SOME USES OF THE MELLIN INTEGRAL TRANSFORM IN THE ANALYSIS OF ALGORITHMS

Philippe Flajolet and *Mireille Regnier*
I.N.R.I.A
Rocquencourt
78150-Le Chesnay (France)

Robert Sedgewick
Department of Computer Science
Brown University
Providence R.I. 02912 (U.S.A.)

ABSTRACT

We informally survey some uses of the Mellin integral transform in the context of the asymptotic evaluation of combinatorial sums arising in the analysis of algorithms.

1. Introduction

The *Mellin transform* is an integral transform that is part of the working kit of analytic number theory where its use is to be traced to Riemann's famous memoir on the distribution of primes. Its usefulness in asymptotic analysis comes from the fact that it relates *asymptotic properties* of a function around 0 and ∞ to the *singularities* of the transformed function.

We propose here to informally explore the uses of the Mellin transform in the context of the average case analysis of algorithms.

Average case analysis of algorithms starts with a counting of certain combinatorial configurations like permutations, words, trees, finite functions, distributions *etc*.... One then computes certain weighted averages of the counting results to determine the expected cost of algorithms. Depending on the particular problem under consideration, the counting results are either obtained explicitly or indirectly accessible through some generating function. Mellin transform techniques are especially valuable when number-theoretic functions (divisor function, functions related to binary representation of integers) and/or periodicities appear.

The Mellin transform of a real valued function $F(x)$ defined over $[0;+\infty[$ is the complex function $F^*(s)$ of the complex variable s given by:

$$F^*(s) = \int_0^\infty F(x)\, x^{s-1}\, dx \qquad (1)$$

also written sometimes:

NATO ASI Series, Vol. F12
Combinatorial Algorithms on Words
Edited by A. Apostolico and Z. Galil

	$F(x)$	$F^*(x)$	
T1.	e^{-x}	$\Gamma(s)$	$0<\mathrm{Re}(s)$
	$e^{-x}-1$	$\Gamma(s)$	$-1<\mathrm{Re}(s)<0$
	$e^{-x}-1+x$	$\Gamma(s)$	$-2<\mathrm{Re}(s)<-1$
	$e^{+x}-1-x-\frac{x^2}{2}$	$\Gamma(s)$	$-3<\mathrm{Re}(s)<-2$
T2.	$\frac{1}{1+x}$	$\pi\operatorname{cosec}\pi s$	$0<\mathrm{Re}(s)<1$
T3.	$\log(1+x)$	$\frac{\pi}{s\sin\pi s}$	$-1<\mathrm{Re}(s)<0$
T4.	$\delta(x)$	$\frac{1}{s}$	$0<\mathrm{Re}(s)$
T5.	$\log x\,\delta(x)$	$\frac{1}{s^2}$	$0<\mathrm{Re}(s)$
T6.	$(1-x)^{m-1}\delta(x)$	$\frac{\Gamma(s)\Gamma(m)}{\Gamma(s+m)}$	$0<\mathrm{Re}(s);0<m$

Table 1: *Some basic transforms. ($\delta(x)$ is defined by $\delta(x) = 1$ if $x\leq 1$ and* 0 *otherwise.)*

$$\mathbf{M}[F(x);s] \quad \text{or} \quad \mathbf{M}[F].$$

The interest of the Mellin transform comes from the combination of two types of properties:

P1 *Asymptotic properties:* under fairly general conditions, one obtains an asymptotic expansion of $F(x)$ from the singularities of its transform:

$$F(x) \sim \pm \sum_{\alpha\in H} Res\,(F^*(s)\,x^{-s}\ ;\ s=\alpha) \tag{2}$$

where H is either a left ($x\to 0$) or right ($x\to\infty$) half-plane. Expansion (2) *is* an asymptotic expansion.

P2 *Functional properties:* somewhat intricate sums often have simple transforms. An important paradigm is that of ***harmonic sums***. A harmonic sum is of the form:

$$F(x) = \sum_k \lambda_k f\,(\mu_k x) \tag{3}$$

and its transform has, at least formally, the factored form

$$F^*(s) = (\sum_k \lambda_k \mu_k^{-s})\, f^*(s) \tag{4}$$

which is simply the product of a (generalized) *Dirichlet series* and the transform of the base function f.

