DAWN PUTNAM SESSION

Putnam 1988 — Papers A and B

Putnam 120 Archive

Contents

Putnam 1988 — Paper A **5**
1988 A1 5
1988 A2 7
1988 A3 9
1988 A4 12
1988 A5 15
1988 A6 17

Putnam 1988 — Paper B **21**
1988 B1 21
1988 B2 23
1988 B3 25
1988 B4 27
1988 B5 30
1988 B6 32

Putnam 1988 — Paper A

1988 A1

1988 A1. Let R be the region consisting of the points (x, y) of the cartesian plane satisfying both $|x| - |y| \leq 1$ and $|y| \leq 1$. Sketch the region R and find its area.

Fields: coordinate geometry, inequalities with absolute values **Position:** A1 — the opener (the highest-scoring problem in the 1985–2000 volume) **Answer:** The area is 6. R is the non-convex hexagon with vertices $(2,1), (-2,1), (-1,0), (-2,-1), (2,-1), (1,0)$: the rectangle $[-2,2] \times [-1,1]$ with two triangular notches of area 1 each removed.

Layer 1 THE WHAT — surface analysis

The natural attempt. You see two absolute values in the first inequality and one in the second, and you do what A-level trained you to do with $|\cdot|$: split into cases. $x \geq 0, y \geq 0$ gives $x - y \leq 1$; $x \geq 0, y < 0$ gives $x + y \leq 1$; $x < 0, y \geq 0$ gives $-x - y \leq 1$; $x < 0, y < 0$ gives $-x + y \leq 1$. Each of these is a half-plane, and you now have to intersect each one with its own quadrant and with the strip $-1 \leq y \leq 1$, then draw all four pieces and glue them together.

This works. It is also where most of the damage happens. Four half-planes, four quadrant restrictions, and a strip: twelve boundary conditions to keep straight, with sign errors waiting in the second and fourth cases. A common slip is to draw $|x| - |y| \leq 1$ as a bounded diamond — the picture of $|x| + |y| \leq 1$ is the one burned into memory — and then report area 2, or to shade the region on the wrong side of the line $x + y = 1$ in the fourth quadrant and get a shape that is not symmetric at all. Even done correctly, the case-by-case sketch gives you four trapezoids whose areas you then add, with no check that the pieces match along the axes.

The model chain. Both defining inequalities are unchanged when $x \mapsto -x$ and when $y \mapsto -y$, because each of x and y appears only through $|x|$ and $|y|$. So R is symmetric in both coordinate axes, and its area is four times the area of $R \cap \{x \geq 0,\ y \geq 0\}$. (The pieces overlap only along the axes, which have zero area.)

In the closed first quadrant $|x| = x$ and $|y| = y$, so the region is

$$T = \{(x, y) :\ x \geq 0,\ 0 \leq y \leq 1,\ x \leq 1 + y\}.$$

For each fixed $y \in [0, 1]$, the horizontal cross-section of T is the segment $0 \leq x \leq 1 + y$, of length $1 + y$. Hence

$$\text{area}(T) = \int_0^1 (1 + y)\, dy = \frac{3}{2}.$$

(Equivalently: T is the trapezoid with vertices $(0,0), (1,0), (2,1), (0,1)$, parallel sides 1 and 2, height 1.) Therefore $\text{area}(R) = 4 \cdot \frac{3}{2} = 6$.

For the sketch, reflect T in both axes. The boundary of R runs along $y = 1$ from $(-2,1)$ to $(2,1)$, down the segment $x = 1 + y$ from $(2,1)$ to $(1,0)$, down $x = 1 - y$ from $(1,0)$ to $(2,-1)$, along $y = -1$ to $(-2,-1)$, and back symmetrically through $(-1,0)$. It is a rectangle $[-2,2] \times [-1,1]$ (area 8) with two triangular bites taken out of its left and right sides, each triangle with vertices $(\pm 1, 0), (\pm 2, 1), (\pm 2, -1)$ and area $\frac{1}{2} \cdot 2 \cdot 1 = 1$. Check: $8 - 2 = 6$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains read the same two inequalities. They split at the first line written. The natural attempt opens with "Case 1: $x \geq 0, y \geq 0 \ldots$" and commits to four cases. The model opens with "Both inequalities depend only on $|x|$ and $|y|$, so R is symmetric in both axes" — and then does exactly one of the four cases.

> *Every variable here sits inside an absolute value. That means flipping the sign of x or of y changes nothing, so the region is symmetric in both axes. I will work in the first quadrant, where the bars disappear, and multiply by four.*

> *I should not trust my picture of $|x| - |y| \leq 1$; the minus sign makes it unbounded. Let me find the first-quadrant piece from the inequality itself: $x \leq 1 + y$ and $0 \leq y \leq 1$ — slice horizontally, length $1 + y$.*

Layer 3 THE WHY

Why is symmetry the natural first move, rather than a shortcut you are expected to have been shown? Go back to what $|x|$ is: the distance from x to 0 on the line. It is by definition blind to sign: $|x| = |-x|$. So a condition written only in terms of $|x|$ and $|y|$ is a condition on the *distances* of a point from the two axes, and the four points $(\pm x, \pm y)$ have the same pair of distances. They must all be in R or all out. Symmetry is not a feature you notice about R after drawing it; it is written into the notation before you draw anything. The phrasing of the problem — every variable wrapped in bars — is the cue.

Once symmetry is on the page, the case analysis is not merely shorter, it is *unnecessary*: the other three cases are the first case reflected, and reflection preserves area. The case-split student is doing the same computation four times with different signs, which is the definition of zombie maths — manipulation undertaken before asking whether the problem has already told you the answer to three quarters of it.

The second decision — slice horizontally rather than vertically — is also forced. In T the variable y has fixed bounds $0 \leq y \leq 1$ and x has a bound depending on y, namely $x \leq 1 + y$. Integrating in the direction of the dependent bound gives one integral; slicing vertically gives two, because the lower boundary of T changes character at $x = 1$ (for $0 \leq x \leq 1$ a vertical slice runs from $y = 0$ to $y = 1$; for $1 \leq x \leq 2$ it runs from $y = x - 1$ to $y = 1$). The inequality $x \leq 1 + y$ is literally a statement about horizontal cross-sections, so read it that way.

The last point is the trap. The shape $|x| + |y| \leq 1$ is a diamond every student has drawn. $|x| - |y| \leq 1$ looks almost the same and is completely different: it says $|x| \leq 1 + |y|$, which grows without bound as $|y|$ grows. On its own it is an unbounded region between the lines $x = \pm(1 + |y|)$; it is the strip $|y| \leq 1$ that cuts it to finite area. Knowing that a sign change can turn a bounded region into an unbounded one is the reason to derive the piece from the inequality, not from memory.

Finally, why the two-way check? The rectangle-minus-triangles computation uses a different decomposition from the trapezoid computation, so agreement at 6 is genuine evidence and costs ten seconds. On an A1, where the danger is a careless slip rather than a missing idea, an independent check is the cheapest insurance available.

Layer $+\alpha$ BACKTESTING

1. The first cue is that x and y appear only as $|x|$ and $|y|$. Write down: "R is symmetric in both axes; area $= 4\times$ first-quadrant area."

2. Reduce to one quadrant, remove the absolute values there, and compute the area by slicing in the direction whose bounds are constant ($y \in [0, 1]$, $0 \leq x \leq 1 + y$).

3. Hesitation occurs when drawing $|x| - |y| \leq 1$, because it resembles the familiar diamond. Reinforce it away by rewriting the inequality as $|x| \leq 1 + |y|$ and reading off that the allowed x-range widens as $|y|$ grows.
4. The trap is importing the picture of $|x| + |y| \leq 1$ (area 2). A secondary trap is a sign error in one of four hand-done cases, producing a non-symmetric region — which symmetry would have exposed immediately.
5. Any condition in which a variable appears only inside $|\cdot|$, or only squared, or only through an even function, hands you a reflection symmetry; use it before doing any casework.
6. Reusable rule: before splitting into cases, check whether the cases are images of one another under a symmetry of the problem. If they are, do one case and transport the answer; then verify the total by an independent decomposition.

▷ **Cue:** every occurrence of a variable is inside an absolute value (or squared, or in an even function) $\longrightarrow$ **Move:** declare the reflection symmetry, reduce to one quadrant or half-line, and multiply at the end
▷ **Cue:** an inequality of the form $x \leq h(y)$ with y in a fixed interval $\longrightarrow$ **Move:** compute the area by horizontal slices: $\int (\text{length of cross-section})\, dy$
▷ **Cue:** an absolute-value region with a minus sign, such as $|x| - |y| \leq c$ $\longrightarrow$ **Move:** rewrite as $|x| \leq c + |y|$ before drawing; do not reuse the diamond picture of $|x| + |y| \leq c$
▷ **Cue:** an area found by one decomposition $\longrightarrow$ **Move:** check it with a second decomposition (bounding rectangle minus pieces)

1988 A2

1988 A2. A not uncommon calculus mistake is to believe that the product rule for derivatives says that $(fg)' = f'g'$. If $f(x) = e^{x^2}$, determine, with proof, whether there exists an open interval (a, b) and a nonzero function g defined on (a, b) such that this wrong product rule is true for x in (a, b).

Fields: differential equations, calculus **Position:** A2 — early, routine once set up **Answer:** Yes. For example $g(x) = e^x\sqrt{2x-1}$ on $(1, 2)$ works. More completely: on an interval not containing $\frac{1}{2}$ the nonzero solutions are exactly $g(x) = Ce^x\sqrt{|2x-1|}$, $C \neq 0$; on an interval containing $\frac{1}{2}$ the only solution is $g \equiv 0$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You read "$(fg)' = f'g'$" and start guessing. g constant? Then $(fg)' = cf'$ but $f'g' = 0$, so no. $g = e^x$? Then $(fg)' = (2x+1)e^{x^2+x}$ and $f'g' = 2xe^{x^2+x}$, so no. $g = e^{x^2}$? $(fg)' = 4xe^{2x^2}$, $f'g' = 4x^2e^{2x^2}$; equal only at $x = 0, 1$. g a polynomial? You try $g = x$ and get $(1 + 2x^2)e^{x^2}$ against $2xe^{x^2}$.

Each guess fails, and each failure is a comparison of two expressions at every x — you are testing candidate functions one at a time against an identity. Some students take the string of failures as evidence and answer "no", trying to prove impossibility from the growth of e^{x^2}. That route has no proof in it, because the answer is yes. Others notice the equation involves g and g' and start manipulating $(fg)' = f'g'$ with f still abstract, getting lost in the general case before using the specific f.

The model chain. Expand the true product rule. For differentiable g on (a, b), the condition $(fg)' = f'g'$ reads

$$f'g + fg' = f'g'.$$

With $f = e^{x^2}$ we have $f' = 2xe^{x^2}$; divide by $e^{x^2} > 0$:

$$2xg + g' = 2xg' \quad \Longleftrightarrow \quad (2x-1)\, g'(x) = 2x\, g(x). \tag{$*$}$$

This is a first-order linear differential equation for g.

Solving it on an interval avoiding $\frac{1}{2}$. There $2x-1 \neq 0$ and $(*)$ says $g'/g = 2x/(2x-1) = 1 + \frac{1}{2x-1}$ wherever $g \neq 0$, suggesting $\ln|g| = x + \frac{1}{2}\ln|2x-1| + c$. To prove it without dividing by g, set $h(x) = g(x)e^{-x}|2x-1|^{-1/2}$. On an interval where $2x-1$ has constant sign $\sigma = \pm 1$, $|2x-1| = \sigma(2x-1)$ and

$$h' = e^{-x}|2x-1|^{-1/2}\Big(g' - g - \frac{g}{2x-1}\Big) = \frac{e^{-x}|2x-1|^{-1/2}}{2x-1}\big((2x-1)g' - 2xg\big) = 0.$$

So h is constant, and $g(x) = Ce^x\sqrt{|2x-1|}$. Conversely every such g satisfies $(*)$: for $x > \frac{1}{2}$, $g = Ce^x(2x-1)^{1/2}$ gives

$$g' = Ce^x\Big((2x-1)^{1/2} + (2x-1)^{-1/2}\Big) = \frac{2x\,Ce^x}{(2x-1)^{1/2}}, \qquad (2x-1)g' = 2x\,g,$$

and the case $x < \frac{1}{2}$ is the same computation with $(1-2x)$. Taking $C = 1$ and $(a,b) = (1,2)$ gives a nonzero g. As a direct check on the original statement: $fg = e^{x^2+x}\sqrt{2x-1}$, so

$$(fg)' = e^{x^2+x}\frac{(2x+1)(2x-1)+1}{\sqrt{2x-1}} = \frac{4x^2e^{x^2+x}}{\sqrt{2x-1}} = 2xe^{x^2}\cdot\frac{2xe^x}{\sqrt{2x-1}} = f'g'.$$

So the answer is **yes**.

Why $\frac{1}{2}$ *must be avoided (not required, but it completes the picture).* If $\frac{1}{2} \in (a,b)$, then $(*)$ at $x = \frac{1}{2}$ gives $g(\frac{1}{2}) = 0$. On each side of $\frac{1}{2}$ the argument above gives $g = C_\pm e^x\sqrt{|2x-1|}$. Differentiability at $\frac{1}{2}$ requires $\lim_{x\to 1/2^+}\frac{g(x)}{x-1/2} = \lim C_+e^x\sqrt{2}\,(x-\frac{1}{2})^{-1/2}$ to be finite, forcing $C_+ = 0$; likewise $C_- = 0$. So $g \equiv 0$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains start from $(fg)' = f'g'$. The natural attempt keeps the left side packed and substitutes candidate g's. The model unpacks $(fg)'$ with the real product rule, $f'g + fg' = f'g'$, and at that line sees an equation linear in g and g' — an ODE to be solved, not an identity to be tested.

