Information Theory by Analytic Methods: Redundancy Rate Problem*

W. Szpankowski
Department of Computer Science
Purdue University
W. Lafayette, IN 47907

July 9, 1999

*Research supported by NSF Grants NCR-9415491 and C-CR 9804760

Outline

1. The Redundancy Rate Problem: Basic definitions
2. Analytic Information Theory
3. Survey of Precise Results for the Redundancy Problem:
   (a) Shannon-Fano Code
   (b) Huffman Code
   (c) Memoryless and Markovian Sources (minimax)
   (d) Renewal Processes
   (e) Lempel-Ziv Code
   (f) Tunstall's Code
4. Analytic Methods at Work:
   (a) Fourier Analysis and Distribution of Sequences Mod 1
   (*Redundancy of Shannon-Fano and Huffman Codes*)
   (b) Tree Generating Functions and Singularity Analysis
   (*Minimax Redundancy of a Memoryless Source*)
   (c) Mellin Transform and Saddle Point Method
   (*Redundancy of a Renewal Process*)
   (d) Analytic Depoissonization
   (*Lempel-Ziv Code*)

Definitions

The **redundancy-rate problem** of universal coding for a class of sources consists in determining by how much the actual code length exceeds the optimal (ideal) code length.

A code
$$C_n : \mathcal{A}^n \to \{0,1\}^*$$
is a mapping from the set $\mathcal{A}^n$ of all sequences of length n over the alphabet $\mathcal{A}$ to the set $\{0,1\}^*$ of binary sequences.

Given a probabilistic source model and a code C_n we let:

- $P(x_1^n)$ be the probability of the message $x_1^n = x_1 \ldots x_n$,
- $L(C_n, x_1^n)$ be the code length for x_1^n,
- Entropy $H_n(P) = -\sum_{x_1^n} P(x_1^n) \lg P(x_1^n)$,
- the "ideal" code length: $-\lg P(x_1^n)$.

Information-theoretic quantities are expressed in binary logarithms written $\lg := \log_2$. We also write $\log := \ln$.

Various Redundancies

The **pointwise redundancy** $R_n(C_n, P; x_1^n)$ and the **average redundancy** $\bar{R}_n(C_n, P)$ are defined as

$$
\begin{aligned}
R_n(C_n, P; x_1^n) &= L(C_n) + \lg P(x_1^n) \\
\bar{R}_n(C_n) &= \mathbf{E}_{X_1^n}[R_n(C_n, P; X_1^n)] \\
&= \mathbf{E}[L(C_n, X_1^n)] - H_n(P)
\end{aligned}
$$

where $\mathbf{E}$ denotes the expectation. The **maximal** redundancy is defined as

$$R^*(C_n, P) = \max_{x_1^n}\{R_n(C_n, P; x_1^n)\}.$$

The pointwise redundancy can be negative, maximal and average redundancy cannot (see next slide).

The **strong redundancy-rate problem** consists in determining for a class $\mathcal{S}$ of source models the rate of growth of the minimax quantities

$$
\begin{aligned}
\bar{R}_n^*(\mathcal{S}) &= \min_{C_n} \max_{P\in\mathcal{S}}\{\bar{R}_n(C_n, P)\}(= o(n)), \\
R_n^*(\mathcal{S}) &= \min_{C_n} \max_{P\in\mathcal{S}}\{R_n^*(C_n, P)\}(= o(n))
\end{aligned}
$$

as $n \to \infty$.

Shannon's Lower Bound

Fact: *For any code, the average code length* $\mathbf{E}[L(C_n, X_1^n)]$ *cannot be smaller than the entropy of the source* $H_n(P)$, *that is,*

$$\mathbf{E}[L(C_n, X_1^n)] \geq H_n(P).$$

Sketch of Proof: Let $K = \sum_{x_1^n} 2^{-L(x_1^n)} \leq 1$, and $L(C_n, x_1^n) := L(C_n)$. Then

$$\begin{aligned} \mathbf{E}[L(C_n, X_1^n)] - H_n(P) &= \sum_{x_1^n \in \mathcal{A}^n} P(x_1^n) L(x_1^n) \\ &+ \sum_{x_1^n \in \mathcal{A}^n} P(x_1^n) \log P(x_1^n) \\ &= \sum_{x_1^n \in \mathcal{A}^n} P(x_1^n) \log \frac{P(x_1^n)}{2^{-L(x_1^n)}/K} - \log K \\ &\geq 0 \end{aligned}$$

since the first term is a divergence and cannot be negative (or $\log x \leq x - 1$) while $K \leq 1$ by Kraft's inequality.

Analytic Information Theory

The redundancy rate problem is typical of a situation where **second-order asymptotics** play a crucial role since the leading term of $L(C_n)$ is known to be nH, where H is the entropy rate. This problem is an ideal candidate for **analytic information theory** that applies analytic tools to information theory.

As argued by Andrew Odlyzko: "*Analytic methods are extremely powerful and when they apply, they often yield estimates of unparalleled precision.*"

In *1997 Shannon Lecture*, Jacob Ziv presented compelling arguments for "backing off" to a certain degree from the (first-order) asymptotic analysis of information systems in order to predict the behaviour of real systems where we always face *finite* (and often small) lengths (of sequences, files, codes, etc.) One way of overcoming these difficulties is to **increase the accuracy of asymptotic analysis** and replace first-order analyses by more **complete asymptotic expansions**, thereby extending their range of applicability to smaller values while providing more accurate analyses (like constructive error bounds, large deviations, local or central limit laws).

Survey: Shannon-Fano Code

Shannon-Fano code assigns code of length $\lceil \lg P(x_1^n) \rceil$ to x_1^n (it is assumed that $P(x_1^n)$ is known).

