GRE Mathematics Subject Test: A Concept-First Preparation Guide

Physics First Series

Copyright © 2026 Physics First Series. All rights reserved.

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in reviews and certain other noncommercial uses permitted by copyright law.

Published by Physics First Series

quantumtutoring.ai

ISBN 979-8-9975551-4-6 (print)
ISBN 979-8-9975551-8-4 (ebook: EPUB & PDF)

First edition, 2026.

10 9 8 7 6 5 4 3 2 1

Contents

Introduction: How to Use This Book **iii**

1 Precalculus: The Language of Everything That Follows **1**

1.1 Functions, Composition, and Inverses 3

1.2 Polynomial and Rational Functions 10

1.3 Exponential and Logarithmic Functions 16

1.4 Trigonometry from the Unit Circle 20

1.5 Analytic Geometry: Lines and Conics 25

1.6 Sequences and Series 29

1.7 Complex Numbers and Euler's Formula 33

Synthesis — Key Results 37

Practice Problems 38

Further Reading 41

2 Calculus I: The Derivative and Its Uses **43**

2.1 Limits and Continuity 45

2.2 The Derivative from Its Definition 49

2.3 The Differentiation Rules 53

2.4 The Chain Rule and Implicit Differentiation 57

2.5 Linear Approximation and Differentials 62

2.6 L'Hôpital's Rule 65

2.7 The Mean Value Theorem 68

2.8 Related Rates and Optimization 71

2.9 Curve Sketching: What f' and f'' Reveal 75

2.10 The Integral and the Fundamental Theorem 78

Synthesis — Key Results (Ch.2) 85

Practice Problems (Ch.2) 86

Further Reading (Ch.2) 89

3 Calculus II: Integration Techniques, Applications, and Infinite Series 91

3.1 Substitution: Running the Chain Rule Backward 93

3.2 Integration by Parts 96

3.3 Trigonometric Integrals and Trigonometric Substitution 99

3.4 Partial Fractions 104

3.5 Improper Integrals 107

3.6 Applications: Areas, Volumes, Arc Length, and More 110

3.7 Sequences 118

3.8 Infinite Series and Convergence Tests 121

3.9 Power Series, Taylor, and Maclaurin 128

3.10 Parametric and Polar Calculus 133

Synthesis — Key Results (Ch.3) 136

Practice Problems (Ch.3) 136

Further Reading (Ch.3) 140

4 Differential Equations 141

4.1 What a Differential Equation Is 142

4.2 First-Order Separable Equations 146

4.3 First-Order Linear Equations: The Integrating Factor 149

4.4 Exact Equations and Substitutions 153

4.5 Modeling with First-Order Equations 157

4.6 Autonomous Equations and Equilibria 161

4.7 Second-Order Linear Equations: The Characteristic Equation 164

4.8 Nonhomogeneous Equations: Finding a Particular Solution 172

4.9 Reading an Equation as a Model 177

Synthesis — Key Results (Ch.4) 180

Practice Problems (Ch.4) 180

Further Reading (Ch.4) 184

5 Linear Algebra **185**

5.1 Vectors, Matrices, and the Product 186

5.2 Systems of Linear Equations 190

5.3 Determinants and Inverses 194

5.4 Vector Spaces, Independence, and Basis 201

5.5 Rank, Nullity, and the Four Subspaces 204

5.6 Linear Transformations and Their Matrices 208

5.7 Eigenvalues, Eigenvectors, and Diagonalization 211

5.8 Orthogonality and Projection 217

Synthesis — Key Results (Ch.5) 222

Practice Problems (Ch.5) 223

Further Reading (Ch.5) 227

6 Number Theory and Abstract Algebra **229**

6.1 Divisibility, Primes, and Unique Factorization 230

6.2 The Euclidean Algorithm and Bézout's Identity 234

6.3 Modular Arithmetic and Congruences 238

6.4 Linear Congruences and the Chinese Remainder Theorem 242

6.5 Fermat's Little Theorem and Euler's Totient 246

6.6 Groups: The Structure Behind Arithmetic 249

6.7 Subgroups, Cyclic Groups, and Generators 252

6.8 Lagrange's Theorem and Cosets 256

6.9 Homomorphisms and Isomorphisms 259

6.10 Rings and Fields 263

Synthesis — Key Results (Ch.6) 267

Practice Problems (Ch.6) 268

Further Reading (Ch.6) 271

7 Additional Topics **273**

7.1 Real Analysis 274

7.2 Point-Set Topology 280

7.3 Complex Analysis 284

7.4 Combinatorics and Discrete Probability 290

7.5 Numerical Analysis 297

7.6 Logic and Set Theory 301

7.7 Algorithms and Complexity 306

Synthesis — Key Results (Ch.7) 308

Practice Problems (Ch.7) 309

Further Reading (Ch.7) 312

Introduction: How to Use This Book

What this book is

This is a concept-first preparation guide for the GRE Mathematics Subject Test—the timed, multiple-choice examination that graduate programs in the mathematical sciences use to gauge undergraduate command of the core curriculum. The test rewards two things at once: a wide net of topics (calculus dominates, but precalculus, linear algebra, abstract algebra, real analysis, probability, and a scattering of "additional topics" all appear), and the ability to move quickly. You have on the order of two and a half minutes per question. Deep but slow will not clear the bar; neither will fast but shallow. The aim of this book is to make the ideas so structurally clear that speed becomes a by-product of understanding rather than a substitute for it.

We are unapologetically *concept-first*. A student who has memorized a hundred integrals but cannot see why substitution is the chain rule read backwards will stall the first time a problem is dressed up unfamiliarly. A student who owns the idea will re-derive what they forgot in seconds. Every chapter is therefore built to install the organizing idea first and let the techniques hang off it.

The Physics First Series method

If you have used another book in the Physics First Series, the furniture will be familiar. Each chapter is engineered around a small number of recurring devices, and it pays to know what each one is *for*.

- **Concept map.** Every chapter opens with a one-page map showing how the chapter's ideas connect. Read it first and return to it last. On the exam you recall *structure* faster than you recall lists.
- **Chapter Road Map and Objectives.** A plain statement of where the chapter goes and exactly what you will be able to *do* when you finish. Treat the objectives as a checklist.

- **Priming Questions.** Before most sections you will meet a boxed question with the instruction to answer *before* reading on. This is not decoration. Retrieval practice—committing to an answer before seeing ours—is one of the best-evidenced ways to make learning stick. Do not skip it.
- **Worked Examples.** Fully worked, in the same voice you should use in your own scratch work: name the idea, plan, compute, then check. The check step is where most exam points are saved.
- **Insight, Misconception, and Pitfall boxes.** The Insight box names the idea worth carrying away. The Misconception and Pitfall boxes name the specific wrong turn students actually take—often the exact wrong answer the test writers planted as a distractor.
- **Dialogues.** A short Socratic exchange that models the reasoning process as a conversation, because that is how understanding usually arrives.
- **Practice Problems.** Thirty per chapter, graded in difficulty. A one-star problem checks a single idea; a three-star problem asks you to combine several under time pressure, the way the real exam does.

Problem-solving strategy — How to read a chapter

First pass: read the concept map, the road map, and the objectives; skim the section headings. Second pass: work through the sections, *always* answering the priming questions and reworking each worked example with the book closed—narrating the Read–Plan–Solve–Check steps aloud. Third pass: the practice problems, timed. If a problem takes more than three minutes, stop, read the solution, and re-solve it a day later. Spacing beats cramming.

Scope and difficulty markers

The Subject Test is roughly one-half single- and multivariable calculus, one-quarter algebra (linear and abstract) and the rest a mix of analysis, differential equations, probability, and topics such as number theory, set theory, and complexity. This book follows that weighting. We begin here, in Chapter 1, with *precalculus*—not because it is heavily tested in its own right, but because every later chapter leans on it. Functions, their inverses, the exponential-logarithm pair, trigonometric identities, and the complex plane are the vocabulary in which calculus, differential equations, and analysis are all written. Time spent making this fluent is repaid with interest.

Throughout, difficulty is marked with stars: $\star$ (single idea, warm-up), $\star\star$ (standard exam level), and $\star\star\star$ (multi-step or time-pressured). A realistic goal for a first pass is to clear the one- and two-star problems cleanly and to understand—if not yet beat the clock on—the three-star ones.

Misconception

"I'll just review formulas." The Subject Test almost never asks you to recite a formula. It asks you to *recognize which idea applies* to an unfamiliar-looking object and then execute quickly. Formula lists are necessary and radically insufficient. This book drills recognition.

A note on the accompanying materials: each chapter has a companion lecture deck and a set of narrated worked sessions built from the same source, so you can study by reading, by watching, or—best—by alternating the two.

Explore the Physics First Foundation and start the adaptive tutor:

Physics First Foundation
physicsfirstfoundation.org

Adaptive AI Tutor
quantumtutoring.ai

CHAPTER 1

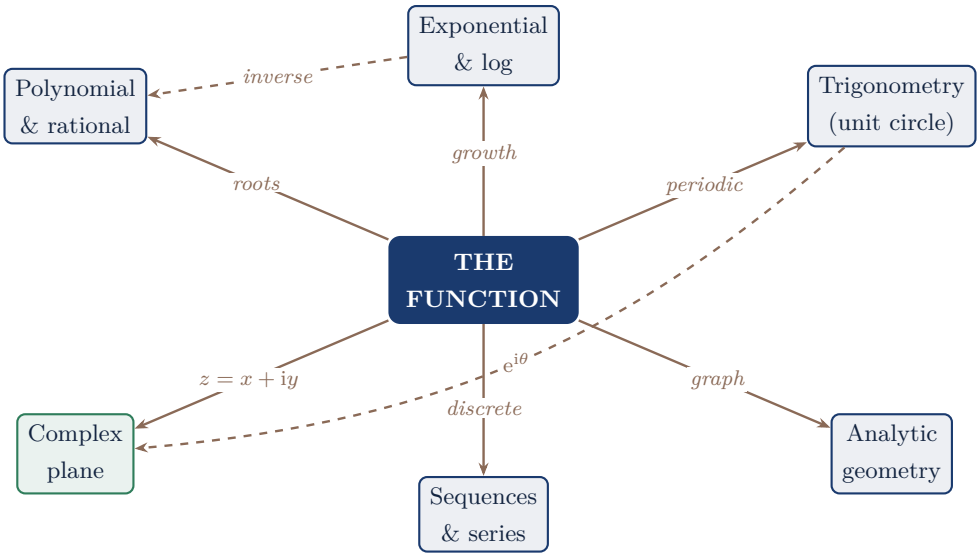
Precalculus

The Language of Everything That Follows

"Mathematics is the art of giving the same name to different things."

— Henri Poincaré

Precalculus is not so much a topic on the Subject Test as the* grammar *in which every other topic is written. This chapter installs the one idea that unifies it all—the* function *—and then six families of functions and the two bridges between them. A function is a rule with a domain; it is invertible exactly when it is one-to-one; polynomials are run by their roots and rational functions by their zeros and poles; the exponential and the logarithm are a single inverse pair that turns multiplication into addition; trigonometry is organized by one picture, the unit circle; conics are read off by completing the square; a geometric series converges exactly when its ratio is smaller than one in size; and the complex plane is just the plane with a multiplication that rotates. Own the recognition moves here and calculus becomes bookkeeping.



