SIMON MARAIS

MATHEMATICS COMPETITION

2019

SOLUTIONS

Problem A1

Consider the sequence $s_1, s_2, s_3, \ldots$ of positive integers defined by

- $s_1 = 2$, and
- for each positive integer n, s_{n+1} is equal to s_n plus the product of the prime factors of s_n.

The first terms of the sequence are $2, 4, 6, 12, 18, 24$.

Prove that the product of the 2019 smallest primes is a term of the sequence.

Solution

More generally, we will prove by induction that the product of the m smallest primes is a term of the sequence, for each positive integer m. One can check that the statement is certainly true for $m = 1, 2, 3$ since $2, 6, 30$ are terms of the sequence.

Let the primes be $p_1 < p_2 < p_3 < \cdots$. Now suppose that $p_1 p_2 \cdots p_k$ is a term of the sequence for some $k \geq 3$. Then the next terms of the sequence are

$$2p_1p_2 \cdots p_k, \quad 3p_1p_2 \cdots p_k, \quad 4p_1p_2 \cdots p_k, \quad 5p_1p_2 \cdots p_k, \quad \ldots.$$

In fact, to obtain successive terms, we keep adding $p_1 p_2 \cdots p_k$ until we obtain a term of the sequence that is divisible by a prime larger than p_k. Since the smallest positive integer not divisible by $p_1, p_2, \ldots, p_k$ is p_{k+1}, this term will be $p_1 p_2 \cdots p_k p_{k+1}$. So we have shown that if the product of the k smallest primes is a term of the sequence, then so is the product of the $k+1$ smallest primes. Combined with the base cases above, this completes the induction.

Problem A2

Consider the operation $*$ that takes a pair of integers and returns an integer according to the rule

$$a * b = a \times (b+1).$$

(a) For each positive integer n, determine all permutations $a_1, a_2, \ldots, a_n$ of the set $\{1, 2, \ldots, n\}$ that maximise the value of

$$(\cdots((a_1 * a_2) * a_3) * \cdots * a_{n-1}) * a_n.$$

(b) For each positive integer n, determine all permutations $b_1, b_2, \ldots, b_n$ of the set $\{1, 2, \ldots, n\}$ that maximise the value of

$$b_1 * (b_2 * (b_3 * \cdots * (b_{n-1} * b_n) \cdots)).$$

Solution

(a) We start by calculating the following expressions.

$$\begin{aligned} a_1 * a_2 &= a_1 \times (a_2+1) \\ (a_1 * a_2) * a_3 &= [a_1 \times (a_2+1)] \times (a_3+1) = a_1(a_2+1)(a_3+1) \\ ((a_1 * a_2) * a_3) * a_4 &= [a_1(a_2+1)(a_3+1)] \times (a_4+1) = a_1(a_2+1)(a_3+1)(a_4+1) \end{aligned}$$

A straightforward induction can then be used to prove more generally that

$$(\cdots((a_1 * a_2) * a_3) * \cdots * a_{n-1}) * a_n = a_1(a_2+1)(a_3+1)\cdots(a_{n-1}+1)(a_n+1).$$

If $a_1, a_2, \ldots, a_n$ is a permutation of $\{1, 2, \ldots, n\}$, then we have the following.

$$\begin{aligned} &(\cdots((a_1 * a_2) * a_3) * \cdots * a_{n-1}) * a_n \\ &= a_1(a_2+1)(a_3+1)\cdots(a_{n-1}+1)(a_n+1) \\ &= \frac{a_1}{a_1+1}(a_1+1)(a_2+1)(a_3+1)\cdots(a_{n-1}+1)(a_n+1) \\ &= \frac{a_1}{a_1+1}(n+1)! \end{aligned}$$

Note that for fixed n, this expression depends only on a_1 and is an increasing function of a_1 for $a_1 > 0$. Given that $a_1 \in \{1, 2, \ldots, n\}$, it is maximised when $a_1 = n$. Therefore, the permutations $a_1, a_2, \ldots, a_n$ of $\{1, 2, \ldots, n\}$ that maximise the value of the expression are precisely those with $a_1 = n$.

