USA TST 2023 Solutions

United States of America — Team Selection Test

Andrew Gu, Evan Chen, and Gopal Goel

64[rd] IMO 2022 Japan and 12[th] EGMO 2023 Slovenia

Contents

§0 Problems

1. There are 2022 equally spaced points on a circular track γ of circumference 2022. The points are labeled A_1, A_2, ..., A_{2022} in some order, each label used once. Initially, Bunbun the Bunny begins at A_1. She hops along γ from A_1 to A_2, then from A_2 to A_3, until she reaches A_{2022}, after which she hops back to A_1. When hopping from P to Q, she always hops along the shorter of the two arcs $\widehat{PQ}$ of γ; if $\overline{PQ}$ is a diameter of γ, she moves along either semicircle.

Determine the maximal possible sum of the lengths of the 2022 arcs which Bunbun traveled, over all possible labellings of the 2022 points.

2. Let ABC be an acute triangle. Let M be the midpoint of side BC, and let E and F be the feet of the altitudes from B and C, respectively. Suppose that the common external tangents to the circumcircles of triangles BME and CMF intersect at a point K, and that K lies on the circumcircle of ABC. Prove that line AK is perpendicular to line BC.

3. Consider pairs (f, g) of functions from the set of nonnegative integers to itself such that

- $f(0) \geq f(1) \geq f(2) \geq \dots \geq f(300) \geq 0$;
- $f(0) + f(1) + f(2) + \dots + f(300) \leq 300$;
- for any 20 nonnegative integers n_1, n_2, ..., n_{20}, not necessarily distinct, we have
$$g(n_1 + n_2 + \dots + n_{20}) \leq f(n_1) + f(n_2) + \dots + f(n_{20}).$$

Determine the maximum possible value of $g(0) + g(1) + \dots + g(6000)$ over all such pairs of functions.

4. For nonnegative integers a and b, denote their *bitwise xor* by $a \oplus b$. (For example, $9 \oplus 10 = 1001_2 \oplus 1010_2 = 0011_2 = 3$.)

Find all positive integers a such that for any integers $x > y \geq 0$, we have
$$x \oplus ax \neq y \oplus ay.$$

5. Let m and n be fixed positive integers. Tsvety and Freyja play a game on an infinite grid of unit square cells. Tsvety has secretly written a real number inside of each cell so that the sum of the numbers within every rectangle of size either $m \times n$ or $n \times m$ is zero. Freyja wants to learn all of these numbers.

One by one, Freyja asks Tsvety about some cell in the grid, and Tsvety truthfully reveals what number is written in it. Freyja wins if, at any point, Freyja can simultaneously deduce the number written in every cell of the entire infinite grid. (If this never occurs, Freyja has lost the game and Tsvety wins.)

In terms of m and n, find the smallest number of questions that Freyja must ask to win, or show that no finite number of questions can suffice.

6. Fix a function $f \colon \mathbb{N} \to \mathbb{N}$ and for any $m, n \in \mathbb{N}$ define
$$\Delta(m,n) = \underbrace{f(f(\dots f}_{f(n)\text{ times}}(m)\dots)) - \underbrace{f(f(\dots f}_{f(m)\text{ times}}(n)\dots)).$$

Suppose $\Delta(m,n) \neq 0$ for any distinct $m, n \in \mathbb{N}$. Show that Δ is unbounded, meaning that for any constant C there exist $m, n \in \mathbb{N}$ with $|\Delta(m,n)| > C$.

§1 Solutions to Day 1

§1.1 USA TST 2023/1, proposed by Kevin Cong

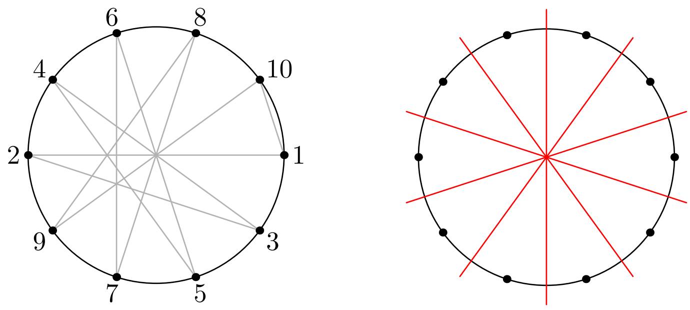
Available online at https://aops.com/community/p26685816.

Problem statement

There are 2022 equally spaced points on a circular track γ of circumference 2022. The points are labeled A_1, A_2, ..., A_{2022} in some order, each label used once. Initially, Bunbun the Bunny begins at A_1. She hops along γ from A_1 to A_2, then from A_2 to A_3, until she reaches A_{2022}, after which she hops back to A_1. When hopping from P to Q, she always hops along the shorter of the two arcs $\widehat{PQ}$ of γ; if $\overline{PQ}$ is a diameter of γ, she moves along either semicircle.

Determine the maximal possible sum of the lengths of the 2022 arcs which Bunbun traveled, over all possible labellings of the 2022 points.

Replacing 2022 with $2n$, the answer is $2n^2 - 2n + 2$.



¶ Construction The construction for $n = 5$ shown on the left half of the figure easily generalizes for all n.

> **Remark.** The validity of this construction can also be seen from the below proof.

¶ First proof of bound Let d_i be the shorter distance from A_{2i-1} to A_{2i+1}.

> **Claim** — The distance of the leg of the journey $A_{2i-1} \to A_{2i} \to A_{2i+1}$ is at most $2n - d_i$.

Proof. Of the two arcs from A_{2i-1} to A_{2i+1}, Bunbun will travel either d_i or $2n - d_i$. One of those arcs contains A_{2i} along the way. So we get a bound of $\max(d_i, 2n - d_i) = 2n - d_i$. □

That means the total distance is at most

$$\sum_{i=1}^{n} (2n - d_i) = 2n^2 - (d_1 + d_2 + \cdots + d_n).$$

Claim — We have
$$d_1 + d_2 + \cdots + d_n \geq 2n - 2.$$

Proof. The left-hand side is the sum of the walk $A_1 \to A_3 \to \cdots \to A_{2n-1} \to A_1$. Among the n points here, two of them must have distance at least $n-1$ apart; the other d_i's contribute at least 1 each. So the bound is $(n-1) + (n-1) \cdot 1 = 2n - 2$. □

¶ **Second proof of bound** Draw the n diameters through the $2n$ arc midpoints, as shown on the right half of the figure for $n = 5$ in red.

Claim (Interpretation of distances) — The distance between any two points equals the number of diameters crossed to travel between the points.

