DAWN PUTNAM SESSION

Putnam 2005 — Papers A and B

Putnam 120 Archive

Contents

Putnam 2005 — Paper A **5**
2005 A1 5
2005 A2 7
2005 A3 9
2005 A4 12
2005 A5 14
2005 A6 16

Putnam 2005 — Paper B **21**
2005 B1 21
2005 B2 23
2005 B3 25
2005 B4 28
2005 B5 31
2005 B6 34

Putnam 2005 — Paper A

2005 A1

2005 A1. Show that every positive integer is a sum of one or more numbers of the form 2^r3^s, where r and s are nonnegative integers and no summand divides another. (For example, $23 = 9 + 8 + 6$.)

Fields: elementary number theory, strong induction **Position:** A1 — the opener **Answer:** Every n admits such a representation; built by strong induction, doubling a representation of $n/2$ when n is even, and combining the largest power of 3 not exceeding n with a doubled representation of the remainder when n is odd.

Layer 1 THE WHAT — surface analysis

The natural attempt. The example $23 = 9 + 8 + 6$ looks greedy: $9 = 3^2$ is the largest 3-smooth-with-this-shape number not exceeding 23 that leaves a workable remainder, and the pattern invites a general greedy algorithm — repeatedly subtract the largest number of the form 2^r3^s not exceeding what remains, and hope the "no summand divides another" condition falls out for free.

Test it. Take $n = 17$. The largest number of the form 2^r3^s not exceeding 17 is $16 = 2^4$. That leaves $17 - 16 = 1 = 2^03^0$. So greedy produces $17 = 16 + 1$. But 1 divides every positive integer, including 16 — the representation is disqualified on the very first case checked outside the worked example. Patching this by forbidding 1 as a summand doesn't fix the underlying problem: greedy has no mechanism that prevents the eventual remainder from being a divisor (or multiple) of an earlier term, because nothing about "take the largest term available" tracks divisibility at all. It tracks size.

A second attempt: strengthen greedy by also disallowing any move that would let the new remainder divide (or be divided by) an already-chosen term. This can be made to work case by case, but now every step requires an auxiliary check against the entire history of the sum so far, and there is no visible reason the check always succeeds — the argument would have to be redone, by exhaustive casework, for every residue of n modulo small powers of 2 and 3. This is workable for a computer but not for a clean induction: it costs a growing amount of bookkeeping precisely where the model chain below costs none.

The model chain. Induct on n, assuming every positive integer less than n has a valid representation. Base case $n = 1 = 2^03^0$.

n **even.** By the induction hypothesis $n/2$ has a representation $n/2 = t_1 + \cdots + t_k$ with no t_i dividing another. Then $n = 2t_1 + \cdots + 2t_k$ is again a valid representation: each $2t_i$ still has the required form $2^{r+1}3^s$, and for positive integers a, b, $a \mid b \iff 2a \mid 2b$, so uniformly doubling every summand preserves "no summand divides another" exactly.

n **odd.** Let $m = \lfloor \log_3 n \rfloor$, so $3^m \le n < 3^{m+1}$. If $3^m = n$, done. Otherwise $n - 3^m$ is even and positive (both n and 3^m are odd), so by the induction hypothesis $(n - 3^m)/2 = s_1 + \cdots + s_k$ has a valid representation, with $(n - 3^m)/2 < n$ so the hypothesis applies. Set

$$n = 3^m + 2s_1 + \cdots + 2s_k.$$

The $2s_i$ pairwise avoid dividing each other, by the same doubling argument as above. It remains to check 3^m against each $2s_i$: since $2s_i$ is even and 3^m is odd, $2s_i$ can never divide the odd number 3^m. For the reverse, note $s_i \leq \sum_j s_j = (n - 3^m)/2$ (all $s_j > 0$), so $2s_i \leq n - 3^m < 3^{m+1} - 3^m = 2 \cdot 3^m$, giving $s_i < 3^m$; since $\gcd(3^m, 2) = 1$, $3^m \mid 2s_i \iff 3^m \mid s_i$, which is impossible for $0 < s_i < 3^m$. So 3^m divides none of the $2s_i$, completing the induction.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains recognise that the size of the summands must be controlled, not just their sum. They split on *how* that control is achieved: the natural attempt tries to police divisibility directly, term by term, against a growing list of committed choices. The model chain never checks divisibility against a list at all — it builds the whole new representation from a smaller, already-trusted one by a single global operation (double everything, or double everything and add one odd term), and lets a structural fact about that operation do the checking in advance.

> *If I halve the problem down to $n/2$ first and only double back up at the very end, the "does one term divide another" question about the new summands is already answered by the old ones — because for positive integers, $a \mid b$ exactly when $2a \mid 2b$. Divisibility is unchanged by uniformly rescaling every term by the same factor. So I never need to compare the new terms to each other directly; I only need to trust the smaller case and rescale it.*

> *For the odd case, if every summand I add besides 3^m is forced to be even, then 3^m — being odd — can never divide any of them, and none of them can ever divide the odd 3^m. The parity split does the "no summand divides another" work automatically, provided I also keep every new summand strictly below 3^m in size, which the choice of m as the* largest *power of 3 not exceeding n guarantees for free.*

Layer 3 THE WHY

Why does uniform doubling preserve "no summand divides another," while greedy subtraction does not? Because divisibility is a multiplicative relation: $a \mid b$ means $b = ka$ for some positive integer k, and this is unaffected by multiplying both a and b by the same constant. Greedy subtraction has no such invariance — the remainder after subtracting the largest term is an arbitrary leftover with no forced relationship to what was already chosen, so nothing prevents it (or a later greedy step) from dividing an earlier term. The model chain's induction, by contrast, is built entirely out of operations that are automatically divisibility-safe: doubling (safe by the identity above) and appending one fixed odd term (safe by the parity split, once the size bound $s_i < 3^m$ is in hand).

The size bound $s_i < 3^m$ is itself not incidental: it is exactly what "$m = \lfloor \log_3 n \rfloor$" buys. Taking the *largest* power of 3 not exceeding n, rather than any power of 3, forces $n - 3^m < 3^{m+1} - 3^m = 2 \cdot 3^m$, which is precisely the inequality that keeps every new summand below 3^m. A smaller choice of power would not close the argument; the maximality is load-bearing, not a convenience.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the phrase "no summand divides another" appearing as a *condition on the sum*, not merely "find a sum." On seeing it, write down: divisibility is preserved under uniform scaling, so any construction built by rescaling an already-valid smaller sum inherits the condition automatically.
2. That cue suggests strong induction with a case split on parity of n: halve-and-double for even n, peel off the largest available power of 3 for odd n (since 3 is the odd prime present in the allowed summand shape $2^r 3^s$).

3. Hesitation occurs right after finding the worked example's pattern $23 = 9 + 8 + 6$, where the pull is to generalise it as a greedy algorithm. This is reinforced away by testing greedy on one adversarial case ($n = 17$) before trusting it — a single counterexample where the leftover is 1 exposes that greedy has no divisibility safeguard built in.

4. The trap is using $1 = 2^0 3^0$ as a leftover summand without checking that it divides everything else in the sum — an easy oversight since 1 satisfies the "form $2^r 3^s$" requirement trivially.

5. Recognise this family faster by treating "no summand divides another" as a signal to look for a scaling operation under which the constraint is automatically invariant, rather than a signal to build a checklist.

6. Reusable rule: when a sum-decomposition problem constrains a relation *between* the summands (divisibility, coprimality, distinctness under some operation), search first for an operation — scaling, shifting, a fixed involution — under which that relation is provably invariant, and build the induction out of that operation. Case-by-case checking against a growing history is the fallback, not the plan.

▷ **Cue:** a decomposition condition constrains a relation between summands (e.g. no summand divides another) $\longrightarrow$ **Move:** find an operation (uniform scaling, parity split) under which that relation is automatically invariant, and induct using only that operation

▷ **Cue:** a worked example suggests a greedy algorithm for a decomposition problem with a between-summand constraint $\longrightarrow$ **Move:** stress-test greedy on one adversarial small case before trusting it — greedy optimises size, not the relational constraint

2005 A2

2005 A2. Let $\mathbf{S} = \{(a, b) \mid a = 1, 2, \ldots, n,\ b = 1, 2, 3\}$. A *rook tour* of $\mathbf{S}$ is a polygonal path made up of line segments connecting points $p_1, p_2, \ldots, p_{3n}$ in sequence such that (i) $p_i \in \mathbf{S}$, (ii) p_i and p_{i+1} are a unit distance apart, for $1 \leq i < 3n$, (iii) for each $p \in \mathbf{S}$ there is a unique i such that $p_i = p$. How many rook tours are there that begin at $(1, 1)$ and end at $(n, 1)$?

Fields: combinatorics, bijective counting **Position:** A2 **Answer:** 2^{n-2} for $n \geq 2$ (and 0 for $n = 1$, since $(1, 1) = (n, 1)$ then, and start and end cannot coincide in a path of 3 distinct points).

Layer 1 THE WHAT — surface analysis

The natural attempt. A rook tour is a Hamiltonian path in the $3 \times n$ grid graph from one corner of the bottom row to the other. The most direct way to attack a Hamiltonian-path count with fixed endpoints is a recursion in n: peel off the last column and see what's left.

Let A_n be the set of tours from $(1, 1)$ to $(n, 1)$, and B_n the set of tours from $(1, 1)$ to $(n, 3)$ — a second family is unavoidable, because a tour arriving at column n must have entered it through either $(n-1, 1)$-$(n, 1)$ or $(n, 2)$-$(n, 1)$, and tracing the forced continuation in each case lands the "problem minus a column" in A_{n-1} in one case and B_{n-1} in the other. Working through both cases (and the mirror argument for B_n) gives, after checking the moves are reversible,

$$|A_n| = |B_n| = 2^{n-2}\left(|A_1| + |B_1|\right), \qquad n \geq 2,$$

with $|A_1| = 0$ (no path of length 3 from $(1, 1)$ back to itself) and $|B_1| = 1$ (the single path $(1, 1)$-$(1, 2)$-$(1, 3)$). This gives $|A_n| = 2^{n-2}$, which is correct — but it took introducing a second, symmetric family B_n purely to make the recursion close, then verifying two mutually-referencing bijections (one for A_n, one for B_n) before the single number $|A_n|$ could be extracted. The bookkeeping doubled to solve a problem that has one asked-for quantity.

The model chain. Encode a tour P by the set $S(P) \subseteq \{1, \ldots, n\}$ of columns at which the path uses a step directly between row 1 or row 3 and row 2 in the specific direction that ends its middle-row visit there — concretely, $i \in S(P)$ exactly when the path contains the directed edge

$(i,1) \to (i,2)$ or $(i,3) \to (i,2)$, read in the direction the tour actually traverses it. Since the tour ends at $(n,1)$, the last step into $(n,1)$ forces $n \in S(P)$ always. Since the tour has $3n$ points and alternates its "column class" in a fixed pattern dictated by the grid's bipartite structure, $S(P)$ is forced to have even size.

Conversely, every subset $S = \{a_1 < a_2 < \cdots < a_{2r} = n\} \subseteq \{1, \ldots, n\}$ of even size containing n determines a unique tour: build the path by visiting the columns in order, alternating which of row 1/row 3 is entered first each time S "switches," with the membership pattern of S pinning down, column by column, exactly one legal way to thread the middle row. This is a bijection between tours and even-size subsets of $\{1, \ldots, n\}$ containing n. The number of such subsets is 2^{n-2} (choose any subset of the remaining $n-1$ elements, discard the odd-size half by symmetry — or directly, $\sum_{k \text{ even}} \binom{n-1}{k} = 2^{n-2}$), giving the count immediately, in one pass, with no recursion to solve.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains agree the object is a Hamiltonian path on a width-3 strip with fixed endpoints. They split on the very first question asked about it. The natural attempt asks: *what happens at the last column, and what problem is left over?* — an $n \to n-1$ question, which forces a second family into existence to close the recursion. The model chain asks instead: *what single piece of data, read off the whole path in one pass, determines it completely?* — an all-at-once question, which turns the count into a subset count with no recursion at all.

> *Instead of asking how the tour behaves at the last column and setting up a recursion in n, I should ask what minimal combinatorial data, read column by column from the whole tour, would let me reconstruct it exactly. If I can name that data and show it ranges freely over a set I can already count, I never need to solve a recursion — I just count the data directly.*

Layer 3 THE WHY

Why does the column-by-column encoding beat the peel-a-column recursion? Because a Hamiltonian path on a width-3 strip has very little freedom *locally*: at each column, the path either threads straight through the middle row on its way past, or it makes a "turn" that uses the middle-row vertex as a hinge between two separate visits to that column. That is a single binary distinction, repeated once per column, and the two families A_n, B_n in the natural attempt were implicitly tracking exactly this binary choice but doing so through the lens of "where does the path end up," which requires tracing the whole tail of the path to know. Reading the same binary choice directly off each column — is this a turning column or not — needs no lookahead and no second family; it is a local, checkable fact about the tour at that column alone. The forced conditions "$n \in S(P)$" and "$|S(P)|$ even" are then just the two boundary/parity constraints coming from the path having to start and end where the problem specifies, not extra casework.

Caveat. The forward direction of the encoding (every tour determines a valid even-size subset containing n) is straightforward from the structure above. The reverse direction — that every even-size subset containing n determines a *unique* legal tour, with no further degrees of freedom hidden in "threading the middle row" — is the part of Kedlaya, Bhargava, and Ng's official first solution that is stated tersely rather than proved in full. I have verified this independently at $n = 1$ (0 tours, matching the empty case), $n = 2$ (exactly 1 tour, checked by direct hand enumeration of the 3×2 grid) and $n = 3$ (the recursion above independently gives $2^{3-2} = 2$, matching the 2 even-size subsets of $\{1,2,3\}$ containing 3, namely $\{1,3\}$ and $\{2,3\}$), and the recursion and the bijection agree in general on the closed form 2^{n-2}. I have not independently re-derived the general injectivity/surjectivity of the column-encoding map beyond these checks and the structural sketch above, so I present it as verified by consistency across two independent proofs plus small cases, not as re-derived

line by line from first principles.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is "$3 \times n$ grid, unit-step path visiting every point exactly once, fixed start and end." On seeing it, write down: this is a Hamiltonian-path count on a thin strip graph, and thin-strip Hamiltonian paths usually have very restricted local behaviour column by column.
2. That cue suggests two live options: set up a recursion in n by classifying the tour's behaviour near one end, or look for a local per-column encoding that turns the whole path into a labelled sequence to count directly.
3. Hesitation sets in once the peel-a-column recursion is found to work: it is correct and gives the right answer, so the pull is to stop there. This is reinforced away by asking explicitly whether the two-family bookkeeping (A_n and B_n) is really tracking one piece of local information that a direct encoding could capture without a recursion.
4. The trap is treating "I found a working recursion" as license to stop looking for a cheaper method — a working method is not automatically the cheapest one, and grid/path counting problems routinely admit column-by-column or bit-string encodings that collapse a two-term recursion into a one-line subset count.
5. Recognise this family faster by treating any Hamiltonian-path count on a width-limited (ladder-like) grid as a standing invitation to look for a local, per-column binary encoding before reaching for a multi-family recursion.
6. Reusable rule: when counting Hamiltonian paths on a thin grid strip with fixed endpoints, first ask whether the path's behaviour at each column reduces to one binary local choice (straight-through vs. turn); if so, encode the whole path as the set/sequence of those choices, identify the boundary constraints forced by the fixed endpoints, and count the resulting combinatorial object directly, before setting up an $n \to n-1$ recursion.

