EDITED BY RICHARD GUY

In this department the MONTHLY *presents easily stated unsolved problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy. Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada T2N 1N4.*

SOME SUSPICIOUSLY SIMPLE SEQUENCES

RICHARD K. GUY

Three problems of varying ages that have been drawn to our attention lately.

Donald E. G. Malm, Department of Mathematical Sciences, Oakland University, Rochester, MI 48063 asks about

Hofstadter's meta-Fibonacci sequence

Doug Hofstadter [**5**; see also **1**, **2**] has proposed the sequence defined by

$$Q(1) = Q(2) = 1, \quad Q(n) = Q(n - Q(n-1)) + Q(n - Q(n-2)), \quad n > 2.$$

If you calculate a few values:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$Q(n)$	1	1	2	3	3	4	5	5	6	6	6	8	8	8	10	9	10	11	11	12
$Q(n+20)$	12	12	12	16	14	14	16	16	16	16	20	17	17	20	21	19	20	22	21	22
$Q(n+40)$	23	23	24	24	24	24	24	32	24	25	30	28	26	30	30	28	32	30	32	32
$Q(n+60)$	32	32	40	33	31	38	35	33	39	40	37	38	40	39	40	39	42	40	41	43
$Q(n+80)$	44	43	43	46	44	45	47	47	46	48	48	48	48	48	48	64	41	52	54	56

you are first struck by the monotonicity, though this breaks down from $Q(15)$ to $Q(16)$. Then from $Q(24)$ to $Q(25)$ it drops by 2; from $Q(31)$ to $Q(32)$ by 3; from $Q(48)$ to $Q(49)$ by 8; from $Q(96)$ to $Q(97)$ by 23 and from $Q(192)$ to $Q(193)$ by 56.

Do these drops become arbitrarily large?

In the main, the values of $Q(n)$ are a little larger than $\frac{1}{2}n$. New maxima include

$$Q(3 \times 2^k) = 2^{k+1}, \quad 0 \leqslant k \leqslant 8.$$

Does this continue? No! $Q(3 \times 2^9) = 808$, a far cry from 2^{10}.

You may also notice strings of constant values,

$$Q(3 \times 2^k - j) = 3 \times 2^{k-1} \quad (1 \leqslant j \leqslant k+1).$$

This is true for $1 \leqslant k \leqslant 5$, but for $k = 6$, $j = 7$, we find that $Q(185) = 94$, not 96, and the pattern soon disappears without trace. The sequence $Q(2^k + 1)$ is intriguing. In the following table:

k	0	1	2	3	4	5	6	7	8
$2^k + 1$	2	3	5	9	17	33	65	129	257
$Q(2^k + 1)$	1	2	3	6	10	17	31	57	106
e_k		0	1	0	2	3	3	5	8

the last line is calculated by $e_k = 2Q(2^{k-1} + 1) - Q(2^k + 1)$. The appearance of Fibonacci numbers is remarkable, though the reversal of 0 and 1 and the repetition of 3 do not inspire confidence. However, a similar table for $Q(3 \cdot 2^k + 1)$:

k	0	1	2	3	4	5	6	7	8
$3 \cdot 2^k + 1$	4	7	13	25	49	97	193	385	769
$Q(3 \cdot 2^k + 1)$	3	5	8	14	24	41	72	129	235
f_k		1	2	2	4	7	10	15	23
$\Delta f_k = e_k$?			1	0	2	3	3	5	8

in which $f_k = 2Q(3 \cdot 2^{k-1} + 1) - Q(3 \cdot 2^k + 1)$, yields, for the first differences of f_k, exactly the same sequence, e_k!! But if you continue either table, the pattern (such as it is) is rudely shattered in each case, and in quite different ways.

Hofstadter's original question was:

Is $Q(n)$ defined for all positive integers n?
I.e., does $Q(n)/n$ never exceed 1?

