Solutions to Math Prize Olympiad 2023

Advantage Testing Foundation / Jane Street

Fall/Winter 2023

§1 Problem 1, proposed by Oleg Kryzhanovsky

We present one of many possible constructions.

Define $c = \left\lfloor \frac{n-6}{3} \right\rfloor$. Start with the identity permutation $p_k = k$, and let $S = \sum_1^n kp_k$. The idea is the following:

Claim — If $p_i = i$ and $p_j = j$ for some $i < j$, then switching to $p_i = j$ and $p_j = i$ decreases S by $(i-j)^2$.

Proof. The change is $i \cdot i + j \cdot j - i \cdot j - j \cdot i = (i-j)^2$. □

Then consider the following set of $c + 3$ possible operations:

- Switching $p_1 = 1$ with $p_2 = 2$ would decrease S by 1.
- Switching $p_3 = 3$ with $p_4 = 4$ would decrease S by 1.
- Switching $p_5 = 5$ with $p_6 = 6$ would decrease S by 1.
- Switching $p_7 = 7$ with $p_9 = 9$ would decrease S by 4.
- Switching $p_{10} = 10$ with $p_{12} = 12$ would decrease S by 4.
- Switching $p_{13} = 13$ with $p_{15} = 15$ would decrease S by 4.
- Switching $p_{16} = 16$ with $p_{18} = 18$ would decrease S by 4.
- ...
- Switching $p_{3c+4} = 3c + 4$ with $p_{3c+6} = 3c + 6$ would decrease S by 4.

If we performed exactly b of the first three operations ($0 \le b \le 3$), and exactly a of the latter c operations ($0 \le a \le c$), then the value of S would decrease by $4a + b$. By varying a and b, we can perform any decrease between 0 and $4c + 3$, respectively.

Now, since $n \ge 2023$, the inequality $n \le 4c + 3$ is clearly true. If k is the remainder when $1^2 + 2^2 + \cdots + n^2$ is divided by n, then it follows we can decrease S by k using this procedure, so that S becomes a multiple of n.

§2 Problem 2, proposed by Anant Mudgal, Sahil Mhaskar, Sutanay Bhattacharya

The answer is that Will cannot guarantee a win; Fitz has a winning strategy.

The strategy for Fitz is as follows:

- Fitz start by setting the x^1 coefficient to 0.
- Thereafter, Fitz pairs the remaining terms into 49 pairs: the pairs are (x^2, x^3), (x^4, x^5), ..., up to (x^{98}, x^{99}). Whenever Will fills a blank for one of these terms with 0 or 1, Fitz fills the blank for the paired term with the same number.

This means that the resulting polynomial must be of the form

$$P(x) = x^{100} + (x+1)(x^{2k_1} + x^{2k_2} + \cdots + x^{2k_m}) + 1$$

for some exponents $1 \le k_1 \le k_2 \le \cdots \le k_m \le 49$, where m is the number of times that Will filled with 1 rather than 0.

Claim — Such a polynomial $P(x)$ only takes positive real values over $\mathbb{R}$ (and in particular has no real zeroes).

Proof. Obviously, if $x \ge 0$ then $P(x) > 0$. So let $x = -r$ for some $r > 0$, so that

$$P(-r) = r^{100} + (1-r)\left(r^{2k_1} + \cdots + r^{2k_m}\right) + 1.$$

If $r \le 1$, then it's clear that $P(r) > 0$ again. For $r > 1$, we instead write

$$P(-r) = (r^{100} - r^{2k_1+1}) + (r^{2k_1} - r^{2k_2+1}) + \cdots + (r^{2k_m} - r) + r + 1$$

which is again positive, since each parenthesized term is positive. □

§3 Problem 3, proposed by Ankan Bhattacharya

The only possible value of k is $k = 100$. More generally, if 100 is replaced by an $n \ge 1$, the answer is $k = n$.

First, we make a cosmetic rewriting. We let V denote the set of vectors of length n whose entries are either 0 or 1, and where all addition is taken modulo 2. (The *zero vector* $\mathbf{0}$ is the one with all components 0; all other elements of V are said to be *nonzero*.) We can identify each number $s \in S$ with a vector of V by taking the i^{th} coordinate to be nonzero if s is divisible by the i^{th} prime. Now the condition that three elements have a square product is equivalent to the corresponding three vectors having sum zero.

