

CPS(A) Match 2022: Solutions and Marking schemes

ISTA, Austria
July 1 – July 4, 2022

Problem 1. Let $k \leq 2022$ be a positive integer. Alice and Bob play a game on a 2022×2022 board. Initially, all cells are white. Alice starts and the players alternate. In her turn, Alice can either color one white cell in red or pass her turn. In his turn, Bob can either color a $k \times k$ square of white cells in blue or pass his turn. Once both players pass, the game ends and the person who colored more cells wins (a draw can occur).

For each $1 \leq k \leq 2022$, determine which player (if any) has a winning strategy.
(David Hruška)

Solution. Answer: For $k \in \{1, 3, 337, 1011\}$ the game ends in a draw, otherwise Alice wins.

Fix k . Number the rows from top to bottom and the columns from left to right using numbers $1, \dots, 2022$. We say that a cell is *critical* if its row-number and column-number are both divisible by k . Note that Bob colors exactly one critical cell in each his turn. Let $c = \lfloor 2022/k \rfloor^2$ be the number of critical cells.

We distinguish 3 cases based on the value $2022/k$.

- (1) $2022/k \notin \mathbb{Z}$. Then Alice can force a win, e.g. by primarily coloring the critical cells (and then coloring the rest): If she follows this strategy then Bob will color a $k \times k$ square in at most $c/2$ of his turns, so in total he will color $\frac{1}{2}c \cdot k^2 < \frac{1}{2} \cdot 2022^2$ cells.
- (2) $2022/k$ is an odd integer. Then Alice can force a win by the same strategy: In this case c is odd, so Bob will color a $k \times k$ square in at most $\lfloor c/2 \rfloor < \frac{1}{2}c$ of his turns, so in total he will color strictly less than $\frac{1}{2}c \cdot k^2 = \frac{1}{2} \cdot 2022^2$ cells.
- (3) $2022/k$ is an even integer. Then both players can guarantee at least a draw: Alice by the above strategy, Bob by always coloring a $k \times k$ square whose bottom right corner cell is critical. Since in this case c is even and each Alice's move prevents Bob from coloring at most 1 such square, using this strategy he will color at least $\frac{1}{2}c \cdot k^2 = \frac{1}{2} \cdot 2022^2$ cells.

Since $2022 = 2 \cdot 3 \cdot 337$, case (3) occurs if and only if $k \in \{1, 3, 337, 1011\}$.

Marking scheme.

Partial scores for otherwise incomplete solutions:

- | | |
|--|------|
| (A ₀) Solving $k = 1$ and $k = 2022$. | +0 p |
| (A ₁) Solving at least one value k with $2 \leq k \leq 2021$. | +1 p |
| (A ₂) Description of a good strategy for Alice in one of (1)-(3), or a good strategy for Bob in (3). | +1 p |
| (A ₃) Correctly solving at least one of the 3 cases (1)-(3). | +1 p |
| (A ₄) Correct answer. | +1 p |

Deductions for essentially complete solutions:

- | | |
|--|------|
| (C) An essentially complete solution. | 7 p |
| (D ₁) Not stating the values 1, 3, 337, 1011 explicitly. | -0 p |
| (D ₂) Incorrect answer. | -1 p |
| (D ₃) Other minor errors (easy to fix). | -1 p |

Define $A = \sum_{i=1}^4 A_i$ and $B = C + D_1 + D_2 + D_3$. Give $\max\{A, B\}$ points.

Problem 2. Find all functions $f: (0, \infty) \rightarrow (0, \infty)$ such that

$$f\left(f(x) + \frac{y+1}{f(y)}\right) = \frac{1}{f(y)} + x + 1$$

for all $x, y > 0$.

(Dominik Burek)

Solution. First, observe that the range of f contains the interval (A, ∞) where $A = \frac{1}{f(1)} + 1$. Indeed, for any $B > A$ we let $x = B - A > 0$ and $y = 1$ and we obtain

$$f\left(f(x) + \frac{y+1}{f(y)}\right) = \frac{1}{f(y)} + x + 1 = B.$$

Second, f is injective. Indeed, if $f(x_1) = f(x_2)$ then

$$\frac{1}{f(y)} + x_1 + 1 = f\left(f(x_1) + \frac{y+1}{f(y)}\right) = f\left(f(x_2) + \frac{y+1}{f(y)}\right) = \frac{1}{f(y)} + x_2 + 1$$

which yields $x_1 = x_2$.