> *The question is "does some g satisfy this relation between g and g′?" That is a differential equation in g. I shouldn't guess solutions; I should write the equation out with the correct product rule and solve it.*

> *After dividing by e^{x^2} I have $(2x-1)g' = 2xg$ — first-order, linear, separable. The only danger is the point where the coefficient of g' vanishes, $x = \frac{1}{2}$, so I pick an interval that avoids it.*

Layer 3 THE WHY

Why is "solve the ODE" natural, rather than a leap? Look at what the problem asks: the existence of a *function* g satisfying a condition that holds *at every point* of an interval and involves g and g'. That is the definition of a differential equation. The phrasing "determine whether there exists … a function g … such that" is the existence question for an ODE, stated in words. Once you recognise the object, the method is the one you already own: separate variables, or use an integrating factor.

The candidate-guessing student has the right instinct (produce an example) aimed at the wrong target. Guessing works when the space of plausible answers is small. Here the answer is $e^x\sqrt{2x-1}$, which nobody guesses; it has to be *derived*, and it can only be derived by treating

the condition as an equation for g. The general principle: if you find yourself testing candidates for a relation between a function and its derivative, stop testing and solve.

Why write the true product rule first? Because $(fg)'$ is the only term in the problem you can compute in general, and until you expand it you cannot see which terms involve g'. Expanded, the equation is $f'g + fg' = f'g'$: the unknown appears only to the first power, with g' on both sides. Collecting g' gives $g'(f - f') = -f'g$, and the coefficient $f - f' = (1 - 2x)e^{x^2}$ is the whole story. Where it is nonzero, the equation can be written $g' = \frac{f'}{f'-f}\, g$, and the existence–uniqueness theorem for linear first-order ODEs (Kedlaya's solution cites it, and stops there) gives a nonzero solution through any nonzero initial value. Where it vanishes, $x = \frac{1}{2}$, the equation degenerates. Kedlaya's solution says "avoid $x = \frac{1}{2}$"; here is where that comes from: it is the zero of the coefficient of the highest derivative, which is exactly where linear ODE theory stops guaranteeing anything.

I prefer to exhibit g explicitly rather than cite the existence theorem. An explicit function with a verification line is a proof any reader can check in thirty seconds, and it needs nothing beyond A-level calculus. The integrating-factor form $h = ge^{-x}|2x - 1|^{-1/2}$ also avoids dividing by g, which the separation-of-variables heuristic does silently; that division is legitimate only where $g \neq 0$, and a rigorous proof should not depend on it.

Finally, the problem's specific f matters in one place only: $f'/f = 2x$ is a polynomial, so $\frac{f'}{f'-f} = \frac{2x}{2x-1}$ is a rational function with an elementary antiderivative. For any f with $f' \neq f$, the same method gives $g = C \exp \int \frac{f'}{f'-f}$; the setter chose e^{x^2} so that the integral closes.

Layer $+\alpha$ BACKTESTING

1. The first cue is "does there exist a function g such that [relation involving g and g'] holds on an interval". Write down the relation with the correct product rule: $f'g + fg' = f'g'$.

2. A linear first-order ODE: collect the g' terms, $(2x - 1)g' = 2xg$, and solve by separation or integrating factor.

3. Hesitation comes at the division by $2x - 1$ and by g. Reinforce it away by choosing an interval that avoids $x = \frac{1}{2}$ and by using the integrating-factor form $h = ge^{-x}|2x - 1|^{-1/2}$, whose derivative is zero without any division by g.

4. The trap is inference from failed guesses: after $g = c, e^x, e^{x^2}, x$ all fail, answering "no". A secondary trap is stating "yes" by quoting the existence theorem without noting the singular point $x = \frac{1}{2}$.

5. Whenever an existence question asks for a function satisfying a pointwise relation between it and its derivatives, it is an ODE; identify its order and linearity before anything else.

6. Reusable rule: to decide whether a function exists satisfying a derivative relation, solve the relation as a differential equation, locate the points where the leading coefficient vanishes, and exhibit an explicit solution on an interval avoiding them; verify it by direct substitution.

▷ **Cue:** "does there exist a function g with [relation in g, g'] on an interval" $\longrightarrow$ **Move:** rewrite as a differential equation and solve; do not test candidate functions

▷ **Cue:** a condition stated with $(fg)'$, $(f/g)'$ or $(f \circ g)'$ $\longrightarrow$ **Move:** expand with the correct rule first so the unknown's derivative is isolated

▷ **Cue:** a linear first-order ODE $p(x)g' = q(x)g$ $\longrightarrow$ **Move:** avoid zeros of p; integrating factor $h = g \exp(-\int q/p)$ has $h' = 0$

▷ **Cue:** separation of variables divides by the unknown $\longrightarrow$ **Move:** replace with an integrating factor so the proof does not assume $g \neq 0$

1988 A3

1988 A3. Determine, with proof, the set of real numbers x for which

$$\sum_{n=1}^{\infty} \left(\frac{1}{n} \csc \frac{1}{n} - 1 \right)^x$$

converges.

Fields: real analysis, series, Taylor expansion **Position:** A3 — mid-paper, routine for anyone who expands **Answer:** The series converges if and only if $x > \frac{1}{2}$. The base behaves like $\frac{1}{6n^2}$, so the terms behave like $6^{-x} n^{-2x}$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You see a series with a parameter and reach for the tests in order. Ratio test: with $a_n = \frac{1}{n} \csc \frac{1}{n} - 1$, you need $\lim (a_{n+1}/a_n)^x$. You can't compute a_{n+1}/a_n without knowing how a_n behaves, and when you do find out, the limit is 1 — the ratio test is inconclusive, as it always is for series comparable to $\sum n^{-p}$. Root test: $a_n^{x/n} \to 1$, inconclusive again.

So you try to understand a_n directly. You know $\csc \frac{1}{n} \to \infty$ and $\frac{1}{n} \to 0$, so $\frac{1}{n} \csc \frac{1}{n}$ is a $0 \cdot \infty$ form; you recall $\frac{\sin t}{t} \to 1$ and conclude $a_n \to 0$. Now you are stuck: "$a_n \to 0$" tells you the terms are small when $x > 0$ and large when $x < 0$, so $x \le 0$ diverges — but for $x > 0$ you need to know *how fast* $a_n \to 0$, and the standard limit $\frac{\sin t}{t} \to 1$ carries no rate. Some students stop here and answer "$x > 0$", which is wrong: at $x = \frac{1}{2}$ the terms are about $\frac{1}{\sqrt{6}\, n}$ and the series diverges like the harmonic series.

The model chain. Write $t = \frac{1}{n} \in (0, 1]$ and $a_n = \frac{t}{\sin t} - 1 = \frac{t - \sin t}{\sin t}$.

Positivity. For $0 < t < \pi$, $0 < \sin t < t$, so $a_n > 0$ and a_n^x is defined for every real x.

Two-sided bound. For $t > 0$ the alternating Taylor bounds give

$$t - \frac{t^3}{6} \;\le\; \sin t \;\le\; t - \frac{t^3}{6} + \frac{t^5}{120}.$$

(Both follow by integrating $\cos s \le 1$ repeatedly: $\cos s \le 1 \Rightarrow \sin t \le t \Rightarrow \cos t \ge 1 - \frac{t^2}{2} \Rightarrow \sin t \ge t - \frac{t^3}{6} \Rightarrow \cos t \le 1 - \frac{t^2}{2} + \frac{t^4}{24} \Rightarrow \sin t \le t - \frac{t^3}{6} + \frac{t^5}{120}$.) Hence for $0 < t \le 1$,

$$\frac{t^3}{6} - \frac{t^5}{120} \;\le\; t - \sin t \;\le\; \frac{t^3}{6}, \qquad t - \frac{t^3}{6} \;\le\; \sin t \;\le\; t.$$

Dividing, with $t \le 1$:

$$a_n \;\ge\; \frac{t^3/6 - t^5/120}{t} = \frac{t^2}{6}\left(1 - \frac{t^2}{20}\right) \;\ge\; \frac{19}{120}\, t^2, \qquad a_n \;\le\; \frac{t^3/6}{t - t^3/6} = \frac{t^2}{6 - t^2} \;\le\; \frac{t^2}{5}.$$

So for every $n \ge 1$,

$$\frac{19}{120} \cdot \frac{1}{n^2} \;\le\; a_n \;\le\; \frac{1}{5} \cdot \frac{1}{n^2}. \tag{†}$$

Also the bounds show $n^2 a_n \to \frac{1}{6}$, since $t^2/(6 - t^2)$ and $\frac{t^2}{6}(1 - t^2/20)$ both equal $\frac{t^2}{6}(1 + O(t^2))$.

Conclusion. Raise (†) to the power x. If $x \ge 0$, a_n^x lies between $(19/120)^x n^{-2x}$ and $(1/5)^x n^{-2x}$; if $x < 0$ the inequalities reverse but a_n^x is still squeezed between two positive constant multiples of n^{-2x}. By the comparison test, $\sum a_n^x$ converges if and only if $\sum n^{-2x}$ converges, that is, if and only if $2x > 1$. The answer is $x > \frac{1}{2}$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains reach "$a_n \to 0$". The natural attempt got there from the limit $\frac{\sin t}{t} \to 1$, which has no rate, and stalls. The model asks for the rate: it expands $\sin t$ to the first term that does not cancel, obtains $a_n \approx \frac{1}{6n^2}$, and compares with a p-series.

> *Knowing $a_n \to 0$ is useless for convergence; I need the order of magnitude. $\frac{t}{\sin t} - 1$ is a difference that cancels at leading order, so I expand $\sin t$ until something survives: $\sin t = t - \frac{t^3}{6} + \cdots$, so $a_n \approx \frac{t^2}{6} = \frac{1}{6n^2}$.*

> *The terms are a constant times n^{-2x}, so this is a p-series in disguise with $p = 2x$. The ratio test will say nothing; compare directly.*

Layer 3 THE WHY

Why is the Taylor expansion the natural first move here, and not a technique you happen to have seen? Start from the question the problem is asking. For a series of positive terms, convergence depends only on how fast the terms go to zero — that is what the comparison test says. So the entire problem is: *what is the size of a_n as $n \to \infty$?* Any tool that does not report a size is irrelevant, and that rules out "$\frac{\sin t}{t} \to 1$" immediately. It tells you $a_n \to 0$ and nothing more.

Then why expand *this* far, to t^3? Because $a_n = \frac{t - \sin t}{\sin t}$ is a difference of two quantities that agree to first order. When you subtract things that agree to leading order, the size of the difference is set by the first term where they disagree. $\sin t$ and t agree at order t and disagree at order t^3, so $t - \sin t \approx \frac{t^3}{6}$, and dividing by $\sin t \approx t$ leaves $\frac{t^2}{6}$. The rule is: expand until the cancellation stops. Here that is one term past the obvious one.

Kedlaya's solution writes $a_n = \frac{1}{6n^2} + O(n^{-4})$ and applies the limit comparison test. Here is where that comes from, and why I have replaced the big-O with explicit constants. The limit comparison test is correct and standard, but an $O(\cdot)$ term hides an unspecified constant and an unspecified threshold "for n large". The alternating bounds on $\sin t$ come from integrating $\cos s \le 1$ up to five times — each step is an A-level integral — and they give bounds valid for every $n \ge 1$, not merely for large n. So (†) is a complete proof with nothing deferred.

Why does the problem have the specific shape $(\cdot)^x$? Because once the base is known to be of size n^{-2}, raising it to the power x gives size n^{-2x}, and the question collapses to the one family of series whose convergence you know exactly: $\sum n^{-p}$ converges iff $p > 1$. Every "for which x does $\sum(\text{something})^x$ converge" problem is secretly asking you to find the exponent q with something $\asymp n^{-q}$; the answer is then $x > 1/q$. The trigonometric dressing is there to make you compute q.

Notice also that the negative and zero values of x need no separate treatment in the model chain. A two-sided bound $c_1 n^{-2} \le a_n \le c_2 n^{-2}$ with $c_1, c_2 > 0$ survives raising to any real power (with the inequalities flipped when $x < 0$), so comparison handles all x at once. The natural attempt splits into "$x \le 0$: terms do not tend to zero" and "$x > 0$: …", which is correct but unnecessary.

Layer $+\alpha$ BACKTESTING

1. The first cue is a positive series with a parameter in the exponent. Write down: "find q with $a_n \asymp n^{-q}$; answer $x > 1/q$."
2. Taylor-expand $\sin t$ at $t = \frac{1}{n}$ until the cancellation in $\frac{t}{\sin t} - 1$ stops, then compare with $\sum n^{-2x}$.
3. Hesitation occurs after deducing $a_n \to 0$, when it is unclear how to proceed. Reinforce it away with the reminder that convergence of positive series is decided by the *rate* of decay, so any limit without a rate is a dead end.
4. The trap is answering "$x > 0$" from "$a_n \to 0$", forgetting that small terms need not have a convergent sum. The boundary case $x = \frac{1}{2}$, where the terms are about $\frac{1}{\sqrt{6}\,n}$, is the harmonic series in disguise.
5. Any expression of the form $F(1/n) - F(0)$ or $\frac{\text{thing}}{\text{thing's leading term}} - 1$ inside a series: expand F until the first nonzero term to get its exact order.

6. Reusable rule: to decide convergence of $\sum a_n^x$ (or $\sum a_n$ with a_n given by a smooth function of $1/n$), establish $c_1 n^{-q} \le a_n \le c_2 n^{-q}$ with explicit positive constants, then invoke the p-series criterion. Do not use the ratio or root test on anything comparable to a power of n.

