PUC

Series: Monografias em Ciência da Computação
Nº 23/77

NOTES ON THE DERIVATION OF ASYMPTOTIC EXPRESSIONS FROM SUMMATIONS

by

Gaston H. Gonnet

Departamento de Informática

Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vivente, 225 — ZC-19
Rio de Janeiro — Brasil

Series: Monografias em Ciência da Computação
Nº 23/77

Series Editor: Michael Challis December 1977

NOTES ON THE DERIVATION OF ASYMPTOTIC EXPRESSIONS FROM SUMMATIONS*

by

Gaston H. Gonnet

*This work was partially financed by FINEP
This paper was submitted for publication elsewhere. As a courtesy to the publishers, it should not be widely distributed.

RESUMO

Este trabalho contém ferramentas matemáticas aplicáveis à derivação de expressões assintóticas de somatórios.

A primeira é uma generalização da fórmula de Euler-Maclaurin a qual resulta efetiva na avaliação do termo constante. A segunda é a aplicação de equações funcionais. Ambos métodos são ilustrados com exemplos tirados da literatura de análise de complexidade.

Palavras chave: Análise assintótica; Análise de algoritmos; constantes de Euler-Maclaurin; Equações funcionais; Função zeta de Riemann; Funções analíticas, Somatórios.

ABSTRACT

This paper contains mathematical tools for the derivation of asymptotic expressions from summations. The first is a generalized form of the Euler-Maclaurin summation formula that proves very effective in the evaluation of the constant term. The second is the application of functional equations. Both methods are illustrated with several examples taken from the complexity analysis literature.

Keywords. Asymptotic analysis; Analysis of algorithms; Euler-Maclaurin formula; Euler-Maclaurin constant; Functional equations; Riemann's zeta function; Analytic functions; Summations.

INTRODUCTION

The analysis of algorithms very often requires the summation of series. It is not frequent, however, that we are able to find closed formulas for these summations. Since the analysis of algorithms is mainly concerned with asymptotic behaviour, it is convenient and desirable to express summations as asymptotic expansions.

There are well known tools to handle this problem, among which the most significant is the Euler-Maclaurin summation formula. There are, however, many cases in which this formula does not produce the required results and we have to resort to an ad-hoc technique.

This paper contains some theoretical and practical aids in this area: a new expression for the Euler-Maclaurin formula, derived from a little known result by Barnes, and the use of functional equations. These are followed by several re-derivations of examples taken from the complexity analysis literature. In the next section we discuss the variation of the Euler-Maclaurin formula that eases the computation of the constant term in the expansion (usually called the Euler-Maclaurin constant). The third section is devoted to examples of application of functional equations to summations.

There are some standard references and bibliography for this material: Bruijn [70], Barnes [05], Knuth [68], Knuth [73], Hardy [49], Whittaker & Watson [27], Abramowitz & Stegun [64]. More detailed references are given within the text. We will present results rather than demonstrations, that can be found and/or derived from the above references. Notation is fairly standard, it is the same as Abramowitz [64]. We denote the Riemann's zeta function by

$$\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$$

and its analytic continuation. The Gamma function, defined by

$$\Gamma(z+1) = \int_0^{\infty} e^{-x} x^z \, dx = z\Gamma(z) = z!$$

and $\psi(z) = \Gamma'(z)/\Gamma(z)$, its logarithmic derivative. It is assumed that the reader is familiar with mathematical analysis. The asymptotic expansions are normally in n, and we assume that $n \to \infty$.

THE EULER-MACLAURIN CONSTANT

The Euler-Maclaurin summation formula is commonly written as

$$\sum_{k=1}^{n-1} f(k) = \int_1^n f(x)\,dx + \sum_{i=1}^{\infty} \left. \frac{B_i\, f^{(i-1)}(x)}{i!} \right|_{x=1}^{x=n}$$

The Euler-Maclaurin constant, denoted by $C(f,a)$, is defined by

$$\sum_{k=1}^{n-1} f(k) = \int_a^n f(x)\,dx + C(f,a) + \sum_{i=1}^{\infty} \frac{B_i\, f^{(i-1)}(n)}{i!}$$

where usually $a = 1$ or $a = 0$.