Now, consider an arbitrary $0 < \delta < \frac{1}{A}$. Write $\delta = \frac{1}{B} - \frac{1}{C}$ for some $C > B > A$. Then $B = f(y_1)$ and $C = f(y_2)$ for some $y_1, y_2 > 0$. Let $x > 0$ be arbitrary and set $x_1 = x$, $x_2 = x + \delta$. We have

$$\frac{1}{f(y_1)} + x_1 + 1 = \frac{1}{f(y_2)} + x_2 + 1,$$

hence

$$f\left(f(x_1) + \frac{y_1+1}{f(y_1)}\right) = f\left(f(x_2) + \frac{y_2+1}{f(y_2)}\right)$$

and, by injectivity,

$$f(x_1) + \frac{y_1+1}{f(y_1)} = f(x_2) + \frac{y_2+1}{f(y_2)}.$$

This rewrites as $f(x + \delta) = f(x) + \varepsilon_\delta$ where $\varepsilon_\delta = \frac{y_1+1}{f(y_1)} - \frac{y_2+1}{f(y_2)}$ depends only on δ . Easy induction yields $f(x + n\delta) = f(x) + n\varepsilon_\delta$ for every integer n satisfying $x + n\delta > 0$.

We prove now that f is linear on $\mathbb{Q} \cap (0, \infty)$. For every integer $k > A$ we have $f(2) = f(1 + k \cdot \frac{1}{k}) = f(1) + k\varepsilon_{1/k}$, hence $\varepsilon_{1/k} = \frac{f(2) - f(1)}{k}$. Let x be a positive rational. Write $x = \frac{m}{n}$ for some integers m, n such that $n > A$. Then

$$\begin{aligned} f(x) &= f\left(1 + \frac{m-n}{n}\right) = f(1) + (m-n)\varepsilon_{1/n} = f(1) + \frac{(m-n)(f(2) - f(1))}{n} \\ &= f(1) + (x-1)(f(2) - f(1)). \end{aligned}$$

We prove now that f is increasing. Suppose otherwise: $f(x) > f(y)$ for some $0 < x < y$. Let $\delta = \frac{y-x}{k}$ where k is an integer so large that $\delta < \frac{1}{A}$. Then

$f(y) = f(x + k\delta) = f(x) + k\varepsilon_\delta$, hence $\varepsilon_\delta = \frac{f(y) - f(x)}{k} < 0$. Then $f(x + n\delta) = f(x) - n\varepsilon_\delta < 0$ for sufficiently large n which contradicts the assumption that the codomain of f is $(0, \infty)$.

Since f is linear on $\mathbb{Q} \cap (0, \infty)$ and increasing on $(0, \infty)$, f is linear on $(0, \infty)$. Letting $f(x) = ax + b$ we obtain

$$a\left(ax + b + \frac{y+1}{ay+b}\right) + b = \frac{1}{ay+b} + x + 1,$$

which immediately gives $a = 1$ and $b = 0$. The function $f(x) = x$ clearly works.

Marking scheme.

(A_0) Showing that among linear functions, the only solution is $f(x) = x$.	+0 p
(A_1) Showing that f is injective.	+1 p
(A_2) Showing that the range of f contains (A, ∞) for some $A > 0$.	+1 p
(A_3) Showing that f is linear on \mathbb{Q}^+ .	+2 p
(A_4) Showing that f is increasing.	+1 p
(L) Showing that f is linear.	+6 p
(C) An essentially complete solution.	7 p
(D) Not checking the solution $f(x) = x$ (stating that it is clear is OK).	-1 p