(b) We start by calculating the following expressions.

$$\begin{aligned} b_1 * b_2 &= b_1 \times (b_2+1) = b_1 + b_1b_2 \\ b_1 * (b_2 * b_3) &= b_1 \times ([b_2 + b_2b_3] + 1) = b_1 + b_1b_2 + b_1b_2b_3 \\ b_1 * (b_2 * (b_3 * b_4)) &= b_1 \times ([b_2 + b_2b_3 + b_2b_3b_4] + 1) = b_1 + b_1b_2 + b_1b_2b_3 + b_1b_2b_3b_4 \end{aligned}$$

A straightforward induction can then be used to prove more generally that

$$b_1 * (b_2 * (b_3 * \cdots * (b_{n-1} * b_n) \cdots)) = b_1 + b_1 b_2 + b_1 b_2 b_3 + \cdots + b_1 b_2 b_3 \cdots b_n.$$

Now suppose that $b_1, b_2, \ldots, b_n$ is a permutation of $\{1, 2, \ldots, n\}$ such that there exists $1 \leq i \leq n-1$ with $b_i < b_{i+1}$. Let $c_1, c_2, \ldots, c_n$ be the permutation of $\{1, 2, \ldots, n\}$ obtained from $b_1, b_2, \ldots b_n$ by swapping the terms b_i and b_{i+1}. Then we have the following.

$$\begin{aligned}
&c_1 * (c_2 * (c_3 * \cdots * (c_{n-1} * c_n) \cdots)) - b_1 * (b_2 * (b_3 * \cdots * (b_{n-1} * b_n) \cdots)) \\
=\ &(c_1 + c_1 c_2 + c_1 c_2 c_3 + \cdots + c_1 c_2 \cdots c_n) - (b_1 + b_1 b_2 + b_1 b_2 b_3 + \cdots + b_1 b_2 \cdots b_n) \\
=\ &c_1 c_2 \cdots c_{i-1} c_i - b_1 b_2 \cdots b_{i-1} b_i \\
=\ &b_1 b_2 \cdots b_{i-1} b_{i+1} - b_1 b_2 \cdots b_{i-1} b_i \\
=\ &b_1 b_2 \cdots b_{i-1} (b_{i+1} - b_i) \\
>\ &0
\end{aligned}$$

The calculation above shows that if $b_1, b_2, \ldots, b_n$ is a permutation of $\{1, 2, \ldots, n\}$ that maximises the value of

$$b_1 * (b_2 * (b_3 * \cdots * (b_{n-1} * b_n) \cdots)),$$

then there cannot be any $1 \leq i \leq n-1$ such that $b_i < b_{i+1}$. However, a permutation that maximises the value of the expression must exist, since there are only finitely many permutations of $\{1, 2, \ldots, n\}$. It follows that the only permutation that maximises the expression is

$$n, n-1, n-2, \ldots, 3, 2, 1,$$

or, in other words, the permutation $b_1, b_2, \ldots, b_n$ with $b_i = n + 1 - i$ for $1 \leq i \leq n$.

Problem A3

For some positive integer n, a coin will be flipped n times to obtain a sequence of n heads and tails. For each flip of the coin, there is probability p of obtaining a head and probability $1-p$ of obtaining a tail, where $0 < p < 1$ is a rational number.

Kim writes all 2^n possible sequences of n heads and tails in two columns, with some sequences in the left column and the remaining sequences in the right column. Kim would like the sequence produced by the coin flips to appear in the left column with probability $\frac{1}{2}$.

Determine all pairs (n, p) for which this is possible.

Solution

Any particular sequence that includes k heads and $n-k$ tails occurs with probability $p^k(1-p)^{n-k}$. So we are looking for (n, p) for which

$$\sum_{k=0}^{n} a_k\, p^k (1-p)^{n-k} = \frac{1}{2},$$

where $a_0, a_1, \ldots, a_n$ are integers satisfying $0 \le a_k \le \binom{n}{k}$. The equation above implies that

$$2\sum_{k=0}^{n} a_k\, p^k (1-p)^{n-k} - 1 = 0.$$

Consider the left side of this equation as a polynomial in p. Observe that it has integer coefficients and constant term $2a_0 - 1 = \pm 1$. The rational root theorem then implies that any rational root p must be of the form $\frac{1}{m}$ for some non-zero integer m. Since we are only interested in $0 < p < 1$, we may furthermore assume that $m \ge 2$. Substituting $p = \frac{1}{m}$ into the equation above yields

$$2\sum_{k=0}^{n} a_k \left(\frac{1}{m}\right)^k \left(\frac{m-1}{m}\right)^{n-k} - 1 = 0 \qquad \Rightarrow \qquad 2\sum_{k=0}^{n} a_k\, (m-1)^{n-k} - m^n = 0.$$

Now consider this equation modulo $m-1$ to deduce that $2a_0 - 1^n \equiv 0 \pmod{m-1}$, from which it follows that $\pm 1 \equiv 0 \pmod{m-1}$. However, this can only hold if $m = 2$, so we conclude that $p = \frac{1}{2}$.