▷ **Cue:** Hamiltonian-path count on a width-3 (or width-k) grid strip with fixed endpoints $\longrightarrow$ **Move:** look for a column-wise local-shape encoding (straight vs. turn) turning the path into a subset or sequence to count directly, before setting up a peel-one-column recursion

▷ **Cue:** a working recursion has been found for a counting problem but required introducing a second, symmetric family to close $\longrightarrow$ **Move:** check whether the second family was only tracking one local binary choice that a direct encoding could capture without a recursion

2005 A3

2005 A3. Let $p(z)$ be a polynomial of degree n all of whose zeros have absolute value 1 in the complex plane. Put $g(z) = p(z)/z^{n/2}$. Show that all zeros of $g'(z) = 0$ have absolute value 1.

Fields: complex analysis, polynomials **Position:** A3 **Answer:** Writing $p(z) = c\prod_{j=1}^{n}(z - r_j)$ with $|r_j| = 1$, one shows $\mathrm{Re}\left(\dfrac{zg'(z)}{g(z)}\right) = \dfrac{|z|^2 - 1}{2}\displaystyle\sum_{j=1}^{n}\frac{1}{|z - r_j|^2}$, which is zero only when $|z| = 1$ (or trivially, when z is itself a root).

Layer 1 THE WHAT — surface analysis

The natural attempt. Differentiate directly. Writing $g(z) = p(z)z^{-n/2}$ and using the product rule,

$$g'(z) = p'(z)z^{-n/2} - \tfrac{n}{2}p(z)z^{-n/2-1} = z^{-n/2-1}\left(zp'(z) - \tfrac{n}{2}p(z)\right).$$

So, away from $z = 0$, $g'(z) = 0 \iff zp'(z) = \frac{n}{2}p(z)$. Writing $p(z) = c\prod_j (z - r_j)$ and taking the logarithmic derivative, $p'(z)/p(z) = \sum_j 1/(z - r_j)$, so the target equation becomes

$$\sum_{j=1}^{n} \frac{z}{z - r_j} = \frac{n}{2}.$$

This is a genuinely useful reduction — both chains reach it — but a strong student's next instinct is to bound the left side directly. Write $\frac{z}{z-r_j} = 1 + \frac{r_j}{z-r_j}$, so the equation says $\sum_j \frac{r_j}{z-r_j} = -\frac{n}{2}$, and try the triangle inequality: $\left|\sum_j \frac{r_j}{z-r_j}\right| \le \sum_j \frac{1}{|z-r_j|}$ (using $|r_j| = 1$). For this to equal $n/2$ forces some constraint relating $|z - r_j|$ to 1 on average — but turning "an inequality on a sum of moduli" into the exact statement $|z| = 1$ requires knowing when the inequality is tight, which depends on all the r_j having aligned phases relative to z, a condition that is not obviously equivalent to $|z| = 1$ and would need to be chased down separately for $|z| > 1$ and $|z| < 1$. The bound is honest work, but it produces a region, not the sharp curve the problem asks for, and closing the gap from "inequality" to "equality of $|z|$" this way is not clearly shorter than starting over.

The model chain. Take the same reduction $g'(z) = 0 \iff zp'(z) = \frac{n}{2}p(z)$ (for $z \neq 0$, $z \neq r_j$), but instead of bounding moduli, symmetrize. Since

$$\frac{zg'(z)}{g(z)} = \frac{zp'(z)}{p(z)} - \frac{n}{2} = \sum_{j=1}^{n} \frac{z}{z - r_j} - \frac{n}{2} = \sum_{j=1}^{n} \left(\frac{1}{2} + \frac{r_j}{z - r_j}\right) = \frac{1}{2}\sum_{j=1}^{n} \frac{z + r_j}{z - r_j},$$

the target equation $zg'(z)/g(z) = 0$ is a statement about a sum of terms of the shape $(z + r)/(z - r)$ with $|r| = 1$. For any complex $r \neq z$,

$$\frac{z + r}{z - r} = \frac{(z + r)(\bar{z} - \bar{r})}{|z - r|^2} = \frac{|z|^2 - |r|^2 + (r\bar{z} - z\bar{r})}{|z - r|^2},$$

and $r\bar{z} - z\bar{r} = 2i\,\mathrm{Im}(r\bar{z})$ is purely imaginary, so

$$\mathrm{Re}\left(\frac{z + r}{z - r}\right) = \frac{|z|^2 - |r|^2}{|z - r|^2}.$$

With $|r_j| = 1$ for every j,

$$\mathrm{Re}\left(\frac{zg'(z)}{g(z)}\right) = \frac{1}{2}\sum_{j=1}^{n} \frac{|z|^2 - 1}{|z - r_j|^2} = \frac{|z|^2 - 1}{2}\sum_{j=1}^{n} \frac{1}{|z - r_j|^2}.$$

If $z \neq r_j$ for every j, the sum $\sum_j 1/|z - r_j|^2$ is a sum of strictly positive real terms, hence strictly positive. So the real part above vanishes if and only if $|z|^2 = 1$. Since $g'(z) = 0$ (with $g(z) \neq 0$) forces $zg'(z)/g(z) = 0$, in particular its real part is 0, so $|z| = 1$. The one remaining case, $z = r_j$ for some j, has $|z| = |r_j| = 1$ automatically. So every zero of g' has $|z| = 1$, in one computation, with no case split on the sign of $|z| - 1$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains reach the identical shared equation $\sum_{j=1}^{n} \frac{z}{z - r_j} = \frac{n}{2}$ — that is where the differentiation and the logarithmic-derivative rewrite end. They split immediately after: the natural attempt applies the triangle inequality to bound the modulus of this sum, aiming to squeeze out a region containing the zeros of g'. The model chain instead rewrites the same sum in the symmetrized form $\frac{1}{2}\sum_j (z + r_j)/(z - r_j)$ and takes a real part, aiming to force an exact equality.

> *I don't want an inequality on $\left|\sum_j z/(z-r_j)\right|$; I want a statement about the real part of something, because "a real part is exactly zero" is the mechanism that pins $|z|$ to precisely 1, not merely into a region. Since every r_j has $|r_j| = 1$, I should rewrite each term as $(z+r_j)/(z-r_j)$, because I already know $\mathrm{Re}[(z+r)/(z-r)]$ has the clean closed form $(|z|^2-|r|^2)/|z-r|^2$ — the modulus condition $|r_j| = 1$ then becomes a literal algebraic cancellation inside that formula, not something I have to estimate.*

Layer 3 THE WHY

Why does the real-part identity succeed where the triangle-inequality bound stalls? Because $\mathrm{Re}[(z+r)/(z-r)] = (|z|^2-|r|^2)/|z-r|^2$ is not a bound but an equality: it converts a comparison of moduli ($|z|$ versus $|r|$) into the sign of a single real number, with no approximation lost along the way. The triangle inequality, by contrast, is lossy by design — it discards the phases of the r_j relative to z, and phase is exactly the information needed to know whether the sum's terms reinforce or cancel. The problem's hypothesis is stated purely in terms of modulus ($|r_j| = 1$ for every j), so the right tool is one that consumes a modulus hypothesis and returns an exact statement, not one that consumes it and returns only a bound.

This is also why the substitution $g(z) = p(z)/z^{n/2}$ is not an arbitrary normalization: dividing by $z^{n/2}$ symmetrizes the logarithmic derivative so that zg'/g becomes exactly the sum $\frac{1}{2}\sum_j (z+r_j)/(z-r_j)$ — a form built entirely from the shape whose real part is a clean function of $|z|^2-|r_j|^2$. Without the $z^{-n/2}$ factor, the analogous computation for p alone gives $zp'(z)/p(z) = \sum_j z/(z-r_j)$, whose real part does not collapse into a single sign-definite quantity in $|z|$; the $n/2$ power is precisely the correction that turns a Gauss–Lucas-type inclusion (critical points lie somewhere in the convex hull of the roots) into an equality (critical points lie exactly on the circle the roots lie on).

Layer $+\alpha$ BACKTESTING

1. The first useful cue is "$g(z) = p(z)/z^{n/2}$" appearing together with "all roots of p have $|z| = 1$." On seeing the half-integer power specifically, write down: this is designed to interact with a logarithmic derivative, since $z^{-n/2}$ turns into a symmetric $-\frac{n}{2}\log z$ term under $\log g$.

2. That cue suggests passing to $g'/g = p'/p - n/(2z)$, writing p'/p as $\sum_j 1/(z-r_j)$ via the root factorization, and multiplying through by z to reach $\sum_j z/(z-r_j) = n/2$.

3. Hesitation occurs right after reaching that shared equation, where the pull is to bound $\left|\sum_j z/(z-r_j)\right|$ by the triangle inequality, since that is the default move on a sum of fractions. It is reinforced away by asking: does the target conclusion want an inequality or an exact equality of $|z|$? An exact equality wants a vanishing real (or imaginary) part, not a bounded modulus.

4. The trap is forgetting to check the case $z = r_j$ separately, since the logarithmic derivative p'/p has a pole exactly there and the sum-of-fractions manipulation silently assumes $z \neq r_j$ throughout — but the case is free, since $|z| = |r_j| = 1$ automatically once z is a root.

5. Recognise this pattern faster: whenever a hypothesis is "all the relevant points lie on a fixed circle $|w| = R$" and the target conclusion is "so does this related point, exactly," reach for $\mathrm{Re}[(z+r)/(z-r)] = (|z|^2-|r|^2)/|z-r|^2$ before reaching for a triangle-inequality bound.

6. Reusable rule: to prove a modulus-equality target ($|z| = R$) from a hypothesis that a family of parameters all satisfy $|r_j| = R$, and the target arises as the vanishing of a sum of terms each depending on one r_j, symmetrize each term into the form $(z+r_j)/(z-r_j)$ and take real parts — this converts the modulus hypothesis into an exact algebraic cancellation, whereas a triangle-inequality bound on the same sum only ever returns a region.

▷ **Cue:** $g(z) = p(z)/z^k$ with all roots of p on a fixed circle $|z| = R$, target: zeros of g' also on $|z| = R$ ⟶ **Move:** pass to the logarithmic derivative $g'/g = p'/p - k/z$, expand $p'/p = \sum_j 1/(z-r_j)$, multiply by z, and symmetrize into $\sum_j (z+r_j)/(z-r_j)$

▷ **Cue:** need to show a modulus is exactly R (not merely bounded), arising from a sum of terms parametrized by points of known modulus R $\longrightarrow$ **Move:** use $\mathrm{Re}[(z+r)/(z-r)] = (|z|^2-|r|^2)/|z-r|^2$ and set the total real part to zero, rather than bounding the sum by the triangle inequality

2005 A4

2005 A4. Let H be an $n \times n$ matrix all of whose entries are ± 1 and whose rows are mutually orthogonal. Suppose H has an $a \times b$ submatrix whose entries are all 1. Show that $ab \leq n$.

Fields: linear algebra (inner products, Hadamard matrices); inequalities **Position:** A4 — mid-paper, but short once the right object is written down. The whole difficulty is choosing which vector to take the length of. **Answer:** A proof: the sum v of the a rows meeting the all-ones block has $|v|^2 = an$ by orthogonality and $|v|^2 \geq a^2 b$ from the block, so $ab \leq n$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You read "± 1 entries, orthogonal rows" and your mind goes to Hadamard matrices: $HH^T = nI$, n is 1, 2 or a multiple of 4, Sylvester's construction. You write down $HH^T = nI$ and try to use it on the block. Permute rows and columns (orthogonality survives both) so that the all-ones block sits in the top-left corner:

$$H = \begin{pmatrix} J_{a\times b} & M \\ * & * \end{pmatrix},$$

where J is all ones and M is $a \times (n-b)$.

Now you look at two rows r_i, r_j among the first a. Their dot product is b (from the block) plus $m_i \cdot m_j$ (from M), and it must be 0. So $m_i \cdot m_j = -b$ for $i \neq j$. Each m_i has length squared $n-b$. You try to squeeze an inequality out of a single pair: by Cauchy–Schwarz, $|m_i \cdot m_j| \leq n-b$, so $b \leq n-b$, i.e. $2b \leq n$. That is true but it is the case $a = 2$ only; it does not see a at all.

So you try three rows, four rows, looking for a pattern. You start counting sign agreements column by column in M: in each column of M, among the a rows, say k entries are $+1$ and $a-k$ are -1. Summing $m_i \cdot m_j$ over pairs… and here the bookkeeping gets heavy. You are, in effect, rediscovering the model argument by hand, one column at a time, and it is easy to get lost before you see that the pair sum collapses.

The model chain. Let $r_1, \ldots, r_a \in \mathbb{R}^n$ be the a rows of H that contain the all-ones submatrix, and let C be the set of its b columns. The hypotheses say:

- $r_i \cdot r_i = n$ (each row has n entries, each ± 1, each squaring to 1);
- $r_i \cdot r_j = 0$ for $i \neq j$ (orthogonality);
- $(r_i)_c = 1$ for every $i \leq a$ and every $c \in C$.

Put $v = r_1 + \cdots + r_a$. Compute $|v|^2$ in two ways.

Algebraically. Expanding the dot product bilinearly,

$$|v|^2 = \sum_{i=1}^{a}\sum_{j=1}^{a} r_i \cdot r_j = \sum_{i=1}^{a} r_i \cdot r_i + \sum_{i\neq j} r_i \cdot r_j = an + 0 = an.$$

Coordinatewise. For $c \in C$, $v_c = \sum_i (r_i)_c = a$. Every coordinate contributes a non-negative amount to $|v|^2 = \sum_c v_c^2$, so discarding the coordinates outside C,

$$|v|^2 \;\geq\; \sum_{c\in C} v_c^2 = b\,a^2.$$

Combining, $an \geq a^2 b$. Since $a \geq 1$, divide by a: $n \geq ab$. ■

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains start in the same place: the a rows through the block, and the fact that their pairwise dot products vanish. They split at (†): "I have $\binom{a}{2}$ equations $r_i \cdot r_j = 0$." The natural attempt uses them *one pair at a time*, which throws away the fact that there are many of them. The model *adds them all up* — and the only natural way to add up all pairwise dot products of a family of vectors is to take the length of their sum.