Consider a binary memoryless source with p denoting the probability of generating 0. For a block of length n, the average redundancy $\bar{R}_n^{SF}$ is

$$\begin{aligned}\bar{R}_n^{SF} &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left(\lceil -\log_2(p^k(1-p)^{n-k}) \rceil \right. \\ &+ \left. \log_2(p^k(1-p)^{n-k}) \right).\end{aligned}$$

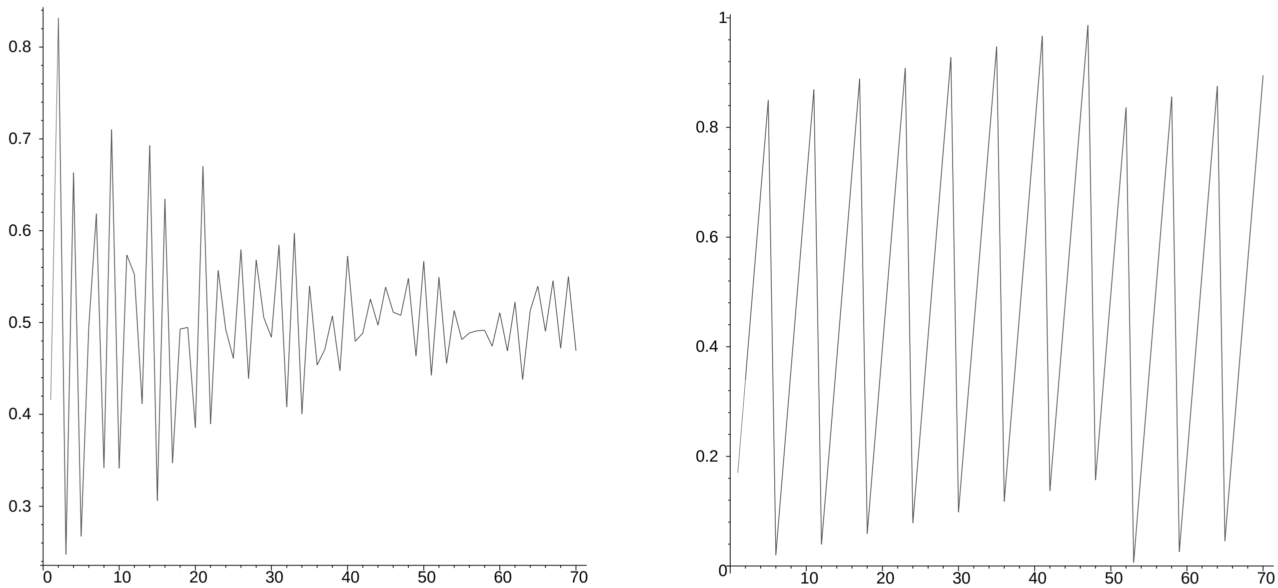
Let

$$\alpha = \log_2\left(\frac{1-p}{p}\right), \quad \beta = \log_2\left(\frac{1}{1-p}\right).$$

For the Shannon–Fano code we prove that its average redundancy is as $n \to \infty$

$$R_n^{SF} = \begin{cases} \frac{1}{2} & \alpha \quad \text{irrational} \\ \frac{1}{2} - \frac{1}{M}\left(\langle Mn\beta \rangle - \frac{1}{2}\right) & \alpha = \frac{N}{M} \end{cases}$$

where $\langle x \rangle = x - \lfloor x \rfloor$ is the fractional part of x, and N, M are integers such that $\gcd(N, M) = 1$.



(a) (b)

Figure 1: Shannon–Fano code redundancy versus block size n for: (a) irrational $\alpha = \log_2(1-p)/p$ with $p = 1/\pi$; (b) rational $\alpha = \log_2(1-p)/p$ with $p = 1/9$.

Survey: Huffman Code

As before we consider a binary memoryless source emitting 0 and 1 with probabilities p and $q = 1 - p$, respectively. For a block of length n, we construct its **Huffman code** (through the associated Huffman tree).

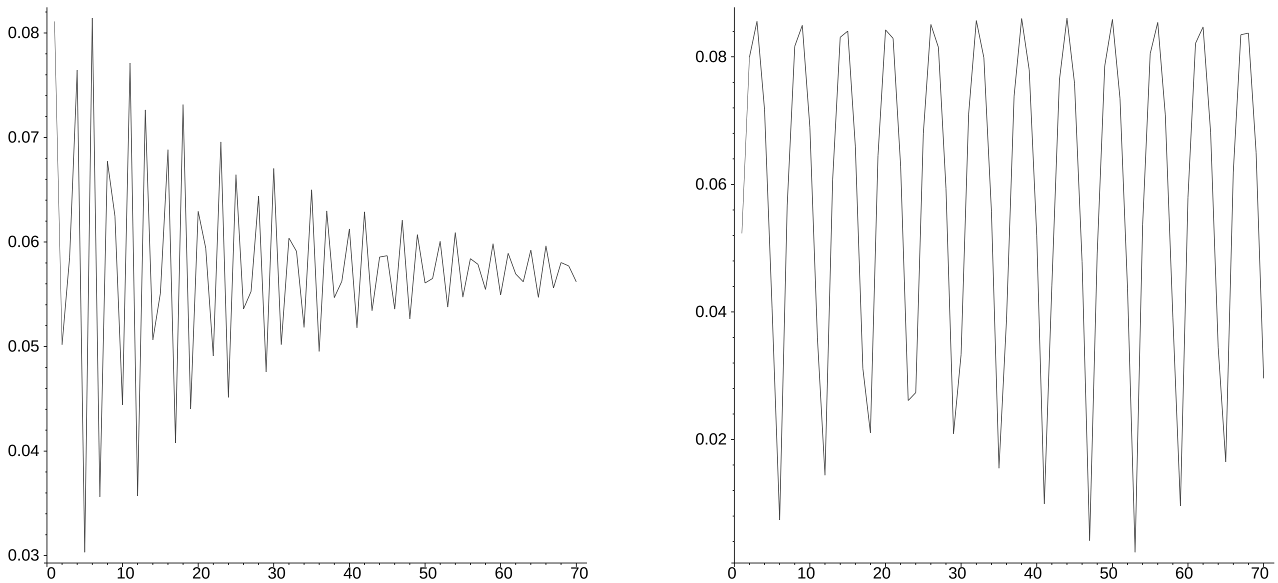
Using Stubley's result (1994), we conclude that the average redundancy $\bar{R}_n^H$ of the Huffman code is

$$\bar{R}_n^H = 1 + \bar{R}_n^{SF} - 2\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} 2^{-\langle \alpha k + \beta n \rangle} + O(\rho^n)$$

where $\rho < 1$ and $\bar{R}_n^{SF}$ is the average redundancy of the Shannon-Fano code. As $n \to \infty$ this becomes

$$R_n^H = \begin{cases} \frac{3}{2} - \frac{1}{\ln 2} = 0.057304\ldots & \alpha \text{ irrational} \\ \frac{3}{2} - \frac{1}{M}\left(\langle \beta M n \rangle - \frac{1}{2}\right) - \frac{1}{M(1-2^{-1/M})} 2^{-\langle n\beta M \rangle / M} & \alpha = \frac{N}{M} \end{cases}$$

with the notation as above.