The conditions of the problem are satisfied for $(n, \frac{1}{2})$ for all positive integers n. For example, Kim can write all sequences that start with a head in the left column and all sequences that start with a tail in the right column.

Problem A4

Suppose that $x_1, x_2, x_3, \ldots$ is a strictly decreasing sequence of positive real numbers such that the series $x_1 + x_2 + x_3 + \cdots$ diverges.

Is it necessarily true that the series

$$\sum_{n=2}^{\infty} \min\left\{x_n, \frac{1}{n \log n}\right\}$$

diverges?

Solution

Throughout, we interpret log as the natural logarithm with base e, although the result is independent of the choice of base, as long as it is a real number greater than 1. We will exhibit below a non-increasing sequence $x_1, x_2, x_3, \ldots$ of positive real numbers such that $x_1 + x_2 + x_3 + \cdots$ diverges, while

$$\sum_{n=2}^{\infty} \min\left\{x_n, \frac{1}{n \log n}\right\}$$

converges. A strictly decreasing sequence that satisfies the same properties is then given by $x_1 + \frac{1}{2}, x_2 + \frac{1}{4}, x_3 + \frac{1}{8}, x_4 + \frac{1}{16}, \ldots$, since $\min\left\{x_n + \frac{1}{2^n}, \frac{1}{n \log n}\right\} \leq \min\left\{x_n, \frac{1}{n \log n}\right\} + \frac{1}{2^n}$.

Begin by choosing any sequence $a_0, a_1, a_2, a_3, \ldots$ of positive integers that satisfies the conditions

- $a_0 = 1$ and $a_1 \geq 3$;
- $a_{k+1} > a_k \log a_k$ for each non-negative integer k;
- $\sum \frac{\log \log a_k}{\log a_k}$ converges.

For example, one can check that the sequence defined by $a_0 = 1$ and $a_k = \lfloor e^{e^k} \rfloor$ for each positive integer k satisfies the conditions. The first condition is trivial to check, while the second condition follows from the deduction below, starting from the well-known inequality $e^x \geq x + 1$.

$$\begin{aligned} e^k \geq k+1 \quad &\Rightarrow \quad 2e^k - 1 \geq e^k + k \quad \Rightarrow \quad e^{k+1} - 1 > e^k + k \\ &\Rightarrow \quad \frac{1}{e} \times e^{e^{k+1}} > e^{e^k} \times e^k \quad \Rightarrow \quad \lfloor e^{e^{k+1}} \rfloor > \lfloor e^{e^k} \rfloor \log \lfloor e^{e^k} \rfloor \end{aligned}$$

The third condition follows from the chain of inequalities below.

$$\sum_{k=1}^{\infty} \frac{\log \log \lfloor e^{e^k} \rfloor}{\log \lfloor e^{e^k} \rfloor} < \sum_{k=1}^{\infty} \frac{\log \log e^{e^k}}{\log(e^{e^k} - 1)} < \sum_{k=1}^{\infty} \frac{k}{\log(e^{e^k} - 1)} < \sum_{k=1}^{\infty} \frac{k}{\log(\frac{1}{e} \times e^{e^k})} < \sum_{k=1}^{\infty} \frac{k}{e^k - 1}$$

The last series converges by the limit comparison test applied to the series $\sum \frac{k}{e^k} = \frac{e}{(e-1)^2}$.