Malm has calculated the first 200,000 members of the sequence. He found only 27 + 11 values of n for which $Q(n)/n$ lies outside the interval [0.4, 0.6]. $Q(n)/n > 0.6$ for

$n =$	1	3	4	6	7	8	9	12	13	15	18	24	31	48	63
$Q(n) =$	1	2	3	4	5	5	6	8	8	10	11	16	20	32	40

$n =$	96	124	192	202	384	394	768	793	860	997	1545	1569
$Q(n) =$	64	78	128	124	256	242	512	492	522	605	928	963

and $Q(n)/n < 0.4$ for

$n =$	193	385	395	433	769	794	801	846	1578	1673	1676
$Q(n) =$	72	135	149	164	278	284	310	335	626	647	662

It is remarkable that $Q(n)/n$ does not stray outside that interval between $n = 1676$ and $n = 200{,}000$. Moreover, in the range $10^5 < n \leqslant 2 \times 10^5$, Malm found only 77 values of n for which $Q(n)/n$ lies outside [0.45, 0.55], and they all occur in the range $100000 < n \leqslant 137958$. Malm later calculated the first million members and found more values of n for which $Q(n)/n$ is outside [0.45, 0.55]: 71 in the interval [202122, 278513], 44 in [394774, 556495] and 27 in [826568, 1000000].

Is there a sense in which $Q(n)$ is asymptotic to $\frac{1}{2}n$?
Do lim sup and lim inf of $Q(n)/n$ exist?

John Isbell, Department of Mathematics, SUNY at Buffalo, Buffalo, NY, 14214 has invented a game [**4**] and wants to know

Can Jack always beat the Giant at Beanstalk?

Beanstalk starts with the choice, by the Giant, of an odd integer, $n_0 > 1$. Thereafter Jack and the Giant play alternately according to the rule

$$n_{i+1} = \begin{cases} n_i/2 & \text{if } n_i \text{ is even,} \\ 3n_i \pm 1 & \text{if } n_i \text{ is odd.} \end{cases}$$

If n_i is even, there is only one option. If n_i is odd, there are just two. The winner is the player who arrives at 1.

This looks suspiciously like the notorious Collatz problem [**3**], [**6**]; does the fact that there is a choice make it any easier to analyse?

Suppose the Giant chooses 3. The play should go as follows, with Jack's moves in parentheses:

3 (10) 5 (16) 8 (4) 2 (1)

and Jack wins. The only alternatives are 3 (8) 4 (2) 1 and the Giant wins, or 3 (10) 5 (14) 7 leading to a much more complicated situation that we'll examine later.

It's clear that Jack wins if he plays to a power of 4, and loses if he plays to an odd power of 2. Jack need never lose! One of his options is always singly even, of shape $4k+2$ $(k>0)$ after which the Giant has to play the odd number $2k+1$ and Jack *keeps control*. However, if he does this, each pair of moves increases n_i to $n_{i+2} \approx 3n_i/2$ and the game will go on forever. His alternative will be a move to $2^{2e}d$ or $2^{2e+1}d$, where d is odd and $e>0$. If $d=1$, we know that the former is a good move, and the latter is a bad one. If $d>1$, the former loses control and the latter retains it. So the only real choice is between $2d$ (**hiding strategy**) and $2^{2e+1}d$ $(e>0$, **greedy strategy**) on those occasions that it occurs. The hiding strategy can't win, but neither is the greedy strategy always successful. In the following game, starting with 7, Jack has four places, underlined, where he can adopt the greedy strategy: the options at other places lose control.

7 (22) 11 (34) 17 (50) 25 (74) 37 (110) 55 (166) 83 (<u>248</u>) 124 (62) 31 (94) 47 (142) 71 (214) 107 (322) 161 (482) 241 (722) 361 (1082) 541 (<u>1624</u>) 812 (406) 203 (<u>608</u>) 304 (152) 76 (38) 19 (<u>56</u>) 28 (14) 7 ...

and we have completed a cycle of 44 moves. We leave the reader to confirm that Jack can win in 65 moves, by a judicious mixture of the hiding and greedy strategies.

But is this true for all odd $n_0 > 1$?

And if it is, how long does Jack take to win? It is true for $n_0 < 1000$. The longest game (assuming that Jack always adopts the best strategy) that we have found has 263 moves and starts with 747. However, since a complete analysis involves a tree search, it is easy to make mistakes.

