ON ASYMPTOTICS OF CERTAIN RECURRENCES ARISING IN UNIVERSAL CODING

Wojciech Szpankowski*
Department of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.

Abstract

Ramanujan's Q-function, Lambert W-function, and the so called "tree function" $T(z)$ defined implicitly by the equation $T(z) = ze^{T(z)}$ found applications in hashing, the birthday paradox problem, random mappings, caching, memory conflicts, and so forth. Recently, several novel applications of these functions to information theory problems such as linear coding and universal portfolios were brought to light. In this paper, we study them in the context of another information theory problem, namely: universal coding which was recently investigated by Shtarkov *et al.* [*Prob. Inf. Trans., 31, 1995*]. We provide asymptotic expansions of certain recurrences studied there which describe the optimal redundancy of universal codes. Our methodology falls under the so called *analytical information theory* that was recently applied to a variety of information theory problems.

1 Introduction

Recently, Jacquet and Szpankowski [7] have proposed to call *analytical information theory* a subarea of information theory that solves problems of information theory by analytical methods, that is, those in which complex analysis plays a primary role. Complex analysis and analytical methods were used for a long time in information theory, most notably in coding and signal processing. However, in recent years there has been a resurgence of interest and a few successful applications of analytical methods to problems on words (strings, sequences, codes) that are at the core of information theory (cf. [5, 6, 9, 10, 14]).

In this paper, we shall continue developing further tools of analytical information theory for coding theory. In a sense, this work is a continuation of our earlier paper [14] that deals with certain sums arising in linear codes. Here, we deal with (multi-alphabet) universal coding. A direct motivation for this research is the recent paper of Shtarkov *et al.* [13] (see also Shtarkov [12]) that designs universal codes in a multi-alphabet environment (e.g., think of constructing the best possible code for texts written in English, French and Polish). The authors of [13] aim at designing an acceptable code such that the maximal redundancy over the alphabets is as small as possible.

*This research was supported in part by NSF Grants NCR-9206315 and NCR-9415491, and in part by NATO Collaborative Grant CGR.950060.

For memoryless sources, the above redundancy turns out to satisfy certain recurrence that is of interest not only to coding but other problems (e.g., hashing, optimal portfolios, caching, memory conflict, etc.). No solution is known for this recurrence, and only an upper bound (cf. [12]) or simple asymptotics (cf. [13]) were derived. Our goal is to solve exactly the recurrence (in terms of certain generating functions) and to establish an asymptotic expansion for it.

We establish the above announced result by analytical methods that resemble our approach from [14]. We apply generating functions, analytical combinatorics and complex asymptotics (e.g., singularity analysis) to solve the problem. In fact, our methodology allows to solve exactly and asymptotically the following recurrence (cf. Section 3 for more details). Under some initial conditions, a sequence x_n^m satisfies for $n \geq 1$ and $m \geq 1$:

$$x_n^m = a_n + \sum_{i=0}^{n} \binom{n}{i} \left(\frac{i}{n}\right)^i \left(1 - \frac{i}{n}\right)^{n-i} (x_i^{m-1} + x_{n-i}^{m-1}) \ , \qquad (1)$$

where a_n is a given sequence (the so called *additive* term), and m is an additional parameter. In this paper we deal with $m \geq 1$ while in [14] we considered mostly $m = 1$ (cf. also [2]).

2 Main Results

Let x^n denote a sequence built over a finite alphabet $\mathcal{A} = \{1, 2, \ldots, M\}$, and let $P(x^n; \omega)$ be the probability of x^n generated by the source ω. We define a uniquely decodable code $\varphi(x^n)$ of length $|\varphi(x^n)| = -\log Q(x^n)$ where $Q(\cdot)$ is an arbitrary probability distribution on $\mathcal{A}^n$. The cumulative *redundancy* of the encoding x^n at the output of a source ω is defined as $\rho(x^n; \varphi_n, \omega) := -\log Q(x^n) + \log P(x^n)$. Finally, let us consider a set of sources Ω, and define the min-max redundancy as $\rho_n(\Omega) := \inf_{\varphi_n} \sup_{\omega \in \Omega} \max_{x^n \in \mathcal{A}^n} \{\rho(x^n; \varphi_n, \omega)\}$.