> *I've got a vectors that are pairwise orthogonal and each of length $\sqrt{n}$. The single cleanest fact about such a family is Pythagoras: the sum has squared length an. And the block tells me the sum has b coordinates each equal to a. So the sum is the vector to look at.*

> *One pair gives me $2b \leq n$ and never mentions a. The bound I want is linear in a, so I need an identity that uses all a rows at once, not a pair.*

Layer 3 THE WHY

Kedlaya's first solution opens with "choose a set of a rows…Hence $\sum_{i,j} r_i \cdot r_j = an$." It does not say why the double sum is the thing to form. Here is where it comes from.

Start from what the hypothesis *is*. "Rows mutually orthogonal" is information about dot products of pairs. "All entries ± 1" is information about each row's length: $|r_i|^2 = n$. So everything the problem gives you lives in the Gram data $r_i \cdot r_j$. The conclusion $ab \leq n$ has to be squeezed out of that Gram data plus the block.

Now ask: what single number is built from the Gram data of a vectors? The squared length of their sum, $\left|\sum r_i\right|^2 = \sum_{i,j} r_i \cdot r_j$. That is not a trick; it is Pythagoras in n dimensions. If you learned "for perpendicular vectors, lengths squared add" for two vectors in the plane, this problem is asking you to apply that sentence to a vectors in $\mathbb{R}^n$. Mutually orthogonal vectors add like legs of a right-angled box: $|r_1 + \cdots + r_a|^2 = |r_1|^2 + \cdots + |r_a|^2 = an$.

Why that sum and not some other combination? Because the block is a statement about *the same coordinates of every one of these rows being* $+1$. On those b coordinates, adding the rows makes the entries pile up to a rather than cancel. Everywhere else you know nothing, but you do not need to: squares are non-negative, so ignoring the rest only makes the length smaller. The sum is the one combination on which the block gives a large, clean lower bound, and on which orthogonality gives an exact value. An upper value meeting a lower bound is an inequality; that is the whole proof.

Look at why the pair-at-a-time attempt stalls. Two orthogonal rows through the block give $m_i \cdot m_j = -b$ with $|m_i|^2 = n - b$, and Cauchy–Schwarz gives $b \leq n - b$. But the m_i are a vectors whose pairwise dot products are all *equally negative*. You cannot have many vectors all pointing "away" from each other by a fixed amount: summing, $0 \leq |m_1 + \cdots + m_a|^2 = a(n-b) - a(a-1)b = a(n-ab)$. That is Stanley's second solution in disguise, and it is the same Pythagoras step applied to the right-hand part of the rows. Either way, the information lives in the sum.

The level of "of course" to reach: *orthogonal ⇒ take the sum, because Pythagoras is the theorem about sums of orthogonal vectors.*

Layer $+\alpha$ BACKTESTING

1. The first cue is "rows are mutually orthogonal". On seeing it, write $r_i \cdot r_j = n\,\delta_{ij}$ immediately — lengths included, since ± 1 entries fix $|r_i|^2 = n$.

2. Orthogonality of a family suggests Pythagoras: compute the squared length of the sum of the relevant vectors, once exactly (by orthogonality) and once by a crude bound (by coordinates).

3. Hesitation occurs after the one-pair calculation gives $2b \le n$: you have a correct but weak inequality and no idea how to bring in a. What reinforces it away is the question "which of my facts mentions a?" — only the count of rows, so the argument must use all a rows together.

4. The trap is the Hadamard-matrix knowledge. Knowing that $n \in \{1, 2\} \cup 4\mathbb{Z}$, Sylvester's construction, $HH^T = nI$ and so on invites you to reach for structure theory, which is irrelevant. Only the Gram data of a rows is used; the other $n - a$ rows never appear.

5. Recognise faster by translating the conclusion: $ab \le n$ is "$a^2 b \le an$", a sum of b squares of size a^2 against a lengths squared of size n. A quantity that looks like (count) $\times$ (square) on one side and (count) $\times$ (length^2) on the other is a squared-length comparison.

6. Decision rule: when you are given many pairwise-orthogonal (or pairwise-negatively-correlated) vectors and must bound how many there are or how aligned they can be, take the squared length of their sum and compare the exact expansion with a coordinate-wise lower bound. More generally: many pairwise constraints are added up by forming the square of a sum.

▷ **Cue:** rows (or vectors) given as mutually orthogonal, with a bound to prove involving how many of them there are $\longrightarrow$ **Move:** let v be the sum of the vectors; $|v|^2 = \sum |r_i|^2$ exactly by Pythagoras

▷ **Cue:** a block of coordinates on which all the chosen vectors agree $\longrightarrow$ **Move:** bound $|v|^2$ below by the contribution of those coordinates alone; the rest are squares and can be dropped

▷ **Cue:** a family of vectors whose pairwise dot products are all equal and negative $\longrightarrow$ **Move:** $0 \le |\sum v_i|^2$ bounds the size of the family

▷ **Cue:** one-pair Cauchy–Schwarz gives a bound missing a parameter $\longrightarrow$ **Move:** the missing parameter is a count, so sum over the whole family instead of a pair

2005 A5

2005 A5. Evaluate $\int_0^1 \frac{\ln(x+1)}{x^2+1}\,dx$.

Fields: real analysis (definite integrals); substitution and symmetry **Position:** A5 — late in the paper, but a one-idea problem: it is hard only for a solver who insists on an antiderivative. **Answer:** $\displaystyle\int_0^1 \frac{\ln(x+1)}{x^2+1}\,dx = \frac{\pi \ln 2}{8}$.

Layer 1 THE WHAT — surface analysis

The natural attempt. You look for an antiderivative. There is a ln and a $\frac{1}{1+x^2}$, so integration by parts is the reflex, with $u = \ln(1+x)$ and $dv = \frac{dx}{1+x^2}$:

$$I = \Big[\ln(1+x)\arctan x\Big]_0^1 - \int_0^1 \frac{\arctan x}{1+x}\,dx = \frac{\pi}{4}\ln 2 - \int_0^1 \frac{\arctan x}{1+x}\,dx.$$

The new integral is no friendlier than the old one. You try parts the other way round and get back where you started. You try partial fractions on $\frac{1}{1+x^2}$ over $\mathbb{C}$, getting $\frac{1}{2i}\left(\frac{1}{x-i} - \frac{1}{x+i}\right)$, and are then facing $\int \frac{\ln(1+x)}{x-i}\,dx$, which is a dilogarithm. You try a series, $\ln(1+x) = \sum_{k\ge1}(-1)^{k-1}x^k/k$, and are left with the moments $J_k = \int_0^1 \frac{x^k}{1+x^2}dx$, which satisfy $J_k + J_{k+2} = \frac{1}{k+1}$; this can be pushed through (Rusin's fourth official solution does exactly that), but it takes a page of alternating-series bookkeeping and a convergence argument.

This is the wall: the integrand has no elementary antiderivative. Every method that hunts for F with $F' = \frac{\ln(1+x)}{1+x^2}$ is doomed or very long. Note the useful by-product of the parts attempt, though: the number $\frac{\pi}{4}\ln 2$ has appeared. Keep it in mind.

The model chain. Step 1: put the $\frac{dx}{1+x^2}$ to work. Substitute $x = \tan\theta$, so $dx = \sec^2\theta\, d\theta$ and $1 + x^2 = \sec^2\theta$. The limits $x = 0, 1$ become $\theta = 0, \frac{\pi}{4}$:

$$I = \int_0^{\pi/4} \ln(1 + \tan\theta)\, d\theta.$$

Step 2: reflect the interval. For any integrable f on $[0, c]$, the substitution $\theta = c - \phi$ gives $\int_0^c f(\theta)\, d\theta = \int_0^c f(c - \phi)\, d\phi$. Take $c = \frac{\pi}{4}$. By the tangent subtraction formula,

$$\tan\left(\tfrac{\pi}{4} - \phi\right) = \frac{1 - \tan\phi}{1 + \tan\phi}, \qquad \text{so} \qquad 1 + \tan\left(\tfrac{\pi}{4} - \phi\right) = \frac{2}{1 + \tan\phi}.$$

Hence

$$I = \int_0^{\pi/4} \ln\frac{2}{1 + \tan\phi}\, d\phi = \int_0^{\pi/4} \ln 2\, d\phi - \int_0^{\pi/4} \ln(1 + \tan\phi)\, d\phi = \frac{\pi}{4}\ln 2 - I.$$

Step 3. So $2I = \frac{\pi}{4}\ln 2$ and $I = \frac{\pi \ln 2}{8}$. (Numerically $I \approx 0.272198$, which agrees.)

Equivalently, without trigonometry (Nelsen's second official solution): substitute $x = \frac{1-u}{1+u}$. Then $1 + x = \frac{2}{1+u}$, $1 + x^2 = \frac{2(1+u^2)}{(1+u)^2}$ and $dx = -\frac{2\,du}{(1+u)^2}$, so $\frac{dx}{1+x^2} = -\frac{du}{1+u^2}$, the limits swap, and the same equation $I = \frac{\pi}{4}\ln 2 - I$ drops out.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains look at the same integrand. They split at the first line: (†) "I need an antiderivative." The natural attempt accepts (†) and goes looking. The model rejects it: it is a *definite* integral with specific limits, and the question is whether those limits carry a symmetry that makes the integral equal to (something known) minus itself.

> *This is a definite integral with limits 0 and 1, and $\frac{\ln(1+x)}{1+x^2}$ doesn't have an antiderivative I recognise. Before I integrate by parts for the third time: is there a substitution that maps $[0, 1]$ to itself and turns the integrand into something plus or minus itself?*

> *$\frac{dx}{1+x^2}$ is $d\theta$. So I'm integrating $\ln(1 + \tan\theta)$ over $[0, \frac{\pi}{4}]$. That interval has a reflection, $\theta \mapsto \frac{\pi}{4} - \theta$, and $\tan(\frac{\pi}{4} - \theta)$ has a formula. Try it.*

Layer 3 THE WHY

Kedlaya's first solution begins "We make the substitution $x = \tan\theta$." Here is why that is forced rather than inspired, and why the reflection that follows is the next thing anyone should try.

Why tan. The factor $\frac{dx}{1+x^2}$ is not decoration; it is the derivative of $\arctan x$, so $\frac{dx}{1+x^2}$ *is* $d(\arctan x)$. Whenever an integrand contains $\frac{dx}{1+x^2}$ multiplied by something awkward, the substitution $\theta = \arctan x$ removes the rational part completely and leaves only the awkward part, now as a function of $\tan\theta$. You are trading a two-factor integrand for a one-factor one. That is the first and cheapest simplification available.

Why reflect. Now the question is what to do with $\int_0^{\pi/4} \ln(1 + \tan\theta)\, d\theta$, which still has no elementary antiderivative. A definite integral that cannot be done by antiderivative can still be *evaluated*, and the standard A-level Further Maths tool for that is the reflection property

$$\int_0^c f(x)\, dx = \int_0^c f(c - x)\, dx,$$

which is nothing more than "the area under a graph doesn't change if you flip the graph left to right". It earns its keep exactly when $f(x)$ and $f(c - x)$ combine into something simple —

classically $\int_0^{\pi/2} \frac{\sin x}{\sin x+\cos x}\,dx$, where adding the two versions gives $\int_0^{\pi/2} 1\,dx$. The question to ask is always: *what does the reflection do to the integrand?* Here $\tan(\frac{\pi}{4}-\theta) = \frac{1-\tan\theta}{1+\tan\theta}$, and adding one gives $\frac{2}{1+\tan\theta}$: the argument of the logarithm is *inverted* and scaled by 2. A logarithm turns inversion into negation and scaling into an additive constant. So reflection sends $\ln(1+\tan\theta)$ to $\ln 2 - \ln(1+\tan\theta)$, and the integral equals a constant minus itself.

Why ln *makes this work.* The reason the problem is posed with $\ln(1+x)$ and not, say, $\sqrt{1+x}$ is precisely that the reflection acts multiplicatively on $1+x$ (it sends $1+x$ to $\frac{2}{1+x}$), and only a logarithm converts that multiplicative relation into the additive one $I = C - I$. If you see ln of something and a symmetry of the domain, check whether the symmetry multiplies or inverts the something.

Why the parts attempt was a hint. Parts produced $\frac{\pi}{4}\ln 2 - \int_0^1 \frac{\arctan x}{1+x}dx$. The number $\frac{\pi}{4}\ln 2$ is $\ln 2$ times the length of the θ-interval, which is exactly the constant that the reflection produces. When an attempt spits out a clean constant plus an ugly integral, it is worth asking whether the ugly integral is secretly the original one.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the factor $\frac{1}{x^2+1}$ together with the limit $1 = \tan\frac{\pi}{4}$. On seeing it, write $x = \tan\theta$ and the new limits $0, \frac{\pi}{4}$.

2. A definite integral over a symmetric-looking interval with no obvious antiderivative suggests the reflection $\theta \mapsto c - \theta$ and the "$I = C - I$" method.

3. Hesitation occurs after integration by parts fails: the solver doubts that a closed form exists and considers series. It is reinforced away by the rule that a Putnam integral with a clean answer and no antiderivative is almost always a symmetry problem, and by the observation that the limits $0, 1$ correspond to the tidy angles $0, \frac{\pi}{4}$.

4. The trap is zombie calculus: parts, partial fractions over $\mathbb{C}$, series with moment recurrences. Each is a legitimate technique, each either loops or runs to a page, and none uses the specific limits — which are the only special feature the problem has.

5. Recognise faster by asking, before any antiderivative hunt: "what map sends my interval to itself, and what does it do to the integrand?" For $[0,1]$ with a $\frac{dx}{1+x^2}$ weight, the answer is the Möbius map $x \mapsto \frac{1-x}{1+x}$, i.e. $\theta \mapsto \frac{\pi}{4} - \theta$.

6. Decision rule: for a definite integral whose integrand has no elementary antiderivative, look for an involution of the interval under which the integrand becomes (known) $\pm$ (itself); a logarithm of a quantity that the involution inverts or rescales is the strongest signal.