The properties above ensure that the sequence $a_0 \log a_0, a_1 \log a_1, a_2 \log a_2, \ldots$ is a strictly increasing sequence of real numbers, such that the only integer appearing in the sequence is $a_0 \log a_0 = 0$. So for each integer $n \geq 1$, there exists a unique integer $k \geq 0$ such that $a_k \log a_k < n < a_{k+1} \log a_{k+1}$. We then define the non-increasing sequence $x_1, x_2, x_3, \ldots$ via

$$x_n = \frac{1}{a_{k+1} \log a_{k+1}}.$$

The number of terms in the sequence that are equal to $\frac{1}{a_{k+1}\log a_{k+1}}$ is greater than the number $a_{k+1}\log a_{k+1} - a_k \log a_k - 1$. Therefore, we have the following.

$$\begin{aligned}\sum_{n=1}^{\infty} x_n &> \sum_{k=0}^{\infty} \frac{a_{k+1}\log a_{k+1} - a_k\log a_k - 1}{a_{k+1}\log a_{k+1}} > \sum_{k=1}^{\infty} \frac{a_{k+1}\log a_{k+1} - a_k\log a_k - 1}{a_{k+1}\log a_{k+1}} \\ &> \sum_{k=1}^{\infty} \frac{a_{k+1}\log a_{k+1} - 2a_k\log a_k}{a_{k+1}\log a_{k+1}} = \sum_{k=1}^{\infty} \left(1 - \frac{2a_k\log a_k}{a_{k+1}\log a_{k+1}}\right) \\ &> \sum_{k=1}^{\infty} \left(1 - \frac{2}{\log a_{k+1}}\right) > \sum_{k=1}^{\infty} \left(1 - \frac{2}{\log a_1}\right).\end{aligned}$$

It follows that the series $x_1 + x_2 + x_3 + \cdots$ diverges.

Note that if $a_k \le n \le a_k \log a_k$, then we have $\min\{x_n, \frac{1}{n\log n}\} = \frac{1}{n\log n}$. So we may deduce the following.

$$\begin{aligned} A = \sum_{k=1}^{\infty} \sum_{n=a_k}^{\lfloor a_k\log a_k\rfloor} \min\left\{x_n, \frac{1}{n\log n}\right\} &= \sum_{k=1}^{\infty} \sum_{n=a_k}^{\lfloor a_k\log a_k\rfloor} \frac{1}{n\log n} \\ &\le \sum_{k=1}^{\infty} \left(\frac{1}{a_k\log a_k} + \int_{a_k}^{a_k\log a_k} \frac{1}{t\log t}\,\mathrm{d}t\right) \\ &= \sum_{k=1}^{\infty} \frac{1}{a_k\log a_k} + \sum_{k=1}^{\infty} \log\log(a_k\log a_k) - \log\log(a_k) \\ &= \sum_{k=1}^{\infty} \frac{1}{a_k\log a_k} + \sum_{k=1}^{\infty} \log\left(1 + \frac{\log\log a_k}{\log a_k}\right)\end{aligned}$$

The first summation converges by construction of the sequence $a_1, a_2, a_3, \ldots$. By the well-known equality $\log(1+x) \le x$, the second summation may be compared with

$$\sum_{k=1}^{\infty} \frac{\log\log a_k}{\log a_k},$$

which converges by construction of the sequence $a_1, a_2, a_3, \ldots$. In particular, the series A converges.

Now note that if $a_k\log a_k < n < a_{k+1}$, then we have $\min\{x_n, \frac{1}{n\log n}\} = x_n = \frac{1}{a_{k+1}\log a_{k+1}}$. So we may deduce the following.

$$\begin{aligned} B = \sum_{k=1}^{\infty} \sum_{n=\lceil a_k\log a_k\rceil}^{a_{k+1}-1} \min\left\{x_n, \frac{1}{n\log n}\right\} &= \sum_{k=1}^{\infty} \sum_{n=\lceil a_k\log a_k\rceil}^{a_{k+1}-1} \frac{1}{a_{k+1}\log a_{k+1}} \\ &< \sum_{k=1}^{\infty} \frac{a_{k+1}}{a_{k+1}\log a_{k+1}} = \sum_{k=1}^{\infty} \frac{1}{\log a_{k+1}} < \sum_{k=1}^{\infty} \frac{\log\log a_{k+1}}{\log a_{k+1}}.\end{aligned}$$

This last summation converges by construction of the sequence $a_1, a_2, a_3, \ldots$. In particular, the series B converges.

It now suffices to observe that the series

$$\sum_{n=a_1}^{\infty} \min\left\{x_n, \frac{1}{n\log n}\right\} = A + B$$

converges, and the desired result follows.