It is well known (cf. [12, 13]) that the minimum of maximal individual "cumulative" redundancy of universal coding for the set of all m-ary memoryless sources Ω_m with known alphabet is equal to $\rho_n(\Omega_m) = \log D_n(m)$, where

$$D_n(m) = \sum_{i=0}^{m} \binom{m}{i} D_n^*(i), \qquad (2)$$

where $D_0^*(1) = 0$, $D_n^*(1) = 1$ for $n \geq 1$, and $D_n^*(i)$ for $i > 1$ satisfies the following recurrence:

$$D_n^*(i) = \sum_{k=1}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} D_{n-k}^*(i-1) \qquad (3)$$

provided that $D_n^*(i) = 0$ for $i > n$. In [13] it was shown that for fixed $m \geq 1$

$$D_n^*(m) = \frac{\sqrt{\pi}}{\Gamma(m/2)} \left(\frac{n}{2}\right)^{\frac{m}{2} - \frac{1}{2}} (1 - O(1/\sqrt{n})$$

where $\Gamma(z)$ is the Euler gamma function.

Our goal is to establish an asymptotic expansion for $D_n^*(m)$, which will further allow us to obtain an asymptotic expansion for $D_n(m)$ and for the redundancy $\rho_n(\Omega_m)$.

Theorem 1 *For fixed $m \geq 1$ the quantity $D_n^*(m)$ attains the following asymptotics*

$$\begin{aligned} D_n^*(m) &= \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{1}{2}} - \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2}-\frac{1}{2})}\left(\frac{2m}{3}\right)\left(\frac{n}{2}\right)^{\frac{m}{2}-1} \\ &+ \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{3}{2}}\left(\frac{3+m(m-2)(8m-5)}{72}\right) + O(n^{\frac{m}{2}-2}) \end{aligned} \tag{4}$$

for large n.

Proof. We shall use certain tools of analytical analysis of algorithms, notably generating functions (cf. [8, 15]) and singularity analysis (cf. [3]). We follow our approach from [14] (cf. also [8, 15]).

Let us introduce a new sequence $\widehat{D}_n^*(m)$ defined as

$$\widehat{D}_n^*(m) = \sum_{k=0}^{n} \binom{n}{k}\left(\frac{k}{n}\right)^k \left(1-\frac{k}{n}\right)^{n-k} D_{n-k}^*(m-1) = D_n^*(m) + D_n^*(m-1) \ . \tag{5}$$

To simplify the above recurrence, we shall work with a special generating functions, namely:

$$\widehat{D}_m^*(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k \widehat{D}_k^*(m) \ .$$

We will also need another special function $B(z)$ defined as (cf. [14])

$$B(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k = \frac{1}{1-T(z)} \ , \tag{6}$$

where $T(z)$ satisfies $T(z) = ze^{T(z)}$ and also

$$T(z) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} z^k \ .$$

The function $T(z)$ is called the "tree function" since it enumerates rooted labeled trees (cf. [4, 8, 14]). It is also related to Lambert's W-function (cf. [1]) defined as a solution of $W(x)\exp(W(x)) = x$ and which can be called from MAPLE. (In fact, $T(z) = -W(-z)$.) Furthermore, it can be obtained from the Ramanujan's Q-function (cf. [4, 8, 14]) which finds many applications in hashing, random mappings, and memory conflict (cf. [1, 4, 14, 15]) for further references).

Let us now return to (5). Observe that it can be re-written as

$$\frac{n^n}{n!}\widehat{D}_n^*(m) = \sum_{k=0}^{n} \frac{k^k}{k!} \cdot \frac{(n-k)^{n-k}}{(n-k)!}\widehat{D}_{n-k}^*(m-1) \ .$$

To translate the above into a recurrence equation on the generating functions just introduced, we multiply both sides of the above by z^n, sum it up, and then by the *convolution formula* of generating functions we find

$$\widehat{D}_m^*(z) = B(z) D_{m-1}^*(z) \ .$$

Thus

$$D_m^*(z) = (B(z) - 1)^{m-1} D_1^*(z) = (B(z) - 1)^m \tag{7}$$

since $D_1^*(z) = B(z) - 1$. Then, the desired $D_n^*(m)$ can be found as

$$D_n^*(m) = \frac{n!}{n^n}[z^n]\left((B(z) - 1)^{m-1} B(z)\right) \tag{8}$$

where $[z^n]f(z)$ is the standard notation for the coefficient of $f(z)$ at z^n.

