On the asymptotic number of $BCK(2)$-terms

Olivier Bodini* Bernhard Gittenberger†

November 9, 2013

Abstract

We investigate the asymptotic number of a particular class of closed lambda-terms. This class is a generalization of a class of terms related to the axiom system BCK which is well known in combinatory logic. We determine the asymptotic number of terms, when their size tends to infinity, up to a constant multiplicative factor and discover a surprising asymptotic behaviour involving an exponential with fractional powers in the exponent.

1 Introduction

A lambda-term is a formal expression built of variables and a quantifier λ which in general occurs more than once and acts on one of the free variables. It can be described by the context-free grammar $T ::= a \mid (T * T) \mid \lambda a.T$ where a is a variable. The concatenation $T_1 * T_2$ of terms is called *application* and putting the string λa in front of a term is called *abstraction*. Each abstraction λa *binds* the variable a in the subsequent term and each variable can only be bound by at most one abstraction. A term where all the variables are bound is called a *closed* lambda-term.

Lambda-terms are the words of the formal language known as lambda-calculus (see [1]) which was invented by Church and Kleene in the 1930s (see [7, 18, 19]) in order to describe computer programs, analyze programming languages or investigate decision problems. Moreover, it is the basis of the programming language LISP. Furthermore, by the Curry-Howard isomorphism (see [22]) lambda-terms can also be seen as proofs of intuitionistic tautologies.

Our original motivation was to gain information on the typical shape of large closed lambda-terms. Therefore the number of closed lambda-terms of a given size has to be known. This turns out to be a very difficult problem (see [4] for a brief discussion on that issue) and therefore we suggest to approach the problem by studying particular subsets $\mathcal{B}_p$ of the set $\mathcal{L}$ of closed lambda-terms which we call $BCK(p)$-terms, $p = 1, 2, \ldots$, for there exists a connection of $\mathcal{B}_1$ to the axiom system BCK in combinatory logic (see [16] for BCK logic and [5] for an explanation of the connection). The set $\mathcal{B}_p$ is the set of closed lambda-terms in which each abstraction binds at most p variables. Thus we have $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{L}$ as well as $\mathcal{B}(\infty) = \mathcal{L}$. In this paper we initiate this study and derive the asymptotic number of terms of size n in $\mathcal{B}_2$, as $n \to \infty$, up to a constant multiplicative factor. It turns out that the enumerative problems given by $\mathcal{B}_p$ are still very difficult and that they exhibit very interesting phenomena involving exponentials of fractional powers of the size which seem to be new in combinatorial enumeration problems. The only example known to us where similar phenomena occur is the asymptotic enumeration of certain Young tableaux [8].

The interest in counting lambda-terms or terms of other logical systems has recently arisen in the literature. E.g. in [10] proofs in propositional systems are counted. David et al. [9] study

*Institut Galilée, Univ. Paris 13, Villetaneuse (France). Email: olivier.bodini@lipn.univ-paris13.fr. Supported by ANR Magnum project (France)

†Institute for Discrete Mathematics and Geometry, Technische Universität Wien, Wiedner Hauptstrasse 8-10/104, A-1040 Wien, Austria. Email: gittenberger@dmg.tuwien.ac.at. Supported by FWF grant SFB F50-03 and ÖAD, grant F04/2012

the number of terms using a different definition of size from that we will use in this paper and analyze terms w.r.t. properties like normalizability. Grygiel and Lescanne [14] aim at generation of terms and derive several recursive formulas related to the number of lambda-terms. Bounds for the number of terms as well as asymptotics for the number of particular terms have been derived in [5]. Further work where terms with certain restrictions are counted are [4] or [6] where the asymptotic number of terms related to the axiom systems BCI and BCK (i.e. our set $\mathcal{B}_1$) has been determined. A comparison of these two systems was also carried out in [13].

We will determine the asymptotic number of $BCK(2)$-terms and give some outlook on $BCK(p)$ for $p > 2$. The combinatorial enumeration problem seems not only of interest for people dealing with lambda-calculus, but is in our opinion also very interesting in its own right. The reason is that there are several difficulties to overcome. First, we are facing unlabelled objects such that the formal specification gives rise to a functional equation for *ordinary* generating function. But their number is so large such that the radius of convergence is zero. Second, the recurrence relation we get for the enumerative sequence is not holonomic since it involves a quadratic term. We can show that omitting the quadratic term, we get another sequence which differs asymptotically from the original one only in the multiplicative factor. The modified sequence is holonomic, but the differential equation for the generating function has an irregular singularity at infinity. This makes it difficult to obtain the asymptotic behaviour of the coefficients.

Plan of the paper: In the next section we will give a precise definition of the objects we want to count and recall some known results. The specifications and translations into equations for generating functions are based on the symbolic method as introduced in [12, Ch. I-III]. Section 3 is devoted to the main result, the derivation of the asymptotic behaviour of $BCK(2)$-terms. A brief discussion of the cases $p > 2$ is presented in Section 4.