Under this notation, the problem statement can be rephrased more naturally as follows:

> The $2^n - 1$ nonzero elements of V need to be colored with k colors (each color used at least once) such that if $a, b, c \in V$ are nonzero and satisfy $a+b+c = \mathbf{0}$, then exactly two different colors are used among them.

¶ Example of valid coloring with exactly n colors. Given each $a \in V$, if the first 1 is the i^{th} coordinate, color a by the i^{th} color.

¶ Proof any coloring uses exactly n colors. We prove this by induction on $n \geq 1$. Since the base case $n = 1$ is obvious, we focus entirely on the inductive step.

The main claim is the following:

Claim (Key claim) — If a, b, and c are the same color and $a + b + c \neq \mathbf{0}$, then $a + b + c$ is also this color.

Proof. Suppose not: let a, b, c be "red" and $a + b + c$ be "blue". Then,

$$\underbrace{b}_{\text{red}} + \underbrace{c}_{\text{red}} = \underbrace{a}_{\text{red}} + \underbrace{(a+b+c)}_{\text{blue}}$$

implies $b + c$ must be blue. Similarly, $c + a$ and $a + b$ are blue. This is a contradiction, because these three elements sum to zero. □

> **Remark.** This main claim has the property that it can also be verified by manually verifying it is true when $n = 3$, and then applying to the general situation by simply looking at "just" the eight elements $\{0, a, b, c, a + b, b + c, c + a, a + b + c\}$.

In particular, this implies:

Claim — Pick any color. The number of times it is used is either 1 or an even number.

Proof. Call the color "red". Suppose there are at least two red elements; say a and b. Note that

$$c := a + b \neq \mathbf{0}$$

is *not* red.

By the key claim, if x has red, then so is $x + c$, and conversely, if $x + c$ has red, then so does x.

It follows that the elements of red split into pairs which sum to c, so there must be an even number of them. □

> **Remark.** Linear algebra experts will recognize that it must actually be 0 or a power of 2 in the above claim, but this isn't needed.

Since $2^n - 1$ is odd, some color must appear exactly once. Suppose that element v has a unique color, and WLOG v has a 1 in the n^{th} component. Then, x and $x + v$ have the same color for all $x \neq \mathbf{0}$; moreover, exactly one of these two has a 0 in the n^{th} components.

This means we can quotient out by v, as follows. Take the $2^{n-1} - 1$ pairs $\{x, x + v\}$ described above. For each pair, take the vector with a 0 in the n^{th} component, and delete this component to get a vector of length $n - 1$. If we do this process, we obtain one (nonzero) vector of length $n - 1$ from each pair, together with a valid coloring with exactly one fewer color (because the unique color for v is the only one that disappeared). That means, by the induction hypothesis, exactly $n - 1$ colors were left. So there were exactly n colors to begin with.

> **Remark.** In fact, this solution shows there is "essentially" only one possible coloring.
>
> To clarify what is meant by "essentially" requires the language of linear algebra. We can view V more abstractly as an $\mathbb{F}_2$-vector space of dimension n, without paying attention to

any particular choice of coordinates. Then the above solution, modified to use this higher language, implies that there *exists* a basis $e_1, \dots, e_n$ of V for which a vector $v \in V$ is colored with the i^{th} color if i is lowest index for which the coefficient of e_i in v is nonzero. Moreover, as we saw, any choice of basis yields a (different) valid coloring.

In particular, in the original problem, the *number* of valid colorings (up to permutations of the colors) turns out to be equal exactly to

$$2^{\binom{n}{2}}(2^1-1)(2^2-1)\dots(2^n-1).$$

because this is the number of ways to pick a basis of an n-dimensional $\mathbb{F}_2$-vector space.

Remark. When $n=3$, a fun corollary of this problem is that any 2-coloring of the points of the Fano plane has a monochromatic line. This fact is used in the niché board game Fire and Ice to ensure that the game can't end in a draw.

¶ Alternative proof that any coloring uses exactly n colors, found by the only contestant who solved this problem. Given two vectors a and b, we define $a \succ b$ to mean that a and b have different colors, and $a+b$ has the same color as a. The heart of this solution is based on the following observation:

Claim — The relation $\succ$ is transitive, meaning if $a \succ b$ and $b \succ c$ then $a \succ c$.

Proof. Suppose a and $a+b$ are "red", while b and $b+c$ is "blue". Then c is not blue, so we consider two cases.