We will first give two examples that illustrate one of the most common problems with the Euler-Maclaurin formula. The first comes from the analysis of several algorithms, e.g. sorting [Knuth 73, 5.2.2]

$$B_s(n) = \sum_{k=1}^{\infty} \frac{e^{-k^2/n}}{k^s}$$

that presents difficulties for $s \neq 0, -1, -2, \ldots$ (any upper limit for the summation of order $O(n^{1/2+\varepsilon})$ or bigger will give the same asymptotic expansion).

The second example comes from the analysis of interpolation sequential search [Gonnet and Rogers 77]. It also frequently appears with the use of the binomial coefficient $\binom{n}{k}$.

$$A_s(n) = \sum_{k=1}^{n-1} [k(n-k)]^{-s}$$

for any non-integer s.

In both cases the main difficulty resides in the evaluation of the Euler-Maclaurin constant. For $A_s(n)$ we are also concerned about the integral, and the value at n that does not show an asymptotic behaviour in n. These two later problems can be solved by summing up to $\lfloor n/2 \rfloor$, using the symmetry of the summand, etc.

From the results of Barnes [05] we derive:

$$C(f,1) = \sum_i a_i \phi(-i) - \sum_i b_i \phi'(-i) + \sum_i c_i \phi''(-i) - \ldots$$

if we can express the function as

$$f(x) = \sum_i a_i x^i + \sum_i b_i \ln(x)\, x^i + \sum_i c_i \ln^2(x)\, x^i + \ldots$$

where the expansion is in ascending powers of x (i may be real), and

$$\phi(i) = \zeta(i) - (i-1)^{-1}; \quad \phi(1) = \gamma; \quad \phi'(1) = -\gamma_1$$

[Briggs & Chowla 55, Abramowitz 64, 23.2.5]. If the integral, in its general expression, converges for $a = 0$ (usually only $a_{-1} = 0$ is required) we obtain the even simpler formula

$$c(f,0) = \sum_i a_i \zeta(-i) - \sum_i b_i \zeta'(-i) + \ldots .$$

For simpler functions, with only one term in the expansion, we find this result in Hardy [49] ($f(x) = x^{-s}$) and Bruijn [70] ($f(x) = \ln(x)$).

With this tool we try the previous examples to find that:

$$B_s(n) = \frac{n^{(1-s)/2}\,\Gamma((1-s)/2)}{2} + \zeta(s) - \frac{\zeta(s-2)}{n} + \frac{\zeta(s-4)}{2n^2} - \ldots$$

($s \neq 1, 3, 5, \ldots$) which has the asymptotic behaviour for which we are looking. For this special case see also Knuth [73, ex. 5.2.2.51]. Before solving for the exceptional s, we are encouraged to generalize further, for example

$$B_{s,t}(n) = \sum_{k=1}^{\infty} \frac{e^{-k^t/n}}{k^s} = \frac{n^{(1-s)/t}\,\Gamma((1-s)/t)}{t} + \zeta(s) - \frac{\zeta(s-t)}{n} + \frac{\zeta(s-2t)}{2n^2} - \cdots$$

provided that $s - mt \neq 1$, $m = 0, 1, 2, \ldots$. The last exception condition (the integral does not converge) can be treated with $C(f,1)$; i.e. integrating from 1 to ∞, the integral becomes the exponential integral, $E_m(1/n)$, and after simplification we derive

$$B_{s,t}(n) = \frac{(-1/n)^m}{tm!}\,[ln(n) + \psi(m+1)] + \zeta(s) - \frac{\zeta(s-t)}{n} + \frac{\zeta(s-2t)}{2n^2} - \cdots$$

with $s - mt = 1$, and $\zeta(s - mt)$ interpreted as γ.

