Solutions to the 63rd William Lowell Putnam Mathematical Competition
Saturday, December 7, 2002

Kiran Kedlaya and Lenny Ng

A–1 By differentiating $P_n(x)/(x^k-1)^{n+1}$, we find that $P_{n+1}(x) = (x^k-1)P_n'(x) - (n+1)kx^{k-1}P_n(x)$; substituting $x=1$ yields $P_{n+1}(1) = -(n+1)kP_n(1)$. Since $P_0(1) = 1$, an easy induction gives $P_n(1) = (-k)^n n!$ for all $n \geq 0$.

Note: one can also argue by expanding in Taylor series around 1. Namely, we have

$$\frac{1}{x^k-1} = \frac{1}{k(x-1)+\cdots} = \frac{1}{k}(x-1)^{-1}+\cdots,$$

so

$$\frac{d^n}{dx^n}\frac{1}{x^k-1} = \frac{(-1)^n n!}{k(x-1)^{-n-1}}$$

and

$$\begin{aligned} P_n(x) &= (x^k-1)^{n+1}\frac{d^n}{dx^n}\frac{1}{x^k-1} \\ &= (k(x-1)+\cdots)^{n+1}\left(\frac{(-1)^n n!}{k}(x-1)^{-n-1}+\cdots\right) \\ &= (-k)^n n! + \cdots. \end{aligned}$$

A–2 Draw a great circle through two of the points. There are two closed hemispheres with this great circle as boundary, and each of the other three points lies in one of them. By the pigeonhole principle, two of those three points lie in the same hemisphere, and that hemisphere thus contains four of the five given points.

Note: by a similar argument, one can prove that among any $n+3$ points on an n-dimensional sphere, some $n+2$ of them lie on a closed hemisphere. (One cannot get by with only $n+2$ points: put them at the vertices of a regular simplex.) Namely, any n of the points lie on a great sphere, which forms the boundary of two hemispheres; of the remaining three points, some two lie in the same hemisphere.

A–3 Note that each of the sets $\{1\},\{2\},\ldots,\{n\}$ has the desired property. Moreover, for each set S with integer average m that does not contain m, $S\cup\{m\}$ also has average m, while for each set T of more than one element with integer average m that contains m, $T\setminus\{m\}$ also has average m. Thus the subsets other than $\{1\},\{2\},\ldots,\{n\}$ can be grouped in pairs, so $T_n - n$ is even.

A–4 (partly due to David Savitt) Player 0 wins with optimal play. In fact, we prove that Player 1 cannot prevent Player 0 from creating a row of all zeroes, a column of all zeroes, or a 2×2 submatrix of all zeroes. Each of these forces the determinant of the matrix to be zero.

For $i,j = 1,2,3$, let A_{ij} denote the position in row i and column j. Without loss of generality, we may assume that Player 1's first move is at A_{11}. Player 0 then plays at A_{22}:

$$\begin{pmatrix} 1 & * & * \\ * & 0 & * \\ * & * & * \end{pmatrix}$$

After Player 1's second move, at least one of A_{23} and A_{32} remains vacant. Without loss of generality, assume A_{23} remains vacant; Player 0 then plays there.

After Player 1's third move, Player 0 wins by playing at A_{21} if that position is unoccupied. So assume instead that Player 1 has played there. Thus of Player 1's three moves so far, two are at A_{11} and A_{21}. Hence for i equal to one of 1 or 3, and for j equal to one of 2 or 3, the following are both true:

(a) The 2×2 submatrix formed by rows 2 and i and by columns 2 and 3 contains two zeroes and two empty positions.

(b) Column j contains one zero and two empty positions.

Player 0 next plays at A_{ij}. To prevent a zero column, Player 1 must play in column j, upon which Player 0 completes the 2×2 submatrix in (a) for the win.

Note: one can also solve this problem directly by making a tree of possible play sequences. This tree can be considerably collapsed using symmetries: the symmetry between rows and columns, the invariance of the outcome under reordering of rows or columns, and the fact that the scenario after a sequence of moves does not depend on the order of the moves (sometimes called "transposition invariance").