(a) (b)

Figure 2: Huffman's code redundancy versus block size n for: (a) irrational $\alpha = \log_2(1-p)/p$ with $p = 1/\pi$; (b) rational $\alpha = \log_2(1-p)/p$ with $p = 1/9$.

The maximum Huffman redundancy is

$$\max\{\bar{R}_n^H\} = 1 - \frac{1 + \log\log 2}{\log 2} = \lg(2(\lg e)/e) = 0.08607\ldots,$$

as $n \to \infty$.

Shtarkov's Mimimax Result

Shtarkov in 1978 proved that the minimax redundancy

$$\lg\left(\sum_{x_1^n}\sup_{\omega\in\mathcal{S}}P(x_1^n,\omega)\right)\le R_n^*(\mathcal{S})\le \lg\left(\sum_{x_1^n}\sup_{\omega\in\mathcal{S}}P(x_1^n,\omega)\right)+1.$$

Indeed, for the **lower bound** Shtarkov considered the following probability distribution

$$q(x_1^n):=\frac{\sup_\omega P(x_1^n,\omega)}{\sum_{x_1^n}\sup_\omega P(x_1^n,\omega)}$$

By Kraft's inequality there exists $\tilde{x}_1^n$ such that (for uniquely decodable codes C_n)

$$-L(C_n)\le \lg q(\tilde{x}_1^n),$$

which implies the lower bound. For the **upper bound**, Shtarkov proposed a code C_n of length

$$L(\widetilde{C}_n)=\left\lceil \lg\big(\sum_{x_1^n}\sup_\omega P(x_1^n),\omega\big)-\lg P(x_1^n)\right\rceil,$$

which gives the desired upper bound.

Finitely Parametrizable Class of Processes

If $\mathcal{M}$ is i.i.d. or the class of Markov chains, or more generally the process belongs to a finitely parametrizable class of dimension K, then Rissanen proved that the average redundancy $\bar{R}_n$ and the minimax redundancy $\bar{R}_n^*$

$$\bar{R}_n(\mathcal{M}) \sim R_n^*(\mathcal{M}) \sim \frac{K}{2} \lg n.$$

as $n \to \infty$. It was also found that the next term of $\bar{R}_n(\mathcal{S})$ and of $R_n^*(\mathcal{S})$ is $O(1)$.

We will prove a full asymptotic expansion of $R_n^*(\mathcal{M})$ for memoryless sources over an m-ary alphabet; e.g.,

$$\begin{aligned} R_n^*(\mathcal{M}) &= \frac{m-1}{2} \lg\left(\frac{n}{2}\right) + \lg\left(\frac{\sqrt{\pi}}{\Gamma(m/2)}\right) + \cdots \\ &+ \frac{\Gamma(m/2)m}{3\Gamma(m/2-1/2)} \cdot \frac{\sqrt{2}}{\sqrt{n}} \\ &+ \left(\frac{3+m(m-2)(2m+1)}{36} - \frac{\Gamma^2(m/2)m^2}{9\Gamma^2(m/2-1/2)}\right) \cdot \frac{1}{n} \\ &+ O\left(\frac{1}{n^{3/2}}\right) \end{aligned}$$

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

Tunstall's Code

Savari and Gallager 1997 and Savari 1998 analyzed Tunstall's variable-to-fixed codes for memoryless and Markovian sources. For memoryless binary source, it was proved that

$$\bar{R}_n(\mathcal{T}) = -\frac{H \lg H + 0.5 h_2}{\lg n}$$

provided $B = \frac{\log p}{\log q}$ is irrational, where $h_2 = p \log^2 p + q \log^2 q$. The case of B rational was not discussed, but one expects some fluctuation in this case.

Renewal Process

Csiszàr and Shields have studied order r Markov renewal sequences in which a 1 is inserted every $T_0, T_1, \ldots$ of 0's, where $\{T_i\}$ is either an i.i.d. or Markov renewal or r-order Markov renewal process. We denote such sources as $\mathcal{R}_r$.

Csiszàr and Shields proved that

$$\bar{R}_n(\mathcal{R}_r) = R^*(\mathcal{R}_r) = \Theta(n^{(r+1)/(r+2)})$$

for $r = 1, 2, \ldots$ which specializes to $\Theta(\sqrt{n})$ when $r = 0$.

We will prove here (Flajolet & Szpankowski 1998) a precise asymptotic expansion of $R_n^*(\mathcal{R}_0)$ for the renewal processes, namely

$$R_n^*(\mathcal{R}_0) = \frac{2}{\log 2}\sqrt{cn} - \frac{5}{8}\lg n + \frac{1}{2}\lg\log n + O(1)$$

where $c = \frac{\pi^2}{6} - 1 \approx 0.645$.

Lempel-Ziv Code

Louchard and Szpankowski 1997, Savari 1997, Wyner 1998, and Jacquet and Szpankowski 1995 proved that the Lempel-Ziv codes in the class of i.i.d. and Markov processes have either rate

- $\Theta(n/\log n)$ for LZ'78
- $\Theta(n \log\log n/\log n)$ for LZ'77 code.

More precisely, for LZ'78 Louchard and Szpankowski 1997 showed that (binary alphabet with 0's generated with probability p)

$$\begin{aligned}\bar{R}_n(\mathcal{LZ}) &= H\left(2-\gamma-\frac{1}{2H}h_2+\omega-\delta(n)\right)\frac{n}{\log n} \\ &+ O\left(\frac{n\log\log n}{\log^2 n}\right)\end{aligned}$$

where $H = -p\log p - q\log q > 0$ is the entropy rate, $\gamma = 0.577\ldots$ is the Euler constant, $h_2 = p\log^2 p + q\log^2 q$, and

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

The function $\delta(x)$ that fluctuates with mean zero and a tiny amplitude for $\log p/\log q$ rational, but satisfies $\lim_{x\to\infty}\delta(x) = 0$ otherwise.