- Suppose c is also red. Then

$$\underbrace{a}_{\text{red}} + \underbrace{c}_{\text{red}} = (\underbrace{a+b}_{\text{red}}) + (\underbrace{b+c}_{\text{blue}})$$

forces $a+c$ to be blue. However, now consider

$$\underbrace{a}_{\text{red}} + (\underbrace{b+c}_{\text{blue}}) = \underbrace{b}_{\text{blue}} + (\underbrace{a+c}_{\text{blue}}) = \underbrace{c}_{\text{red}} + (\underbrace{a+b}_{\text{red}}).$$

There is no way to color $a+b+c$, giving a contradiction.

- Suppose c is instead a third color, say "green". Then

$$\underbrace{a}_{\text{red}} + \underbrace{c}_{\text{green}} = (\underbrace{a+b}_{\text{red}}) + (\underbrace{b+c}_{\text{blue}})$$

forces $a+c$ to be red, so $a \succ c$. □

Remark. This claim also has the same property as the main claim of the first solution, where its truth for just $n=3$ implies the result for all n.

Since $\succ$ is transitive but anti-reflexive (we never have $v \succ v$), there exists a particular element v which is $\succ$-minimal, meaning $v \not\succ v'$ for every other v'. Let's say it is "black". Then there cannot be any other black element w, since otherwise we would have $v \prec (v+w)$. In other words, we have shown there exists an element with a unique color. We can now proceed as in the first solution above.

§4 Problem 4, proposed by Holden Mui

The proof is divided into three main steps, in ascending order of difficulty:

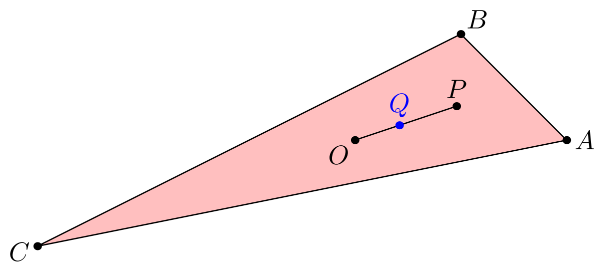
Step 1 Show that if $\overline{OP}$ contains a lattice point, the no marine triangle contains P.

Step 2 Show that if $\overline{OP}$ has rational slope but contains no lattice points, then P is contained inside some marine triangle.

Step 3 Show that if $\overline{OP}$ has irrational slope but contains no lattice points, then P is contained inside some marine triangle.

Step 3 is the most difficult step and we provide several approaches for it.

¶ Proof of Step 1, using Pick's theorem. It would be sufficient to prove that a marine triangle cannot contain both O and a second lattice point $Q \neq O$.



Indeed, this follows by *Pick's theorem*: any triangle which contains both O and Q would have area at least $\frac{3}{2} + 2 - 1 = \frac{5}{2}$, so cannot be marine.

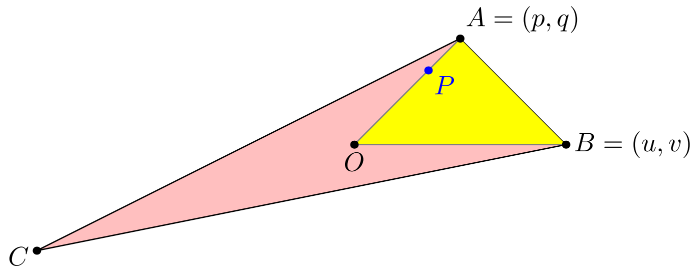
¶ Common boilerplate for Steps 2 and 3. Before proceeding, we remark that in any marine triangle ABC, because O is the centroid, it follows that $[AOB] = [BOC] = [COA] = 1/2$. This gives an equivalent definition of marine triangle via areas.

We will assume henceforth without loss of generality that the coordinates of $P = (u, v)$ satisfy $0 < u < v$. This follows by eliminating several edge cases in turn:

- We may assume P lies in the first quadrant by rotating quadrants.
- We also henceforth discard the case where P lies on one of the coordinate axes. These cases are easy to deal with by hand; we can use the marine triangle with vertices at $(1, 0)$, $(0, 1)$ and $(-1, -1)$.
- If P lies on the line $x = y$, we may similarly resolve the cases manually. (Note the triangle with vertices $(1, 1)$, $(1, 0)$, $(-2, -1)$ is marine.)
- After that, we may reflect P along the line $x = y$ to assume that its y-coordinate is strictly greater than its x-coordinate.