Note (due to Paul Cheng): one can reduce Determinant Tic-Tac-Toe to a variant of ordinary tic-tac-toe. Namely, consider a tic-tac-toe grid labeled as follows:

A_{11}	A_{22}	A_{33}
A_{23}	A_{31}	A_{12}
A_{32}	A_{13}	A_{21}

Then each term in the expansion of the determinant occurs in a row or column of the grid. Suppose Player 1 first plays in the top left. Player 0 wins by playing first in the top row, and second in the left column. Then there are only one row and column left for Player 1 to threaten, and Player 1 cannot already threaten both on the third move, so Player 0 has time to block both.

A–5 It suffices to prove that for any relatively prime positive integers r,s, there exists an integer n with $a_n = r$ and $a_{n+1} = s$. We prove this by induction on $r+s$, the case $r+s=2$ following from the fact that $a_0 = a_1 = 1$. Given r and s not both 1 with $\gcd(r,s) = 1$, we must have $r \neq s$. If $r > s$, then by the induction hypothesis we have $a_n = r-s$ and $a_{n+1} = s$ for some n; then $a_{2n+2} = r$ and $a_{2n+3} = s$. If $r < s$, then we have $a_n = r$ and $a_{n+1} = s-r$ for some n; then $a_{2n+1} = r$ and $a_{2n+2} = s$.

Note: a related problem is as follows. Starting with the sequence

$$\frac{0}{1}, \frac{1}{0},$$

repeat the following operation: insert between each pair $\frac{a}{b}$ and $\frac{c}{d}$ the pair $\frac{a+c}{b+d}$. Prove that each positive rational number eventually appears.

Observe that by induction, if $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive terms in the sequence, then $bc-ad=1$. The same holds for consecutive terms of the n-th *Farey sequence*, the sequence of rational numbers in $[0,1]$ with denominator (in lowest terms) at most n.

A–6 The sum converges for $b=2$ and diverges for $b \geq 3$. We first consider $b \geq 3$. Suppose the sum converges; then the fact that $f(n) = nf(d)$ whenever $b^{d-1} \leq n \leq b^d - 1$ yields

$$\sum_{n=1}^{\infty} \frac{1}{f(n)} = \sum_{d=1}^{\infty} \frac{1}{f(d)} \sum_{n=b^{d-1}}^{b^d-1} \frac{1}{n}. \tag{1}$$

However, by comparing the integral of $1/x$ with a Riemann sum, we see that

$$\begin{aligned} \sum_{n=b^{d-1}}^{b^d-1} \frac{1}{n} &> \int_{b^{d-1}}^{b^d} \frac{dx}{x} \\ &= \log(b^d) - \log(b^{d-1}) = \log b, \end{aligned}$$

where log denotes the natural logarithm. Thus (1) yields

$$\sum_{n=1}^{\infty} \frac{1}{f(n)} > (\log b) \sum_{n=1}^{\infty} \frac{1}{f(n)},$$

a contradiction since $\log b > 1$ for $b \geq 3$. Therefore the sum diverges.

For $b=2$, we have a slightly different identity because $f(2) \neq 2f(2)$. Instead, for any positive integer i, we have

$$\sum_{n=1}^{2^i-1} \frac{1}{f(n)} = 1 + \frac{1}{2} + \frac{1}{6} + \sum_{d=3}^{i} \frac{1}{f(d)} \sum_{n=2^{d-1}}^{2^d-1} \frac{1}{n}. \tag{2}$$

Again comparing an integral to a Riemann sum, we see that for $d \geq 3$,

$$\begin{aligned} \sum_{n=2^{d-1}}^{2^d-1} \frac{1}{n} &< \frac{1}{2^{d-1}} - \frac{1}{2^d} + \int_{2^{d-1}}^{2^d} \frac{dx}{x} \\ &= \frac{1}{2^d} + \log 2 \\ &\leq \frac{1}{8} + \log 2 < 0.125 + 0.7 < 1. \end{aligned}$$

Put $c = \frac{1}{8} + \log 2$ and $L = 1 + \frac{1}{2} + \frac{1}{6(1-c)}$. Then we can prove that $\sum_{n=1}^{2^i-1} \frac{1}{f(n)} < L$ for all $i \geq 2$ by induction on i. The case $i=2$ is clear. For the induction, note that by (2),

$$\begin{aligned} \sum_{n=1}^{2^i-1} \frac{1}{f(n)} &< 1 + \frac{1}{2} + \frac{1}{6} + c \sum_{d=3}^{i} \frac{1}{f(d)} \\ &< 1 + \frac{1}{2} + \frac{1}{6} + c\frac{1}{6(1-c)} \\ &= 1 + \frac{1}{2} + \frac{1}{6(1-c)} = L, \end{aligned}$$

as desired. We conclude that $\sum_{n=1}^{\infty} \frac{1}{f(n)}$ converges to a limit less than or equal to L.

