \text{if } \sigma_a=s_0=0,\\ C_4\,n^{s_0}\cdot(1+o(1)) & \text{if } \sigma_a<s_0 \text{ and } s_0>0,\\ C_5n^{s_0}(\log n)^{\alpha_0}(\log\log n)^{\beta_0}\cdot(1+o(1)) & \text{if } \sigma_a=s_0>0,\\ C_6\,(\log n)^{\alpha_0}(\log\log n)^{\beta_0}(1+o(1)) & \text{if } \sigma_a=0 \text{ and } s_0<0,\\ C_7\,n^{\sigma_a}(\log n)^{\alpha_0-1}(\log\log n)^{\beta_0}\cdot(1+o(1)) & \text{if } \sigma_a>s_0 \text{ and } \sigma_a>0,\end{cases}\tag{14}$$

where the explicitly computable constants $C_1,C_2,C_3,C_4,C_5,C_6,C_7$ are positive and C_2' is real. Furthermore if α_0 is contained the set $\{0,-1,-2,\ldots\}$ (and if $\beta_0>0$) then we have to replace the factor $(\log\log n)^{\beta_0}$ by $(\log\log n)^{\beta_0-1}$ in (14) if

if $\sigma_a=s_0=0$ and $\alpha_0\le -2$,
if $\sigma_a=s_0>0$ and $\alpha_0\le -1$,
if $\sigma_a=0$, $s_0<0$, and $\alpha_0\le -1$, and
if $\sigma_a>s_0$ and $\sigma_a>0$.

In all other cases there is no change in (14).

(ii) *If $\log(1/p_1),\ldots,\log(1/p_m)$ are rationally related and if in the case $s_0=\sigma_a$ the Fourier series*

$$\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{\widetilde{A}(s_0+2\pi ik/L)}{s_0+2\pi ik/L}e^{2\pi ikx/L}\tag{15}$$

is convergent for $x\in\mathbb{R}$ and represents an integrable function, then $T(n)$ behaves as in the irrationally related case with the following two exceptions:

$$T(n)=\begin{cases} C_2\log n+\Psi_2(\log n)+o(1) & \text{if } \sigma_a<s_0 \text{ and } s_0=0,\\ \Psi_4(\log n)\,n^{s_0}\cdot(1+o(1)) & \text{if } \sigma_a<s_0 \text{ and } s_0>0,\end{cases}\tag{16}$$

where C_2 is positive and $\Psi_2(t),\Psi_4(t)$ are periodic functions with period L.

Remark 1. We should point out that the periodic functions $\Psi_2(t)$ and $\Psi_4(t)$ that appear in the second part of Theorem 1 have (usually) countably many discontinuities and, thus, have no absolutely convergent Fourier series. This makes the analysis actually more challenging. We will show in Section 4.4 that $\Psi(t)$ has building blocks of the form

$$\lambda^{-t} \sum_{n\geq 1} B_n \frac{\lambda^{\lfloor t-\frac{\log n}{L} \rfloor+1}}{\lambda - 1}$$

for some $\lambda > 1$ and a sequence B_n such that the series $\sum_{n\geq 1} B_n \lambda^{-(\log n)/L}$ is absolutely convergent.[1] This function is periodic (with period L) and of bounded variation. Consequently, it has a convergent Fourier series but it is discontinuous for $t = \{\log n/L\}$, $n \geq 1$, where, as usual $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of a real number x.

Remark 2. The condition (15) for $\widetilde{A}(s)$ looks artificial. However, it is really needed in the proof in order to control the polar singularities of $\widetilde{T}(s)$ at s_k, $k \in \mathbb{Z} \setminus \{0\}$. Nevertheless it is no real restriction in practice. As shown in Appendix A the condition (15) is satisfied for sequences of the form $a_n = n^\sigma (\log n)^\alpha$.

Remark 3. By linearity the superposition principle applies. Hence we can combine Theorem 1 and the Akra-Bazzi theorem [2] to recurrences (5), where

$$a_n = c\, n^\sigma (\log n)^\alpha + O\left(n^{\sigma_1} (\log n)^{\alpha_1}\right)$$

where $\sigma_1 < \sigma$ or $\sigma_1 = \sigma$ and $\alpha_1 < \alpha$. Let $T_0(n)$ be the solution of (5) when $a_n 0 = n^\sigma (\log n)^\alpha$. Then

$$T(n) \sim T_0(n).$$

for large n.

2.3 Applications

We first illustrate our theorem on a few simple divide and conquer recurrences before in the next subsection presenting a detailed analysis of Boncelet's algorithm. Several of these examples are also discussed in [25], where the growth order of $T(n)$ is determined.
Example 1. Consider the recurrence

$$T(n) = 2\,T(\lfloor n/2 \rfloor) + 3\,T(\lfloor n/6 \rfloor) + n \log n.$$

The Dirichlet series $\widetilde{A}(s) = \sum (a_{n+2} - a_{n+1}) n^{-s}$ corresponding to the sequence $a_n = n \log n$ has $\sigma_a = 1$ as the abscissa of absolute convergence. Furthermore the equation

$$2 \cdot 2^{-s} + 3 \cdot 6^{-s} = 1$$

has the (real) solution $s_0 = 1.402\ldots > 1$. It is also easy to check that $\log(1/2)$ and $\log(1/6)$ are irrationally related. Namely, if $\log(1/2)/\log(1/6)$ were rational, say a/b then it would follow that $2^b = 6^a$. However, this equation has no non-zero integer solution. Hence by (14) Case 4, we obtain

$$T(n) \sim C n^{s_0} \qquad (n \to \infty)$$

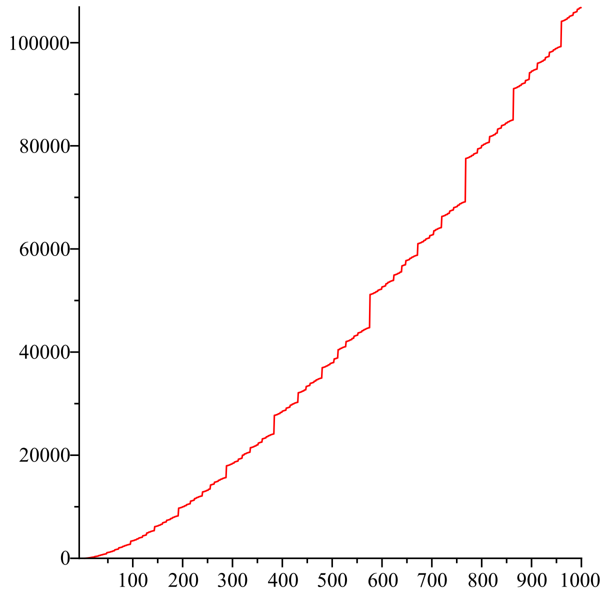

Figure 2: $T(n)$ versus n from Example 1.

for some constant $C > 0$ as shown in Figure 2.

Example 2. Next consider the recurrence

$$T(n) = 2\,T(\lfloor n/2 \rfloor) + \frac{8}{9}T(\lfloor 3n/4 \rfloor) + \frac{n2}{\log n}.$$

Here $\sigma_a = s_0 = 2$ and we are (again) in the irrationally related case. Now Theorem 8 implies that $\widetilde{A}(s)$ has a singular representation of the form

$$\widetilde{A}(s) = s \log \frac{1}{s-2} + G(s)$$

for some function $G(s)$ that is analytic for $\Re(s) > 1$. Consequently, by (14) Case 6 we have

$$T(n) \sim Cn2 \log \log n \qquad (n \to \infty)$$

for some constant $C > 0$.

Example 3. Consider now

$$T(n) = T(\lfloor n/2 \rfloor) + \log n.$$

Here we have $\sigma_a = s_0 = 0$. Since $m = 1$ we are also in the rational case. By Theorem 8 of Appendix A $\widetilde{A}(s)$ has a singular representation of the form

$$\widetilde{A}(s) = \frac{1}{s} + G(s)$$

[1]One can derive a complicated explicit formula for $\Psi(t)$ but it does not provide any new insight.

with some $G(s)$ that is analytic for $\Re(s) > -1$. Recall that Theorem 9 assures that condition (15) is satisfied. Hence, by (14) case 3 we obtain

$$T(n) \sim C(\log n)2 \qquad (n \to \infty)$$

for some constant $C > 0$.

Example 4. The recurrence

$$T(n) = \frac{1}{2}T(\lfloor n/2 \rfloor) + \frac{1}{n}$$

is not covered by Theorem 1 since a_n is decreasing. Hence, $T(n)$ is not increasing, either. However, we can apply the proof methods of Theorem 1.[2] Formally we have $\sigma_a = s_0 = -1 < 0$ and, since $m = 1$, we are in the rationally related case. It follows that

$$T_n = C\frac{\log n}{n} + \frac{\Psi(\log n)}{n} + o\left(\frac{1}{n}\right)$$

for some constant $C > 0$ and a periodic function $\Psi(t)$.

Example 5. The recurrence

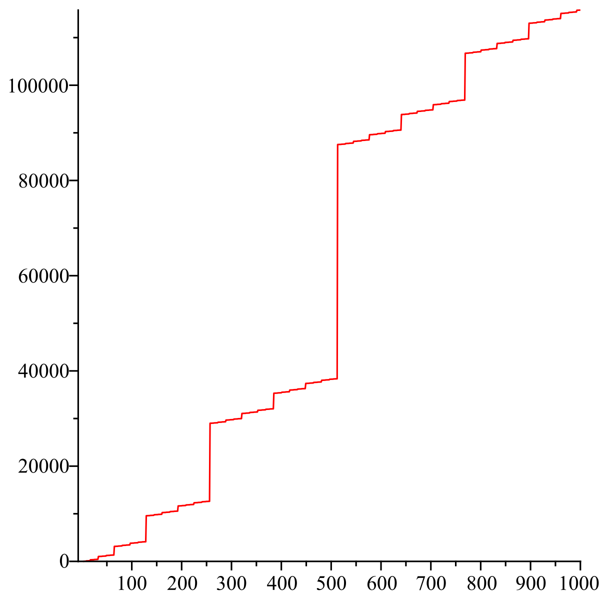

Figure 3: $T(n)$ versus n from Example 5.

$$T(n) = 3T(\lceil n/2 \rceil) + n$$

is related to the Karatsuba algorithm [21, 22]. Here we have $s_0 = (\log 3)/(\log 2) = 1.5849\ldots$ and $s_0 > \sigma_a = 1$. Furthermore, since $m = 1$, we are in the rationally related case. Hence, we have

$$T(n) = \Psi(\log n)\, n^{\frac{\log 3}{\log 2}} \cdot (1 + o(1)) \qquad (n \to \infty)$$

[2]Actually the formal calculations are the same.

with some periodic function $\Psi(t)$, as shown in Figure 3.

In a similar manner, the Strassen algorithm for matrix multiplications results in the following recurrence

$$T(n) = 7T(\lceil n/2 \rceil) + n2.$$

Again, here we have $m = 1$, $s_0 = \log 7/\log 2 \approx 2.81$ and $\sigma_a = 2$, thus

$$T(n) = \Psi(\log n)\, n^{\frac{\log 7}{\log 2}} \cdot (1 + o(1)) \qquad (n \to \infty)$$

with some periodic function $\Psi(t)$.

Example 6. The recurrences

$$\begin{aligned}
T(n) &= T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + n - 1,\\
Y(n) &= Y(\lfloor n/2 \rfloor) + Y(\lceil n/2 \rceil) + \lfloor n/2 \rfloor,\\
U(n) &= U(\lfloor n/2 \rfloor) + U(\lceil n/2 \rceil) + n - \frac{\lfloor n/2 \rfloor}{\lceil n/2 \rceil + 1} + \frac{\lceil n/2 \rceil}{\lfloor n/2 \rfloor + 1} \lfloor n/2 \rfloor
\end{aligned}$$

are related to Mergesort (see [13]). For all three recurrences we have $\sigma_a = s_0 = 1$ and we are (again) in the rationally related case. Hence, we obtain asymptotic expansions of the form

$$C\, n \log n + n\Psi(\log n) + o(n) \qquad (n \to \infty),$$

where $C = 1/\log 2$ for $T(n)$ and $U(n)$ and $C = 1/(2\log 2)$ for $Y(n)$, and $\Psi(t)$ is a periodic function.

3 Boncelet's Arithmetic Coding Algorithm

We present a novel application of our analytic approach to discrete divide and conquer recurrences. We compute the redundancy of a new and practical variable-to-fixed compression algorithm due to Boncelet [4]. To recall, a variable-to-fixed length encoder partitions the source string, say over an m-ary alphabet $\mathcal{A}$, into a concatenation of variable-length phrases. Each phrase belongs to a given dictionary of source strings. We represent a uniquely parsable dictionary by a *complete parsing tree*, i.e., a tree in which every internal node has all m children nodes. The dictionary entries correspond to the *leaves* of the associated parsing tree. The encoder represents each parsed string by the fixed length binary code word corresponding to its dictionary entry. There are several well known variable-to-fixed algorithms; e.g., Tunstall and Khodak schemes (cf. [10, 20, 31]). Boncelet's algorithm, described next, is a practical and computationally fast algorithm that becomes more and more popular. Therefore, we compare its redundancy to the (asymptotically) optimal Tunstall's algorithm.