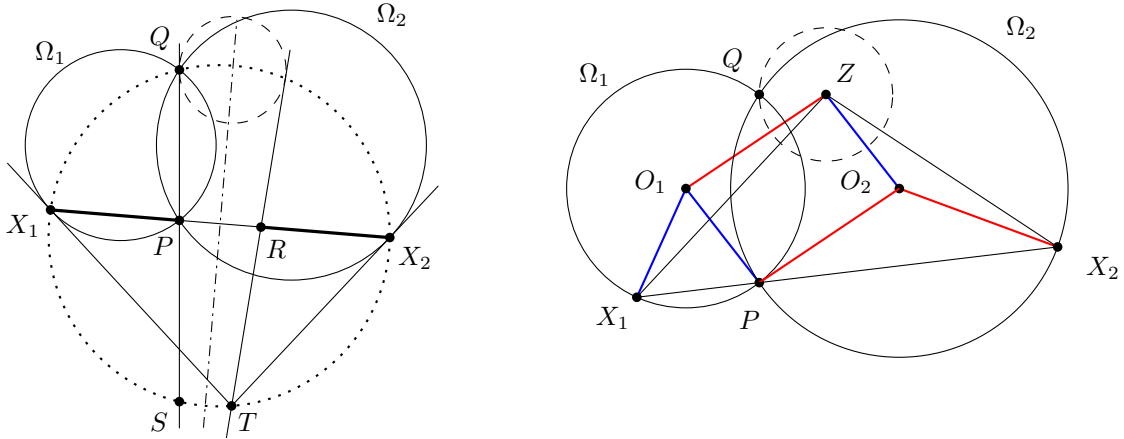
Define $A = \sum_{i=1}^4 A_i$. Give $\max\{A, L, C + D\}$ points.

Problem 3. Circles Ω_1 and Ω_2 with different radii intersect at two points, denote one of them by P . A variable line ℓ passing through P intersects the arc of Ω_1 which is outside of Ω_2 at X_1 , and the arc of Ω_2 which is outside of Ω_1 at X_2 . Let R be the point on segment X_1X_2 such that $X_1P = RX_2$. The tangent to Ω_1 through X_1 meets the tangent to Ω_2 through X_2 at T . Prove that line RT is tangent to a fixed circle, independent of the choice of ℓ . (Josef Tkadlec)

Solution. Denote the other intersection of Ω_1 and Ω_2 by Q . First we angle-chase that points Q, X_1, T, X_2 are concyclic: Indeed,

$$180^\circ - \angle X_2TX_1 = \angle TX_1X_2 + \angle X_1X_2T = \angle X_1QP + \angle PQX_2 = \angle X_1QX_2.$$

Moreover, if we denote by S the second intersection of PQ and the circumcircle of QX_1TX_2 then $ST \parallel X_1X_2$: Indeed, $\angle X_1QS = \angle TX_1X_2$. Hence RT is a reflection of PQ about the perpendicular bisector of X_1X_2 and it suffices to prove that this perpendicular bisector passes through a fixed point not on PQ , independent of the choice of ℓ – all lines RT will then be tangent to a circle with center at this fixed point that is tangent to PQ .



Let O_1, O_2 be the centers of Ω_1, Ω_2 and r_1, r_2 their radii. We claim that the desired fixed point is the fourth vertex Z of parallelogram O_1PO_2Z . Since $ZO_2 = O_1P = r_1 = X_1O_1$ and $O_2X_2 = r_2 = O_2P = O_1Z$, it suffices to prove that $\angle X_2O_2Z = \angle X_1O_1Z$. And this is straightforward angle-chasing again: E.g. by looking at a (possibly self-intersecting) pentagon $ZO_1X_1X_2O_2$ it suffices to show $\angle X_1O_1Z + \angle ZO_2X_2 = 360^\circ$ and we indeed have

$$\begin{aligned} 540^\circ - \angle X_1O_1Z + \angle ZO_2X_2 &= \angle O_1ZO_2 + \angle O_2X_2X_1 + \angle X_2X_1O_1 \\ &= \angle O_2PO_1 + \angle X_2PO_2 + \angle O_1PX_1 = 180^\circ. \end{aligned}$$

Another Solution. We sketch another way to finish the solution after proving that line TR is the reflection of PQ about the perpendicular bisector of X_1X_2 .

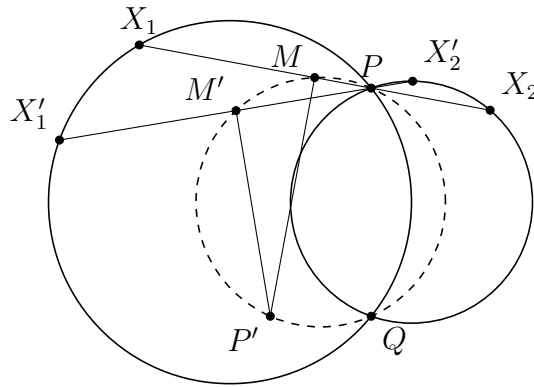
Claim. All the midpoints of the segments X_1X_2 lay on a circle.