Learning Objectives

You will be able to:

1. state a function's domain and range and decide whether it is invertible;
2. build and unbuild compositions, and compute inverses on the correct domain;
3. locate polynomial roots with the factor and rational-root theorems, and read a rational graph from its zeros, poles, and holes;
4. move fluently between exponential and logarithmic form and use the log laws;
5. evaluate trigonometric functions from the unit circle and apply the core identities;
6. classify a conic and extract its features by completing the square;
7. sum arithmetic and geometric series and decide geometric convergence;
8. compute in $\mathbb{C}$, convert to polar form, and use Euler's formula and De Moivre.

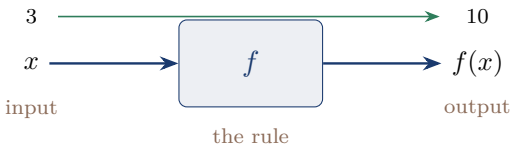
Did You Know? *how the word "function" entered mathematics*

Leibniz coined *functio* in the 1670s for a quantity depending on a curve, but the modern idea—a rule pairing each input with exactly one output, independent of any formula—was not pinned down until Dirichlet in 1837. The Subject Test lives in Dirichlet's world: a function need not have a tidy formula, and the questions that trip students hardest are the ones testing the *definition*, not a computation.

Functions, Composition, and Inverses

A *function* is a rule that assigns to each input *exactly one* output. We write $f : A \to B$, call A the *domain* (the allowed inputs) and B the codomain, and call the set of outputs actually attained the *range*. This "exactly one output" clause is the whole content of the idea, and it is worth saying aloud what it forbids: an input may not lead to two different outputs. Everything else in this section—inverses, compositions, the graphs you will read—follows from taking that clause seriously, especially the part everyone forgets: a function is its rule *together with* its domain. Change the domain and you have changed the function, sometimes turning a rule that had no inverse into one that does.

To evaluate a function is to substitute a value for every occurrence of the input symbol. The symbol is a placeholder: $f(x) = x^2$ and $f(t) = t^2$ are the same function, and this is exactly why $f(x+h)$ makes sense—wherever an x sits in the rule, drop in the entire quantity $x + h$. This placeholder discipline is the seed of the derivative in Chapter 2, where the whole game is comparing $f(x+h)$ to $f(x)$.



A function as a machine: each input enters the rule and leaves as exactly one output. Here $f(x) = x^2 + 1$ sends the input 3 to the output 10.

What & why. What a function is, when it can be undone, and how compositions chain rules together—the vocabulary the rest of the book assumes.

Solve. Read the rule *and* its domain; check one-to-one before claiming an inverse; track the domain through every composition.

Means. Each input gives exactly one output.

Approach Read · Plan · Solve · Check

Read. Decide whether the task asks you to evaluate, invert, or compose, and record each function *together with* its domain.

Plan.

1. To evaluate, substitute the entire input expression for every occurrence of the variable.
2. To invert, confirm the function is one-to-one, then set $y = f(x)$, solve for x, and rename.

3. To compose, feed the inner function into the outer and impose *both* domain conditions.

Solve. Carry out each substitution or algebra step one line at a time, keeping the domain in view.

Check. Test a numerical point; for an inverse confirm $f(f^{-1}(x)) = x$, and for a composition hunt for the hidden domain exclusion.

Common traps — every wrong answer is one of these

Distributes f across a sum.
Writes $f(a+b) = f(a) + f(b)$. A function is a rule applied to the *whole* input, not a quantity that distributes.

Confuses inverse with reciprocal.
Reads $f^{-1}(x)$ as $1/f(x)$. The inverse undoes f; the reciprocal is a different thing entirely.

Claims an inverse without one-to-one.
Inverts a function that fails the horizontal line test. Restrict the domain until each output comes from a single input first.

Drops the inner domain in a composition.
1. Keeps only the outer function's restrictions.
2. Forgets the inner function's own domain.
3. Fix: an x is allowed only if it lies in the domain of g *and* $g(x)$ lies in the domain of f.

Reverses the composition order.
Treats $f \circ g$ as $g \circ f$. The function written first is applied *last*; order changes the result.

Evaluating at a compound input. For $f(x) = x^2 - 3x$, find $f(a+1)$.

Read.

Substitute the entire quantity $a+1$ for every x; do not distribute f across the sum.

Plan.

Substitute $a+1$ for every x, then expand:

$$\begin{aligned} f(a+1) &= (a+1)^2 - 3(a+1) && \text{substitute the whole input} \\ &= a^2 + 2a + 1 - 3a - 3 && \text{expand } (a+1)^2 \text{ and } -3(a+1) \end{aligned}$$

Solve.

$f(a+1) = a^2 - a - 2$.

Check.

Sanity check at $a = 1$: directly $f(2) = 4 - 6 = -2$, and the formula gives $1 - 1 - 2 = -2$. Agree.

The difference quotient

The combination

$$\frac{f(x+h) - f(x)}{h} \tag{1.1}$$

measures the average rate of change of f over a step of width h. It is the single most important expression in the whole book, because its limit as $h \to 0$ is the derivative. Learning to simplify it now—so that the h in the denominator cancels—removes the only obstacle to Chapter 2.

Simplifying a difference quotient. For $f(x) = x^2$, simplify $\frac{f(x+h)-f(x)}{h}$.

Read.

Expand $f(x+h)$, subtract $f(x)$, then cancel the common factor h.

Plan.

Form the numerator, then divide by h:

$$\begin{aligned}
f(x+h) - f(x) &= (x+h)^2 - x^2 && \text{write both terms} \\
&= x^2 + 2xh + h^2 - x^2 && \text{expand the square} \\
&= 2xh + h^2 && \text{the } x^2 \text{ terms cancel} \\
\frac{2xh + h^2}{h} &= \frac{h(2x+h)}{h} && \text{factor } h \text{ from the numerator} \\
&= 2x + h && \text{cancel } h,\ h \neq 0
\end{aligned}$$

Solve.

$\frac{f(x+h)-f(x)}{h} = 2x + h$.

Check.

As $h \to 0$ this tends to $2x$, which we will later call $f'(x)$; the cancellation of h is what makes that limit finite. ✓

Even, odd, and reading symmetry

A function is *even* if $f(-x) = f(x)$ (its graph is mirror-symmetric across the y-axis) and *odd* if $f(-x) = -f(x)$ (symmetric through the origin). The test is purely mechanical: replace x by $-x$ and compare. Symmetry is worth spotting because it halves the work of graphing and shows up in integrals later.

Classifying symmetry. Determine whether $f(x) = x^3 - 4x$ is even, odd, or neither.

Read.

Compute $f(-x)$ and compare it to $f(x)$ and to $-f(x)$.

Plan.

Substitute $-x$ and compare with $f(x)$:

$$\begin{aligned} f(-x) &= (-x)^3 - 4(-x) && \text{replace } x \text{ by } -x \\ &= -x^3 + 4x && \text{simplify} \\ &= -(x^3 - 4x) = -f(x) && \text{factor out } -1 \end{aligned}$$

Solve.

$f(-x) = -f(x)$, so f is *odd*.

Check.

Odd matches the graph: a cubic through the origin with point symmetry. Every purely odd-power polynomial (with no constant term) is odd. ✓

Average rate of change

The *average rate of change* of f between $x = a$ and $x = b$ is the slope of the secant line joining $(a, f(a))$ and $(b, f(b))$, namely $\frac{f(b)-f(a)}{b-a}$—the difference quotient (1.1) with $a = x$ and $b = x + h$. It is the discrete cousin of the derivative: shrink the interval to a single point and the secant becomes the tangent. Reading a rate off two points is a recurring exam task, especially when f is given by a table or a word problem.

Average rate of change. Find the average rate of change of $f(x) = x^2 + 1$ on the interval $[1, 4]$.

Read.

Evaluate f at both endpoints and divide the change in output by the change in input.

Plan.

Evaluate at the endpoints, then form the slope:

$$f(1) = 1^2 + 1 = 2, \qquad f(4) = 4^2 + 1 = 17 \quad \text{the two output values}$$

$$\frac{f(4) - f(1)}{4 - 1} = \frac{17 - 2}{3} \quad \text{change in output over change in input}$$

$$= \frac{15}{3} \quad \text{simplify}$$

Solve.

The average rate of change is 5.

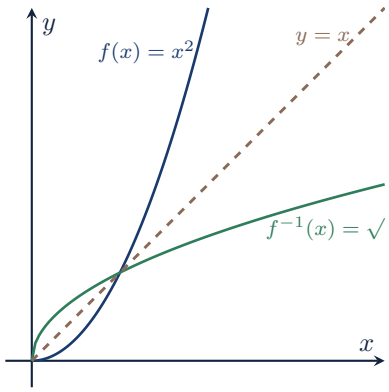
Check.

This is the slope of the secant from $(1, 2)$ to $(4, 17)$; the instantaneous rate $f'(x) = 2x$ of Chapter 2 averages to 5 over this interval, matching its midpoint value $2(2.5) = 5$. ✓

Injective, surjective, and the inverse

Definition 1.1.1 (One-to-one, onto, bijective)**.** f is *injective* (one-to-one) if different inputs give different outputs: $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. It is *surjective* (onto) if its range is all of B. It is *bijective* if both. A function has a genuine inverse f^{-1} *exactly when* it is bijective.

Graphically, injectivity is the *horizontal line* test: no horizontal line meets the graph more than once. When the inverse exists, its graph is the reflection of the graph of f across the line $y = x$, because "$y = f(x)$" and "$x = f^{-1}(y)$" say the same thing. To find an inverse algebraically, write $y = f(x)$, solve for x in terms of y, and then rename—this three-step habit never varies.



Inverse as reflection. On $[0,\infty)$, $f(x)=x^2$ and $f^{-1}(x)=\sqrt{x}$ are mirror images across the dashed line $y=x$: swapping the roles of input and output reflects the graph.

Finding an inverse. Find f^{-1} for $f(x)=\frac{2x-1}{x+3}$.

Read.

Set $y=f(x)$, then solve for x in terms of y; the result, renamed, is f^{-1}.

Plan.

Set $y=f(x)$ and solve for x:

$$
\begin{aligned}
y &= \frac{2x-1}{x+3} && \text{start from } y=f(x)\\
y(x+3) &= 2x-1 && \text{clear the denominator}\\
yx+3y &= 2x-1 && \text{distribute}\\
yx-2x &= -1-3y && \text{collect the } x\text{-terms}\\
x(y-2) &= -1-3y && \text{factor out } x\\
x &= \frac{-1-3y}{y-2}=\frac{3y+1}{2-y} && \text{divide, then tidy signs}
\end{aligned}
$$

Solve.

$f^{-1}(x)=\frac{3x+1}{2-x}$.

Check.

Check a point: $f(1)=\frac{1}{4}$, and $f^{-1}\left(\frac{1}{4}\right)=\frac{3/4+1}{2-1/4}=\frac{7/4}{7/4}=1$. It inverts.

A domain restriction rescues an inverse. Why does $g(x)=x^2$ have no inverse on $\mathbb{R}$, but does on $[0,\infty)$?

Read.

Test one-to-one via the horizontal line test, then restrict the domain to fix it.

Plan.

On $\mathbb{R}$, both $x=2$ and $x=-2$ give $g=4$, so g is not one-to-one—the horizontal line $y=4$ meets the parabola twice. Restrict the domain to $x\geq 0$: now each output comes from one input, and solving $y=x^2$ with $x\geq 0$ gives $x=\sqrt{y}$.

Solve.

On $[0,\infty)$, $g^{-1}(y)=\sqrt{y}$; on all of $\mathbb{R}$ there is no inverse.

Check.

This is why $\sqrt{x^2} = |x|$, not x: the square root returns the *non-negative* root. Forgetting this is the classic error.

