Solutions to the 83rd William Lowell Putnam Mathematical Competition
Saturday, December 3, 2022

Manjul Bhargava, Kiran Kedlaya, and Lenny Ng

A1 Write $f(x) = \ln(1+x^2)$. We show that $y = ax+b$ intersects $y = f(x)$ in exactly one point if and only if (a,b) lies in one of the following groups:

- $a = b = 0$
- $|a| \geq 1$, arbitrary b
- $0 < |a| < 1$, and $b < \ln(1-r_-)^2 - |a|r_-$ or $b > \ln(1-r_+)^2 - |a|r_+$, where
$$r_\pm = \frac{1 \pm \sqrt{1-a^2}}{a}.$$

Since the graph of $y = f(x)$ is symmetric under reflection in the y-axis, it suffices to consider the case $a \geq 0$: $y = ax+b$ and $y = -ax+b$ intersect $y = f(x)$ the same number of times. For $a = 0$, by the symmetry of $y = f(x)$ and the fact that $f(x) > 0$ for all $x \neq 0$ implies that the only line $y = b$ that intersects $y = f(x)$ exactly once is the line $y = 0$.

We next observe that on $[0,\infty)$, $f'(x) = \frac{2x}{1+x^2}$ increases on $[0,1]$ from $f'(0) = 0$ to a maximum at $f'(1) = 1$, and then decreases on $[1,\infty)$ with $\lim_{x\to\infty} f'(x) = 0$. In particular, $f'(x) \leq 1$ for all x (including $x < 0$ since then $f'(x) < 0$) and $f'(x)$ achieves each value in $(0,1)$ exactly twice on $[0,\infty)$.

For $a \geq 1$, we claim that any line $y = ax+b$ intersects $y = f(x)$ exactly once. They must intersect at least once by the intermediate value theorem: for $x \ll 0$, $ax+b < 0 < f(x)$, while for $x \gg 0$, $ax+b > f(x)$ since $\lim_{x\to\infty} \frac{\ln(1+x^2)}{x} = 0$. On the other hand, they cannot intersect more than once: for $a > 1$, this follows from the mean value theorem, since $f'(x) < a$ for all x. For $a = 1$, suppose that they intersect at two points (x_0,y_0) and (x_1,y_1). Then
$$1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\int_{x_0}^{x_1} f'(x)\,dx}{x_1 - x_0} < 1$$
since $f'(x)$ is continuous and $f'(x) \leq 1$ with equality only at one point.

Finally we consider $0 < a < 1$. The equation $f'(x) = a$ has exactly two solutions, at $x = r_-$ and $x = r_+$ for $r_\pm$ as defined above. If we define $g(x) = f(x) - ax$, then $g'(r_\pm) = 0$; g' is strictly decreasing on $(-\infty, r_-)$, strictly increasing on (r_-, r_+), and strictly decreasing on (r_+,∞); and $\lim_{x\to-\infty} g(x) = \infty$ while $\lim_{x\to\infty} g(x) = -\infty$. It follows that $g(x) = b$ has exactly one solution for $b < g(r_-)$ or $b > g(r_+)$, exactly three solutions for $g(r_-) < b < g(r_+)$, and exactly two solutions for $b = g(r_\pm)$. That is, $y = ax+b$ intersects $y = f(x)$ in exactly one point if and only if $b < g(r_-)$ or $b > g(r_+)$.

A2 The answer is $2n-2$. Write $p(x) = a_n x^n + \cdots + a_1 x + a_0$ and $p(x)^2 = b_{2n}x^{2n} + \cdots + b_1 x + b_0$. Note that $b_0 = a_0^2$ and $b_{2n} = a_n^2$. We claim that not all of the remaining $2n-1$ coefficients $b_1, \dots, b_{2n-1}$ can be negative, whence the largest possible number of negative coefficients is $\leq 2n-2$. Indeed, suppose $b_i < 0$ for $1 \leq i \leq 2n-1$. Since $b_1 = 2a_0a_1$, we have $a_0 \neq 0$. Assume $a_0 > 0$ (or else replace $p(x)$ by $-p(x)$). We claim by induction on i that $a_i < 0$ for $1 \leq i \leq n$. For $i = 1$, this follows from $2a_0a_1 = b_1 < 0$. If $a_i < 0$ for $1 \leq i \leq k-1$, then
$$2a_0a_k = b_k - \sum_{i=1}^{k-1} a_i a_{k-i} < b_k < 0$$
and thus $a_k < 0$, completing the induction step. But now $b_{2n-1} = 2a_{n-1}a_n > 0$, contradiction.

It remains to show that there is a polynomial $p(x)$ such that $p(x)^2$ has $2n-2$ negative coefficients. For example, we may take
$$p(x) = n(x^n+1) - 2(x^{n-1} + \cdots + x),$$
so that
$$\begin{aligned} p(x)^2 = n^2(x^{2n} + x^n + 1) - 2n(x^n+1)(x^{n-1} + \cdots + x) \\ + (x^{n-1} + \cdots + x)^2. \end{aligned}$$

For $i \in \{1, \dots, n-1, n+1, \dots, n-1\}$, the coefficient of x^i in $p(x)^2$ is at most $-2n$ (coming from the cross term) plus $-2n+2$ (from expanding $(x^{n-1} + \cdots + x)^2$), and hence negative.