Boncelet describes his algorithm in terms of a parsing tree. For fixed n (representing the number of leaves in the parsing tree and hence also the number of distinct phrases), the algorithm in each step creates two subtrees of predetermined number of leaves (phrases). Thus at the root, n is split into two subtrees with the number of leaves, respectively, equal to $n_1 = \left\lfloor p_1 n + \frac{1}{2} \right\rfloor$ and $n_2 = \left\lfloor p_2 n + \frac{1}{2} \right\rfloor$. This continues recursively until only 1 or 2 leaves are left. Note that this splitting procedure does not assure that $n_1 + n_2 = n$. For example if $p_1 = \frac{3}{8}$ and $p_2 = \frac{5}{8}$, then $n = 4$ would be split into $n_1 = 2$ and $n_2 = 3$. Therefore, we propose to modify the splitting as follows $n_1 = \lfloor p_1 n + \delta \rfloor$ and $n_2 = \lceil p_2 n - \delta \rceil$ for some $\delta \in (0, 1)$ that satisfies $2p_1 + \delta < 2$.

Let $\{v_1, \dots v_n\}$ denote phrases of the Boncelet code that correspond to the paths from the root to leaves of the parsing tree, and let $\ell(v_1), \dots, \ell(v_n)$ be the corresponding phrase lengths. Observe that while the parsing tree in the Boncelet's algorithm is fixed, a randomly generated sequence is partitioned into random length phrases. Therefore, one can talk about the probabilities of phrases denoted as $P(v_1), \dots, P(v_n)$. Here we restrict the analysis to a binary alphabet and denote the probabilities by $p := p_1$ and $q := p_2 = 1 - p$.

For sequences generated by a binary memoryless source, we aim at understanding the probabilistic behavior of the phrase length that we denote as D_n. Its probability generating function is defined as

$$C(n, y) = \mathbb{E}\, y^{D_n}$$

which can also be represented as

$$C(n, y) = \sum_{j=1}^{n} P(v_j) y^{\ell(v_j)}.$$

The Boncelet's splitting procedure leads to the following recurrence on $C(n, y)$ for $n \geq 2$

$$C(n, y) = p\, y\, C\left(\lfloor pn + \delta \rfloor, y\right) + q\, y\, C\left(\lceil qn - \delta \rceil, y\right) \tag{17}$$

with initial conditions $C(0, y) = 0$ and $C(1, y) = 1$.

Next let $d(n)$ denote the average phrase length

$$\mathbb{E}\, D_n := d(n) = \sum_{j=1}^{n} P(v_j)\, \ell(v_j)$$

which is also given by $d(n) = C'(n, 1)$ (where the derivative is taken with respect to y) and satisfies the recurrence

$$d(n) = 1 + p_1 d\left(\lfloor p_1 n + \delta \rfloor\right) + p_2 d\left(\lceil p_2 n - \delta \rceil\right) \tag{18}$$

with $d(0) = d(1) = 0$. This recurrence falls exactly under our general divide and conquer recurrence, hence Theorem 1 applies.

Theorem 2. *Consider a binary memoryless source with positive probabilities $p_1 = p$ and $p_2 = q$ and the entropy rate $H = p \log(1/p) + q \log(1/q)$. Let $d(n) = \mathbb{E}\, D_n$ denote the expected phrase length of the binary Boncelet code.*

(i) *If the ratio $(\log p)/(\log q)$ is irrational, then*

$$d(n) = \frac{1}{H} \log n - \frac{\alpha}{H} + o(1), \tag{19}$$

where

$$\alpha = \overline{E}'(0) - \overline{G}'(0) - H - \frac{H_2}{2H}, \tag{20}$$

$H_2 = p \log 2p + q \log 2q$, and $\overline{E}'(0)$ and $\overline{G}'(0)$ are the derivatives at $s = 0$ of the Dirichlet series defined in (27) of Section 4.

(ii) *If $(\log p)/(\log q)$ is rational, then*

$$d(n) = \frac{1}{H} \log n - \frac{\alpha + \Psi(\log n)}{H} + O(n^{-\eta}), \tag{21}$$

where $\Psi(t)$ is a periodic function and $\eta > 0$.

For practical data compression algorithms, it is important to achieve low redundancy defined as the excess of the code length over the optimal code length nH. For variable-to-fixed codes, the average redundancy is expressed as [10, 27]

$$R_n = \frac{\log n}{\mathbb{E}\, D_n} - H = \frac{\log n}{d(n)} - H$$

since every phrase of average length $d(n)$ requires $\log n$ bits to point to a dictionary entry. Our previous results imply immediately the following corollary.

Corollary 1. *Let R_n denote the redundancy of the binary Boncelet code with positive probabilities $p_1 = p$ and $p_2 = q$.*
(i) *If the ratio $(\log p)/(\log q)$ is irrational, then*

$$R_n = \frac{H\alpha}{\log n} + o\left(\frac{1}{\log n}\right). \tag{22}$$

with α defined in (20).
(ii) *If $(\log p)/(\log q)$ is rational, then*

$$R_n = \frac{H(\alpha + \tilde{\Psi}(\log n))}{\log n} + o\left(\frac{1}{\log n}\right). \tag{23}$$

where $\tilde{\Psi}(t)$ is a periodic function.

We should compare the redundancy of Boncelet's algorithm to asymptotically optimal Tunstall algorithm. From [10, 27] we know that the redundancy of the Tunstall code is

$$R_n^T = \frac{H}{\log n}\left(-\log H - \frac{H_2}{2H}\right) + o\left(\frac{1}{\log n}\right)$$

(provided that $(\log p)/(\log q)$ is irrational; in the rational case there is also a periodic term in the leading asymptotics). This should be compared to the redundancy of the Boncelet algorithm.

Example. Consider $p = 1/3$ and $q = 2/3$. Then one computes

$$\begin{aligned}
\alpha = &\sum_{m\ge 1} \frac{d(m+2) - d(m+1)}{3}\left(\log\left\lceil 3m + \frac{5}{2}\right\rceil - \log(3m)\right) \\
&+ 2\sum_{m\ge 1} \frac{d(m+2) - d(m+1)}{3}\left(\log\left\lfloor \frac{3}{2}m + \frac{5}{4}\right\rfloor - \log(\frac{3m}{2})\right) \\
&+ \frac{\log 3}{3} - H - \frac{H_2}{2H} \approx 0.0518
\end{aligned}$$

while for the Tunstall code $-\log H - \frac{H_2}{2H} \approx 0.0496$.

We also have to observe that the leading constant of the $\log n$-term equals $1/H$. This follows analytically from the initial conditions $d(0) = d(1) = 0$ and the fact that the function $1 - p^{s+1} - q^{s+1}$ has derivative H at $s = 0$. On the other hand, if we use as an a-priori information that $d(n) \sim c \log n$ then by comparing the asymptotic expansion on the left and right hand side of (18) it also follows that $c = 1/H$.

Finally, we deal with the limiting distribution of the phrase length D_n. The proof is presented in Section 5.

Theorem 3. *Consider a biased memoryless source (i.e., $p \neq q$) generating a sequence of length n parsed by the Boncelet algorithm. The phrase length D_n satisfies the central limit law, that is,*

$$\frac{D_n - \frac{1}{H}\log n}{\sqrt{\left(\frac{H_2}{H3} - \frac{1}{H}\right)\log n}} \to N(0,1),$$

where $N(0,1)$ denotes the standard normal distribution, and

$$\mathbb{E}\, D_n = \frac{\log n}{H} + O(1), \qquad \mathbb{V}\mathrm{ar}\, D_n \sim \left(\frac{H_2}{H3} - \frac{1}{H}\right)\log n$$

for $n \to \infty$.

We observe that the phrase length D_n follows the same central limit law as the Tunstall algorithm [10].

4 Analysis and Asymptotics

We prove here a general asymptotic solution of the divide and conquer recurrence (cf. Theorem 1). We first derive the appropriate Dirichlet series and apply Tauberian theorem for the irrationally related case, then discuss the Perron-Mellin formula, and finally finish the proof of Theorem 1 for the rationally related case.

4.1 Dirichlet Series

As discussed in the previous section, the proof makes use of the Dirichlet series

$$\widetilde{T}(s) = \sum_{n=1}^{\infty} \frac{T(n+2) - T(n+1)}{n^s},$$

where we apply Tauberian theorems and the Mellin-Perron formula to obtain asymptotics for $T(n)$ from a singularity analysis of $\widetilde{T}(s)$.

By partial summation and using a-priori upper bounds for the sequence $T(n)$, it follows that $\widetilde{T}(s)$ converges (absolutely) for $s \in \mathbb{C}$ with $\Re(s) > \max\{s_0, \sigma_a, 0\}$, where s_0 is the real solution of the equation (10), and σ_a is the abscissa of absolute convergence of $\tilde{A}(s)$.

Next we apply the recurrence relation (5) to $\widetilde{T}(s)$. To simplify our presentation, we assume that $\overline{b}_j = 0$, that is, we consider only the floor function on the right hand side of the recurrence (5); those parts that contain the ceiling function can be handled in the same way. We thus obtain

$$\widetilde{T}(s) = \widetilde{A}(s) + \sum_{j=1}^{m} b_j \sum_{n=1}^{\infty} \frac{T\left(\lfloor h_j(n+2)\rfloor\right) - T\left(\lfloor h_j(n+1)\rfloor\right)}{n^s}.$$

For $k \geq 1$ set

$$n_j(k) := \max\{n \geq 1 : h_j(n+1) < k+2\}.$$

By definition it is clear that $n_j(k+1) \geq n_j(k)$ and

$$n_j(k) = \frac{n}{p_j} + O\left(k^{1-\delta}\right). \tag{24}$$

Furthermore by setting

$$G_j(s) = \sum_{n\ge 1, h_j(n+2)<3} \frac{T\left(\lfloor h_j(n+2)\rfloor\right) - T\left(\lfloor h_j(n+1)\rfloor\right)}{n^s}$$

we obtain

$$\sum_{n=1}^{\infty} \frac{T\left(\lfloor h_j(n+2)\rfloor\right) - T\left(\lfloor h_j(n+1)\rfloor\right)}{n^s} = G_j(s) + \sum_{k=1}^{\infty} \frac{T(k+2)-T(k+1)}{n_j(k)^s}.$$

We now compare the last sum to $p_j^s \widetilde{T}(s)$:

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{T(k+2)-T(k+1)}{n_j(k)^s} &= \sum_{k=1}^{\infty} \frac{T(k+2)-T(k+1)}{(k/p_j)^s} \\ &\quad - \sum_{k=1}^{\infty} (T(k+2)-T(k+1)) \left(\frac{1}{(k/p_j)^s} - \frac{1}{n_j(k)^s} \right) \\ &= p_j^s \widetilde{T}(s) - E_j(s),\end{aligned}$$

where

$$E_j(s) = \sum_{k=1}^{\infty} (T(k+2)-T(k+1)) \left(\frac{1}{(k/p_j)^s} - \frac{1}{n_j(k)^s} \right). \tag{25}$$

Defining

$$E(s) = \sum_{j=1}^{m} b_j E_j(s) \quad \text{and} \quad G(s) = \sum_{j=1}^{m} b_j G_j(s)$$

we finally obtain the relation

$$\widetilde{T}(s) = \frac{\widetilde{A}(s) + G(s) - E(s)}{1 - \sum_{j=1}^{m} b_j\, p_j^s}. \tag{26}$$

As mentioned above, (almost) the same procedure applies if some of the $\overline{b}_j$ are positive, that is, the ceiling function also appear in the recurrence equation. The only difference to (26) is that we arrive at a representation of the form

$$\widetilde{T}(s) = \frac{\widetilde{A}(s) + \overline{G}(s) - \overline{E}(s)}{1 - \sum_{j=1}^{m} (b_j + \overline{b}_j)\, p_j^s}, \tag{27}$$

with a slightly modified functions $\overline{G}(s)$ and $\overline{E}(s)$, however, they have the same analyticity properties as in (26).

By our previous assumptions, we know the analytic behaviors of $\widetilde{A}(s)$ and $\left(1 - \sum_{j=1}^{m} (b_j + \overline{b}_j)\, p_j^s\right)^{-1}$: $\widetilde{A}(s)$ has a pole-like singularity at $s = \sigma_a$ (if $\sigma_a \ge s_0$) and a proper continuation to a complex domain that contains the (punctuated) line $\Re(s) = \sigma_a$, $s \ne \sigma_a$. On the other hand,

$$\frac{1}{1 - \sum_{j=1}^{m} (b_j + \overline{b}_j)\, p_j^s}$$

has a polar singularity at $s = s_0$ (and infinitely many other poles on the line $\Re(s) = s_0$ if the numbers $\log(1/p_j)$ are rationally related), and also a meromorphic continuation to a complex domain that contains the line $\Re(s) = s_0$. Furthermore, $G(s)$ is an entire function. It suffices to discuss $E_j(s)$. First observe (24) implies

$$\frac{1}{(k/p_j)^s} - \frac{1}{n_j(k)^s} = O\left(\frac{1}{(k/p_j)^{\Re(s)+\delta}}\right).$$

By partial summation (and by using again the a-priori estimates), it follows immediately that the series

$$\sum_{k=1}^{\infty} (T(k+2) - T(k+1)) \frac{1}{(k/p_j)^{\Re(s)+\delta}}$$

converges for $\Re(s) > \max\{s_0, \sigma_a, 0\} - \delta$. Since $T(n)$ is an increasing sequence, this implies (absolute) convergence of the series $E_j(s)$, just representing an analytic function in this region, too.