2 $BCK(p)$-terms and known facts

First of all, we have to specify precisely the objects we want to count: We count closed lambda terms up to α-conversion (see [1, Ch. 2]), that is, only the relation between variables and quantifiers are significant, not the variable names themselves. In this context, we prefer another viewpoint to look at lambda-terms:

Definition 1. A *$BCK(p)$-term* is a directed acyclic graph constructed as follows: Start with a Motzkin tree, i.e. a rooted plane tree where each node has 0, 1 or 2 children. This tree is the *underlying tree* of a lambda-term. The nodes with 0 children are called *leaves*,[1] those having 1 child *unary nodes* and those with 2 children *binary nodes* of the lambda-term. Assume that all the edges are directed away from the root. Then add further directed edges, called *pointers*, each having a unary node as start vertex and as end vertex a leaf which belongs to the subtree rooted at the start vertex. Moreover, every unary node must be the start vertex of at most p pointers.

The *size* of a $BCK(p)$-term is defined as the total number of its vertices.

Remark 1. The correspondence between terms and graphs is obvious: Binary nodes correspond to applications, unary nodes to abstractions (which bind at most p leaves) and leaves to variables. Pointers represent the binding relation.

Remark 2. The size of a lambda-term as defined above satisfies the following recursive relation:

$$|x| = 1, \quad |\lambda x.T| = 1 + |T|, \quad |(S * T)| = 1 + |S| + |T|.$$

Here x denotes a variable and S, T denote subterms. That means that each of the variables, abstractions and applications of a term contribute 1 to its size.

[1] Note the slight abuse of the terminology: We add further edges such that the resulting graph does not possess any end nodes. Still, when speaking of lambda-terms we always keep the underlying tree in mind, so that such a wording does make sense.

Now, we will recall some known results. We are interested in the numbers $f_n^{(p)}$ of $BCK(p)$-terms of size n and in order to derive an asymptotic expression for $f_n^{(p)}$ we will work with the generating function

$$F_p(z) = \sum_{n\geq 0} f_n^{(p)} z^n.$$

Since the emphasis of this paper is on $BCK(2)$, set $f_n := f_n^{(2)}$ and $F(z) := F_2(z)$.

In [5] the following results for $F_p(z)$ have been shown:

Proposition 1 ([5]). *Let $F_p(z)$ be the generating function for the counting sequence of $BCK(p)$-terms. Then $F_p(z) = Y(z/(1-z))$ where $Y(z)$ is the unique power series $Y(z) = \sum_{n\geq 0} Y_n z^n$ with nonnegative coefficients which satisfies*

$$Y(z) = \sum_{l=1}^{p} C_{l-1} z^{2l} + zY(z)^2 + \left(\sum_{l=1}^{p} \Delta_l\right) Y(z) \tag{1}$$

with

$$\Delta_p = \sum_{l=1}^{p} \sum_{\sum_i s_i = l;\, \sum_i i s_i = p} \binom{l}{s_1, \ldots, s_p} \left(\prod_{m=1}^{p} \binom{2m}{m}^{s_m}\right) \frac{z^{l+2p+1} \left(\frac{\partial}{\partial z}\right)^l}{l!}$$

Proposition 2 ([5]). *Let $M(z,u)$ denote the generating function of Motzkin trees where z marks the size (i.e. the total number of nodes) and u marks the number of leaves. This function is given by the unique power series solution of $M(z,u) = uz + zM(z,u) + zM(z,u)^2$, that is*

$$M(z,u) = \frac{1 - z - \sqrt{(1-z)^2 - 4uz^2}}{2z}. \tag{2}$$

Then $F_p(z)$ is given as the solution of

$$F_p(z) = z[u^p]\frac{M(z,u)}{1-u} + zF_p(z)^2 + z[u^p]\frac{1}{1-u} F_p\left(\frac{z}{1-2zM(z,u)}\right). \tag{3}$$

We will also utilize the following auxiliary result for fast growing sequences:

Lemma 1 ([5]). *Let $n_0 \in \mathbb{N}$ and $A(z) = \sum_{n\geq n_0} a_n z^n$ be a power series with positive coefficients (from index n_0 on). Assume that there exists $\gamma > 0$ with $a_{n+1}/a_n = \Omega(n^\gamma)$ as $n \to \infty$. Then $[z^n]A(z)^2 = 2a_{n_0}a_{n-n_0}(1 + O(n^{-\gamma}))$ as $n \to \infty$. If we want the second order term, we take the next two terms, and so on.*

Remark 3. In [5] this result was shown for $\gamma \leq 1$. But an obvious modification of the proof shows that this condition can be relaxed to $\gamma > 0$.