Proof. Angle chasing or spiral similarity lemma gives that $\triangle X_1QX_2 \sim \triangle X_1'QX_2'$. Therefore the spiral similarity that takes $\triangle X_1QX_2$ to $\triangle X_1'QX_2'$, also takes M to M' . Finally, angle chasing gives

$$\angle MPM' = \angle X_1PX_1' = \angle X_1QX_1' = \angle MQM'. \quad \square$$

Next, consider the antipodal point to P on (MPQ) . Call it P' . Then by Thales theorem $\angle P'MP = 90$. Thus all perpendicular bisectors intersect at P' .

Finally, construct the circle with center P' tangent to PQ and denote it by Ω . Then all the lines ℓ are tangent to Ω :



Indeed, since MP' is a diameter of Ω and RT is the reflection of PQ across MP' , RT has to also be a tangent.

Marking scheme.

- (A_1) Proving that Q, X_1, X_2, T are concyclic. +1 p
- (A_2) Proving that $ST \parallel X_1X_2$ (or that PQ is the reflection of TR across the perpendicular bisector of X_1X_2). +1 p
- (A_3) Reducing the problem to proving that the bisector of X_1X_2 passes through a fixed point. 3 p
- (B) Proving that the bisector of X_1X_2 passes through a fixed point. +2 p
- (C) An essentially complete solution. 7 p
- (D) Not excluding the scenario “ ℓ passes through a fixed point” (which is the case when the two circles are congruent). -1 p

Define $A = \max\{A_1 + A_2, A_3\}$. Give $\max\{A + B, C + D\}$ points.

Problem 4. Given a positive integer n , denote by $\tau(n)$ the number of positive divisors of n , and by $\sigma(n)$ the sum of all positive divisors of n . Find all positive integers n satisfying

$$\sigma(n) = \tau(n) \cdot \lceil \sqrt{n} \rceil.$$

(Here, $\lceil x \rceil$ denotes the smallest integer not less than x .) (Michael Reitmair)

Solution. Answer: $n \in \{1, 3, 5, 6\}$.

We consider two cases:

- (1) n is a square: $n = 1$ is a solution, so assume $n > 1$. Squares have an odd number of positive divisors, so we write $\tau(n) = 2k + 1$ with $k > 0$ and $d_{k+1} = \sqrt{n}$. Furthermore, $\lceil \sqrt{n} \rceil = \sqrt{n}$, so the given equation is equivalent to

$$d_1 + \cdots + d_k + d_{k+2} + \cdots + d_{2k+1} = 2k \cdot \sqrt{n}$$

For $i \in \{1, \dots, k\}$, d_i and d_{2k+2-i} are complementary divisors, i.e. $d_i \cdot d_{2k+2-i} = n$. Thus, we obtain an equivalent equation:

$$\begin{aligned} & (d_1 - 2\sqrt{n} + d_{2k+1}) + (d_2 - 2\sqrt{n} + d_{2k}) + \cdots + (d_k - 2\sqrt{n} + d_{k+2}) = 0 \\ \Leftrightarrow & \left(\sqrt{d_1} - \sqrt{d_{2k+1}} \right)^2 + \left(\sqrt{d_2} - \sqrt{d_{2k}} \right)^2 + \cdots + \left(\sqrt{d_k} - \sqrt{d_{k+2}} \right)^2 = 0. \end{aligned}$$

Since squares are non-negative and d_1, \dots, d_{2k+1} are pairwise distinct, the equation has no solutions.

