65th International Mathematical Olympiad

Bath, United Kingdom, 10th–22nd July 2024

PROBLEMS

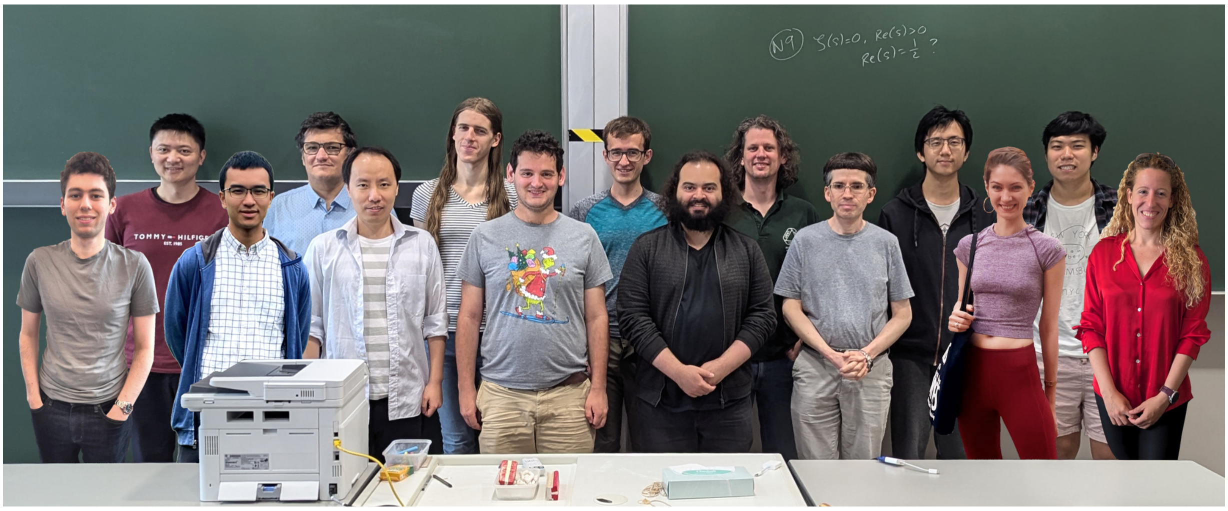
WITH SOLUTIONS

Contributing Countries

The Organising Committee and the Problem Selection Committee of IMO 2024 thank the following 63 countries for contributing 229 problem proposals:

> Algeria, Australia, Azerbaijan, Bangladesh, Belarus, Brazil, Bulgaria, Canada, China, Colombia, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Estonia, France, Georgia, Germany, Ghana, Greece, Hong Kong, India, Indonesia, Ireland, Iran, Israel, Japan, Kazakhstan, Kosovo, Latvia, Lithuania, Luxembourg, Malaysia, Mexico, Moldova, Netherlands, New Zealand, Norway, Peru, Poland, Portugal, Romania, Senegal, Serbia, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Syria, Taiwan, Thailand, Tunisia, Türkiye, Uganda, Ukraine, U.S.A., Uzbekistan.

Problem Selection Committee



Aron Thomas, Yinghua Ai, Andrew Ng, Géza Kós, Ivan Guo, Alice Carlotti, James Aaronson, Sam Bealing, Adrian Agisilaou, James Cranch, Joseph Myers (chair), Harvey Yau, Maria-Romina Ivan, Michael Ren, Elisa Lorenzo García

Problems

Day 1

Problem 1. Determine all real numbers α such that, for every positive integer n, the integer

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor$$

is a multiple of n. (Note that $\lfloor z \rfloor$ denotes the greatest integer less than or equal to z. For example, $\lfloor -\pi \rfloor = -4$ and $\lfloor 2 \rfloor = \lfloor 2.9 \rfloor = 2$.)

(*Colombia*)

Problem 2. Determine all pairs (a, b) of positive integers for which there exist positive integers g and N such that

$$\gcd(a^n + b,\, b^n + a) = g$$

holds for all integers $n \geqslant N$. (Note that $\gcd(x, y)$ denotes the greatest common divisor of integers x and y.)

(*Indonesia*)

Problem 3. Let $a_1, a_2, a_3, \ldots$ be an infinite sequence of positive integers, and let N be a positive integer. Suppose that, for each $n > N$, a_n is equal to the number of times a_{n-1} appears in the list $a_1, a_2, \ldots, a_{n-1}$.

Prove that at least one of the sequences $a_1, a_3, a_5, \ldots$ and $a_2, a_4, a_6, \ldots$ is eventually periodic.

(An infinite sequence $b_1, b_2, b_3, \ldots$ is *eventually periodic* if there exist positive integers p and M such that $b_{m+p} = b_m$ for all $m \geqslant M$.)

(*Australia*)

Day 2

Problem 4. Let ABC be a triangle with $AB < AC < BC$. Let the incentre and incircle of triangle ABC be I and ω, respectively. Let X be the point on line BC different from C such that the line through X parallel to AC is tangent to ω. Similarly, let Y be the point on line BC different from B such that the line through Y parallel to AB is tangent to ω. Let AI intersect the circumcircle of triangle ABC again at $P \neq A$. Let K and L be the midpoints of AC and AB, respectively.

Prove that $\angle KIL + \angle YPX = 180°$.

(Poland)

Problem 5. Turbo the snail plays a game on a board with 2024 rows and 2023 columns. There are hidden monsters in 2022 of the cells. Initially, Turbo does not know where any of the monsters are, but he knows that there is exactly one monster in each row except the first row and the last row, and that each column contains at most one monster.

Turbo makes a series of attempts to go from the first row to the last row. On each attempt, he chooses to start on any cell in the first row, then repeatedly moves to an adjacent cell sharing a common side. (He is allowed to return to a previously visited cell.) If he reaches a cell with a monster, his attempt ends and he is transported back to the first row to start a new attempt. The monsters do not move, and Turbo remembers whether or not each cell he has visited contains a monster. If he reaches any cell in the last row, his attempt ends and the game is over.

Determine the minimum value of n for which Turbo has a strategy that guarantees reaching the last row on the n^{th} attempt or earlier, regardless of the locations of the monsters.

(Hong Kong)

Problem 6. Let $\mathbb{Q}$ be the set of rational numbers. A function $f\colon \mathbb{Q} \to \mathbb{Q}$ is called *aquaesulian* if the following property holds: for every $x, y \in \mathbb{Q}$,

$$f(x + f(y)) = f(x) + y \qquad \text{or} \qquad f(f(x) + y) = x + f(y).$$

Show that there exists an integer c such that for any aquaesulian function f there are at most c different rational numbers of the form $f(r) + f(-r)$ for some rational number r, and find the smallest possible value of c.

(Japan)

Solutions

Day 1

Problem 1. Determine all real numbers α such that, for every positive integer n, the integer

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor$$

is a multiple of n. (Note that $\lfloor z \rfloor$ denotes the greatest integer less than or equal to z. For example, $\lfloor -\pi \rfloor = -4$ and $\lfloor 2 \rfloor = \lfloor 2.9 \rfloor = 2$.)

(Colombia)

Answer: All even integers satisfy the condition of the problem and no other real number α does so.

Solution 1. First we will show that even integers satisfy the condition. If $\alpha = 2m$ where m is an integer then

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor = 2m + 4m + \cdots + 2mn = mn(n+1)$$

which is a multiple of n.

Now we will show that they are the only real numbers satisfying the conditions of the problem. Let $\alpha = k + \epsilon$ where k is an integer and $0 \leqslant \epsilon < 1$. Then the number

$$\begin{aligned}\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor &= k + \lfloor \epsilon \rfloor + 2k + \lfloor 2\epsilon \rfloor + \cdots + nk + \lfloor n\epsilon \rfloor \\ &= \frac{kn(n+1)}{2} + \lfloor \epsilon \rfloor + \lfloor 2\epsilon \rfloor + \cdots + \lfloor n\epsilon \rfloor\end{aligned}$$

has to be a multiple of n. We consider two cases based on the parity of k.

Case 1: k is even.

Then $\frac{kn(n+1)}{2}$ is always a multiple of n. Thus

$$\lfloor \epsilon \rfloor + \lfloor 2\epsilon \rfloor + \cdots + \lfloor n\epsilon \rfloor$$

also has to be a multiple of n.

We will prove that $\lfloor n\epsilon \rfloor = 0$ for every positive integer n by strong induction. The base case $n = 1$ follows from the fact that $0 \leqslant \epsilon < 1$. Let us suppose that $\lfloor m\epsilon \rfloor = 0$ for every $1 \leqslant m < n$. Then the number

$$\lfloor \epsilon \rfloor + \lfloor 2\epsilon \rfloor + \cdots + \lfloor n\epsilon \rfloor = \lfloor n\epsilon \rfloor$$

has to be a multiple of n. As $0 \leqslant \epsilon < 1$ then $0 \leqslant n\epsilon < n$, which means that the number $\lfloor n\epsilon \rfloor$ has to be equal to 0.

The equality $\lfloor n\epsilon \rfloor = 0$ implies $0 \leqslant \epsilon < 1/n$. Since this has to happen for all n, we conclude that $\epsilon = 0$ and then α is an even integer.

Case 2: k is odd.

We will prove that $\lfloor n\epsilon\rfloor = n-1$ for every natural number n by strong induction. The base case $n=1$ again follows from the fact that $0 \leqslant \epsilon < 1$. Let us suppose that $\lfloor m\epsilon\rfloor = m-1$ for every $1 \leqslant m < n$. We need the number

$$\begin{aligned}\frac{kn(n+1)}{2} + \lfloor\epsilon\rfloor + \lfloor 2\epsilon\rfloor + \cdots + \lfloor n\epsilon\rfloor &= \frac{kn(n+1)}{2} + 0 + 1 + \cdots + (n-2) + \lfloor n\epsilon\rfloor \\ &= \frac{kn(n+1)}{2} + \frac{(n-2)(n-1)}{2} + \lfloor n\epsilon\rfloor \\ &= \frac{k+1}{2}n^2 + \frac{k-3}{2}n + 1 + \lfloor n\epsilon\rfloor\end{aligned}$$

to be a multiple of n. As k is odd, we need $1+\lfloor n\epsilon\rfloor$ to be a multiple of n. Again, as $0 \leqslant \epsilon < 1$ then $0 \leqslant n\epsilon < n$, so $\lfloor n\epsilon\rfloor = n-1$ as we wanted.

