

**Solutions to the 81st William Lowell Putnam Mathematical Competition**  
**Saturday, February 20, 2021**

Manjul Bhargava, Kiran Kedlaya, and Lenny Ng

A1 The values of  $N$  that satisfy (ii) and (iii) are precisely the numbers of the form  $N = (10^a - 10^b)/9$  for  $0 \leq b < a \leq 2020$ ; this expression represents the integer with  $a$  digits beginning with a string of 1's and ending with  $b$  0's. A value  $N$  of this form is divisible by  $2020 = 2^2 \cdot 5 \cdot 101$  if and only if  $10^b(10^{a-b} - 1)$  is divisible by each of  $3^2$ ,  $2^2 \cdot 5$ , and  $101$ . Divisibility by  $3^2$  is a trivial condition since  $10 \equiv 1 \pmod{9}$ . Since  $10^{a-b} - 1$  is odd, divisibility by  $2^2 \cdot 5$  occurs if and only if  $b \geq 2$ . Finally, since  $10^2 \equiv -1 \pmod{101}$ , we see that  $10^{a-b}$  is congruent to  $10$ ,  $-1$ ,  $-10$ , or  $1 \pmod{101}$  depending on whether  $a - b$  is congruent to  $1$ ,  $2$ ,  $3$ , or  $0 \pmod{4}$ ; thus  $10^{a-b} - 1$  is divisible by  $101$  if and only if  $a - b$  is divisible by  $4$ .

It follows that we need to count the number of  $(a, b)$  with  $2 \leq b < a \leq 2020$  with  $4 \mid a - b$ . For given  $b$ , there are  $\lfloor \frac{2020-b}{4} \rfloor$  possible values of  $a$ . Thus the answer is

$$504 + 504 + 504 + 503 + 503 + 503 + 503 + \cdots + 1 + 1 + 1 + 1 \\ = 4(504 + 503 + \cdots + 1) - 504 = 504 \cdot 1009 = 508536.$$

A2 The answer is  $4^k$ .

**First solution.** Let  $S_k$  denote the given sum. Then, with the convention that  $\binom{n}{-1} = 0$  for any  $n \geq 0$ , we have for  $k \geq 1$ ,

$$S_k = \sum_{j=0}^k 2^{k-j} \left[ \binom{k-1+j}{j} + \binom{k-1+j}{j-1} \right] \\ = 2 \sum_{j=0}^{k-1} 2^{k-1-j} \binom{k-1+j}{j} + \binom{2k-1}{k} + \sum_{j=1}^k 2^{k-j} \binom{k-1+j}{j-1} \\ = 2S_{k-1} + \binom{2k-1}{k} + \sum_{j=0}^{k-1} 2^{k-j-1} \binom{k+j}{j} \\ = 2S_{k-1} + S_k/2$$

and so  $S_k = 4S_{k-1}$ . Since  $S_0 = 1$ , it follows that  $S_k = 4^k$  for all  $k$ .

**Second solution.** Consider a sequence of fair coin flips  $a_1, a_2, \dots$  and define the random variable  $X$  to be the index of the  $(k+1)$ -st occurrence of heads. Then

$$P[X = n] = \binom{n-1}{k} 2^{-n};$$

writing  $n = k + j + 1$ , we may thus rewrite the given sum as

$$2^{2k+1} P[X \leq 2k + 1].$$

It now suffices to observe that  $P[X \leq 2k + 1] = \frac{1}{2}$ : we have  $X \leq 2k + 1$  if and only if there are at least  $k + 1$

heads among the first  $2k + 1$  flips, and there are exactly as many outcomes with at most  $k$  heads.

**Third solution.** (by Pankaj Sinha) The sum in question in the coefficient of  $x^k$  in the formal power series

$$\sum_{j=0}^k 2^{k-j} (1+x)^{k+j} = 2^k (1+x)^k \sum_{j=0}^k 2^{-j} (1+x)^j \\ = 2^k (1+x)^k \frac{1 - (1+x)^{k+1}/2^{k+1}}{1 - (1+x)/2} \\ = \frac{2^{k+1}(1+x)^k - (1+x)^{2k+1}}{1-x} \\ = (2^{k+1}(1+x)^k - (1+x)^{2k+1})(1+x+\cdots).$$

This evidently equals

$$2^{k+1} \sum_{j=0}^k \binom{k}{j} - \sum_{j=0}^k \binom{2k+1}{j} = 2^{k+1}(2^k) - \frac{1}{2} 2^{2k+1} \\ = 2^{2k+1} - 2^{2k} = 2^{2k} = 4^k.$$

**Remark.** This sum belongs to a general class that can be evaluated mechanically using the *WZ method*. See for example the book  $A = B$  by Petvoksek–Wilf–Zeilberger.

A3 The series diverges. First note that since  $\sin(x) < x$  for all  $x > 0$ , the sequence  $\{a_n\}$  is positive and decreasing, with  $a_1 = 1$ . Next, we observe that for  $x \in [0, 1]$ ,  $\sin(x) \geq x - x^3/6$ : this follows from Taylor's theorem with remainder, since  $\sin(x) = x - x^3/6 + (\sin c)x^4/24$  for some  $c$  between  $0$  and  $x$ .

We now claim that  $a_n \geq 1/\sqrt{n}$  for all  $n \geq 1$ ; it follows that  $\sum a_n^2$  diverges since  $\sum 1/n$  diverges. To prove the claim, we induct on  $n$ , with  $n = 1$  being trivial. Suppose that  $a_n \geq 1/\sqrt{n}$ . To prove  $\sin(a_n) \geq 1/\sqrt{n+1}$ , note that since  $\sin(a_n) \geq \sin(1/\sqrt{n})$ , it suffices to prove that  $x - x^3/6 \geq (n+1)^{-1/2}$  where  $x = 1/\sqrt{n}$ . Squaring both sides and clearing denominators, we find that this is equivalent to  $(n+1)(6n-1)^2 \geq 36n^3$ , or  $24n^2 - 11n + 1 \geq 0$ . But this last inequality is true since  $24n^2 - 11n + 1 = (3n-1)(8n-1)$ , and the induction is complete.

A4 The answer is  $1/e$ . We first establish a recurrence for  $w(N)$ . Number the squares  $1$  to  $N+2$  from left to right. There are  $2(N-1)$  equally likely events leading to the first new square being colored black: either we choose one of squares  $3, \dots, N+1$  and color the square to its left, or we choose one of squares  $2, \dots, N$  and color the square to its right. Thus the probability of square  $i$  being

the first new square colored black is  $\frac{1}{2(N-1)}$  if  $i = 2$  or  $i = N + 1$  and  $\frac{1}{N-1}$  if  $3 \leq i \leq N$ . Once we have changed the first square  $i$  from white to black, then the strip divides into two separate systems, squares 1 through  $i$  and squares  $i$  through  $N + 2$ , each with first and last square black and the rest white, and we can view the remaining process as continuing independently for each system. Thus if square  $i$  is the first square to change color, the expected number of white squares at the end of the process is  $w(i-2) + w(N+1-i)$ . It follows that

$$\begin{aligned} w(N) &= \frac{1}{2(N-1)}(w(0) + w(N-1)) + \\ &\quad \frac{1}{N-1} \left( \sum_{i=3}^N (w(i-2) + w(N+1-i)) \right) \\ &\quad + \frac{1}{2(N-1)}(w(N-1) + w(0)) \end{aligned}$$

and so

$$(N-1)w(N) = 2(w(1) + \dots + w(N-2)) + w(N-1).$$

If we replace  $N$  by  $N-1$  in this equation and subtract from the original equation, then we obtain the recurrence

