DAWN PUTNAM SESSION

Putnam 1973 — Papers A and B

Putnam 120 Archive

Contents

Putnam 1973 — Paper A **5**
1973 A1 5
1973 A2 7
1973 A3 10
1973 A4 12
1973 A5 14
1973 A6 17

Putnam 1973 — Paper B **21**
1973 B1 21
1973 B2 23
1973 B3 26
1973 B4 29
1973 B5 31
1973 B6 34

Putnam 1973 — Paper A

1973 A1

1973 A1. (a) Let ABC be any triangle. Let X, Y, Z be points on the sides BC, CA, AB respectively. Suppose the distances $BX \leq XC$, $CY \leq YA$, $AZ \leq ZB$. Show that the area of triangle XYZ is $\geq (1/4)$(area of triangle ABC). (b) Let ABC be any triangle, and let X, Y, Z be points on the sides BC, CA, AB respectively (but without any assumption about the ratios of the distances BX/XC, etc.). Using (a) or by any other method, show: one of the three corner triangles AZY, BXZ, CYX has an area $\leq$ area of triangle XYZ.

Fields: geometry, triangle areas, ratio substitution, monotonicity/convexity **Position:** A1 (opener)
Answer: (a) $\text{area}(XYZ)/\text{area}(ABC) \geq 1/4$, equality exactly at the medial triangle; (b) always true, with the surviving corner triangle located by a two-case argument.

Layer 1 THE WHAT — surface analysis

The natural attempt. The three inequalities $BX \leq XC, CY \leq YA, AZ \leq ZB$ read like a constraint region, and the instinct with a constrained-optimisation-flavoured inequality is to reach for calculus: fix B, C on the x-axis and A somewhere above, write $X = (x, 0)$, Y, Z as points determined by two more real parameters, expand $\text{area}(XYZ)$ by the shoelace formula in absolute coordinates, and try to minimise it subject to the three inequality constraints via Lagrange multipliers or by checking the boundary of the feasible region. This is not a foolish plan – it is exactly what "minimise an area subject to constraints" trained you to do – but it immediately drowns the problem in a triangle's worth of irrelevant data. The area of XYZ in absolute coordinates depends on the specific shape of ABC (its side lengths, its angles), and yet the answer the problem wants, $1/4$, does not depend on the shape of ABC at all. A solver who presses on with coordinates for, say, a 3-4-5 triangle can numerically confirm the bound and even locate the minimum at the midpoints by trial, but generalising that numerical fact to an arbitrary triangle by pure coordinate algebra means carrying A's two free coordinates through the whole computation, and the constraint set stops looking like three independent one-dimensional intervals and starts looking like a genuinely three-dimensional optimisation with a shape-dependent objective. The proof that results, if it is ever finished, will not obviously generalise to part (b), where the constraints disappear entirely and a different comparison (against the corner triangles, not against ABC) is wanted.

The model chain. The area of a triangle does not care about the individual lengths of two sides once you fix the angle between them and look at a *ratio*: for points Z on AB and Y on AC, $\text{area}(AZY) = \frac{1}{2} AZ \cdot AY \sin A$, and dividing by $\text{area}(ABC) = \frac{1}{2} AB \cdot AC \sin A$ cancels $\sin A$ and both absolute lengths, leaving $\text{area}(AZY)/\text{area}(ABC) = (AZ/AB)(AY/AC)$ – a quantity depending only on where Z and Y sit *proportionally* along their sides, not on the shape of ABC at all. This is the substitution that dissolves the whole problem: let

$$a = \frac{BX}{BC}, \qquad b = \frac{CY}{CA}, \qquad c = \frac{AZ}{AB}, \qquad a, b, c \in [0, 1].$$

The hypothesis of (a) is exactly $a, b, c \leq 1/2$. By the cancellation above,

$$\frac{\text{area}(BXZ)}{\text{area}(ABC)} = a(1-c), \quad \frac{\text{area}(CYX)}{\text{area}(ABC)} = b(1-a), \quad \frac{\text{area}(AZY)}{\text{area}(ABC)} = c(1-b),$$

(each corner triangle uses the two ratios measured *from its own vertex*), and since the three corner triangles plus XYZ tile ABC,

$$f(a,b,c) := \frac{\text{area}(XYZ)}{\text{area}(ABC)} = 1 - a(1-c) - b(1-a) - c(1-b) = 1 - a - b - c + ab + bc + ca.$$

This f is shape-independent – it is the whole problem, reduced to a three-variable algebra exercise. Now $\partial f/\partial a = -1 + b + c$. On the domain $b, c \in [0, 1/2]$ this is $\leq -1 + 1/2 + 1/2 = 0$, so f is non-increasing in a for every fixed b, c in range; the same holds for b and c by the cyclic symmetry of f. A function of three variables that is non-increasing in each, separately, on a box is minimised at the corner of the box where every variable is largest – here $a = b = c = 1/2$, the midpoints of every side. There, $f = 1 - 3/2 + 3/4 = 1/4$. Since f can only be larger elsewhere in the box, $\text{area}(XYZ)/\text{area}(ABC) = f(a,b,c) \geq f(1/2, 1/2, 1/2) = 1/4$ for all $a, b, c \in [0, 1/2]$, which is exactly (a).

For (b), the restriction $a, b, c \leq 1/2$ is dropped; a, b, c now range freely over $[0, 1]$, and f can fall all the way to 0 (e.g. $a = 1, b = 0, c = 0$ collapses XYZ onto a point). Two cases exhaust all of $[0, 1]^3$. **Case A: a, b, c all lie on the same side of $1/2$.** If all $\leq 1/2$, part (a) already gives $f \geq 1/4$. If all $\geq 1/2$, the same derivative computation gives $\partial f/\partial a = -1 + b + c \geq -1 + 1/2 + 1/2 = 0$, so f is now non-*decreasing* in each variable, minimised at the same corner $a = b = c = 1/2$, again giving $f \geq 1/4$. Either way $f \geq 1/4$, and since the three corner-area fractions sum to $1 - f$, if all three exceeded f their sum would exceed $3f \geq 3/4 \geq 1 - f$ – forcing $1 - f > 3f$, i.e. $f < 1/4$, a contradiction. So at least one corner triangle has area fraction $\leq f$. **Case B: mixed** – some cyclically adjacent pair, say $a \geq 1/2$ and $b \leq 1/2$ (the case the archive of solutions singles out, with c totally unconstrained), is handled directly: $f - b(1-a) = c(a + b - 1) + (1-a)(1-2b)$. If $a + b \geq 1$ this is minimised over $c \in [0, 1]$ at $c = 0$, giving $(1-a)(1-2b) \geq 0$ since $a \leq 1$, $b \leq 1/2$; if $a + b < 1$ it is minimised at $c = 1$, giving $b(2a - 1) \geq 0$ since $a \geq 1/2$. Either way $f \geq b(1-a) = \text{area}(CYX)/\text{area}(ABC)$, so corner CYX is the one that does not exceed XYZ – for *any* value of c whatsoever. The other mixed sub-cases are the same computation with the roles of (a, b, c) and the three corner triangles cycled.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both approaches start from the same three inequalities. They split at the decision of what variable to optimise in: the natural attempt keeps X, Y, Z as points with absolute coordinates tied to a specific triangle ABC; the model immediately replaces them with the three ratios $a = BX/BC$, $b = CY/CA$, $c = AZ/AB$ and never looks at a coordinate again.

> *The area of ABC is going to cancel out of every ratio I care about anyway, so I should never have written it down in absolute terms – the whole problem lives in the three numbers $a, b, c \in [0, 1]$, and I want the formula for* $\text{area}(XYZ)/\text{area}(ABC)$ *purely in terms of those three numbers.*

Layer 3 THE WHY

The reduction to ratios is not a trick of convenience; it is forced by what "area of a triangle with one shared angle" actually is. area $= \frac{1}{2}(\text{side})(\text{side})\sin(\text{included angle})$ is a bilinear form in the two side-lengths once the angle is fixed, so any ratio of two such areas sharing that angle is automatically a ratio of side-length products, with the angle and any common scale cancelling identically. The problem's own hypotheses – inequalities between lengths on the *same* side, like $BX \leq XC$ – are already scale-free statements about a single ratio; converting to $a = BX/BC \in [0, 1/2]$ just writes the hypothesis in the units it was born in. Once this is

seen, the appearance of $f(a,b,c) = 1 - a - b - c + ab + bc + ca$ is not a fortunate simplification but the *only* possible shape of the answer: it must be a polynomial that is affine (degree ≤ 1) in each of a, b, c separately, because each corner-area fraction is a product of exactly one ratio from each of two adjacent sides, and XYZ's fraction is 1 minus the sum of those three products. Affine-in-each-variable functions on a box are governed entirely by their behaviour at the $2^3 = 8$ corners – this is why a full partial-derivative argument was overkill in spirit even though it was the fastest way to see it: monotonicity in each variable, on an interval, is precisely what "affine in that variable" gives you for free, no calculus needed beyond noticing the sign of the coefficient of a (namely $b + c - 1$) does not depend on a itself.

Part (b)'s case split is likewise not an artificial partition but the only structure available: the real line splits into exactly two rays at $1/2$, and three numbers distributed between two classes must either all land in one class or split 2-1 – there is no third shape of case, which is why exactly two arguments (the monotonicity-plus-pigeonhole argument, and the direct algebraic identity) suffice to cover every triangle and every choice of X, Y, Z.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the phrase "distances $BX \leq XC$" attached to a claim about *areas*: on seeing a length inequality feeding an area conclusion, write down "convert every length constraint to a ratio along its own side" before touching coordinates.

2. That cue suggests parametrising the whole configuration by the three ratios $a, b, c \in [0,1]$ and expressing every area in the picture (all four sub-triangles) as a polynomial in a, b, c, rather than solving for positions.

3. Hesitation sets in right after writing $\text{area}(XYZ)/\text{area}(ABC) = 1 - a - b - c + ab + bc + ca$ and needing to minimise it: the temptation is to reach for Lagrange multipliers on a three-variable function. Reinforce against it by checking the function's degree in each variable separately – degree 1 in a means the minimum over an interval in a is always at an endpoint, so no calculus beyond a sign check is needed.

4. The trap in part (b) is treating it as requiring an entirely new argument because the constraint $a, b, c \leq 1/2$ has vanished; in fact the correct response is to case-split on whether the *same* constraint holds after relabelling around $1/2$, reusing (a)'s machinery almost unchanged in the "all-same-side" case and needing only one new (but structurally identical) computation for the mixed case.

5. Recognise this faster next time by treating any area ratio involving points that divide two sides sharing a vertex as an instant invitation to write it as (ratio)×(ratio), before any other manipulation.

6. Decision rule: whenever a geometry problem's hypotheses are stated as ratios or inequalities between lengths on the same segment, translate the entire configuration into those ratios as free parameters in $[0,1]$ immediately; any resulting area expression will then be affine in each parameter separately, and its extrema over a box are found by checking corners, not by calculus.

1973 A2

1973 A2. Consider an infinite series whose nth term is $\pm(1/n)$, the $\pm$ signs being determined according to a pattern that repeats periodically in blocks of eight. [There are 2^8 possible patterns, of which two examples are $++-----++$ and $+----+---$. The first example generates the series $1 + (1/2) - (1/3) - (1/4) - (1/5) - (1/6) + (1/7) + (1/8) + (1/9) + (1/10) - \cdots$.] (a) Show that a sufficient condition for the series to be conditionally convergent is that there be four "+" signs and four "−" signs in the block of eight. (b) Is this sufficient condition also necessary? [Here "convergent" means "convergent

to a finite limit."]

Fields: infinite series, conditional convergence, asymptotic expansion of partial sums **Position:** A2
Answer: (a) Yes, a 4-4 split of signs in each block is sufficient. (b) Yes, it is also necessary: any block with an unequal number of $+$ and $-$ signs makes the series diverge.

Layer 1 THE WHAT — surface analysis

The natural attempt. The series is $\sum \pm 1/n$ with a periodic sign pattern, and the reflex on seeing "conditionally convergent, sign pattern" is the alternating series test – but that test needs the signs to alternate one-for-one, and here they come in blocks of up to eight of the same sign in a row (e.g. $++---++$ has four minuses back to back), so it does not apply directly. The next reflex is to try to bound the partial sums crudely: group the series into blocks of eight and note $1/n$ is roughly constant across a block of eight consecutive n's when n is large, so a block with a 4-4 split "roughly cancels." This is true in spirit and is exactly the content of part (a), but turning "roughly cancels" into a proof invites the trap of stopping at first order: writing $1/(8j+i) \approx 1/(8j)$ for each of the eight terms in block j and concluding the block sum is approximately $(1/(8j))\sum_i s_i$, which is exactly zero when the signs are balanced. That looks like it finishes the sufficiency direction in one line – but "approximately zero" is not a proof of convergence, and worse, it gives no purchase at all on part (b): if the block sum is only shown to be *approximately* zero, there is no way to tell whether the actual (nonzero) discrepancy accumulates to a divergent series or not, for either the balanced or the unbalanced case. The first-order approximation has thrown away exactly the term that decides the problem.

The model chain. Since $u_n = \pm 1/n \to 0$, the full sequence of partial sums S_N converges if and only if the subsequence S_{8m} (partial sums at multiples of 8) converges: for $8m < N < 8(m+1)$, $|S_N - S_{8m}| \le \max_{8m<n\le N} |u_n| \to 0$, so any oscillation between consecutive block-boundaries vanishes in the limit. It is therefore enough to understand $D_j := S_{8(j+1)} - S_{8j} = \sum_{i=1}^{8} \frac{s_i}{8j+i}$, where $s_1, \dots, s_8 \in \{+1, -1\}$ is the fixed sign pattern, since $S_{8m} = S_8 + \sum_{j=1}^{m-1} D_j$ and convergence of S_{8m} is exactly convergence of $\sum_j D_j$.

The exact (not approximate) identity $\dfrac{1}{8j+i} = \dfrac{1}{8j} - \dfrac{i}{8j(8j+i)}$ splits D_j into a leading term carrying all the sign information and a genuinely small remainder:

$$D_j = \frac{1}{8j}\sum_{i=1}^{8} s_i \;-\; \sum_{i=1}^{8} \frac{s_i\, i}{8j(8j+i)}, \qquad \text{where } \left|\sum_{i=1}^{8} \frac{s_i\, i}{8j(8j+i)}\right| \le \frac{1+2+\cdots+8}{(8j)^2} = \frac{36}{64j^2}.$$

Let $C = \sum_{i=1}^{8} s_i$, an even integer in $[-8, 8]$ (even because the pattern has 8 signs). This identity is the whole problem: $D_j = \dfrac{C}{8j} + O(1/j^2)$, exactly, with an explicit, absolutely summable remainder.

Sufficiency (a). If the block has four $+$ and four $-$ signs, $C = 0$, so $D_j = O(1/j^2)$ with the bound above, hence $\sum_j |D_j|$ is bounded by a convergent multiple of the p-series $\sum 1/j^2$. So $\sum_j D_j$ converges absolutely, S_{8m} converges, and by the reduction above the full series converges. It is only *conditionally* convergent, not absolutely, since $\sum |u_n| = \sum 1/n$ diverges regardless of the sign pattern.

