MathCastle

Putnam

Complete course booklet • 23 topics

Contents

1 Everything the Putnam Tests **4**

1.1 The Syllabus, Honestly 4

2 Integrals I: Feynman's Trick **5**

2.1 Differentiating Under the Integral 5

2.2 Proofs & Why It Matters 6

2.3 Going Deeper: Worked Problems 6

3 Integrals II: King's Rule, Queen's Rule, and Glasser **7**

3.1 Symmetry Substitutions 7

3.2 Proofs & Why It Matters 8

3.3 Going Deeper: Worked Problems 9

4 Integrals III: Frullani, Beta, Gamma, Weierstrass & Series **10**

4.1 The Named Integrals 10

4.2 Proofs & Why It Matters 11

4.3 Going Deeper: Worked Problems 11

4.4 More Named Results 12

4.5 Proofs & Why It Matters 13

5 Sequences & Series **13**

5.1 Summing and Testing 13

5.2 Proofs & Why It Matters 15

5.3 Going Deeper: Worked Problems 15

6 Hard Sequences: Closed Forms, Rates & Periodicity **16**

6.1 Sequences the Putnam Actually Sets 16

6.2 Proofs & Why It Matters 17

6.3 Going Deeper: Worked Problems 17

7 Limits, Continuity & the Big Theorems **17**
7.1 The Theorems That Do the Work . . . 18
7.2 Proofs & Why It Matters . . . 18
7.3 Going Deeper: Worked Problems . . . 19

8 Differential Equations on the Putnam **19**
8.1 The ODE Toolkit the Exam Assumes . . . 19
8.2 Proofs & Why It Matters . . . 21

9 Inequalities & Convexity **21**
9.1 The Inequality Toolkit . . . 21
9.2 Proofs & Why It Matters . . . 22
9.3 Going Deeper: Worked Problems . . . 22

10 Linear Algebra I: Determinants & Structured Matrices **23**
10.1 Determinants the Putnam Way . . . 23
10.2 Proofs & Why It Matters . . . 24
10.3 Going Deeper: Worked Problems . . . 26

11 Linear Algebra II: Spectra, Cayley–Hamilton & Finite Fields **26**
11.1 Spectral Reasoning . . . 26
11.2 Proofs & Why It Matters . . . 28
11.3 Going Deeper: Worked Problems . . . 29

12 Polynomials **30**
12.1 Working With Roots You Never Find . . . 30
12.2 Proofs & Why It Matters . . . 31
12.3 Going Deeper: Worked Problems . . . 31

13 Number Theory **32**
13.1 Proof-Level Number Theory . . . 32
13.2 Proofs & Why It Matters . . . 33
13.3 Going Deeper: Worked Problems . . . 33

14 Combinatorics & Generating Functions **34**
14.1 Counting With Structure . . . 34
14.2 Proofs & Why It Matters . . . 35
14.3 Going Deeper: Worked Problems . . . 35

15 Probability & Expectation **36**
15.1 Expectation Without Distributions . . . 36
15.2 Proofs & Why It Matters . . . 37
15.3 Going Deeper: Worked Problems . . . 37

16 Abstract Algebra Basics **38**
16.1 The Definitions the Putnam Assumes . . . 38
16.2 Proofs & Why It Matters . . . 39
16.3 Going Deeper: Worked Problems . . . 40

17 Functional Equations & Recurrences **40**
17.1 Solving for Functions . . . 40
17.2 Proofs & Why It Matters . . . 41
17.3 Going Deeper: Worked Problems . . . 42

18 Multivariable Calculus on the Putnam **42**
18.1 The Moves That Recur . . . 42
18.2 Why It Works . . . 44

19 Proof Techniques for the Putnam **44**
19.1 The Three Forms, at Putnam Level . . . 44

20 Invariance and Symmetry **45**
20.1 Finding the Invariant . . . 45

21 Extremal Principle and Working Backwards **46**
21.1 The Extremal Principle . . . 46

22 Strategy & Writing a 10 **47**
22.1 How to Take the Exam . . . 47

23 Formulas & Methods **48**
23.1 Analysis . . . 48
23.2 Algebra . . . 49
23.3 Combinatorics, Probability & Geometry . . . 49

1 Everything the Putnam Tests

The complete map of the exam, and where each piece lives in this course

1.1 The Syllabus, Honestly

What the exam is

Twelve problems, two three-hour sessions of six, each scored 0–10. The median score is usually 0 or 1; a 20 typically ranks in the top few hundred nationwide.

Almost all scores on a problem are 0, 1, 9, or 10 — graders reward COMPLETE, rigorous solutions and give almost nothing for partial ideas. A-1 and B-1 are accessible to anyone who has finished this course; A-6 and B-6 are research-adjacent. The winning strategy is not to attempt everything: it is to find the two or three problems you can finish, write them so cleanly that a skeptical grader cannot object, and treat the rest as bonus.

The six pools

ANALYSIS · LINEAR ALGEBRA · ALGEBRA & POLYNOMIALS · NUMBER THEORY · COMBINATORICS & PROBABILITY · GEOMETRY. Every exam samples all six.

Analysis: sequences and series, limits as Riemann sums or derivatives, the intermediate and mean value theorems, Taylor with remainder, and — the Putnam signature — definite integrals with no antiderivative. Linear algebra: structured determinants, eigenvalues and traces, rank, Cayley–Hamilton, matrices over finite fields. Algebra and polynomials: Vieta and Newton, roots of unity, interpolation, functional equations, the definitions of groups and rings. Number theory: orders, valuations, quadratic residues, Diophantine equations, divisor sums. Combinatorics and probability: bijections, generating functions, recurrences, inclusion–exclusion, expectation by indicators, random walks. Geometry: convexity, lattice points, complex numbers, vectors.

The map of this course

Topics 2–4 are the integrals. Topics 5–7 the rest of analysis. Topics 8–9 linear algebra. Topics 10–16 algebra, number theory, combinatorics, probability, abstract algebra, functional equations, geometry. Topic 17 is how to write a 10.

Every unit ends with a practice pool of hand-verified Putnam-level problems; the integral units share the non-elementary pool with the Calculus course. The Linear Algebra and Multivariable courses on this site are the prerequisites this course assumes.

How to use this course

Read a topic, then do its practice pool with a timer — forty minutes per problem, the exam pace. Write full solutions on paper even for numeric answers: the Putnam grades proofs, and the habit of justifying every step is the skill being trained here. Return to the Proofs sections until you can reproduce each argument from memory.

2 Integrals I: Feynman's Trick

Differentiation under the integral sign — the tool that makes impossible integrals routine

2.1 Differentiating Under the Integral

The method in one line

Embed the integral in a family $I(a) = \int f(x,a)\,dx$, differentiate in the parameter, integrate back, and fix the constant from a value of a where $I(a)$ is known.

Why it works: $\dfrac{d}{da}\int f(x,a)\,dx = \int \dfrac{\partial f}{\partial a}\,dx$ whenever f and $\partial_a f$ are continuous (Leibniz rule). The parameter is chosen so that $\partial_a f$ is ELEMENTARY even though f is not — a logarithm in a denominator, an oscillating factor, a power that refuses to integrate. Then $I'(a)$ is a routine integral, $I(a)$ is its antiderivative in a, and one known value pins the constant. The technique is Richard Feynman's name for a trick already in Leibniz; on the Putnam it appears in some form most years.

The model example: $\int_0^1 \frac{x^a-1}{\ln x}dx$

Set $I(a) = \displaystyle\int_0^1 \frac{x^a-1}{\ln x}\,dx$, with $I(0) = 0$.

Step 1 — differentiate in a: $\partial_a \dfrac{x^a-1}{\ln x} = \dfrac{x^a \ln x}{\ln x} = x^a$ — the logarithm cancels. Step 2 — $I'(a) = \int_0^1 x^a\,dx = \dfrac{1}{a+1}$. Step 3 — integrate back: $I(a) = \ln(a+1) + C$, and $I(0) = 0$ gives $C = 0$. Step 4 — so $\int_0^1 \frac{x^2-1}{\ln x}dx = \ln 3$, and in general $\ln(a+1)$. Check: at $a = 1$ the integrand is $\frac{x-1}{\ln x} \approx 1$ near $x = 1$ and $\to 0$ near 0, so a value near $\ln 2 \approx 0.69$ is plausible ✓.

The Dirichlet integral: $\int_0^\infty \frac{\sin x}{x}dx = \frac{\pi}{2}$

The integrand has no elementary antiderivative; the parameter is a DAMPING factor.

Step 1 — define $I(t) = \int_0^\infty e^{-tx}\dfrac{\sin x}{x}\,dx$ for $t \geq 0$; the target is $I(0)$, and $I(t) \to 0$ as $t \to \infty$ (the exponential crushes everything). Step 2 — $I'(t) = -\int_0^\infty e^{-tx}\sin x\,dx = -\dfrac{1}{1+t^2}$ — a Laplace transform, computable by two integrations by parts or by taking the imaginary part of $\int e^{-(t-i)x}dx = \frac{1}{t-i}$. Step 3 — integrate: $I(t) = -\arctan t + C$, and $I(\infty) = 0$ forces $C = \frac{\pi}{2}$. Step 4 — $I(0) = \frac{\pi}{2}$. (The interchange of limit and integral at $t = 0$ needs care — Abel's theorem or dominated convergence — which is exactly the kind of sentence a Putnam grader wants to see.)

Where the parameter hides

If there is no parameter, invent one: replace a constant by a ($e^{-x^2} \to e^{-ax^2}$), an exponent by a

($x^2 \to x^a$), or multiply by e^{-tx} to make a divergent-looking integral converge. Then ask which choice makes $\partial_a f$ elementary. That question, not the calculus, is the whole art.

Try it: a parameter in the exponent

Evaluate $\displaystyle\int_0^\infty \frac{e^{-x}-e^{-2x}}{x}\,dx$ by Feynman's trick.

Answer. Set $I(a)=\int_0^\infty \frac{e^{-x}-e^{-ax}}{x}dx$ with $I(1)=0$. Then $I'(a)=\int_0^\infty e^{-ax}dx=\frac{1}{a}$, so $I(a)=\ln a$ and $I(2)=\ln 2$. (This is also Frullani's theorem, topic 4 — two names for one mechanism.)

2.2 Proofs & Why It Matters

Proof: the Leibniz rule

If f and $\partial_a f$ are continuous on $[c,d]\times[a_0,a_1]$, then $\dfrac{d}{da}\int_c^d f(x,a)\,dx=\int_c^d \partial_a f(x,a)\,dx$.

Form the difference quotient $\dfrac{I(a+h)-I(a)}{h}=\int_c^d \dfrac{f(x,a+h)-f(x,a)}{h}dx$. By the mean value theorem in a, the integrand equals $\partial_a f(x,\xi_x)$ for some ξ_x between a and $a+h$. Continuity of $\partial_a f$ on the compact rectangle makes it uniformly continuous, so $\partial_a f(x,\xi_x)\to\partial_a f(x,a)$ UNIFORMLY in x as $h\to 0$ — and uniform convergence lets the limit pass inside the integral. ■ For infinite intervals one needs a dominating integrable function (dominated convergence); Putnam solutions should say so in one sentence.

Proof: $\int_0^\infty e^{-tx}\sin x\,dx=\frac{1}{1+t^2}$

The Laplace transform that powers the Dirichlet integral.

Integrate by parts twice with $u=\sin x$ then $u=\cos x$: $J=\int_0^\infty e^{-tx}\sin x\,dx=\left[-\frac{e^{-tx}}{t}\sin x\right]_0^\infty+\frac{1}{t}\int_0^\infty e^{-tx}\cos x\,dx=\frac{1}{t}\left(\left[-\frac{e^{-tx}}{t}\cos x\right]_0^\infty-\frac{1}{t}J\right)=\frac{1}{t^2}-\frac{1}{t^2}J$. Solve: $J(1+\frac{1}{t^2})=\frac{1}{t^2}$, so $J=\frac{1}{1+t^2}$. ■ Significance: the same "bounce-back and solve" is the engine of $\int e^x\sin x$ from first-year calculus — Feynman's trick is built from techniques you already own.

2.3 Going Deeper: Worked Problems

Worked: the Gaussian with a parameter

Compute $\displaystyle\int_0^\infty e^{-x^2}\cos(2bx)\,dx$ as a function of b.

Step 1 — let $I(b)$ be the integral; $I(0)=\frac{\sqrt{\pi}}{2}$ (half the Gaussian). Step 2 — differentiate: $I'(b)=-\int_0^\infty 2xe^{-x^2}\sin(2bx)\,dx$. Step 3 — integrate by parts with $u=\sin 2bx$, $dv=2xe^{-x^2}dx$ (so

$v = -e^{-x^2}$): $I'(b) = \left[e^{-x^2}\sin 2bx\right]_0^\infty - 2b\int_0^\infty e^{-x^2}\cos 2bx\,dx = -2bI(b)$. Step 4 — a differential equation! $I'/I = -2b$ gives $I(b) = Ce^{-b^2}$, and $I(0)$ fixes C: $I(b) = \frac{\sqrt{\pi}}{2}e^{-b^2}$. Step 5 — the lesson: sometimes I' is expressed back in terms of I, and Feynman's trick hands you an ODE instead of a direct integral.

Worked: $\int_0^\pi \ln(1 - 2a\cos x + a^2)\,dx$ **for** $|a| < 1$

A Putnam classic with the parameter already present.

Step 1 — $I'(a) = \int_0^\pi \dfrac{2a - 2\cos x}{1 - 2a\cos x + a^2}dx$. Step 2 — the Weierstrass substitution (topic 4) or the identity $\dfrac{2a - 2\cos x}{1 - 2a\cos x + a^2} = \dfrac{1}{a}\left(1 - \dfrac{1 - a^2}{1 - 2a\cos x + a^2}\right)$ reduces it to the standard $\int_0^\pi \dfrac{dx}{1 - 2a\cos x + a^2} = \dfrac{\pi}{1 - a^2}$. Step 3 — so $I'(a) = \dfrac{1}{a}(\pi - \pi) = 0$ for $|a| < 1$. Step 4 — I is constant, and $I(0) = \int_0^\pi \ln 1\,dx = 0$: the integral is 0 for every $|a| < 1$. (For $|a| > 1$, factor out a^2 to get $2\pi\ln|a|$.) A whole family of integrals evaluated by showing the derivative vanishes.

3 Integrals II: King's Rule, Queen's Rule, and Glasser

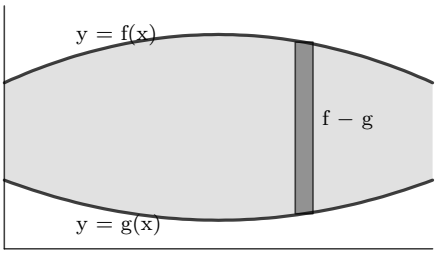
Substitutions that map the interval onto itself, then add

3.1 Symmetry Substitutions

King's rule

$\int_a^b f(x)\,dx = \int_a^b f(a + b - x)\,dx$ — reflect the interval onto itself. Then ADD the two forms.

Reflect the interval onto itself and add the two copies.



The substitution $u = a + b - x$ runs the interval backward with $du = -dx$; flipping the limits back absorbs the sign. Nothing is evaluated yet — the power comes from adding the reflected copy to the original: $2I = \int_a^b [f(x) + f(a + b - x)]\,dx$, and the bracket often collapses. On $[0, \frac{\pi}{2}]$, sine and cosine trade places; on $[0, \pi]$, sin is fixed and cos flips sign; a stray factor x becomes $a + b - x$, so

$x + (a + b - x) = a + b$ is constant. Model: $I = \int_0^{\pi/2} \frac{\sin x}{\sin x + \cos x} dx$; reflected, $\frac{\cos x}{\cos x + \sin x}$; sum $= 1$; $I = \frac{\pi}{4}$.

Queen's rule and Jack's rule

$\int_0^{2a} f = 2\int_0^a f$ if $f(2a - x) = f(x)$, and $= 0$ if $f(2a - x) = -f(x)$. Likewise $\int_{-a}^{a} f = 2\int_0^a f$ for even f, 0 for odd f.

Both are King's rule applied to half the interval: split $[0, 2a]$ at a, reflect the right half onto the left, and compare. Use them to FOLD an interval before doing any real work: $\int_0^{\pi} \sin^3 x\, dx = 2\int_0^{\pi/2} \sin^3 x\, dx = \frac{4}{3}$, and any odd integrand over a symmetric interval is 0 on sight ($\int_{-1}^{1} x^3 e^{x^2} dx = 0$ — no computation). Half the Putnam's "obviously zero" integrals are Jack's rule.

Glasser's master theorem

$\int_{-\infty}^{\infty} f\left(x - \sum_i \frac{a_i}{x - b_i}\right) dx = \int_{-\infty}^{\infty} f(x)\, dx$ for $a_i > 0$. Simplest case: $\int f(x - \frac{1}{x}) = \int f$.

The map $u = x - \frac{1}{x}$ is increasing on $(-\infty, 0)$ and on $(0, \infty)$, covering all of $\mathbb{R}$ on each half-line. Splitting the integral at 0 and substituting on each half, the two Jacobian pieces $\frac{dx}{du}$ ADD to exactly 1 — the identity $\sum_{\text{branches}} \frac{dx}{du} = 1$ is the content of the theorem. So $\int_{-\infty}^{\infty} e^{-(x-1/x)^2} dx = \int e^{-u^2} du = \sqrt{\pi}$: a hideous integrand, evaluated by recognizing its shape. The companion on $(0, \infty)$ is $x \to \frac{1}{x}$ alone, which maps the half-line to itself.

Recognizing the shape

Look for a function of $x - \frac{c}{x}$ (Glasser), an interval whose endpoints add to a constant appearing inside a trig function (King), or a symmetric interval with an even/odd integrand (Jack). The substitution is never the hard part; noticing that it applies is. Train the reflex: every definite integral, before any technique, ask "what does $x \to a + b - x$ do to this?"

Try it: King's rule with a stray x

Evaluate $\int_0^{\pi} \frac{x \sin x}{1 + \cos^2 x}\, dx$.