$$w(N) = w(N-1) + \frac{w(N-2)}{N-1}.$$

We now claim that  $w(N) = (N+1) \sum_{k=0}^{N+1} \frac{(-1)^k}{k!}$  for  $N \geq 0$ . To prove this, we induct on  $N$ . The formula holds for  $N = 0$  and  $N = 1$  by inspection:  $w(0) = 0$  and  $w(1) = 1$ . Now suppose that  $N \geq 2$  and  $w(N-1) = N \sum_{k=0}^N \frac{(-1)^k}{k!}$ ,  $w(N-2) = (N-1) \sum_{k=0}^{N-1} \frac{(-1)^k}{k!}$ . Then

$$\begin{aligned} w(N) &= w(N-1) + \frac{w(N-2)}{N-1} \\ &= N \sum_{k=0}^N \frac{(-1)^k}{k!} + \sum_{k=0}^{N-1} \frac{(-1)^k}{k!} \\ &= (N+1) \sum_{k=0}^{N-1} \frac{(-1)^k}{k!} + \frac{N(-1)^N}{N!} \\ &= (N+1) \sum_{k=0}^{N+1} \frac{(-1)^k}{k!} \end{aligned}$$

and the induction is complete.

Finally, we compute that

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{w(N)}{N} &= \lim_{N \rightarrow \infty} \frac{w(N)}{N+1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = \frac{1}{e}. \end{aligned}$$

**Remark.** AoPS user pieater314159 suggests the following alternate description of  $w(N)$ . Consider the

numbers  $\{1, \dots, N+1\}$  all originally colored white. Choose a permutation  $\pi \in S_{N+1}$  uniformly at random. For  $i = 1, \dots, N+1$  in succession, color  $\pi(i)$  black in case  $\pi(i+1)$  is currently white (regarding  $i+1$  modulo  $N+1$ ). After this, the expected number of white squares remaining is  $w(N)$ .

**Remark.** Andrew Bernoff reports that this problem was inspired by a similar question of Jordan Ellenberg (disseminated via Twitter), which in turn was inspired by the final question of the 2017 MATHCOUNTS competition. See <http://bit-player.org/2017/counting-your-chickens-before-theyre-pecked> for more discussion.

A5 The answer is  $n = F_{4040} - 1$ . In both solutions, we use freely the identity

$$F_1 + F_2 + \dots + F_{m-2} = F_m - 1 \quad (1)$$

which follows by a straightforward induction on  $m$ . We also use the directly computed values

$$a_1 = a_2 = 2, a_3 = a_4 = 3. \quad (2)$$

**First solution.** (by George Gilbert)

We extend the definition of  $a_n$  by setting  $a_0 = 1$ .

**Lemma 1.** For  $m > 0$  and  $F_m \leq n < F_{m+1}$ ,

$$a_n = a_{n-F_m} + a_{F_{m+1}-n-1}. \quad (3)$$

*Proof.* Consider a set  $S$  for which  $\sum_{k \in S} F_k = n$ . If  $m \in S$  then  $S \setminus \{m\}$  gives a representation of  $n - F_m$ , and this construction is reversible because  $n - F_m < F_{m-1} \leq F_m$ . If  $m \notin S$ , then  $\{1, \dots, m-1\} \setminus S$  gives a representation of  $F_{m+1} - n - 1$ , and this construction is also reversible. This implies the desired equality.  $\square$

**Lemma 2.** For  $m \geq 2$ ,

$$a_{F_m} = a_{F_{m+1}-1} = \left\lfloor \frac{m+2}{2} \right\rfloor.$$

*Proof.* By (2), this holds for  $m = 2, 3, 4$ . We now proceed by induction; for  $m \geq 5$ , given all preceding cases, we have by Lemma 1 that

$$a_{F_m} = a_0 + a_{F_{m-1}-1} = 1 + \left\lfloor \frac{m}{2} \right\rfloor = \left\lfloor \frac{m+2}{2} \right\rfloor$$

$$a_{F_{m+1}-1} = a_{F_{m-1}-1} + a_0 = a_{F_m}. \quad \square$$

Using Lemma 2, we see that  $a_n = 2020$  for  $n = F_{4040} - 1$ .

**Lemma 3.** For  $F_m \leq n < F_{m+1}$ ,  $a_n \geq a_{F_m}$ .

*Proof.* We again induct on  $m$ . By Lemma 2, we may assume that

$$1 \leq n - F_m \leq (F_{m+1} - 2) - F_m = F_{m-1} - 2. \quad (4)$$

By (2), we may also assume  $n \geq 6$ , so that  $m \geq 5$ . We apply Lemma 1, keeping in mind that

$$(n - F_m) + (F_{m+1} - n - 1) = F_{m-1} - 1.$$

If  $\max\{n - F_m, F_{m+1} - n - 1\} \geq F_{m-2}$ , then one of the summands in (5) is at least  $a_{F_{m-2}}$  (by the induction hypothesis) and the other is at least 2 (by (4) and the induction hypothesis), so

$$a_n \geq a_{F_{m-2}} + 2 = \left\lfloor \frac{m+4}{2} \right\rfloor.$$

Otherwise,  $\min\{n - F_m, F_{m+1} - n - 1\} \geq F_{m-3}$  and so by the induction hypothesis again,

$$a_n \geq 2a_{F_{m-3}} = 2 \left\lfloor \frac{m-1}{2} \right\rfloor \geq 2 \frac{m-2}{2} \geq \left\lfloor \frac{m+2}{2} \right\rfloor. \quad \square$$

Combining Lemma 2 and Lemma 3, we deduce that for  $n > F_{4040} - 1$ , we have  $a_n \geq a_{F_{4040}} = 2021$ . This completes the proof.

**Second solution.** We again start with a computation of some special values of  $a_n$ .

**Lemma 1.** For all  $m \geq 1$ ,

$$a_{F_m-1} = \left\lfloor \frac{m+1}{2} \right\rfloor$$

*Proof.* We proceed by induction on  $m$ . The result holds for  $m = 1$  and  $m = 2$  by (2). For  $m > 2$ , among the sets  $S$  counted by  $a_{F_m-1}$ , by (1) the only one not containing  $m-1$  is  $S = \{1, 2, \dots, m-2\}$ , and there are  $a_{F_m-F_{m-1}-1}$  others. Therefore,

$$\begin{aligned} a_{F_m-1} &= a_{F_m-F_{m-1}-1} + 1 \\ &= a_{F_{m-2}-1} + 1 = \left\lfloor \frac{m-1}{2} \right\rfloor + 1 = \left\lfloor \frac{m+1}{2} \right\rfloor. \quad \square \end{aligned}$$

Given an arbitrary positive integer  $n$ , define the set  $S_0$  as follows: start with the largest  $k_1$  for which  $F_{k_1} \leq n$ , then add the largest  $k_2$  for which  $F_{k_1} + F_{k_2} \leq n$ , and so on, stopping once  $\sum_{k \in S_0} F_k = n$ . Then form the bitstring

$$s_n = \cdots e_1 e_0, \quad e_k = \begin{cases} 1 & k \in S_0 \\ 0 & k \notin S_0; \end{cases}$$

note that no two 1s in this string are consecutive. We can thus divide  $s_n$  into segments

$$t_{k_1, \ell_1} \cdots t_{k_r, \ell_r} \quad (k_i, \ell_i \geq 1)$$

where the bitstring  $t_{k, \ell}$  is given by

$$t_{k, \ell} = (10)^k (0)^\ell$$

(that is,  $k$  repetitions of 10 followed by  $\ell$  repetitions of 0). Note that  $\ell_r \geq 1$  because  $e_1 = e_0 = 0$ .