Shields' Result

Shields proved that there is no function $\rho(n) = o(n)$ which is a weak redundancy rate bound for the class of all ergodic processes.

ANALYTIC METHODS: Fourier Analysis and Sequence Distribution Modulo 1

We consider here redundancy of the **Shannon-Fano** block code and the **Huffman** block code for a memoryless source generating a block of length n with the binomial distribution.

Let $p(k) = p^k(1-p)^{n-k}$, where p is the probability of generating 0. Redundancy of the **Shannon-Fano** code is

$$\begin{aligned}\bar{R}_n^{SF} &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \left(\lceil -\lg p(k) \rceil + \lg p(k)\right) \\ &= 1 - \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \langle \alpha k + \beta n \rangle\end{aligned}$$

$\langle x \rangle = x - \lfloor x \rfloor$ is the fractional part of x and

$$\begin{aligned}\alpha &= \lg\left(\frac{1-p}{p}\right), \\ \beta &= \lg\left(\frac{1}{1-p}\right).\end{aligned}$$

Shannon-Fano Code – Irrational Case

Throughout, we shall use the following Fourier series; for real x

$$\begin{aligned}\langle x\rangle &= \frac{1}{2}-\sum_{m=1}^{\infty}\frac{\sin 2\pi m x}{m\pi}\\ &= \frac{1}{2}-\sum_{m\in\mathbf{Z}-\{0\}} c_m e^{2\pi i m x}, \qquad c_m=-\frac{i}{2\pi m},\end{aligned}$$

where $\mathbf{Z}$ is the set of all integers. Hereafter, we shall write $\sum_{m\neq 0} := \sum_{m\in\mathbf{Z}-\{0\}}$.

Irrational Case:

$$\begin{aligned}\bar{R}_n^{SF} &= \frac{1}{2}+\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\sum_{m\neq 0} c_m e^{2\pi i m(\alpha k+\beta n)}\\ &= \frac{1}{2}+\sum_{m\neq 0} c_m e^{2\pi i m\beta n}\left(p e^{2\pi i m\alpha}+q\right)^n .\end{aligned}$$

How to prove that the last term is $o(1)$ for α irrational?

Bernoulli Uniformly Distributed Sequences Modulo 1

Definition 1. [B-u.d. mod 1] *A sequence $x_n \in \mathbf{R}$ is said to be Bernoulli uniformly distributed modulo 1 (in short: B-u.d. mod 1) if for $0 < p < 1$*

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \chi_I(\langle x_k \rangle) = \lambda(I)$$

holds for every interval $I \subset \mathbf{R}$, where $\chi_I(x_n)$ is the characteristic function of I (i.e., it equals to 1 if $x_n \in I$ and 0 otherwise) and $\lambda(I)$ is the Lebesgue measure of I.

Theorem 1. *Let $0 < p < 1$ be a fixed real number and suppose that the sequence x_n is B-uniformly distributed modulo 1. Then for every Riemann integrable function $f : [0,1] \to \mathbf{R}$ we have*

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f(\langle x_k + y \rangle) = \int_0^1 f(t)\,dt,$$

where the convergence is uniform for all shifts $y \in \mathbf{R}$.

Proof. Standard; cf. Drmota and Tichy (1997) or Kuipers and Niederreiter (1974).

Weyl's Criterion

Theorem 2. [Weyl's Criterion] *A sequence x_n is B-u.d. mod 1 if and only if*

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{2\pi i m x_k} = 0$$

holds for all non-zero $m \in \mathbf{Z} - \{0\}$.

Proof. The proof again is standard. Basically, it is based on the fact that by Weierstrass's *approximation theorem* every Riemann integrable function f of period 1 can be uniformly approximated by a trigonometric polynomial (i.e., a finite combination of functions of the type $e^{2\pi i m x}$).

Finishing the Irrational Case

In our case, we must show that $\langle \alpha k \rangle$ is B-u.d. mod 1. By Weyl's criterion

$$\begin{aligned} \lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} e^{2\pi i m (k\alpha)} &= \lim_{n\to 0} \left(p e^{2\pi i m \alpha} + q \right)^n \\ &= 0 \end{aligned}$$

provided α is irrational. Hence, by the previous theorem, with $f(t) = t$ and $y = \beta n$, we immediately obtain

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \langle \alpha k + \beta n \rangle = \int_0^1 t dt = \frac{1}{2}.$$

This proves that for α irrational

$$R_n^{SF} = \frac{1}{2} + o(1).$$

It can be proved (thanks to M. Drmota) that for almost all irrational α the rate of convergence in the above is $O\left(\frac{\log^{1+\delta} n}{\sqrt{n}}\right)$ for some $\delta > 0$.

Shannon-Fano Redundancy – Rational Case

Now we assume that $\alpha = N/M$ where N, M are integers such that $\gcd(N, M) = 1$. Denote $p_{n,k} = \binom{n}{k} p^k q^{n-k}$. We proceed as follows:

$$
\begin{aligned}
S_n &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \left\langle k\frac{N}{M} + \beta n \right\rangle \\
&= \sum_{\ell=0}^{M-1} \sum_{m:\ k=\ell+mM\le n} p_{n,k} \left\langle \ell\frac{N}{M} + N + \beta n \right\rangle \\
&= \sum_{\ell=0}^{M-1} \left\langle \frac{\ell}{M} + \beta n \right\rangle \sum_{m:\ k=\ell+mM\le n} p_{n,k}.
\end{aligned}
$$

We will prove that the last sum is well approximated by $1/M$.