In the second example, assuming that n is even, we can write

$$A_s(n) = 2\sum_{k=1}^{\frac{n}{2}-1} [k(n-k)]^{-s} + (2/n)^{-2s} .$$

Computing the integral from 0 ($a = 0$) we obtain

$$A_s(n) = \frac{(n/2)^{1-2s}\sqrt{\pi}\,\Gamma(1-s)}{\Gamma(3/2-s)} + 2C(f,0) + 2\sum_{i=1}^{\infty} \frac{B_i\, f^{(i-1)}(n/2)}{i!} + (2/n)^{-2s}$$

The term in B_1 cancels with the last summand. The function is symmetric with respect to $y = n/2$, consequently all odd derivatives vanish and since

$$f(x) = [x(n-x)]^{-s} = n^{-s}(x^{-s} + \frac{sx^{1-s}}{n} + \frac{s(s+1)x^{2-s}}{2n^2} + \ldots)$$

we derive the final expression:

$$A_s(n) = \frac{(n/2)^{1-2s}\sqrt{\pi}\,\Gamma(1-s)}{\Gamma(3/2-s)} + 2n^{-s}[\zeta(s) + \frac{s\zeta(s-1)}{n} + \frac{s(s+1)\,\zeta(s-2)}{2n^2} + \ldots]$$

which has the desired asymptotic behaviour in n. Surprisingly this formula works, not only for the even, but for all n, and for all $s \neq 1, 2, 3, \ldots$

A primary analysis of the insertion of the last keys in a hashing table [Brent 73] requires the evaluation of

$$A(\alpha) = \sum_{k\geq 0} \alpha^{k(k+1)/2} = 1 + \alpha + \alpha^3 + \alpha^6 + \alpha^{10} + \ldots$$

when $\alpha \to 1^-$. We can write this summation as

$$A(\alpha) = 1 + \sum_{k\geq 1} e^{k(k+1)\ln(\alpha)/2} .$$

Using the modified Euler-Maclaurin formula we find

$$A(\alpha) = 1 + \int_{-1/2}^{\infty} e^{x(x+1)\ln(\alpha)/2}\, dx - \int_{-1/2}^{0} e^{x(x+1)\ln(\alpha)/2}\, dx + \zeta(0) +$$

$$+ \frac{(\zeta(-2) + \zeta(-1))\ \ln(\alpha)}{2n} + \ldots$$

$$= 1 + \alpha^{-1/8}\left[\sqrt{\frac{\pi}{2\ \ln(\alpha)}} - \frac{1}{2} - \frac{\ln(\alpha)}{48} - \frac{\ln^2(\alpha)}{1280}\right] - \frac{1}{2}$$

$$- \frac{\ln(\alpha)}{24} + \frac{\ln^2(\alpha)}{480} + O(\ln^3(\alpha)).$$

After simplification, we obtain

$$A(\alpha) = \alpha^{-1/8} \sqrt{\frac{\pi}{-2\ \ln(\alpha)}} + O(\ln^3(\alpha)) .$$

Using a different argument we can show that the first is the only term in the asymptotic expansion.

Further examples are the harmonic numbers

$$H_n = \sum_{k=1}^{n} 1/k$$

and $n!$ where besides the well known expansions we find the constants,

$\gamma = \phi(1)$ and $\sqrt{2\pi} = e^{-\zeta'(0)}$ respectively. The so-called Glaisher's constant, [Glaisher 1877, Knuth 68], i.e. A in the equation

$$1^1.2^2.3^3.4^4. \ldots n^n = A\, n^{(n^2+n+1/6)/2}\, e^{-n^2/4}(1 + O(n^{-1})) ,$$

is found to be with our scheme:

$$A = e^{-\zeta'(-1)+1/12} = 1.2842\ 2712\ 9100 \ldots$$

Alternating series, or summations over the odd integers etc., are treated systematically by solving first for the general expression

$$\sum_{k=1}^{n} f(ak) = F(a,n)$$

We derive then

$$\sum_{k=1}^{n} (-1)^k f(k) = F(a,n) - 2F(2a, \lfloor n/2 \rfloor) ,$$

$$\sum_{\substack{k=1 \\ k \text{ odd}}}^{n} f(k) = F(a,n) - F(2a, \lfloor n/2 \rfloor) , \text{ etc.}$$

FUNCTIONAL EQUATION APPROACH

In some cases we find sums for which we can easily verify a functional equation. The general solution of the functional equation provides the result we are seeking.