The combination of (2) and (3) allows for derivation of very many asymptotic expansions. The generality of expansions (2) permits in particular to attack expressions whose expansions involve non trivial *periodicity phenomena* which correspond to poles of the transform function in expansion (2) that have a non-zero imaginary part.

2. Basic properties of the Mellin transform

Let $F(x)$ be piecewise continuous on $[0;\infty]$ and assume F satisfies:

$$F(x) = O(x^{\alpha}) \quad (x \to 0) \quad ; \quad F(x) = O(x^{\beta}) \quad (x \to \infty) \quad .$$

Then the Mellin transform of $F(x)$ is defined in the strip:

$$-\alpha < \mathrm{Re}(s) < -\beta \, .$$

This strip is called the ***fundamental strip*** of F^* and is sometimes denoted by $<\alpha;\beta>$. It is empty unless $\beta<\alpha$. Thus the Mellin transform of a function is only defined if its ***order at infinity*** is smaller than its ***order at zero***.

Table 1 summarizes some basic transforms while Table 2 describes the main functional properties of the Mellin transform

	$F(x)$	$F^*(s)$	
P1	$F(x)$	$\int_0^{\infty} F(x)x^{s-1}\,dx$	definition
P2	$\frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty} F^*(s)s\,ds$	$F^*(s)$	inversion th.
P3	$F(ax)$	$a^{-s}F^*(s)$	$a>0$
P4	$x^{\nu}F(x^{\mu})$	$\frac{1}{\beta}F^*(\frac{s+\nu}{\mu})$	
P5	$\frac{d}{dx}F(x)$	$-(s-1)F^*(s-1)$	
P6	$F(x)\log x$	$\frac{d}{ds}F^*(s)$	
P7	$\sum\lambda_k f(\mu_k x)$	$(\sum\lambda_k\mu_k^{-s})f^*(s)$	harmonic sum

Table 2: *Basic functional properties of the Mellin transform.*

2.1. Direct properties

The characteristic property of Mellin transforms is: *local properties* at 0 and ∞ are reflected by the *singularities* of the transform function.

Theorem 1: *Assume $F(x)$ has the following expansions:*

$$F(x) \sim \sum_{k>0} c_k x^{\alpha_k} \qquad (x \to 0)$$

$$F(x) \sim \sum_{k>0} d_k x^{\beta_k} \qquad (x \to \infty)$$

where the α_k form an increasing sequence that tends to $+\infty$ and the β_k form a decreasing sequence that tends to $-\infty$.

Then the transform $F^*(s)$ of $F(x)$ is meromorphic in the whole complex plane with simple poles at points $-\alpha_k$ and $-\beta_k$ with:

$$F^*(s) \sim \frac{c_k}{s+\alpha_k} \quad (s \to -\alpha_k) \ ,$$

$$F^*(s) \sim -\frac{d_k}{s+\beta_k} \quad (s \to -\beta_k) .$$

Proof: (Sketch) To translate the expansion at 0, consider the transform:

$$F_o^*(s) = \int_0^1 F(x)\, dx .$$

Formally applying this partial transform to the asymptotic expansion leads to (see the transform of the δ function in Table 1)

$$F_o^*(s) \sim \sum_k \frac{c_k}{s+\alpha_k} .$$

Theorem 1 admits a number of extensions: the α_k only need to be complex numbers whose sequence of real parts tends to $\pm\infty$. Also partial asymptotic expansions may be similarly translated: then the transform function is meromorphic in an extended strip (not necessarily the whole of the complex plane). Finally generalized expansions of the form $c_k(\log x)$ for a sequence of polynomials $c_k(t)$ may be allowed. The meromorphy result still holds but then $F^*(s)$ will have a pole of order $deg(c_k)-1$ at $-\alpha_k$ (obvious analogues are true for expansions at ∞).

In other words the asymptotic behaviour of F at 0 (resp. ∞) is reflected by the singularities of F^* in a ***left half-plane*** (resp. ***right half-plane***) w.r.t. the fundamental strip.