Proof. Clear. □

With this in mind, call a diameter *critical* if it is crossed by all $2n$ arcs.

Claim — At most one diameter is critical.

Proof. Suppose there were two critical diameters; these divide the circle into four arcs. Then all $2n$ arcs cross both diameters, and so travel between opposite arcs. But this means that points in two of the four arcs are never accessed — contradiction. □

Claim — Every diameter is crossed an even number of times.

Proof. Clear: the diameter needs to be crossed an even number of times for the loop to return to its origin. □

This immediately implies that the maximum possible total distance is achieved when one diameter is crossed all $2n$ times, and every other diameter is crossed $2n - 2$ times, for a total distance of at most
$$n \cdot (2n - 2) + 2 = 2n^2 - 2n + 2.$$

§1.2 USA TST 2023/2, proposed by Kevin Cong

Available online at https://aops.com/community/p26685484.

Problem statement

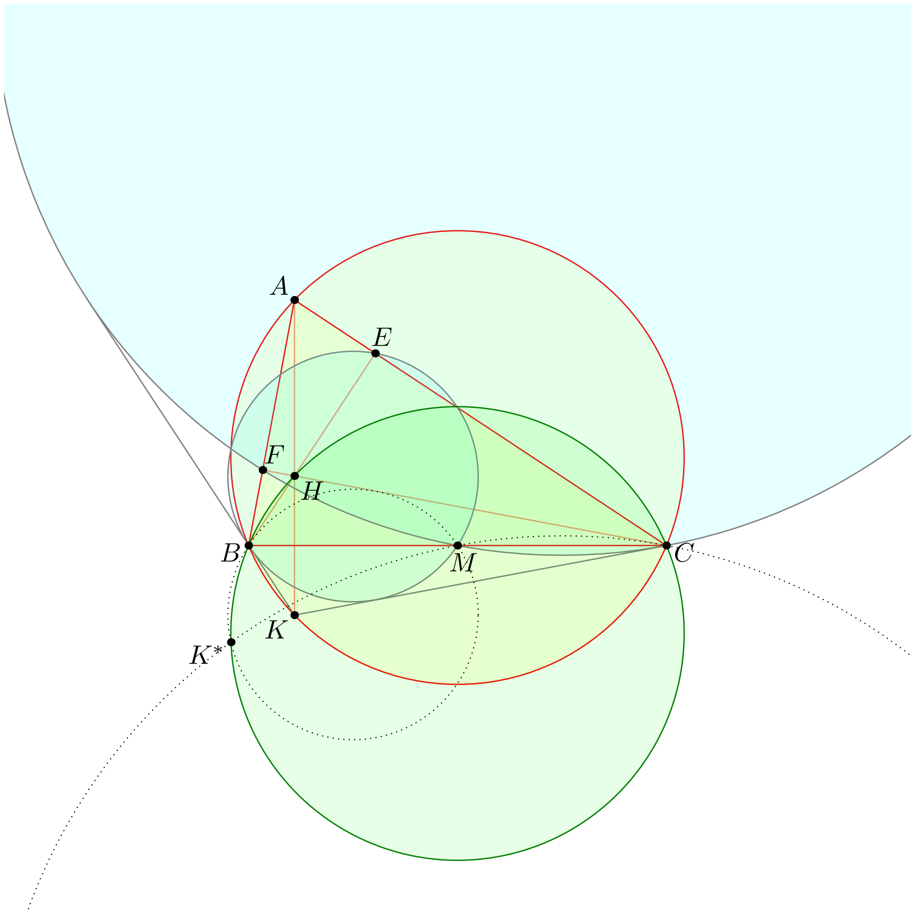
Let ABC be an acute triangle. Let M be the midpoint of side BC, and let E and F be the feet of the altitudes from B and C, respectively. Suppose that the common external tangents to the circumcircles of triangles BME and CMF intersect at a point K, and that K lies on the circumcircle of ABC. Prove that line AK is perpendicular to line BC.

We present several distinct approaches.

¶ Inversion solution submitted by Ankan Bhattacharya and Nikolai Beluhov Let H be the orthocenter of $\triangle ABC$. We use inversion in the circle with diameter $\overline{BC}$. We identify a few images:

- The circumcircles of $\triangle BME$ and $\triangle CMF$ are mapped to lines BE and CF.
- The common external tangents are mapped to the two circles through M which are tangent to lines BE and CF.
- The image of K, denoted K^*, is the second intersection of these circles.
- The assertion that K lies on (ABC) is equivalent to K^* lying on (BHC).

However, now K^* is simple to identify directly: it's just the reflection of M in the bisector of $\angle BHC$.



In particular, $\overline{HK^*}$ is a symmedian of $\triangle BHC$. However, since K^* lies on (BHC), this means $(HK^*; BC) = -1$.

Then, we obtain that $\overline{BC}$ bisects $\angle HMK^* \equiv \angle HMK$. However, K also lies on (ABC), which forces K to be the reflection of H in $\overline{BC}$. Thus $\overline{AK} \perp \overline{BC}$, as wanted.

¶ Solution with coaxial circles (Pitchayut Saengrungkongka) Let H be the orthocenter of $\triangle ABC$. Let Q be the second intersection of $\odot(BME)$ and $\odot(CMF)$. We first prove the following well-known properties of Q.

Claim — Q is the Miquel point of $BCEF$. In particular, Q lies on both $\odot(AEF)$ and $\odot(ABC)$.

Proof. Follows since $BCEF$ is cyclic with M being the circumcenter. □

Claim — $A(Q, H; B, C) = -1$.

Proof. By the radical center theorem on $\odot(AEF)$, $\odot(ABC)$, and $\odot(BCEF)$, we get that AQ, EF, and BC are concurrent. Now, the result follows from a well-known harmonic property. □

Now, we get to the meat of the solution. Let the circumcircle of $\odot(QMK)$ meet BC again at $T \neq M$. The key claim is the following.

Claim — QT is tangent to $\odot(BQC)$.

Proof. We use the "forgotten coaxiality lemma".

$$\begin{aligned}\frac{BT}{TC} &= \frac{TB \cdot TM}{TC \cdot TM} \\ &= \frac{\operatorname{pow}(T, \odot(BME))}{\operatorname{pow}(T, \odot(CMF))} \\ &= \frac{\operatorname{pow}(K, \odot(BME))}{\operatorname{pow}(K, \odot(CMF))} \\ &= \left(\frac{r_{\odot(BME)}}{r_{\odot(CMF)}}\right)^2 \\ &= \left(\frac{BQ/\sin\angle QMB}{CQ/\sin\angle QMC}\right)^2 \\ &= \frac{BQ^2}{CQ^2},\end{aligned}$$

implying the result. □

To finish, let O be the center of $\odot(ABC)$. Then, from the claim, $\angle OQT = 90^\circ = \angle OMT$, so O also lies on $\odot(QMTK)$. Thus, $\angle OKT = 90^\circ$, so KT is also tangent to $\odot(ABC)$ as well. This implies that $QBKC$ is harmonic quadrilateral, and the result follows from the second claim.