The nonnegative assumption can be obtained by rotating the quadrants; the assumption $u \leq v$ can be done by flipping over the line $y = x$.

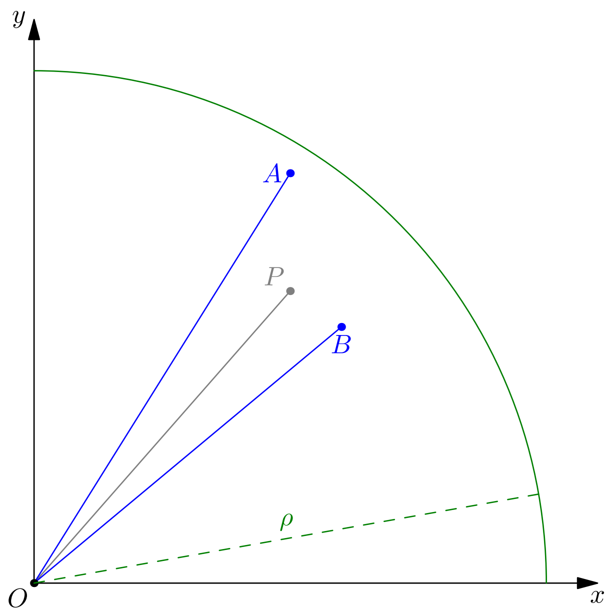
¶ Proof of Step 2. Suppose $\overline{OP}$ has positive rational slope $\frac{p}{q}$, with $1 \le p \le q$ and $\gcd(p,q) = 1$. Let $A = (p,q)$, so that A lies on the extension of ray $\overline{OP}$ past A. Use Bezout's lemma to find nonnegative integers u, v such that $pv - qu = 1$; then if $B = (u,v)$ the triangle OAB is seen to have area $1/2$.



Let $C = -(A+B) = (-(p+u), -(q+v))$. Then $\triangle ABC$ is the desired marine triangle.

¶ First proof of Step 3 using Pick's theorem again. Say a lattice point is *reduced* if its coordinates are coprime positive integers. Also, let m be the slope of line $\overline{OP}$, which is irrational.

Fix a positive number ρ to be determined later. Consider the quarter-disk centered at O of radius ρ in the first quadrant. We let $A = A(\rho)$ be the reduced lattice point with the minimal slope in this disk that lies above the line OP (the slope of $\overline{OA}$ thus gives best rational over-approximation of m among lattice points in this disk). Similarly $B = B(\rho)$ is the reduced lattice point with maximal slope in this disk that lies below the line OP.



Claim — If ρ is chosen large enough, then $\triangle OAB$ contains point P.

Proof. It is sufficient that $\min(OA, OB) > OP$.

Let S denote the finite set of lattice points within distance OP from the origin. Since m is irrational, there exist rational numbers which are larger than m but less than the slopes of the lines joining O to any point in S above line OP. So once ρ is large enough that our quarter-disk contains a point corresponding to such an approximation, then $OA > OP$ will automatically be true.

The proof for $OB > OP$ is similar again for large enough ρ. $\square$

Claim — For any ρ, $\triangle OAB$ has no lattice points inside it and hence has area $1/2$.

Proof. Indeed, if it did, then any lattice point we obtained would give a better approximation to m lying inside triangle OAB and hence inside our quarter-disk, contradicting the definition of A or B! Then the area claim follows by Pick theorem. $\square$

Choose large enough ρ for these claims and then let $C = -(A+B)$ again. Then $\triangle ABC$ is the desired marine triangle.

¶ Second proof of Step 3 using Farey sequences. This proof is not so different from the previous proof, except that instead of Pick's theorem we appeal to the theory of Farey sequences. Specifically, consider consecutive fractions $\frac{p}{q}$ and $\frac{r}{s}$ in the Farey sequence such that

$$\frac{p}{q} < m < \frac{r}{s}.$$

Consecutive fractions in the Farey sequence always obey $|ps - qr| = 1$, so if we set $A = (p,q)$ and $B = (r,s)$ then $\triangle AOB$ has area $1/2$, and upon setting $C = -(A+B) = (-(p+r), -(q+s))$ we would be done. Thus, like in the previous solution, it would be enough that P lies inside AOB.

To do this, it suffices to simply ensure that $\min(q,s) > OP$. Since m is irrational, this follows again from the properties of Farey sequences; the set of rational fractions whose reduced denominators exceed OP is still dense in $\mathbb{R}$.

