INMO 2021

Official Solutions

Problem 1. Suppose $r \geq 2$ is an integer, and let $m_1, n_1, m_2, n_2, \cdots, m_r, n_r$ be $2r$ integers such that

$$|m_i n_j - m_j n_i| = 1$$

for any two integers i and j satisfying $1 \leq i < j \leq r$. Determine the maximum possible value of r.

Solution. Let m_1, n_1, m_2, n_2 be integers satisfying $m_1n_2 - m_2n_1 = \pm 1$. By changing the signs of m_2, n_2 if need be, we may assume that

$$m_1n_2 - m_2n_1 = 1.$$

If m_3, n_3 are integers satisfying $m_1n_3 - m_3n_1 = \pm 1$, again we may assume (by changing their signs if necessary) that

$$m_1n_3 - m_3n_1 = 1.$$

So $m_1(n_2 - n_3) = n_1(m_2 - m_3)$.

As m_1, n_1 are relatively prime, m_1 divides $m_2 - m_3$; say, $m_2 - m_3 = m_1a$ for some integer a. Thus, we get $n_2 - n_3 = n_1a$. In other words,

$$m_3 = m_2 - m_1a \ , \ n_3 = n_2 - n_1a.$$

Now, if $m_2n_3 - n_2m_3 = \pm 1$, we get

$$\pm 1 = m_2(n_2 - n_1a) - n_2(m_2 - m_1a) = (m_1n_2 - m_2n_1)a = a.$$

Thus, $m_3 = m_2 - m_1a = m_2 \pm m_1, n_3 = n_2 - n_1a = n_2 \pm n_1$.

Now if we were to have another pair of integers m_4, n_4 such that

$$m_1n_4 - n_1m_4 = \pm 1,$$

we may assume that $m_1n_4 - n_1m_4 = 1$. As seen above, $m_4 = m_2 \mp m_1, n_4 = n_2 \mp n_1$. But then

$$m_3n_4 - n_3m_4 = (m_2 \pm m_1)(n_2 \mp n_1) - (n_2 \pm n_1)(m_2 \mp m_1) = \pm 2.$$

Therefore, there can be only 3 pairs of such integers.

That there do exist many sets of 3 pairs is easy to see; for instance, $(1,0)$, $(1,1)$,$(0,1)$ is such a triple. □

Alternate Solution. It is clear that r can be 3 due to the valid solution $m_1 = 1, n_1 = 1, m_2 = 1, n_2 = 2, m_3 = 2, n_3 = 3$.
If possible, let $r > 3$. We observe that:

$$m_1n_2n_3 - m_2n_1n_3 = \pm n_3$$
$$m_2n_3n_1 - m_3n_2n_1 = \pm n_1$$
$$m_3n_1n_2 - m_1n_3n_2 = \pm n_2$$

Adding, we get $\pm n_1 \pm n_2 \pm n_3 = 0$; which forces at least one of n_1, n_2, n_3 to be even; WLOG let n_1 be even.

Repeating the argument for indices $2, 3, 4$, we deduce that at least one of n_2, n_3, n_4 is even; WLOG let n_2 be even. This leads to a contradiction, since $|m_1n_2 - m_2n_1| = 1$ cannot be even. Hence $r > 3$ is not possible, as claimed. □

Problem 2. Find all pairs of integers (a, b) so that each of the two cubic polynomials

$$x^3 + ax + b \text{ and } x^3 + bx + a$$

has all the roots to be integers.

Solution. The only such pair is $(0, 0)$, which clearly works. To prove this is the only one, let us prove an auxiliary result first.