3 The asymptotic number of $BCK(2)$-terms

Propositions 1 and 2 give rise to recurrence relations for f_n. The first one yields a finite recurrence for the coefficients of $Y(z)$. However, the recurrence is quadratic and, moreover, a substitution $z \to z/(1-z)$ is required afterwards which gives an unpleasant recurrence for f_n involving infinitely many terms. The same problem arises in Proposition 2 since again there is a quadratic term and translating the coefficient extraction in the last term gives again a recurrence with infinitely many terms. The quadratic term makes the recurrence unamenable to the wealth of methods developed for holonomic sequences. Therefore it seems rather hopeless to get a nice recurrence from this. The best shape that we can reach as follows: The sequence f_n is the solution of the recurrence

relation

$$l_n = m_n + \sum_{k=2}^{n-3} l_k l_{n-k-1} + l_{n-1} + (2n-6) l_{n-3} + \sum_{k=4}^{n-2} \left((2k-8)n^2 - (3k^2-11k-6)n + k(k^2-3k-6)\right) l_{n-k},$$

where m_n is itself given by the recurrence

$$(n-2)^2 m_n = (9n-26)m_{n-1} + (n-3)(n-6)m_{n-2}$$

and the initial condition of these two recurrences are $m_1 = 0, m_2 = 1$ and $l_0 = 0, l_1 = 0, l_2 = 1, l_3 = 2, l_4 = 4, l_5 = 13$.

The polynomial $\delta_{n,k} = (2k-8)n^2 - (3k^2-11k-6)n + k(k^2-3k-6)$ can be obtained from the following formula given in [5] and using creative telescoping methods [24]:

$$\delta_{n,k} = \begin{cases} \sum_{t=0}^{\min\left(2, \lfloor \frac{n-l-1}{2} \rfloor\right)} \sum_{r=0}^{t} \frac{r 2^r \binom{l-1+r}{r} (n-l-2-r)!}{t!\,(t-r)!\,(n-l-1-2t)!} & \text{if } 1 \le l < n-1, \\ 1 & \text{if } l = n-1. \end{cases}$$

To overcome this difficulty we will utilize another sequence which differs from f_n asymptotically only by a multiplicative constant. Let $G(z)$ be the generating function defined by

$$G(z) = z[u^2]\frac{M(z,u)}{1-u} + z[u^2]\frac{1}{1-u} G\left(\frac{z}{1-2zM(z,u)}\right). \tag{4}$$

So we omitted only the quadratic term. The recurrence for the coefficients of $G(z)$ is now a linear one which turns out to be holonomic. Moreover, we will eventually show that this modification does not seriously affect our result: indeed, the difference is only a multiplicative constant in the asymptotic result. So we start with the examination of $g_n = [z^n]G(z)$

Lemma 2. *The coefficients $g_n = [z^n]G(z)$ satisfy the recurrence relation*

$$g_n = 4g_{n-1} - 6g_{n-2} + (2n-2)g_{n-3} - (4n-15)g_{n-4} + (2n^2-14n+20)g_{n-5} - (2n^2-22n+84)g_{n-6}, \quad \text{for } n \ge 6, \tag{5}$$

with initial condition $g_0 = g_1 = 0, g_2 = 1, g_3 = 2, g_4 = 4, g_5 = 12$.

Proof. If we extract the coefficients in Equation (4) and use that the partial derivatives of $M(z,u)$ satisfy $M(z,0) = 0$, $M_u(z,0) = z/(1-z)$ and $M_{uu}(z,0) = 2z^3/(1-z)^3$, then we obtain

$$G(z) = \frac{z^2(2z^2-2z+1)}{(1-z)^3} + zG(z) + \frac{4z^6(3-2z}{(1-z)^3}G'(z) + \frac{4z^7}{(1-z)^2}G''(z).$$

This is (after multiplying with the common denominator $(1-z)^3$) a holonomic differential equation from which we can directly extract a recurrence relation for the coefficients of $G(z)$. This is done for instance using Maple's gfun package [21]. In that way we obtained (5). □

Remark 4. Let us also remark that generating the first few terms and using guess-and-prove by means of gfun we can obtain several recurrence relations for the coefficients of $G(z)$. Indeed, this depends on the preassigned parameters of the guess-and-prove procedure (which are the bounds for the order of the recurrence itself as well as the order of its coefficients). If we aim at finding a smallest order recurrence, we get

$$\frac{n}{2}g_{n+5} = (n^3-n)g_n - (n^2+n)g_{n+1} + (n^2 + \frac{3n}{2} + \frac{1}{2})g_{n+2} - (\frac{3n}{2}+1)g_{n+3} + (\frac{3n}{2}+\frac{1}{2})g_{n+4}.$$

But for the arguments we use in the following, the recurrence given in (5) was more convenient.

Intuitively, the recurrence tells us that the fifth term on the right-hand side is dominant and thus the exponential growth rate should be $(n!)^{2/5}$. The following lemmas make this precise.

