THEORETICAL COMPUTER SCIENCE, 144, 1995

ASYMPTOTIC BEHAVIOR OF THE LEMPEL-ZIV PARSING SCHEME AND DIGITAL SEARCH TREES*

January 30, 1995

Philippe Jacquet†
INRIA
Rocquencourt
78153 Le Chesnay Cedex
France
jacquet@blagny.inria.fr

Wojciech Szpankowski‡
Department of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.
spa@cs.purdue.edu

Abstract

The Lempel-Ziv parsing scheme finds a wide range of applications, most notably in data compression and algorithms on words. It partitions a sequence of length n into variable phrases such that a new phrase is the shortest substring not seen in the past as a phrase. The parameter of interest is the number M_n of phrases that one can construct from a sequence of length n. In this paper, for the memoryless source with *unequal* probabilities of symbols generation we derive the limiting distribution of M_n which turns out to be normal. This proves a long standing open problem. In fact, to obtain this result we solved another open problem, namely, that of establishing the limiting distribution of the internal path length in a digital search tree. The latter is a consequence of an asymptotic solution of a multiplicative differential-functional equation often arising in the analysis of algorithms on words. Interestingly enough, our findings are proved by a combination of probabilistic techniques such as renewal equation and uniform integrability, and analytical techniques such as Mellin transform, differential-functional equations, de-Poissonization, and so forth. In concluding remarks we indicate a possibility of extending our results to Markovian models.

---

*Preliminary version of this paper was presented in *26-th ACM Symposium on the Theory of Computing*, Montreal, 1994.

†This work was partly supported by the ESPRIT Basic Research Action No. 7141 (ALCOM II).

‡This research was primarily done while the author was visiting INRIA in Rocquencourt, France. The author wishes to thank INRIA (projects ALGO, MEVAL and REFLECS) for a generous support. In addition, support was provided by NSF Grants NCR-9206315, CCR-9201078 and INT-8912631, and in part by NATO Collaborative Grant 0057/89.

1. INTRODUCTION

The primary motivation for this work is the desire to understand the asymptotic behavior of the fundamental parsing algorithm on words due to Lempel and Ziv [27]. It partitions a word into phrases (blocks) of variable sizes such that a new block is the shortest subword not seen in the past as a phrase. For example, the string 110010100010001000 is parsed into (1)(10)(0)(101)(00)(01)(000)(100).

These parsing algorithms play a crucial rôle in universal data compression schemes and their numerous applications such as efficient transmission of data (cf. [26, 27]), estimation of entropy (cf. [25]), discriminating between information sources (cf. [7]), test of randomness, estimating the statistical model of individual sequences (cf. [16]), and so forth. The parameters of interest for these applications are: the number of phrases, the number of phrases of a given size, the size of a phrase, the length of a sequence built from a given number of phrases, etc. But, by all means the most important parameter is the number of phrases: This parameter is used to obtain the compression ratio in a universal data compression (cf. [3]), while its distribution is needed in the analysis of other parameters of the Lempel-Ziv scheme (e.g., redundancy rate [21], length of a phrase [14], and so forth).

In this paper, we study the distribution of the number of phrases M_n constructed form a word of a fixed length n in a probabilistic framework. We assume that the word is generated by a probabilistic memoryless binary source (extension to finite non-binary alphabet is simple). That is: *symbols are generated in an independent manner with "0" and "1" occurring respectively with probability* p *and* $q = 1 - p$. If $p = q = 0.5$, then such a probabilistic model will be further called the *symmetric Bernoulli* model; otherwise we refer to the *asymmetric Bernoulli* model.

Aldous and Shields [1] attested that obtaining the limiting distribution of the number of phrases is a difficult problem. They solved it only for the *symmetric* Bernoulli model. The authors of [1] wrote: "It is natural to conjecture that asymptotic normality holds for a larger class of processes ... . But in view of the difficulty of even the simplest case (i.e., the fair coin-tossing case we treat here) we are not optimistic about finding a general result. We believe the difficulty of our normality result is intrinsic ... ." We settle the conjecture of [1] in the affirmative for the *asymmetric* Bernoulli model, and in concluding remarks we indicate a possibility of extending our findings to Markovian models. Actually, we do more, and provide solutions to some other problems that have been open up-to-date, namely: the limiting distribution for internal path lengths in digital trees (cf. [1, 11, 15]), the number of parsings of given length built from a fixed number of words (cf. [7]), and probabilistic behavior of the Lempel-Ziv code redundancy (cf. [16, 21]).

All of these problems are solved in a uniform manner by a combination of probabilistic and analytical methods. We apply the *renewal equation* (cf. [2]) to reduce the problem of finding the number of phrases in the Lempel-Ziv scheme to another problem on digital search trees, namely that of finding the limiting distribution of the internal path length in a digital search tree built from a *fixed* number of independent words.

The reader is referred to [11, 15] for discussion and definition of the digital trees. However, for the reader's convenience we show in Figure 1 the digital search tree associated with the word mentioned at the beginning of this section. In particular, the root of the tree is empty (i.e., we start parsing with an empty phrase). All other phrases of the Lempel-Ziv parsing algorithm are stored in internal nodes. When a new phrase is created, the search starts at the root and proceeds down the tree as directed by the input symbols exactly in the same manner as in the digital tree construction, that is, symbol "0" in the input string means move to the right and "1" means proceed to the left. The search is completed when a branch is taken from an existing tree node to a new node that has not been visited before. Then, the edge and the new node are added to the tree. The phrases created in such a way are stored directly into the nodes of the tree. In passing, we note that for a word of fixed length, n, the size of the associated digital search tree is *random*, and this fact gives a new twist to the analysis of digital trees (cf. also [14]).

Second-order properties, such as limiting distributions and large deviation results of the Lempel-Ziv scheme, have been scarcely discussed in the past with a notable exception of the work of Aldous and Shields [1] who studied the symmetric model. Recently, Louchard and Szpankowski [14] obtained the limiting distribution of a randomly selected phrase length in the Lempel-Ziv scheme.

On the other hand, digital search trees (built from a *fixed* number of independent words!) have been quite thoroughly investigated in the past (cf. [4, 5, 10, 11, 13, 14, 22]). In particular, Knuth [11], and Flajolet and Sedgewick [4] introduced analytical methods in the analysis of digital search trees. This was continued by Flajolet and Richmond [5], Louchard [13], and Szpankowski [22]. None of these papers, however, deals with second order properties of the internal path length in digital search trees, which is the main object of our study. Only very recently, Kirschenhofer, Prodinger and Szpankowski [10] obtained an asymptotic expression for the variance of the internal path length in the *symmetric* Bernoulli model (in fact, this allowed to close the gap in the Aldous and Shields analysis by yielding the leading term in the variance of the number of phrases in the Lempel-Ziv parsing scheme). The authors of [10], however, did not extend their results to the asymmetric model. We not only provide such an

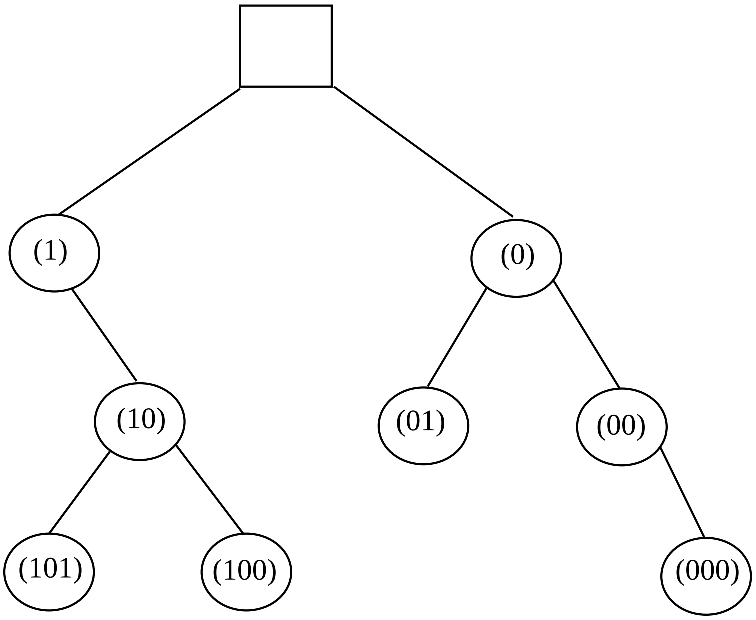

Figure 1: A digital tree representation of the Lempel-Ziv parsing for the string 11001010001000100...

extension, but we carry out this analysis to obtain the limiting distribution for the internal path length.

The paper is organized as follows. In the next section we present all our main findings concerning digital search trees and the Lempel-Ziv scheme. All proofs are delayed till Section 3 which is of its own interest. In this section, we present a methodology that leads to an asymptotic solution of a functional-differential equation that often arises in other problems of engineering and science.

2. MAIN RESULTS

We recall that the associated tree constructed during the course of the Lempel-Ziv parsing algorithm (cf. Figure 1) is a digital tree built from a *random* number of words (phrases). We apply the renewal equation to show that the limiting distribution of the internal path length in a digital search tree built from a *fixed* number of independent words directly implies the limiting distribution of the number of phrases M_n in the Lempel-Ziv parsing scheme. In the sequel, we first carry out our analysis for digital trees, and then provide necessary tools to derive the limiting distribution of M_n.

Let us first consider a digital search tree built from m statistically *independent* words (of possibly infinite length) each generated according to the Bernoulli model. We leave the root empty, and store the next word in the first available node, as discussed above (cf. [11, 15]). Let $D_m(i)$ be the length of a path from the root to the ith node containing this word. In fact, observe that $D_m(i) = D_i(i)$ for $m \geq i$ since the position of the ith node does not depend on words inserted after it. We define the internal path length as $L_m = \sum_{i=1}^{m} D_i(i)$.

Hereafter, we shall consistently use n as the length of a single word to be parsed according to the Lempel-Ziv scheme, and m as the number of words used to construct an independent digital search tree.

We infer probabilistic behavior of L_m from its generating function. Thus, we apply "analytical approach" to the problem. Define for complex u and z the following generating functions $L_m(u) = Eu^{L_m}$ and $L(z,u) = \sum_{m=0}^{\infty} L_m(u) z^m/(m!)$. We also set $\widetilde{L}(z,u) = L(z,u)e^{-z}$ which can be interpreted as the generating function of the internal path length in a family of digital search trees built from a random number of words which is Poisson distributed with mean z. Observe that this is a standard poissonization trick but disguised in a generating function form. One expects a simpler equation for $L(z,u)$ than for $L_m(u)$, and it turns to be true, as seen below.

The generating functions $L_m(u)$ and $L(z,u)$ satisfy some recurrences that we discuss next. Let $R_m = k$ be the number of words that start with "0". These words will create the right subtree of the digital search tree. Clearly, $\Pr\{R_m = k\} = \binom{m}{k} p^k q^{m-k}$, and the lengths of the internal paths in the right and the left subtrees are respectively L_k and L_{m-k}. Thus, *conditioning* on $\{R_m = k\}$ we have $L_{m+1} = m + L_k + L_{m-k}$, and finally $Eu^{L_{m+1}} = u^m E_{R_m}\left(Eu^{L_k} Eu^{L_{m-k}} | R_m\right)$ for u complex. Certainly, this implies the following

$$L_{m+1}(u) = u^m \sum_{k=0}^{m} \binom{m}{k} p^k q^{m-k} L_k(u) L_{m-k}(u) \qquad m \geq 0 \tag{1}$$

with $L_0(u) = 1$. Hence, also

$$\frac{\partial L(z,u)}{\partial z} = L(pzu, u) L(qzu, u) \tag{2}$$

with $L(z,0) = 1$. This is our basic functional-differential equation that we solve asymptotically to obtain the limiting distribution of L_m. We observe that the above equation is of a multiplicative form which makes the problem hard.

In section 3 we prove our main results concerning a digital search tree, which is stated below.

Theorem 1A. *Consider a digital search tree built from m independent words under the* **asymmetric** *Bernoulli model.*

(i) *Asymptotically the average value* EL_m *and the variance* Var L_m *become*

$$EL_m = \frac{m}{h}\left(\log m + \frac{h_2}{2h} + \gamma - 1 - \alpha + \delta_0(\log m)\right)$$
$$+ \frac{1}{h}\left(\log m + \frac{h_2}{2h} - \gamma - \log p - \log q + \alpha\right) + O(1) \quad (3)$$
$$\text{Var } L_m = c_2 m \log m + O(m) \quad (4)$$

where $h = -p\log p - q\log q$ *is the entropy of the alphabet,* $\gamma = 0.577\ldots$ *is the Euler constant,* $h_2 = p\log^2 p + q\log^2 q$*, and* $c_2 = (h_2 - h^2)/h^3$*,*

$$\alpha = -\sum_{k=1}^{\infty} \frac{p^{k+1}\log p + q^{k+1}\log q}{1 - p^{k+1} - q^{k+1}} , \quad (5)$$

and $\delta_0(\log m)$ *is a fluctuating functions for* $\log p/\log q$ *rational with small amplitude, and zero otherwise (cf. [8, 18, 22]).*

(ii) *Furthermore:*

$$\frac{L_m - EL_m}{\sqrt{\text{Var } L_m}} \xrightarrow{d} N(0,1) \quad (6)$$

where $\xrightarrow{d}$ *denote the convergence "in distribution", and* $N(0,1)$ *is the standard normal distribution. In fact, a stronger result holds: Let* $c_1 = 1/h$*. For a complex* ϑ *in a neighbourhood of zero, and for any* $\varepsilon > 0$

$$e^{-\vartheta c_1 m \log m} E e^{\vartheta L_m} = e^{c_2 \frac{\vartheta^2}{2} m \log m}\left(1 + O(1/m^{1/2+\varepsilon})\right) . \quad (7)$$

(iii) *There exist positive constants* A *and* $\mu < 1$ *such that uniformly in* k

$$\Pr\left\{\left|\frac{L_m - EL_m}{\sqrt{\text{Var } L_m}}\right| > k\right\} < A\mu^k \quad (8)$$

for large m*.* ■

Actually, our analytical approach also works for the symmetric Bernoulli model. We need, however, in this case to refine the method to obtain the leading term in the asymptotics of the variance. Fortunately, this was recently done by Kirschenhofer, Prodinger and Szpankowski [10] who proved that

$$\text{Var } L_m^{sym} \sim (C + \delta(\log_2 m))m \quad (9)$$

where $C = 0.26600\ldots$ and $\delta(x)$ is a fluctuating function with small amplitude (cf. Theorem 1B). In the above, we write L_m^{sym} for the internal path length in the symmetric case. We have the following result.