Answer. Reflect: $I = \int_0^{\pi} \frac{(\pi - x)\sin x}{1 + \cos^2 x} dx$. Add: $2I = \pi \int_0^{\pi} \frac{\sin x}{1 + \cos^2 x} dx$. Substitute $u = \cos x$: $\pi \int_{-1}^{1} \frac{du}{1 + u^2} = \pi \cdot \frac{\pi}{2}$. So $I = \frac{\pi^2}{4}$.

3.2 Proofs & Why It Matters

Proof: Glasser's theorem in the simplest case

Show $\displaystyle\int_{-\infty}^{\infty} f\left(x-\tfrac{1}{x}\right)dx = \int_{-\infty}^{\infty} f(u)\,du.$

On $x>0$ set $u=x-\frac{1}{x}$; solving, $x=\frac{u+\sqrt{u^2+4}}{2}$, an increasing bijection $(0,\infty)\to\mathbb{R}$ with $\frac{dx}{du}=\frac{1}{2}+\frac{u}{2\sqrt{u^2+4}}$. On $x<0$ the same u gives $x=\frac{u-\sqrt{u^2+4}}{2}$, also a bijection onto $\mathbb{R}$, with $\frac{dx}{du}=\frac{1}{2}-\frac{u}{2\sqrt{u^2+4}}$. Adding the two contributions to the integral: $\int f(u)\left[\frac{1}{2}+\frac{u}{2\sqrt{u^2+4}}+\frac{1}{2}-\frac{u}{2\sqrt{u^2+4}}\right]du=\int f(u)\,du$. ■ Significance: the general theorem with $\sum\frac{a_i}{x-b_i}$ works identically — the $n+1$ branches' Jacobians always sum to 1 (a partial-fractions identity). It is a favorite because the integrand looks impossible and the solution is three lines.

Proof: the log-sine integral

Show $\displaystyle\int_0^{\pi/2}\ln(\sin x)\,dx=-\frac{\pi}{2}\ln 2$ using only King and Queen.

King gives $I=\int_0^{\pi/2}\ln(\cos x)dx$. Adding: $2I=\int_0^{\pi/2}\ln(\sin x\cos x)dx=\int_0^{\pi/2}\ln\frac{\sin 2x}{2}dx=\int_0^{\pi/2}\ln(\sin 2x)dx-\frac{\pi}{2}\ln 2$. In the remaining integral substitute $u=2x$: $\frac{1}{2}\int_0^{\pi}\ln(\sin u)du$, and Queen's rule folds $[0,\pi]$ to give $\frac{1}{2}\cdot 2I=I$. Hence $2I=I-\frac{\pi}{2}\ln 2$ and $I=-\frac{\pi}{2}\ln 2$. ■ The integral was never computed; it was trapped in an equation with itself.

3.3 Going Deeper: Worked Problems

Worked: $\int_0^{\infty}\frac{dx}{1+x^4}$ by the $1/x$ reflection

The reflection on $(0,\infty)$ produces a second form to add.

Step 1 — substitute $x\to\frac{1}{x}$: $I=\int_0^{\infty}\dfrac{x^2\,dx}{1+x^4}$. Step 2 — add: $2I=\int_0^{\infty}\dfrac{1+x^2}{1+x^4}dx$. Step 3 — divide numerator and denominator by x^2: $\dfrac{1+x^{-2}}{x^2+x^{-2}}=\dfrac{1+x^{-2}}{(x-x^{-1})^2+2}$. Step 4 — with $u=x-\frac{1}{x}$ (so $du=(1+x^{-2})dx$ and u runs over $\mathbb{R}$): $2I=\int_{-\infty}^{\infty}\dfrac{du}{u^2+2}=\dfrac{\pi}{\sqrt{2}}$. Step 5 — $I=\dfrac{\pi}{2\sqrt{2}}=\dfrac{\pi\sqrt{2}}{4}$. Cross-check with the general formula $\int_0^{\infty}\frac{dx}{1+x^n}=\frac{\pi/n}{\sin(\pi/n)}$ at $n=4$: $\frac{\pi/4}{\sqrt{2}/2}=\frac{\pi\sqrt{2}}{4}$ ✓.

Worked: $\int_0^1\frac{\ln(1+x)}{1+x^2}\,dx$

A Putnam-style integral where the reflection is a TANGENT substitution.

Step 1 — substitute $x=\tan\theta$: $I=\int_0^{\pi/4}\ln(1+\tan\theta)\,d\theta$. Step 2 — King on $[0,\frac{\pi}{4}]$: $\theta\to\frac{\pi}{4}-\theta$, and $1+\tan(\frac{\pi}{4}-\theta)=1+\dfrac{1-\tan\theta}{1+\tan\theta}=\dfrac{2}{1+\tan\theta}$. Step 3 — so $I=\int_0^{\pi/4}\left[\ln 2-\ln(1+\tan\theta)\right]d\theta=\frac{\pi}{4}\ln 2-I$. Step 4 — $I=\dfrac{\pi\ln 2}{8}$. Two substitutions, no antiderivative — and the answer, $\frac{\pi}{8}\ln 2\approx 0.272$, is one of the most famous values in the subject.

4 Integrals III: Frullani, Beta, Gamma, Weierstrass & Series

The named results that finish what symmetry and parameters start

4.1 The Named Integrals

Frullani's theorem

$$\int_0^\infty \frac{f(ax)-f(bx)}{x}\,dx = \big(f(0)-f(\infty)\big)\ln\frac{b}{a}$$ for continuous f with finite limits at 0 and ∞.

Every integral of the form "difference of two scalings, divided by x" is Frullani: $\int_0^\infty \frac{e^{-x}-e^{-3x}}{x}dx = \ln 3$, $\int_0^\infty \frac{\cos x-\cos 2x}{x}dx = \ln 2$, $\int_0^\infty \frac{\arctan 2x-\arctan x}{x}dx = \frac{\pi}{2}\ln 2$. The $\frac{1}{x}$ makes each piece diverge alone; only the difference converges, and the value depends on f only through its two endpoint values — which is why it looks like magic and is really Feynman's trick in disguise.

Gamma and Beta

$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx$, with $\Gamma(n+1) = n!$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. $B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx = \dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

Gamma interpolates the factorial ($\Gamma(s+1) = s\Gamma(s)$ by one integration by parts) and evaluates every $\int_0^\infty x^n e^{-ax}dx = \frac{n!}{a^{n+1}}$. Beta evaluates every polynomial-times-polynomial integral on $[0,1]$ in one line ($\int_0^1 x^3(1-x)^2dx = \frac{3!\,2!}{6!} = \frac{1}{60}$) and, via $x = \sin^2\theta$, every $\int_0^{\pi/2}\sin^m\theta\cos^n\theta\,d\theta = \frac{1}{2}B(\frac{m+1}{2},\frac{n+1}{2})$ — the Wallis integrals are Beta values. The reflection formula $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin\pi s}$ is why $\int_0^\infty \frac{dx}{1+x^n} = \frac{\pi/n}{\sin(\pi/n)}$.

Weierstrass substitution and series expansion

$t = \tan\frac{x}{2}$ rationalizes any $\int R(\sin x,\cos x)\,dx$. Expanding the integrand in a series and swapping $\sum$ and $\int$ turns hard integrals into known sums.

Weierstrass: $\sin x = \frac{2t}{1+t^2}$, $\cos x = \frac{1-t^2}{1+t^2}$, $dx = \frac{2\,dt}{1+t^2}$ — so $\int_0^\pi \frac{dx}{2+\cos x} = \int_0^\infty \frac{2\,dt}{3+t^2} = \frac{\pi}{\sqrt{3}}$. Series: $\int_0^1 \frac{\ln(1+x)}{x}dx = \sum\frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}$, and $\int_0^1 \ln x\ln(1-x)dx = \sum\frac{1}{n(n+1)^2} = 2-\frac{\pi^2}{6}$. The swap is legal under uniform convergence on the interval — say so.

Which tool, in what order

Symmetry first (it costs nothing). Then look for a parameter (Feynman). Then match a NAMED shape: difference-over-x is Frullani; $x^{\text{power}}e^{-x}$ is Gamma; $x^a(1-x)^b$ is Beta; a rational function of sine and cosine is Weierstrass; a log or arctan over a simple denominator is a series. Only then reach for parts or partial fractions. The Putnam rewards recognition, not endurance.

Try it: Beta in disguise

Evaluate $\displaystyle\int_0^{\pi/2} \sin^5\theta\cos^3\theta\,d\theta$.

Answer. With $x=\sin^2\theta$ this is $\frac{1}{2}B(3,2)=\frac{1}{2}\cdot\frac{2!\,1!}{4!}=\frac{1}{24}$. (Or substitute $u=\sin\theta$, $\cos^2\theta=1-u^2$: $\int_0^1 u^5(1-u^2)du=\frac{1}{6}-\frac{1}{8}=\frac{1}{24}$ ✓.)

4.2 Proofs & Why It Matters

Proof: Frullani's theorem

Assume f continuous with $f(0)$, $f(\infty)$ finite and $0<a<b$.

Write $\dfrac{f(ax)-f(bx)}{x}=-\displaystyle\int_a^b f'(tx)\,dt$ (fundamental theorem in t; assume f' continuous for the clean version). Then $\int_0^\infty \frac{f(ax)-f(bx)}{x}dx=-\int_a^b\left(\int_0^\infty f'(tx)\,dx\right)dt=-\int_a^b\frac{f(\infty)-f(0)}{t}dt=(f(0)-f(\infty))\ln\frac{b}{a}$, swapping the order of integration (Fubini, justified by absolute convergence). ■ Significance: the proof IS Feynman's trick — the $\frac{1}{x}$ was hiding an integral over a parameter. Once you see that, Frullani stops being a formula to memorize.

Proof: $\Gamma(\frac{1}{2})=\sqrt{\pi}$ and the Gaussian integral

Two famous constants, one computation.

Substituting $x=u^2$ in $\Gamma(\frac{1}{2})=\int_0^\infty x^{-1/2}e^{-x}dx$ gives $2\int_0^\infty e^{-u^2}du=\int_{-\infty}^\infty e^{-u^2}du=I$. Square it and go polar: $I^2=\iint e^{-(x^2+y^2)}dA=\int_0^{2\pi}\int_0^\infty e^{-r^2}r\,dr\,d\theta=2\pi\cdot\frac{1}{2}=\pi$. So $I=\sqrt{\pi}=\Gamma(\frac{1}{2})$. ■ No one-variable method evaluates this integral; the plane does. Every normal-distribution constant in statistics traces back to this square.

4.3 Going Deeper: Worked Problems

Worked: a Frullani with cosines

Evaluate $\displaystyle\int_0^\infty\frac{\cos x-\cos 2x}{x}\,dx$.

Step 1 — recognize $f(t)=\cos t$ with $a=1$, $b=2$; but $\cos t$ has no limit at ∞, so the plain theorem does not apply. Step 2 — the fix: Frullani holds if $f(\infty)$ is replaced by the CESRO mean $\lim_{T\to\infty}\frac{1}{T}\int_0^T f$, which for cosine is 0. Step 3 — the value is $(f(0)-0)\ln\frac{2}{1}=\ln 2$. Step 4 — verify with the damping trick: $\int_0^\infty e^{-sx}\frac{\cos x-\cos 2x}{x}dx=\frac{1}{2}\ln\frac{s^2+4}{s^2+1}\to\ln 2$ as $s\to 0$ ✓. Knowing the theorem's exact hypotheses — and how to repair them — is the difference between a 10 and a 1.

Worked: a series-expansion integral from a Putnam

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

Step 1 — $\ln(1+x) = \sum_{n\ge1} \frac{(-1)^{n+1}x^n}{n}$ on $[0,1]$ (convergent at $x=1$ by the alternating series test; uniform on $[0,1]$ by Abel). Step 2 — divide by x and integrate term by term: $\sum_{n\ge1} \frac{(-1)^{n+1}}{n}\int_0^1 x^{n-1}dx = \sum \frac{(-1)^{n+1}}{n^2}$. Step 3 — this is $\eta(2) = (1-2^{1-2})\zeta(2) = \frac{1}{2}\cdot\frac{\pi^2}{6} = \frac{\pi^2}{12}$. Step 4 — numeric check: $1-\frac14+\frac19-\frac{1}{16}+\cdots \approx 0.822$ and $\frac{\pi^2}{12}\approx 0.822$ ✓.

4.4 More Named Results

Dirichlet variants by parts

$\int_0^\infty \frac{\sin^2 x}{x^2}dx = \int_0^\infty \frac{1-\cos x}{x^2}dx = \frac{\pi}{2}$ — one integration by parts each, landing on $\int_0^\infty \frac{\sin x}{x}dx$.

With $u=\sin^2 x$ and $dv = x^{-2}dx$ the boundary term vanishes and the remainder is $\int \frac{\sin 2x}{x}dx$, which rescales to Dirichlet. The same move on $1-\cos x$ gives Dirichlet directly. A whole family — $\int_0^\infty \left(\frac{\sin x}{x}\right)^n dx$ — follows by repeated parts and is a Putnam regular for $n=3$ ($\frac{3\pi}{8}$).

Log-trig integrals and the tangent bridge

$\int_0^{\pi/2} \ln\cos x\,dx = -\frac{\pi}{2}\ln 2$, hence $\int_0^\infty \frac{\ln(1+x^2)}{1+x^2}dx = \pi\ln 2$ (substitute $x=\tan\theta$) and $\int_0^{\pi/2}\ln\tan x\,dx = 0$.

The tangent substitution converts every rational-times-log integrand on $(0,\infty)$ into a log-trig integral on $(0,\frac{\pi}{2})$, where King and Queen apply. Feynman handles the arctangent cousin: $\int_0^\infty \frac{\arctan x}{x(1+x^2)}dx = \frac{\pi}{2}\ln 2$ via the parameter $\arctan(ax)$.

Fresnel, the reflection formula, and Wallis

$\int_0^\infty \sin(x^2)dx = \int_0^\infty \cos(x^2)dx = \frac12\sqrt{\frac{\pi}{2}}$; $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin\pi s}$; $\prod_{n\ge1}\frac{4n^2}{4n^2-1} = \frac{\pi}{2}$; $\int_0^1 \frac{\ln x}{1-x}dx = -\frac{\pi^2}{6}$.

Fresnel is the Mellin integral $\int_0^\infty t^{s-1}\sin t\,dt = \Gamma(s)\sin\frac{\pi s}{2}$ at $s=\frac12$. The reflection formula is why $\int_0^\infty \frac{x^{s-1}}{1+x}dx = \frac{\pi}{\sin\pi s}$ and hence $\int_0^\infty \frac{dx}{1+x^n} = \frac{\pi/n}{\sin(\pi/n)}$. Wallis comes from the sine-power reduction formula squeezed between consecutive terms; $\int_0^1 \frac{\ln x}{1-x}$ is the series method landing on $\zeta(2)$.

A note on rigor

Every result on this shelf involves an interchange — of limit and integral, of sum and integral, of the order of a double integral. On the Putnam, one sentence naming the justification (dominated convergence

with a named dominating function, uniform convergence on the interval, absolute convergence for Fubini) is the difference between a 10 and a 7. Say it every time.

Try it: from Dirichlet to its cube

Given $\int_0^\infty \frac{\sin x}{x}dx = \frac{\pi}{2}$, evaluate $\int_0^\infty \frac{\sin^3 x}{x^3}dx$.

Answer. Use $\sin^3 x = \frac{3\sin x - \sin 3x}{4}$ and integrate by parts twice to reduce $\int \frac{\sin^3 x}{x^3}$ to $\int \frac{\sin x - (\text{rescaled})}{x}$ terms; the result is $\frac{3\pi}{8}$.

4.5 Proofs & Why It Matters

Proof: $\int_0^\infty \frac{\sin^2 x}{x^2}dx = \frac{\pi}{2}$

Parts onto Dirichlet.

Let $u = \sin^2 x$, $dv = x^{-2}dx$, $v = -x^{-1}$: $\int_0^\infty \frac{\sin^2 x}{x^2}dx = \left[-\frac{\sin^2 x}{x}\right]_0^\infty + \int_0^\infty \frac{2\sin x\cos x}{x}dx$. The boundary term is 0 at both ends ($\sin^2 x \sim x^2$ near 0; bounded over x at infinity). The remaining integrand is $\frac{\sin 2x}{x}$; substituting $t = 2x$ gives $\int_0^\infty \frac{\sin t}{t}dt = \frac{\pi}{2}$. ■ Significance: the equality of $\int \frac{\sin x}{x}$ and $\int \frac{\sin^2 x}{x^2}$ is not a coincidence but a one-line consequence of parts — and the pattern extends to the Fourier transform of a triangle function.

Proof: the reflection formula $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$ for $0 < s < 1$

Via the Beta function and a contour-free evaluation.

$\Gamma(s)\Gamma(1-s) = \Gamma(1)B(s, 1-s) = \int_0^1 t^{s-1}(1-t)^{-s}dt$; substitute $t = \frac{x}{1+x}$ to get $\int_0^\infty \frac{x^{s-1}}{1+x}dx$. This last integral equals $\frac{\pi}{\sin \pi s}$ — provable without complex analysis by expanding $\frac{1}{1+x}$ as a geometric series on $(0,1)$ and using $x \to \frac{1}{x}$ on $(1,\infty)$, which yields $\sum_{k\geq 0}(-1)^k\left(\frac{1}{s+k} + \frac{1}{1-s+k}\right)$, the partial-fraction expansion of $\frac{\pi}{\sin \pi s}$. ■ Significance: $s = \frac{1}{2}$ gives $\Gamma(\frac{1}{2})^2 = \pi$ — the Gaussian integral yet again — and the formula evaluates every $\int_0^\infty \frac{dx}{1+x^n}$.

5 Sequences & Series

Convergence, telescoping, power series, and Stolz–Cesro

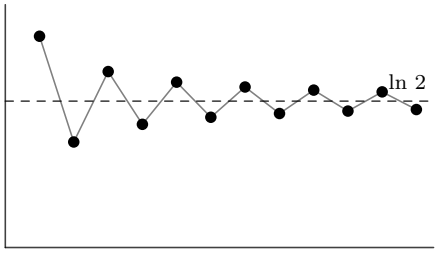
5.1 Summing and Testing

The three series to know by sight

$\sum_{n\geq 0} x^n = \frac{1}{1-x}$, $\sum_{n\geq 1}\frac{x^n}{n} = -\ln(1-x)$, $\sum_{n\geq 0}\frac{x^n}{n!} = e^x$ — and everything you get by

differentiating or integrating them.