Useful Lemma

Lemma 1. *For fixed $\ell \leq M$ and M, there exist $\rho < 1$ such that*

$$\sum_{m:\ k=\ell+mM\leq n} \binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{M} + O(\rho^n).$$

Proof. Let $\omega_k = e^{2\pi i k/M}$ for $k = 0, 1, \ldots, M-1$ be the Mth root of unity. It is well known that

$$\frac{1}{M}\sum_{k=0}^{M-1} \omega_k^n = \begin{cases} 1 & \text{if } M|n \\ 0 & \text{otherwise.} \end{cases}$$

where $M|n$ means that M divides n. In view of this, we can write

$$\begin{aligned} \sum_{m:\ k=\ell+mM\leq n} \binom{n}{k} p^k q^{n-k} &= \frac{1 + (p\omega_1 + q)^{n-\ell} + .. + (p\omega_{M-1} +}{M} \\ &= \frac{1}{M} + O(\rho^n), \end{aligned}$$

since $|(p\omega_r + q)| = p^2 + q^2 + 2pq\cos(2\pi r/M) < 1$ for $r \neq 0$.

Finishing the Rational Case

Continuing the derivation and using the above lemma we obtain

$$\begin{aligned} S_n &= \frac{1}{M}\sum_{\ell=0}^{M-1}\left(\frac{1}{2}-\sum_{m\neq 0} c_m e^{2\pi i m(\ell/M+\beta n)}\right) \\ &= \frac{1}{2}-\sum_{m\neq 0} c_m e^{2\pi i m n\beta}\frac{1}{M}\sum_{\ell=0}^{M-1} e^{2\pi i m\frac{\ell}{M}} \\ &= \frac{1}{2}-\frac{1}{M}\sum_{m=kM\neq 0} c_{kM} e^{2\pi i k M\beta n} \\ &= \frac{1}{2}-\frac{1}{M}\left(\frac{1}{2}-\langle \beta n M\rangle\right). \end{aligned}$$

Huffman Redundancy

We only need to analyze

$$T_n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} 2^{-\langle \alpha k + \beta n \rangle}.$$

The **irrational** case is easy since a direct application of our previous result, with $f(t) = 2^{-t}$ a nd $y = \beta n$, yields

$$\lim_{n\to\infty} T_n = \int_0^1 2^{-t} dt = \frac{1}{2 \log 2}$$

for α irrational.

Huffman Redundancy – Rational Case

We could use Fourier series again, but instead we generalize our previous approach in the following theorem (proposed by M. Drmota):

Theorem 3. *Let* $0 < p < 1$ *be a fixed real number and suppose that* $\alpha = \frac{N}{M}$ *is a rational number with* $\gcd(N, M) = 1$. *Then, for every bounded function* $f : [0,1] \to \mathbf{R}$ *we have*

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} f(\langle k\alpha + y \rangle) \quad = \quad \frac{1}{M} \sum_{l=0}^{M-1} f\left(\frac{l}{M} + \frac{\langle My \rangle}{M} \right)$$

uniformly for all $y \in \mathbf{R}$ *and some* $\rho < 1$.

Setting $f(t) = 2^{-t}$ we obtain

$$\begin{aligned} T_n &= \frac{1}{M} \sum_{l=0}^{M-1} 2^{-(l/M) - (\langle M\beta n \rangle / M)} + O(\rho^n) \\ &= \frac{1}{M} 2^{-\langle M\beta n \rangle / M} \frac{1 - 2^{-M/M}}{1 - 2^{-1/M}} + O(\rho^n) \\ &= \frac{1}{2M(1 - 2^{-1/M})} 2^{-\langle M\beta n \rangle / M} + O(\rho^n) \end{aligned}$$

for $\rho < 1$.

ANALYTIC METHODS: Tree Generating Function and Singularity Analysis

We consider here the minimax redundancy for a memoryless source over an m-ary alphabet $\mathcal{A}(m)$. Shtarkov's result implies that

$$R_n^* = \log D_n(m)$$

where $D_n(m)$ satisfies

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

where $D_0^*(1) = 0$, $D_n^*(1) = 1$ for $n \geq 1$, and for $i > 1$ we have

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

Derivation of the Recurrence on $D_n^*(m)$

Observe first that one can write

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

where $\mathcal{A}(i)$ represents a subset of $\mathcal{A}$ consisting of i symbols. Indeed, we count separately sequences consisting of only symbols form $\mathcal{A}(i)$.

To derive the recurrence of $D_n^*(i)$ we argue as follows: Consider an alphabet $\mathcal{A}(i-1)$ and assume that these $i-1$ symbols of $\mathcal{A}(i-1)$ 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

$$\sup_{\omega} P^*(x_1^n, \omega) = \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k}$$

and establishes the recurrence:

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

Main Result

Theorem 4. *For fixed $m \geq 1$ the quantity $D_n^*(m)$ attains the following asymptotics*

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

for large n.

Corollary 1. *For $m \geq 2$*

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

for large n.

Sketch of Proof of Theorem 1

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

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

2. Observe that

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

This is a convolution of two sequences $\{k^k/k!\}$ and $\{k^k/k! D_k^*(m-1)$.

3. Define

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

The **tree function** $T(z)$ is defined as a solution to

$$T(z) = ze^{T(z)}$$

which is 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. 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 which finds many applications in hashing, random mappings, and memory conflict.

Related to the tree function is

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

which we need in the analysis.

4. From the recurrence and our definitions we obtain

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

Thus

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

since $D^*_1(z) = B(z) - 1$. So

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

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

5. Properties of $T(z)$. The tree function has an algebraic singularity at $z = e^{-1}$. This can be seen if we view the functional equation of $T(z)$ is a definition of an implicit function of

$$z(T) = Te^{-T}.$$

This function achieves its maximum value $z = e^{-1}$ at $T = 1$, and by the **implicit-function theorem** it can not be inverted. Thus, it has an algebraic singularity at this point.

6a. Asymptotics. We know that

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

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

$$B(z) = \frac{1}{\sqrt{2(1-ez)}} + \frac{1}{3} - \frac{\sqrt{2}}{24}\sqrt{(1-ez)} + \frac{4}{135}(1-ez) + \cdots$$

Singularity analysis of Flajolet and Odlyzko, allows to compute separately the coefficients for every function involved in the above asymptotic expansion. For example,

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

6b. Thus

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

as desired.