Problem B1

Determine all pairs (a, b) of real numbers with $a \leq b$ that maximise the integral

$$\int_a^b e^{\cos x} \left(380 - x - x^2\right) \, \mathrm{d}x.$$

Solution

Write the integrand as

$$f(x) = e^{\cos x} \left(380 - x - x^2\right) = -e^{\cos x}(x + 20)(x - 19),$$

and observe that it is continuous. Since $e^{\cos x}$ is positive for all real x, the integrand $f(x)$ is positive for $-20 < x < 19$ and negative for $x < -20$ or $x > 19$. The essential idea is to interpret the integral as the signed area under the graph of $f(x)$, lying between $x = a$ and $x = b$. Hence, it is maximised when the interval of integration (a, b) includes all positive values of $f(x)$ and does not include any negative values. So the answer is given by $(a, b) = (-20, 19)$.

More precisely, let

$$F(x) = \int_0^x e^{\cos t} \left(380 - t - t^2\right) \, \mathrm{d}t.$$

We will repeatedly use the fact that if $f(x) > 0$ for $x \in (a, b)$ with $a < b$, then $\int_a^b f(x) \, \mathrm{d}x > 0$. Similarly, if $f(x) < 0$ for $x \in (a, b)$ with $a < b$, then $\int_a^b f(x) \, \mathrm{d}x < 0$.

We claim that the integral is maximised only for $(a, b) = (-20, 19)$, in which case the integral is positive. The proof is divided into the following three cases.

- If $a \geq 19$ and $a \leq b$, then $\int_a^b f(x) \, \mathrm{d}x < 0$, since $f(x) < 0$ for $x \in (a, b)$.
- If $b \leq -20$ and $a \leq b$, then $\int_a^b f(x) \, \mathrm{d}x < 0$, since $f(x) < 0$ for $x \in (a, b)$.
- Otherwise, we have $a < 19$ and $b > -20$ with $a \leq b$. In this case, we have

$$\begin{aligned}
\int_{-20}^{19} f(x) \, \mathrm{d}x - \int_a^b f(x) \, \mathrm{d}x &= [F(19) - F(-20)] - [F(b) - F(a)] \\
&= [F(19) - F(b)] + [F(a) - F(-20)] \\
&= \int_b^{19} f(x) \, \mathrm{d}x + \int_{-20}^a f(x) \, \mathrm{d}x \\
&\geq 0.
\end{aligned}$$

  The inequality holds since the first integral in the penultimate line is positive for all $-20 < b < 19$, as well as for $b > 19$. Similarly, the second integral in the penultimate line is positive for all $a < -20$, as well as for $-20 < a < 19$. Therefore,

$$\int_{-20}^{19} f(x) \, \mathrm{d}x \geq \int_a^b f(x) \, \mathrm{d}x,$$

  with equality if and only if $(a, b) = (-20, 19)$. This proves our claim.

Problem B2

For each odd prime p, prove that the integer

$$1! + 2! + 3! + \cdots + p! - \left\lfloor \frac{(p-1)!}{e} \right\rfloor$$

is divisible by p.

(Here, e denotes the base of the natural logarithm and $\lfloor x \rfloor$ denotes the largest integer that is less than or equal to x.)

Solution

The Taylor series expansion

$$e^x = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

has infinite radius of convergence, so we can write

$$\frac{1}{e} = e^{-1} = \frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots.$$

Multiply both sides of this equation by $(p-1)!$ and use the fact that p is odd to obtain

$$\frac{(p-1)!}{e} = \frac{(p-1)!}{0!} - \frac{(p-1)!}{1!} + \frac{(p-1)!}{2!} - \frac{(p-1)!}{3!} + \cdots + \frac{(p-1)!}{(p-1)!} - \frac{1}{p} + \frac{1}{p(p+1)} - \cdots.$$

The sequence $\frac{1}{p}, \frac{1}{p(p+1)}, \frac{1}{p(p+1)(p+2)}, \ldots$ is decreasing and converges to 0, so the alternating series

$$\frac{1}{p} - \frac{1}{p(p+1)} + \frac{1}{p(p+1)(p+2)} - \cdots$$

converges to a number between $\frac{1}{p} - \frac{1}{p(p+1)} = \frac{1}{p+1}$ and $\frac{1}{p}$. In particular, it converges to a number between 0 and 1. Therefore, we have

$$\left\lfloor \frac{(p-1)!}{e} \right\rfloor = \frac{(p-1)!}{0!} - \frac{(p-1)!}{1!} + \frac{(p-1)!}{2!} - \frac{(p-1)!}{3!} + \cdots - \frac{(p-1)!}{(p-2)!}.$$