For  $a = 1, \dots, k$  and  $b = 0, \dots, \lfloor (\ell-1)/2 \rfloor$ , we can replace  $t_{k, \ell}$  with the string of the same length

$$(10)^{k-a} (0) (1)^{2a-1} (01)^b 10^{\ell-2b}$$

to obtain a new bitstring corresponding to a set  $S$  with  $\sum_{k \in S} F_k = n$ . Consequently,

$$a_n \geq \prod_{i=1}^r \left( 1 + k_i \left\lfloor \frac{\ell_i + 1}{2} \right\rfloor \right). \quad (5)$$

For integers  $k, \ell \geq 1$ , we have

$$1 + k \left\lfloor \frac{\ell + 1}{2} \right\rfloor \geq k + \left\lfloor \frac{\ell + 1}{2} \right\rfloor \geq 2.$$

Combining this with repeated use of the inequality

$$xy \geq x + y \quad (x, y \geq 2),$$

we deduce that

$$a_n \geq \sum_{i=1}^r \left( k_i + \left\lfloor \frac{\ell_i + 1}{2} \right\rfloor \right) \geq \left\lfloor \frac{1 + \sum_{i=1}^r (2k_i + \ell_i)}{2} \right\rfloor.$$

In particular, for any even  $m \geq 2$ , we have  $a_n > \frac{m}{2}$  for all  $n \geq F_m$ . Taking  $m = 4040$  yields the desired result.

**Remark.** It can be shown with a bit more work that the set  $S_0$  gives the unique representation of  $n$  as a sum of distinct Virahanka–Fibonacci numbers, no two consecutive; this is commonly called the *Zeckendorf representation* of  $n$ , but was first described by Lekkerkerker. Using this property, one can show that the lower bound in (5) is sharp.

A6 The smallest constant  $M$  is  $\pi/4$ .

We start from the expression

$$f_N(x) = \sum_{n=0}^N \frac{1}{2} \left( \frac{2}{2n+1} - \frac{1}{N+1} \right) \sin((2n+1)x). \quad (6)$$

Note that if  $\sin(x) > 0$ , then

$$\begin{aligned} \sum_{n=0}^N \sin((2n+1)x) &= \frac{1}{2i} \sum_{n=0}^N (e^{i(2n+1)x} - e^{-i(2n+1)x}) \\ &= \frac{1}{2i} \left( \frac{e^{i(2N+3)x} - e^{ix}}{e^{2ix} - 1} - \frac{e^{-i(2N+3)x} - e^{-ix}}{e^{-2ix} - 1} \right) \\ &= \frac{1}{2i} \left( \frac{e^{i(2N+2)x} - 1}{e^{ix} - e^{-ix}} - \frac{e^{-i(2N+2)x} - 1}{e^{-ix} - e^{ix}} \right) \\ &= \frac{1}{2i} \frac{e^{i(2N+2)x} + e^{-i(2N+2)x} - 2}{e^{ix} - e^{-ix}} \\ &= \frac{2 \cos((2N+2)x) - 2}{2i(2i \sin(x))} \\ &= \frac{1 - \cos((2N+2)x)}{2 \sin(x)} \geq 0. \end{aligned}$$

We use this to compare the expressions of  $f_N(x)$  and  $f_{N+1}(x)$  given by (6). For  $x \in (0, \pi)$  with  $\sin((2N+3)x) \geq 0$ , we may omit the summand  $n = N+1$  from  $f_{N+1}(x)$  to obtain

$$\begin{aligned} & f_{N+1}(x) - f_N(x) \\ & \geq \frac{1}{2} \left( \frac{1}{N+1} - \frac{1}{N+2} \right) \sum_{n=0}^N \sin((2n+1)x) \geq 0. \end{aligned}$$

For  $x \in (0, \pi)$  with  $\sin((2N+3)x) \leq 0$ , we may insert the summand  $n = N+1$  into  $f_{N+1}(x)$  to obtain

$$\begin{aligned} & f_{N+1}(x) - f_N(x) \\ & \geq \frac{1}{2} \left( \frac{1}{N+1} - \frac{1}{N+2} \right) \sum_{n=0}^{N+1} \sin((2n+1)x) \geq 0. \end{aligned}$$

In either case, we deduce that for  $x \in (0, \pi)$ , the sequence  $\{f_N(x)\}_N$  is nondecreasing.

Now rewrite (6) as

$$f_N(x) = \sum_{n=0}^N \frac{\sin((2n+1)x)}{2n+1} - \frac{1 - \cos((2N+2)x)}{4(N+1)\sin(x)} \quad (7)$$

and note that the last term tends to 0 as  $N \rightarrow \infty$ . Consequently,  $\lim_{N \rightarrow \infty} f_N(x)$  equals the sum of the series

$$\sum_{n=0}^{\infty} \frac{1}{2n+1} \sin((2n+1)x),$$

which is the Fourier series for the ‘‘square wave’’ function defined on  $(-\pi, \pi]$  by

$$x \mapsto \begin{cases} -\frac{\pi}{4} & x \in (-\pi, 0) \\ \frac{\pi}{4} & x \in (0, \pi) \\ 0 & x = 0, \pi \end{cases}$$

and extended periodically. Since this function is continuous on  $(0, \pi)$ , we deduce that the Fourier series converges to the value of the function; that is,

$$\lim_{N \rightarrow \infty} f_N(x) = \frac{\pi}{4} \quad (x \in (0, \pi)).$$

This is enough to deduce the desired result as follows. Since

$$f_N(x+2\pi) = f_N(x), \quad f_N(-x) = -f_N(x),$$

it suffices to check the bound  $f_N(x) \leq \pi$  for  $x \in (-\pi, \pi]$ . For  $x = 0, \pi$  we have  $f_N(x) = 0$  for all  $N$ . For  $x \in (-\pi, 0)$ , the previous arguments imply that

$$0 \geq f_0(x) \geq f_1(x) \geq \dots$$

For  $x \in (0, \pi)$ , the previous arguments imply that

$$0 \leq f_0(x) \leq f_1(x) \leq \dots \leq \frac{\pi}{4}$$

and the limit is equal to  $\pi/4$ . We conclude that  $f_N(x) \leq M$  holds for  $M = \pi/4$  but not for any smaller  $M$ , as desired.

**Remark.** It is also possible to replace the use of the convergence of the Fourier series with a more direct argument; it is sufficient to do this for  $x$  in a dense subset of  $(0, \pi)$ , such as the rational multiples of  $\pi$ .

Another alternative (described at <https://how-did-i-get-here.com/2020-putnam-a6/>) is to deduce from (7) and a second geometric series computation (omitted here) that

$$\begin{aligned} f'_N(x) &= \sum_{n=0}^N \cos((2n+1)x) - \frac{d}{dx} \left( \frac{1 - \cos((2N+2)x)}{4(N+1)\sin(x)} \right) \\ &= \frac{\sin((2N+2)x)}{2\sin(x)} \\ &\quad - \frac{(2N+2)\sin((2N+2)x) - \cos(x)(1 - \cos((N+2)x))}{4(N+1)\sin(x)^2} \\ &= \frac{\cos(x)(1 - \cos((N+2)x))}{4(N+1)\sin(x)^2}, \end{aligned}$$

which is nonnegative for  $x \in (0, \pi/2]$  and nonpositive for  $x \in [\pi/2, \pi)$ . This implies that  $f_N(x)$  always has a global maximum at  $x = \pi/2$ , so it suffices to check the convergence of the Fourier series for the square wave at that point. This reduces to the Madhava–Gregory–Newton series evaluation

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \arctan(1) = \frac{\pi}{4}.$$