Necessity (b). If the block does not split 4-4, then $C \neq 0$ (an even nonzero integer). Write $D_j = C/(8j) + r_j$ where $|r_j| \le 36/(64j^2)$, so $\sum_j r_j$ converges absolutely. Then $\sum_j D_j$ converges if and only if $\sum_j C/(8j) = \frac{C}{8}\sum_j 1/j$ converges – but this is $C/8$ times the harmonic series, which diverges (to $+\infty$ or $-\infty$ according to the sign of C). A convergent series plus a divergent one is divergent, so $\sum_j D_j$ diverges, $S_{8m} \to \pm\infty$, and the original series diverges. Hence the 4-4 condition is not just sufficient but necessary: every unbalanced block of eight destroys convergence.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both paths agree that the series should be grouped into blocks of eight and that $1/n$ is "almost constant" across a block. They split at how far that approximation is pushed: the natural attempt stops at the leading-order approximation $1/(8j+i) \approx 1/(8j)$ and declares the balanced case "approximately zero," which cannot distinguish convergence from divergence or address necessity at all. The model instead writes the *exact* identity $1/(8j+i) = 1/(8j) - i/[8j(8j+i)]$, isolating the same leading term $C/(8j)$ but keeping the remainder as an explicit, provably small quantity rather than discarding it.

> *"Approximately zero" is not a mathematical object I can sum – if I want to know whether $\sum_j D_j$ converges, I need D_j written as an exact sum of a term I understand completely (its size and its dependence on the sign pattern) plus a remainder I can bound absolutely, not just estimate to leading order.*

Layer 3 THE WHY

The reason the crude first-order approximation fails is not a technical oversight; it is that convergence of $\sum_j D_j$ is a question about the *size* of D_j relative to $1/j$, and "approximately zero" carries no size information beyond "smaller than $1/j$" – which is compatible with D_j behaving like $1/j^{1.001}$ (divergent, since $\sum 1/j^{1.001}$... in fact that converges, but the point stands that "small" alone doesn't pin down convergence) or like $1/j^2$ (convergent) or like a divergent-but-sign-cancelling series. The problem is precisely built so that the leading-order term either vanishes exactly (when $C = 0$) or does not (when $C \neq 0$), and the entire dichotomy between convergence and divergence in this problem lives in that single integer $C = \sum s_i$ and the single fact that $\sum 1/j$ diverges while $\sum 1/j^2$ converges. This is why the exact expansion $1/(8j+i) = 1/(8j) - i/(8j(8j+i))$ – not an asymptotic "$\sim$" or "$O(\cdot)$" claim reached by intuition – is unavoidable: it is the unique algebraic identity that separates D_j into "the part governed by C" and "the part that is absolutely summable no matter what," and the problem's answer (necessary *and* sufficient) is only accessible once that separation is exact rather than approximate. In deeper terms, this is the general principle that a periodic perturbation of a harmonic-type series converges exactly when its per-period sum of *signs* (not magnitudes) vanishes – the harmonic series is the borderline case between $\sum 1/j^\alpha$ converging ($\alpha > 1$) and diverging ($\alpha \leq 1$), and $C \neq 0$ pushes the block sum down to exactly that $\alpha = 1$ borderline, while $C = 0$ drops it to $\alpha = 2$, safely convergent.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that the sign pattern has period 8 while the terms are $1/n$: on seeing "periodic sign pattern with slowly varying magnitude," write down "reduce to the subsequence of partial sums at block boundaries, since the terms within a block all go to zero together."
2. That cue suggests expanding the block sum $D_j = \sum_i s_i/(8j+i)$ around its leading behaviour in j, using the exact identity $1/(8j+i) = 1/(8j) - i/[8j(8j+i)]$ rather than an asymptotic symbol.
3. Hesitation sets in exactly at "the block sum is approximately zero when balanced" – it feels like the proof is done. Reinforce against it by asking: does this approximation, as stated, distinguish a convergent remainder from a divergent one? If the answer isn't an explicit bound on the remainder, the argument is not yet a proof.
4. The trap is treating $O(1/j^2)$-type reasoning as free once you've spotted the cancellation of leading terms; the actual work is producing an exact identity with a provably bounded remainder, since "leading terms cancel" alone cannot rule out the remainder itself summing to something divergent.
5. Recognise this faster next time by treating any series $\sum \pm f(n)$ with a periodic sign pattern and slowly varying f as an invitation to compute the exact per-period sum via a first-order-

with-error expansion of f, then classify convergence by whether the leading coefficient (here $C = \sum s_i$) vanishes.

6. Decision rule: for a periodic-sign series with terms of order $1/n^\alpha$, the block sum over a period of length L satisfies $D_j = \frac{C}{Lj^\alpha} + O(1/j^{\alpha+1})$ where C is the sum of signs in one period; the series converges (conditionally, if $\alpha = 1$) exactly when $C = 0$, and diverges otherwise – so both sufficiency and necessity fall out of the same single computation, never two separate arguments.

1973 A3

1973 A3. Let n be a fixed positive integer and let $b(n)$ be the minimum value of $k + \dfrac{n}{k}$ as k is allowed to range through all positive integers. Prove that $b(n)$ and $\sqrt{4n+1}$ have the same integer part. [The "integer part" of a real number is the greatest integer which does not exceed it.]

Fields: number theory, floor function, AM-GM, discrete optimisation **Position:** A3 **Answer:** $\lfloor b(n) \rfloor = \lfloor \sqrt{4n+1} \rfloor$ for every positive integer n; both equal $2m$ for $m^2 \le n < m^2 + m$ and $2m+1$ for $m^2 + m \le n < (m+1)^2$.

Layer 1 THE WHAT — surface analysis

The natural attempt. $b(n) = \min_{k \in \mathbb{Z}^+}(k + n/k)$ looks like a calculus problem wearing a number theory costume: the real function $g(x) = x + n/x$ has $g'(x) = 1 - n/x^2 = 0$ at $x = \sqrt{n}$, with $g(\sqrt{n}) = 2\sqrt{n}$, so the strong first instinct is that $b(n)$ should be "close to" $2\sqrt{n}$ and the problem is really about controlling the gap between the continuous minimiser $\sqrt{n}$ and the nearest integers. This instinct is correct in spirit but dangerous in execution: it invites trying to directly compare $2\sqrt{n}$ (not $\sqrt{4n+1}$, which is what the problem actually names) against $b(n)$, and then attempting to bound $|b(n) - 2\sqrt{n}|$ by Taylor-expanding g near $\sqrt{n}$. That approach can be made to work, but it is fighting the problem on the wrong terrain: it produces an inequality of the form $b(n) < 2\sqrt{n} + O(1/\sqrt{n})$ which is the right flavour but the wrong target, since the problem asks for an *exact* match of integer parts against $\sqrt{4n+1}$, not $2\sqrt{n}$, and $\sqrt{4n+1} \neq 2\sqrt{n}$ in general (they only agree in the limit). A solver who fixes their sights on $2\sqrt{n}$ can spend a long time trying to explain why the problem insisted on writing $\sqrt{4n+1}$ instead of the more natural $2\sqrt{n}$, without ever discovering that $4n+1$ is not decoration – it is the exact quantity that makes $b(m^2+m) = 2m+1$ land as an *equality* of integer values, not an approximation.

The model chain. Two structural facts about b do all the work, and neither requires estimating $\sqrt{n}$ at all.

Fact 1: b is strictly increasing. Let $k(n)$ be an integer achieving the minimum defining $b(n)$. Since $b(n-1)$ is a minimum over *all* positive integers k, in particular

$$b(n-1) \;\le\; k(n) + \frac{n-1}{k(n)} \;<\; k(n) + \frac{n}{k(n)} \;=\; b(n),$$

the middle inequality strict because $k(n) > 0$. So $b(n-1) < b(n)$ for every $n \ge 1$: b is strictly increasing on the positive integers.

Fact 2: the exact values $b(m^2) = 2m$ and $b(m^2+m) = 2m+1$. By AM-GM, for every positive real k, $k + m^2/k \ge 2\sqrt{m^2} = 2m$, with equality exactly at $k = m$ – and $k = m$ is a positive integer, so the *integer* minimum equals the real minimum here: $b(m^2) = 2m$. For $n = m^2 + m$: at $k = m$, the value is $m + (m^2+m)/m = m + m + 1 = 2m+1$; at $k = m+1$, it is $(m+1) + (m^2+m)/(m+1) = (m+1) + m = 2m+1$ as well. Since $k + n/k$ is convex in k (as a real function) and its real minimiser $\sqrt{m^2+m}$ lies strictly between m and $m+1$ (because

$m^2 < m^2 + m < (m+1)^2$), the two integers flanking the real minimiser are m and $m+1$, and for a convex function the integer minimum is attained at one of these two flanking integers. Both give $2m+1$, so $b(m^2+m) = 2m+1$ exactly.

Combining the two facts: since b is strictly increasing, $b(m^2) = 2m$, and $b(m^2+m) = 2m+1$, every n with $m^2 \le n < m^2+m$ satisfies $2m = b(m^2) \le b(n) < b(m^2+m) = 2m+1$, so $\lfloor b(n) \rfloor = 2m$; and every n with $m^2+m \le n < (m+1)^2$ satisfies $2m+1 = b(m^2+m) \le b(n) < b((m+1)^2) = 2m+2$, so $\lfloor b(n) \rfloor = 2m+1$.

Now compute $\lfloor \sqrt{4n+1} \rfloor$ on exactly the same two ranges, and the reason $\sqrt{4n+1}$ (not $2\sqrt{n}$) appears becomes visible. For $m^2 \le n \le m^2+m-1$: $4n+1$ ranges over $[4m^2+1,\ 4m^2+4m-3]$, and this entire interval lies strictly between $(2m)^2 = 4m^2$ and $(2m+1)^2 = 4m^2+4m+1$, so $\lfloor \sqrt{4n+1} \rfloor = 2m$ throughout. For $m^2+m \le n \le (m+1)^2-1$: $4n+1$ ranges over $[4m^2+4m+1,\ 4m^2+8m+1]$ – note the left endpoint is *exactly* $(2m+1)^2$, so $\sqrt{4(m^2+m)+1} = 2m+1$ exactly, matching $b(m^2+m) = 2m+1$ exactly – and the whole interval lies in $[(2m+1)^2, (2m+2)^2)$ since $4m^2+8m+1 < 4m^2+8m+4$, so $\lfloor \sqrt{4n+1} \rfloor = 2m+1$ throughout. The two floors agree on every n, term for term, over the same case boundaries.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both approaches begin from the same observation, that $b(n)$ should track $2\sqrt{n}$ near the real-valued critical point $x = \sqrt{n}$. They split at the very next step: the natural attempt tries to bound the *size* of the gap between $b(n)$ and $2\sqrt{n}$ using estimates (Taylor expansion, inequalities), aiming for an approximate comparison; the model instead pins down $b(n)$ at two *exact*, algebraically clean values – $n = m^2$ and $n = m^2+m$ – and lets strict monotonicity fill in every n between them without any estimation at all.

> *I don't need to know $b(n)$ for every n – I only need to know it exactly at the values of n where the answer changes, and since b is strictly increasing, pinning it down exactly at $n = m^2$ and $n = m^2+m$ (the two values where $\lfloor \sqrt{4n+1} \rfloor$ jumps) automatically pins down its floor everywhere in between.*

Layer 3 THE WHY

The choice of $n = m^2$ and $n = m^2+m$ as the two anchor points is not a guess; it is forced by asking where $\lfloor \sqrt{4n+1} \rfloor$ itself changes value, since that is the function we are trying to match. $\sqrt{4n+1} = 2m$ exactly would require $4n+1 = 4m^2$, impossible since $4n+1$ is odd and $4m^2$ is even – so the floor jumps to $2m$ strictly *after* $\sqrt{4n+1}$ passes $2m$, i.e. at the smallest n with $4n+1 > 4m^2$, which is $n = m^2$ (giving $4n+1 = 4m^2+1$, the least odd number exceeding $4m^2$). The next jump, to $2m+1$, happens where $4n+1 = (2m+1)^2 = 4m^2+4m+1$ exactly – an odd perfect square, so this one *is* achieved exactly, at $n = m^2+m$. These are precisely the two values at which the model chain evaluates b. This is the real content of the problem: $4n+1$ was chosen, instead of the more symmetric $4n$, exactly so that its square root is an odd integer at $n = m^2+m$ and strictly between consecutive even integers elsewhere – the "+1" is not noise, it is what makes $\sqrt{4n+1}$ an integer exactly where the discrete problem $b(n)$ also produces an integer, at the boundary case $k = m, k = m+1$ tying.

The AM-GM step is likewise not incidental: $k + n/k \ge 2\sqrt{n}$ is AM-GM applied to the two positive numbers k and n/k, whose product is the constant n – this is why the real-valued minimum of $g(x) = x + n/x$ is exactly $2\sqrt{n}$, attained when $k = n/k$, i.e. $k = \sqrt{n}$. When $\sqrt{n}$ happens to be an integer ($n = m^2$), the discrete and continuous problems coincide exactly, which is why $b(m^2) = 2m$ needs no convexity argument at all – only when $\sqrt{n}$ falls strictly between two integers (as at $n = m^2+m$, where $\sqrt{n} = \sqrt{m^2+m} \in (m, m+1)$) does convexity have to do any work, and even then only to justify that the true integer minimum lies at one of the two integers flanking the real minimiser, never further away.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the shape $k + n/k$ itself: on seeing a sum of a variable and a constant divided by it, write down "AM-GM gives a real minimum of $2\sqrt{n}$ at $k = \sqrt{n}$" immediately, before worrying about integrality.
2. That cue suggests asking exactly when $\sqrt{n}$ is an integer (giving an exact discrete minimum for free) and, when it is not, invoking convexity to restrict the integer search to the two integers flanking $\sqrt{n}$.
3. Hesitation sets in when trying to match $b(n)$ against $\sqrt{4n+1}$ rather than the more familiar $2\sqrt{n}$: the two expressions look like they should be interchangeable. Reinforce against it by testing $n = m^2 + m$ directly – $\sqrt{4n+1} = 2m+1$ exactly there, while $2\sqrt{n} = 2\sqrt{m^2+m}$ is irrational, showing the problem's chosen expression is exact where the natural one is only approximate.
4. The trap is trying to bound $|b(n) - 2\sqrt{n}|$ with estimates, chasing an inequality-flavoured proof for what is actually an equality-flavoured problem: both $b(n)$ and $\lfloor\sqrt{4n+1}\rfloor$ are piecewise constant in n, so the correct tool is pinning down exact values at the (few) points where each piece changes, not estimating a continuous approximation.
5. Recognise this faster next time by treating any "prove these two integer-valued (or floor-valued) functions of n agree" problem as a search for the exact points where each one changes value, rather than as an estimation problem.
6. Decision rule: to show $\lfloor f(n)\rfloor = \lfloor g(n)\rfloor$ for all n, when f is a discrete optimisation (here $b(n) = \min_k(k + n/k)$) and g is a closed form, first show f is monotonic in n, then locate the exact integers n where g's floor changes value, and verify f attains exactly those integer values there – monotonicity plus exact agreement at finitely many anchor points forces agreement of the floor everywhere between them.

1973 A4

1973 A4. How many zeros does the function $f(x) = 2^x - 1 - x^2$ have on the real line? [By a "zero" of a function f, we mean a value x_0 in the domain of f (here the set of all real numbers) such that $f(x_0) = 0$.]