- (2) n is not a square: Then no divisor of n is its own complementary divisor, so n has an even number of positive divisors and we write $\tau(n) = 2k$. We shall prove the following inequality for n sufficiently large:

$$d_1 + \cdots + d_{2k} > 2k \cdot \lceil \sqrt{n} \rceil$$

First, note that $\sqrt{n} + 1 \geq \lceil \sqrt{n} \rceil$, so it suffices to show

$$d_1 + \cdots + d_{2k} > 2k \cdot (\sqrt{n} + 1).$$

For $i \in \{1, \dots, k\}$, d_i and d_{2k+1-i} are complementary divisors, so $d_i \cdot d_{2k+1-i} = n$. Hence we obtain an equivalent inequality:

$$\begin{aligned} & (d_1 - 2\sqrt{n} + d_{2k}) + (d_2 - 2\sqrt{n} + d_{2k-1}) + \cdots + (d_k - 2\sqrt{n} + d_{k+1}) > 2k \\ \Leftrightarrow & \left(\sqrt{d_1} - \sqrt{d_{2k}} \right)^2 + \left(\sqrt{d_2} - \sqrt{d_{2k-1}} \right)^2 + \cdots + \left(\sqrt{d_k} - \sqrt{d_{k+1}} \right)^2 > 2k. \end{aligned}$$

But for n sufficiently large, the term $\left(\sqrt{d_1} - \sqrt{d_{2k}} \right)^2 = (\sqrt{n} - 1)^2$ is already bigger than the right-hand side: Indeed, the numbers $\lfloor \frac{n}{2} \rfloor + 1, \lfloor \frac{n}{2} \rfloor + 2, \dots, n - 1$ do not divide n , so $2k = \tau(n) \leq \lfloor \frac{n}{2} \rfloor + 2 \leq \frac{n}{2} + 2$. As $\frac{n}{2}$ grows faster than $2\sqrt{n}$ for large n , we will get the desired estimate. Indeed, this happens for $n \geq 20$, since

$$\begin{aligned} \frac{n}{2} &= \frac{\sqrt{n}}{2} \cdot \sqrt{n} \geq \frac{\sqrt{20}}{2} \cdot \sqrt{n} = 2\sqrt{n} + (\sqrt{5} - 2)\sqrt{n} \geq \\ &\geq 2\sqrt{n} + (\sqrt{5} - 2)\sqrt{20} = 2\sqrt{n} + 10 - 2\sqrt{5} > 2\sqrt{n} + 1, \end{aligned}$$

the last inequality resulting from $\sqrt{5} < 9/4$.

Finally, it remains to manually check all non-squares $n \in \{2, 3, \dots, 19\}$. Thus, we obtain solutions 3, 5, and 6, and the complete set of solutions is $\{1, 3, 5, 6\}$.

Another Solution. (for case n not a square, sketch). Let $\tau(n) = 2k$ and $\lceil \sqrt{n} \rceil = s < \sqrt{n} + 1$. We show that for $n \geq 8$ we have $\sigma(n) > 2k \cdot s$, thus it remains to check $n \in \{2, 3, 5, 6, 7\}$ of which only $n \in \{3, 5, 6\}$ work.

Pairing the divisors d_1, \dots, d_{2k} up such that product in each pair is constant, we get $d_{2k} + d_1 > d_{2k-1} + d_2 > \dots > d_{k+1} + d_k$, where each sum is an integer. Moreover, even the smallest sum satisfies $d_{k+1} + d_k > 2\sqrt{n} > 2(s-1)$, and since both sides are integers we get $d_{k+1} + d_k \geq 2s - 1$. It remains to show that for $n \geq 8$ we have $n + 1 = d_{2k} + d_1 > 2s + 1$, i.e. $n > 2s$. This is simple algebra: it suffices to show $n > 2(\sqrt{n} + 1)$, i.e. $(\sqrt{n} - 1)^2 > 3$, which is true for $n > (\sqrt{3} + 1)^2 = 4 + 2\sqrt{3} \doteq 7.5$. (As a slightly weaker bound, for $n \geq 9$ (thus $\sqrt{n} \geq 3$) we have $n \geq 2\sqrt{n} + \sqrt{n} > 2(\sqrt{n} + 1) > 2s$).

Marking scheme.