Lemma If α, β, γ are reals so that $\alpha + \beta + \gamma = 0$ and $|\alpha|, |\beta|, |\gamma| \geq 2$, then

$$|\alpha\beta + \beta\gamma + \gamma\alpha| < |\alpha\beta\gamma|.$$

Proof. Some two of these reals have the same sign; WLOG, suppose $\alpha\beta > 0$. Then $\gamma = -(\alpha + \beta)$, so by substituting this,

$$|\alpha\beta + \beta\gamma + \gamma\alpha| = |\alpha^2 + \beta^2 + \alpha\beta|, \ |\alpha\beta\gamma| = |\alpha\beta(\alpha + \beta)|.$$

So we simply need to show $|\alpha\beta(\alpha + \beta)| > |\alpha^2 + \beta^2 + \alpha\beta|$. Since $|\alpha| \geq 2$ and $|\beta| \geq 2$, we have

$$\begin{aligned} |\alpha\beta(\alpha + \beta)| &= |\alpha||\beta(\alpha + \beta)| \geq 2|\beta(\alpha + \beta)|, \\ |\alpha\beta(\alpha + \beta)| &= |\beta||\alpha(\alpha + \beta)| \geq 2|\alpha(\alpha + \beta)|. \end{aligned}$$

Adding these and using triangle inequality,

$$\begin{aligned} 2|\alpha\beta(\alpha + \beta)| &\geq 2|\beta(\alpha + \beta)| + 2|\alpha(\alpha + \beta)| \geq 2|\beta(\alpha + \beta) + \alpha(\alpha + \beta)| \\ &\geq 2(\alpha^2 + \beta^2 + 2\alpha\beta) > 2(\alpha^2 + \beta^2 + \alpha\beta) \\ &= 2|\alpha^2 + \beta^2 + \alpha\beta|. \end{aligned}$$

Here we have used the fact that $\alpha^2 + \beta^2 + 2\alpha\beta = (\alpha + \beta)^2$ and $\alpha^2 + \beta^2 + \alpha\beta = \left(\alpha + \frac{\beta}{2}\right)^2 + \frac{3\beta^2}{4}$ are both nonnegative. This proves our claim. □

For our main problem, suppose the roots of $x^3 + ax + b$ are the integers r_1, r_2, r_3 and the roots of $x^3 + bx + a$ are the integers s_1, s_2, s_3. By Vieta's relations, we have

$$\begin{gathered} r_1 + r_2 + r_3 = 0 = s_1 + s_2 + s_3 \\ r_1r_2 + r_2r_3 + r_3r_1 = a = -s_1s_2s_3 \\ s_1s_2 + s_2s_3 + s_3s_1 = b = -r_1r_2r_3 \end{gathered}$$

If all six of these roots had an absolute value of at least 2, by our lemma, we would have

$$|b| = |s_1s_2 + s_2s_3 + s_3s_1| < |s_1s_2s_3| = |r_1r_2 + r_2r_3 + r_3r_1| < |r_1r_2r_3| = |b|,$$

which is absurd. Thus at least one of them is in the set $\{0, 1, -1\}$; WLOG, suppose it's r_1.

1. If $r_1 = 0$, then $r_2 = -r_3$, so $b = 0$. Then the roots of $x^3 + bx + a = x^3 + a$ are precisely the cube roots of $-a$, and these are all real only for $a = 0$. Thus $(a, b) = (0, 0)$, which is a solution.
2. If $r_1 = \pm 1$, then $\pm 1 \pm a + b = 0$, so a and b can't both be even. If $a = -s_1s_2s_3$ is odd, then s_1, s_2, s_3 are all odd, so $s_1 + s_2 + s_3$ cannot be zero. Similarly, if b is odd, we get a contradiction.

The proof is now complete. □

Alternate Solution. The only such pair is $(0, 0)$, which clearly works. Let us prove this is the only one. In what follows, we use $\nu_2(n)$ to denote the largest integer k so that $2^k | n$ for any non-zero $n \in \mathbb{Z}$.

If one of the cubics has 0 as a root, say the first one, then $0^3 + 0 \cdot a + b = 0$, so $b = 0$. Then the roots of $x^3 + bx + a = x^3 + a$ are precisely the cube roots of $-a$, and these are all real only for $a = 0$. Thus $(a, b) = (0, 0)$.