▷ **Cue:** a factor $\frac{1}{1+x^2}$ multiplying something awkward, with limits that are tangents of nice angles $\longrightarrow$ **Move:** substitute $x = \tan\theta$; the rational factor disappears into $d\theta$

▷ **Cue:** a definite integral with no elementary antiderivative and a clean-looking expected answer $\longrightarrow$ **Move:** apply $\int_0^c f(x)\,dx = \int_0^c f(c-x)\,dx$ and add or subtract the two versions

▷ **Cue:** ln(something) where a symmetry of the domain inverts or rescales the something $\longrightarrow$ **Move:** the symmetry turns ln into constant minus ln, giving $I = C - I$

▷ **Cue:** integration by parts returns a clean constant plus a new ugly integral $\longrightarrow$ **Move:** test whether the ugly integral is the original in disguise

2005 A6

2005 A6. Let n be given, $n \geq 4$, and suppose that $P_1, P_2, \ldots, P_n$ are n randomly, independently and uniformly, chosen points on a circle. Consider the convex n-gon whose vertices are the P_i. What is the probability that at least one of the vertex angles of this polygon is acute?

Fields: probability (geometric probability, uniform spacings); inclusion–exclusion; circle geometry
Position: A6 — the last problem of the morning paper. Hard because "at least one" invites a union

bound that overcounts, and the fix needs a structural fact about how many acute angles a cyclic polygon can have. **Answer:** $\dfrac{n(n-2)}{2^{n-1}}$.

Layer 1 THE WHAT — surface analysis

The natural attempt. Normalise the circumference to 1. You start with a single vertex V. Its angle is an inscribed angle, subtending the arc between its two neighbours that does not contain V; it is acute exactly when that arc is less than half the circle, i.e. when the other $n-1$ points lie in an arc of length less than $\frac{1}{2}$ that avoids V. Rotate so that V is at 0; the other points are $n-1$ independent uniforms on $(0,1)$ and the condition is that their range max $-$ min is below $\frac{1}{2}$. For m independent uniforms, $\Pr(\text{range} \le t) = mt^{m-1} - (m-1)t^m$, so with $m = n-1$, $t = \frac{1}{2}$:

$$\Pr(\text{angle at } V \text{ acute}) = \frac{n-1}{2^{n-2}} - \frac{n-2}{2^{n-1}} = \frac{n}{2^{n-1}}.$$

Now "at least one": there are n vertices, so multiply by n and get $\frac{n^2}{2^{n-1}}$. For $n = 4$ that is 2. A probability of 2 is the wall. You have computed the *expected number* of acute angles, not the probability that there is one; polygons with two acute angles have been counted twice. To repair it you need to know how often, and in what configurations, acute angles coincide — and full inclusion–exclusion over n events looks hopeless.

The model chain. *Step 1: angles in terms of gaps.* With probability 1 the points are distinct. Going counterclockwise, let $g_1, \dots, g_n$ be the arc lengths between consecutive vertices (the gaps), indices mod n, so $g_i \ge 0$ and $\sum g_i = 1$. Let V_i be the vertex between gaps g_i and g_{i+1}. Its neighbours cut off, on the side away from V_i, an arc of length $1 - g_i - g_{i+1}$; the inscribed angle at V_i is half the corresponding central angle, namely $\pi(1 - g_i - g_{i+1})$. Hence

$$\angle V_i \text{ acute} \iff g_i + g_{i+1} > \tfrac{1}{2}. \qquad (A_i)$$

Step 2: the distribution of the gaps. By rotation invariance, measure positions from P_n. The other $n-1$ points are independent uniforms on $[0,1)$; their order statistics $0 < u_1 < \cdots < u_{n-1} < 1$ have constant density $(n-1)!$ on that region. The gaps are u_1, $u_2 - u_1, \dots, 1 - u_{n-1}$, a linear change of variables with Jacobian 1. So $(g_1, \dots, g_n)$ is uniform on the simplex $\{g_i \ge 0, \sum g_i = 1\}$, with density $(n-1)!$ in the coordinates $g_1, \dots, g_{n-1}$. This density is symmetric in all n gaps, so any statement about a particular block of consecutive gaps has the same probability wherever the block sits. A consequence used twice below: the set of k non-negative coordinates with sum s has $(k-1)$-dimensional volume $\frac{s^{k-1}}{(k-1)!}$.

Step 3: at most two acute angles, and then adjacent. If $\{i, i+1\}$ and $\{j, j+1\}$ are disjoint, A_i and A_j cannot both hold, since then $g_i + g_{i+1} + g_j + g_{j+1} > 1$. For $n \ge 4$, among any three distinct indices two are non-adjacent (three mutually adjacent indices mod n need $n = 3$), so no three A_i hold together. Inclusion–exclusion therefore stops after the pair terms, and the only pairs that can co-occur are the n adjacent pairs $\{A_i, A_{i+1}\}$:

$$\Pr\Big(\bigcup_i A_i\Big) = n \Pr(A_1) - n \Pr(A_1 \cap A_2).$$

Step 4: $\Pr(A_1)$. A_1 says $r := g_3 + \cdots + g_n < \frac{1}{2}$. Integrating out the gaps, r has density $(n-1)! \cdot \frac{r^{n-3}}{(n-3)!} \cdot (1-r) = (n-1)(n-2)\, r^{n-3}(1-r)$ (the factor $1-r$ is the length of the segment $g_1 + g_2 = 1 - r$). So

$$\Pr(A_1) = (n-1)(n-2) \int_0^{1/2} (r^{n-3} - r^{n-2})\, dr = \frac{n-1}{2^{n-2}} - \frac{n-2}{2^{n-1}} = \frac{n}{2^{n-1}},$$

matching the natural attempt.

Step 5: $\Pr(A_1 \cap A_2)$. Write $a = g_1$, $b = g_2$, $c = g_3$, $r = g_4 + \cdots + g_n$ (a sum of $n-3 \geq 1$ gaps). Then $A_1 \iff a+b > \frac{1}{2} \iff c+r < \frac{1}{2}$, and $A_2 \iff a+r < \frac{1}{2}$. The joint density of (a,c,r) is $(n-1)! \cdot \frac{r^{n-4}}{(n-4)!} = (n-1)(n-2)(n-3)\, r^{n-4}$ on $a,c,r \geq 0$, $a+c+r \leq 1$ (b is then determined). The event is $a < \frac{1}{2} - r$, $c < \frac{1}{2} - r$ with $r < \frac{1}{2}$, which lies inside $a+c+r \leq 1$. So

$$\Pr(A_1 \cap A_2) = (n-1)(n-2)(n-3) \int_0^{1/2} r^{n-4} \left(\tfrac{1}{2} - r\right)^2 dr.$$

With $r = x/2$ this is $\frac{(n-1)(n-2)(n-3)}{2^{n-1}} \int_0^1 x^{n-4}(1-x)^2\, dx = \frac{(n-1)(n-2)(n-3)}{2^{n-1}} \cdot \frac{2}{(n-1)(n-2)(n-3)} = \frac{2}{2^{n-1}}$.

Step 6. $\Pr(\text{some acute angle}) = n \cdot \frac{n}{2^{n-1}} - n \cdot \frac{2}{2^{n-1}} = \frac{n(n-2)}{2^{n-1}}$.

Checks. For $n = 4$ this is 1: a cyclic quadrilateral has opposite angles summing to π, so one of each opposite pair is acute unless both are right angles, an event of probability 0. A Monte Carlo run (2×10^5 trials each) gives 0.937, 0.750, 0.377 for $n = 5, 6, 8$ against 0.9375, 0.75, 0.375.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains compute $\Pr(\text{angle at } V \text{ acute}) = \frac{n}{2^{n-1}}$. They split at (†): "multiply by n". The natural attempt multiplies and hits a probability of 2. The model stops before multiplying and asks how the events for different vertices overlap — which is a question about the *gaps*, not about the points.

> *Multiplying by n gives the expected number of acute angles. That only equals the probability if two acute angles never happen together, and at $n = 4$ I get 2, so they do. How many acute angles can a polygon inscribed in a circle actually have?*

> *An inscribed angle only depends on the arc opposite it, so the angle at a vertex is a function of the two gaps next to it: acute means those two gaps add to more than half the circle. Two disjoint pairs of gaps can't both do that. So acute angles come at most two at a time, and only next to each other — inclusion–exclusion has only n pair terms.*

Layer 3 THE WHY

Kedlaya's first solution opens with "The angle at a vertex P is acute if and only if all of the other points lie on an open semicircle," and then argues that acute angles must be consecutive by a covering argument with semicircles. Both facts are right, but the reason they are true, and the reason one should go looking for them, is the gap reformulation. Here is where it comes from.

Why gaps. The problem is rotation-invariant: rotating all points does not change any angle. So the positions carry one wasted degree of freedom, and the natural coordinates are the differences between consecutive points — the gaps. That is not a device; it is what you get by quotienting out the symmetry. Then the inscribed angle theorem, which says an inscribed angle is half the central angle on the same arc, makes each vertex angle a linear function of two gaps: $\pi(1 - g_i - g_{i+1})$. Once angles are linear in gaps and gaps sum to 1, a statement like "acute" becomes a linear inequality on a simplex, and the whole problem becomes a volume computation.

Why the overlap structure is forced. "At least one" is a union, and the only exact tool for a union is inclusion–exclusion. Inclusion–exclusion over n events looks unmanageable only until you ask which intersections are possible. With a budget of total length 1, you cannot have two disjoint pairs each spending more than half. That single budget constraint kills every intersection except adjacent pairs — it is the same pigeonhole fact as "two disjoint arcs each longer than a semicircle cannot fit on a circle," which is exactly Kedlaya's covering argument, now visible at a glance.

Why $n \geq 4$. For $n = 3$, every pair of indices is adjacent and all three events can overlap; indeed every triangle has at least two acute angles, so the probability is 1, not the formula's $\frac{3}{4}$. The hypothesis $n \geq 4$ is precisely the condition for "three distinct indices include a non-adjacent pair". When a hypothesis looks arbitrary, find where the proof uses it; here that is Step 3.

Why the numbers are powers of 2. Every event is "some block of gaps sums to less than $\frac{1}{2}$", and the volume of a scaled simplex scales by a power of the scale factor. Halving produces $2^{-(n-1)}$ and the polynomial prefactors come from the Beta integrals. Seeing $\frac{1}{2}$ in the event and expecting $2^{-(n-1)}$ in the answer is a useful sanity check while computing.

Layer $+\alpha$ BACKTESTING

1. The first cue is "points on a circle" together with "vertex angles". Write down the inscribed angle theorem in gap form: angle at V_i equals $\pi(1 - g_i - g_{i+1})$, so acute iff $g_i + g_{i+1} > \frac{1}{2}$.

2. Rotation invariance plus a condition on angles suggests spacings (gaps), which are uniform on a simplex; "at least one" suggests inclusion–exclusion, to be pruned by a budget argument.

3. Hesitation occurs after the union bound gives 2 for $n = 4$. It is reinforced away by asking the structural question "how many of these events can hold at once?" before attempting any further computation.

4. The trap is the linearity-of-expectation reflex: $n \cdot \Pr(\text{one vertex})$ is the expected number of acute angles, not the probability of at least one. A second trap is working with the points' positions instead of the gaps, which makes the adjacency structure of the events invisible.

5. Recognise faster by testing a small case first: $n = 4$ must give 1 (opposite angles of a cyclic quadrilateral are supplementary), which immediately exposes the overcount and tells you acute angles must pair up.

6. Decision rule: for "at least one of n events", first find the maximum number that can occur simultaneously and which ones; if a conserved total (length, mass, probability) forbids large overlaps, inclusion–exclusion truncates to a few explicit terms, and symmetry makes each term a single computation.

▷ **Cue:** random points on a circle and a question about angles or arcs $\longrightarrow$ **Move:** pass to the gaps between consecutive points; they are uniform on the simplex $\sum g_i = 1$

▷ **Cue:** an inscribed angle in a polygon on a circle $\longrightarrow$ **Move:** it equals π times the arc opposite it, which is 1 minus the two adjacent gaps

▷ **Cue:** "probability that at least one of n symmetric events occurs" and $n \cdot \Pr(\text{one})$ exceeds 1 $\longrightarrow$ **Move:** ask how many events can occur simultaneously; use a budget constraint to kill most intersections, then truncated inclusion–exclusion

▷ **Cue:** several quantities each required to exceed half of a fixed total $\longrightarrow$ **Move:** disjoint ones cannot coexist; only overlapping ones can

▷ **Cue:** a hypothesis like $n \geq 4$ with no visible role $\longrightarrow$ **Move:** locate the step that fails for the excluded case and check the formula there

Putnam 2005 — Paper B

2005 B1

2005 B1. Find a nonzero polynomial $P(x, y)$ such that $P(\lfloor a \rfloor, \lfloor 2a \rfloor) = 0$ for all real numbers a. (Note: $\lfloor v \rfloor$ is the greatest integer less than or equal to v.)

Fields: number theory (floor function identities) **Position:** B1 – opener, an algebraic identity dressed as an existence problem **Answer:** Yes: $P(x, y) = (y - 2x)(y - 2x - 1)$, since $\lfloor 2a \rfloor - 2\lfloor a \rfloor \in \{0, 1\}$ for every real a.

Layer 1 THE WHAT — surface analysis

The natural attempt. The instinct on seeing "find a polynomial P" is to start generating data and fit a curve to it. Tabulate a few values: $a = 0.3$ gives $(\lfloor a \rfloor, \lfloor 2a \rfloor) = (0, 0)$; $a = 0.6$ gives $(0, 1)$; $a = 1.2$ gives $(1, 2)$; $a = 1.7$ gives $(1, 3)$; $a = 2.4$ gives $(2, 4)$. The pattern "y is close to $2x$" suggests trying $P(x, y) = y - 2x$ first. This fails immediately: at $a = 0.6$, $y - 2x = 1 - 0 = 1 \neq 0$. So y is not always $2x$; sometimes it is $2x + 1$. Rather than stopping to ask what governs the choice between the two, the natural attempt escalates the polynomial: posit a general form, say $P(x, y) = Ay^2 + Bxy + Cx^2 + Dy + Ex + F$, plug in six or more of the tabulated pairs, and solve the resulting linear system in $A, \ldots, F$ for a polynomial that vanishes on all of them. This is a legitimate method – curve-fitting through enough points does constrain the coefficients – but it is expensive, and worse, it is blind. A student doing this has no assurance that a low-degree polynomial passing through five or six sampled pairs actually vanishes on *every* real a; the fit could easily be an artefact of an unlucky (or lucky) choice of sample points, and nothing in the method certifies completeness. The approach can be made to work with enough care, but it treats the problem as numerical interpolation rather than asking what actually constrains $\lfloor 2a \rfloor$ in terms of $\lfloor a \rfloor$.

