1996

On asymptotics of Certain Recurrences Arising in Multi-Alphabet Universal Coding

Wojciech Szpankowski
Purdue University, spa@cs.purdue.edu

Report Number:
96-086

ON ASYMPTOTICS OF CERTAIN RECURRENCES ARISING IN MULTI-ALPHABET UNIVERSAL CODING

Wojciech Szpankowski

Department of Computer Science
Purdue University
West Lafayette, IN 47907

CSD-TR 96-086
December 1996

ON ASYMPTOTICS OF CERTAIN RECURRENCES ARISING IN MULTI-ALPHABET UNIVERSAL CODING

January 4, 1997

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

Abstract

Ramanujan's Q-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: multi-alphabet 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 successfully to a variety of information theory problems.

Key Words: Source coding, multi-alphabet universal coding, redundancy, minimum description length, analytical information theory, singularity analysis, Ramanujan's Q-function, Lambert's W-function.

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

1 Introduction

Recently, Jacquet and Szpankowski [7] propose 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, 11, 12, 18, 19, 20]).

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 [19] that deals with certain sums arising in linear codes. Here, we deal with (multi-alphabet) universal coding [9, 14, 17]). A direct motivation for this research is the recent paper of Shtarkov *et al.* [18] (see also Shtarkov [17]) 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 [18] and [17] aim at designing an acceptable code such that the maximal redundancy over the alphabets is as small as possible.

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. [17]) or simple asymptotics ([18]) were derived. Our goal is to solve exactly the recurrence (in terms of certain generating functions) and to establish a full asymptotic expansion for it. We also aspire to present a mathematical tool to deal with this type of recurrences.

We establish the above announced result by analytical methods that resemble our approach from [19]. 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 two types of recurrences (cf. Section 4 for more details). Under some initial conditions, the sequence x_n satisfies for $n \geq 1$

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

or x_n^m for $m \geq 1$

$$x_n^m = a_n + \sum_{i=0}^{n} \binom{n}{i} \left(\frac{i}{n}\right)^i \left(1 - \frac{i}{n}\right)^{n-i} (x_i^{m-1} + x_{n-i}^{m-1}) \tag{2}$$

where a_n is a given sequence (the so called *additive* term), and m in (2) is an additional parameter, that is, x_n^m is a double-indexed sequence. In this paper we deal with recurrences of type (2) while in [19] we were more in the framework of (1) (cf. also [2]).

2 Main Results

In this section, we formulate precisely the problem and present our main results. Throughout, we shall try to use the notation from [18].

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 *redundancy* for 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

$$\rho_n(\Omega) := \inf_{\varphi_n} \sup_{\omega \in \Omega} \max_{x^n \in \mathcal{A}^n} \{\rho(x^n; \varphi_n, \omega)\} \ .$$

In [17] it is proved that

$$\rho_n(\Omega) = \log D_n(\Omega) \tag{3}$$

where

$$D_n(\Omega) = \sum_{x^n \in \Omega} P^*(x^n; \Omega) \qquad \text{with} \qquad P^*(x^n; \Omega) := \sup_{\omega \in \Omega} P(x^n; \omega) \ . \tag{4}$$

Furthermore, the optimal code φ_n, called the *maximal probability code*, is of length

$$|\varphi^*(x^n; \Omega)| = -\log\left(\frac{P^*(x^n; \Omega)}{D_n(\Omega)}\right) \ . \tag{5}$$

Hereafter, we restrict ourselves to **memoryless sources**. Then, it is not difficult to observe that

$$P^*(x^n; \Omega) := \prod_{a \in \mathcal{A}} \left(\frac{t_a}{n}\right)^{t_a} \tag{6}$$

where t_a is the number of $a \in \mathcal{A}$ occurrences in x^n. Indeed, the above is a consequence of a simple optimization. For example, for a binary alphabet this follows from

$$\max_{0 \le p \le 1} \{p^i (1-p)^{n-i}\} = \left(\frac{i}{n}\right)^i \left(1 - \frac{i}{n}\right)^{n-i} \ .$$

Our goal is to provide precise asymptotics of $D_n(m) := D_n(\Omega_m)$ for memoryless sources where Ω_m is a set of sources that generates *only* $m \le M$ symbols of the alphabet $\mathcal{A} =$

$\{1, 2, \ldots, M\}$. We should observe, however, that when $|\Omega_M| = 1$, then clearly $D_n(M) = 1$, and the maximal probability code coincides with the well-known Shannon-Fano code. We shall see below that a quite different situation arises when $|\Omega_m| > 1$ for some $m < M$.