¶ A third proof of Steps 2 and 3 simultaneously using shear transformations. This approach, submitted by the original author, avoids working with P and instead defines two *shear transformations* as follows:

$$\begin{aligned} s_1 &\colon (x,y) \mapsto (x-y, y) \\ s_2 &\colon (x,y) \mapsto (x, y-x). \end{aligned}$$

The main claim is the following:

Claim — Let P be any point. If $s_1(P)$ lies strictly inside a marine triangle, then so does P. The same is true for $s_2(P)$.

Proof. The shear transformation s_1 has an inverse given by $s_1^{-1}\colon (x,y) \mapsto (x+y, y)$. This inverse preserves the origin and preserves (signed) areas. So if the marine triangle ABC contains $s_1(P)$, then the triangle with vertices $s_1^{-1}(A)$, $s_1^{-1}(B)$, $s_1^{-1}(C)$ is marine and contains the point P.

The proof for s_2 is exactly the same with $s_2^{-1}\colon (x,y) \mapsto (x, x+y)$. $\square$

We also have the following simple observation.

Claim — If P is such that $\overline{OP}$ contains no lattice points (other than O), then there are no lattice points on the line segments joining O to $s_1(P)$ or $s_2(P)$ either.

Proof. The shear transformations (and their inverses) preserve collinearity, betweenness, and send lattice points to lattice points. □

Suppose $P = (x, y)$ is such that $\overline{OP}$ contains no lattice points. Assume WLOG that $0 \le x \le y$ to begin. We apply shear transformations to P in the following algorithm:

Step 1 If $x = 0$ or $x+y < 1$, stop. Otherwise, apply s_1 to P (that is, repeatedly subtract x from the second coordinate) until $y < x$ but still $y \ge 0$. Then go to Step 2.

Step 2 If $y = 0$ or $x+y < 1$, stop. Otherwise, apply s_2 to P (that is, repeatedly subtract y from the first coordinate) until $x < y$ but still $x \ge 0$. Then go back to Step 1.

For example, if $P = (3.14, 20.23)$ then the operation goes

$$(3.14, 20.23) \xmapsto{\text{Step 1}} (3.14, 1.39) \xmapsto{\text{Step 2}} (0.36, 1.39) \xmapsto{\text{Step 1}} (0.36, 0.31)$$

and then terminates. And the algorithm will always terminate eventually, because of the following.

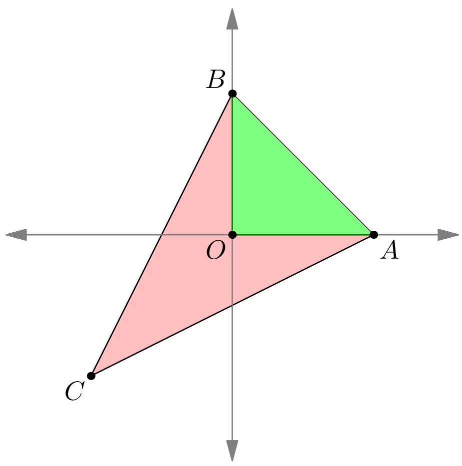
Claim — If $P = (x, y)$ before Step 1, $0 < x \le y$, and $P = (x, y')$ after Step 1, then

$$\frac{x + y'}{x + y} < \frac{2}{3}.$$

In other words, each of Step 1 or Step 2 decreases the sum of coordinates by at least a factor of $\frac{2}{3}$.

Proof. We know $y' = y + nx$ for some $n \ge 1$, and $x > y' \ge 0$. So we need to check $\frac{x+y}{x+y+nx} < \frac{2}{3}$ which is obvious. □

Consider the situation after the algorithm terminated, resulting in a final point $Q = (x, y)$. We claim that this point Q is contained inside the marine triangle with vertices $A = (1, 0)$, $B = (0, 1)$, $C = (-1, -1)$.



We consider three cases.

- If $x = 0$, then because $\overline{OQ}$ also doesn't contain lattice points, we have $y < 1$. Hence Q lies on segment OB.
- If $y = 0$, we similarly have $x < 1$ and Q lies on segment OA.

- Otherwise, the algorithm must have terminated because $x + y < 1$. So in this case Q lies inside the triangle OAB.

Thus Q lies inside this marine triangle. Having completed all cases, we're done.