In order to recover (asymptotically) $T(n)$ from $\widetilde{T}(s)$ we apply several different techniques discussed in the next subsection. The main analytic tools are Tauberian theorems (of Wiener-Ikehara which discussed in detail in Appendix B) and the Mellin-Perron formula (Theorem 4).

4.2 Tauberian Theorems

We are now ready to prove several parts of Theorem 1 with the help of Tauberian theorems of Wiener-Ikehara type (see Appendix B). We recall that such theorems apply in general to the so-called Mellin-Stieltjes transform

$$\int_{1-}^{\infty} v^{-s}\, d\overline{c}(v) = s \int_{1}^{\infty} \overline{c}(v) v^{-s-1}\, dv$$

of a non-negative and non-decreasing function $\overline{c}(v)$. If $c(n)$ is a sequence of non-negative numbers, then the Dirichlet series $C(s) = \sum_{n\ge 1} c(n) n^{-s}$ is just the Mellin-Stieltjes transform of the function $\overline{c}(v) = \sum_{n\le v} c(n)$:

$$C(s) = \sum_{n\ge 1} c(n) n^{-s} = \int_{1-}^{\infty} v^{-s}\, d\overline{c}(v) = s \int_{1}^{\infty} \overline{c}(v) v^{-s-1}\, dv.$$

Informally, a Tauberian theorem is a correspondence between the singular behaviour of $\frac{1}{s}C(s)$ and the asymptotic behaviour of $\overline{c}(v)$. In the context of Tauberian theorems of Wiener-Ikehara type one assumes that $C(s)$ continues analytically to a proper region, has only one (real) singularity s_0 on the *critical line* $\Re(s) = s_0$, and the singularity is of special type (for example a polar or algebraic singularity, see Appendix B).

We recall that $\widetilde{T}(s)$ is the Dirichlet series of the sequence $c(n) = T(n+2) - T(n+1)$. Hence

$$T(n) = \overline{c}(n-2) + T(2).$$

Consequently, if we know the asymptotic behaviour of $\overline{c}(v)$ we also find that of $T(n)$ (which is more or less the same).

We recall that $\widetilde{T}(s)$ is given by (27). Hence the dominant singularity of $\frac{1}{s}\widetilde{T}(s)$ is either zero, or induced by the singular behaviour of $\widetilde{A}(s)$, or induced by the zeros of the denominator

$$1 - \sum_{j=1}^{m} (b_j + \overline{b}_j) p_j^s.$$

Here it is essential to assume that the $\log p_j$ are *irrationally related*. Precisely in this case the denominator has only the real zero s_0 on the line $\Re(s) = s_0$. Hence Tauberian theorems can be applied in the irrationally related case if $s_0 \geq \sigma_a$. (For the rational case we will apply a different approach to cover the case $s_0 \geq \sigma_a$.)

Our conclusions for the proof of the first part of Theorem 1 are summarized as follows:

1. $\sigma_a < 0$ and $s_0 < 0$:
   This is indeed a trivial case, since the dominant singularity is at $s = 0$ and the series $\widetilde{T}(s)$ converges for $s = 0$:
   $$\widetilde{T}(0) = \sum_{n\geq 1} (T(n+2) - T(n+1)),$$
   hence
   $$T(n) = C_1 + o(1),$$
   where $C_1 = T(2) + \widetilde{T}(0)$.

2. $\sigma_a < s_0$ and $s_0 = 0$:
   We can apply directly a proper version of the Wiener-Ikehara theorem (Theorem 11 of Appendix B) that proves
   $$T(n) = C_2 \log n \cdot (1 + o(1)).$$
   Observe, that $s = 0$ is a double pole of $\frac{1}{s}\widetilde{T}(s)$ that induces the $\log n$-term in the asymptotic expansion. Note that this does not prove the full version that is stated in Theorem 1. By applying Theorem 5 (that is based on a more refined analysis) we also arrive at an asymptotic expansion of the form
   $$T(n) = C_2 \log n + C_2' + o(1).$$

3. $\sigma_a = s_0 = 0$:
   In this case the dominant singular term of $\frac{1}{s}\widetilde{T}(s)$ is given by
   $$C\, \frac{(\log(1/s))^{\beta_0}}{s^{\alpha_0+2}} \quad \text{with} \quad C = \frac{-g_0(0)}{\sum_{j=1}^m (b_j + \overline{b}_j) \log p_j}$$
   Hence, an application of Theorem 12 of Appendix B provides the asymptotic leading term for $T(n)$. Recall that we have to handle separately the case when α_0 is contained in the set $\{-2, -3, \ldots\}$ (and $\beta_0 > 0$). In this case, only logarithmic singularities remain.

4. $\sigma_a < s_0$ and $s_0 > 0$:
   Here the classical version of the Wiener-Ikehara theorem (Theorem 10 of Appendix B) applies. Note again that it is crucial that the denominator has only one pole on the line $\Re(s) = s_0$.

5. $\sigma_a = s_0 > 0$:
   Here the function $\frac{1}{s}\widetilde{T}(s)$ has the dominant singular term
   $$C\, \frac{(\log(1/(s-\sigma_a)))^{\beta_0}}{(s-\sigma_a)^{\alpha_0+1}}$$
   for some constant $C > 0$ (and there are no other singularities on the line $\Re(s) = s_0$). Thus, an application of Theorem 12 of Appendix B provides the asymptotic leading term for $T(n)$. Observe that we have to handle separately the case when α_0 is contained in the set $\{-1, -2, \ldots\}$ (and $\beta_0 > 0$).

6. $\sigma_a = 0$ and $s_0 < 0$:
   The analysis of this case is very close to the previous one. The dominant singular term of $\frac{1}{s}\widetilde{T}(s)$ is of the form
   $$C\,\frac{(\log(1/s))^{\beta_0}}{s^{\alpha_0+1}}.$$
7. $\sigma_a > s_0$ and $\sigma_a > 0$:
   In this case the singular behavior of $\widetilde{A}(s)$ dominates the asymptotic behavior of $\frac{1}{s}\widetilde{T}(s)$. An application of Theorem 12 of Appendix B provides the asymptotic leading term of $T(n)$.

4.3 Mellin-Perron Formula

One disadvantage of the use of Tauberian theorems is that they provide (usually) only the asymptotic leading term and no error terms. In order to provide error terms or second order terms one has to use more refined methods. In the framework of Dirichlet series we can apply the Mellin-Perron formula that we recall next (in fact, it follows from Lemma 2 below).

Below we shall use Iverson's notation $[\![P]\!]$ which is 1 if P is a true proposition and 0 else.

Theorem 4 (see [3])**.** *For a sequence $c(n)$ define the Dirichlet series*
$$C(s) = \sum_{n=1}^{\infty} \frac{c(n)}{n^s}$$
and assume that the abscissa of absolute convergence σ_a is finite or $-\infty$. Then for all $\sigma > \sigma_a$ and all $x > 0$
$$\sum_{n<x} c(n) + \frac{c(\lfloor x \rfloor)}{2}[\![x \in \mathbb{Z}]\!] = \lim_{T\to\infty} \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} C(s)\frac{x^s}{s}\,ds.$$

Note that – similarly to the Tauberian theorems – the Mellin-Perron formula enables us to obtain precise information about the function $\overline{c}(v) = \sum_{n\ge v} c(n)$ if we know the behaviour of $\frac{1}{s}C(s)$. In our context we have $c(n) = T(n+2) - T(n)$, that is,
$$T(n) = T(2) + \lim_{T\to\infty} \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \widetilde{T}(s)\frac{(n-\frac{3}{2})^s}{s}\,ds \tag{28}$$
with
$$\widetilde{T}(s) = \sum_{n=1}^{\infty} \frac{T(n+2) - T(n+1)}{n^s}.$$

As a first application we apply the Mellin-Perron formula of Theorem 4 for Dirichlet series of the form
$$C(s) = \sum_{n\ge 1} c(n) n^{-s} = \frac{B(s)}{1 - \sum_{j=1}^m b_j p_j^s}, \tag{29}$$
where we assume that the $\log p_j$ are *not* rationally related and where $B(s)$ is analytic in a region that contains the real zero s_0 of the denominator. This theorem can be also applied to the proof of some parts of Theorem 1; in particular for the (irrationally related) cases

if $\sigma_a < 0$ and $s_0 < 0$,
if $\sigma_a < s_0$ and $s_0 = 0$, and
if $\sigma_a < s_0$ and $s_0 > 0$.

Note that Theorem 5 provides a second order term in the case $\sigma_a < s_0 = 0$, see also Remark 5.

Theorem 5. *Suppose that $0 < p_j < 1$, $1 \le j \le m$, are given such that $\log p_j$, $1 \le j \le m$, are not rationally related and let s_0 denote the real solution of the equation*

$$\sum_{j=1}^{m} b_j p_j^s = 1,$$

where $b_j > 0$, $1 \le j \le m$. Let $C(s) = \sum_{n\ge1} c(n)n^{-s}$ be a Dirichlet series with non-negative coefficients $c(n)$ that has a representation of the form (29), that is,

$$C(s) = \sum_{n\ge1} c(n)n^{-s} = \frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s}$$

where $B(s)$ is an analytic function for $\Re(s) \ge s_0 - \eta$ for some $\eta > 0$ and is bounded in this region. Then

$$\sum_{n\le v} c(n) = \begin{cases} \dfrac{B(0)}{1-\sum_{j=1}^m b_j} + o(1) & \text{if } s_0 < 0, \\ \dfrac{B(0)}{H(0)}\log v + \dfrac{B'(0)+B(0)H_2/H}{H} + o(1) & \text{if } s_0 = 0, \\ \dfrac{B(s_0)}{-\sum_{j=1}^m b_j p_j^{s_0}\log p_j} v^{s_0}\left(1+o(1)\right) & \text{if } s_0 > 0 \end{cases},$$

where $H(s) = -\sum_{j=1}^m b_j p_j^s \log p_j$ with $H = H(0)$, and $H_2(s) = \sum_{j=1}^m b_j p_j^s (\log p_j)2$ with $H_2 = H_2(0)$.

Proof. We will use the Mellin-Perron formula of Theorem 4, however, we cannot use it directly, since there are convergence problems. Namely, if we shift the line of integration $\Re(s) = c > s_0$ to the left (to $\Re(s) = \sigma < s_0$) and collect residues we obtain (with $\mathcal{Z} = \{s \in \mathbb{C} : \sum_{j=1}^m b_j p_j^s = 1\}$)

$$\begin{aligned} \sum_{n\le v} c(n) &= \lim_{T\to\infty} \sum_{s'\in\mathcal{Z},\ \Re(s')<\sigma, |\Im(s')|<T} \mathrm{Res}(C(s)\frac{v^s}{s},\ s=s') \\ &\quad + \frac{1}{2\pi i}\lim_{T\to\infty}\int_{\sigma-iT}^{\sigma+iT} \frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s}\frac{v^s}{s}\,ds \\ &= \lim_{T\to\infty} \sum_{s'\in\mathcal{Z},\ \Re(s')<\sigma, |\Im(s')|<T} \frac{B(s')v^{s'}}{s'H(s')} \\ &\quad + \frac{1}{2\pi i}\lim_{T\to\infty}\int_{\sigma-iT}^{\sigma+iT} \frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s}\frac{v^s}{s}\,ds \end{aligned}$$

provided that the series of residues converges and the limit $T \to \infty$ of the last integral exists. The problem is that neither the series nor the integral above are absolutely convergent since the integrand is only of order $1/s$. We have to introduce the auxiliary function

$$\overline{c}_1(v) = \int_0^v \left(\sum_{n\le w} c(n)\right) dw$$

which is also given by

$$\overline{c}_1(v) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} C(s)\frac{v^{s+1}}{s(s+1)}\,ds = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s}\frac{v^{s+1}}{s(s+1)}\,ds,$$

for $c > s_0$. Note that there is no need to consider the limit $T \to \infty$ in this case since the series and the integral are now absolutely convergent. Hence, the above procedure works without any convergence problem. We shift the line of integration to $\Re(s) = \sigma < \min\{-1, s_0\}$. In order to make the presentation of our analysis slightly easier we additionally assume that the region of analyticity of $B(s)$ is large enough such that all zeros in $\mathcal{Z}$ have real part $> \sigma$. Then we have to consider the sum of residues

$$\sum_{s' \in \mathcal{Z},} \operatorname{Res}\left(C(s)\, \frac{v^{s+1}}{s(s+1)},\ s = s' \right) =$$

$$= \sum_{s' \in \mathcal{Z}} \frac{B(s')}{s'(s'+1)H(s')} v^{s'+1}.$$

For $\sigma < 0$ or $\sigma < -1$ the residues at $s = 0$ and $s = 1$ are respectively

$$\frac{B(0)}{1 - \sum_{j=1}^m b_j} v, \quad -\frac{B(-1)}{1 - \sum_{j=1}^m b_j p_j^{-1}},$$

and the integral is

$$\frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} C(s) \frac{v^{s+1}}{s(s+1)}\, ds = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{B(s)}{1 - \sum_{j=1}^m b_j p_j^s} \frac{v^{s+1}}{s(s+1)}\, ds = O(v^{1+\sigma}).$$