¶ Solution by Luke Robitaille Let Q be the second intersection of $\odot(BME)$ and $\odot(CMF)$. We use the first two claims of the previous solution. In particular, $Q \in \odot(ABC)$. We have the following claim.

Claim (Also appeared in ISL 2017 G7) — We have $\measuredangle QKM = \measuredangle QBM + \measuredangle QCM$.

Proof. Let KQ and KM meet $\odot(BME)$ again at Q' and M'. Then, by homothety, $\measuredangle Q'QM' = \measuredangle QCM$, so

$$\begin{aligned}\measuredangle QKM &= \measuredangle Q'QM' + \measuredangle QM'M \\ &= \measuredangle QCM + \measuredangle QBM,\end{aligned}$$

as desired. □

Now, we extend KM to meet $\odot(ABC)$ again at Q_1. We have

$$\begin{aligned}\measuredangle Q_1QB = \measuredangle Q_1KB &= \measuredangle Q_1KQ + \measuredangle QCB \\ &= \measuredangle MKQ + \measuredangle QKB \\ &= (\measuredangle MBQ + \measuredangle MCQ) + \measuredangle QCB \\ &= \measuredangle CBQ,\end{aligned}$$

implying that $QQ_1 \parallel BC$. This implies that $QBKC$ is harmonic quadrilateral, so we are done.

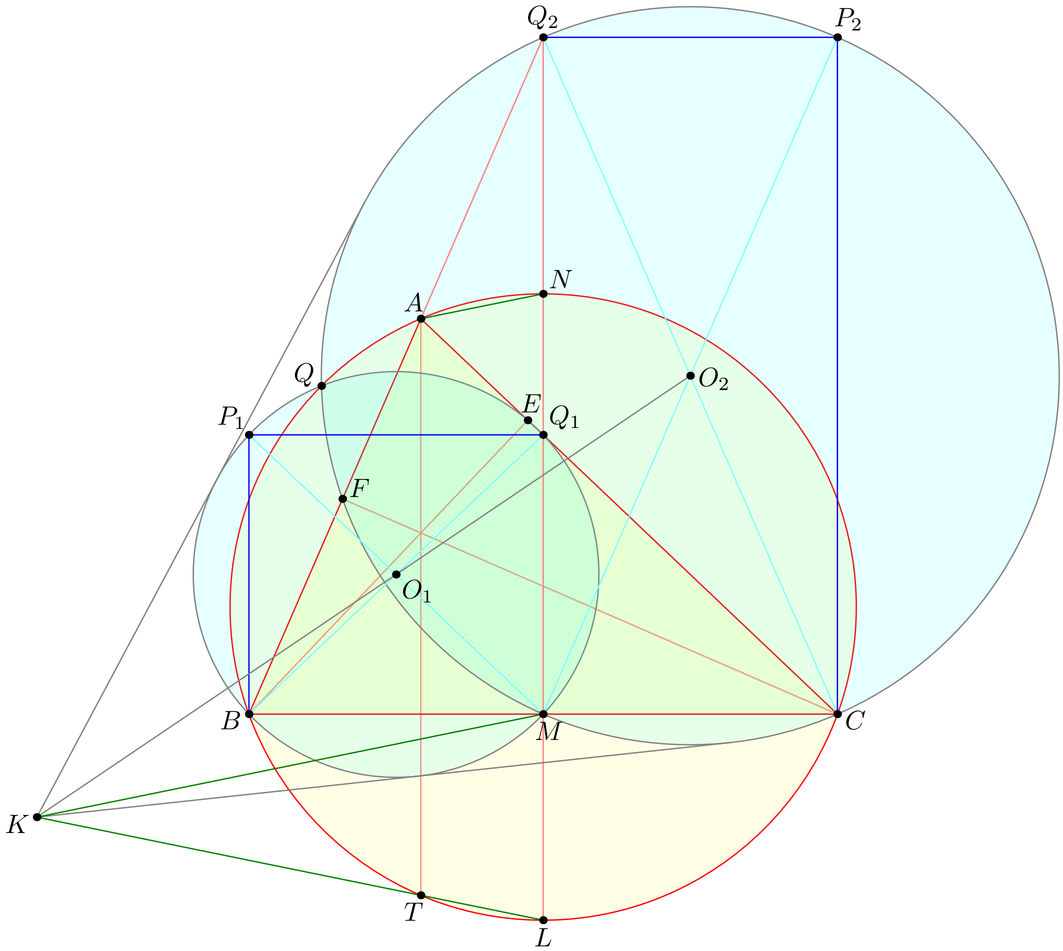
¶ Synthetic solution due to Andrew Gu (Harvard 2026) Define O_1 and O_2 as the circumcenters of (BME) and (CMF). Let T be the point on (ABC) such that $\overline{AT} \perp \overline{BC}$. Denote by L the midpoint of minor arc $\widehat{BC}$.

We are going to ignore the condition that K lies on the circumcircle of ABC, and prove the following unconditional result:

Proposition

The points T, L, K are collinear.

This will solve the problem because if K is on the circumcircle of ABC, it follows $K = T$ or $K = L$; but $K = L$ can never occur since O_1 and O_2 are obviously on different sides of line LM so line LM must meet O_1O_2 inside segment O_1O_2, and K lies outside this segment.



We now turn to the proof of the main lemma. Let P_1 and P_2 be the antipodes of M on these circles.

Claim — Lines AC and LM meet at the antipode Q_1 of B on (BME), so that BP_1Q_1M is a rectangle. Similarly, lines AB and LM meet at the antipode Q_2 of C on (CMF), so that CP_2Q_2M is a rectangle.

Proof. Let $Q_1' = \omega_1 \cap AC \neq E$. Then $\measuredangle BMQ_1' = \measuredangle BEQ_1' = 90^\circ$ hence $Q_1' \in LM$. The other half of the lemma follows similarly. $\square$

From this, it follows that $P_1Q_1 = BM = \frac{1}{2}BC = MC = P_2Q_2$. Letting r_1 denote the radius of ω_1 (and similarly for ω_2), we deduce that $CQ_1 = BQ_1 = 2r_1$.

Claim — $KM = KL$.