▷ **Cue:** a series of terms built from a smooth function evaluated at $1/n$ $\longrightarrow$ **Move:** substitute $t = \frac{1}{n}$ and Taylor-expand in t to find the order of the terms

▷ **Cue:** a difference whose two parts agree at leading order, such as $\frac{t}{\sin t} - 1$ or $1 - \cos t$ $\longrightarrow$ **Move:** expand until the first term that does not cancel; that term gives the size

▷ **Cue:** a positive series with a real parameter as an exponent, $\sum a_n^x$ $\longrightarrow$ **Move:** find q with $a_n \asymp n^{-q}$; the series converges iff $qx > 1$

▷ **Cue:** terms that behave like a power of n $\longrightarrow$ **Move:** skip the ratio and root tests (they return 1); use comparison with a p-series

▷ **Cue:** a limit-comparison argument with a big-O remainder $\longrightarrow$ **Move:** replace with explicit two-sided bounds from alternating Taylor inequalities, valid for all n

1988 A4

1988 A4. (a) If every point of the plane is painted one of three colors, do there necessarily exist two points of the same color exactly one inch apart?

(b) What if "three" is replaced by "nine"?

Justify your answers.

Fields: combinatorial geometry, graph colouring **Position:** A4 — mid-paper; (a) needs one idea, (b) needs one construction with honest numbers **Answer:** (a) Yes. Two points at distance $\sqrt{3}$ are forced to share a colour (they are the far vertices of two unit equilateral triangles on a common side), and an isosceles triangle with sides $\sqrt{3}, \sqrt{3}, 1$ then forces two points at distance 1 to share a colour. (b) No. Tile the plane by half-open squares of side $\frac{2}{3}$ and colour them periodically with a 3×3 pattern of nine colours.

Layer 1 THE WHAT — surface analysis

The natural attempt. Part (a). You suppose a bad colouring exists — three colours, no two points at distance 1 the same colour — and start looking for a contradiction. The first figure everyone draws is a unit equilateral triangle: its three vertices are pairwise at distance 1, so they use all three colours. Good, but that is not a contradiction; it is a constraint. So you add more triangles: a hexagonal patch of the triangular lattice, a regular hexagon with its centre, and you try to three-colour it. And it *can* be three-coloured: the triangular lattice $\{a + b\omega\}$ coloured by $(a + b)$ mod 3 has no two adjacent lattice points the same colour. Every lattice-only picture you draw is colourable, so the search stalls. The student now tries rotating a triangle about a vertex, or adds random extra unit segments, and the diagram becomes a mess of arcs with no organising principle. Time goes.

Part (b). You want a nine-colouring with no monochromatic unit pair, so you reach for a grid, and the first grid is unit squares in a 3×3 colour pattern. It fails inside a single square: the diagonal of a unit square is $\sqrt{2} > 1$, so $(0.05, 0.05)$ and $(0.05 + \frac{1}{\sqrt{2}}, 0.05 + \frac{1}{\sqrt{2}}) \approx (0.757, 0.757)$ lie in one square at distance exactly 1. So you shrink the squares — to side 0.4, say — and now the other side bites: two squares of the same colour have only two squares between them, a gap of $2 \times 0.4 = 0.8 < 1$, and a point in one is at distance exactly 1 from some point in the other. Without writing the two inequalities down side by side, you oscillate between too big and too small.

The model chain. **(a) Yes.** Suppose, for contradiction, that the plane is coloured with three colours so that any two points at distance exactly 1 receive different colours.

Step 1: two points at distance $\sqrt{3}$ have the same colour. Let $|AB| = \sqrt{3}$. Place $A = (0,0)$, $B = (\sqrt{3},0)$ and put $P = (\frac{\sqrt{3}}{2}, \frac{1}{2})$, $Q = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$. Then

$$|AP| = |AQ| = |BP| = |BQ| = \sqrt{\tfrac{3}{4} + \tfrac{1}{4}} = 1, \qquad |PQ| = 1.$$

So APQ and BPQ are unit equilateral triangles. In each, the three vertices are pairwise at distance 1, so they carry three different colours; with only three colours available, A must carry the unique colour not used by P and Q, and so must B. Hence A and B have the same colour.

Step 2: the contradiction. Take any triangle CDE with $|CD| = |CE| = \sqrt{3}$ and $|DE| = 1$ (it exists since $1 < 2\sqrt{3}$; for instance $C = (0,0)$, $D, E = (\sqrt{11}/2, \pm\frac{1}{2})$, so $|CD|^2 = \frac{11}{4} + \frac{1}{4} = 3$). By Step 1, C and D share a colour and C and E share a colour, so D and E share a colour. But $|DE| = 1$. Contradiction.

(b) No. For a point (x,y) set

$$i = \left\lfloor \tfrac{3x}{2} \right\rfloor, \qquad j = \left\lfloor \tfrac{3y}{2} \right\rfloor,$$

so (x,y) lies in the half-open square $S_{ij} = [\frac{2}{3}i, \frac{2}{3}(i+1)) \times [\frac{2}{3}j, \frac{2}{3}(j+1))$. The squares S_{ij} partition the plane. Give (x,y) the colour (i mod 3, j mod 3), one of nine colours.

Let $P = (x,y)$ and $P' = (x',y')$ have the same colour, with indices (i,j) and (i',j').

- If $(i,j) = (i',j')$, both lie in one half-open square of side $\frac{2}{3}$, so $|x - x'| < \frac{2}{3}$ and $|y - y'| < \frac{2}{3}$, giving $|PP'| < \frac{2}{3}\sqrt{2} = \frac{\sqrt{8}}{3} < 1$ (since $8 < 9$).
- Otherwise $i \equiv i'$ and $j \equiv j' \pmod 3$ with $(i,j) \neq (i',j')$, so, say, $i' \geq i + 3$ (the other cases are the same with roles or coordinates exchanged). Then $x' \geq \frac{2}{3}(i+3)$ and $x < \frac{2}{3}(i+1)$, so $x' - x > \frac{4}{3}$ and $|PP'| > \frac{4}{3} > 1$.

Either way $|PP'| \neq 1$. So nine colours do not force a monochromatic unit pair.

Layer 2 THE HOW — the bifurcation

Bifurcating point. In (a), both chains start from "a unit equilateral triangle uses all three colours". The natural attempt treats that as a local constraint and tries to tile more triangles. The model asks what the constraint *implies about a pair of points that are not at distance* 1: two triangles sharing the side PQ force their apexes A, B to match. The moment you have a rule "distance $\sqrt{3} \Rightarrow$ same colour", you are hunting for a unit segment whose ends are both joined by $\sqrt{3}$ to a common point. In (b), the split is at whether the square size is chosen by the two inequalities or by trial.

> *Three colours and a triangle that uses all three means the colouring is rigid: once two vertices are fixed, the third is determined. So two triangles sharing an edge force their far vertices to agree. That gives me a new distance, $\sqrt{3}$, at which colours are* equal. *Equal-colour relations chain; different-colour relations do not. Now I need a unit segment whose ends are both at distance $\sqrt{3}$ from one point.*

> *For the nine-colouring I have two enemies: two points in the same square, and two points in different squares of the same colour. The first needs side$\cdot\sqrt{2} < 1$; the second needs the gap between same-coloured squares, which is twice the side, to exceed* 1. *So the side must lie in $(\frac{1}{2}, \frac{1}{\sqrt{2}})$. Pick $\frac{2}{3}$.*

Layer 3 THE WHY

Kedlaya's solution to (a) begins "if A and B are two points at distance $\sqrt{3}$…". Here is where $\sqrt{3}$ comes from.

The hypothesis you are refuting is a list of *inequalities* of colours: $c(X) \neq c(Y)$ whenever $|XY| = 1$. Inequalities do not compose: $c(X) \neq c(Y)$ and $c(Y) \neq c(Z)$ tells you nothing about X and Z. A contradiction has to be an equation clashing with an inequality, so you must *manufacture an equation*. With only three colours, the pigeonhole principle is the only source of equations: if a point must differ from two points that differ from each other, it is pinned to the third colour. That is exactly the unit equilateral triangle, and pinning happens to the apex over a fixed base. Two apexes over the same base — the rhombus of two equilateral triangles — are therefore forced equal, and the distance between them is the long diagonal of that rhombus, $2 \cdot \frac{\sqrt{3}}{2} = \sqrt{3}$. So $\sqrt{3}$ is not a guess; it is the only distance the three-colour pigeonhole can produce from one base.

Now equality *does* compose. "Same colour at distance $\sqrt{3}$" applied twice from a common point C gives same colour for any two points on the circle of radius $\sqrt{3}$ about C. That circle has chords of every length up to $2\sqrt{3}$, in particular 1. So the contradiction is automatic once the equation exists. (The seven-point graph formed this way, two rhombi hinged at C, is the Moser spindle.) The reason lattice pictures stall is also now visible: in the triangular lattice the $\sqrt{3}$-equality relation only connects points already consistent with the $(a + b)$ mod 3 colouring. You need to rotate one rhombus about C, off the lattice, and the circle argument tells you by how much: until the far vertices are 1 apart.

For (b), the design is dictated by the definition. A colour class is a union of regions; the only ways a colour class can contain a unit pair are *within* a region (so each region needs diameter < 1) or *across* two regions of that colour (so same-coloured regions need gaps > 1). A periodic $k \times k$ pattern of squares of side s puts same-coloured squares $(k-1)s$ apart. With $k = 3$: $s\sqrt{2} < 1$ and $2s > 1$. The interval $(\frac{1}{2}, \frac{1}{\sqrt{2}}) \approx (0.5, 0.707)$ is nonempty, so nine colours work; with $k = 2$ you would need $s > 1$ and $s < 0.707$, impossible — that is why the problem says nine rather than four. The half-open squares are not fussiness: they make the regions a genuine partition so each point has exactly one colour.

Layer $+\alpha$ BACKTESTING

1. The first cue is "three colors" together with the equilateral triangle, whose three vertices are pairwise at the forbidden distance. Write down: "each unit equilateral triangle is rainbow; so the apex is determined by the base".

2. It suggests building forced *equalities*: two triangles on a common base force their apexes (distance $\sqrt{3}$) to match; then chain equalities to reach a unit segment. For (b): a periodic square colouring, with side fixed by two inequalities.

3. Hesitation comes after the first triangle, when a constraint does not yet look like progress. Reinforce with: "inequalities never contradict each other alone; I need an equality, and three colours give equalities by pigeonhole".

4. The trap in (a) is staying on the triangular lattice, which is three-colourable, so no lattice-only figure can ever work. The trap in (b) is unit squares, whose diagonal $\sqrt{2}$ exceeds 1.

5. Whenever a problem forbids a relation (different colours, non-adjacent, distinct values), look for the smallest configuration where the pigeonhole principle turns the prohibitions into an *equality*, then compose equalities.

6. Reusable rule: to refute a colouring with k colours, find a gadget that pins one vertex's colour (a K_k plus one vertex adjacent to all of it), extract a forced-equal distance, and use rotation about a point to realise the forbidden distance between two forced-equal points. To construct a colouring, list every way a monochromatic bad pair can arise and turn each into an inequality on the design parameter.

▷ **Cue:** k colours and a configuration of k points pairwise at the forbidden distance $\longrightarrow$ **Move:** the configuration is rainbow; a further point adjacent to $k-1$ of them is forced to the remaining colour

▷ **Cue:** a hypothesis consisting only of "different" relations and a demand for a contradiction

$\longrightarrow$ **Move:** manufacture an "equal" relation first (pigeonhole), because equalities compose and inequalities do not

▷ **Cue:** two points forced to share a colour at distance d $\longrightarrow$ **Move:** rotate about a common point: all points on a circle of radius d share its colour, and the circle has chords of every length up to $2d$

▷ **Cue:** construct a colouring of the plane avoiding a distance $\longrightarrow$ **Move:** periodic tiling; require tile diameter below the distance and same-colour tile gap above it, then solve for the tile size

1988 A5

1988 A5. Prove that there exists a *unique* function f from the set R^+ of positive real numbers to R^+ such that

$$f(f(x)) = 6x - f(x)$$

and

$$f(x) > 0$$

for all $x > 0$.

Fields: functional equations, linear recurrences **Position:** A5 — short once seen; the difficulty is entirely in knowing what to do with "unique" **Answer:** $f(x) = 2x$. For each fixed x, the orbit $x, f(x), f(f(x)), \ldots$ satisfies $a_{n+2} = 6a_n - a_{n+1}$, whose general solution is $c_1 2^n + c_2(-3)^n$; positivity of every term forces $c_2 = 0$, hence $f(x) = 2x$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You guess a linear solution. Put $f(x) = kx$; then $f(f(x)) = k^2x$ and the equation becomes $k^2x = 6x - kx$, so $k^2 + k - 6 = 0$, i.e. $(k-2)(k+3) = 0$. Since f must be positive, $k = 2$, and $f(x) = 2x$ works: $f(f(x)) = 4x = 6x - 2x$. Existence is done in two lines.

Uniqueness is where the natural attempt goes wrong. The equation gives no obvious way to pin f at a point. Students try: is f injective? (Yes: $f(x) = f(y)$ gives $6x - f(x) = f(f(x)) = f(f(y)) = 6y - f(y)$, so $x = y$.) Is it increasing? Is it continuous? None of this is given, and you can spend a long time proving regularity that the problem does not need. A second, more productive grind is to squeeze f between linear bounds. Positivity of $f(f(x))$ gives $f(x) < 6x$. Feeding a bound $f(y) \le by$ back in: $6x - f(x) = f(f(x)) \le bf(x)$, so $f(x) \ge \frac{6}{1+b}x$; feeding a bound $f(y) \ge ay$ in: $6x - f(x) \ge af(x)$, so $f(x) \le \frac{6}{1+a}x$. Starting from $a_0 = 0$ this produces

$$0 < \tfrac{6}{7}x \ \ldots \le f(x) \le \ \ldots \ 6x,$$

and the bounds are iterates of $t \mapsto \frac{6}{1+t}$, which does converge to its fixed point 2 (the composite $t \mapsto \frac{6(1+t)}{7+t}$ is increasing with unique positive fixed point 2). So this route *works* — but it costs a page of bookkeeping, a convergence argument for the bounds, and it hides why 2 appears.

The model chain. *Existence.* $f(x) = 2x$ maps R^+ to R^+ and $f(f(x)) = 4x = 6x - 2x$.