So suppose none of the roots are zero. Take the cubic $x^3 + ax + b$, and suppose its roots are x, y, z. We cannot have $\nu_2(x) = \nu_2(y) = \nu_2(z)$; indeed, if we had $x = 2^k x_1, y = 2^k y_1, z = 2^k z_1$ for odd x_1, y_1, z_1, then

$$0 = x + y + z = 2^k(x_1 + y_1 + z_1).$$

But $x_1 + y_1 + z_1$ is odd, and hence non-zero, so this cannot happen.

Thus we can assume WLOG that $\nu_2(x) > \nu_2(y)$. Then the third root is $-(x+y)$. Similarly, the three roots of $x^3 + bx + a$ can be written as $p, q, -(p + q)$ where $\nu_2(p) > \nu_2(q)$. By Vieta's relations,

$$\begin{gathered} xy - x(x + y) - y(x + y) = -(x^2 + xy + y^2) = a = pq(p + q) \\ pq - p(p + q) - q(p + q) = -(p^2 + pq + q^2) = b = xy(x + y) \end{gathered}$$

Suppose $x = 2^k x_1$ and $y = 2^\ell y_1$ for odd x_1, y_1 and $k > \ell$; in particular $k > 0$. Then

$$xy(x+y) = 2^k x_1 \cdot 2^\ell y_1 \cdot (2^k x_1 + 2^\ell y_1) = 2^{k+2\ell} x_1 y_1 (2^{k-\ell} x_1 + y_1).$$

Here $x_1 y_1 (2^{k-\ell} x_1 + y_1)$ is clearly odd, so $\nu_2(xy(x+y)) = k + 2\ell$.

Also,

$$x^2 + xy + y^2 = 2^{2k} x_1^2 + 2^k x_1 \cdot 2^\ell y_1 + 2^{2\ell} y_1^2 = 2^{2\ell} \left(2^{2k-2\ell} x_1^2 + 2^{k-\ell} x_1 y_1 + y_1^2\right).$$

Again, all the terms in the second factor are even except y_1^2, so the entire factor is odd. This means $\nu_2(x^2 + xy + y^2) = 2\ell$. Therefore

$$\nu_2(xy(x+y)) > \nu_2(x^2 + xy + y^2).$$

Similarly, one may show

$$\nu_2(pq(p+q)) > \nu_2(p^2 + pq + q^2).$$

But then

$$\nu_2(b) = \nu_2(xy(x+y)) > \nu_2(x^2 + xy + y^2) = \nu_2(pq(p+q)) > \nu_2(p^2 + pq + q^2) = \nu_2(b).$$

Here we have used the fact that $\nu_2(n) = \nu_2(-n)$ for any integer n. But this is a contradiction, proving our claim. □

Problem 3. Betal marks 2021 points on the plane such that no three are collinear, and draws all possible line segments joining these. He then chooses any 1011 of these line segments, and marks their midpoints. Finally, he chooses a line segment whose midpoint is not marked yet, and challenges Vikram to construct its midpoint using **only** a straightedge. Can Vikram always complete this challenge?

Note: A straightedge is an infinitely long ruler without markings, which can only be used to draw the line joining any two given distinct points.

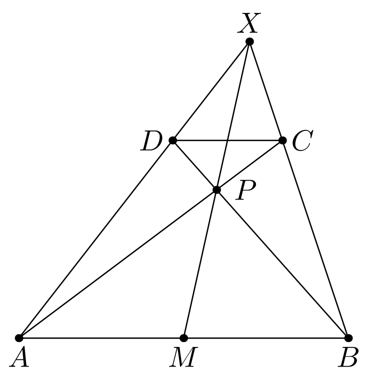
Solution. The answer is 'yes'. To prove this, we will first prove two lemmas:

Lemma 1 Given any two points A, B, their midpoint M, and any point C, Vikram can draw a line parallel to AB through C.