Composition and its domain

The *composition* $(f \circ g)(x) = f(g(x))$ feeds the output of g into f. Its domain is the set of x in the domain of g whose output $g(x)$ also lies in the domain of f —two conditions, and the exam almost always hides its trap in the second one. Order matters: in general $f \circ g \neq g \circ f$.

Composition with a hidden exclusion. For $f(x) = \frac{1}{x-2}$ and $g(x) = \sqrt{x}$, find $f \circ g$ and its domain.

Read.

Substitute g into f, then impose *both* domain conditions.

Plan.

Substitute g into f, then impose both domain conditions:

$$\begin{aligned} (f \circ g)(x) &= f(\sqrt{x}) && \text{feed } g \text{ into } f \\ &= \frac{1}{\sqrt{x} - 2} && \text{apply the rule of } f \end{aligned}$$

Condition one, from g: $x \geq 0$. Condition two, from f: the denominator $\sqrt{x} - 2 \neq 0$, i.e. $\sqrt{x} \neq 2$, i.e. $x \neq 4$.

Solve.

$(f \circ g)(x) = \frac{1}{\sqrt{x}-2}$ with domain $[0, 4) \cup (4, \infty)$.

Check.

The excluded point $x = 4$ is invisible in the formula—you find it only by tracking the inner function into the outer one.

Order matters in composition. For $f(x) = x^2$ and $g(x) = x + 3$, compute $(f \circ g)(x)$ and $(g \circ f)(x)$ and confirm they differ.

Read.

Substitute the inner function into the outer, once each way, then compare.

Plan.

Substitute each way and compare:

$$\begin{aligned} (f \circ g)(x) &= f(x+3) = (x+3)^2 && \text{feed } g \text{ into } f \\ &= x^2 + 6x + 9 && \text{expand} \\ (g \circ f)(x) &= g(x^2) = x^2 + 3 && \text{feed } f \text{ into } g \end{aligned}$$

These are not the same polynomial.

Solve.

$(f \circ g)(x) = x^2 + 6x + 9$ but $(g \circ f)(x) = x^2 + 3$.

Check.

Composition is not commutative: which function acts last changes the result. Test at $x = 1$: $f \circ g$ gives 16, $g \circ f$ gives 4. ✓

Problem-solving strategy — Finding a domain from the formula

Scan the rule for the three things that can go wrong and exclude the inputs that trigger them: (1) a *denominator* equal to zero; (2) an *even root* of a negative quantity; (3) a *logarithm* of a non-positive quantity. Everything else is allowed. For a composition, apply the scan to the inner function first, then to the outer.

Misconception

f^{-1} is the inverse function, not the reciprocal. $f^{-1}(x)$ means the function that undoes f; it is *not* $1/f(x)$. The notation collides with exponents (compare $\sin^{-1} x$ with $\sin^2 x$), so read it from context and never assume $f^{-1}(x) = (f(x))^{-1}$.

Polynomial and Rational Functions

A polynomial of degree n, $p(x) = a_n x^n + \cdots + a_1 x + a_0$ with $a_n \neq 0$, is completely governed by its *roots*. The Fundamental Theorem of Algebra guarantees exactly n of them in the complex numbers, counted with multiplicity, and two elementary theorems turn "find the roots" from guesswork into a short, finite search.

Factor and Remainder Theorems. $(x - c)$ divides $p(x)$ exactly when $p(c) = 0$. More generally, the remainder on dividing $p(x)$ by $(x - c)$ equals $p(c)$—so you can test a candidate root by a single evaluation.

Rational Root Theorem. If p has integer coefficients, any rational root written in lowest terms as r/s has r dividing the constant term a_0 and s dividing the leading coefficient a_n. This produces a finite list of candidates to test.

The quadratic case deserves its own closed form, and deriving it once shows where every term comes from—it is nothing but completing the square on the general quadratic.

Derivation. *the quadratic formula by completing the square*

Start from $ax^2 + bx + c = 0$ with $a \neq 0$ and divide by a: $x^2 + \frac{b}{a}x + \frac{c}{a} = 0$. Move the constant and complete the square by adding the square of half the x-coefficient, $\left(\frac{b}{2a}\right)^2$, to both sides:

$$x^2 + \tfrac{b}{a}x + \left(\tfrac{b}{2a}\right)^2 = \left(\tfrac{b}{2a}\right)^2 - \tfrac{c}{a}, \qquad \text{so} \qquad \left(x + \tfrac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Take the square root and isolate x: $x + \frac{b}{2a} = \pm\frac{\sqrt{b^2-4ac}}{2a}$, hence $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$. The quantity $b^2 - 4ac$ under the root is the *discriminant*: positive gives two real roots, zero a repeated root, negative a conjugate pair of complex roots.

The strategy for higher degree is always the same: list the rational candidates, evaluate to find one true root c, divide out $(x - c)$ to drop the degree, and repeat until what remains is a quadratic you can finish by factoring or the formula above.

Approach Read · Plan · Solve · Check

Read. Decide whether you are finding the roots of a polynomial or reading a rational graph, and note the degree(s) and the leading and constant coefficients.

Plan.

1. For roots, list the rational candidates r/s (r divides the constant, s divides the leading coefficient).
2. Evaluate to find one true root c, divide out $(x - c)$ to drop the degree, and repeat until a quadratic remains.
3. For a rational function, factor numerator and denominator to lowest terms, then read intercepts, poles, and end behavior.

Solve. Carry out each division or factoring one step at a time, keeping any complex conjugate pair together.

Check. Multiply the factors back (match the leading term and the constant), and confirm the number of roots equals the degree, counted with multiplicity.

Common traps — every wrong answer is one of these

Calls a cancelled factor an asymptote.
A factor common to numerator and denominator is a *hole*, not a vertical asymptote; reduce to lowest terms before reading poles.

Forgets the conjugate root.
With real coefficients, a non-real root forces its conjugate; a real cubic cannot have exactly one complex root.

Misreads end behavior.
1. Compares only the leading coefficients.
2. Ignores the degrees.
3. Fix: compare degrees first — $m < n$ gives $y = 0$, $m = n$ gives the coefficient ratio, $m = n+1$ gives a slant asymptote.

Truncates the rational-root list.
Tests only integer divisors of the constant and forgets that s must divide the leading coefficient, missing fractional roots such as $\frac{1}{2}$.

Miscounts multiplicity.
Assumes n *distinct* roots. A degree-n polynomial has n roots only when a repeated root is counted according to its multiplicity.

Full factorization by the rational-root search. Factor $p(x) = 2x^3 - 3x^2 - 3x + 2$ completely.

Read.

List candidates r/s with $r \mid 2$, $s \mid 2$; find one root; divide; finish the quadratic.

Plan.

List candidates and test one, then divide:

$$\begin{aligned} \text{candidates: } \pm 1,\ \pm 2,\ \pm \tfrac{1}{2} \qquad & r \mid 2 \text{ and } s \mid 2 \\ p(2) = 2(8) - 3(4) - 3(2) + 2 \qquad & \text{test } x = 2 \\ = 16 - 12 - 6 + 2 = 0 \qquad & \text{so } (x-2) \text{ is a factor} \end{aligned}$$

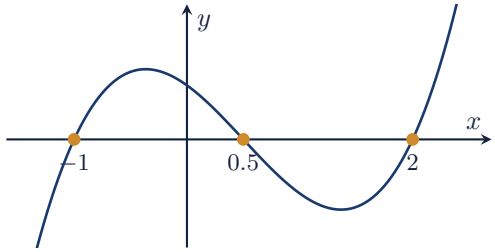
Dividing p by $(x-2)$ leaves $2x^2 + x - 1$, which factors as $(2x-1)(x+1)$.

Solve.

$p(x) = (x-2)(2x-1)(x+1)$, with roots $x = 2,\ \tfrac{1}{2},\ -1$.

Check.

Three real roots for a cubic—consistent with degree 3. Multiply the leading terms: $x \cdot 2x \cdot x = 2x^3$. ✓



The cubic $p(x) = 2x^3 - 3x^2 - 3x + 2$ crosses the axis at its three roots $x = -1, \tfrac{1}{2}, 2$ (orange). Reading the roots off the factored form is far faster than plotting.

Building a polynomial from complex roots. Find the monic cubic with real coefficients whose roots include 1 and $3 + 2\mathrm{i}$.

Read.

Real coefficients force conjugate pairs; multiply the three linear factors, pairing the conjugates first.

Plan.

Real coefficients require the conjugate $3 - 2\mathrm{i}$ to also be a root. Multiply the conjugate pair

first, then the remaining factor:

$$\begin{aligned}
(x-(3+2\mathrm{i}))(x-(3-2\mathrm{i})) &= (x-3)^2-(2\mathrm{i})^2 && \text{difference of squares}\\
&= (x-3)^2+4 && \text{since } (2\mathrm{i})^2=-4\\
&= x^2-6x+13 && \text{expand}\\
(x-1)(x^2-6x+13) &= x^3-6x^2+13x-x^2+6x-13 && \text{multiply by } (x-1)\\
&= x^3-7x^2+19x-13 && \text{collect like terms}
\end{aligned}$$

Solve.

$p(x)=x^3-7x^2+19x-13$.

Check.

Pairing conjugates first keeps the arithmetic real throughout—never multiply the complex factors separately. ✓

For a *rational* function $r(x)=p(x)/q(x)$ written in lowest terms, the graph is read off from where numerator and denominator vanish. Zeros of p give x-intercepts; zeros of q give vertical asymptotes; and the behavior far out is set by comparing degrees. If a factor is common to both—so it cancels—you get a *hole*, not an asymptote, at that x. This distinction is a favorite trap.

End behavior of $r=p/q$, with $\deg p=m$ and $\deg q=n$: if $m<n$, horizontal asymptote $y=0$; if $m=n$, horizontal asymptote $y=(\text{lead } p)/(\text{lead } q)$; if $m=n+1$, a slant asymptote (divide to find it); if $m>n+1$, no linear asymptote.

Reading a rational graph. Describe every intercept and asymptote of $r(x)=\frac{x^2-1}{x^2-4}$.

Read.

Factor top and bottom; zeros of the top are intercepts, zeros of the bottom are poles; compare degrees for end behavior.

Plan.

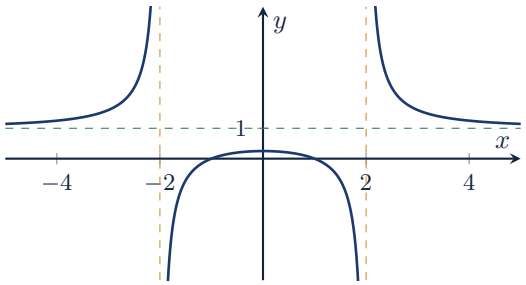
Factor: $r(x)=\frac{(x-1)(x+1)}{(x-2)(x+2)}$. No factor cancels, so there are no holes. Numerator zeros: $x=\pm1$ (x-intercepts). Denominator zeros: $x=\pm2$ (vertical asymptotes). Degrees are equal ($m=n=2$), so the horizontal asymptote is the ratio of leading coefficients, $1/1$.

Solve.

Intercepts at $x=\pm1$; vertical asymptotes at $x=\pm2$; horizontal asymptote $y=1$.

Check.

At $x = 0$, $r(0) = \frac{-1}{-4} = \frac{1}{4}$, a sensible finite y-intercept between the asymptotes. ✓



The graph of $r(x) = \frac{x^2-1}{x^2-4}$: intercepts at $x = \pm 1$, vertical asymptotes (dashed, orange) at $x = \pm 2$, and the horizontal asymptote $y = 1$ (dashed, green).

A slant asymptote. Find the slant asymptote of $r(x) = \frac{2x^2+1}{x-3}$.