Proof. I first claim that $\overline{CL}$ is the external bisector of $\angle Q_1CQ_2$; this follows from

$$\measuredangle Q_1CL = \measuredangle ACL = \measuredangle ABL = \measuredangle Q_2BL = \measuredangle Q_2CL.$$

The external angle bisector theorem then gives an equality of directed ratios

$$\frac{LQ_1}{LQ_2} = \frac{|CQ_1|}{|CQ_2|} = \frac{|BQ_1|}{|CQ_2|} = \frac{2r_1}{2r_2} = \frac{r_1}{r_2}$$

Let the reflection of M over K be P; then P lies on $\overline{P_1P_2}$ and

$$\frac{PP_1}{PP_2} = \frac{2KO_1}{2KO_2} = \frac{KO_1}{KO_2} = \frac{r_1}{r_2} = \frac{LQ_1}{LQ_2}$$

where again the ratios are directed. Projecting everything onto line LM, so that P_1 lands at Q_1 and P_2 lands at Q_2, we find that the projection of P must land exactly at L. $\square$

Claim — Line KM is an external angle bisector of $\angle O_1MO_2$.

Proof. Because $\frac{KO_1}{KO_2} = \frac{r_1}{r_2} = \frac{MO_1}{MO_2}$. $\square$

To finish, note that we know that $\overline{MP_1} \parallel \overline{CQ_1} \equiv \overline{AC}$ and $\overline{MP_2} \parallel \overline{BQ_2} \equiv \overline{AB}$, meaning the angles $\angle O_1MO_2$ and $\angle CAB$ have parallel legs. Hence, if N is the antipode of L, it follows that $\overline{MK} \parallel \overline{AN}$. Now from $MK = KL$ and the fact that $ANLT$ is an isosceles trapezoid, we deduce that $\overline{LT}$ and $\overline{LK}$ are lines in the same direction (namely, the reflection of $MK \parallel AN$ across $\overline{BC}$), as needed.

¶ Complex numbers approach with Apollonian circles, by Carl Schildkraut We use complex numbers. As in the first approach, we will ignore the hypothesis that K lies on (ABC).

Let $Q := (AH) \cap (ABC) \cap (AEF) \neq A$ be the Miquel point of $BFEC$ again. Construct the point T on (ABC) for which $AT \perp BC$; note that $T = -\frac{bc}{a}$. This time the unconditional result is:

Proposition

We have Q, M, T, K are concyclic (or collinear) on an Apollonian circle of $\overline{O_1O_2}$.

This will solve the original problem since once K lies on (ABC) it must be either Q or T. But since K is not on (BME), $K \neq Q$, it will have to be T.

We now prove the proposition. Suppose (ABC) is the unit circle and let $A = a$, $B = b$, $C = c$. Let $H = a + b + c$ be the orthocenter of $\triangle ABC$. By the usual formulas,

$$E := \frac{1}{2}\left(a + b + c - \frac{bc}{a}\right).$$

Let O_1 be the center of (BME) and O_2 be the center of (CMF).

Claim (Calculation of the Miquel point) — We have $Q = \frac{2a+b+c}{a(\frac{1}{a}+\frac{1}{b}+\frac{1}{c})+1}$.

Proof. We now compute that $Q = q$ satisfies $\overline{q} = 1/q$ (since Q is on the unit circle) and $\frac{q-h}{q-a} \in i\mathbb{R}$ (since $AQ \perp QH$), which expands to

$$0 = \frac{q-h}{q-a} + \frac{1/q - \overline{h}}{1/q - 1/a} = \frac{q-h}{q-a} - \frac{a(1-q\overline{h})}{q-a}.$$

This solves to $q = \frac{h+a}{a\overline{h}+1} = \frac{2a+b+c}{a\overline{h}+1}$. □

Claim (Calculation of O_1 and O_2) — We have $O_1 = \frac{b(2a+b+c)}{2(a+b)}$ and $O_2 = \frac{c(2a+b+c)}{2(a+c)}$.

Proof. We now compute O_1 and O_2. For $x, y, z \in \mathbb{C}$, let $\operatorname{Circum}(x, y, z)$ denote the circumcenter of the triangle defined by vertices x, y, and z in $\mathbb{C}$. We have

$$\begin{aligned} O_1 &= \operatorname{Circum}(B, M, E) \\ &= b + \frac{1}{2}\operatorname{Circum}\left(0, c-b, \frac{(a-b)(b-c)}{b}\right) \\ &= b - \frac{b-c}{2b}\operatorname{Circum}(0, b, b-a) \\ &= b - \frac{b-c}{2b}\left(b - \operatorname{Circum}(0, b, a)\right) \\ &= b - \frac{b-c}{2b}\left(b - \frac{ab}{a+b}\right) = b - \frac{b(b-c)}{2(a+b)} = \frac{b(2a+b+c)}{2(a+b)}. \end{aligned}$$

Similarly, $O_2 = \frac{c(2a+b+c)}{2(a+c)}$. □

We are now going to prove the following:

Claim — We have
$$\frac{TO_1}{TO_2} = \frac{MO_1}{MO_2} = \frac{QO_1}{QO_2}.$$

Proof. We now compute

$$MO_1 = BO_1 = \left|b - \frac{b(2a+b+c)}{2(a+b)}\right| = \left|\frac{b(b-c)}{2(a+b)}\right| = \frac{1}{2}\left|\frac{b-c}{a+b}\right|$$

and

$$QO_1 = \left|r - \frac{b(2a+b+c)}{2(a+b)}\right| = \left|1 - \frac{b(a+h)}{2(a+b)r}\right| = \left|1 - \frac{b(a\overline{h}+1)}{2(a+b)}\right| = \left|\frac{a - \frac{ab}{c}}{2(a+b)}\right| = \frac{1}{2}\left|\frac{b-c}{a+b}\right|.$$

This implies both (by symmetry) that $\frac{MO_1}{MO_2} = \frac{QO_1}{QO_2} = \left|\frac{a+c}{a+b}\right|$ and that Q is on (BME) and (CMF). Also,

$$\frac{TO_1}{TO_2} = \frac{\left|\frac{b(2a+b+c)}{2(a+b)} + \frac{bc}{a}\right|}{\left|\frac{c(2a+b+c)}{2(a+c)} + \frac{bc}{a}\right|} = \left|\frac{\frac{b(2a^2+ab+ac+2ac+2bc)}{2a(a+b)}}{\frac{c(2a^2+ab+ac+2ab+2bc)}{2a(a+c)}}\right| = \left|\frac{a+c}{a+b}\right| \cdot \left|\frac{2a^2+2bc+ab+3ac}{2a^2+2bc+3ab+ac}\right|;$$

if $z = 2a^2 + 2bc + ab + 3ac$, then $a^2bc\overline{z} = 2a^2 + 2bc + 3ab + ac$, so the second term has magnitude 1. This means $\frac{TO_1}{TO_2} = \frac{MO_1}{MO_2} = \frac{QO_1}{QO_2}$, as desired. □

To finish, note that this common ratio is the ratio between the radii of these two circles, so it is also $\frac{KO_1}{KO_2}$. By Apollonian circles the points $\{Q, M, T, K\}$ lie on a circle or a line.