Uniqueness. Let $f : \mathrm{R}^+ \to \mathrm{R}^+$ satisfy $f(f(x)) = 6x - f(x)$ for all $x > 0$. Fix $x > 0$ and define the orbit

$$a_0 = x, \qquad a_{n+1} = f(a_n) \quad (n \ge 0).$$

Every a_n is positive, because f takes positive values, so each a_n lies in the domain and the sequence is defined for all n. Applying the functional equation at the point a_n:

$$a_{n+2} = f(f(a_n)) = 6a_n - f(a_n) = 6a_n - a_{n+1}, \qquad \text{i.e.} \qquad a_{n+2} + a_{n+1} - 6a_n = 0.$$

This is a linear recurrence with constant coefficients and characteristic polynomial $t^2 + t - 6 = (t-2)(t+3)$, whose roots 2 and -3 are distinct. Choose c_1, c_2 with

$$c_1 + c_2 = a_0, \qquad 2c_1 - 3c_2 = a_1$$

(possible since the determinant is $-5 \neq 0$). The sequence $b_n = c_1 2^n + c_2(-3)^n$ satisfies the same recurrence and agrees with a_n at $n = 0, 1$, so by induction $a_n = b_n$ for all n.

Suppose $c_2 \neq 0$. Then

$$\frac{a_n}{c_2(-3)^n} = 1 + \frac{c_1}{c_2}\Big(-\frac{2}{3}\Big)^n \longrightarrow 1,$$

so for all large n, a_n has the sign of $c_2(-3)^n$, which alternates. That contradicts $a_n > 0$ for all n. Hence $c_2 = 0$, $a_n = c_1 2^n$, and in particular $f(x) = a_1 = 2a_0 = 2x$. Since $x > 0$ was arbitrary, $f(x) = 2x$ for every $x > 0$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains find $k = 2$ and $k = -3$ from the linear guess. They split at what they do with the equation afterwards. The natural attempt reads $f(f(x)) = 6x - f(x)$ as a statement about the *function* f and tries to prove properties of f. The model reads it as a statement about three consecutive terms of a *sequence*, x, $f(x)$, $f(f(x))$, and iterates.

> *The equation only ever evaluates f along the chain $x \mapsto f(x) \mapsto f(f(x))$. So fix one x and follow that chain. Along it, the equation is a second-order linear recurrence: $a_{n+2} = 6a_n - a_{n+1}$.*

> *The recurrence has roots 2 and -3. The -3 part flips sign every step and eventually dominates, but every term must stay positive. So the -3 coefficient is zero, and the orbit is pure 2^n.*

Layer 3 THE WHY

Kedlaya's solution begins: "Fix $x > 0$ and consider the sequence $a_0 = x$, $a_{n+1} = f(a_n)$." Here is where that comes from.

Look at what the equation constrains. It relates $f(f(x))$, $f(x)$ and x — the points x, $f(x)$, $f^2(x)$ of a single orbit. It never compares f at two unrelated inputs, such as x and $x + 1$. So the equation cannot tell you anything about how f behaves *across* different orbits; everything it says is *along* an orbit. The natural object is therefore the orbit, and on an orbit the equation, applied at every point, becomes

$$a_{n+2} = 6a_n - a_{n+1}.$$

That is the whole of the information in the hypothesis. Iterating is not a trick; it is extracting all the constraint there is.

Why is the answer forced rather than merely suggested? A second-order linear recurrence is determined by two initial values, and its solution space is two-dimensional, spanned by 2^n and $(-3)^n$ (distinct roots of the characteristic polynomial). The function f chooses $a_1 = f(x)$; that is one free parameter per orbit, equivalent to choosing c_2. Positivity is an *infinite* list of conditions, one for each n, and the only thing that survives all of them is $c_2 = 0$, because $|-3| > |2|$ means the $(-3)^n$ component wins in the long run and its sign alternates. So the uniqueness really lives in the *codomain* condition $f(x) > 0$, applied infinitely often. If the problem were posed for $f : \mathbb{R} \to \mathbb{R}$ with no sign condition, $f(x) = -3x$ would be a second solution ($f(f(x)) = 9x = 6x - (-3x)$) — which is why the problem states positivity twice.

This also explains the squeezing route of the natural attempt. The map $t \mapsto \frac{6}{1+t}$ that improves the bounds is exactly the recurrence rewritten for the ratio $r_n = a_{n+1}/a_n$ and run backwards: dividing $a_{n+2} = 6a_n - a_{n+1}$ by a_{n+1} gives $r_{n+1} = \frac{6}{r_n} - 1$, equivalently $r_n = \frac{6}{1+r_{n+1}}$, and the positive fixed point of that is 2. The squeeze is the recurrence in disguise, done the

long way. The characteristic polynomial $t^2 + t - 6$ is the same polynomial as the linear guess $k^2 + k - 6$; the linear guess found the eigenvalues, and the orbit argument proves that the eigenvalue -3 is excluded.

Layer $+\alpha$ BACKTESTING

1. The first cue is $f(f(x))$ expressed linearly in x and $f(x)$. On seeing it, write $a_0 = x$, $a_1 = f(x)$, $a_2 = f(f(x))$ and the relation $a_2 = 6a_0 - a_1$.
2. It suggests iterating to a linear recurrence and solving it with the characteristic polynomial.
3. Hesitation occurs after finding $f(x) = 2x$, when "prove uniqueness" seems to demand regularity of f (continuity, monotonicity) that is not given. Reinforce it away: uniqueness for an iterative equation usually comes from one orbit and one sign or growth condition, with no regularity needed.
4. The trap is proving only that the *linear* solutions are $2x$ and $-3x$ and calling that uniqueness. The linear guess says nothing about nonlinear f.
5. Any functional equation in which f is composed with itself and the rest is linear in x, $f(x)$, $f(f(x))$, ... is a linear recurrence on orbits. The coefficients of the equation are the coefficients of the characteristic polynomial.
6. Reusable rule: when a hypothesis only ever links x to $f(x)$ to $f(f(x))$, fix x and study the orbit; solve the orbit's recurrence in closed form; then use sign, boundedness or integrality conditions, which apply at every step, to kill all modes but one.

▷ **Cue:** a functional equation involving $f(f(x))$, $f(x)$ and x linearly $\longrightarrow$ **Move:** fix x, set $a_n = f^n(x)$, and write the equation as a linear recurrence in n
▷ **Cue:** a linear recurrence whose terms must all be positive (or bounded) $\longrightarrow$ **Move:** write the closed form; a root of larger modulus with negative (or modulus > 1) value must have coefficient zero
▷ **Cue:** "prove uniqueness" for a functional equation with no regularity hypothesis $\longrightarrow$ **Move:** look for an iteration or orbit argument, not for continuity or monotonicity
▷ **Cue:** the linear guess $f(x) = kx$ gives a polynomial in k $\longrightarrow$ **Move:** that polynomial is the characteristic polynomial of the orbit recurrence; use it to name the modes to be excluded

1988 A6

1988 A6. If a linear transformation A on an n-dimensional vector space has $n + 1$ eigenvectors such that any n of them are linearly independent, does it follow that A is a scalar multiple of the identity? Prove your answer.

Fields: linear algebra **Position:** A6 — the hardest slot on paper A, but a short problem; the difficulty is choosing which n of the $n + 1$ vectors to use as a basis and what to do with the last one **Answer:** Yes. Write the extra eigenvector in the basis of the other n; every coefficient is nonzero (else n of the vectors would be dependent), and applying A forces every eigenvalue to equal the extra one.

Layer 1 THE WHAT — surface analysis

The natural attempt. "Does it follow?" invites a hunt for a counterexample first, so you try $n = 2$ with a non-scalar diagonal matrix, $A = \mathrm{diag}(1, 2)$. Its eigenvectors are the nonzero multiples of e_1 and of e_2 and nothing else. Any three of them include two multiples of the same e_i, which are dependent. So no counterexample here, and every non-scalar diagonal matrix behaves the same way. A non-diagonalisable matrix does worse (fewer independent eigenvectors). The evidence points to "yes".

To prove it, the zombie route is to write $A = (a_{ij})$ in an arbitrary basis, write each eigenvector x_k in coordinates, and try to solve $Ax_k = \lambda_k x_k$ for the n^2 entries. For $n = 2$ this is already four unknowns plus three eigenvalues plus six coordinates, and the independence hypothesis —

"any two of the three are independent" — becomes three 2×2 determinants being nonzero. You can grind this out for $n = 2$, but it does not generalise, and the determinants never interact with the eigen-equations in any visible way.

A stronger student notices that $x_1, \ldots, x_n$ are independent, hence a basis of eigenvectors, so A is diagonalisable: in that basis $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Now the eigenvalues must be shown equal, and the student stalls, because the diagonal form alone does not force that — the information is in x_{n+1}, which has not been used yet.

The model chain. Let the eigenvectors be $x_1, \ldots, x_{n+1}$ (nonzero, by definition of eigenvector) with $Ax_i = \lambda_i x_i$. The vectors $x_1, \ldots, x_n$ are n independent vectors in an n-dimensional space, so they form a basis, and

$$x_{n+1} = \sum_{i=1}^{n} \alpha_i x_i$$

for unique scalars $\alpha_1, \ldots, \alpha_n$.

Every α_i is nonzero. If $\alpha_k = 0$ for some k, then x_{n+1} is a combination of the $n-1$ vectors x_i with $i \neq k$, $i \leq n$, so the n vectors $\{x_1, \ldots, x_{n+1}\} \setminus \{x_k\}$ are linearly dependent, contrary to hypothesis.

Apply A two ways. On one hand $Ax_{n+1} = \lambda_{n+1} x_{n+1} = \sum_i \lambda_{n+1} \alpha_i x_i$. On the other, by linearity, $Ax_{n+1} = \sum_i \alpha_i A x_i = \sum_i \lambda_i \alpha_i x_i$. Subtracting,

$$\sum_{i=1}^{n} (\lambda_i - \lambda_{n+1}) \alpha_i \, x_i = 0.$$

Since $x_1, \ldots, x_n$ are independent, $(\lambda_i - \lambda_{n+1})\alpha_i = 0$ for each i, and since $\alpha_i \neq 0$, $\lambda_i = \lambda_{n+1}$ for every i. Write λ for the common value.

Conclusion. A and λI agree on the basis $x_1, \ldots, x_n$, hence, being linear, agree everywhere: $A = \lambda I$. (For $n = 1$ the argument still runs: $x_2 = \alpha_1 x_1$ with $\alpha_1 \neq 0$ because $x_2 \neq 0$.)

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both the stronger natural attempt and the model reach "$x_1, \ldots, x_n$ is an eigenbasis, so $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$". They split at the next line. The natural attempt looks at the diagonal matrix for a reason the λ_i should agree. The model turns to the one vector not yet used and writes it in the basis: $x_{n+1} = \sum \alpha_i x_i$, then applies A.

> *I have used n of the eigenvectors to build a basis. The hypothesis is about $n+1$ of them, and "any n" is a statement about all the subsets, so the leftover vector x_{n+1} must carry the information. Write it in the basis I have and apply A to both sides.*

> *"Any n are independent" says, in coordinates, that no coordinate of x_{n+1} is zero. An eigenvector of a diagonal matrix with all coordinates nonzero forces all the diagonal entries to be equal.*

Layer 3 THE WHY

Kedlaya's Solution 2 (due to Lenny Ng) opens by writing $x_{n+1} = \sum \alpha_i x_i$. His Solution 1 instead compares traces in the $n+1$ different bases. Here is where both come from, and why the first is the one to reach for.

Start with the definition. An eigenvector v with eigenvalue μ is a vector on which A acts as the scalar μ. A scalar multiple of the identity is the map that acts by the *same* scalar on every vector. So the question is: can an operator act as a scalar on each of $n+1$ lines in general position without acting as one scalar everywhere? The obstruction is plain once you ask what A does to a *sum* of eigenvectors with different eigenvalues. If $Au = \mu u$ and $Av = \nu v$, then $A(u+v) = \mu u + \nu v$, which is a multiple of $u+v$ only if $\mu = \nu$ (when u, v are independent).

A vector that is a genuine mixture of eigenvectors from different eigenvalues is never an eigenvector. So the extra eigenvector, if it is a mixture of all n basis eigenvectors, can exist only if all their eigenvalues agree.

That leaves one question: is x_{n+1} a genuine mixture of *all* of $x_1, \ldots, x_n$, i.e. are all the α_i nonzero? That is exactly what "any n of them are linearly independent" says, once translated. The hypothesis is not decorative; it is precisely the statement that x_{n+1} has full support in the basis $x_1, \ldots, x_n$. This translation — "general position" equals "no zero coordinates" — is the step the natural attempt was missing, and it is why the extra vector is the thing to write down.

The trace proof works for the same reason but more expensively. Each choice of n of the vectors is an eigenbasis, giving $\operatorname{tr} A = S - \lambda_i$ where $S = \sum_{j=1}^{n+1} \lambda_j$; invariance of trace under change of basis forces all λ_i equal. It is correct and short, but it relies on a basis-invariance fact that has nothing to do with the question, and it hides the mechanism.

A third viewpoint, useful for generalising: eigenvectors for distinct eigenvalues are independent, so the eigenspaces E_μ form a direct sum and $\sum_\mu \dim E_\mu \leq n$. Distribute the $n+1$ vectors among the eigenspaces. If some eigenspace receives all $n+1$, it contains n independent vectors, so it is the whole space and A is scalar. Otherwise every eigenspace receives at most n vectors, which are then independent, so eigenspace E_μ receives at most $\dim E_\mu$ of them, and the total is at most $n < n+1$ — a contradiction. That is a pigeonhole on dimensions, and it shows the count $n+1$ is sharp: n eigenvectors in general position say nothing (every diagonalisable operator has them).

Layer $+\alpha$ BACKTESTING

1. The first cue is "$n+1$ vectors in an n-dimensional space". On seeing it, write: "n of them are a basis; express the last in that basis: $x_{n+1} = \sum \alpha_i x_i$."