Alternating partial sums straddle the limit.



partial sums of $1 - \frac{1}{2} + \frac{1}{3} - \dots$

Differentiate the geometric series: $\sum nx^{n-1} = \frac{1}{(1-x)^2}$, so $\sum \frac{n}{2^n} = 2$ and $\sum \frac{n+1}{3^n} = \frac{9}{4}$. Integrate it: $\sum \frac{x^n}{n}$, so $\sum \frac{1}{n2^n} = \ln 2$. Add the classical constants $\zeta(2) = \frac{\pi^2}{6}$, $\eta(2) = \frac{\pi^2}{12}$, $\sum \frac{(-1)^{n+1}}{n} = \ln 2$, $\sum \frac{(-1)^n}{2n+1} = \frac{\pi}{4}$. A Putnam series problem is usually one of these wearing a costume.

Telescoping and partial fractions

Write the term as $g(n) - g(n+1)$ (or a lagged version) and watch everything cancel except the ends.

$\sum \frac{1}{n(n+1)} = 1$; $\sum \frac{1}{n(n+2)} = \frac{1}{2}(1+\frac{1}{2}) = \frac{3}{4}$ (lag 2: two front terms survive); $\sum \frac{1}{n(n+1)(n+2)} = \frac{1}{2}\cdot\frac{1}{1\cdot 2} = \frac{1}{4}$. Products telescope too: $\prod \left(1 - \frac{1}{n^2}\right) = \prod \frac{(n-1)(n+1)}{n^2} \to \frac{1}{2}$. The tail must go to 0 — say so.

Stolz–Cesro and Riemann sums

If $b_n \uparrow \infty$ and $\dfrac{a_{n+1} - a_n}{b_{n+1} - b_n} \to L$, then $\dfrac{a_n}{b_n} \to L$. And $\dfrac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \to \int_0^1 f$.

Stolz–Cesro is L'Hôpital for sequences: it turns a ratio of sums into a ratio of terms. $\dfrac{\sqrt{1} + \cdots + \sqrt{n}}{n^{3/2}} \to \dfrac{\sqrt{n}}{\frac{3}{2}\sqrt{n}} = \frac{2}{3}$ — the same answer the Riemann sum $\int_0^1 \sqrt{x}\,dx$ gives. Sums of $f(k/n)$ over k, divided by n, are integrals; sums $\sum_{k=1}^{n} \frac{1}{n+k} = \frac{1}{n}\sum \frac{1}{1+k/n} \to \ln 2$.

Convergence tests, ranked by usefulness on the Putnam

Comparison and limit comparison first; the ratio test for anything with factorials or exponentials; the integral test for $\sum \frac{1}{n^p}$-type tails; the alternating series test for sign-alternating terms with decreasing magnitude; Cauchy condensation for slowly diverging logs ($\sum \frac{1}{n \ln n}$ diverges). Root test rarely. Absolute convergence licenses rearranging and swapping with integrals — always state which kind of convergence you have.

Try it: telescoping with a lag

Compute $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+3)}$.

Answer. $\frac{1}{3}\left(\frac{1}{n} - \frac{1}{n+3}\right)$ telescopes with lag 3: the surviving front terms are $1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$, so the sum is $\frac{1}{3} \cdot \frac{11}{6} = \frac{11}{18}$.

5.2 Proofs & Why It Matters

Proof: Stolz–Cesro

If b_n is strictly increasing to ∞ and $\frac{a_{n+1}-a_n}{b_{n+1}-b_n} \to L$, then $\frac{a_n}{b_n} \to L$.

Fix $\varepsilon > 0$; for $n \geq N$, $(L-\varepsilon)(b_{n+1} - b_n) < a_{n+1} - a_n < (L+\varepsilon)(b_{n+1} - b_n)$. Sum these from N to $n-1$ (telescoping!): $(L-\varepsilon)(b_n - b_N) < a_n - a_N < (L+\varepsilon)(b_n - b_N)$. Divide by $b_n \to \infty$: the terms a_N/b_n and b_N/b_n vanish, leaving $L - \varepsilon \leq \liminf \frac{a_n}{b_n} \leq \limsup \frac{a_n}{b_n} \leq L + \varepsilon$. ■ Significance: it is the discrete mean value theorem, and it proves Cesro means preserve limits ($b_n = n$) — a tool for every "average of a sequence" problem.

Proof: $\zeta(2) = \frac{\pi^2}{6}$ via a double integral

The Basel problem, by a route the Putnam would accept.

$\zeta(2) = \sum \frac{1}{n^2} = \sum \int_0^1 \int_0^1 (xy)^{n-1} dx\, dy = \iint_{[0,1]^2} \frac{dx\, dy}{1-xy}$ (geometric series, swap by monotone convergence). Rotate coordinates 45° ($u = \frac{x+y}{\sqrt{2}}$, $v = \frac{y-x}{\sqrt{2}}$); the integrand becomes $\frac{1}{1-\frac{u^2-v^2}{2}}$ and the inner integral is an arctangent; two elementary trigonometric integrals later the total is $\frac{\pi^2}{6}$ (Beukers–Calabi–Kolk). ■ The lesson for the exam: sums become integrals ($\frac{1}{n} = \int_0^1 x^{n-1} dx$) and integrals become sums; move freely between them.

5.3 Going Deeper: Worked Problems

Worked: a series from a Putnam via generating functions

Compute $\displaystyle\sum_{n=1}^{\infty} \frac{n^2}{2^n}$.

Step 1 — from $\sum x^n = \frac{1}{1-x}$: differentiate and multiply by x twice. $\sum n x^n = \frac{x}{(1-x)^2}$; then $\sum n^2 x^n = x\frac{d}{dx}\frac{x}{(1-x)^2} = \frac{x(1+x)}{(1-x)^3}$. Step 2 — at $x = \frac{1}{2}$: $\frac{\frac{1}{2}\cdot\frac{3}{2}}{\frac{1}{8}} = 6$. Step 3 — check the first terms: $\frac{1}{2} + 1 + \frac{9}{8} + 1 + \frac{25}{32} + \cdots \approx 4.4$ after five terms, climbing toward 6 ✓. The operator "$x\frac{d}{dx}$" multiplies coefficients by n; apply it as many times as the power of n demands.

Worked: convergence with a parameter

For which real p does $\displaystyle\sum_{n\geq 2}\frac{1}{n(\ln n)^p}$ converge?

Step 1 — the terms decrease, so Cauchy condensation applies: the series converges iff $\sum 2^k\cdot\frac{1}{2^k(\ln 2^k)^p}=\sum\frac{1}{(k\ln 2)^p}$ does. Step 2 — that is a p-series in k: converges iff $p>1$. Step 3 — so $\sum\frac{1}{n\ln n}$ diverges (barely!), $\sum\frac{1}{n(\ln n)^2}$ converges. The integral test says the same: $\int\frac{dx}{x(\ln x)^p}$ with $u=\ln x$ becomes $\int\frac{du}{u^p}$. Logarithmic factors change convergence only at the razor's edge $p=1$.

6 Hard Sequences: Closed Forms, Rates & Periodicity

Squaring tricks, reciprocal tricks, Lucas closed forms, Stolz rates, and maps of finite order

6.1 Sequences the Putnam Actually Sets

The squaring and reciprocal tricks

For $a_{n+1}=a_n+\frac{1}{a_n}$, square: $a_{n+1}^2=a_n^2+2+a_n^{-2}$, so $a_n^2\approx 2n$. For $a_{n+1}=a_n-a_n^2$, take reciprocals: $\frac{1}{a_{n+1}}-\frac{1}{a_n}=\frac{1}{1-a_n}\to 1$, so $na_n\to 1$.

Recurrences that are not linear usually become tame after a transformation chosen to make the increments simple: square when the recurrence adds a reciprocal, take reciprocals when it subtracts a square, take logarithms when it multiplies. After the transformation, Stolz–Cesro (the discrete L'Hôpital) reads off the growth rate: if $b_{n+1}-b_n\to L$ then $b_n/n\to L$. The Putnam then asks for $\lfloor a_{100}\rfloor$ or $\lim na_n$ — questions that are hopeless by direct iteration and one line after the trick.

Closed forms from $\alpha+\alpha^{-1}$

If $a_0=\alpha+\alpha^{-1}$ and $a_{n+1}=a_n^2-2$, then $a_n=\alpha^{2^n}+\alpha^{-2^n}$. With $a_0=3$ this is a Lucas number, $L_{2^{n+1}}$: $3,7,47,2207,\ldots$

Because $(\alpha^k+\alpha^{-k})^2-2=\alpha^{2k}+\alpha^{-2k}$, the map $x\mapsto x^2-2$ is the doubling map in disguise (compare $2\cos 2\theta=(2\cos\theta)^2-2$). This is the engine behind the Lucas–Lehmer primality test and behind Putnam problems that ask about $\prod(1+1/a_k)$ or a_n mod m: a closed form makes them telescope or cycle. The trigonometric twin, $a_{n+1}=\sqrt{2+a_n}$, has closed form $2\cos\frac{\pi}{2^{n+1}}$ and limit 2.

Maps of finite order and periodic sequences

$x\mapsto\frac{1}{1-x}$ has order 3; $x\mapsto\frac{1+x}{1-x}$ has order 4; $x\mapsto\frac{x+1}{x-1}$ is an involution. A recurrence $x_{n+1}=g(x_n)$ with g of order k is periodic with period dividing k.

Mbius maps $x\mapsto\frac{ax+b}{cx+d}$ compose like the matrices $\begin{pmatrix}a&b\\c&d\end{pmatrix}$, so the order of the map is the order of

the matrix up to scalars — which is why $\frac{1+x}{1-x}$ (the matrix of a 45° rotation) has order 4. When a Putnam problem asks for x_{2024}, compute a few terms, spot the period, and PROVE it by composing the map with itself.

Generating functions for linear recurrences

A linear recurrence $a_{n+2} = pa_{n+1} + qa_n$ has generating function $\frac{\text{(polynomial)}}{1-px-qx^2}$; evaluating at a point inside the radius of convergence sums series like $\sum F_n/2^n = 2$ instantly. The radius is $1/|\text{largest root}|$ — for Fibonacci, $1/\varphi \approx 0.618$, so $x = \frac{1}{2}$ is safely inside.

Try it: a rate by Stolz

If $a_1 = 1$ and $a_{n+1} = a_n + \frac{1}{a_n^2}$, find $\lim a_n^3/n$.

Answer. Cube: $a_{n+1}^3 = a_n^3 + 3 + 3a_n^{-3} + a_n^{-6}$, so $a_{n+1}^3 - a_n^3 \to 3$ and by Stolz $a_n^3/n \to 3$.

6.2 Proofs & Why It Matters

Proof: Stolz–Cesro

If b_n is strictly increasing with $b_n \to \infty$ and $\frac{a_{n+1}-a_n}{b_{n+1}-b_n} \to L$, then $\frac{a_n}{b_n} \to L$.

Fix $\varepsilon > 0$ and N with $L - \varepsilon < \frac{a_{k+1}-a_k}{b_{k+1}-b_k} < L + \varepsilon$ for $k \geq N$. Multiply by the positive $b_{k+1} - b_k$ and sum from N to $n-1$: $(L-\varepsilon)(b_n - b_N) < a_n - a_N < (L+\varepsilon)(b_n - b_N)$. Divide by b_n and let $n \to \infty$: since $b_N/b_n \to 0$ and $a_N/b_n \to 0$, every limit point of a_n/b_n lies in $[L-\varepsilon, L+\varepsilon]$. ■ Significance: it turns "the increments tend to L" into "the average tends to L" — Cesro means, the rate of $a_{n+1} = a_n - a_n^2$, and every $\lim \frac{1}{n^{p+1}} \sum k^p$ are one application each.

6.3 Going Deeper: Worked Problems

Worked: $\lfloor a_{100} \rfloor$ for $a_{n+1} = a_n + 1/a_n$

With $a_1 = 1$, find $\lfloor a_{100} \rfloor$.

Step 1 — square: $a_{n+1}^2 = a_n^2 + 2 + a_n^{-2}$, so $a_{100}^2 = 199 + \sum_{k=1}^{99} a_k^{-2}$. Step 2 — lower bound $a_{100}^2 > 199$. Step 3 — upper bound: from Step 1, $a_k^2 \geq 2k - 1$, so the correction is at most $1 + \sum_{k=2}^{99} \frac{1}{2k-1} < 1 + \frac{1}{2}\ln 197 + \frac{1}{2} < 5$. Step 4 — $199 < a_{100}^2 < 204$, so $14.1 < a_{100} < 14.3$ and the floor is 14. Every step is an inequality you can state precisely — that is what makes it a Putnam answer rather than a calculator answer.

7 Limits, Continuity & the Big Theorems

IVT, MVT, Taylor with remainder, and limits in disguise

7.1 The Theorems That Do the Work

Limits that are derivatives, integrals, or exponentials in disguise

$n(2^{1/n}-1) \to \ln 2$ (a derivative); $\frac{1}{n}\sum f(\frac{k}{n}) \to \int_0^1 f$ (an integral); $(1+\frac{a}{n})^n \to e^a$ (an exponential); $\frac{(n!)^{1/n}}{n} \to \frac{1}{e}$ (a log of a Riemann sum).

The first reflex for any strange limit: rewrite it as a difference quotient, a Riemann sum, or exp(log). The second: Taylor expand — $\frac{\sin x - x}{x^3} \to -\frac{1}{6}$ because $\sin x = x - \frac{x^3}{6} + \cdots$, and $\lim(\cos x)^{1/x^2} = e^{-1/2}$ because $\ln\cos x \approx -\frac{x^2}{2}$. L'Hôpital is the tool of last resort: it is slow and it hides the structure.

IVT, MVT, Rolle, and Taylor with remainder

Continuous f on $[a,b]$ takes every value between $f(a)$ and $f(b)$. Differentiable f has $f(b)-f(a) = f'(c)(b-a)$ for some $c \in (a,b)$. And $f(x) = \sum_{k<n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}(\xi)}{n!}(x-a)^n$.

Existence of roots is IVT; uniqueness is monotonicity (a derivative that keeps one sign) — $x = \cos x$ has exactly one solution, $e^x = 3x$ exactly two. MVT converts a statement about VALUES into one about a DERIVATIVE at some unknown point; Rolle is the case $f(a) = f(b)$, which proves that between two roots of f lies a root of f' (so a polynomial with n real roots has $n-1$ real critical points). Taylor with the Lagrange remainder gives INEQUALITIES: $\cos x \geq 1 - \frac{x^2}{2}$, $e^x \geq 1+x$, $\ln(1+x) \leq x$ — each a one-line remainder-sign argument.

Continuity versus uniform continuity, and why graders care

On a closed bounded interval, continuous functions are bounded, attain their extrema, and are uniformly continuous — three facts a Putnam solution may cite by name. On open or infinite intervals none of them is automatic, and the counterexamples ($\frac{1}{x}$ on $(0,1)$, x^2 on $\mathbb{R}$) are exactly the ones that appear in the hard problems.

Try it: count the solutions

How many real solutions does $x^3 - 3x + 1 = 0$ have?

Answer. $f' = 3x^2 - 3$ vanishes at ± 1; $f(-1) = 3 > 0$ and $f(1) = -1 < 0$ with $f \to \mp\infty$ at $\mp\infty$: three sign changes, hence three real roots (and a cubic has at most three). They are $2\cos\frac{2\pi k}{9}$-type numbers — not needed for the count.

7.2 Proofs & Why It Matters

Proof: the mean value theorem from Rolle

Rolle: if f is continuous on $[a,b]$, differentiable on (a,b), and $f(a) = f(b)$, then $f'(c) = 0$ for some c.

Rolle: f attains a max and min on $[a,b]$ (extreme value theorem). If both occur at endpoints then f is constant and any c works; otherwise an interior extremum has $f'(c) = 0$ (Fermat: the difference

quotients from the two sides have opposite signs). MVT: apply Rolle to $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$, which has $g(a) = g(b) = f(a)$; $g'(c) = 0$ gives $f'(c) = \frac{f(b)-f(a)}{b-a}$. ■ Significance: the MVT is how every inequality between a function and its derivative is proved, and how Taylor's remainder is derived (apply it to the right auxiliary function).

Proof: $\frac{(n!)^{1/n}}{n} \to \frac{1}{e}$

A limit that looks like it needs Stirling and does not.

$\ln \frac{(n!)^{1/n}}{n} = \frac{1}{n}\sum_{k=1}^{n} \ln k - \ln n = \frac{1}{n}\sum_{k=1}^{n} \ln \frac{k}{n}$, a Riemann sum for $\int_0^1 \ln x\, dx$. The integrand is unbounded at 0, so justify: the sum is a right-endpoint sum of a monotone function, hence squeezed between $\int_0^1 \ln x\, dx = -1$ and $\int_{1/n}^1 \ln x\, dx + \frac{1}{n}\ln\frac{1}{n} \to -1$. Exponentiate: the limit is e^{-1}. ■ Significance: "log, Riemann sum, exponentiate" handles every limit of a geometric-mean shape — and handling the endpoint singularity honestly is what earns the 10.

7.3 Going Deeper: Worked Problems

Worked: an MVT inequality

Prove $\ln(1+x) \geq \dfrac{x}{1+x}$ for $x > 0$, and find $\lim_{n\to\infty} n\ln\left(1+\frac{1}{n}\right)$.