The model chain. Write $a = m + t$ where $m = \lfloor a \rfloor \in \mathbb{Z}$ and $t \in [0, 1)$ is the fractional part. Then $2a = 2m + 2t$, and since $0 \leq t < 1$ we have $0 \leq 2t < 2$. By definition of the floor function applied to the integer-plus-remainder decomposition of $2a$,

$$\lfloor 2a \rfloor = 2m + \lfloor 2t \rfloor.$$

Because $2t$ ranges over $[0, 2)$, $\lfloor 2t \rfloor$ can only be 0 (when $0 \leq t < 1/2$) or 1 (when $1/2 \leq t < 1$). Hence, writing $x = \lfloor a \rfloor$ and $y = \lfloor 2a \rfloor$,

$$y - 2x = \lfloor 2t \rfloor \in \{0, 1\}$$

for every real a, with no exceptions. A quantity known to take only the values 0 or 1 is exactly the set of roots of $z(z - 1) = 0$. Substituting $z = y - 2x$ gives $P(x, y) = (y - 2x)(y - 2x - 1)$, a nonzero polynomial that vanishes identically on every pair $(\lfloor a \rfloor, \lfloor 2a \rfloor)$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains start from the same tabulated pairs and both notice that y is either $2x$ or $2x + 1$. They split at the moment of deciding what the tabulated data is *evidence*

of. The natural attempt treats the pairs as points to be passed to a curve-fitting procedure – an unknown polynomial with unknown coefficients, solved for numerically. The model treats the same pairs as symptoms of an exact algebraic fact about $y - 2x$, to be derived directly from what floor and doubling do to the fractional part of a, before a single coefficient is ever written down.

> *The pairs (x, y) I am tabulating are not scattered points to be fit by an arbitrary curve – they are constrained by an exact relationship $y - 2x \in \{0, 1\}$ that I can derive directly from the definition of the floor function applied to $2a = 2m + 2t$, so I should derive that relationship first and write down the polynomial only once I know exactly what finite set of values $y - 2x$ takes.*

> *A quantity I know is always an integer confined to a finite set $\{k_1, \ldots, k_r\}$ is automatically a root of $\prod_i (z - k_i)$ – I do not need to search for a polynomial by curve-fitting; I need to pin down the finite set, and the polynomial writes itself.*

Layer 3 THE WHY

The identity $a = \lfloor a \rfloor + \{a\}$, with $0 \leq \{a\} < 1$, is not a trick imported for this problem – it is the literal definition of the floor function, the statement of what "greatest integer less than or equal to a" means when you also track the remainder. Multiplying the interval $[0, 1)$ by the positive constant 2 sends it to $[0, 2)$: this is elementary interval scaling, nothing more. Applying the floor function's own definition to a quantity known to lie in $[0, 2)$ forces that quantity's floor to be 0 or 1 – there is no other integer it could be. So the fact $y - 2x \in \{0, 1\}$ is two uses of the same definition, chained through a multiplication by a positive constant; nothing about it is specific to the numbers 1 and 2 in the problem, and the identical argument would produce a corresponding two-term (or, for a scaling by n, n-term) product for $\lfloor na \rfloor - n\lfloor a \rfloor$.

The second half of the "why" is the move from "z lies in a finite set" to "z is a root of a specific polynomial." This is also not a trick: it is the defining property of a root run backwards. A polynomial's zero set can be exactly prescribed by multiplying out $(z - s)$ for each s in the target set, and this is the standard device for converting a case-based fact about integers – here, a two-case fact – into a single polynomial identity that holds unconditionally, with no case distinction visible in the final answer. The official solution states the bound $2m \leq \lfloor 2a \rfloor \leq 2m + 1$ and the resulting polynomial in two lines without separating these two ideas; the archive's job is to notice that the first line is nothing but the definition of floor applied twice, and the second line is nothing but the definition of a root applied to a two-element set.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that a few tabulated pairs $(\lfloor a \rfloor, \lfloor 2a \rfloor)$ show y is always either $2x$ or $2x + 1$, never anything else. On seeing this, write down immediately "the difference $y - 2x$ is confined to a 2-element set," not "let me fit a curve."

2. That cue suggests the method: once a quantity is known to be confined to a finite set of integers, express it as a root of the product of $(z - k)$ over that set, rather than searching for a polynomial that merely happens to pass through sampled points.

3. Hesitation occurs at turning "$y - 2x \in \{0, 1\}$" into a usable polynomial identity, particularly for a student who has only ever built polynomials from formulas, not from set membership. It is reinforced away by recalling the elementary fact that an integer $z \in \{0, 1\}$ satisfies $z^2 = z$, i.e. $z(z - 1) = 0$ – a fact from basic algebra/logic, not analysis, and one worth having on permanent recall for any "confined to $\{0, 1\}$" situation.

4. The trap is treating a floor-function existence problem as a numerical interpolation problem: sampling enough (x, y) pairs and solving a linear system for polynomial coefficients, a method that can be made to work but never certifies, by itself, that the resulting polynomial vanishes on every real a rather than merely the sampled ones.

5. Recognisable faster next time by treating any expression built from $\lfloor a \rfloor$ and $\lfloor ca \rfloor$ (for a fixed constant c) as a decomposition problem first: write $a = \lfloor a \rfloor + \{a\}$, scale the fractional part by c, and read off exactly which integers the resulting floor can be. This settles such problems in one or two lines before any polynomial is written.

6. Reusable decision rule: if a quantity built from an existence/construction requirement is provably confined to a finite set of integers $\{k_1, \ldots, k_r\}$, the required polynomial identity is always $\prod_i (z - k_i) = 0$; the actual mathematical content of the problem is establishing the finite set, never the polynomial itself.

▷ **Cue:** a floor of a scaled real number, e.g. $\lfloor ca \rfloor$, appears alongside $\lfloor a \rfloor$ $\longrightarrow$ **Move:** write $a = \lfloor a \rfloor + \{a\}$ with $0 \leq \{a\} < 1$, scale the fractional part by c, and read off the finite set of integer values the floor of the scaled term can take

▷ **Cue:** a quantity is known to lie in a finite integer set $\{k_1, \ldots, k_r\}$ $\longrightarrow$ **Move:** encode this as the polynomial identity $\prod_i (z - k_i) = 0$, rather than hunting for a polynomial by fitting sampled points

2005 B2

2005 B2. Find all positive integers $n, k_1, \ldots, k_n$ such that $k_1 + \cdots + k_n = 5n - 4$ and

$$\frac{1}{k_1} + \cdots + \frac{1}{k_n} = 1.$$

Fields: number theory (Cauchy-Schwarz / AM-HM inequality, finite case analysis) **Position:** B2
Answer: Yes: $(n, k_1, \ldots, k_n) = (1;\, 1)$, $(3;$ a permutation of $2, 3, 6)$, and $(4;\, 4, 4, 4, 4)$ – no other solutions exist.

Layer 1 THE WHAT — surface analysis

The natural attempt. The problem hands you a sum and a sum of reciprocals that must simultaneously equal specific numbers, and the reflex is to start testing n directly. $n = 1$ is immediate: $k_1 = 5(1) - 4 = 1$ and $1/k_1 = 1$, so it works. For $n = 2$: $k_1 + k_2 = 6$ with $1/k_1 + 1/k_2 = 1$; try the divisor pairs of 6 by hand – $(2, 4)$ gives $1/2 + 1/4 = 3/4 \neq 1$; $(3, 3)$ gives $2/3 \neq 1$ – neither works, but this is discovered only by checking every pair that sums to 6. For $n = 3$: $k_1 + k_2 + k_3 = 11$ with $1/k_1 + 1/k_2 + 1/k_3 = 1$; there are nine unordered partitions of 11 into three positive parts, and checking all of them by hand (e.g. $(1, 1, 9), (1, 2, 8), (1, 3, 7), (1, 4, 6), (1, 5, 5), (2, 2, 7), (2, 3, 6), (2, 4, 5), (3, 3, 5), (3, 4, 4)$) is tedious but doable, eventually landing on $(2, 3, 6)$. The trouble is what this method does *not* tell you: nothing in the process of checking $n = 1, 2, 3$ by exhaustive partition-listing gives any reason to stop at $n = 4$, or indeed at any particular n. The number of partitions of $5n - 4$ into n positive parts grows quickly, and a student proceeding this way is running an unbounded manual search with no certificate that all solutions have been found once the checking gets tedious enough to abandon.

The model chain. The hypotheses supply $\sum k_i = 5n - 4$ and $\sum 1/k_i = 1$ for the *same* list of positive integers – exactly the pairing bounded below by the Cauchy-Schwarz (equivalently AM-HM) inequality:

$$\left(\sum_{i=1}^{n} k_i \right) \left(\sum_{i=1}^{n} \frac{1}{k_i} \right) \geq n^2,$$

with equality if and only if all k_i are equal. Substituting the given values, $(5n - 4)(1) \geq n^2$, i.e.

$$n^2 - 5n + 4 \leq 0 \iff (n - 1)(n - 4) \leq 0 \iff 1 \leq n \leq 4.$$

This reduces the problem to exactly four finite cases, checked as follows (taking $k_1 \leq \cdots \leq k_n$ without loss of generality). $n = 1$: forces $k_1 = 1$, which works. $n = 2$: from $1/k_1 + 1/k_2 = 1$, $k_1 + k_2 = k_1 k_2$; combined with $k_1 + k_2 = 6$ this gives $k_1 k_2 = 6$ too, so k_1, k_2 are roots of $t^2 - 6t + 6 = 0$, whose discriminant 12 is not a perfect square – no integer solution. $n = 3$: $k_1 + k_2 + k_3 = 11$ with $k_1 \leq k_2 \leq k_3$ forces $3k_1 \leq 11$, so $k_1 \in \{2, 3\}$ (since $k_1 = 1$ forces $\sum 1/k_i > 1$, impossible for $n > 1$). For $k_1 = 2$: $k_2 + k_3 = 9$, $1/k_2 + 1/k_3 = 1/2 \Rightarrow k_2 k_3 = 2(k_2 + k_3) = 18$, giving $t^2 - 9t + 18 = 0$, roots $3, 6$ – the solution $(2, 3, 6)$. For $k_1 = 3$: $k_2 + k_3 = 8$, $1/k_2 + 1/k_3 = 2/3 \Rightarrow k_2 k_3 = (3/2)(8) = 12$, giving $t^2 - 8t + 12 = 0$, roots $2, 6$ – but $k_2 \geq k_1 = 3$ rules out $k_2 = 2$, so no valid solution here. Hence $n = 3$ gives only $(2, 3, 6)$ up to order. $n = 4$: here $5n - 4 = 16 = n^2$ exactly, so the Cauchy-Schwarz inequality is an equality, which forces all k_i equal: $k_1 = k_2 = k_3 = k_4 = 4$, and indeed $4 \cdot (1/4) = 1$ checks. No case search is needed for $n = 4$ at all.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both paths eventually resolve the same four values of n by hand. They split before either of them touches a single partition: the natural attempt starts testing $n = 1, 2, 3, \ldots$ directly and enumerating partitions of $5n - 4$ into n parts without first asking whether n itself is bounded. The model asks, on reading the two hypotheses together, "haven't I seen a sum and a reciprocal-sum given together before?" and reaches for Cauchy-Schwarz/AM-HM before writing down a single explicit partition.

> *I am given both $\sum k_i$ and $\sum 1/k_i$ as explicit numbers – that is exactly the pairing that Cauchy-Schwarz (equivalently AM-HM) bounds below by n^2, so I should test that inequality against the two given numbers before enumerating anything by hand.*

> *Ordering $k_1 \leq k_2 \leq k_3$ and using $3k_1 \leq k_1 + k_2 + k_3 = 11$ bounds the smallest term to $k_1 \leq 3$, cutting the $n = 3$ case search down to two values of k_1 instead of listing all nine partitions of 11 into three positive parts.*

> *$5n - 4 = n^2$ exactly when $n = 4$, so at $n = 4$ the AM-HM inequality holds with equality – and equality in Cauchy-Schwarz forces every k_i to be equal. I do not need to search partitions of 16 into 4 parts at all; the equality condition finishes the case for free.*

Layer 3 THE WHY

A sum and a sum of reciprocals of the same finite list, handed to you as two separate hypotheses, is not an incidental pairing – it is exactly the input the Cauchy-Schwarz inequality is built to consume. Applying Cauchy-Schwarz to the vectors $(\sqrt{k_i})_i$ and $(1/\sqrt{k_i})_i$ gives $(\sum k_i)(\sum 1/k_i) \geq (\sum 1)^2 = n^2$, equivalently the AM-HM inequality $\frac{\sum k_i}{n} \geq \frac{n}{\sum 1/k_i}$. This is not a trick imported from outside the problem; it is the defining behaviour of the arithmetic and harmonic means of the same list, and the problem's two hypotheses are literally the numerator and denominator that inequality is stated about. Recognising the match is the "of course" step, once the shape "sum" and "sum of reciprocals of the same list" is seen as a named object rather than two independent conditions to satisfy by search.

The official solution states the inequality and deduces $n \leq 4$ in a single sentence, then proceeds to case-check $n = 1, \ldots, 4$ without commenting on why the equality case at $n = 4$ needs no separate algebra: the reason is that $5n - 4 = n^2$ exactly at $n = 4$ is not a coincidence to be verified after the fact, it is precisely the condition under which the bounding inequality used to derive $n \leq 4$ becomes tight, and Cauchy-Schwarz's own equality condition – proportionality of the two vectors, here forcing all k_i equal – answers the $n = 4$ case without any further computation. The archive's contribution here is to notice that the same inequality that bounds n

also, for free, solves the extremal case of n it produces; that link is not stated explicitly in the terse official write-up, but it is exactly why $n = 4$ is the easiest of the four cases rather than a fourth independent piece of casework.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that the problem hands you $\sum k_i$ and $\sum 1/k_i$ as two explicit numbers for the same list. On seeing both a sum and its reciprocal-sum specified together, write down Cauchy-Schwarz/AM-HM immediately as a candidate bound, before any case search.
2. That cue suggests testing the given numbers against $(\sum k_i)(\sum 1/k_i) \geq n^2$ to bound the number of variables before enumerating a single partition.
3. Hesitation occurs in moving from "$n \leq 4$" to actually resolving each of the four cases, especially at $n = 3$ where many partitions of 11 exist. Reinforce it away by ordering $k_1 \leq \cdots \leq k_n$ and bounding the smallest term via $nk_1 \leq \sum k_i$, which slashes the case list before any brute enumeration is needed.
4. The trap is treating "find all" as license for unbounded manual search over every n and every partition, when the two given constraints already bound n to at most four values via a single inequality.
5. Recognisable faster next time: whenever a problem specifies both a sum and a sum of reciprocals of the same list, reach for Cauchy-Schwarz/AM-HM before writing a single explicit case, and check immediately whether the given numbers make the inequality an equality.
6. Reusable rule: a sum and a reciprocal-sum given together for the same list is the signature of an AM-HM/Cauchy-Schwarz bound; use it first to bound the number of variables (or their size), and check whether the given numbers force the equality case, which resolves that case without further algebra.