2. It suggests applying A to that expression and comparing coefficients using independence of the basis.

3. Hesitation occurs after diagonalising, when $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ does not visibly force anything. Reinforce it away with: "I have not used every hypothesis yet; the unused vector is the lever."

4. The trap is the phrase "does it follow?", which sends you hunting for counterexamples among diagonal matrices, and the zombie route of solving for all n^2 matrix entries.

5. Whenever a hypothesis says "any n of these $n+1$ vectors are independent" (general position), translate it at once: in the basis of any n of them, the remaining vector has every coordinate nonzero.

6. Reusable rule: to show a linear map is forced into a special form, take a basis adapted to it (an eigenbasis), express every remaining given vector in that basis, apply the map, and compare coefficients. Zero coefficients are where freedom hides; hypotheses that rule them out are the ones doing the work.

▷ **Cue:** $n+1$ vectors in an n-dimensional space, any n independent $\longrightarrow$ **Move:** take n as a basis; the last has all coordinates nonzero in that basis

▷ **Cue:** an eigenvector that is a combination of eigenvectors from a basis $\longrightarrow$ **Move:** apply A both ways and compare coefficients: every basis vector with a nonzero coefficient shares its eigenvalue

▷ **Cue:** "show A is a scalar multiple of the identity" $\longrightarrow$ **Move:** show A acts by the same scalar on every vector of some basis

▷ **Cue:** several eigenvectors distributed among eigenspaces $\longrightarrow$ **Move:** eigenspaces form a direct sum, so $\sum \dim E_\mu \leq n$; pigeonhole the vectors against the dimensions

Putnam 1988 — Paper B

1988 B1

1988 B1. A *composite* (positive integer) is a product ab with a and b not necessarily distinct integers in $\{2, 3, 4, \dots\}$. Show that every composite is expressible as $xy + xz + yz + 1$, with x, y, z positive integers.

Fields: number theory, algebraic factorisation **Position:** B1 – opening problem of the second session (with 1988 A1, the highest-scoring problem of the year) **Answer:** Take $(x, y, z) = (a - 1,\ b - 1,\ 1)$: then $xy + xz + yz + 1 = (x + 1)(y + 1) = ab$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You read "composite" and "show that every ... is expressible", and your number-theory reflexes fire. You start listing: $4 = ?$, $6 = ?$, $8 = ?$, $9 = ?$. For 4 you need $xy + xz + yz = 3$ with positive integers, so $x = y = z = 1$. For 6 you need $xy + xz + yz = 5$: $(1, 1, 2)$ gives $1 + 2 + 2 = 5$. For 8 you need 7: $(1, 1, 3)$ gives $1 + 3 + 3 = 7$. For 9 you need 8: $(1, 2, 2)$ gives $2 + 2 + 4 = 8$. The data works, but it is not telling you a pattern – you are solving a fresh Diophantine equation each time. So you switch strategy and try to attack the expression head-on: it is symmetric in x, y, z, so perhaps it is $(x + 1)(y + 1)(z + 1)$ minus something? Expanding, $(x + 1)(y + 1)(z + 1) = xyz + (xy + xz + yz) + (x + y + z) + 1$, so

$$xy + xz + yz + 1 = (x + 1)(y + 1)(z + 1) - xyz - (x + y + z).$$

That is true, and useless: you now have a product minus a mess, and you wanted a product. Some students next try to prove the statement by induction on the composite, or to split into n even / n odd. Each of these burns time on an expression you have not yet looked at properly.

The model chain. Write the target as the thing the problem hands you: $n = ab$ with $a, b \geq 2$. We want positive integers x, y, z with $xy + xz + yz + 1 = ab$. Group the expression by one variable:

$$xy + xz + yz + 1 = xy + z(x + y) + 1.$$

If $z = 1$ this is $xy + x + y + 1 = (x + 1)(y + 1)$. So set

$$x = a - 1, \qquad y = b - 1, \qquad z = 1.$$

Since $a, b \geq 2$, both $x = a - 1 \geq 1$ and $y = b - 1 \geq 1$ are positive integers, and $z = 1$ is a positive integer. Then

$$xy + xz + yz + 1 = (a - 1)(b - 1) + (a - 1) + (b - 1) + 1 = ab,$$

which is the composite. That is the whole proof.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both paths start from the same expression. The natural attempt treats $xy + xz + yz + 1$ as a symmetric block to be manipulated as a whole, or as an equation to be solved number by number. The model breaks the symmetry at once: it asks whether some *special value* of one variable makes the expression factor, because the thing to be produced – a composite – is by definition a product of two factors.

> *The target is a product ab of two numbers, and I have three free variables. I only need two of them. What value of the third one makes $xy + xz + yz + 1$ collapse into a product of two brackets?*

> *$xy + x + y + 1$ is $(x+1)(y+1)$ – that is the factorisation I know. It appears exactly when $z = 1$. And $a, b \geq 2$ is precisely $x, y \geq 1$, so the constraint in the definition is the positivity constraint in the conclusion.*

Layer 3 THE WHY

Why is "set $z = 1$" not a trick? Go back to the definition the problem puts in italics. A composite is *defined* as ab with $a, b \geq 2$. So what you must manufacture is a product of two factors, each at least 2. The problem gives you three variables and one equation. There is a count mismatch in your favour: three unknowns to hit a target that has only two degrees of freedom, a and b. When you have more freedom than you need, the first move is to spend the surplus – fix one variable at a convenient value – and see what the remaining expression becomes. The question then is purely which value, and the answer comes from matching shapes: in two variables the only familiar "$\cdots + 1$" factorisation is

$$(x+1)(y+1) = xy + x + y + 1,$$

and $xy + z(x+y) + 1$ matches it term for term exactly when $z = 1$.

The official solution in Kedlaya–Poonen–Vakil is one line: "Substituting $z = 1$ yields $(x+1)(y+1)$." Here is where that line comes from: the definition of composite names a two-factor product, a three-variable expression has one variable to spare, and the spare is spent to make the expression look like a two-factor product. The "$+1$" in the expression is itself a cue – it is the constant term of $(x+1)(y+1)$ – and so is the threshold 2 in $\{2, 3, 4, \dots\}$, which is $1 + 1$, the smallest value of $x + 1$ for a positive integer x. The problem's numbers are shaped to fit this substitution; nothing about it needs luck.

Why does the natural attempt stall? Because listing $4, 6, 8, 9$ treats each composite as a separate equation to be solved, and the solutions it finds, such as $(1, 1, 2)$ for 6 or $(1, 2, 2)$ for 9, do not all have $z = 1$ in the order you happened to write them. The symmetry of the expression hides the pattern: $(1, 1, 2)$ is the same as $(2, 1, 1)$, and read that way $x = 2, y = 1, z = 1$ gives $(3)(2) = 6$. Had you insisted on writing each solution with a 1 in the last slot, the data would have shown you the substitution. And the $(x+1)(y+1)(z+1)$ expansion fails because it uses all three variables symmetrically, when the target is a product of only two things.

A remark on what this problem is sitting on top of: the book notes that all but finitely many *primes* are also of the form $xy + xz + yz + 1$ (for instance $13 = 2 \cdot 2 + 2 \cdot 2 + 2 \cdot 2 + 1$). A direct search up to 2000 confirms that the only non-representable positive integers there are $1, 2, 3, 5, 7, 11, 19, 23, 31, 43, 59, 71, 79, 103, 131, 191, 211, 331, 463$, matching the list in the book's remark; the book reports that ruling out a further exception above 10^{11} depends on the Generalised Riemann Hypothesis. The composite case is the easy corner of a deep problem, and the easy corner is exactly the part where the expression factors.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the italicised definition: a composite is ab with $a, b \geq 2$. On seeing it, write down $n = ab$, $a, b \geq 2$, and the goal $xy + xz + yz + 1 = ab$.

2. It suggests specialisation: three variables, a two-parameter target, so fix one variable so that the expression becomes a product of two brackets.

3. Hesitation occurs on seeing the symmetric expression, where the instinct is to keep it symmetric. It is reinforced away by asking how many parameters the target really has (two) and how many you are given (three).

4. The trap is treating "show every composite is expressible" as a number-theory classification problem – listing cases, inducting, splitting by parity – when it is an algebraic identity

problem. A secondary trap is forgetting to check that $x = a - 1$ is *positive*; that is exactly what $a \geq 2$ is for.

5. Whenever the target set is defined as a product (composites, non-squarefree numbers, reducible polynomials) and you are given more free variables than the product has factors, look for a value of the surplus variable that makes the expression factor.

6. The reusable rule: count the degrees of freedom of the target and of the expression; spend any surplus by fixing variables at the value that turns the expression into the target's shape, and use the target's constraints (here ≥ 2) to check the range of the substitution.

▷ **Cue:** an expression of the form $xy + x + y + 1$, or one that becomes it on fixing a variable $\longrightarrow$ **Move:** factor as $(x+1)(y+1)$ – Simon's favourite factoring trick

▷ **Cue:** show every member of a set defined as a product ab is of some form with more free variables than factors $\longrightarrow$ **Move:** fix the surplus variables so the form factors into the right number of brackets, then match brackets to factors

▷ **Cue:** the problem's lower bound on a parameter is 2 and the conclusion wants positive integers $\longrightarrow$ **Move:** expect the substitution $x = a - 1$; the bound $a \geq 2$ is exactly $x \geq 1$

▷ **Cue:** a symmetric expression in three variables must hit a two-parameter target $\longrightarrow$ **Move:** break the symmetry deliberately by setting one variable to a constant

1988 B2

1988 B2. Prove or disprove: If x and y are real numbers with $y \geq 0$ and $y(y+1) \leq (x+1)^2$, then $y(y-1) \leq x^2$.

Fields: inequalities, algebra **Position:** B2 – early second-session problem **Answer:** True: the implication holds for all real x and all $y \geq 0$.

Layer 1 THE WHAT — surface analysis

The natural attempt. "Prove or disprove" makes you suspicious, so you first hunt for a counterexample. You try y large and x just big enough for the hypothesis: with $x + 1 = \sqrt{y(y+1)}$, is $x^2 \geq y(y-1)$? Numerically, $y = 10$: $\sqrt{110} \approx 10.488$, so $x \approx 9.488$, $x^2 \approx 90.02 \geq 90$. Barely true. $y = 100$: $x \approx 99.4988$, $x^2 \approx 9900.0025 \geq 9900$. Barely true again. So it is probably true, and tight at infinity. Now you try to prove it by taking square roots. From $y(y+1) \leq (x+1)^2$ you get $|x+1| \geq \sqrt{y(y+1)}$. You want $|x| \geq \sqrt{y(y-1)}$ (when $y \geq 1$). Now the signs bite. If $x + 1 \geq 0$ then $x \geq \sqrt{y^2+y} - 1$, and you need $\sqrt{y^2+y} - 1 \geq \sqrt{y^2-y}$, which after squaring (both sides non-negative, which you must check for $y \geq 1$) becomes $2y + 1 \geq 2\sqrt{y^2+y}$, and squaring again, $4y^2 + 4y + 1 \geq 4y^2 + 4y$. True. If $x + 1 < 0$ then $-x \geq 1 + \sqrt{y^2+y}$, even bigger, fine. And $0 \leq y < 1$ separately. It works – this is essentially the book's Solution 2 – but it is two layers of square roots, two squarings each needing a sign justification, and three cases, for an inequality between two quadratics.

The model chain. If $0 \leq y \leq 1$ then $y(y-1) \leq 0 \leq x^2$ and there is nothing to prove. So assume $y > 1$. Compare the hypothesis and the conclusion term by term. Their left sides differ by

$$y(y+1) - y(y-1) = 2y,$$

and their right sides differ by

$$(x+1)^2 - x^2 = 2x + 1.$$

Hence, using the hypothesis,

$$y(y-1) = y(y+1) - 2y \leq (x+1)^2 - 2y = x^2 + (2x + 1 - 2y).$$

Case $x \leq y - \frac{1}{2}$. Then $2x + 1 - 2y \leq 0$, so $y(y-1) \leq x^2$.

Case $x > y - \frac{1}{2}$. Since $y > 1$, $y - \frac{1}{2} > 0$, so $x > y - \frac{1}{2} > 0$ and squaring preserves the inequality:

$$x^2 > \left(y - \tfrac{1}{2}\right)^2 = y^2 - y + \tfrac{1}{4} > y(y-1).$$

In both cases $y(y-1) \le x^2$. The statement is true.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both paths agree the claim is probably true after testing large y. They split at the first manipulation. The natural attempt takes square roots of the hypothesis to isolate x, which brings in $|x+1|$, sign cases and nested radicals. The model does not isolate anything: it writes the conclusion's left side as the hypothesis's left side minus $2y$, i.e. it *subtracts* the two inequalities' ingredients, and the remaining gap $2x + 1 - 2y$ is linear.

> *The hypothesis and the conclusion are the same two quadratics with everything shifted by one:* $y(y+1)$ *against* $y(y-1)$, $(x+1)^2$ *against* x^2. *So I should not solve either one – I should subtract and see what the shift costs on each side.*

> *After subtracting, the leftover is* $2x + 1 - 2y$, *which is linear. Its sign splits the plane by the line* $x = y - \frac{1}{2}$. *On one side I am done; on the other side* x *is big, and a big* x *makes* x^2 *large without any hypothesis at all.*

Layer 3 THE WHY

Why is subtraction the natural move rather than a lucky one? Look at the shape of the statement. Put $f(t) = t(t-1)$ and $g(t) = t^2$. Then

$$y(y+1) = f(y+1), \qquad y(y-1) = f(y), \qquad (x+1)^2 = g(x+1), \qquad x^2 = g(x).$$

So the problem reads: if $f(y+1) \le g(x+1)$ then $f(y) \le g(x)$. It is the *same* comparison at two points one unit apart along the diagonal direction $(1,1)$. Whenever two statements are the same expression evaluated at shifted arguments, the information that connects them is the *finite difference*: $f(y+1) - f(y) = 2y$ and $g(x+1) - g(x) = 2x + 1$. Quadratics have linear first differences, and that is the axiom-level reason the model's leftover is linear and the case split is along a straight line. Taking square roots, by contrast, destroys the polynomial structure that makes differences cheap: $\sqrt{y(y+1)} - \sqrt{y(y-1)}$ is not a nice object, which is why the natural attempt needed two squarings to recover what subtraction gives in one line.