Thus, we obtain

$$\overline{c}_1(v) = \frac{B(s_0)}{(s_0+1)H(s_0)} (1 + Q(\log v)) v^{1+s_0} + O(v^{1+s_0-\eta})$$

for some $\eta > 0$, where

$$Q(x) = \sum_{s' \in \mathcal{Z} \setminus \{s_0\}} \frac{2H(s_0)B(s')}{s'(s'+1)H(s')B(s_0)} e^{-x(s'+1)}.$$

It is easy to show that $Q(x) \to 0$ as $x \to \infty$ (cf. also with [28, Lemma 4] and [30]). Indeed, suppose that $\varepsilon > 0$ is given. Then there exists $S_0 = S_0(\varepsilon) > 0$ such that

$$\sum_{s' \in \mathcal{Z},\, |s'| > S_0} \left| \frac{2B(s')H(s_0)}{s'(s'+1)H(s')B(s_0)} \right| < \frac{\varepsilon}{2}.$$

Further, since $\Re(s') < s_0$ for all $s' \in \mathcal{Z} \setminus \{s_0\}$, and by the assumption of irrationality zeros are not on the critical line $\Re(s) = s_0$ (except the real one), it follows that there exists $x_0 = x_0(\varepsilon) > 0$ with

$$\left| \sum_{s' \in \mathcal{Z} \setminus \{s_0\},\, |s'| \le S_0} \frac{2B(s')H(s_0)}{s'(s'+1)H(s')B(s_0)} e^{-x(s'+1)} \right| < \frac{\varepsilon}{2}$$

for $x \ge x_0$. Hence $|Q(x)| < \varepsilon$ for $x \ge x_0(\varepsilon)$.

Note that we cannot obtain the rate of convergence for $Q(x)$. This means that we just get

$$\overline{c}_1(v) \sim \frac{B(s_0)}{(s_0+1)H(s_0)} \cdot v^{1+s_0}$$

as $v \to \infty$. However, since, $\sum_{n \le v} c(n)$ is monotonely increasing in v (by assumption) it also follows that

$$\sum_{n \le v} c(n) \sim \frac{B(s_0)}{H(s_0)} v^{s_0},$$

compare with the case $s_0 = 0$ that we discuss next.

Now suppose that $s_0 = 0$ which means that $C(s)$ has a double pole as $s = 0$. We can almost use the same analysis as above and obtain the asymptotic expansion

$$\overline{c}_1(v) \sim \frac{B(0)}{H} v \log v + \frac{B'(0) - B(0) + B(0)H_2/H)}{H} v.$$

It is now an easy exercise to derive from this expansion the final result

$$\sum_{n \le v} c(n) = \frac{B(0)}{H} \log v + \frac{B'(0) + B(0)H_2/H}{H} + o(1) \tag{30}$$

in the following way. For simplicity we write $\overline{c}_1(v) = C_1 v \log v + C_2 v + o(v)$. By the assumption

$$|\overline{c}_1(v) - C_1 v \log v + C_2 v| \le \varepsilon v$$

for $v \ge v_0$. Set $v' = \varepsilon^{1/2} v$, then by monotonicity we obtain (for $v \ge v_0$)

$$\begin{aligned}
\sum_{n \le v} c(n) &\le \frac{\overline{c}_1(v+v') - \overline{c}_1(v)}{v'} \le \frac{1}{v'} \left(C_1(v+v') \log(v+v') + C_2(v+v') - \right. \\
&\quad \left. C_1 v \log v - C_2 v\right) + \varepsilon \frac{2v+v'}{v'} \\
&= C_1 \log(v+v') + C_2 + C_1 \frac{v}{v'} \log\left(1 + \frac{v'}{v}\right) + \varepsilon \frac{2v+v'}{v'} \\
&= C_1 \log v + C_2 + C_1 + O\left(\varepsilon^{1/2}\right),
\end{aligned}$$

where the O-constant is an absolute one. In a similar manner, we obtain the corresponding lower bound (for $v \ge v_0 + v_0^{1/2}$). Hence, it follows that

$$\left| \sum_{n \le v} c(n) - C_1 \log v - C_1 - C_2 \right| \le C' \varepsilon^{1/2}$$

for $v \ge v_0 + v_0^{1/2}$. This proves $\sum_{n \le v} c(n) = C_1 \log v + C_1 + C_2 + o(1)$ and consequently (30). □

Remark 5. The advantage of the preceding proof is its flexibility. For example, we can apply the procedure for multiple poles and are able to derive asymptotic expansions of the form

$$\sum_{n \le v} c(n) = \sum_{j=0}^{K} A_j \frac{(\log v)^j}{j!} v^{s_0} + o(v^{s_0}).$$

Furthermore we can derive asymptotic expansions that are uniform in an additional parameter when we have some control on the singularities in terms of the additional parameter. We will use this generalization in the proof of the central limit theorem for the phrase lengths of the Boncelet code (Theorem 3).

In principle it is also possible to obtain bounds for the error terms. However, they depend heavily on Diophantine approximation properties of the vector $(\log p_1, \ldots \log p_m)$, see [16].

4.4 The Rationally Related Case

Unfortunately, the previous method generally is not applicable when there are several poles (or infinitely many poles) on the line $\Re(s) = s_0$. This means that we cannot use the above procedure when the $\log p_j$ are rationally related. The reason is that it does not follow *automatically* that an asymptotic expansion of the form

$$\overline{c}_1(v) = \int_0^v \overline{c}(w)\, dw \sim \Psi_1(\log v) \cdot v^{s_0+1}$$

implies

$$\overline{c}(v) \sim \Psi(\log v) \cdot v^{s_0}$$

for certain periodic functions Ψ and Ψ_1, even if $\overline{c}(v)$ is non-negative and non-decreasing.

Therefore we will apply an alternative approach which is – in some sense – more direct and applies *only* in this case, but it proves a convergence result for $c(v)$ of the form

$$\overline{c}(v) = \sum_{n \le v} c(n) \sim \Psi(\log v)\, v^{s_0}$$

even for a periodic functions $\Psi(t)$ with countably many discontinuities.

Suppose that $\log p_j = -n_j L$ for coprime integers n_j and a real number $L > 0$. Then the equation $1 - \sum_{j=1}^m b_j p_j^s$ with the only real solution s_0 becomes an algebraic equation

$$1 - \sum_{j=1}^m b_j z^{n_j} = 0 \qquad \text{with } z = e^{-Ls}.$$

with a single (dominating) real root $z_0 = e^{-Ls_0}$. We can factor this polynomial as

$$1 - \sum_{j=1}^m b_j z^{n_j} = (1 - e^{Ls_0} z) P(z)$$

and obtain also a partial fraction decomposition of the form

$$\frac{1}{1 - \sum_{j=1}^m b_j z^{n_j}} = \frac{1/P(e^{-Ls_0})}{1 - e^{Ls_0} z} + \frac{Q(z)}{P(z)}.$$

Therefore, it is natural in this context to consider Mellin-Perron integrals of the form

$$\frac{1}{2\pi i} \lim_{T\to\infty} \int_{c-iT}^{c+iT} \frac{B(s)}{1 - e^{-Ls}\lambda} \frac{x^s}{s}\, ds$$

for some complex number $\lambda \neq 0$ and a Dirichlet series $B(s)$. The corresponding result is stated below in Theorem 6

For the proof of Theorem 6 we need the following two lemmas. The first lemma (Lemma 2) is also the basis of the proof of the Mellin-Perron formula (cf. [3, 30]). For the reader's convenience we provide a short proof of Lemma 2.

Lemma 2. *Suppose that a and c are positive real numbers. Then*

$$\begin{aligned}
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} a^s\frac{ds}{s} - 1\right| &\leq \frac{a^c}{\pi T\log a} \qquad (a>1),\\
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} a^s\frac{ds}{s}\right| &\leq \frac{a^c}{\pi T\log(1/a)} \qquad (0<a<1),\\
\left|\frac{1}{2\pi i}\int_{c-iT}^{c+iT} a^s\frac{ds}{s} - \frac{1}{2}\right| &\leq \frac{C}{T} \qquad (a=1).
\end{aligned}$$

Proof. Suppose first that $a > 1$. By considering the contour integral of the function $F(s) = a^s/s$ around the rectangle with vertices $-A-iT, c-iT, c+iT, -A+iT$ and letting $A \to \infty$ one directly obtains the representation

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT} a^s\frac{ds}{s} = \mathrm{Res}(a^s/s; s=0) + \frac{1}{2\pi i}\int_{-\infty}^{c}\frac{a^{x+iT}}{x+iT}\,dx + \frac{1}{2\pi i}\int_{-\infty}^{c}\frac{a^{x-iT}}{x-iT}\,dx.$$

Since

$$\left|\frac{1}{2\pi i}\int_{-\infty}^{c}\frac{a^{x\pm iT}}{x\pm iT}\,dx\right| \leq \frac{a^c}{\pi T\log a}$$

we directly obtain the bound in the case $a > 1$.

The case $0 < a < 1$ can be handled in the same way. And finally, in the case $a = 1$ the integral can be explicitly calculated (and estimated). □

Lemma 3. *Suppose that L is a positive real number, λ a complex number different from 0 and 1, and c a real number with $c > \frac{1}{L}\log|\lambda|$. Then we have for all real numbers $x > 1$*

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{1}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = \frac{\lambda^{\lfloor\frac{\log x}{L}\rfloor+1}-1}{\lambda-1} - \frac{1}{2}\lambda^{\lfloor\frac{\log x}{L}\rfloor}[\![\log x/L\in\mathbb{Z}]\!]. \tag{31}$$

Proof. By assumption we have $|\lambda e^{-Ls}| < 1$. Thus, by using a geometric series expansion we get for all $x > 1$ such that $\log x/L$ is not an integer

$$\begin{aligned}
\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{1}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds &= \sum_{k\geq 0}\lambda^k\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(\frac{x}{e^{Lk}}\right)^s\frac{ds}{s}\\
&= \sum_{k\leq\frac{\log x}{L}}\lambda^k + O\left(\frac{1}{T}\sum_{k\geq 0}\frac{|\lambda|^k\left(\frac{x}{e^{Lk}}\right)^c}{\left|\log\left(\frac{x}{e^{Lk}}\right)\right|}\right)\\
&= \frac{\lambda^{\lfloor\frac{\log x}{L}\rfloor+1}-1}{\lambda-1} + O\left(\frac{1}{T}\frac{x^c}{1-\frac{1}{e^{Lc}}|\lambda|}\right).
\end{aligned}$$

In the second line above we use the first part of Lemma 2 replacing the integral by 1 plus the error term. Similarly we can proceed if $\log x/L$ is an integer. Of course, this implies (31). □

Theorem 6. *Let L be a positive real number, λ be a non-zero complex number, and suppose that*

$$B(s) = \sum_{n\geq 1} B_n n^{-s}$$

is a Dirichlet series that is absolutely convergent for $\Re(s) > \frac{1}{L}\log|\lambda| - \eta$ *for some* $\eta > 0$*. Then*

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{B(s)}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = \sum_{n\geq 1} B_n \frac{\lambda^{\left\lfloor\frac{\log(x/n)}{L}\right\rfloor+1}}{\lambda-1} - \frac{1}{2}\sum_{n\geq 1} B_n \lambda^{\left\lfloor\frac{\log(x/n)}{L}\right\rfloor}[\![\log(x/n)/L \in \mathbb{Z}]\!] + O\left(x^{\frac{1}{L}\log|\lambda|-\eta}\right). \tag{32}$$

if $|\lambda| > 1$*, and*

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{B(s)}{1-e^{-Ls}}\frac{x^s}{s}\,ds = \sum_{n\geq 1} B_n\left(\left\lfloor\frac{\log(x/n)}{L}\right\rfloor+1\right) - \frac{1}{2}\sum_{n\geq 1} B_n[\![\log(x/n)/L \in \mathbb{Z}]\!] + O\left(x^{-\eta}\right). \tag{33}$$

if $\lambda = 1$.

Proof. We split the integral into an infinite sum of integrals according to the series $B(s) = \sum_{n\geq 1} B_n n^{-s}$ and apply (31) for each term by replacing x by x/n.