Step 1 — MVT on ln over $[1, 1+x]$: $\ln(1+x) - \ln 1 = \frac{1}{c}\cdot x$ for some $c \in (1, 1+x)$, and $\frac{1}{c} > \frac{1}{1+x}$, giving $\ln(1+x) > \frac{x}{1+x}$. Step 2 — the same MVT also gives $\ln(1+x) < x$ (since $\frac{1}{c} < 1$). Step 3 — squeeze with $x = \frac{1}{n}$: $\frac{1/n}{1+1/n} < \ln(1+\frac{1}{n}) < \frac{1}{n}$, multiply by n: $\frac{n}{n+1} < n\ln(1+\frac{1}{n}) < 1$, so the limit is 1. Step 4 — exponentiate to recover $(1+\frac{1}{n})^n \to e$, now PROVED rather than assumed.

Worked: Taylor's remainder bounds an error

Show $\left|\cos x - \left(1 - \frac{x^2}{2}\right)\right| \leq \dfrac{x^4}{24}$ for all real x.

Step 1 — Taylor at 0 to order 3 with Lagrange remainder: $\cos x = 1 - \frac{x^2}{2} + 0\cdot x^3 + \frac{\cos^{(4)}(\xi)}{4!}x^4$ for some ξ between 0 and x. Step 2 — $\cos^{(4)} = \cos$, and $|\cos\xi| \leq 1$, so the remainder is at most $\frac{x^4}{24}$ in absolute value. Step 3 — in fact $\cos\xi \geq 0$ is not guaranteed, but the BOUND is; and since the next omitted term is $+\frac{x^4}{24}\cos\xi$, we also get $\cos x \geq 1 - \frac{x^2}{2}$ whenever $\cos\xi \geq 0$ — e.g. for $|x| \leq \frac{\pi}{2}$. The remainder term is a precise, citable error bar.

8 Differential Equations on the Putnam

Separable and linear equations, uniqueness by integrating factor, blow-up, and Grnwall

8.1 The ODE Toolkit the Exam Assumes

Separable and first-order linear equations

Separable: $y' = f(x)g(y) \Rightarrow \int \frac{dy}{g(y)} = \int f(x)dx$. Linear: $y' + P(x)y = Q(x)$ is solved by the integrating factor $\mu = e^{\int P}$, which makes $(\mu y)' = \mu Q$.

The Putnam never asks you to solve a hard ODE; it asks you to recognize that a functional condition IS an ODE, solve the easy ODE, and then argue about existence, uniqueness, or a bound. $y' = y^2$ with $y(0) = 1$ gives $y = \frac{1}{1-x}$ — and the point of the problem is the blow-up at $x = 1$: no solution exists on all of $\mathbb{R}$. Likewise $f' = 1 + f^2$ forces $f = \tan$, which cannot be defined on an interval longer than π.

Uniqueness by integrating factor, and Grnwall

If $f' = f$ then $(fe^{-x})' = 0$, so $f = Ce^x$. If $f' \le f$ and $f(0) = 1$, then fe^{-x} is non-increasing, so $f(x) \le e^x$ for $x \ge 0$ (Grnwall).

Multiplying by e^{-x} converts a differential equation into "a derivative is zero" and a differential INEQUALITY into "a function is monotone" — the two-line arguments graders want in place of "obviously the solution is e^x". For second-order equations the analogous tool is an energy: for $y'' + y = 0$, $E = y^2 + y'^2$ has $E' = 2y'(y + y'') = 0$, so a solution with $y(0) = y'(0) = 0$ has $E \equiv 0$ and is identically zero — uniqueness for free.

Constant-coefficient linear equations

$y'' + py' + qy = 0$ has solutions e^{rx} for the roots r of $r^2 + pr + q = 0$ (with xe^{rx} for a double root, and $e^{ax}\cos bx$, $e^{ax}\sin bx$ for $r = a \pm bi$).

The solution space is two-dimensional (a solution is determined by $y(0), y'(0)$ — that is the uniqueness theorem), so two independent solutions span everything. The same characteristic-root method solves linear recurrences, which is why the "Sequences" and "Differential Equations" units share a proof.

What the Putnam does with ODEs

Typical asks: show a function satisfying an $f'(x) = f(x+1)$-type condition has a specific form; bound $f(2)$ given $f' \le f$; show a solution of $y' = y^2 + x$ cannot exist on all of $\mathbb{R}$ (compare with $y' \ge y^2$ after some point and use blow-up); compute $\int_0^1 f$ for the unique f with $f' = f$, $f(0) = 1$. Existence and uniqueness (Picard) may be cited; blow-up must be argued.

Try it: an integrating factor

Solve $y' + \frac{1}{x}y = x$ with $y(1) = 1$ and evaluate $y(2)$.

Answer. $\mu = x$: $(xy)' = x^2$, so $xy = \frac{x^3}{3} + C$, $C = \frac{2}{3}$, and $y(2) = \frac{1}{2}\left(\frac{8}{3} + \frac{2}{3}\right) = \frac{5}{3}$.

8.2 Proofs & Why It Matters

Proof: Grnwall's inequality

If $u'(x) \le \beta u(x)$ on $[0, T]$ with $u(0) = u_0$, then $u(x) \le u_0 e^{\beta x}$.

Let $v(x) = u(x)e^{-\beta x}$. Then $v' = (u' - \beta u)e^{-\beta x} \le 0$, so v is non-increasing and $v(x) \le v(0) = u_0$, i.e. $u(x) \le u_0 e^{\beta x}$. ■ Significance: Grnwall is how uniqueness of ODE solutions is actually proved (apply it to the square of the difference of two solutions), and it is the model for every "differential inequality ⇒ growth bound" argument.

Proof: blow-up for $y' \ge y^2$

If y is differentiable with $y' \ge y^2$ and $y(0) = c > 0$, then y cannot exist on $[0, 1/c]$.

Since $y' \ge y^2 \ge 0$, y is non-decreasing, so $y > 0$ wherever it exists and $\left(-\frac{1}{y}\right)' = \frac{y'}{y^2} \ge 1$. Integrating from 0 to x: $-\frac{1}{y(x)} + \frac{1}{c} \ge x$, i.e. $\frac{1}{y(x)} \le \frac{1}{c} - x$, which is impossible once $x \ge \frac{1}{c}$ (the left side is positive). ■ Significance: this comparison is the standard way to show that solutions of $y' = y^2 + (\text{anything} \ge 0)$ have finite lifetime — a favorite Putnam "prove that no such function exists on all of $\mathbb{R}$."

9 Inequalities & Convexity

AM–GM, Cauchy–Schwarz, Jensen, rearrangement, and integral inequalities

9.1 The Inequality Toolkit

The core four

AM–GM: $\frac{\sum a_i}{n} \ge \sqrt[n]{\prod a_i}$. Cauchy–Schwarz: $(\sum a_i b_i)^2 \le \sum a_i^2 \sum b_i^2$. Jensen: for convex f, $f(\frac{\sum x_i}{n}) \le \frac{\sum f(x_i)}{n}$. Rearrangement: similarly ordered sequences maximize $\sum a_i b_i$.

Each comes with an EQUALITY CASE — all equal for AM–GM, proportional sequences for Cauchy–Schwarz, all equal for strictly convex Jensen — and the equality case tells you which inequality to use: if the answer "should" be attained at $a = b = c$, try AM–GM or Jensen. Cauchy–Schwarz in Engel form, $\sum \frac{a_i^2}{b_i} \ge \frac{(\sum a_i)^2}{\sum b_i}$, proves $(a+b+c)(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}) \ge 9$ in one line. Power means ($QM \ge AM \ge GM \ge HM$) are Jensen for x^p.

Convexity and its consequences

A function with $f'' \ge 0$ lies above its tangent lines and below its chords. Both facts are inequalities you can quote.

Above tangents: $e^x \ge 1+x$, $\ln x \le x-1$, $\sin x \le x$ for $x \ge 0$ (concave). Below chords: the average of f at points is at least f at the average (Jensen). Tangent-line trick for olympiad-style sums: bound

each term $f(x_i)$ below by the tangent at the equality point, then sum — the constraint $\sum x_i = \text{const}$ makes the linear terms collapse. Integral versions: $\int_0^1 f \geq f(\frac{1}{2})$ for convex f (Hermite–Hadamard), and $\left(\int fg\right)^2 \leq \int f^2 \int g^2$ (Cauchy–Schwarz for integrals).

Weighted and integral forms are the same theorems

Cauchy–Schwarz is an inner-product statement, so it holds for sums, weighted sums, and integrals alike — $\left(\int_a^b fg\right)^2 \leq \int_a^b f^2 \int_a^b g^2$ — and Jensen holds for any probability weighting: $f(E[X]) \leq E[f(X)]$. A Putnam analysis problem that looks new is often Cauchy–Schwarz with an integral in place of a sum.

Try it: locate the equality case first

Minimize $x + \dfrac{4}{x}$ for $x > 0$, then minimize $\dfrac{x^2+2}{\sqrt{x^2+1}}$ over all real x.

Answer. AM–GM: $x + \frac{4}{x} \geq 2\sqrt{4} = 4$ at $x = 2$. For the second, set $t = \sqrt{x^2+1} \geq 1$: the expression is $t + \frac{1}{t} \geq 2$ at $t = 1$, i.e. $x = 0$ — and $t = 1$ IS in the domain, so the minimum 2 is attained.

9.2 Proofs & Why It Matters

Proof: Cauchy–Schwarz by a quadratic

For real vectors $\mathbf{a}, \mathbf{b}$: $(\mathbf{a} \cdot \mathbf{b})^2 \leq |\mathbf{a}|^2 |\mathbf{b}|^2$.

For every real t, $0 \leq |\mathbf{a} - t\mathbf{b}|^2 = |\mathbf{a}|^2 - 2t(\mathbf{a} \cdot \mathbf{b}) + t^2|\mathbf{b}|^2$. A quadratic in t that is never negative has discriminant ≤ 0: $4(\mathbf{a} \cdot \mathbf{b})^2 - 4|\mathbf{a}|^2|\mathbf{b}|^2 \leq 0$. Equality iff the quadratic has a root, i.e. $\mathbf{a} = t\mathbf{b}$. ■ The identical argument with $\int$ in place of $\sum$ proves the integral version — one proof, every inner product space.

Proof: Jensen's inequality

For convex f and weights $w_i \geq 0$ summing to 1: $f(\sum w_i x_i) \leq \sum w_i f(x_i)$.

Let $m = \sum w_i x_i$ and take a supporting line at m: convexity gives $f(x) \geq f(m) + s(x - m)$ for all x (with $s = f'(m)$ when differentiable). Apply at each x_i, multiply by w_i, and add: $\sum w_i f(x_i) \geq f(m) + s\sum w_i(x_i - m) = f(m) + 0$. ■ Significance: AM–GM is Jensen for ln (concave), the power-mean inequality is Jensen for x^p, and entropy inequalities in information theory are Jensen for $x \ln x$ — the tangent-line argument is the master proof.

9.3 Going Deeper: Worked Problems

Worked: an integral inequality

For continuous $f > 0$ on $[0,1]$, show $\displaystyle\int_0^1 f \cdot \int_0^1 \frac{1}{f} \geq 1$.

Step 1 — Cauchy–Schwarz for integrals with $\sqrt{f}$ and $\frac{1}{\sqrt{f}}$: $\left(\int_0^1 \sqrt{f} \cdot \frac{1}{\sqrt{f}}\right)^2 \leq \int_0^1 f \cdot \int_0^1 \frac{1}{f}$. Step 2 — the left side is $\left(\int_0^1 1\right)^2 = 1$. Step 3 — equality iff $\sqrt{f}$ and $\frac{1}{\sqrt{f}}$ are proportional, i.e. f constant. This is the integral version of $(\sum a_i)(\sum \frac{1}{a_i}) \geq n^2$ — the same inequality, the same proof, an integral in place of a sum.

Worked: the tangent-line trick

For positive a, b, c with $a + b + c = 3$, show $\displaystyle\sum \frac{a}{a^2+1} \leq \frac{3}{2}$.

Step 1 — the equality case is $a = b = c = 1$, where each term is $\frac{1}{2}$. Step 2 — find the tangent line to $g(x) = \frac{x}{x^2+1}$ at $x = 1$: $g'(x) = \frac{1-x^2}{(x^2+1)^2}$, so $g'(1) = 0$ and the tangent is the horizontal line $y = \frac{1}{2}$. Step 3 — claim $g(x) \leq \frac{1}{2}$ for all $x > 0$: indeed $\frac{x}{x^2+1} \leq \frac{1}{2} \iff 2x \leq x^2 + 1 \iff (x-1)^2 \geq 0$ ✓. Step 4 — sum: $\sum g(a) \leq \frac{3}{2}$. Here the constraint was not even needed — but in general the tangent line has a slope s, and $\sum s(x_i - 1) = s(\sum x_i - 3) = 0$ is where the constraint enters.

10 Linear Algebra I: Determinants & Structured Matrices

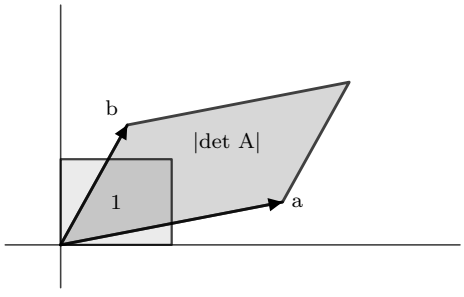
Rank-one updates, Vandermonde, circulants, tridiagonals

10.1 Determinants the Putnam Way

Identity plus rank one

$\det(I + \mathbf{u}\mathbf{v}^\mathsf{T}) = 1 + \mathbf{v}^\mathsf{T}\mathbf{u}$ (the matrix determinant lemma). So $\det(aI + bJ) = a^{n-1}(a + nb)$ for the all-ones matrix J.

The determinant is the area scale factor.



Proof idea: $\mathbf{u}\mathbf{v}^\mathsf{T}$ has rank 1 with the single nonzero eigenvalue $\mathbf{v}^\mathsf{T}\mathbf{u}$ (eigenvector $\mathbf{u}$), so $I + \mathbf{u}\mathbf{v}^\mathsf{T}$ has eigenvalues $1 + \mathbf{v}^\mathsf{T}\mathbf{u}$ and 1 (repeated). The 5×5 matrix with 2 on the diagonal and 1 elsewhere is

$I + J$: determinant 6. Matrices with entries $a_i + b_j$ have rank ≤ 2 (entries $i + j$: rank exactly 2); entries $a_i b_j$ give rank 1. Recognize low rank and the determinant is 0 before any computation.

Vandermonde, circulant, tridiagonal

Vandermonde: $\det[x_i^{j-1}] = \prod_{i<j}(x_j - x_i)$. Circulant: eigenvalues are the DFT of the first row, $\lambda_k = \sum_j c_j \omega^{jk}$. Tridiagonal Toeplitz ($a$ diagonal, b off): eigenvalues $a + 2b\cos\frac{k\pi}{n+1}$.

Vandermonde is why $n + 1$ points determine a degree-n polynomial (the interpolation matrix is invertible iff the x_i are distinct). Circulants are diagonalized by the Fourier matrix, so their determinants are products of polynomial values at roots of unity — how det of "a on the diagonal, b next to it, cyclically" becomes $\prod_k (a + b\omega^k + b\omega^{-k})$. Tridiagonal determinants satisfy a three-term recurrence $D_n = aD_{n-1} - b^2 D_{n-2}$, which is why their eigenvalues are cosines.

Row operations are your first move

Before any formula, subtract rows: a determinant whose rows are arithmetic progressions collapses to zero after two subtractions; one whose entries are $\min(i, j)$ or $\max(i, j)$ becomes triangular after subtracting adjacent rows ($\det[\min(i, j)] = 1$, $\det[\max(i, j)] = (-1)^{n+1} n$). The Putnam determinant is almost always a two-row-operations problem in disguise.

Try it: rank on sight

What is the rank of the 6×6 matrix with entries $m_{ij} = (i - j)^2$?

Answer. Expand: $(i - j)^2 = i^2 - 2ij + j^2$ — a sum of THREE rank-one matrices ($i^2 \cdot 1$, $i \cdot j$, $1 \cdot j^2$), so rank ≤ 3; and it is exactly 3 (the three are independent for $n \geq 3$). Any matrix whose entries are a polynomial in i and j of total degree d has rank at most $d + 1$.

10.2 Proofs & Why It Matters

Proof: the matrix determinant lemma

$\det(I + \mathbf{u}\mathbf{v}^\mathsf{T}) = 1 + \mathbf{v}^\mathsf{T}\mathbf{u}$, and more generally $\det(A + \mathbf{u}\mathbf{v}^\mathsf{T}) = \det A\,(1 + \mathbf{v}^\mathsf{T} A^{-1}\mathbf{u})$ for invertible A.

Block-matrix proof: the identity $\begin{pmatrix} I & 0 \\ \mathbf{v}^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} I + \mathbf{u}\mathbf{v}^\mathsf{T} & \mathbf{u} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ -\mathbf{v}^\mathsf{T} & 1 \end{pmatrix} = \begin{pmatrix} I & \mathbf{u} \\ 0 & 1 + \mathbf{v}^\mathsf{T}\mathbf{u} \end{pmatrix}$ holds by direct multiplication; take determinants of both sides (the outer factors are unitriangular, determinant 1; the right side is block triangular) to get $\det(I + \mathbf{u}\mathbf{v}^\mathsf{T}) = 1 + \mathbf{v}^\mathsf{T}\mathbf{u}$. For the general form factor $A + \mathbf{u}\mathbf{v}^\mathsf{T} = A(I + A^{-1}\mathbf{u}\mathbf{v}^\mathsf{T})$. ■ Significance: every "all entries equal except the diagonal" determinant, and every rank-one perturbation, is one line — $\det(aI + bJ) = a^{n-1}(a + nb)$ is the case $\mathbf{u} = b\mathbf{1}$, $\mathbf{v} = \mathbf{1}$.

Proof: the Vandermonde determinant

$\det\left[x_i^{j-1}\right]_{i,j=1}^{n} = \prod_{1\le i<j\le n}(x_j - x_i)$.

View the determinant as a polynomial P in the variables $x_1, \dots, x_n$. If $x_i = x_j$ two rows coincide and $P = 0$, so $(x_j - x_i)$ divides P for every pair; these $\binom{n}{2}$ linear factors are pairwise coprime, so their product Q divides P. Both P and Q have total degree $0+1+\cdots+(n-1) = \binom{n}{2}$, hence $P = cQ$ for a constant c; comparing the coefficient of the monomial $x_2 x_3^2 \cdots x_n^{n-1}$ (the main diagonal term in P, and the product of the leading terms in Q) gives $c = 1$. ■ Significance: Vandermonde is nonzero iff the nodes are distinct, which is exactly why polynomial interpolation through distinct points is unique — Lagrange interpolation is this determinant's invertibility.