Lemma 3. *For $n \geq 2$ we have $g_{n+1}/g_n \geq 2$.*

Proof. Note that the functional equation for $G(z)$ admits a combinatorial interpretation: We are counting $BCK(2)$-terms which do not have a binary node as their root. Now take a term of size n and choose one pointer which connects a unary node u and a leaf l. Attach a new unary root node r on top of the term. This gives a term of size $n+1$. Deleting the pointer $u \to l$ and inserting a pointer $r \to l$ instead yields another term of size $n+1$. Thus each term of size n corresponds to at least two terms of size $n+1$. □

Corollary 1. *For $n \geq 6$ we have $g_n \geq n^2 g_{n-5}$.*

Proof. The result is immediate from the previous lemma: Indeed we have

$$\begin{aligned} 4g_{n-1} - 6g_{n-2} &\geq 2g_{n-2} \geq 0, \\ (2n-2)g_{n-3} - (4n-15)g_{n-4} &\geq 11g_{n-4} \geq 0, \\ (2n^2-14+20)g_{n-5} - (2n^2-22n+84)g_{n-6} &\geq (n^2-7n+10)g_{n-5} + (8n-64)g_{n-6} \geq n^2 g_{n-5}. \end{aligned}$$ □

Corollary 2. *For sufficiently large n we have $g_n \geq (n!)^{2/5}$.*

Thus we are faced with a rapidly growing sequence, as expected.

Lemma 4. *For every $\varepsilon > 0$ the sequence $(g_n)_{n\geq 0}$ satisfies $g_n \leq (n!)^{2/5}(2+e)^{n/5}$*

Proof. We choose an arbitrary $\varepsilon > 0$ and substitute

$$v_n = v_n(\varepsilon) := \frac{g_n}{(n!)^{2/5}(2+\varepsilon)^{n/5}}$$

Then (5) becomes

$$\begin{aligned} v_n &= \sum_{k=1}^{6} a_k(n) v_{n-k} \\ &= \frac{4}{n^{2/5}(2+\varepsilon)^{1/5}} v_{n-1} - \frac{6}{(n)_2^{2/5}(2+\varepsilon)^{2/5}} v_{n-2} + \frac{2n-2}{(n)_3^{2/5}(2+\varepsilon)^{3/5}} v_{n-3} - \frac{4n-15}{(n)_4^{2/5}(2+\varepsilon)^{4/5}} v_{n-4} \\ &\quad + \frac{2n^2-14n+20}{(n)_5^{2/5}(2+\varepsilon)} v_{n-5} - \frac{2n^2-22n+84}{(n)_6^{2/5}(2+\varepsilon)^{6/5}} v_{n-6} \end{aligned}$$

Now it easy to prove that $0 \leq v_n \leq 1$ by induction. □

In order to get information on the sequence $(f_n)_{n\geq 0}$ one might be tempted to add the quadratic term to the recurrence, but this is not feasible. But it may serve as a heuristic. Indeed putting the quadratic term and substituting as before, this term becomes $(2+\varepsilon)^{-1/5} \sum_{l=1}^{n-1} \frac{1}{n\binom{n-1}{l}} v_l v_{n-1-l}$ which is obviously very small. This indicates that the quadratic term is not essential for the asymptotic behaviour of f_n. We will show that it actually only affects the multiplicative factor of the asymptotic expression. The sequences start with

$$\begin{aligned} (f_n)_{n\geq 0} &= (0, 0, 1, 2, 4, 13, 40, 127, 459, 1671, 6284, \dots) \\ (g_n)_{n\geq 0} &= (0, 0, 1, 2, 4, 12, 35, 110, 390, 1384, 5101, \dots) \end{aligned}$$

and Figure 1 depicts the behaviour of the quotient for the first 5000 terms.

Since we are facing a holonomic sequence we can use again the rich toolbox available for such sequences. Due to the rapid growth of g_n, $G(z)$ is not an analytic function. Thus we will work with its Borel transform

$$H(z) = \sum_{n\geq 0} \frac{g_n}{n!} z^n$$

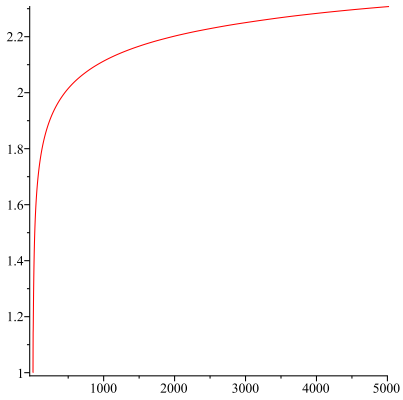

Figure 1: The quotient f_n/g_n.

which is an entire function. Using Maple's gfun package again we obtain a differential equation for $H(z)$.