A3 **First solution.** We view the sequence $a_1, a_2, \dots$ as lying in $\mathbb{F}_p^\times \subset \mathbb{F}_p$. Then the sequence is determined by the values of a_1 and a_2, via the recurrence $a_{n+2} = (1+a_{n+1})/a_n$. Using this recurrence, we compute
$$\begin{gathered} a_3 = \frac{1+a_2}{a_1}, \, a_4 = \frac{1+a_1+a_2}{a_1a_2}, \\ a_5 = \frac{1+a_1}{a_2}, \, a_6 = a_1, \, a_7 = a_2 \end{gathered}$$
and thus the sequence is periodic with period 5. The values for a_1 and a_2 may thus be any values in $\mathbb{F}_p^\times$ provided that $a_1 \neq p-1$, $a_2 \neq p-1$, and $a_1 + a_2 \neq p-1$. The number of choices for $a_1, a_2 \in \{1, \dots, p-2\}$ such that $a_1 + a_2 \neq p-1$ is thus $(p-2)^2 - (p-2) = (p-2)(p-3)$.

Since p is not a multiple of 5, $(p-2)(p-3)$ is a product of two consecutive integers $a, a+1$, where $a \not\equiv 2 \pmod{5}$. Now $0 \cdot 1 \equiv 0$, $1 \cdot 2 \equiv 2$, $3 \cdot 4 \equiv 2$, and

$4 \cdot 0 \equiv 0 \pmod 5$. Thus the number of possible sequences $a_1, a_2, \dots$ is 0 or 2 (mod 5), as desired.

Second solution. Say that a sequence is *admissible* if it satisfies the given conditions. As in the first solution, any admissible sequence is 5-periodic.

Now consider the collection S of possible 5-tuples of numbers mod p given by (a_1,a_2,a_3,a_4,a_5) for admissible sequences $\{a_n\}$. Each of these 5-tuples in S comes from a unique admissible sequence, and there is a 5-periodic action on S given by cyclic permutation: $(a,b,c,d,e) \to (b,c,d,e,a)$. This action divides S into finitely many orbits, and each orbit either consists of 5 distinct tuples (if a,b,c,d,e are not all the same) or 1 tuple (a,a,a,a,a). It follows that the number of admissible sequences is a multiple of 5 plus the number of constant admissible sequences.

Constant admissible sequences correspond to nonzero numbers $a \pmod p$ such that $a^2 \equiv 1 + a \pmod p$. Since the quadratic $x^2 - x - 1$ has discriminant 5, for $p > 5$ it has either 2 roots (if the discriminant is a quadratic residue mod p) or 0 roots mod p.

A4 The expected value is $2e^{1/2} - 3$.

Extend S to an infinite sum by including zero summands for $i > k$. We may then compute the expected value as the sum of the expected value of the i-th summand over all i. This summand occurs if and only if $X_1, \dots, X_{i-1} \in [X_i, 1]$ and $X_1, \dots, X_{i-1}$ occur in nonincreasing order. These two events are independent and occur with respective probabilities $(1 - X_i)^{i-1}$ and $\frac{1}{(i-1)!}$; the expectation of this summand is therefore

$$\begin{aligned}
&\frac{1}{2^i (i-1)!} \int_0^1 t(1-t)^{i-1}\,dt \\
&\quad = \frac{1}{2^i(i-1)!} \int_0^1 ((1-t)^{i-1} - (1-t)^i)\,dt \\
&\quad = \frac{1}{2^i(i-1)!} \left(\frac{1}{i} - \frac{1}{i+1} \right) = \frac{1}{2^i(i+1)!}.
\end{aligned}$$

Summing over i, we obtain

$$\sum_{i=1}^\infty \frac{1}{2^i (i+1)!} = 2 \sum_{i=2}^\infty \frac{1}{2^i i!} = 2 \left(e^{1/2} - 1 - \frac{1}{2} \right).$$

A5 We show that the number in question equals 290. More generally, let $a(n)$ (resp. $b(n)$) be the optimal final score for Alice (resp. Bob) moving first in a position with n consecutive squares. We show that

$$\begin{aligned}
a(n) &= \left\lfloor \frac{n}{7} \right\rfloor + a\left(n - 7 \left\lfloor \frac{n}{7} \right\rfloor \right), \\
b(n) &= \left\lfloor \frac{n}{7} \right\rfloor + b\left(n - 7 \left\lfloor \frac{n}{7} \right\rfloor \right),
\end{aligned}$$

and that the values for $n \leq 6$ are as follows:

n	0	1	2	3	4	5	6
$a(n)$	0	1	0	1	2	1	2
$b(n)$	0	1	0	1	0	1	0

Since $2022 \equiv 6 \pmod 7$, this will yield $a(2022) = 2 + \lfloor \frac{2022}{7} \rfloor = 290$.

We proceed by induction, starting with the base cases $n \leq 6$. Since the number of odd intervals never decreases, we have $a(n), b(n) \geq n - 2 \lfloor \frac{n}{2} \rfloor$; by looking at the possible final positions, we see that equality holds for $n = 0,1,2,3,5$. For $n = 4,6$, Alice moving first can split the original interval into two odd intervals, guaranteeing at least two odd intervals in the final position; whereas Bob can move to leave behind one or two intervals of length 2, guaranteeing no odd intervals in the final position.