Let us first derive a recurrence on $D_n(m)$. Following [17, 18] we observe that $D_n^*(1) = 1$ and

$$D_n(m) = \sum_{i=1}^{m} \binom{m}{i} D_n^*(i) \tag{7}$$

where $D_n^*(i)$ satisfies the following recurrence:

$$D_n^*(i) = \sum_{k=1}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} D_{n-k}^*(i-1) \ . \tag{8}$$

provided that $D_n^*(i) = 0$ whenever $i > n$.

For the completeness, we present a simple proof of (7)–(8). Observe first that one can write (4) as follows

$$\begin{aligned} D_n(\Omega_m) &= \sum_{i=1}^{m} \binom{m}{i} \sum_{x^n \in \mathcal{A}(i)} P^*(x^n; \Omega_i) \\ &= \sum_{i=1}^{m} \binom{m}{i} D_n^*(i) \end{aligned}$$

where $\mathcal{A}(i)$ represents a subset of $\mathcal{A}$ consisting of i symbols. Clearly, the last equation is (7). To derive (8), we consider an alphabet $\mathcal{A}(i-1)$ and assume that these $i-1$ symbols of $\mathcal{A}$ occur on $n-k$ positions of x^n. Thus, we deal with $D_{n-k}^*(i-1)$. On the remaining k positions we place the ith symbol with the (optimal) probability $\left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k}$, and this proves (8).

We now can summarize our main results. The first theorem deals with asymptotics of $D_n(m)$ and it provides a precise estimate of the redundancy $\rho_n(\Omega_m)$ expressed by (3).

Theorem 1 *For fixed $m \geq 1$ and large n, $D_n(m)$ attains the following asymptotic expansion:*

$$\begin{aligned} D_n(m) &= \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})} \left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{1}{2}} + \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2}-\frac{1}{2})} \left(\frac{m+3}{3}\right) \left(\frac{n}{2}\right)^{\frac{m}{2}-1} \\ &+ \sqrt{\pi} \left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{3}{2}} \left(\frac{3+(m-2)(2m^2+13m+24)}{72\Gamma(\frac{m}{2})} + \frac{\delta_{m,1}}{2\Gamma(\frac{1}{2})}\right) + O(n^{\frac{m}{2}-2}) \end{aligned} \tag{9}$$

where $\delta_{m,1}$ is the Kronecker delta (i.e., $\delta_{m,1} = 1$ for $m = 1$ and zero otherwise). The above implies that for $m \geq 2$

$$\rho_n(\Omega_m) = \log D_n(\Omega_m) = \frac{m-1}{2} \log\left(\frac{n}{2}\right) + \log\left(\frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\right) \tag{10}$$

$$
\begin{aligned}
&+ \quad \frac{\Gamma(\frac{m}{2})(m+3)}{3\Gamma(\frac{m}{2}-\frac{1}{2})} \cdot \frac{1}{\sqrt{n}} \\
&+ \quad \left(\frac{6+(m-2)(4m^2+26m+48)}{144} - \frac{\Gamma^2(\frac{m}{2})(m+3)^2}{18\Gamma^2(\frac{m}{2}-\frac{1}{2})} \right) \cdot \frac{1}{n} \\
&+ \quad O\left(\frac{1}{n^{3/2}}\right)
\end{aligned}
$$

where $\Gamma(z)$ *is the Euler gamma function (cf. [1]).*

We should observe that a precise asymptotic expansion for the redundancy $\rho_n(\Omega_m)$ is necessary for large or unbounded alphabets (e.g., alphabet of an image is 256 pixels and often is of the same order of magnitude as its dimensions). In such cases, one may expect a significant contribution from the lower order terms of $\rho_n(\Omega_m)$, as explicitly shown in (10). In passing, we also point out that our methodology from the next sections allows to extract a full asymptotic expansion for $D_n(m)$ (in the theorem above we only shown the first four terms).

The authors of [18] next consider the *minimum description length* defined as the length of

$$|\varphi^{**}(x^n)| = -\log\left(\frac{P^*(x^n)}{D_n(m)}\right)$$

where $P^*(x^n)$ is given in (5). It is further argued in [18] that for the maximal probability code one should replace the above with

$$|\tilde{\varphi}(x^n)| = -\log\left(\frac{\alpha_m}{\binom{M}{m}} \frac{P^*(x^n)}{D_n^*(m)}\right)$$

where $\alpha_m < 1$ is a certain coefficient (see [18] for details) and $D_n^*(m)$ is the quantity defined in (8). In order to compare the above two formulæ one needs a precise evaluation of $D_n(m) - D_n^*(m)$ which is offered in the next theorem.