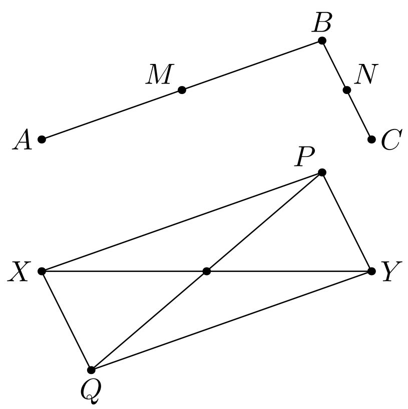
Proof. If C is on line AB we are already done. If not, extend BC to X as shown, draw $P = AC \cap XM$, and then draw $D = BP \cap AX$. We claim CD is the desired line. Indeed, using Ceva's theorem on triangle ABX and the fact $AM = MB$, we see that

$$\frac{AM}{MB} \cdot \frac{BC}{CX} \cdot \frac{XD}{DA} = 1 \implies \frac{XC}{CB} = \frac{XD}{DA}.$$

This means $CD \parallel AB$. □



Lemma 1



Lemma 2

Lemma 2 Given two non-parallel segments AB, BC and their midpoints M, N, Vikram can draw the midpoint of any other segment XY.

Proof. Assume first XY is not parallel to AB or BC. Using lemma 1, draw lines ℓ_1 and ℓ_2 through X parallel to AB and BC respectively, and similarly draw m_1 and m_2 through Y parallel to AB and BC respectively. If we draw $P = \ell_1 \cap m_2$ and $Q = \ell_2 \cap m_1$, then $XPYQ$ is a parallelogram, so intersecting PQ and XY gives the midpoint of XY.

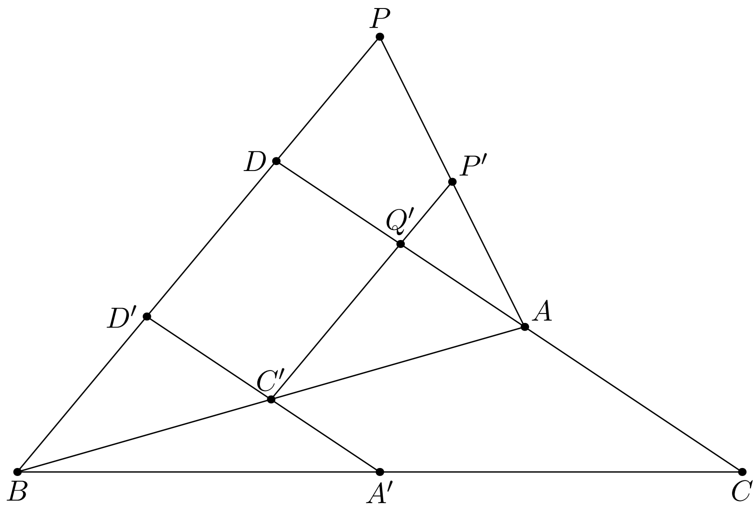
As for the remaining case, one can draw AC and construct the midpoint P of AC by the construction described above. Since XY can be parallel to at most one of the sides AB, BC and AC, we can pick the two non-parallel sides, and use the above construction to draw the midpoint of XY. □

Now for the main problem, note that if no two of the 1011 chosen segments share an endpoint, then we have at least $2 \cdot 1011 = 2022$ distinct endpoints, a contradiction. Thus there must be two segments AB and BC which have their midpoints marked. Since no three of the chosen 2021 points were collinear, AB and BC are not parallel, so using lemma 2, Vikram can construct the midpoint of any other segment, in particular, the segment chosen by Betal. □

Alternate Solution As in the previous solution, note that there exist AB and AC whose midpoints C' and B' are marked. Using the straightedge, Vikram can draw the two medians AC' and AB' and obtain their intersection, the centroid G of $\triangle ABC$. Now intersecting AG with BC gives A', the midpoint of BC.

Lemma Given a point P not on AB, AC, Vikram can draw the midpoint of PA.

Proof. If $PB || AC$ and $PC || AB$, then $PBAC$ is a parallelogram, in which case A' constructed above is the midpoint of PA. Without loss of generality, we may assume $PB \nparallel AC$.