Caveat. Statement read from the MAA OCR, where the exponent was garbled ("2′′′"); it is restored as 2^x. Identity check: the book's solution under 1973 A-4 answers "three; at 0, 1 and some $x > 1$" using $f'''(x) = (\log 2)^3 2^x \neq 0$, which solves exactly this statement. All values below were checked by hand and numerically (the third zero is $x \approx 4.2577$).

Fields: real analysis (zeros of functions, Rolle's theorem) **Position:** A4 — mid-paper by position, but one of the easier problems of 1973; the difficulty is entirely in proving there are *no more* zeros **Answer:** Exactly three: $x = 0$, $x = 1$, and one zero in $(4, 5)$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You test small integers, because 2^x and x^2 are both easy at integers. $f(0) = 1 - 1 - 0 = 0$ and $f(1) = 2 - 1 - 1 = 0$: two zeros found in thirty seconds. $f(2) = 4 - 1 - 4 = -1$, $f(3) = 8 - 1 - 9 = -2$, $f(4) = 16 - 1 - 16 = -1$, $f(5) = 32 - 1 - 25 = 6$. The sign change between 4 and 5 gives a third zero by the intermediate value theorem. For negative x you note $2^x < 1$, so $f(x) < -x^2 < 0$: nothing there.

Then the student has to say why there is nothing else on $[0, \infty)$, and the natural move is to study $f'(x) = 2^x \ln 2 - 2x$ and try to find where it vanishes. That equation, $2^x \ln 2 = 2x$, has no closed-form solution. The student sketches, estimates, argues that "f goes down then

up", and ends with a picture rather than a proof. Worse, the picture is misleading: between 0 and 1 the graph rises slightly ($f(\frac{1}{2}) = \sqrt{2} - 1 - \frac{1}{4} \approx 0.164$), so "down then up" is false, and a sketch-based count of turning points is exactly where marks go.

The model chain. *Existence of three zeros.* $f(0) = 0$, $f(1) = 0$, $f(4) = -1 < 0$, $f(5) = 6 > 0$; f is continuous, so there is $x_3 \in (4, 5)$ with $f(x_3) = 0$. These are three distinct zeros.

At most three zeros. f is infinitely differentiable, and

$$f'(x) = 2^x \ln 2 - 2x, \qquad f''(x) = 2^x (\ln 2)^2 - 2, \qquad f'''(x) = 2^x (\ln 2)^3 > 0.$$

Suppose f had four distinct zeros $a_1 < a_2 < a_3 < a_4$. By Rolle's theorem on each $[a_i, a_{i+1}]$, f' has three distinct zeros, one in each open interval (a_i, a_{i+1}), hence strictly ordered. Applying Rolle again, f'' has two distinct zeros, and once more, f''' has a zero. But $f''' > 0$ everywhere. Contradiction. So f has exactly three real zeros.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains agree on finding the three zeros. They split at the upper bound. The natural attempt tries to locate the turning points of f by solving $f'(x) = 0$ — a transcendental equation. The model does not locate anything: it differentiates until the derivative has a fixed sign, and counts.

> *I cannot solve $2^x \ln 2 = 2x$, and I do not need to. I need a count, not a location. Differentiate until the x^2 is gone: $f''' = (\ln 2)^3 2^x$ never vanishes, so f has at most three zeros.*

> *Negative x is also covered by the same count — I do not even need the separate argument $2^x < 1$. Three zeros found, at most three possible: done.*

Layer 3 THE WHY

Why is "differentiate three times" the natural move, and not a trick? Look at what f is made of: an exponential minus a polynomial of degree 2. Differentiation kills a polynomial of degree 2 in three steps and leaves the exponential alive forever (it only picks up factors of $\ln 2$). So after three derivatives you are left with a pure exponential, and a pure exponential has no zeros. The structure of the function tells you how many derivatives to take: the degree of the polynomial part plus one.

Why does a fixed-sign derivative bound the number of zeros? This is Rolle's theorem read as a counting statement. Between two zeros of a differentiable function there is a zero of its derivative. So

$$\#\{\text{zeros of } g\} \leq \#\{\text{zeros of } g'\} + 1,$$

applied three times: $\#\{\text{zeros of } f\} \leq \#\{\text{zeros of } f'''\} + 3 = 3$. You are not asked where the zeros are, you are asked how many. A counting question deserves a counting tool, and Rolle is the counting tool of calculus.

The competitor who is stuck on $f' = 0$ has confused two questions. "Where does f turn?" needs a solution of $2^x \ln 2 = 2x$. "How many times can f cross zero?" needs only an inequality. That confusion is the zombie-maths reflex: faced with a function, find its stationary points, whatever the question was. Before starting that computation, ask what the question needs. Here it needs an upper bound, and upper bounds on counts come from Rolle or from convexity.

There is an equivalent convexity reading worth having. f'' is strictly increasing, so it vanishes at most once, so f' is decreasing then increasing and has at most two zeros, so f is monotone on at most three intervals and has at most three zeros. Same argument, told upward instead of downward.

Finally, why must you find the third zero at all? Because the upper bound alone gives "at most three", and $x = 0, 1$ alone give "at least two". The trap is to stop at the two zeros

you can see and assume 2^x "eventually wins" without checking that it first loses. The table $f(2), f(3), f(4) < 0 < f(5)$ is the cheapest possible proof that it does.

Layer $+\alpha$ BACKTESTING

1. The first cue is "how many zeros" applied to exponential minus polynomial. Write down: the zeros you can see (0 and 1), and $f'''(x) = (\ln 2)^3 2^x > 0$.
2. Rolle's theorem as a zero-counting inequality, plus the intermediate value theorem for existence.
3. Hesitation comes when $f'(x) = 0$ will not solve. It is reinforced away by asking "count or location?" — a count never needs the location of the turning points.
4. Two traps: stopping at the two visible zeros, and a sketch that claims "f decreases then increases" (false on $[0, 1]$). The table of integer values exposes the hidden third zero in $(4, 5)$.
5. Any "how many solutions" question for e^x-type minus polynomial of degree d: differentiate $d + 1$ times and count. Recognise it from the shape of the function, not from the numbers.
6. For a count of zeros, get the lower bound by sign changes and the upper bound by Rolle: if $g^{(k)}$ has no zeros, g has at most k. Never solve for stationary points when the question only asks how many.

▷ **Cue:** "how many zeros" for exponential minus a polynomial of degree d ⟶ **Move:** differentiate $d + 1$ times; the result has fixed sign, so at most $d + 1$ zeros
▷ **Cue:** you cannot solve $f'(x) = 0$ in closed form ⟶ **Move:** ask whether you need the location or only the count; for a count use Rolle
▷ **Cue:** an upper bound on zeros is known and you have found fewer ⟶ **Move:** tabulate at integers and look for a sign change
▷ **Cue:** f'' strictly monotone ⟶ **Move:** f' has at most two zeros, f at most three: convexity count

1973 A5

1973 A5. A particle moves in 3-space according to the equations:

$$\frac{dx}{dt} = yz, \qquad \frac{dy}{dt} = zx, \qquad \frac{dz}{dt} = xy.$$

[Here $x(t), y(t), z(t)$ are real-valued functions of the real variable t.] Show that:
(a) If two of $x(0), y(0), z(0)$ equal zero, then the particle never moves.
(b) If $x(0) = y(0) = 1, z(0) = 0$, then the solution is:

$$x = \sec t, \quad y = \sec t, \quad z = \tan t;$$

whereas if $x(0) = y(0) = 1, z(0) = -1$, then

$$x = 1/(t+1), \quad y = 1/(t+1), \quad z = -1/(t+1).$$

(c) If at least two of the values $x(0), y(0), z(0)$ are different from zero, then either the particle moves to infinity at some finite time in the future, or it came from infinity at some finite time in the past. [A point (x, y, z) in 3-space "moves to infinity" if its distance from the origin approaches infinity.]

Caveat. Statement read from the MAA OCR with the derivative fractions, sec, tan and the initial values (1, −1, read as "I" in places) restored. Identity check: the book's solution under 1973 A-5 uses $xx' = yy' = zz' = xyz$ and the uniqueness theorem, and says part (b)

"was intended as a hint for part (c)", which matches this statement. The book's part (c) is a sketch ("set $c_3 = 0$", "let time move in the direction which makes $|z|$ increase"); the case analysis below is my own and was verified step by step, including the degenerate case $y^2 \equiv z^2$ that the sketch does not separate out.

Fields: differential equations (autonomous systems, first integrals, finite-time blow-up) **Position:** A5 — a three-part problem whose first two parts are routine; (c) needs a conserved quantity and a case split **Answer:** (a) the constant function is a solution and solutions are unique; (b) direct substitution plus uniqueness; (c) $x^2 - z^2$ and $y^2 - z^2$ are constant, which reduces the system to a scalar equation $u' = \pm\sqrt{(u^2+a)(u^2+b)}$ or $u' = u^2 - a$, and both blow up in finite time in at least one direction.

Layer 1 THE WHAT — surface analysis

The natural attempt. Parts (a) and (b) go smoothly. For (a), with $y(0) = z(0) = 0$, the constant triple $(x(0), 0, 0)$ satisfies all three equations, and "the solution is unique because the right-hand side is a polynomial" finishes it. For (b), you differentiate: $(\sec t)' = \sec t \tan t = yz$ and $(\tan t)' = \sec^2 t = xy$; $(1/(t+1))' = -1/(t+1)^2 = yz$ and $(-1/(t+1))' = 1/(t+1)^2 = xy$.

For (c) the natural attempt is to try to solve the system. You notice from (b) that $x = y$ in both examples and guess that setting $x = y$ reduces the system to $x' = xz$, $z' = x^2$. That pair still does not separate. You try differentiating again: $z'' = 2xx' = 2x^2 z = 2zz'$, so $z' = z^2 + C$, a Riccati equation that blows up. Good — but only on the invariant set $x = y$. For general initial data there is no reason for $x = y$, and the attempt stalls: the student has a solved special case and no bridge to the general case.

The model chain. *Preliminaries.* The right-hand side (yz, zx, xy) is a polynomial, hence locally Lipschitz; by the Picard–Lindelöf theorem each initial point has a unique maximal solution. Three symmetries map solutions to solutions: (i) any permutation of (x, y, z) (the system is symmetric); (ii) negating any two coordinates, e.g. $(x, y, z) \mapsto (x, -y, -z)$ (each equation has either both or neither side negated); (iii) $(x, y, z)(t) \mapsto -(x, y, z)(-t)$, which reverses time.

(a) If $y(0) = z(0) = 0$, the constant $(x(0), 0, 0)$ is a solution; by uniqueness it is the solution. The other cases follow by symmetry (i).

(b) Substitution, as above, and uniqueness make each the solution (on $|t| < \pi/2$, resp. $t > -1$).

(c) Step 1: first integrals. Multiply the equations by x, y, z:

$$xx' = yy' = zz' = xyz.$$

So $x^2 - z^2 = a$ and $y^2 - z^2 = b$ are constant along every solution. By (i), assume $z(0)^2 \le y(0)^2 \le x(0)^2$; then $a \ge b \ge 0$, and for all t,

$$x^2 = z^2 + a, \qquad y^2 = z^2 + b.$$

Step 2: case $b > 0$. Then $x^2, y^2 \ge b > 0$ always, so x and y never vanish and keep their signs. By (ii) (negate (x, z) or (y, z) as needed) assume $x, y > 0$. Then

$$z' = xy = \sqrt{(z^2+a)(z^2+b)} =: g(z), \qquad g(z) \ge z^2 + b > 0.$$

This is a separable scalar equation: $t = \int_{z(0)}^{z} d\zeta / g(\zeta)$. Since $0 < 1/g(\zeta) \le 1/(\zeta^2 + b)$,

$$T_+ = \int_{z(0)}^{\infty} \frac{d\zeta}{g(\zeta)} \le \frac{\pi}{\sqrt{b}} < \infty,$$

and z increases to $+\infty$ as $t \uparrow T_+$. (Explicitly, the inverse of $z \mapsto \int_{z(0)}^{z} d\zeta/g$ gives z on $[0, T_+)$, and $x = \sqrt{z^2+a}$, $y = \sqrt{z^2+b}$ then satisfy $x' = zz'/x = zy$, $y' = zx$; by uniqueness this is our

solution.) The distance from the origin is at least $|z| \to \infty$: the particle moves to infinity at time T_+. (It also came from infinity at a finite past time, by the same integral to $-\infty$.)

Step 3: case $b = 0$. Now $y^2 = z^2$ for all t. If $z(0) = 0$ then $y(0) = 0$ and at most one coordinate is non-zero, excluded by hypothesis. So $z(0) \neq 0$. Moreover z never vanishes: if $z(t_1) = 0$ then $y(t_1) = 0$, and by (a) applied at time t_1 the solution is constant with $z \equiv 0$, contradicting $z(0) \neq 0$. Hence $y = \varepsilon z$ with a fixed sign ε; by (ii) (negate (x, y)) assume $\varepsilon = 1$. Then

$$x' = yz = z^2 = x^2 - a, \qquad x^2 = z^2 + a > a.$$

If $x(0) > \sqrt{a}$: x increases, and it reaches $+\infty$ at time $\int_{x(0)}^{\infty} d\xi/(\xi^2 - a) < \infty$ (the integrand is continuous and positive on $[x(0), \infty)$ and $\sim \xi^{-2}$). If $x(0) < -\sqrt{a}$: $x' > 0$ still, so going backwards in time x decreases, reaching $-\infty$ at time $-\int_{-\infty}^{x(0)} d\xi/(\xi^2 - a) > -\infty$: the particle came from infinity. Either way $|x| \to \infty$. (With $a = 0$ this is $x = 1/(1/x(0) - t)$; the second example of (b) is $x(0) = 1$ after the sign change, blowing up at $t = -1$.)

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains verify (a) and (b). At (c) the natural attempt tries to *solve* the system, and can only do it on the special set $x = y$ suggested by (b). The model looks for quantities the flow *preserves*, and finds $x^2 - z^2$ and $y^2 - z^2$ from the line $xx' = yy' = zz'$.

> *Three unknowns and I cannot solve the system. What does the motion conserve? Each right-hand side is the product of the other two, so if I multiply $x' = yz$ by x I get xyz — and the same for y and z. So $xx' = yy' = zz'$, and the differences of squares are constant.*

> *Two constants of motion in three dimensions: the orbit lies on a curve, so this is really a one-variable problem. Write everything in terms of the smallest coordinate z and look at z'.*

> *Growth like $u' \geq u^2$ blows up in finite time, because $\int^{\infty} du/u^2$ converges. So I only need a lower bound of the form $z' \geq z^2 + b$.*

Layer 3 THE WHY

Why look for conserved quantities at all? Because "moves to infinity at some finite time" is a statement about one scalar, the distance, and the system has three. An autonomous system in $\mathbb{R}^3$ that you cannot solve is made tractable by reducing dimension, and the only way to reduce dimension without solving is a first integral — a function $H(x, y, z)$ with $\frac{d}{dt}H = 0$ along solutions. Each first integral confines the orbit to a surface; two of them confine it to a curve, and on a curve the motion is a scalar ODE, which you can always separate.