Theorem 2 *For fixed* $m \geq 1$ $D_n^*(m)$ *attains the following asymptotics for large* n

$$
\begin{aligned}
D_n^*(m) &= \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{1}{2}} - \frac{\sqrt{\pi}}{\Gamma(\frac{m}{2}-\frac{1}{2})}\left(\frac{2m-3}{3}\right)\left(\frac{n}{2}\right)^{\frac{m}{2}-1} \qquad (11)\\
&+ \quad \sqrt{\pi}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{3}{2}} \frac{(m-1)(8m^2-37m+45)}{72\Gamma(\frac{m}{2})} + O(n^{\frac{m}{2}-2}) \ .
\end{aligned}
$$

In particular, for $m \geq 2$

$$
\begin{aligned}
D_n(m) - D_n^*(m) &= \frac{\sqrt{\pi} m}{\Gamma(\frac{m}{2}-\frac{1}{2})}\left(\frac{n}{2}\right)^{\frac{m}{2}-1} \qquad (12)\\
&+ \quad \sqrt{\pi}\left(\frac{n}{2}\right)^{\frac{m}{2}-\frac{3}{2}} \frac{12m(-m^2+9m-14)}{144\Gamma(\frac{m}{2})} + O(n^{\frac{m}{2}-2}) \ .
\end{aligned}
$$

In particular, our asymptotic expression (12) should be compared to a cruder $D_n^*(m) = D_n(m)\left(1 - O(n^{-\frac{1}{2}})\right)$ proposed in [18]. As before, we can derive a full asymptotic expansion for $D_n(m) - D_n^*(m)$ using the approach from the next section.

3 Analysis

In this section we prove Theorems 1 and 2 using certain tools of analytical analysis of algorithms, notably generating functions (cf. [16, 20]) and singularity analysis (cf. [3]). We shall follow our approach from [19] (cf. also [20]).

We start with $D_n^*(m)$ given by (8). Let us introduce a new sequence $\widehat{D}_n^*(m)$ defined as

$$\begin{aligned} \widehat{D}_n^*(m) &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k} D_{n-k}^*(m-1) && (13)\\ &= D_n^*(m) + D_n^*(m-1) \ . && (14) \end{aligned}$$

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 a special function $B(z)$ defined as (cf. [19])

$$B(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} z^k = \frac{1}{1 - T(z)} \qquad (15)$$

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. [10]). It is also related to Lambert's W-function 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]) which finds many applications in hashing, random mappings, and memory conflict (cf. [19, 20]) for further references). We shall discuss some properties of $T(z)$ and $B(z)$ below.

Let us now return to (13). 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

$$\widehat{D}^*_m(z) = B(z)D^*_{m-1}(z) \tag{16}$$

Finally, using (14) and (16) we obtain

$$D^*_m(z) = (B(z)-1)^{m-1}D^*_1(z) = B(z)(B(z)-1)^{m-1} \ . \tag{17}$$

Then, the desired $D^*_n(m)$ can be found as

$$D^*_n(m) = \frac{n!}{n^n}[z^n]\left((B(z)-1)^{m-1}B(z)\right) \tag{18}$$

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. This, as it is well-known, depends on 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 [8, 16, 19, 20] that $T(z)$ has a singularity at $z = e^{-1}$, and around this point it has the following expansion:

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

Using MAPLE, we can also expand $B(z)$ around $z = e^{-1}$ leading to

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

In a similar manner, we can obtain asymptotic expansions of $(B(z)-1)^m$ for any m. For example, MAPLE gives

$$\begin{aligned}
(B(z)-1)^2 &= \frac{1}{2(1-ez)} + \frac{2}{3}\frac{\sqrt{2}}{\sqrt{1-ez}} + \frac{13}{36} - \frac{23}{270}\sqrt{1-ez} + O(1-ez) \\
(B(z)-1)^3 &= -\frac{1}{4}\frac{\sqrt{2}}{(1-ez)^{3/2}} - \frac{1}{1-ez} - \frac{29}{48}\frac{\sqrt{2}}{\sqrt{1-ez}} - \frac{23}{270} + O(\sqrt{1-ez}) \\
(B(z)-1)^4 &= \frac{1}{4}\frac{1}{(1-ez)^2} + \frac{2}{3}\frac{\sqrt{2}}{(1-ez)^{3/2}} + \frac{5}{4}\frac{1}{1-ez} + \frac{107}{270}\frac{\sqrt{2}}{\sqrt{1-ez}} + O(1)
\end{aligned}$$

and so forth.