§1.3 USA TST 2023/3, proposed by Sean Li

Available online at https://aops.com/community/p26685437.

Problem statement

Consider pairs (f, g) of functions from the set of nonnegative integers to itself such that

- $f(0) \geq f(1) \geq f(2) \geq \cdots \geq f(300) \geq 0$;
- $f(0) + f(1) + f(2) + \cdots + f(300) \leq 300$;
- for any 20 nonnegative integers n_1, n_2, …, n_{20}, not necessarily distinct, we have
$$g(n_1 + n_2 + \cdots + n_{20}) \leq f(n_1) + f(n_2) + \cdots + f(n_{20}).$$

Determine the maximum possible value of $g(0) + g(1) + \cdots + g(6000)$ over all such pairs of functions.

Replace $300 = \frac{24 \cdot 25}{2}$ with $\frac{s(s+1)}{2}$ where $s = 24$, and 20 with k. The answer is $115440 = \frac{ks(ks+1)}{2}$. Equality is achieved at $f(n) = \max(s - n, 0)$ and $g(n) = \max(ks - n, 0)$. To prove
$$g(n_1 + \cdots + n_k) \leq f(n_1) + \cdots + f(n_k),$$
write it as
$$\max(x_1 + \cdots + x_k, 0) \leq \max(x_1, 0) + \cdots + \max(x_k, 0)$$
with $x_i = s - n_i$. This can be proven from the $k = 2$ case and induction.

It remains to show the upper bound. For this problem, define a *partition* to be a nonincreasing function $p : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ such that $p(n) = 0$ for some n. The sum of p is defined to be $\sum_{n=0}^{\infty} p(n)$, which is finite under the previous assumption. Let $L = \mathbb{Z}_{\geq 0}^2$. The *Young diagram* of the partition is the set of points
$$\mathcal{P} := \{(x, y) \in L : y < p(x)\}.$$

The number of points in $\mathcal{P}$ is equal to the sum of p. The *conjugate* of a partition defined as
$$p_*(n) = \text{the number of } i \text{ for which } p(i) > n.$$
This is a partition with the same sum as p. Geometrically, the Young diagrams of p and p_* are reflections about $x = y$.

Since each $g(n)$ is independent, we may maximize each one separately for all n and assume that
$$g(n) = \min_{n_1 + \cdots + n_k = n} (f(n_1) + \cdots + f(n_k)). \qquad (*)$$

The conditions of the problem statement imply that $f\left(\frac{s(s+1)}{2}\right) = 0$. Then, for any $n \leq k\frac{s(s+1)}{2}$, there exists an optimal combination $(n_1, \ldots, n_k)$ in (*) where all n_i are at most $\frac{s(s+1)}{2}$, by replacing any term in an optimum greater than $\frac{s(s+1)}{2}$ by $\frac{s(s+1)}{2}$ and shifting the excess to smaller terms (because f is nonincreasing). Therefore we may extend f to a partition by letting $f(n) = 0$ for $n > \frac{s(s+1)}{2}$ without affecting the relevant values of g. Then (*) implies that g is a partition as well.

The problem can be restated as follows: f is a partition with sum $\frac{s(s+1)}{2}$, and g is a partition defined by (*). Find the maximum possible sum of g. The key claim is that the problem is the same under conjugation.

Claim — Under these conditions, we have

$$g_*(n) = \min_{n_1+\cdots+n_k=n} (f_*(n_1) + \cdots + f_*(n_k)).$$

Proof. Let $\mathcal{F}$ and $\mathcal{G}$ be the Young diagrams of f and g respectively, and $\overline{\mathcal{F}} = L \setminus \mathcal{F}$ and $\overline{\mathcal{G}} = L \setminus \mathcal{G}$ be their complements. The lower boundary of $\overline{\mathcal{F}}$ is formed by the points $(n, f(n))$ for $i \in \mathbb{Z}_{\geq 0}$. By the definition of g, the lower boundary of $\overline{\mathcal{G}}$ consists of points $(n, g(n))$ which are formed by adding k points of $\overline{\mathcal{F}}$. This means

$$\overline{\mathcal{G}} = \underbrace{\overline{\mathcal{F}} + \cdots + \overline{\mathcal{F}}}_{k\ \overline{\mathcal{F}}\text{'s}}$$

where $+$ denotes set addition. This definition remains invariant under reflection about $x = y$, which swaps f and g with their conjugates. $\square$

Let A be the sum of g. We now derive different bounds on A. First, by Hermite's identity

$$n = \sum_{i=0}^{k-1} \left\lfloor \tfrac{n+i}{k} \right\rfloor$$

we have

$$\begin{aligned} A &= \sum_{n=0}^{\infty} g(n) \\ &\leq \sum_{n=0}^{\infty} \sum_{i=0}^{k-1} f\left(\left\lfloor \tfrac{n+i}{k} \right\rfloor\right) \\ &= k^2 \sum_{n=0}^{\infty} f(n) - \frac{k(k-1)}{2} f(0) \\ &= k^2 \frac{s(s+1)}{2} - \frac{k(k-1)}{2} f(0). \end{aligned}$$

By the claim, we also get the second bound $A \leq k^2 \frac{s(s+1)}{2} - \frac{k(k-1)}{2} f_*(0)$.

For the third bound, note that $f(f_*(0)) = 0$ and thus $g(kf_*(0)) = 0$. Moreover,

$$g(qf_*(0) + r) \leq q \cdot f(f_*(0)) + (k-q-1)f(0) + f(r) = (k-q-1)f(0) + f(r),$$

so we have

$$\begin{aligned} A &= \sum_{\substack{0 \leq q < k \\ 0 \leq r < f_*(0)}} g(qf_*(0) + r) \\ &\leq \frac{k(k-1)}{2} f_*(0) f(0) + k \sum_{0 \leq r < f_*(0)} f(r) \\ &= \frac{k(k-1)}{2} f_*(0) f(0) + k \frac{s(s+1)}{2}. \end{aligned}$$

Now we have three cases:

- If $f(0) \ge s$ then
$$A \le k^2 \frac{s(s+1)}{2} - \frac{k(k-1)}{2} f(0) \le \frac{ks(ks+1)}{2}.$$
- If $f_*(0) \ge s$ then
$$A \le k^2 \frac{s(s+1)}{2} - \frac{k(k-1)}{2} f_*(0) \le \frac{ks(ks+1)}{2}.$$
- Otherwise, $f(0)f_*(0) \le s^2$ and
$$A \le \frac{k(k-1)}{2} f_*(0) f(0) + k \frac{s(s+1)}{2} \le \frac{ks(ks+1)}{2}.$$

In all cases, $A \le \frac{ks(ks+1)}{2}$, as desired.