This implies that $1 - \frac{1}{n} \leqslant \epsilon < 1$ for all n which is absurd. So there are no other solutions in this case.

Solution 2. As in Solution 1 we check that for even integers the condition is satisfied. Then, without loss of generality we can assume $0 \leqslant \alpha < 2$. We set $S_n = \lfloor\alpha\rfloor + \lfloor 2\alpha\rfloor + \cdots + \lfloor n\alpha\rfloor$.

Notice that

$$S_n \equiv 0 \pmod{n} \qquad (1)$$

$$S_n \equiv S_n - S_{n-1} = \lfloor n\alpha\rfloor \pmod{n-1} \qquad (2)$$

Since $\gcd(n, n-1) = 1$, (1) and (2) imply that

$$S_n \equiv n\lfloor n\alpha\rfloor \pmod{n(n-1)}. \qquad (3)$$

In addition,

$$0 \leqslant n\lfloor n\alpha\rfloor - S_n = \sum_{k=1}^{n}\Big(\lfloor n\alpha\rfloor - \lfloor k\alpha\rfloor\Big) < \sum_{k=1}^{n}\Big(n\alpha - k\alpha + 1\Big) = \frac{n(n-1)}{2}\alpha + n. \qquad (4)$$

For n large enough, the RHS of (4) is less than $n(n-1)$. Then (3) forces

$$0 = S_n - n\lfloor n\alpha\rfloor = \sum_{k=1}^{n}\Big(\lfloor n\alpha\rfloor - \lfloor k\alpha\rfloor\Big) \qquad (5)$$

for n large enough.

Since $\lfloor n\alpha\rfloor - \lfloor k\alpha\rfloor \geqslant 0$ for $1 \leqslant k \leqslant n$, we get from (5) that, for all n large enough, all these inequalities are equalities. In particular $\lfloor\alpha\rfloor = \lfloor n\alpha\rfloor$ for all n large enough, which is absurd unless $\alpha = 0$.

Comment. An alternative ending to the previous solution is as follows.

By definition we have $S_n \leqslant \alpha\frac{n(n+1)}{2}$, on the other hand (5) implies $S_n \geqslant \alpha n^2 - n$ for all n large enough, so $\alpha = 0$.

Solution 3. As in other solutions, without loss of generality we may assume that $0 \leqslant \alpha < 2$. Even integers satisfy the condition, so we assume $0 < \alpha < 2$ and we will derive a contradiction.

By induction on n, we will simultaneously show that

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor = n^2, \tag{6}$$

$$\text{and} \qquad \frac{2n-1}{n} \leqslant \alpha < 2. \tag{7}$$

The base case is $n = 1$: If $\alpha < 1$, consider $m = \left\lceil \frac{1}{\alpha} \right\rceil > 1$, then

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor m\alpha \rfloor = 1$$

is not a multiple of m, so we deduce (7). Hence, $\lfloor \alpha \rfloor = 1$ and (6) follows.

For the induction step: assume the induction hypothesis to be true for n, then by (7)

$$2n + 1 - \frac{1}{n} \leqslant (n+1)\alpha < 2n + 2.$$

Hence,

$$n^2 + 2n \leqslant \lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor + \lfloor (n+1)\alpha \rfloor = n^2 + \lfloor (n+1)\alpha \rfloor < n^2 + 2n + 2.$$

So, necessarily $\lfloor (n+1)\alpha \rfloor = 2n + 1$ and

$$\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor + \lfloor (n+1)\alpha \rfloor = (n+1)^2$$

in order to obtain a multiple of $n + 1$. These two equalities give (6) and (7) respectively.

Finally, we notice that condition (7) being true for all n gives a contradiction.

Solution 4. As in other solutions without loss of generality we will assume that $0 < \alpha < 2$ and derive a contradiction. For each n, we define

$$b_n = \frac{\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor}{n},$$

which is a nonnegative integer by the problem condition and our assumption. Note that

$$\lfloor (n+1)\alpha \rfloor \geqslant \lfloor \alpha \rfloor, \lfloor 2\alpha \rfloor, \ldots, \lfloor n\alpha \rfloor \quad \text{and} \quad \lfloor (n+1)\alpha \rfloor > \lfloor \alpha \rfloor$$

for all $n > \frac{1}{\alpha}$. It follows that $b_{n+1} > b_n \implies b_{n+1} \geqslant b_n + 1$ for $n > \frac{1}{\alpha}$. Thus, for all such n,

$$b_n \geqslant n + C$$

where C is a fixed integer. On the other hand, the definition of b_n gives

$$b_n = \frac{\lfloor \alpha \rfloor + \lfloor 2\alpha \rfloor + \cdots + \lfloor n\alpha \rfloor}{n} \leqslant \frac{\alpha + 2\alpha + \cdots + n\alpha}{n} = \frac{\alpha}{2}(n+1),$$

which is a contradiction for sufficiently large n.

Solution 5. First consider the case in which $\alpha = \frac{p}{q}$ is a rational number with $p,\, q \in \mathbb{Z}$, $q > 0$ and $\gcd(p, q) = 1$. We decompose a real number $x = \lfloor x \rfloor + \{x\}$ as a sum of its integer and fractional parts.

Let us consider $n = kq$ with $k \in \mathbb{N}$, then

$$kq \mid S_{kq} = \sum_{k=1}^{kq} \lfloor k\alpha \rfloor = \sum_{k=1}^{kq} (k\alpha - \{k\alpha\}) = \frac{kq(kq+1)}{2}\alpha - k\frac{q(q-1)}{2q} = \frac{k}{2}(kqp + p - q + 1).$$

In particular, $2q \mid (kqp + p - q + 1)$ for all $k \in \mathbb{N}$. Thus, $2 \mid p$ and $2q \mid p - q + 1$. Hence, $p = q - 1 + 2qm$ for some $m \in \mathbb{Z}$. Now, $n = 2$ gives $S_2 = 2m + 4m + \left\lfloor \frac{2(q-1)}{q} \right\rfloor$, which is not a multiple of 2 unless $q = 1$. So the only rational numbers that can satisfy the condition of the problem are the even integers. As in previous solutions, we check that they do satisfy it.

Finally, for irrational α we compare S_{2n} and S_{2n+2} to get:

$$2 \mid S_{2n+2} - S_{2n} = \lfloor (2n+2)\alpha \rfloor + \lfloor (2n+1)\alpha \rfloor \equiv \lfloor (2n+2)\{\alpha\} \rfloor + \lfloor (2n+1)\{\alpha\} \rfloor + \lfloor \alpha \rfloor \pmod 2$$

for all $n \in \mathbb{N}$. But the Equidistribution Theorem (the sequence α, 2α, 3α, $\ldots$ mod 1 is uniformly distributed on $[0, 1)$ when α is an irrational number) implies that we can find an n such that the two numbers $\lfloor (2n+2)\{\alpha\} \rfloor$ and $\lfloor (2n+1)\{\alpha\} \rfloor$ have different parity (for even $\lfloor \alpha \rfloor$) or the same parity (for odd $\lfloor \alpha \rfloor$), so no irrational α satisfies the condition of the problem.

Problem 2. Determine all pairs (a, b) of positive integers for which there exist positive integers g and N such that

$$\gcd(a^n + b,\, b^n + a) = g$$

holds for all integers $n \geqslant N$. (Note that $\gcd(x, y)$ denotes the greatest common divisor of integers x and y.)

(Indonesia)

Answer: The only solution is $(a, b) = (1, 1)$.

Solution 1. It is clear that we may take $g = 2$ for $(a, b) = (1, 1)$. Supposing that (a, b) satisfies the conditions in the problem, let N be a positive integer such that $\gcd(a^n + b, b^n + a) = g$ for all $n \geqslant N$.

Lemma. We have that $g = \gcd(a, b)$ or $g = 2\gcd(a, b)$.

Proof. Note that both $a^N + b$ and $a^{N+1} + b$ are divisible by g. Hence

$$a(a^N + b) - (a^{N+1} + b) = ab - b = b(a - 1)$$

is divisible by g. Analogously, $a(b - 1)$ is divisible by g. Their difference $a - b$ is then divisible by g, so g also divides $a(b - 1) + a(a - b) = a^2 - a$. All powers of a are then congruent modulo g, so $a + b \equiv a^N + b \equiv 0 \pmod{g}$. Then $2a = (a + b) + (a - b)$ and $2b = (a + b) - (a - b)$ are both divisible by g, so $g \mid 2\gcd(a, b)$. On the other hand, it is clear that $\gcd(a, b) \mid g$, thus proving the Lemma. □

Let $d = \gcd(a, b)$, and write $a = dx$ and $b = dy$ for coprime positive integers x and y. We have that

$$\gcd\left((dx)^n + dy, (dy)^n + dx\right) = d \gcd\left(d^{n-1}x^n + y, d^{n-1}y^n + x\right),$$

so the Lemma tells us that

$$\gcd\left(d^{n-1}x^n + y, d^{n-1}y^n + x\right) \leqslant 2$$

for all $n \geqslant N$. Defining $K = d^2xy + 1$, note that K is coprime to each of d, x, and y. By Euler's theorem, for $n \equiv -1 \pmod{\varphi(K)}$ we have that

$$d^{n-1}x^n + y \equiv d^{-2}x^{-1} + y \equiv d^{-2}x^{-1}(1 + d^2xy) \equiv 0 \pmod{K},$$

so $K \mid d^{n-1}x^n + y$. Analogously, we have that $K \mid d^{n-1}y^n + x$. Taking such an n which also satisfies $n \geqslant N$ gives us that

$$K \mid \gcd(d^{n-1}x^n + y, d^{n-1}y^n + x) \leqslant 2.$$

This is only possible when $d = x = y = 1$, which yields the only solution $(a, b) = (1, 1)$.

Solution 2. After proving the Lemma, one can finish the solution as follows.

For any prime factor p of $ab + 1$, p is coprime to a and b. Take an $n \geqslant N$ such that $n \equiv -1 \pmod{p - 1}$. By Fermat's little theorem, we have that

$$a^n + b \equiv a^{-1} + b = a^{-1}(1 + ab) \equiv 0 \pmod{p},$$
$$b^n + a \equiv b^{-1} + a = b^{-1}(1 + ab) \equiv 0 \pmod{p},$$

then p divides g. By the Lemma, we have that $p \mid 2\gcd(a, b)$, and thus $p = 2$. Therefore, $ab + 1$ is a power of 2, and a and b are both odd numbers.