To extract the coefficients at z^n of (18), 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. Thus we use the above, and the following (for more details see [3]):

$$\begin{aligned}
[z^n]\left(\frac{1}{\sqrt{1-ez}}\right) &= \frac{e^n}{\sqrt{\pi n}}\left(1-\frac{1}{8n}+O(n^{-2})\right) , \\
[z^n]\left(\sqrt{1-ez}\right) &= -\frac{e^n}{\sqrt{\pi n^3}}\left(\frac{1}{2}+\frac{3}{16n}O(n^{-2}\right) \\
[z^n]\left(\frac{1}{1-ez}\right) &= e^n , \\
\frac{n!}{n^n} &= e^{-n}\sqrt{2\pi n}\left(1+\frac{1}{12n}+O(n^{-2})\right) .
\end{aligned}$$

Plugging this into (18), one can derive the asymptotic expansion for $D_n^*(m)$. Actually, we use directly MAPLE (more precisely, `equivalent` program thanks to B. Salvy [15]) to extract coefficients of $(B(z)-1)^{m-1}B(z)$. This leads to formula (11) of Theorem 2.

It remains now to derive asymptotics of $D_n(m)$ given by (7). By (17) we obtain

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

Thus,

$$D_m(z)=\frac{B(z)}{B(z)-1}\left(B^m(z)-1\right)$$

Clearly, we can obtain an asymptotic expansion for $D_n(m)$ in a similar manner as above, that is,

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

Using MAPLE, singularity analysis and the asymptotic expansion of $B^m(z)$ as shown above, we finally find formula (9) of Theorem 1. For $m=1$ we also observe that

$$\frac{n!}{n^n}[z^n]\frac{B(z)}{(B(z)-1)}=\frac{-1}{n+1},$$

for $n\geq 2$, by Lagrange inversion formula. To prove (10) we additionally need the following:

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

as $x\to 0$. Then, (12) of Theorem 2 follows.

4 Conclusion

In this paper we provide a precise asymptotic expansion for $D_n(m)$ that plays a crucial role in the multi-alphabet universal coding. In fact, $\log D_n(m)$ represents the redundancy of the maximal probability code.

We further observe that $D_n^*(m)$ is closely related to $D_n(m)$ which satisfies recurrence (8). This recurrence falls under our general recurrence (2) from the introduction. In fact, using the same arguments as in Section 3 one can solve recurrence (2) in terms of generating functions, namely:

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

where

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

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.

In a similar manner, one can solve recurrence (1) which in terms of the above generating functions becomes

$$X(z) = \frac{A(z)}{1 - 2B(z)} \,.$$

Again, using the asymptotic expansion of $B(z)$ (and possibly $A(z)$) around $z = e^{-1}$ together with the singularity analysis will lead to an asymptotic expansion of x_n.

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.

References

[1] M. Abramowitz, and I. Stegun (Eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, John Wiley & Sons, New York 1972.

[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] L. Guibas and A. Odlyzko, Maximal Prefix-Synchronized Codes, *SIAM J. Appl. Math*, 35, 401-418, 1978.

[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] R. Krichevsky, *Universal Compression and Retrieval*, Kluwer Academic Publishers, Dordrecht 1994.

[10] J.H. van Lint, *Introduction to Coding Theory*, Springer-Verlag, New York 1982

[11] G. Louchard and W. Szpankowski, Average Profile and Limiting Distribution for a Phrase Size in the Lempel-Ziv Parsing Algorithm , *IEEE Trans. Information Theory*, 41, 478-488, 1995.

[12] G. Louchard and W. Szpankowski, On the Average Redundancy Rate of the Lempel-Ziv Code, *IEEE Trans. Information Theory*, 43, 1–7, 1997.

[13] A. Odlyzko, Asymptotic Enumeration, in *Handbook of Combinatorics*, Vol. II, (Eds. R. Graham, M. Götschel and L. Lovász), Elsevier Science, 1063-1229, 1995.

[14] B. Ryabko, Twice-Universal Coding, *Problems of Information Transmission*, 173–177, 1984.

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

[16] R. Sedgewick and P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison-Wesley Publishing Company, Reading Mass., 1995.

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

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

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

[20] W. Szpankowski, Techniques of the Average Case Analysis of Algorithms, in *Handbook on Algorithms and Theory of Computation* (Ed. M. Atallah), CRC 1997.