Higher Secondary

Problem Corner

Kohli's Number 2997

SHAILESH SHIRALI

Consider the following function f defined from the set of positive integers $\mathbb{N}$ into itself:

$$f(n) = 111 \times \text{the sum of the digits of } n. \tag{1}$$

(Note that we work throughout in base 10.)

For example,

$$f(23) = 111 \times 5 = 555, \quad f(2345) = 111 \times 14 = 1554.$$

In the article [1], Uttkarsh Kohli describes a curious property of this function when it is iterated. Namely, if we start with any positive integer n and compute the sequence

$$n, \quad f(n), \quad f(f(n)), \quad f(f(f(n))), \quad \ldots, \tag{2}$$

then after just a few steps we will reach the number 2997. Moreover, once we reach that number (2997), we stay there.

A comment is needed here regarding the notation. The expression $f(f(f(n)))$ looks quite awkward, and the succeeding terms, $f(f(f(f(n))))$, $f(f(f(f(f(n)))))$, ... look more awkward still, with more and more closing brackets that start to resemble the layers of an onion. Some mathematicians prefer to write $f \circ f(n)$ in place of $f(f(n))$, $f \circ f \circ f(n)$ in place of $f(f(f(n)))$, $f \circ f \circ f \circ f(n)$ in place of $f(f(f(f(n))))$, and so on, with '$\circ$' denoting the function composition symbol. This is certainly far more pleasant to the eye!

We give a proof of Kohli's assertion here. To start with, we claim that:

$$\textit{If } n < 10^k, \textit{ then } f(n) \le 999k. \tag{3}$$

To see why this is true, note that among all numbers less than 10^k, the number with the largest sum of digits is $10^k - 1$, which is made up entirely of 9's. The sum of the digits of this number is $9k$. Hence the claim.

Keywords: Iteration, base 10, function, logarithmic function, fixed point, Kohli's number

Since $999 < 1000$, the above claim implies the following.

$$\text{If } n < 10^k, \text{ then } f(n) < 1000\, k. \tag{4}$$

This may be stated in another way as follows.

$$\text{For all positive integers } n, \quad f(n) < 1000 \log_{10} n. \tag{5}$$

Next, we would like to find a number M such that

$$\text{If } n > M, \text{ then } n > 1000 \log_{10} n. \tag{6}$$

(For, if we find such a number, then we have: if $n > M$, then $1000 \log_{10} n < n$ and also $f(n) < 1000 \log_{10} n$, which means that if $n > M$, then $f(n) < n$.) To find M, we must study the behaviour of the following function g (defined for $x > 1$) as x grows indefinitely large:

$$g(x) = \frac{x}{\log_{10} x}. \tag{7}$$

It is easy to verify via differentiation that $g(x)$ decreases for $1 < x < e$, takes its minimum value at $x = e$, and then steadily rises for $x > e$. (The 'steady rise' should not come as a surprise, considering the behaviour of the logarithmic function, which cuts even extremely large numbers down to manageable size.) The following table of values illustrates this assertion.

x	3	10	10^2	10^3	10^4	10^5
$x/\log_{10} x$	6.3	10	50	333.3	2500	20000

Computations reveal that $g(x)$ crosses the value 1000 roughly around $x = 3555$. As this number is less than 4000, we can safely take $M = 4000$ and thus state the following:

$$\text{If } n > 4000, \text{ then } 1000 \log_{10} n < n. \tag{8}$$

Next, observe that the number below 4000 with the largest f-value is 3999, whose f-value is $30 \times 111 = 3330$, and note that this number itself is below 4000. Combining this observation with (8), we obtain the following two important results:

$$\left.\begin{array}{ll} \bullet & \text{If } n > 4000, \text{ then } f(n) < n. \\ \bullet & \text{If } n \leq 4000, \text{ then } f(n) \leq 4000. \end{array}\right\} \tag{9}$$

Why are these two results important? They imply that even if we start with extremely large values of n, the sequence of iterates

$$n, \quad f(n), \quad f(f(n)), \quad f(f(f(n))), \quad \ldots, \tag{10}$$

is *strictly decreasing* till we reach a number below 4000. (The first result in (9) guarantees this.) Once we do reach a number below 4000, the sequence of iterates is no longer strictly decreasing or strictly increasing, but the numbers stay below 4000. (This is guaranteed by the second result in (9).)

This means that if we wish to study the behaviour of iterates of the function f, it suffices to restrict our attention to the set S_0 of integers between 1 and 4000. Results (9) imply the following important result:

$$\text{If } n \in S_0, \text{ then } f(n) \in S_0. \tag{11}$$

Combining the assertions in (10) and (11), we have the following claim:

For any positive integer n, however large, repeated applications of f will ultimately yield numbers in S_0, and once we reach S_0, we never leave it.

Next, note that the definition of f implies that for any n, $f(n)$ is a multiple of 111 (and therefore also a multiple of 3). This means that by applying f to all the numbers in S_0, the resulting set will be a subset of the set of multiples of 111 within S_0, i.e., a subset of the following set:

$$S_1 = \{111,\ 222,\ 333,\ 444,\ \ldots,\ 3663,\ 3774,\ 3885,\ 3996\}\,. \tag{12}$$

Set S_1 has 36 elements, but we have listed only the first 4 and the last 4 elements. The following claim should now be clear:

> For any positive integer n, however large, repeated applications of f will ultimately yield numbers in S_1, and once we reach S_1, we never leave it.

Now consider the second iterate $f(f(n))$. As there is again a multiplication by the factor 111, it follows that:

$$\textit{For any } n,\ f(f(n)) \textit{ is a multiple of } 9. \tag{13}$$

This implies that by applying f to all the numbers in S_1, the resulting set will be a subset of the set of multiples of 333 within S_1, i.e., a subset of the following set:

$$S_2 = \{333,\ 666,\ 999,\ 1332,\ 1665,\ 1998,\ 2331,\ 2664,\ 2997,\ 3330,\ 3663,\ 3996\}\,. \tag{14}$$

Consequently, we can now claim the following:

> For any positive integer n, however large, repeated applications of f will ultimately yield numbers in S_2, and once we reach S_2, we never leave it.

The progression should now be clear. The f-values of the numbers in S_2 form the following set:

$$S_3 = \{999,\ 1998,\ 2997\}\,. \tag{15}$$

We can now claim the following:

> For any positive integer n, however large, repeated applications of f will ultimately yield numbers in S_3, and once we reach S_3, we never leave it.

A quick check shows that the f-values of the numbers in S_3 are all the same, because all three numbers have the same sum of digits (namely, 27). This common f-value is 2997. We can therefore claim the following:

> For any positive integer n, however large, repeated applications of f will ultimately yield the number 2997. Once we reach this number, no further changes take place.

The claim that "no further changes take place" is true because $f(2997) = 2997$. This is sometimes expressed by saying that 2997 is a *fixed point* of the function f. (A 'fixed point' of a function h is any number x such that $h(x) = x$.)

The claim made at the start of the article thus stands proved. Kohli's number is a genuine constant!

References

1. Uttkarsh Kohli, "The Mystical Number 2997 (Kohli's Number)" from https://azimpremjiuniversity.edu.in/SitePages/resources-ara-issue-no-8-november-2020-mystical-number.aspx

SHAILESH SHIRALI is Director of Sahyadri School (KFI), Pune, and Head of the Community Mathematics Centre in Rishi Valley School (AP). He has been closely involved with the Math Olympiad movement in India. He is the author of many mathematics books for high school students, and serves as Chief Editor for *At Right Angles*. He may be contacted at shailesh.shirali@gmail.com.