Lemma 5. *The function $H'(z)$ is a solution of the differential equation*

$$y^{(iv)} = 4y''' + (2z-6)y'' + (2z^2-4z+8)y' + (-2z^2+8z-5)y \tag{6}$$

with initial conditions $H(0) = H'(0) = 0, H''(0) = 1, H'''(0) = 2, H^{(iv)}(0) = 4$.

Proposition 3. *There exists a $\delta > 0$ such that, as $z \to \infty$ in the sector*

$$S = \{z \in \mathbb{C} \mid |\arg(z)| < \delta\}, \tag{7}$$

we have

$$H(z) = \frac{1}{z}\exp\left(\frac{3\cdot 2^{1/3}}{5}z^{5/3} + 2^{-4/3}z^{4/3} + z - \frac{2^{1/3}}{27}z^{2/3} + \frac{2^{5/3}}{81}z^{1/3} + O(1)\right). \tag{8}$$

Proof. Setting

$$\mathbf{y} = \begin{pmatrix} y \\ y' \\ y'' \\ y''' \\ y^{(iv)} \end{pmatrix}$$

the differential equation (6) can be rewritten in the form

$$\begin{aligned} z^{-2}\mathbf{y}' = A(z)\mathbf{y} &= (A_0 + z^{-1}A_1 + z^{-2}A_2)\mathbf{y} \\ &= \left(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{pmatrix} + \frac{1}{z}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 8 & -4 & 2 & 0 \end{pmatrix} + \frac{1}{z^2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -5 & 8 & -6 & 4 \end{pmatrix}\right)\mathbf{y}(z) \end{aligned} \tag{9}$$

According to [23, Ch. 10] this is a differential equation having an irregular singularity at $z = \infty$ and being of rank 3. Unfortunately, it does not belong to the simple class where all the eigenvalues of the leading matrix A_0 are distinct. In the simple case, we could obtain the asymptotic expansion for $\mathbf{y}(z)$ valid in a sector S with $\delta = \pi/6$, according to the theory developed in [23, Ch. 12]. The matrix J_0 in Jordan canonical form corresponding to A_0 has one block of dimension two and two blocks of dimension one, all of them with the same eigenvalue 0. The complicated case, i.e. only one distinct eigenvalue, is treated in [23, Ch. 19]. What is necessary in that case is a transformation such that A_0 is in Jordan's canonical form and nilpotent. This involves only a constant transformation matrix. Next we have to apply a shearing transformation of the form $\mathbf{y} = P(z)\mathbf{u}$ with

$$P(z) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & z^{-g} & 0 & 0 \\ 0 & 0 & z^{-2g} & 0 \\ 0 & 0 & 0 & z^{-3g} \end{pmatrix}.$$

This transforms our differential equation (9) into $z^{-2}\mathbf{u}'(z) = B(z)\mathbf{u}(z)$ with

$$B(z) = P^{-1}(z)A(z)P(z) - z^{-2}P'(z)P(z).$$

The parameter $g = g_1/p$, $\gcd(g_1, p) = 1$, is the minimal positive number such that $\lim_{z\to+\infty} z^g B(z)$. By construction, g is a rational number, but not necessarily an integer. In our case we obtain $g = 1/2$ and thus $p = 2$. Therefore the matrix $B(z)$ contains now fractional powers of the independent variable z. In order to be able to apply the method of [23, Ch 12] we have to get rid of this. If we have to apply the transformation $z = p^{g-3}t^p = 2^{-5/2}t^2$ and get the new differential equation

$$t^{-4}\mathbf{v}'(t) = C(t)\mathbf{v}(t) = \sum_{l\geq 0} C_l t^{-l}.$$

If we were lucky, we would be in the simple case now where either all eigenvalues of C_0 are distinct and [23, Theorem 12.3] directly applies or we would have at least two different eigenvalues such that the system of differential equations can be decomposed into two or more independent (systems of) equations of lower dimension. Neither of this the case here. We are facing the most difficult situation which Wasow calls problem of type (E). What to do then is explained in detail in [23, Sec. 19.4]. It amounts to iteration of the steps carried out above and then by [23, Lemma 19.4] we know that after finitely many steps we arrive at a simple case.

In our case, the first two steps yields $g = 1/2$ and $g = 1/3$, respectively. It is not necessary to carry out all these. The important issue is the consequence of all these computations: In each iteration we have to apply a similarity transformation of the form $z = \alpha t^p$, yielding eventually a differential equation for which we know the asymptotic behaviour of its solutions inside a sector S of the form (7). Each similarity transformation introduces a factor p among the powers of z such that we obtain eventually an asymptotic behaviour of the form $\exp(Q(z^{1/q})z^r \sum_{j\geq 0} Q_j(\log z)z^{-js}$ (Structure theorem for irregular singularities, cf. [12, Theorem VIII.7]) where q is the product of all the p's introduced by the similarity transformations, r is a complex number, s is a positive rational number and Q, Q_j are polynomials. The asymptotic expression is valid in the sector S with $\delta = \pi/(6q)$. In our case we know that q is an integer multiple of 6.