Another important property of Mellin transforms is their smallness towards $i\infty$. From the Riemann-Lebesgue lemma, one has:

Theorem 2: *If $F(x)$ is infinitely differentiable over $[0;+\infty[$, then for s in the fundamental strip, one has for all $m>0$:*

$$\lim_{|s|\to\infty} s^m F^*(s) = 0 .$$

2.2. Inverse properties

One can come back from F^* to F by means of the following important inversion theorem:

Theorem 3: *For any real c inside the fundamental strip, one has:*

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

This theorem is crucial for proving a converse of Theorem 1, namely:

Theorem 4: *Let $<-\alpha;-\beta>$ be the fundamental strip of F^*, and assume that F^* is small towards $i\infty$. If $F^*(s)$ is meromorphic for $-L \le \mathrm{Re}(s) \le -\alpha$ with finitely many simple poles at $-\alpha_1 > -\alpha_2 > \cdots > -L$, with residue c_k at $-\alpha_k$, then $F(x)$ admits the asymptotic expansion:*

$$F(x) \sim \sum_{k\ge 1} c_k x^{\alpha_k} + O(x^L) .$$

If $F^*(s)$ is meromorphic for $-\beta \le \mathrm{Re}(s) \le -M$ with finitely many simple poles at $-\beta_1 < -\beta_2 < \cdots < -M$, with residue d_k at $-\beta_k$, then $F(x)$ admits the asymptotic expansion:

$$F(x) \sim -\sum_{k \ge 1} c_k x^{\beta_k} + O(x^M) .$$

Proof: Start with the inversion theorem to express $F(x)$ and evaluate the integral by shifting the line of integration to the left until the line $\mathrm{Re}(s) = -L$ only taking residues into account.

In this way, one obtains a form:

$$F(x) = \sum_k Res\,(F^*(s)x^{-s}; s = -\alpha_k) + \frac{1}{2i\pi} \int_{-L-i\infty}^{-L+i\infty} F^*(s) x^{-s}\, ds .$$

The other result about the expansion of F at $+\infty$ is proved similarly by moving the line of integration to the right.

Observation: The same process can be applied to multiple poles and be extended to cases where F^* has infinitely many poles in a finite width strip, provided F^* remains small along some horizontal lines towards $i\infty$. One observes that:

1. A simple pole of the form $s = \sigma + i\tau$ will give a term in the asymptotic expansion of F of order

$$x^{-s} = x^{-\sigma} x^{-i\tau} = x^{-\sigma} e^{-i\tau \log x}$$

Poles of F^* farther west contribute smaller terms to the expansion of F at 0. Poles of F^* farther east contribute smaller terms to the expansion of F at ∞.

Poles with non-zero imaginary parts correspond to asymptotic fluctuations.

2. A pole of F^* of order k at s contributes a term of the form (P_{k-1} being a polynomial of degree $k-1$):

$$P_{k-1}(\log x) x^{-s} .$$

Or:

Multiple poles introduce factors that are powers of $\log x$ *in asymptotic expansions.*

3. Full asymptotic expansions may or may not be convergent (for sums of residues in either a left or a right half plane). In most of the cases dealt with here, they are *not convergent*, thus only asymptotic.

3. Harmonic sums

Let f be a smooth enough function. Then a sum of the form:

$$F(x) = \sum_k \lambda_k f(\mu_k x) \tag{7}$$

is called a *harmonic sum*. Function f is called the *base function*; the coefficients λ_k are the *amplitudes* and the μ_k are the *frequencies*. Applying the basic functional property of the Mellin transform to (7), we find formally:

$$F^*(s) = \omega(s) f^*(s) \tag{8}$$

where ω is the (generalized) Dirichlet series:

$$\omega(s) = \sum_k \lambda_k \mu_k^{-s} \,. \tag{9}$$

The transformed equation (8) is valid inside the intersection of the fundamental domain of f^* and of the domain of absolute convergence of ω.
A pair of sequences $\{\lambda_k\}$, $\{\mu_k\}$ is said to be an ***arithmetic pair*** iff the series (8) is meromorphic in the whole complex plane.

Mellin transform techniques are well suited to the treatment of harmonic sums associated to arithmetic pairs amplitudes-frequencies when the base function is smooth (the exponential function in many applications).