Isbell, Selfridge, Erdős and others have reminded me from time to time of an old problem considered by Mahler and Popken [**7**]:

What is the least number of ones needed to represent *n*, using only + and × (and parentheses)?

If we denote this number by $f(n)$, then, since the last operation is either multiplication or addition, we can define $f(n)$ recursively by the formula

$$f(n) = \min_{\substack{d \mid n \\ 2 \leqslant d \leqslant \sqrt{n} \\ 1 \leqslant e \leqslant n/2}} \{ f(d) + f(n/d), f(e) + f(n-e) \}.$$

However, this is somewhat less than useful. When the minimum is realized by the second member, is e ever greater than 1? Isbell and Myerson suggest that $n = (3^{70}+3^{30}+3^{10}+1)^2+6$ might be such an example, with ¿$f(n) = 431$?

Is it always true that, for a prime p, $f(p) = 1 + f(p-1)$? And $f(2p) = \min\{2+f(p), 1+f(2p-1)\}$? The minimum is usually attained by the former, but also occasionally by the latter, e.g., for $p = 23$, 41 and 59. Is it always true that

$$3 + f(p) \leqslant 1 + f(3p-1)?$$

Here is a table of the first 100 values of $f(n)$:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$f(n)$	1	2	3	4	5	5	6	6	6	7	8	7	8	8	8	8	9	8	9	9
$f(n+20)$	9	10	11	9	10	10	9	10	11	10	11	10	11	11	11	10	11	11	11	11
$f(n+40)$	12	11	12	12	11	12	13	11	12	12	12	12	13	11	12	12	12	13	14	12
$f(n+60)$	13	13	12	12	13	13	14	13	14	13	14	12	13	13	13	13	14	13	14	13
$f(n+80)$	12	13	14	13	14	14	14	14	15	13	14	14	14	15	14	13	14	14	14	14

Selfridge uses induction to show that 3^k is the largest solution of $f(x) = 3k$, and that $3^k \pm 3^{k-1}$ are the largest solutions of $f(x) = 3k \pm 1$. If we write lb and lt for **binary** and **ternary log**, i.e., logs to base 2 and 3 respectively, we have

$$f(n) \geqslant 3 \operatorname{lt} n.$$

In fact $f(2^a 3^b) = 2a + 3b$ for $a = 0, 1$ and 2.

For what larger a is this true?

It is true for $2^a 3^b \leqslant 216$. But $2^a 3^b$ is not the largest n for which $f(n) = 2a + 3b$. For example $f(96) = 13$, but so is $f(108)$.

Are larger n always of the form $2^{a-3c} 3^{b+2c}$?

Selfridge asks:

Is there an a for which $f(2^a) < 2a$?

We could find such a 2^a if its $\lceil a \operatorname{lt} 2 \rceil$ ternary digits had an average size of less than $2/\operatorname{lt} 2 - 3 \approx 0.1699250014$, but the expectation of this happening must be very small.

If we write n in binary, $n = \sum_{i=0}^{k} a_i 2^i$, where $0 \leqslant a_i \leqslant 1 = a_k$, then

$$n = a_0 + (1+1)(a_1 + (1+1)(a_2 + (\cdots (1+1)(a_{k-1} + 1 + 1) \cdots))),$$

where there are k parentheses $(1+1)$ and d ones, where $d(\leqslant k)$ is one less than the number of digits 1 in the binary expansion, $d = a_0 + a_1 + \cdots + a_{k-1}$, and

$$f(n) \leqslant 2k + d \leqslant 3k \leqslant 3 \operatorname{lb} n < 4.755 \operatorname{lt} n,$$

since $n \geqslant 2^k$.

Isbell notes that, for a set of density 1, namely those n which are nearly normal in the scale of 6,

$$f(n) < \left(\frac{19}{3} + \varepsilon\right) \log_6 n < 3.8833 \operatorname{lt} n.$$

Here 19/3 is the average number of ones needed to convert an expression for k into an expression for $6k + l$, $0 \leqslant l \leqslant 5$. Similarly, working in duodecimal, there's a set of density 1 for which

$$f(n) < \left(\frac{26}{3} + \varepsilon\right) \log_{12} n < 3.8317 \operatorname{lt} n.$$

Some other bases $2^a 3^b$ will yield further small improvements, e.g., base 24 gives $f(n) < 3.817 \operatorname{lt} n$ for almost all n.