This is the cheapness principle applied to inequalities: before solving for a variable, ask what the hypothesis and conclusion *have in common*. Here they share everything except a unit shift. Kedlaya–Poonen–Vakil's Solution 1 opens with "$y(y-1) = y(y+1) - 2y$"; the hint in the book is "Subtract and factor." Here is where that opening comes from: the two sides of the problem are translates of each other, and the translation is measured by subtraction.

The case $x > y - \frac{1}{2}$ also deserves a WHY, because it looks like it ignores the hypothesis – and it does. Complete the square: $y(y-1) = (y - \frac{1}{2})^2 - \frac{1}{4}$. So the conclusion is $x^2 \ge (y - \frac{1}{2})^2 - \frac{1}{4}$, which holds automatically whenever $|x| \ge y - \frac{1}{2}$. The hypothesis is only needed in the region where x is small relative to y, and the linear leftover identifies exactly that region. Similarly, the book's third solution (a monotonicity argument along $t \mapsto (x+t, y+t)$) is the continuous version of the same idea: follow the diagonal shift and track how $f - g$ changes.

Finally, why did the counterexample hunt come out "barely true"? With $u = y - \frac{1}{2}$ the hypothesis is $(u+1)^2 - (x+1)^2 \le \frac{1}{4}$ and the conclusion is $u^2 - x^2 \le \frac{1}{4}$. Along the hyperbola-like boundary, $u - x \to 0$ as $y \to \infty$, so the two differences $2(u - x)$ separating them tend to 0. The implication is tight asymptotically; that is why it is asked as "prove or disprove".

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that hypothesis and conclusion are built from the same pieces shifted by one: $y(y\pm1)$, $(x+1)^2$ versus x^2. On seeing it, write down the two differences $y(y+1)-y(y-1)=2y$ and $(x+1)^2-x^2=2x+1$.

2. It suggests subtracting: express the conclusion's left side through the hypothesis, $y(y-1)=y(y+1)-2y\le x^2+(2x+1-2y)$, and split on the sign of the linear remainder.

3. Hesitation occurs at "prove or disprove", where time drains into testing. Two quick numerical tests at large y are enough to commit to "true"; commit and look for structure.

4. The trap is taking square roots: it introduces $|x+1|$, sign cases, and squarings that each need a non-negativity check. A second trap is forgetting the case $0\le y\le 1$, where the square-root approach needs $\sqrt{y(y-1)}$ of a negative number.

5. Whenever the hypothesis is the conclusion with its arguments shifted (or scaled), compute the difference between the two versions of each side before doing anything else; for polynomials the difference has lower degree.

6. The reusable rule: compare hypothesis and conclusion structurally first; if they are the same expressions at shifted arguments, subtract, and split cases on the sign of the (lower-degree) remainder.

▷ **Cue:** hypothesis and conclusion are the same polynomial comparison at arguments shifted by one $\longrightarrow$ **Move:** subtract: take finite differences of each side, which drop the degree by one

▷ **Cue:** an inequality between quadratics, with the temptation to take square roots $\longrightarrow$ **Move:** avoid roots; subtract or complete the square so the comparison stays polynomial

▷ **Cue:** the remaining condition after subtraction is linear in the variables $\longrightarrow$ **Move:** split into two cases along that line; in the far case check whether the conclusion holds without the hypothesis

▷ **Cue:** "prove or disprove" and numerical tests come out barely true $\longrightarrow$ **Move:** the statement is probably true but tight at infinity; stop testing and look for the exact identity

1988 B3

1988 B3. For every n in the set $\mathrm{N}=\{1,2,\dots\}$ of positive integers, let r_n be the minimum value of $|c-d\sqrt{3}|$ for all nonnegative integers c and d with $c+d=n$. Find, with proof, the smallest positive real number g with $r_n\le g$ for all $n\in\mathrm{N}$.

Fields: number theory (Diophantine approximation), real analysis (density of $n\alpha$ mod 1) **Position:** B3 – middle of the second session **Answer:** $g=\dfrac{1+\sqrt{3}}{2}\approx 1.3660$; it is a supremum, never attained.

Layer 1 THE WHAT — surface analysis

The natural attempt. You compute. With $c=n-d$, you try every d for small n:

$$r_1=1,\ r_2=\sqrt{3}-1\approx 0.732,\ r_3=2-\sqrt{3}\approx 0.268,\ r_4=3-\sqrt{3}\approx 1.268,\ r_5\approx 0.464,\ r_6\approx 0.536,\ r_7\approx 1.196,\dots$$

The values jump around with no visible order. You look for the largest: $r_1=1$, then $r_4\approx 1.268$, then (if you are patient) $r_{15}=10-5\sqrt{3}\approx 1.340$, then $r_{56}=36-20\sqrt{3}\approx 1.359$. The record keeps creeping up. At this point you either guess from the records (and the guesses 1, $3-\sqrt{3}$, $4/3$ are all wrong), or you try to bound r_n directly: for each n, you want some d with $|n-d-d\sqrt{3}|$ small, and you start estimating $d\approx n/(1+\sqrt{3})$ by rounding, but you keep the two terms c and $d\sqrt{3}$ separate and the bound you get is loose, something like $r_n\le 1+\sqrt{3}$. Now you have an upper bound that is too big and a sequence of lower bounds that do not obviously converge anywhere.

The model chain. *Step 1: one variable.* Since $c = n - d$ with $0 \le d \le n$,

$$c - d\sqrt{3} = n - d(1+\sqrt{3}).$$

Put $\lambda = 1+\sqrt{3}$. Then $r_n = \min_{0\le d\le n} |n - d\lambda| = \lambda \min_{0\le d\le n} \left|\frac{n}{\lambda} - d\right|$. Because $0 < n/\lambda < n$, the integer nearest to n/λ lies in $\{0, 1, \ldots, n\}$, so the restriction $0 \le d \le n$ costs nothing:

$$r_n = \lambda \, \|n\beta\|, \qquad \beta = \frac{1}{\lambda} = \frac{\sqrt{3}-1}{2},$$

where $\|t\|$ denotes the distance from t to the nearest integer.

Step 2: the bound. Since $\|t\| \le \frac{1}{2}$ for every real t, $r_n \le \lambda/2 = \frac{1+\sqrt{3}}{2} =: g$ for all n. (Equality would need $n\beta \in \mathbb{Z} + \frac{1}{2}$, i.e. $\beta = (2k+1)/(2n)$ rational; β is irrational, so in fact $r_n < g$ always.)

Step 3: sharpness. It remains to show that no smaller g' works, i.e. for every $\varepsilon > 0$ some n has $\|n\beta\| > \frac{1}{2} - \varepsilon$. Write $\{t\} = t - \lfloor t \rfloor$. Choose an integer $N > 1/\varepsilon$. The $N+1$ numbers $\{\beta\}, \{2\beta\}, \ldots, \{(N+1)\beta\}$ lie in $[0,1)$, which is the union of the N intervals $[i/N, (i+1)/N)$; by pigeonhole two of them, $\{p\beta\}$ and $\{q\beta\}$ with $p < q$, share an interval. Let $m = q - p \ge 1$. Then $m\beta$ differs from an integer by less than $1/N$, and not by 0 since β is irrational, so either $\{m\beta\} = \delta$ or $\{m\beta\} = 1 - \delta$ for some $0 < \delta < 1/N < \varepsilon$.

If $\{m\beta\} = \delta$, then $\{km\beta\} = k\delta$ for every k with $k\delta < 1$; these points march up from 0 in steps shorter than ε, so some k has $k\delta \in (\frac{1}{2} - \varepsilon, \frac{1}{2})$. If $\{m\beta\} = 1 - \delta$, then $\{km\beta\} = 1 - k\delta$ while $k\delta < 1$, marching down from 1 in steps shorter than ε, so some k has $1 - k\delta \in (\frac{1}{2}, \frac{1}{2} + \varepsilon)$. Either way $n = km$ satisfies $\|n\beta\| > \frac{1}{2} - \varepsilon$, hence $r_n > g - \lambda\varepsilon$. As ε is arbitrary, no $g' < g$ bounds every r_n.

Therefore the smallest g is $\dfrac{1+\sqrt{3}}{2}$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both paths start by setting $c = n - d$. The natural attempt then keeps thinking of c and $d\sqrt{3}$ as two separate quantities being balanced. The model writes $c - d\sqrt{3} = n - d(1+\sqrt{3})$ and reads it as "n minus a multiple of one fixed step λ" – an arithmetic progression through n with common difference λ, which must straddle 0. From that line on the upper bound is one sentence and the rest is a density question.

> *The constraint $c + d = n$ kills a variable. Then $c - d\sqrt{3}$ is $n - d(1+\sqrt{3})$: as d goes up by one, the value drops by exactly $1+\sqrt{3}$. A progression with step $1+\sqrt{3}$ that runs from n down to $-n\sqrt{3}$ must pass within half a step of 0. So $r_n \le (1+\sqrt{3})/2$.*

> *The question asks for the* smallest *g, and my records $1, 1.268, 1.340, 1.359$ are creeping towards 1.366 without reaching it. So the answer is a supremum, and the proof of sharpness has to be a limiting argument: I need $n/(1+\sqrt{3})$ to come as close as I like to a half-integer, which is density of multiples of an irrational mod 1.*

Layer 3 THE WHY

Why is the substitution to one step natural? Because the problem's constraint $c + d = n$ is a line, and on a line the expression $c - d\sqrt{3}$ is an affine function of one parameter. Affine functions of an integer parameter are arithmetic progressions, and the only structural fact about an arithmetic progression with step s that crosses 0 is that some term lies within $s/2$ of 0. That is the whole upper bound. The natural attempt stalls because it never reduces to one parameter: it keeps two terms, c and $d\sqrt{3}$, and so never sees that there is a single step size $\lambda = 1+\sqrt{3}$ controlling everything. The appearance of $1+\sqrt{3}$ in the answer is itself a cue – the answer is half the step.

Kedlaya–Poonen–Vakil phrase the first half as: "the sequence n, $(n-1)-\sqrt{3}, \ldots, -n\sqrt{3}$ is an arithmetic sequence with common difference $-2g$." Here is where that comes from: $c+d=n$ is a constraint, constraints are for eliminating variables, and eliminating c leaves a progression. The "$2g$" is not guessed; it is λ, and g is defined as half of it because half the step is the worst-case distance to 0.

Why is density the right tool for the second half, and why is it forced? The question asks for the *smallest* g that bounds every r_n, i.e. $\sup_n r_n$. Rewriting $r_n = \lambda\|n\beta\|$ shows that r_n depends only on the fractional part of $n\beta$. So $\sup r_n = \lambda/2$ exactly when $\{n\beta\}$ can be pushed arbitrarily close to $\frac{1}{2}$ – a statement about where the multiples of β land mod 1. Irrationality of β is the only information available about β, and it is exactly what the pigeonhole argument uses: two multiples fall close together, their difference is a tiny nonzero step δ (nonzero only because $\beta \notin \mathbb{Q}$), and repeating a tiny step visits every region of $[0,1)$. If $\sqrt{3}$ were replaced by a rational number the fractional parts would be periodic, the supremum would be a maximum over finitely many residues, and it could be strictly less than half the step; so irrationality is not decoration here but the reason the answer is $\lambda/2$ and not something smaller. It also explains why the numerical records never reach 1.366: the supremum is not attained.

The "pigeonhole on fractional parts" argument is Dirichlet's approximation principle, and the book's remark adds the stronger fact that $\{n\alpha\}$ is equidistributed (Weyl). Only density is needed here; do not reach for more machinery than the claim requires.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the constraint $c+d=n$ inside a minimum of $|c-d\sqrt{3}|$. On seeing it, write $c-d\sqrt{3} = n-d(1+\sqrt{3})$.

2. It suggests reading the values as an arithmetic progression with step $\lambda = 1+\sqrt{3}$ crossing 0, giving $r_n \le \lambda/2$, and then rewriting $r_n = \lambda\|n/\lambda\|$ to turn sharpness into a density question.

3. Hesitation occurs at "smallest g", when small cases do not produce g. What reinforces it away is recognising that a bound approached but not reached is a supremum, so the proof must be a limit argument, not an example.

4. The trap is guessing g from the largest value found in small cases (for example $3-\sqrt{3}$ from $n=4$). A second trap is forgetting the range $0 \le d \le n$ – here it is harmless because $0 < n/\lambda < n$, but that must be said.

5. Whenever an expression is linear in an integer parameter with an irrational coefficient, its distance to a fixed target is governed by fractional parts of multiples of an irrational; worst-case bounds are half a step, and they are approached (not attained) by density.

6. The reusable rule: use each linear constraint to remove a variable until the expression is a progression in one integer; bound it by half the step; for sharpness with an irrational step, invoke pigeonhole density of $\{k\alpha\}$.

▷ **Cue:** a minimum of a linear expression in integers subject to a linear constraint like $c+d=n$ ⟶ **Move:** eliminate a variable; the values form an arithmetic progression, so some term is within half a step of the target

▷ **Cue:** "find the smallest g with $r_n \le g$ for all n" and numerical records creep upward ⟶ **Move:** the answer is a supremum that may not be attained; prove the bound, then prove approach by a limiting argument

▷ **Cue:** an expression depending on $n\alpha$ with α irrational, and you need it near a specific value mod 1 ⟶ **Move:** pigeonhole $N+1$ multiples into N intervals to get a tiny nonzero step, then iterate that step across $[0,1)$

▷ **Cue:** the answer involves half of a natural quantity like $(1+\sqrt{3})/2$ ⟶ **Move:** look for a step of size $1+\sqrt{3}$; half-steps are worst-case rounding errors

1988 B4

1988 B4. Prove that if $\sum_{n=1}^{\infty} a_n$ is a convergent series of positive real numbers, then so is $\sum_{n=1}^{\infty} (a_n)^{n/(n+1)}$.