In order to proceed, we only consider the positive real line. Here, once knowing more or less what kind of asymptotic behaviour we can expect, we do bootstrapping. By the growth rate of f_n we can expect that the dominant term of $Q(z^{-1/q})$ should be of order $z^{5/3}$. Plugging this into Maple and bootstrapping further we get (8) which must indeed be the correct asymptotic expression for the whole sector S. □

Knowing the asymptotic behaviour of the function, the next step is to derive the asymptotic behaviour of its coefficients. Due to the rapid growth of the function [20, Sec. 12] suggests the application of the saddle point method. Indeed the function fits well into the framework of Hayman admissible functions [15].

Lemma 6. *The function $H(z)$ is Hayman-admissible.*

Proof. On the arcs, $\{z = re^{i\phi} \mid |\phi| \le r^{-2/3}\}$ we can use the asymptotic expression provided by Proposition 3 and obtain

$$H(re^{i\phi}) = H(r)\exp\left(i2^{1/3}r^{5/3}\phi - \frac{5\cdot 2^{1/3}}{6}r^{5/3}\phi^2 + O(r^{5/3}\phi^3)\right)$$

Since we have a power series with positive and monotone coefficients, the maximum of $|H(z)|$ on the arcs $\{z = re^{i\phi} \mid r^{-2/3} \le \phi \le 2\pi - r^{-2/3}\}$ must be attained on the boundary points. But, for r sufficiently large, these points lie inside the sector S given in Proposition 3 and thus they fall in the asymptotic regime given by (8). This implies that for $|\phi| > r^{-2/3}$ we have that

$$|H(re^{i\phi})| = O\left(H(r)\exp\left(-\frac{5\cdot 2^{1/3}}{6}r^{1/3}\right)\right) = o\left(\frac{H(r)}{r^{5/3}}\right). \tag{10}$$

since the second logarithmic derivative of $H(z)$, $B(z) = \frac{\mathrm{d}\log H(z)}{\mathrm{d}\log z}$, satisfies $B(r) \sim r^{5/3}$, we can rewrite Eq. (10) in the form $|H(re^{i\phi})| = o(H(r)/B(r))$ which suffices to prove our lemma. □

Proposition 4. *The function $H(z)/z^n$ has a uniquely determined saddle point z_n on the positive real axis which satisfies*

$$z_n \sim 2^{-1/5}n^{3/5} - \frac{2^{1/5}}{5}n^{2/5} - \frac{13\cdot 2^{3/5}}{50}n^{1/5} + \frac{1}{5} + \frac{41\cdot 2^{2/5}}{625}n^{-1/5} + \frac{8069\cdot 2^{4/5}}{31250}n^{-2/5},$$

as $n \to \infty$.

Proof. The saddle point is the positive solution of the equation $zH'(z)/H(z) = n$. From (8) we can derive easily the assertion. □

Now we are able to determine the desired asymptotic behaviour. We are unable to find the constant. Even numerically this seems a nontrivial task, since the convergence rate is very slow and the determination of many values of f_n computationally difficult. A comparison between f_n and the asymptotic expression for the first 10000 terms of the sequence is given in Figure 2.

Remark 5. The curve in Figure 2 gives the impression that we can hope for some "nice" constant like $1/e$. But unfortunately, the convergence is extremely slow. Indeed, Manuel Kauers [17] was able to compute the first 500000 terms and got $A \approx 0.922388909933584583$.

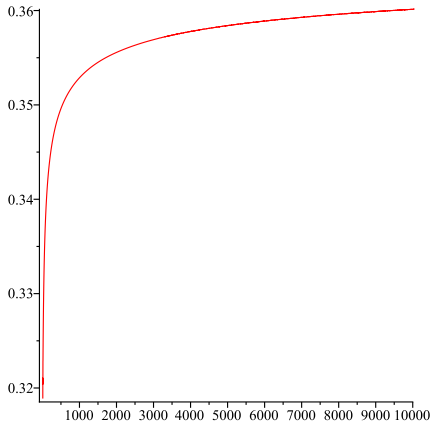

Figure 2: The quotient between f_n and the asymptotic expression (11) without the constant A.

Proposition 5. *The quotient g_n/f_n of the coefficients of the solution of the linearized equation* (4) *and those of the original equation* (2)*, for $p = 2$, tends to a constant.*

Proof. First of all we note that by construction we obviously have $g_n \leq f_n$. The idea of the prove is now as follows: We construct a sequence $(l_n)_{n\geq 0}$ such that $g_n \leq f_n \leq l_n$. Furthermore, $(l_n)_{n\geq 0}$ should be a D-finite sequence satisfying $g_n \sim Cl_n$ for some constant $C \leq 1$. The proof will be completed by showing that there is an increasing sequence of holonomic functions $G^k(z)$ (the exponent k is nothing but an index) tending to $F(z)$ and such that $g_n^k \sim C_k g_n$. So, C_k is increasing and bounded by C. Thus, it admit a limit which is the expected constant.