There are conflicting conjectures:

¿ For large n, $(3 + \varepsilon) \operatorname{lt} n$ ones suffice ?

¿ There are infinitely many n, perhaps a set of positive density for which $(3 + c) \operatorname{lt} n$ ones are needed, for some $c > 0$?

References

1. P. Erdős and R. L. Graham, Old and New Problems and Results in Combinatorial Number Theory, L'Enseignement Math., Geneva, 1980, p. 83.

2. Richard K. Guy, Unsolved Problems in Number Theory, Springer-Verlag, New York, 1981, Problems E31 and F26.

3. Richard K. Guy, Don't try to solve these problems! this MONTHLY, 90 (1983) 35–41.

4. Richard K. Guy, John Isbell's game of Beanstalk and John Conway's game of Beans-Don't-Talk, Math. Mag.

5. Douglas Hofstadter, Gödel, Escher, Bach: an Eternal Golden Braid, Basic Books, New York, 1979, p. 137.

6. Jeffery C. Lagarias, The $3x + 1$ problem and its generalizations, this MONTHLY, 92 (1985) 3–23.

7. K. Mahler and J. Popken, On a maximum problem in arithmetic (Dutch), Nieuw Arch. Wiskunde, (3) 1 (1953) 1–15; MR 14, 852e.

NOTES

EDITED BY SABRA S. ANDERSON, SHELDON AXLER, AND J. ARTHUR SEEBACH, JR.

For instructions about submitting Notes for publication in this department see the inside front cover.

THE TIETZE EXTENSION THEOREM AND THE OPEN MAPPING THEOREM

SANDY GRABINER
Department of Mathematics, Pomona College, Claremont, CA 91711

The Open Mapping Theorem says that a bounded linear transformation from one Banach space onto another must be an open mapping, while the Tietze Extension Theorem says that a bounded continuous function can always be extended from a closed subset of a normal topological space to the entire space. The two theorems sound as though they are unrelated, but the central arguments in the standard proofs of these two theorems are in fact quite similar. In each proof we can first crudely approximate a continuous function or an element of an abstract Banach space. We then inductively construct better approximations, by, at each stage, approximating the error from the previous stage. This process produces an infinite series that converges exactly to the original element. For the Open Mapping Theorem, one proves that if T is a bounded linear operator and B is an open ball about the origin for which the closure $\overline{T(B)}$ contains a ball about the origin, then $T(B)$ already contains such a ball. In fact, as many people have noticed, the usual proof does not use the fact that $T(B)$ is actually dense in some ball, but only that it gets "close enough" to elements in this ball. The following lemma gives a precise statement of what one can prove. Furthermore, this precise statement is strong enough to give the Tietze Extension Theorem.

APPROXIMATION LEMMA. *Suppose that T is a bounded linear operator from a (real or complex) Banach space E to a Banach space F, and suppose also that m and r are positive numbers with $r < 1$. If for each y in F, there is an x_0 in E with $\|x_0\| \leqslant m\|y\|$ and $\|y - Tx_0\| \leqslant r\|y\|$, then there is also an x in E with $Tx = y$ and $\|x\| \leqslant m\|y\|/(1 - r)$.*

Proof. For simplicity take $\|y\| = 1$. Applying the hypothesis to $y - Tx_0$ in place of y, we find an x_1 with $\|x_1\| \leqslant rm$ and $\|y - T(x_0 + x_1)\| \leqslant r^2$. Proceeding inductively, we obtain a sequence $\{x_n\}_0^\infty$ with

$$\|x_n\| \leqslant r^n m, \tag{1}$$

and

$$\|y - T(x_0 + x_1 + \cdots x_n)\| \leqslant r^{n+1}. \tag{2}$$

Formula (1) implies that the series $\Sigma_0^\infty x_n$ converges absolutely to a vector x of norm less than or equal to $m/(1 - r)$. Letting n go to infinity in formula (2) then shows that $y = Tx$.