ASYMPTOTIC ESTIMATION OF A SUM OF DIGITS

Harald Riede
University of Koblenz-Landau, Mathematisches Institut, Rheinau 1, 56075 Koblenz, Germany
(*Submitted May 1996-Final Revision September 1996*)

Let $s(k)$ denote the sum of the base 10 digits of $k \in \mathbf{N}$. For natural $x \geq 2$ and arbitrary fixed exponent $m \in \mathbf{N}$, it will be shown that

$$\frac{1}{x} \cdot \sum_{k=1}^{x-1} s(k)^m = \left(\frac{9}{2} \lg x\right)^m + O((\lg x)^{m-1}).$$

Here, "lg" denotes the base 10 log function. It is obvious that this formula can be generalized on arbitrary p-adic systems. The case $m=1$ has been treated in [1], $m=2$ in [2]; there the general case is exhibited as an *open problem*. The proof given now is based on induction.

I wish to thank Harald Scheid, University of Wuppertal, Germany, who drew my attention to certain unsolved arithmetical problems, the above among them.

1. THE ASSUMPTION

Let A_x for $x = 2, 3, \ldots$ be the arithmetic function

$$A_x(m) = \sum_{k=0}^{x-1} s(k)^m, \ m \in \mathbf{N}_0 \ (= \{0, 1, 2, \ldots\}).$$

I denote the above assertion in the following manner,

$$A_x(i) = x\left(\frac{9}{2} \lg x\right)^i + d_i(x) \cdot x(\lg x)^{i-1}, \ x \geq 2, \tag{1}$$

with certain bounded functions $d_i(x)$, i.e.,

$$|d_i(x)| \leq d_i \text{ for all } x, \tag{2}$$

and assume that it is valid for $i = 1, \ldots, m-1$. The validity for $i=1$ is guaranteed in [1] and the validity for $i=m$ will be deduced now in several steps.

2. A REDUCTION FORMULA FOR A_{10x}

The *binomial product* $B * C$ of two arithmetical functions is defined by

$$B * C(m) = \sum_{k=0}^{m} \binom{m}{k} B(k) \cdot C(m-k).$$

First, I will show that

$$A_{10x} = A_{10} * A_x. \tag{3}$$

$$A_{10x}(m) = \sum_{k=0}^{x-1} \sum_{i=0}^{9} s(10k+i)^m = \sum_{k=0}^{x-1} \sum_{i=0}^{9} (s(k)+i)^m = \sum_{k=0}^{x-1} \sum_{i=0}^{9} \sum_{j=0}^{m} \binom{m}{j} s(k)^{m-j} i^j$$

$$= \sum_k \sum_j \left(s(k)^{m-j} \binom{m}{j} \sum_{i=0}^{9} i^j \right) = \sum_k \sum_j s(k)^{m-j} \binom{m}{j} A_{10}(j) = \sum_j \left(\binom{m}{j} A_{10}(j) \sum_k s(k)^{m-j} \right)$$

$$= \sum_j \binom{m}{j} A_{10}(j) \cdot A_x(m-j) = (A_{10} * A_x)(m).$$

3. ESTIMATION OF THE REMAINDER

Let x have the decomposition $10y+z$ with $z<10$. Suppose $R_x = A_x - A_{10y}$. In the case $z=0$ we have $R_x = 0$, otherwise

$$R_x(m) = \sum_{i=0}^{z-1} s(10y+i)^m.$$

If $n+1$ denotes the number of digits of x, then

$$R_x(m) \le z \cdot ((n+1) \cdot 9)^m \le 9^{m+1} \cdot (n+1)^m.$$

Let $(a_n \ldots a_0)$ be the decimal representation of x and $x_k = (a_n \ldots a_k)$ (especially $x_0 = x$, $x_n = a_n$), then, in particular,

$$R_{x_k}(m) \le 9^{m+1}(n-k+1)^m, \; k = 0, 1, \ldots, n. \tag{4}$$

4. A DECOMPOSITION OF $A_x(m)$

One can verify immediately that

$$A_x = 10^n A_{x_n} + \sum_{k=1}^{n} (10^{k-1} A_{10x_k} - 10^k A_{x_k}) + \sum_{k=0}^{n-1} (10^k A_{x_k} - 10^k A_{10x_{k+1}})$$

$$\overset{(3)}{=} 10^n A_{a_n} + \sum_{k=1}^{n} 10^{k-1} (A_{10} * A_{x_k} - 10 A_{x_k}) + \sum_{k=0}^{n-1} 10^k R_{x_k},$$

$$A_x(m) = 10^n A_{a_n}(m) + \sum_{k=1}^{n} \left(10^{k-1} \sum_{i=1}^{m} \binom{m}{i} A_{10}(i) A_{x_k}(m-i) \right) + \sum_{k=0}^{n-1} 10^k R_{x_k}(m)$$

$$= \underbrace{10^n A_{a_n}(m)}_{U} + \underbrace{\sum_{i=1}^{m} \left(\underbrace{\binom{m}{i} \frac{A_{10}(i)}{10}}_{c_i} \underbrace{\sum_{k=1}^{n} 10^k A_{x_k}(m-i)}_{W_{m-i}} \right)}_{W} + \underbrace{\sum_{k=0}^{n-1} 10^k R_{x_k}(m)}_{V}$$

The expressions U, V, and W shall be treated now one after another.