We start with the construction of $(l_n)_{n\geq 0}$. The growth rate of $(g_n)_{n\geq 0}$ given by Corollary 2 and Lemma 4 imply that we can apply Lemma 1 to (3) with $\gamma = 2/5$. We obtain (keeping in mind that $f_0 = f_1 = 0$)

$$[z^n]zF(z)^2 = \sum_{i=1}^{n-1} f_i f_{n-1-i} \sim 2f_2 f_{n-3} = 2f_{n-3} < 3f_{n-3}.$$

Thus we consider the functional equation

$$G^{\uparrow}(z) = z[u^2]\frac{M(z,u)}{1-u} + 3z^3 G^{\uparrow}(z) + z[u^2]\frac{1}{1-u}G^{\uparrow}\left(\frac{z}{1-2zM(z,u)}\right).$$

which determines the generating function of a holonomic sequence. Now we can proceed as before, derive a differential equation:

$$\begin{aligned}\frac{d^6}{dz^6}H^{\uparrow}(z) = &-3H^{\uparrow}(z) - (4z-17)\frac{d}{dz}H^{\uparrow}(z) - \left(2z^2-12z+18\right)\frac{d^2}{dz^2}H^{\uparrow}(z)\\ &+ \left(2z^2-4z+13\right)\frac{d^3}{dz^3}H^{\uparrow}(z) + (2z-6)\frac{d^4}{dz^4}H^{\uparrow}(z) + 4\frac{d^5}{dz^5}H^{\uparrow}(z)\end{aligned}$$

for the Borel transform of $G^{\uparrow}(z)$, apply Wasow's theory to obtain the asymptotic behaviour of $H^{\uparrow}(z)$ in some sector and then apply the saddle point method to obtain the asymptotic behaviour of the sequence $g_n^{\uparrow}$ itself. As expected, we obtain, up to a constant, the same asymptotic behaviour as for $g_n = n! \cdot h_n$, given in (11).

The final task is to consider

$$G^k(z) = z[u^2]\frac{M(z,u)}{1-u} + \sum_{n=0}^{k} f_n z^n \cdot G^k(z) + z[u^2]\frac{1}{1-u}G^k\left(\frac{z}{1-2zM(z,u)}\right).$$

Clearly, $G^k(z)$ is holonomic and increasingly tends to $F(z)$ when $k \to \infty$. Now, if we look at the Borel transform $H^k(z)$ of $G^k(z)$, we notice the extra terms $\sum_{n=0}^{k} f_n z^n \cdot G^k(z)$ have no effect on the asymptotics up to a constant (this is a long but straightforward way). So, $h_n^k \sim C_k h_n$ and we conclude that the limit of the C_k exists and corresponds to the expected constant. □

Theorem 1. *The asymptotic number of $BCK(2)$-terms of size n satisfies*

$$f_n \sim An^{2n/5}2^{n/5}e^{-2n/5}\exp\left(2^{-8/5}n^{4/5} + \frac{7\cdot 2^{4/5}}{15}n^{3/5} - \frac{17\cdot 2^{1/5}}{75}n^{2/5} - \frac{41\cdot 2^{3/5}}{500}n^{1/5}\right)n^{-\frac{3}{5}} \quad (11)$$

where A is some positive constant.

Proof. Applying the saddle point method to $H(z)$ and using (8) we obtain the asymptotic behaviour of $h_n = [z^n]H(z)$ up to a multiplicative constant. Furthermore, by Proposition 5 we know that

$$\frac{f_n}{g_n} = \frac{f_n}{n!h_n} \to \text{const.}$$

From this we immediately arrive at the assertion. □

4 On the asymptotic number of $BCK(p)$-terms for $p > 2$

The approach we used to determine the asymptotic number of $BCK(2)$-terms works for $p > 2$ as well. But the computations are getting very involved already for $p = 3$. Let $G_p(z)$ denote the generating function solving the functional equation given in Proposition 2, but with the quadratic term removed, and $H_p(z)$ its Borel transform. Experiments with higher values of p indicate that the differential equation for $H_p(z)$ is a holonomic equation of order $2p+1$ with coefficients being polynomials of order up to p.

Conjecture 1. *The function $H_p(z)$ behaves in a sector S of the shape (7) asymptotically as follows:*

$$H_p(z) = Q_p(z) \exp\left(\frac{p+1}{2^p+1} \left(\frac{2^p}{p!}\right)^{1/(p+1)} z^{(2p+1)/(p+1)} + \sum_{m=1}^{2p} a_m z^{m/(p+1)} + O(1) \right)$$

Therefore this involves a bootstrapping approach to determine $2p+1$ constants, then calculating the first $(2p+1)$ terms of the saddle point asymptotic and finally carrying out the saddle point method. With these heavy computations we quickly reach limits of computer memory resources. Therefore a clever computer algebra approach is necessary to proceed further.