We now proceed to the induction step. Suppose that $n \geq 7$ and the claim is known for all $m < n$. In particular, this means that $a(m) \geq b(m)$; consequently, it does not change the analysis to allow a player to pass their turn after the first move, as both players will still have an optimal strategy which involves never passing.

It will suffice to check that

$$a(n) = a(n-7) + 1, \qquad b(n) = b(n-7) + 1.$$

Moving first, Alice can leave behind two intervals of length 1 and $n-3$. This shows that

$$a(n) \geq 1 + b(n-3) = a(n-7) + 1.$$

On the other hand, if Alice leaves behind intervals of length i and $n-2-i$, Bob can choose to play in either one of these intervals and then follow Alice's lead thereafter (exercising the pass option if Alice makes the last legal move in one of the intervals). This shows that

$$\begin{aligned}
a(n) &\leq \max\{\min\{a(i) + b(n-2-i), \\
&\qquad b(i) + a(n-2-i)\} : i = 0, 1, \dots, n-2\} \\
&= a(n-7) + 1.
\end{aligned}$$

Moving first, Bob can leave behind two intervals of lengths 2 and $n-4$. This shows that

$$b(n) \leq a(n-4) = b(n-7) + 1.$$

On the other hand, if Bob leaves behind intervals of length i and $n-2-i$, Alice can choose to play in either one of these intervals and then follow Bob's lead thereafter (again passing as needed). This shows that

$$\begin{aligned}
b(n) &\geq \min\{\max\{a(i) + b(n-2-i), \\
&\qquad b(i) + a(n-2-i)\} : i = 0, 1, \dots, n-2\} \\
&= b(n-7) + 1.
\end{aligned}$$

This completes the induction.

A6 **First solution.** The largest such m is n. To show that $m \geq n$, we take

$$x_j = \cos \frac{(2n+1-j)\pi}{2n+1} \qquad (j = 1, \dots, 2n).$$

It is apparent that $-1 < x_1 < \cdots < x_{2n} < 1$. The sum of the lengths of the intervals can be interpreted as

$$\begin{aligned}
&-\sum_{j=1}^{2n}((-1)^{2n+1-j}x_j)^{2k-1}\\
&= -\sum_{j=1}^{2n}\left(\cos(2n+1-j)\left(\pi+\frac{\pi}{2n+1}\right)\right)^{2k-1}\\
&= -\sum_{j=1}^{2n}\left(\cos\frac{2\pi(n+1)j}{2n+1}\right)^{2k-1}.
\end{aligned}$$

For $\zeta = e^{2\pi i(n+1)/(2n+1)}$, this becomes

$$\begin{aligned}
&= -\sum_{j=1}^{2n}\left(\frac{\zeta^j+\zeta^{-j}}{2}\right)^{2k-1}\\
&= -\frac{1}{2^{2k-1}}\sum_{j=1}^{2n}\sum_{l=0}^{2k-1}\binom{2k-1}{l}\zeta^{j(2k-1-2l)}\\
&= -\frac{1}{2^{2k-1}}\sum_{l=0}^{2k-1}\binom{2k-1}{l}\sum_{j=1}^{2n}\zeta^{j(2k-1-2l)}\\
&= -\frac{1}{2^{2k-1}}\sum_{l=0}^{2k-1}\binom{2k-1}{l}(-1) = 1,
\end{aligned}$$

using the fact that $\zeta^{2k-1-2l}$ is a *nontrivial* root of unity of order dividing $2n+1$.

To show that $m \leq n$, we use the following lemma. We say that a multiset $\{x_1,\dots,x_m\}$ of complex numbers is *inverse-free* if there are no two indices $1 \leq i \leq j \leq m$ such that $x_i + x_j = 0$; this implies in particular that 0 does not occur.

Lemma. *Let $\{x_1,\dots,x_m\},\{y_1,\dots,y_n\}$ be two inverse-free multisets of complex numbers such that*

$$\sum_{i=1}^{m}x_i^{2k-1} = \sum_{i=1}^{n}y_i^{2k-1} \qquad (k = 1,\dots,\max\{m,n\}).$$

Then these two multisets are equal.

Proof. We may assume without loss of generality that $m \leq n$. Form the rational functions

$$f(z) = \sum_{i=1}^{m}\frac{x_iz}{1-x_i^2z^2}, \quad g(z) = \sum_{i=1}^{n}\frac{y_iz}{1-y_i^2z^2};$$

both $f(z)$ and $g(z)$ have total pole order at most $2n$. Meanwhile, by expanding in power series around $z = 0$, we see that $f(z) - g(z)$ is divisible by z^{2n+1}. Consequently, the two series are equal.