▷ **Cue:** a sum $\sum k_i$ and a sum of reciprocals $\sum 1/k_i$ of the same finite list are both given as explicit numbers $\longrightarrow$ **Move:** apply Cauchy-Schwarz/AM-HM: $(\sum k_i)(\sum 1/k_i) \geq n^2$, to bound n (or the size of the k_i) before any case enumeration

▷ **Cue:** the given numbers make the AM-HM inequality an equality (i.e. the two given values multiply to exactly n^2) $\longrightarrow$ **Move:** invoke the equality condition of Cauchy-Schwarz directly – all terms equal – instead of solving that case by explicit algebra

2005 B3

2005 B3. Find all differentiable functions $f : (0, \infty) \to (0, \infty)$ for which there is a positive real number a such that

$$f'\left(\frac{a}{x}\right) = \frac{x}{f(x)}$$

for all $x > 0$.

Fields: analysis (functional/differential equations, involution substitution) **Position:** B3 **Answer:** Yes: $f(x) = cx^d$ for constants $c, d > 0$, subject to the constraint $c^2 d\, a^{d-1} = 1$ for the given a – which is solvable for any $d \neq 1$ by choosing a appropriately, but forces $c = 1$ exactly when $d = 1$.

Layer 1 THE WHAT — surface analysis

The natural attempt. The equation swaps x for a/x in a suggestive way, and a strong first instinct is to guess that f is a power function and check it. Try $f(x) = cx^d$ for constants $c, d > 0$. Then $f'(x) = cd\, x^{d-1}$, so $f'(a/x) = cd\,(a/x)^{d-1} = cd\, a^{d-1} x^{1-d}$. On the other side, $x/f(x) = x/(cx^d) = x^{1-d}/c$. Matching powers of x (both sides are x^{1-d}, so this is automatic) and equating coefficients gives $cd\, a^{d-1} = 1/c$, i.e. $c^2 d\, a^{d-1} = 1$. For $d \neq 1$ this can always be solved for a suitable $a > 0$ given any $c, d > 0$; for $d = 1$, $a^{d-1} = a^0 = 1$ regardless of a, so the

constraint collapses to $c^2 = 1$, forcing $c = 1$. This is a genuinely successful piece of work – it produces the entire family of solutions and even catches the $d = 1$ exception. But it stops at exhibiting examples. Nothing in "guess a form, plug it in, solve for the constants" rules out some other, non-power-law function also satisfying the same equation; the ansatz method can prove sufficiency (these functions work) but has no built-in mechanism for proving necessity (these are the *only* functions that work), which is what "find all" actually demands.

The model chain. Substitute a/x for x throughout the given equation. Since $a/(a/x) = x$, this yields a second, genuinely new equation:

$$f'(x) = \frac{a}{x\,f(a/x)}. \tag{$*$}$$

Differentiate $(*)$ with respect to x. Writing $D(x) = xf(a/x)$ so that $f'(x) = a/D(x)$,

$$D'(x) = f(a/x) - \frac{a}{x}f'(a/x),$$

and substituting the *original* equation $f'(a/x) = x/f(x)$ removes the last occurrence of f' at a/x:

$$D'(x) = f(a/x) - \frac{a}{f(x)}.$$

Using $(*)$ again to write $f(a/x) = a/(xf'(x))$ and $D(x) = a/f'(x)$, and substituting throughout, the algebra collapses (after multiplying out and simplifying) to a second-order ODE purely in $f(x)$:

$$xf(x)f''(x) + f(x)f'(x) = xf'(x)^2.$$

Dividing by $f(x)^2$:

$$\frac{xf''(x)}{f(x)} + \frac{f'(x)}{f(x)} - \frac{xf'(x)^2}{f(x)^2} = 0.$$

The left side is exactly $\dfrac{d}{dx}\left(\dfrac{xf'(x)}{f(x)}\right)$, by the quotient rule applied directly to $xf'(x)/f(x)$. So $xf'(x)/f(x)$ is constant, say equal to d:

$$\frac{f'(x)}{f(x)} = \frac{d}{x} \implies \ln f(x) = d\ln x + C \implies f(x) = cx^d, \quad c = e^C > 0.$$

This proves *every* solution has the power-law form; substituting back into the original equation, as in the natural attempt, then pins down exactly which (c, d, a) are admissible: $c^2 d\, a^{d-1} = 1$, giving $d > 0$ free (with a chosen accordingly) except $d = 1$, which forces $c = 1$.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains notice the equation's argument swap $x \leftrightarrow a/x$ and both make productive use of it. They split on how many times that observation is used. The natural attempt uses the swap once, informally, to motivate trying a power-law ansatz, then verifies the ansatz directly. The model uses the swap twice and formally: once to literally substitute a/x for x inside the given equation and generate a second, independent equation $(*)$; and a second time, after differentiating $(*)$, to re-substitute the *original* equation back in and eliminate every remaining evaluation at a/x, leaving an equation in x alone.

> *Guessing $f(x) = cx^d$ and checking it satisfies the equation is not a proof that every solution has this form – to get uniqueness I need to turn the functional equation into a differential equation relating f and its derivatives at a* single *point x, not at the two related points x and a/x.*

> *The given equation relates f' at a/x to a quantity at x; substituting a/x in place of x throughout the equation is legitimate precisely because $a/(a/x) = x$ sends the argument back to where it started – so I get a second, genuinely different equation, and now I have two relations in the two related unknowns $f(x)$ and $f(a/x)$, which is enough to eliminate one of them.*

> *After differentiating the substituted equation, $f'(a/x)$ reappears – and I already have a formula for exactly that, from the original equation itself, so substituting it back in removes the last evaluation at a/x and leaves one self-contained equation purely in x.*

> *The combination $xf''(x)/f(x) + f'(x)/f(x) - xf'(x)^2/f(x)^2$, once everything is divided through by $f(x)^2$, is exactly the derivative of $xf'(x)/f(x)$ by the quotient rule – so instead of solving a second-order ODE by some other means, I should integrate this recognised total derivative directly.*

Layer 3 THE WHY

The deep reason the substitution $x \leftrightarrow a/x$ can be used twice, and closes up into a clean single-variable ODE, is that $x \mapsto a/x$ is an involution: applying it twice returns x exactly. Any equation built from an involutive substitution, when that substitution is applied to the equation itself and the result differentiated and re-substituted, must eventually reduce to an equation in one variable, because the two evaluation points (x and its image under the involution) keep swapping back into each other rather than generating an infinite chain of new points. This is not a device specific to this problem; it is the general method for any functional or differential equation whose two arguments are related by an involution ($x \leftrightarrow 1/x$, $x \leftrightarrow -x$, $x \leftrightarrow a/x$, and so on): substitute the involution once to double the number of equations available, then eliminate.

The recognition that $xf''/f + f'/f - xf'^2/f^2$ is an exact derivative is itself not a leap but the quotient rule run in reverse: computing $\frac{d}{dx}\left(\frac{xf'}{f}\right) = \frac{f'}{f} + \frac{xf''}{f} - \frac{xf'^2}{f^2}$ directly, by the ordinary quotient rule on xf'/f, reproduces exactly the left side once the ODE is divided by f^2. So once the eliminated ODE is in hand, spotting this shape is mechanical differentiation, not insight to be admired.

The official solution performs the substitute–differentiate–eliminate sequence in four terse lines without pausing to justify why re-substituting removes $f(a/x)$; the justification is exactly that $a/(a/x) = x$, so the original equation and the equation obtained by substituting a/x for x are both statements about the same unknown function f, evaluated at the same pair of points $\{x, a/x\}$ – and two independent relations between f, f' at those two points are, algebraically, enough to eliminate one point entirely and land on a self-contained statement about the other. The $d = 1$ exception, credited in the official remarks to Brad Rodgers without derivation, follows the same principle as the rest of the argument once made explicit: $a^{d-1} = a^0 = 1$ for every a exactly when $d = 1$, so the constraint $c^2 d\, a^{d-1} = 1$ loses its dependence on a at that single value of d and collapses to $c^2 = 1$ – there is no freedom left to tune a to absorb an arbitrary c, whereas for every other d, a^{d-1} varies continuously and can be set to whatever value $1/(c^2 d)$ requires.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the equation's argument swap: f' is evaluated at a/x while everything else is evaluated at x. On seeing an argument swap like $x \leftrightarrow a/x$, ask immediately whether substituting a/x for x throughout the equation produces a second, independent, usable equation.
2. That cue suggests the method: substitute the involution to double the equations, then differentiate one of them and re-substitute the other to eliminate the "other" evaluation point, reducing everything to a single ODE in $f(x)$ alone.

3. Hesitation occurs at recognising $xf''(x)/f(x) + f'(x)/f(x) - xf'(x)^2/f(x)^2$ as an exact derivative after dividing by $f(x)^2$. It is reinforced away by writing $g(x) = xf'(x)/f(x)$ and computing $g'(x)$ directly via the quotient rule to see that it reproduces exactly this expression, rather than trying to guess an integrating factor from a blank page.

4. The trap is stopping at a successfully verified ansatz – $f(x) = cx^d$ works for a suitable a – and treating that as the complete solution, when "find all" requires proving no other function can satisfy the equation.

5. Recognisable faster next time: any functional or differential equation whose two arguments are related by an involution ($x \leftrightarrow 1/x$, $x \leftrightarrow -x$, $x \leftrightarrow a/x$) should be re-derived with the involution substituted into the equation itself before attempting an ansatz. Ansatz-and-check only ever proves sufficiency, never necessity.

6. Reusable rule: to prove a functional equation's solution set is exactly an ansatz family, do not stop once the ansatz checks out. Substitute the equation's own symmetry into itself to generate a second relation, then eliminate cross-evaluations by differentiating and back-substituting until a single self-contained ODE in one variable remains, and solve that ODE for the general form.

▷ **Cue:** a functional or differential equation relates f (or f') at x to f (or f') at an involution of x, such as a/x, $1/x$, or $-x$ $\longrightarrow$ **Move:** substitute the involution into the equation itself to generate a second equation, then differentiate and re-substitute to eliminate cross-evaluations, landing on a self-contained ODE in one variable

▷ **Cue:** after algebraic elimination, an expression of the form $xf''/f + f'/f - xf'^2/f^2$ appears $\longrightarrow$ **Move:** recognise it as the derivative of xf'/f via the quotient rule and integrate directly, rather than attacking the second-order ODE by other means

2005 B4

2005 B4. For positive integers m and n, let $f(m,n)$ denote the number of n-tuples $(x_1, x_2, \ldots, x_n)$ of integers such that $|x_1| + |x_2| + \cdots + |x_n| \leq m$. Show that $f(m,n) = f(n,m)$.

Fields: combinatorics (lattice-point counting, stars and bars, double counting); recurrences; generating functions **Position:** B4 — mid-paper; genuinely approachable once you refuse to count the whole set in one go **Answer:** $f(m,n) = \sum_{k\geq 0} 2^k \binom{m}{k}\binom{n}{k}$, which is visibly symmetric in m and n.

Layer 1 THE WHAT — surface analysis

The natural attempt. The strong student starts by computing. $f(m,1)$: integers x with $|x| \leq m$, so $2m+1$. $f(1,n)$: tuples with $\sum |x_i| \leq 1$, which is the zero tuple plus one coordinate equal to ± 1, so $2n+1$. Symmetric. Good. $f(2,2)$: $|x| + |y| \leq 2$ is a diamond with $1+4+8 = 13$ points. $f(2,3)$ and $f(3,2)$ both come out as 25 after some effort.

Then the attempt tries to find a formula. The natural first cut is by the *value of the sum*: let $s = |x_1| + \cdots + |x_n|$ and count tuples with sum exactly s, then add up $s = 0, 1, \ldots, m$. For fixed s, the number of n-tuples of *nonnegative* integers summing to s is $\binom{s+n-1}{n-1}$, but the signs mess this up: a coordinate equal to 0 has one sign, a nonzero one has two. So the student writes "$2^n\binom{s+n-1}{n-1}$", sees it overcounts zeros, and starts correcting by inclusion–exclusion on which coordinates are zero. After a page of this they have something like

$$f(m,n) = \sum_{s=0}^{m} \sum_{j} (-1)^j \cdots$$

with no visible symmetry at all. The wall is real: slicing by s treats m as a range of sums and n as a number of slots, two completely different roles, and a formula built that way has no reason to look symmetric even though it is.

The model chain. **Step 1: slice by the number of nonzero coordinates.** For $k \geq 0$, let $f(m,n,k)$ count the n-tuples with $\sum |x_i| \leq m$ and exactly k nonzero entries. Then $f(m,n) = \sum_k f(m,n,k)$.

Step 2: count $f(m,n,k)$ by three independent choices.

- Which k positions are nonzero: $\binom{n}{k}$ ways.
- The sign of each nonzero entry: 2^k ways.
- The absolute values $a_1, \ldots, a_k \geq 1$ (in the chosen positions, in order) with $a_1 + \cdots + a_k \leq m$.

These choices determine the tuple and are determined by it, so $f(m,n,k)$ is the product of the three counts.

Step 3: the third count is $\binom{m}{k}$. Map $(a_1, \ldots, a_k)$ to the partial sums $b_j = a_1 + \cdots + a_j$. Because every $a_j \geq 1$, we get $1 \leq b_1 < b_2 < \cdots < b_k \leq m$. Conversely any k-element subset $\{b_1 < \cdots < b_k\} \subseteq \{1, \ldots, m\}$ gives back $a_1 = b_1$, $a_j = b_j - b_{j-1} \geq 1$, with $a_1 + \cdots + a_k = b_k \leq m$. This is a bijection, so the count is $\binom{m}{k}$ (zero if $k > m$). For $k = 0$ it is $1 = \binom{m}{0}$, the empty tuple of absolute values.

Step 4: conclude.

$$f(m,n) = \sum_{k \geq 0} 2^k \binom{n}{k} \binom{m}{k},$$

and each summand is unchanged when m and n are swapped. Hence $f(m,n) = f(n,m)$. □

Check: $f(2,3) = 1 + 2 \cdot 3 \cdot 2 + 4 \cdot 3 \cdot 1 = 25$, matching the hand count.

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains agree that $f(m,n)$ must be split into pieces that can each be counted. They split at the choice of *what to split by*.

(N) By the total $s = \sum |x_i|$: this uses m as an upper limit of a sum over s, and n as the number of parts.

(M) By the number k of nonzero coordinates: this makes n enter only through "choose k of n positions" and m enter only through "choose k of m cut points".