Theorem 1B. *For the* **symmetric** *Bernoulli model the following analogue of (7) holds*

$$e^{-\vartheta m \log_2 m} E e^{\vartheta L_m^{sym}} = e^{\frac{\vartheta^2}{2} m(C+\delta(\log_2 m))} \left(1 + O(1/m^{1/2+\varepsilon})\right) \ , \tag{10}$$

for any $\varepsilon > 0$, *where* $\delta(x)$ *is periodic continuous function of period* 1, *mean* 0 *and very small amplitude* ($< 10^{-6}$). *Similarly, uniformly in* k

$$\Pr\left\{ \left| \frac{L_m^{sym} - E L_m^{sym}}{\sqrt{m(C + \delta(\log_2 m))}} \right| > k \right\} < A\mu^k \tag{11}$$

with $\mu < 1$. ■

Equipped with Theorem 1, we now can attack the main problem, namely the limiting distribution of the number of phrases M_n in the Lempel-Ziv parsing scheme. Fortunately, the problem can be reduced to the one discussed in Theorem 1 through the so called *renewal equation* which we introduce next.

We recall that $D_i(i) = D_i(m)$ is the length of the ith phrase in the Lempel-Ziv parsing scheme obtained from a fixed number, m, of words (cf. [7, 14]), that is, the depth of the ith node in the digital tree built from these m words. Fix now n, and start partitioning the sequence of length n into phrases. Clearly, $D_1(1) = 1$. After obtaining the second phrase, we check whether $D_1(1) + D_2(2) > n$ or not. If yes, then $M_n = 1$, otherwise we continue the process. It should be clear by now that the number of phrases M_n can be computed from the following relationship

$$M_n = \max\{m : \ L_m = \sum_{i=1}^{m} D_i(i) \le n\} \ . \tag{12}$$

The above equation is known as the *renewal equation* (cf. [2]). We also observe that it directly implies the following

$$\Pr\{M_n > m\} = \Pr\{L_m \le n\} \ , \tag{13}$$

which is useful in some computations. The following result is due to Billingsley [2] (cf. Theorem 17.3).

Lemma 2. [Billingsley] *Let* M_n *and* L_m *satisfy the relationship (12), and assume* $D_i(i)$ *are positive random variables. Then*

$$\frac{L_m - \mu_m}{\sigma_m} \xrightarrow{d} N(0,1) \ , \tag{14}$$

implies

$$\frac{M_n - n/(\mu_n/n)}{\sigma_n(\mu_n/n)^{-3/2}} \xrightarrow{d} N(0,1) \tag{15}$$

where μ_m and σ_m are positive constants (in our case they can be asymptotically interpreted as the mean and the variance of L_m). ∎

Theorem 1 and Lemma 2 are next used to obtain the following result that solves the open problem left in Aldous and Shields [1].

Theorem 3. *In the asymmetric Bernoulli model, define $Z_n = \frac{M_n - EM_n}{\sqrt{\mathrm{Var} M_n}}$. Then:*

(i) *The sequence of random variables Z_n converges weakly (i.e., in distribution) to $N(0,1)$. In addition, for all $r \geq 0$ the sequence $(Z_n)^r$ is uniformly integrable. Thus, all moments of Z_n exist and converge to the appropriate moments of the normal distribution. In particular,*

$$EM_n \sim \frac{nh}{\log(n)} \tag{16}$$

$$\mathrm{Var}\ M_n \sim \frac{c_2 h^3 n}{\log^2 n} \tag{17}$$

where $c_2 = (h_2 - h^2)/h^3$.

(ii) *For any $\varepsilon > 0$, there exist an integer $n_0 \geq 1$ such that for all $n > n_0$*

$$\Pr\{|M_n - EM_n| > \varepsilon EM_n\} \leq A \exp\left(-a\varepsilon\sqrt{n}\right) \tag{18}$$

for some positive constants $A, a > 0$.

(iii) *The above results are also true for the symmetric model if one replaces the variance by*

$$\mathrm{Var}\ M_n^{sym} \sim \frac{n(C + \delta(\log_2 n))}{\log_2^3 n} \tag{19}$$

where the constant $C = 0.26600\ldots$ and $\delta(x)$ is defined in Theorem 1B. In (ii) one must replace $\sqrt{n}$ by $\sqrt{n/\log n}$.

Proof. The weak convergence of Z_n of part (i) follows directly from Lemma 2, while the uniform integrability of $(Z_n)^r$ for any $r \geq 0$ is a consequence of the large deviation result (cf. 18), which is proved next.

For part (ii), we will basically show that uniform integrability of L_m naturally translates into uniform integrability of M_n. To proceed we need some new notation. Define $\Lambda(m) = EL_m$ for all integer m, and let for all $y \geq 0$ the function $\Lambda(y)$ be a linear interpolation of $\Lambda(m)$ between integer points. (Actually, we could define $\Lambda(m)$ as an approximation of EL_m only up to $O(m)$ term in (3).) Now, let $\Lambda^{-1}(x)$ be the inverse function of $\Lambda(y)$ defined for $x \geq 0$. Note that $\Lambda(x) \sim \frac{x \log x}{h}$ and $\Lambda^{-1}(x) \sim \frac{hx}{\log x}$ for $x \to \infty$. As is easy to check, the function $\Lambda(\cdot)$ is convex, hence $\Lambda^{-1}(\cdot)$ is concave (this should be at least clear for $\Lambda(x)$ approximated by the first two terms in (3) which are convex functions). Now we refer to (8) of Theorem 1A so that we can find positive constants A and α such that for all m and $y \geq 0$, the following two inequalities hold:

$$\Pr\{L_m < \Lambda(m)(1-y)\} \leq A\exp\left(-\alpha y\sqrt{\Lambda(m)}\right) \tag{20}$$

$$\Pr\{L_m > \Lambda(m)(1+y)\} \leq A\exp\left(-\alpha y\sqrt{\Lambda(m)}\right) \tag{21}$$

Note that in the first inequality we can relax y to $y \leq 1$ since L_m cannot be negative.

We will not directly show that Z_n is uniformly integrable but rather that the sequence of random variables

$$Y_n = \frac{M_n - \Lambda^{-1}(n)}{\sqrt{n}\Lambda^{-1}(n)}$$

is uniformly integrable. Due to the asymptotic expansions of $\Lambda^{-1}(n)$ the uniform integrability of Z_n will follow.

We shall prove separately (18) for $M_n > (1+\varepsilon)EM_n$ and $M_n < (1-\varepsilon)EM_n$. For the former, let us first consider inequality (20) with $y \leq 1$. We refer to the fundamental identity (13) to obtain

$$\Pr\{M_{\lfloor\Lambda(m)(1-y)\rfloor} \geq m\} = \Pr\{L_m < \Lambda(m)(1-y)\}\ .$$

But, (20) implies

$$\Pr\{M_n \geq \Lambda^{-1}(n/(1-y))\} \leq A\exp\left(-\alpha\frac{y\sqrt{n}}{\sqrt{1-y}}\right)\ .$$

Using concavity of $\Lambda^{-1}(\cdot)$ and $y \leq 1$, we also have $\Lambda^{-1}(\frac{n}{1-y}) \leq \frac{\Lambda^{-1}(n)}{1-y}$, therefore $\Pr\{M_n \geq \Lambda^{-1}(n/(1-y))\} \geq \Pr\{M_n \geq \Lambda^{-1}(n)/(1-y)\}$. Thus,

$$\Pr\{M_n \geq (1-y)^{-1}\Lambda^{-1}(n)\} \leq A\exp\left(-\alpha\frac{y\sqrt{n}}{\sqrt{1-y}}\right)\ .$$

Setting $1+\frac{x}{\sqrt{n}} = \frac{1}{1-y}$ for $x \geq 0$ we finally obtain

$$\Pr\{M_n \geq (1+\frac{x}{\sqrt{n}})\Lambda^{-1}(n)\} \leq A\exp\left(-\alpha\frac{x}{\sqrt{1+\frac{x}{\sqrt{n}}}}\right) \leq A\exp\left(-\alpha\frac{x}{\sqrt{1+x}}\right)$$

which proves (18) for positive values of $x > 0$ after setting $\varepsilon = x/\sqrt{n}$.

To prove the uniform integrability for $M_n < (1-\varepsilon)EM_n$, we use (21) with $y \geq 0$. By the same arguments as above, we find

$$\Pr\{M_n < \Lambda^{-1}(n(1+y)^{-1})\} \leq A\exp\left(-\alpha\frac{y\sqrt{n}}{\sqrt{1+y}}\right)\ .$$

By concavity of $\Lambda^{-1}(\cdot)$ we obtain $\Lambda^{-1}(n/(1+y)) \geq \Lambda^{-1}(n)/(1+y)$ since $y \geq 0$. Therefore, $\Pr\{M_n < \Lambda^{-1}(n/(1+y))\} \geq \Pr\{M_n < (\Lambda^{-1}(n)/(1+y)\}$ and thus:

$$\Pr\{M_n < (1+y)^{-1}\Lambda^{-1}(n)\} \leq A \exp\left(-\alpha \frac{y\sqrt{n}}{\sqrt{1+y}}\right) .$$

Setting $1 - \frac{x}{\sqrt{n}} = \frac{1}{1+y}$ for $x \geq 0$ (but smaller than $\sqrt{n}$ because M_n cannot be negative) we obtain

$$\Pr\{M_n < (1 - \frac{x}{\sqrt{n}})\Lambda^{-1}(n)\} \leq A \exp\left(-\alpha \frac{x}{\sqrt{1-\frac{x}{\sqrt{n}}}}\right) \leq A \exp(-\alpha x)$$

which proves the second part of (18) for $M_n < (1-\varepsilon)EM_n$. Observe that the uniform integrability of $(Z_n)^r$ $(r \geq 0)$ follows, which further implies that $EM_n \sim \Lambda^{-1}(n)$ and Var $M_n \sim$ Var $L_n \cdot (\Lambda^{-1}(n)/n)^3$. This proves (16) and (17). ■

Theorem 1 and Theorem 2 have several important consequences for data compression, coding theory, and so forth. We will discuss only two applications of our results, namely the number of parsings of given length (cf. [7]), and a large deviation estimate of the Lempel-Ziv code redundancy (cf. [16, 21]).

Let us start with the problem posed by Gilbert and Kadota [7], namely: How many parsings of total length n one can construct from m words? For example, for $m = 2$ we have four parsings of length three, namely: (0)(00), (0)(01), (1)(10) and (1)(11), and two parsings of length two, namely: (0)(1) and (1)(0). Thus, let $F_m(n)$ be the number of parsings built from m words of total length n, and let $F_m(x) = \sum_{n=0}^{\infty} F_m(n)x^n$ be its generating function. Note that (cf. [7])

$$F_{m+1}(x) = x^m \sum_{k=0}^{m} \binom{m}{k} F_k(x)F_{m-k}(x). \tag{22}$$

The next result is a direct consequence of Theorem 1B and it answers a problem of [7].

Corollary 4A. *The number of parsings built from m words of total length n is*

$$F_m(n) = 2^n \Pr\{L_m^{sym} = n\} . \tag{23}$$

In particular, for $n = m\log_2 m + O(\sqrt{m})$ we obtain asymptotically

$$F_m(n) \sim \frac{2^n}{\sqrt{2\pi(C+\delta(\log_2 n))m}} \exp\left(-\frac{(n - m\log_2 m)^2}{2(C+\delta(\log_2 n))m}\right) \tag{24}$$

where C and $\delta(x)$ are defined in Theorem 1B. If $n = (1+\varepsilon)m\log_2 m$, then the large deviation result (11) must be used.

Proof. Consider the recurrence (22) and note that $F_m(x/2) = L_m^{sym}(x)$. The rest follows from Theorem 1B and the analysis of Section 3. ■

Finally, we consider the redundancy rate R_n of the Lempel-Ziv code. It is defined as (cf. [16, 21])

$$R_n = \frac{(M_n+1)\log M_n - nh}{n} \ . \tag{25}$$

(Note that $(M_n+1)\log M_n$ is the length of the Lempel-Ziv code while nh is the length of the optimal code.) The redundancy rate R_n is a measure of the additional cost in using the Lempel-Ziv code instead of the optimal one predicted by the *Source Coding Theorem* (cf. [3]). It is known [16] that for the Bernoulli model $ER_n = O(\log\log n/\log n)$, but very little seems to be known about probabilistic behavior of R_n (e.g., how quickly the Lempel-Ziv code achieves the optimal rate). The next results provides some insights into this problem.

Corollary 4B. *For $r \ll h$ and large n*

$$\Pr\{R_n > r\} = \Pr\{(M_n+1)\log M_n > n(h+r)\} \le A\exp\left(-ar\sqrt{\frac{n}{(h-r)}}\right) \tag{26}$$

for some constants $A, a > 0$.

Proof. It is a simple consequence of (18) in Theorem 3(ii), and an asymptotic solution of $(M_n+1)\log M_n = n(h+r)$. Details are left for the interested reader. ■

The functional equation studied here has many other applications in the analysis of algorithms on words (e.g., external path length of a trie or PATRICIA trie). For more details the reader is referred to our preliminary version of this paper [9].

3. ANALYSIS AND PROOFS

This section provides all necessary details required to prove our main finding which is Theorem 1. We shall adhere to the following plan:

1. We first analyze the Poisson model (cf. Section 3.1) that is characterized by the exponential bivariate generating function $L(z,u)$ satisfying (2).

2. We transform the multiplicative equation (2) into an additive functional equation by considering $\log L(z,u)$. For this we need the existence of $\log L(z,u)$ in some domain. We shall prove that there is a convex cone (cf. Definition 1 of Sec. 3.1) around the real axis of z and a real neighbourhood $\mathcal{U}(1)$ of $u=1$ such that for some $\kappa(u)$ we have $\log L(z,u) = \Theta(z^{\kappa(u)})$ (cf. Theorem 5 and proof in Section 3.2).

3. Next, we use the Taylor expansion of $\log L(z,u)$ in the convex cone to show that for large z the generating function $L(z,u)$ appropriately normalized converges to the generating function of the normal distribution (cf. Theorem 6 of Sec. 3.1).

4. To prove the above we must know precise asymptotics for the average and the variance of the internal path length in the Poisson model (cf. Theorem 6).

5. The final effort is to de-Poissonize the above results, that is, to transform the normal distribution of the Poisson model into the normal distribution of the Bernoulli model (cf. Theorem 9 and proof in Section 3.3).

3.1. A Streamlined Analysis

The goal of this section is to establish Theorem 1 for the **Poisson model** which is summarized in Theorem 6 and Corollary 8 below.

We shall consistently use the notation from Section 2. In particular, we write $L_m(u)$ to denote the generating function of the internal path length, and $L(z,u)$ for the bivariate generating function. We assume the function equation (2) holds, that is,

$$\frac{\partial L(z,u)}{\partial z} = L(pzu,u)L(qzu,u) \tag{27}$$

with $L(z,0)=1$.