5. ESTIMATION OF U AND V

$U = 10^n A_{a_n}(m) = 10^n R_{a_n}(m) \overset{(4)}{\le} 10^n \cdot 9^{m+1}$ and, since $10^n \le x < 10^{n+1}$, we have $U = O(x)$. Furthermore,

$$V = \sum_{k=0}^{n-1} 10^k R_{x_k}(m) \overset{(4)}{\le} 9^{m+1} \sum_{k=0}^{n-1} 10^k (n-k+1)^m.$$

Since the power series $\sum_k k^m z^k$ has radius of convergence 1, it is particularly convergent for $z = 1/10$; hence,

$$\sum_{k=0}^{n} 10^k (n-k+1)^m = 10^{n+1} \sum_{k=1}^{n+1} k^m \left(\frac{1}{10}\right)^k = O(x). \tag{5}$$

Thus, $V = O(x)$.

6. DECOMPOSITION AND ESTIMATION OF THE W_i

With respect to the assumption under induction, we obtain, for $i \le m-1$,

$$W_i = \sum_{i=1}^{n} 10^k A_{x_k}(i) = \sum_{k=1}^{n} 10^k \left(x_k \left(\frac{9}{2} \lg x_k\right)^i + d_i(x_k) \cdot x_k \cdot (\lg x_k)^{i-1} \right)$$

$$= \left(\frac{9}{2}\right)^i \underbrace{\sum_{k=1}^{n} 10^k x_k (\lg x_k)^i}_{G_i} + \underbrace{\sum_{k=1}^{n} d_i(x_k) \cdot 10^k x_k (\lg x_k)^{i-1}}_{G_i^*}.$$

Let $y_k = (a_k \ldots a_0)$. Then $10^k x_k = \underbrace{(a_n \ldots a_k 0 \ldots 0)}_{n+1 \text{ digits}} = x - (a_{k-1} \ldots a_0) = x - y_{k-1}$, so we have

$$G_i = \sum_{k=1}^{n} (x - y_{k-1})(\lg x_k)^i = x \sum_{k=1}^{n} (\lg x_k)^i - \sum_{k=1}^{n} y_{k-1} (\lg x_k)^i.$$

The two sums herein shall now be estimated separately:

a) We have $(n-k)^i = (\lg 10^{n-k})^i \le (\lg x_k)^i < (\lg 10^{n-k+1})^i = (n-k+1)^i$; hence,

$$\sum_{k=1}^{n} (n-k)^i \le \sum_{k} (\lg x_k)^i < \sum_{k=1}^{n} (n-k+1)^i = n^i + \sum_{k=1}^{n} (n-k)^i.$$

Since

$$\sum_{k=1}^{n} (n-k)^i = \sum_{k=1}^{n-1} k^i = \frac{n^{i+1}}{i+1} + O(n^i),$$

we see that, for arbitrary $i \in \mathbf{N}$,

$$\sum_{k=1}^{n} (\lg x_k)^i = \frac{n^{i+1}}{i+1} + O(n^i).$$

b) $\displaystyle\sum_{k=1}^{n} y_{k-1} (\lg x_k)^i \le \sum_{k} 10^k (\lg 10^{n-k+1})^i = \sum_{k} 10^k (n-k+1)^i \overset{(5)}{=} O(x).$

Putting the two parts together, we have

$$G_i = x \cdot \frac{n^{i+1}}{i+1} + O(x \cdot n^i),$$

particularly with respect to (2): $|G_i^*| \le d_i G_{i-1} = O(x \cdot n^i)$; therefore,

$$W_i = \left(\frac{9}{2}\right)^i G_i + G_i^* = \left(\frac{9}{2}\right)^i x \cdot \frac{n^{i+1}}{i+1} + O(x \cdot n^i) \text{ for all } i \le m-1.$$

Now it is easily seen that

$$W = \sum_{i=1}^{m} \binom{m}{i} \frac{A_{10}(i)}{10} W_{m-i} = m \frac{A_{10}(1)}{10} W_{m-1} + O(x \cdot n^{m-1})$$
$$= m \cdot \frac{9}{2} \cdot \left(\frac{9}{2}\right)^{m-1} x \cdot \frac{n^m}{m} + O(x \cdot n^{m-1}) = \left(\frac{9}{2}\right)^m x \cdot n^m + O(x \cdot n^{m-1}).$$

And, finally,

$$A_x(m) = \left(\frac{9}{2}\right)^m x \cdot n^m + O(x \cdot n^{m-1}).$$

From this, the initial assertion is deduced immediately.

Often a solved problem procreates a new problem. Here is an open question: Does the given asymptotic estimation hold even for arbitrary *real* $m \geq 1$? The reader is invited to prove or disprove this result.

REFERENCES

1. P. H. Cheo & S. C. Yien. "A Problem on the k-adic Representation of Positive integers." *Acta Math. Sinica* **5** (1955):433-48.
2. R. E. Kennedy & C. N. Cooper. "An Extension of a Theorem by Cheo and Yien Concerning Digital Sums." *The Fibonacci Quarterly* **29.2** (1991):145-49.

AMS Classification Numbers: 10J06, 10L99, 10A99

❖❖❖