Using the straightedge, one can mark the points $D = PB \cap AC$ and $PB \cap A'C' = D'$. Since $CA || A'C'$, we have

$$\frac{BD'}{D'D} = \frac{BC'}{C'A} = 1,$$

so D' is the midpoint of BD. Now in $\triangle ABD$, two midpoints C' and D' are known, so the midpoint of Q' of AD can be constructed using the centroid construction outlined before. Let $P' = C'Q' \cap PA$; this exists as $C'Q' || BP \nparallel AP$. As before, $C'P' || BP$, so

$$\frac{AP'}{P'P} = \frac{AC'}{C'B} = 1,$$

which means P' is the desired midpoint of PA. □

Now suppose we need to find the midpoint of PQ. If P, Q are different points from A, then one can draw the midpoints of AP and AQ using the lemma. Then by using the centroid of $\triangle APQ$, one can find the midpoint of PQ as we did for BC. If P or Q is A, the above lemma immediately yields the required midpoint. □

Problem 4 A Magician and a Detective play a game. The Magician lays down cards numbered from 1 to 52 face-down on a table. On each move, the Detective can point to two cards and inquire if the numbers on them are consecutive. The Magician replies truthfully. After a finite number of moves the Detective points to two cards. She wins if the numbers on these two cards are consecutive, and loses otherwise.

Prove that the Detective can guarantee a win if and only if she is allowed to ask at least 50 questions.

Solution. *Strategy for the Detective:* Pick a card A and compare against all others except one. If he ever gets a "Yes", that pair works; else the remaining card is consecutive with A. This process takes at most 50 queries.

Strategy for the Magician: We show that it is not always possible to obtain a "Yes" in 50 turns, hence showing that 49 turns are not enough to figure out a consecutive pair. It is enough to conjure a labelling of cards for which denying all 50 inquiries is valid.

Replace 52 by any $N > 3$. Think of the cards as vertices of a complete graph K_N. Delete all edges joining vertices which correspond to pairs of cards the Detective inquired about. We will show that deleting any $N-2$ edges of K_N still leaves a graph that admits a path containing all the vertices. Labelling all cards along this path as 1 to N would finish. Several proofs of this claim are possible. We present three of them.

Proof 1.

For any two vertices a and b, since $\deg a + \deg b \geq 2(N-1) - (N-2) = N$, they share a common neighbour. Hence the graph is connected.

Pick the longest path $\mathcal{P} : u = u_0 \to u_1 \to \cdots \to u_k = v$. All neighbours of u and v must remain within the path, else we could get a longer path. Let u have x neighbours $\{u_{i_1}, u_{i_2}, \ldots, u_{i_x}\}$ with $1 = i_1 < i_2 < \cdots < i_x \leq k$. Let v have y neighbours $\{u_{j_1}, \ldots, u_{j_y}\}$. Since $x + y \geq n$, we see that $i_s = j_r + 1$ for some r and s. Thus there exists i such that $u \to u_{i+1}$ and $u_i \to v$ are edges. Thus the path is a cycle

$$\mathcal{C} = u_{i+1} \to u_0 \to u_1 \cdots \to u_i \to v \to u_{k-1} \cdots \to u_{i+1}.$$

Suppose a vertex w is not in the path $\mathcal{P}$. By connectedness, we have a path $\mathcal{P}'$ from w to some vertex of $\mathcal{P}$. Continue along this path via the cycle $\mathcal{C}$ to obtain a path longer than $\mathcal{P}$; contradiction! Thus the graph has a path of length $N-1$, as desired. □

Proof 2.

Pick the longest cycle $\mathcal{C} = v_1 \to \cdots \to v_k \to v_1$. Note that any vertex w not in the cycle can be incident to no more than $\frac{k}{2}$ of the vertices in it; else there exists i such that wv_i and wv_{i+1} (indices mod k) are edges, so we can put w in to get a longer cycle. Thus our graph is missing at least $\frac{1}{2}k(N-k)$ edges. So $2(N-2) \geq k(N-k)$. Clearly $k > 2$ so we see that $k \in \{N-2, N-1\}$.

Case 1. $k = N-1$. Pick the leftover w outside $\mathcal{C}$. Not all edges from w to the cycle are missing (since only $N-2$ are missing in total), so follow an edge from w to $\mathcal{C}$ and continue along $\mathcal{C}$ to get a path of length $N-1$.