However, we can uniquely recover the multiset $\{x_1,\dots,x_m\}$ from $f(z)$: f has poles at $\{1/x_1^2,\dots,1/x_m^2\}$ and the residue of the pole at $z = 1/x_i^2$ uniquely determines both x_i (i.e., its sign) and its multiplicity. Similarly, we may recover $\{y_1,\dots,y_n\}$ from $g(z)$, so the two multisets must coincide. □

Now suppose by way of contradiction that we have an example showing that $m \geq n+1$. We then have

$$1^{2k-1} + \sum_{i=1}^{n}x_{2i-1}^{2k-1} = \sum_{i=1}^{n}x_{2i}^{2k-1} \qquad (k = 1,\dots,n+1).$$

By the lemma, this means that the multisets $\{1,x_1,x_3,\dots,x_{2n-1}\}$ and $\{x_2,x_4,\dots,x_{2n}\}$ become equal after removing pairs of inverses until this becomes impossible. However, of the resulting two multisets, the first contains 1 and the second does not, yielding the desired contradiction.

Remark. One can also prove the lemma using the invertibility of the Vandermonde matrix

$$(x_i^j)_{i=0,\dots,n;j=0,\dots,n}$$

for $x_0,\dots,x_n$ pairwise distinct (this matrix has determinant $\prod_{0\leq i<j\leq n}(x_i - x_j) \neq 0$). For a similar argument, see Proposition 22 of: M. Bhargava, Galois groups of random integer polynomials and van der Waerden's conjecture, arXiv:2111.06507.

Remark. The solution for $m = n$ given above is not unique (see below). However, it does become unique if we add the assumption that $x_i = -x_{2n+1-i}$ for $i = 1,\dots,2n$ (i.e., the set of intervals is symmetric around 0).

Second solution. (by Evan Dummit) Define the polynomial

$$p(x) = (x-x_1)(x+x_2)\cdots(x-x_{2n-1})(x+x_{2n})(x+1);$$

by hypothesis, $p(x)$ has $2n+1$ distinct real roots in the interval $[-1,1)$. Let s_k denote the k-th power sum of $p(x)$; then for any given m, the desired condition is that $s_{2k-1} = 0$ for $k = 1,\dots,m$. Let e_k denote the k-th elementary symmetric function of the roots of $p(x)$; that is,

$$p(x) = x^{2n+1} + \sum_{i=k}^{2n+1}(-1)^k e_k x^{2n+1-k}.$$

By the Girard–Newton identities,

$$(2k-1)e_{2k-1} = s_1e_{2k-2} - s_2e_{2k-2} + \cdots - s_{2k}e_1;$$

hence the desired condition implies that $e_{2k-1} = 0$ for $k = 1,\dots,m$.

If we had a solution with $m = n+1$, then the vanishing of $e_1,\dots,e_{2k+1}$ would imply that $p(x)$ is an odd polynomial (that is, $p(x) = -p(x)$ for all x), which in turn would imply that $x = 1$ is also a root of p. Since we have already identified $2n+1$ other roots of p, this yields a contradiction.

By the same token, a solution with $m = n$ corresponds to a polynomial $p(x)$ of the form $xq(x^2)+a$ for some polynomial $q(x)$ of degree n and some real number a (necessarily equal to $q(1)$). It will thus suffice to choose $q(x)$

so that the resulting polynomial $p(x)$ has roots consisting of -1 plus $2n$ distinct values in $(-1,1)$. To do this, start with any polynomial $r(x)$ of degree n with n distinct positive roots (e.g., $r(x) = (x-1)\cdots(x-n)$). The polynomial $xr(x^2)$ then has $2n+1$ distinct real roots; consequently, for $\varepsilon > 0$ sufficiently small, $xr(x^2)+\varepsilon$ also has $2n+1$ distinct real roots. Let $-\alpha$ be the smallest of these roots (so that $\alpha > 0$); we then take $q(x) = r(x\sqrt{\alpha})$ to achieve the desired result.

Remark. Brian Lawrence points out that one can also produce solutions for $m = n$ by starting with the degenerate solution

$$-a_{n-1}, \dots, -a_1, 0, a_1, \dots, a_{n-1}, 1$$

(where $0 < a_1 < \cdots < a_{n-1} < 1$ but no other conditions are imposed) and deforming it using the implicit function theorem. More precisely, there exists a differentiable parametric solution $x_1(t), \dots, x_{2n}(t)$ with $x_i(t) = x_{2n-i}(t)$ for $i = 1, \dots, n-1$ specializing to the previous solution at $t = 0$, such that $x_i'(0) \neq 0$ for $i = n, \dots, 2n$; this is because the Jacobian matrix

$$J = ((2k-1)x_i(0)^{2k-2})_{i=n,\dots,2n; k=1,\dots,n}$$

(interpreting 0^0 as 1) has the property that every maximal minor is nonzero (these being scaled Vandermonde matrices). In particular we may normalize so that $x_{2n}'(0) < 0$, and then evaluating at a small positive value of t gives the desired example.

In the proof that $m = n+1$ cannot occur, one can similarly use the implicit function theorem (with some care) to reduce to the case where $\{|x_1|, \dots, |x_{2n}|\}$ has cardinality $n+1$. This can be extended to a complete solution, but the details are rather involved.

B1 We prove that $b_k k!$ is an odd integer for all $k \geq 0$.

First solution. Since $e^{P(x)} = \sum_{n=0}^\infty \frac{(P(x))^n}{n!}$, the number $k!\,b_k$ is the coefficient of x^k in

$$(P(x))^k + \sum_{n=0}^{k-1} \frac{k!}{n!}(P(x))^n.$$

In particular, $b_0 = 1$ and $b_1 = a_1$ are both odd.