B1 Note that

$$(1-x)(1-x^2)(1-x^4)\dots(1-x^{1024}) = \sum_{k=0}^{2047} (-1)^{d(k)} x^k$$

and

$$x^{2016}(1-x)(1-x^2)\dots(1-x^{16}) = \sum_{k=2016}^{2047} (-1)^{d(k)} x^k.$$

Applying  $x \frac{d}{dx}$  to both sides of each of these two equations three times, and then setting  $x = 1$ , shows that

$$\sum_{k=0}^{2047} (-1)^{d(k)} k^3 = \sum_{k=2016}^{2047} (-1)^{d(k)} k^3 = 0,$$

and therefore

$$\sum_{k=1}^{2015} (-1)^{d(k)} k^3 = 0.$$

Hence we may write

$$\begin{aligned}
S &= \sum_{k=2016}^{2020} (-1)^{d(k)} k^3 \\
&= \sum_{k=0}^4 (-1)^{d(k)} (k+2016)^3 \\
&\equiv (-4)^3 + (-1)(-3)^3 + (-1)(-2)^3 + (1)(-1)^3 \\
&= -64 + 27 + 8 - 1 \\
&\equiv -30 \equiv 1990 \pmod{2020}.
\end{aligned}$$

**Remark.** The function  $d(n)$  appears in the OEIS as sequence A000120.

B2 We refer to this two-player game, with  $n$  holes and  $k$  pegs, as the  $(n, k)$ -game. We will show that Alice has a winning strategy for the  $(n, k)$ -game if and only if at least one of  $n$  and  $k$  is odd; otherwise Bob has a winning strategy.

We reduce the first claim to the second as follows. If  $n$  and  $k$  are both odd, then Alice can move the  $k$ -th peg to the last hole; this renders the last hole, and the peg in it, totally out of play, thus reducing the  $(n, k)$ -game to the  $(n-1, k-1)$ -game, for which Alice now has a winning strategy by the second claim. Similarly, if  $n$  is odd but  $k$  is even, then Alice may move the first peg to the  $(k+1)$ -st hole, removing the first hole from play and reducing the  $(n, k)$ -game to the  $(n-1, k)$  game. Finally, if  $n$  is even but  $k$  is odd, then Alice can move the first peg to the last hole, taking the first and last holes, and the peg in the last hole, out of play, and reducing the  $(n, k)$ -game to the  $(n-2, k-1)$ -game.

We now assume  $n$  and  $k$  are both even and describe a winning strategy for the  $(n, k)$ -game for Bob. Subdivide the  $n$  holes into  $n/2$  disjoint pairs of adjacent holes. Call a configuration of  $k$  pegs *good* if for each pair of holes, both or neither is occupied by pegs, and note that the starting position is good. Bob can ensure that after each of his moves, he leaves Alice with a good configuration: presented with a good configuration, Alice must move a peg from a pair of occupied holes to a hole in an unoccupied pair; then Bob can move the other peg from the first pair to the remaining hole in the second pair, resulting in another good configuration. In particular, this ensures that Bob always has a move to make. Since the game must terminate, this is a winning strategy for Bob.

B3 Let  $f(\delta)$  denote the desired expected value of  $Z$  as a function of  $\delta$ . We prove that  $f(\delta) = 1 - \log(\delta)$ , where  $\log$  denotes natural logarithm.

For  $c \in [0, 1]$ , let  $g(\delta, c)$  denote the expected value of  $Z$  given that  $x_1 = c$ , and note that  $f(\delta) = \int_0^1 g(\delta, c) dc$ . Clearly  $g(\delta, c) = 1$  if  $c < \delta$ . On the other hand, if  $c \geq \delta$ , then  $g(\delta, c)$  is 1 more than the expected value of  $Z$  would be if we used the initial condition  $x_0 = c$  rather than  $x_0 = 1$ . By rescaling the interval  $[0, c]$  linearly to

$[0, 1]$  and noting that this sends  $\delta$  to  $\delta/c$ , we see that this latter expected value is equal to  $f(\delta/c)$ . That is, for  $c \geq \delta$ ,  $g(\delta, c) = 1 + f(\delta/c)$ . It follows that we have

$$\begin{aligned}
f(\delta) &= \int_0^1 g(\delta, c) dc \\
&= \delta + \int_\delta^1 (1 + f(\delta/c)) dc = 1 + \int_\delta^1 f(\delta/c) dc.
\end{aligned}$$

Now define  $h : [1, \infty) \rightarrow \mathbb{R}$  by  $h(x) = f(1/x)$ ; then we have

$$h(x) = 1 + \int_{1/x}^1 h(cx) dc = 1 + \frac{1}{x} \int_1^x h(c) dc.$$

Rewriting this as  $xh(x) - x = \int_1^x h(c) dc$  and differentiating with respect to  $x$  gives  $h(x) + xh'(x) - 1 = h(x)$ , whence  $h'(x) = 1/x$  and so  $h(x) = \log(x) + C$  for some constant  $C$ . Since  $h(1) = f(1) = 1$ , we conclude that  $C = 1$ ,  $h(x) = 1 + \log(x)$ , and finally  $f(\delta) = 1 - \log(\delta)$ . This gives the claimed answer.

B4 The answer is  $\frac{1}{4040}$ . We will show the following more general fact. Let  $a$  be any nonzero number and define  $q(\mathbf{v}) = 1 + \sum_{j=1}^{2n-1} a^{s_j}$ ; then the average of  $\frac{1}{q(\mathbf{v})}$  over all  $\mathbf{v} \in V_n$  is equal to  $\frac{1}{2n}$ , independent of  $a$ .

Let  $W_n$  denote the set of  $(2n)$ -tuples  $\mathbf{w} = (w_1, \dots, w_{2n})$  such that  $n$  of the  $w_i$ 's are equal to  $+1$  and the other  $n$  are equal to  $-1$ . Define a map  $\phi : W_n \rightarrow W_n$  by  $\phi(w_1, w_2, \dots, w_{2n}) = (w_2, \dots, w_{2n}, w_1)$ ; that is,  $\phi$  moves the first entry to the end. For  $\mathbf{w} \in W_n$ , define the *orbit* of  $\mathbf{w}$  to be the collection of elements of  $W_n$  of the form  $\phi^k(\mathbf{w})$ ,  $k \geq 1$ , where  $\phi^k$  denotes the  $k$ -th iterate of  $\phi$ , and note that  $\phi^{2n}(\mathbf{w}) = \mathbf{w}$ . Then  $W_n$  is a disjoint union of orbits. For a given  $\mathbf{w} \in W_n$ , the orbit of  $\mathbf{w}$  consists of  $\mathbf{w}, \phi(\mathbf{w}), \dots, \phi^{m-1}(\mathbf{w})$ , where  $m$  is the smallest positive integer with  $\phi^m(\mathbf{w}) = \mathbf{w}$ ; the list  $\phi(\mathbf{w}), \dots, \phi^{2n}(\mathbf{w})$  runs through the orbit of  $\mathbf{w}$  completely  $2n/m$  times, with each element of the orbit appearing the same number of times.

Now define the map  $f : W_n \rightarrow V_n$  by  $f(\mathbf{w}) = \mathbf{v} = (s_0, \dots, s_{2n})$  with  $s_j = \sum_{i=1}^j w_i$ ; this is a one-to-one correspondence between  $W_n$  and  $V_n$ , with the inverse map given by  $w_j = s_j - s_{j-1}$  for  $j = 1, \dots, 2n$ . We claim that for any  $\mathbf{w} \in W_n$ , the average of  $\frac{1}{q(\mathbf{v})}$ , where  $\mathbf{v}$  runs over vectors in the image of the orbit of  $\mathbf{w}$  under  $f$ , is equal to  $\frac{1}{2n}$ . Since  $W_n$  is a disjoint union of orbits,  $V_n$  is a disjoint union of the images of these orbits under  $f$ , and it then follows that the overall average of  $\frac{1}{q(\mathbf{v})}$  over  $\mathbf{v} \in V_n$  is  $\frac{1}{2n}$ .