As before, by $\widetilde{L}(z,u) = L(z,u)e^{-z}$ we denote the moment generating function in the Poisson model. Let also $\widetilde{X}(z)e^z = L'_u(z,1)$ and $\widetilde{V}(z)e^z = L''_{uu}(z,1) + L'_u(z,1) - (L'_u(z,1))^2$ where $L'_u(z,1)$ and $L''_{uu}(z,1)$ denote the first and the second derivative of $L(z,u)$ with respect to u at $u=1$. We observe that $\widetilde{X}(z)$ and $\widetilde{V}(z)$ are the mean and the variance of L_m in the Poisson model. We shall need precise asymptotics of $L(z,u)$, $\widetilde{X}(z)$ and $\widetilde{V}(z)$ as $z\to\infty$ in a cone around the real axis of z for u real and positive in a neighbourhood $\mathcal{U}(1)$ of $u=1$.

The domain of z we shall work with is a *convex cone* $\mathcal{C}(D,\delta)$ defined as follows:

Definition 1. Convex Cone. *The set $\mathcal{C}(D,\delta)$ of $z=x+iy$ is called the convex cone if for $\delta<1$ and $x>0$ we have $|y|\le Dx^\delta$ for some $D>0$.* ■

The crucial part of our proof relies on proving the existence of the logarithm of $L(z,u)$ in a convex cone and real positive u. For this, we need a precise bound for $L(z,u)$ in such a cone. Let $\kappa(u)$ be a solution of the following equation

$$(pu)^{\kappa(u)} + (qu)^{\kappa(u)} = 1\ . \tag{28}$$

It is easy to notice that $\kappa(u) = 1 + h(u-1) + O((u-1)^2)$. In Section 3.2 we prove the following result which is the "heart" of our asymptotic analysis of the function equation (27).

Theorem 5. (i) *There exists a convex cone $\mathcal{C}(D,\delta)$, a neighbourhood $\mathcal{U}(1)$ of $u=1$, and a constant ξ such that for $|z|>\xi$ the logarithm of $L(z,u)$ exists and* $\log L(z,u) = \Theta(z^{\kappa(u)})$.

(ii) *In addition, under the same hypothesis as above, for any $\beta > 0$ and all $l \geq 1$*

$$\frac{\partial^l}{\partial^l u} \log L(z,u) = O(z^{\kappa(u)+\beta}) \tag{29}$$

for $z \in \mathcal{C}(D,\delta)$ and $z \to \infty$. ■

Now we are ready to formulate our main result concerning the Poisson model.

Theorem 6. *In the Poisson model for $(z,u) \in \mathcal{C}[D,\delta] \times \mathcal{U}(1)$ the following holds for any $\beta > 0$ and $z \to \infty$ (with $u = e^t$ for some real t in the vicinity of zero)*

$$\log L(z, e^t) = z + \widetilde{X}(z)t + \widetilde{V}(z)\frac{t^2}{2} + O(t^3 z^{\kappa(u)+\beta}) , \tag{30}$$

where the mean $\widetilde{X}(z)$ and the variance $\widetilde{V}(z)$ become asymptotically

$$\widetilde{X}(z) = \frac{z}{h}\log z + \frac{z}{h}\left(\gamma - 1 + \frac{h_2}{2h} - \alpha - \delta_1(\log z)\right) + O(1) , \tag{31}$$

and

$$\widetilde{V}(z) = \frac{z\log^2 z}{h^2} + \frac{2z\log z}{h^3}\left(\gamma h + h_2 - \frac{h^2}{2} - \alpha h - h\delta_1(\log z)\right) + O(z) , \tag{32}$$

where

$$\alpha = -\sum_{k=1}^{\infty} \frac{p^{k+1}\log p + q^{k+1}\log q}{1 - p^{k+1} - q^{k+1}} . \tag{33}$$

The function $\delta_1(x)$ is identically equal to zero when $\frac{\log p}{\log q}$ is irrational, while for $\frac{\log p}{\log q} = \frac{r}{t}$ where r, t are integers such that $\gcd(r,t) = 1$, the function $\delta_1(\log z)$ is fluctuating with a small amplitude as given by the formula below

$$\delta_1(\log z) = \sum_{\substack{\ell=-\infty \\ \ell \neq 0}}^{\infty} \frac{\Gamma(s_0^\ell)Q(-2)}{Q(s_0^\ell - 1)} \exp\left(-\frac{2\pi i\ell r}{\log p}\log z\right) , \tag{34}$$

where

$$Q(s) = \prod_{k=0}^{\infty}(1 - p^{-s+k} + q^{-s+k}) , \tag{35}$$

and $s_0^\ell = -1 + 2\pi i\ell r/\log p$ for $\ell = \mathcal{Z} = \pm 1, \pm 2, \ldots$ is a solution of $1 = p^{-s} + q^{-s}$.

Proof. Using Taylor's expansion of $\log L(z,u)$ we obtain

$$\log L(z,u) = \log L(z,1) + (u-1)\frac{\partial}{\partial u}\log L(z,1) + \frac{(u-1)^2}{2}\frac{\partial^2}{\partial u^2}\log L(z,1) + R(z,u)$$

with $R(z,u)$ being the remainder term of the following form (cf. [19])

$$R(z,u) = \frac{1}{2}\int_1^u (v-1)^2 \frac{\partial^3}{\partial v^3}\log L(z,v)\mathrm{d}v .$$

Due to Theorem 5 the error term is $O(z^{\kappa(u)+\beta}(u-1)^3)$. Now it suffices to note that $\log L(z,1) = z$ and to substitute $u = e^t$ for t in the vicinity of zero to obtain (30).

The remaining part of the proof is devoted to deriving the asymptotics (31) and (32) of $\widetilde{X}(z)$ and $\widetilde{V}(z)$. Since we need several terms of such asymptotic expansions (to prove our main result concerning the Bernoulli model; cf. Lemma 10 below) we use the Mellin transform method. The reader may familiarize himself with the technique from [15].

Consider first the mean $\widetilde{X}(z)$. Direct differentiation of our basic equation (27) leads to the following recurrence

$$\widetilde{X}(z) + \widetilde{X}'(z) = \widetilde{X}(zp) + \widetilde{X}(zq) + z \ . \tag{36}$$

Let $X(s)$ denote the Mellin transform of $\widetilde{X}(z)$, that is, $X(s) = \int_0^\infty \widetilde{X}(z) z^{s-1} dz$. It can be easily proved that it exists in the strip $\Re(s) \in (-2,-1)$. Observe that the Mellin transform of $\widetilde{X}'(z) - z$ is also defined in $\Re(s) \in (-2,-1)$. Then, (36) translates into

$$X(s) - (s-1)X(s-1) = p^{-s}X(s) + q^{-s}X(s) \tag{37}$$

in terms of the Mellin transform.

To solve the functional equation (37) we make a substitution $X(s) = \gamma(s)\Gamma(s)$ where $\Gamma(s)$ is the gamma function (cf. [19, 11]), and $\gamma(s)$ satisfies the following recurrence

$$\gamma(s) - \gamma(s-1) = p^{-s}\gamma(s) + q^{-s}\gamma(s) \ .$$

After some algebra one obtains

$$\gamma(s) = \prod_{k=0}^{\infty} \frac{1 - p^{k+2} - q^{k+2}}{1 - p^{-s+k} - q^{-s+k}} = \frac{Q(-2)}{Q(s)} \tag{38}$$

for $\Re(s) \in (-2,-1)$. Applying the Cauchy residue theorem (cf. [19]) to the above (i.e., inverting the Mellin transform) one proves (31). In fact, the calculation are almost exactly the same as the ones done in [22], so they are omitted (the fluctuating function $\delta_1(x)$ for $\log p/\log q$ rational is derived below).

The variance is more intricate, as already seen in [10]. We observe (after quite tedious algebra) that $\widetilde{W}(z) = \widetilde{V}(z) - \widetilde{X}(z)$ satisfies the following recurrence

$$\widetilde{W}(z) + \widetilde{W}'(z) = \widetilde{W}(zp) + \widetilde{W}(zq) + 2zp\widetilde{X}'(zp) + 2zq\widetilde{X}'(zq) + (\widetilde{X}'(z))^2 \ . \tag{39}$$

This functional equation is harder to solve due to the last term for which there is no closed form expression for the Mellin transform. But, fortunately, we can prove that the last term contributes $O(z)$ and we need only terms up to $O(z)$ (to recover the leading terms in the variance in the Bernoulli model as indicated by (49) of Lemma 10). Indeed, let $\widetilde{W}(z) = \widetilde{W}_1(z) + \widetilde{W}_2(z)$ where

$$\widetilde{W}_1(z) + \widetilde{W}_1'(z) = \widetilde{W}_1(zp) + \widetilde{W}_1(zq) + 2zp\tilde{X}'(zp) + 2zq\tilde{X}'(zq) \tag{40}$$

$$\widetilde{W}_2(z) + \widetilde{W}_2'(z) = \widetilde{W}_2(zp) + \widetilde{W}_2(zq) + (\tilde{X}'(z))^2 \; . \tag{41}$$

Using the tools of Section 3.2, we will prove in the appendix the following simple result.

Lemma 7A. *A solution $\widetilde{W}_2(z)$ of (41) satisfies $\widetilde{W}_2(z) = O(z)$ for $z \to \infty$.* ■

We concentrate now on solving (40) for $\widetilde{W}_1(z)$. Its Mellin transform $W_1(s)$ becomes

$$W_1(s) - (s-1)W_1(s-1) = (p^{-s} + q^{-s})W_1(s) - 2(p^{-s} + q^{-s})sX(s) \; .$$

After the substitution $W_1(s) = \Gamma(s)\beta(s)$ we find

$$\beta(s)(1 - p^{-s} - q^{-s}) - \beta(s-1) = -2(p^{-s} + q^{-s})s\gamma(s) \; .$$

Solving it, we obtain

$$\begin{aligned} W_1(s) &= -2\Gamma(s)\sum_{k=0}^{\infty}\frac{(p^{-s}+q^{-s})(s-k)\gamma(s-k)}{\prod_{m=0}^{k}(1-p^{-s+m}-q^{-s+m})} \\ &= -2\Gamma(s)\sum_{k=0}^{\infty}\frac{(p^{-s}+q^{-s})(s-k)Q(-2)}{\prod_{m=0}^{k}(1-p^{-s+m}-q^{-s+m})Q(s-k)} \; . \end{aligned} \tag{42}$$

We must now find the reverse Mellin transform of $W(s)$, that is

$$\widetilde{W}_1(z) = \frac{1}{2\pi i}\int_{-3/2-i\infty}^{-3/2+i\infty} \beta(s)\Gamma(s)z^{-s}ds \; . \tag{43}$$

Clearly, the Cauchy residue theorem is the simplest way to estimate the above integral. Note that the poles of the function under the integral are roots of $1 = p^{-s+k} + q^{-s+k}$ for $k = 0, 1, \ldots$. We need the following result, the detailed proof of which can be found in Jacquet [8] and/or Schachinger [20].

Lemma 7B. *Let s_k^ℓ for $k = 0, 1, \ldots$ and $\ell = \mathcal{Z} = 0, \pm 1, \pm 2, \ldots$ be solutions of*

$$p^{-s+k} + q^{-s+k} = 1$$

where $p + q = 1$ and s is complex.

(i) *For all $\ell \in \mathcal{Z} = \{0, \pm 1, \pm 2, \ldots\}$ and $k = 0, 1, \ldots$*

$$-1 + k \le \Re(s_k^\ell) \le \sigma_0 + k \tag{44}$$

where σ_0 is a positive solution of $1+q^{-s}=p^{-s}$. Furthermore,

$$\frac{(2\ell-1)\pi}{\log p} \le \Im(s_k^\ell) \le \frac{(2\ell+1)\pi}{\log p} \ .$$

(ii) *If $\Re(s_k^\ell) = -1+k$ and $\Im(s_k^\ell) \neq 0$, then $\log p/\log q$ must be rational. More precisely, if $\frac{\log p}{\log q} = \frac{r}{t}$ where* $\gcd(r,t)=1$ *for $r,t \in \mathcal{Z}$, then*

$$s_k^\ell = -1 + k + \frac{2\ell r\pi i}{\log p} \tag{45}$$

for all $\ell \in \mathcal{Z}$. ∎

Having the above, we can continue our investigation of $\widetilde{W}_1(z)$ given by (43). As in [4, 11, 22] we conclude that the dominating pole of the Cauchy integral is $s_0^0=-1$, and for $\log p/\log q = r/t$ (rational) we also must consider the poles $s_0^\ell = -1+2\ell r\pi i/\log p$. Actually, we can assess that the remaining poles for $k \neq 0$ contribute only $O(\log z)$ which is negligible compared to $O(z)$ contributed by $\widetilde{W}_2(z)$ (cf. Lemma 7A). We first consider $s_0^0=-1$ (irrespectively whether $\log p/\log q$ is rational or not). Thus, we must evaluate the residue of $f(s) = -2a(s)\cdot b(s)$ where

$$\begin{aligned} a(s) &= z^{-s} s\Gamma(s)(p^{-s}+q^{-s})(1-p^{-s}+q^{-s})^{-2} \ , \\ b(s) &= Q(-2)\prod_{m=1}^{\infty}(1-p^{-s+m}-q^{-s+m})^{-1} = Q(-2)/Q(s-1) \ . \end{aligned}$$

Using MAPLE, we find the following expansion for $a(s)$ around $s_0^0=-1$

$$a(s) = -\frac{z\log z}{h^2}\frac{1}{(s+1)^2} + \frac{z\log z}{(s+1)}\left(\frac{\log z}{2h^2} + \frac{1}{h^2} + \frac{\gamma-1}{h^2} + \frac{h_2}{h^3} - \frac{1}{h}\right) + O(1) \ .$$

Also, by Lemma 2.3 of [4] (cf. also Lemma 3.1 of [22]) we find that

$$b(s) = 1 - \alpha(s+1) + O((s+1)^2) \ ,$$

where α is defined in (33). Thus, taking the product of the above, and computing the coefficient of $(s+1)^{-1}$ we obtain the desired residue. This leads to the dominating term in (32) except the function $\delta_1(x)$ coming from the poles s_0^ℓ for $\ell \neq 0$.

Let now $\ell \neq 0$. If $\log p/\log q$ is irrational, then $\Re(s_0^\ell) = \sigma_0 > -1$, thus this pole only contributes $O(z^{-\sigma_0}\log z)$ and can be safely ignored in comparison with $O(z)$ coming from $\widetilde{W}_2(z)$.