> *The answer is supposed to be symmetric, so I want a decomposition in which m and n each appear in the same way. Slicing by the sum makes m a summation limit and n a number of slots — they can't match. What statistic of a tuple sees n and m as equals? The number of nonzero entries: I choose k of the n slots, and I choose k values whose partial sums are k of the numbers* $1, \ldots, m$. *Both are "choose k".*

> *Zeros and signs are fighting each other — a zero has one sign, a nonzero entry has two. Rather than correct for that by inclusion–exclusion, I'll fix which entries are zero up front, and then every remaining entry has exactly two signs and a positive size.*

▷ **Cue:** a counting formula must be shown symmetric in two parameters that play visibly different roles in the definition ⟶ **Move:** look for a decomposition by a statistic k under which each parameter contributes a factor of the same shape (e.g. $\binom{m}{k}\binom{n}{k}$)

▷ **Cue:** counting integer tuples with absolute values, where zeros and nonzeros behave differently under sign ⟶ **Move:** condition first on the set of nonzero positions; then signs contribute a clean 2^k

▷ **Cue:** ordered k-tuples of positive integers with sum $\leq m$ $\longrightarrow$ **Move:** partial sums give a bijection with k-subsets of $\{1,\dots,m\}$: the count is $\binom{m}{k}$

▷ **Cue:** "show $F(m,n) = F(n,m)$" for a counting function $\longrightarrow$ **Move:** alternatively find a recurrence symmetric in m,n with symmetric boundary values, or a generating function symmetric in x,y

Layer 3 THE WHY

Why is "number of nonzero coordinates" the natural slice, and not a lucky one? Start from what the problem is asking. A statement $f(m,n) = f(n,m)$ is a claim that two different-looking descriptions define the same number. The definition treats n as a *dimension* (how many slots) and m as a *size* (how much total absolute value). To prove symmetry you need a formula in which the dimension and the size are indistinguishable. So the real question is: in what sense is a "size budget of m" like "m slots"?

Stars and bars answers that. A budget of m spent in positive integer chunks is the same thing as choosing where to cut a row of m units — that is, choosing a subset of $\{1,\dots,m\}$. The budget *is* a set of m positions. Now both m and n are sets of positions, and the tuple is described by choosing k of each: k of the n coordinate slots are switched on, and k of the m budget positions are cut points. The only thing that isn't "choose k" is the signs, and those contribute 2^k, which depends on k alone. That is why the slice is by k: it is the one number that both halves of the description share.

Kedlaya's first solution begins "Define $f(m,n,k)$ as the number of n-tuples … exactly k of $x_1,\dots,x_n$ are nonzero." Here is where that comes from: the moment you try to count, you hit the sign asymmetry between zero and nonzero entries (that is precisely where the natural attempt drowned in inclusion–exclusion). The fix for any quantity that behaves differently on two classes is to condition on the class first. Conditioning on the zero set cures the sign problem *and* produces the symmetric formula; the two needs point at the same move.

The official solution states the stars-and-bars count $\binom{m}{k}$ with a sentence about "adding stars at the end". The partial-sum bijection in Step 3 is the same thing written so you can check it: the constraint $a_j \geq 1$ is exactly what makes the partial sums strictly increasing, and $\sum a_j \leq m$ is exactly $b_k \leq m$. Each hypothesis is used once.

There are two other routes worth knowing, because they transfer. Kuperberg's: split tuples with $\sum |x_i| \leq m$ into those with sum $< m$ (that is $f(m-1,n)$), those with sum $= m$ and $x_n \geq 0$ (drop x_n: $f(m,n-1)$), and those with sum $= m$ and $x_n < 0$ (drop x_n, the remaining sum is $m - |x_n| \leq m-1$: $f(m-1,n-1)$). So $f(m,n) = f(m-1,n) + f(m,n-1) + f(m-1,n-1)$, a recurrence symmetric in m,n, with $f(m,0) = f(0,n) = 1$; induction on $m+n$ finishes. Martin's generating function $\sum f(m,n)x^m y^n = 1/(1-x-y-xy)$ is the same fact packed into one symmetric expression. All three proofs are exhibiting a structure in which m and n enter identically; the combinatorial one is the cheapest.

Caveat. Verified by direct enumeration: for all $1 \leq m,n \leq 5$, the brute-force count of $f(m,n)$ equals $\sum_k 2^k \binom{m}{k}\binom{n}{k}$ and equals $f(n,m)$.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is the word "Show that $f(m,n) = f(n,m)$" against a definition where m and n play unequal roles. On seeing it, write: *I need a formula where m and n enter the same way.*

2. Decompose by a statistic shared by both roles; here the number k of nonzero coordinates, giving $\binom{n}{k} \cdot 2^k \cdot \binom{m}{k}$.

3. Hesitation comes when the signs of zero entries break the clean 2^n factor. Reinforce: a quantity behaving differently on two classes means condition on the class first, not correct afterwards.

4. The trap is slicing by the value of the sum s and producing an asymmetric inclusion–exclusion formula that is true but useless for the symmetry.

5. Faster recognition: when asked to prove a hidden symmetry of a count, ask immediately *which description would make each parameter a "choose k from a set of that size"*. Budgets become sets via stars and bars.

6. Reusable decision rule: *to prove a symmetry $F(m,n) = F(n,m)$, find a decomposition in which each summand is symmetric — typically by the size k of a common sub-choice — or a recurrence and boundary conditions that are symmetric; never expect symmetry to appear from a formula whose derivation treated the parameters differently.*

2005 B5

2005 B5. Let $P(x_1,\dots,x_n)$ denote a polynomial with real coefficients in the variables $x_1,\dots,x_n$, and suppose that

$$\left(\frac{\partial^2}{\partial x_1^2}+\cdots+\frac{\partial^2}{\partial x_n^2}\right)P(x_1,\dots,x_n)=0 \quad \text{(identically)}$$

and that

$$x_1^2+\cdots+x_n^2 \text{ divides } P(x_1,\dots,x_n).$$

Show that $P=0$ identically.

Fields: algebra (polynomial rings, homogeneous decomposition); multivariable calculus (the Laplacian); harmonic polynomials **Position:** B5 — hard; the official solution is six lines, but the six lines are all decisions **Answer:** P must be 0: writing a homogeneous $P = Q^m R$ with $Q \nmid R$, one gets $\nabla^2 P = Q^{m-1}\big(Q\,\nabla^2 R + 2m(n+2m-2+2e)R\big)$ with $e = \deg R$, forcing $Q \mid R$.

Layer 1 THE WHAT — surface analysis

Notation throughout: $Q = x_1^2+\cdots+x_n^2$ and $\nabla^2 = \sum_i \partial^2/\partial x_i^2$.

The natural attempt. The analyst's reflex: $\nabla^2 P = 0$ means P is harmonic, and harmonic functions obey the maximum principle — on a bounded region, a harmonic function attains its maximum and minimum on the boundary. If P vanished on a boundary, P would vanish inside. So the student looks for where P vanishes. $Q \mid P$ means $P = QS$, so P vanishes wherever Q does. Over $\mathbb{R}$, $Q = 0$ only at the origin. One point is not the boundary of anything. Wall.

The second reflex is to write $P = QS$ and expand:

$$\nabla^2(QS) = S\,\nabla^2 Q + 2\nabla Q\cdot\nabla S + Q\,\nabla^2 S = 2nS + 4\sum_i x_i\frac{\partial S}{\partial x_i} + Q\,\nabla^2 S.$$

Setting this to zero gives $2nS + 4\sum x_i S_{x_i} = -Q\nabla^2 S$. The student stares at the middle term $\sum x_i\,\partial S/\partial x_i$ and does not know what to do with it; tries small cases ($n=1$: $P = x^2 S(x)$ with $P''=0$ forces $\deg P \le 1$, so $P=0$; fine, but it doesn't generalise); and stalls. The expansion was correct. What was missing is a way to make that middle term into a multiple of S.

The model chain. **Step 1: reduce to homogeneous** P**.** Write $P = \sum_d P_d$ with P_d homogeneous of degree d. ∇^2 sends a homogeneous polynomial of degree d to one of degree $d-2$, so $\nabla^2 P = \sum_d \nabla^2 P_d$ is the homogeneous decomposition of $\nabla^2 P$; since it is 0, every $\nabla^2 P_d = 0$. If $P = QS$, write $S = \sum_d S_d$; since Q is homogeneous of degree 2, QS_{d-2} is the degree-d part of P, so $P_d = QS_{d-2}$ and $Q \mid P_d$. Thus each P_d satisfies both hypotheses, and it suffices to prove every such homogeneous polynomial is 0.

Step 2: extract all factors of Q**.** Suppose, for contradiction, $P \neq 0$ is homogeneous of degree d, harmonic, and divisible by Q. Since $P \neq 0$, the power of Q dividing P is bounded (by $d/2$); let

$m \geq 1$ be the largest, and write $P = Q^m R$ with $Q \nmid R$. R is homogeneous of degree $e = d - 2m$ (a factor of a homogeneous polynomial is homogeneous), and $R \neq 0$.

Step 3: Euler's identity. For R homogeneous of degree e, $R(tx) = t^e R(x)$; differentiating in t at $t = 1$ gives $\sum_i x_i \, \partial R/\partial x_i = eR$.

Step 4: compute $\nabla^2(Q^m R)$. First, $\partial_i Q^m = 2m x_i Q^{m-1}$ and

$$\partial_i^2 Q^m = 2m Q^{m-1} + 4m(m-1) x_i^2 Q^{m-2}, \qquad \text{so} \qquad \nabla^2 Q^m = \big(2mn + 4m(m-1)\big) Q^{m-1}.$$

By the product rule $\nabla^2(FG) = G\nabla^2 F + 2\nabla F \cdot \nabla G + F\nabla^2 G$,

$$\nabla^2(Q^m R) = \big(2mn + 4m(m-1)\big) Q^{m-1} R + 4m Q^{m-1} \sum_i x_i \frac{\partial R}{\partial x_i} + Q^m \nabla^2 R.$$

Using Step 3,

$$0 = \nabla^2 P = Q^{m-1}\Big(Q\,\nabla^2 R + c\,R\Big), \qquad c = 2m\,(n + 2m - 2 + 2e).$$

Step 5: contradiction. The polynomial ring has no zero divisors and $Q^{m-1} \neq 0$, so $Q\nabla^2 R = -cR$. Since $m \geq 1$, $n \geq 1$, $e \geq 0$, we have $n + 2m - 2 + 2e \geq 1$, so $c > 0$. Hence $R = -c^{-1} Q \nabla^2 R$ is divisible by Q, contradicting the maximality of m. So $P = 0$. □

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains write P as Q times something and apply the product rule for ∇^2. They split at the term $\sum_i x_i \, \partial S/\partial x_i$.

(N) The natural attempt treats S as a general polynomial, so the Euler term is a new, unrelated polynomial and the equation relates three different things.

(M) The model first makes everything homogeneous, so the Euler term is e times the cofactor, and the equation collapses to "Q divides a nonzero multiple of the cofactor".

A second, smaller split: the model pulls out the *maximal* power of Q, which is what turns "Q divides the cofactor" into a contradiction rather than an infinite regress.

> *The term $\sum x_i \, \partial_i S$ is Euler's operator. On a homogeneous polynomial of degree e it is multiplication by e. Both hypotheses respect degree — ∇^2 drops it by 2, Q raises it by 2 — so I lose nothing by assuming P homogeneous, and then that term becomes a number times S.*

> *Once I get "Q divides the cofactor", I could peel off another Q and repeat. That's a descent; to make it terminate cleanly, take the largest power Q^m at the start, and then "Q divides R" is the contradiction.*

> *The only way this could go wrong is if the coefficient in front of R were zero. Check its sign: $2m(n + 2m - 2 + 2e)$, and every piece is nonnegative with $n \geq 1$. Positive. That positivity is where real coefficients and characteristic zero are secretly used.*

▷ **Cue:** a condition on a polynomial built from a homogeneous operator (like ∇^2) and a homogeneous polynomial (like $x_1^2 + \cdots + x_n^2$) ⟶ **Move:** split into homogeneous components; the conditions hold componentwise

▷ **Cue:** an expression $\sum_i x_i \, \partial F/\partial x_i$ appears ⟶ **Move:** it is Euler's operator: if F is homogeneous of degree e, replace it by eF

▷ **Cue:** "Q divides P" and you are trying to show $P = 0$ by descent $\longrightarrow$ **Move:** write $P = Q^m R$ with m maximal and $Q \nmid R$; aim to prove $Q \mid R$

▷ **Cue:** an identity of the form $Q \cdot A = c \cdot R$ in a polynomial ring $\longrightarrow$ **Move:** if the constant c is nonzero, $Q \mid R$; check the constant's sign carefully

Layer 3 THE WHY

Why homogeneous? Go to the definitions. ∇^2 maps each monomial of degree d to a combination of monomials of degree $d-2$; multiplying by Q maps degree d to degree $d+2$. Both operations commute with the scaling $x \mapsto tx$ up to a power of t. When every operator in a problem respects a grading, the problem splits along that grading, and each piece is a smaller copy of the same problem with one extra property for free. You do not need to be told to decompose; you need only ask "does every hypothesis respect degree?" and notice the answer is yes.

Why does homogeneity kill the troublesome term? Because the troublesome term is $\nabla Q \cdot \nabla S = 2\sum x_i\, \partial_i S$, and $\nabla Q = 2x$ is the radial vector field. Differentiating along the radial direction is the same as differentiating with respect to scale, and a homogeneous function's rate of change with respect to scale is its degree times itself. That is Euler's identity, and it is Step 3 above: it is literally the derivative of $R(tx) = t^e R(x)$ at $t = 1$. So the appearance of $x \cdot \nabla$ in the natural attempt is itself the cue: Q is the squared radius, its gradient points radially, and the tool for radial derivatives is homogeneity.

Kedlaya's first solution begins "Since Q is homogeneous, P is divisible by Q if and only if each of the homogeneous components of P is divisible by Q. It is thus sufficient to solve the problem in case P itself is homogeneous." Here is where that comes from: the paragraph above. The next line, "Suppose that we have a factorization $P = Q^m R \ldots$ not divisible by Q", comes from asking what a single application of the identity gives you — one extra factor of Q — and wanting to avoid iterating.

Caveat. The official first solution's displayed computation has two slips: it calls R "homogeneous of degree d" where it has degree $d - 2m$, and it gives the coefficient as $2mn + 4md$, omitting the term $4m(m-1)$ that comes from differentiating Q^m twice. Neither affects the conclusion, since the correct coefficient $2m(n + 2m - 2 + 2e)$ is still positive. The identity $\nabla^2(Q^m R) = Q^m \nabla^2 R + 2m(n + 2m - 2 + 2e)Q^{m-1}R$ was verified symbolically for $n \le 3$, $m \le 3$, $e \le 3$ with generic R, and the statement of the problem was checked directly by solving the linear system $\nabla^2(QS) = 0$ for generic homogeneous S in $n = 3$ variables, total degree 2 to 6: the only solution is $S = 0$.