Why these particular integrals? The equations have a very specific shape: x' equals the product of the *other* two variables. Multiplying by the missing variable completes the product, $xx' = xyz$, and does so identically for all three equations. It is the same structure as the Euler equations for a free rigid body, where the conserved quantities are also combinations of squares. When every equation of a system is "derivative of one variable = something missing only that variable", multiply by the missing variable and compare.

Why is finite-time blow-up the expected outcome? A scalar equation $u' = h(u)$ gives $t = \int du/h(u)$, and the time to reach infinity is $\int^{\infty} du/h(u)$. For $h(u) \sim u^2$ this integral converges: quadratic growth means the particle covers infinitely much space in finite time. The Riccati equation $u' = u^2$ with solution $1/(c - t)$ is the model, and part (b) hands you exactly that model ($1/(t + 1)$ and $\tan t$). The examiners are telling you what to expect in (c); the official solution says in so many words that (b) "was intended as a hint".

Why does the hypothesis "at least two non-zero" appear? Because it is exactly the condition that the initial point is not an equilibrium. The equilibria of the system are the points where $yz = zx = xy = 0$, i.e. where at least two coordinates vanish — part (a). The theorem says: every non-equilibrium orbit blows up. In the proof, the hypothesis is used in precisely one place: in the case $y^2 \equiv z^2$ it rules out $y = z = 0$. That is where to look for it, and a proof that never uses a hypothesis has a hole.

Finally, the case split $b > 0$ versus $b = 0$ is forced, not decorative. When $b > 0$, x and y are bounded away from 0, so xy has a fixed sign and z is monotone. When $b = 0$, y and z can have the same size and $z' = xy$ no longer has a clean lower bound in terms of z alone; you switch to x, whose equation $x' = \pm(x^2 - a)$ is the Riccati model once more. The book's sketch treats both at once ("let time move in the direction which makes $|z|$ increase"); writing it out, the case $b = 0$ is where direction genuinely depends on the sign of $x(0)$, and that is why the conclusion is "future *or* past".

Layer $+\alpha$ BACKTESTING

1. The cue is the shape $x' = yz$, $y' = zx$, $z' = xy$: each derivative is the product of the other two. Write down $xx' = yy' = zz' = xyz$ immediately, and the constants $x^2 - z^2$, $y^2 - z^2$.

2. Reduce by first integrals to a scalar autonomous ODE, separate variables, and test convergence of $\int^{\infty} du/h(u)$.

3. Hesitation comes after (b), when the explicit solutions do not generalise. Reinforce with the question "what is conserved?" before "what is the solution?".

4. The traps: solving only on the invariant set $x = y$; forgetting uniqueness in (a) and (b); and missing the degenerate case $y^2 \equiv z^2$, where the hypothesis is actually needed and where the blow-up may be in the past only.

5. Any system where each equation is "u_i' = expression not involving u_i" invites multiplying by u_i (or by a suitable factor) to produce a common right-hand side. Finite-time blow-up is recognised by superlinear growth: $u' \gtrsim u^{1+\delta}$.

6. When a system cannot be solved, count its first integrals; with $n - 1$ of them in $\mathbb{R}^n$ the motion is one-dimensional and separable. Then blow-up in finite time is decided by one improper integral.

▷ **Cue:** $x' = yz$, $y' = zx$, $z' = xy$ (each derivative is the product of the others) ⟶ **Move:** multiply by the missing variable: $xx' = yy' = zz'$, so differences of squares are constant

▷ **Cue:** an autonomous system in $\mathbb{R}^n$ you cannot solve ⟶ **Move:** find first integrals to reduce to a scalar ODE

▷ **Cue:** "moves to infinity in finite time" ⟶ **Move:** show $u' \geq c\,u^2$ (or superlinear) and use $\int^{\infty} du/u^2 < \infty$

▷ **Cue:** earlier part gives explicit solutions like $\tan t$ or $1/(t+1)$ ⟶ **Move:** the later part expects Riccati-type blow-up; aim for $u' = u^2 + c$

▷ **Cue:** a hypothesis "at least two non-zero" in a dynamics problem ⟶ **Move:** it excludes equilibria; find the single case of the proof where it is used

1973 A6

1973 A6. Prove that it is impossible for seven distinct straight lines to be situated in the euclidean plane so as to have at least six points where exactly three of these lines intersect and at least four points where exactly two of these lines intersect.

Caveat. Statement read from the MAA OCR; it came through clean. Identity check: the

book's solution under 1973 A-6 counts the $\binom{7}{2} = 21$ pairs of lines against $6 \cdot 3 + 4 \cdot 1 = 22$, which solves exactly this statement.

Fields: combinatorics (double counting, incidence geometry) **Position:** A6 — last problem of the morning, but a short one once the right quantity is counted; its difficulty is the reflex to draw rather than count **Answer:** Count pairs of lines: each pair meets in at most one point, a triple point uses 3 pairs and a double point uses 1, so the configuration needs $6 \cdot 3 + 4 = 22$ pairs, but there are only $\binom{7}{2} = 21$.

Layer 1 THE WHAT — surface analysis

The natural attempt. The student draws. Seven lines, try to make six triple points: start with a triangle and its medians, or the Fano-plane picture (seven lines, seven triple points — but the Fano plane is not realisable in the real plane, a fact the student may half-remember). Each drawing gives some triple points and a scatter of double points, and the student tries to push the count of triple points up while keeping four double points. After several drawings the best found is something like six triple points with only three ordinary crossings, or five triple points with many double points. The student is now convinced the configuration is impossible, but has no proof: failing to draw it is not an argument. The attempts may also drift into projective geometry or Sylvester–Gallai-type facts, which are true and irrelevant.

The model chain. Call a point where exactly three of the lines meet a *triple point*, and one where exactly two meet a *double point*. Let $\mathcal{P}$ be the set of unordered pairs of the seven lines; $|\mathcal{P}| = \binom{7}{2} = 21$.

Assign to each intersection point X the set $\mathcal{P}(X)$ of pairs of lines that both pass through X. If exactly three lines pass through X, then $|\mathcal{P}(X)| = \binom{3}{2} = 3$; if exactly two, $|\mathcal{P}(X)| = 1$.

For distinct points $X \neq Y$, the sets $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ are disjoint: a pair of distinct lines in the plane meets in at most one point, so it cannot pass through both X and Y.

Suppose there were six triple points $X_1, \ldots, X_6$ and four double points $Y_1, \ldots, Y_4$ (take any six and any four if there are more). These ten points are distinct (a point cannot be both a triple and a double point), so

$$21 = |\mathcal{P}| \;\geq\; \sum_{i=1}^{6} |\mathcal{P}(X_i)| + \sum_{j=1}^{4} |\mathcal{P}(Y_j)| = 6 \cdot 3 + 4 \cdot 1 = 22,$$

a contradiction. Hence no such configuration exists.

Layer 2 THE HOW — the bifurcation

Bifurcating point. The chains split at the very first step. The natural attempt treats the problem as a construction to be attempted in a picture. The model treats "impossible" as a demand for a quantity that is too small, and counts the pairs of lines.

> *"Prove it is impossible" with specific numbers — seven, six, four. The numbers are the argument. Something is being overspent. What is there a fixed supply of? Seven lines give $\binom{7}{2} = 21$ pairs, and each pair meets at most once.*

> *A point where three lines meet eats $\binom{3}{2} = 3$ pairs; a point where two meet eats one. Different points cannot share a pair. $6 \cdot 3 + 4 = 22 > 21$.*

Layer 3 THE WHY

Why count pairs of lines, and not, say, points or lines? Start from the definition of the objects. A "point where exactly three lines intersect" is defined by the lines through it; it is an incidence, not a location. The one axiom about lines that the problem allows you to use is Euclid's: two

distinct lines meet in at most one point. Read the other way round, that axiom says each *pair* of lines is responsible for at most one intersection point. So pairs of lines are the currency, intersection points are the purchases, and the question is whether the budget covers the bill. That is not a trick: it is the definition of an intersection point translated into a count.

Why is it natural to expect a counting proof? Because of the form of the statement. "Impossible" plus explicit small numbers (7, 6, 4) almost always means an inequality that fails by a little. Here it fails by one: 22 versus 21. The "at least four double points" clause is there precisely to push the total past 21 — with only three double points you get exactly 21, and the count gives no contradiction. When an impossibility statement has a clause that looks arbitrary, ask what the count would be without it; the clause is usually what tips the inequality.

Why does drawing fail as a strategy? A drawing can only prove existence. To prove non-existence you need a quantity that every configuration must satisfy, and pictures do not produce quantities. The picture-first reflex is the geometric version of zombie maths: it feels like progress, it costs the whole time allowance, and it cannot reach the goal even when it "succeeds".

The underlying tool is double counting of incidences: count the set of (pair of lines, intersection point) incidences in two ways — by pairs (each contributes at most one) and by points (a point on exactly k lines contributes $\binom{k}{2}$). The resulting identity for a configuration of n lines is

$$\sum_X \binom{k_X}{2} \;\le\; \binom{n}{2},$$

with equality when no two lines are parallel. This problem is one instance of that identity; so are many olympiad statements about points and lines, and the same inequality is the start of the Sylvester–Gallai and Szemerédi–Trotter circle of results.

Layer $+\alpha$ BACKTESTING

1. The first cue is "prove that it is impossible" together with specific small numbers of lines and intersection points. Write down $\binom{7}{2} = 21$ and the cost $\binom{k}{2}$ of a point on k lines.
2. Double counting of incidences between pairs of lines and intersection points.
3. Hesitation comes from the urge to draw and from half-remembered theorems (Fano plane, Sylvester–Gallai). It is reinforced away by asking "what is there a fixed supply of?".
4. The trap is spending the time on constructions. A smaller trap is forgetting to state that distinct intersection points use disjoint sets of pairs; that is the step that needs the axiom.
5. Impossibility statements with explicit small numbers are counting problems: find the resource whose total is fixed and show the demand exceeds it. An arbitrary-looking extra clause ("at least four double points") is the one that tips the count.
6. To prove a configuration cannot exist, identify the objects that the definitions are built from (here pairs of lines), count what each feature consumes, and compare with the supply. Never try to prove non-existence by failing to construct.

▷ **Cue:** "prove it is impossible" with explicit small numbers $\longrightarrow$ **Move:** find a fixed supply and show the demand exceeds it
▷ **Cue:** points where exactly k of n lines meet $\longrightarrow$ **Move:** each costs $\binom{k}{2}$ of the $\binom{n}{2}$ pairs; distinct points use disjoint pairs
▷ **Cue:** an arbitrary-looking extra condition in an impossibility statement $\longrightarrow$ **Move:** compute the count without it; the condition is what breaks the inequality
▷ **Cue:** two distinct lines meet in at most one point $\longrightarrow$ **Move:** each pair of lines is charged to at most one intersection point

Putnam 1973 — Paper B

1973 B1

1973 B1. Let $a_1, a_2, \ldots, a_{2n+1}$ be a set of integers such that, if any one of them is removed, the remaining ones can be divided into two sets of n integers with equal sums. Prove $a_1 = a_2 = \cdots = a_{2n+1}$.

Fields: number theory (2-adic descent, congruences) **Position:** B1 – opening problem of the second session **Answer:** No numerical answer; the conclusion is that all $2n+1$ integers must be equal.

Layer 1 THE WHAT — surface analysis

The natural attempt. The phrase "removed... split into two equal-sum halves" screams parity, so you write it down immediately: if a_i is removed, the remaining $2n$ integers split into two sets of n with equal sums T, so their total is $2T$ – even. That total is $S - a_i$, where $S = a_1 + \cdots + a_{2n+1}$. So $S - a_i$ is even for every i, which means every a_i has the same parity (namely the parity of S). This is real progress and it was free – one line.

Now what? You have a genuine necessary condition, but "all the same parity" is a much weaker statement than "all equal," and it is not obvious how to squeeze more out of the hypothesis. You try to find a counterexample to make sure you are not about to waste time proving something false: take $n = 2$ (five integers), all even, not all equal – say $0, 0, 0, 0, 4$. Remove the 4: remaining $0, 0, 0, 0$ splits into $\{0, 0\}$ and $\{0, 0\}$, sums equal. Good, consistent with the hypothesis so far. But now remove a 0: remaining $0, 0, 0, 4$ must split into two pairs with equal sums. The three ways to split $\{0, 0, 0, 4\}$ into two pairs give sums $(0, 4)$, $(0, 4)$, $(0, 4)$ – never equal. So this attempted counterexample fails, which is reassuring, but you have only checked one example by hand; you have not shown *why* every all-same-parity, non-equal configuration must fail, and hand-checking is not a proof.

The model chain. Having established that all a_i are congruent mod 2, ask: is the balancing hypothesis itself preserved if I shrink the integers? Let $r \in \{0, 1\}$ be the common residue, and set $b_i = (a_i - r)/2$, which is an integer for every i. Take any subset A of the indices with $|A| = n$, and its complement A' among the remaining $2n$ indices (after removing some a_k), also of size n. Then

$$\sum_{i \in A} a_i = \sum_{i \in A} (2b_i + r) = 2 \sum_{i \in A} b_i + nr,$$

and likewise for A'. So $\sum_{i \in A} a_i = \sum_{i \in A'} a_i$ if and only if $\sum_{i \in A} b_i = \sum_{i \in A'} b_i$ – the nr terms cancel exactly *because both subsets have size n*. This means the b_i satisfy the identical hypothesis: removing any one of them, the rest split into two size-n subsets of equal sum.

But that is exactly the original problem, restated for the b_i. Apply the one-line parity argument again: all b_i share a common parity. Translating back, this says $a_i = 2b_i + r$ with the b_i all $\equiv$ some fixed residue mod 2, i.e. all $a_i \equiv r \pmod 4$. Nothing stops the same substitution from being applied to the b_i in turn, and again, and again: at every stage k, all a_i are congruent modulo 2^k.

Now invoke the one fact about integers that closes the argument: if two integers agree modulo 2^k for every $k \geq 1$, they are equal. This is not a limiting or topological statement in disguise – it is finite: a nonzero integer $d = a_i - a_j$ has a finite number of factors of 2 (its 2-adic valuation $v_2(d)$ is some finite number), so $2^{v_2(d)+1} \nmid d$, contradicting "$2^k \mid d$ for all k" once $k > v_2(d)$. Hence $d = 0$ for every pair i, j, i.e. $a_1 = a_2 = \cdots = a_{2n+1}$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains reach the same first line: $S - a_i$ even for every i, so all a_i share a parity. They split immediately after. The natural attempt treats this as one fact to be combined with something else (extremal reasoning, direct case-checking) and starts testing examples. The model chain instead asks whether the same one-line argument can be re-run on a rescaled copy of the problem.

> *I derived "all a_i have the same parity" using nothing but the definition of the hypothesis – so if I subtract off that common residue and halve, do the rescaled integers satisfy the exact same hypothesis? If yes, I have not used up the problem, I have shrunk it, and I can run the identical argument again.*

> *The reason the constant r cancels when I check whether the rescaled integers balance is that both subsets in the partition have size exactly n – if the two sides of the split were different sizes, the nr term would not cancel and the induction would break. So the fact that the problem insists on* two sets of n*, not just any two-part split, is not incidental; it is the hinge the whole descent depends on.*

Layer 3 THE WHY

Why does the parity step alone feel insufficient, and why is that feeling correct? Because "all a_i are congruent mod 2" is consistent with infinitely many non-equal configurations (any all-even or all-odd set satisfies it), so parity mod 2 cannot by itself force equality; it only rules out roughly half of all possible integer tuples. The natural attempt's discomfort after one step is the right instinct – one congruence relation is never going to pin down an infinite set to a single point, no matter how many integers are involved.