Case 2. $k = N-2$. Pick the leftover a, b outside $\mathcal{C}$. It is clear that both of them have edges to the cycle and ab is also an edge (since $k(N-k) = 2(N-2)$ in this case). So starting at a, going to b, to some vertex of $\mathcal{C}$ and following along $\mathcal{C}$ gives us a path of length $N-1$.

The proof is complete. □

Proof 3.

The idea is to prove the stronger claim by induction on $N \geq 3$: a graph on N vertices with $\binom{N-1}{2} + 2$ edges has a cycle of length N. Deleting the extra edge will give a path of length $N-1$ through all the vertices.

The base case $N = 3$ is trivial. Suppose it holds for all $k \leq N$, we prove it for $N+1$. Since $\frac{2\left(2+\binom{N}{2}\right)}{N+1} > N-2$ we see that some vertex v has degree either $N-1$ or N.

Case 1. If degree of v is $N-1$. Then we have an edge $e = uv$ missing among all the edges through v. Delete v along with all the edges through it in the graph. The induced graph has a cycle of length N. Pick two consecutive vertices that are not u, and append v between them.

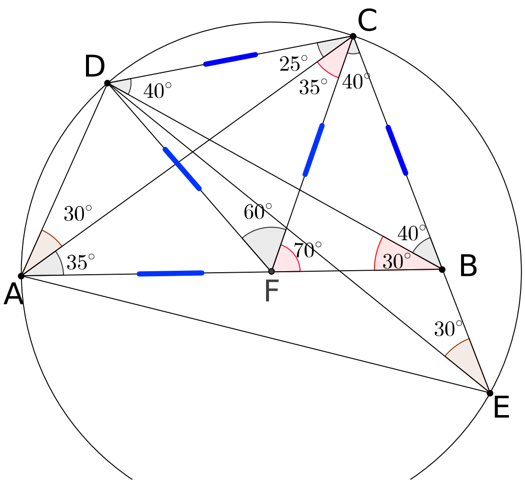
Case 2. If degree of v is N. Delete v along with all its edges. Add an arbitrarily chosen extra edge to the graph so obtained. By induction hypothesis, this resulting graph has a cycle of length N. If removing the extra edge does not disrupt the cycle, append v anywhere between two consecutive vertices. If it does break the cycle, use v to connect the vertices it joined.

The induction is complete. $\square$

Problem 5 In a convex quadrilateral $ABCD$, $\angle ABD = 30°$, $\angle BCA = 75°$, $\angle ACD = 25°$ and $CD = CB$. Extend CB to meet the circumcircle of triangle DAC at E. Prove that $CE = BD$.

Solution. First we show that $\angle DEC = 30°$. Choose a point F on AB such that $CF = CB$. Join FC and FD. Observe that $\angle DCB = 75° + 25° = 100°$. Since $CD = CB$, we have $\angle CDB = \angle CBD = 40°$. Therefore $\angle CBF = 40° + 30° = 70°$. This gives $\angle CFB = 70°$.

Since $CD = CB = CF$, we have the isosceles triangle CDF. But $\angle BCF = 40°$. Hence $\angle FCD = 60°$. Therefore we have an equilateral triangle CDF. This means $FD = FC = CD$ and $\angle DFC = 60°$.

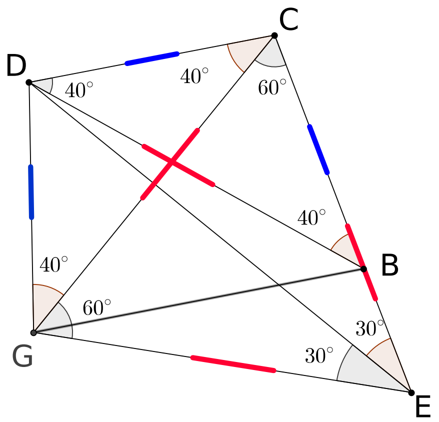


Observe that $\angle AFC = 110°$ and $\angle FCA = 35°$. Hence $\angle FAC = 35°$. This means $FA = FC = FD$. Thus F is the circumcentre of $\triangle ADC$. This implies that

$$\angle CAD = \frac{\angle CFD}{2} = 30°.$$

Therefore $\angle DEC = \angle DAC = 30°$. Now concentrate on triangle DCE.