First assume that $\log(x/n)/L$ is not an integer for $n \geq 1$. Hence, if $x > ne^{Lk}$, then we have

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{\frac{1}{n^s}}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = \frac{\lambda^{\left\lfloor\frac{\log(x/n)}{L}\right\rfloor+1}-1}{\lambda-1} + O\left(\frac{1}{T}\sum_{k\geq 0}\frac{|\lambda|^k\left(\frac{x}{e^{Lk}n}\right)^c}{\left|\log\left(\frac{x}{e^{Lk}n}\right)\right|}\right),$$

and if $x < ne^{Lk}$, then we just have

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{\frac{1}{n^s}}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = O\left(\frac{1}{T}\sum_{k\geq 0}\frac{|\lambda|^k\left(\frac{x}{e^{Lk}n}\right)^c}{\left|\log\left(\frac{x}{e^{Lk}n}\right)\right|}\right).$$

Further, for given x there are only finitely many pairs (k, n) with

$$\left|\frac{x}{e^{Lk}n} - 1\right| < \frac{1}{2}.$$

Hence, the series

$$\sum_{n\geq 1}\sum_{k\geq 0} B_n \frac{|\lambda|^k\left(\frac{x}{e^{Lk}n}\right)^c}{\left|\log\left(\frac{x}{e^{Lk}n}\right)\right|}$$

is convergent. Consequently we get

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{\sum_{n\geq 1} B_n n^{-s}}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds =$$

$$\frac{1}{\lambda-1}\sum_{n<x} B_n \left(\lambda^{\left\lfloor \frac{\log(x/n)}{L}\right\rfloor+1}-1\right)+O(1)$$

(and a similar expression if there are integers $n\ge 1$ for which $\log(x/n)/L$ is an integer). Finally, since

$$\sum_{n<x} B_n = O\left(n^{\frac{1}{L}\log|\lambda|-\eta}\right)$$

and

$$\sum_{n>x} B_n n^{-\frac{1}{L}\log|\lambda|} = O\left(x^{-\eta}\right)$$

it follows that

$$\frac{1}{\lambda-1}\sum_{n<x} B_n \left(\lambda^{\left\lfloor \frac{\log(x/n)}{L}\right\rfloor+1}-1\right) = \frac{1}{\lambda-1}\sum_{n\ge 1} B_n \left(\lambda^{\left\lfloor \frac{\log(x/n)}{L}\right\rfloor+1}\right)+O\left(n^{\frac{1}{L}\log|\lambda|-\eta}\right)$$

(and similarly if there are integers $n\ge 1$ for which $\log(x/n)/L$ is an integer). This proves (35).

If $\lambda=1$ we first observe that (31) rewrites to

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{1}{1-e^{-Ls}}\frac{x^s}{s}\,ds = \left\lfloor \frac{\log x}{L}\right\rfloor + 1 - \frac{1}{2}[\![\log x/L\in\mathbb{Z}]\!].$$

Now the proof of (33) is very similar to that of (35). □

Remark 6. The representations (32) and (33) have nice interpretations. When $|\lambda|>1$ set

$$\Psi(t)=\lambda^{-t}\sum_{n\ge 1} B_n \frac{\lambda^{\left\lfloor t-\frac{\log n}{L}\right\rfloor+1}}{\lambda-1} - \frac{\lambda^{-t}}{2}\sum_{n\ge 1} B_n \lambda^{\left\lfloor t-\frac{\log n}{L}\right\rfloor}[\![t-\log n/L\in\mathbb{Z}]\!].$$

Then $\Psi(t)$ is a periodic function of bounded variation with period 1, that has (usually) countably many discontinuities for $t=\{\log n/L\}$, $n\ge 1$, and

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{B(s)}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = x^{\frac{1}{L}\log\lambda}\,\Psi\left(\frac{\log x}{L}\right)+O\left(x^{\frac{1}{L}\log|\lambda|-\eta}\right).$$

Formally this representation also follows by adding the residues of

$$B(s)/(1-e^{-Ls}\lambda)$$

at $s=s_0+2k\pi i/L$ ($k\in\mathbb{Z}$) which are the zeros of $1-e^{Ls}\lambda=0$. This means that in both cases the leading asymptotic follows from a *formal residue calculus.*

Furthermore, if we go back to the original problem, where we have to discuss a function of the form

$$\frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s},$$

for $\log p_j$ rationally related, then we have

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{B(s)}{1-\sum_{j=1}^m b_j p_j^s}\frac{x^s}{s}\,ds = x^{s_0}\,\Psi\left(\frac{\log x}{L}\right)+O\left(x^{s_0-\eta}\right).$$

As mentioned above we split up the integral with the help of a partial fraction decomposition of the rational function

$$\frac{1}{1-\sum_{j=1}^{m} b_j z^{n_j}}.$$

The *leading term* can be handled directly with the help of Theorem 3. The remaining terms one again uses (31) and obtains (finally) a second error term of order $O\left(x^{s_0-\eta}\right)$.

Remark 7. If $\lambda = 1$ then the situation is even simpler. Set

$$C = \frac{1}{L}\sum_{n\ge 1} B_n$$

and

$$\tilde{\Psi}(t) = \sum_{n\ge 1} B_n \left(-\left\{t-\frac{\log n}{L}\right\}+1\right) - \frac{1}{2}\sum_{n\ge 1} B_n [\![t-\frac{\log n}{L}\in\mathbb{Z}]\!] - \frac{1}{L}\sum_{n\ge 1} B_n \log n.$$

Then $\tilde{\Psi}(t)$ is a periodic function with period 1, that has (usually) countably many discontinuities for $t = \{\log n/L\}$, $n \ge 1$, and we have

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT} \frac{B(s)}{1-e^{-Ls}}\frac{x^s}{s}\,ds = C\log x + \tilde{\Psi}\left(\frac{\log x}{L}\right) + O\left(x^{-\eta}\right).$$

Hence, by applying the same partial fraction decomposition as above we also obtain (if $s_0 = 0$ and if the $\log p_j$ are rationally related)

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT} \frac{B(s)}{1-\sum_{j=1}^{m} b_j p_j^s}\frac{x^s}{s}\,ds = C\log x + \tilde{\Psi}\left(\frac{\log x}{L}\right) + O\left(x^{-\eta}\right).$$

Remark 8. There is also an immediate generalization of (32) to functions of the form

$$B(s) = \sum_{n\ge 1} B_n\left(\frac{1}{n^s} - \frac{1}{(n+h_n)^s}\right), \tag{34}$$

where $(h_n)_{n\ge 1}$ is a sequence that is bounded by $h_n = O(n^{1-\delta})$ for some $\delta > 0$ and where the series converges absolutely for $\Re(s) > \frac{1}{L}\log|\lambda| - \eta$ for some $\eta > 0$. Here we have

$$\begin{aligned}\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT} \frac{B(s)}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds &= \frac{1}{1-\lambda^{-1}}\sum_{n\ge 1} B_n\left(\lambda^{\lfloor\frac{\log(x/n)}{L}\rfloor} - \lambda^{\lfloor\frac{\log(x/(n+h_n))}{L}\rfloor}\right) \\ &\quad -\frac{1}{2}\sum_{n\ge 1} B_n \lambda^{\lfloor\frac{\log(x/n)}{L}\rfloor}[\![\log(x/n)/L\in\mathbb{Z}]\!] \\ &\quad +\frac{1}{2}\sum_{n\ge 1} B_n \lambda^{\lfloor\frac{\log(x/(n+h_n))}{L}\rfloor}[\![\log(x/(n+h_n))/L\in\mathbb{Z}]\!] + O(1).\end{aligned} \tag{35}$$

Again if we define the 1-periodic function

$$\begin{aligned}\overline{\Psi}(t) &= \frac{\lambda^{-t}}{1-\lambda^{-1}}\sum_{n\ge 1} B_n\left(\lambda^{\lfloor t-\frac{\log n}{L}\rfloor} - \lambda^{\lfloor t\frac{\log(n+h_n)}{L}\rfloor}\right) - \frac{\lambda^{-t}}{2}\sum_{n\ge 1} B_n\lambda^{\lfloor t\frac{\log n}{L}\rfloor}[\![t-\log n/L\in\mathbb{Z}]\!] \\ &+ \frac{\lambda^{-t}}{2}\sum_{n\ge 1} B_n\lambda^{\lfloor t-\frac{\log(n+h_n)}{L}\rfloor}[\![t-\log(n+h_n)/L\in\mathbb{Z}]\!],\end{aligned}$$

then

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{B(s)}{1-e^{-Ls}\lambda}\frac{x^s}{s}\,ds = x^{\frac{1}{L}\log\lambda}\,\overline{\Psi}\left(\frac{\log x}{L}\right)+O\left(x^{\frac{1}{L}\log|\lambda|-\eta}\right).$$

Summing up, we can handle all parts of $\widetilde{T}(s)$ (given by (27)) with the help of these techniques if $s_0 > \sigma_a$ and $s_0 \ge 0$. (Recall that $G'(s)$ is a finite Dirichlet series and $E'(s)$ is a finite sum of function of the form (34).)

4.5 Finishing the Proof

It remains to complete the proof of Theorem 1 in the *rationally related* case. Actually we only have to (re)consider the cases, where $s_0 \ge \sigma_a$. Namely, if $\sigma_a > s_0$ then the zeros of the equation (10) do not contributed to the leading analytic behaviour of $\widetilde{T}(s)$ and we can apply proper Tauberian theorems. In what follows we comment on the differences in the cases of interest.

2. $\sigma_a < s_0$ and $s_0 = 0$:
   This case is basically handled in Theorem 3, in particular see Remarks 7 and 8.

3. $\sigma_a = s_0 = 0$:
   In this case we apply proper generalizations of Tauberian theorems. Recall that in this case the dominant singular term of $\frac{1}{s}\widetilde{T}(s)$ is given by

   $$C\,\frac{(\log(1/s))^{\beta_0}}{s^{\alpha_0+2}}$$

   and there are infinitely many simple poles at $s = 2\pi ik/L$ ($k \in \mathbb{Z}\setminus\{0\}$). Of course we have $\alpha_0 \ge 0$, otherwise the sequence a_n would not be non-decreasing. Here we need a slightly modified version of Theorem 11 or Theorem 12, resp., which is easy to establish. The proof just requires that the Fourier series (15) converges and represents an integrable function, see Remark 9. However, this property does not effect the asymptotic leading term, it is only required in the proof.

4. $\sigma_a < s_0$ and $s_0 > 0$:
   Here we apply Theorem 6, see also Remark 6.

5. $\sigma_a = s_0$ and $s_0 > 0$:
   This case is very similar to Case 3.

5 Proof of Theorem 3

Finally, we briefly discuss the proof of Theorem 3 for the non-symmetric binary case (biased memoryless source). For simplicity, we shall write p for p_1 and q for $p_2 = 1 - p \ne p_1$.

We recall that $C(n, y)$ satisfies the recurrence (17) with initial conditions $C(0, y) = 0$ and $C(1, y) = 1$. It is clear that for every fixed positive real number y we can apply Theorem 1. However, we have to be careful since we need an asymptotic representation for $C(n, y)$ uniformly for y in an interval that contains 1 in its interior. Note that $C(n, 1) = 1$.

For the proof of Theorem 3, one has to consider the Dirichlet series

$$C(s, y) = \sum_{n=1}^{\infty}\frac{C(n+2, y) - C(n+1, y)}{n^s}.$$

For simplicity we just consider here the case $y > 1$. (The case $y \le 1$ can be handled in a similar way.) Then $C(s, y)$ converges for $\Re(s) > s_0(y)$, where $s_0(y)$ denotes the real zero of the equation $y(p^{s+1} + q^{s+1}) = 1$. We find

$$C(s,y) = \frac{(y-1) - \tilde{E}(s,y)}{1 - y(p^{s+1} + q^{s+1})},$$

where

$$\begin{aligned}\tilde{E}(s,y) &= py \sum_{k=1}^{\infty} (C(k+2,y)) - C(k+1,y)) \left(\frac{1}{(k/p)^s} - \frac{1}{\left(\left\lceil \frac{k+2-\delta}{p} \right\rceil - 2\right)^s} \right) \\ &+ qy \sum_{k=1}^{\infty} (C(k+2,y)) - C(k+1,y)) \left(\frac{1}{(k/q)^s} - \frac{1}{\left(\left\lfloor \frac{k+1+\delta}{q} \right\rfloor - 1\right)^s} \right)\end{aligned}$$

converges for $\Re(s) > s_0(y) - 1$ and satisfies $\tilde{E}(0, y) = 0$ and $\tilde{E}(s, 1) = 0$.