Now suppose $k \geq 2$; we want to show that b_k is odd. The coefficient of x^k in $(P(x))^k$ is a_1^k. It suffices to show that the coefficient of x^k in $\frac{k!}{n!}(P(x))^n$ is an even integer for any $n < k$. For k even or $n \leq k-2$, this follows immediately from the fact that $\frac{k!}{n!}$ is an even integer. For k odd and $n = k-1$, we have

$$\begin{aligned}\frac{k!}{(k-1)!}(P(x))^{k-1} &= k(a_1x + a_2x^2 + \cdots)^{k-1} \\ &= k(a_1^{k-1}x^{k-1} + (k-1)a_1^{k-2}a_2x^k + \cdots)\end{aligned}$$

and the coefficient of x^k is $k(k-1)a_1^{k-2}a_2$, which is again an even integer.

Second solution. Let G be the set of power series of the form $\sum_{n=0}^\infty c_n \frac{x_n}{n!}$ with $c_0 = 1, c_n \in \mathbb{Z}$; then G forms a group under formal series multiplication because

$$\left(\sum_{n=0}^\infty c_n \frac{x^n}{n!}\right)\left(\sum_{n=0}^\infty d_n \frac{x^n}{n!}\right) = \sum_{n=0}^\infty e_n \frac{x^n}{n!}$$

with

$$e_n = \sum_{m=0}^n \binom{n}{m} c_m d_{n-m}.$$

By the same calculation, the subset H of series with $c_n \in 2\mathbb{Z}$ for all $n \geq 1$ is a subgroup of G.

We have $e^{2x} \in H$ because $\frac{2^n}{n!} \in 2\mathbb{Z}$ for all $n \geq 1$: the exponent of 2 in the prime factorization of $n!$ is

$$\sum_{i=1}^\infty \left\lfloor \frac{n}{2^i} \right\rfloor < \sum_{i=1}^\infty \frac{n}{2^i} = n.$$

For any integer $k \geq 2$, we have $e^{x^k} \in H$ because $\frac{(nk)!}{n!} \in 2\mathbb{Z}$ for all $n \geq 1$: this is clear if $k = 2, n = 1$, and in all other cases the ratio is divisible by $(n+1)(n+2)$.

We deduce that $e^{P(x)-x} \in H$. By writing $e^{P(x)}$ as $e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$ times an element of H, we deduce that $k!b_k$ is odd for all $k \geq 0$.

Third solution. (by David Feldman) We interpret $e^{P(x)}$ using the exponential formula for generating functions. For each j, choose a set S_j consisting of $|a_j|$ colors. Then b_k is a weighted count over set partitions of $\{1, \dots, k\}$, with each part of size j assigned a color in S_j, and the weight being $(-1)^i$ where i is the number of parts of any size j for which $a_j < 0$.

Since we are only looking for the parity of b_k, we may dispense with the signs; that is, we may assume $a_j \geq 0$ for all j and forget about the weights.

Choose an involution on each S_j with at most one fixed point; this induces an involution on the partitions, so to find the parity of b_k we may instead count fixed points of the involution. That is, we may assume that $a_j \in \{0,1\}$.

Let T_k be the set of set partitions in question with the all-singletons partition removed; it now suffices to exhibit a fixed-point-free involution of T_k. To wit, for each partition in T_k, there is a smallest index $i \in \{1, \dots, k-1\}$ for which i and $i+1$ are not both singletons; we define an involution by swapping the positions of i and $i+1$.

B2 The possible values of n are 1 and 7.

Clearly the set $S = \{0\}$ works. Suppose that $S \neq \{0\}$ is a finite set satisfying the given condition; in particular, S does not consist of a collection of collinear vectors, since otherwise $\{v \times w : v, w \in S\} = \{0\}$. We claim that S cannot contain any nonzero vector v with $\|v\| \neq 1$. Suppose otherwise, and let $w \in S$ be a vector not collinear with v. Then S must contain the nonzero vector $u_1 = v \times w$, as well as the sequence of vectors u_n defined inductively by $u_n = v \times u_{n-1}$. Since each u_n is orthogonal to v by construction, we have $\|u_n\| = \|v\| \|u_{n-1}\|$ and so $\|u_n\| = \|v\|^{n-1} \|u_1\|$. The sequence $\|u_n\|$ consists of all distinct numbers and thus S is infinite, a contradiction. This proves the claim, and so every nonzero vector in S is a unit vector.

Next note that any pair of vectors $v, w \in S$ must either be collinear or orthogonal: by the claim, v, w are both unit vectors, and if v, w are not collinear then $v \times w \in S$ must be a unit vector, whence $v \perp w$. Now choose any pair of non-collinear vectors $v_1, v_2 \in S$, and write $v_3 = v_1 \times v_2$. Then $\{v_1, v_2, v_3\}$ is an orthonormal basis of $\mathbb{R}^3$, and it follows that all of these vectors are in S: 0, v_1, v_2, v_3, $-v_1 = v_3 \times v_2$, $-v_2 = v_1 \times v_3$, and $-v_3 = v_2 \times v_1$. On the other hand, S cannot contain any vector besides these seven, since any other vector w in S would have to be simultaneously orthogonal to all of v_1, v_2, v_3.