To prove the claim, we compute the average of  $\frac{1}{q(f(\phi^k(\mathbf{w})))}$  over  $k = 1, \dots, 2n$ ; since  $\phi^k(\mathbf{w})$  for  $k = 1, \dots, 2n$  runs over the orbit of  $\mathbf{w}$  with each element in the orbit appearing equally, this is equal to the desired average. Now if we adopt the convention that the indices in  $w_i$  are considered mod  $2n$ , so that  $w_{2n+i} = w_i$  for all  $i$ , then the  $i$ -th entry of  $\phi^k(\mathbf{w})$  is  $w_{i+k}$ ; we can

then define  $s_j = \sum_{i=1}^j w_i$  for all  $j \geq 1$ , and  $s_{2n+i} = s_i$  for all  $i$  since  $\sum_{i=1}^{2n} w_i = 0$ . We now have

$$q(f(\phi^k(\mathbf{w}))) = \sum_{j=1}^{2n} a^{\sum_{i=1}^j w_{i+k}} = \sum_{j=1}^{2n} a^{s_{j+k}-s_k} = a^{-s_k} \sum_{j=1}^{2n} a^{s_j}.$$

Thus

$$\sum_{k=1}^{2n} \frac{1}{q(f(\phi^k(\mathbf{w})))} = \sum_{k=1}^{2n} \frac{a^{s_k}}{\sum_{j=1}^{2n} a^{s_j}} = 1,$$

and the average of  $\frac{1}{q(f(\phi^k(\mathbf{w})))}$  over  $k = 1, \dots, 2n$  is  $\frac{1}{2n}$ , as desired.

**B5 First solution.** (by Mitja Mastnak) It will suffice to show that for any  $z_1, z_2, z_3, z_4 \in \mathbb{C}$  of modulus 1 such that  $|3 - z_1 - z_2 - z_3 - z_4| = |z_1 z_2 z_3 z_4|$ , at least one of  $z_1, z_2, z_3$  is equal to 1.

To this end, let  $z_1 = e^{\alpha i}, z_2 = e^{\beta i}, z_3 = e^{\gamma i}$  and

$$f(\alpha, \beta, \gamma) = |3 - z_1 - z_2 - z_3|^2 - |1 - z_1 z_2 z_3|^2.$$

A routine calculation shows that

$$\begin{aligned} f(\alpha, \beta, \gamma) &= 10 - 6\cos(\alpha) - 6\cos(\beta) - 6\cos(\gamma) \\ &\quad + 2\cos(\alpha + \beta + \gamma) + 2\cos(\alpha - \beta) \\ &\quad + 2\cos(\beta - \gamma) + 2\cos(\gamma - \alpha). \end{aligned}$$

Since the function  $f$  is continuously differentiable, and periodic in each variable,  $f$  has a maximum and a minimum and it attains these values only at points where  $\nabla f = (0, 0, 0)$ . A routine calculation now shows that

$$\begin{aligned} \frac{\partial f}{\partial \alpha} + \frac{\partial f}{\partial \beta} + \frac{\partial f}{\partial \gamma} &= 6(\sin(\alpha) + \sin(\beta) + \sin(\gamma) - \sin(\alpha + \beta + \gamma)) \\ &= 24 \sin\left(\frac{\alpha + \beta}{2}\right) \sin\left(\frac{\beta + \gamma}{2}\right) \sin\left(\frac{\gamma + \alpha}{2}\right). \end{aligned}$$

Hence every critical point of  $f$  must satisfy one of  $z_1 z_2 = 1$ ,  $z_2 z_3 = 1$ , or  $z_3 z_1 = 1$ . By symmetry, let us assume that  $z_1 z_2 = 1$ . Then

$$f = |3 - 2\operatorname{Re}(z_1) - z_3|^2 - |1 - z_3|^2;$$

since  $3 - 2\operatorname{Re}(z_1) \geq 1$ ,  $f$  is nonnegative and can be zero only if the real part of  $z_1$ , and hence also  $z_1$  itself, is equal to 1.

**Remark.** If  $z_1 = 1$ , we may then apply the same logic to deduce that one of  $z_2, z_3, z_4$  is equal to 1. If  $z_1 = z_2 = 1$ , we may factor the expression

$$3 - z_1 - z_2 - z_3 - z_4 + z_1 z_2 z_3 z_4$$

as  $(1 - z_3)(1 - z_4)$  to deduce that at least three of  $z_1, \dots, z_4$  are equal to 1.

**Second solution.** We begin with an ‘‘unsmoothing’’ construction.

**Lemma 1.** Let  $z_1, z_2, z_3$  be three distinct complex numbers with  $|z_j| = 1$  and  $z_1 + z_2 + z_3 \in [0, +\infty)$ . Then there exist another three complex numbers  $z'_1, z'_2, z'_3$ , not all distinct, with  $|z'_j| = 1$  and

$$z'_1 + z'_2 + z'_3 \in (z_1 + z_2 + z_3, +\infty), \quad z_1 z_2 z_3 = z'_1 z'_2 z'_3.$$

*Proof.* Write  $z_j = e^{i\theta_j}$  for  $j = 1, 2, 3$ . We are then trying to maximize the target function

$$\cos \theta_1 + \cos \theta_2 + \cos \theta_3$$

given the constraints

$$0 = \sin \theta_1 + \sin \theta_2 + \sin \theta_3$$

$$* = \theta_1 + \theta_2 + \theta_3$$

Since  $z_1, z_2, z_3$  run over a compact region without boundary, the maximum must be achieved at a point where the matrix

$$\begin{pmatrix} \sin \theta_1 & \sin \theta_2 & \sin \theta_3 \\ \cos \theta_1 & \cos \theta_2 & \cos \theta_3 \\ 1 & 1 & 1 \end{pmatrix}$$

is singular. Since the determinant of this matrix computes (up to a sign and a factor of 2) the area of the triangle with vertices  $z_1, z_2, z_3$ , it cannot vanish unless some two of  $z_1, z_2, z_3$  are equal. This proves the claim.  $\square$

For  $n$  a positive integer, let  $H_n$  be the *hypocycloid curve* in  $\mathbb{C}$  given by

$$H_n = \{(n-1)z + z^{-n+1} : z \in \mathbb{C}, |z| = 1\}.$$

In geometric terms,  $H_n$  is the curve traced out by a marked point on a circle of radius 1 rolling one full circuit along the interior of a circle of radius 1, starting from the point  $z = 1$ . Note that the interior of  $H_n$  is not convex, but it is *star-shaped*: it is closed under multiplication by any number in  $[0, 1]$ .

**Lemma 2.** For  $n$  a positive integer, let  $S_n$  be the set of complex numbers of the form  $w_1 + \dots + w_n$  for some  $w_1, \dots, w_n \in \mathbb{C}$  with  $|w_j| = 1$  and  $w_1 \cdots w_n = 1$ . Then for  $n \leq 4$ ,  $S_n$  is the closed interior of  $H_n$  (i.e., including the boundary).

*Proof.* By considering  $n$ -tuples of the form  $(z, \dots, z, z^{-n+1})$ , we see that  $H_n \subseteq S_n$ . It thus remains to check that  $S_n$  lies in the closed interior of  $H_n$ . We ignore the easy cases  $n = 1$  (where  $H_1 = S_1 = \{1\}$ ) and  $n = 2$  (where  $H_2 = S_2 = [-2, 2]$ ) and assume hereafter that  $n \geq 3$ .