Suppose first that we are in the irrational case. Then by the Wiener-Ikehara theorem only the residue at $s_0(y)$ contributes to the main asymptotic leading term. (Recall that we consider the case $y > 1$). We thus have

$$\begin{aligned}C(n,y) &\sim \operatorname{Res}\left(\frac{((y-1) - \tilde{E}(s,y))(n-3/2)^s}{s(1 - y(p^{s+1} + q^{s+1}))}; s = s_0(y) \right) \\ &= \frac{((y-1) - \tilde{E}(s_0(y),y))(n-3/2)^{s_0(y)}}{-s_0(y)(\log(p)p^{s_0(y)+1} + \log(q)q^{s_0(y)+1}))}(1 + o(1)).\end{aligned}$$

The essential but non-trivial observation is that this asymptotic relation holds uniform for y in an interval around 1. In order to make this precise we can use the Mellin-Perron formula from Theorem 4

$$C(n,y) = C(2,y) + \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} C(s,y) \frac{\left(n - \frac{3}{2}\right)^s}{s} \, ds$$

and apply the methods presented in the proof of Theorem 5, compare also with [10]. We observe that the sum of residues (that is denoted by $Q(x)$ in the proof of Theorem 5) converges to 0 uniformly in y. This follows from the fact that the zeros of the equation $y(p^{s+1} + q^{s+1}) = 1$ vary continuously in y. Hence, if y in contained in some (compact) interval Y and $s_{\mathrm{nr}}(y)$ denotes one of the non-real zeros, then

$$\min_{y \in Y} \Re(s_{\mathrm{nr}}(y) - s_0(y)) > 0.$$

Hence we find

$$C(n,y) = (1 + O(y-1))n^{s_0(y)}(1 + o(1))$$

uniformly for real y that are contained in an interval around 1. Finally by using the local expansion

$$s_0(y) = \frac{y-1}{H} + \left(\frac{H_2}{2H3} - \frac{1}{H} \right) (y-1)2 + O((y-1)3), \tag{36}$$

and by setting $y = e^{t/(\log n)^{1/2}}$ we obtain

$$
\begin{aligned}
n^{s_0(y)} &= \exp\left(\log n\left(\frac{y-1}{H} - \left(\frac{1}{H} - \frac{H_2}{2H3}\right)(y-1)2 + O(|z-1|^3)\right)\right)\\
&= \exp\left(\frac{1}{H}t\sqrt{\log n} + \frac{1}{H}\frac{t2}{2} - \left(\frac{1}{H} - \frac{H_2}{2H3}\right)t2 + O(t3/\sqrt{\log n})\right)\\
&= \exp\left(\frac{1}{H}t\sqrt{\log n} + \left(\frac{H_2}{H3} - \frac{1}{H}\right)\frac{t2}{2} + O(t3/\sqrt{\log n})\right),
\end{aligned}
$$

and consequently

$$
\mathbb{E}\left[e^{D_n t/\sqrt{\log n}}\right] = C\left(n, e^{t/\sqrt{\log n}}\right) = \exp\left(\frac{1}{H}t\sqrt{\log n} + \left(\frac{H_2}{H3} - \frac{1}{H}\right)\frac{t2}{2}\right)(1+o(1)).
$$

Hence, we arrive at

$$
\begin{aligned}
\mathbb{E}\left[e^{t(D_n - \frac{1}{H}\log n)/\sqrt{\log n}}\right] &= e^{-(t/H)\sqrt{\log v}}\mathbb{E}\left[e^{D_n t/\sqrt{\log n}}\right] \\
&= e^{\frac{t2}{2}\left(\frac{H_2}{H3} - \frac{1}{H}\right)} + o(1).
\end{aligned} \tag{37}
$$

By the convergence theorem for the Laplace transform (see [30]) this proves the normal limiting distribution as $n \to \infty$ and also convergence of (centralized) moments.

In the rational case we can use a similar procedure. However, we have to use a proper variation of the proof of Theorem 3, from which we obtain estimates that are uniform in (real) y. Formally, we just have to add the residues coming from the zeros $s_k(y) = s_0(y) + k2\pi i/L$ for $k \neq 0$ (where $L > 0$ is the largest real number such that $\log(1/p)$ and $\log(1/q)$ are integer multiples of L). These terms lead to an additional contribution of the form

$$
\sum_{k \in \mathbb{Z}\setminus\{0\}} \frac{((y-1) - \tilde{E}(s_k(y), y))(n-3/2)^{s_k(y)}}{-s_k(y)(\log(p)p^{s_k(y)+1} + \log(q)q^{s_k(y)+1}))} = O(|y-1|n^{s_0(y)}).
$$

Since $(y-1) - \tilde{E}(s_k(y), y) = O(|y-1|)$ it follows that these additional terms are bounded by $O(|y-1|n^{s_0(y)})$. Hence, if we set $y = e^{t/(\log n)^{1/2}}$ this term is asymptotically negligible and the central limit theorem follows also in the rational case.

References

[1] M. Abramowitz, and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1964.

[2] M. Akra and L. Bazzi, On the Solution of Linear Recurrence Equations, *Computational Optimization and Applications*, 10,195-201, 1998.

[3] T. M. Apostol, Introduction to analytic number theory, Undergraduate Texts in Mathematics, Springer, New York, 1976.

[4] C. G. Boncelet, Block arithmetic coding for source compression, *IEEE Trans. Inform. Theory*, IT-39, no.5, pp.1546-1554, Sep 1993.

[5] Y.K. Cheung, P. Flajolet, M. Golin, and C.Y. Lee, Multidimensional Divide-and-Conquer and Weighted Digital Sums, *ANALCO*, 2008.

[6] V. Choi and M. J. Golin, Lopsided trees. I. Analyses. *Algorithmica*, 31, 3, 240-290, 2001.

[7] T. Cormen, C. Leierson, and R. Rivest, *Introduction to Algorithms*, MIT Press, Cambridge, Mass., 1990.

[8] H. Delange, Généralisation du théorème de Ikehara, *Ann. Sci. Éc. Norm. Supér.*, 71, 213–242, 1954.

[9] H. Delange, Sur la fonction sommatoire de la fonction "Somme des Chiffres." *Enseignement Math.* 21, 31–47, 1975.

[10] M. Drmota, Y. Reznik, and W. Szpankowski, Tunstall Code, Khodak Variations, and Random Walks *IEEE Trans. Information Theory*, 56, 2010.

[11] P. Erdős, A. Hildebrand, A. Odlyzko, P. Pudaite, and B. Reznick, The Asymptotic Behavior of a Family of Sequences, *Pacific Journal of Mathematics*, 126, 227-241, 1987.

[12] G. Fayolle, P. Flajolet, and M. Hofri, On a functional equation arising in the analysis of a protocol for a multi-access broadcast channel, *Adv. Appl. Prob.*, 18, 441-472, 1986.

[13] P. Flajolet, and M. Golin, Mellin Transforms and Asymptotics: The Mergesort Recurrence, *Acta Informatica*, 31, 673–696, 1994.

[14] P. Flajolet, X. Gourdon, and P. Dumas, Mellin Transforms and Asymptotics: Harmonic sums, *Theoretical Computer Science*, 144, 3–58, 1995.

[15] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2008.

[16] P. Flajolet, M. Roux, and B. Vallee, Digital Trees and Memoryless Sources: from Arithmetics to Analysis, *Proc. AofA'10*, Vienna, 2010.

[17] P. J. Grabner and J. M. Thuswaldner, *Analytic Continuation of a Class of Dirichlet Series*, Abh. Math. Sem. Univ. Hamburg **66** (1996), 281–287.

[18] H-K. Hwang, Distribution of the number of factors in random ordered factorizations of integers, *Journal of Number Theory* ,81, 61–92, 2000.

[19] H.-K. Hwang and S. Janson, *A central limit theorem for random ordered factorizations of integers*, manuscript, 2010.

[20] G. L. Khodak, Connection Between Redundancy and Average Delay of Fixed-Length Coding, *All-Union Conference on Problems of Theoretical Cybernetics* (Novosibirsk, USSR, 1969) 12 (in Russian)

[21] A. Karatsuba, and Yu. Ofman, Multiplication of Many-Digital Numbers by Automatic Computers, *Doklady Akad. Nauk SSSR* 145, 293-294, 1962 (translation in *Physics-Doklady* 7, 595-596, 1963).

[22] D.E. Knuth, *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

[23] J. Korevaar, A Century of complex Tauberian theory, *Bull. Amer. Math. Soc.* 39 (2002), 475–531.

[24] T.W. Körner, *Fourier analysis*, Second edition, Cambridge University Press, Cambridge, 1989.

[25] T. Leighton, *Notes on Better Master Theorems for Divide-and-Conquer Recurrences*, Manuscript. Massachusetts Institute of Technology, 1996.

[26] S. Roura, Improved Master Theorems for Divide-and-Conquer Recurrences, *Journal of the ACM*, 48, 170 – 205, 2001.

[27] S. A. Savari and Robert G. Gallager; Generalized Tunstall codes for sources with memory, IEEE Trans. Info. Theory, vol. IT-43, pp. 658 - 668, March 1997.

[28] W. Schachinger, Limiting distributions for the costs of partial match retrievals in multidimensional tries. *Random Structures and Algorithms*, 17, no. 3-4, 428-459, 2001.

[29] R. Sedgewick, and P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison-Wesley, Reading, MA, 1995.

[30] W. Szpankowski, Average Case Analysis of Algorithms on Sequences, Wiley, New York, 2001.

[31] B. P. Tunstall, Synthesis of Noiseless Compression Codes, Ph.D. dissertation, (Georgia Inst. Tech., Atlanta, GA, 1967)

A Analytic Continuation of Dirichlet Series

Dirichlet series of special sequences are frequently used in the present paper. In particular we are interested in the Dirichlet series of sequences of the form

$$c(n) = n^\sigma (\log n)^\alpha.$$

It is clear that the Dirichlet series $C(s) = \sum_{n\ge 1} c(n) n^{-s}$ converges (absolutely) for complex s with $\Re(s) > \sigma + 1$. We also know that the abscissa of absolute convergence is given by $\sigma_a = \sigma + 1$. However, it is not immediate that $C(s)$ has a certain analytic continuation to a larger region (that does not contain the singularity $s = \sigma_a$. Nevertheless, such continuation properties do hold (see [17]).

Theorem 7. *Suppose that σ and α are real numbers and let $C(s)$ be the Dirichlet series*

$$C(s) = \sum_{n\ge 2} n^\sigma (\log n)^\alpha n^{-s}.$$

(i) *If α is not a negative integer, then $C(s)$ can be represented as*

$$C(s) = \frac{\Gamma(\alpha+1)}{(s-\sigma-1)^{\alpha+1}} + G(s),$$

where $G(s)$ is an entire function.

(ii) *If $\alpha = -k$ is a negative integer, then we have*

$$C(s) = \frac{(-1)^k}{(k-1)!}(s-\sigma-1)^{k-1}\log(s-\sigma+1) + G(s),$$

where $G(s)$ is an entire function.

Proof. We do not provide a full proof but sketch the arguments from [17] where even a slightly more general situation was considered. Furthermore it is sufficient to consider the case $\sigma = 0$.

First it follows form the Euler Maclaurin summation that $C(s)$ can be represented (for $\Re(s) > 1$) as

$$C(s) = \int_2^\infty \frac{(\log v)^\alpha}{v^s}\,dv + \frac{(\log 2)^\alpha}{2^{s+1}} + \int_2^\infty \left(\{v\} - \frac{1}{2}\right)\left(\alpha(\log v)^{\alpha-1} - s(\log v)^\alpha\right) v^{-s-1}\,dv,$$

where the second integral on the right hand side represents a function that is analytic for $\Re(s) > 0$. Furthermore, by using the substitution $z = (s-1)\log v$ the first integral can be rewritten as

$$\int_2^\infty \frac{(\log v)^\alpha}{v^s}\,dv = (s-1)^{-\alpha-1}\int_{(s-1)\log 2}^\infty z^\alpha e^{-z}\,dz.$$

The latter integral is precisely the incomplete Γ-function.

If α is not a negative integer, then [1]

$$\int_w^\infty z^\alpha e^{-z}\,dz = \Gamma(\alpha+1) - w^{\alpha+1}\sum_{m=0}^\infty \frac{(-1)^m}{m!}\frac{w^m}{(m+\alpha+1)}$$

and if $\alpha = -k$ is a negative integer, then [1]

$$\int_w^\infty z^{-k} e^{-z}\,dz = \Gamma_{k-1}(-k+1) + \frac{(-1)^k}{(k-1)!}\log(w) - w^{\alpha+1}\sum_{m=0,\, m\ne k-1}^\infty \frac{(-1)^m}{m!}\frac{w^m}{(m+\alpha+1)},$$

where $\Gamma_k(z) = \Gamma(z) - (-1)^k/(k!(k+z))$. Hence the conclusion follows. □

Note that the above method is quite flexible. For example, if

$$c(n) = n^\sigma(\log n)^\alpha + O(n^{\sigma-\delta})$$

for some $\delta > 0$, then we obtain a similar representation except that $G(s)$ is not any more an entire function but a function that is analytic for $\Re(s) > \sigma + 1 - \delta$.

It is now easy to apply Theorem 7 to sequences of the form

$$c(n) = a_{n+2} - a_{n+1},$$

where

$$a_n = n^\sigma(\log n)^\alpha.$$

Theorem 8. *Suppose that $a_n = n^\sigma (\log n)^\alpha$, where σ and α are real numbers, and let $\widetilde{A}(s)$ be the Dirichlet series*

$$\widetilde{A}(s) = \sum_{n\ge 1} \frac{a_{n+2}-a_{n+1}}{n^s}.$$

(i) *If α is not a negative integer, then $\widetilde{A}(s)$ can be represented as*

$$\widetilde{A}(s) = \sigma\frac{\Gamma(\alpha+1)}{(s-\sigma)^{\alpha+1}} + \frac{\Gamma(\alpha+1)}{(s-\sigma)^{\alpha}} + G(s),$$

where $G(s)$ is analytic for $\Re(s) > \sigma - 1$.
(ii) *If $\alpha = -k$ is a negative integer, then we have*

$$\begin{aligned}\widetilde{A}(s) &= \sigma\frac{(-1)^k}{(k-1)!}(s-\sigma)^{k-1}\log(s-\sigma)\\ &\quad + \frac{k(-1)^k}{(k-1)!}(s-\sigma)^{k}\log(s-\sigma) + G(s),\end{aligned}$$

where $G(s)$ is analytic for $\Re(s) > \sigma - 1$.