Read.

The numerator degree is one more than the denominator, so divide; the polynomial part is the slant asymptote.

Plan.

Divide $2x^2 + 1$ by $x - 3$. Long division: $2x^2 \div x = 2x$; $2x(x - 3) = 2x^2 - 6x$; subtract to get $6x + 1$. Then $6x \div x = 6$; $6(x - 3) = 6x - 18$; subtract to get remainder 19. So $r(x) = 2x + 6 + \frac{19}{x-3}$. As $x \to \pm\infty$ the remainder term vanishes.

Solve.

Slant asymptote $y = 2x + 6$; also a vertical asymptote at $x = 3$.

Check.

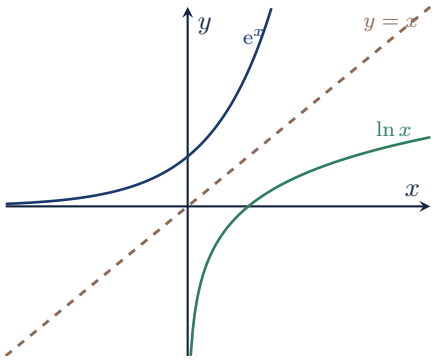
The remainder $19/(x - 3) \to 0$ far out, so the graph hugs the line $y = 2x + 6$—which is what "slant asymptote" means. ✓

Misconception

Cancelled factor = hole, not asymptote. $\frac{x^2-1}{x-1} = x + 1$ for $x \neq 1$: a removable hole at $x = 1$, with no asymptote there. The test writers count on you calling $x = 1$ an asymptote out of habit—always reduce to lowest terms first.

Exponential and Logarithmic Functions

For a base $b > 0$ with $b \neq 1$, the exponential $f(x) = b^x$ is a one-to-one function from $\mathbb{R}$ onto the positive reals $(0, \infty)$. Because it is one-to-one, it has an inverse, and that inverse is precisely the logarithm base b: the statement $b^x = y$ and the statement $x = \log_b y$ carry identical information. The logarithm exists for one reason—to isolate an unknown that is stuck in an exponent—and everything else about it follows from the exponent laws it inherits. Multiplying powers adds exponents; reading that backwards, the log of a product is the sum of the logs. The natural base $\mathrm{e} \approx 2.71828$ is the one whose growth rate equals its own value, and we write $\ln = \log_{\mathrm{e}}$.



Exponential and logarithm as an inverse pair: e^x (navy) and $\ln x$ (green) are reflections across $y = x$. The exponential's horizontal asymptote $y = 0$ becomes the logarithm's vertical asymptote $x = 0$.

Log laws (for positive arguments and a valid base):

$$\log_b(xy) = \log_b x + \log_b y, \qquad \log_b \tfrac{x}{y} = \log_b x - \log_b y,$$
$$\log_b(x^r) = r \log_b x, \qquad \text{change of base} \quad \log_b x = \tfrac{\ln x}{\ln b}. \tag{1.2}$$

Derivation. *the change-of-base formula*

Let $y = \log_b x$, which means $b^y = x$ by definition. Take the natural log of both sides: $\ln(b^y) = \ln x$. By the power law the left side is $y \ln b$, so $y \ln b = \ln x$ and therefore $y = \frac{\ln x}{\ln b}$. Since y was $\log_b x$, we have proved $\log_b x = \frac{\ln x}{\ln b}$—any base can be rewritten in terms of the natural log, which is why a calculator with only ln suffices for every logarithm.

What & why. Exponential growth and its inverse, the logarithm; the log laws are the only tools you need to solve equations where the unknown sits in an exponent.

Solve. To solve $b^{(\cdots)} = k$, take a logarithm of both sides and bring the exponent down.

Means. Log and exponential are one idea seen from two sides.

Approach Read · Plan · Solve · Check

Read. Identify whether the unknown sits in an exponent (a logarithm frees it) or inside a logarithm (rewrite in exponential form), and note the base.

Plan.

1. To free an exponent, take a logarithm of both sides and bring it down with the power law.
2. To solve a log equation, combine to a single logarithm, drop the log, then test every root against the domain.
3. Watch for an equation that is secretly a quadratic in b^x — substitute $u = b^x$.

Solve. Apply one log law per line, keeping every argument positive.

Check. Back-substitute into the *original* equation and discard any root that makes an argument non-positive.

Common traps — every wrong answer is one of these

Splits the log of a sum.
Writes $\log(x + y) = \log x + \log y$. The sum law is for the log of a *product*; there is no identity for the log of a sum.

Keeps extraneous roots.
Combining logs can enlarge the domain and admit false roots; every candidate must keep the original arguments positive.

Misplaces the power-law exponent.
$\log(x^r) = r \log x$, not $(\log x)^r$; the exponent descends as a coefficient.

Rejects a legal substitution value. 1. Substitutes $u = \mathrm{e}^x$.

2. Discards a *positive* root, thinking it illegal.
3. Fix: only $u \le 0$ is impossible, since $\mathrm{e}^x > 0$; every positive u is legal.

Forgets the base restriction.
Allows $b = 1$ or $b \le 0$. A logarithm requires a positive base other than 1.

Solving for an exponent. Solve $3^{\,2x+1} = 7$ exactly, then estimate.

Read.

Take ln of both sides so the exponent comes down as a coefficient; isolate x.

Plan.

Take ln of both sides and isolate x:

$$\begin{aligned} \ln\!\left(3^{2x+1}\right) &= \ln 7 && \text{apply } \ln \\ (2x+1)\ln 3 &= \ln 7 && \text{power law brings the exponent down} \\ 2x+1 &= \frac{\ln 7}{\ln 3} && \text{divide by } \ln 3 \\ x &= \frac{1}{2}\left(\frac{\ln 7}{\ln 3} - 1\right) && \text{solve for } x \end{aligned}$$

Solve.

$x = \frac{1}{2}\left(\frac{\ln 7}{\ln 3} - 1\right) \approx \frac{1}{2}(1.771 - 1) \approx 0.386.$

Check.

Back-substitute the estimate: $3^{2(0.386)+1} = 3^{1.771} \approx 7$. ✓

An exponential equation that is secretly a quadratic. Solve $\mathrm{e}^{2x} - 5\mathrm{e}^{x} + 6 = 0$.

Read.

Substitute $u = \mathrm{e}^{x}$ (note $u > 0$), solve the quadratic, then undo the substitution.

Plan.

Let $u = \mathrm{e}^{x}$, so $\mathrm{e}^{2x} = (\mathrm{e}^{x})^2 = u^2$:

$$\begin{aligned} u^2 - 5u + 6 &= 0 && \text{the equation in } u \\ (u-2)(u-3) &= 0 && \text{factor} \\ u = 2 \;\text{ or }\; u &= 3 && \text{both positive, so both legal} \end{aligned}$$

Undo the substitution: $\mathrm{e}^{x} = 2$ gives $x = \ln 2$, and $\mathrm{e}^{x} = 3$ gives $x = \ln 3$.

Solve.

$x = \ln 2$ or $x = \ln 3$.

Check.

Had a root been negative it would be rejected, because $\mathrm{e}^{x} > 0$ for every real x.

A log equation with an extraneous root. Solve $\ln x + \ln(x-3) = \ln 10$.

Read.

Combine the left side with the product law, drop the logs, solve, then discard roots outside the domain.

Plan.

Combine the logs, drop them, and solve:

$$\begin{aligned} \ln\big(x(x-3)\big) &= \ln 10 && \text{product law} \\ x(x-3) &= 10 && \ln \text{ is one-to-one} \\ x^2 - 3x - 10 &= 0 && \text{expand and set to zero} \\ (x-5)(x+2) &= 0 && \text{factor} \\ x = 5 \quad &\text{or} \quad x = -2 && \text{two candidates} \end{aligned}$$

Check the domain: $\ln x$ needs $x > 0$ and $\ln(x-3)$ needs $x > 3$. Only $x = 5$ survives; $x = -2$ is extraneous.

Solve.

$x = 5$.

Check.

Combining logs can enlarge the domain and admit false roots, so every solution must be tested against the *original* equation. ✓

Doubling time. A culture grows as $P(t) = P_0\, \mathrm{e}^{0.2t}$ with t in hours. How long until it doubles?

Read.

Set $P = 2P_0$, cancel P_0, and solve the exponential with a logarithm.

Plan.

Set $P = 2P_0$ and solve for t:

$$\begin{aligned} P_0 \mathrm{e}^{0.2t} &= 2P_0 && \text{doubling condition} \\ \mathrm{e}^{0.2t} &= 2 && \text{cancel } P_0 \\ 0.2t &= \ln 2 && \text{take } \ln \\ t &= \frac{\ln 2}{0.2} = 5 \ln 2 && \text{solve for } t \end{aligned}$$

Solve.

$t = 5 \ln 2 \approx 3.47$ hours.

Check.

Doubling time depends only on the rate 0.2, not on the starting amount P_0—the hallmark of exponential growth. ✓

Expanding with the log laws. Write $\ln \frac{x^3\sqrt{y}}{z^2}$ as a sum and difference of simple logarithms.

Read.

Apply the quotient law to split top from bottom, then the product and power laws on each piece.

Plan.

Apply the laws in order — quotient, product, then power:

$$\begin{aligned} \ln \frac{x^3\sqrt{y}}{z^2} &= \ln(x^3\sqrt{y}) - \ln(z^2) && \text{quotient law} \\ &= \ln(x^3) + \ln(y^{1/2}) - \ln(z^2) && \text{product law on the numerator} \\ &= 3\ln x + \tfrac{1}{2}\ln y - 2\ln z && \text{power law lowers each exponent} \end{aligned}$$

Solve.

$\ln \frac{x^3\sqrt{y}}{z^2} = 3\ln x + \frac{1}{2}\ln y - 2\ln z$.

Check.

Each rule was applied once, in order product/quotient then power—the reverse of how you would recombine them. ✓

Misconception

The log of a sum does not split. $\log(x+y)$ is *not* $\log x + \log y$; the sum law in (1.2) is for the log of a *product*. There is no useful identity for the log of a sum—leave it intact.

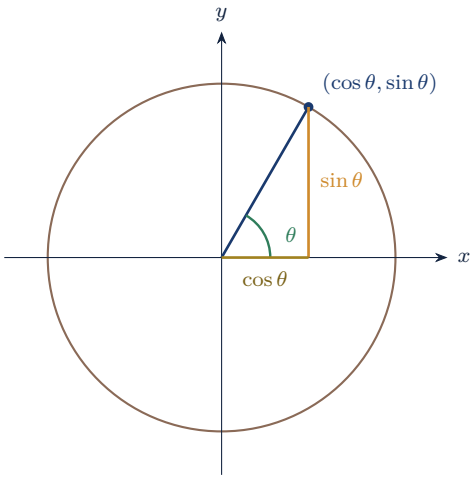
Trigonometry from the Unit Circle

Place an angle θ at the origin, measured counterclockwise from the positive x-axis, and follow its terminal ray to where it meets the unit circle. That point is $(\cos\theta, \sin\theta)$—this single definition gives all six trigonometric functions, explains their signs in each quadrant, and makes their periodicity

(2π for sine and cosine, π for tangent) visible at a glance. From the same picture the Pythagorean identity is immediate, because the point lies on the unit circle:

$$\cos^2\theta + \sin^2\theta = 1. \tag{1.3}$$

You should not try to memorize the identity zoo. Memorize a few parents and the two moves that breed the rest.



The unit circle. The terminal point of angle θ has coordinates $(\cos\theta, \sin\theta)$; its height is $\sin\theta$ (orange) and its horizontal offset is $\cos\theta$ (gold). Because the point lies on the circle, $\cos^2\theta + \sin^2\theta = 1$.