Sketch of Proof of Corollary 1

From our definition we have

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

Thus

$$D_m(z) = B^m(z) - 1.$$

Then, $D_n(m) = \frac{n!}{n^n}[z^n](B^m(z)-1)$. We found:

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

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.

General Recurrence

Our method allows to solve a general recurrence of the following form:

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

where a_n is a given sequence (the so called $additive$ term), and m is an additional parameter.

Indeed, the above leads to

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

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.

ANALYTIC METHODS: Mellin Transform and Saddle Point Approach

Csiszár and Shields studied the minimax redundancy of the **renewal process** defined as:
Let $T_1, T_2 \ldots$ be a sequence of i.i.d. positive-valued random variables with distribution $Q(j) = \Pr\{T_i = j\}$ over nonnegative integers $j \geq 0$. The process $T_0, T_0 + T_1, T_0 + T_1 + T_2, \ldots$ is called the renewal process which is stationary if T_0 is chosen properly. With such a renewal process we associate a **binary renewal sequence** in which the positions of the 1's are at the renewal epoches $T_0, T_0 + T_1, T_0 + T_1 + T_2, \ldots$.

By Shtarkov's method, to study the redundancy R_n^* we should evaluate

$$R_n = \lg \left(\sum_{x_1^n} \sup_Q P(x_1^n) \right)$$

A Simple Lemma

Lemma 2. *Define* $r_n = 2^{R_n}$. *Then*

$$
\begin{aligned}
r_n &= \sum_{k=0}^{n} r_{n,k} \\
r_{n,k} &= \sum_{\mathcal{P}(n,k)} \binom{k}{k_0 \cdots k_{n-1}} \left(\frac{k_0}{k}\right)^{k_0} \left(\frac{k_1}{k}\right)^{k_1} \cdots \left(\frac{k_{n-1}}{k}\right)^{k_{n-1}}
\end{aligned}
$$

where $\mathcal{P}(n,k)$ *denotes the partition of* n *into* k *terms, that is*

$$
\begin{aligned}
n &= k_0 + 2k_1 + \cdots + nk_{n-1}, \\
k &= k_0 + \cdots + k_{n-1}.
\end{aligned}
$$

Proof. Observe that the renewal sequence x_1^n can be represented as

$$x_1^n = 0^{\alpha_1} 1 0^{\alpha_2} 1 \cdots 1 0^{\alpha_n} 1 \underbrace{0 \cdots 0}_{k^*}$$

where $0 \le \alpha_i \le n$ for $i = 1, \ldots, n$. Let k_m be the number of i such that $\alpha_i = m$, where $m = 0, 1, \ldots, n-1$.

Then

$$P(x_1^n) = Q^{k_0}(0)Q^{k_1}(1)\cdots Q^{k_{n-1}}(n-1)Q^*(k^*)$$

subject to $Q(0) + Q(1) + \cdots + Q(n-1) \leq 1$, where

$$k_0 + 2k_1 + \cdots + nk_{n-1} + k^* = n,$$

and $Q^*(k^*) = \Pr\{T_1 \geq k^*\}$. This is a simple optimization problem with constraints that can be easily solved leading to

$$\sup_Q P(x_1^n) = \left(\frac{k_0}{k_0 + \cdots + k_{n-1}}\right)^{k_0} \cdots \left(\frac{k_{n-1}}{k_0 + \cdots + k_{n-1}}\right)^{k_{n-1}}$$

which proves the lemma.

Re-Formulation of the Problem

A difficulty of finding asymptotics of r_n stems from the factor $k!/k^k$ present in the definition of $r_{n,k}$. We circumvent this problem by analyzing a related pair of sequences, namely s_n and $s_{n,k}$ that are defined as

$$\begin{cases} s_n & = & \sum_{k=0}^{n} s_{n,k} \\ s_{n,k} & = & e^{-k} \sum_{\mathcal{P}(n,k)} \frac{k^{k_0}}{k_0!} \cdots \frac{k^{k_{n-1}}}{k_{n-1}!}. \end{cases}$$

The translation from s_n to r_n is most conveniently expressed in probabilistic terms. Introduce the random variable K_n whose probability distribution is $s_{n,k}/s_n$, that is,

$$\varpi_n : \qquad \Pr\{K_n = k\} = \frac{s_{n,k}}{s_n},$$

where ϖ_n denotes the distribution. Then Stirling's formula yields

$$\begin{aligned} \frac{r_n}{s_n} &= \sum_{k=0}^{n} \frac{r_{n,k}}{s_{n,k}} \frac{s_{n,k}}{s_n} = \mathbf{E}[(K_n)! K_n^{-K_n} e^{-K_n}] \\ &= \mathbf{E}[\sqrt{2\pi K_n}] + O(\mathbf{E}[K_n^{-\frac{1}{2}}]). \end{aligned}$$

Thus, the problem of finding r_n reduces to asymptotic evaluations of s_n, $\mathbf{E}[\sqrt{K_n}]$ and $\mathbf{E}[K_n^{-\frac{1}{2}}]$.

Fundamental Lemmas

The heart of the matter is the following lemma which provides the necessary estimates.

Lemma 3. *Let* $\mu_n = \mathbf{E}[K_n]$ *and* $\sigma_n^2 = \mathbf{Var}(K_n)$, *where* K_n *has the distribution* ϖ_n *defined above. The following holds*

$$
\begin{aligned}
s_n &\sim \exp\left(2\sqrt{cn} - \frac{7}{8}\log n + d + o(1)\right) \\
\mu_n &= \frac{1}{4}\sqrt{\frac{n}{c}}\log\frac{n}{c} + o(\sqrt{n}) \\
\sigma_n^2 &= O(n\log n) = o(\mu_n^2),
\end{aligned}
$$

where $c = \pi^2/6 - 1$, $d = -\log 2 - \frac{3}{8}\log c - \frac{3}{4}\log \pi$.

By Chebyshev's we also have:

Lemma 4. *For large* n

$$
\begin{aligned}
\mathbf{E}[\sqrt{K_n}] &= \mu_n^{1/2}(1 + o(1)) \\
\mathbf{E}[K_n^{-\frac{1}{2}}] &= o(1).
\end{aligned}
$$

where $\mu_n = \mathbf{E}[K_n]$.