From the preceding sections, we see that provided $\omega(s)$ is not too ill behaved at $i\infty$, the asymptotics of harmonic sums can be easily determined.

Example 1: *Harmonic numbers.*
Harmonic numbers are particular harmonic sums. The n-th harmonic number is classically defined by

$$H_n = \sum_{k=1}^{n} \frac{1}{k} \,.$$

It can be extended to a smooth real function by introducing:

$$h(x) = \sum_{k\geq 0} \left[\frac{1}{k} - \frac{1}{x+k} \right]$$

so that $h(n) = H_n$. The fundamental strip of h^* is $<-1;0>$ and rewriting $h(x)$ under the form:

$$h(x) = \sum_n \frac{1}{n} \left(\frac{x}{n}\right)\left(1+\frac{x}{n}\right)^{-1}$$

we see that it is a harmonic sum. The Dirichlet series associated to amplitudes and frequencies is:

$$\omega(s) = \sum_{n\geq 1} \frac{1}{n} \left(\frac{1}{n}\right)^{-s}$$

that is $\omega(s) = \zeta(1-s)$ where ζ is the classical Riemann zeta function. The transform of the base function is obtained from Table 1, so that:

$$h^*(s) = \frac{\pi}{\sin \pi s} \zeta(1-s) \tag{10}$$

for $-1 < Re(s) < 0$. Thus we can choose $c = 1/2$ in the inversion theorem and get:

$$h(x) = \frac{1}{2i\pi} \int_{-1/2-i\infty}^{-1/2+i\infty} \frac{\pi}{\sin \pi s} \zeta(1-s)\, ds$$

The asymptotic behaviour of h towards $+\infty$ is obtained by moving the line of integration to the right. One first encounters a double pole at $s=0$ corresponding to a logarithmic dominant term and then poles at some of the integers; whence the classical expansion (the B_k are the Bernoulli numbers):

$$h(x) \sim \log x + \gamma + \frac{1}{2n} + \sum_{k\geq 2} \frac{(-1)^k B_k}{k\, n^k} \tag{11}$$

Example 2: *Special Euler-Maclaurin summations.*
Assume $f(x)$ is exponentially small towards ∞, and admits around 0 an expansion of the form:

$$f(x) \sim \sum_{k\geq 0} c_k x^{\alpha_k} \tag{12}$$

where $\alpha_0=0<\alpha_1< \cdots$. Consider the sum:

$$F(x) = \sum_{n\geq 1} f(nx) . \tag{13}$$

for which an expansion as $x \to 0$ is sought. The transform of F is:

$$F^*(s) = f^*(s)\zeta(s)$$

for $\mathrm{Re}(s)>1$. From Section 2.1, the poles of f^* are at the $-\alpha_k$; $\zeta(s)$ has a unique pole at $s=1$ with residue 1 so that:

$$F^*(s) \sim \frac{1}{s-1} f^*(1) \quad (s \to 1) \tag{14}$$

thus putting everything together, we find an Euler-Maclaurin-Barnes expansion:

$$F(x) \sim \frac{1}{x}\int_0^\infty f(t)\,dt + \frac{1}{2}f(0) + \sum_{k\geq 1} c_k \zeta(-\alpha_k) x^{\alpha_k} . \tag{15}$$

Cases where the expansion of f also involves logarithmic terms can be dealt with similarly (multiple poles appear that introduce the derivatives of the zeta function at $-\alpha_k$), yielding formulae of Gonnet.

Example 3: *Powers-of-two sums and periodicities.*
Assume here for f the same conditions as in the previous example and consider the sum:

$$F(x) = \sum_{k\geq 0} f(2^k x) \tag{16}$$

for which an asymptotic expansion is sought as $x \to 0$. The transform of F is for $\mathrm{Re}(s)>0$:

$$F^*(s) = \frac{f^*(s)}{1-2^{-s}} .$$

It has a double pole at $s=0$ if $f(0)\neq 0$ so that we need two terms in the expansion of f^* at 0. Writing:

$$f^*(s) = \frac{f(0)}{s} + \int_0^1 (f(x)-f(0))\,x^{s-1}\,dx + \int_1^\infty f(x) x^{s-1}\,dx$$

we see that around 0:

$$f^*(s) = \frac{f(0)}{s} + \gamma_f + O(s) ,$$

where:

$$\gamma_f = \int_0^1 (f(x)-f(0))\,\frac{dx}{x} + \int_1^\infty f(x)\,\frac{dx}{x}. \tag{17}$$