The analysis of radix exchange sorting [Knuth 73, 5.2.2] requires the computation of

$$T_n = \sum_{j\geq 1} (2^j e^{-n/2^j} - 2^j + n)$$

Looking for a functional equation, we find by inspection that

$$T_{2n} = 2(e^{-n} - 1 + n) + 2T_n .$$

The general solution of the above is

$$T_n = n\, log_2(n) + 2 + nP(log_2(n)) + O(e^{-n})$$

$$= n\, log_2(n) + 2 + \left(\frac{\gamma - 1}{ln\ 2} - \frac{1}{2}\right)n + nP^*(log_2(n)) + O(e^{-n})$$

where $P(x)$ and $P^*(x)$ are periodic functions with period 1. The latter result, derived in Knuth [73], gives an explicit expression for $P^*(x)$ which oscillates around 0.

Similarly, the analysis of optimal-minimax hashing [Gonnet 77-1] requires the evaluation of

$$Q(n) = \sum_{k\geq 0} (1 - e^{-ne^{-k}})$$

which is closely related to the derivative of the above, i.e. T'_n. $Q(n)$ follows the functional equation

$$Q(n) = Q(ne) - 1 + e^{-n} ,$$

that has a general solution

$$Q(n) = ln(n) + P(ln\ n) + O(e^{-n})$$

$$= ln(n) + \gamma + \frac{1}{2} + P^*(ln\ n) + O(e^{-n})$$

In the last two examples the functions $P^*(x)$ are computed to oscillate "centred" around 0. Under this condition they are negligible for practical purposes. We also find, going back to the original summations, that all derivatives of these functions are bounded.

The analysis of another method for inserting the last elements in a hash table [Gonnet and Munro 77] requires the evaluation of

$$D(\alpha) = \sum_{k \geq 0} \alpha^{2^k - 1} = 1 + \alpha + \alpha^3 + \alpha^7 + \alpha^{15} + \ldots$$

when $\alpha \to 1^-$. By inspection we find the functional equation

$$\alpha D(\alpha^2) = D(\alpha) - 1$$

that has a general solution

$$\alpha D(\alpha) = lg(1-\alpha) + \alpha\left(1 - \frac{1}{2\ ln\ 2}\right)(1-\alpha) + \alpha\left(\frac{4}{3} - \frac{17}{24\ ln\ 2}\right)(1-\alpha)^2$$

$$+ O((1-\alpha)^3) + P(lg(-\ lg\ \alpha))$$

where $P(x)$ is periodic with period 1, and we can write

$$P(x) = -0.332746176 \ldots + P^*(x)$$

so that

$$|P^*(x)| \leq 0.00000316 \ldots$$

Further examples can be found in [Gonnet 77-2].

We can conclude from these examples that a) some summations, very difficult to approach with the usual tools, (especially those doubly exponential) become quite easy to solve; b) we are normally left with an unknown periodic function. It is desirable to decompose this function into a constant and a periodic component which oscillates around 0. In the examples analyzed the periodic contribution is small enough to be ignored for practical purposes.

APPENDIX. SPECIAL VALUES OF THE ZETA FUNCTIONS

$$\zeta(s) = 2(2\pi)^{s-1} \sin(\tfrac{1}{2}\pi s)\, \Gamma(1-s)\, \zeta(1-s)$$

$$\zeta(2m) = \frac{(2\pi)^{2m}(-1)^{m-1}}{2\,(2m)!} B_{2m} \qquad m = 1, 2, 3, \ldots$$

$$\zeta(-2m) = 0$$

$$\zeta(1-2m) = -\frac{B_{2m}}{2m}$$

$$\zeta'(-2m) = -\frac{(2m)!\,\zeta(2m+1)}{2(2\pi)^{2m}}$$

$$\zeta(0) = -\frac{1}{2}$$

$$\zeta'(0) = -\frac{1}{2}\ln(2\pi)$$

$$\frac{\zeta'(\frac{1}{2})}{\zeta(\frac{1}{2})} = \frac{\ln(8\pi) + \gamma + \pi/2}{2}$$