Note: the above argument proves that the sum for $b=2$ is at most $L < 2.417$. One can also obtain a lower bound by the same technique, namely $1 + \frac{1}{2} + \frac{1}{6(1-c')}$ with $c' = \log 2$. This bound exceeds 2.043. (By contrast, summing the first 100000 terms of the series only yields a lower bound of 1.906.) Repeating the same arguments with $d \geq 4$ as the cutoff yields the upper bound 2.185 and the lower bound 2.079.

B–1 The probability is $1/99$. In fact, we show by induction on n that after n shots, the probability of having made any number of shots from 1 to $n-1$ is equal to $1/(n-1)$. This is evident for $n=2$. Given the result for n, we see that the probability of making i shots after $n+1$ attempts is

$$\begin{aligned} \frac{i-1}{n}\frac{1}{n-1} + \left(1 - \frac{i}{n}\right)\frac{1}{n-1} &= \frac{(i-1)+(n-i)}{n(n-1)} \\ &= \frac{1}{n}, \end{aligned}$$

as claimed.

B–2 (Note: the problem statement assumes that all polyhedra are connected and that no two edges share more than one face, so we will do likewise. In particular, these are true for all convex polyhedra.) We show that in fact the first player can win on the third move. Suppose the polyhedron has a face A with at least four edges. If the first player plays there first, after the second player's first move there will be three consecutive faces B, C, D adjacent to A which are all unoccupied. The first player wins by playing in C; after the second player's second

move, at least one of B and D remains unoccupied, and either is a winning move for the first player.

It remains to show that the polyhedron has a face with at least four edges. (Thanks to Russ Mann for suggesting the following argument.) Suppose on the contrary that each face has only three edges. Starting with any face F_1 with vertices v_1, v_2, v_3, let v_4 be the other endpoint of the third edge out of v_1. Then the faces adjacent to F_1 must have vertices v_1, v_2, v_4; v_1, v_3, v_4; and v_2, v_3, v_4. Thus v_1, v_2, v_3, v_4 form a polyhedron by themselves, contradicting the fact that the given polyhedron is connected and has at least five vertices. (One can also deduce this using Euler's formula $V - E + F = 2 - 2g$, where V, E, F are the numbers of vertices, edges and faces, respectively, and g is the genus of the polyhedron. For a convex polyhedron, $g = 0$ and you get the "usual" Euler's formula.)

Note: Walter Stromquist points out the following counterexample if one relaxes the assumption that a pair of faces may not share multiple edges. Take a tetrahedron and remove a smaller tetrahedron from the center of an edge; this creates two small triangular faces and turns two of the original faces into hexagons. Then the second player can draw by signing one of the hexagons, one of the large triangles, and one of the small triangles. (He does this by "mirroring": wherever the first player signs, the second player signs the other face of the same type.)

B–3 The desired inequalities can be rewritten as

$$1 - \frac{1}{n} < \exp\left(1 + n \log\left(1 - \frac{1}{n}\right)\right) < 1 - \frac{1}{2n}.$$

By taking logarithms, we can rewrite the desired inequalities as

$$\begin{aligned} -\log\left(1 - \frac{1}{2n}\right) &< -1 - n\log\left(1 - \frac{1}{n}\right) \\ &< -\log\left(1 - \frac{1}{n}\right). \end{aligned}$$

Rewriting these in terms of the Taylor expansion of $-\log(1-x)$, we see that the desired result is also equivalent to

$$\sum_{i=1}^{\infty} \frac{1}{i2^i n^i} < \sum_{i=1}^{\infty} \frac{1}{(i+1)n^i} < \sum_{i=1}^{\infty} \frac{1}{in^i},$$

which is evident because the inequalities hold term by term.