Construct an equilateral triangle ECG with CE as base, on the side of B. Join GD.



We have $\angle CGE = \angle GCE = \angle CEG = 60°$ and $CE = EG = GC$. Since $\angle CED = 30°$, we get $\angle GED = 30°$. Thus ED is the angle bisector of the isosceles triangle GEC. This implies that ED is also the pependicular bisector of GC. Thus D is on the pependicular bisector of GC. Therefore $DC = DG$ and hence $\angle DGC = \angle DCG$.

But $\angle DCG = 100° - 60° = 40°$. This implies that $\angle DGC = 40°$ and hence $\angle CDG = 100°$.

Consider the quadrilateral $GBCD$. We have $DG = DC = CB$, $\angle GDC = 100° = \angle DCB$. It is an isosceles trapezium. (or we can show that $\triangle GDC \cong \triangle BCD$.) Therefore $DB = GC$. But $GC = CE$. Thus we get $DB = CE$. $\square$

Alternate Solution As in the previous solution, one shows that F is the circumcenter of $\triangle ADC$. since E lies on this circumcircle, this means FE is equal to all of the sides FA, FD, FC and thus also to CD and CB. Now CDB and FCE are both isosceles triangles with base angles $40°$, and they have $CD = FC$, so they are in fact congruent. This directly implies $CE = BD$, as required. □

Problem 6. Let $\mathbb{R}[x]$ be the set of all polynomials with real coefficients, and let $\deg P$ denote the degree of a nonzero polynomial P. Find all functions $f : \mathbb{R}[x] \to \mathbb{R}[x]$ satisfying the following conditions:

- f maps the zero polynomial to itself,
- for any non-zero polynomial $P \in \mathbb{R}[x]$, $\deg f(P) \leq 1 + \deg P$, and
- for any two polynomials $P, Q \in \mathbb{R}[x]$, the polynomials $P - f(Q)$ and $Q - f(P)$ have the same set of real roots.

Solution.

Answer

We have $f(p) = p \forall p \in \mathbb{R}[x]$, or $f(p) = -p \forall p \in \mathbb{R}[x]$. These clearly satisfy the given conditions.

Proof

Claim 1 For all $p \in \mathbb{R}[x]$, $f(f(p)) = p$.

Proof. Using condition 3 on the polynomials p and $f(p)$, we see that $p - f(f(p))$ has the same set of real roots as $f(p) - f(p) = 0$, which is $\mathbb{R}$. Therefore $p - f(f(p))$ is identically zero. □

Note that this implies f is bijective. In what follows, $p \sim q$ will mean that p and q have the same set of real roots. Note that putting $f(q)$ for q in condition 2 gives $p - q \sim f(p) - f(q)$ for all p, q (call this statement $(\star)$). In particular, putting $q = 0$ here, $p \sim f(p)$ for all p (call this $(\star\star)$).

Claim 2 For all non-zero $p \in \mathbb{R}[x]$, $\deg p - 1 \leq \deg f(p) \leq \deg p + 1$.

Proof. The right inequality is simply condition 2. Now using condition 2 on the polynomial $f(p)$, we see that $\deg f(f(p)) \leq \deg f(p) + 1$ which gives $\deg f(p) \geq \deg p - 1$ because of claim 1. □

Claim 3 For all $p \in \mathbb{R}[x]$, $\deg f(p) = \deg p$.

Proof. Note that nonzero constant polynomials have no root, so by $(\star\star)$, their image must have no root. This is impossible if that image has degree 1; so by condition 2, the image has degree 0, i.e., is a constant polynomial. First consider the case when $\deg p$ is even; assume for now the leading coefficient of p is positive. That means $p(x) \to \infty$ for $x \to \pm\infty$, so it has a global minimum, say C. Then the polynomial $p + k$ ($k > C$) has no real roots. Using $(\star)$ on p and the constant polynomial $-k$, we see that $f(p) - f(-k)$ has no roots. But this is impossible if $\deg f(p)$ is odd (since $f(-k)$ is a constant), so by claim 2, we must $\deg f(p) = \deg p$. A similar argument holds if p has negative leading coefficient.