$$\zeta(\tfrac{1}{2} - i) = \frac{(-1)^{\lfloor\frac{i+1}{2}\rfloor}\, 1.3 \ldots (2i-1)}{(4\pi)^{i}}\, \zeta(\tfrac{1}{2} + i)$$

$$\frac{-2m\,\zeta'(1-2m)}{B_{2m}} = \ln(2\pi) - \psi(2m) - \frac{\zeta'(2m)}{\zeta(2m)}$$

$$\sum_{m=2}^{\infty} \frac{(-z)^m \zeta(m)}{m} = \gamma z + \ln(\Gamma(z+1))$$

$$\sum_{m=2}^{\infty} (-z)^{m-1} \zeta(m) = -\gamma - \psi(z+1)$$

$$\sum_{m=-\infty}^{\infty} \zeta(a+bm)z^m = -\int_0^{\infty} \frac{dx}{(1 - z/x^b)x^a} = \begin{cases} \dfrac{z^{(1-a)/b}}{b}\,\pi \cot\left(\dfrac{a-1}{b}\pi\right) & (z > 0) \\[2ex] \dfrac{z^{(1-a)/b}}{b}\,\pi \csc\left(\dfrac{a-1}{b}\pi\right) & (z < 0) \end{cases}$$

The following table shows some selected values of $\zeta(s)$ and its derivatives, rounded to 12 decimal places

s	ζ(s)	ζ'(s)	ζ"(s)
-3	0.0083 3333 3333	0.0053 7857 6358	-0.0142 5786 21
-5/2	0.0085 1692 8778	-0.0062 6573 6369	-0.0337 9771 804
-2	0.0	-0.0304 4845 7058	-0.0657 6351 619
-3/2	-0.0254 8520 1890	-0.0763 0925 5320	-0.1241 5334 3793
-1	-0.0833 3333 3333	-0.1654 2114 3700	-0.2502 0442 4110
-1/2	-0.2078 8622 4977	-0.3608 5433 9600	-0.5962 2917 6765
0	-0.5	-0.9189 3853 3205	-2.0063 5645 5909
1/2	-1.4603 5450 8810	-3.9226 4613 9209	-16.0083 5701 3929
1	0.5772 1566 4902	0.0728 1584 5484	-0.0096 9036 3193 ($\phi^{(n)}(1)$)
3/2	2.6123 7534 8685	-3.9322 3973 7431	15.9895 5637 1226
2	1.6449 3406 6848	-0.9375 4825 4316	1.9892 8023 4299
5/2	1.3414 8725 7251	-0.3873 4195 0326	0.5819 6892 7042
3	1.2020 5690 3160	-0.1981 2624 2885	0.2397 4691 7305

APPENDIX. SUMMATION OF BERNOULLI NUMBERS

Naive application of the Euler-Maclaurin formula may lead to a constant that is defined by a series involving the Bernoulli numbers. These series are, in general, only semi-convergent [Hardy 49], and we need some extra tool to sum them.

From the asymptotic expansion of $\psi(z)$ we find the identity:

$$\sum_{i=\infty}^{\infty} (-1)^i \frac{B_i \, \Gamma(i+k+1)}{\Gamma(i+1)} = \psi^{(k)}(1) = \Gamma(k+2)\,\zeta(k+2) \qquad \text{(A.1)}$$

Taking limits for $k = 2, 1, 0, -1, -2, -3, \ldots$ we derive:

$$\sum_{i=0}^{\infty} (-1)^i B_i (i+1)(i+2) = 6\zeta(4) = 6.4939\ 3940\ 2266\ \ldots$$

$$\sum_{i=0}^{\infty} (-1)^i B_i (i+1) = 2\zeta(3) = 2.4041\ 1380\ 5319\ \ldots$$