Function F^* also has simple poles at $s=\chi_k$, $k \in \mathbb{Z}/\{0\}$ with $\chi_k = \dfrac{2ik\pi}{\log 2}$, as well as at the points $-\alpha_k$, whence the general expansion as $x \to 0$:

$$F(x) \sim -f(0)\log_2 x + \frac{1}{2}f(0) + \frac{\gamma_f}{\log 2} + P(\log_2 x) + \sum_{k\geq 1} \frac{c_k}{1-2^{\alpha_k}} x^{\alpha_k},$$

with:

$$P(u) = \frac{1}{\log 2} \sum_{k\in \mathbb{Z}/\{0\}} f^*\left(\frac{2ik\pi}{\log 2}\right) e^{-2ik\pi u/\log 2}. \tag{18}$$

Example 4: *Perron's and other number-theoretic formulae*

The framework of harmonic sum is also a natural setting for formulae relating ordinary generating functions to Dirichlet series as well as for Perron's formula expressing partial sums of coefficients of Dirichlet series. This last formula follows from the application of our methods to sums of the form:

$$F(x) = \sum_{k\geq 1} \lambda_k \delta(kx)$$

where function delta is defined in Table 1.

4. Harmonic power sums

We only give brief indications here on the idea underlying the asymptotic treatment of *harmonic power sums* referring the reader to M. Regnier's thesis for a precise statement of theorems, and validity conditions.

A harmonic power sum is a sum of the form:

$$F(x) = \sum_{k\geq 1} \lambda_k f(\mu_k x)^{\gamma_k}. \tag{19}$$

We assume here that f and g are smooth enough functions. From (19), taking transforms, we see that one has at least formally:

$$F^*(s) = \sum_{k\geq 1} \lambda_k \mu_k^{-s} f^*_{\gamma_k}(s) \tag{20}$$

where f^*_γ is the transform:

$$f^*_\gamma(s) = \int_0^\infty f^\gamma(x)\, x^{s-1}\, dx. \tag{21}$$

Our treatment concerns cases where the $\gamma_k \to \infty$. In that case, it is natural to expect from (21) that the behaviour of the sum F is determined by the behaviour of (21) as γ gets large.

Rewritting (21) in an exponential form, and assuming without loss of generality that $f(0)=1$, we should expect

$$f^*_\gamma(s) = \int_0^\infty e^{\gamma \log f(x)} x^{s-1}\, dx,$$

to be approximable for large γ by the *Laplace method for integrals*. If $\log f(x)$ admits around 0 an asymptotic expansion of the form $-c\, x^\alpha + O(x^\beta)$, then one can prove under certain conditions that F^* is "approximated" by a function

$$\omega_1(s) = \frac{1}{\alpha}\Gamma\left(\frac{s}{\alpha}\right) c^{-s/\alpha} \sum_{k\geq 1} \lambda_k\, \mu_k^{-s} \gamma_k^{-s/\alpha}$$

in the sense that $F^*(s)-\omega_1(s)$ is analytic in a larger strip than $F^*(s)$. Whence, for F an asymptotic expansion of the form:

$$F(x) \sim \sum_{s_0 \in \mathbf{S}} Res(\omega_1(s)x^{-s} \; ; \; s=s_0) \tag{22}$$

where the sum is extended to poles s_0 in some strip $\mathbf{S}$ of the complex s-plane.

With these techniques one can estimate sums appearing in the analysis of the *Extendible Hashing* algorithm used for storing large files on disk while maintaining direct access in a dynamic context. For instance, under a Poisson model, the expected size of an Extendible Hashing directory formed with n records is $D(n)$ where:

$$D(x) = \sum_{k \geq 0} (1 - f_b(x 2^{-k})^{2^k}) .$$

and $f_b(x)$ is $e^{-x} \sum_{k=0}^{b} \frac{x^k}{k!}$. Function $D(x)$ is reducible to a harmonic power sum.