Now we assume that $\log p/\log q = r/t$ (rational) for some integers $r,t \in \mathcal{Z}$. By Lemma 7B we know that $s_0^\ell = -1+2\pi i\ell r/\log p$ for $\ell \in \mathcal{Z}$. The residue of the function under the integral (43) becomes

$$\Delta_1(\log z) = \sum_{\substack{\ell=-\infty \\ \ell\neq 0}}^{\infty} \frac{\Gamma(s_0^\ell)Q(-2)e^{-2\pi i\ell r\log_p z}}{h^2(s_0^\ell)Q(s_0^\ell-1)} \ ,$$

where $h(s_0^\ell) = -p^{-s_0^\ell}\log p - q^{-s_0^\ell}\log q$. But $h(s_0^\ell) = h = -p\log p - q\log q$ since $p^{2\ell r i\pi/\log p} = q^{2\ell r i\pi/\log p} = 1$. Thus, $\Delta_1(x) = h^{-2}\delta_1(x)$, and after some algebra this completes the proof of Theorem 6. ∎

As a simple consequence of Theorem 6 we obtain the following corollary that completes the proof of Theorem 1 for the Poisson model, that is, it establishes the limiting distribution of the internal path length in the Poisson model.

Corollary 8. *For any $\varepsilon > 0$ the following holds*

$$\widetilde{L}\left(z, e^{t/\sqrt{\widetilde{V}(z)}}\right) e^{-t\widetilde{X}(z)/\sqrt{\widetilde{V}(z)}} = e^{t^2/2}\left(1 + O\left(1/m^{1/2+\varepsilon}\right)\right) \ , \tag{46}$$

i.e., the internal path length is normally distributed with parameters $\widetilde{X}(z)$ and $\widetilde{V}(z)$. Moreover, the moments $\widetilde{L}(z)$ converge to the appropriate moments of the normal distribution.

Proof. It follows directly from (30)–(32). ∎

The main problem that remains to be settled is to de-Poissonize Corollary 8, that is, to obtain results for the original Bernoulli model. This work is of tauberian type, and needs subtle arguments. In Section 3.3 we prove the following result.

Theorem 9. De-Poissonization. *Consider the Bernoulli model, and let $\widetilde{X}(m)$ and $\widetilde{V}(m)$ be the values of the mean and the variance of the Poisson model at $z = m$. Then, for any $\beta > 0$ and real t in the vicinity of zero*

$$L_m\left(e^{t/\sqrt{m}}\right)\exp\left(-\frac{\widetilde{X}(m)}{\sqrt{m}}t - \frac{\widetilde{V}(m) - m(\widetilde{X}'(m))^2}{2m}t^2\right) = 1 + O(1/m^{1/2-\beta}) \tag{47}$$

for large m. ∎

The rest is easy. Let $X_m = \widetilde{X}(m) = c_1 m\log m + O(m)$ and $V_m = \widetilde{V}(m) - m^2(\widetilde{X}'(m))^2 = c_2 m\log m + O(m)$. We shall prove below that $EL_m \sim X_m$, and $\mathrm{Var}\ L_m \sim V_m$, and furthermore that L_m converges to the normal distribution with parameters X_m and V_m which completes the proof of Theorem 1. More formally:

Lemma 10. *There exists $\varepsilon > 0$ such that*

$$L_m\left(e^{t/\sqrt{V_m}}\right)\exp\left(-tX_m/\sqrt{V_m}\right) = e^{t^2/2}\left(1 + O\left(1/m^{1/2+\varepsilon}\right)\right) \tag{48}$$

with

$$\mathrm{Var}\ L_m = V_m + O(m) = \widetilde{V}(m) - m(\widetilde{X}'(m))^2 + O(m) = c_2 m\log m + O(m)\ . \tag{49}$$

Proof. Observe that $\ell_m(t) = L_m(e^{t/\sqrt{V_m}})e^{-X_m t/\sqrt{V_m}}$ is the Laplace transform of a random variable $\ell_m = (L_m - X_m)/\sqrt{V_m}$. From Theorem 9 we know that for any real t the moment generating function $\ell_m(t)$ converges to $e^{t^2/2}$, that is, to the standard normal distribution. But, clearly the convergence is also true for any complex t since $|\ell_m(t)| \le \ell_m(|t|) + \ell_m(-|t|)$, so Theorem 9 implies also the convergence in moments. Hence

$$\begin{aligned}\lim_{m\to\infty} \mathrm{E}(\ell_m) &= \lim_{m\to\infty} \frac{\mathrm{E}L_m - X(m)}{\sqrt{V_m}} = 0 \\ \lim_{m\to\infty} \mathrm{Var}(\ell_m) &= \lim_{m\to\infty} \frac{\mathrm{Var}\ L_m}{V_m} = 1\ .\end{aligned}$$

Observe that (49) follows directly from the above and (47) (cf. [8, 18]). ∎

In passing, we note that the large deviation result of Theorem 1A follows directly from Lemma 10. Indeed, (48) implies that the moment generating function of $Z_m = (L_m - EL_m)/\sqrt{\mathrm{Var} L_m}$ is uniformly bounded in the vicinity of $t = 0$. Thus, as in Flajolet and Soria [6] we immediately conclude (8) of Theorem 1A(iii).

3.2 Asymptotic Solution of the Functional Equation

The ultimate goal of this section is to prove the two parts (i) and (ii) of Theorem 5 above, which justify the procedure used above in Theorem 6 to obtain equation (30). We found, however, working with $L(z,u)$ rather inconvenient due to the fact that its exponential growth makes it hard to control the function even in a small domain. Therefore, we introduce a polynomial kernel which is a function $f(z,u)$ defined as

$$f(z,u) = \frac{L(z,u)}{L'_z(z,u)} = \frac{L(z,u)}{L(quz,u)L(puz,u)}\ . \tag{50}$$

Note that $f^{-1}(z,u) = \frac{d}{dz}\log L(z,u)$.

In the first part of this section, we shall work with $f(z,u)$ as a function of z, so we often simplify the notation to $f(z)$. We also write $f'(z)$ to denote the derivative of $f(z)$ with respect to z. Clearly, the kernel $f(z)$ satisfies the following differential equation

$$f'(z) = 1 - \left(\frac{pu}{f(puz)} + \frac{qu}{f(quz)}\right) f(z)\ . \tag{51}$$

Since $L(0,u) = 1$ and $L(z,1) = e^z$ we also have $f(0,u) = 1$ and $f(z,1) = 1$.

Our first goal will then be the following technical theorems that are used to prove part (i) of Theorem 5.

Theorem 11A. *There exist a neighbourhood $\mathcal{U}(1)$ of $u = 1$, a constant ξ, and non-negative $a(u)$ and $A(u)$ such that for all real $x > \xi$ and $u \in \mathcal{U}(1)$ the following holds: $a(u)x^{1-\kappa(u)} \le f(x) \le A(u)x^{1-\kappa(u)}$ where $a(u), A(u) \to 1$ as $u \to 1$.* ∎

Theorem 11B. *For $\delta < 1$ and any $\beta > 0$ (with $\delta + \beta < 1$) there exists a neighbourhood $\mathcal{U}(1)$, a constant ξ, and a convex cone $\mathcal{C}(D, \delta)$ such that for a constant $B(u)$ the following is true $|f'(z)| \le B(u)|z|^{-\kappa(u)+\beta}$ for $|z| > \xi$.* ∎

Provided the above theorems are proved (cf. Sec. 3.2A-B), we immediately obtain the following corollary:

Corollary 11. *There exists a polynomial cone $\mathcal{C}(D, \delta)$, a neighbourhood $\mathcal{U}(1)$ of real positive $u = 1$, and a constant ξ such that, for $|z| > \xi$, uniformly in (z, u) for some $a(u)$ and $A(u)$ (independent of z) the following holds*

$$a(u)|z^{1-\kappa(u)}| \le |f(z, u)| \le A(u)|z^{1-\kappa(u)}| \tag{52}$$

with $a(u), A(u) \to 1$ as $u \to 1$.

Proof. Observe that

$$f(z) = f(\Re(z)) + \int_{\Re(z)}^{z} f'(x)\mathrm{d}x \ .$$

Thus, in view of Theorems 11A and 11B, we obtain $f(z) - f(\Re(z)) = \Theta(z^{-\kappa(u)+\delta+\beta})$ which leads to (52) provided $\delta + \beta < 1$. ∎

Having shown the above corollary, we can present a **proof of Theorem 5(i)**. Indeed, it suffices to observe that $\log L(z, u) = \int_0^z f^{-1}(x)\mathrm{d}x$.

Our next goal will be establishing another result presented below as Theorem 12 which is used to prove part (ii) of Theorem 5. For this we need an estimate on higher derivatives of $f(z, u)$ with respect to u. Define

$$f^{\langle l,k\rangle}(z, u) = \frac{\partial}{\partial z^l \partial u^k} f(z, u) \ .$$

We prove in Section 3.2C the following result.

Theorem 12. *For any $\beta > 0$ and $\delta < 1$, there exists $\mathcal{U}(1)$ and a convex cone $\mathcal{C}(D, \delta)$ such that for $(z, u) \in \mathcal{C}[D, \delta] \times \mathcal{U}(1)$: $f_k^{\langle l,k\rangle}(z, u) = O(z^{-\kappa(u)+\beta-l+1})$.* ∎

Observe now that the **proof of Theorem 5 (ii)** follows directly from the above. Indeed, the derivative of $\partial^k L(z, u)/\partial u^k$ with respect to z is equal to $\partial^k (f(z))^{-1}/\partial u^k$, thus

$$\frac{\partial}{\partial u^k} \log L(z, u) = \int_0^z \frac{\partial}{\partial u^k} f^{-1}(x, u)\mathrm{d}x \ .$$

But, the kth derivative of $f^{-1}(z, u)$ with respect to u is a sum of terms like $f^{\langle 0,k\rangle}(z)(f(z))^{-2}$ and terms like $(f^{\langle 0,1\rangle}(z))^k (f(z))^{-k-1}$ (cf. Sec. 3.2C). By Theorem 12 the former term is of order $O(z^{\kappa(u)-1+\beta})$ while the latter is $O(z^{\kappa(u)-1+k\beta})$. Since $\beta > 0$ is an arbitrary positive number, the proof of Theorem 5 is finally completed.

In view of the above, we can concentrate on establishing Theorems 11A, 11B and 12. The proof of these theorems will themselves depend on a set of five **basic facts** which we state and prove in the remainder of this subsection. In the next subsection 3.2A we prove 11A; in the following subsection 3.3B we prove 11B, and in the final subsection 3.2C we prove Theorem 12, thus completing all the steps of Theorem 5.

We now formulate our five **basic facts** used to prove Theorems 11 and 12.

Fact 1. *Consider a differential inequality of the form $f'(z) \le b(z) - g(z)f(z)$. Let $G(z)$ be the primitive function of $g(z)$. Then for any z and z_0*

$$f(z) \le f(z_0)e^{G(z_0)-G(z)} + \int_{z_0}^{z} b(x)e^{G(x)-G(z)}\mathrm{d}x \ .$$

Proof. Note that $(f(z)\exp(G(z)))' \le b(z)\exp(G(z))$. Thus, integrating this over (z_0, z) we establish the claim. ■

Fact 2. *For all a and d in a compact set such that $a, d > 0$ and $x \to \infty$ we have the following for some constant $A > 0$*

$$\begin{aligned}
\int_0^x y^{a-1}\exp\left(\frac{y^d - x^d}{d}\right)\mathrm{d}y &= x^{a-d}(1+O(1/x)) \ , \\
\int_0^x (x^a - y^a)\exp\left(\frac{y^d - x^d}{Ad}\right)\mathrm{d}y &= aAx^{a+1-2d}(1+O(1/x)) \ , \\
\int_0^x e^{ay^d}\mathrm{d}y &= \frac{e^{ax^d}}{adx^{d-1}}(1+O(1/x)) \ .
\end{aligned}$$

In particular,

$$\int_0^x \exp\left(\frac{y^d - x^d}{Ad}\right)\mathrm{d}y = Ax^{1-d}(1+O(1/x)) \ .$$

Proof. These asymptotic formulas seem to be well known, however, for completeness we provide a sketch of proofs for the first two formulas. Let

$$I(x) = \int_0^x y^{a-1}\exp\left(\frac{y^d - x^d}{d}\right)\mathrm{d}y \ .$$

Using the Taylor expansion of the form $y^b = x^b + (y-x)bx^{b-1}(1+O(x^{-1}))$, one obtains

$$\begin{aligned}
I(x) &= \int_0^x \Big(x^{a-1} + (y-x)(a-1)x^{a-2}(1+O(1/x))\Big)\exp\Big((y-x)x^{d-1}(1+O(1/x))\Big)\,dy \\
&= (1+O(1/x))\int_0^x x^{a-1}\exp(-yx^{d-1})\mathrm{d}y \\
&= x^{a-d}(1+O(1/x))\int_0^\infty e^{-y}\mathrm{d}y \ .
\end{aligned}$$

To prove the second formula, let

$$J(x) = \int_0^x (x^a - y^a) \exp\left(\frac{y^d - x^d}{d}\right) \mathrm{d}y \ .$$

Using again the same Taylor expansion as above, one obtains

$$\begin{aligned}
J(x) &= \int_0^x \Big((y-x)ax^{a-1}(1+O(1/x))\Big) \exp\Big((y-x)x^{d-1}(1+O(1/x))\Big)\, dy \\
&= ax^{a-1}(1+O(1/x)) \int_0^x y \exp(-yx^{d-1}) \mathrm{d}y \\
&= x^{a+1-2d}(1+O(1/x)) \int_0^\infty ye^{-y} \mathrm{d}y \ ,
\end{aligned}$$

and this completes the proof. ■

Fact 3. (i) *Let $y_m(x)$ be a sequence of nonnegative continuous functions of x satisfying the recurrence inequality $y_{m+1} \le y_m(x) \cdot F(x, y_m)$. If $F(x,y)$ is continuous in (x,y), and for all $y \in [0, y_0(0)]$ and $\varepsilon > 0$ we have $F(0,y) < 1 - \varepsilon$, then there exists a neighbourhood $\mathcal{U}(0)$ of $x = 0$ such that $y_m(x)$ uniformly decreases to zero with an exponential rate; more precisely $y_m(x) = O((1-\varepsilon)^m)$.*

(ii) *Under the same hypotheses as above, let now $y_{m+1}(x) \le B(x) + y_m(x) \cdot F(x, y_m)$. If for a small neighbourhood $\mathcal{U}(0)$ of $x = 0$ we have $F(x,y) < 1 - \varepsilon$ for $\varepsilon > 0$ and $\max\{B(x)(1 + 1/\varepsilon), y_0(x)\} < D$, where $D > 0$, then the sequence $y_m(x)$ is uniformly bounded for $x \in \mathcal{U}(0)$, that is, $y_m(x) \le D$.*