Now observe that for all non-negative integers n,

$$n! \equiv (1-p)(2-p)(3-p)\cdots(n-p) \equiv (-1)^n (p-1)(p-2)\cdots(p-n) \pmod{p}.$$

By substituting these congruences into the denominators of the previous formula, we obtain

$$\begin{aligned} \left\lfloor \frac{(p-1)!}{e} \right\rfloor &\equiv \frac{(p-1)!}{1} + \frac{(p-1)!}{p-1} + \frac{(p-1)!}{(p-1)(p-2)} + \cdots + \frac{(p-1)!}{(p-1)(p-2)\cdots 2} \\ &\equiv (p-1)! + (p-2)! + (p-3)! + \cdots + 1! \pmod{p}. \end{aligned}$$

The cancellations required to pass from the first to the second line are valid since p is prime. We conclude the proof with the observation that $p! \equiv 0 \pmod{p}$.

Problem B3

Let G be a finite simple graph and let k be the largest number of vertices of any clique in G. Suppose that we label each vertex of G with a non-negative real number, so that the sum of all such labels is 1. Define the *value of an edge* to be the product of the labels of the two vertices at its ends. Define the *value of a labelling* to be the sum of the values of the edges.

Prove that the maximum possible value of a labelling of G is $\frac{k-1}{2k}$.

(A *finite simple graph* is a graph with finitely many vertices, in which each edge connects two distinct vertices and no two edges connect the same two vertices. A *clique* in a graph is a set of vertices in which any two are connected by an edge.)

Solution

First, we show that there exists a labelling of G whose value is $\frac{k-1}{2k}$. Choose any largest clique and label each vertex in the clique with the number $\frac{1}{k}$, and label all other vertices with the number 0. Then the value of each edge in the clique is $\frac{1}{k^2}$, and the value of all other edges is 0. Since there are $\frac{k(k-1)}{2}$ edges in the clique, the value of the labelling is $\frac{k(k-1)}{2} \times \frac{1}{k^2} = \frac{k-1}{2k}$.

Next, we consider an arbitrary labelling of G and show that its value is at most $\frac{k-1}{2k}$. Suppose that there are vertices u and v not joined by an edge, such that both are labelled with positive numbers. Let a and b be the labels of the vertices u and v, respectively. Without loss of generality, let us assume that the sum of the labels of the vertices adjacent to u is greater than or equal to the sum of the labels of the vertices adjacent to v. It follows that changing the labels of u and v to $a + b$ and 0, respectively, does not decrease the value of the labelling.

One can repeat the procedure described in the previous paragraph until any two vertices with positive labels are joined by an edge. Note that the process must terminate, since the number of vertices labelled with 0 increases at each step. The value of our original labelling of G is at most the value of the new labelling obtained as a result of this process.

Now let t be the number of vertices with positive labels at this stage and let the labels be $a_1, a_2, \ldots, a_t$. Since the t vertices form a clique, we must have $t \leq k$. The value of the labelling is

$$\sum_{1 \leq i < j \leq t} a_i a_j = \frac{1}{2}(a_1 + a_2 + \cdots + a_t)^2 - \frac{1}{2}(a_1^2 + a_2^2 + \cdots + a_t^2).$$

By the definition of a labelling, we have $a_1 + a_2 + \cdots + a_t = 1$. By the quadratic mean–arithmetic mean inequality (or the Cauchy–Schwarz inequality), we have

$$a_1^2 + a_2^2 + \cdots + a_t^2 \geq \frac{1}{t}(a_1 + a_2 + \cdots + a_t)^2 = \frac{1}{t}.$$

Hence, the value of the labelling satisfies

$$\frac{1}{2}(a_1 + a_2 + \cdots + a_t)^2 - \frac{1}{2}(a_1^2 + a_2^2 + \cdots + a_t^2) \leq \frac{1}{2} - \frac{1}{2t} \leq \frac{1}{2} - \frac{1}{2k} = \frac{k-1}{2k}.$$

It follows that the value of any labelling of G is at most $\frac{k-1}{2k}$.

Problem B4

The following problem is open in the sense that no solution is currently known to part (b). A proof of part (a) will be awarded 3 points.

A *binary string* is a sequence, each of whose terms is 0 or 1. A set $\mathcal{B}$ of binary strings is defined inductively according to the following rules.