The resolution is to notice that the hypothesis is not really "a fact about these particular integers" – it is a statement about a family closed under an operation, namely $a_i \mapsto (a_i - r)/2$. Whenever a hypothesis is invariant under a shrinking map like this, proving a single instance of the conclusion (here, a congruence mod 2) is not the end of the proof; it is the base case of an induction that the map itself hands you for free. The problem did not need to say "prove this holds mod every power of 2" – once you see the affine self-similarity, that stronger statement is available at no extra cost, because the *same* one-line argument closes each stage.

Finally, why does "congruent mod 2^k for all k" force equality, rather than merely a very small difference? Because the integers are not real numbers: there is no integer with infinitely many factors of 2 except 0 itself. This is the discrete analogue of the fact that a nonzero real number has a well-defined order of magnitude – here the "size" being controlled is the 2-adic valuation, and driving it above every finite bound is only possible for the zero difference. This is the axiom actually doing the work at the very end: not a limiting argument, but the elementary fact that factorization of a fixed nonzero integer terminates.

Layer $+\alpha$ BACKTESTING

1. The cue is "remove any one, the rest split into two sets of n with equal sums." On seeing it, write down immediately: the sum of the remaining $2n$ integers is twice the common value, hence even, for every choice of removed element.

2. That cue suggests a parity (mod 2) argument on the a_i, but the deeper method it should suggest is asking whether the same hypothesis, and hence the same argument, survives an affine rescaling of the integers.

3. Hesitation occurs right after "all a_i have the same parity," when that single congruence feels too weak to imply full equality; it is reinforced away by asking explicitly, "does subtracting the common residue and halving preserve the exact same balancing hypothesis?" – checking that the subset sizes (n and n) are what make the constant term cancel.

4. The trap is stopping at one congruence and reaching for an unrelated tool (extremal case-checking, explicit small counterexamples) instead of recognising the hypothesis as self-similar under rescaling, which would hand you the full induction for free.

5. Any time a combinatorial hypothesis about equal-sized, equal-sum subsets survives being stated for a rescaled version of the same data (shift by a constant, divide by a fixed integer), recognise that as an invitation to induct on the number of times the rescaling can be applied.

6. Reusable rule: if a balance/equal-sum hypothesis on integers is invariant under $x \mapsto (x - r)/m$ for a fixed integer $m > 1$ (because the two sides of every relevant partition have matching size, so the constant term cancels identically), then a single application giving "all values congruent mod m" is only the base case; iterate to get congruence mod m^k for every k, and conclude the values are equal, since no two distinct integers agree modulo arbitrarily high powers of a fixed $m > 1$.

▷ **Cue:** a condition forces $S - a_i$ (or some fixed combination minus each term) to be even for every choice of removed/distinguished term $\longrightarrow$ **Move:** conclude all the a_i are congruent mod 2

▷ **Cue:** a combinatorial equal-subset-sum hypothesis is stated on integers a_i, and both sides of every relevant split have the same size $\longrightarrow$ **Move:** check whether the hypothesis is invariant under $a_i \mapsto (a_i - r)/2$ (constant shift, then halve) – if so, treat the first congruence found as the base case of an induction, not the answer

▷ **Cue:** two integers are shown congruent modulo 2^k for every k (or modulo m^k for fixed $m > 1$) $\longrightarrow$ **Move:** conclude they are equal, since a nonzero integer has only finitely many factors of 2 (finite 2-adic valuation)

1973 B2

1973 B2. Let $z = x + iy$ be a complex number with x and y rational and with $|z| = 1$. Show that the number $|z^{2n} - 1|$ is rational for every integer n.

Fields: complex numbers, number theory (rationality) **Position:** B2 **Answer:** $|z^{2n} - 1| = 2|\operatorname{Im}(z^n)|$, and this is rational because $\operatorname{Im}(z^n)$ is a $\mathbb{Z}$-polynomial expression in the rationals x, y.

Layer 1 THE WHAT — surface analysis

The natural attempt. You want $|z^{2n} - 1|$, and the standard move for a modulus is to multiply by the conjugate: $|w|^2 = w\bar{w}$ for any complex w. With $w = z^{2n} - 1$,

$$|z^{2n} - 1|^2 = (z^{2n} - 1)(\overline{z^{2n}} - 1) = |z^{2n}|^2 - z^{2n} - \overline{z^{2n}} + 1 = 1 - 2\operatorname{Re}(z^{2n}) + 1,$$

using $|z| = 1$ so $|z^{2n}| = |z|^{2n} = 1$. So $|z^{2n} - 1|^2 = 2 - 2\operatorname{Re}(z^{2n})$, and $|z^{2n} - 1| = \sqrt{2 - 2\operatorname{Re}(z^{2n})}$. This is genuinely elegant progress: you have reduced a modulus of a complex expression to a real quantity built entirely from $\operatorname{Re}(z^{2n})$.

Now show $\operatorname{Re}(z^{2n})$ is rational: since $z = x + iy$ with $x, y \in \mathbb{Q}$, and z^{2n} is a polynomial in x, y with integer coefficients (expand $(x + iy)^{2n}$ by the binomial theorem and collect the terms with even powers of i – these are real and contribute to Re), $\operatorname{Re}(z^{2n})$ is a rational number for $n \geq 0$; a symmetric argument via $z^{-1} = \bar{z}$ (valid since $|z| = 1$) handles negative n. So $\operatorname{Re}(z^{2n}) \in \mathbb{Q}$, hence $2 - 2\operatorname{Re}(z^{2n}) \in \mathbb{Q}$.

And now you are stuck, or ought to feel stuck: $|z^{2n} - 1| = \sqrt{2 - 2\operatorname{Re}(z^{2n})}$, and the square root of a rational number is not, in general, rational. You have proved the *argument of the square*

root is rational, which is not the same statement as the problem asks for. Something extra about this particular rational number – not an arbitrary one – has to force it to be a perfect square of a rational. Grinding out $\mathrm{Re}(z^{2n})$ explicitly as a messy binomial sum in x, y will not, by inspection, reveal why $2 - 2\,\mathrm{Re}(z^{2n})$ happens to be such a square.

The model chain. The fix is to stop working with z^{2n} directly and instead introduce the "half-angle" object $w = z^n = u + iv$ (real u, v), so that $z^{2n} = w^2$ and $|w| = |z|^n = 1$, i.e. $u^2 + v^2 = 1$. Now compute $|w^2 - 1|^2$ directly in terms of u, v rather than via Re alone:

$$w^2 - 1 = (u^2 - v^2 - 1) + 2uvi, \qquad |w^2 - 1|^2 = (u^2 - v^2 - 1)^2 + (2uv)^2.$$

Substitute $u^2 - 1 = -v^2$ (from $u^2 + v^2 = 1$) into the first term: $u^2 - v^2 - 1 = -v^2 - v^2 = -2v^2$. So

$$|w^2 - 1|^2 = 4v^4 + 4u^2v^2 = 4v^2(u^2 + v^2) = 4v^2,$$

using $u^2 + v^2 = 1$ again in the last step. Hence $|z^{2n} - 1| = |w^2 - 1| = 2|v| = 2|\,\mathrm{Im}(z^n)|$ – an *exact* identity, no approximation, and manifestly a perfect square times 4 before the square root is even taken.

(This is consistent with the natural attempt's formula: $2 - 2\,\mathrm{Re}(w^2) = 2 - 2(u^2 - v^2) = 2 - 2u^2 + 2v^2$, and since $1 - u^2 = v^2$, this is $2v^2 + 2v^2 = 4v^2$ as well – the two computations agree, but the second form makes the perfect-square structure visible for free, while the first hides it inside a generic-looking rational number.)

It remains only to show $v = \mathrm{Im}(z^n)$ is rational. Since $x, y \in \mathbb{Q}$, expand $z^n = (x + iy)^n$ by the binomial theorem for $n \geq 0$:

$$z^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k}(iy)^k,$$

and every term with k odd contributes $\pm i \cdot (\text{rational})$ to the sum (since $i^k = \pm i$ for odd k), while every term with k even contributes a rational real number. Collecting the odd-k terms gives $\mathrm{Im}(z^n)$ as an integer-coefficient polynomial evaluated at the rationals x, y (with an overall factor absorbing the sign of i^k), hence $\mathrm{Im}(z^n) \in \mathbb{Q}$. For $n < 0$, $|z| = 1$ gives $z^{-1} = \bar{z} = x - iy$, which again has rational real and imaginary parts, so the same binomial argument applied to $(z^{-1})^{|n|} = z^n$ shows $\mathrm{Im}(z^n) \in \mathbb{Q}$ for negative n too. Therefore $|z^{2n} - 1| = 2|\,\mathrm{Im}(z^n)|$ is rational for every integer n.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains agree completely up through $|z^{2n} - 1|^2 = 2 - 2\,\mathrm{Re}(z^{2n})$ (equivalently $4v^2$, once you know to look for it). They split at what to do with that expression: the natural attempt tries to certify rationality of $\mathrm{Re}(z^{2n})$ directly and stalls at the square root; the model chain instead reroutes through $w = z^n$ and computes $|w^2 - 1|^2$ from scratch in terms of $u, v = \mathrm{Re}(w), \mathrm{Im}(w)$, which exposes the perfect-square structure $4v^2$ that was always there but invisible in the $\mathrm{Re}(z^{2n})$-only form.

> *I have shown the quantity under the square root is rational, but rational numbers are not closed under square roots – I have proved a necessary condition, not the conclusion, and I need some extra structural fact about* this specific *rational number that forces it to be a perfect square.*

> *The modulus $|z^{2n} - 1|$ only involves z^{2n}, which is a perfect square, $(z^n)^2$ – so instead of expanding the exponent $2n$ directly, I should introduce $w = z^n$ and see what $|w^2 - 1|$ looks like purely in terms of $w = u + iv$ with $u^2 + v^2 = 1$; this rewrites the target as a function of the halved exponent, where the unit-circle constraint has a chance to force a perfect square.*

Layer 3 THE WHY

Why does $\mathrm{Re}(z^{2n})$ being rational fail to finish the problem, even though it feels like exactly the right quantity to control? Because rationality of c says nothing about rationality of $\sqrt{c}$ in general ($c = 2$ is the standard counterexample) – the square-root operation is not rational-preserving, only field operations ($+, -, \times, \div$) are. Proving $2 - 2\,\mathrm{Re}(z^{2n}) \in \mathbb{Q}$ therefore proves membership in the wrong algebraic category for what the problem is asking; you need membership in $\mathbb{Q}$ *after* the square root, which is a strictly stronger and different claim.

Why does routing through $w = z^n$ resolve this, rather than being a cosmetic change of variable? Because the constraint $|z| = 1$ is not merely "some real number is 1" – it is the single algebraic relation $u^2 + v^2 = 1$ that is available to convert between $\mathrm{Re}(w^2) = u^2 - v^2$ and $v^2 = \mathrm{Im}(w)^2$. The identity $2 - 2(u^2 - v^2) = 4v^2$ uses this relation twice (once to replace u^2 by $1 - v^2$), and the result is that the quantity under the square root is revealed to be 4 times a square, not because of a coincidence but because $|w^2 - 1|$ for w on the unit circle is a completely determined function of the chord from w^2 to 1, which classical geometry already tells you is twice the sine of the half-angle – i.e. $|w^2 - 1| = 2|\sin\theta|$ if $w = e^{i\theta}$, and $\sin\theta = \mathrm{Im}(w) = v$ exactly. The algebra is reproducing a geometric fact: the chord between two points on the unit circle separated by angle 2θ has length $2|\sin\theta|$, and this length is manifestly a single real quantity, not the square root of an arbitrary one. Seeing this ahead of time (or discovering it through the u, v computation) is what tells you $4v^2$, not some unrelated-looking rational, must appear.

Finally, why is $\mathrm{Im}(z^n) \in \mathbb{Q}$ itself not in doubt? Because $\mathbb{Q}[i]$ (or more simply, the real and imaginary parts of a product of numbers with rational real/imaginary parts) is closed under multiplication: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$ keeps you inside pairs of rationals at every step, so any finite power of $x + iy$ with $x, y \in \mathbb{Q}$ has rational real and imaginary parts by an immediate induction – no trigonometric identity or angle-addition formula is needed, only that $\mathbb{Q}$ is closed under $+, -, \times$.

Layer $+\alpha$ BACKTESTING

1. The cue is a modulus of a complex expression built from a number known to be rational-and-on-the-unit-circle: $|z^{2n} - 1|$ with $z = x + iy$, $x, y \in \mathbb{Q}$, $|z| = 1$. On seeing it, write down $|w|^2 = w\bar{w}$ to convert the modulus into a real expression, and separately note that $z^{2n} = (z^n)^2$ is a perfect square, inviting the substitution $w = z^n$.

2. That cue suggests reducing the target to $2 - 2\,\mathrm{Re}(\text{something})$ via conjugation, but the deeper method it should suggest is working with the halved exponent $w = z^n$ from the start, so the unit-circle relation $u^2 + v^2 = 1$ can be applied to expose a perfect square rather than a generic rational.

3. Hesitation occurs right after reducing to $\sqrt{2 - 2\,\mathrm{Re}(z^{2n})}$ and confirming $\mathrm{Re}(z^{2n}) \in \mathbb{Q}$, when the target still has an unresolved square root; it is reinforced away by asking, "have I actually shown the quantity under the root is a perfect square of a rational, or only that it is rational?"

4. The trap is mistaking "rational" for "perfect square of a rational" – the natural attempt's computation of $\mathrm{Re}(z^{2n}) \in \mathbb{Q}$ is correct and necessary but is silently the wrong target, since $\sqrt{\cdot}$ does not preserve rationality.

5. Whenever a problem's modulus involves z^{2n} or any even power, immediately substitute the halved exponent $w = z^n$ and re-derive the modulus formula in terms of $\mathrm{Re}(w), \mathrm{Im}(w)$ directly, rather than expanding the even power first and hoping the perfect-square structure survives being buried in a single real number.

6. Reusable rule: to prove $|f(z)|$ is rational when z has rational real/imaginary parts and lies on the unit circle, never settle for showing the value under a square root is rational – rewrite the modulus, using $|z| = 1$, as an expression that is manifestly $(\text{rational})^2$ before taking the root; the unit-circle identity $u^2 + v^2 = 1$ is exactly the tool for converting a real-part expression into an imaginary-part-squared expression (or vice versa) to make this happen.

