REGULARITY OF A FUNCTION RELATED TO THE 2-ADIC LOGARITHM

JAN-CHRISTOPH SCHLAGE-PUCHTA

MSC-Index: 68Q45, 68Q70

For a function $f : \mathbb{N} \to X$ mapping the positive integers to some set X, define the q-kernel $K_q(f)$ as the set of functions $\{f_{k,\ell} : k \in \mathbb{N}, 0 \leq \ell < q^k\}$, where $f_{k,\ell}(n) = f(q^k n + \ell)$. The q-kernel is related to the concept of q-automaticity by the following criterion due to Eilenberg [2] (see also [1, Theorem 6.6.2]).

Theorem 1. *A function f is q-automatic if and only if $K_q(f)$ is finite.*

The notion of q-regularity generalizes the concept of q-automaticity in the case that X is the set of integers. A function f is called q-regular if $K_q(f)$ is contained in a finitely generated $\mathbb{Z}$-module.

Motivated by work of Lengyel [3] on the 2-adic logarithm, Allouche and Shallit [1, Problem 16.7.4] asked whether the function

$$f(n) = \min_{k \geq n}(k - \nu_2(k)), \tag{1}$$

where $\nu_2(k)$ is the 2-adic valuation, is 2-regular or not. Here we give a negative answer to this question. More precisely, we show the following.

Theorem 2. *The functions $f_{k,0} : n \mapsto f(2^k n)$ are $\mathbb{Q}$-linearly independent.*

For the proof we need the following simple statements concerning f.

Proposition 1. (1) *We have $f(n) = n - \mathcal{O}(\log n)$.*
(2) *For $n = (2^{\ell+2} - 3)2^m$ we have $f(n) = \min\big(n - m, n - m - \ell - 2 + 3 \cdot 2^m\big)$.*

Proof. (1) We trivially have the bound $f(n) \leq n$. On the other hand we have $\nu_2(k) \leq \frac{\log k}{\log 2}$, and hence $f(n) \geq \min_{k \geq n} k - \frac{\log k}{\log 2}$. Since the derivative of the function $t - \frac{\log t}{\log 2}$ is $1 - \frac{1}{t \log 2}$, which is positive for $t \geq 2$, for $n \geq 2$ the minimum is attained for $k = n$ and we conclude $f(n) \geq n - \frac{\log n}{\log 2}$, and the first claim is proven.

(2) We want to show that as k runds over all integers ≥n the minimum in (1) is attained at $k = n$ or at $k = 2^{\ell+m+2} = n + 3 \cdot 2^m$. From this our claim follows by computing the value of $k - \nu_2(k)$ at these two positions. Assume first that $k \geq n$ is not divisible by 2^{m+1}. Then we have $k - \nu_2(k) \geq n - \nu_2(k) \geq n - m$, which is what we want to have. Next assume that $\nu_2(k) > m$ and $k < 2^{\ell+m+2}$. Then $k = (2^{\ell+2} - 2)2^m$, that is, $\nu_2(k) = m + 1$, and we have $k - \nu_2(k) = (n + 2^m) - (m + 1) \geq n - m$, which is also consistent with our claim. For $k = 2^{\ell+m+2}$ we have $k - \nu_2(k) = n - m - \ell - 2 + 3 \cdot 2^m$, and thus it remains to consider the range $k > 2^{\ell+m+2}$. For $2^{\ell+m+2} < k < 2^{\ell+m+3}$ we have $k - \nu_2(k) \geq 2^{\ell+m+2} + 1 - (\ell+m+1) > 2^{\ell+m+2} - (\ell + m + 2)$, and hence this range cannot contribute to the minimum. Finally, if $k \geq 2^{\ell+m+3}$, then $k - \nu_2(k) \geq k - \frac{\log k}{\log 2} \geq 2^{\ell+m+3} - (\ell+m+3) > 2^{\ell+m+2} - (\ell+m+2)$, and this range is of no importance as well. Hence, the second claim follows as well. □

We now turn to the proof of the theorem. Assume the family of functions $(f_{k,0})_{k\geq 0}$ was linearly dependent. Then there exist rational numbers $\lambda_0, \dots, \lambda_p$, not all 0, such that

$$\sum_{j=0}^{p} \lambda_j f(2^j n) = 0 \tag{2}$$

holds for all integers n. Evaluating this equation asymptotically for $n \to \infty$ we see that the left hand side is $n \cdot \left(\sum_{j=0}^{p} 2^j \lambda_j\right) + \mathcal{O}(\log n)$. This expression can only vanish identically if

$$\sum_{j=0}^{p} 2^j \lambda_j = 0. \tag{3}$$

Let j_0 be the least integer satisfying $\lambda_{j_0} = 0$. Then define $\ell = 3 \cdot 2^{j_0} - 1$, and put $n = 2^\ell - 3$ into (2). We have

$$n - j_0 > n - j_0 - \ell - 2 + 3 \cdot 2^{j_0} = n - j_0 - 1.$$

On the other hand we have

$$n - j < n - j - \ell - 2 + 3 \cdot 2^j = n - j - 1 - (j - j_0) + 3 \cdot (2^j - 2^{j_0})$$

for all $j > j_0$, hence, by the second part of the proposition relation (2) becomes

$$\lambda_{j_0}(2^{j_0} n - j_0 - \ell - 2 + 3 \cdot 2^{j_0}) + \sum_{j=j_0}^{p} \lambda_j (2^j n - j) = 0. \tag{4}$$

Finally we put $n' = 2^{\ell+1} - 3$ into (2). The same computation as the one used for n yields the equation

$$\lambda_{j_0}(2^{j_0} n' - j_0 - \ell - 3 + 3 \cdot 2^{j_0}) + \sum_{j=j_0}^{p} \lambda_j (2^j n' - j) = 0. \tag{5}$$

Note that the difference between (4) and (5) is that n is replaced by n', and -2 is replaced by -3. If we take the difference of (4) and (5), we therefore obtain

$$\lambda_{j_0}(2^{j_0}(n' - n) + 1) + \sum_{j=j_0}^{p} \lambda_j 2^j (n' - n) = 0.$$

If we now multiply (3) by $(n - n')$, and subtract the result from the last equation, all that remains is $\lambda_{j_0} = 0$. But j_0 was chosen subject to the condition $\lambda_{j_0} \neq 0$. Hence, the initial assumption that not all λ_j are 0 is wrong, and we conclude that there is no linear relation among the functions $f_{k,0}$.

The reader might wonder why we restricted our attention to the functions $f_{k,0}$. Essentially the same method of proof can be used to show that the dimension of the linear span $\langle f_{k,0}, f_{k_1}, \dots, f_{k,2^k-1}\rangle$ tends to infinity with k. However, things become notationally more involved, since these functions are no longer linearly independent. In fact, we have $f_{k,a} = f_{k,a+1}$ for every odd a and many more identities like this, that is, these functions are not even different, and to give a lower bound for the dimension we have to choose a suitable subset.

References

[1] J.-P. Allouche, J. Shallit, *Automatic Sequences*, Cambridge University Press, Cambridge, 2003.
[2] Eilenberg, *Automata, Languages, and Machines*, Academic Press, New York, 1974.
[3] T. Lengyel, Characterizing the 2-adic order of the logarithm. *Fibonacci Quart.* **32** (1994), 397–401.