Now if $\deg p$ is odd, then $\deg f(p)$ cannot be even, otherwise $q = f(p)$ would be an even degree polynomial whose image $f(q) = f(f(p)) = p$ has odd degree, contradicting the last paragraph. Thus $\deg f(p)$ is odd, and using claim 2, we infer than $\deg f(p) = \deg p$. □

We call a polynomial p *ninth-grade* if all $\deg p$ roots of p are real and distinct. Clearly for any ninth-grade p, p and $f(p)$ have the roots and same degree, so $f(p) = c_p p$ for some non-zero $c_p \in \mathbb{R}$.

Claim 4 Given any non-constant $q \in \mathbb{R}[x]$, we can choose r with degree bigger than q so that both r and $q - r$ are ninth-grade.

Proof. Assume that all real roots of q are inside the interval $[a, b]$. Now choose a number n which has the same parity as $\deg q$ and is bigger than $\deg q$, and choose numbers $c_1 = a < c_2 < \cdots < c_{n-1} < c_n = b$. Consider the polynomial $p = k(x - c_1)(x - c_2) \cdots (x - c_n)$, so that k has the same sign as the leading coefficient of q (value of k will be chosen later). Clearly p has alternating signs on the intervals $(-\infty, c_1)$, $(c_1, c_2), \cdots, (c_{n-1}, c_n)$, (c_n, ∞), and has the same sign as q outside $[a, b]$. Let $k_1, k_2, \cdots, k_{n-1}$ be the extrema of p on the intervals $[c_1, c_2], \cdots [c_{n-1}, c_n]$ in that order, and suppose they are attained at $x_1, \cdots x_{n-1}$. Make $|k|$ large enough so that $|k_i| > \max_{x \in [a,b]} |q(x)|$ for all i. Then $p + q$ has degree n, and has alternating signs at $a - \epsilon, x_1, \cdots, x_n, b + \epsilon$ for $\epsilon > 0$, so it has exactly n distinct roots. Now it is enough to take $r = -p$. $\square$

Claim 5 For any $q \in \mathbb{R}[x]$, $f(q) = c_q q$ for some non-zero real c_q.

Proof. We have already proved this for ninth-grade polynomials. Take ninth-grade r so that $q - r$is ninth grade and $n = \deg(q - r) > \deg q$. Then $q - r \sim f(q) - f(r) = f(q) - c_r r$. Since $q - r$ is ninth-grade and has the same degree as $f(q) - c_r r$, $q - r = c(f(q) - c_r r) = cf(q) - c_1 r$ for non-zero reals c, c_1. Comparing the leading term (which belongs to r) on both sides, $c_1 = 1$, therefore $q = cf(q) \implies f(q) = c_q q$. $\square$

Claim 6 For any $p, q \in \mathbb{R}[x]$, $c_p = c_q$.

Proof. We note that for any two polynomials p, q if $p - q$ has a real root which is not a root of p, then $c_p = c_q$. Indeed, if s is a root of $p - q$ (meaning $p(s) = q(s) \neq 0$), then it's also a root of $f(p) - f(q) = c_p p - c_q q$, so that $c_p p(s) = c_q q(s) \implies c_p = c_q$.

Now for any two p, q, choose odd N such that $N > \max\{\deg p, \deg q\}$. Then the polynomial $r = x^N$ is such that $r - p$ and $r - q$ both have real roots, so $c_q = c_r = c_p$. $\square$

Claim 6 clearly means there is $c \in \mathbb{R}$ so that $f(p) = cp$ for all $p \in \mathbb{R}[x]$. Using the fact $f(f(p)) = p$, we see that the only possibilities are $c = 1$ or $c = -1$, completing the proof. $\square$