Remark. One can estimate the answer to be around $k^2 \frac{s(s+1)}{2}$ by observing the set addition operation "dilates" $\mathcal{F}$ by a factor of k, but significant care is needed to sharpen the bound.

§2 Solutions to Day 2

§2.1 USA TST 2023/4, proposed by Carl Schildkraut

Available online at https://aops.com/community/p26896062.

Problem statement

For nonnegative integers a and b, denote their *bitwise xor* by $a \oplus b$. (For example, $9 \oplus 10 = 1001_2 \oplus 1010_2 = 0011_2 = 3$.)

Find all positive integers a such that for any integers $x > y \geq 0$, we have

$$x \oplus ax \neq y \oplus ay.$$

Answer: the function $x \mapsto x \oplus ax$ is injective if and only if a is an even integer.

¶ Even case First, assume $\nu_2(a) = k > 0$. We wish to recover x from $c := x \oplus ax$. Notice that:

- The last k bits of c coincide with the last k bits of x.
- Now the last k bits of x give us also the last $2k$ bits of ax, so we may recover the last $2k$ bits of x as well.
- Then the last $2k$ bits of x give us also the last $3k$ bits of ax, so we may recover the last $3k$ bits of x as well.
- ...and so on.

¶ Odd case Conversely, suppose a is odd. To produce the desired collision:

Claim — Let n be any integer such that $2^n > a$, and define

$$x = \underbrace{1\dots1}_{n} = 2^n - 1, \qquad y = 1\underbrace{0\dots0}_{n}1 = 2^n + 1.$$

Then $x \oplus ax = y \oplus ay$.

Proof. Let P be the binary string for a, zero-padded to length n, and let Q be the binary string for $a-1$, zero-padded to length n, Then let R be the bitwise complement of Q. (Hence all three of are binary strings of length n.) Then

$$\begin{aligned} ax = \overline{QR} &\implies x \oplus ax = \overline{QQ} \\ ay = \overline{PP} &\implies y \oplus ay = \overline{QQ}. \end{aligned}$$

We're done. □

§2.2 USA TST 2023/5, proposed by Nikolai Beluhov

Available online at https://aops.com/community/p26896130.

Problem statement

Let m and n be fixed positive integers. Tsvety and Freyja play a game on an infinite grid of unit square cells. Tsvety has secretly written a real number inside of each cell so that the sum of the numbers within every rectangle of size either $m \times n$ or $n \times m$ is zero. Freyja wants to learn all of these numbers.

One by one, Freyja asks Tsvety about some cell in the grid, and Tsvety truthfully reveals what number is written in it. Freyja wins if, at any point, Freyja can simultaneously deduce the number written in every cell of the entire infinite grid. (If this never occurs, Freyja has lost the game and Tsvety wins.)

In terms of m and n, find the smallest number of questions that Freyja must ask to win, or show that no finite number of questions can suffice.

The answer is the following:

- If $\gcd(m, n) > 1$, then Freyja cannot win.
- If $\gcd(m, n) = 1$, then Freyja can win in a minimum of $(m-1)^2 + (n-1)^2$ questions.

First, we dispose of the case where $\gcd(m, n) > 1$. Write $d = \gcd(m, n)$. The idea is that any labeling where each $1 \times d$ rectangle has sum zero is valid. Thus, to learn the labeling, Freyja must ask at least one question in every row, which is clearly not possible in a finite number of questions.

Now suppose $\gcd(m, n) = 1$. We split the proof into two halves.

¶ Lower bound Clearly, any labeling where each $m \times 1$ and $1 \times m$ rectangle has sum zero is valid. These labelings form a vector space with dimension $(m-1)^2$, by inspection. (Set the values in an $(m-1) \times (m-1)$ square arbitrarily and every other value is uniquely determined.)

Similarly, labelings where each $n \times 1$ and $1 \times n$ rectangle have sum zero are also valid, and have dimension $(n-1)^2$.

It is also easy to see that no labeling other than the all-zero labeling belongs to both categories; labelings in the first space are periodic in both directions with period m, while labelings in the second space are periodic in both directions with period n; and hence any labeling in both categories must be constant, ergo all-zero.

Taking sums of these labelings gives a space of valid labelings of dimension $(m-1)^2 + (n-1)^2$. Thus, Freyja needs at least $(m-1)^2 + (n-1)^2$ questions to win.

¶ Proof of upper bound using generating functions, by Ankan Bhattacharya We prove:

Claim (Periodicity) — Any valid labeling is doubly periodic with period mn.

Proof. By Chicken McNugget, there exists N such that N and $N+1$ are both nonnegative integer linear combinations of m and n.

Then both $mn \times N$ and $mn \times (N+1)$ rectangles have zero sum, so $mn \times 1$ rectangles have zero sum. This implies that any two cells with a vertical displacement of mn are equal; similarly for horizontal displacements. $\square$

With that in mind, consider a valid labeling. It naturally corresponds to a generating function

$$f(x,y) = \sum_{a=0}^{mn-1} \sum_{b=0}^{mn-1} c_{a,b} x^a y^b$$

where $c_{a,b}$ is the number in (a,b).

The generating function corresponding to sums over $n \times m$ rectangles is

$$f(x,y)(1+x+\cdots+x^{m-1})(1+y+\cdots+y^{n-1}) = f(x,y) \cdot \frac{x^m-1}{x-1} \cdot \frac{y^n-1}{y-1}.$$

Similarly, the one for $m \times n$ rectangles is

$$f(x,y) \cdot \frac{x^n-1}{x-1} \cdot \frac{y^m-1}{y-1}.$$

Thus, the constraints for f to be valid are equivalent to

$$f(x,y) \cdot \frac{x^m-1}{x-1} \cdot \frac{y^n-1}{y-1} \quad \text{and} \quad f(x,y) \cdot \frac{x^n-1}{x-1} \cdot \frac{y^m-1}{y-1}$$

being zero when reduced modulo $x^{mn}-1$ and $y^{mn}-1$, or, letting $\omega = \exp(2\pi i/mn)$, both terms being zero when powers of ω are plugged in.