Parents to own. Alongside the Pythagorean identity (1.3), the angle-sum formulas are the other parents:

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta, \qquad \cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta. \tag{1.4}$$

Dividing the Pythagorean identity by $\cos^2\theta$ gives $\tan^2\theta + 1 = \sec^2\theta$; the angle-sum formulas at $\beta = \alpha$ give $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta = 1 - 2\sin^2\theta$.

Derivation. *the double-angle formulas from the angle-sum formulas*

Set $\beta = \alpha$ in the sine sum formula (1.4): $\sin(2\alpha) = \sin(\alpha+\alpha) = \sin\alpha\cos\alpha + \cos\alpha\sin\alpha = 2\sin\alpha\cos\alpha$. Do the same in the cosine sum formula: $\cos(2\alpha) = \cos\alpha\cos\alpha - \sin\alpha\sin\alpha = \cos^2\alpha - \sin^2\alpha$. Replacing $\cos^2\alpha$ with $1 - \sin^2\alpha$ (Pythagorean identity (1.3)) gives the alternate form $\cos 2\alpha = 1 - 2\sin^2\alpha$, and replacing $\sin^2\alpha$ instead gives $\cos 2\alpha = 2\cos^2\alpha - 1$. You never memorize three cosine double-angle formulas—you carry one and swap with the Pythagorean

identity.

Special-angle values to have instantly. Every exact evaluation on the exam builds from these first-quadrant values (angles in radians):

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$
$\sin\theta$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$

Values at $\frac{\pi}{2}$ and beyond come from the same reference angles with signs set by the quadrant. Memorize the pattern sin: $\frac{\sqrt{0}}{2}, \frac{\sqrt{1}}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{2}, \frac{\sqrt{4}}{2}$ across $0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}$ and read cosine backwards.

Approach Read · Plan · Solve · Check

Read. Identify the target angle or the trigonometric equation, and decide which parent identity — Pythagorean or angle-sum — applies.

Plan.

1. For an exact value, write the angle as a sum or difference of special angles and apply an angle-sum formula.
2. For an equation, reduce to a single trigonometric function, factor, and read *every* angle in the interval off the unit circle.
3. Use a double-angle formula, swapping via the Pythagorean identity, when the argument is 2θ.

Solve. Substitute special values one line at a time, simplifying to a single fraction.

Check. Verify numerically, or confirm the answers satisfy a Pythagorean identity.

Common traps — every wrong answer is one of these

Reports only the first solution.
A reference angle usually yields two angles per revolution; list every solution in the requested interval, not just the first one you see.

Works in degrees where radians are required.
The clean derivative $\frac{\mathrm{d}}{\mathrm{d}\theta}\sin\theta = \cos\theta$ and the arc-length reading of an angle hold only in radians.

Memorizes three cosine double-angle forms.
Carry one form and swap with the Pythagorean identity rather than confusing $\cos^2 - \sin^2$, $1 - 2\sin^2$, and $2\cos^2 - 1$.

Accepts an impossible value. 1. Solves for $\sin\theta$ or $\cos\theta$.

2. Keeps a value of magnitude above 1.
3. Fix: discard any $|\sin\theta| > 1$ or $|\cos\theta| > 1$ — no angle produces it.

Splits $\sin(\alpha + \beta)$.
Sine of a sum is *not* $\sin\alpha + \sin\beta$; use the angle-sum formula.

An exact value by subtracting angles. Find $\cos 15°$ exactly.

Read.

Write $15°$ as a difference of special angles, then apply the cosine angle-difference formula.

Plan.

Write $15° = 45° - 30°$ and apply the cosine angle-difference formula:

$$\begin{aligned}
\cos 15° &= \cos(45° - 30°) && \text{split into special angles} \\
&= \cos 45° \cos 30° + \sin 45° \sin 30° && \text{angle-difference formula} \\
&= \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{3}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{1}{2} && \text{substitute special values} \\
&= \frac{\sqrt{6}}{4} + \frac{\sqrt{2}}{4} && \text{multiply}
\end{aligned}$$

Solve.

$\cos 15° = \frac{\sqrt{6}+\sqrt{2}}{4}$.

Check.

Numerically $\frac{\sqrt{6}+\sqrt{2}}{4} \approx \frac{2.449+1.414}{4} \approx 0.966$, and $\cos 15° \approx 0.966$. ✓

Using a double angle. Given $\sin\theta = \frac{3}{5}$ with θ in the first quadrant, find $\sin 2\theta$ and $\cos 2\theta$.

Read.

Find $\cos\theta$ from the Pythagorean identity, then substitute into the double-angle formulas.

Plan.

Find $\cos\theta$ from the Pythagorean identity, then substitute:

$$\cos^2\theta = 1 - \tfrac{9}{25} = \tfrac{16}{25} \qquad \text{Pythagorean identity}$$
$$\cos\theta = \tfrac{4}{5} \qquad \text{first quadrant, so positive}$$
$$\sin 2\theta = 2\sin\theta\cos\theta = 2\cdot\tfrac{3}{5}\cdot\tfrac{4}{5} = \tfrac{24}{25} \qquad \text{double-angle for sine}$$
$$\cos 2\theta = 1 - 2\sin^2\theta = 1 - \tfrac{18}{25} = \tfrac{7}{25} \qquad \text{double-angle for cosine}$$

Solve.

$\sin 2\theta = \frac{24}{25}$, $\cos 2\theta = \frac{7}{25}$.

Check.

Check the new Pythagorean identity: $\left(\frac{24}{25}\right)^2 + \left(\frac{7}{25}\right)^2 = \frac{576+49}{625} = 1$. ✓

Problem-solving strategy — Solving a trigonometric equation

(1) Reduce to a single trig function if you can, using identities. (2) Solve the resulting algebraic equation for that function (often by factoring a quadratic in sin or cos). (3) Read *all* angles in the requested interval off the unit circle—one reference angle usually yields two solutions per revolution. (4) Discard any value outside the range of sine or cosine (nothing has sin > 1).

Solving a trigonometric equation on an interval. Solve $2\sin^2\theta - \sin\theta - 1 = 0$ for $\theta \in [0, 2\pi)$.

Read.

Treat $\sin\theta$ as the variable and factor; then read *all* angles in the interval from the unit circle.

Plan.

Factor as a quadratic in $s = \sin\theta$:

$$2s^2 - s - 1 = (2s+1)(s-1) = 0 \qquad \text{factor}$$
$$\sin\theta = -\tfrac{1}{2} \quad \text{or} \quad \sin\theta = 1 \qquad \text{set each factor to zero}$$

On $[0, 2\pi)$, $\sin\theta = -\frac{1}{2}$ at $\theta = \frac{7\pi}{6}$ and $\theta = \frac{11\pi}{6}$; and $\sin\theta = 1$ at $\theta = \frac{\pi}{2}$.

Solve.

$\theta = \frac{\pi}{2},\ \frac{7\pi}{6},\ \frac{11\pi}{6}$.

Check.

Three solutions—dropping either of the $-\frac{1}{2}$ angles is the standard mistake, so count intersections with the whole circle, not just the first one you see.

Did You Know? *why calculus insists on radians*

An angle in radians is the arc length it subtends on the unit circle, so a full turn is 2π. This is not a stylistic choice: the clean derivative rule $\frac{\mathrm{d}}{\mathrm{d}\theta}\sin\theta = \cos\theta$ holds *only* in radians. In degrees the same derivative carries an ugly factor of $\pi/180$. Train yourself to think in radians now and every later chapter runs smoother.

Analytic Geometry: Lines and Conics

A line is fixed by its slope $m = \Delta y/\Delta x$ and one point, giving the point-slope form $y - y_0 = m(x - x_0)$; two lines are parallel when their slopes are equal and perpendicular when the slopes multiply to -1. The conics—circle, ellipse, parabola, hyperbola—each have a clean *standard form* from which every feature is read directly, and the one skill that gets you there from a messy expanded equation is *completing the square*. Nearly every "identify this conic" or "find the center" question on the exam is a completing-the-square question wearing a disguise.

Standard forms. Circle $(x-h)^2 + (y-k)^2 = r^2$; ellipse $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$; hyperbola $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$; parabola $(y-k)^2 = 4p(x-h)$. For an ellipse with $a > b$, the foci lie at distance $c = \sqrt{a^2 - b^2}$ from the center along the major axis.

Derivation. *the circle equation from the distance formula*

The distance between (x, y) and (h, k) is $\sqrt{(x-h)^2 + (y-k)^2}$ (the Pythagorean theorem applied to the horizontal and vertical legs). A circle of radius r centered at (h, k) is by definition the set of points at distance exactly r from the center: $\sqrt{(x-h)^2 + (y-k)^2} = r$. Squaring both sides removes the root and gives the standard form $(x-h)^2 + (y-k)^2 = r^2$. Every conic's standard form is, in the same way, a geometric distance condition rewritten in coordinates.

Approach Read · Plan · Solve · Check

Read. Identify the conic from its squared terms: both squares present with equal positive coefficients is a circle, both positive but unequal an ellipse, opposite signs a hyperbola, only one square a parabola.

Plan.

1. Group the x-terms and the y-terms.
2. Complete the square in each, adding the same amounts to the other side.
3. Match the result to a standard form and read off center, radius, or foci.

Solve. Complete one square at a time, balancing both sides as you go.

Check. Confirm the right-hand side is positive for a real circle or ellipse, and that c comes out real.

Common traps — every wrong answer is one of these

Adds to one side only.
Completing the square adds a constant that must be added to *both* sides; changing only the left alters the equation.

Confuses a^2 and b^2 for the foci.
The larger denominator is a^2 and marks the major axis; read it before computing c.

Misreads the sign of the shift.
$(x-h)^2$ centers at $+h$; a term written $(x+3)^2$ means $h=-3$, not $+3$.

Swaps the ellipse and hyperbola c-formula. 1. Writes $c^2=a^2-b^2$ for a hyperbola.
2. Gets an imaginary c.
3. Fix: a hyperbola uses $c^2=a^2+b^2$, an ellipse $c^2=a^2-b^2$.

Forgets the perpendicular-slope test.
Perpendicular lines have slopes multiplying to -1, not slopes that are equal or merely opposite in sign.

Completing the square to identify a conic. Identify the curve $x^2+y^2-6x+4y+9=0$ and give its center and radius.

Read.

Group x-terms and y-terms, complete the square in each, and match a standard form.

Plan.

Group, complete each square, and match a standard form:

$$\begin{aligned} (x^2-6x)+(y^2+4y) &= -9 && \text{group } x\text{- and } y\text{-terms} \\ (x^2-6x+9)+(y^2+4y+4) &= -9+9+4 && \text{add 9 and 4 to both sides} \\ (x-3)^2+(y+2)^2 &= 4 && \text{write as perfect squares} \end{aligned}$$

Solve.

A circle with center $(3,-2)$ and radius $r=\sqrt{4}=2$.

Check.

The right side, 4, is positive, so a real circle exists; a negative value there would signal no real graph. ✓

Extracting the foci of an ellipse. Find the foci and eccentricity of $\frac{(x-1)^2}{25}+\frac{(y+2)^2}{9}=1$.

Read.

Read h,k,a,b off the standard form; the larger denominator is a^2 and sets the major axis; then $c=\sqrt{a^2-b^2}$.

Plan.