Our goal, however, is to find an asymptotic expansion for $D_n^*(m)$ for large n and fixed m. It is well–known that such asymptotics can be extracted by studying singularities of the associated generating function. In our case, singularities of $B(z)$ and singularities of $T(z)$ play a role. Let us recall from [1, 8, 14, 15] that $T(z)$ has a singularity at $z = e^{-1}$, and around this point it has the following expansion (see [1] for a full asymptotic expansion):

$$T(z) - 1 = -\sqrt{2(1 - ez)} + \frac{2}{3}(1 - ez) - \frac{11\sqrt{2}}{36}(1 - ez)^{3/2} + \frac{43}{135}(1 - ez)^2 + O((1 - ez)^{5/2})\ .$$

Then, $B(z)$ can also be expanded around $z = e^{-1}$ leading to

$$B(z) = \frac{1}{\sqrt{2(1 - ez)}} + \frac{1}{3} - \frac{\sqrt{2}}{24}\sqrt{(1 - ez)} + \frac{4}{135}(1 - ez) - \frac{23\sqrt{2}}{1728}(1 - ez)^{3/2} + O((1 - ez)^2)\ .$$

To extract the coefficients at z^n of (8), we shall use the *singularity analysis* and *transfer theorems* of Flajolet and Odlyzko [3] which allow us to compute separately the coefficients for every function involved in the asymptotic expansion. Actually, we use directly MAPLE (more precisely, `equivalent` program thanks to B. Salvy [11]) to extract coefficients of $(B(z) - 1)^m$. In fact, `equivalent` directly gives

$$\begin{aligned} e^{-n}[z^n](B(z) - 1)^m &= \frac{n^{\frac{m}{2}-1}}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} - \frac{n^{\frac{m}{2}-\frac{3}{2}}}{2^{\frac{m}{2}-\frac{1}{2}}}\left(\frac{2m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\right) \\ &+ \frac{n^{\frac{m}{2}-2}}{2^{\frac{m}{2}}}\left(\frac{m(m-2)(8m-5)}{36\Gamma(\frac{m}{2})}\right) + O(n^{\frac{m}{2}-\frac{5}{2}}) \end{aligned}$$

as desired. ■

As a simple corollary to the above theorem, we obtain an asymptotic expansion of the redundancy $\rho_n(\Omega_m)$.

Corollary 1 *For $m \geq 2$*

$$\begin{aligned} \rho_n(\Omega_m) = \log D_n(\Omega_m) &= \frac{m-1}{2}\log\left(\frac{n}{2}\right) + \log\left(\frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\right) \\ &+ \frac{\Gamma(\frac{m}{2})m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\cdot\frac{\sqrt{2}}{\sqrt{n}} \\ &+ \left(\frac{3 + m(m-2)(2m+1)}{36} - \frac{\Gamma^2(\frac{m}{2})m^2}{9\Gamma^2(\frac{m}{2}-\frac{1}{2})}\right)\cdot\frac{1}{n} + O\left(\frac{1}{n^{3/2}}\right) \end{aligned} \tag{9}$$

for large n.

Proof: By (7) we obtain

$$D_m(z) = \sum_{i=1}^{m} \binom{m}{i} D_i^*(z) = \sum_{i=1}^{m} \binom{m}{i} (B(z)-1)^i \ .$$

Thus, $D_m(z) = B^m(z) - 1$. Then, $D_n(m) = \frac{n!}{n^n}[z^n](B^m(z)-1)$. Using MAPLE, singularity analysis and the asymptotic expansion of $B^m(z)$ prove the theorem. In fact, `equivalent` directly produces

$$\begin{aligned} e^{-n}[z^n]B^m(z) &= \frac{n^{\frac{m}{2}-1}}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} + \frac{n^{\frac{m}{2}-\frac{3}{2}}}{2^{\frac{m}{2}-\frac{1}{2}}}\left(\frac{m}{3\Gamma(\frac{m}{2}-\frac{1}{2})}\right) \\ &+ \frac{n^{\frac{m}{2}-2}}{2^{\frac{m}{2}}}\left(\frac{m(m-2)(2m+1)}{36\Gamma(\frac{m}{2})}\right) + O(n^{\frac{m}{2}-\frac{5}{2}}) \ . \end{aligned}$$