To restate the constraints one final time, we need

$$f(\omega^a, \omega^b) \cdot \frac{\omega^{am}-1}{\omega^a-1} \cdot \frac{\omega^{bn}-1}{\omega^b-1} = f(\omega^a, \omega^b) \cdot \frac{\omega^{an}-1}{\omega^a-1} \cdot \frac{\omega^{bm}-1}{\omega^b-1} = 0$$

for all $a, b \in \{0, \dots, mn-1\}$.

This implies $f(\omega^a, \omega^b) = 0$ for most choices of (a,b). If it does not, we need

$$\frac{\omega^{am}-1}{\omega^a-1} \cdot \frac{\omega^{bn}-1}{\omega^b-1} = \frac{\omega^{an}-1}{\omega^a-1} \cdot \frac{\omega^{bm}-1}{\omega^b-1} = 0.$$

This happens when (at least) one fraction in either product is zero.

- If the first fraction is zero, then either $n \mid a$ and $a > 0$, or $m \mid b$ and $b > 0$.
- If the second fraction is zero, then either $m \mid a$ and $a > 0$, or $n \mid b$ and $b > 0$.

If the first condition holds in both cases, then $mn \mid a$, but $0 < a < mn$, a contradiction. Thus if $n \mid a$, then we must have $n \mid b$, and similarly if $m \mid a$ then $m \mid b$.

The former case happens $(m-1)^2$ times, and the latter case happens $(n-1)^2$ times. Thus, at most $(m-1)^2 + (n-1)^2$ values of $f(\omega^a, \omega^b)$ are nonzero, It follows that the dimension of the space of valid labelings is at most $(m-1)^2 + (n-1)^2$, as desired.

¶ Explicit version of winning algorithm by Freyja, from author Suppose that $\gcd(m, n) = 1$ and $m \le n$. Let $[a, b]$ denote the set of integers between a and b inclusive.

Let Freyja ask about all cells (x, y) in the two squares

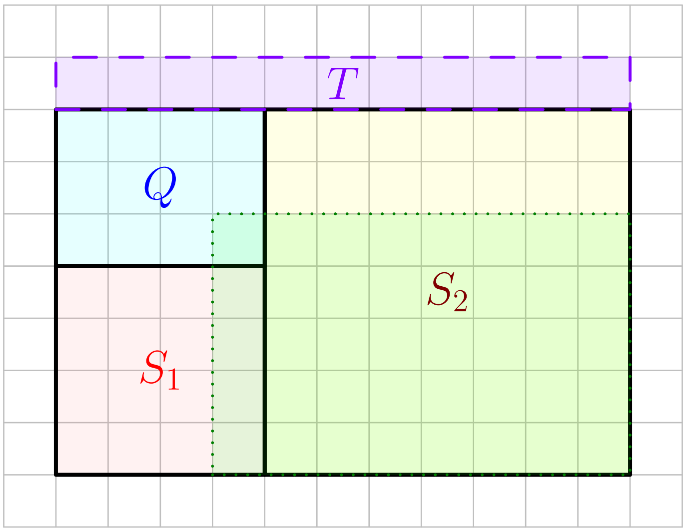
$$S_1 = [1, m-1] \times [1, m-1]$$
$$S_2 = [m, m+n-2] \times [1, n-1].$$

In the beginning, one by one, Freyja determines all values inside of the rectangle $Q := [1, m-1] \times [m, n-1]$. To that end, on each step she considers some rectangle with m rows and n columns such that its top left corner is in Q and all of the other values in it have been determined already. In this way, Freyja uncovers all of Q, starting with its lower right corner and then proceeding upwards and to the left.

Thus Freyja can learn all numbers inside of the rectangle

$$R := [1, m+n-2] \times [1, n-1] = Q \cup S_1 \cup S_2.$$

See the figure below for an illustration for $(m, n) = (5, 8)$. The first cell of Q is uncovered using the dotted green rectangle.



$(m, n) = (5, 8)$

We need one lemma:

Lemma

Let m and n be positive integers with $\gcd(m, n) = 1$. Consider an unknown sequence of real numbers $z_1, z_2, \dots, z_s$ with $s \ge m+n-2$. Suppose that we know the sums of all contiguous blocks of size either m or n in this sequence. Then we can determine all individual entries in the sequence as well.

Proof. By induction on $m+n$. Suppose, without loss of generality, that $m \le n$. Our base case is $m = 1$, which is clear. For the induction step, set $\ell = n - m$. Each contiguous block of size ℓ within $z_1, z_2, \dots, z_{s-n}$ is the difference of two contiguous blocks of sizes m and n within the original sequence. By the induction hypothesis for ℓ and m, it follows that we can determine all of $z_1, z_2, \dots, z_{s-n}$. Then we determine the remaining z_i as well, one by one, in order from left to right, by examining on each step an appropriate contiguous block of size m. $\square$

Let T be the rectangle $[1, m+n-2] \times \{n\}$. By looking at appropriate rectangles of sizes $m \times n$ and $n \times m$ such that their top row is contained within T and all of their other rows are contained within R, Freyja can learn the sums of all contiguous blocks of values of sizes m and n within T. By the Lemma, it follows that Freyja can uncover all of T.

In this way, with the help of the Lemma, Freyja can extend her rectangular area of knowledge both upwards and downwards. Once its height reaches $m+n-2$, by the same method she will be able to extend it to the left and right as well. This allows Freyja to determine all values in the grid. Therefore, $(m-1)^2 + (n-1)^2$ questions are indeed sufficient.

Remark. The ideas in the solution also yield a proof of the following result:

> Let m and n be relatively prime positive integers. Consider an infinite grid of unit square cells coloured in such a way that every rectangle of size either $m \times n$ or $n \times m$ contains the same multiset of colours. Then the colouring is either doubly periodic with period length m or doubly periodic with period length n.

(Here, "doubly periodic with period length s" means "both horizontally and vertically periodic with period length s".)

Here is a quick sketch of the proof. Given two positive integers i and j with $1 \le i, j \le m-1$, we define

$$f_{ij}(x,y) \coloneqq \begin{cases} +1 & \text{when } (x,y) \equiv (0,0) \text{ or } (i,j) \pmod m; \\ -1 & \text{when } (x,y) \equiv (0,j) \text{ or } (i,0) \pmod m; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Define $g_{i,j}$ similarly, but with $1 \le i, j \le n-1$ and "mod m" everywhere replaced by "mod n". First we show that if a linear combination $h := \sum \alpha_{i,j} f_{i,j} + \sum \beta_{i,j} g_{i,j}$ of the $f_{i,j}$ and $g_{i,j}$ contains only two distinct values, then either all of the $\alpha_{i,j}$ vanish or all of the $\beta_{i,j}$ do. It follows that each colour, considered in isolation, is either doubly periodic with period length m or doubly periodic with period length n. Finally, we check that different period lengths cannot mix.