- The binary string 1 is in $\mathcal{B}$.
- If $s_1, s_2, \ldots, s_n$ is in $\mathcal{B}$ with n odd, then both $s_1, s_2, \ldots, s_n, 0$ and $0, s_1, s_2, \ldots, s_n$ are in $\mathcal{B}$.
- If $s_1, s_2, \ldots, s_n$ is in $\mathcal{B}$ with n even, then both $s_1, s_2, \ldots, s_n, 1$ and $1, s_1, s_2, \ldots, s_n$ are in $\mathcal{B}$.
- No other binary strings are in $\mathcal{B}$.

For each positive integer n, let b_n be the number of binary strings in $\mathcal{B}$ of length n.

(a) Prove that there exist constants $c_1, c_2 > 0$ and $1.6 < \lambda_1, \lambda_2 < 1.9$ such that $c_1\lambda_1^n < b_n < c_2\lambda_2^n$ for all positive integers n.

(b) Determine $\liminf\limits_{n\to\infty} \sqrt[n]{b_n}$ and $\limsup\limits_{n\to\infty} \sqrt[n]{b_n}$.

Solution to part (a)

Each string in the set $\mathcal{B}$ is constructed by starting with the string 1 and alternately concatenating 0s and 1s, on either end of the string. We refer to each such concatenation as a *move*. A move that adds a term on the left of the string is called a *left-move*; similarly, a move that adds a term on the right of the string is called a *right-move*. Let $\mathcal{B}_n$ denote the set of binary strings in $\mathcal{B}$ of length n.

Lower bound

We will prove that there exists a constant $c_1 > 0$ such that

$$b_n > c_1\phi^n,$$

where $\phi = \frac{1+\sqrt{5}}{2} > 1.6$ is the golden ratio.

Let $\mathcal{A}_n \subseteq \mathcal{B}_n$ be the set of binary strings in $\mathcal{B}$ of length n whose leftmost and rightmost terms differ. Let $a_n = |\mathcal{A}_n|$ so that we have the obvious inequality $b_n \geq a_n$. We claim that for all integers $n \geq 2$,

$$a_{n+1} \geq a_n + a_{n-1}.$$

The claim will follow from the fact that there exists an injective function $f : \mathcal{A}_n \cup \mathcal{A}_{n-1} \to \mathcal{A}_{n+1}$. Given a string $s \in \mathcal{A}_n$, there is a unique move that produces a string $s' \in \mathcal{A}_{n+1}$. In this case, we define $f(s) = s'$. In other words, for $0t1 \in \mathcal{A}_n$ and $1t0 \in \mathcal{A}_n$, we define

$$f(0t1) = \begin{cases} 00t1, & \text{if } n+1 \text{ is even,} \\ 0t11, & \text{if } n+1 \text{ is odd,} \end{cases} \qquad f(1t0) = \begin{cases} 1t00 & \text{if } n+1 \text{ is even,} \\ 11t0 & \text{if } n+1 \text{ is odd.} \end{cases}$$

On the other hand, for for $0t1 \in \mathcal{A}_{n-1}$ and $1t0 \in \mathcal{A}_{n-1}$, we define

$$f(0t1) = 10t10 \in \mathcal{A}_{n+1} \qquad \text{and} \qquad f(1t0) = 01t01 \in \mathcal{A}_{n+1}.$$

Note that the order in which the outer two terms are added is determined by the parity of n.

It is clear that f is injective on each of $\mathcal{A}_n$ and $\mathcal{A}_{n-1}$. Moreover, we have $f(\mathcal{A}_n) \cap f(\mathcal{A}_{n-1}) = \emptyset$, because elements of $f(\mathcal{A}_n)$ begin or end with a repeated term while elements of $f(\mathcal{A}_{n-1})$ do not. So the claim that $a_{n+1} \geq a_n + a_{n-1}$ follows.

We have $a_{n+1} \geq a_n + a_{n-1}$ for $n \geq 2$ and one can calculate the initial values $a_1 = 2$ and $a_2 = 3$. It then follows that the sequence $a_n \geq f_n$, where the sequence $f_1, f_2, f_3, \dots$ is defined by $f_1 = 2$, $f_2 = 3$ and $f_{n+1} = f_n + f_{n-1}$ for $n \geq 2$. By the standard theory of homogeneous linear recursions, we have $f_n = a\phi^n + b\overline{\phi}^n$, where $\phi = \frac{1+\sqrt{5}}{2}$ and $\bar{\phi} = \frac{1-\sqrt{5}}{2}$ are the roots of the characteristic polynomial $x^2 - x - 1$ of the recursion. Since the sequence $f_1, f_2, f_3, \dots$ is increasing and $-1 < \bar{\phi} < 0$, it must be the case that $a > 0$. So we have $b_n \geq a_n \geq f_n \geq a\phi^n + b$ and it follows that there exists $c_1 > 0$ such that $b_n > c_1\phi^n$.