Finally, we additionally observe that

$$\log(1 + a\sqrt{x} + bx + cx^{3/2}) = a\sqrt{x} + (b - \frac{1}{2}a^2)x + O(x^{3/2})$$

as $x \to 0$, and this completes the proof. ■

In particular, our asymptotic expressions for $D_n(m)$ and $D_n^*(m)$ should be compared to a cruder $D_n^*(m) = D_n(m)\left(1 - O(n^{-\frac{1}{2}})\right)$ derived in [13]. As an additional checking, we observe that for $m = 1$ both $D_n(1) = D_n^*(1) = 1$, as needed. Furthermore, for $m = 2$ we obtain $D_n(2) = \sqrt{n\pi/2} + 2/3 + \sqrt{\pi}/(12n) + O(n^{-3/2})$ and $D_n^*(2) = \sqrt{n\pi/2} - 4/3 + \sqrt{\pi}/(12n) + O(n^{-3/2})$, while for $m = 3$ we have: $D_n(3) = n + \sqrt{n\pi/2} + 2/3 + O(1/n)$ and $D_n^*(3) = n - \sqrt{2n\pi} + 5/3 + O(1/n)$, as easy to check by a direct computation.

3 Conclusion

In this paper, we present an asymptotic expansion for $D_n^*(m)$ that finds applications in universal coding, universal portfolio, hashing, random mapping, etc. (cf. [1, 2, 4, 14]). In fact, $\log D_n(m)$ represents the redundancy of the maximal probability code.

We observe that $D_n^*(m)$ satisfying recurrence (3) falls under the general recurrence (1) from the introduction. In fact, using the same arguments as in the proof of Theorem 1 one can solve recurrence (1) in terms of generating functions, namely:

$$X_m(z) = A(z) + 2B(z)X_{m-1}(z)$$

where

$$X_m(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k x_k^m \ , \qquad A(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k a_k \ .$$

This last recurrence can be solved by telescoping in terms of m, and then the singularity analysis will provide an asymptotic expansion, as discussed above.

ACKNOWLEDGEMENT

I thank P. Kirschenhofer and H. Prodinger for comments, and H-K. Hwang for pointing out some misprints in the earlier version of this paper. Finally, I am in debt to an anonymous referee for his/her comments, and D. Nogin for an efficient handing of this paper.

References

[1] R. Corless, G. Gonnet, D. Hare, D. Jeffrey and D. Knuth, On the Lambert W Function, *Adv. Computational Mathematics*, 5, 329–359, 1996.

[2] T. Cover and E. Ordentlich, Universal Portfolios with Side Information, *IEEE Trans. Information Theory*, 42, 348–363, 1996.

[3] P. Flajolet and A. Odlyzko, Singularity Analysis of Generating Functions, *SIAM J. Disc. Methods*, 3, 216-240, 1990.

[4] P. Flajolet, P. Grabner, P. Kirschenhofer, and H. Prodinger, On Ramanujan's Q-function, *J. Comp. and Appl. Math.*, 58, 103-116, 1995.

[5] P. Jacquet, Analytic Information Theory in Service of Queueing with Aggregated Exponential On/Off Arrivals, *Proc. 35-th Allerton Conference*, Monticello, 1997.

[6] P. Jacquet and W. Szpankowski, Asymptotic Behavior of the Lempel-Ziv Parsing Scheme and Digital Search Trees, *Theoretical Computer Science*, 144, 161–197, 1995.

[7] P. Jacquet and W. Szpankowski, Analytical Information Theory: Entropy Computations, Purdue University CSD-TR-96-085, 1996.

[8] D. Knuth, *The Art of Computer Programming: Fundamental Algorithms,* vol. 1., Addison-Wesley, Reading 1973.

[9] C. Krattenthaler and P. Slater, Asymptotic Redundancies for Universal Quantum Coding, preprint (cf. `http://radon.mat.univie.ac.at/People/kratt/papers.html`).

[10] G. Louchard and W. Szpankowski, On the Average Redundancy Rate of the Lempel-Ziv Code, *IEEE Trans. Information Theory*, 43, 2–8, 1997.

[11] B. Salvy, Examples of Automatic Asymptotic Expansions, *SIGSAM Bulletin*, 25, 4–17, 1991.

[12] Y. Shtarkov, Universal Sequential Coding of Single Messages, *Problems of Information Transmission*, 23, 175–186, 1987.

[13] Y. Shtarkov, T. Tjalkens and F.M. Willems, Multi-alphabet Universal Coding of Memoryless Sources, *Problems of Information Transmission*, 31, 114-127, 1995

[14] W. Szpankowski, On Asymptotics of Certain Sums Arising in Coding Theory, *IEEE Trans. Information Theory*, 41, 2087–2090, 1995

[15] W. Szpankowski, Techniques of the Average Case Analysis of Algorithms, in *Handbook on Algorithms and Theory of Computation* (Ed. M. Atallah), CRC 1998.