Fields: analysis; series; inequalities **Position:** B4 — middle of the paper; the whole problem is one case split, and the case split is forced by one line of algebra **Answer:** True. Put $b_n = a_n^{n/(n+1)}$. If $a_n \geq 2^{-(n+1)}$ then $b_n = a_n \cdot a_n^{-1/(n+1)} \leq 2a_n$; if $a_n < 2^{-(n+1)}$ then $b_n < 2^{-n}$. So $b_n \leq 2a_n + 2^{-n}$ for every n, and $\sum b_n \leq 2\sum a_n + 1 < \infty$.

Layer 1 THE WHAT — surface analysis

The natural attempt. Convergence of a positive series with terms tied to a convergent series: the student reaches for the comparison test, and wants $b_n = a_n^{n/(n+1)} \leq Ca_n$. Write the ratio:

$$\frac{b_n}{a_n} = a_n^{-1/(n+1)}.$$

Since $a_n \to 0$, $a_n < 1$ eventually, so this ratio is > 1 — the new terms are *bigger* than the old ones. Is the ratio at least bounded? Try $a_n = 1/n^2$: the ratio is $n^{2/(n+1)} \to 1$, fine. Try $a_n = e^{-n^2}$: the ratio is $e^{n^2/(n+1)} \to \infty$. So no inequality $b_n \leq Ca_n$ holds in general, and the plain comparison test is dead.

The next reflexes are the ratio and root tests. The root test on $\sum b_n$ needs $b_n^{1/n} = a_n^{1/(n+1)}$ to be eventually below some $r < 1$, but $\sum a_n$ convergent says nothing of the kind (for $a_n = 1/n^2$, $a_n^{1/(n+1)} \to 1$). The ratio test is worse: nothing is known about a_{n+1}/a_n. Then come attempts to use monotonicity, or to assume a_n is decreasing "without loss of generality" (it is not without loss: rearranging changes which exponent each term receives). The wall: every standard test either needs information the hypothesis does not give, or needs a uniform comparison that the example e^{-n^2} destroys.

The model chain. Let $b_n = a_n^{n/(n+1)}$. Fix n and split into two cases according to the size of a_n against the threshold $2^{-(n+1)}$.

Case 1: $a_n \geq 2^{-(n+1)}$. Then $a_n^{1/(n+1)} \geq \frac{1}{2}$, since $x \mapsto x^{1/(n+1)}$ is increasing on $(0, \infty)$. Hence

$$b_n = \frac{a_n}{a_n^{1/(n+1)}} \leq 2a_n.$$

Case 2: $a_n < 2^{-(n+1)}$. Since $x \mapsto x^{n/(n+1)}$ is increasing on $(0, \infty)$,

$$b_n < \left(2^{-(n+1)}\right)^{n/(n+1)} = 2^{-n}.$$

In both cases $0 < b_n \leq 2a_n + 2^{-n}$. The partial sums of $\sum b_n$ are increasing and bounded above by $2\sum_{n\geq 1} a_n + \sum_{n\geq 1} 2^{-n} = 2\sum a_n + 1$, so $\sum b_n$ converges. □

Same idea, one line. Weighted AM–GM (Young's inequality) with weights $\frac{n}{n+1}, \frac{1}{n+1}$ gives, for any $c > 0$,

$$a_n^{n/(n+1)} = (ca_n)^{\frac{n}{n+1}} \left(c^{-n}\right)^{\frac{1}{n+1}} \leq \tfrac{n}{n+1}\, ca_n + \tfrac{1}{n+1}\, c^{-n} \leq ca_n + c^{-n}.$$

With $c = 2$ this is the same bound without cases. The case split is the version you can find in the arena; the AM–GM line is the version you write up once you know what you are looking for.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains write down the ratio $b_n/a_n = a_n^{-1/(n+1)}$ and see that it can be unbounded. (N) reads this as "comparison fails" and moves on to other tests. (M) reads it as "comparison fails *only when* a_n *is tiny*" and asks how tiny, and what b_n is in that case.

> *The exponent only hurts me when a_n is very small. But when a_n is very small, b_n is also small in absolute terms. So I don't need one comparison — I need two: compare with a_n when a_n is not too small, and with a fixed convergent series when it is.*

> *How small is "too small"? The bad factor is $a_n^{-1/(n+1)}$. That is bounded by 2 exactly when $a_n \geq 2^{-(n+1)}$. So the threshold has to be geometric in n — and conveniently a geometric sequence is summable.*

- ▷ **Cue:** a term-by-term comparison $b_n \leq Ca_n$ fails only on terms where a_n is very small (or very large) $\longrightarrow$ **Move:** split each term into cases by a threshold ε_n and bound $b_n \leq Ca_n + (\text{something summable})$
- ▷ **Cue:** an exponent that depends on n, like $a_n^{1/(n+1)}$ or $a_n^{n/(n+1)}$ $\longrightarrow$ **Move:** choose a geometric threshold $c^{-(n+1)}$, because its $(n+1)$th root is the constant $1/c$
- ▷ **Cue:** you want $x^\theta \leq (\text{linear in } x) + (\text{constant})$ for $0 < \theta < 1$ $\longrightarrow$ **Move:** weighted AM–GM / Young: $x^\theta y^{1-\theta} \leq \theta x + (1-\theta)y$

Layer 3 THE WHY

Kedlaya's hint for this problem is one line: divide the terms according to whether $a_n \geq 1/2^{n+1}$. Here is where that threshold comes from.

Start from what the hypothesis actually gives you. "$\sum a_n$ converges, $a_n > 0$" is information about sums, not about individual terms beyond $a_n \to 0$. So the only tool that can transfer it to $\sum b_n$ is a term-by-term inequality $b_n \leq (\text{const})\, a_n + (\text{summable})$. The ratio, root and ratio-type tests all need extra structure on a_n that you have not been given; that is why the natural attempt dies on them. The comparison test is the only honest route, and it must be repaired, not replaced.

Now unpack the failure. $b_n = a_n \cdot a_n^{-1/(n+1)}$, and the multiplier $a_n^{-1/(n+1)}$ is large precisely when a_n is small *on the scale of* n: an $(n+1)$th root pulls any number that is not exponentially small towards 1. Solve the condition directly: $a_n^{-1/(n+1)} \leq 2 \iff a_n \geq 2^{-(n+1)}$. So the threshold is not guessed; it is the solution set of "the multiplier is harmless". The only thing chosen is the constant 2, and any constant $c > 1$ works because $\sum c^{-n}$ converges.

Then the other case is free. Below the threshold, monotonicity of $x^{n/(n+1)}$ gives $b_n < (2^{-(n+1)})^{n/(n+1)} = 2^{-n}$ — a summable series that does not depend on (a_n) at all. That is the structural fact the whole problem rests on: a function x^θ with $0 < \theta < 1$ is worse than linear only near 0, and near 0 its values are themselves small. Every "x^θ versus x" estimate is a split at the point where the two cross, and the crossing point here is where $a_n = 2^{-(n+1)}$ gives $b_n = 2^{-n} = 2a_n$ — the two bounds meet exactly at the threshold. That meeting is why the split is natural and not a trick.

Young's inequality is the same fact with the case split hidden inside convexity: log is concave, so a geometric mean is at most the arithmetic mean. The Blueprint standard asks: what makes this move "of course"? The definition of b_n as a weighted geometric mean of a_n (weight $\frac{n}{n+1}$) and 1 (weight $\frac{1}{n+1}$) — a geometric mean with a mostly-a_n weight wants to be compared with the arithmetic mean with the same weights, and the free constant c lets you rebalance.

Layer $+\alpha$ BACKTESTING

1. The first cue is the n-dependent exponent $n/(n+1)$ on a term of a convergent positive series. On seeing it write $b_n/a_n = a_n^{-1/(n+1)}$ and ask when this is bounded.

2. Comparison test, repaired by a case split: $b_n \leq 2a_n$ above a geometric threshold, $b_n \leq 2^{-n}$ below it.

3. Hesitation occurs after the counterexample $a_n = e^{-n^2}$ kills $b_n \leq Ca_n$. Reinforce it away with: a failed uniform comparison tells you *where* it fails; bound that region separately.

4. The trap is abandoning comparison for the ratio or root test, neither of which can use the hypothesis. A second trap is assuming (a_n) decreasing, which is not given and not harmless because each term carries its own exponent.
5. Whenever a power x^θ with $0 < \theta < 1$ must be controlled by x, remember the two regimes cross at one point; name the point and split there.
6. Reusable decision rule: *if a term-by-term comparison fails, locate the terms where it fails, solve for the threshold that makes the bad factor bounded, and show that below the threshold the term is dominated by a fixed summable sequence. The bound becomes $b_n \le Ca_n + \varepsilon_n$ with $\sum \varepsilon_n < \infty$.*

1988 B5

1988 B5. For positive integers n, let M_n be the $2n+1$ by $2n+1$ skew-symmetric matrix for which each entry in the first n subdiagonals below the main diagonal is 1 and each of the remaining entries below the main diagonal is -1. Find, with proof, the rank of M_n. (According to one definition, the rank of a matrix is the largest k such that there is a $k \times k$ submatrix with nonzero determinant.)

One may note that

$$M_1 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 0 & -1 & -1 & 1 & 1 \\ 1 & 0 & -1 & -1 & 1 \\ 1 & 1 & 0 & -1 & -1 \\ -1 & 1 & 1 & 0 & -1 \\ -1 & -1 & 1 & 1 & 0 \end{pmatrix}.$$

Fields: linear algebra; matrices; parity **Position:** B5 — late in the paper, but the bound splits into two short halves once you see them **Answer:** The rank is $2n$. The all-ones vector is in the kernel, so the rank is at most $2n$; the top-left $2n \times 2n$ submatrix is congruent mod 2 to $J - I$, whose determinant $-(2n-1)$ is odd, so that minor is nonzero and the rank is at least $2n$.

Layer 1 THE WHAT — surface analysis

The natural attempt. The student computes. M_1: the rows add to zero and the first two rows are independent, so rank 2. M_2: row-reduce the 5×5; after a page of elimination the rank is 4. The guess "rank $2n$" is on the table within ten minutes.

Proving it: the student tries Gaussian elimination on general M_n. Subtract row 1 from row 2, and the pattern of ± 1s shifts by one place with a wrap-around; subtract again, and the entries become $0, \pm 2$ in positions that depend on whether the index has passed n. After two or three steps the bookkeeping of which entries are 0, ± 1, ± 2 runs away. The upper bound is also in doubt: the student may remember that an odd skew-symmetric matrix is singular ($\det M = \det M^T = \det(-M) = (-1)^{2n+1} \det M$), which gives rank $\le 2n$, but then wants the *exact* determinant of a $2n \times 2n$ minor to prove the lower bound. Computing that determinant exactly — for small n it comes out as 1 — is a genuine Pfaffian calculation, and the arena clock runs out. The wall is the belief that "nonzero determinant" requires knowing the determinant.

The model chain. Index rows and columns by $0, 1, \ldots, 2n$. The description (and the displayed examples) say the entry in row i, column j depends only on $d = (j - i) \bmod (2n+1)$:

$$(M_n)_{ij} = \begin{cases} 0, & d = 0, \\ -1, & 1 \le d \le n, \\ +1, & n+1 \le d \le 2n. \end{cases}$$

(Check against the statement: below the diagonal, $j < i$, $d = 2n + 1 - (i - j)$, so the first n subdiagonals $i - j \in \{1, \dots, n\}$ have $d \in \{n+1, \dots, 2n\}$ and entry $+1$; the rest have entry -1. Above the diagonal, skew-symmetry gives the negatives, matching $d \in \{1, \dots, n\} \mapsto -1$ and $d \in \{n+1, \dots, 2n\} \mapsto +1$.)

Upper bound. Each row contains every nonzero residue d exactly once, hence n entries equal to -1 and n equal to $+1$. So $M_n \mathbf{1} = \mathbf{0}$ where $\mathbf{1} = (1, \dots, 1)^T$. A nonzero kernel vector means M_n is singular, so every $(2n+1) \times (2n+1)$ minor vanishes and rank $M_n \le 2n$.

Lower bound. Let A be the submatrix on rows and columns $0, \dots, 2n-1$. Every off-diagonal entry of A is $\pm 1 \equiv 1 \pmod 2$ and every diagonal entry is 0, so

$$\det A \equiv \det(J - I) \pmod 2,$$

where J is the $2n \times 2n$ all-ones matrix (the determinant is a polynomial with integer coefficients in the entries, so it respects congruences). J has eigenvalue $2n$ on $\mathbf{1}$ and 0 on the $(2n-1)$-dimensional space of vectors with coordinate sum 0; so $J - I$ has eigenvalues $2n - 1$ (once) and -1 ($2n - 1$ times), and

$$\det(J - I) = (2n-1)(-1)^{2n-1} = -(2n-1),$$

an odd integer. Hence $\det A$ is odd, in particular nonzero, and rank $M_n \ge 2n$.