By Lemma 1, for each ray emanating from the the origin, the extreme intersection point of  $S_n$  with this ray (which exists because  $S_n$  is compact) is achieved by some tuple  $(w_1, \dots, w_n)$  with at most two distinct values. For  $n = 3$ , this immediately implies that this point lies on  $H_n$ . For  $n = 4$ , we must also consider tuples consisting of two pairs of equal values; however, these only give rise to points in  $[-4, 4]$ , which are indeed contained in  $H_4$ .  $\square$

Turning to the original problem, consider  $z_1, \dots, z_4 \in \mathbb{C}$  with  $|z_j| = 1$  and

$$3 - z_1 - z_2 - z_3 - z_4 + z_1 z_2 z_3 z_4 = 0;$$

we must prove that at least one  $z_j$  is equal to 1. Let  $z$  be any fourth root of  $z_1 z_2 z_3 z_4$ , put  $w_j = z_j/z$ , and put  $s = w_1 + \dots + w_4$ . In this notation, we have

$$s = z^3 + 3z^{-1},$$

where  $s \in S_4$  and  $z^3 + 3z^{-1} \in H_4$ . That is,  $s$  is a boundary point of  $S_4$ , so in particular it is the extremal point of  $S_4$  on the ray emanating from the origin through  $s$ . By Lemma 1, this implies that  $w_1, \dots, w_4$  take at most two distinct values. As in the proof of Lemma 2, we distinguish two cases.

– If  $w_1 = w_2 = w_3$ , then

$$w_1^{-3} + 3w_1 = z^3 + 3z^{-1}.$$

From the geometric description of  $H_n$ , we see that this forces  $w_1^{-1} = z$  and hence  $z_1 = 1$ .

– If  $w_1 = w_2$  and  $w_3 = w_4$ , then  $s \in [-4, 4]$  and hence  $s = \pm 4$ . This can only be achieved by taking  $w_1 = \dots = w_4 = \pm 1$ ; since  $s = z^3 + 3z^{-1}$  we must also have  $z = \pm 1$ , yielding  $z_1 = \dots = z_4 = 1$ .

**Remark.** With slightly more work, one can show that Lemma 2 remains true for all positive integers  $n$ . The missing extra step is to check that for  $m = 1, \dots, n-1$ , the hypocycloid curve

$$\{mz^{n-m} + (n-m)z^{-m} : z \in \mathbb{C}, |z| = 1\}$$

is contained in the filled interior of  $H_n$ . In fact, this curve only touches  $H_n$  at points where they both touch the unit circle (i.e., at  $d$ -th roots of unity for  $d = \gcd(m, n)$ ); this can be used to formulate a corresponding version of the original problem, which we leave to the reader.

**B6 First solution.** Define the sequence  $\{a_k\}_{k=0}^{\infty}$  by  $a_k = \lfloor k(\sqrt{2}-1) \rfloor$ . The first few terms of the sequence  $\{(-1)^{a_k}\}$  are

$$1, 1, 1, -1, -1, 1, 1, 1, -1, -1, 1, 1, 1, \dots$$

Define a new sequence  $\{c_i\}_{i=0}^{\infty}$  given by  $3, 2, 3, 2, 3, \dots$ , whose members alternately are the lengths of the clusters of consecutive 1's and the lengths of the clusters of consecutive  $-1$ 's in the sequence  $\{(-1)^{a_k}\}$ . Then for any  $i$ ,  $c_0 + \dots + c_i$  is the number of nonnegative integers  $k$  such that  $\lfloor k(\sqrt{2}-1) \rfloor$  is strictly less than  $i+1$ , i.e., such that  $k(\sqrt{2}-1) < i+1$ . This last condition is equivalent to  $k < (i+1)(\sqrt{2}+1)$ , and we conclude that

$$\begin{aligned} c_0 + \dots + c_i &= \lfloor (i+1)(\sqrt{2}+1) \rfloor + 1 \\ &= 2i + 3 + \lfloor (i+1)(\sqrt{2}-1) \rfloor. \end{aligned}$$

Thus for  $i > 0$ ,

$$c_i = 2 + \lfloor (i+1)(\sqrt{2}-1) \rfloor - \lfloor i(\sqrt{2}-1) \rfloor. \quad (8)$$

Now note that  $\lfloor (i+1)(\sqrt{2}-1) \rfloor - \lfloor i(\sqrt{2}-1) \rfloor$  is either 1 or 0 depending on whether or not there is an integer  $j$  between  $i(\sqrt{2}-1)$  and  $(i+1)(\sqrt{2}-1)$ ; this condition is equivalent to  $i < j(\sqrt{2}+1) < i+1$ . That is, for  $i > 0$ ,

$$c_i = \begin{cases} 3 & \text{if } i = \lfloor j(\sqrt{2}+1) \rfloor \text{ for some integer } j, \\ 2 & \text{otherwise;} \end{cases} \quad (9)$$

by inspection, this also holds for  $i = 0$ .

Now we are asked to prove that

$$\sum_{k=0}^n (-1)^{a_k} \geq 1 \quad (10)$$

for all  $n \geq 1$ . We will prove that if (10) holds for all  $n \leq N$ , then (10) holds for all  $n \leq 4N$ ; since (10) clearly holds for  $n = 1$ , this will imply the desired result.

Suppose that (10) holds for  $n \leq N$ . To prove that (10) holds for  $n \leq 4N$ , it suffices to show that the partial sums

$$\sum_{i=0}^m (-1)^{a_i} c_i$$

of the sequence  $\{(-1)^{a_k}\}$  are positive for all  $m$  such that  $c_0 + \dots + c_{m-1} < 4N+3$ , since these partial sums cover all clusters through  $a_{4N}$ . Now if  $c_0 + \dots + c_{m-1} < 4N+3$ , then since each  $c_i$  is at least 2, we must have  $m < 2N+2$ . From (9), we see that if  $m$  is odd, then

$$\begin{aligned} \sum_{i=0}^m (-1)^{a_i} c_i &= \sum_{i=0}^m (-1)^{a_i} (c_i - 2) \\ &= \sum_j (-1)^{\lfloor j(\sqrt{2}+1) \rfloor} = \sum_j (-1)^{a_j} \end{aligned}$$

where the sum in  $j$  is over nonnegative integers  $j$  with  $j(\sqrt{2}+1) < m$ , i.e.,  $j < m(\sqrt{2}-1)$ ; since  $m(\sqrt{2}-1) < m/2 < N+1$ ,  $\sum_j (-1)^{a_j}$  is positive by the induction hypothesis. Similarly, if  $m$  is even, then  $\sum_{i=0}^m (-1)^{a_i} c_i = c_m + \sum_j (-1)^{a_j}$  and this is again positive by the induction hypothesis. This concludes the induction step and the proof.

**Remark.** More generally, using the same proof we can establish the result with  $\sqrt{2}-1$  replaced by  $\sqrt{n^2+1}-n$  for any positive integer  $n$ .

**Second solution.** For  $n \geq 0$ , define the function

$$f(n) = \sum_{k=1}^n (-1)^{\lfloor k(\sqrt{2}-1) \rfloor}$$

with the convention that  $f(0) = 0$ .

Define the sequence  $q_0, q_1, \dots$  by the initial conditions

$$q_0 = 0, q_1 = 1$$

and the recurrence relation

$$q_j = 2q_{j-1} + q_{j-2}.$$

This is OEIS sequence A000129; its first few terms are

$$0, 1, 2, 5, 12, 29, 70, \dots$$

Note that  $q_j \equiv j \pmod{2}$ .