If $(a, b) \neq (1, 1)$, then $ab + 1$ is divisible by 4, hence $\{a, b\} = \{-1, 1\} \pmod{4}$. For odd $n \geqslant N$, we have that

$$a^n + b \equiv b^n + a \equiv (-1) + 1 = 0 \pmod{4},$$

then $4 \mid g$. But by the Lemma, we have that $\nu_2(g) \leqslant \nu_2(2\gcd(a, b)) = 1$, which is a contradiction. So the only solution to the problem is $(a, b) = (1, 1)$.

Solution 3. In fact, considering $a^n + b$ and $b^n + a$ modulo $ab + 1$ is sufficient for the solution, without the Lemma.

Let again $K = ab + 1$, which is coprime to both a and b. By Euler's theorem, for $n = k \cdot \varphi(K) - 1$ we have

$$a(a^n + b) = (a^k)^{\varphi(K)} + ab \equiv 1 + ab = K \equiv 0 \pmod{K},$$

so $K \mid a^n + b$ and similarly $K \mid b^n + a$. Hence K is a common divisor of $a^n + b$ and $b^n + a$. If n is sufficiently large, the greatest common divisor is supposed to be g, so $K \mid g$.

Now, for sufficiently large $n = k \cdot \varphi(K)$ we have $K \mid g \mid a^n + b$ and therefore

$$0 \equiv a^n + b = (a^k)^{\varphi(K)} + b \equiv 1 + b \pmod{K},$$

so $K = ab + 1 \mid b + 1$ and similarly $ab + 1 \mid a + 1$, which is possible only for $a = b = 1$,

Problem 3. Let $a_1, a_2, a_3, \dots$ be an infinite sequence of positive integers, and let N be a positive integer. Suppose that, for each $n > N$, a_n is equal to the number of times a_{n-1} appears in the list $a_1, a_2, \dots, a_{n-1}$.

Prove that at least one of the sequences $a_1, a_3, a_5, \dots$ and $a_2, a_4, a_6, \dots$ is eventually periodic.

(An infinite sequence $b_1, b_2, b_3, \dots$ is *eventually periodic* if there exist positive integers p and M such that $b_{m+p} = b_m$ for all $m \geqslant M$.)

(Australia)

Solution 1. Let $M > \max(a_1, \dots, a_N)$. We first prove that some integer appears infinitely many times. If not, then the sequence contains arbitrarily large integers. The first time each integer larger than M appears, it is followed by a 1. So 1 appears infinitely many times, which is a contradiction.

Now we prove that every integer $x \geqslant M$ appears at most $M - 1$ times. If not, consider the first time that any $x \geqslant M$ appears for the M^{th} time. Up to this point, each appearance of x is preceded by an integer which has appeared $x \geqslant M$ times. So there must have been at least M numbers that have already appeared at least M times before x does, which is a contradiction.

Thus there are only finitely many numbers that appear infinitely many times. Let the largest of these be k. Since k appears infinitely many times there must be infinitely many integers greater than M which appear at least k times in the sequence, so each integer $1, 2, \dots, k-1$ also appears infinitely many times. Since $k+1$ doesn't appear infinitely often there must only be finitely many numbers which appear more than k times. Let the largest such number be $l \geqslant k$. From here on we call an integer x *big* if $x > l$, *medium* if $l \geqslant x > k$ and *small* if $x \leqslant k$. To summarise, each small number appears infinitely many times in the sequence, while each big number appears at most k times in the sequence.

Choose a large enough $N' > N$ such that $a_{N'}$ is small, and in $a_1, \dots, a_{N'}$:

- every medium number has already made all of its appearances;
- every small number has made more than $\max(k, N)$ appearances.

Since every small number has appeared more than k times, past this point each small number must be followed by a big number. Also, by definition each big number appears at most k times, so it must be followed by a small number. Hence the sequence alternates between big and small numbers after $a_{N'}$.

Lemma 1. Let g be a big number that appears after $a_{N'}$. If g is followed by the small number h, then h equals the amount of small numbers which have appeared at least g times before that point.

Proof. By the definition of N', the small number immediately preceding g has appeared more than $\max(k, N)$ times, so $g > \max(k, N)$. And since $g > N$, the g^{th} appearance of every small number must occur after a_N and hence is followed by g. Since there are k small numbers and g appears at most k times, g must appear exactly k times, always following a small number after a_N. Hence on the h^{th} appearance of g, exactly h small numbers have appeared at least g times before that point. □

Denote by $a_{[i,j]}$ the subsequence $a_i, a_{i+1}, \dots, a_j$.

Lemma 2. Suppose that i and j satisfy the following conditions:

(a) $j > i > N' + 2$,

(b) a_i is small and $a_i = a_j$,

(c) no small value appears more than once in $a_{[i,j-1]}$.

Then a_{i-2} is equal to some small number in $a_{[i,j-1]}$.

Proof. Let $\mathcal{I}$ be the set of small numbers that appear at least a_{i-1} times in $a_{[1,i-1]}$. By Lemma 1, $a_i = |\mathcal{I}|$. Similarly, let $\mathcal{J}$ be the set of small numbers that appear at least a_{j-1} times in $a_{[1,j-1]}$. Then by Lemma 1, $a_j = |\mathcal{J}|$ and hence by (b), $|\mathcal{I}| = |\mathcal{J}|$. Also by definition, $a_{i-2} \in \mathcal{I}$ and $a_{j-2} \in \mathcal{J}$.

Suppose the small number a_{j-2} is not in $\mathcal{I}$. This means a_{j-2} has appeared less than a_{i-1} times in $a_{[1,i-1]}$. By (c), a_{j-2} has appeared at most a_{i-1} times in $a_{[1,j-1]}$, hence $a_{j-1} \leqslant a_{i-1}$. Combining with $a_{[1,i-1]} \subset a_{[1,j-1]}$, this implies $\mathcal{I} \subseteq \mathcal{J}$. But since $a_{j-2} \in \mathcal{J} \setminus \mathcal{I}$, this contradicts $|\mathcal{I}| = |\mathcal{J}|$. So $a_{j-2} \in \mathcal{I}$, which means it has appeared at least a_{i-1} times in $a_{[1,i-1]}$ and one more time in $a_{[i,j-1]}$. Therefore $a_{j-1} > a_{i-1}$.

By (c), any small number appearing at least a_{j-1} times in $a_{[1,j-1]}$ has also appeared $a_{j-1}-1 \geqslant a_{i-1}$ times in $a_{[1,i-1]}$. So $\mathcal{J} \subseteq \mathcal{I}$ and hence $\mathcal{I} = \mathcal{J}$. Therefore, $a_{i-2} \in \mathcal{J}$, so it must appear at least $a_{j-1} - a_{i-1} = 1$ more time in $a_{[i,j-1]}$. □

For each small number a_n with $n > N' + 2$, let p_n be the smallest number such that $a_{n+p_n} = a_i$ is also small for some i with $n \leqslant i < n + p_n$. In other words, $a_{n+p_n} = a_i$ is the first small number to occur twice after a_{n-1}. If $i > n$, Lemma 2 (with $j = n + p_n$) implies that a_{i-2} appears again before a_{n+p_n}, contradicting the minimality of p_n. So $i = n$. Lemma 2 also implies that $p_n \geqslant p_{n-2}$. So $p_n, p_{n+2}, p_{n+4}, \ldots$ is a nondecreasing sequence bounded above by $2k$ (as there are only k small numbers). Therefore, $p_n, p_{n+2}, p_{n+4}, \ldots$ is eventually constant and the subsequence of small numbers is eventually periodic with period at most k.

Note. Since every small number appears infinitely often, Solution 1 additionally proves that the sequence of small numbers has period k. The repeating part of the sequence of small numbers is thus a permutation of the integers from 1 to k. It can be shown that every permutation of the integers from 1 to k is attainable in this way.

Solution 2. We follow Solution 1 until after Lemma 1. For each $n > N'$ we keep track of how many times each of 1, 2, …, k has appeared in $a_1, \ldots, a_n$. We will record this information in an updating $(k+1)$-tuple

$$(b_1, b_2, \ldots, b_k; j)$$

where each b_i records the number of times i has appeared. The final element j of the $(k+1)$-tuple, also called the *active* element, represents the latest small number that has appeared in $a_1, \ldots, a_n$.

As n increases, the value of $(b_1, b_2, \ldots, b_k; j)$ is updated whenever a_n is small. The $(k+1)$-tuple updates deterministically based on its previous value. In particular, when $a_n = j$ is small, the active element is updated to j and we increment b_j by 1. The next big number is $a_{n+1} = b_j$. By Lemma 1, the next value of the active element, or the next small number a_{n+2}, is given by the number of b terms greater than or equal to the newly updated b_j, or

$$|\{i \mid 1 \leqslant i \leqslant k,\ b_i \geqslant b_j\}|. \tag{1}$$

Each sufficiently large integer which appears $i+1$ times must also appear i times, with both of these appearances occurring after the initial block of N. So there exists a global constant C such that $b_{i+1} - b_i \leqslant C$. Suppose that for some r, $b_{r+1} - b_r$ is unbounded from below. Since the value of $b_{r+1} - b_r$ changes by at most 1 when it is updated, there must be some update where $b_{r+1} - b_r$ decreases and $b_{r+1} - b_r < -(k-1)C$. Combining with the fact that $b_i - b_{i-1} \leqslant C$ for all i, we see that at this particular point, by the triangle inequality

$$\min(b_1, \ldots, b_r) > \max(b_{r+1}, \ldots, b_k). \tag{2}$$

Since $b_{r+1} - b_r$ just decreased, the new active element is r. From this point on, if the new active element is at most r, by (1) and (2), the next element to increase is once again from $b_1, \ldots, b_r$. Thus only $b_1, \ldots, b_r$ will increase from this point onwards, and b_k will no longer increase, contradicting the fact that k must appear infinitely often in the sequence. Therefore $|b_{r+1} - b_r|$ is bounded.

Since $|b_{r+1} - b_r|$ is bounded, it follows that each of $|b_i - b_1|$ is bounded for $i = 1, \ldots, k$. This means that there are only finitely many different states for $(b_1 - b_1, b_2 - b_1, \ldots, b_k - b_1; j)$. Since the next active element is completely determined by the relative sizes of $b_1, b_2, \ldots, b_k$ to each other, and the update of b terms depends on the active element, the active element must be eventually periodic. Therefore the small numbers subsequence, which is either $a_1, a_3, a_5, \ldots$ or $a_2, a_4, a_6, \ldots$, must be eventually periodic.