Proof. This follows from the simple fact that

$$\begin{aligned}a_{n+2}-a_{n+1} &= \sigma n^{\sigma-1}(\log n)^\alpha\left(1+O(n^{-1})\right)\\ &\quad + \alpha n^{\sigma-1}(\log n)^{\alpha-1}\left(1+O(n^{-1})\right).\end{aligned}$$

□

Note that Theorem 8 is even more flexible than Theorem 7. For example, we can also consider sequences of the form $a_n = (\lfloor \rho n + \tau \rfloor)^\sigma$ for some ρ with $0 < \rho < 1$ (or similarly defined sequences). In this case one could argue, as in Section 4.1, that

$$\widetilde{A}(s) = \rho^s B(s) + R(s),$$

where $B(s)$ is the Dirichlet series of the differences $(n+2)^\sigma - (n+1)^\sigma$ and $R(s)$ is analytic for $\Re(s) > \sigma - 1$.

Finally we show that condition (15) of Theorem 1 is satisfied for sequences $a_n = n^\sigma(\log n)^\alpha$.

Theorem 9. *Suppose that $a_n = n^\sigma (\log n)^\alpha$ and let $\widetilde{A}(s)$ denote the corresponding Dirichlet series. Then the Fourier series*

$$\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\widetilde{A}(\sigma + 2\pi i k/L)}{\sigma + 2\pi i k/L} e^{2\pi i k x/L} \tag{A.1}$$

is convergent for $x \in \mathbb{R}$ and represents an integrable function.

Proof. We restrict ourselves to the case $\sigma = 1$, which means that the sequence $a_{n+2} - a_{n+1}$ consists (mainly) of the two terms $(\log n)^\alpha$ and $(\log n)^{\alpha-1}$. To simplify the presentation, we only discuss the function

$$A(s) = \sum_{n\ge 2} (\log n)^\alpha n^{-s}$$

instead of $\widetilde{A}(s)$ (and neglect the error terms, since they be handled easily).

Following the proof of Theorem 7 we have to discuss the three integrals

$$\begin{aligned} A_1(s) &= \int_2^\infty \frac{(\log v)^\alpha}{v^s}\, dv, \\ A_2(s) &= \int_2^\infty \left(\{v\} - \frac{1}{2}\right) \alpha (\log v)^{\alpha-1} v^{-s-1}\, dv, \\ A_3(s) &= s \int_2^\infty \left(\{v\} - \frac{1}{2}\right) (\log v)^\alpha v^{-s-1}\, dv. \end{aligned}$$

Let us start with $A_3(s)$ which we represent as

$$A_3(s) = s \int_0^\infty v^{-s} h(v)\, dv,$$

where $h(v) = 0$ for $0 \le v < 2$ and $h(v)/v$ is of bounded variation on $[2, \infty)$. (Note that in our case, $h(v)$ is not continuous if v is an integer.) Set

$$F(x) = L \sum_{m \in \mathbb{Z}} h(e^{x+mL}).$$

Then $F(x)$ is periodic (with period L) and also of bounded variation. Hence it has a convergent Fourier series with Fourier coefficients (see [24])

$$\begin{aligned} f_k &= \frac{1}{L} \int_0^L F(x) e^{-2\pi i k x/L}\, dx = \int_0^L \sum_{m \in \mathbb{Z}} h(e^{x+mL}) e^{-(x+mL)2\pi i k/L}\, dx \\ &= \int_{-\infty}^\infty h(e^x)\, e^{-x 2\pi i k/L}\, dx = \int_0^\infty h(v)\, v^{-(1+2\pi i k/L)}\, dv \\ &= \frac{A_3(1+2\pi i k/L)}{1+2\pi i k/L}. \end{aligned}$$

Consequently the Fourier series with Fourier coefficients $A_3(1+2\pi i k/L)/(1+2\pi i k/L)$ is convergent. Furthermore, it is integrable, since the set of discontinuities of $F(x)$ is countable and $F(x)$ equals its Fourier series at all points of continuity (here we use the fact that $f_k = O(1/k)$).

Similarly we can handle $A_2(s)$. We represent it as

$$A_2(s) = \int_2^\infty \overline{h}(v) v^{-s}\, dv,$$

where $\overline{h}(v)/v$ is of bounded variation on $[2, \infty)$. Here the corresponding periodic function is given by

$$\overline{F}(x) = L \int_2^\infty \overline{h}(v) \frac{e^{-L\{(x-\log v)/L\}}}{v(1-e^{-L})}\, dv.$$

Finally, we have to consider $A_1(s)$. By Theorem 7 we know that $A_1(s)$ has an analytic continuation to the slit region $\mathbb{C} \setminus (-\infty, 1]$. In particular it follows that the limit

$$\lim_{\varepsilon \to 0+} A_1(1 + \varepsilon + 2\pi i k/L)$$

exists and equals to (the analytically continued value) $A_1(1+2\pi ik/L)$. By partial integration it follows that (for real t)

$$\int_2^\infty \frac{(\log v)^\alpha}{v^{1+\varepsilon+it}}\,dv = \frac{(\log 2)^\alpha}{\varepsilon+it}2^{-\varepsilon-it} + \frac{\alpha}{\varepsilon+it}\int_2^\infty \frac{(\log v)^{\alpha-1}}{v^{1+\varepsilon+it}}\,dv.$$

This implies that

$$A_1(1+it) = O\left(\frac{1}{t}\right).$$

Consequently the Fourier series with Fourier coefficients $A_1(1+2\pi ik/L)/(1+2\pi ik/L)$, $k\neq 0$, converges absolutely. This completes the proof of the Theorem. □

B Tauberian Theorems

The main analytic problem in the present paper is to obtain asymptotic information on the partial sums

$$\overline{c}(v) = \sum_{n\le v} c(n)$$

from analytic properties of the Dirichlet series

$$C(s) = \sum_{n\ge 1} c(n)n^{-s}.$$

The classical Tauberian theorem of Wiener-Ikehara, as presented in Theorem 10, is a very strong tool in this context. Actually it applies to the Mellin-Stieltjes transforms (see [23]) that is closely related to Dirichlet series:

$$C(s) = \sum_{n\ge 1} c(n)n^{-s} = \int_{1-}^\infty v^{-s}\,d\overline{c}(v).$$

Theorem 10 (Wiener-Ikehara; cf. [23]). *Let $\overline{c}(v)$ be non-negative and non-decreasing on $[1,\infty)$ such that the Mellin-Stieltjes transform*

$$C(s) = \int_{1-}^\infty v^{-s}\,d\overline{c}(v) = s\int_1^\infty \overline{c}(v)v^{-s-1}\,dv$$

exists for $\Re(s)>1$. Suppose that for some constant $A_0>0$, the analytic function

$$F(s) = \frac{1}{s}C(s) - \frac{A_0}{s-1} \qquad (\Re(s)>1)$$

has a continuous extension to the closed half-plane $\Re(s)\ge 1$. Then

$$\overline{c}(v) \sim A_0 v$$

as $v\to\infty$.

Theorem 10 is quite flexible. For example, it is sufficient to assume that $\overline{c}(v)(\log v)^\alpha$ is non-decreasing for some real α (and $v\ge 2$). Furthermore it is clear that it generalizes directly to the case when $C(s)$ converges for $\Re(s)>s_0$ and has a continuous extension to the closed half-plane $\Re(s)\ge s_0$ (for $s_0\ge 0$). It also applies if $C(s)$ behaves like a pole of higher order for $s\to\ s_0$, however, the asymptotic result has to be adjusted accordingly.

Theorem 11. *Let $\overline{c}(v)$ be non-negative and non-decreasing on $[1,\infty)$ such that the Mellin-Stieltjes transform $C(s)$ exists for $\Re(s) > s_0$ for some $s_0 \ge 0$ and suppose that there exist real constants $A_0, \ldots, A_K$ (with $A_K > 0$) such that*

$$\tilde{F}(s) = \frac{1}{s}C(s) - \sum_{j=0}^{K} \frac{A_j}{(s-s_0)^{j+1}} \tag{B.1}$$

has a continuous extension to the closed half-plane $\Re(s) \ge s_0$. Then we have

$$\overline{c}(v) \sim \frac{A_K}{K!}(\log v)^K v^{s_0} \qquad (v \to \infty). \tag{B.2}$$

We indicate how Theorem 11 can be deduced from (a slight variation of) Theorem 10 when $K = 2$ and $s_0 = 1$. Let

$$\frac{1}{s}C(s) = \int_1^\infty \overline{c}(v)v^{-s-1}\,ds = \frac{A_1}{(s-1)2} + \frac{A_0}{s-1} + \tilde{F}(s)$$

with some $A_1 > 0$ and some function $\tilde{F}(s)$ that is analytic for $\Re(s) > 1$ and has a continuous extension to the half plane $\Re(s) \ge 1$. By subtracting $A_0/(s-1)$ and by splitting up the integral into two parts we obtain

$$\begin{aligned}\int_2^\infty (\overline{c}(v) - A_0 v)\, v^{-s-1}\,dv &= \frac{A_1}{(s-1)2} + \tilde{F}(s) \\ &\quad - \int_1 2\,(\overline{c}(v) - A_0 v)\, v^{-s-1}\,dv\end{aligned}$$

Hence, by integrating with respect to s (from 2 to s) we have

$$\begin{aligned}\int_2^\infty \left(\frac{\overline{c}(v) - A_0 v}{\log v}\right) v^{-s-1}\,dv &= \frac{A_1}{s-1} - A_1 - \int_2^s \tilde{F}(t)\,dt + \int_2^\infty \left(\frac{\overline{c}(v) - A_0 v}{\log v}\right) v^{-3}\,dv \\ &\quad + \int_2^s \int_1 2(\overline{c}(v) - A_0 v)v^{-t-1}\,dv\,dt.\end{aligned}$$

We can apply a slight generalization of Theorem 10 to $(\overline{c}(v) - A_0 v)/\log v$. Note that the right hand side is of the form $A_1/(s-1) + \overline{F}(s)$, where $\overline{F}(s)$ has a continuous continuation to the half plane $\Re(s) \ge 1$. The point is that the function $(\overline{c}(v) - A_0 v)/\log v$ is not necessarily non-negative and non-decreasing. However, there is certainly a constant $C > 0$ such that $(\overline{c}(v) - A_0 v)/\log v + Cv \ge 0$, and A_1 on the right hand side can be replaced by $A_1 + C$. Furthermore, the proof of Theorem 10 has some flexibility. As mentioned above the proof of Theorem 10 can be easily modified so that it also applies to a function of the form $(\overline{c}(v) - A_0 v)/\log v$, where it is only assumed that $\overline{c}(v)$ is non-decreasing [23].

Note that the cases $s_0 > 0$ and $s_0 = 0$ of Theorem 1 have to be handled separately.[3] Furthermore, the case $s_0 < 0$ is not applicable in this setting. Namely if $c(v) > 0$ and non-decreasing, then $C(s)$ cannot converge for s with $\Re(s) < 0$. Note also that we cannot expect a more precise asymptotic expansion in this generality. For example if $\overline{c}(v) = (A_1 \log v + A_0 + \sin(\log 2v))v$ with $A_1 > 2$. Then $\overline{c}(v)$ is positive and non-decreasing, so (B.1) is satisfied but we do not have $\overline{c}(v) = (A_1 \log v + A_0 + o(1))v$.

[3] The approach we present works for $s_0 > 0$. For $s_0 = 0$ we have to adjust parts of the proof of Theorem 10.

Remark 9. The above mentioned proof method of Theorem 11 also applies to situations, where $\frac{1}{s}C(s)$ has a representation of the form

$$\frac{1}{s}C(s) = \int_1^\infty \overline{c}(v)v^{-s-1}\,ds = \frac{A_1}{(s-1)2} + \sum_{m\in\mathbb{Z}} \frac{A_{0,m}}{s+im\tau-1} + \tilde{F}(s)$$

with some $A_1 > 0$ and some function $\tilde{F}(s)$ that is analytic for $\Re(s) > 1$ and has a continuous extension to the half plane $\Re(s) \geq 1$. Furthermore we have to assume that the Fourier series

$$\sum_{m\in\mathbb{Z}} A_{0,m}e^{im\tau x}$$

is convergent and represents an integrable function. Note that this condition corresponds to the condition (15) in Theorem 1. Under these assumptions the previous proof works, too, and it follows that $\overline{c}(v) \sim A_1 v \log v$.

This kind of reasoning is precisely what is needed in Section 4.5, where we completed the proof of Theorem 1 in the rationally related case.