We now observe that the fractions  $q_{j-1}/q_j$  are the *convergents* of the continued fraction expansion of  $\sqrt{2}-1$ . This implies the following additional properties of the sequence.

– For all  $j \geq 0$ ,

$$\frac{q_{2j}}{q_{2j+1}} < \sqrt{2} - 1 < \frac{q_{2j+1}}{q_{2j+2}}.$$

– There is no fraction  $r/s$  with  $s < q_j + q_{j+1}$  such that  $\frac{r}{s}$  separates  $\sqrt{2}-1$  from  $q_j/q_{j-1}$ . In particular, for  $k < q_j + q_{j+1}$ ,

$$\lfloor k(\sqrt{2}-1) \rfloor = \left\lfloor \frac{kq_{j-1}}{q_j} \right\rfloor$$

except when  $j$  is even and  $k \in \{q_j, 2q_j\}$ , in which case they differ by 1.

We use this to deduce a “self-similarity” property of  $f(n)$ .

**Lemma 1.** *Let  $n, j$  be nonnegative integers with  $q_j \leq n < q_{j+1}$ .*

(a) *If  $j$  is even, then*

$$f(n) = f(q_j) - f(n - q_j).$$

(b) *If  $j$  is odd, then*

$$f(n) = f(n - q_j) + 1.$$

*Proof.* If  $j$  is even, then

$$\begin{aligned} f(n) &= f(q_j) + \sum_{k=q_j+1}^n (-1)^{\lfloor k(\sqrt{2}-1) \rfloor} \\ &= f(q_j) + \sum_{k=q_j+1}^n (-1)^{\lfloor kq_{j-1}/q_j \rfloor} + * \end{aligned}$$

where  $*$  equals 2 if  $n \geq 2q_j$  (accounting for the term  $k = 2q_j$ ) and 0 otherwise. Continuing,

$$\begin{aligned} f(n) &= f(q_j) + \sum_1^{n-q_j} (-1)^{q_{j-1} + \lfloor kq_{j-1}/q_j \rfloor} + * \\ &= f(q_j) - \sum_1^{n-q_j} (-1)^{q_{j-1} + \lfloor k(\sqrt{2}-1) \rfloor} \\ &= f(q_j) - f(n - q_j). \end{aligned}$$

If  $j$  is odd, then

$$\begin{aligned} f(n) &= f(n - q_j) + \sum_{k=n-q_j+1}^n (-1)^{\lfloor k(\sqrt{2}-1) \rfloor} \\ &= f(n - q_j) - 2 + \sum_{k=n-q_j+1}^n (-1)^{\lfloor kq_{j-1}/q_j \rfloor}. \end{aligned}$$

Since

$$\lfloor (k + q_j)q_{j-1}/q_j \rfloor \equiv \lfloor kq_{j-1}/q_j \rfloor \pmod{2},$$

we also have

$$f(n) = f(n - q_j) + \sum_{k=1}^{q_j} (-1)^{\lfloor kq_{j-1}/q_j \rfloor}.$$

In this sum, the summand indexed by  $q_j$  contributes 1, and the summands indexed by  $k$  and  $q_j - k$  cancel each other out for  $k = 1, \dots, q_j - 1$ . We thus have

$$f(n) = f(n - q_j) + 1$$

as claimed.  $\square$

From Lemma 1, we have

$$f(q_{2j}) = f(q_{2j} - 2q_{2j-1}) + 2 = f(q_{2j-2}) + 2.$$

By induction on  $j$ ,  $f(q_{2j}) = 2j$  for all  $j \geq 0$ ; by similar logic, we have  $f(n) \leq f(q_{2j}) = 2j$  for all  $n \leq q_{2j}$ . We can now apply Lemma 1 once more to deduce that  $f(n) \geq 0$  for all  $j$ .

**Remark.** As a byproduct of the first solution, we confirm the equality of two sequences that were entered separately in the OEIS but conjectured to be equal: A097509 (indexed from 0) matches the definition of  $\{c_i\}$ , while A276862 (indexed from 1) matches the characterization of  $\{c_{i-1}\}$  given by (8).

**Remark.** We can confirm an additional conjecture from the OEIS by showing that in the notation of the first solution, the sequence  $a(n) = c_{n+1}$  indexed from 1 equals A082844: “Start with 3,2 and apply the rule  $a(a(1) + a(2) + \dots + a(n)) = a(n)$ , fill in any undefined terms with  $a(t) = 2$  if  $a(t-1) = 3$  and  $a(t) = 3$  if  $a(t-1) = 2$ .” We first verify the recursion. By (10),

$$\begin{aligned} a(1) + \dots + a(n) &= c_0 + \dots + c_{n+1} - c_0 - c_1 \\ &= \lfloor (n+2)(\sqrt{2}+1) \rfloor - 4. \end{aligned}$$

From (9), we see that  $a(a(1) + \dots + a(n) + 3) = 3$ . Consequently, exactly one of  $a(a(1) + \dots + a(n))$  or  $a(a(1) + \dots + a(n) + 1)$  equals 3, and it is the former if and only if

$$\lfloor (n+2)(\sqrt{2}+1) \rfloor - 3 = \lfloor (n+1)(\sqrt{2}+1) \rfloor,$$

i.e., if and only if  $a(n) = c_{n+1} = 3$ .



We next check that the definition correctly fills in values not determined by the recursion. If  $a(n) = 3$ , then  $a(a(1) + \dots + a(n) + 1) = 2$  because no two consecutive values can both equal 3; by the same token,  $a(n+1) = 2$  and so there are no further values to fill in. If  $a(n) = 2$ , then  $a(a(1) + \dots + a(n) + 1) = 3$  by the previous paragraph; this in turn implies  $a(a(1) + \dots + a(n) + 2) = 2$ , at which point there are no further values to fill in.

**Remark.** We can confirm an additional conjecture from the OEIS by showing that in the notation of the first solution, the sequence  $\{c_i\}$  equals A245219. This depends on some additional lemmas.

**Lemma 2.** *Let  $k$  be a positive integer. Then*

$$\left\{i(\sqrt{2}-1)\right\} < \left\{k(\sqrt{2}-1)\right\} \quad (i=0, \dots, k-1)$$

if and only if  $k = q_{2j}$  or  $k = q_{2j} + q_{2j-1}$  for some  $j > 0$ .

*Proof.* For each  $j > 0$ , we have

$$\frac{q_{2j-2}}{q_{2j-1}} < \frac{q_{2j}}{q_{2j+1}} = \frac{q_{2j-1} + 2q_{2j-2}}{q_{2j} + 2q_{2j-1}} < \sqrt{2} - 1 < \frac{q_{2j+1}}{q_{2j+2}} < \frac{q_{2j-1}}{q_{2j}}.$$

We also have

$$\frac{q_{2j-2}}{q_{2j-1}} < \frac{q_{2j}}{q_{2j+1}} = \frac{q_{2j-1} + 2q_{2j-2}}{q_{2j} + 2q_{2j-1}} < \frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}} < \frac{q_{2j-1}}{q_{2j}}.$$

Moreover,  $\frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}}$  cannot be less than  $\sqrt{2} - 1$ , or else it would be a better approximation to  $\sqrt{2} - 1$  than the convergent  $q_{2j}/q_{2j+1}$  with  $q_{2j+1} > q_{2j} + q_{2j-1}$ . By the same token,  $\frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}}$  cannot be a better approximation to  $\sqrt{2} - 1$  than  $q_{2j+1}/q_{2j+2}$ . We thus have

$$\frac{q_{2j}}{q_{2j+1}} < \sqrt{2} - 1 < \frac{q_{2j+1}}{q_{2j+2}} < \frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}} < \frac{q_{2j-1}}{q_{2j}}.$$