Proof: the tridiagonal recurrence and its eigenvalues

For T_n with a on the diagonal and b, c on the off-diagonals, $\det T_n = a \det T_{n-1} - bc \det T_{n-2}$; the eigenvalues of the symmetric case ($b = c$) are $a + 2b\cos\frac{k\pi}{n+1}$, $k = 1, \dots, n$.

Expand along the first row: the $(1,1)$ cofactor is $\det T_{n-1}$; the $(1,2)$ entry b has a cofactor whose first column contains only c (in its first row), and expanding that along its first column leaves $-bc \det T_{n-2}$. For the eigenvalues, try $\mathbf{x} = (\sin\theta, \sin 2\theta, \dots, \sin n\theta)$: the identity $\sin(j-1)\theta + \sin(j+1)\theta = 2\cos\theta \sin j\theta$ shows $T_n\mathbf{x} = (a + 2b\cos\theta)\mathbf{x}$ provided the boundary terms vanish, i.e. $\sin(n+1)\theta = 0$, which gives $\theta = \frac{k\pi}{n+1}$. ■ Significance: T_n with $a = 2$, $b = c = -1$ is the discrete Laplacian; its determinant $n+1$ and its spectrum $2 - 2\cos\frac{k\pi}{n+1}$ appear in resistor networks, random walks, and the Putnam every few years.

Proof: the matrix determinant lemma

$\det(I + \mathbf{u}\mathbf{v}^{\mathsf{T}}) = 1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}$.

The block identity $\begin{pmatrix} I & 0 \\ \mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix}\begin{pmatrix} I + \mathbf{u}\mathbf{v}^{\mathsf{T}} & \mathbf{u} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I & 0 \\ -\mathbf{v}^{\mathsf{T}} & 1 \end{pmatrix} = \begin{pmatrix} I & \mathbf{u} \\ 0 & 1 + \mathbf{v}^{\mathsf{T}}\mathbf{u} \end{pmatrix}$ has determinant $1 + \mathbf{v}^{\mathsf{T}}\mathbf{u}$ on the right and $\det(I + \mathbf{u}\mathbf{v}^{\mathsf{T}})$ on the left (the outer factors are unit triangular). ■ Significance: with A in place of I, $\det(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}) = \det A\,(1 + \mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u})$ — the tool behind Sherman–Morrison updates in numerical linear algebra and half the Putnam determinant problems.

Proof: the Vandermonde determinant

$\det[x_i^{j-1}]_{i,j=1}^{n} = \prod_{1\le i<j\le n}(x_j - x_i)$.

View the determinant as a polynomial in x_n: it has degree $n-1$ and vanishes whenever $x_n = x_i$ for $i < n$ (two equal rows), so it equals $c\prod_{i<n}(x_n - x_i)$ with c independent of x_n; comparing the coefficient of x_n^{n-1} (expand along the last row) shows c is the $(n-1)\times(n-1)$ Vandermonde. Induct. ■ Significance: the product is nonzero iff the x_i are distinct — polynomial interpolation is unique, and the roots-of-unity Vandermonde is (up to scale) the unitary Fourier matrix.

10.3 Going Deeper: Worked Problems

Worked: a tridiagonal determinant by recurrence

Let D_n be the determinant of the $n \times n$ matrix with 2 on the diagonal and 1 on both adjacent diagonals. Find D_n.

Step 1 — expand along the first row: $D_n = 2D_{n-1} - 1 \cdot D_{n-2}$ (the second term comes from the 1 in position $(1,2)$, whose minor expands to D_{n-2}). Step 2 — $D_1 = 2$, $D_2 = 3$; the recurrence $D_n = 2D_{n-1} - D_{n-2}$ has characteristic root 1 (double), so $D_n = A + Bn$. Step 3 — fit: $A + B = 2$, $A + 2B = 3$ gives $D_n = n + 1$. Step 4 — cross-check with the eigenvalue formula: $\prod_{k=1}^{n} \left(2 + 2\cos\frac{k\pi}{n+1}\right) = \prod 4\cos^2\frac{k\pi}{2(n+1)} = n + 1$, a known product identity ✓.

Worked: a determinant that is secretly a polynomial

Compute $\det \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix}$ and use it to show three points (a, a^2), (b, b^2), (c, c^2) on a parabola are never collinear.

Step 1 — Vandermonde: the determinant is $(b-a)(c-a)(c-b)$. Step 2 — three points (x_i, y_i) are collinear iff $\det[1, x_i, y_i] = 0$ (the rows are then linearly dependent — the standard area formula). Step 3 — with $y_i = x_i^2$ that determinant is the Vandermonde, nonzero for distinct a, b, c. Hence no three distinct points of a parabola are collinear. Step 4 — the same argument shows a degree-n polynomial's graph meets no line in more than n points — Vandermonde is the algebraic form of "a polynomial has at most n roots."

11 Linear Algebra II: Spectra, Cayley–Hamilton & Finite Fields

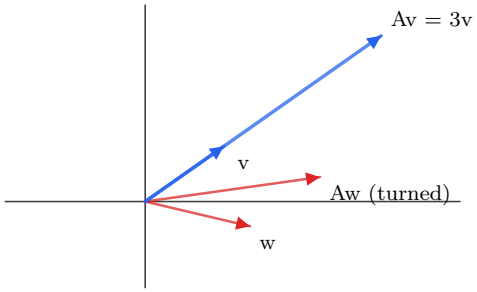
Eigenvalues do the arithmetic; counting matrices over F_p

11.1 Spectral Reasoning

Eigenvalues carry every polynomial fact

If A has eigenvalues λ_i then $p(A)$ has eigenvalues $p(\lambda_i)$: $\det(A^2 + I) = \prod(\lambda_i^2 + 1)$, $\operatorname{tr}(A^k) = \sum \lambda_i^k$, $\det e^A = e^{\operatorname{tr} A}$.

An eigenvector keeps its direction; A only scales it.



A 3×3 matrix with eigenvalues $1, 2, 3$ has $\det(A^2 + I) = 2 \cdot 5 \cdot 10 = 100$ — no entries needed. Trace and determinant are the first two symmetric functions of the spectrum; Newton's identities convert power sums $\operatorname{tr}(A^k)$ into the characteristic polynomial and back. If $\operatorname{tr}(A^k) = 0$ for $k = 1..n$, then all eigenvalues vanish and A is nilpotent — a classic Putnam lemma.

Cayley–Hamilton and minimal polynomials

Every matrix satisfies its own characteristic polynomial: $p_A(A) = 0$. For 2×2: $A^2 - (\operatorname{tr} A)A + (\det A)I = 0$.

Consequences: A^n is a combination of $I, A, \ldots, A^{n-1}$, so powers reduce; A^{-1} is a polynomial in A when $\det A \neq 0$; and a matrix with $A^2 = A$ (idempotent) or $A^2 = I$ (involution) or $A^k = 0$ (nilpotent) has eigenvalues confined to the roots of the corresponding polynomial — $\{0, 1\}$, $\{\pm 1\}$, $\{0\}$. Rank of an idempotent equals its trace. These one-liners decide many "show that" problems.

Matrices over finite fields: count

$|GL_n(\mathbb{F}_q)| = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$ — pick columns one at a time, each outside the span of the previous.

$|GL_2(\mathbb{F}_2)| = 3 \cdot 2 = 6$ ($\cong S_3$); $|GL_2(\mathbb{F}_3)| = 8 \cdot 6 = 48$; $|SL_n| = |GL_n|/(q-1)$. The number of k-dimensional subspaces of $\mathbb{F}_q^n$ is the Gaussian binomial $\binom{n}{k}_q$. Over $\mathbb{F}_2$, "how many symmetric / nilpotent / invertible matrices" problems all reduce to this counting-by-columns template.

Real symmetric matrices are the nice case

Real eigenvalues, orthogonal eigenvectors, diagonalizable — so $\mathbf{x}^{\mathsf{T}} A \mathbf{x}$ on the unit sphere ranges over $[\lambda_{\min}, \lambda_{\max}]$ (the Rayleigh quotient). Positive definiteness is "all eigenvalues positive," tested by leading principal minors. Half of Putnam linear algebra is recognizing that the matrix in front of you is symmetric and applying the spectral theorem.

Try it: Cayley–Hamilton reduces a power

For $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, express A^5 as $aA + bI$.

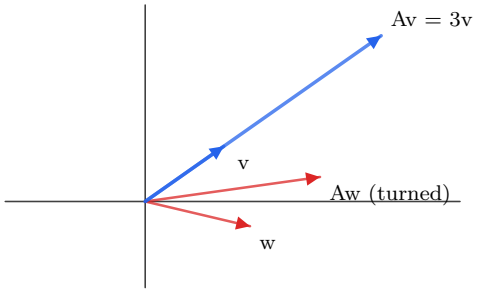
Answer. $\operatorname{tr} = 1$, $\det = -1$: $A^2 = A+I$. Then $A^3 = A^2+A = 2A+I$, $A^4 = 2A^2+A = 3A+2I$, $A^5 = 3A^2+2A = 5A+3I$. The coefficients are Fibonacci numbers — because A is the Fibonacci matrix.

11.2 Proofs & Why It Matters

Proof: $\operatorname{tr} A^k = 0$ for $k = 1, \dots, n$ forces A nilpotent

If an $n \times n$ complex matrix has $\operatorname{tr} A^k = 0$ for all $1 \le k \le n$, then every eigenvalue is 0.

An eigenvector keeps its direction; A only scales it.



Let the distinct nonzero eigenvalues be $\lambda_1, \dots, \lambda_m$ with multiplicities $c_1, \dots, c_m \ge 1$. The hypotheses say $\sum_i c_i \lambda_i^k = 0$ for $k = 1, \dots, m$ (we only need $m \le n$ of them): a homogeneous linear system in the unknowns c_i whose coefficient matrix $[\lambda_i^k]$ is a Vandermonde matrix in distinct nonzero nodes times $\operatorname{diag}(\lambda_i)$ — invertible. So all $c_i = 0$, contradicting $c_i \ge 1$ unless $m = 0$. ■ Significance: traces of powers determine the spectrum (Newton's identities in matrix clothing); the converse — a nilpotent matrix has all power-traces zero — is immediate from triangularization.

Proof: Cayley–Hamilton via the adjugate

Every square matrix satisfies its own characteristic polynomial: $p_A(A) = 0$.

Let $p_A(t) = \det(tI - A) = t^n + c_{n-1}t^{n-1} + \cdots + c_0$ and let $B(t) = \operatorname{adj}(tI - A)$, a matrix whose entries are polynomials in t of degree $\le n-1$; write $B(t) = \sum_{k=0}^{n-1} B_k t^k$ with constant matrices B_k. The adjugate identity $(tI - A)B(t) = p_A(t)I$ holds for every scalar t, so comparing coefficients of t^k on both sides gives $B_{k-1} - AB_k = c_k I$ (with $B_{-1} = B_n = 0$). Multiply the kth equation by A^k on the left and add over k: the left side telescopes to 0, the right side is $\sum c_k A^k = p_A(A)$. ■ (The one-line "substitute $t = A$ into $\det(tI - A)$" is NOT a proof — t is a scalar there.) Significance: A^n is a combination of $I, A, \dots, A^{n-1}$, so powers, inverses, and e^A all live in an n-dimensional space; that is what makes "compute A^{100}" tractable.

Proof: the spectral theorem for real symmetric matrices

A real symmetric matrix has real eigenvalues and an orthonormal basis of eigenvectors.

Real eigenvalues: if $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \ne 0$ complex, then $\bar{\mathbf{x}}^{\mathsf{T}} A\mathbf{x} = \lambda|\mathbf{x}|^2$, and taking the conjugate

transpose of the same scalar (using $A^{\mathsf{T}} = A$ real) gives $\bar{\lambda}|\mathbf{x}|^2$; hence $\lambda = \bar{\lambda}$. Orthogonality: for eigenvectors with $\lambda \neq \mu$, $\lambda\,\mathbf{y}^{\mathsf{T}}\mathbf{x} = \mathbf{y}^{\mathsf{T}}A\mathbf{x} = (A\mathbf{y})^{\mathsf{T}}\mathbf{x} = \mu\,\mathbf{y}^{\mathsf{T}}\mathbf{x}$, so $\mathbf{y}^{\mathsf{T}}\mathbf{x} = 0$. Full basis: take a unit eigenvector $\mathbf{q}_1$ (it exists since the characteristic polynomial has a real root); A maps the orthogonal complement $\mathbf{q}_1^{\perp}$ to itself (if $\mathbf{x} \perp \mathbf{q}_1$ then $\mathbf{q}_1^{\mathsf{T}}A\mathbf{x} = (A\mathbf{q}_1)^{\mathsf{T}}\mathbf{x} = \lambda_1\mathbf{q}_1^{\mathsf{T}}\mathbf{x} = 0$), and restricted there it is again symmetric; induct on the dimension. ■ Significance: $\mathbf{x}^{\mathsf{T}}A\mathbf{x} = \sum \lambda_i c_i^2$ in the eigenbasis, which is the Rayleigh-quotient principle and the reason every quadratic-form extremum on the Putnam is "largest or smallest eigenvalue".

Proof: Cayley–Hamilton for diagonalizable matrices, and why it extends

$p_A(A) = 0$ where $p_A(\lambda) = \det(\lambda I - A)$.

If $A = PDP^{-1}$ with D diagonal, then $p_A(A) = P\,p_A(D)P^{-1}$, and $p_A(D)$ is diagonal with entries $p_A(\lambda_i) = 0$. So $p_A(A) = 0$. Diagonalizable matrices are dense in all matrices (perturb the eigenvalues to be distinct), and $A \mapsto p_A(A)$ is continuous (a polynomial in the entries), so the identity persists for every matrix. ■ Significance: the density argument is itself a Putnam technique — prove an algebraic identity on a dense set, then pass to the limit.

Proof: trace-zero powers force nilpotence

If $\operatorname{tr}(A^k) = 0$ for $k = 1, \ldots, n$ (over $\mathbb{C}$), then A is nilpotent.

Let the distinct nonzero eigenvalues be $\mu_1, \ldots, \mu_m$ with multiplicities $c_j \geq 1$. The hypotheses say $\sum_j c_j\mu_j^k = 0$ for $k = 1..n$, in particular for $k = 1..m$. The matrix of this linear system in the unknowns c_j is $[\mu_j^k]$, a Vandermonde times $\operatorname{diag}(\mu_j)$, invertible for distinct nonzero μ_j. So all $c_j = 0$ — contradiction unless $m = 0$: every eigenvalue is 0, and A is nilpotent. ■ Vandermonde reappears as the engine of a spectral argument — the two linear algebra topics are one subject.

11.3 Going Deeper: Worked Problems

Worked: an eigenvalue argument for a "show that"

Let A be a real $n \times n$ matrix with $A^3 = A$. Show $\operatorname{rank}(A) = \operatorname{tr}(A^2)$.

Step 1 — A satisfies $x^3 - x = x(x-1)(x+1)$, which has distinct roots, so A is diagonalizable with eigenvalues in $\{-1, 0, 1\}$. Step 2 — in the eigenbasis, A^2 is diagonal with entries $\lambda_i^2 \in \{0, 1\}$: it is 1 exactly for the nonzero eigenvalues. Step 3 — so $\operatorname{tr}(A^2)$ counts nonzero eigenvalues, which (diagonalizable!) equals the rank. ■ The whole proof is: a polynomial with distinct roots forces diagonalizability, and then everything is a count.

Worked: counting matrices over $\mathbb{F}_2$

How many 3×3 matrices over $\mathbb{F}_2$ are invertible, and what fraction is that?

Step 1 — columns one at a time: the first is any nonzero vector ($2^3 - 1 = 7$), the second avoids the 2 vectors in the span of the first ($8 - 2 = 6$), the third avoids the 4 in the span of the first two

$(8-4=4)$. Step 2 — $7\cdot 6\cdot 4=168$ (this group is $GL_3(\mathbb{F}_2)$, the second-smallest nonabelian simple group). Step 3 — fraction: $\frac{168}{512}\approx 0.33$; as $n\to\infty$ the fraction of invertible matrices over $\mathbb{F}_2$ tends to $\prod_{k\geq 1}(1-2^{-k})\approx 0.289$. A random binary matrix is singular with probability about 0.71.

12 Polynomials

Vieta, Newton, roots of unity, interpolation, integer polynomials

12.1 Working With Roots You Never Find

Vieta and Newton's identities

For $x^n-e_1x^{n-1}+e_2x^{n-2}-\cdots$, the e_k are the elementary symmetric functions of the roots, and the power sums obey $p_k=e_1p_{k-1}-e_2p_{k-2}+\cdots\pm ke_k$.

So $\sum r_i^2=e_1^2-2e_2$ always (= 4 for x^4-2x^2+3, whose roots are complex), and any symmetric expression in the roots is a polynomial in the coefficients. The reverse direction — given p_1,p_2,p_3, recover e_1,e_2,e_3 and hence the roots — solves systems like $x+y+z=6$, $x^2+y^2+z^2=14$, $x^3+y^3+z^3=36$.

Roots of unity and the filter

$\sum_{k=0}^{n-1}\omega^{jk}=n$ if $n\mid j$, else 0. Hence $\sum_{k\equiv r\,(n)}a_k=\frac{1}{n}\sum_j\omega^{-jr}P(\omega^j)$.

The filter extracts every nth coefficient of a polynomial (or generating function): the number of subsets of $\{1..N\}$ with size divisible by 3, the sum of $\binom{n}{3k}$, the coefficient sums of $(1+x+x^2)^n$ at spacing 3. Also $x^n-1=\prod(x-\omega^k)$ gives $\prod_{k=1}^{n-1}(1-\omega^k)=n$ and the trig products $\prod\sin\frac{k\pi}{n}=\frac{n}{2^{n-1}}$.