▷ **Cue:** a modulus $|w^2 - 1|$ is wanted for w on the unit circle $\longrightarrow$ **Move:** compute it directly as $2|\operatorname{Im}(w)|$ via $u^2 + v^2 = 1$, rather than via $\sqrt{2 - 2\operatorname{Re}(w^2)}$ left unresolved

▷ **Cue:** an even power z^{2n} appears inside a modulus, with z on the unit circle $\longrightarrow$ **Move:** substitute $w = z^n$ first and redo the modulus computation in terms of w, since even powers are perfect squares and the unit-circle relation converts $\operatorname{Re}(w^2)$ into $\operatorname{Im}(w)^2$

▷ **Cue:** a real or imaginary part of z^n is needed for $z = x + iy$ with x, y rational $\longrightarrow$ **Move:** note $\mathbb{Q}[i]$-type closure: products of numbers with rational real/imaginary parts again have rational real/imaginary parts, so any integer power of z does too, by induction – no trigonometric identity required

1973 B3

1973 B3. Consider an integer $p > 1$ with the property that the polynomial $x^2 - x + p$ takes prime values for all integers x in the range $0 \le x < p$. (Examples: $p = 5$ and $p = 41$ have this property.) Show that there is exactly one triple of integers a, b, c satisfying the conditions:

$$b^2 - 4ac = 1 - 4p, \qquad 0 < a \le c, \qquad -a \le b < a.$$

Fields: number theory (quadratic forms, discriminants) **Position:** B3 **Answer:** The unique triple is $(a, b, c) = (1, -1, p)$.

Caveat. The source scan of the official solution's final bounding inequality is garbled at the line establishing $x < p$ (it prints as "1 b 1 $\le a \le J < 4p - -0/3$"). The argument below reconstructs that bound from scratch and verifies it in full; it agrees with the officially stated conclusion but the specific chain of inequalities is this archive's own reconstruction, not a transcription.

Layer 1 THE WHAT — surface analysis

The natural attempt. Existence first, since it is cheap: try the smallest plausible triple, $a = 1$. Then $b^2 - 4c = 1 - 4p$, and with $b = -1$ this gives $1 - 4c = 1 - 4p$, so $c = p$. Check the constraints: $0 < a = 1 \le c = p$ since $p > 1$, and $-a = -1 \le b = -1 < a = 1$. So $(1, -1, p)$ works, and existence is done in two lines.

For uniqueness, the natural instinct is to treat $b^2 - 4ac = 1 - 4p$ as a Diophantine constraint and bound a, b, c using nothing but the ordering $0 < a \le c$ and $-a \le b < a$, hoping size alone pins the triple down. Suppose (a, b, c) is any solution. Since $b^2 \equiv 1 \pmod 4$ (as $b^2 - 4ac = 1 - 4p \equiv 1 \pmod 4$), b must be odd. Since $|b| \le a \le c$ and $ac = \dfrac{b^2 + 4p - 1}{4}$, you can chase inequalities: $ac \ge a^2 \ge b^2$, so $\dfrac{b^2 + 4p - 1}{4} \ge b^2$, giving $4p - 1 \ge 3b^2$, i.e. $b^2 \le \dfrac{4p - 1}{3}$. This bounds $|b|$, hence bounds a and, through $ac = \dfrac{b^2 + 4p - 1}{4}$, bounds c too – a genuine finite search space. You could now, in principle, enumerate all integer pairs (a, c) with ac equal to a specific number and $a \le c$, and check which give a valid b. But this misses something: the bound produced this way shows there are only finitely many candidate triples, not that there is exactly one. For a general integer $N = ac$, there can be several factorisations $a \le c$ with $ac = N$ – nothing so far has ruled out, say, $a = 2, c = N/2$ alongside $a = 1, c = N$. The size bound alone caps the search; it does not close it.

The model chain. The missing ingredient is the one piece of the hypothesis not yet used: the polynomial $x^2 - x + p$ is prime-valued for every integer x with $0 \le x < p$. Write $|b| = 2x - 1$

for a positive integer x (legitimate since b is odd, so $|b| \geq 1$ is odd). Then $b^2 = (2x-1)^2 = 4x^2 - 4x + 1$, and $b^2 - 4ac = 1 - 4p$ rearranges to

$$4ac = 4x^2 - 4x + 1 + 4p - 1 = 4x^2 - 4x + 4p \quad \Longrightarrow \quad ac = x^2 - x + p.$$

This is exactly the polynomial from the hypothesis, evaluated at x. Provided $0 \leq x < p$, the hypothesis says ac is *prime*. A prime has exactly two positive divisor pairs, $1 \times$ (itself) and (itself) $\times 1$; since $0 < a \leq c$, the pair with $a \leq c$ forces $a = 1$ and $c = ac = x^2 - x + p$.

With $a = 1$, the constraint $-a \leq b < a$ becomes $-1 \leq b < 1$; since b is odd, $b = 0$ is impossible, so $b = -1$. Then $|b| = 1$ gives $x = (|b| + 1)/2 = 1$, and $c = x^2 - x + p = 1 - 1 + p = p$. This recovers exactly the triple found by inspection: $(a, b, c) = (1, -1, p)$, and shows it is forced, not merely one option among several.

It remains to justify the hypothesis's precondition $0 \leq x < p$ – the bound the natural attempt was reaching for, but needs now to close rigorously rather than merely to shrink the search. From $|b| \leq a \leq c$ and $ac = x^2 - x + p$: $ac \geq a^2 \geq b^2 = (2x-1)^2 = 4x^2 - 4x + 1$, so

$$x^2 - x + p \geq 4x^2 - 4x + 1 \quad \Longrightarrow \quad p \geq 3x^2 - 3x + 1.$$

Now check this forces $x < p$. For $x = 1$: the inequality reads $p \geq 1$, true since $p > 1$, and $x = 1 < p$ holds since $p > 1$. For $x \geq 2$: $3x^2 - 3x + 1 = (x+1) + 3x^2 - 4x \geq (x+1) + (12 - 8) = x + 5 > x + 1$ when $x \geq 2$ (checking $3x^2 - 4x \geq 4$ for $x \geq 2$: at $x = 2$, $12 - 8 = 4$, and the quadratic $3x^2 - 4x$ is increasing for $x \geq 2$), so $p \geq 3x^2 - 3x + 1 > x + 1 > x$. Either way $x < p$, so $0 \leq x < p$ holds and the hypothesis legitimately applies to declare ac prime. This closes the gap the natural attempt's size bound left open: the primality of ac, not sheer smallness, is what rules out every factorisation except $a = 1$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains reach the identity $ac = x^2 - x + p$ (after substituting $|b| = 2x - 1$) and both derive a genuine size bound on x. They split on what that bound is *for*: the natural attempt uses it to shrink the search space of triples, planning to check candidates by hand; the model chain uses it only to certify that x lies in the range $[0, p)$ where the problem's own hypothesis – $x^2 - x + p$ prime – can be legally invoked, and then lets primality itself force $a = 1$ outright.

> *I've built the equation $ac = x^2 - x + p$, and this is not an arbitrary Diophantine equation – it is literally the polynomial the problem told me is prime-valued on $0 \leq x < p$. I haven't used that hypothesis at all yet; the moment I do, $a \leq c$ and ac prime forces $a = 1$ immediately, with no search required.*

> *The bound I'm deriving on x isn't for narrowing a search – it's a precondition I have to verify before I'm allowed to call ac prime at all, since the hypothesis only promises primality for x strictly less than p.*

Layer 3 THE WHY

Why does bounding a, b, c by size alone fail to finish the uniqueness proof, even though it correctly shrinks the space of candidates to a finite one? Because "finite" is not "one": a composite number N typically has several factorisations $a \leq c$ with $ac = N$, and nothing about the size bound $b^2 \leq (4p - 1)/3$ distinguishes a correct factorisation from an incorrect one – it only says how big the numbers involved can be. Finiteness is necessary for a proof by exhaustive check to even be feasible, but it supplies no mechanism for eliminating the wrong candidates; only some additional structural fact can do that.

That structural fact is hiding in plain sight: the problem does not merely give you a discriminant equation, it hands you, as a hypothesis about p, a guarantee that a specific quadratic is prime-valued on a specific range. The substitution $|b| = 2x - 1$ is not an arbitrary parametrisation – it is forced by "$b^2 \equiv 1 \pmod 4$" (only odd squares are $\equiv 1$ mod 4), and it is chosen precisely so that $b^2 - 4ac = 1 - 4p$ collapses into $ac = x^2 - x + p$, matching the hypothesis's polynomial exactly. Once you see that the discriminant equation, correctly parametrised, reproduces the given polynomial verbatim, using the hypothesis stops being optional: it is the only fact strong enough to pin a two-factor product down to a unique ordered pair, because primality (unlike mere smallness) has exactly one factorisation with $a \leq c$ and $a > 0$.

Why must $0 \leq x < p$ hold before this can be invoked? Because the hypothesis is a conditional statement – "for x in this range, the value is prime" – and applying a conditional outside its stated domain is not a shortcut, it is an error; a value of $x^2 - x + p$ for $x \geq p$ need not be prime at all (indeed for $p = 5$, $x^2 - x + p$ at $x = 5$ gives 25, not prime). So the size bound the natural attempt reached for is not decorative – it is the logical hinge that licenses the single strongest step in the whole proof.

There is a deeper structural reason this whole configuration exists: the triple (a, b, c) with $b^2 - 4ac = 1 - 4p$ satisfying $0 < a \leq c$, $-a \leq b < a$ is exactly a *reduced* representative of a binary quadratic form $at^2 + bt + c$ of discriminant $1 - 4p$, in the classical Gauss reduction theory of quadratic forms.

Caveat. The fact that reduced forms of a given discriminant biject with equivalence classes of forms, and that "class number 1" (a unique reduced form) is precisely the condition under which $x^2 - x + p$ is prime for all $0 \leq x < p$ (these are Euler's "lucky numbers", connected to the discriminants $1 - 4p \in \{-3, -7, -11, -19, -43, -67, -163\}$), is classical quadratic-form / class-number theory well beyond A-level. The elementary argument above is self-contained and does not depend on this theory, but a reader who wants the full picture of *why* such p are so rare should know this is the reason.

Layer $+\alpha$ BACKTESTING

1. The cue is $b^2 - 4ac = 1 - 4p$ together with an explicit hypothesis that $x^2 - x + p$ is prime on $0 \leq x < p$. On seeing "prove uniqueness of a triple satisfying a discriminant equation, given a primality hypothesis about a specific related polynomial," write down: the discriminant equation must be re-parametrised until it visibly reproduces that exact polynomial.

2. That cue suggests substituting $|b| = 2x - 1$ (forced by b odd, from $b^2 \equiv 1 \pmod 4$) to turn $b^2 - 4ac = 1 - 4p$ into $ac = x^2 - x + p$, then directly invoking the given primality hypothesis rather than pursuing size bounds as an end in themselves.

3. Hesitation occurs right after deriving the bound $b^2 \leq (4p - 1)/3$ and realising it only shows finitely many candidates remain; it is reinforced away by asking, "have I used the problem's actual hypothesis about $x^2 - x + p$ being prime, or only the shape of the discriminant equation?"

4. The trap is treating the given hypothesis (primality of $x^2 - x + p$ on $0 \leq x < p$) as background flavour text justifying the existence of such p, rather than as the operative tool that collapses "ac equals this number" into "ac is prime, hence $a = 1$."

5. Whenever a problem states a special property of a specific object (a prime-valued polynomial, a distinguished sequence) and then asks you to prove uniqueness of some other structure, actively look for an algebraic substitution that makes the other structure's defining equation reproduce that specific object verbatim.

6. Reusable rule: a size bound derived from an ordering hypothesis (like $0 < a \leq c$) is often not the finishing move but the precondition that licenses invoking a separately given hypothesis (like a primality guarantee on a bounded range); check what conditional statement in the problem your bound unlocks before concluding the bound itself proves uniqueness.

- ▷ **Cue:** a discriminant equation $b^2 - 4ac = 1 - 4p$ is given alongside $b \equiv 1 \pmod 2$ forced by $b^2 \equiv 1 \pmod 4$ $\longrightarrow$ **Move:** substitute $|b| = 2x - 1$ to convert the equation into $ac = x^2 - x + p$
- ▷ **Cue:** a problem's hypothesis promises a specific polynomial is prime-valued on a bounded range, and a derived equation matches that polynomial exactly $\longrightarrow$ **Move:** invoke the primality hypothesis directly: with $a \le c$ both positive, ac prime forces $a = 1$
- ▷ **Cue:** a size bound (from an ordering like $0 < a \le c$) only shows finitely many candidates remain, not uniqueness $\longrightarrow$ **Move:** check what conditional hypothesis in the problem that bound is a precondition for, and use it to eliminate all but one candidate structurally rather than by exhaustive search

1973 B4

1973 B4. (a) On $[0,1]$, let f have a continuous derivative satisfying $0 < f'(x) \le 1$. Also suppose that $f(0) = 0$. Prove that

$$\left[\int_0^1 f(x)\,dx\right]^2 \ge \int_0^1 [f(x)]^3\,dx.$$

[Hint: Replace the inequality by one involving the inverse function to f.]
(b) Show an example in which equality occurs.

Caveat. The MAA OCR drops the displayed inequality in the statement of (a). It is restored here from the book's own solution, whose theorem reads $\left[\int_0^1 f\right]^2 \ge \int_0^1 f^3$ under the same hypotheses, and it is the standard wording of 1973 B4. The book's theorem is stated with $0 \le f' \le 1$; the problem says $0 < f' \le 1$. The proof below works under either; only the equality case changes (with $0 \le f'$ the zero function also gives equality).

Fields: real analysis; integral inequalities; monotonicity via the derivative **Position:** B4 — mid-paper, one idea deep but the idea is not the one the eye reaches for **Answer:** Put the upper limit t in place of 1; the difference $F(t) = \left(\int_0^t f\right)^2 - \int_0^t f^3$ has $F'(t) = f(t)\,G(t)$ with $G(t) = 2\int_0^t f - f(t)^2$ and $G' = 2f(1 - f') \ge 0$, so $F(1) \ge 0$. Equality holds for $f(x) = x$, and (under $f' > 0$) only for it.

Layer 1 THE WHAT — surface analysis

The natural attempt. You see an integral squared on one side and an integral of a power on the other, and the reflex is Cauchy–Schwarz or Jensen. Try it. Cauchy–Schwarz on $[0,1]$ gives

$$\left(\int_0^1 f\right)^2 \le \int_0^1 f^2,$$

which points the wrong way: it bounds the left-hand side *above*. Jensen with the convex function $u \mapsto u^3$ gives $\int f^3 \ge (\int f)^3$, again a lower bound on the right-hand side, the wrong direction. Hölder only reshuffles the same inequalities.

The student then tries $f(x) = cx$ to see what is going on: the left side is $c^2/4$, the right is $c^3/4$, and the inequality is $c^2 \ge c^3$, true exactly because $c \le 1$. So the hypothesis $f' \le 1$ is doing all the work, and no averaging inequality ever uses a derivative. That is the wall: the general inequalities you know are blind to the one hypothesis that makes this true. Try $f = 2x$ and it is false. Any proof that does not use $f' \le 1$ must be wrong.

The model chain. Write $I(t) = \int_0^t f(x)\,dx$ for $t \in [0,1]$. Since $f(0) = 0$ and $f' > 0$, f is strictly increasing, so $f(t) > 0$ for $t > 0$.

Step 1: make the upper limit variable. Define

$$F(t) = I(t)^2 - \int_0^t f(x)^3\,dx.$$

Then $F(0) = 0$ and, by the fundamental theorem of calculus (legitimate since f is continuous),

$$F'(t) = 2I(t)f(t) - f(t)^3 = f(t)\left[2I(t) - f(t)^2\right] = f(t)\,G(t), \qquad G(t) := 2I(t) - f(t)^2.$$

Step 2: the bracket has a sign, by the same trick. $G(0) = 0 - 0 = 0$, and since f' is continuous,

$$G'(t) = 2f(t) - 2f(t)f'(t) = 2f(t)\big[1 - f'(t)\big] \;\geq 0,$$

because $f(t) \geq 0$ and $f'(t) \leq 1$. So G is non-decreasing from $G(0) = 0$, hence $G \geq 0$ on $[0,1]$.