From this, we see that

$$\left\{q_{2j}(\sqrt{2}-1)\right\} < \left\{(q_{2j} + q_{2j-1})(\sqrt{2}-1)\right\} < \left\{q_{2j+2}(\sqrt{2}-1)\right\}.$$

It will now suffice to show that for  $q_{2j} < k < q_{2j} + q_{2j-1}$ ,

$$\left\{k(\sqrt{2}-1)\right\} < \left\{q_{2j}(\sqrt{2}-1)\right\}$$

while for  $q_{2j} + q_{2j-1} < k < q_{2j+2}$ ,

$$\left\{k(\sqrt{2}-1)\right\} < \left\{(q_{2j} + q_{2j-1})(\sqrt{2}-1)\right\}.$$

The first of these assertion is an immediate consequence of the ‘‘best approximation’’ property of the convergent  $q_{2j-1}/q_{2j}$ . As for the second assertion, note that for  $k$  in this range, no fraction with denominator  $k$  can lie strictly between  $\frac{q_{2j}}{q_{2j+1}}$  and  $\frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}}$  because these fractions are consecutive terms in a Farey sequence (that is, their difference has numerator 1 in lowest terms); in particular, such a fraction cannot be a better upper approximation to  $\sqrt{2} - 1$  than  $\frac{q_{2j-1} + q_{2j-2}}{q_{2j} + q_{2j-1}}$ .  $\square$

**Lemma 3.** *For  $j > 0$ , the sequence  $c_0, \dots, c_{j-1}$  is palindromic if and only if*

$$j = q_{2i+1} \quad \text{or} \quad j = q_{2i+1} + q_{2i+2}$$

for some nonnegative integer  $i$ . (That is,  $j$  must belong to one of the sequences A001653 or A001541.) In particular,  $j$  must be odd.

*Proof.* Let  $j$  be an index for which  $\{c_0, \dots, c_{j-1}\}$  is palindromic. In particular,  $c_{j-1} = c_0 = 3$ , so from (9), we see that  $j-1 = \lfloor k(\sqrt{2}+1) \rfloor$  for some  $k$ . Given this, the sequence is palindromic if and only if

$$\lfloor i(\sqrt{2}+1) \rfloor + \lfloor (k-i)(\sqrt{2}+1) \rfloor = \lfloor k(\sqrt{2}+1) \rfloor \quad (i=0, \dots, k),$$

or equivalently

$$\left\{i(\sqrt{2}-1)\right\} + \left\{(k-i)(\sqrt{2}-1)\right\} = \left\{k(\sqrt{2}-1)\right\} \quad (i=0, \dots, k)$$

where the braces denote fractional parts. This holds if and only if

$$\left\{i(\sqrt{2}-1)\right\} < \left\{k(\sqrt{2}-1)\right\} \quad (i=0, \dots, k-1),$$

so we may apply Lemma 2 to identify  $k$  and hence  $j$ .  $\square$

**Lemma 4.** *For  $j > 0$ , if there exists a positive integer  $k$  such that*

$$(c_0, \dots, c_{j-2}) = (c_k, \dots, c_{k+j-2}) \text{ but } c_{j-1} \neq c_{k+j-1},$$

then

$$j = q_{2i+1} \quad \text{or} \quad j = q_{2i+1} + q_{2i+2}$$

for some nonnegative integer  $i$ . In particular,  $j$  is odd and (by Lemma 3) the sequence  $(c_0, \dots, c_{j-1})$  is palindromic.

*Proof.* Since the sequence  $\{c_i\}$  consists of 2s and 3s, we must have  $\{c_{j-1}, c_{k+j-1}\} = \{2, 3\}$ . Since each pair of 3s is separated by either one or two 2s, we must have  $c_{j-2} = 2$ ,  $c_{j-3} = 3$ . In particular, by (9) there is an integer  $i$  for which  $j-3 = \lfloor (i-1)(\sqrt{2}+1) \rfloor$ ; there is also an integer  $l$  such that  $k = \lfloor l(\sqrt{2}+1) \rfloor$ . By hypothesis, we have

$$\lfloor (h+l)(\sqrt{2}+1) \rfloor = \lfloor h(\sqrt{2}+1) \rfloor + \lfloor l(\sqrt{2}+1) \rfloor$$

for  $h = 0, \dots, i-1$  but not for  $h = i$ . In other words,

$$\left\{(h+l)(\sqrt{2}-1)\right\} = \left\{h(\sqrt{2}-1)\right\} + \left\{l(\sqrt{2}-1)\right\}$$

for  $h = 0, \dots, i-1$  but not for  $h = i$ . That is,  $\{h(\sqrt{2}-1)\}$  belongs to the interval  $(0, 1 - \{l(\sqrt{2}-1)\})$  for  $h = 0, \dots, i-1$  but not for  $h = i$ ; in particular,

$$\left\{h(\sqrt{2}-1)\right\} < \left\{i(\sqrt{2}-1)\right\} \quad (h=0, \dots, i-1),$$

so we may apply Lemma 2 to identify  $i$  and hence  $j$ .  $\square$

The sequence A245219 is defined as the sequence of coefficients of the continued fraction of  $\sup\{b_i\}$  where  $b_1 = 1$  and for  $i > 1$ ,

$$b_{i+1} = \begin{cases} b_i + 1 & \text{if } i = \lfloor j\sqrt{2} \rfloor \text{ for some integer } j; \\ 1/b_i & \text{otherwise.} \end{cases}$$

It is equivalent to take the supremum over values of  $i$  for which  $b_{i+1} = 1/b_i$ ; by Beatty's theorem, this occurs precisely when  $i = \lfloor j(2 + \sqrt{2}) \rfloor$  for some integer  $j$ . In this case,  $b_i$  has continued fraction

$$[c_{j-1}, \dots, c_0].$$

Let  $K$  be the real number with continued fraction  $[c_0, c_1, \dots]$ ; we must show that  $K = \sup\{b_i\}$ . In one direction, by Lemma 3, there are infinitely many values of  $i$  for which  $[c_{j-1}, \dots, c_0] = [c_0, \dots, c_{j-1}]$ ; the corresponding values  $b_i$  accumulate at  $K$ , so  $K \leq \sup\{b_i\}$ .

In the other direction, we show that  $K \geq \sup\{b_i\}$  as

follows. It is enough to prove that  $K \geq b_i$  when  $i = \lfloor j(2 + \sqrt{2}) \rfloor$  for some integer  $j$ .

- If  $c_0, \dots, c_{j-1}$  is palindromic, then Lemma 3 implies that  $j$  is odd; that is, the continued fraction  $[c_{j-1}, \dots, c_0]$  has odd length. In this case, replacing the final term  $c_0 = c_{j-1}$  by the larger quantity  $[c_{j-1}, c_j, \dots]$  increases the value of the continued fraction.
- If  $c_0, \dots, c_{j-1}$  is not palindromic, then there is a least integer  $k \in \{0, \dots, j-1\}$  such that  $c_k \neq c_{j-1-k}$ . By Lemma 3, the sequence  $c_0, c_1, \dots$  has arbitrarily long palindromic initial segments, so the sequence  $(c_{j-1}, \dots, c_{j-1-k})$  also occurs as  $c_h, \dots, c_{h+k}$  for some  $h > 0$ . By Lemma 4,  $k$  is even and  $c_k = 3 > 2 = c_{j-1-k}$ ; hence in the continued fraction for  $b_i$ , replacing the final segment  $c_{j-1-k}, \dots, c_0$  by  $c_k, c_{k+1}, \dots$  increases the value.