ON A CONSTANT ARISING IN THE ANALYSIS OF BIT COMPARISONS IN QUICKSELECT

PETER J. GRABNER† AND HELMUT PRODINGER‡

Abstract. A (real) constant that appears as the factor of the leading term of the average number of bit comparisons required by quickselect, and was originally given in terms of complex numbers, is expressed using real numbers alone. A further representation is derived which is converging very quickly.

Methods include residue calculus and the Euler-MacLaurin summation formula.

Fill and Nakama in their study [1] worked out the following constant as the factor of the leading term of the average number of bit comparisons required by quickselect:

$$c = \frac{28}{9} + \frac{17-6\gamma}{9\log 2} - \frac{4}{\log 2}\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{\zeta(1-\chi_k)\Gamma(1-\chi_k)}{\Gamma(4-\chi_k)(1-\chi_k)},$$

with $\chi_k = \frac{2\pi ik}{\log 2}$.

In this short note we want to elaborate on the sum appearing in this constant (which is real, although it is presented using complex numbers); we rewrite it using the formula $\Gamma(x+1) = x\Gamma(x)$ and the symmetry $\chi_{-k} = -\chi_k$ as

$$S = \sum_{k\neq 0} \frac{\zeta(1+\chi_k)}{(3+\chi_k)(2+\chi_k)(1+\chi_k)^2}.$$

We note here that a similar analysis has been carried out in [2].

Using residue calculus we derive

$$S = \frac{\log 2}{2\pi i}\left[\int_{(\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2}\frac{ds}{2^s-1} - \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2}\frac{ds}{2^s-1}\right] + \frac{17}{36} - \frac{\gamma}{6} + \frac{1}{12}\log 2.$$

(We use the notation $\int_{(a)} \dots$ for $\int_{a-i\infty}^{a+i\infty} \dots$.)

Date: August 1, 2007.

Key words and phrases. Bit comparisons, quickselect, residues, Euler-MacLaurin formula.

† This author is supported by the Austrian Science Foundation FWF, project S9605, part of the Austrian National Research Network “Analytic Combinatorics and Probabilistic Number Theory”.

‡This author is funded by the NRF grant 2053748.

The first integral vanishes since the line of integration can be shifted to the right, where the integrand tends to zero; thus we have

$$S = -\frac{\log 2}{2\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2} \frac{ds}{2^s - 1} + \frac{17}{36} - \frac{\gamma}{6} + \frac{1}{12}\log 2.$$

Expanding $-\frac{1}{2^s-1}$ into a geometric series (note that $\Re s < 0$), we obtain

$$S = \log 2 \sum_{\ell=0}^{\infty} \frac{1}{2\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2} 2^{\ell s}\, ds + \frac{17}{36} - \frac{\gamma}{6} + \frac{1}{12}\log 2.$$

We now analyse the integral

$$\frac{1}{2\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2} 2^{\ell s}\, ds.$$

First we shift the line of integration to the right and expand the Riemann zeta function into a series to obtain

$$\frac{1}{2\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2} 2^{\ell s}\, ds = -\frac{1}{6} + \sum_{n=1}^{\infty} \frac{1}{n} \frac{1}{2\pi i} \int_{(1)} \frac{1}{(3+s)(2+s)(1+s)^2} \left(\frac{n}{2^\ell}\right)^{-s} ds.$$

The integral occurring inside the sum equals

$$\frac{1}{2\pi i} \int_{(1)} \frac{x^{-s}}{(3+s)(2+s)(1+s)^2}\, ds = \begin{cases} 0 & \text{for } x > 1, \\ -\frac{x^3}{4} + x^2 - \frac{3x}{4} - \frac{1}{2}x\log x & \text{for } 0 < x \le 1. \end{cases}$$

(The justification is again residue calculus. In our previous paper [2] we discussed several similar examples in detail.)

For $x > 0$ we set

$$\Phi(x) = -\frac{x^3}{4} + x^2 - \frac{3x}{4} - \frac{1}{2}x\log x.$$

Thus we have

$$\frac{1}{2\pi i} \int_{(-\frac{1}{2})} \frac{\zeta(1+s)}{(3+s)(2+s)(1+s)^2} 2^{\ell s}\, ds = -\frac{1}{6} + \sum_{n=1}^{2^\ell} \frac{1}{n} \Phi\left(\frac{n}{2^\ell}\right)$$

and consequently

$$S = \log 2 \sum_{\ell=0}^{\infty} \left(\sum_{n=1}^{2^\ell} \frac{1}{n} \Phi\left(\frac{n}{2^\ell}\right) - \frac{1}{6} \right) + \frac{17}{36} - \frac{\gamma}{6} + \frac{1}{12}\log 2.$$