Integer polynomials and interpolation

For $P\in\mathbb{Z}[x]$: $a-b\mid P(a)-P(b)$. Lagrange: the degree-$<n$ polynomial through n points is $\sum_i y_i\prod_{j\neq i}\frac{x-x_j}{x_i-x_j}$. Finite differences: a degree-d polynomial has constant dth differences.

The divisibility fact kills "find an integer polynomial with $P(1)=2$, $P(4)=7$" ($3\nmid 5$). Interpolation with consecutive integer nodes is fastest by finite differences ($P(0),P(1),P(2)=1,2,5$ gives second difference 2, so $P(3)=10$). Chebyshev polynomials $T_n(\cos\theta)=\cos n\theta$ handle "polynomial bounded by 1 on $[-1,1]$" extremal problems.

The remainder trick

To find $x^{100}\bmod(x^2-1)$, do not divide: write $x^{100}=Q(x)(x^2-1)+ax+b$ and plug in the roots $x=\pm 1$ of the divisor to get two equations for a,b. For a repeated root, differentiate first. Remainders modulo a polynomial are found by evaluating, never by long division.

Try it: Newton in reverse

Real x, y satisfy $x + y = 5$ and $x^3 + y^3 = 35$. Find xy.

Answer. $x^3 + y^3 = (x + y)^3 - 3xy(x + y)$: $35 = 125 - 15xy$, so $xy = 6$ — and indeed $\{x, y\} = \{2, 3\}$.

12.2 Proofs & Why It Matters

Proof: Newton's identity for p_2 and the general pattern

$p_2 = e_1^2 - 2e_2$, and in general $p_k - e_1 p_{k-1} + e_2 p_{k-2} - \cdots + (-1)^k k e_k = 0$.

$e_1^2 = (\sum r_i)^2 = \sum r_i^2 + 2\sum_{i<j} r_i r_j = p_2 + 2e_2$ — rearrange. For the general identity, expand $\prod_i (1 - r_i t) = \sum (-1)^k e_k t^k$, take the logarithmic derivative, and compare coefficients of t^{k-1} with $-\sum_i \frac{r_i}{1 - r_i t} = -\sum_k p_k t^{k-1}$. ■ Significance: the generating-function proof is a template — many Putnam identities are "compare coefficients after a log-derivative."

Proof: the roots-of-unity filter

For $P(x) = \sum a_k x^k$ and $\omega = e^{2\pi i/n}$: $\sum_{n|k} a_k = \frac{1}{n} \sum_{j=0}^{n-1} P(\omega^j)$.

$\frac{1}{n} \sum_j P(\omega^j) = \sum_k a_k \cdot \frac{1}{n} \sum_j \omega^{jk}$, and the inner average is 1 when $n \mid k$ (every term is 1) and 0 otherwise (a geometric series with ratio $\omega^k \neq 1$ summing to $\frac{\omega^{kn} - 1}{\omega^k - 1} = 0$). ■ Significance: this is the discrete Fourier transform in its simplest form; shifting by ω^{-jr} picks out the residue class r. It is how one counts subsets by size mod n — a Putnam favorite.

12.3 Going Deeper: Worked Problems

Worked: subsets by size mod 3

How many subsets of a 10-element set have size divisible by 3?

Step 1 — the answer is $\sum_{3|k} \binom{10}{k}$, the filter applied to $P(x) = (1+x)^{10}$. Step 2 — $\frac{1}{3}\left[P(1) + P(\omega) + P(\omega^2)\right]$ with $\omega = e^{2\pi i/3}$: $P(1) = 1024$; $1 + \omega = -\omega^2 = e^{i\pi/3}$, so $P(\omega) = e^{10i\pi/3} = e^{-2i\pi/3}$; similarly $P(\omega^2) = e^{2i\pi/3}$. Step 3 — $P(\omega) + P(\omega^2) = 2\cos\frac{2\pi}{3} = -1$. Step 4 — total $\frac{1024 - 1}{3} = 341$. Check: $\binom{10}{0} + \binom{10}{3} + \binom{10}{6} + \binom{10}{9} = 1 + 120 + 210 + 10 = 341$ ✓.

Worked: a polynomial determined by its values at integers

A polynomial P of degree n satisfies $P(k) = \dfrac{k}{k+1}$ for $k = 0, 1, \ldots, n$. Find $P(n+1)$.

Step 1 — consider $Q(x) = (x+1)P(x) - x$, degree $n+1$, vanishing at $x = 0, 1, \ldots, n$: so $Q(x) = c\,x(x-1)\cdots(x-n)$. Step 2 — evaluate at $x = -1$: $Q(-1) = 0 \cdot P(-1) + 1 = 1 = c(-1)(-2)\cdots(-n-$

$1) = c(-1)^{n+1}(n+1)!$, giving $c = \frac{(-1)^{n+1}}{(n+1)!}$. Step 3 — at $x = n+1$: $Q(n+1) = c(n+1)! = (-1)^{n+1}$, so $(n+2)P(n+1) - (n+1) = (-1)^{n+1}$ and $P(n+1) = \dfrac{n+1+(-1)^{n+1}}{n+2}$ — equal to 1 for odd n and $\frac{n}{n+2}$ for even n. The auxiliary polynomial with known roots is THE technique for "values at consecutive integers."

13 Number Theory

Orders, valuations, quadratic residues, Diophantine equations, divisor sums

13.1 Proof-Level Number Theory

Orders, primitive roots, and CRT

The order of a mod n divides every k with $a^k \equiv 1$, hence divides $\varphi(n)$. Compute orders mod composites prime-power by prime-power and take the lcm.

Last two digits of 3^{2026}: the order of 3 mod 100 is 20, and $2026 \equiv 6$, so $3^6 = 729 \to 29$. Mod 2^k ($k \geq 3$) the equation $x^2 \equiv 1$ has FOUR solutions, so $x^2 \equiv 1 \pmod{24}$ has $4 \cdot 2 = 8$ — a standard trap. Primitive roots exist exactly mod $2, 4, p^k, 2p^k$.

Valuations: Legendre, Kummer, LTE

$v_p(n!) = \sum_{k \geq 1} \lfloor n/p^k \rfloor$; $v_p\binom{n}{k}$ is the number of carries adding $k+(n-k)$ in base p; $v_p(a^n - b^n) = v_p(a-b) + v_p(n)$ for odd $p \mid a-b$.

$v_5(1000!) = 200 + 40 + 8 + 1 = 249$. $v_2\binom{2026}{1013}$ is the number of 1-bits of 1013. LTE: $v_3(2^{243}+1) = 1 + 5 = 6$ (the $a^n + b^n$ form for odd n). Valuations turn "largest power dividing" into arithmetic, and they are the language of most Putnam number theory write-ups.

Divisor sums and multiplicative functions

τ, σ, φ are multiplicative: compute on prime powers and multiply. $\sigma(360) = 15 \cdot 13 \cdot 6 = 1170$; $\sum_{d \mid n} \varphi(d) = n$.

Sums over divisors are attacked by swapping the order of summation: $\sum_{n \leq N} \tau(n) = \sum_{d \leq N} \lfloor N/d \rfloor$. Mbius inversion recovers f from $F(n) = \sum_{d \mid n} f(d)$. Quadratic residues: $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$, $\left(\frac{2}{p}\right) = 1$ iff $p \equiv \pm 1 \pmod 8$, and reciprocity for the rest.

Diophantine reflexes

Factor (SFFT: $xy + ax + by = c$ becomes $(x+b)(y+a) = c + ab$); bound (compare growth rates, squeeze between consecutive squares); take mod a small number (squares are $0, 1 \bmod 4$, cubes $0, \pm 1 \bmod 9$); descend (Vieta jumping, infinite descent). Nearly every Putnam Diophantine problem yields to one of

these four, and knowing which to try first is a matter of pattern practice.

Try it: a valuation

Find the largest k such that 2^k divides $\binom{100}{50}$.

Answer. Kummer: count carries adding $50+50$ in base 2. $50 = 110010_2$; doubling a binary number produces one carry per 1-bit, and 50 has three 1-bits: $k = 3$. (Legendre confirms: $v_2(100!) - 2v_2(50!) = 97 - 2 \cdot 47 = 3$.)

13.2 Proofs & Why It Matters

Proof: lifting the exponent (odd prime case)

For odd prime p with $p \mid a-b$ and $p \nmid ab$: $v_p(a^n - b^n) = v_p(a-b) + v_p(n)$.

Two lemmas. (i) If $p \nmid n$ then $v_p(a^n - b^n) = v_p(a-b)$: since $a^n - b^n = (a-b)\sum a^{n-1-i}b^i$ and the sum is $\equiv na^{n-1} \not\equiv 0 \pmod p$. (ii) $v_p(a^p - b^p) = v_p(a-b) + 1$: write $a = b + kp^t$ and expand a^p binomially — the terms beyond the first two carry extra factors of p, and the $\binom{p}{2}$ term supplies exactly one more (this is where p odd is used). Induct on $v_p(n)$, peeling off one factor of p at a time with (ii) and finishing with (i). ■ Significance: LTE is the exact accounting of how a prime divides a difference of powers — it replaces pages of casework in olympiad and Putnam solutions alike.

Proof: $\sum_{d \mid n} \varphi(d) = n$

A one-line counting proof of a multiplicative identity.

Sort the fractions $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$ by their reduced denominator d (a divisor of n). Exactly $\varphi(d)$ of them reduce to denominator d (the numerators coprime to d). Summing over d counts all n fractions. ■ Significance: the argument proves more — that the multiplicative group mod p is cyclic (count elements by order, compare with $\sum \varphi(d) = p-1$), i.e. primitive roots exist.

13.3 Going Deeper: Worked Problems

Worked: an order computation via CRT

Find the order of 3 modulo 1000.

Step 1 — $1000 = 8 \cdot 125$; the order mod 1000 is the lcm of the orders mod 8 and mod 125. Step 2 — mod 8: $3^2 = 9 \equiv 1$, order 2. Step 3 — mod 125: $\varphi(125) = 100$, so the order divides 100; test the maximal proper divisors 50 and 20. $3^{20} \equiv 1 + 20 \cdot \ldots$ — compute: $3^5 = 243 \equiv -7$, $3^{10} \equiv 49$, $3^{20} \equiv 49^2 = 2401 \equiv 26 \neq 1$; $3^{50} = (3^{10})^5 \equiv 49^5$, and $49^2 \equiv 26$, $49^4 \equiv 26^2 = 676 \equiv 51$, $49^5 \equiv 51 \cdot 49 = 2499 \equiv -1 \neq 1$. So the order mod 125 is 100. Step 4 — $\operatorname{lcm}(2, 100) = 100$. The method: never search; use φ to list candidates and test only the maximal proper divisors.

Worked: a Diophantine equation by bounding

Find all positive integers n such that $n^2 + 3n + 5$ is a perfect square.

Step 1 — squeeze between consecutive squares: for $n \geq 1$, $(n+1)^2 = n^2+2n+1 < n^2+3n+5$ and $(n+2)^2 = n^2+4n+4 > n^2+3n+5$ iff $n > 1$. Step 2 — so for $n \geq 2$ the expression lies strictly between two consecutive squares: never a square. Step 3 — check $n = 1$: $1+3+5 = 9 = 3^2$ ✓. Answer: only $n = 1$. Bounding by neighboring squares (or cubes) is the first thing to try on "when is this polynomial a perfect power."

14 Combinatorics & Generating Functions

Bijections, recurrences, generating functions, inclusion–exclusion, extremal arguments

14.1 Counting With Structure

Bijections before formulas

Pair objects with objects. Subsets of $\{1..10\}$ with even sum: toggle the element 1 — an involution swapping parity — so exactly half, 512.

Dyck paths $\leftrightarrow$ balanced parentheses $\leftrightarrow$ triangulations (Catalan, $C_n = \frac{1}{n+1}\binom{2n}{n}$, $C_5 = 42$, proved by the reflection bijection). Compositions of $n \leftrightarrow$ subsets of $n-1$ gaps (2^{n-1}). Stars and bars: nonnegative solutions of $x_1 + \cdots + x_k = n \leftrightarrow$ arrangements of n stars and $k-1$ bars. A bijection is a proof; a formula is a claim.

Recurrences and generating functions

Set up a_n by the first choice; solve by characteristic equation or by the generating function $A(x) = \sum a_n x^n$, which turns the recurrence into algebra.

Domino tilings of $2 \times n$: $t_n = t_{n-1} + t_{n-2}$ (Fibonacci; $t_{10} = 89$). Derangements: $D_n = (n-1)(D_{n-1} + D_{n-2})$, or by inclusion–exclusion $n! \sum \frac{(-1)^k}{k!}$ ($D_6 = 265$). Generating functions multiply when choices are independent: $(1 + x + x^2 + \cdots)^k = (1-x)^{-k}$ is stars and bars, $\prod(1 + x^i)$ counts partitions into distinct parts, and Vandermonde $\sum \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}$ is $(1+x)^a(1+x)^b$.

Inclusion–exclusion and extremal arguments

$|A_1 \cup \cdots \cup A_n| = \sum |A_i| - \sum |A_i \cap A_j| + \cdots$. And: the largest, smallest, or extreme object has the least room to misbehave.

PIE counts surjections, derangements, and "no two adjacent" arrangements. Extremal: among finitely many points not all collinear, a pair at minimum distance... ; in a tournament, the vertex with maximum out-degree...; the longest path in a graph has endpoints whose neighbors all lie on it. Pigeonhole with $kn + 1$ objects forces $k + 1$ in a box; the Erds–Szekeres theorem ($n^2 + 1$ reals

contain a monotone subsequence of length $n+1$) is pigeonhole plus extremal, and it is a Putnam regular.

Graph theory at Putnam level

Degree sum $= 2|E|$; a tree on n vertices has $n-1$ edges; bipartite iff no odd cycle; Mantel ($\leq n^2/4$ edges triangle-free) and Turn; Hall's marriage condition; Euler's $V-E+F=2$ for planar graphs (with $E \leq 3V-6$). Model anything with "pairs" or "relations" as a graph and these become available.

Try it: inclusion–exclusion

How many integers in $\{1,\ldots,1000\}$ are divisible by none of $2,3,5$?

Answer. $1000-(500+333+200)+(166+100+66)-33 = 1000-1033+332-33 = 266$. (Approximately $1000\cdot\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5} = 266.7$ ✓.)

14.2 Proofs & Why It Matters

Proof: the Catalan reflection

Lattice paths from $(0,0)$ to (n,n) never rising above $y=x$ number $\frac{1}{n+1}\binom{2n}{n}$.

All paths: $\binom{2n}{n}$. A BAD path touches the line $y=x+1$; reflect the portion after its first touch across that line — it now ends at $(n-1,n+1)$. This is a bijection between bad paths and all paths to $(n-1,n+1)$, of which there are $\binom{2n}{n-1}$. Good paths: $\binom{2n}{n}-\binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n}$. ■ Significance: the reflection principle also solves the ballot problem and gambler's ruin — one bijection, three theorems.

Proof: Erds–Szekeres

Any sequence of n^2+1 distinct reals has a monotone subsequence of length $n+1$.

Label each term a_i with (u_i, d_i) = lengths of the longest increasing and longest decreasing subsequence ENDING at a_i. For $i<j$: if $a_i < a_j$ then $u_j > u_i$; if $a_i > a_j$ then $d_j > d_i$ — so the labels are all distinct. If every $u_i, d_i \leq n$ there are only n^2 possible labels for n^2+1 terms: pigeonhole. ■ Significance: the proof is the template "assign labels that are forced to be distinct, then count labels" — a pigeonhole pattern that recurs across Putnam combinatorics.

14.3 Going Deeper: Worked Problems

Worked: a generating-function count

How many ways can 20 be written as an ordered sum of 1s, 2s, and 3s?

Step 1 — let c_n count such compositions; the first part is 1, 2, or 3: $c_n = c_{n-1}+c_{n-2}+c_{n-3}$ (tribonacci), $c_0=1$, $c_1=1$, $c_2=2$. Step 2 — the generating function is $\sum c_n x^n = \dfrac{1}{1-x-x^2-x^3}$

(each composition is a word in the "alphabet" $x + x^2 + x^3$). Step 3 — run the recurrence: $1, 1, 2, 4, 7, 13, 24, 44, 81, 149, 274, 504, 927, 1705, 3136, 5768, 10609, 19513, 35890, 66012, 121415$: $c_{20} = 121415$. Step 4 — growth rate ≈ 1.839, the real root of $x^3 = x^2 + x + 1$ — the generating function's smallest pole.

Worked: an extremal argument

In a tournament (every pair of teams plays once, no ties), show some team T has the property that every other team either lost to T or lost to a team that lost to T.

Step 1 — take T with the MAXIMUM number of wins. Step 2 — suppose some team S is not reachable in ≤ 2 steps: S beat T, and S beat every team that T beat. Step 3 — then S has at least (wins of T) $+1$ wins — the teams T beat, plus T itself — contradicting maximality. ■ Choosing the extreme object made the contradiction one sentence long; that is the extremal principle at its cleanest.

15 Probability & Expectation

Indicators, linearity, geometric probability, random walks, coupon collecting

15.1 Expectation Without Distributions

Linearity and indicators

$E[\sum X_i] = \sum E[X_i]$ with NO independence assumption. Write a count as a sum of 0/1 indicators, and expectation becomes a sum of probabilities.

Expected fixed points of a random permutation: $n \cdot \frac{1}{n} = 1$. Expected number of adjacent equal letters, expected number of records in a random sequence (H_n), expected inversions ($\binom{n}{2}/2$) — all indicator sums. The distribution may be a nightmare; the expectation is a one-liner. Coupon collector: waits are geometric with expectations $\frac{n}{n-k}$, summing to nH_n ($\frac{25}{3}$ for $n = 4$).

Geometric probability and random walks

Probability = favorable measure ÷ total measure; expectations become integrals. A ± 1 walk returns to 0 after $2n$ steps with probability $\binom{2n}{n}/4^n \approx \frac{1}{\sqrt{\pi n}}$.