Step 3: conclude. $F'(t) = f(t)G(t) \geq 0$, so F is non-decreasing from $F(0) = 0$, and $F(1) \geq 0$. That is (a).

Step 4: equality. If $F(1) = 0$ then, since F is non-decreasing with $F(0) = F(1) = 0$, $F \equiv 0$ and $F' \equiv 0$. For $t > 0$, $f(t) > 0$, so $G(t) = 0$ on $(0,1]$, hence $G' \equiv 0$ there, so $2f(t)(1 - f'(t)) = 0$ and $f'(t) = 1$ on $(0,1]$; by continuity of f', on $[0,1]$. With $f(0) = 0$, $f(x) = x$. Conversely $f(x) = x$ gives $\left(\frac{1}{2}\right)^2 = \frac{1}{4} = \int_0^1 x^3\,dx$. So (b): $f(x) = x$, and it is the only example.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains start from the same observation: the inequality is about two numbers, $\left(\int_0^1 f\right)^2$ and $\int_0^1 f^3$. The natural attempt treats them as fixed numbers and reaches for an inequality between numbers. The model replaces the 1 by t and turns two numbers into two functions that start equal (both are 0 at $t = 0$). At that line the paths split: comparing numbers needs a global inequality, and none uses f'; comparing functions that start together only needs a comparison of their derivatives, and differentiation is exactly where f' enters.

> *Both sides are integrals from* 0*, and both vanish if the upper limit is* 0*. So I let the upper limit be* t *and compare the rates at which the two sides grow. If the left grows at least as fast for every* t*, I am done.*

> *The derivative is* $f(t)$ *times* $2\int_0^t f - f(t)^2$*. That bracket is zero at* $t = 0$ *too. Same move again: differentiate it. Now* f' *appears, and* $f' \leq 1$ *is exactly what makes it non-negative.*

- ▷ **Cue:** an inequality between two integrals over $[0,a]$, both sides vanishing when $a = 0$ $\longrightarrow$ **Move:** replace a by a variable t, differentiate the difference, and prove the derivative has a sign
- ▷ **Cue:** a hypothesis on f' that a standard averaging inequality (Cauchy–Schwarz, Jensen) cannot see $\longrightarrow$ **Move:** the proof must differentiate something; find the quantity whose derivative contains f'
- ▷ **Cue:** test family $f = cx$ turns the inequality into $c^2 \geq c^3$ $\longrightarrow$ **Move:** the inequality is sharp at the extreme allowed by the derivative bound; expect equality at $f' \equiv 1$ and a monotonicity proof

Layer 3 THE WHY

Why is "make the limit variable" the natural thought rather than a trick? Go to what the problem is. The right-hand side at $f(x) = x$ is $\int_0^1 x^3\,dx$ and the left is $\left(\int_0^1 x\,dx\right)^2$. Their equality is the continuous shadow of Nicomachus' identity

$$1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2.$$

How is that identity proved? By induction on n: add the next term to both sides and check that the increments match, $S_n^2 - S_{n-1}^2 = n\,(2S_{n-1} + n) = n^3$, using $2S_{n-1} = n(n-1)$. Induction on the upper limit of a sum *is* differentiation with respect to the upper limit of an integral. So the problem is "prove Nicomachus as an inequality for sequences whose steps are at most 1", and the method for Nicomachus is the method here. Once you see that the equality case is a sum-of-cubes identity, varying the upper limit is not a trick; it is the only proof of that identity anyone learns.

Why does the bracket $G(t) = 2\int_0^t f - f(t)^2$ appear, and why is it non-negative? In the discrete version, the step is $S_n^2 - S_{n-1}^2 = a_n(2S_{n-1} + a_n)$ against a_n^3, so you need $2S_{n-1} + a_n \geq a_n^2$: twice the area under the staircase is at least the square of its height. In the continuous version, $2\int_0^t f \geq f(t)^2$ says the same thing, and it is geometric. Write $y = f(x)$. Since $f' \leq 1$, f rises no faster than the line $y = x$, so its graph up to height $f(t)$ encloses at least as much area as the triangle with legs $f(t)$ and $f(t)$, which is $\frac{1}{2}f(t)^2$. Differentiating G is the calculus way of saying "slope at most 1 means area at least half-square".

Where does the hint come from? The book's second solution uses the inverse $g = f^{-1}$ on $[0, c]$, $c = f(1)$. Substituting $x = g(y)$ gives $\int_0^1 f = \int_0^c y\,g'(y)\,dy$, so

$$\left(\int_0^1 f\right)^2 = 2\int_0^c\!\!\int_0^z y\,g'(y)\,z\,g'(z)\,dy\,dz \;\geq\; 2\int_0^c z\,g'(z)\left[\int_0^z y\,dy\right]dz = \int_0^c z^3 g'(z)\,dz = \int_0^1 f^3,$$

using $g' \geq 1$ (the inverse of a function with slope at most 1 has slope at least 1) and the symmetry of the integrand across $y = z$ to write the square as twice the triangle. This is the same content: the inner integral $\int_0^z y\,g'(y)\,dy$ is the area $\int_0^{g(z)} f$, and the bound $\geq z^2/2$ is exactly $G \geq 0$. The hint is optional. The variable-limit proof is shorter and uses nothing beyond A-level calculus.

Why must $f' \leq 1$ enter only through a derivative? Because the inequality is not invariant under scaling. Replacing f by cf multiplies the left by c^2 and the right by c^3. Any inequality that is homogeneous (Cauchy–Schwarz, Jensen, Hölder) cannot distinguish $c = 1$ from $c = 2$, and at $c = 2$ the claim is false. Scaling-sensitivity is the signal that a size bound is essential, and here the only size bound is on f'.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that both sides are integrals from 0 and $f(0) = 0$. On seeing it, write $F(t) = \left(\int_0^t f\right)^2 - \int_0^t f^3$ and $F(0) = 0$.

2. Monotonicity: show $F' \geq 0$. When F' factors as $f \cdot G$ with $G(0) = 0$, apply the same method to G.

3. Hesitation comes after the first derivative: $F' = f(2\int_0^t f - f^2)$ looks no simpler than the original. Reinforce it away by checking the bracket at $t = 0$; it vanishes, which is the signal to differentiate again, and the second derivative brings in f'.

4. The trap is Cauchy–Schwarz and Jensen, which point the wrong way and cannot use the hypothesis $f' \leq 1$. The test $f = 2x$ shows that any proof ignoring the derivative bound proves something false.

5. Try $f = cx$ first. If the inequality becomes $c^k \geq c^m$ with $k < m$, a bound on the size of f or f' is essential, and the proof will be a derivative argument with equality at the extreme.

6. Reusable decision rule: *to prove an inequality between two quantities that agree at a starting point, turn the fixed endpoint into a variable and compare derivatives; if the derivative comparison is itself a new inequality that holds with equality at the start, repeat. The continuous version of "induction on n" is "differentiate in the upper limit".*

1973 B5

1973 B5. (a) Let z be a solution of the quadratic equation

$$az^2 + bz + c = 0$$

and let n be a positive integer. Show that z can be expressed as a rational function of z^n, a, b, c.

(b) Using (a) or by any other means, express x as a rational function of x^3 and $x + (1/x)$. (Display your answer explicitly in a clearly visible form.)

[By a rational function of several variables, we mean a quotient of polynomials in those variables, the polynomials having rational numbers as coefficients, and the denominator being not identically zero. Thus to obtain x as a rational function of $u = x^2$ and $v = x + (1/x)$, we could write $x = (u+1)/v$.]

Fields: algebra; polynomial remainders; linear recurrences **Position:** B5 — late in the paper but the lightest of the last three once the reduction is seen **Answer:** (a) Reducing z^n modulo the quadratic gives $z^n = p_n z + q_n$ with p_n, q_n polynomials in $-b/a, -c/a$ and $p_n \not\equiv 0$, so $z = (z^n - q_n)/p_n$. (b) With $v = x + 1/x$: $x = \dfrac{x^3 + v}{v^2 - 1} = \dfrac{x^3 + x + 1/x}{(x + 1/x)^2 - 1}$.

Layer 1 THE WHAT — surface analysis

The natural attempt. For (a) the first instinct is the quadratic formula: $z = \left(-b \pm \sqrt{b^2 - 4ac}\right)/(2a)$. That is not rational in a, b, c; the square root has to be removed, and the only thing available to remove it is z^n. The student raises the formula to the nth power, gets a binomial expansion full of $\sqrt{b^2 - 4ac}$ terms, and tries to solve for the root. For $n = 2$ it works after some effort; for general n it is a mess.

For (b) the natural attempt is to play with $v = x + 1/x$ directly. Cube it:

$$v^3 = x^3 + 3x + \frac{3}{x} + \frac{1}{x^3} = x^3 + \frac{1}{x^3} + 3v, \qquad \text{so} \qquad x^3 + \frac{1}{x^3} = v^3 - 3v.$$

Now $1/x^3$ is known in terms of x^3 and v. But every such identity is symmetric under $x \leftrightarrow 1/x$, and x is not. Powers of v alone never isolate x. That is the wall, and it is instructive: from v you can only get what x and $1/x$ share.

The model chain. *(a)* Assume $a \neq 0$ (otherwise the equation is not a quadratic). Put $r = -b/a$ and $s = -c/a$, so that $z^2 = rz + s$. Define polynomials $p_k, q_k \in \mathbb{Q}[r, s]$ by

$$p_0 = 0,\ q_0 = 1, \qquad p_1 = 1,\ q_1 = 0, \qquad p_k = rp_{k-1} + sp_{k-2}, \quad q_k = rq_{k-1} + sq_{k-2}\ (k \geq 2).$$

Claim: $z^k = p_k z + q_k$ for all $k \geq 0$. True for $k = 0, 1$. If it holds for $k-1$ and $k-2$, then multiplying $z^2 = rz + s$ by z^{k-2} gives $z^k = rz^{k-1} + sz^{k-2} = (rp_{k-1} + sp_{k-2})z + (rq_{k-1} + sq_{k-2})$.

Claim: $p_n \not\equiv 0$ as a polynomial for $n \geq 1$. By induction every p_k with $k \geq 1$ has non-negative integer coefficients and $p_k(1, 0) = 1$ (at $r = 1, s = 0$ the recursion gives $p_k = p_{k-1}$). So p_n is not the zero polynomial.

Hence $z = \dfrac{z^n - q_n(r, s)}{p_n(r, s)}$. Substitute $r = -b/a$, $s = -c/a$ and multiply top and bottom by a^{n-1} (each p_n, q_n has degree at most $n - 1$): the result is $F(z^n, a, b, c)/G(a, b, c)$ with F, G integer polynomials. G is not identically zero, since at $a = 1$ it equals $p_n(-b, -c)$, a nonzero polynomial. This is the required rational function, valid whenever $G(a, b, c) \neq 0$.

(b) Let $v = x + 1/x$. Then $x^2 - vx + 1 = 0$: this is (a) with $a = 1, b = -v, c = 1$, so $r = v$, $s = -1$. The recursion gives $p_2 = v$, $q_2 = -1$, then

$$p_3 = vp_2 - p_1 = v^2 - 1, \qquad q_3 = vq_2 - q_1 = -v,$$

i.e. $x^3 = (v^2-1)x - v$. Directly: $x^3 = x \cdot x^2 = x(vx-1) = v(vx-1) - x = (v^2-1)x - v$. Therefore

$$\boxed{x = \frac{x^3+v}{v^2-1} = \frac{x^3 + \left(x+\frac{1}{x}\right)}{\left(x+\frac{1}{x}\right)^2 - 1}}$$

Check at $x = 2$: $v = \frac{5}{2}$, $(8+\frac{5}{2})/(\frac{25}{4}-1) = \frac{21}{2}\cdot\frac{4}{21} = 2$. As an identity: $(v^2-1)x - v = x^3 + 2x + \frac{1}{x} - x - x - \frac{1}{x} = x^3$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains start from $az^2 + bz + c = 0$. The natural attempt *solves* it (quadratic formula) and then fights the square root. The model *uses* it as a rewriting rule, $z^2 \to rz + s$, and never solves anything. At that line the paths split. Solving introduces an irrationality that must be cancelled later; rewriting keeps everything inside polynomials in z with coefficients in a, b, c, and every such polynomial collapses to a linear one.

> *I do not want the roots; I want a relation. The equation says z^2 is a linear expression in z. Then so is z^3, and z^4, and z^n. So $z^n = pz + q$ with p, q built from a, b, c, and a linear equation in z I can solve by division.*

> *For (b): x^3 is the only piece of data that is not symmetric in x and $1/x$. So x must come out as a linear function of x^3. Write x^3 as (something) x + (something) using $x^2 = vx - 1$.*

- ▷ **Cue:** "express z in terms of z^n" where z satisfies a polynomial equation $\longrightarrow$ **Move:** reduce z^n modulo the polynomial (use the equation as a rewriting rule) to get $z^n =$ lower-degree expression in z, then solve
- ▷ **Cue:** a quadratic relation $z^2 = rz + s$ and a question about z^n $\longrightarrow$ **Move:** set up $z^n = p_n z + q_n$ with the two-term recurrence $p_n = rp_{n-1} + sp_{n-2}$
- ▷ **Cue:** given only $x + 1/x$ and one other quantity, recover x $\longrightarrow$ **Move:** $x + 1/x$ is symmetric under $x \leftrightarrow 1/x$; the other quantity must carry the asymmetry, and x will be linear in it over the symmetric functions

Layer 3 THE WHY

Why is "rewrite, don't solve" the natural move? Go to the definition of what a root is. The equation $az^2 + bz + c = 0$ is the only thing you know about z, and it says one thing: z^2 can be traded for $rz + s$. Every higher power can therefore be traded down, one degree at a time, until only 1 and z remain. That is polynomial division: $z^n = Q(z)(z^2 - rz - s) + (p_n z + q_n)$, and the first term vanishes at a root. The recurrence is division by $z^2 - rz - s$ done one step at a time. The quadratic formula, by contrast, answers a question the problem did not ask. The problem asks for a rational expression, and the formula is the one non-rational thing in sight.

Why must the answer be *linear* in z^n? Think of the field $\mathbb{Q}(a, b, c)$ of all rational functions of the coefficients, and the bigger field obtained by adjoining z. Since z satisfies a quadratic over the smaller field, every element of the bigger field is $\alpha z + \beta$ with α, β rational in a, b, c. In particular z^n is. So z^n and z are related by a linear equation, and solving a linear equation needs one division. You do not need field theory to use this; you need to know that "degree 2" means "everything is $\alpha z + \beta$".