There are even more general versions by Delange [8] that cover singularities of algebraic-logarithmic type that we state next. Note that this theorem requires an analytic continuation property and not only a continuity property.

Theorem 12 (Delange [8])**.** *Let $\overline{c}(v)$ be non-negative and non-decreasing on $[1,\infty)$ such that the Mellin-Stieltjes transform $C(s)$ exists for $\Re(s) > s_0$ for some $s_0 > 0$ and suppose that there exist functions $\tilde{F}(s)$, $g_0(s),\ldots,g_J(s)$ that are analytic in a region that contains half plane $\Re(s) \geq s_0$ such that*

$$\frac{1}{s}C(s) = g_0(s)\frac{\left(\log\frac{1}{s-s_0}\right)^{\beta_0}}{(s-s_0)^{\alpha_0}} + \sum_{j=1}^{J} g_j(s)\frac{\left(\log\frac{1}{s-s_0}\right)^{\beta_j}}{(s-s_0)^{\alpha_j}} + \tilde{F}(s),$$

where $g_0(s_0) \neq 0$, β_j are non-negative integers, α_0 is real but not a negative integer when it is non-zero, and $\alpha_1,\ldots,\alpha_J$ are complex numbers with $\Re(\alpha_j) < \alpha_0$. Furthermore $\beta_0 > 0$ if α_0 is contained in the set $\{0,-1,-2,\ldots\}$. Then, as $v \to \infty$,

$$\overline{c}(v) \sim \frac{g_0(s_0)}{\Gamma(\alpha_0)}(\log v)^{\alpha_0-1}(\log\log v)^{\beta_0}v^{s_0} \tag{B.3}$$

if α_0 is not contained in the set $\{0,-1,-2,\ldots\}$ and

$$\overline{c}(v) \sim (-1)^{\alpha_0}(-\alpha_0)!\beta_0 g_0(s_0)(\log v)^{\alpha_0-1}(\log\log v)^{\beta_0-1}v^{s_0} \tag{B.4}$$

if α_0 is contained in the set $\{0,-1,-2,\ldots\}$ and $\beta_0 > 0$.

Interestingly, Theorem 10 generalizes – partly – to the case, where there are infinitely many *poles* on the line $\Re(s) = s_0$, where one obtains a fluctuating factor in the asymptotic expansion.

The drawback of this generalization is that it only applies if the appearing periodic function has an absolutely convergent Fourier series. Unfortunately we cannot apply it in the present context, since the appearing periodic functions have discontinuities. Anyway, we could not find such a theorem in the literature, so we present it here.

Theorem 13. *Let $\overline{c}(v)$ be non-negative and non-decreasing on $[1, \infty)$ such that the Mellin-Stieltjes transform $C(s)$ exists for $\Re(s) > s_0$, where $s_0 > 0$. Assume that the function*

$$\tilde{F}(s) = \frac{1}{s}C(s) - \sum_{m\in\mathbb{Z}} \frac{A_m}{s - s_0 - im\tau}, \qquad \text{(B.5)}$$

with some real $\tau > 0$ and real coefficients A_m, where $A_0 > 0$, has a continuous extension to the closed half-plane $\Re(s) \geq s_0$. Furthermore assume that the Fourier series

$$\Psi(x) = \sum_{m\in\mathbb{Z}} A_m\, e^{im\tau x}$$

is absolutely convergent and has bounded derivative. Then

$$\overline{c}(v) \sim \Psi(\log v)\, v^{s_0} \qquad (v \to \infty). \qquad \text{(B.6)}$$

The proof is an extension of the approach from [23]. For the reader's convenience we give it here. Let

$$K_\lambda(t) = \frac{1 - \cos(\lambda t)}{\pi\lambda t2} = \frac{\lambda}{2\pi}\left(\frac{\sin(\lambda t/2)}{\lambda t/2}\right)2$$

denote the Fejer kernel.

Lemma 4. *Let $\kappa > 0$ and*

$$h(t) = \sum_{m\in\mathbb{Z}} A_m e^{im\tau t}$$

be an absolutely convergent Fourier series with bounded derivative. Then

$$\int_0^\infty K_\lambda(u - t)h(t)\, dt = h(u) + o(1) \qquad (\lambda \to \infty) \qquad \text{(B.7)}$$

uniformly for $u \geq 1$.

Proof. Let $m_0 = m_0(\varepsilon)$ be defined by

$$\sum_{|m|>m_0(\varepsilon)} |A_m| < \varepsilon$$

and suppose that for $\lambda_0 = \lambda_0(\varepsilon) >> \kappa m_0(\varepsilon)$ we have

$$\sum_{|m|\leq m_0(\varepsilon)} |mA_m| < \varepsilon\, \lambda_0(\varepsilon).$$

Furthermore we note that for $u \geq 1$

$$\int_u^\infty K_\lambda(t)\, dt = O\left(\frac{1}{\lambda}\right) \qquad (\lambda \to \infty).$$

Consequently it follows for $\lambda \ge \lambda_0(\varepsilon)$

$$\begin{aligned}\int_0^\infty K_\lambda(u-t)h(t)\,dt &= \sum_{m\in\mathbb{Z}} A_m \int_0^\infty K_\lambda(u-t)e^{im\tau t}\,dt \\ &= \sum_{m\in\mathbb{Z}} A_m \int_{-\infty}^\infty K_\lambda(t)e^{im\tau(u-t)}\,dt + O\left(\int_u^\infty K_\lambda(t)\,dt\right) \\ &= \sum_{m\in\mathbb{Z}} A_m e^{im\kappa u}\hat{K}_\lambda(\kappa m) + O\left(\frac{1}{\lambda}\right) \\ &= \sum_{m\in\mathbb{Z}} A_m e^{im\kappa u}\left(1-\frac{|\kappa m|}{\lambda}\right) + O\left(\frac{1}{\lambda}\right) \\ &= \sum_{|m|\le\lambda/\kappa} A_m e^{im\kappa u} + O\left(\frac{1}{\lambda}\sum_{|m|\le\lambda/\kappa}|mA_m|\right) + O\left(\frac{1}{\lambda}\right)\end{aligned}$$

Since $\lambda_0 > \kappa m_0$ we have

$$\left|h(u) - \sum_{|m|\le\lambda/\kappa} A_m e^{im\kappa u}\right| \le \sum_{|m|>m_0(\varepsilon)} |A_m| < \varepsilon.$$

Furthermore

$$\begin{aligned}\frac{1}{\lambda}\sum_{|m|\le\lambda/\kappa}|mA_m| &\le \frac{1}{\lambda}\sum_{|m|\le m_0}|mA_m| + \frac{1}{\lambda}\sum_{m_0<|m|\lambda/\kappa\le}|mA_m| \\ &< \varepsilon + \sum_{|m|>m_0}|A_m| \\ &< 2\varepsilon.\end{aligned}$$

Of course this proves (B.7). □

Lemma 5. *Let $\ell(t)$ be non-negative for $t \ge 0$ such that the Laplace transform*

$$L(z) = \int_0^\infty \ell(t)e^{-zt}\,dt$$

exists for $\Re(z) > 0$. Suppose further that there exists a bounded and integrable function $h(t)$ (for $t \ge 0$) with the property that

$$G(z) = L(z) - H(z)$$

has a continuous extension to the closed half-plane $\Re(z) \ge 0$, where $H(z)$ denotes the Laplace transform of $h(t)$. Then

$$\lim_{u\to\infty}\left(\int_0^\infty K_\lambda(u-t)\ell(t)\,dt - \int_0^\infty K_\lambda(u-t)h(t)\,dt\right) = 0$$

Proof. Let $\hat{K}_\lambda(y) = \max\{1-|y|/\lambda, 0\}$ denote the Fourier transform of $K_\lambda(t)$ which is non-negative and has support $[-\lambda,\lambda]$. Then we have for $x > 0$

$$\begin{aligned}\int_0^\infty K_\lambda(u-t)\ell(t)e^{ixt}\,dt &= \int_0^\infty K_\lambda(u-t)h(t)e^{ixt}\,dt \\ &\quad + \frac{1}{2\pi}\int_{-\lambda}^\lambda \hat{K}_\lambda(y)G(x+iy)e^{iuy}\,dy.\end{aligned}$$

By assumption the right hand side has a finite limit as $x \to 0+$. Hence, by the monotone convergence theorem it follows that $K_\lambda(u-t)\ell(t)$ is integrable over $(0,\infty)$ and it follows that

$$\int_0^\infty K_\lambda(u-t)\ell(t)\,dt = \int_0^\infty K_\lambda(u-t)h(t)\,dt + \frac{1}{2\pi}\int_{-\lambda}^{\lambda} \hat{K}_\lambda(y)G(iy)e^{iuy}\,dy.$$

Finally, the Riemann-Lebesgue lemma implies

$$\lim_{u\to\infty} \frac{1}{2\pi}\int_{-\lambda}^{\lambda} \hat{K}_\lambda(y)G(iy)e^{iuy}\,dy = 0.$$

This proves the lemma. □

With the help of these preliminaries we prove Theorem 11.

Proof of Theorem 11. We set $\ell(t) = e^{-s_0 t}a(e^t)$. Then for $\Re(z) > 0$

$$L(z) = \int_0^\infty \ell(t)e^{-zt}\,dt = \int_1^\infty a(v)v^{-(s_0+z)-1}\,dv = \frac{A(s_0+z)}{s_0+z}.$$

Furthermore observe that the Laplace transform of $\Psi(t)$ (for $\Re(z) > 0$) is given by

$$H(z) = \int_0^\infty \Psi(t)e^{-tz}\,dt = \sum_{m\in\mathbb{Z}} \frac{A_m}{z - im\tau}$$

Hence, by assumption the function

$$G(z) = L(z) - H(z) = \frac{A(s_0+z)}{s_0+z} - \sum_{m\in\mathbb{Z}} \frac{A_m}{z-im\tau}$$

has a continuous extension to the half-plane $\Re(z) \ge 0$. Consequently by Lemma 5

$$\lim_{u\to\infty}\left(\int_0^\infty K_\lambda(u-t)e^{-s_0 t}a(e^t)\,dt - \int_0^\infty K_\lambda(u-t)\Psi(t)\,dt\right) = 0$$

Since $\Psi(t)$ is bounded it also follows that the second integral is uniformly bounded in λ and u. Hence

$$\limsup_{u\to\infty} \int_0^\infty K_\lambda(u-t)\ell(t)\,dt \le C$$

for some constant that is uniform in λ. Since $a(v)$ is positive and non-decreasing it follows that

$$\int_0^\infty K_\lambda(u-t)\ell(t)\,dt \ge \ell(u-1/\sqrt{\lambda})e^{-2s_0/\sqrt{\lambda}} \int_{-1/\sqrt{\lambda}}^{1/\sqrt{\lambda}} K_\lambda(t)\,dt = \ell(u-1/\sqrt{\lambda})\left(1+O(1/\sqrt{\lambda})\right)$$

and consequently

$$\limsup_{u\to\infty} \ell(u-1/\sqrt{\lambda}) \le C\left(1+O(1/\sqrt{\lambda})\right).$$

This shows that $\ell(t)$ is a bounded function.

Now, for given $\varepsilon > 0$ choose $\lambda_0 = \lambda_0(\varepsilon) >> 1/\varepsilon 2$ such that

$$\left|\int_0^\infty K_{\lambda_0}(u-t)\Psi(t)\,dt - \Psi(u)\right| < \varepsilon.$$

Since $\Psi(t)$ has bounded derivative we also have $|\Psi(u) - \Psi(u - 1/\sqrt{\lambda_0})| \le C/\sqrt{\lambda_0} \le C\varepsilon$. Putting these estimates together it follows that

$$\limsup_{u\to\infty} \left(\ell(u - 1/\sqrt{\lambda_0}) \left(1 + O(1/\sqrt{\lambda_0})\right) - \Psi(u - 1/\sqrt{\lambda_0})\right) \le (1 + C)\varepsilon$$

and consequently

$$\limsup_{u\to\infty} (\ell(u) - \Psi(u)) \le 0.$$

Similarly we obtain estimates from below. We just have to observe that

$$\begin{aligned}\int_0^\infty K_\lambda(u-t)\ell(t)\,dt &\le \ell(u + 1/\sqrt{\lambda})e^{2s_0/\sqrt{\lambda}} \int_{-1/\sqrt{\lambda}}^{1/\sqrt{\lambda}} K_\lambda(t)\,dt + O\left(\int_{1/\sqrt{\lambda}} K_\lambda(t)\,dt\right) \\ &= \ell(u + 1/\sqrt{\lambda})\left(1 + O(1/\sqrt{\lambda})\right) + O\left(1/\sqrt{\lambda}\right)\end{aligned}$$

and obtain in the same way

$$\liminf_{u\to\infty} (\ell(u) - \Psi(u)) \ge 0.$$

Hence, $\ell(u) = \Psi(u) + o(1)$ and consequently $a(v) = (\Psi(\log v) + o(1))\, v^{s_0}$. Finally, since $a(v)$ is non-decreasing we have $\min \Psi(u) > 0$ and consequently $a(v) \sim \Psi(\log v) v^{s_0}$. □