Upper bound

We will prove that there exists a constant $c_2 > 0$ such that

$$b_n < c_2\psi^n,$$

where $\psi < 1.9$ is the unique real root of $p(x) = x^3 - x^2 - x - 1$. The stationary points of $p(x)$ occur when $p'(x) = 3x^2 - 2x - 1 = 0$, which has solutions $x = -\frac{1}{3}$ and $x = 1$. Since $p(-\frac{1}{3}) = -\frac{22}{27} < 0$ and $p(1) = -2 < 0$, the cubic $p(x)$ does indeed have a unique real root ψ. Since $p(1.8) = -0.208 < 0$ and $p(1.9) = 0.349 > 0$, we have $1.8 < \psi < 1.9$.

Let $\mathcal{D}_n$ be the set of binary strings on the alphabet $\{L, R\}$ of length n that do not contain $RRLL$ as a substring, and let $d_n = |\mathcal{D}_n|$. We claim that for all integers $n \geq 2$,

$$b_n \leq d_{n-1}.$$

Consider the function $g : \mathcal{D}_{n-1} \to \mathcal{B}_n$ defined as follows. Interpret an element of $\mathcal{D}_{n-1}$ as a sequence of left-moves and right-moves and apply these in order to the string 1.

We now apply the following simple observation. Given a string in $\mathcal{B}$, if we perform two left-moves followed by two right-moves, then we obtain the same result as if we perform two right-moves followed by two left-moves.

Suppose that a string in $\mathcal{B}_n$ is produced by a sequence of moves applied to the string 1 and record the types of these moves via a binary string on the alphabet $\{L, R\}$ in the natural way. By repeatedly replacing any occurrence of $RRLL$ in the string with $LLRR$, we arrive at a sequence of moves that produces s without ever performing two right-moves followed by two left-moves. It follows that the function $g : \mathcal{D}_{n-1} \to \mathcal{B}_n$ is surjective, so the claim that $b_n \leq d_{n-1}$ follows.

Now given $w \in \mathcal{D}_{n+1}$, deleting the final term necessarily gives an element of $\mathcal{D}_n$. Conversely, given $w \in \mathcal{D}_n$, the strings wL and wR both belong to $\mathcal{D}_{n+1}$, unless w ends in RRL, in which case wL does not belong to $\mathcal{D}_{n+1}$.

If w ends in RRL, then we may write $w = w'RRL$, where $w' \in \mathcal{D}_{n-3}$. Conversely, if $w' \in \mathcal{D}_{n-3}$, then $w = w'RRL \in \mathcal{D}_n$, so strings in $\mathcal{D}_n$ ending in RRL are in one-to-one correspondence with strings in $\mathcal{D}_{n-3}$. Putting these observations together proves that the sequence $d_1, d_2, d_3, \ldots$ satisfies

$$d_{n+1} = 2d_n - d_{n-3}.$$

By the standard theory of homogeneous linear recursions, we have

$$d_n = A_1\psi_1^n + A_2\psi_2^n + A_3\psi_3^n + A_4\psi_4^n,$$

where $\psi_1\psi_2, \psi_3, \psi_4$ are the roots of the characteristic polynomial

$$x^4 - 2x^3 + 1 = (x-1)(x^3 - x^2 - x - 1) = (x-1)\,p(x).$$

Let us assume without loss of generality that $\psi_1 = \psi$ is the unique real root of $p(x)$ and that $\psi_4 = 1$. Then ψ_2 and ψ_3 must be complex conjugates. Since we have $\psi_1\psi_2\psi_3\psi_4 = 1$, it follows that $|\psi_2| < 1$ and $|\psi_3| < 1$. Therefore, for some constant A, we have

$$b_n \leq b_{n+1} \leq d_n = A_1\psi_1^n + A_2\psi_2^n + A_3\psi_3^n + A_4\psi_4^n < A_1\psi^n + A.$$

It follows that there exists a constant $c_2 > 0$ such that $b_n < c_2\psi^n$.