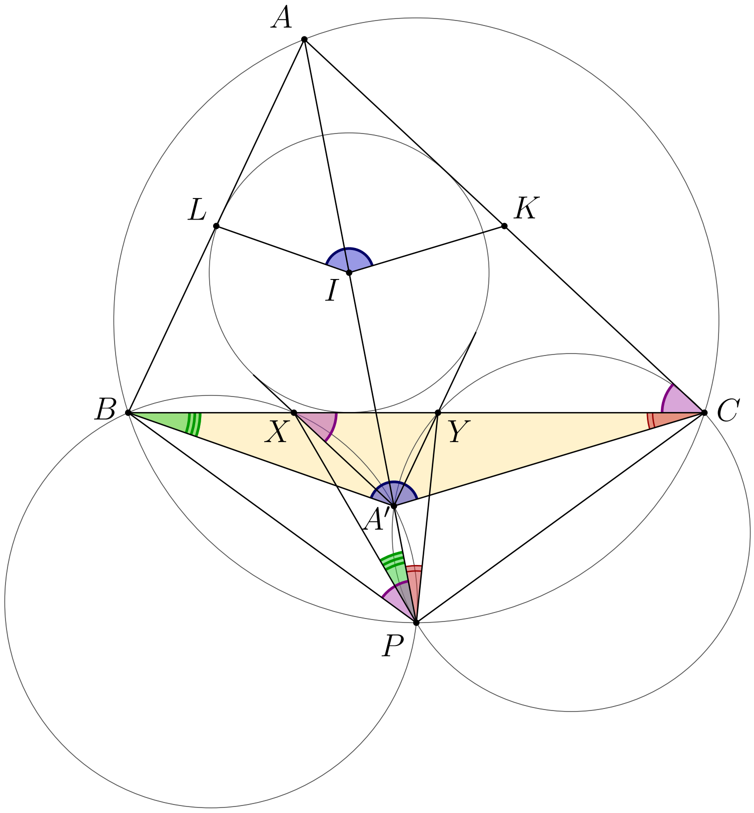
Day 2

Problem 4. Let ABC be a triangle with $AB < AC < BC$. Let the incentre and incircle of triangle ABC be I and ω, respectively. Let X be the point on line BC different from C such that the line through X parallel to AC is tangent to ω. Similarly, let Y be the point on line BC different from B such that the line through Y parallel to AB is tangent to ω. Let AI intersect the circumcircle of triangle ABC again at $P \neq A$. Let K and L be the midpoints of AC and AB, respectively.

Prove that $\angle KIL + \angle YPX = 180^\circ$.

(Poland)

Solution 1. Let A' be the reflection of A in I, then A' lies on the angle bisector AP. Lines $A'X$ and $A'Y$ are the reflections of AC and AB in I, respectively, and so they are the tangents to ω from X and Y. As is well-known, $PB = PC = PI$, and since $\angle BAP = \angle PAC > 30^\circ$, $PB = PC$ is greater than the circumradius. Hence $PI > \frac{1}{2}AP > AI$; we conclude that A' lies in the interior of segment AP.

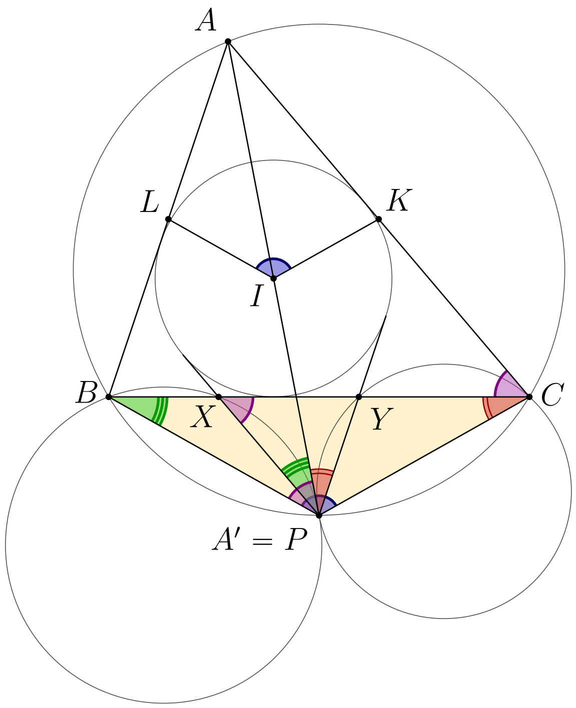


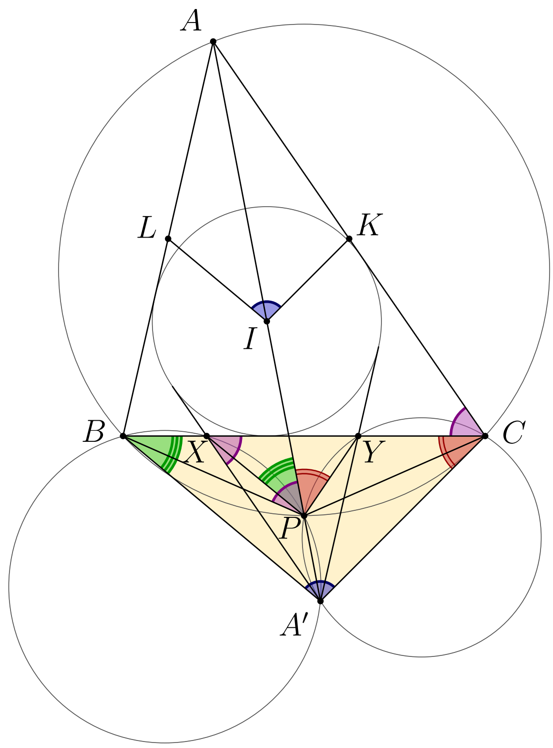
We have $\angle APB = \angle ACB$ in the circumcircle and $\angle ACB = \angle A'XC$ because $A'X \parallel AC$. Hence, $\angle APB = \angle A'XC$, and so quadrilateral $BPA'X$ is cyclic. Similarly, it follows that $CYA'P$ is cyclic.

Now we are ready to transform $\angle KIL + \angle YPX$ to the sum of angles in triangle $A'CB$. By a homothety of factor 2 at A we have $\angle KIL = \angle CA'B$. In circles $BPA'X$ and $CYA'P$ we have $\angle APX = \angle A'BC$ and $\angle YPA = \angle BCA'$, therefore

$$\angle KIL + \angle YPX = \angle CA'B + \big(\angle YPA + \angle APX\big) = \angle CA'B + \angle BCA' + \angle A'BC = 180^\circ.$$

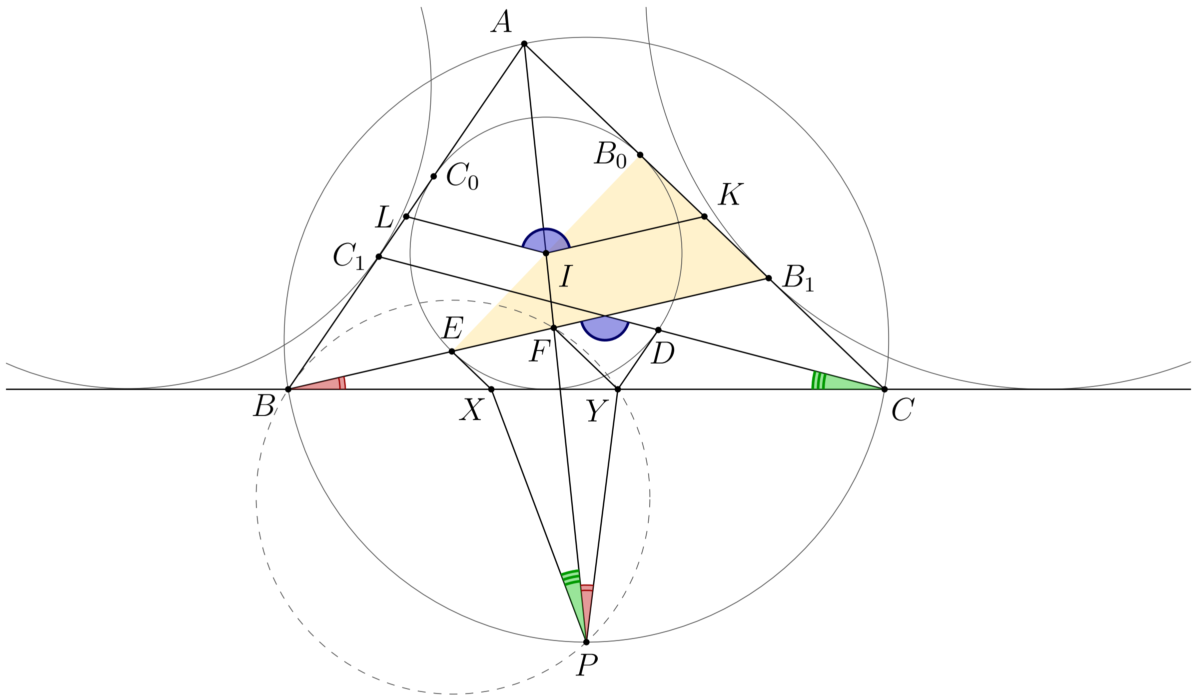
Comment. The constraint $AB < AC < BC$ was added by the Problem Selection Committee in order to reduce case-sensitivity. Without that, there would be two more possible configurations according to the possible orders of points A, P and A', as shown in the pictures below. The solution for these cases is broadly the same, but some extra care is required in the degenerate case when A' coincides with P and line AP is a common tangent to circles BPX and CPY.





Solution 2. Let $BC = a$, $AC = b$, $AB = c$ and $s = \frac{a+b+c}{2}$, and let the radii of the incircle, B-excircle and C-excircle be r, r_b and r_c, respectively. Let the incircle be tangent to AC and AB at B_0 and C_0, respectively; let the B-excircle be tangent to AC at B_1, and let the C-excircle be tangent to AB at C_1. As is well-known, $AB_1 = s - c$ and $\text{area}(\triangle ABC) = rs = r_c(s-c)$.

Let the line through X, parallel to AC be tangent to the incircle at E, and the line through Y, parallel to AB be tangent to the incircle at D. Finally, let AP meet BB_1 at F.