(iii) *Under the same hypothesis as in (i), the solution $y_m(x)$ of $y_{m+1}(x) \le \max\{y_m(x), y_m(x) \cdot F(x, y_m(x))\}$ uniformly decays to zero. Similarly, under the assumptions of (ii) $y_m(x)$ satisfying $y_{m+1}(x) \le \max\{y_m(x), B(x) + y_m(x)F(x, y_m(x))\}$ is uniformly bounded.*

Proof. These results are direct consequences of the fixed point theorem. For completeness, we show how to prove part (ii). Note that the recurrence has the following solution

$$y_{m+1} \le B(x)\left(1 + \sum_{k=1}^m F(x, y_k)\right) + y_0(x) \prod_{k=0}^m F(x, y_k) \ .$$

which implies (ii). Using mathematical induction one concludes part (iii) from (ii). ■

Fact 4. *Let us consider a neighbourhood $\mathcal{U}(1)$ of $u = 1$ such that $0 < \max\{pu, qu\} \le \nu < 1$. Define for $m = 1, 2, \ldots$ a sequence of increasing compact domains $\mathcal{D}_m$ as $\mathcal{D}_m = \{z : \ \Re(z) \in [\xi, \nu^{-m}]\}$ with $\xi > \nu$ (cf. Figure 2). Then,*

$$z \in \mathcal{D}_{m+1} - \mathcal{D}_m \qquad \Rightarrow \qquad puz, \ quz \in \mathcal{D}_m \ .$$

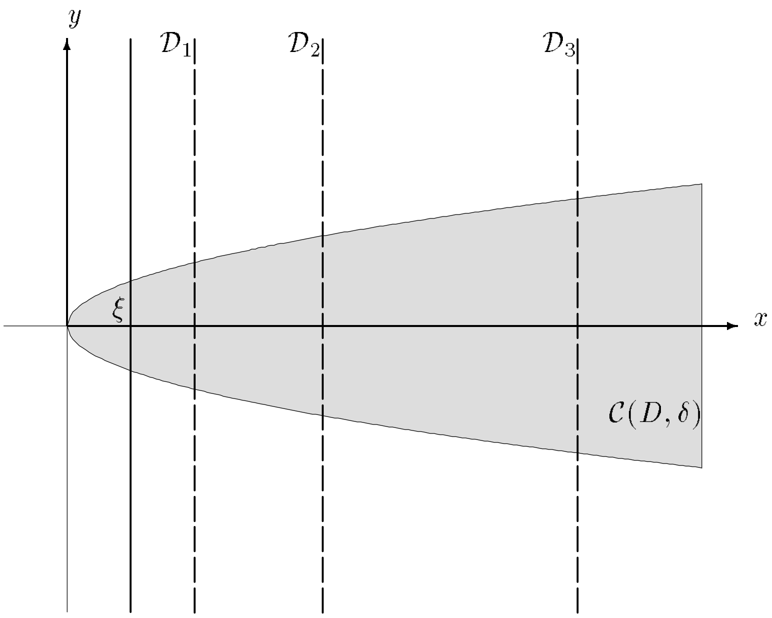

Figure 2: The convex cone $\mathcal{C}(D, \delta)$ and the domains $\mathcal{D}_m$ defined in Fact 4.

for all $z = x + yj$. ■

Proof. If $z \in \mathcal{D}_{m+1} - \mathcal{D}_m$, then $\xi\nu^{-m} \leq x \leq \xi\nu^{-(m+1)}$, and hence $\max\{pu, qu\}x \leq \xi\nu^{-m}$. ■

Fact 5. *Let $f(z)$ be an analytic function defined on a convex cone $\mathcal{C}(D, \delta)$ on which $f(z) = O(z^\alpha)$ for some α when $z \to \infty$. Let $f^{\langle k \rangle}(z) = \frac{\mathrm{d}^k}{\mathrm{d}z^k} f(z)$. Then, for all $k \geq 1$ there exists a smaller convex cone $\mathcal{C}(D', \delta)$ with $D' < D$ such that: $f^{\langle k \rangle}(z) = O(z^{\alpha - k\delta})$.*

Proof. By Cauchy formula (cf. [19])

$$f^{\langle k \rangle}(z) = \frac{k!}{2i\pi} \oint \frac{f(w)}{(w - z)^{k+1}} \mathrm{d}x$$

where the integration is along a complex circle around z. Fix $z \in \mathcal{C}(D, \delta)$, and let z also belong to the boundary of another (smaller) cone $\mathcal{C}(D', \delta)$ (i.e., $z = x + iy$ with $y = D'x^\delta$ where $D' < D$) as shown in Figure 3. We now consider a circle of integration that is the largest possible but still contained in the bigger cone $\mathcal{C}(D, \delta)$ (cf. Figure 3). Note that the circle we are working with has the radius smaller than $(D - D')|z|^\delta$. Thus, using the fact that $|f(w)| \leq B|w|^\alpha$ for some B and complex w, we finally obtain

$$|f^{\langle k \rangle}(z)| \leq (k!)B \frac{|z|^\alpha}{(|z|^\delta(D - D'))^k}$$

which completes the proof. ■

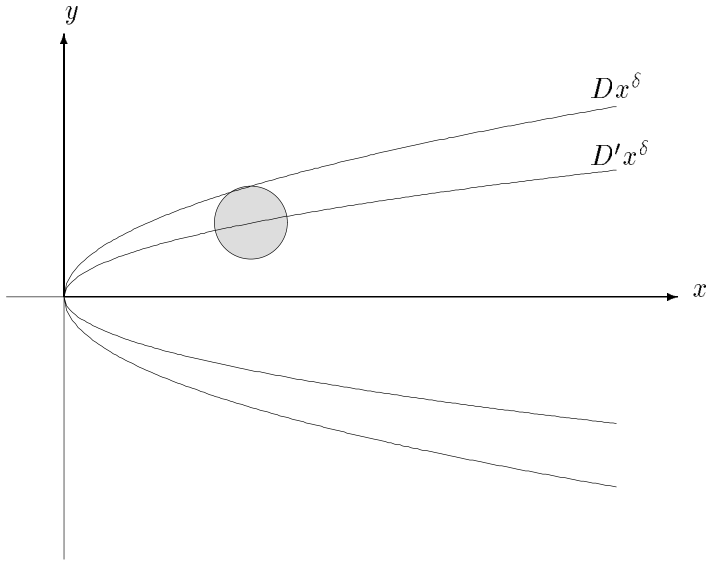

Figure 3: Illustration to the proof of Fact 5.

A. Analysis on the Real Line

We now prove Theorem 11A. For simplicity we write a and A for $a(u)$ and $A(u)$, and κ for $\kappa(u)$. Let A_m and a_m be upper and lower bounds for $f(x)x^{\kappa(u)-1}$ on the domain $\mathcal{D}_m$ (restricted to the real line) defined in Fact 4. If we prove that both a_m and A_m are bounded (with respect to m), then we establish our result. We concentrate on the upper bound since the lower bound mimics the proof of the upper bound.

We use induction with respect to m. Clearly, A_1 is bounded on $\mathcal{D}_1 \cap \mathcal{C}(D,\delta)$ since $L(z,u)$ is nonzero on this set for some $\mathcal{U}(1)$ due to compactness of $\mathcal{D}_1 \cap \mathcal{C}(D,\delta)$ and $L(z,1) = e^z$. So, let now x denote a real number belonging to $\mathcal{D}_{m+1} - \mathcal{D}_m$, and assume that $f(pux) \leq A_m(pux)^{1-\kappa}$ and $f(qux) \leq A_m(qux)^{1-\kappa}$. Observe now that (50) and the above implies $f'(x) \leq 1 - x^{\kappa-1}A_m^{-1}f(x)$. Thus by Fact 2 with $z_0 = \nu x$, we have

$$f(x) \leq f(\nu x)\exp\left(-\frac{(1-\nu^\kappa)x^\kappa}{\kappa A_m}\right) + \int_{\nu x}^{x} \exp\left(\frac{y^\kappa - x^\kappa}{\kappa A_m}\right) \mathrm{d}y \ .$$

Multiplying the above by $x^{\kappa-1}$, we obtain $A_{m+1} \leq \max\{A_m, A'_{m+1}\}$ where

$$A'_{m+1} \leq A_m \max_{x\in[\nu^{-m},\nu^{-m-1}]} \left\{\nu^{-\kappa}\exp(-\frac{(1-\nu^\kappa)x^\kappa}{\kappa A_m}) + \int_{\nu x}^{x} \exp\left(\frac{y^\kappa - x^\kappa}{\kappa A_m}\right) \mathrm{d}y\right\} \ .$$

We proceed in two steps. First, we let $v_m = A_m\nu^{\kappa m}$, and prove that v_m exponentially decays to zero. Secondly, we prove that A_m is an increasing sequence uniformly bounded from the above.

We know, by Fact 2, that

$$\int_{\nu x}^{x} \exp\left(\frac{y^\kappa - x^\kappa}{\kappa}\right) \mathrm{d}y = x^{1-\kappa}\left(1 - e^{-x^\kappa(1-\nu^\kappa)}\right)(1 + O(1/x)) = x^{1-\kappa}\eta(x) ,$$

where $\eta(x) = 1 + O(1/x)$. Thus

$$A_{m+1} \le \max\left\{A_m, A_m \nu^{-\kappa} \exp\left(-\frac{\nu^{-\kappa m}}{A_m \kappa} + A_m \eta\left(\frac{\nu^{-m}}{A_m^{1/\kappa}}\right)\right)\right\} . \tag{53}$$

which can be reduced to

$$v_{m+1} \le v_m F(u, v_m) \tag{54}$$

where $F(u, v_m) = \max\{\nu^\kappa, \exp(-(v_m\kappa)^{-1}) + \nu^\kappa \eta(v_m^{-1/\kappa})$. Note that $F(\cdot,\cdot)$ is of the form already discussed in Fact 3. Clearly, there exists m_0 such that $F(1, v_{m_0}) < 1$ for $m > m_0$, which implies exponential decay of v_m. To see this it is enough to observe that $F(x, 0) = \nu^\kappa < 1$ and $F(\cdot,\cdot)$ is a continuous function (for $u = 1$ one also has $A_m = 1$, as desired). Thus by (i) of Fact 3, the exponential decay of v_m follows.

Now we return to (53) to get

$$A_{m+1} \le A_m \max\left\{1, \left(\nu^{-\kappa} \exp(-(v_m\kappa)^{-1}) + \eta(v_m^{-1/\kappa})\right)\right\} ,$$

and then $A_{m+1} \le (1 + O(v_m^{1/\kappa}))A_m$. Thus

$$A_{m+1} \le A_0 \prod_{j=0}^{m}\left(1 + O(v_j^{1/\kappa})\right) < \infty , \tag{55}$$

which proves that A_m are uniformly bounded by A_∞. Clearly, as $\mathcal{U}(1)$ becomes smaller and smaller (i.e., $u \to 1$) the constant $A_\infty(u)$ tends to 1.

The lower bound can be shown along the same lines. In particular, we derive $a_{m+1} \ge \min\{a_m, a'_m\}$ where

$$a'_{m+1} \ge a_m \min_{x \in [\nu^{-m}, \nu^{-m-1}]} \left\{\nu^{-\kappa} \exp\left(-\frac{(1-\nu^\kappa)x^\kappa}{\kappa a_m}\right) + \int_{\nu x}^{x} \exp\left(\frac{y^\kappa - x^\kappa}{\kappa a_m}\right) \mathrm{d}y\right\} .$$

Since $a_m \nu^{\kappa m} \le v_m$, we get $a_{m+1} \ge (1 - O(v_m))a_m$. This gives the desired result. We also have the uniform lower bound $a_\infty(u)$ for a_m tending to 1 as $u \to 1$. ■

B. Analysis on the Complex Plane

Now we prove Theorem 11B which extends Theorem 11A to the complex plane (more precisely: to a convex cone). We need the following preliminary lemma.

Lemma 13. *Let A, a and $\mathcal{U}(1)$ be defined as in Theorem 11A, and let $z = x + jy$ be such that $z \in \mathcal{C}(D, \delta)$. If $|f'(z)| \le Bx^{-\kappa(u)+\beta}$ for some $\beta > 0$, and $|y| \le x^{1-\beta}$ with $\delta < 1 - \beta$, then*

$$
\begin{aligned}
(a - Bx^{\beta}|\frac{y}{x}|)x^{1-\kappa(u)} \le \quad & |f(z)| && \le (A + Bx^{\beta}|\frac{y}{x}|)x^{1-\kappa(u)} \;, && (56)\\
& |\Im(f(z))| && \le Bx^{1-\kappa(u)+\beta}|\frac{y}{x}| \;, && (57)\\
x^{1-\kappa(u)}(a - Bx^{\beta}|\frac{y}{x}|) \le \quad & \Re(f(z)) && \le x^{1-\kappa(u)}(A + Bx^{\beta}|\frac{y}{x}|) \;, && (58)\\
& |\Im(f^{-1}(z))| && \le \frac{Bx^{\kappa(u)-1}x^{\beta}|\frac{y}{x}|}{(a - Bax^{\beta}|\frac{y}{x}|)^2} \;, && (59)\\
\frac{x^{\kappa(u)-1}(a - Bx^{\beta}|\frac{y}{x}|)}{(A + Bx^{\beta}|\frac{y}{x}|)^2 + (Bx^{\beta}|\frac{y}{x}|)^2} \le \quad & \Re(f^{-1}(z)) && \le \frac{x^{\kappa(u)-1}(A + Bx^{\beta}|\frac{y}{x}|)}{(a - Bax^{\beta}|\frac{y}{x}|)^2} \;. && (60)
\end{aligned}
$$

Proof. The proof is a straightforward application of $f(z) = f(x) + \int_x^z f'(y)\mathrm{d}y$ and the following estimate $|\int_x^z f'(y)\mathrm{d}y| \le |z - x| \max_{y\in[x,z]}\{f'(y)\}$ (cf. [19]). For example, the left-hand side of (58) can be proved as follows:

$$
\Re(f(z)) \ge f(x) - \left|\int_x^z f'(y)\mathrm{d}y\right| \ge ax^{1-\kappa} - |y|Bx^{-\kappa+\beta} \;.
$$

The last two inequalities are direct consequences of the previous ones and

$$
\frac{1}{f(z)} = \frac{\Re(f(z))}{\Re^2(f(z)) + \Im^2(f(z))} - i\frac{\Im(f(z))}{\Re^2(f(z)) + \Im^2(f(z))} \;.
$$

This completes the proof. ∎

Equipped in this result, we proceed to the proof of Theorem 11B. The proof is by induction over the domains $\mathcal{D}_m$ defined in Fact 4. We have already proved in Theorem 11A that in a neighbourhood $\mathcal{U}(1)$ there exist $a(u)$ and $A(u)$ satisfying Theorem 11A such that $a(u), A(u) \to 1$ as $u \to 1$. We further denote these quantities as a and A. We consider a convex cone such that the domains $\mathcal{D}_m$ inside such a cone form a compact set (cf. Fig.2). Finally, we assume throughout the proof that $\delta < 1 - \beta$.