Main Result

In summary, r_n and s_n are related by

$$\begin{aligned} r_n &= s_n \mathbf{E}[\sqrt{2\pi K_n}](1 + o(1)) \\ &= s_n \sqrt{2\pi \mu_n}(1 + o(1)). \end{aligned}$$

This leads to

Theorem 5. [Flajolet and Szpankowski 1998] *Consider the class of renewal processes as defined above. The minimax redundancy ρ_n attains the following asymptotics*

$$R_n^*(\mathcal{R}_0) = \frac{2}{\log 2}\sqrt{cn} - \frac{5}{8}\lg n + \frac{1}{2}\lg\log n + O(1)$$

where $c = \frac{\pi^2}{6} - 1 \approx 0.645$

Proof of the Fundamental Lemma

1. We start by introducing the well-known "tree function" $T(z)$ defined as the solution of

$$T(z) = ze^{T(z)}$$

that is analytic at 0. The function $T(z)$ satisfies, by the Lagrange inversion theorem,

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

2. Next define the function $\beta(z)$ as

$$\beta(z) = \sum_{k=0}^{\infty} \frac{k^k}{k!} e^{-k} z^k.$$

One has (e.g., by Lagrange inversion again or otherwise)

$$\beta(z) = \frac{1}{1 - T(ze^{-1})}.$$

3. The quantities s_n and $s_{n,k}$ have generating functions,

$$S_n(u) = \sum_{k=0}^{\infty} s_{n,k} u^k, \qquad S(z,u) = \sum_{n=0}^{\infty} S_n(u) z^n.$$

Then, since $s_{n,k}$ involves convolutions of sequences of the form $k^k/k!$, we have

$$\begin{aligned} S(z,u) &= \sum_{\mathcal{P}_{n,k}} z^{1k_0+2k_1+\cdots} \left(\frac{u}{e}\right)^{k_0+\cdots+k_{n-1}} \frac{k^{k_0}}{k_0!} \cdots \frac{k^{k_{n-1}}}{k_{n-1}!} \\ &= \prod_{i=1}^{\infty} \beta(z^i u). \end{aligned}$$

4. To compute the moments μ_n and $\mathbf{E}[K_n(K_n-1)]$ we use the following formulas

$$\begin{aligned} s_n &= [z^n]S(z,1), \\ \mu_n &= \frac{[z^n]S'_u(z,1)}{[z^n]S(z,1)}, \\ \mathbf{E}[K_n(K_n-1)] &= \frac{[z^n]S''_{uu}(z,1)}{[z^n]S(z,1)} \end{aligned}$$

where $[z^n]F(z)$ denotes the coefficient at z^n of $F(z)$, $S'_u(z,1)$ and $S''_{uu}(z,1)$ represent the first and the second derivative of $S(z,u)$ at $u=1$.

Mellin Asymptotics

5. We deal here with

$$S(z,1) = \prod_{i=1}^{\infty} \beta(z^i).$$

The behaviour of the generating function $S(z,1)$ as $z \to 1$ is an essential ingredient of the analysis.

5a. The singularity of the tree function $T(z)$ at $z = e^{-1}$ is of the square-root type, that is, near $z = 1$, $\beta(z)$ admits the singular expansion :

$$\beta(z) = \frac{1}{\sqrt{2(1-z)}} + \frac{1}{3} - \frac{\sqrt{2}}{24}\sqrt{(1-z)} + O(1-z).$$

5b. We now turn to the infinite product asymptotics as $z \to 1^-$, with z real. Let $L(z) = \log S(z,1)$ and $z = e^{-t}$, so that

$$L(e^{-t}) = \sum_{k=1}^{\infty} \log \beta(e^{-kt}).$$

Mellin transform techniques provide an expansion of $L(e^{-t})$ around $t = 0$ (or equivalently $z = 1$) since the sum falls under the *harmonic sum* paradigm.

Mellin Properties

(M1) Direct and Inverse Mellin Transforms. Let c belong to the *fundamental strip* defined below.

$$f^*(s) := \mathcal{M}(f(x); s) = \int_0^\infty f(x) x^{s-1} dx$$

then

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s} ds.$$

(M2) Fundamental Strip. The Mellin transform of $f(x)$ exists in the *fundamental strip* $\Re(s) \in (-\alpha, -\beta)$, where

$$f(x) = O(x^\alpha) \qquad (x \to 0), \qquad\qquad f(x) = O(x^\beta) \qquad (x \to \infty).$$

(M3) Harmonic Sum Property. By linearity and the scale rule $\mathcal{M}(f(ax); s) = a^{-s}\mathcal{M}(f(x); s)$,

$$f(x) = \sum_{k \geq 0} \lambda_k g(\mu_k x)$$

then

$$f^*(s) = g^*(s) \sum_{k \geq 0} \lambda_k \mu_k^{-s}.$$

(M4) MAPPING PROPERTIES (Asymptotic expansion of $f(x)$ and singularities of $f^*(s)$).

$$f(x) = \sum_{(\xi,k)\in A} c_{\xi,k} x^{\xi} (\log x)^k + O(x^M)$$

then

$$f^*(s) \asymp \sum_{(\xi,k)\in A} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}}.$$

(i) *Direct Mapping.* Assume that $f(x)$ admits as $x \to 0^+$ the asymptotic expansion of the above for some $-M < -\alpha$ and $k > 0$. Then for $\Re(s) \in (-M, -\beta)$, the transform $f^*(s)$ satisfies the singular expansion of above.

(ii) *Converse Mapping.* Assume that $f^*(s) = O(|s|^{-r})$ with $r > 1$, as $|s| \to \infty$ and that $f^*(s)$ admits the singular expansion above for $\Re(s) \in (-M, -\alpha)$. Then $f(x)$ satisfies the asymptotic expansion of above at $x = 0^+$.