It is well-known that points B, E, and B_1 are collinear by the homothety between the incircle and the B-excircle, and $BE \parallel IK$ because IK is a midline in triangle B_0EB_1. Similarly, it follows that C, D, and C_1 are collinear and $CD \parallel IL$. Hence, the problem reduces to proving $\angle YPA = \angle CBE$ (and its symmetric counterpart $\angle APX = \angle DCB$ with respect to the vertex C), so it suffices to prove that $FYPB$ is cyclic. Since $ACPB$ is cyclic, that is equivalent to $FY \parallel B_1C$ and $\frac{BF}{FB_1} = \frac{BY}{YC}$.

By the angle bisector theorem we have

$$\frac{BF}{FB_1} = \frac{AB}{AB_1} = \frac{c}{s-c}.$$

The homothety at C that maps the incircle to the C-excircle sends Y to B, so

$$\frac{BC}{YC} = \frac{r_c}{r} = \frac{s}{s-c}.$$

So,

$$\frac{BY}{YC} = \frac{BC}{YC} - 1 = \frac{s}{s-c} - 1 = \frac{c}{s-c} = \frac{BF}{FB_1},$$

which completes the solution.

Solution 3. We claim triangles AIK and BYP are similar. Similarly it will follow that triangles AIL and CXP are similar so we would get

$$\angle KIL + \angle YPX = \angle KIA + \angle AIL + \angle YPX = \angle XYP + \angle PXY + \angle YPX = 180^\circ.$$

To prove this claim, first observe that

$$\angle IAK = \angle PAC = \angle PBC = \angle PBY.$$

By considering projections of points I and Y onto AB we get $AI = r/\sin(A/2)$ and $BY = 2r/\sin B$ (as the tangent from Y to ω distinct from BC is parallel to AB). Also applying extended sine rule we get,

$$\frac{AI}{AK} = \frac{r/\sin(A/2)}{R\sin B} \quad \text{and} \quad \frac{BY}{BP} = \frac{2r/\sin B}{2R\sin(A/2)}$$

which are equal. Combining these gives the similarity.

Solution 4. Let A' be the reflection of A in I. As in Solution 1, we show that the lines AB, AC, $A'X$, and $A'Y$ form a rhombus. Let T be the intersection of lines AI and BC and let $\mathcal{T}$ denote the negative homothety with centre D and factor α taking A to A'. We have that

$$\mathcal{T}(B) = \mathcal{T}(AB \cap BC) = \mathcal{T}(AB) \cap \mathcal{T}(BC) = A'Y \cap BC = Y,$$

and similarly $\mathcal{T}(C) = X$.

Let $A'' = \mathcal{T}(A')$ then

$$\angle XA''Y = \angle \mathcal{T}^{-1}(X)\mathcal{T}^{-1}(A'')\mathcal{T}^{-1}(Y) = \angle CA'B = \angle KIL$$

where the last step comes from a homothety of factor 2 at A. As A' lies on the opposite side of BC to A, A'' and P lie on opposite sides of BC. Thus, the above shows that $\angle KIL + \angle YPX = 180^\circ$ is equivalent to $YPXA''$ being cyclic. Computing powers at point T and using properties of the homothety we get

$$TA'' \cdot TP = T\mathcal{T}^2(A) \cdot TP = \alpha^2 TA \cdot TP = \alpha^2 TB \cdot TC = T\mathcal{T}(B) \cdot T\mathcal{T}(C) = TY \cdot TX$$

which gives the result.

Comment 1. Letting $\tilde{P} = \mathcal{T}^{-1}(P)$ and noting $\mathcal{T}^{-1}(YPXA'') = B\tilde{P}CA'$, we could also finish by showing $B\tilde{P}CA'$ is cyclic which follows from

$$TA' \cdot T\tilde{P} = T\mathcal{T}(A) \cdot T\mathcal{T}^{-1}(P) = TA \cdot TP = TB \cdot TC.$$

Comment 2. We can also use the above approach to show that $PA'XB$ is cyclic by

$$TX \cdot TB = T\mathcal{T}(C) \cdot TB = \alpha TC \cdot TB = \alpha TA \cdot TP = T\mathcal{T}(A) \cdot TP = TA' \cdot TP.$$

Solution 5. Let IX intersect CP at Q.

Claim. Triangle CXQ is similar to $\triangle AIB$ (and Q lies on circle BIC).

Proof. Since the tangent at X to ω, distinct from BC, is parallel to AC, and that XI bisects the two tangents from X,

$$\angle CXI = \frac{1}{2}(180^\circ - \angle ACB) = 90^\circ - \frac{1}{2}\angle ACB.$$

Using this

$$\angle CQX = \angle CXI - \angle BCP = 90^\circ - \frac{1}{2}\angle ACB - \angle BAP = 90^\circ - \frac{1}{2}\left(\angle ACB + \angle BAC\right) = \frac{1}{2}\angle CBA.$$

(Which is enough to show Q lies on circle BIC.) Combining this with $\angle XCQ = \angle BCP = \angle BAI$ gives the similarity. □

Note that P is the centre of circle BIC so P is the midpoint of CQ. This shows that P and L are similar points in the two triangles in the Claim since they are so

$$\angle PXY = \angle PXC = \angle AIL.$$

Similarly $\angle XYP = \angle KIA$ so

$$\angle KIL + \angle YPX = \angle KIA + \angle AIL + \angle YPX = \angle XYP + \angle PXY + \angle YPX = 180^\circ.$$

Solution 6. Let C_1 be where the C-excircle touches side AB. Let the C-mixtilinear incircle be tangent to side BC at X' and circle ABC at T_C.

Claim. $X = X'$.

Proof. From a well-known configuration, $\angle X'IC = 90^\circ$. As in Solution 5, we can show that $\angle CXI = 90^\circ - \frac{1}{2}\angle ACB$ so

$$\angle XIC = 180^\circ - \angle CXI - \angle ICX = 180^\circ - \left(90^\circ - \frac{1}{2}\angle ACB\right) - \frac{1}{2}\angle ACB = 90^\circ.$$

Hence $\angle XIC = \angle X'IC = 90^\circ$ so $X = X'$ as claimed. □

The homothety at T_C that sends the C-mixtilinear incircle to circle ABC sends X to P since the tangents at both points to the respective circles are parallel to BC. Hence P, X, and T_C are collinear.

An inversion centred at C with radius $\sqrt{CA \cdot CB}$ composed with a reflection in the bisector of $\angle ACB$ swaps C_1 and T_C so $\angle ACT_C = \angle C_1CB$. Hence,

$$\angle APX = \angle APT_C = \angle ACT_C = \angle C_1CB.$$

We can finish as in Solution 2.

Solution 7. Let N be the midpoint of BI.

Claim. PI and PB are tangent to circle BXI.

Proof. As in Solution 5, we can show that $\angle CXI = 90^\circ - \frac{1}{2}\angle ACB$ so

$$\angle BIX = \angle CXI - \angle XBI = 90^\circ - \frac{1}{2}\angle ACB - \frac{1}{2}CBA = \frac{1}{2}\angle BAC = \angle PAC = \angle PBX$$

which shows PB is tangent to circle BXI. P is the centre of circle BIC so $PB = PI$ so PI is also tangent to circle BXI. □

The Claim shows that PX is the X-symmedian in triangle BXI so

$$\angle IXN = 180^\circ - \angle BXP = \angle PXY.$$

Also, from the angle chase in the Claim, we get $\angle BAI = \angle BIX$. Combining this with $\angle XBI = \angle IBA$ we get $\triangle ABI \sim \triangle IBX$. L and N are similar points in these triangles so $\angle IXN = \angle AIL$. Therefore, $\angle PXY = \angle AIL$. Similarly, $\angle XYP = \angle KIA$ so

$$\angle KIL + \angle YPX = (\angle AIL + \angle KIA) + \angle YPX = \angle PXY + \angle XYP + \angle YPX = 180^\circ.$$

Problem 5. Turbo the snail plays a game on a board with 2024 rows and 2023 columns. There are hidden monsters in 2022 of the cells. Initially, Turbo does not know where any of the monsters are, but he knows that there is exactly one monster in each row except the first row and the last row, and that each column contains at most one monster.

Turbo makes a series of attempts to go from the first row to the last row. On each attempt, he chooses to start on any cell in the first row, then repeatedly moves to an adjacent cell sharing a common side. (He is allowed to return to a previously visited cell.) If he reaches a cell with a monster, his attempt ends and he is transported back to the first row to start a new attempt. The monsters do not move, and Turbo remembers whether or not each cell he has visited contains a monster. If he reaches any cell in the last row, his attempt ends and the game is over.

Determine the minimum value of n for which Turbo has a strategy that guarantees reaching the last row on the n^{th} attempt or earlier, regardless of the locations of the monsters.

(*Hong Kong*)

Comment. One of the main difficulties of solving this question is in determining the correct expression for n. Students may spend a long time attempting to prove bounds for the wrong value for n before finding better strategies.

Students may incorrectly assume that Turbo is not allowed to backtrack to squares he has already visited within a single attempt. Fortunately, making this assumption does not change the answer to the problem, though it may make it slightly harder to find a winning strategy.

Answer: The answer is $n = 3$.

Solution. First we demonstrate that there is no winning strategy if Turbo has 2 attempts.

Suppose that $(2, i)$ is the first cell in the second row that Turbo reaches on his first attempt. There can be a monster in this cell, in which case Turbo must return to the first row immediately, and he cannot have reached any other cells past the first row.

Next, suppose that $(3, j)$ is the first cell in the third row that Turbo reaches on his second attempt. Turbo must have moved to this cell from $(2, j)$, so we know $j \neq i$. So it is possible that there is a monster on $(3, j)$, in which case Turbo also fails on his second attempt. Therefore Turbo cannot guarantee to reach the last row in 2 attempts.

Next, we exhibit a strategy for $n = 3$. On the first attempt, Turbo travels along the path

$$(1,1) \to (2,1) \to (2,2) \to \cdots \to (2,2023).$$

This path meets every cell in the second row, so Turbo will find the monster in row 2 and his attempt will end.