Read the standard form, then compute c:

$$\begin{aligned} (h,k)=(1,-2),\quad a^2=25,\quad b^2=9 && \text{since } 25>9,\ a \text{ is horizontal} \\ c^2=a^2-b^2=25-9=16 && \text{ellipse focal relation} \\ c=4 && \text{take the square root} \end{aligned}$$

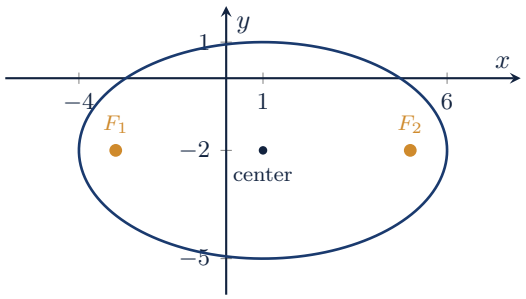
The foci sit $c=4$ to each side of the center along the horizontal major axis.

Solve.

Foci at $(1\pm4,-2)=(5,-2)$ and $(-3,-2)$; eccentricity $e=c/a=4/5$.

Check.

$0<e<1$ confirms an ellipse (a circle would give $e=0$, a parabola $e=1$). ✓



The ellipse $\frac{(x-1)^2}{25}+\frac{(y+2)^2}{9}=1$ *with center* $(1,-2)$ *and foci* $F_1(-3,-2)$, $F_2(5,-2)$ *at distance* $c=4$ *along the major axis.*

A hyperbola and its asymptotes. Put $9x^2-4y^2-54x+72=0$ into standard form and give its center and asymptotes.

Read.

Group and complete the square in x; there is no linear y-term, so only x shifts. Divide to reach standard form.

Plan.

Group the x-terms, complete the square, and divide to standard form:

$$\begin{aligned} 9(x^2-6x)-4y^2+72&=0 && \text{group and factor the } x\text{-terms}\\ 9\big((x-3)^2-9\big)-4y^2+72&=0 && \text{since } x^2-6x=(x-3)^2-9\\ 9(x-3)^2-4y^2&=9 && \text{simplify the constants}\\ \frac{(x-3)^2}{1}-\frac{y^2}{9/4}&=1 && \text{divide by } 9 \end{aligned}$$

Center $(3,0)$, $a^2=1$, $b^2=\frac{9}{4}$. Asymptotes pass through the center with slopes $\pm b/a=\pm\frac{3}{2}$.

Solve.

Center $(3,0)$; asymptotes $y=\pm\frac{3}{2}(x-3)$.

Check.

The minus sign between the squared terms is the signature of a hyperbola; an ellipse would have a plus. ✓

Vertex and focus of a parabola. Put $y=x^2-4x+7$ into vertex form and give its vertex and axis of symmetry.

Read.

Complete the square in x to reach the form $y = (x-h)^2 + k$; then (h, k) is the vertex.

Plan.

Complete the square in x:

$$\begin{aligned} y &= (x^2 - 4x) + 7 && \text{group the } x\text{-terms} \\ &= (x^2 - 4x + 4) - 4 + 7 && \text{add and subtract } \left(\tfrac{-4}{2}\right)^2 = 4 \\ &= (x-2)^2 + 3 && \text{write as a perfect square} \end{aligned}$$

Solve.

Vertex $(2, 3)$; axis of symmetry $x = 2$; opens upward.

Check.

Since the coefficient of $(x-2)^2$ is positive, the vertex is the minimum, and $y \geq 3$ for all x—a quick way to read the range straight off vertex form. ✓

Sequences and Series

A sequence is just a function whose domain is the positive integers. Two patterns account for most exam questions. An *arithmetic* sequence adds a fixed difference d each step, $a_n = a_1 + (n-1)d$, and its partial sum is $S_n = \frac{n}{2}(a_1 + a_n)$—the count of terms times their average. A *geometric* sequence multiplies by a fixed ratio r each step, $a_n = a_1 r^{\,n-1}$, with partial sum $S_n = a_1 \frac{1-r^n}{1-r}$. The subtle point, and the one the exam probes, is convergence of the infinite geometric series: it settles to a finite value *only* when the terms shrink fast enough, which for a geometric series means the ratio is smaller than one in size.

Derivation. *the geometric series sum*

Let $S = a + ar + ar^2 + \cdots + ar^{n-1}$ be the partial sum. Multiply through by r: $rS = ar + ar^2 + \cdots + ar^n$. Subtract the second line from the first; every middle term cancels, leaving $S - rS = a - ar^n$, i.e. $S(1-r) = a(1-r^n)$, so $S = a\frac{1-r^n}{1-r}$. Now let $n \to \infty$: if $|r| < 1$ then $r^n \to 0$ and the sum approaches $\frac{a}{1-r}$; if $|r| \geq 1$ the term r^n does not die and the series diverges. The whole convergence rule falls out of this one subtraction.

Infinite geometric series.

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \quad \text{when } |r| < 1; \tag{1.5}$$

otherwise the series diverges. Terms tending to zero is *necessary* for convergence but not sufficient—the harmonic series $\sum 1/n$ has terms going to zero yet diverges.

Approach Read · Plan · Solve · Check

Read. Classify the sequence as arithmetic (a fixed difference d) or geometric (a fixed ratio r), and note the first term.

Plan.

1. Find the number of terms n from the last-term formula ($a_n = a_1 + (n-1)d$ or $a_n = a_1 r^{n-1}$).
2. For an arithmetic sum use $S_n = \frac{n}{2}(a_1 + a_n)$; for a finite geometric sum use $S_n = a_1 \frac{1-r^n}{1-r}$.
3. For an infinite geometric series use $\frac{a}{1-r}$ — but only when $|r| < 1$.

Solve. Solve for n first, then substitute into the sum formula one step at a time.

Check. For an infinite series confirm $|r| < 1$; for a finite sum, verify against the "count times average" picture.

Common traps — every wrong answer is one of these

Sums a divergent geometric series.
Applies $\frac{a}{1-r}$ when $|r| \geq 1$. The closed form for the *infinite* sum requires $|r| < 1$; otherwise it diverges.

Miscounts the terms. 1. Uses $a_n = a_1 + (n-1)d$.

2. Drops the $(n-1)$, or forgets both endpoints are counted.
3. Fix: solve the last-term equation carefully — the count includes the first and last terms.

Reads terms$\to 0$ as convergence.
Terms shrinking to zero is necessary but not sufficient; the harmonic series has terms going to zero yet diverges.

Confuses the two families.
Check whether consecutive terms differ by a constant *difference* (arithmetic) or a constant *ratio* (geometric) before choosing a formula.

Misplaces the exponent.
A geometric term is $a_1 r^{n-1}$, not $a_1 r^n$; the first term uses the zeroth power of r.

An arithmetic sum. Sum $3 + 7 + 11 + \cdots + 99$.

Read.

Identify a_1 and d, find the number of terms n, then apply $S_n = \frac{n}{2}(a_1 + a_n)$.

Plan.

With $a_1 = 3$ and $d = 4$, find n from the last term, then sum:

$$\begin{aligned} 99 &= 3 + 4(n-1) && \text{last term } a_n = 99 \\ 4(n-1) &= 96 && \text{subtract 3} \\ n &= 25 && \text{so there are 25 terms} \\ S_{25} &= \tfrac{25}{2}(3+99) = \tfrac{25}{2} \cdot 102 = 25 \cdot 51 && \text{count times average} \end{aligned}$$

Solve.

$S = 1275$.

Check.

The sum is the number of terms (25) times the average of first and last ($\frac{3+99}{2} = 51$)—Gauss's trick, and a fast mental check. ✓

A finite geometric sum. Evaluate $2 + 6 + 18 + \cdots + 486$.

Read.

Identify a_1 and r, find n from the last term, then apply $S_n = a_1 \frac{1-r^n}{1-r}$.

Plan.

With $a_1 = 2$ and $r = 3$, find n from the last term, then sum:

$$\begin{aligned} 2 \cdot 3^{n-1} &= 486 && \text{last term } a_n = 486 \\ 3^{n-1} &= 243 = 3^5 && \text{divide by 2} \\ n &= 6 && \text{so } n - 1 = 5 \\ S_6 &= 2 \cdot \frac{1 - 3^6}{1 - 3} = 2 \cdot \frac{-728}{-2} = 2 \cdot 364 && \text{geometric sum formula} \end{aligned}$$

Solve.

$S = 728$.

Check.

Because $|r| = 3 > 1$, this *finite* sum is large and grows fast; the closed form still applies—only the *infinite* sum would diverge. ✓

A repeating decimal as a geometric series. Write $0.\overline{27} = 0.272727\ldots$ as a fraction in lowest terms.

Read.

Express the repeating block as an infinite geometric series, sum with $a/(1-r)$, then reduce.

Plan.

Write the repeating block as a geometric series, then sum:

$$0.\overline{27} = \frac{27}{100} + \frac{27}{100^2} + \cdots \qquad a = \tfrac{27}{100},\ r = \tfrac{1}{100}$$

$$\frac{a}{1-r} = \frac{27/100}{1 - 1/100} = \frac{27/100}{99/100} \qquad \text{sum with } |r| < 1$$

$$= \frac{27}{99} \qquad \text{simplify, then reduce by 9}$$

Solve.

$0.\overline{27} = \frac{3}{11}$.

Check.

Divide back: $3 \div 11 = 0.2727\ldots$ ✓ Every repeating decimal is a geometric series in disguise.

Connection

Toward Chapter 3. The finite geometric sum you just derived is the seed of the *power series* that dominate the calculus chapters: $\frac{1}{1-x} = \sum_{n\geq 0} x^n$ for $|x| < 1$ is literally the infinite geometric series (1.5) with $a = 1$, $r = x$. Later, Taylor series extend this idea to write e^x, $\sin x$, and $\cos x$ as infinite polynomials—and Euler's formula (next section) is the moment those three series collide.

Complex Numbers and Euler's Formula

Write a complex number as $z = x + \mathrm{i}y$ with $\mathrm{i}^2 = -1$; its conjugate is $\bar{z} = x - \mathrm{i}y$ and its modulus $|z| = \sqrt{x^2 + y^2}$, so that $z\bar{z} = |z|^2$. The single most useful mental image is geometric: the number z is the point (x, y) in the plane, and *multiplication rotates*. Multiplying by i turns a point 90° counterclockwise; more generally, multiplication adds angles and multiplies lengths. That is exactly what the polar form $z = r(\cos\theta + \mathrm{i}\sin\theta) = r\mathrm{e}^{\mathrm{i}\theta}$ records, with $r = |z|$ and $\theta = \arg z$.

Division is handled by the conjugate: to simplify $\frac{1}{z}$ or $\frac{w}{z}$, multiply top and bottom by $\bar{z}$, turning the denominator into the real number $|z|^2$.

Approach Read · Plan · Solve · Check

Read. Decide whether the task is arithmetic (add, multiply, divide) or a power or root; for powers and roots, plan to use polar form.

Plan.

1. For division, multiply numerator and denominator by the conjugate of the denominator.
2. For a power, convert to $r\mathrm{e}^{\mathrm{i}\theta}$, apply De Moivre, then convert back.
3. For n-th roots, take the n-th root of the modulus and split the argument into n equally spaced values.

Solve. Simplify with $\mathrm{i}^2 = -1$ one step at a time, keeping modulus and argument separate.

Check. Confirm a divided result has a real denominator, or raise a root back to its power to recover the original.

Common traps — every wrong answer is one of these

Leaves i in the denominator.
Multiply by the conjugate; the product $z\bar{z} = |z|^2$ is real, so there is never a reason to keep i downstairs.

Finds only one root.
An n-th root has n values spaced $360°/n$ apart; add full turns to the argument before dividing by n.