The inner sum can now be simplified using the known summation formulæ for power sums

$$\sum_{n=1}^{2^\ell} \frac{1}{n}\Phi\left(\frac{n}{2^\ell}\right) - \frac{1}{6} = -\frac{1}{2} + 3\cdot 2^{-\ell-3} - \frac{1}{3}2^{-2\ell-3} - \frac{1}{2^{\ell+1}}\sum_{n=1}^{2^\ell} \log\frac{n}{2^\ell},$$

which simplifies the expression for S to

$$S = -\frac{\log 2}{2}\sum_{\ell=0}^{\infty}\left(2^{-\ell}\sum_{n=1}^{2^\ell}\log\frac{n}{2^\ell} + 1\right) + \frac{17-6\gamma}{36} + \frac{7}{9}\log 2$$

and furthermore

$$c = 2\sum_{\ell=0}^{\infty}\left(1 + 2^{-\ell}\sum_{n=1}^{2^\ell}\log\frac{n}{2^\ell}\right). \tag{1}$$

While this representation of the constant c is aesthetically appealing, its convergence is very poor. However, we are able to rework the formula, so that we can compute some 50 digits without much effort: Applying the Euler-MacLaurin summation formula to the remaining inner sum yields

$$\begin{aligned} 2^{-\ell}\sum_{n=1}^{2^\ell}\log\frac{n}{2^\ell} + 1 = \ell 2^{-\ell-1}\log 2 + 2^{-\ell-1}\log(2\pi) + \sum_{j=1}^{m}\frac{B_{2j}}{(2j-1)2j}\frac{1}{2^{2j\ell}} \\ - \frac{(2m)!}{2^{(2m+1)\ell}}\int_1^\infty \frac{P_{2m+1}(\{2^\ell t\})}{t^{2m+1}}\,dt \end{aligned}$$

for any $m \in \mathbb{N}$. The remainder term

$$\frac{1}{2}\frac{(2m)!}{2^{(2m+1)\ell}}\int_1^\infty \frac{P_{2m+1}(\{2^\ell t\})}{t^{2m+1}}\,dt$$

can be bounded by $\frac{1}{2}(2m-1)!2^{-(2m+1)\ell}\|P_{2m+1}\|$. For computing c we split the summation at L to obtain

$$\begin{aligned}
c &= 2\sum_{\ell=0}^{L}\left(2^{-\ell}\sum_{n=1}^{2^\ell}\log\frac{n}{2^\ell}+1\right)\\
&\quad+2\sum_{\ell=L+1}^{\infty}\Bigg(\ell 2^{-\ell-1}\log 2+2^{-\ell-1}\log(2\pi)+\sum_{j=1}^{m}\frac{B_{2j}}{(2j-1)2j}\frac{1}{2^{2j\ell}}\\
&\qquad\qquad-\frac{(2m)!}{2^{(2m+1)\ell}}\int_1^\infty\frac{P_{2m+1}(\{2^\ell t\})}{t^{2m+1}}\,dt\Bigg)\\
&= 2\sum_{\ell=0}^{L}\left(2^{-\ell}\sum_{n=1}^{2^\ell}\log\frac{n}{2^\ell}+1\right)\\
&\quad+(L+2)2^{-L}\log 2+2^{-L}\log(2\pi)+\sum_{j=1}^{m}\frac{B_{2j}}{j(2j-1)}\frac{2^{-2jL}}{2^{2j}-1}+R_{m,L},
\end{aligned}$$

where

$$|R_{m,L}|\le(2m-1)!\frac{2^{-(2m+1)L}}{2^{2m+1}-1}\|P_{2m+1}\|.$$

Taking $m=7$ and $L=13$ gives an error less than 10^{-53}. Thus we can give the following approximation for c

$$c = 5.27937\ 82410\ 80958\ 37386\ 56270\ 37785\ 81538\ 08364\ 11493\ 49031\dots.$$

Remark. In the paper [1], there are a few constants similar to c, which can all be treated in a similar manner.

References

[1] J. A. Fill and T. Nakama, *Analysis of the expected number of bit comparisons required by quickselect*, arXiv (2007), no. arXiv:0706.2437v1.

[2] P. J. Grabner and H. Prodinger, *Sorting algorithms for broadcast communications: Mathematical analysis*, Theor. Comput. Sci. **289** (2002), 51–67.

(P. Grabner) Institut für Analysis und Computational Number Theory, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria

E-mail address: peter.grabner@tugraz.at

(H. Prodinger) Department of Mathematics, University of Stellenbosch, 7602 Stellenbosch, South Africa

E-mail address: hproding@sun.ac.za