If the monster in the second row is not on the edge of the board (that is, it is in cell $(2, i)$ with $2 \leqslant i \leqslant 2022$), then Turbo takes the following two paths in his second and third attempts:

$$(1,i-1) \to (2,i-1) \to (3,i-1) \to (3,i) \to (4,i) \to \cdots \to (2024,i).$$
$$(1,i+1) \to (2,i+1) \to (3,i+1) \to (3,i) \to (4,i) \to \cdots \to (2024,i).$$

The only cells that may contain monsters in either of these paths are $(3, i-1)$ and $(3, i+1)$. At most one of these can contain a monster, so at least one of the two paths will be successful.

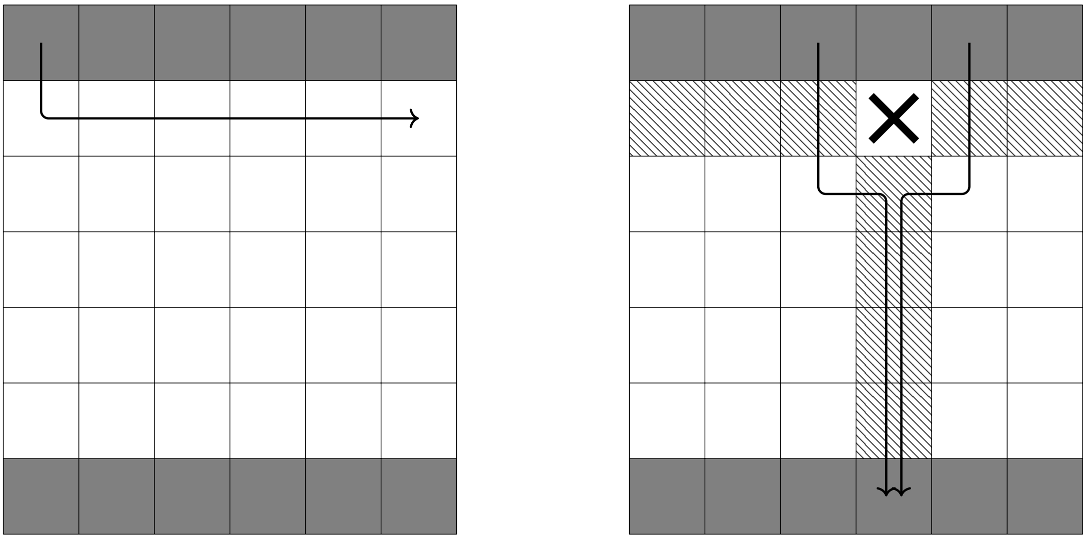


Figure 1: Turbo's first attempt, and his second and third attempts in the case where the monster on the second row is not on the edge. The cross indicates the location of a monster, and the shaded cells are cells guaranteed to not contain a monster.

If the monster in the second row is on the edge of the board, without loss of generality we may assume it is in $(2,1)$. Then, on the second attempt, Turbo takes the following path:

$$(1,2) \to (2,2) \to (2,3) \to (3,3) \to \cdots \to (2022,2023) \to (2023,2023) \to (2024,2023).$$

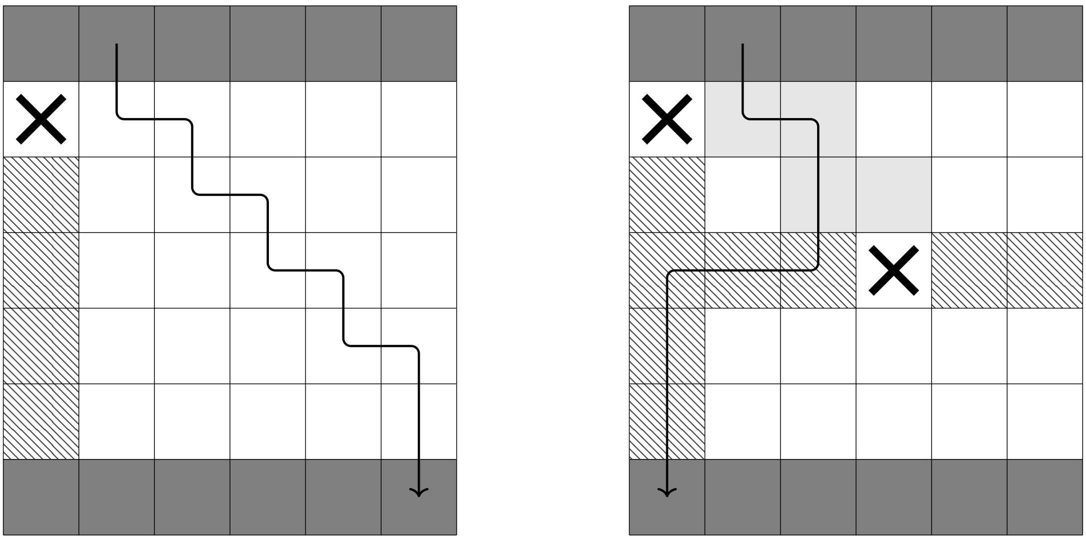


Figure 2: Turbo's second and third attempts in the case where the monster on the second row is on the edge. The light gray cells on the right diagram indicate cells that were visited on the previous attempt. Note that not all safe cells have been shaded.

If there are no monsters on this path, then Turbo wins. Otherwise, let (i,j) be the first cell on which Turbo encounters a monster. We have that $j=i$ or $j=i+1$. Then, on the third attempt, Turbo takes the following path:

$$\begin{aligned}(1,2) &\to (2,2) \to (2,3) \to (3,3) \to \cdots \to (i-2,i-1) \to (i-1,i-1)\\ &\to (i,i-1) \to (i,i-2) \to \cdots \to (i,2) \to (i,1)\\ &\to (i+1,1) \to \cdots \to (2023,1) \to (2024,1).\end{aligned}$$

Now note that

- The cells from $(1, 2)$ to $(i - 1, i - 1)$ do not contain monsters because they were reached earlier than (i, j) on the previous attempt.
- The cells (i, k) for $1 \leqslant k \leqslant i - 1$ do not contain monsters because there is only one monster in row i, and it lies in (i, i) or $(i, i + 1)$.
- The cells $(k, 1)$ for $i \leqslant k \leqslant 2024$ do not contain monsters because there is at most one monster in column 1, and it lies in $(2, 1)$.

Therefore Turbo will win on the third attempt.

Comment. A small variation on Turbo's strategy when the monster on the second row is on the edge is possible. On the second attempt, Turbo can instead take the path

$$\begin{aligned}(1, 2023) &\to (2, 2023) \to (2, 2022) \to \cdots \to (2, 3) \to (2, 2) \to (2, 3) \to \cdots \to (2, 2023)\\ &\to (3, 2023) \to (3, 2022) \to \cdots \to (3, 4) \to (3, 3) \to (3, 4) \to \cdots \to (3, 2023)\\ &\to \cdots\\ &\to (2022, 2023) \to (2022, 2022) \to (2022, 2023)\\ &\to (2023, 2023)\\ &\to (2024, 2023).\end{aligned}$$

If there is a monster on this path, say in cell (i, j), then on the third attempt Turbo can travel straight down to the cell just left of the monster instead of following the path traced out in the second attempt.

$$\begin{aligned}(1, j - 1) &\to (2, j - 1) \to \cdots \to (i - 1, j - 1) \to (i, j - 1)\\ &\to (i, j - 2) \to \cdots \to (i, 2) \to (i, 1)\\ &\to (i + 1, 1) \to \cdots \to (2023, 1) \to (2024, 1).\end{aligned}$$

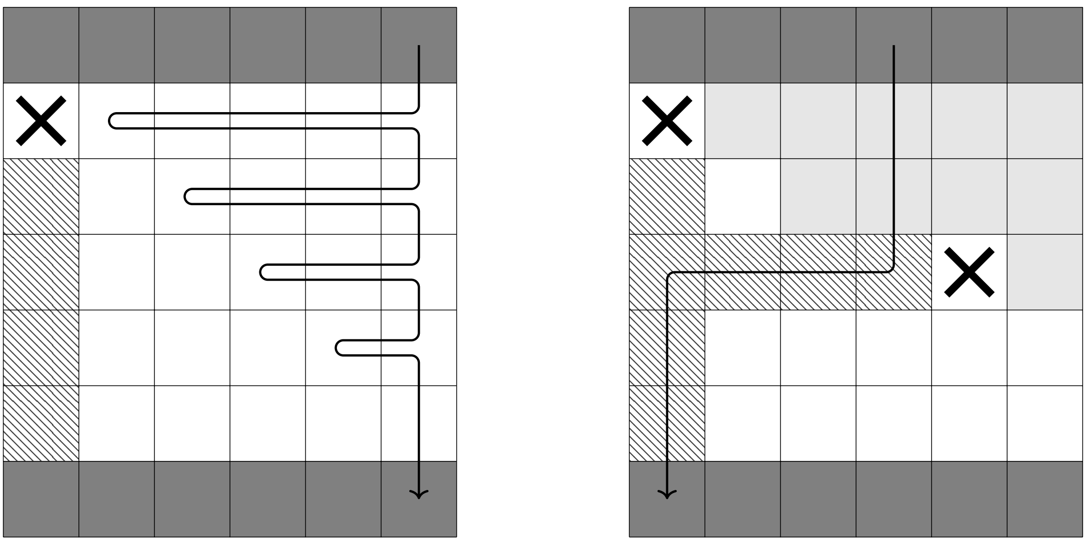


Figure 3: Alternative strategy for Turbo's second and third attempts.

Problem 6. Let $\mathbb{Q}$ be the set of rational numbers. A function $f\colon \mathbb{Q} \to \mathbb{Q}$ is called *aquaesulian* if the following property holds: for every $x, y \in \mathbb{Q}$,

$$f(x + f(y)) = f(x) + y \qquad \text{or} \qquad f(f(x) + y) = x + f(y).$$

Show that there exists an integer c such that for any aquaesulian function f there are at most c different rational numbers of the form $f(r) + f(-r)$ for some rational number r, and find the smallest possible value of c.

(Japan)

Answer: The smallest value is $c = 2$.

Common remarks. Suppose that f is a function satisfying the condition of the problem. We will use the following throughout all solutions.

- $a \sim b$ if either $f(a) = b$ or $f(b) = a$,
- $a \to b$ if $f(a) = b$,
- $P(x, y)$ to denote the proposition that either $f(x + f(y)) = f(x) + y$ or $f(f(x) + y) = x + f(y)$,
- $g(x) = f(x) + f(-x)$.

With this, the condition $P(x, y)$ could be rephrased as saying that $x + f(y) \sim f(x) + y$, and we are asked to determine the maximum possible number of elements of $\{g(x) \mid x \in \mathbb{Q}\}$.