Thus any set $S \neq \{0\}$ satisfying the given condition must be of the form $\{0, \pm v_1, \pm v_2, \pm v_3\}$ where $\{v_1, v_2, v_3\}$ is an orthonormal basis of $\mathbb{R}^3$. It is clear that any set of this form does satisfy the given condition. We conclude that the answer is $n = 1$ or $n = 7$.

B3 The answer is yes. Let $R_0, B_0 \subset \mathbb{R}^+$ be the set of red and blue numbers at the start of the process, and let R_n, B_n be the set of red and blue numbers after n steps. We claim that $R_2 = \mathbb{R}^+$.

We first note that if $y \in B_1$, then $y/2 \in R_1$. Namely, the numbers y and $2y$ must be of opposite colors in the original coloring, and then $3y/2$ must be of the same color as one of y or $2y$.

Now suppose by way of contradiction that $x \in B_2$. Then of the four numbers $x, 2x, 3x, 4x$, every other number must be in R_1 and the other two must be in B_1. By the previous observation, $2x$ and $4x$ cannot both be in B_1; it follows that $2x, 4x \in R_1$ and $x, 3x \in B_1$. By the previous observation again, $x/2$ and $3x/2$ must both be in R_1, but then $x = 3x/2 - x/2$ is in R_2, contradiction. We conclude that $R_2 = \mathbb{R}^+$, as desired.

B4 The values of n in question are the multiples of 3 starting with 9. Note that we interpret "distinct" in the problem statement to mean "pairwise distinct" (i.e., no two equal). See the remark below.

We first show that such a sequence can only occur when n is divisible by 3. If d_1 and d_2 are the common differences of the arithmetic progressions $\{x_m, x_{m+1}, x_{m+2}\}$ and $\{x_{m+1}, x_{m+2}, x_{m+3}\}$ for some m, then $d_2 \in \{d_1, 2d_1, d_1/2\}$. By scaling we may assume that the smallest common difference that occurs is 1; in this case, all of the common differences are integers. By shifting, we may assume that the x_i are themselves all integers. We now observe that any three consecutive terms in the sequence have pairwise distinct residues modulo 3, forcing n to be divisible by 3.

We then observe that for any $m \geq 2$, we obtain a sequence of the desired form of length $3m + 3 = (2m - 1) + 1 + (m + 1) + 2$ by concatenating the arithmetic progressions

$$\begin{gathered}(1, 3, \dots, 4m-3, 4m-1),\\ 4m-2, (4m, 4m-4, \dots, 4, 0), 2.\end{gathered}$$

We see that no terms are repeated by noting that the first parenthesized sequence consists of odd numbers; the second sequence consists of multiples of 4; and the remaining numbers 2 and $4m - 2$ are distinct (because $m \geq 2$) but both congruent to 2 mod 4.

It remains to show that no such sequence occurs with $n = 6$. We may assume without loss of generality that the smallest common difference among the arithmetic progressions is 1 and occurs for $\{x_1, x_2, x_3\}$; by rescaling, shifting, and reversing the sequence as needed, we may assume that $x_1 = 0$ and $(x_2, x_3) \in \{(1, 2), (2, 1)\}$. We then have $x_4 = 3$ and

$$(x_5, x_6) \in \{(4, 5), (-1, -5), (-1, 7), (5, 4), (5, 7)\}.$$

In none of these cases does $\{x_5, x_6, 0\}$ form an arithmetic progression.

Remark. If one interprets "distinct" in the problem statement to mean "not all equal", then the problem becomes simpler: the same argument as above shows that n must be a multiple of 3, in which case a suitable repetition of the sequence $-1, 0, 1$ works.

B5 **First solution.** The answer is $p \leq 1/4$. We first show that $p > 1/4$ does not satisfy the desired condition. For $p > 1/3$, $P(0, 1) = 1 - 2p < p = P(1, 1)$. For $p = 1/3$, it is easily calculated (or follows from the next calculation) that $P(0, 1, 2) = 1/9 < 2/9 = P(1, 1, 2)$. Now suppose $1/4 < p < 1/3$, and consider $(b, a_1, a_2, a_3, \dots, a_n) = (1, 1, 2, 4, \dots, 2^{n-1})$. The only solution to

$$X_1 + 2X_2 + \cdots + 2^{n-1} X_n = 0$$

with $X_j \in \{0, \pm 1\}$ is $X_1 = \cdots = X_n = 0$; thus $P(0, 1, 2, \dots, 2^{2n-1}) = (1 - 2p)^n$. On the other hand, the solutions to

$$X_1 + 2X_2 + \cdots + 2^{n-1} X_n = 1$$

with $X_j \in \{0, \pm 1\}$ are

$$\begin{gathered}(X_1, X_2, \dots, X_n) = (1, 0, \dots, 0), (-1, 1, 0, \dots, 0),\\ (-1, -1, 1, 0, \dots, 0), \dots, (-1, -1, \dots, -1, 1),\end{gathered}$$

and so

$$\begin{aligned}&P(1, 1, 2, \dots, 2^{n-1})\\ &= p(1 - 2p)^{n-1} + p^2(1 - 2p)^{n-2} + \cdots + p^n\\ &= p \frac{(1 - 2p)^n - p^n}{1 - 3p}.\end{aligned}$$

It follows that the inequality $P(0,1,2,\dots,2^{n-1}) \geq P(1,1,2,\dots,2^{n-1})$ is equivalent to

$$p^{n+1} \geq (4p-1)(1-2p)^n,$$

but this is false for sufficiently large n since $4p-1>0$ and $p<1-2p$.