Let B_m be an upper bound of $|f'(z)|x^{\kappa(u)-\beta}$ where $x = \Re(z)$. If we prove that B_m are uniformly bounded, then our theorem is true, since $|z| = O(\Re(z))$ in our convex cone. Clearly, B_1 is bounded, so we proceed as before by induction. Below, we write $x = \Re(z)$ and $y = \Im(z)$.

Let $g(z) = \frac{pu}{f(puz)} + \frac{qu}{f(quz)}$ so that $f'(z) = 1 - g(z)f(z)$, and let $G(z)$ be the primitive function of $g(z)$. Using Fact 2, we just have (with $z_0 = \nu z$)

$$
f(z) = f(\nu z)\exp(G(\nu z) - G(z)) + \int_{\nu}^{1} \exp(G(tz) - G(z))z\mathrm{d}t \;.
$$

Differentiating the above, and after some elementary algebra, we obtain

$$
f'(z) = r(z)\exp(G(\nu z) - G(z)) + \int_{\nu}^{1} (g(tz) - g(z))\exp(G(tz) - G(z))z\mathrm{d}t \qquad (61)
$$

with $r(z) = (\nu g(\nu z) - g(z))f(\nu z) + \nu f'(\nu z) + (1-\nu)$.

Our next task is to estimate various terms in (61) to get a suitable recurrence for B_m. This is possible since by Fact 4 $putz$ and $qutz$ belong to $\mathcal{D}_m$ for all $t \in [\nu, 1]$ if $z \in \mathcal{D}_{m+1}$, and thus we can invoke the induction hypothesis.

Let us start with an estimate for $g(tz)$. Using (59) and (60) of Lemma 13, we immediately obtain

$$\begin{aligned} \Im(g(tz)) &\leq (tx)^{\kappa-1+\beta} B_m |\frac{y}{x}| F_1(a, A, B_m) , \\ \Re(g(tz)) &\geq (tx)^{\kappa-1}(A - B_m x^{\beta}|\frac{y}{x}|)| F_1(a, A, B_m) , \end{aligned}$$

where $F_1(a, A, B_m) = a^{-2}(1-B_m)^{-2}$ and $F_2(a, A, B_m) = 1/((A+B_m)^2 + B_m^2)$ are rational functions of a, A, B_m such that $F_1(1,1,0) = F_2(1,1,0) = 1$. More precisely, we have $\lim_{u\to 1} F_1(a(u), A(u), B_m(u)) = \lim_{u \to 1} F_2(a(u), A(u), B_m(u)) = 1$.

We now estimate $\Re(G(z) - G(tz))$. Observe that for $z = x + iy$

$$\begin{aligned} \Re(G(z) - G(tz)) &= \int_t^1 \Re(g(z\theta)z)\mathrm{d}\theta = \int_t^1 (x\Re(g(\theta z)) - y\Im(g(\theta z))\,\mathrm{d}\theta \\ &\geq x^{\kappa}\left((a - B_m)F_2(a, A, B_m) - x^{-1}|y|B_m F_1(a, A, B_m)\right)\int_t^1 \theta^{\kappa-1}\mathrm{d}\theta \\ &\geq x^{\kappa}\frac{1-t^{\kappa}}{\kappa}F_3(a, A, B_m) \end{aligned} \tag{62}$$

where $F_3(a, A, B_m) = (a - B_m)F_2(a, A, B_m) - B_m F_1(a, A, B_m)$ being a rational function of a, A, B_m such that $F_3(1,1,0) = 1$.

Now, we are ready to give estimate on the terms of the right hand side in (61). We start with an estimate of $r(z)$. We refer to the previous estimate of $g(tz)$ to get the following:

$$|\nu g(\nu(z)) - g(z)| \leq (1 + \nu^{\kappa-1})x^{\kappa-1}F_1(a, A, B_m).$$

We also have

$$|f(\nu z)| \leq x^{1-\kappa}(A + x^{\beta}|\frac{y}{x}|B_m) ,$$

hence

$$|(\nu g(\nu(z)) - g(z))f(\nu z)| \leq (1 + \nu^{\kappa-1})F_4(a, A, B_m)$$

with $F_4(a, A, B_m) = F_1(a, A, B_m)(A + B_m)$. We easily check that $F_4(1,1,0) = 1$. Since by the hypothesis $|f'(\nu z)| \leq (\nu x)^{-\kappa+\beta}B_m$, hence $|r(z)| \leq B_m(\nu x)^{-\kappa+\beta} + R(u, B_m)$ with $R(u, B_m) = (1+\nu^{\kappa-1})F_4(a, A, B_m) + 1 - \nu$.

Estimating the integral on the right hand side of (61) is more intricate to deal with, and needs careful computations. First, we estimate $g(tz) - g(z)$ under the integral in (61). We take advantage of the identity $g(tz)-g(z)=\int_t^1 g'(\theta z)z\mathrm{d}\theta$. We formally have $-g'(z)=\frac{(pu)^2f'(puz)}{f^2(puz)}+\frac{(qu)^2f'(quz)}{f^2(quz)}$. We refer to the estimate $|f'(z)|\le B_m x^{-\kappa+\beta}$ and $|f(z)|^{-1}\le x^{\kappa-1}F_1(a,A,B_m)$ valid for $z\in\mathcal{D}_m$ to get

$$|g(tz)-g(z)|\le B_m x^{\kappa-\beta+1}\frac{1-t^{\beta+\kappa-2}}{\beta+\kappa+2}((pu)^{\kappa+\beta}+(qu)^{\kappa+\beta})(F_1(a,A,B_m))^2\ . \tag{63}$$

We observe that $(pu)^{\kappa+\beta}+(qu)^{\kappa+\beta}=\mu<1$. Using the estimate on $\Re(G(tz)-G(z))$ established in (62), one shows

$$\begin{aligned}\left|\int_\nu^1(g(zt)-g(z))\exp(G(zt)-G(z))z\mathrm{d}t\right| &\le \\ B_m\mu(F_1(a,A,B_m))^2\int_{\nu x}^{x}\frac{x^{\beta+\kappa-1}-y^{\beta+\kappa-1}}{\beta+\kappa-1} &\cdot \exp\left(\frac{y^\kappa-x^\kappa}{\kappa}F_3(a,A,B_m)\right)\mathrm{d}y\ ,\end{aligned}$$

and then by Fact 2

$$\left|\int_\nu^1(g(zt)-g(z))\exp(G(zt)-G(z))z\mathrm{d}t\right|\le\mu B_m x^{\beta-\kappa}(1+O(1/x))F_5(a,A,B_m))$$

with $F_5(a,A,B_m)=(F_1(a,A,B_m))^2(F_3(a,A,B_m))^{-1}$.

Putting everything together we finally obtain

$$\begin{aligned}|f'(z)| &\le R(u,B_m)\exp(-(1-\nu^\kappa)x^\kappa F_3(a,A,B_m)) \\ &+ B_m x^{-\kappa+\beta}\left(\mu F_5(a,A,B_m)+\nu^{-\kappa+\beta}\exp(-(1-\nu^\kappa)x^\kappa F_3(a,A,B_m))\right)\ .\end{aligned}$$

We use this global estimate to carry out the recurrence for B_{m+1} which becomes

$$B_{m+1}\le\max\{B_m,B'_{m+1}\} \tag{64}$$

with

$$\begin{aligned}B'_{m+1} &\le B_m\left(\mu F_5(a,A,B_m)+\nu^{-\kappa+\beta}\exp(-(1-\nu^\kappa)\nu^{-\kappa m}F_3(a,A,B_m))\right)+ \\ &+ R(u,B_m)\exp(-(1-\nu^\kappa)x^\kappa F_3(a,A,B_m))\ .\end{aligned} \tag{65}$$

Let us choose (u,B_m) in a compact neighbourhood of $(1,0)$ such that for example $F_3(a,A,B_m)>1/2$. We can rewrite (65) as:

$$B'_{m+1}\le B_m(\mu F_5(a,A,B_m)+\xi_m)+\xi'_m$$

where $\xi_m=\exp(-(1-\nu^\kappa)\nu^{\kappa m}/2)$ and $\xi'_m=\nu^{\kappa m}\xi_m$ are two sequences tending to 0 as $m\to\infty$. For any $\varepsilon>0$ we can take m_0 large enough such that for all $m\ge m_0$:

$$B'_{m+1}\le B_m(\mu F_5(a,A,B_m)+\varepsilon)+\varepsilon$$

In summary, we have

$$B_{m+1} \leq \max\{B_m, B_m \cdot F_6(u, B_m) + \varepsilon\}$$

with $F_6(1,0) < \mu + \varepsilon < 1$. Since $B_m(1) = 0$ and $\varepsilon > 0$ can be made as small as needed in a neighbourhood $\mathcal{U}(1)$ of $u = 1$, by Fact 3(iii) the sequence B_m is uniformly bounded. Note that $B_\infty(u)$ is continuous and that $B_\infty(1) = 0$. This completes the proof of Theorem 11B.

In passing, we observe that we have also proved that $f^{-1}(z) = O(z^{\kappa(u)-1})$ and $g(z) = O(z^{\kappa(u)-1})$.

C. Finishing the Proof of Theorem 5.

In this subsection we establish Theorem 12. We start with a simple result.

Lemma 14. *Let $f^{\langle l\rangle}(z)$ be the lth derivative of $f(z)$. Then, for any $\beta > 0$ and δ (with $\beta + \delta < 1$), there exists $\mathcal{U}(1)$, a constant ξ, and a convex cone $\mathcal{C}(D, \delta)$ such that $f^{\langle l\rangle}(z) = O(z^{-\kappa(u)+\beta-l+1})$ for $|z| > \xi$.*

Proof. The corollary was already proved for $l = 0$ and $l = 1$. For arbitrary l, applying Fact 5 to $f'(z)$ yields $f^{\langle l\rangle}(z) = O(z^{-\kappa(u)+\beta-(l-1)\delta})$. Set now $\delta > 1-\varepsilon$ (with $\varepsilon > 0$) and $\beta' = \beta+(l-1)\varepsilon$ to prove the corollary due to arbitrariness of β. ■

Now, we are ready to prove Theorem 12 that is repeated below for the reader's convenience. We recall, that we use the following notation

$$f^{\langle l,k\rangle}(z, u) = \frac{\partial}{\partial z^l \partial u^k} f(z, u) \ .$$

Theorem 12. *For any $\beta > 0$ and $\delta < 1$, there exists $\mathcal{U}(1)$ and a convex cone $\mathcal{C}(D, \delta)$ such that for $(z, u) \in \mathcal{C}[D, \delta] \times \mathcal{U}(1)$: $f_k^{\langle l,k\rangle}(z, u) = O(z^{-\kappa(u)+\beta-l+1})$.*

Proof. Observe that it suffices to show that $f^{\langle 0,k\rangle}(z, u) = O(z^{1-\kappa(u)+\beta})$ since by Fact 5 we obtain $f^{\langle l,k\rangle}(z, u) = O(z^{-\kappa(u)+\beta-l+1})$ for any $\beta > 0$ as in the proof of Lemma 14.

To prove $f^{\langle 0,k\rangle}(z, u) = O(z^{1-\kappa(u)+\beta})$ we proceed by double induction: one with respect to k and the other with respect to increasing domains $\mathcal{D}_m$ as described in Fact 4.

For $k = 0$, our claim is true by Theorem 11B (in fact, in this case $\beta = 0$). So, we assume now that our theorem is true for all $i < k$ and all $l \geq 0$.

After taking the derivative with respect to u, our basic functional equation (51) is transformed into

$$\frac{\partial^2 f(z, u)}{\partial z \partial u} \quad = \quad f^{\langle 1,1\rangle}(z, u) = \left(\frac{puf^{\langle 0,1\rangle}(puz, u)}{(f(puz, u))^2} + \frac{quf^{\langle 0,1\rangle}(quz, u)}{(f(quz, u))^2} \right) f(z, u)$$

$$
\begin{aligned}
&+ \quad f(z,u)\left(\frac{p^2uzf^{\langle 1,0\rangle}(puz,u)-pf(puz,u)}{(f(puz,u))^2}+\frac{q^2uzf^{\langle 1,0\rangle}(quz,u)-qf(quz,u)}{(f(puz,u))^2}\right)\\
&- \quad (\frac{pu}{f(puz,u)}+\frac{qu}{f(quz,u)})f^{\langle 0,1\rangle}(z,u)\ .
\end{aligned}
$$

This formula suggests the following general scheme

$$
f^{\langle 1,k\rangle}(z,u)=b_k(z)+a_k(z)-g(z)f^{\langle 0,k\rangle}(z,u) \tag{66}
$$

where $a_k(z)$ being of the form $\frac{R_1(z)}{(f(puz))^{k+1}}+\frac{R_2(z)}{(f(puz))^{k+1}}$ and $R_i(z)$'s are polynomials of degree $k+1$ with terms of the form as $z^l f^{\langle l,i\rangle}(z,u)$ at points z, puz and quz for i and l strictly smaller than k. Furthermore,

$$
b_k(z)=\left(\frac{pu}{(f(puz,u))^2}f^{\langle 0,k\rangle}(puz,u)+\frac{qu}{(f(quz,u))^2}f^{\langle 0,k\rangle}(quz,u)\right)f(z,u)\ .
$$

We can easily estimate $a_k(z)$ and $b_k(z)$. For the former, we just note that by the induction assumption for $i\le k-1$ we have $z^l f^{\langle l,i\rangle}(z,u)=O(z^{1-\kappa(u)+\beta})$, hence $a_k(z)=O(1)$. For the latter, we use the induction with respect to the increasing domains $\mathcal{D}_m$ as in the previous proofs. Thus, after elementary calculus we obtain $|b_k(z)|\le\mu\alpha(u)x^{\beta}$ due to $(pu)^{\kappa(u)+\beta}+(qu)^{\kappa(u)+\beta})=\mu<1$, where $\alpha(u)\to 1$ as $u\to 1$.

Let now C_m be the upper bound on $|f^{\langle 0,k\rangle}(z,u)z^{\kappa(u)-\beta-1}|$ over the domains $\mathcal{D}_m$. As before, we shall prove that C_m are uniformly bounded for all m which will complete the proof. To develop a recursion for C_m we apply Fact 2 to the differential equation (66). One derives

$$
f^{\langle 0,k\rangle}(z,u)=f(\nu z,u)e^{G(\nu z)-G(z)}+\int_{\nu z}^{z}(a_k(x)+b_k(x))e^{G(x)-G(z)}\mathrm{d}x\ . \tag{67}
$$

In sequel, we estimate the terms of the above equation in order to obtain a recurrence on C_m.