Two uniform points on $[0, 1]$: $E|X - Y| = \frac{1}{3}$, $E[\max] = \frac{2}{3}$, $P(|X - Y| < \frac{1}{4}) = \frac{7}{16}$ (unit-square pictures). Broken-stick triangle: $\frac{1}{4}$. Buffon's needle: $\frac{2\ell}{\pi d}$. Random walk returns: $\binom{4}{2}/16 = \frac{3}{8}$ after four steps; the $\frac{1}{\sqrt{\pi n}}$ decay is why $\sum P(\text{return at } 2n) = \infty$ and the 1D walk is recurrent, while in 3D it is not.

Conditioning and first-step analysis

For expected times and absorption probabilities, condition on the first step: $E_i = 1 + \sum_j p_{ij} E_j$

— a small linear system.

Expected rolls until two consecutive sixes: 42. Expected steps for a walker advancing with probability $\frac{2}{3}$ else staying, to go 3 steps: $3 \cdot \frac{3}{2} = 4.5$. Gambler's ruin from k with target N (fair coin): ruin probability $1 - \frac{k}{N}$, expected duration $k(N-k)$. Markov-chain steady states (the Linear Algebra course) are the long-run version of the same conditioning.

When independence is not there

Indicators handle dependence for expectations, but NOT for variances: $\operatorname{Var}(\sum X_i) = \sum \operatorname{Var}(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j)$. Putnam problems that ask for a variance or a probability of a specific count usually intend inclusion–exclusion or a generating function, not linearity.

Try it: an indicator sum

A fair coin is flipped 10 times. What is the expected number of times a head is immediately followed by a tail?

Answer. Nine adjacent pairs, each HT with probability $\frac{1}{4}$: expectation $\frac{9}{4}$. The pairs overlap and are dependent — linearity does not care.

15.2 Proofs & Why It Matters

Proof: linearity of expectation

$E[X+Y] = E[X] + E[Y]$ for ANY random variables with finite expectations.

For discrete variables, $E[X+Y] = \sum_{x,y}(x+y)P(X=x, Y=y) = \sum_x x \sum_y P(X=x, Y=y) + \sum_y y \sum_x P(X=x, Y=y) = \sum_x xP(X=x) + \sum_y yP(Y=y)$ — the joint distribution marginalizes, and no independence was used. ■ Significance: the ENTIRE indicator method rests on this; it is worth stating in a solution that independence is not required, because graders check that you know.

Proof: the coupon collector expectation

Collecting all n equally likely coupons takes nH_n draws on average.

After k distinct coupons are held, each draw is new with probability $\frac{n-k}{n}$, independently of the past; the number of draws until the next new one is geometric with mean $\frac{n}{n-k}$. Sum over $k = 0, \ldots, n-1$ (linearity): $n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\right) = nH_n$. ■ Significance: $nH_n \approx n \ln n$ is the universal cost of "cover everything by random sampling" — hashing, epidemic spreading, and randomized algorithms all hit this bound.

15.3 Going Deeper: Worked Problems

Worked: expected value of a maximum

Three numbers are chosen independently and uniformly from $[0, 1]$. Find the expected value of the largest.

Step 1 — the CDF of the max: $P(\max \le t) = t^3$. Step 2 — for a nonnegative variable, $E[M] = \int_0^1 P(M > t)\,dt = \int_0^1 (1 - t^3)dt = 1 - \frac{1}{4} = \frac{3}{4}$. Step 3 — general pattern: $E[\text{max of } n] = \frac{n}{n+1}$ and $E[\min] = \frac{1}{n+1}$; the n order statistics split $[0, 1]$ into $n + 1$ pieces of equal expected length. Step 4 — the tail-integral formula $E[X] = \int_0^\infty P(X > t)\,dt$ is the fastest route to expectations of maxima, minima, and waiting times.

Worked: gambler's ruin by first-step analysis

A fair-coin gambler starts with \$3 and stops at \$0 or \$10. Find the probability of reaching \$10 and the expected number of bets.

Step 1 — let p_k be the probability of reaching 10 from k: $p_k = \frac{1}{2}p_{k-1} + \frac{1}{2}p_{k+1}$, $p_0 = 0$, $p_{10} = 1$. The recurrence forces p_k linear: $p_k = \frac{k}{10}$, so $p_3 = 0.3$. Step 2 — let E_k be the expected duration: $E_k = 1 + \frac{1}{2}E_{k-1} + \frac{1}{2}E_{k+1}$ with $E_0 = E_{10} = 0$; the particular solution of the constant inhomogeneity is $-k^2$, so $E_k = k(10 - k)$ and $E_3 = 21$. Step 3 — sanity: from the midpoint $E_5 = 25$ is the longest expected game; near the boundaries games end fast. Step 4 — conditioning on one step turned a random process into two linear recurrences — the whole method.

16 Abstract Algebra Basics

Groups, orders, Lagrange, cyclic groups, rings and fields at the Putnam level

16.1 The Definitions the Putnam Assumes

Groups and orders

A group is a set with an associative operation, an identity, and inverses. The ORDER of an element is the least k with $g^k = e$; Lagrange: the order of a subgroup (hence of any element) divides the order of the group.

In $\mathbb{Z}_n$ (addition mod n) the element a has order $\frac{n}{\gcd(a,n)}$: $6 \in \mathbb{Z}_{15}$ has order 5. Cyclic groups of order n have exactly one subgroup per divisor of n ($\mathbb{Z}_{12}$ has 6 subgroups) and exactly $\varphi(d)$ elements of order d ($\mathbb{Z}_8$ has 2 of order 4). Consequences of Lagrange: a group of prime order is cyclic; $a^{|G|} = e$ for all a — which IS Fermat–Euler when $G = (\mathbb{Z}/n)^\times$.

Recurring Putnam groups

Symmetric groups S_n (permutations, cycle structure, parity), matrix groups $GL_n(\mathbb{F}_q)$, units mod n, dihedral groups, and "the set of all maps satisfying " that turns out to be a group.

A permutation's order is the lcm of its cycle lengths — a Putnam question asks for the maximum order in S_n (Landau's function) or the number of elements of a given order. Any finite subset of a group closed under the operation is a subgroup (finite cancellation forces inverses). If every element satisfies $g^2 = e$ the group is abelian. These small lemmas are quotable and graders expect them.

Rings, fields, and polynomials over them

A field has division by nonzero elements; $\mathbb{F}_p = \mathbb{Z}/p$ and its extensions $\mathbb{F}_{p^k}$ are the finite ones. A polynomial of degree n over a FIELD has at most n roots.

Over $\mathbb{Z}/n$ with n composite that fails ($x^2 - 1$ has four roots mod 8) — the root-count bound is a FIELD property. Wilson's theorem $(p-1)! \equiv -1$ is "pair each element with its inverse"; $x^{p-1} - 1 = \prod_{a \neq 0}(x - a)$ in $\mathbb{F}_p[x]$ is the same fact as a polynomial identity. Frobenius $x \mapsto x^p$ is a ring homomorphism in characteristic p: $(a+b)^p = a^p + b^p$.

How much abstract algebra the Putnam really wants

Definitions, Lagrange, cyclic-group facts, and comfort with S_n and matrix groups. Sylow theorems and classification results almost never appear; when a problem says "group," the intended solution usually uses only closure, associativity, and counting.

Try it: Lagrange in action

Show that in any group of order 15, an element of order 3 and an element of order 5 commute if their product has order 15... simpler: how many elements of order 15 does $\mathbb{Z}_{15}$ have?

Answer. $\varphi(15) = 8$ — the generators $1, 2, 4, 7, 8, 11, 13, 14$. (And indeed every group of order 15 is cyclic, by a Sylow argument you will not need on the exam.)

16.2 Proofs & Why It Matters

Proof: Lagrange's theorem

For a subgroup H of a finite group G, $|H|$ divides $|G|$.

The left cosets gH partition G: two cosets either coincide or are disjoint (if $g_1H \cap g_2H \neq \emptyset$ then $g_2^{-1}g_1 \in H$, forcing $g_1H = g_2H$), and every element lies in its own coset. Each coset has exactly $|H|$ elements (the map $h \mapsto gh$ is a bijection $H \to gH$). So $|G| = (\text{number of cosets}) \cdot |H|$. ■ Significance: Fermat's little theorem, Euler's theorem, and "a group of prime order is cyclic" are all corollaries — Lagrange is the one group-theory proof every Putnam competitor should be able to write from memory.

Proof: a polynomial over a field has at most deg roots

The property that separates fields from rings like $\mathbb{Z}/8$.

If $P(a) = 0$, divide: $P(x) = (x - a)Q(x)$ with $\deg Q = \deg P - 1$ (polynomial division needs only that the leading coefficient of $x - a$ is invertible). Any other root $b \neq a$ satisfies $(b - a)Q(b) = 0$,

and in a FIELD $b - a \neq 0$ is invertible, so $Q(b) = 0$. Induct on the degree. ■ Over $\mathbb{Z}/8$, $(b - a)$ can be a zero divisor ($2 \cdot 4 = 0$), and $x^2 - 1$ has the four roots $1, 3, 5, 7$. Significance: this is why $\mathbb{F}_p^\times$ is cyclic and why Lagrange interpolation works — uniqueness of polynomials from values is a field theorem.

16.3 Going Deeper: Worked Problems

Worked: orders in the symmetric group

What is the largest possible order of an element of S_{10}?

Step 1 — the order of a permutation is the lcm of its cycle lengths, which sum to at most 10. Step 2 — maximize lcm subject to the sum: try $2 + 3 + 5 = 10$, lcm 30; $3 + 7 = 10$, lcm 21; $4 + 5 = 9$, lcm 20; $1 + 2 + 3 + 4$? sum 10, lcm 12; $2 + 3 + 5$ wins. Step 3 — the maximum is 30 (Landau's function $g(10) = 30$). Step 4 — the same reasoning gives the number of elements of order 2 in S_n (products of disjoint transpositions) and answers every "orders in S_n" Putnam question: it is a partition problem in disguise.

Worked: a finite subset closed under multiplication

Let G be a group and $H \subseteq G$ a nonempty FINITE subset closed under the group operation. Show H is a subgroup.

Step 1 — take $h \in H$; the powers $h, h^2, h^3, \ldots$ all lie in H (closure) and H is finite, so $h^i = h^j$ for some $i < j$, whence $h^{j-i} = e$: the identity is in H. Step 2 — with $k = j - i \geq 1$, h^{k-1} is in H (it is e if $k = 1$, otherwise a positive power) and $h \cdot h^{k-1} = e$: the inverse of h lies in H. Step 3 — closure, identity, inverses: a subgroup. ■ Finiteness converted "closed under multiplication" into "closed under inverses" — infinite counterexample: the positive integers inside $(\mathbb{Z}, +)$.

17 Functional Equations & Recurrences

Substitution, symmetry, Cauchy's equation, periodicity, fixed points

17.1 Solving for Functions

The substitution playbook

Plug in 0, 1, $-x$, $\frac{1}{x}$, $1 - x$, $y = x$; use the equation to build a second equation, then eliminate.

$f(x) + 2f(1 - x) = 3x$: substitute $x \to 1 - x$ and solve the 2×2 system — $f(x) = 2 - 3x$. $f(x) + 2f(\frac{1}{x}) = 3x$: substitute $x \to \frac{1}{x}$ — $f(x) = \frac{2}{x} - x$. Always VERIFY the candidate in the original equation; the elimination only shows that IF a solution exists it has this form.

Cauchy's equation and its relatives

$f(x + y) = f(x) + f(y)$ has only $f(x) = cx$ on $\mathbb{Q}$ (and on $\mathbb{R}$ under any regularity: continuous,

monotone, bounded on an interval). $f(xy) = f(x)f(y)$, $f(x+y) = f(x)f(y)$, $f(xy) = f(x) + f(y)$ reduce to it by logs and exponentials.

On $\mathbb{Q}$: $f(n) = nf(1)$ by induction, then $f(\frac{p}{q}) = \frac{p}{q}f(1)$ by $qf(\frac{p}{q}) = f(p)$ — so $f(1) = 7$ gives $f(\frac{3}{7}) = 3$. A functional equation with an extra term ($f(x+y) = f(x) + f(y) + xy$) is Cauchy after subtracting a particular solution ($\frac{x^2}{2}$). Multiplicative equations: $f(12) = f(2)^2 f(3)$ when the equation holds for ALL pairs.

Periodicity, involutions, and iteration

Compose the function with itself; find the order. $f(x) = \frac{1}{1-x}$ has $f^{(3)} = x$, so $f^{(2026)} = f$ and $f^{(2026)}(2) = -1$.

Equations like $f(f(x)) = x$ (involutions), $f(x+1) = -f(x)$ (period 2), or $f(x+a) = \frac{1+f(x)}{1-f(x)}$ (a tangent-addition disguise, period $4a$) are solved by computing a few iterates. Fixed points constrain everything: if $f(f(x)) = x$ and f is continuous and increasing then $f(x) = x$. For recurrences $a_{n+1} = g(a_n)$, the fixed points of g and the sign of g' there decide convergence.

Injectivity and surjectivity: the hidden lemmas

If $f(f(x)) = x + c$ then f is injective (compose to cancel) and its range covers a translate of the real line. Showing f is injective lets you cancel it on both sides of an equation; showing it is surjective lets you choose an x with $f(x) =$ anything. Half of Putnam functional-equation solutions are one such lemma followed by a substitution.

Try it: build the second equation

Find all $f : \mathbb{R} \setminus \{0, 1\} \to \mathbb{R}$ with $f(x) + f\left(\frac{1}{1-x}\right) = x$.

Answer. The map $x \to \frac{1}{1-x}$ has order 3; apply it twice more to get three equations in $f(x), f(\frac{1}{1-x}), f(1-\frac{1}{x})$, and solve: $f(x) = \frac{x^3-x+1}{2x(x-1)}$. Verify by substitution.

17.2 Proofs & Why It Matters

Proof: Cauchy's equation on $\mathbb{Q}$

If $f(x+y) = f(x) + f(y)$ for all rationals, then $f(x) = cx$ with $c = f(1)$.

$f(0) = 2f(0)$ gives $f(0) = 0$; $f(-x) = -f(x)$ from $f(x) + f(-x) = f(0)$. Induction gives $f(nx) = nf(x)$ for positive integers n, hence for all integers. For $x = \frac{p}{q}$: $qf(\frac{p}{q}) = f(p) = pf(1)$, so $f(\frac{p}{q}) = \frac{p}{q}f(1)$. ■ Significance: extending to $\mathbb{R}$ needs a regularity hypothesis — continuity at one point, monotonicity, or boundedness on an interval — and without it there are wild solutions (a Hamel basis). A Putnam solution must name which hypothesis it uses.

Proof: a continuous involution with no fixed... an increasing involution is the identity

If f is continuous, strictly increasing, and $f(f(x)) = x$, then $f(x) = x$.

Suppose $f(a) > a$ for some a. Apply the increasing f: $f(f(a)) > f(a)$, i.e. $a > f(a)$ — contradiction. Symmetrically $f(a) < a$ is impossible. ■ Significance: monotonicity plus an involution equation collapses to the identity — the reason problems with $f(f(x)) = x$ always specify "decreasing" or drop monotonicity when they want interesting solutions like $f(x) = \frac{1}{x}$ or $c - x$.

17.3 Going Deeper: Worked Problems

Worked: a recurrence by fixed points

Let $a_1 = 1$ and $a_{n+1} = \sqrt{2 + a_n}$. Show the sequence converges and find the limit.

Step 1 — fixed points of $g(x) = \sqrt{2+x}$: $x^2 = 2 + x$ gives $x = 2$ (the positive root). Step 2 — monotone and bounded: by induction $a_n < 2$ (since $a_n < 2 \Rightarrow a_{n+1} < \sqrt{4} = 2$) and $a_{n+1} > a_n$ (since $\sqrt{2+a} > a$ iff $a^2 - a - 2 < 0$ iff $a < 2$). Step 3 — the monotone convergence theorem gives a limit L, and continuity of g gives $L = g(L)$, so $L = 2$. Step 4 — rate: $g'(2) = \frac{1}{4}$, so the error shrinks by a factor of 4 per step — a contraction, which is the general convergence criterion.

Worked: a functional equation with an extra term

Find all $f : \mathbb{Q} \to \mathbb{Q}$ with $f(x + y) = f(x) + f(y) + 2xy$ and $f(1) = 3$.

Step 1 — guess the particular solution from the extra term: x^2 satisfies $(x + y)^2 = x^2 + y^2 + 2xy$. Step 2 — set $g(x) = f(x) - x^2$; then $g(x + y) = g(x) + g(y)$: Cauchy on $\mathbb{Q}$, so $g(x) = cx$ with $c = g(1) = f(1) - 1 = 2$. Step 3 — $f(x) = x^2 + 2x$. Step 4 — verify: $(x + y)^2 + 2(x + y) = x^2 + 2x + y^2 + 2y + 2xy$ ✓, and $f(1) = 3$ ✓. Subtracting a particular solution to reach Cauchy is the standard reduction.

18 Multivariable Calculus on the Putnam

Double integrals by symmetry, Jacobians that collapse integrands, Lagrange multipliers, and Green's theorem in disguise

18.1 The Moves That Recur

Symmetrize before integrating

Over a region symmetric under $x \leftrightarrow y$, $\iint f(x, y) = \iint f(y, x)$; average the two. Over a region symmetric under $x \to -x$, odd parts vanish.

The Putnam almost never wants a double integral computed by brute force. The integrand is chosen so that a symmetry of the region turns it into something trivial: $\iint_{[0,1]^2} \frac{x}{x+y}\, dA$ averages with its

mirror to $\iint \frac{1}{2} = \frac{1}{2}$. Whenever the region is a square, a disk, or a simplex, list its symmetries FIRST and apply each to the integrand.

Change variables to match the integrand

With $u = x + y$, $v = x - y$ (Jacobian $\frac{1}{2}$), or polar coordinates, or $u = xy$, $v = y/x$: choose the substitution that makes the integrand a function of one new variable.