Note: David Savitt points out that the upper bound can be improved from $1/(ne)$ to $2/(3ne)$ with a slightly more complicated argument. (In fact, for any $c > 1/2$, one has an upper bound of $c/(ne)$, but only for n above a certain bound depending on c.)

B–4 Use the following strategy: guess $1, 3, 4, 6, 7, 9, \ldots$ until the target number n is revealed to be equal to or lower than one of these guesses. If $n \equiv 1 \pmod 3$, it will be guessed on an odd turn. If $n \equiv 0 \pmod 3$, it will be guessed on an even turn. If $n \equiv 2 \pmod 3$, then $n+1$ will be guessed on an even turn, forcing a guess of n on the next turn. Thus the probability of success with this strategy is $1335/2002 > 2/3$.

Note: for any positive integer m, this strategy wins when the number is being guessed from $[1, m]$ with probability $\frac{1}{m}\lfloor \frac{2m+1}{3} \rfloor$. We can prove that this is best possible as follows. Let a_m denote m times the probability of winning when playing optimally. Also, let b_m denote m times the corresponding probability of winning if the objective is to select the number in an even number of guesses instead. (For definiteness, extend the definitions to incorporate $a_0 = 0$ and $b_0 = 0$.)

We first claim that $a_m = 1 + \max_{1 \le k \le m}\{b_{k-1} + b_{m-k}\}$ and $b_m = \max_{1 \le k \le m}\{a_{k-1} + a_{m-k}\}$ for $m \ge 1$. To establish the first recursive identity, suppose that our first guess is some integer k. We automatically win if $n = k$, with probability $1/m$. If $n < k$, with probability $(k-1)/m$, then we wish to guess an integer in $[1, k-1]$ in an even number of guesses; the probability of success when playing optimally is $b_{k-1}/(k-1)$, by assumption. Similarly, if $n < k$, with probability $(m-k)/m$, then the subsequent probability of winning is $b_{m-k}/(m-k)$. In sum, the overall probability of winning if k is our first guess is $(1 + b_{k-1} + b_{m-k})/m$. For optimal strategy, we choose k such that this quantity is maximized. (Note that this argument still holds if $k = 1$ or $k = m$, by our definitions of a_0 and b_0.) The first recursion follows, and the second recursion is established similarly.

We now prove by induction that $a_m = \lfloor (2m+1)/3 \rfloor$ and $b_m = \lfloor 2m/3 \rfloor$ for $m \ge 0$. The inductive step relies on the inequality $\lfloor x \rfloor + \lfloor y \rfloor \le \lfloor x + y \rfloor$, with equality when one of x, y is an integer. Now suppose that $a_i = \lfloor (2i+1)/3 \rfloor$ and $b_i = \lfloor 2i/3 \rfloor$ for $i < m$. Then

$$\begin{aligned} 1 + b_{k-1} + b_{m-k} &= 1 + \left\lfloor \frac{2(k-1)}{3} \right\rfloor + \left\lfloor \frac{2(m-k)}{3} \right\rfloor \\ &\le \left\lfloor \frac{2m}{3} \right\rfloor \end{aligned}$$

and similarly $a_{k-1} + a_{m-k} \le \lfloor (2m+1)/3 \rfloor$, with equality in both cases attained, e.g., when $k = 1$. The inductive formula for a_m and b_m follows.

B–5 (due to Dan Bernstein) Put $N = 2002!$. Then for $d = 1, \ldots, 2002$, the number N^2 written in base $b = N/d - 1$ has digits $d^2, 2d^2, d^2$. (Note that these really are digits because $2(2002)^2 < (2002!)^2/2002 - 1$.)

Note: one can also produce an integer N which has base b digits $1, *, 1$ for n different values of b, as follows. Choose c with $0 < c < 2^{1/n}$. For m a large positive integer, put $N = 1 + (m+1)\cdots(m+n)\lfloor cm \rfloor^{n-2}$. For m sufficiently large, the bases

$$b = \frac{N-1}{(m+i)m^{n-2}} = \prod_{j \ne i}(m+j)$$

for $i = 1, \ldots, n$ will have the properties that $N \equiv 1 \pmod{b}$ and $b^2 < N < 2b^2$ for m sufficiently large.