5. Other methods

We only mention briefly here two other uses of Mellin techniques:

A. Complex Mellin inversion and singularity analysis of generating functions.

B. The so-called "Rice" formula, actually a classical formula from the calculus of finite differences.

A. It is known in many cases that the behaviour of a function around its (dominant) *singularities* determines the asymptotic behaviour of its coefficients. However in most cases (except for some scarce cases of application of Tauberian theorems) an asymptotic expansion of the function around its singularity *in the complex plane* is required. For instance it is known that the generating function of the quantity E_n representing the total number of registers to evaluate optimally all binary trees of size n is:

$$E(z) = \frac{1-u^2}{u} \sum_{p \geq 1} \frac{u^{2^p}}{1-u^{2^p}} \tag{23}$$

where:

$$u = u(z) = \frac{1-\sqrt{1-4z}}{1+\sqrt{1-4z}} .$$

The singularity of $E(z)$ is at $z=1/4$, where $u \to 1$ as $z \to 1/4$. The asymptotic behaviour of (23) when z is in a neighbourhood of $1/4$ in the complex plane (u in a neighbourhood of 1) determines the asymptotics of the E_n. Setting $u=e^{-x}$, the sum in (23) becomes:

$$F(x) = \sum_{p \geq 1} \frac{e^{-x 2^p}}{1-e^{-x 2^p}} , \tag{24}$$

typically a harmonic sum. whose transform is :

$$\frac{\Gamma(s)\zeta(s)}{2^s-1} .$$

The asymptotics of F for x in a complex neighbourhood of 0 can still be determined by appealing to ***complex Mellin inversion*** formulae, from which the asymptotics of E_n is found. (See a forthcoming paper of Flajolet and Prodinger for details).

The method also applies to the analysis of the expected height of general planted plane trees and of odd-even merge sorting networks. It allows derivation of full asymptotic expansions rather easily.

B. Let a_n be a sequence of numbers; their *Poisson generating function* is defined as:

$$a(x) = e^{-x} \sum_{n \ge 0} a_n \frac{x^n}{n!} . \tag{25}$$

Many applications require determining asymptotic properties of the a_n from their Poisson generating function.

One way of proceeding is to consider the *Mellin transform* of (25) which is easily found to involve a *Newton series:*

$$a^*(s) = \Gamma(s) \sum_{n \ge 0} a_n \frac{s(s+1)(s+2)\cdots(s+n-1)}{n!} . \tag{26}$$

Thus if everything goes well, a_n is nothing but the n-th difference of $\alpha(s) \equiv a^*(s)/\Gamma(s)$:

$$a_0 = \alpha(0) \; ; \; a_1 = \alpha(0)-\alpha(-1) \; ; \; a_2 = \alpha(0)-2\alpha(-1)+\alpha(-2) \quad \cdots . \tag{27a}$$

From there, the a_n can be recovered by a classical formula from the calculus of finite differences and the theory of Newton series:

$$a_n = \frac{n!}{2i\pi} \int_\Lambda \frac{a^*(s)}{\Gamma(s)} \frac{ds}{s(s+1)(s+2)\cdots(s+n)} \tag{27b}$$

for a small contour around points $0,-1,-2,-3,\cdots$. This method was used by S.O. Rice as pointed out by Knuth. We call it the *Poisson-Mellin-Newton cycle*. A similar method is used to analyze digital search trees (see a forthcoming paper of Flajolet and Sedgewick). Notice that Knuth's use of the method is restricted to cases where an explicit form for the a_n has been obtained (by expanding generating functions) while our approach is more general since it accounts directly for the shape of the integrand in (27b). (See below an example of application to "trie sums").

6. Sample applications

We shall restrict ourselves here to presenting a few applications of the above methods: to carry propagation, to interpolation search, and to trie sums.

6.1. Interpolation search

Interpolation search is a way of searching an ordered table using the value of the record to be found in order to calculate an address at which it is likely to be found. It has been shown by several authors that the expected cost of such a search in a file of size n is $\log_2 \log_2 n + O(1)$. We propose here to reexamine some of the steps of a derivation of that result by Gonnet.