Therefore rank $M_n = 2n$. □

Second dress (the circulant route). M_n is a circulant, so $v_j = (1, \omega, \omega^2, \dots, \omega^{2n})^T$ with $\omega = e^{2\pi i j/(2n+1)}$ is an eigenvector with eigenvalue $\lambda_j = -\sum_{k=1}^{n} \omega^k + \sum_{k=n+1}^{2n} \omega^k$. For $j = 0$ this is 0. For $1 \le j \le 2n$, write $\theta = 2\pi j/(2n+1)$; since $\omega^{2n+1-k} = \overline{\omega^k}$ the second sum is the conjugate of the first, so $\lambda_j = -2i \sum_{k=1}^{n} \sin k\theta$, and the standard sine-sum identity with $(n + \frac{1}{2})\theta = \pi j$ gives

$$\sum_{k=1}^{n} \sin k\theta = \frac{\sin^2(\pi j/2) - \sin^2(\theta/4)}{\sin(\theta/2)},$$

which equals $\cos^2(\theta/4)/\sin(\theta/2)$ for odd j and $-\sin^2(\theta/4)/\sin(\theta/2)$ for even j; neither vanishes because $0 < \theta/4 < \pi/2$. So exactly one of the $2n+1$ distinct eigenvectors has eigenvalue 0, and the rank is $2n$. This is the route of Kedlaya's hint; it is correct but costs a trigonometric identity that the parity argument does not need.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains reach "rank $\le 2n$" quickly. They split at the lower bound. (N) says: to show a $2n \times 2n$ minor is nonzero I must compute it, and starts eliminating. (M) says: I do not need the value, only that it is not zero, and reduces mod 2.

> *I don't need to know this determinant. I only need to know it isn't zero. What is the cheapest property of an integer that rules out zero? Being odd. And mod 2 every ± 1 is the same, so the signs — the only messy part of this matrix — vanish.*

> *Each row has n plus-ones and n minus-ones. So every row kills the all-ones vector. That's the kernel, and that's the upper bound for free.*

- ▷ **Cue:** a matrix with integer entries whose determinant you need only to be nonzero ⟶ **Move:** reduce modulo a small prime (usually 2) and compute the determinant of the simpler matrix there
- ▷ **Cue:** an integer matrix whose off-diagonal entries are all ± 1 and whose diagonal is even ⟶ **Move:** mod 2 it is $J - I$ (or J); use the eigenvalues of J: m once and 0 otherwise

▷ **Cue:** every row of a matrix has the same entries in some order summing to zero $\longrightarrow$ **Move:** the all-ones vector is in the kernel

▷ **Cue:** an entry depends only on $j - i$ modulo the size $\longrightarrow$ **Move:** the matrix is circulant: its eigenvectors are the Fourier vectors $(1, \omega, \omega^2, \dots)$

Layer 3 THE WHY

Why is reducing mod 2 natural, not a trick? Go to the definition the problem itself gives you: rank is the largest k with a $k \times k$ submatrix of *nonzero* determinant. The problem has told you, in its own brackets, that the question is a nonvanishing question. An integer is nonzero as soon as it is nonzero modulo anything, and the cheapest modulus is 2. So the definition of rank, applied to an integer matrix, is an invitation to work mod 2. The exact determinant (a perfect square, being the square of a Pfaffian; it is 1 in every case I checked, $n \leq 7$) is information you were never asked for.

Why does mod 2 make this particular matrix easy? Because all the difficulty in M_n is in the *signs*: which entries are $+1$ and which -1 is what the long description of subdiagonals is about. Mod 2, $+1 = -1$, so the description collapses to "zero on the diagonal, one elsewhere". That matrix, $J - I$, is the most symmetric matrix there is, and its eigenvalues are forced by the definition of J: J has rank one, its image is spanned by $\mathbf{1}$, and $J\mathbf{1} = m\mathbf{1}$. Parity erases exactly the structure that made the problem look hard.

Why is the upper bound natural? Read the statement for the cue "the first n subdiagonals are 1, the rest -1": each row has $2n$ off-diagonal entries, split n and n. Equal numbers of $+1$ and -1 in every row is the definition of $M\mathbf{1} = \mathbf{0}$. (Skew-symmetry of odd order gives singularity too; either reason suffices, and noticing both is a check.)

Why does Kedlaya's official route use eigenvectors? Because the matrix is circulant, and the circulant structure is the deepest fact about it: the Fourier vectors diagonalise every circulant. That route gives all the eigenvalues, which is more than asked. The archive's point is cost: the parity argument proves the same bound with no trigonometry, and it transfers to non-circulant matrices, where the Fourier route does not.

Layer $+\alpha$ BACKTESTING

1. The first cue is the problem's own definition: "a $k \times k$ submatrix with nonzero determinant". Write down: *upper bound by a kernel vector; lower bound by one nonzero minor, and nonzero can be checked mod* 2.

2. Row sums give the kernel vector $\mathbf{1}$; reduction mod 2 turns the leading $2n \times 2n$ minor into $\det(J - I)$.

3. Hesitation comes after the guess $2n$, when the student starts eliminating the general matrix. Reinforce it away with: do I need this determinant's value, or only that it is nonzero?

4. The trap is the sign pattern: it invites exact elimination or an exact Pfaffian, both long and error-prone. A second trap is trying the parity trick on the full $(2n+1) \times (2n+1)$ matrix: there $\det(J - I) = 2n$ is even, as it must be, and mod 2 proves nothing. The $2n \times 2n$ block works because $2n - 1$ is odd.

5. Any time an integer quantity must merely be shown nonzero — a determinant, a count, a coefficient — try the smallest modulus first.

6. Reusable decision rule: *rank $= r$ is two separate claims; prove "$\leq r$" by exhibiting kernel vectors and "$\geq r$" by exhibiting one $r \times r$ minor that is nonzero modulo a prime. Choose the prime to erase the messiest feature of the matrix.*

1988 B6

1988 B6. Prove that there exist an infinite number of ordered pairs (a, b) of integers such that for every positive integer t, the number $at + b$ is a triangular number if and only if t is a

triangular number. (The triangular numbers are the $t_n = n(n+1)/2$ with n in $\{0, 1, 2, \dots\}$.)

Fields: number theory; quadratic forms; perfect squares **Position:** B6 — last on the paper, but short once the right translation is made; the difficulty is the "only if" direction, not the construction **Answer:** For each $k \geq 1$ the pair $(a, b) = \big((2k+1)^2,\ k(k+1)/2\big)$ works, since $8(at+b)+1 = (2k+1)^2(8t+1)$ and a non-negative integer T is triangular exactly when $8T+1$ is a perfect square. The first pair is $(9, 1)$.

Layer 1 THE WHAT — surface analysis

The natural attempt. The student tests small triangular numbers $0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, \dots$ and looks for a linear map sending them into themselves. Pattern-hunting finds

$$9t_m + 1 = \frac{9m^2+9m+2}{2} = \frac{(3m+1)(3m+2)}{2} = t_{3m+1}.$$

So $9 \cdot 1 + 1 = 10$, $9 \cdot 3 + 1 = 28$, $9 \cdot 6 + 1 = 55$: triangular in, triangular out. More generally, expanding t_{cm+k} and asking for the m-dependence to be a multiple of t_m forces $c = 2k+1$ and gives $t_{(2k+1)m+k} = (2k+1)^2 t_m + t_k$. That is infinitely many pairs for the "if" direction, and the student feels close.

Now the "only if": if $9t+1$ is triangular, must t be? Suppose $9t+1 = t_N$. The student tries to show $N \equiv 1 \pmod 3$ so that $N = 3m+1$ and then $t = t_m$. Work mod 3: $t_N \equiv 1 \pmod 9$ is needed, so $N(N+1) \equiv 2 \pmod{18}$… and the case analysis on N modulo 9 or 18 starts. It can be finished for $a = 9$ with patience, but for general k it becomes a residue analysis modulo $2(2k+1)^2$ with the answer hidden in the arithmetic. The wall is that "triangular" is not a condition you can manipulate directly; every step is a fresh computation with $N(N+1)/2$.

The model chain. *Lemma.* A non-negative integer T is triangular if and only if $8T+1$ is a perfect square.

Proof. If $T = n(n+1)/2$ then $8T+1 = 4n^2+4n+1 = (2n+1)^2$. Conversely, if $8T+1 = s^2$ with $s \geq 0$, then s^2 is odd, so $s = 2n+1$ for some integer $n \geq 0$, and $T = (s^2-1)/8 = n(n+1)/2$. □

Construction. Fix an integer $k \geq 1$ and put $a = (2k+1)^2$, $b = t_k = k(k+1)/2$. For any positive integer t, $at+b$ is a positive integer, and

$$8(at+b)+1 = 8(2k+1)^2 t + 4k^2 + 4k + 1 = (2k+1)^2(8t+1).$$

If. If t is triangular, $8t+1 = s^2$ by the lemma, so $8(at+b)+1 = \big((2k+1)s\big)^2$ and $at+b$ is triangular.

Only if. If $at+b$ is triangular, $(2k+1)^2(8t+1) = u^2$ for some integer $u \geq 0$. Then $8t+1 = \big(u/(2k+1)\big)^2$ is the square of a rational number and is an integer. A rational number whose square is an integer is itself an integer (write it in lowest terms p/q; $q^2 \mid p^2$ with $\gcd(p,q) = 1$ forces $q = 1$). So $8t+1$ is a perfect square and t is triangular.

Infinitely many. Distinct k give distinct values of a, so the pairs $\big((2k+1)^2, t_k\big)$, $k = 1, 2, 3, \dots$, are infinitely many distinct ordered pairs of integers. □

Remark. The identity $t_{(2k+1)m+k} = (2k+1)^2 t_m + t_k$ that the natural attempt discovered is the same fact: $2\big((2k+1)m+k\big)+1 = (2k+1)(2m+1)$. The square root of $8T+1$ is $2n+1$, and the map multiplies it by $2k+1$. (The pair $k = 0$ gives $(1, 0)$, which also works but is the identity; the problem's "infinite number" is not threatened either way.)

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains find $9t_m + 1 = t_{3m+1}$, or the general version. They split at the converse. (N) tries to invert the triangular map directly by residue analysis on N. (M) translates "triangular" into "$8T+1$ is a square", after which the converse is one line.

> *Triangular numbers are values of a quadratic, so completing the square should turn them into squares. $n(n+1)/2 = T$ means $(2n+1)^2 = 8T+1$. Now "triangular iff triangular" is "square iff square".*

> *What preserves squareness in both directions? Multiplying by a nonzero square: c^2m is a square exactly when m is. So I want $8(at+b)+1 = c^2(8t+1)$; comparing coefficients, $a = c^2$ and $8b+1 = c^2$, so c is odd and b is itself triangular.*

- ▷ **Cue:** a condition "$T = q(n)$ for some integer n" with q quadratic $\longrightarrow$ **Move:** complete the square: it becomes "some linear function of T is a perfect square" (here $8T+1 = (2n+1)^2$)
- ▷ **Cue:** you need a map that preserves a property in both directions ("if and only if") $\longrightarrow$ **Move:** translate the property into one that has an obvious two-way-preserving operation, then pull that operation back
- ▷ **Cue:** $c^2m = u^2$ with integers $\longrightarrow$ **Move:** m is a perfect square, because a rational whose square is an integer is an integer

Layer 3 THE WHY

Kedlaya's hint is the single sentence that an integer t is triangular if and only if $8t+1$ is a square. Here is where that comes from, and why it is not a piece of trivia you had to have memorised.

Start from the definition: T is triangular when $T = n(n+1)/2$ for some $n \geq 0$. That is "T is a value of a quadratic polynomial at an integer". The only general tool for deciding whether a number is a value of a quadratic is the one from A-level: complete the square, which turns n^2+n into $(n+\frac{1}{2})^2 - \frac{1}{4}$. Clearing denominators gives $(2n+1)^2 = 8T+1$. So the $8T+1$ criterion is not a fact to recall; it is what the quadratic formula says about $n^2+n-2T=0$ — the discriminant $1+8T$ must be a perfect square for n to be an integer. The circle example of the Blueprint has the same shape: the definition mentions a quadratic, so the object of interest is its discriminant.

Once triangular numbers are "T with $8T+1$ square", the problem asks for an affine map $t \mapsto at+b$ with

$$8t+1 \text{ square} \iff 8(at+b)+1 \text{ square}.$$

In the world of squares the two-way-preserving operation is obvious: scale by a square. Scaling by a non-square fails in one direction ($2 \cdot 8 = 16$ is square while 8 is not), and adding constants destroys squareness in both. So the search collapses to $8(at+b)+1 = c^2(8t+1)$, and matching coefficients forces $a = c^2$ and $8b+1 = c^2$. The second equation says c is odd, $c = 2k+1$, and then $b = t_k$. Nothing was guessed: the pairs are the complete family of solutions of the ansatz.

Why does the natural attempt stall on the converse? Because it works with $N(N+1)/2$, a product, where divisibility by 3 or 9 can be split between the two factors in several ways; that is what the residue case analysis is tracking. The translation replaces the product by a single square $(2N+1)^2$, where divisibility is governed by one number. Unique factorisation makes squares easy and products of consecutive integers awkward — that is the structural reason the square form is the right coordinate.

The "only if" also has a rigour point worth naming: $c^2m = u^2$ gives $m = (u/c)^2$, a rational square, and you need the step that an integer which is a rational square is an integer square. Without it the proof has a hole, however obvious it feels.

Layer $+\alpha$ BACKTESTING

1. The first cue is the definition $t_n = n(n+1)/2$, a quadratic in n. On seeing it write $8t_n+1 = (2n+1)^2$.
2. Translate triangularity into squareness of $8T+1$; look for an affine map that becomes multiplication by a square.

3. Hesitation comes after $9t_m + 1 = t_{3m+1}$ is found and the converse refuses to yield to residues. Reinforce it away with: an "if and only if" about a property is easiest where the property has a two-way-preserving operation — find that coordinate first.

4. The trap is proving only the "if" direction and believing the problem finished, or grinding the converse with congruences on N. A smaller trap is omitting the step that a rational whose square is an integer is an integer.

5. Any condition "is a value of a quadratic at an integer" should be rewritten at once as "a certain discriminant is a perfect square".

6. Reusable decision rule: *before constructing maps that preserve a property, change coordinates so the property becomes as simple as possible (here: "is a perfect square"); then the preserving maps are the obvious ones in the new coordinates, and both directions of the "iff" come together.*