Note (due to Russ Mann): one can also give a "nonconstructive" argument. Let N be a large positive integer. For $b \in (N^2, N^3)$, the number of 3-digit base-b palindromes in the range $[b^2, N^6 - 1]$ is at least

$$\left\lfloor \frac{N^6 - b^2}{b} \right\rfloor - 1 \geq \frac{N^6}{b^2} - b - 2,$$

since there is a palindrome in each interval $[kb, (k+1)b - 1]$ for $k = b, \ldots, b^2 - 1$. Thus the average number of bases for which a number in $[1, N^6 - 1]$ is at least

$$\frac{1}{N^6} \sum_{b=N^2+1}^{N^3-1} \left(\frac{N^6}{b} - b - 2 \right) \geq \log(N) - c$$

for some constant $c > 0$. Take N so that the right side exceeds 2002; then at least one number in $[1, N^6 - 1]$ is a base-b palindrome for at least 2002 values of b.

B–6 We prove that the determinant is congruent modulo p to

$$x \prod_{i=0}^{p-1} (y + ix) \prod_{i,j=0}^{p-1} (z + ix + jy). \tag{3}$$

We first check that

$$\prod_{i=0}^{p-1} (y + ix) \equiv y^p - x^{p-1} y \pmod{p}. \tag{4}$$

Since both sides are homogeneous as polynomials in x and y, it suffices to check (4) for $x = 1$, as a congruence between polynomials. Now note that the right side has $0, 1, \ldots, p-1$ as roots modulo p, as does the left side. Moreover, both sides have the same leading coefficient. Since they both have degree only p, they must then coincide.

We thus have

$$\begin{aligned}
&x \prod_{i=0}^{p-1} (y + ix) \prod_{i,j=0}^{p-1} (z + ix + jy) \\
&\quad \equiv x(y^p - x^{p-1} y) \prod_{j=0}^{p-1} ((z + jy)^p - x^{p-1}(z + jy)) \\
&\quad \equiv (xy^p - x^p y) \prod_{j=0}^{p-1} (z^p - x^{p-1} z + jy^p - jx^{p-1} y) \\
&\quad \equiv (xy^p - x^p y)((z^p - x^{p-1} z)^p \\
&\qquad - (y^p - x^{p-1} y)^{p-1} (z^p - x^{p-1} z)) \\
&\quad \equiv (xy^p - x^p y)(z^{p^2} - x^{p^2-p} z^p) \\
&\qquad - x(y^p - x^{p-1} y)^p (z^p - x^{p-1} z) \\
&\quad \equiv xy^p z^{p^2} - x^p y z^{p^2} - x^{p^2-p+1} y^p z^p + x^{p^2} y z^p \\
&\qquad - xy^{p^2} z^p + x^{p^2-p+1} y^p z^p + x^p y^{p^2} z - x^{p^2} y^p z \\
&\quad \equiv xy^p z^{p^2} + yz^p x^{p^2} + zx^p y^{p^2} \\
&\qquad - xz^p y^{p^2} - yx^p z^{p^2} - zy^p x^{p^2},
\end{aligned}$$

which is precisely the desired determinant.

Note: a simpler conceptual proof is as follows. (Everything in this paragraph will be modulo p.) Note that for any integers a, b, c, the column vector $[ax + by + cz, (ax + by + cz)^p, (ax + by + cz)^{p^2}]$ is a linear combination of the columns of the given matrix. Thus $ax + by + cz$ divides the determinant. In particular, all of the factors of (3) divide the determinant; since both (3) and the determinant have degree $p^2 + p + 1$, they agree up to a scalar multiple. Moreover, they have the same coefficient of $z^{p^2} y^p x$ (since this term only appears in the expansion of (3) when you choose the first term in each factor). Thus the determinant is congruent to (3), as desired.

Either argument can be used to generalize to a corresponding $n \times n$ determinant, called a Moore determinant; we leave the precise formulation to the reader. Note the similarity with the classical Vandermonde determinant: if A is the $n \times n$ matrix with $A_{ij} = x_i^j$ for $i, j = 0, \ldots, n-1$, then

$$\det(A) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$