Forgets to raise the modulus.
1. Applies De Moivre to the angle.
2. Leaves the modulus unchanged.
3. Fix: $(r\mathrm{e}^{\mathrm{i}\theta})^n = r^n\mathrm{e}^{\mathrm{i}n\theta}$ — the modulus is raised to the power too.

Mishandles i^2.

Treats i^2 as $+1$. The defining fact is $\mathrm{i}^2 = -1$, and it drives every complex computation.

Adds moduli when multiplying.

Multiplication *multiplies* the moduli and *adds* the arguments — not the other way around.

Dividing complex numbers. Write $\frac{2+3\mathrm{i}}{1-\mathrm{i}}$ in the form $x + \mathrm{i}y$.

Read.

Multiply numerator and denominator by the conjugate of the denominator, then simplify using $\mathrm{i}^2 = -1$.

Plan.

Multiply by the conjugate $1 + \mathrm{i}$, then simplify numerator and denominator:

$$\frac{2+3\mathrm{i}}{1-\mathrm{i}} = \frac{2+3\mathrm{i}}{1-\mathrm{i}} \cdot \frac{1+\mathrm{i}}{1+\mathrm{i}} \qquad \text{multiply by the conjugate}$$

$$= \frac{(2+3\mathrm{i})(1+\mathrm{i})}{(1-\mathrm{i})(1+\mathrm{i})} \qquad \text{combine}$$

$$(2+3\mathrm{i})(1+\mathrm{i}) = 2 + 2\mathrm{i} + 3\mathrm{i} + 3\mathrm{i}^2 = -1 + 5\mathrm{i} \qquad \text{numerator; } \mathrm{i}^2 = -1$$

$$(1-\mathrm{i})(1+\mathrm{i}) = 1 - \mathrm{i}^2 = 1 + 1 = 2 \qquad \text{denominator is real}$$

$$\frac{2+3\mathrm{i}}{1-\mathrm{i}} = \frac{-1+5\mathrm{i}}{2} \qquad \text{assemble}$$

Solve.

$\frac{2+3\mathrm{i}}{1-\mathrm{i}} = -\frac{1}{2} + \frac{5}{2}\mathrm{i}$.

Check.

The conjugate trick always makes the denominator real; there is never a reason to leave i downstairs. ✓

Euler and De Moivre.

$$\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta, \tag{1.6}$$

and therefore

$$\left(\cos\theta + \mathrm{i}\sin\theta\right)^n = \cos n\theta + \mathrm{i}\sin n\theta. \tag{1.7}$$

In polar form, $r_1\mathrm{e}^{\mathrm{i}\theta_1} \cdot r_2\mathrm{e}^{\mathrm{i}\theta_2} = r_1 r_2\,\mathrm{e}^{\mathrm{i}(\theta_1+\theta_2)}$: multiply the moduli, add the arguments.

Derivation. *De Moivre's theorem from Euler's formula*

Euler's formula (1.6) says $\cos\theta + \mathrm{i}\sin\theta = \mathrm{e}^{\mathrm{i}\theta}$. Raise both sides to the nth power. The right side follows the exponent law: $\left(\mathrm{e}^{\mathrm{i}\theta}\right)^n = \mathrm{e}^{\mathrm{i}n\theta}$. Apply Euler's formula again to $\mathrm{e}^{\mathrm{i}n\theta}$: it equals $\cos n\theta + \mathrm{i}\sin n\theta$. Therefore $\left(\cos\theta + \mathrm{i}\sin\theta\right)^n = \cos n\theta + \mathrm{i}\sin n\theta$, which is exactly De Moivre's theorem. Powers of complex numbers are just repeated addition of the argument—rotation done n times.

A high power via polar form. Compute $(1+\mathrm{i})^8$.

Read.

Convert $1+\mathrm{i}$ to polar form, raise to the power with De Moivre, then convert back.

Plan.

Convert to polar form, then apply De Moivre:

$$
\begin{aligned}
|1+\mathrm{i}| = \sqrt{1^2+1^2} = \sqrt{2}, \quad \arg = \tfrac{\pi}{4} &\qquad \text{modulus and argument}\\
1+\mathrm{i} = \sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4} &\qquad \text{polar form}\\
(1+\mathrm{i})^8 = (\sqrt{2})^8\,\mathrm{e}^{\mathrm{i}\cdot 8\pi/4} = 16\,\mathrm{e}^{\mathrm{i}\,2\pi} &\qquad \text{raise modulus and angle}\\
= 16 &\qquad \text{since } \mathrm{e}^{\mathrm{i}2\pi} = 1
\end{aligned}
$$

Solve.

$(1+\mathrm{i})^8 = 16$.

Check.

A full 2π of rotation lands back on the positive real axis, so the answer is real and positive—and 16 is far quicker than an eightfold binomial expansion. ✓

Roots of unity. Find all solutions of $z^3 = 1$.

Read.

Write 1 in polar form allowing full turns, take the cube root of modulus and split the argument into three equal parts.

Plan.

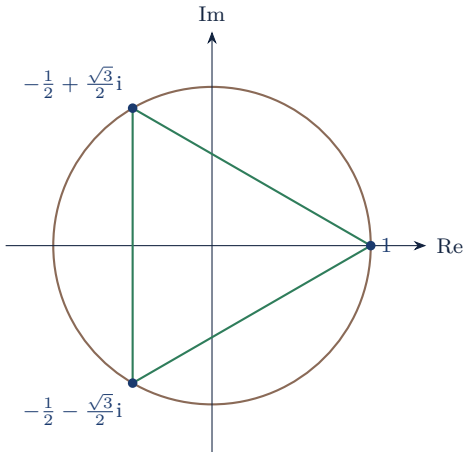
Write $1 = \mathrm{e}^{\mathrm{i}\cdot 2\pi k}$ for integer k. Taking cube roots, $z = \mathrm{e}^{\mathrm{i}\cdot 2\pi k/3}$ for $k = 0, 1, 2$ (further k repeat). These are $z = 1$ (k=0), $z = \mathrm{e}^{\mathrm{i}2\pi/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}\mathrm{i}$ (k=1), and $z = \mathrm{e}^{\mathrm{i}4\pi/3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}\mathrm{i}$ (k=2).

Solve.

$z = 1,\ -\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}.$

Check.

The three roots are equally spaced by 120° on the unit circle—the nth roots of unity always sit at the vertices of a regular n-gon. ✓



The three cube roots of unity sit at the vertices of an equilateral triangle on the unit circle, spaced 120° *apart—the geometry that De Moivre's theorem* (1.7) *predicts.*

Square roots of a complex number. Find both square roots of i.

Read.

Write i in polar form, halve the argument for the root, and add a full turn to get the second root.

Plan.

Write i in polar form, then halve the argument:

$$\begin{aligned} \mathrm{i} &= \mathrm{e}^{\mathrm{i}\pi/2} && \text{modulus } 1,\ \text{argument } \tfrac{\pi}{2} \\ \sqrt{\mathrm{i}} &= \mathrm{e}^{\mathrm{i}\pi/4} = \cos\tfrac{\pi}{4} + \mathrm{i}\sin\tfrac{\pi}{4} && \text{halve argument, root the modulus} \\ &= \tfrac{\sqrt{2}}{2} + \tfrac{\sqrt{2}}{2}\mathrm{i} && \text{evaluate} \end{aligned}$$

The second root sits half a turn (π) away: $\mathrm{e}^{\mathrm{i}(\pi/4+\pi)} = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathrm{i}$.

Solve.

$\sqrt{\mathrm{i}} = \pm\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathrm{i}\right).$

Check.

Check by squaring: $\left(\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}\mathrm{i}\right)^2 = \frac{1}{2}+2\cdot\frac{1}{2}\mathrm{i}+\frac{1}{2}\mathrm{i}^2 = \frac{1}{2}+\mathrm{i}-\frac{1}{2} = \mathrm{i}$. ✓ Every nonzero complex number has exactly two square roots, half a turn apart.

Connection

One idea, three costumes. The exponential of §1.3 (growth), the unit circle of §1.4 (rotation), and complex multiplication here are the same object: $\mathrm{e}^{\mathrm{i}\theta}$ walks around the unit circle at unit speed. Euler's formula (1.6) is the seam where these three sections meet, and it is the reason complex numbers are the natural language for anything that oscillates.

Did You Know? *Euler's identity, "the most beautiful equation"*

Set $\theta = \pi$ in Euler's formula (1.6): $\mathrm{e}^{\mathrm{i}\pi} = \cos\pi + \mathrm{i}\sin\pi = -1$, so $\mathrm{e}^{\mathrm{i}\pi}+1 = 0$. This single line ties together the additive identity 0, the multiplicative identity 1, the circle constant π, the growth constant e, and the imaginary unit i—five of the most important numbers in mathematics, with each appearing exactly once. It is routinely voted the most beautiful equation in mathematics, and it is simply Euler's formula seen at a half-turn.

Synthesis — Key Results You Must Own

1. A function is rule *plus* domain; invertible $\iff$ bijective; the inverse reflects the graph across $y = x$. Track the domain through every composition.
2. Degree $n \Rightarrow n$ complex roots; real coefficients $\Rightarrow$ conjugate pairs; find roots with the factor and rational-root theorems.
3. Rational function: intercepts from the numerator, vertical asymptotes from the denominator, end behavior from degree comparison; cancelled factors are holes.
4. b^x and $\log_b$ are an inverse pair; every log law is an exponent law reversed; solve exponential equations by taking a logarithm.
5. The unit circle defines all six trig functions; own $\sin^2+\cos^2 = 1$ and the angle-sum formulas and derive the rest.
6. Complete the square to classify any conic and read its center, radius, or foci.
7. A geometric series sums to $a/(1-r)$ exactly when $|r| < 1$.

8. $e^{i\theta} = \cos\theta + i\sin\theta$; multiplication rotates and scales; De Moivre gives powers and roots.

Practice Problems

Difficulty: $\star$ warm-up, $\star\star$ standard, $\star\star\star$ challenge. Target under three minutes each on a timed pass; the worked examples above model the Read–Plan–Solve–Check habit to use.

Functions, composition, inverses

1 $\star$.

State the domain of $f(x) = \frac{\sqrt{x+3}}{x-1}$.

2 $\star$.

If $f(x) = 3x - 4$, find $f^{-1}(x)$ and verify $f(f^{-1}(x)) = x$.

3 $\star\star$.

Let $f(x) = \frac{x}{x+1}$. Compute $(f \circ f)(x)$ and simplify.

4 $\star\star$.

For what real a is $g(x) = x^2 - 2ax$ one-to-one on $[1, \infty)$?

Hint: Where is the vertex of the parabola?

5 $\star\star\star$.

Suppose f is odd and invertible. Prove that f^{-1} is also odd.

Polynomial and rational functions

6 $\star$.

Find all rational roots of $x^3 - 4x^2 + x + 6$.

7 ⋆.

Divide $x^3 - 2x + 1$ by $x - 1$ and state the remainder.

8 ⋆⋆.

Give the equations of all asymptotes of $r(x) = \frac{2x^2+1}{x-3}$.

9 ⋆⋆.

A cubic with real coefficients and leading coefficient 1 has roots 1 and $2 - \mathrm{i}$. Write it in expanded form.

10 ⋆⋆⋆.

Determine the values of k for which $x^4 + kx^2 + 1 = 0$ has no real solutions.

Exponentials and logarithms

11 ⋆.

Solve $5^x = 125^{x-2}$.

12 ⋆.

Simplify $\log_2 40 - \log_2 5$.

13 ⋆⋆.

Solve $\ln(x) + \ln(x - 3) = \ln(10)$.

Hint: Check for extraneous roots against the domain of ln.