Now suppose $p \leq 1/4$; we want to show that for arbitrary $a_1,\dots,a_n$ and $b \neq 0$, $P(0,a_1,\dots,a_n) \geq P(b,a_1,\dots,a_n)$. Define the polynomial

$$f(x) = px + px^{-1} + 1 - 2p,$$

and observe that $P(b,a_1,\dots,a_n)$ is the coefficient of x^b in $f(x^{a_1})f(x^{a_2})\cdots f(x^{a_n})$. We can write

$$f(x^{a_1})f(x^{a_2})\cdots f(x^{a_n}) = g(x)g(x^{-1})$$

for some real polynomial g: indeed, if we define $\alpha = \frac{1-2p+\sqrt{1-4p}}{2p} > 0$, then $f(x) = \frac{p}{\alpha}(x+\alpha)(x^{-1}+\alpha)$, and so we can use

$$g(x) = \left(\frac{p}{\alpha}\right)^{n/2}(x^{a_1}+\alpha)\cdots(x^{a_n}+\alpha).$$

It now suffices to show that in $g(x)g(x^{-1})$, the coefficient of x^0 is at least as large as the coefficient of x^b for any $b \neq 0$. Since $g(x)g(x^{-1})$ is symmetric upon inverting x, we may assume that $b>0$. If we write $g(x) = c_0x^0 + \cdots + c_mx^m$, then the coefficients of x^0 and x^b in $g(x)g(x^{-1})$ are $c_0^2+c_1^2+\cdots+c_m^2$ and $c_0c_b + c_1c_{b+1} + \cdots + c_{m-b}c_m$, respectively. But

$$\begin{aligned}
&2(c_0c_b + c_1c_{b+1} + \cdots + c_{m-b}c_m) \\
&\leq (c_0^2+c_b^2) + (c_1^2+c_{b+1}^2) + \cdots + (c_{m-b}^2+c_m^2) \\
&\leq 2(c_0^2+\cdots+c_m^2),
\end{aligned}$$

and the result follows.

Second solution. (by Yuval Peres) We check that $p \leq 1/4$ is necessary as in the first solution. To check that it is sufficient, we introduce the following concept: for X a random variable taking finitely many integer values, define the *characteristic function*

$$\varphi_X(\theta) = \sum_{\ell\in\mathbb{Z}} P(X=\ell)e^{i\ell\theta}$$

(i.e., the expected value of $e^{iX\theta}$, or the Fourier transform of the probability measure corresponding to X). We use two evident properties of these functions:

- If X and Y are independent, then $\varphi_{X+Y}(\theta) = \varphi_X(\theta) + \varphi_Y(\theta)$.
- For any $b \in \mathbb{Z}$,

  $$P(X=b) = \frac{1}{2}\int_0^{2\pi} e^{-ib\theta}\varphi_X(\theta)\,d\theta.$$

  In particular, if $\varphi_X(\theta) \geq 0$ for all θ, then $P(X=b) \leq P(X=0)$.

For $p \leq 1/4$, we have

$$\varphi_{X_k}(\theta) = (1-2p) + 2p\cos(\theta) \geq 0.$$

Hence for $a_1,\dots,a_n \in \mathbb{Z}$, the random variable $S = a_1X_1 + \cdots + a_nX_n$ satisfies

$$\varphi_S(\theta) = \prod_{k=1}^n \varphi_{a_kX_k}(\theta) = \prod_{k=1}^n \varphi_{X_k}(a_k\theta) \geq 0.$$

We may thus conclude that $P(S=b) \leq P(S=0)$ for any $b \in \mathbb{Z}$, as desired.

B6 The only such functions are the functions $f(x) = \frac{1}{1+cx}$ for some $c \geq 0$ (the case $c=0$ giving the constant function $f(x)=1$). Note that we interpret $\mathbb{R}^+$ in the problem statement to mean the set of positive real numbers, excluding 0.

For convenience, we reproduce here the given equation:

$$f(xf(y)) + f(yf(x)) = 1 + f(x+y) \qquad (1)$$

We first prove that

$$\lim_{x\to 0^+} f(x) = 1. \qquad (2)$$

Set

$$L_- = \liminf_{x\to 0^+} f(x), \quad L_+ = \limsup_{x\to 0^+} f(x).$$

For any fixed y, we have by (1)

$$\begin{aligned}
L_+ &= \limsup_{x\to 0^+} f(xf(y)) \\
&\leq \limsup_{x\to 0^+} (1+f(x+y)) = 1 + f(y) < \infty.
\end{aligned}$$

Consequently, $xf(x) \to 0$ as $x \to 0^+$. By (2) with $y=x$,

$$\begin{aligned}
2L_+ &= \limsup_{x\to 0^+} 2f(xf(x)) \\
&= \limsup_{x\to 0^+} (1+f(2x)) = 1 + L_+ \\
2L_- &= \liminf_{x\to 0^+} 2f(xf(x)) \\
&= \liminf_{x\to 0^+} (1+f(2x)) = 1 + L_-
\end{aligned}$$

and so $L_- = L_+ = 1$, confirming (2).