What exactly is p_n, and when does division fail? If the roots are $z_1 \neq z_2$, then $z_i^n = p_n z_i + q_n$ for both roots (the recurrence does not know which root it is applied to), and subtracting gives

$$p_n = \frac{z_1^n - z_2^n}{z_1 - z_2}.$$

So $p_n = 0$ exactly when $z_1^n = z_2^n$, which is exactly when z^n cannot tell the two roots apart. Example: $z^2 + 1 = 0$, $n = 2$: $p_2 = r = 0$, and indeed $i^2 = (-i)^2$, so no formula could

recover z from z^2. This is why the problem's definition insists only that the denominator is *not identically* zero: the formula is a polynomial identity, valid off the set where the question has no answer. The official solution proves $p_n \not\equiv 0$ by noting its coefficients are positive; the evaluation $p_n(1,0) = 1$ used above is a quicker witness of the same fact.

Why does x^3 have to be the asymmetric input in (b)? The map $x \mapsto 1/x$ fixes v. Any rational function of v alone is fixed too, so it cannot equal x (which is not fixed unless $x = \pm 1$). The problem supplies x^3 precisely to break the symmetry, and the example in the problem's own bracket ($x = (u+1)/v$ with $u = x^2$) is the $n = 2$ case of the same method: $x^2 = vx - 1$ gives $x = (x^2+1)/v$. The bracket is a worked instance of part (a), handed to you. Read it as the method, not as notation.

The recurrence $p_n = rp_{n-1} + sp_{n-2}$ is the same one that produces Fibonacci numbers ($r = s = 1$) and Chebyshev polynomials ($r = 2\cos\theta$, $s = -1$, where $p_n = \sin n\theta / \sin\theta$). Here $r = v$, $s = -1$, and $p_n(v)$ is the Chebyshev polynomial of the second kind in $v/2$: $p_3 = v^2 - 1$ is $U_2(v/2)$. Every problem that says "$x + 1/x$" is a Chebyshev problem in disguise, with $x = e^{i\theta}$ and $v = 2\cos\theta$.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is "z is a solution of $az^2 + bz + c = 0$" combined with "z^n". On seeing it, write $z^2 = rz + s$ and then $z^n = p_n z + q_n$.

2. Reduction modulo the quadratic, organised as a two-term linear recurrence for the coefficients.

3. Hesitation is over whether p_n could vanish, which would wreck the division. Reinforce it away with $p_n = (z_1^n - z_2^n)/(z_1 - z_2)$: it vanishes only when z^n genuinely fails to determine z, and as a polynomial it is nonzero because $p_n(1,0) = 1$.

4. The trap is the quadratic formula, which brings in a square root that the answer must not contain. In (b) the trap is manipulating v alone, which can only ever produce functions symmetric in x and $1/x$.

5. Whenever a quantity satisfies a polynomial equation of degree d, every polynomial in it is a combination of $1, z, \ldots, z^{d-1}$. When asked to express z through some other function of z, reduce that function to this form and solve the resulting low-degree equation.

6. Reusable decision rule: *do not solve an algebraic equation when you only need to compute with its root; use it as a rewriting rule that lowers degree. For a quadratic, $z^n = p_n z + q_n$ with p_n the Lucas-type sequence $(z_1^n - z_2^n)/(z_1 - z_2)$.*

1973 B6

1973 B6. On the domain $0 \le \theta \le 2\pi$:

(a) Prove that $\sin^2\theta \cdot \sin(2\theta)$ takes its maximum at $\pi/3$ and $4\pi/3$ (and hence its minimum at $2\pi/3$ and $5\pi/3$).

(b) Show that

$$\left|\sin^2\theta \left[\sin^3(2\theta) \cdot \sin^3(4\theta) \cdots \sin^3(2^{n-1}\theta)\right] \sin(2^n\theta)\right|$$

takes its maximum at $\theta = \pi/3$. (The maximum may also be attained at other points.)

(c) Derive the inequality:

$$\sin^2\theta \cdot \sin^2(2\theta) \cdot \sin^2(4\theta) \cdots \sin^2(2^n\theta) \le (3/4)^n.$$

Fields: trigonometry; optimisation; telescoping products **Position:** B6 — last problem of the paper; the difficulty is seeing why the exponents in (b) are $2, 3, \ldots, 3, 1$ **Answer:** With $h(x) = \sin^2 x \sin 2x$, $|h| \le 3\sqrt{3}/8$ with equality exactly when $x \equiv \pm\pi/3 \pmod{\pi}$. The expression in (b) is $\prod_{k=0}^{n-1} |h(2^k\theta)| \le (3\sqrt{3}/8)^n$, attained at $\theta = \pi/3$ because doubling maps $\pm\pi/3$ to $\mp\pi/3$ mod π. Raising to the $2/3$ power and using $|\sin| \le 1$ on the two end factors gives (c).

Layer 1 THE WHAT — surface analysis

The natural attempt. The natural attempt starts at (c), the inequality you actually want, and tries to bound the product directly. Bounding each factor by 1 gives nothing. Pairing neighbours is the next idea: with $s = \sin^2 x$,

$$\sin^2 x \, \sin^2 2x = 4 \sin^4 x \cos^2 x = 4s^2(1-s) \le 4 \cdot \tfrac{4}{27} = \tfrac{16}{27},$$

maximised at $s = \frac{2}{3}$. Splitting the $n+1$ factors of (c) into disjoint consecutive pairs gives roughly $(16/27)^{(n+1)/2} = (0.7698\ldots)^{n+1}$ per pair of factors. The target is $(3/4)^n$, about 0.75 per factor. The pairing loses, and it loses for a structural reason: the pair $(x, 2x)$ wants $\sin^2 x = \frac{2}{3}$, but then the next pair $(4x, 8x)$ is evaluated at an angle the first pair did not control, and there is no single θ at which all pairs are simultaneously at their maxima. The student has a bound that is true, too weak, and no idea how to tighten it.

For (a) the student usually differentiates $\sin^2\theta \sin 2\theta$ with the product rule and gets it right; that part is not where the difficulty lies.

The model chain. Let $h(x) = \sin^2 x \, \sin 2x = 2\sin^3 x \cos x$.

(a) $h'(x) = 2(3\sin^2 x \cos^2 x - \sin^4 x) = 2\sin^2 x\,(3\cos^2 x - \sin^2 x) = 2\sin^2 x\,(4\cos^2 x - 1)$. On $[0, 2\pi]$, $h' = 0$ when $\sin x = 0$ ($x = 0, \pi, 2\pi$, where $h = 0$) or $\cos x = \pm\frac{1}{2}$ ($x = \pi/3, 2\pi/3, 4\pi/3, 5\pi/3$). The values are

$$h(\tfrac{\pi}{3}) = h(\tfrac{4\pi}{3}) = \tfrac{3}{4} \cdot \tfrac{\sqrt{3}}{2} = \tfrac{3\sqrt{3}}{8}, \qquad h(\tfrac{2\pi}{3}) = h(\tfrac{5\pi}{3}) = -\tfrac{3\sqrt{3}}{8}.$$

A continuous function on a closed interval attains its extrema at critical points or endpoints, so the maximum $3\sqrt{3}/8$ is at $\pi/3, 4\pi/3$ and the minimum $-3\sqrt{3}/8$ at $2\pi/3, 5\pi/3$. Since h has period π and is odd, this holds for all real x:

$$|h(x)| \le \tfrac{3\sqrt{3}}{8}, \qquad \text{with equality iff } x \equiv \pm\tfrac{\pi}{3} \pmod{\pi}. \tag{$*$}$$

(b) Call the expression $E_n(\theta)$. Consider

$$\prod_{k=0}^{n-1} h(2^k\theta) = \prod_{k=0}^{n-1} \sin^2(2^k\theta)\, \sin(2^{k+1}\theta).$$

Count the exponent of $\sin(2^j\theta)$: for $j = 0$ it is 2 (from $k = 0$); for $1 \le j \le n-1$ it is 2 (from $k = j$) plus 1 (from $k = j-1$), which is 3; for $j = n$ it is 1 (from $k = n-1$). So $E_n(\theta) = \prod_{k=0}^{n-1} |h(2^k\theta)|$, and by $(*)$,

$$E_n(\theta) \le \left(\tfrac{3\sqrt{3}}{8}\right)^n.$$

At $\theta = \pi/3$: $2^k\pi/3 \equiv \pi/3$ or $2\pi/3 \pmod{\pi}$ for every k (doubling swaps these two classes: $2 \cdot \frac{\pi}{3} = \frac{2\pi}{3}$ and $2 \cdot \frac{2\pi}{3} = \frac{4\pi}{3} \equiv \frac{\pi}{3}$). Both classes are $\equiv \pm\pi/3$, so every factor equals $3\sqrt{3}/8$ and $E_n(\pi/3) = (3\sqrt{3}/8)^n$. Hence the maximum of E_n is attained at $\pi/3$.

(c) From (b), $E_n(\theta)^{2/3} \le \left((3\sqrt{3}/8)^2\right)^{n/3} = (27/64)^{n/3} = (3/4)^n$. Now

$$E_n^{2/3} = |\sin\theta|^{4/3} \prod_{k=1}^{n-1} \sin^2(2^k\theta)\, |\sin 2^n\theta|^{2/3},$$

and the product P in (c) is

$$P = \sin^2\theta \prod_{k=1}^{n-1} \sin^2(2^k\theta)\, \sin^2(2^n\theta) = E_n^{2/3} \cdot |\sin\theta|^{2/3}\, |\sin 2^n\theta|^{4/3} \le E_n^{2/3} \le (3/4)^n,$$

since $|\sin| \le 1$. (For $n = 1$ the middle product is empty and the same computation holds.)

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains want to bound a product of sines at doubling angles by bounding small blocks. The natural attempt chooses blocks that are *disjoint* (pairs $x, 2x$) with equal weights. The model chooses blocks that *overlap* (each factor $h(2^k\theta)$ shares $\sin(2^k\theta)$ with the previous block) and weights them $2:1$. At that line the paths split. Disjoint blocks can each be maximised but not at the same θ, so the bound is not attained and is too weak. Overlapping blocks of the right shape are all maximised at one θ, so the product bound is sharp.

> *Part (b) has exponents* $2, 3, 3, \ldots, 3, 1$. *That is* $2+1$ *in the middle: each middle sine is squared by its own block and appears once more at the end of the previous block. So (b) is a product of copies of* $\sin^2 x \sin 2x$ *at* $x = \theta, 2\theta, 4\theta, \ldots$, *which is (a) applied n times.*

> *A product of maxima is attained only if one* θ *maximises every factor. Where does doubling send* $\pi/3$*? To* $2\pi/3$*, which is* $-\pi/3$ *mod* π*, and* $|h|$ *is the same there. The orbit never leaves the maximum set. That is why* $\pi/3$.

> *Part (c) has every exponent equal to* 2*; (b) has average exponent close to* 3. *Raise (b) to the power* $2/3$ *and the middle exponents become* 2*; the ends are off, and* $|\sin| \leq 1$ *fixes them in the right direction.*

- ▷ **Cue:** a product $\prod f(2^k\theta)$ of trigonometric factors at doubling angles $\longrightarrow$ **Move:** look for a two-term block $g(x)\,g'(2x)$ whose maximiser is carried to another maximiser by $x \mapsto 2x$; the product then telescopes into copies of the block
- ▷ **Cue:** given exponents in a multi-part problem that look arbitrary $(2, 3, \ldots, 3, 1)$ $\longrightarrow$ **Move:** read them as a sum of overlapping blocks; the earlier part of the problem tells you which block
- ▷ **Cue:** a bound for a product of factors with the wrong exponents $\longrightarrow$ **Move:** raise to a fractional power to match the middle exponents; handle the end factors with $|\sin| \leq 1$

Layer 3 THE WHY

Why the block $\sin^2 x \sin 2x$ and not something else? Work out which blocks can possibly be maximised all along a doubling orbit. Take a general block $|\sin x|^a\,|\sin 2x|^b$ and take logs: the derivative of $a \ln|\sin x| + b \ln|\sin 2x|$ is

$$a\cot x + 2b\cot 2x = a\cot x + b(\cot x - \tan x) = (a+b)\cot x - b\tan x,$$

so the maximum is at $\tan^2 x = (a+b)/b$. For the doubling orbit to stay at maxima, the maximiser x_0 must satisfy $2x_0 \equiv \pm x_0 \pmod{\pi}$ (because $|\sin|^a|\sin 2\cdot|^b$ has period π and is even, the maximising set mod π is $\{\pm x_0\}$). That forces $3x_0 \equiv 0$ or $x_0 \equiv 0 \pmod{\pi}$, so $x_0 = \pi/3$, $\tan^2 x_0 = 3$, and $a = 2b$. The ratio $2:1$ is not a choice; it is the only ratio for which the argument works, and $\pi/3$ is the only angle. The problem's exponents are forced by the dynamics of $x \mapsto 2x$ on the circle. This is the WHY of the whole problem, and the official solution does not state it.

Why does $\pi/3$ keep coming back? Because it is a periodic point of doubling. Modulo π the map $x \mapsto 2x$ sends $\pi/3 \to 2\pi/3 \to \pi/3$, a 2-cycle, and modulo sign ($|h|$ is even) it is a fixed point: $2\cdot\frac{\pi}{3} \equiv -\frac{\pi}{3}$. The expression in (b) is maximised by making every factor maximal, and the only way to do that at all n simultaneously is to sit on an orbit that doubling cannot move off the maximum set. The equation $2x \equiv -x \pmod{\pi}$ is the whole reason.

Why does (c) come out as $(3/4)^n$? At $\theta = \pi/3$ every $\sin^2(2^k\theta) = \frac{3}{4}$, so the product in (c) is $(3/4)^{n+1}$. The bound $(3/4)^n$ is therefore sharp to within one factor of $\frac{3}{4}$, the price paid for the two end factors. The pairing attempt could not reach 0.75 per factor because a disjoint pair is

maximised at $\sin^2 x = \frac{2}{3}$, an angle that is not a periodic point of doubling, so the bound $16/27$ per pair is never attained along an orbit. The number $3/4$ is $\sin^2(\pi/3)$; it is the value on the invariant orbit, and that is why it is the constant.

The official solution says "(a) Simple calculus" and proves (b) by induction, noting that the ratio of consecutive expressions is $|\sin^2 2^n\theta \, \sin 2^{n+1}\theta| = |h(2^n\theta)|$. That ratio *is* the block. The induction and the product decomposition above are the same argument; the product form makes visible that every factor is maximised at once, which is the only reason the induction closes.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the pattern of exponents $2, 3, \ldots, 3, 1$ in (b). On seeing it, write the expression as $\prod_{k=0}^{n-1} |\sin^2(2^k\theta)\sin(2^{k+1}\theta)|$.

2. Bound each block by (a), then check that one θ makes every block maximal.

3. Hesitation is at "the product of maxima need not be attained". Reinforce it away by computing the orbit of $\pi/3$ under doubling mod π: it is $\{\pi/3, 2\pi/3\}$, and $|h|$ is maximal on both.

4. The trap is disjoint pairing in (c), which yields a true but weaker constant (≈ 0.77 per factor) because the pair maximisers do not lie on a single doubling orbit. The other trap is doing (c) from scratch instead of using (b); the problem is a chain, and each part is the tool for the next.

5. For any product over $x, 2x, 4x, \ldots$ (or any iterated map), ask which point is fixed or periodic under the map, and build blocks maximised there. Weights are forced by requiring the block's maximiser to be that point.

6. Reusable decision rule: *to bound a product along an orbit of a map, split it into overlapping blocks whose common maximiser is a periodic point of the map; the bound is then attained and therefore sharp. Unequal exponents in the statement are the fingerprint of such overlapping blocks.*