- | | |
|--|------|
| (A ₁) Correct answer. | +1 p |
| (A ₂) Solving the case “ n is a square”. | +2 p |
| (A ₃) Solving the case “ $n > x$ is not a square”, for some $x \leq 100$. | +3 p |
| (C) An essentially complete solution. | 7 p |
| (D) An incorrect answer. | -1 p |

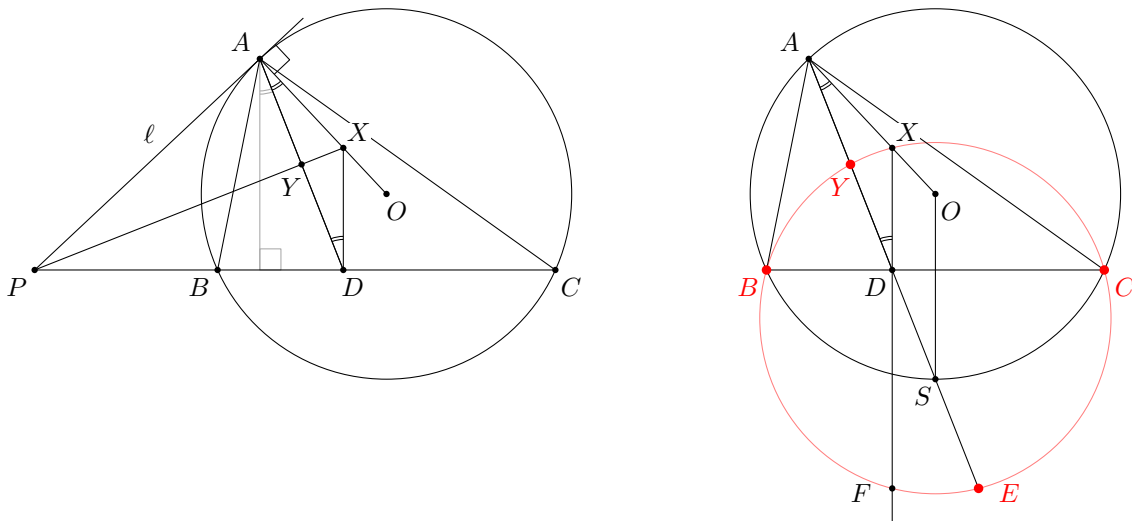
Define $A = \max\{A_1 + A_2, A_3\}$. Give $\max\{A, C + D\}$ points.

Problem 5. Let ABC be a triangle with $AB < AC$ and circumcenter O . The angle bisector of $\angle BAC$ meets the side BC at D . The line through D perpendicular to BC meets the segment AO at X . Furthermore, let Y be the midpoint of segment AD . Prove that points B, C, X, Y are concyclic. (Karl Czakler)

Solution. First, we angle-chase that the triangle ADX is X -isosceles: By isogonal conjugation of circumcenter and orthocenter, the angle bisector AD also bisects the angle between AO and the A -altitude. Since DX is parallel to this altitude, the angles $\angle XDA$ and $\angle DAX$ are equal.

Next, let ℓ be the tangent to the circle (ABC) through A and let P be its intersection with BC . Since $PAXD$ is a kite and XY is the perpendicular bisector of its diagonal AD , the line XY passes through P too.

Finally, from the right triangle PAX and by the power of A with respect to the circle (ABC) we obtain $PY \cdot PX = PA^2 = PB \cdot PC$, thus $BCXY$ is cyclic.



Another Solution. Denote by S the midpoint of arc BC not containing A . It is well known that S lies on line AD , and that OS is the perpendicular bisector of segment BC . Let E be the reflection of D across the point S , and let F be the reflection of E across the perpendicular bisector of BC .

Since the quadrilateral $ABSC$ is inscribed, we have $DB \cdot DC = DA \cdot DS$, and furthermore, $DS = \frac{1}{2}DE$ and $DA = 2DY$ by definition. Thus, $DB \cdot DC = DY \cdot DE$ as well, hence B, Y, C and E are concyclic. Since $BCEF$ is an isosceles trapezoid, point F lies on this circle too.

As in the first solution we angle-chase that triangle DXA is X -isosceles. Thus, $\angle DYX = \angle EYX = 90^\circ$, and since $\angle EFX = 90^\circ$ as well, point X lies on the circumcircle of triangle EYF , as do points B and C , finishing the proof.

Another Solution.

Claim: $XY \perp AD$.

Let $S = AD \cap \odot ABC$. Then $XD \parallel OS$, implying that $\triangle AXD \sim \triangle AOS$. Hence $\triangle AOS$ is isosceles and we are done.