We next confirm that

$$f(x) \geq 1 \text{ for all } x>0 \Longrightarrow f(x) = 1 \text{ for all } x>0. \qquad (3)$$

Suppose that $f(x) \geq 1$ for all $x>0$. For $0 < c \leq \infty$, put $S_c = \sup\{f(x) : 0 < x \leq c\}$; for $c < \infty$, (2) implies that $S_c < \infty$. If there exists $y>0$ with $f(y)>1$, then from (1) we have $f(x+y) - f(xf(y)) = f(yf(x)) - 1 \geq 0$; hence

$$S_c = S_{(c-y)f(y)} \qquad \left(c \geq c_0 = \frac{yf(y)}{f(y)-1}\right)$$

and (since $(c-y)f(y)-c_0 = f(y)(c-c_0)$) iterating this construction shows that $S_\infty = S_c$ for any $c > c_0$. In any case, we deduce that

$$f(x) \geq 1 \text{ for all } x > 0 \Longrightarrow S_\infty < \infty. \tag{4}$$

Still assuming that $f(x) \geq 1$ for all $x > 0$, note that from (1) with $x = y$,

$$f(xf(x)) = \frac{1}{2}(1+f(2x)).$$

Since $xf(x) \to 0$ as $x \to 0^+$ by (2) and $xf(x) \to \infty$ as $x \to \infty$, $xf(x)$ takes all positive real values by the intermediate value theorem. We deduce that $2S_\infty \leq 1 + S_\infty$ and hence $S_\infty = 1$; this proves (3).

We may thus assume hereafter that $f(x) < 1$ for some $x > 0$. We next check that

$$\lim_{x\to\infty} f(x) = 0. \tag{5}$$

Put $I = \inf\{f(x) : x > 0\} < 1$, choose $\varepsilon \in (0, (1-I)/2)$, and choose $y > 0$ such that $f(y) < I + \varepsilon$. We then must have $xf(x) \neq y$ for all x, or else

$$1 + I \leq 1 + f(2x) = 2f(y) < 2I + 2\varepsilon,$$

contradiction. Since $xf(x) \to 0$ as $x \to 0^+$ by (2), we have $\sup\{xf(x) : x > 0\} < \infty$ by the intermediate value theorem, yielding (5).

By (2) plus (5), $f^{-1}(1/2)$ is nonempty and compact. We can now simplify by noting that if $f(x)$ satisfies the original equation, then so does $f(cx)$ for any $c > 0$; we may thus assume that the least element of $f^{-1}(1/2)$ is 1, in which case we must show that $f(x) = \frac{1}{1+x}$.

We next show that

$$\lim_{x\to\infty} xf(x) = 1. \tag{6}$$

For all $x > 0$, by (1) with $y = x$,

$$f(xf(x)) = \frac{1}{2}(1+f(2x)) > \frac{1}{2} = f(1), \tag{7}$$

so in particular $xf(x) \neq 1$. As in the proof of (5), this implies that $xf(x) < 1$ for all $x > 0$. However, by (5) and (7) we have $f(xf(x)) \to \frac{1}{2}$ as $x \to \infty$, yielding (6).

By substituting $y \mapsto xy$ in (1),

$$f(xf(xy)) + f(xyf(x)) = 1 + f(x+xy).$$

Taking the limit as $x \to \infty$ and applying (6) yields

$$f(1/y) + f(y) = 1. \tag{8}$$

Combining (1) with (8) yields

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

Multiply both sides by $xf(y)$, then take the limit as $x \to \infty$ to obtain

$$\begin{aligned} 1 &= \lim_{x\to\infty} xf(y)f(x+y) + \lim_{x\to\infty} xf(y)f\left(\frac{1}{yf(x)}\right) \\ &= f(y) + \lim_{x\to\infty} xf(y)yf(x) \\ &= f(y) + yf(y) \end{aligned}$$

and solving for $f(y)$ now yields $f(y) = \frac{1}{1+y}$, as desired.

Remark. Some variants of the above approach are possible. For example, once we have (5), we can establish that f is monotone decreasing as follows. We first check that

$$f(x) < 1 \text{ for all } x > 0. \tag{9}$$

Suppose by way of contradiction that $f(x) = 1$ for some x. By (1),

$$f(2x) + 1 = 2f(xf(x)) = 2f(x) = 2$$

and so $f(2x) = 1$. It follows that $f^{-1}(1)$ is infinite, contradicting (5).

We next check that

$$x < y \Longrightarrow f(x) > f(y). \tag{10}$$

For $x < y$, by substituting $x \mapsto y - x$ in (1) we obtain

$$\begin{aligned} 1 + f(y) &= f(xf(y-x)) + f((y-x)f(x)) \\ &< 1 + f((y-x)f(x)), \end{aligned}$$

whence $f((y-x)f(x)) > f(y)$. Because $(y-x)f(x) \to 0$ as $x \to y^-$ and $(y-x)f(x) \to y$ as $x \to 0^+$, $(y-x)f(x)$ takes all values in $(0,y)$ as x varies over $(0,y)$; this proves (10).