From (62) we have

$$
|\exp(G(\nu z)-G(z))|\le\exp\left(-x^{\kappa(u)}(1-\nu^{\kappa(u)})F_3^{-1}(a,A,B_m)\right)
$$

where $F_3(a,A,B_m)$ was defined below (62). Furthermore, by Fact 2 we also have the following

$$
\begin{aligned}
\int_{\nu z}^{z}a_k(x)e^{G(x)-G(z)}\mathrm{d}x &\le \rho_1(u)x^{1-\kappa(u)}\ ,\\
\int_{\nu z}^{z}b_k(x)e^{G(x)-G(z)}\mathrm{d}x &\le C_m\mu\alpha_2(u)x^{\beta+1-\kappa(u)}(1+O(1/x))
\end{aligned}
$$

where $\rho_1(u),\alpha_1(u)<\infty$ do not depend on z, and $\alpha_1(u)\to 1$ as $u\to 1$.

Putting everything together, we finally obtain

$$
C_{m+1}\le\max\{C_m,C'_m\}
$$

where

$$C'_{m+1} \le \rho_1(u)\nu^{m\beta} + C_m\left(\mu\alpha_2(u) + \nu^{1+\beta-\kappa(u)}\exp\left(-\rho_2(u)\nu^{-\kappa(u)m}(1-\nu^{\kappa(u)})\right)\right)(1+O(1/x))$$

for some $0 < \rho_2(u) < \infty$ and $\mu < 1$. Clearly, the last recurrence can be re-written as

$$C_{m+1} \le \max\{C_m, c(d) + F(d, C_m)C_m\}$$

for some functions, $c(\cdot)$ and $F(\cdot,\cdot)$ such that $F(0,y) < 1$ and $c(d) < \infty$ where d is the diameter of $\mathcal{U}(1)$. As in Section 3.2B, we can now use Fact 3(iii) to show that C_m are uniformly bounded for some $\mathcal{U}(1)$, and this completes the proof of Theorem 12. ■

3.3 De-Poissonization

In this subsection, we prove the de-Poissonization Theorem 9 which will complete the proof of Theorem 1. We use the Cauchy formula

$$L_m(u) = \frac{m!}{2i\pi}\oint L(z,u)\frac{\mathrm{d}z}{z^{m+1}} , \tag{68}$$

where the integration is over a circle with the center at the origin and radius m. We split this circle into two non-overlapping arcs $\mathcal{A}_m(\theta) \cup \overline{\mathcal{A}}_m(\theta) = \{z : |z| = m\}$ where $w = me^{i\theta}$ is a point of the circle as shown in Figure 4. (We use standard notation for polar coordinates, that is, for $z = \rho e^{i\theta}$ we set $\rho = |z|$ and $\theta = \arg(z)$.) More precisely, for $w = me^{i\theta}$ $(\theta > 0)$ the arc $\mathcal{A}_m(\theta)$ is defined as $\mathcal{A}_m(\theta) = \{v : |v| = m \;\&\; -\theta \le \arg(v) \le \theta\}$, and $\overline{\mathcal{A}}_m(\theta) = \{v : |v| = m\} - \mathcal{A}_m(\theta)$ (cf. Figure 4).

Our proof of the de-poissonization is based on the ideas already used in Jacquet [8], and Rais *et al.* [17]. Namely, we shall show that the main contribution of the Cauchy formula (68) comes from the integration over the arc $\mathcal{A}_m(\theta)$ while the remaining contribution over $\overline{\mathcal{A}}_m(\theta)$ is exponentially small.

To proceed along these lines, we need upper bounds for $L(z,u)$ over the arcs $\mathcal{A}_m(\theta)$ and $\overline{\mathcal{A}}_m(\theta)$ for some $w = me^{i\theta}$ on the circle of integration. In Theorem 5 we have already developed a suitable bound over the first arc, so we need only a bound for $|L(z,u)|$ for $z \in \overline{\mathcal{A}}_m(\theta)$. We denote such a bound by $\overline{L}(w,u)$ for $w = me^{i\theta}$, that is, $\max_{v\in\overline{\mathcal{A}}_m(\theta)}\{|L(v,u)|\} \le \overline{L}(w,u)$. In passing, we observe that for $w = x$ real (i.e., $\theta = 0$) the arc $\overline{\mathcal{A}}_m(0)$ coincides with the whole circle of integration (of radius m), and $\overline{L}(x,u) = L(x,u)$. Also, for any complex w we have $\overline{L}(w,1) = |e^w|$.

Before we formulate our result, we must introduce some new notation. Let $\alpha(\theta)$ and $\mu(u)$ be two positive functions of θ and u in a neighbourhoods $\mathcal{U}_\theta(0)$ and $\mathcal{U}_u(1)$ respectively of 0

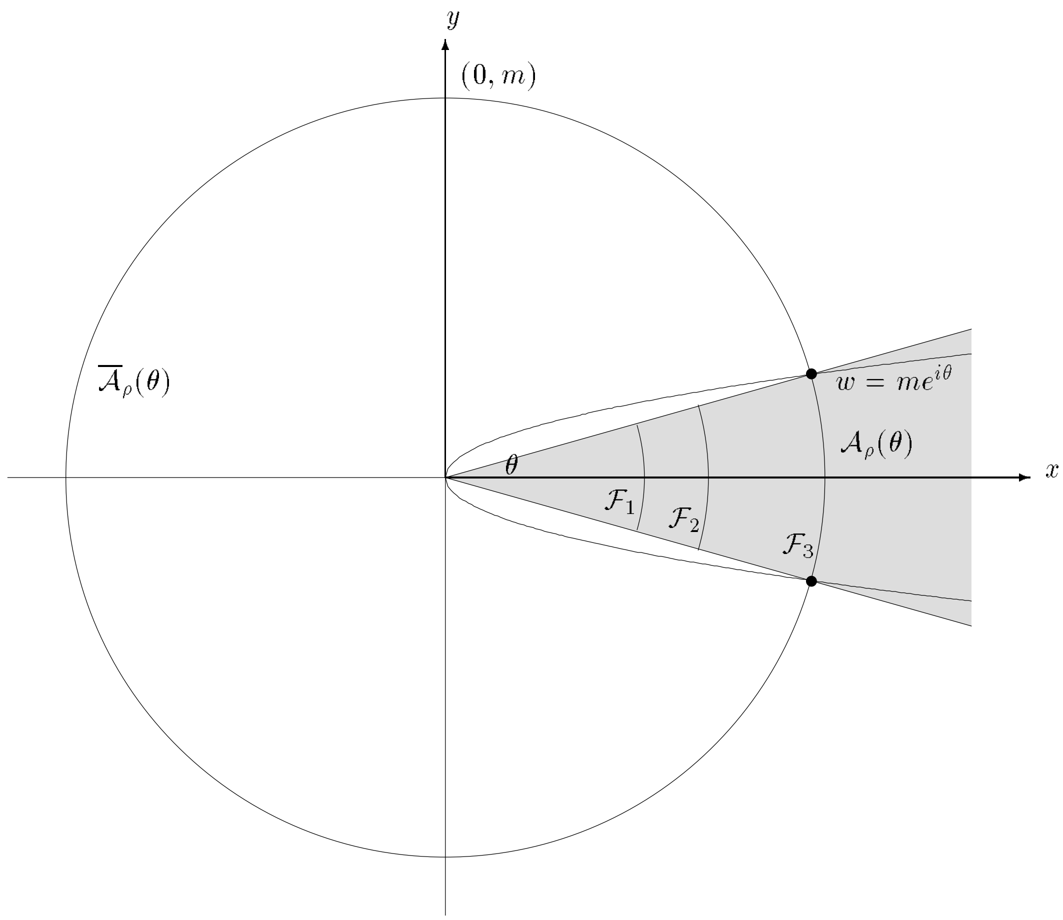

Figure 4: Illustration of de-Poissonization.

and 1. Actually, we want $\alpha(0) = 1$ and $\mu(1) = 0$, and, if $\theta \neq 0$, then $\cos\theta < \alpha(\theta) < 1$, and if $u \neq 1$, then $1 + \mu(u) > \kappa(u)$.

We prove the following theorem that provides the desired bound for $L(z, u)$ over the arc $\overline{\mathcal{A}}_m(\theta)$.

Theorem 15. *There exist neighbourhoods $\mathcal{U}_\theta(0)$ of $\theta = 0$ and $\mathcal{U}_u(1)$ of $u = 1$ such that for some $w = \rho e^{i\theta}$*

$$\max_{v \in \overline{\mathcal{A}}_\rho(\theta)} \{|L(v, u)|\} \leq \overline{L}(\rho e^{i\theta}, u) \leq \alpha(\theta) \exp\left(\alpha(\theta)\rho^{1+\mu(u)}\right) \ . \tag{69}$$

when $\rho \to \infty$, provided $\alpha(\theta) = 1 - \theta^2/4$ and $1 + \mu(u) > \kappa(u)$.

Proof. The proof is by induction over increasing domains as already discussed in Fact 4. However, since we work now with polar coordinates we redefine them. Let $\mathcal{F}_m = \{z = \rho e^{i\theta} : \ \theta \in \mathcal{U}_\theta(0) \quad \& \quad \rho \in [\xi\rho_0, \nu^{-m}\rho_0]\}$ such that $0 < \max\{pu, qu\} \leq \nu < 1$ and $\xi > \nu$ (cf.

Figure 4), Moreover, we require $\rho_0 > 1$ such that for all $\rho > \xi\rho_0$ we have $\alpha(\theta)\exp(\rho\alpha(\theta)) > \exp(\rho\cos\theta)$ for $\theta \neq 0$ and for some (small) $\mathcal{U}_\theta(0)$ and $\mathcal{U}_u(1)$.

Take now such a small $\mathcal{U}_u(1)$ that $\overline{L}(\rho e^{i\theta}, u) < \alpha(\theta)\exp\left(\alpha(\theta)\rho^{1+\mu(u)}\right)$ holds for $u \neq 1$ or $\theta \neq 0$ (i.e., for $z \in \mathcal{F}_1$). This is possible due to our choice of ρ_0. Now, we assume (69) is true for all $m' \leq m$ and we prove it also holds for $m+1$. Let $z \in \mathcal{F}_{m+1} - \mathcal{F}_m$. From our basic functional equation we have for $z_0 = \rho_0 e^{i\theta}$

$$L(z,u) = L(z_0,u) + \int_{z_0}^{z} L(pux,u)L(qux,u)\mathrm{d}x \ ,$$

which for $w = \rho e^{i\theta}$ and $w_0 = \rho_0 e^{i\theta}$ translates into

$$\overline{L}(\rho e^{i\theta}, u) \leq \overline{L}(\rho_0 e^{i\theta}, u) + \int_{\rho_0}^{\rho} \overline{L}(puxe^{i\theta}, u))\overline{L}(quxe^{i\theta}, u)\mathrm{d}x \ .$$

Since pux and qux both belong to $\mathcal{F}_m$ we have by the induction hypothesis

$$\overline{L}(puxe^{i\theta}, u))\overline{L}(quxe^{i\theta}, u) \leq \alpha^2(\theta)\exp\left(\alpha(\theta)\left((pu)^{1+\mu(u)} + (qu)^{1+\mu(u)}\right)x^{1+\mu(u)}\right) \ .$$

Observe now that due to $1+\mu(u) > \kappa(u)$ we have $(pu)^{1+\mu(u)} + (qu)^{1+\mu(u)} < 1$ by the definition of $\kappa(u)$. Therefore,

$$\begin{aligned}
\overline{L}(\rho e^{i\theta}, u) &\leq \overline{L}(\rho_0 e^{i\theta}, u) + \int_{\rho_0}^{\rho} \alpha^2(\theta)\exp\left(\alpha(\theta)x^{1+\mu(u)}\right)\mathrm{d}x \\
&\leq \overline{L}(\rho_0 e^{i\theta}, u) + \int_{\rho_0}^{\rho} \alpha^2(\theta)\left(\frac{x}{\rho_0}\right)^{\mu(u)}\exp\left(\alpha(\theta)x^{1+\mu(u)}\right)\mathrm{d}x \ ,
\end{aligned}$$

which, after integration by parts, leads to

$$\begin{aligned}
\overline{L}(\rho e^{i\theta}, u) &\leq \overline{L}(\rho_0 e^{i\theta}, u) + \frac{\alpha(\theta)}{\rho_0^{\mu(u)}(1+\mu(u))}\left(\exp\left(\alpha(\theta)\rho^{1+\mu(u)}\right) - \exp\left(\alpha(\theta)\rho_0^{1+\mu(u)}\right)\right) \\
&\leq \overline{L}(\rho_0 e^{i\theta}, u) + \alpha(\theta)\exp\left(\alpha(\theta)\rho^{1+\mu(u)}\right) - \alpha(\theta)\exp\left(\alpha(\theta)\rho_0^{1+\mu(u)}\right)
\end{aligned}$$

since in $\mathcal{U}_u(1)$ we can always choose $\rho_0 > 1$ such that $\rho_0^{\mu(u)}(1+\mu(u)) > 1$. Finally, by induction $\overline{L}(\rho_0 e^{i\theta}, u) \leq \alpha(\theta)\exp\left(\alpha(\theta)\rho_0^{1+\mu(u)}\right)$ so

$$\overline{L}(\rho e^{i\theta}, u) \leq \alpha(\theta)\exp\left(\alpha(\theta)\rho^{1+\mu(u)}\right)$$

in $\mathcal{F}_{m+1} - \mathcal{F}_m$, hence also in $\mathcal{F}_{m+1}$, and this completes the proof of Theorem 15. ■

Finally, we are ready to finish the proof of our main result Theorem 1, by completing the proof of the de-Poissonization, namely Theorem 9. To recall, we want to prove the following (cf. (47)

$$L_m\left(e^{t/\sqrt{m}}\right)\exp\left(-\frac{\widetilde{X}(m)}{\sqrt{m}}t - \frac{\widetilde{V}(m) - m(\widetilde{X}'(m))^2}{2m}t^2\right) = 1 + O(1/m^{1/2-\beta}) \ .$$

We now split the Cauchy formula (68) into two parts, namely

$$L_m(u) = I_m(u) + E_m(u)$$

with $I_m(u)$ being the part of the integration over $\mathcal{A}_m(\theta)$ and $E_m(u)$ the integration over $\overline{\mathcal{A}}_m(\theta)$ for some $w = me^{i\theta}$ belonging to the circle of integration and lying on the boundary of a convex cone $\mathcal{C}(D, \delta)$. More precisely, we set

$$I_m(u) = \frac{m!m^{-m}}{2\pi} \int_{-Dm^{\delta-1}}^{Dm^{\delta-1}} L(me^{i\theta}, u)e^{-im\theta} \mathrm{d}\theta \ , \tag{70}$$

$$E_m(u) = \frac{m!m^{-m}}{2\pi} \int_{|\theta|\in[Dm^{\delta-1},\pi]} L(me^{i\theta}, u)e^{-im\theta} \mathrm{d}\theta \ . \tag{71}$$

We compute the above integrals separately.