Consider the inversion around A with radius $\sqrt{AB \cdot AC}$, followed by a reflection over the angle bisector of $\angle BAC$. This transformation has the following known properties:

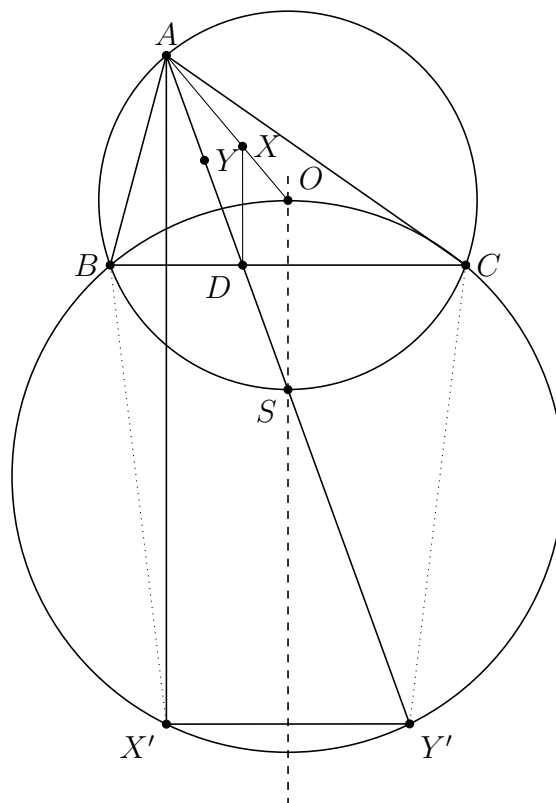
- $B \leftrightarrow C$.
- $BC \leftrightarrow \odot ABC$.

So D is sent to S . Furthermore, since Y is the midpoint of AD , D' has to be the midpoint of Y' . Thus Y' is the reflection of A across S .

It is well known that AO becomes the altitude through A after reflection over the angle bisector of $\angle BAC$. Together with the fact that

$$90 = \angle AYX = \angle AX'Y'$$

one can conclude that X' is the projection from Y' to the altitude of A . Thus $X'Y' \parallel BC$. OS is the perpendicular bisector of BC . Due to Thales theorem $X'S = Y'S$, meaning OS is also the perpendicular bisector of $X'Y'$. Finally, one can see that $X'Y'CB$ is an isosceles trapezoid and therefore has a circumcircle.



Marking scheme.

- | | |
|---|------|
| (A ₁) Proving that $\triangle ADX$ is X -isosceles (implicitly is enough). | +2 p |
| (A ₂) Proving that BC , XY , and ℓ are concurrent. | 3 p |
| (C) An essentially complete solution. | 7 p |
| (D) Using configuration-sensitive angle-chasing. | -0 p |
| (X) An analytical / complex numbers solution that misses a substantial amount of computation. | 0 p |

Give $\max\{A_1, A_2, C + D\}$ points.

Problem 6. Consider 26 letters A, \dots, Z . A *string* is a finite sequence consisting of those letters. We say that a string s is *nice* if it contains each of the 26 letters at least once, and each permutation of letters A, \dots, Z occurs in s as a subsequence the same number of times. Prove that:

- (a) There exists a nice string.
- (b) Any nice string contains at least 2022 letters.

(Here, a permutation π of the 26 letters is as a *subsequence* of a string s if there exist 26 indices $i_1 < i_2 < \dots < i_{26}$ such that $\pi = s_{i_1}s_{i_2}\dots s_{i_{26}}$.) (Václav Rozhoň)

Solution. Let $n(t, s)$ be the number of occurrences of t in s as a subsequence.

Part (a). We will construct a nice string by the following inductive process. First, let $s_1 = AB\dots Z$. Next, for each $1 \leq i < 26$ we define s_{i+1} from s_i as follows. For a string s and a function $\pi: \{A, B, \dots, Z\} \rightarrow \{A, B, \dots, Z\}$ permuting the alphabet, we let the π -*version* of s be the string $\pi(s) = \pi(s_1)\pi(s_2)\dots\pi(s_{|s|})$. Consider all $26!$ permuting functions $\pi_1, \pi_2, \dots, \pi_{26!}$; we set

$$s_{i+1} = \pi_1(s_i)\pi_2(s_i)\dots\pi_{26!}(s_i).$$

That is, s_{i+1} is the concatenation of all versions of s_i in an arbitrary order.