On the other hand, if m and n are not relatively prime, then there exist infinitely many non-isomorphic valid colourings. Furthermore, when $\gcd(m,n) = 2$, there exist valid colourings which are not horizontally periodic; and, when $\gcd(m,n) \ge 3$, there exist valid colourings which are neither horizontally nor vertically periodic.

§2.3 USA TST 2023/6, proposed by Maxim Li

Available online at https://aops.com/community/p26896222.

Problem statement

Fix a function $f\colon \mathbb{N} \to \mathbb{N}$ and for any $m, n \in \mathbb{N}$ define

$$\Delta(m,n) = \underbrace{f(f(\dots f}_{f(n)\text{ times}}(m)\dots)) - \underbrace{f(f(\dots f}_{f(m)\text{ times}}(n)\dots)).$$

Suppose $\Delta(m,n) \neq 0$ for any distinct $m, n \in \mathbb{N}$. Show that Δ is unbounded, meaning that for any constant C there exist $m, n \in \mathbb{N}$ with $|\Delta(m,n)| > C$.

Suppose for the sake of contradiction that $|\Delta(m,n)| \le N$ for all m, n. Note that f is injective, as

$$f(m) = f(n) \implies \Delta(m,n) = 0 \implies m = n,$$

as desired.

Let G be the "arrow graph" of f, which is the directed graph with vertex set $\mathbb{N}$ and edges $n \to f(n)$. The first step in the solution is to classify the structure of G. Injectivity implies that G is a disjoint collection of chains (infinite and half-infinite) and cycles. We have the following sequence of claims that further refine the structure.

Claim — The graph G has no cycles.

Proof. Suppose for the sake of contradiction that $f^k(n) = n$ for some $k \ge 2$ and $n \in \mathbb{N}$. As m varies over $\mathbb{N}$, we have $|\Delta(m,n)| \le N$, so $f^{f(n)}(m)$ can only take on some finite set of values. In particular, this means that

$$f^{f(n)}(m_1) = f^{f(n)}(m_2)$$

for some $m_1 \neq m_2$, which contradicts injectivity. $\square$

Claim — The graph G has at most $2N + 1$ chains.

Proof. Suppose we have numbers $m_1, \dots, m_k$ in distinct chains. Select a positive integer $B > \max\{f(m_1), \dots, f(m_k)\}$. Now,

$$\left|\Delta\left(m_i, f^{B-f(m_i)}(1)\right)\right| \le N \implies \left|f^B(1) - f^{f^{B-f(m_i)+1}(1)}(m_i)\right| \le N.$$

Since the m_is are in different chains, we have that $f^{f^{B-f(m_i)+1}(1)}(m_i)$ are distinct for each i, which implies that $k \le 2N + 1$, as desired. $\square$

Claim — The graph G consists of exactly one half-infinite chain.

Proof. Fix some $c \in \mathbb{N}$. Call an element of $\mathbb{N}$ *bad* if it is not of the form $f^k(c)$ for some $k \ge 0$. It suffices to show that there are only finitely many bad numbers.

Since there are only finitely many chains, $f^{f(c)}(n)$ achieves all sufficiently large positive integers, say all positive integers at least M. Fix A and B such that $B > A \ge M$.

If $f^{f(c)}(n) \in [A, B]$, then $f^{f(n)}(c) \in [A-N, B+N]$, and distinct n generate distinct $f^{f(n)}(c)$ due to the structure of G. Therefore, we have at least $B-A+1$ good numbers in $[A-N, B+N]$, so there are at most $2N$ bad numbers in $[A-N, B+N]$.

Varying B, this shows there are at most $2N$ bad numbers at least $A-N$. $\square$

Let c be the starting point of the chain, so every integer is of the form $f^k(c)$, where $k \geq 0$. Define a function $g\colon \mathbb{Z}_{\geq 0} \to \mathbb{N}$ by

$$g(k) \coloneqq f^k(c).$$

Due to the structure of G, g is a bijection. Define

$$\delta(a,b) \coloneqq \Delta(f^a(c), f^b(c)) = g(g(b+1)+a) - g(g(a+1)+b),$$

so the conditions are equivalent to $|\delta(a,b)| \leq N$ for all $a, b \in \mathbb{Z}_{\geq 0}$ and $\delta(a,b) \neq 0$ for $a \neq b$, which is equivalent to $g(a+1) - a \neq g(b+1) - b$ for $a \neq b$. This tells us that $g(x) - x$ is injective for $x \geq 1$.

Lemma

For all M, there exists a nonnegative integer x with $g(x) \leq x - M$.

Proof. Assume for the sake of contradiction that $g(x) - x$ is bounded below. Fix some large positive K. Since $g(x) - x$ is injective, there exists B such that $g(x) - x \geq K$ for all $x \geq B$. Then $\min\{g(B+1), g(B+2), \dots\} \geq B + K$, while $\{g(0), \dots, g(B)\}$ only achieve $B+1$ values. Thus, at least $K-1$ values are not achieved by g, which is a contradiction. $\square$

Now pick B such that $g(B) + N \leq B$ and $g(B) > N$. Note that infinitely many such B exist, since we can take M to be arbitrarily small in the above lemma. Let

$$t = \max\{g^{-1}(g(B)-N), g^{-1}(g(B)-N+1), \dots, g^{-1}(g(B)+N)\}.$$

Note that $g(t) \leq g(B) + N \leq B$, so we have

$$|\delta(t-1, B-g(t))| = |g(B) - g(t-1+g(B+1-g(t)))| \leq N,$$

so

$$t - 1 + g(B+1-g(t)) \in \{g^{-1}(g(B)-N), g^{-1}(g(B)-N+1), \dots, g^{-1}(g(B)+N)\},$$

so by the maximality of t, we must have $g(B+1-g(t)) = 1$, so $B + 1 - g(t) = g^{-1}(1)$. We have $|g(t) - g(B)| \leq N$, so

$$|(B - g(B)) + 1 - g^{-1}(1)| \leq N.$$

This is true for infinitely many values of B, so infinitely many values of $B - g(B)$ (by injectivity of $g(x) - x$), which is a contradiction. This completes the proof.