Why is the conclusion not a triviality? Because it fails in general characteristic: over a field of characteristic p, $(x_1^2 + \cdots + x_n^2)^p = x_1^{2p} + \cdots + x_n^{2p}$ is harmonic. In the computation, that is $c = 2m(\ldots)$ becoming 0 mod p. So the positivity check in Step 5 is not decoration — it is the entire content of "real coefficients".

Two other proofs are worth carrying away. Stanley's: if $P = \sum c_\alpha x^\alpha$, then applying the differential operator $P(\partial)$ to P has constant term $\sum_\alpha \alpha!\, c_\alpha^2$ (the official solution omits the harmless factor $\alpha! = \alpha_1! \cdots \alpha_n!$); but $P(\partial) = S(\partial)\nabla^2$, so $P(\partial)P = S(\partial)(\nabla^2 P) = 0$. A sum of positive multiples of squares vanishing forces every $c_\alpha = 0$. This is an inner-product argument: $\langle F, G\rangle = F(\partial)G|_0$ is positive definite, and multiplication by Q is adjoint to ∇^2, so harmonic polynomials are orthogonal to everything divisible by Q — including themselves. Elkies's solution is the same fact as a dimension count, the decomposition of homogeneous polynomials into Q^j times harmonics.

The natural attempt's maximum-principle idea deserves a last word. Over $\mathbb{R}$ it cannot see Q's divisibility at all, because divisibility is algebraic information that the real zero set (one point) does not record. When a hypothesis is algebraic, reach for algebraic tools first.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is that both hypotheses are *homogeneous*: ∇^2 lowers degree by exactly 2, Q has degree exactly 2. On seeing it, write: *WLOG P homogeneous of degree d.*
2. Homogeneity suggests Euler's identity $\sum x_i \partial_i R = eR$, which converts the cross term of the product rule into a multiple of R; combined with a maximal factorisation $P = Q^m R$, this gives a descent contradiction.
3. Hesitation occurs at the cross term $\nabla Q \cdot \nabla S$, which looks like new information. It is reinforced away by recognising $\nabla Q = 2x$ as the radial field and $x \cdot \nabla$ as Euler's operator.
4. The trap is the maximum principle: harmonic plus a vanishing condition sounds like it should give zero, but over $\mathbb{R}$ the zero set of Q is a single point and the analytic tool has nothing to act on.
5. Recognise faster by checking, for every operator and every polynomial in the hypotheses, how it interacts with degree. If all of them shift degree by fixed amounts, decompose first.
6. Reusable decision rule: *when every hypothesis respects a grading, reduce to a single graded piece; when $x \cdot \nabla$ appears, replace it by the degree; and when a divisibility conclusion would allow infinite descent, factor out the maximal power at the start so that one more factor is a contradiction.*

2005 B6

2005 B6. Let S_n denote the set of all permutations of the numbers $1, 2, \ldots, n$. For $\pi \in S_n$, let $\sigma(\pi) = 1$ if π is an even permutation and $\sigma(\pi) = -1$ if π is an odd permutation. Also, let $\nu(\pi)$ denote the number of fixed points of π. Show that

$$\sum_{\pi \in S_n} \frac{\sigma(\pi)}{\nu(\pi) + 1} = (-1)^{n+1} \frac{n}{n+1}.$$

Fields: combinatorics (signed permutation sums); linear algebra (determinants, eigenvalues of rank-one perturbations); calculus ($\frac{1}{k+1} = \int_0^1 x^k \, dx$) **Position:** B6 — the last problem, and hard to start; once the two translations below are made it collapses to a four-line computation **Answer:** The sum equals $\int_0^1 \det J_x \, dx$, where J_x has x on the diagonal and 1 elsewhere; $\det J_x = (x + n - 1)(x - 1)^{n-1}$, and the integral is $(-1)^{n+1} \frac{n}{n+1}$.

Layer 1 THE WHAT — surface analysis

The natural attempt. The combinatorialist groups permutations by their set of fixed points. A permutation with exactly i fixed points is a choice of the i fixed points ($\binom{n}{i}$ ways) and a derangement of the other $n - i$. Its sign is the sign of that derangement. So

$$\sum_{\pi} \frac{\sigma(\pi)}{\nu(\pi) + 1} = \sum_{i=0}^{n} \binom{n}{i} \frac{E_{n-i}}{i+1}, \qquad E_k = \sum_{\text{derangements } \delta \text{ of } k} \sigma(\delta).$$

Now the student needs E_k, the signed count of derangements. Small cases: $E_0 = 1$ (empty permutation), $E_1 = 0$, $E_2 = -1$ (one transposition), $E_3 = 2$ (two 3-cycles, both even). Guess $E_k = (-1)^{k-1}(k-1)$. To prove it: every π is determined by its fixed set T and a derangement of the complement, and the sum of σ over all of S_k is 0 for $k \geq 2$ (the even and odd permutations are equinumerous), so $\sum_j \binom{k}{j} E_{k-j} = [k \leq 1]$. Inverting this binomial transform — inclusion–exclusion — gives $E_k = \sum_j (-1)^{k-j} \binom{k}{j} [j \leq 1] = (-1)^k + (-1)^{k-1} k = (-1)^{k-1}(k-1)$.

Then the sum becomes

$$\sum_{i=0}^{n} \binom{n}{i} \frac{(-1)^{n-i-1}(n-i-1)}{i+1},$$

and with $n-i-1 = n-(i+1)$ and $\frac{1}{i+1}\binom{n}{i} = \frac{1}{n+1}\binom{n+1}{i+1}$ this splits as $\sum(-1)^{n-i-1}\binom{n}{i}\cdot\frac{n}{i+1}$ minus $\sum(-1)^{n-i-1}\binom{n}{i}$, both of which are alternating binomial sums equal to known values; the answer drops out.

This attempt *works*. It is Kedlaya's second solution. But it costs a guess, an inversion formula, and two binomial identities — three separate places to make a sign error on the last problem of the paper — and none of it explains why the answer is so simple.

The model chain. **Step 1: turn the fraction into an integral.** For an integer $k \geq 0$, $\frac{1}{k+1} = \int_0^1 x^k\,dx$. Hence

$$\sum_{\pi\in S_n}\frac{\sigma(\pi)}{\nu(\pi)+1} = \int_0^1 F_n(x)\,dx, \qquad F_n(x) = \sum_{\pi\in S_n}\sigma(\pi)\,x^{\nu(\pi)}$$

(a finite sum, so the interchange is immediate).

Step 2: recognise F_n as a determinant. Let J_x be the $n\times n$ matrix with $(J_x)_{ii} = x$ and $(J_x)_{ij} = 1$ for $i \neq j$. The Leibniz formula gives

$$\det J_x = \sum_{\pi\in S_n}\sigma(\pi)\prod_{i=1}^{n}(J_x)_{i,\pi(i)}.$$

The factor $(J_x)_{i,\pi(i)}$ is x when $\pi(i) = i$ and 1 otherwise, so the product is $x^{\nu(\pi)}$. Thus $F_n(x) = \det J_x$.

Step 3: compute the determinant. $J_x = (x-1)I + U$, where U is the all-ones matrix. U has rank 1, with $U\mathbf{1} = n\mathbf{1}$ ($\mathbf{1}$ the all-ones vector) and $Uv = 0$ for every v in the $(n-1)$-dimensional space $\{v : \sum v_i = 0\}$. These together span $\mathbb{R}^n$, so U has eigenvalues n (once) and 0 ($n-1$ times), and J_x has eigenvalues $x-1+n$ (once) and $x-1$ ($n-1$ times) with the same eigenvectors. The determinant is the product of eigenvalues:

$$\det J_x = (x+n-1)(x-1)^{n-1}.$$

Step 4: integrate. Substitute $y = 1-x$, so $(x-1)^{n-1} = (-1)^{n-1}y^{n-1}$ and $x+n-1 = n-y$:

$$\int_0^1 (x+n-1)(x-1)^{n-1}\,dx = (-1)^{n-1}\int_0^1 (n-y)y^{n-1}\,dy = (-1)^{n-1}\Big(1-\frac{1}{n+1}\Big) = (-1)^{n+1}\frac{n}{n+1}. \qquad \square$$

Layer 2 THE HOW — the bifurcation

Bifurcating point. Both chains must deal with two awkward features of the summand: the $\sigma(\pi)$ and the $\frac{1}{\nu(\pi)+1}$. They split at the first line.

(N) Organise the sum by fixed-point set and compute signed derangement numbers.

(M) Replace $\frac{1}{\nu+1}$ by $\int_0^1 x^\nu\,dx$; then $\sum\sigma(\pi)x^{\nu(\pi)}$ is a determinant.

> *$\frac{1}{k+1}$ is awkward to sum, but x^k is easy: $\frac{1}{k+1} = \int_0^1 x^k\,dx$. So the sum is the integral of $\sum_\pi \sigma(\pi)x^{\nu(\pi)}$ — a generating polynomial for fixed points, weighted by sign.*

> *A sum over all permutations of $\sigma(\pi)$ times a product that depends on π — that's the Leibniz formula for a determinant. I need a matrix whose entry at $(i,\pi(i))$ is x when $\pi(i) = i$ and 1 otherwise: x on the diagonal, 1 off it.*

> *x on the diagonal, 1 elsewhere is (all-ones) plus $(x-1)I$. The all-ones matrix has rank one, so I know its eigenvalues without computing anything.*

- ▷ **Cue:** a sum over permutations weighted by the sign $\sigma(\pi)$ $\longrightarrow$ **Move:** look for a determinant: $\sum_\pi \sigma(\pi) \prod_i a_{i,\pi(i)}$
- ▷ **Cue:** a summand containing $\frac{1}{k+1}$ (or $\frac{1}{k+a}$) with k a nonnegative integer statistic $\longrightarrow$ **Move:** write $\frac{1}{k+1} = \int_0^1 x^k\,dx$ and sum the generating polynomial first
- ▷ **Cue:** a matrix that is a constant times I plus a rank-one matrix $\longrightarrow$ **Move:** eigenvalues are the constant ($n-1$ times) and the constant plus the trace of the rank-one part
- ▷ **Cue:** a statistic counting fixed points, $\nu(\pi)$, inside a permutation sum $\longrightarrow$ **Move:** put a variable x on the diagonal of a matrix: $x^{\nu(\pi)}$ is what the diagonal contributes

Layer 3 THE WHY

Why integrate? Because the denominator $\nu(\pi)+1$ is the only thing standing between this sum and a polynomial identity. A sum of $c_\pi x^{\nu(\pi)}$ is a polynomial, and polynomials are what algebra hands you in closed form; a sum of $c_\pi/(\nu(\pi)+1)$ is not. The identity $\int_0^1 x^k\,dx = \frac{1}{k+1}$ is the definition of the integral of a monomial read backwards. Whenever a combinatorial sum has $\frac{1}{k+1}$ attached to a statistic k, the statistic wants to be an exponent, and the $\frac{1}{k+1}$ is an integral waiting to happen. The problem's phrasing even puts the "+1" there for you: $\frac{1}{\nu+1}$, not $\frac{1}{\nu}$ — exactly what $\int_0^1$ produces, and a signal that this was built backwards from an integral.

Why a determinant? Go to the definition. The determinant is, by definition, $\sum_\pi \sigma(\pi) \prod_i a_{i,\pi(i)}$. The sign of a permutation appears in very few places in mathematics, and the determinant is the dominant one. So the moment $\sigma(\pi)$ appears, ask: can the rest of the summand be written as $\prod_i a_{i,\pi(i)}$? Here the rest is $x^{\nu(\pi)}$, and $\nu(\pi)$ counts the i with $\pi(i)=i$ — exactly the i for which the product picks a diagonal entry. So put x on the diagonal and 1 elsewhere. The matrix is forced by the statistic; there is no choice to make.

Kedlaya's first solution begins "Let J_x be the matrix with x's on the diagonal and 1's elsewhere." Here is where that comes from: the two paragraphs above, in reverse order. The official write-up states the matrix first and reveals the integral last; the thinking happens the other way round.

Why is the determinant easy? J_x differs from a multiple of the identity by the all-ones matrix U, and $U = \mathbf{1}\mathbf{1}^\mathsf{T}$ has rank one. A rank-one matrix uv^T kills everything orthogonal to v and has one nonzero eigenvalue, equal to $v^\mathsf{T}u$ — which is its trace. Here the trace is n. Adding $(x-1)I$ shifts every eigenvalue by $x-1$. If eigenvalues feel beyond A-level, the same answer comes from row operations: add every row of J_x to the first, making the first row all equal to $x+n-1$; factor that out; subtract the new first row from each other row, leaving an upper triangular matrix with diagonal $1, x-1, \ldots, x-1$.

What does the natural attempt teach? That the fixed-point grouping is the determinant expansion done by hand. $E_k = (-1)^{k-1}(k-1)$ is $\det(U-I)$ for the $k\times k$ all-ones matrix U, i.e. $F_k(0)$; the binomial convolution is expanding $\det((x-1)I+U)$ in powers of x. The attempt rebuilds, from scratch and with sign risk at every step, a structure that linear algebra hands over for free. That is the zombie maths this problem punishes: correct, laborious, and blind to the one-line identification that makes it unnecessary.

Caveat. Verified by exhaustive computation in exact rational arithmetic: for $n = 1, \ldots, 7$ the sum over all of S_n equals $(-1)^{n+1}\frac{n}{n+1}$.

Layer $+\alpha$ BACKTESTING

1. The first useful cue is $\sigma(\pi)$ summed over all of S_n. On seeing it, write the Leibniz formula $\det A = \sum_\pi \sigma(\pi) \prod a_{i,\pi(i)}$ and ask what matrix produces the rest of the summand.

2. The second cue, $\frac{1}{\nu(\pi)+1}$, suggests $\int_0^1 x^\nu \, dx$; together they suggest integrating $\det J_x$ over $[0,1]$.
3. Hesitation occurs at the fraction: it does not look like a product of matrix entries. It is reinforced away by the rule that $\frac{1}{k+1}$ attached to an integer statistic is $\int_0^1 x^k$.
4. The trap is grouping by fixed points and computing signed derangement numbers. It succeeds, but through a guessed formula, an inversion, and two alternating binomial sums on the hardest problem of the paper, with a sign error available at every stage.
5. Recognise faster by keeping a short list of where $\sigma(\pi)$ lives — determinants, alternating sums, Vandermonde-type identities — and checking it the instant a sign of a permutation appears.
6. Reusable decision rule: *when a sum carries a sign $\sigma(\pi)$, try to realise it as a determinant; when it carries $\frac{1}{k+1}$ for an integer statistic k, integrate x^k over $[0,1]$. Turn the statistic into an exponent, sum the polynomial by structure, then integrate.*