Next, we prove that s_{26} is a nice string. For $1 \leq i \leq 26$, consider the set T_i of all strings of length i containing unique letters from $\{A, \dots, Z\}$. We prove that the number of occurrences of each string $t \in T_i$ in s_i is the same which, in particular, implies that s_{26} is nice. We prove this by induction; it holds for $i = 1$, in fact $n(\ell, s_1) = 1$ for all letters ℓ . For $i > 1$, consider any two strings $t, t' \in T_i$ and we show that $n(t, s_i) = n(t', s_i)$. We consider separately the occurrences of t, t' in s_i that are fully contained in some $\pi_j(s_{i-1})$, $1 \leq j \leq 26!$, and those that are not.

The number of occurrences of t in s_i of the first type is equal to $\sum_{1 \leq j \leq 26!} n(t, \pi_j(s_{i-1}))$. This expression does not change its value if we change t to t' , since it is a sum over all π -versions of s_{i-1} and so it remains the same when we rename the letters.

The number of occurrences of t in s_i of the second type is a sum over all the ways how to distribute the i letters of t in different versions of s_{i-1} without distributing all letters to some $\pi_j(s_{i-1})$. For each such way, if we use t_j for the subsequence of t that we distribute to $\pi_j(s_{i-1})$, we get the number of such occurrences of t in s by computing the product $\prod_{1 \leq j \leq 26!} n(t_j, \pi_j(s_{i-1}))$ (we set $n(\emptyset, s) = 1$). Since for all j we have $|t_j| < |t|$, by induction replacing t by t' does not change the value of any term in the product, and hence the number of all occurrences of the second type is the same too.

Part (b). We will prove that for any nice string s we have $\max_{\ell \in \{A, \dots, Z\}} n(\ell, s) > 2022$.

To do so, we first prove that if we remove all occurrences of a letter, say, Z , from s , the resulting string s' is still nice (under the natural generalization of a string being nice for general alphabets). This is because the number of occurrences of each new 25-letter permutation, e.g. $AB\dots Y$, can be computed as

$$n(AB\dots Y, s') = \frac{n(ZAB\dots Y, s) + n(AZB\dots Y, s) + \dots + n(AB\dots YZ, s)}{n(Z, s)}$$

and all terms in the numerator on the right hand side have the same value for all permutations of A, B, \dots, Y . Iterating this argument (e.g. by straightforward induction) we see that restricting a nice string s to any (nonzero) number of letters gives a nice string.

Next, consider any nice string s' over an alphabet \mathcal{A} consisting of $|\mathcal{A}| = p$ letters, for p a prime. The number of occurrences of any permutation on p letters in s

needs to be exactly $\prod_{\ell \in \mathcal{A}} n(\ell, s')/p!$, so in particular for at least one $\ell \in \mathcal{A}$ we have $p|n(\ell, s')$.

Note that this property needs to hold for *any* subsequence of s that we get by dropping all occurrences of all but some p different letters from it. In particular, there can be at most $p-1$ letters of s such that $p \nmid n(\ell, s)$, for any prime $1 < p < 26$.

Consider the primes 2, 3, 5, 7, 11. As $26 - 1 - 2 - 4 - 6 - 10 = 3 \geq 1$, there exists a letter $\ell \in \{A, \dots, Z\}$ such that each of these five primes divides $n(\ell, s)$. Hence, $n(\ell, s) \geq 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310 > 2022$, as needed.

Remark. In the final step of part (b), one can get a better estimate by using AM-GM inequality instead of asserting existence of one letter whose number of occurrences (“frequency”) is divisible by lot of primes. That is, look at the frequencies $\{n(\ell, s) \mid \ell \in \{A, \dots, Z\}\}$ of the 26 letters in s . Then for any prime $p < 26$, at least $26 - (p-1) = 27 - p$ of those frequencies are multiples of p . By AM-GM we thus get

$$|s| = \sum_{\ell \in \{A, \dots, Z\}} n(\ell, s) \geq 26 \cdot \sqrt[26]{2^{25} 3^{24} \dots 23^4} > 500\,000.$$

Marking scheme.

- | | |
|-----------------------|------|
| (A) Solving part (a). | +4 p |
| (B) Solving part (b). | +3 p |

Give $A + B$ points.