By a sequence of analytic and probabilistic arguments, Gonnet proves that the probability that more than k probes are required is majorized by the quantity:

$$p_k(t) = \prod_{i=1}^{k} (1-\frac{1}{2}e^{-t2^{-i}})$$

where $e^{-t} = \frac{8}{\pi n}$ $(t = \log \pi n / 8)$. Thus the expected number of probes is itself majorized by the function:

$$F(t) = \sum_{k=0}^{\infty} p_k(t) .$$

To evaluate F asymptotically for small t, introduce the infinite product:

$$Q(t) = \prod_{i=1}^{\infty}(1-\frac{1}{2}e^{-t2^{-i}})$$

so that $F(t)$ becomes:

$$F(t) = \frac{1}{Q(t)} \sum_{k=0}^{\infty} Q(t2^{-k}) ,$$

and the sum is a typical harmonic sum to which methods of Section 2 apply. One obtains in this way an asymptotic expansion of the form:

$$F(t) = K \log_2 t + P(\log_2 t) + \cdots \qquad (t \to \infty) \tag{28}$$

where P is a periodic function whose Fourier coefficients have simple expressions in terms of the values of Q^* at points $\chi_k = \frac{2ik\pi}{\log 2}$.

Equation (28) leads to the loglog result for interpolation search since t increases as $\log n$.

6.2. Carry propagation

The following problem arises in a work by Knuth relative to certain binary adders. What is the expected length of the longest sequence of (consecutive) *ones* in a random 0–1 string of length n?

Let $p_{n,k}$ denote the probability that this longest sequence has length k. Then it is known (see Feller's book) that:

$$p_k(x) = \sum_{n\geq 0} p_{n,k} x^n = \frac{1-2^{-k-1}x^{k+1}}{1-x+2^{-k-2}x^{k+2}} .$$

It is easily seen that $p_k(x)$ has a unique pole of smallest modulus that satisfies:

$$\rho_k = 1 + 2^{-k-2} + O(k2^{-2k}) .$$

Numerically tight approximations for the probabilities that also prove sufficient for asymptotic analysis are obtained by retaining only the contribution in $p_{n,k}$ coming from the dominant pole. This leads to the approximations:

$$\begin{aligned} p_{n,k} &\approx \rho_k^{-n-1} \\ &\approx (1+2^{-k-2})^{-n-1} \\ &\approx e^{-n2^{-k-2}} . \end{aligned}$$

Using the above approximation in the expression $R_n = \sum_k 1-p_{n,k}$ representing the expected length of the longest run of ones in a random binary string of length n (this use can be justified rigorously), one is led to the approximation $R_n \sim F(n)$ where $F(x)$ is a familiar harmonic sum:

$$F(x) = \sum_k 1-e^{-x2^{-k-2}}$$

to which previously developed methods apply.

6.3. Trie sums

We refer the reader to vol. 3 of Knuth's *The Art of Computer Programming* for the origin of the problem (see the analysis of radix exchange sort, pp. 131 *et sq*. and the analysis of trie searching).

The question is to determine the asymptotic behaviour of a sequence of numbers $\{a_n\}$ whose exponential generating function $\sum_{n\geq 0} a_n x^n / n!$ satisfies the functional equation:

$$a(x) = 2e^{x/2}a(\frac{x}{2}) + x(e^x-1)\ . \tag{30}$$

Given the initial conditions $a_0 = a_1 = 0$, one can solve (30) by iteration and take coefficients of the solution from which there results:

$$a_n = n \sum_{k\geq 0}[1-(1-\frac{1}{2^k})^{n-1}]\ . \tag{31}$$

One can also notice from (30), that the Poisson generating function $b(x)$ of the a_n satisfies the simpler equation:

$$b(x) = 2b(\frac{x}{2}) + x(1-e^{-x})\ , \tag{32}$$

from which an alternative expression of the a_n follows by identifying coefficients.

We propose now a brief exposition of three different approaches to the asymptotic evaluation of the a_n.