5 Conclusion

We have determined the asymptotic number $BCK(2)$-terms and presented a method which in principle works works also for higher values of p. The sequence exhibits a very interesting behaviour involving fraction powers of the size n in the exponent and relying on differential equation with irregular singularities. Even the asymptotic behaviour of the generating function involves exponentials of fraction powers of the independent variable which seems to be new in combinatorial enumeration. This is, for instance, not the case in the enumeration of Young tableaux [8], where the counting sequence contains fractional powers of the size, but the asymptotic expression for the generating function involves integer powers. Other enumeration problems where powers of $\sqrt{n}$ appear in the enumeration sequence are presented in [2, 3]. Recently, Fill et al. [11] presented an enumeration problem the solution of which involves fractional powers not being a multiple of $1/2$. Maybe this problem fits as well into the framework presented here and it might be possible to solve what the authors called a "herculean task", i.e. getting the full asymptotic expansion of the considered sequence.

References

[1] Henk P. Barendregt. *The Lambda Calculus*, volume 103 of *Studies in Logic and the Foundations of Mathematics. Its syntax and semantics.* North-Holland Publishing Co., Amsterdam, revised edition, 1984.

[2] Edward A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Math.*, 10:217–223, 1974.

[3] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[4] Olivier Bodini, Danièle Gardy, and Bernhard Gittenberger. Lambda-terms of bounded unary height. In *ANALCO, workshop on ANALytic COmbinatorics*, San Francisco (USA), January 2011.

[5] Olivier Bodini, Danièle Gardy, Bernhard Gittenberger, and Alice Jacquot. Enumeration of generalized BCI lambda-terms. 2013. Submitted.

[6] Olivier Bodini, Danièle Gardy, and Alice Jacquot. Asymptotics and random sampling for BCI and BCK lambda terms. *TCS*, 2013. To appear.

[7] Alonzo Church. An unsolvable problem of elementary number theory. *Amer. J. Math.*, 58(2):345–363, 1936.

[8] Frédéric Chyzak, Marni Mishna, and Bruno Salvy. Effective scalar products of D-finite symmetric functions. *J. Combin. Theory Ser. A*, 112(1):1–43, 2005.

[9] René David, Katarzyna Grygiel, Jakub Kozik, Christophe Raffalli, Guillaume Theyssier, and Marek Zaionc. Asymptotically almost all λ-terms are strongly normalizing. *Log. Methods Comput. Sci.*, 9(1):1:02, 30, 2013.

[10] René David and Marek Zaionc. Counting proofs in propositional logic. *Arch. Math. Logic*, 48(2):185–199, 2009.

[11] James Allen Fill, Svante Janson, and Mark Daniel Ward. Partitions with distinct multiplicities of parts: on an "unsolved problem" posed by Herbert Wilf. *Electron. J. Combin.*, 19(2):Paper 18, 6, 2012.

[12] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 2009.

[13] Katarzyna Grygiel, Paweł Idziak, and Marek Zaionc. How big is BCI fragment of BCK logic. *J. Logic Comput.*, 23(3):673–691, 2013.

[14] Katarzyna Grygiel and Pierre Lescanne. Counting and generating lambda terms. *J. Functional Programming*, 2013. To appear.

[15] Walter K. Hayman. A generalisation of Stirling's formula. *J. Reine Angew. Math.*, 196:67–95, 1956.

[16] Kiyoshi Iséki and Shôtarô Tanaka. An introduction to the theory of BCK-algebras. *Math. Japon.*, 23(1):1–26, 1978/79.

[17] Manuel Kauers. private communication.

[18] Stephen C. Kleene. A theory of positive integers in formal logic. Part I. *Amer. J. Math.*, 57(1):153–173, 1935.

[19] Stephen C. Kleene. A theory of positive integers in formal logic. Part II. *Amer. J. Math.*, 57(2):219–244, 1935.

[20] Andrew M. Odlyzko. Asymptotic enumeration methods. In *Handbook of Combinatorics, Vol. 1, 2*, pages 1063–1229. Elsevier, Amsterdam, 1995.

[21] Bruno Salvy and Paul Zimmerman. Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.*, 20(2):163–177, June 1994.

[22] Morten Heine Sørensen and Paweł Urzyczyn. *Lectures on the Curry-Howard Isomorphism.* Amsterdam: Elsevier, 2006.

[23] Wolfgang Wasow. *Asymptotic Expansions for Ordinary Differential Equations*. Dover, Toronto, 1987.

[24] Doron Zeilberger. The method of creative telescoping. *J. Symbolic Comput.*, 11(3):195–204, 1991.