A Jacobian is not a nuisance factor — it is the reason the trick works. The polar r is what makes $\int e^{-x^2}$ computable; the $\frac{1}{2}$ from $u = x + y$, $v = x - y$ is what turns $\iint_{[0,1]^2} |x - y|\, dA$ into a one-variable integral. The other classic: $\iint$ over the triangle $x, y \geq 0$, $x + y \leq 1$ of $f(x + y)$ equals $\int_0^1 t f(t)\, dt$, because the level sets $x + y = t$ have length proportional to t.

Lagrange multipliers, then check the boundary

To extremize f on $g = c$: solve $\nabla f = \lambda \nabla g$ together with $g = c$. On a compact region also test the boundary; on an unbounded one, argue that f blows up (or decays) away from the candidates.

The Putnam version usually has symmetric f and g, so the candidates are points with several equal coordinates — but the maximum is not always at the fully symmetric point. Compute f at every candidate. For a sum $\sum x_i^2$ with $\sum x_i$ fixed, the extremes are "all equal" and "one variable takes everything", and Lagrange tells you which is which.

Green, Stokes and the divergence theorem as identities

A line integral around a closed curve equals a double integral of a curl; the flux of $\mathbf{r}/|\mathbf{r}|^3$ through any surface enclosing the origin is 4π; the area of a region is $\frac{1}{2} \oint (x\, dy - y\, dx)$.

These appear on the Putnam as identities to recognize, not theorems to apply mechanically. The area formula turns "area of a polygon with vertices" into the shoelace formula; the flux fact makes a hideous surface integral equal 4π or 0 depending only on whether the surface encloses the origin.

Where it goes wrong

Forgetting the Jacobian entirely, or applying Lagrange multipliers where $\nabla g = 0$ (the method finds nothing there) or on a region with a boundary you never tested.

Try it: symmetry

Compute $\displaystyle\iint_{[0,1]^2} \frac{x^2}{x^2 + y^2}\, dA.$

Answer. Average with the $x \leftrightarrow y$ mirror: $\frac{1}{2} \iint \frac{x^2+y^2}{x^2+y^2} = \frac{1}{2} \cdot 1 = \frac{1}{2}$. No polar coordinates needed.

18.2 Why It Works

A change of variables $(u, v) \mapsto (x, y)$ sends a tiny $du \times dv$ rectangle to a tiny parallelogram spanned by the vectors $(x_u, y_u)\,du$ and $(x_v, y_v)\,dv$. The area of that parallelogram is the absolute value of the determinant $x_u y_v - x_v y_u$ times $du\,dv$. Summing over all tiny rectangles — which is what an integral does — the determinant appears as the scale factor between the two coordinate systems. In polar coordinates the parallelogram is (nearly) an $r\,d\theta \times dr$ rectangle, which is why $dA = r\,dr\,d\theta$.

19 Proof Techniques for the Putnam

Direct proof, contradiction, and induction — as the graders want to see them

19.1 The Three Forms, at Putnam Level

Direct proofs that name their theorems

A direct proof at this level is a chain of cited results: "by the mean value theorem there is $c \in (a, b)$ with", "since the series converges absolutely we may rearrange". Each citation must come with its hypotheses checked.

The Putnam grades exposition. A correct idea written as "clearly the maximum is at $x = 1$" earns 1 or 2 out of 10; the same idea with the derivative computed, the sign checked on each side, and the endpoints compared earns 10. Write the hypothesis of every theorem you use, in the sentence where you use it. The graders are looking for exactly that sentence.

Contradiction with a minimal counterexample

Assume the claim fails; take the SMALLEST failure (smallest n, shortest sequence, least degree). Show a smaller failure exists. This proves claims about all positive integers without an explicit induction.

Minimal-counterexample proofs are induction in a different costume, and they are the natural form when the "smaller case" is not $n - 1$ but something you construct — a polynomial of lower degree, a subset with fewer elements, a fraction with a smaller denominator. State explicitly which well-ordered quantity you are minimizing.

Induction that is actually strong

When $P(n + 1)$ depends on several earlier cases, assume $P(k)$ for ALL $k \le n$. Check as many base cases as the step reaches back. For two-variable claims, induct on one variable with the other fixed, or on the sum.

A frequent Putnam pattern: prove a recurrence-defined sequence has a closed form (strong induction on n), or that a polynomial identity holds for all degrees (induction on degree, using the case of one lower degree after differentiating or dividing by $x - r$). Write the inductive hypothesis with an explicit letter — "assume the statement for all $k \le n$" — and point to the exact place it is used.

Sentences graders reward

"We proceed by strong induction on n." "Suppose for contradiction that no such x exists; let n_0 be the least" "By the intermediate value theorem applied to the continuous function g on $[0,1]$" "The case $n = 1$ holds since"

Try it: minimal counterexample

Prove that every integer $n \geq 2$ has a prime factor.

Answer. Suppose not, and let $n_0 \geq 2$ be the least integer with no prime factor. Then n_0 is not prime, so $n_0 = ab$ with $1 < a < n_0$. By minimality a has a prime factor p, and $p \mid a \mid n_0$ — contradiction.

20 Invariance and Symmetry

Quantities that survive every move, and symmetries that solve for you

20.1 Finding the Invariant

Parity, sums mod k, and colorings

For a process (a game, an operation repeated on a board or a set), compute a quantity after each move: the parity of a count, a sum modulo k, a weighted sum, a determinant, a product of signs. If it never changes, states with different values are mutually unreachable.

The Putnam disguises invariants inside algebra: "the operation replaces (a, b) by $(a + b, ab)$ show some value is never reached" is an invariant problem, and the invariant is often a symmetric function of the pair (here $a + b - ab$ is not preserved, but $(a - 1)(b - 1)$ maps to $(a + b - 1)(ab - 1)$ — try several candidates). The systematic search: write the move algebraically and test the elementary symmetric functions, discriminants and determinants.

Monovariants

A quantity that always DEcreases (or increases) and is bounded proves that a process terminates, and — if it is an integer — how many steps at most. Potential functions, total energy, number of inversions.

When the question is "show the process stops" rather than "show the state is unreachable", look for a monovariant. The classic: n people around a table pass coins; the sum of squares of holdings strictly decreases with each legal move, so the process ends. Strict decrease of a non-negative integer is a proof of termination in one sentence.

Symmetry that solves the problem

If the problem is invariant under a transformation (swapping variables, $x \to 1 - x$, rotation

of a configuration), the answer or the extremal configuration is often invariant too — or the transformation pairs up terms that cancel.

Symmetry arguments give integrals for free (King's rule), pair off terms in sums (roots of unity, $k \leftrightarrow n-k$), and settle games (the second player mirrors). The failure mode is assuming the extremum is at the symmetric point without proof: symmetry says the symmetric point is a critical point, not that it is the maximum. Check.

Where it goes wrong

An "invariant" that is preserved by most moves but not all — verify it on EVERY kind of move. And an invariant shows unreachability only; to show a state IS reachable you must construct the path.

Try it: an invariant

On a board are the numbers $1, 2, \ldots, 2n$. A move erases two numbers a, b and writes $|a-b|$. Show the last number has the same parity as n.

Answer. The parity of the sum of the numbers on the board is invariant: $a+b \equiv |a-b| \pmod 2$. The initial sum is $1+2+\cdots+2n = n(2n+1) \equiv n \pmod 2$, so the last number is congruent to n mod 2. (Check $n=1$: the numbers $1, 2$ leave 1, odd.)

21 Extremal Principle and Working Backwards

Choose the largest, the smallest, the closest — and reason from the answer toward the start

21.1 The Extremal Principle

Pick the extreme object

Among finitely many objects, choose the one maximizing or minimizing something — the largest number, the point farthest from a line, the longest chain, the polynomial of least degree. Its extremality forces a property, or a "more extreme" object contradicts the choice.

The extremal principle is existence-by-optimization. It proves that some pair of points is at minimum distance, that some vertex has minimum degree, that some solution has minimal x — and then the argument works with that special object. On the Putnam it is often the whole proof: "let M be the maximum of the a_i; then the equation at index i gives $M \le \ldots$", and the inequality closes.

Vieta jumping

For an equation quadratic in each variable, take a solution with minimal $x+y$; Vieta gives a second root, producing a smaller solution unless the minimal one is a fixed base case. Classic: if $\frac{a^2+b^2}{ab+1}$ is an integer, it is a perfect square.

Vieta jumping is the extremal principle applied to a family of solutions, with the quadratic formula supplying the "smaller solution". Write the equation as a quadratic in a, note the other root $a\prime = kb - a = \frac{b^2-k}{a}$ is an integer, show $0 \le a\prime < a$, and descend until $a\prime = 0$, which forces $k = b^2$.

Working backwards

Start from what you want and ask what would suffice. "$f(n) = n^2$ would follow if f is multiplicative and $f(p) = p^2$" Reduce the goal to a chain of sufficient conditions until one is something you can prove directly; then write it forwards.

Working backwards finds the proof; it is not the proof. The written solution must run from hypotheses to conclusion. The habit that makes it safe: at each backwards step, note whether the implication is reversible ("it suffices to show" is fine; "this is equivalent to" needs both directions).

Where it goes wrong

Choosing an extremum that may not exist (a maximum over an infinite set) — say why it exists (finiteness, compactness, well-ordering). And in Vieta jumping, forgetting to prove the new root is non-negative and an integer.

Try it: extremal

Positive reals $a_1, \dots, a_n$ satisfy $a_i = \frac{a_{i-1}+a_{i+1}}{2}$ for all i (indices mod n). Show all a_i are equal.

Answer. Let $M = \max a_i = a_j$. Then $a_{j-1} + a_{j+1} = 2M$ with both terms $\le M$, so both equal M. Repeat around the cycle: every $a_i = M$.

22 Strategy & Writing a 10

Choosing problems, managing time, and writing proofs graders cannot dent

22.1 How to Take the Exam

Triage in the first ten minutes

Read all six problems. Rank them. Commit to the two you find most approachable; touch the rest only after those are WRITTEN UP.

A-1/B-1 and A-2/B-2 are designed to be solvable with the material in this course; the median competitor still scores 0 on them by starting problem 3 first or by leaving a correct idea unwritten. Forty minutes of clean work on an accessible problem is worth more than three hours of near-misses. If a problem asks for a value, find it numerically or by small cases FIRST, then prove it — knowing the answer transforms the search.

What a 10 looks like

A complete argument with every nontrivial step justified, edge cases handled, and the conclusion stated. A correct answer with a hand-wave earns 1.

Name the theorems you use ("by the mean value theorem on $[1, 1+x]$"). State hypotheses when you invoke them (continuity for IVT, absolute convergence for rearranging, a dominating function for swapping limit and integral). Handle the equality case of every inequality and the base case of every induction. Write in complete sentences; a solution that reads like a proof is graded like one.

Small cases, then generalize

Compute $n = 1, 2, 3$. Guess the pattern. Then prove it — by induction, by a bijection, by an invariant, or by an identity you now know to look for.

Most Putnam problems hide a clean answer ($\frac{\pi}{4}$, $n+1$, $\frac{1}{e}$, $\binom{2n}{n}$); small cases reveal it, and knowing the target usually reveals the method. Symmetry, extremal choices, and generating functions are the three most common "second ideas" after the pattern is found.

The last five minutes

Re-read each written solution as a hostile grader: Is the base case there? Did you use a hypothesis you never stated? Did you divide by something that could be zero? Is the final answer boxed and consistent with your small cases? Those five minutes routinely turn a 7 into a 10.

23 Formulas & Methods

The Putnam reference sheet: analysis, algebra, and the named integrals in one place

23.1 Analysis

Limits, series, integrals

Limits as derivatives ($n(a^{1/n} - 1) \to \ln a$), as Riemann sums ($\frac{1}{n}\sum f(\frac{k}{n}) \to \int_0^1 f$), as exponentials ($(1 + \frac{a}{n})^n \to e^a$); Stolz–Cesro; Taylor with Lagrange remainder; the three series ($\sum x^n$, $\sum \frac{x^n}{n}$, $\sum \frac{x^n}{n!}$) and their derivatives; $\zeta(2) = \frac{\pi^2}{6}$, $\eta(2) = \frac{\pi^2}{12}$; telescoping with lag; Cauchy condensation; Abel summation $\sum a_n b_n = A_N b_N - \sum A_n(b_{n+1} - b_n)$.

Rewrite limits as one of the three shapes; expand in Taylor series to second order; swap sums and integrals only after stating the justification (absolute/uniform convergence, monotone or dominated convergence).

The named integrals

Feynman: $I'(a) = \int \partial_a f$; Dirichlet $\int_0^\infty \frac{\sin x}{x} = \frac{\pi}{2}$ and $\int_0^\infty \frac{\sin^2 x}{x^2} = \frac{\pi}{2}$; Gaussian $\int e^{-x^2} = \sqrt{\pi}$

and $\int_0^\infty e^{-x^2}\cos 2bx = \frac{\sqrt{\pi}}{2}e^{-b^2}$; King $\int_a^b f = \int_a^b f(a+b-x)$; Queen/Jack folding; Glasser $\int f(x-\frac{1}{x}) = \int f$; the $1/x$ reflection; Frullani $\int_0^\infty \frac{f(ax)-f(bx)}{x} = (f(0)-f(\infty))\ln\frac{b}{a}$; Gamma $\Gamma(n+1) = n!$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$; Beta $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$; $\int_0^\infty \frac{dx}{1+x^n} = \frac{\pi/n}{\sin(\pi/n)}$; log-sine $\int_0^{\pi/2} \ln\sin x = -\frac{\pi}{2}\ln 2$; $\int_0^1 \frac{\ln(1+x)}{1+x^2} = \frac{\pi}{8}\ln 2$; $\int_0^\infty \frac{\ln(1+x^2)}{1+x^2} = \pi \ln 2$; Fresnel $\int_0^\infty \sin x^2 = \frac{1}{2}\sqrt{\frac{\pi}{2}}$; Wallis $\prod \frac{4n^2}{4n^2-1} = \frac{\pi}{2}$; Weierstrass $t = \tan\frac{x}{2}$; series expansion and swap.

Order of attack: symmetry ($x \to a+b-x$, $x \to \frac{1}{x}$) parameter (Feynman) named shape (Frullani, Beta, Gamma, Dirichlet) substitution (tangent, Weierstrass) series. Never start with parts.

Inequalities and convexity

AM–GM, Cauchy–Schwarz (sums and integrals), Jensen, rearrangement, power means; Hermite–Hadamard $f(\frac{a+b}{2}) \le \frac{1}{b-a}\int_a^b f \le \frac{f(a)+f(b)}{2}$ for convex f; tangent-line trick; MVT-derived bounds ($e^x \ge 1+x$, $\ln x \le x-1$, $\sin x \le x$).

Identify the equality case first; convert integral inequalities to their sum versions mentally, then apply the same proof.

23.2 Algebra

Linear algebra

Matrix determinant lemma $\det(I+\mathbf{u}\mathbf{v}^\mathsf{T}) = 1+\mathbf{v}^\mathsf{T}\mathbf{u}$ and $\det(aI+bJ) = a^{n-1}(a+nb)$; Vandermonde $\prod_{i<j}(x_j-x_i)$; circulant eigenvalues by DFT; tridiagonal Toeplitz eigenvalues $a+2b\cos\frac{k\pi}{n+1}$; $\det p(A) = \prod p(\lambda_i)$, $\mathrm{tr}A^k = \sum \lambda_i^k$; Cayley–Hamilton; rank–nullity, $\mathrm{rank}(AB) \le \min$; polynomial-entry matrices have rank $\le \deg+1$; $|GL_n(\mathbb{F}_q)| = \prod(q^n-q^k)$; spectral theorem and Rayleigh quotient; $\mathrm{tr}A^k = 0$ for all $k \le n$ implies nilpotent.

Row-reduce two rows before any expansion. Look for identity-plus-low-rank. Move to the spectrum for anything involving powers or polynomials of a matrix.

Polynomials, number theory, abstract algebra

Vieta and Newton; roots of unity filter; auxiliary polynomial with known roots for values at integers; Chebyshev; Eisenstein; $a-b \mid P(a)-P(b)$. Orders and CRT; LTE; Legendre/Kummer; QR supplements and reciprocity; Wilson; multiplicative functions and $\sum_{d\mid n}\varphi(d) = n$; bounding between squares. Lagrange's theorem; cyclic-group counts ($\varphi(d)$ elements of order d, one subgroup per divisor); orders in S_n as lcm of cycle lengths; a polynomial over a field has at most deg roots; finite closed subsets are subgroups.

For values of a polynomial at consecutive integers, build the auxiliary polynomial. For group questions, use only closure, Lagrange, and counting — nothing heavier is expected.

23.3 Combinatorics, Probability & Geometry

Combinatorics and probability

Bijections; PIE; derangements; Catalan and the reflection principle; generating functions and $x\frac{d}{dx}$; Erds–Szekeres; Mantel/Turn, Hall, Euler; extremal and pigeonhole; games (P/N, Nim, strategy stealing). Linearity with indicators; tail-integral $E[X] = \int P(X > t)dt$; order statistics $E[\max] = \frac{n}{n+1}$; coupon collector nH_n; gambler's ruin $\frac{k}{N}$, duration $k(N-k)$; random walk returns $\binom{2n}{n}/4^n$; first-step analysis.

Write indicator sums before distributions. For "expected time," write first-step equations. For random geometry, draw the region and integrate.

Geometry

Pick's theorem; gcd $+1$ lattice points on a segment; Minkowski; Gauss circle count $\pi r^2 + O(r)$; complex numbers (rotation, collinearity via real ratios, equilateral criterion $a + \omega b + \omega^2 c = 0$); point-to-line/plane distances; convexity (Helly, Carathodory); continuous-motion IVT arguments.

Count lattice points by rows; prove existence by sweeping a parameter and invoking the IVT; reach for complex numbers when rotations appear.

The universal problem-solving loop

Read twice. Restate the goal in your own words. List the givens. Pick a representation. Try small cases. Look for symmetry, an invariant, or an extremal object. Compute, then CHECK against a second method or a sanity bound.

At AMC 10/12 and AIME level, the second method is what separates a 4-second guess from a 12-second certainty: compute a probability two ways (complement and direct), a length via two theorems, a count via a recurrence and a formula. Casework must be organized by a stated criterion so nothing is double-counted; algebra must be checked by substituting back.