We start with (71). From Theorem 15 we have

$$|E_m(u)| \leq m!m^{-m} \exp\left(\alpha(Dm^{\delta-1})m^{1+\mu(u)}\right) \ .$$

Now, by Stirling's formula: $m! = m^m e^{-m}\sqrt{2\pi m}(1 + O(1/m))$, and after some algebra we obtain (setting $\alpha(\theta) = 1 - \theta^2/4$ and $1 + \mu(u) > \kappa(u)$)

$$E_m(e^{t/\sqrt{m}}) = \exp\left(-0.25 \cdot D^2 m^{2\delta-1} + O(\sqrt{m}\log m)\right) \ .$$

Thus, as $m \to \infty$ we have $E_m(e^{t/\sqrt{m}}) \to 0$ exponentially fast as long as $\delta > 3/4$. By Theorem 6, $\widetilde{X}(m)/\sqrt{m} = O(\sqrt{m}\log m)$ and $(\widetilde{V}(m) - m(\widetilde{X}'(m))^2)/m = O(\log m)$, so

$$\lim_{m\to\infty} E_m(e^{t/\sqrt{m}}) \exp\left(-t\frac{\widetilde{X}(m)}{\sqrt{m}} - t^2\frac{\widetilde{V}(m) - m(\widetilde{X}'(m))^2}{2m}\right) = O(e^{-m^{\varepsilon}})$$

for some $\varepsilon > 0$ and $\delta > 3/4$.

Now, we turn our attention to the evaluation of $I_m(u)$ defined in (70). Let us examine the following expression

$$J_m(t) = \frac{m^m e^{-m}\sqrt{2\pi m}}{m!} I_m\left(e^{t/\sqrt{m}}\right) \exp\left(-t\frac{\widetilde{X}(m)}{\sqrt{m}} - t^2\frac{\widetilde{V}(m)}{2m}\right) \ .$$

After a change of variables (i.e., $\Theta = \theta/\sqrt{m}$) we obtain

$$J_m(t) = \frac{1}{\sqrt{2\pi}} \int_{-Dm^{\delta-1/2}}^{Dm^{\delta-1/2}} L(me^{i\theta/\sqrt{m}}, e^{t/\sqrt{m}}) \exp\left(-\frac{\widetilde{X}(m)t}{\sqrt{m}} - \frac{\widetilde{V}(m)t^2}{2m} - m(i\theta + 1)\right) \mathrm{d}\theta \ .$$

We now assume that $\delta > 1/2$. Then, by Theorem 6 for any $\beta > 0$ and some $\mathcal{U}_u(1)$

$$\begin{aligned}
\log(L(me^{i\theta}, e^t)) &= me^{i\theta} + \widetilde{X}(me^{i\theta})t + \frac{\widetilde{V}(me^{i\theta})}{2}t^2 + O(m^{1+\beta}t^3) \ , \\
\widetilde{X}(me^{i\theta}) &= \widetilde{X}(m) + mi\theta\widetilde{X}'(m) + O(\theta^2 m^{1+\beta}) \\
\widetilde{V}(me^{i\theta}) &= \widetilde{V}(m) + O(\theta m^{1+\beta}) \ ,
\end{aligned}$$

and $me^{i\theta} - m - mi\theta = -m\theta^2/2 + O(m\theta^3)$. Thus,

$$\log(L(me^{i\theta}, e^t)) - \widetilde{X}(m)t - \widetilde{V}(m)\frac{t^2}{2} - m(i\theta+1) = \widetilde{X}'(m)mi\theta t - m\frac{\theta^2}{2} + O(m^{1+\beta}(|\theta|+|t|)^3) \quad (72)$$

which proves that

$$\begin{aligned} L(me^{i\theta/\sqrt{m}}, e^{t/\sqrt{m}}) \exp\left(-\frac{\widetilde{X}(m)t}{\sqrt{m}} - \frac{\widetilde{V}(m)t^2}{2m} - m(i\theta+1)\right) &- \exp\left(it\widetilde{X}'(m)\theta - \frac{\theta^2}{2}\right) \\ &= O(1/m^{1/2-\beta}) \end{aligned}$$

provided $m^{1+\beta}m^{-3/2} \to 0$, that is, $\beta < 1/2$. Furthermore, since $\Re(me^{i\theta} - m - mi\theta) \le -m\theta^2/4$ it is easy to see that for any $\varepsilon < 1/4$ uniformly in t on a compact set and m large, we have

$$\left| L\left(me^{i\theta/\sqrt{m}}, e^{t/\sqrt{m}}\right) \exp\left(-\frac{\widetilde{X}(m)t}{\sqrt{m}} - \frac{\widetilde{V}(m)t^2}{2m} - m(i\theta+1)\right)\right| \le \exp\left(\varepsilon + \varepsilon|\theta| - (0.25-\varepsilon)\theta^2\right)$$

which, by the dominated convergence theorem, leads finally to

$$J_m(t) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{it\widetilde{X}'(m)\theta - \frac{\theta^2}{2}} d\theta = O(1/m^{1/2-\beta}) .$$

To complete the proof of Theorem 9 it suffices to observe that the above integral is equal to $\exp(-(\widetilde{X}'(m))^2 t^2/2)$. After multiplying the above by $\exp(-(\widetilde{X}'(m))^2 t^2/2)$ we obtain (47). This completes the proof of Theorem 9, and also the proof of Theorem 1.

In passing, we note that in the course of these derivation we establish a relationship between the mean $\widetilde{X}(z)$ and the variance $\widetilde{V}(z)$ of the Poisson model and the mean EL_m and the variance $\mathrm{Var}(L_m)$ of the Bernoulli model. For example, a refinment of the above leads to (cf. Lemma 10)

$$\mathrm{Var}(L_m) \sim \widetilde{V}(m) - m(\widetilde{X}'(m))^2 + 0.5m^2(\widetilde{X}''(m))^2 .$$

Thus, after this long trip we completed the proof of Theorem 1A (and also Theorem 1B if one "borrows" the variance result from [10]).

4. CONCLUDING REMARKS

In this paper we settle in the affirmative the conjecture of Aldous and Shields [1] concerning the limiting distribution of the number of phrases in the Lempel-Ziv parsing scheme. Our result was proved for the asymmetric Bernoulli model. One can wonder whether the proposed technique can be extended to prove similar result for the Markovian model. We think we can answer in the affirmative, and in these concluding remarks we briefly sketch our line of arguments. We hope to publish a rigorous proof in a forthcoming paper.

Assume a binary alphabet $\Sigma = \{0,1\}$, and let symbols of all strings be generated according a Markov chain with the transition probability p_{ij} $(i,j \in \Sigma)$. We denote by π_i $(i \in \Sigma)$ the stationary probability of the Markov chain. Finally, let $L^i(u,z)$ and $L(z,u)$ denote respectively the conditional (under the condition that all the strings start with character $i \in \Sigma$) and unconditional Poisson generating functions of the internal path length in a digital search tree. It is not difficult to notice that these functions satisfy the following differential-functional equations

$$\begin{aligned} \frac{\partial L(z,u)}{\partial z} &= L^0(\pi_0 zu, u) L^1(\pi_1 zu, u) && (73) \\ \frac{\partial L^i(z,u)}{\partial z} &= L^0(p_{i0} zu, u) L^1(p_{i1} zu, u) \quad , \quad i \in \Sigma && (74) \end{aligned}$$

The clue to our analysis is – as in the case of the Bernoulli model – establishing the growth rate of $L(z,u)$ which should lead the limiting distribution in the Poisson model. As in Section 3.2, instead of analyzing directly $L^i(z,u)$ (which we expect to have exponential growth) we deal with new functions defined as below

$$f^i(z,u) = \frac{L^i(z,u)}{L^0(p_{i0}uz, u) L^1(p_{i1}uz, u)} \quad , \quad i \in \Sigma \ .$$

We claim that $f^i(z,u) = O(z^{1-\kappa(u)})$ where this time the function $\kappa(u)$ satisfies the following equation

$$\begin{vmatrix} (p_{00}u)^{\kappa(u)} - 1 & (p_{01}u)^{\kappa(u)} \\ (p_{10}u)^{\kappa(u)} & (p_{11}u)^{\kappa(u)} - 1 \end{vmatrix} = 0$$

where $|A|$ denote the determinant of a matrix A. Once we have this result, we can reason along the lines of Section 3 to prove our claim. Clearly, the analysis is much more technical and more challenging. We promise to return to it!

APPENDIX: Proof of Lemma 7A

We prove Lemma 7A. More generally, let

$$v(z) + v'(z) = v(zp) + v(zq) + g(z) \tag{75}$$

be a differential-functional equation of $v(z)$ such that $g(z) = O(\log^2 z)$ for $|z| \to \infty$. We prove that $v(z) = O(z)$.

The proof is by induction over the increasing domains $\mathcal{D}_m$ as described in Fact 4. From Fact 1 we conclude that (75) has the following solution

$$v(z) = v(\nu z) e^{-z(1-\nu)} + e^{-z} \int_{\nu z}^{z} e^x \left(v(px) + v(qx) + g(x) \right) \mathrm{d}x \ .$$

Let now V_m be the upper bound on $|v(z)x^{-1}|$ over the domain $\mathcal{D}_m$ in the convex cone $\mathcal{C}(D,\delta)$, where $z = x + iy$. From the induction hypothesis one obtains the following recurrence

$$V_{m+1} \leq \max\{V_m, V'_m\}$$

where

$$V'_{m+1} \leq V_m \left(1 + \nu e^{-\nu^{-m}(1-\nu)}\right) + \nu^m m^2 \log^2 \nu \ .$$

The above recurrence falls under the pattern discussed in Fact 3(iii), hence by the same arguments as in Section 3.2, we show that V_m are uniformly bounded. Lemma 7A is proved.

ACKNOWLEDGEMENT

It is a pleasure to acknowledge several enlighting discussions that the authors had with Philippe Flajolet, INRIA, on the topic of this paper. The authors also thank Philippe for his long standing support and generosity in sharing his knowledge. "Merci au bon docteur Flajolet!"

We also thank D. Aldous, University of California, Berkeley, for pointing to us the renewal equation, and P. Kirschenhofer and W. Schachinger, Technischen Universität Wien, for discussions about the variance and the fluctuating function arising in Theorem 6. Finally, conscientious referees helped us to eliminate some typos and and to improve overall presentation of the paper. Thanks!

References

[1] D. Aldous and P. Shields, A Diffusion Limit for a Class of Random-Growing Binary Trees, *Probab. Th. Rel. Fields*, 79, 509-542 (1988).

[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York 1968.

[3] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley&Sons, New York 1991.

[4] P. Flajolet and R. Sedgewick, Digital Search Trees Revisited, *SIAM J. Computing*, 15, 748-767 (1986).

[5] P. Flajolet and B. Richmond, Generalized Digital Trees and Their Difference-Differential Equations, *Random Structures & Algorithms*, 3, 305-320 (1992).

[6] P. Flajolet and M. Soria, General Combinatorial Schemas: Gaussian Limit Distributions and Exponential Tails, *Discrete Mathematics*, 114, 159-180 (1993).

[7] E. Gilbert and T. Kadota, The Lempel-Ziv Algorithm and Message Complexity, *IEEE Trans. Information Theory*, 38, 1839-1842 (1992).

[8] P. Jacquet, *Contribution de l'Analyse d'Algorithmes a l'Evaluation de Protocoles de Communication*, Thèse Université de Paris Sud-Orsay, 1989.

[9] P. Jacquet and W. Szpankowski, A Functional Equation Arising in the Analysis of Algorithms, Proc. 26th STOC, 780-789, Monteral 1994.

[10] P. Kirschenhofer, H. Prodinger and W. Szpankowski, Digital Search Trees Again Revisited: The Internal Path Length Perspective, *SIAM J. Computing*, 23, (1994).

[11] D. Knuth, *The Art of Computer Programming. Sorting and Searching,* Addison-Wesley (1973).

[12] A. Lempel and J. Ziv, On the Complexity of Finite Sequences, *IEEE Information Theory* 22, 1, 75-81 (1976).

[13] G. Louchard, Exact and Asymptotic Distributions in Digital and Binary Search Trees, *RAIRO Theoretical Inform. Applications*, 21, 479-495 (1987).

[14] G. Louchard and W. Szpankowski, Average Profile and Limiting Distribution for a Phrase Size in the Lempel-Ziv Parsing Algorithm, *IEEE Trans. Information Theory*, to appear.

[15] H. Mahmoud, *Evolution of Random Search Trees*, John Wiley & Sons, New York 1992.

[16] E. Plotnik, M.J. Weinberger, and J. Ziv, Upper Bounds on the Probability of Sequences Emitted by Finite-State Sources and on the Redundancy of the Lempel-Ziv Algorithm, *IEEE Trans. Information Theory*, 38, 66-72 (1992).

[17] B. Rais, P. Jacquet and W. Szpankowski, A Limiting Distribution for the Depth in PATRICIA Tries, *SIAM J. Discrete Mathematics*, 6, 197-213 (1993).

[18] M. Regnier and P. Jacquet, New results on the Size of Tries, *IEEE Trans. Information Theory*, 35, 203-205 (1989).

[19] R. Remmert, *Theory of Complex Functions*, Springer-Verlag, New York 1991.

[20] W. Schachinger, *Beiträge zur Analyse von Datenstrukturen zur Digitalen Suche*, Dissertation Technische Universität Wien, 1993

[21] P. Shields, Universal Redundancy Rates Do Not Exist, *IEEE Information Theory*, 39, 520-524 (1993).

[22] W. Szpankowski, A Characterization of Digital search Trees From the Successful Search Viewpoint, *Theoretical Computer Science*, 85, 117-134 (1991).

[23] W. Szpankowski, A Generalized Suffix Tree and Its (Un)Expected Asymptotic Behaviors, *SIAM J. Computing*, 22, 1176-1198 (1993).

[24] W. Szpankowski, Asymptotic Properties of Data Compression and Suffix Trees, *IEEE Trans. Information Theory*, 39, 1647-1659 (1993).

[25] J. Ziv, Coding Theorem for Individual Sequences, *IEEE Trans. Information Theory*, 24, 405-412 (1978).

[26] J. Ziv and A. Lempel, A Universal Algorithm for Sequential Data Compression, *IEEE Trans. Information Theory*, 23, 3, 337-343 (1977).

[27] J. Ziv and A. Lempel, Compression of Individual Sequences via Variable-Rate Coding, *IEEE Trans. Information Theory*, 24, 530-536 (1978).