$$\sum_{i=0}^{\infty} (-1)^i B_i = \frac{\pi^2}{6}$$

$$\sum_{i=1}^{\infty} \frac{(-1)^i B_i}{i} = \gamma$$

$$\sum_{i=2}^{\infty} \frac{(-1)^i B_i}{i(i-1)} = 1 - \frac{1}{2} \ln(2\pi)$$

$$\sum_{i=3}^{\infty} \frac{(-1)^i B_i}{i(i-1)(i-2)} = -\zeta'(-1) - \frac{1}{6} = -0.0012\ 4552\ 2966\ \ldots$$

$$\sum_{i=4}^{\infty} \frac{(-1)^i B_i}{i(i-1)(i-2)(i-3)} = \frac{\zeta'(-2)}{2} + \frac{1}{72} = -0.0013\ 3533\ 9640\ \ldots$$

Taking the derivative of formula (A.1) with respect to k we find another formula involving the Bernoulli numbers, namely:

$$\sum_{i=0}^{\infty} \frac{(-1)^i B_i \psi(i+k+1)\Gamma(i+k+1)}{\Gamma(i+1)} = \psi(k+2)\ \Gamma(k+2)\ \zeta(k+2) + \Gamma(k+2)\ \zeta'(k+2)$$

from which we derive for $k = 1, 0, -1, -2, \ldots$

$$\sum_{i=0}^{\infty} (-1)^i B_i H_{i+1}(i+1) = 3\zeta(3) + 2\zeta'(3) = 3.2099\ 1822\ 3707\ \ldots$$

$$\sum_{i=0}^{\infty} (-1)^i B_i H_i = \frac{\pi^2}{6} + \zeta'(2) = 0.7073\ 8581\ 2532\ \ldots$$

$$\sum_{i=1}^{\infty} \frac{(-1)^i B_i H_{i-1}}{i} = -\gamma_1 = -\zeta''(0) - \frac{\pi^2}{24} + \frac{\gamma^2}{2} - \frac{\ln^2(2\pi)}{2} = 0.0728\ 1584\ 5483\ \ldots$$

$$\sum_{i=2}^{\infty} \frac{(-1)^i B_i H_{i-2}}{i(i-1)} = 1 + \frac{\zeta''(0)}{2} = -0.0031\ 7822\ 7954\ \ldots$$

REFERENCES

Abramowitz, M. and Stegun, I. A. Handbook of Mathematical Functions, Dover, New York, 1964.

Barnes, E. W. "The Maclaurin Sum Formula", Proceedings of the London Mathematical Society (2), 3 (1905), pp. 253-272.

Brent, R. P. "Reducing the Retrieval Time of Scatter Storage Techniques", CACM 16, 2 (Feb. 1973), pp. 105-109.

Briggs, W. E. and Chowla, S. The Power Series Coefficients of $\zeta(s)$, A.M.M. 62, 5 (May 1955), pp. 323-325.

de Bruijn, N. G. Asymptotic Methods in Analysis, North Holland, Amsterdam, 1970.

Glaisher, J. W. L. On the Product $1^1.2^2.3^3... n^n$", Messenger Math, 7(1877), pp. 43-47.

Gonnet, G. H. and Rogers, L. D. "The Interpolation Sequential Search Algorithm", Information Processing Letters 6, 4 (Aug. 1977), pp. 136-139.

Gonnet, G. H. and Munro, J. I. "The Analysis of an Improved Hashing Technique", Proceedings of the Ninth Annual ACM Symposium on the Theory of Computing, Boulder, Colorado (May 1977).

Gonnet, G. H. (1) "Average Lower Bounds for Open-Addressing Hash Coding", Proceedings of the Conference on Theoretical Computer Science, Waterloo, Ontario, Canada (Aug 1977), pp. 159-162.

Gonnet, G. H. (2) "Interpolation and Interpolation Hash Searching" Doctoral Dissertation, University of Waterloo, CS-RR-77-02 (Jan 1977).

Hardy, G. H. Divergent Series, Oxford, Clarendon Press, 1949.

Knuth, D. E. The Art of Computer Programming, Vol. I & Vol. III, Ch 1,5,6. Addison Wesley, Don Mills, 1968 & 1973.

Whittaker, E. T. and Watson, G. N. A Course of Modern Analysis, Cambridge, University Press, 1927.