Solution 1. We begin by providing an example of a function f for which there are two values of $g(x)$. We take the function $f(x) = \lfloor x \rfloor - \{x\}$, where $\lfloor x \rfloor$ denotes the floor of x (that is, the largest integer less than or equal to x) and $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x.

First, we show that f satisfies $P(x, y)$. Given $x, y \in \mathbb{Q}$, we have

$$\begin{aligned} f(x) + y &= \lfloor x \rfloor - \{x\} + \lfloor y \rfloor + \{y\} = (\lfloor x \rfloor + \lfloor y \rfloor) + (\{y\} - \{x\}); \\ x + f(y) &= \lfloor x \rfloor + \{x\} + \lfloor y \rfloor - \{y\} = (\lfloor x \rfloor + \lfloor y \rfloor) + (\{x\} - \{y\}). \end{aligned}$$

If $\{x\} < \{y\}$, then we have that the fractional part of $f(x) + y$ is $\{y\} - \{x\}$ and the floor is $\lfloor x \rfloor + \lfloor y \rfloor$, so $f(x) + y \to x + f(y)$. Likewise, if $\{x\} > \{y\}$, then $x + f(y) \to f(x) + y$. Finally, if $\{x\} = \{y\}$, then $f(x) + y = x + f(y) = \lfloor x \rfloor + \lfloor y \rfloor$ is an integer. In all cases, the relation P is satisfied.

Finally, we observe that if x is an integer then $g(x) = 0$, and if x is not an integer then $g(x) = -2$, so there are two values for $g(x)$ as required.

Now, we prove that there cannot be more than two values of $g(x)$. $P(x, x)$ tells us that $x + f(x) \sim x + f(x)$, or in other words, for all x,

$$f(x + f(x)) = x + f(x). \tag{1}$$

We begin with the following lemma.

Lemma 1. f is a bijection, and satisfies

$$f(-f(-x)) = x. \tag{2}$$

Proof. We first prove that f is injective. Suppose that $f(x_1) = f(x_2)$; then $P(x_1, x_2)$ tells us that $f(x_1)+x_2 \sim f(x_2)+x_1$. Without loss of generality, suppose that $f(x_1)+x_2 \to f(x_2)+x_1$.

But $f(x_1) = f(x_2)$, so $f(f(x_1) + x_2) = f(f(x_2) + x_2) = f(x_2) + x_2$ by (1). Therefore, $f(x_2) + x_1 = f(x_2) + x_2$, as required.

Now, (1) with $x = 0$ tells us that $f(f(0)) = f(0)$ and so by injectivity $f(0) = 0$.

Applying $P(x, -f(x))$ tells us that $0 \sim x + f(-f(x))$, so either $0 = f(0) = x + f(-f(x))$ or $f(x + f(-f(x))) = 0$ which implies that $x + f(-f(x)) = 0$ by injectivity. Either way, we deduce that $x = -f(-f(x))$, or $x = f(-f(-x))$ by replacing x with $-x$.

Finally, note that bijectivity follows immediately from (2). □

Since f is bijective, it has an inverse, which we denote f^{-1}. Rearranging (2) (after replacing x with $-x$) gives that $f(-x) = -f^{-1}(x)$. We have $g(x) = f(x) + f(-x) = f(x) - f^{-1}(x)$.

Suppose $g(x) = u$ and $g(y) = v$, where $u \neq v$ are both nonzero. Define $x' = f^{-1}(x)$ and $y' = f^{-1}(y)$; by definition, we have

$$x' \to x \to x' + u$$
$$y' \to y \to y' + v.$$

Putting in $P(x', y)$ gives $x+y \sim x'+y'+v$, and putting in $P(x, y')$ gives $x+y \sim x'+y'+u$. These are not equal since $u \neq v$, and $x + y$ may have only one incoming and outgoing arrow because f is a bijection, so we must have either $x' + y' + u \to x + y \to x' + y' + v$ or the same with the arrows reversed. Swapping (x, u) and (y, v) if necessary, we may assume without loss of generality that this is the correct direction for the arrows.

Also, we have $-x' - u \to -x \to -x'$ by Lemma 1. Putting in $P(x + y, -x' - u)$ gives $y \sim y' + v - u$, and so $y' + v - u$ must be either $y' + v$ or y'. This means u must be either 0 or v, and this contradicts our assumption about u and v.

Comment. Lemma 1 can also be proven as follows. We start by proving that f must be surjective. Suppose not; then, there must be some t which does not appear in the output of f. $P(x, t - f(x))$ tells us that $t \sim x + f(t - f(x))$, and so by assumption $f(t) = x + f(t - f(x))$ for all x. But setting $x = f(t) - t$ gives $t = f(t - f(f(t) - t))$, contradicting our assumption about t.

Now, choose some t such that $f(t) = 0$; such a t must exist by surjectivity. $P(t, t)$ tells us that $f(t) = t$, or in other words $t = 0$ and $f(0) = 0$. The remainder of the proof is the same as the proof given in Solution 1.

Solution 2. We again start with Lemma 1, and note $f(0) = 0$ as in the proof of that lemma.

$P(x, -f(y))$ gives $x+f(-f(y)) \sim f(x)-f(y)$, and using (2) this becomes $x-y \sim f(x)-f(y)$. In other words, either $f(x - y) = f(x) - f(y)$ or $x - y = f(f(x) - f(y))$. In the latter case, we deduce that

$$\begin{aligned} f(-(x-y)) &= f(-f(f(x)-f(y))) \\ f(y-x) &= f(-f(f(x)-f(y))) \\ &= f(y) - f(x). \end{aligned}$$

Thus, $f(y) - f(x)$ is equal to either $f(y - x)$ or $-f(x - y)$. Replacing y with $x + d$, we deduce that $f(x + d) - f(x) \in \{f(d), -f(-d)\}$.

Now, we prove the following claim.

Claim 1. For any $n \in \mathbb{Z}_{>0}$ and $d \in \mathbb{Q}$, we have that either $g(d) = 0$ or $g(d) = \pm g(d/n)$.

In particular, if $g(d/n) = 0$ then $g(d) = 0$.

Proof. We first prove that if $g(d/n) = 0$ then $g(d) = 0$. Suppose that $g(d/n) = 0$. Then $f(d/n) = -f(-d/n)$ and so $f(x + d/n) - f(x) = f(d/n)$ for any x. Applying this repeatedly, we deduce that $f(x + d) - f(x) = nf(d/n)$ for any x. Applying this with $x = 0$ and $x = -d$ and adding gives $f(d) + f(-d) = 0$, so $g(d) = 0$, and in particular the claim is true whenever $g(d) = 0$.

Now, select $n \in \mathbb{Z}_{>0}$ and $d \in \mathbb{Q}$ such that $g(d) \neq 0$, and observe that we must have $g(d/n) \neq 0$. Observe that for any $k \in \mathbb{Z}$ we have that $f(kd/n) - f((k-1)d/n) \in \{f(d/n), -f(-d/n)\}$. Let A_i be the number of $k \in \mathbb{Z}$ with $i - n < k \leqslant i$ such that this difference equals $f(d/n)$.

We deduce that for any $i \in \mathbb{Z}$,

$$\begin{aligned} f(id/n) - f(id/n - d) &= \sum_{i-n<k\leqslant i} f(kd/n) - f((k-1)d/n) \\ &= A_i f(d/n) - (n - A_i)f(-d/n) \\ &= -nf(-d/n) + A_i g(d/n). \end{aligned}$$

Since $g(d/n)$ is nonzero, this is a nonconstant linear function of A_i. However, there are only two possible values for $f(id/n) - f(id/n - d)$, so there must be at most two possible values for A_i as i varies. And since $A_{i+1} - A_i \in \{-1, 0, 1\}$, those two values must differ by 1 (if there are two values).

Now, we have

$$\begin{aligned} f(d) - f(0) &= -nf(-d/n) + A_n g(d/n), \qquad \text{and} \\ f(0) - f(-d) &= -nf(-d/n) + A_0 g(d/n). \end{aligned}$$

Subtracting these (using the fact that $f(0) = 0$) we obtain

$$\begin{aligned} f(d) + f(-d) &= (A_n - A_0)g(d/n) \\ &= \pm g(d/n), \end{aligned}$$

where the last line follows from the fact that $g(d)$ is nonzero. □

It immediately follows that there can only be one nonzero number of the form $g(x)$ up to sign; to see why, if $g(d)$ and $g(d')$ are both nonzero, then for some $n, n' \in \mathbb{Z}_{>0}$ we have $d/n = d'/n'$. But

$$g(d) = \pm g(d/n) = \pm g(d').$$

Finally, suppose that for some d, d' we have $g(d) = c$ and $g(d') = -c$ for some nonzero c. So we have

$$f(d) + f(-d) - f(d') - f(-d') = 2c$$

which rearranges to become $(f(d) - f(d')) - (f(-d') - f(-d)) = 2c$.

Each of the bracketed terms must be equal to either $f(d - d')$ or $-f(d' - d)$. However, they cannot be equal since c is nonzero, so $g(d - d') = f(d - d') + f(d' - d) = \pm 2c$. This contradicts the assertion that $g(-x) = \pm c$ for all x.

Comment. After establishing Claim 1, Solution 2 may also be finished as follows. We will prove that Claim 1 may be improved to $g(d) = 0$ or $g(d) = g(d/n)$, in other words we may exclude the case that $g(d) = -g(d/n) \neq 0$. This will imply that there can be at most one nonzero number of the form $g(x)$. Suppose that d and n are as in the hypothesis of Claim 1.

Let B_i be the number of $k \in \mathbb{Z}$ with $i - n + 1 < k \leqslant i$ such that $f(kd/n) - f((k-1)d/n)$ equals $f(d/n)$. As in the proof in Solution 2, it follows that for any $i \in \mathbb{Z}$,

$$f(id/n) - f(id/n - (n-1)d/n) = -(n-1)f(-d/n) + B_i g(d/n).$$

There are two possible values for $f(id/n) - f(id/n - (n-1)d/n)$ so there can be at most two values for B_i as i varies, which must differ by 1 (if there are two values). However, we have

$$\begin{aligned} f(d) - f(d/n) &= -(n-1)f(-d/n) + B_n g(d/n), \qquad \text{and} \\ f(-d/n) - f(-d) &= -(n-1)f(-d/n) + B_{-1} g(d/n). \end{aligned}$$

Subtracting these gives that $g(d) - g(d/n) = (B_n - B_{-1})g(d/n)$, and hence it is impossible to have $g(d/n) = -g(d) \neq 0$ as that would force $B_n - B_{-1} = -2$.