Classical approach based on exponential approximations

This is Knuth's approach following suggestions by De Bruijn. it starts by using repeatedly the exponential approximation:

$$(1-a)^n \approx e^{-an} \tag{33}$$

which in the context of (31) leads to (after somewhat unpleasant real analysis):

$$a_n = n\,F(n) + o(n) \quad ; \quad F(x) = \sum_{k\geq 0}[1-e^{-x2^k}]\ . \tag{34}$$

Function $F(x)$ being clearly a harmonic sum, an asymptotic expansion involving periodicities appears, as usual:

$$F(x) = \log_2 x + P(\log_2 x) + \cdots \tag{35}$$

This method does not seem to allow for derivation of full asymptotic expansions, due to the limitations of real analysis methods involved in the derivation of approximation (34)

Direct approach

This consists in observing that the the a_n *are* simply expressible in terms of a harmonic sum; one has $a_{n+1}/n+1 = 1+G(n)$, where G is given by:

$$G(x) = \sum_{k\geq 1}[1-\exp(-\mu_k x)] \quad ; \quad \mu_k = \log(1-\frac{1}{2^k})^{-1} . \tag{36}$$

Thus $G^*(s) = \mu(s)\Gamma(s)$, for $-1<\mathrm{Re}(s)<0$, with:

$$\mu(s) = \sum_{k\geq 1}\mu_k^{-s} . \tag{37}$$

To obtain an analytic continuation of $\mu(s)$, notice that:

$$\begin{aligned}(\log(1-u)^{-1})^{-s} &= u^{-s}\,(1+\frac{u}{2}+\frac{u^2}{3}+\frac{u^3}{4}+\cdots)^{-s} \\ &= u^{-s}\sum_{j\geq 0}c_j(s)u^j .\end{aligned}$$

Thus using this expansion in the definition of μ, we find the meromorphic approximation of μ:

$$\mu(s) \approx \sum_{j=0}^{\infty}c_j(s)\,\frac{2^{s-j}}{1-2^{s-j}} \tag{38}$$

(convergence is not implied by this equation; only taking k terms of the sum will give a function that differs from $\mu(s)$ by a function analytic for $\mathrm{Re}(s)<k$).

Thus we see that $\mu(s)$ is meromorphic in the whole of the complex plane, with singularities that obtain from (38). Hence $G(x)$ admits a full asymptotic expansion of the form:

$$G(x) = \log_2 x \; + \; \sum_{j\geq 0}P_j(\log_2 x)\,x^{-j} \tag{39}$$

where the P_j are periodic with period 1 and with Fourier coefficients expressible in terms of values of the gamma function and the polynomials c_j.

The Poisson-Mellin-Newton cycle

One starts there from Equation (32) giving a functional equation for the Poisson generating function of the a_n; the Mellin transform of $b(x)$ is defined for $-2<\mathrm{Re}(s)<-1$ and satisfies the the transform equation:

$$b^*(s) = 2^{1+s}b^*(s) - s\Gamma(s) , \tag{40}$$

so that solving, we get:

$$b^*(s) = \frac{s\Gamma(s)}{1-2^{1+s}} . \tag{41}$$

From the "cycle" explained above, we see that:

$$a_n = \frac{n!}{2i\pi}\int_{\Lambda}(1-2^{1+s})^{-1}\,\frac{ds}{(s+1)(s+2)..(s+n)} \tag{42}$$

the integrand being exactly $b^*(s)/\Gamma(s)$ multiplied by the standard weight function $[s(s+1)\cdots s(n)]^{-1}$; there Λ is a skinny curve encircling the points $-2,\cdots,-n$. Moving the contour of integration and taking residues into account leads to closed form expressions and asymptotic expansions of the a_n that have a lot of similarities with Mellin inversion expansions.

7. Conclusion

We have presented here some applications of powerful Mellin transform techniques, trying to show that there are a few well recognizable cases of applications.

Applications to analysis of algorithms can be found in the following domains:

Tree parameters: height of trees; register allocation; height of random walks and analysis of dynamic data structures (stack "histories").

Trie parameters: radix exchange sort; trie searching; digital search trees; dynamic and extendible hashing; communication protocols.

Carry propagation and binary adders; odd-even merge sorting networks; interpolation search; longest probe sequence in hashing; probabilistic counting; approximate counting *etc*...

We do not have space here for a complete exposition or an even partial bibliography, for which the reader is referred to a forthcoming paper [1].

Bibliography

[1] P. Flajolet, M. Regnier, R. Sedgewick: "Mellin transform techniques and the analysis of algorithms", in preparation (1984).