Continuation of the Proof

6. The Mellin transform $L^*(s) = \mathcal{M}(L(e^{-t}); s)$ of $L(e^{-t})$ is computed by the harmonic sum property (M3). For $\Re(s) \in (1, \infty)$, the transform evaluates to

$$L^*(s) = \zeta(s)\Lambda(s)$$

where $\zeta(s) = \sum_{n \geq 1} n^{-s}$ is the Riemann zeta function, and

$$\Lambda(s) = \int_0^\infty \log \beta(e^{-t}) t^{s-1} dt.$$

7. By the direct mapping property (M4), the expansion of $\beta(z)$ at $z = 1$ implies

$$\log \beta(e^{-t}) = -\frac{1}{2} \log t - \frac{1}{2} \log 2 + O(\sqrt{t}),$$

so that, collecting local expansions,

$$\Lambda(s) \asymp \big(\Lambda(1)\big)_{s=1} + \left(\frac{1}{2}\frac{1}{s^2} - \frac{1}{2}\frac{\log 2}{s}\right)_{s=0}.$$

Then

$$L^*(s) \asymp \left(\frac{\Lambda(1)}{s-1}\right)_{s=1} + \left(-\frac{1}{4s^2} - \frac{\log \pi}{4s}\right)_{s=0}.$$

8. An application of the converse mapping property (M4) allows us to come back to the original function,

$$L(e^{-t}) = \frac{\Lambda(1)}{t} + \frac{1}{4}\log t - \frac{1}{4}\log\pi + O(\sqrt{t}),$$

which translates in

$$L(z) = \frac{\Lambda(1)}{1-z} + \frac{1}{4}\log(1-z) - \frac{1}{4}\log\pi - \frac{1}{2}\Lambda(1) + O(\sqrt{1-z}).$$

where

$$\begin{aligned} c = \Lambda(1) &= -\int_0^1 \log(1 - T(x/e))\frac{dx}{x} \\ &= \frac{\pi^2}{6} - 1. \end{aligned}$$

9. In summary, we just proved that, as $z \to 1^-$,

$$S(z,1) = e^{L(z)} = a(1-z)^{\frac{1}{4}} \exp\left(\frac{c}{1-z}\right)(1 + o(1)),$$

where $a = \exp(-\frac{1}{4}\log\pi - \frac{1}{2}c)$.

10. It remains to collect the information gathered on $S(z,1)$ and recover $s_n = [z^n]S(z,1)$ asymptotically. The inversion is provided by the Cauchy coefficient formula, that is,

$$s_n = \frac{1}{2\pi i} \oint \frac{S(z,1)}{z^{n+1}} dz$$

where the integration path is any simple loop around 0 inside the unit disk.

11. To estimate s_n we use the following lemma that is based on an application of the saddle point method summarized on the next few slides.

Lemma 5. *For positive $A > 0$, and reals B and C, define $f(z) = f_{A,B,C}(z)$ as*

$$f(z) = \exp\left(\frac{A}{1-z} + B\log\frac{1}{1-z} + C\log\left(\frac{1}{z}\log\frac{1}{1-z}\right)\right).$$

Then, the nth Taylor coefficient of $f_{A,B,C}(z)$ satisfies asymptotically, for large n,

$$\begin{aligned}[z^n]f_{A,B,C}(z) &= 2\sqrt{An} + \frac{1}{2}\left(B - \frac{3}{2}\right)\log n \\ &\quad + C\log\log\sqrt{\frac{n}{A}} \\ &\quad - \frac{1}{2}\log\left(4\pi e^{-A}/\sqrt{A}\right) + o(1).\end{aligned}$$

Saddle Point Method

Input: A function $g(z)$ analytic in $|z| < R$ $(0 < R < +\infty)$ with nonnegative Taylor coefficients and "fast growth" as $z \to R^-$. Let $h(z) := \log g(z) - (n+1)\log z$.

Output: The asymptotic formula for $g_n := [z^n]g(z)$ derived from the Cauchy coefficient integral

$$g_n = \frac{1}{2i\pi}\int_\gamma g(z)\,\frac{dz}{z^{n+1}} = \frac{1}{2i\pi}\int_\gamma e^{h(z)}\,dz$$

where γ is a loop around $z = 0$.

(S1). SADDLE POINT CONTOUR. *Require that* $g'(z)/g(z) \to +\infty$ *as* $z \to R^-$. Let $r = r(n)$ be the unique positive root of the saddle point equation

$$h'(r) = 0 \qquad \text{or} \qquad \frac{rg'(r)}{g(r)} = n+1,$$

so that $r \to R$ as $n \to \infty$. The integral above is evaluated on $\gamma = \{z \mid |z| = r\}$.

(S2). BASIC SPLIT. *Require that* $h'''(r)^{1/3}h''(r)^{-1/2} \to 0$. Define $\varphi = \varphi(n)$ called the "range" of the saddle point by

$$\varphi = \left| h'''(r)^{-1/6} h''(r)^{-1/4} \right|,$$

so that $\varphi \to 0$, $h''(r)\varphi^2 \to \infty$, and $h'''(r)\varphi^3 \to 0$. Split $\gamma = \gamma_0 \cup \gamma_1$, where $\gamma_0 = \{z \in \gamma \mid |\arg(z)| \leq \varphi\}$, $\gamma_1 = \{z \in \gamma \mid |\arg(z)| \geq \varphi\}$.

(S3) ELIMINATION OF TAILS. *Require that* $|g(re^{i\theta})| \leq |g(re^{i\varphi})|$ *on* γ_1. Then, the tail integral satisfies the trivial bound,

$$\left| \int_{\gamma_1} e^{h(z)}\, dz \right| = O\left(|e^{-h(re^{i\varphi})}| \right).$$

(S4) LOCAL APPROXIMATION. *Require that* $h(re^{i\theta}) - h(r) - \frac{1}{2}r^2\theta^2h''(r) = O(|h'''(r)\varphi^3|)$ *on* γ_0. Then, the central integral is asymptotic to a complete Gaussian integral, and

$$\frac{1}{2i\pi}\int_{\gamma_0} e^{h(z)}\,dz = \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}\left(1 + O(|h'''(r)\varphi^3|)\right).$$

(S5) COLLECTION. Requirements $(S1)$, $(S2)$, $(S3)$, $(S4)$, imply the estimate:

$$[z^n]g(z) = \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}\left(1 + O(|h'''(r)\varphi^3|)\right) \sim \frac{g(r)r^{-n}}{\sqrt{2\pi h''(r)}}.$$