Solution 3. As in Solution 1, we start by establishing Lemma 1 as above, and write $f^{-1}(x) = -f(-x)$ for the inverse of f, and $g(x) = f(x) - f^{-1}(x)$.

We now prove the following.

Lemma 2. If $g(x) \neq g(y)$, then $g(x+y) = \pm(g(x) - g(y))$.

Proof. Assume x and y are such that $g(x) \neq g(y)$. Applying $P(x, f^{-1}(y))$ gives $x + y \sim f(x) + f^{-1}(y)$, and applying $P(f^{-1}(x), y)$ gives $x + y \sim f^{-1}(x) + f(y)$.

Observe that

$$\begin{aligned} (f(x) + f^{-1}(y)) - (f^{-1}(x) + f(y)) &= (f(x) - f^{-1}(x)) - (f(y) - f^{-1}(y)) \\ &= g(x) - g(y). \end{aligned}$$

By assumption, $g(x) \neq g(y)$, and so $f(x) + f^{-1}(y) \neq f^{-1}(x) + f(y)$. Since f is bijective, this means that these two values must be $f(x+y)$ and $f^{-1}(x+y)$ in some order, and so $g(x+y) = f(x+y) - f^{-1}(x+y)$ must be their difference up to sign, which is either $g(x) - g(y)$ or $g(y) - g(x)$. □

Claim. If x and q are rational numbers such that $g(q) = 0$ and n is an integer, then $g(x + nq) = g(x)$.

Proof. If $g(b) = 0$ and $g(a) \neq g(a+b)$, then the lemma tells us that $g(b) = \pm(g(a+b) - g(a))$, which contradicts our assumptions. Therefore, $g(a) = g(a+b)$ whenever $g(b) = 0$.

A simple induction then gives that $g(nb) = 0$ for any positive integer n, and $g(nb) = 0$ for negative n as $g(x) = g(-x)$. The claim follows immediately. □

Lemma 3. There cannot be both positive and negative elements in the range of g.

Proof. Suppose that $g(x) > 0$ and $g(y) < 0$. Let $\mathcal{S}$ be the set of numbers of the form $mx + ny$ for integers m, n. We first show that $g(\mathcal{S})$ has infinitely many elements. Indeed, suppose $g(\mathcal{S})$ is finite, and let $a \in \mathcal{S}$ maximise g and $b \in \mathcal{S}$ maximise $-g$. Then $a + b \in \mathcal{S}$, and $g(a+b) = g(a) - g(b)$ or $g(b) - g(a)$. In the first case $g(a+b) > g(a)$ and in the second case $g(a+b) < g(b)$; in either case we get a contradiction.

Now, we show that there must exist some nonzero rational number q with $g(q) = 0$. Indeed, suppose first that $a + f(a) = 0$ for all a. Then $g(a) = f(a) + f(-a) = 0$ for all a, and so g takes no nonzero value. Otherwise, there is some a with $a + f(a) \neq 0$, and so (1) yields that $f(q) = 0$ for $q = a + f(a) \neq 0$. Noting that $f(-q) = 0$ from Lemma 1 tells us that $g(q) = 0$, as required.

Now, there must exist integers s and s' such that $xs = qs'$ and integers t and t' such that $yt = qt'$. The claim above gives that the value of $g(mx + ny)$ depends only on the values of m mod s and n mod t, so $g(mx + ny)$ can only take finitely many values. □

Finally, suppose that $g(x) = u$ and $g(y) = v$ where $u \neq v$ have the same sign. Assume $u, v > 0$ (the other case is similar) and assume $u > v$ without loss of generality.

$P(f^{-1}(x), -y)$ gives $x - y \sim f^{-1}(x) - f^{-1}(y) = f(x) - f(y) - (u - v)$, and $P(x, -f(y))$ gives $x - y \sim f(x) - f(y)$. $u - v$ is nonzero, so $f(x-y)$ and $f^{-1}(x-y)$ must be $f(x) - f(y) - (u-v)$ and $f(x) - f(y)$ in some order, and since $g(x-y)$ must be nonnegative, we have

$$f(x) - f(y) - (u - v) \to x - y \to f(x) - f(y).$$

Then, $P(x - y, f^{-1}(y))$ tells us that $(x - y) + y \sim (f(x) - f(y)) + (f(y) - v)$, so $x \sim f(x) - v$, contradicting either $v \neq u$ or $v > 0$.

Comment. Lemma 2 also follows from $f(x+d) - f(x) \in \{f(d), -f(-d)\}$ as proven in Solution 2. Indeed, we also have $f(-x) - f(-x-d) \in \{f(d), -f(-d)\}$, and then subtracting the second from the first we get $g(x+d) - g(x) \in \{g(d), -g(d), 0\}$. Replacing $x+d$ and x with x and $-y$ gives the statement of Lemma 2.

Comment. It is possible to prove using Lemma 2 that g must have image of the form $\{0, c, 2c\}$ if it has size greater than 2. Indeed, if $g(x) = c$ and $g(y) = d$ with $0 < c < d$, then $g(x+y) = d - c$ as it must be nonnegative, and $g(y) = g((x+y) + (-x)) = |d - 2c|$ provided that $d \neq 2c$.

However, it is not possible to rule out $\{0, c, 2c\}$ based entirely on the conclusion of Lemma 2; indeed, the function given by

$$g(x) = \begin{cases} 0, & \text{if } x = 2n \text{ for } n \in \mathbb{Z}; \\ 2, & \text{if } x = 2n+1 \text{ for } n \in \mathbb{Z}; \\ 1, & \text{if } x \notin \mathbb{Z}. \end{cases}$$

satisfies the conclusion of Lemma 2 (even though there is no function f giving this choice of g).

Note. Solution 1 actually implies that the result also holds over $\mathbb{R}$. The proposal was originally submitted and evaluated over $\mathbb{Q}$ as it is presented here, and the Problem Selection Committee believes that this form is more suitable for the competition because it allows for more varied and interesting approaches once Lemma 1 has been established. Even the variant here defined over $\mathbb{Q}$ was found to be fairly challenging.

Solution 4. As in the other solutions we establish Lemma 1, and as in Solution 2 we deduce that $f(x) - f(y)$ equals either $f(x-y)$ or $-f(y-x)$.

From $f(x-y) + f(y-x) = g(x-y)$ we deduce that $f(x) - f(y) - f(x-y)$ equals either 0 or $-g(x-y)$, and so replacing x with $x+y$ gives

$$f(x+y) - f(x) - f(y) \in \{0, -g(y)\}. \tag{3}$$

Swapping x and y gives that if $g(x) \neq g(y)$ then we must have $f(x+y) = f(x) + f(y)$.

Then,

$$\begin{aligned} g(x+y) &= f(x+y) + f((-x) + (-y)) \\ &= f(x) + f(y) + f(-x) + f(-y) \\ &= g(x) + g(y), \end{aligned}$$

where we used that $g(x) = g(-x)$ in the second line.

If $g(x) \neq g(y)$ are both nonzero, then $g(x+y) = g(x) + g(y)$ is not equal to $g(-y)$, so $g(x) = g(x+y) + g(-y) = g(x) + 2g(y)$ which is a contradiction.

Solution 5. As in Solution 4, we establish Lemma 1 and (3).

From $P(x, f^{-1}(y))$ we deduce that $x + y \sim f(x) + f(y) - g(y)$, which implies that either $f(x+y) = f(x) + f(y) - g(y)$ or $f(x+y) - g(x+y) = f^{-1}(x+y) = f(x) + f(y) - g(y)$. Rearranging, we deduce that $f(x+y) - f(x) - f(y) \in \{-g(y), g(x+y) - g(y)\}$.

Suppose that $g(x) \neq g(y)$ are both nonzero. As in Solution 4, we deduce from (3) that $f(x+y) - f(x) - f(y) = 0$. Since $g(y) \neq 0$, we must in fact have that $g(x+y) = g(y)$, and by symmetry we also have that $g(x+y) = g(x)$. This contradicts $g(x) \neq g(y)$.

Solution 6. As in the other solutions we establish Lemma 1, and as in Solution 2 we deduce that $f(x) - f(y)$ equals either $f(x-y)$ or $-f(y-x)$.

Lemma 4. For any x and y, at least one of the following equalities holds.

$$g(f(x)) = g(y)$$
$$g(f(x)) = 0$$
$$g(f(y)) = g(x)$$
$$g(f(y)) = 0$$

Proof. If x and y satisfy that $f(f(x)+y) = x + f(y)$, then $f(f(x)) - f(f(x)+y)$ equals either $f(-y)$ or $-f(y)$, so either $f(f(x)) + f(y) = x + f(y)$ or $f(f(x)) - f(-y) = x + f(y)$. So $f(f(x)) - x$ equals 0 or $g(y)$; in particular, $g(f(x))$ equals 0 or $g(y)$.

Otherwise, $f(x) + y = f(x + f(y))$, in which case $g(f(y))$ equals 0 or $g(x)$. □

Applying Lemma 4 with $y = x$ we get that $g(f(x))$ equals either $g(x)$ or 0. In the latter case, $f(f(x)) = x$, so $f(x) = f^{-1}(x)$, so $g(x) = 0$. In other words, if $g(x) \neq 0$, then $g(f(x)) = g(x)$.

This means that Lemma 4 reduces to the assertion that at for any x and y, either $g(x) = 0$, $g(y) = 0$ or $g(x) = g(y)$, as required.

Solution 7. As in the other solutions we establish Lemma 1, and as in Solution 2 we deduce that either $f(x) - f(y) = f(x-y)$ or $f(y) - f(x) = f(y-x)$.

Suppose that x and y have $g(x)$ and $g(y)$ both nonzero. Without loss of generality, we have $f(x) - f(y) = f(x-y)$. From $P(x-y, f^{-1}(y))$, we deduce that $x \sim f(x-y) + f^{-1}(y) = f(x) - g(y)$.

If $f(x) = f(x) - g(y)$, then $g(y) = 0$ which is a contradiction. Otherwise, $f^{-1}(x) = f(x) - g(y)$, from which we deduce that $g(x) = g(y)$; in other words, there is at most one nonzero value in the image of g.