14 ⋆⋆.

Solve $4^x - 3 \cdot 2^x - 4 = 0$.

15 ⋆⋆⋆.

If $\log_a b = 3$ and $\log_b c = 2$, find $\log_a(a^2 c)$ as a number.

Trigonometry

16 ⋆.

Evaluate $\sin\frac{5\pi}{6}$ and $\cos\frac{5\pi}{6}$ exactly.

17 ⋆.

Find $\tan 75^\circ$ exactly using an angle-sum formula.

18 ⋆⋆.

Solve $\cos 2\theta = \cos\theta$ for $\theta \in [0, 2\pi)$.

19 ⋆⋆.

Given $\sin\theta = \frac{3}{5}$ with θ in the second quadrant, find $\sin 2\theta$.

20 ⋆⋆⋆.

Show that $\frac{\sin 3\theta}{\sin\theta} = 3 - 4\sin^2\theta$ wherever $\sin\theta \neq 0$.

Analytic geometry

21 ⋆.

Find the center and radius of $x^2 + y^2 + 8x - 2y + 8 = 0$.

22 ⋆⋆.

Identify the conic $9x^2 - 4y^2 - 54x + 72 = 0$ and give its center.

23 ⋆⋆.

Find the foci of $\frac{x^2}{16} + \frac{y^2}{7} = 1$.

24 ⋆⋆⋆.

The line $y = mx$ is tangent to the circle $(x-4)^2 + y^2 = 4$. Find all such m.

Hint: The distance from the center to the line equals the radius at tangency.

Sequences, series, complex numbers

25 ⋆.

Sum the arithmetic series $3 + 7 + 11 + \cdots + 99$.

26 ⋆.

Evaluate $\displaystyle\sum_{n=0}^{\infty} \left(\tfrac{2}{3}\right)^n$.

27 ⋆⋆.

Write $0.\overline{45}$ as a fraction in lowest terms.

28 ⋆⋆.

Compute $(1 - \mathrm{i}\sqrt{3})^6$ using polar form.

29 ⋆⋆.

Find all solutions of $z^4 = -16$ and describe their positions on the circle of radius 2.

30 ⋆⋆⋆.

For which real x does $\displaystyle\sum_{n=1}^{\infty} (2x-1)^n$ converge, and to what?

Further Reading

For the concept-first spirit of this chapter, the opening chapters of a rigorous calculus text (functions and the real line) repay the effort and lead directly into Chapter 2. For trigonometry and complex numbers as a single subject, the early chapters of a complex-variables book make Euler's formula feel inevitable rather than magical.

CHAPTER 2

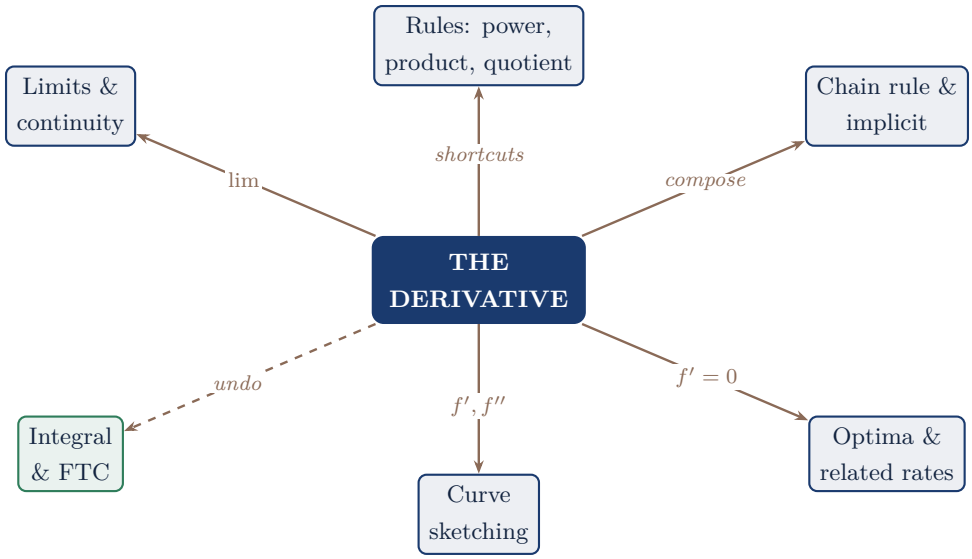
Calculus I

The Derivative and Its Uses

"The calculus was the first achievement of modern mathematics, and it is difficult to overestimate its importance."

— John von Neumann

Differential calculus is the study of* instantaneous rate of change*, and its entire machinery grows from a single limit. The derivative $f'(x)$ is the limit of the average rate of change $\frac{f(x+h)-f(x)}{h}$ as $h \to 0$: the slope of the tangent line, the velocity of a moving point, the sensitivity of an output to its input. Once that limit is in hand, a short table of rules—power, product, quotient, and above all the chain rule—lets you differentiate anything built from the precalculus families of Chapter 1 without ever returning to the limit. The payoff is a toolkit for the questions the Subject Test asks constantly: where is a function largest or smallest, how fast is one quantity changing when another is, and what does a graph do. We add the workhorse techniques the exam assumes—linear approximation, L'Hôpital's rule, the Mean Value Theorem—and close by meeting the derivative's inverse operation, the integral, and the theorem that ties the two together.



Learning Objectives

You will be able to:

1. evaluate limits, including one-sided limits, indeterminate forms, and limits at infinity, and test continuity;
2. compute derivatives from the limit definition and, more quickly, from the rules;
3. apply the chain rule, implicit differentiation, and logarithmic differentiation;
4. approximate values with the tangent line and resolve indeterminate limits with L'Hôpital's rule;
5. state and apply the Mean Value Theorem;
6. solve related-rates and optimization problems by translating words into a derivative;
7. locate extrema and inflection points and use them to sketch a curve;
8. state the Fundamental Theorem of Calculus, integrate by substitution, and compute areas between curves and net change.

Problem-solving strategy — How the exam uses this chapter

Expect fast derivative computations (power, product, quotient, chain), one or two related-rates or optimization word problems, a curve-analysis question that reads f' and f'', and a definite integral or an area. Speed comes from recognizing which rule a disguised expression needs—train the recognition, not just the mechanics.

Limits and Continuity

A *limit* answers the question: as the input x approaches some value a, what single value does the output $f(x)$ approach? Crucially, the limit is about the trend near a, not the value *at* a—a function can approach a height it never actually attains, or attains somewhere else. We write $\lim_{x\to a} f(x) = L$. For the elementary functions of Chapter 1, limits at points where the function is defined are found by direct substitution; the interesting cases are those where substitution gives the *indeterminate form* $\frac{0}{0}$, a signal to do algebra first—factor and cancel, rationalize, or combine fractions—until the offending factor disappears.

What & why. What a limit is and how to evaluate one when direct substitution fails—the foundation the derivative is built on.

Solve. Substitute; if you get $\frac{0}{0}$, factor or rationalize to cancel the vanishing factor, then substitute again.

Means. The limit is the value the outputs approach, whether or not it is attained.

Approach Read · Plan · Solve · Check

Read. Identify the target value $x \to a$ and the function, then substitute a directly as a first move.

Plan.

1. If direct substitution gives a number, that number is the limit.
2. If it gives $\frac{0}{0}$, do algebra to remove the vanishing factor — factor and cancel, rationalize with a conjugate, or combine fractions.
3. For a one-sided or piecewise limit, evaluate from each side; for a limit at infinity, divide by the highest power of x.

Solve. Carry out the algebra one line at a time, then substitute a into the simplified expression.

Check. The limit is the value the outputs approach near a — whether or not $f(a)$ is defined. For a two-sided limit, confirm the left and right sides agree.

Common traps — every wrong answer is one of these

Calls $\frac{0}{0}$ the answer.
Reads the indeterminate form $\frac{0}{0}$ as 0, as 1, or as "undefined" and stops. $\frac{0}{0}$ is a signal to do more algebra, not a value — factor, rationalize, or combine first.

Cancels terms instead of factors.
Strikes out matching pieces that are added rather than multiplied (such as the x's in $\frac{x+3}{x}$). You may only cancel a factor common to the entire numerator and denominator.

Confuses the limit with the value.
Assumes $\lim_{x\to a} f(x) = f(a)$ even when f has a hole or jump at a. The limit describes the trend near a, and can exist where $f(a)$ does not.

Ignores the two sides.
1. Evaluates the limit from only one direction.
2. Reports that as the two-sided limit.
3. Fix: the two-sided limit exists only when the left-hand and right-hand limits are equal.

A $\frac{0}{0}$ limit by factoring. Evaluate $\lim_{x\to 3} \dfrac{x^2-9}{x-3}$.

Read.

Substitute first; the indeterminate $\frac{0}{0}$ tells you to factor and cancel the common factor.

Plan.

Direct substitution gives $\frac{9-9}{3-3} = \frac{0}{0}$, which is indeterminate, so factor and cancel:

$$\frac{x^2-9}{x-3} = \frac{(x-3)(x+3)}{x-3} \qquad \text{factor the numerator (difference of squares)}$$

$$= x+3 \qquad \text{cancel the common factor } (x-3),\ x \neq 3$$

$$\lim_{x\to 3}(x+3) = 6 \qquad \text{now substitute } x = 3$$

Solve.

The limit is 6.

Check.

The original function is undefined at $x = 3$, yet the limit exists—the graph has a removable hole at $(3, 6)$, exactly the point the limit names.

A $\frac{0}{0}$ limit by rationalizing. Evaluate $\lim_{x\to 0} \dfrac{\sqrt{x+4}-2}{x}$.

Read.

Substitution gives $\frac{0}{0}$; multiply by the conjugate to clear the root, then cancel.

Plan.

Multiply numerator and denominator by the conjugate $\sqrt{x+4}+2$:

$$\frac{\sqrt{x+4}-2}{x}\cdot\frac{\sqrt{x+4}+2}{\sqrt{x+4}+2}=\frac{(x+4)-4}{x\big(\sqrt{x+4}+2\big)}=\frac{x}{x\big(\sqrt{x+4}+2\big)}=\frac{1}{\sqrt{x+4}+2}.$$

Now substitute $x=0$: $\frac{1}{\sqrt{4}+2}=\frac{1}{4}$.

Solve.

The limit is $\frac{1}{4}$.

Check.

Rationalizing is the reflex whenever a root sits inside a $\frac{0}{0}$ form; it moves the vanishing factor to where it can cancel.

One-sided limits and continuity

Sometimes the approach from the left and the approach from the right disagree. The *one-sided limits* $\lim_{x\to a^-} f(x)$ and $\lim_{x\to a^+} f(x)$ record each; the two-sided limit exists only when they are equal. This is exactly the tool for piecewise functions, whose rule changes at a break point. A function is *continuous* at a if three things agree: $f(a)$ is defined, $\lim_{x\to a} f(x)$ exists, and the two are equal. Intuitively, the graph has no break, jump, or hole there—you can draw through a without lifting the pen. Continuity matters because the major theorems of calculus (the Intermediate Value Theorem, the Extreme Value Theorem, the Mean Value Theorem) all require it.

Continuity of a piecewise function. Find the value of k that makes $f(x)=\begin{cases}x^2+1 & x\le 2\\ kx-1 & x>2\end{cases}$ continuous at $x=2$.

Read.

Match the left and right one-sided limits to the value at the break point.

Plan.

From the left, $\lim_{x\to 2^-} f(x)=2^2+1=5$, which also equals $f(2)$. From the right, $\lim_{x\to 2^+} f(x)=k(2)-1=2k-1$. Continuity requires these to be equal: $2k-1=5$, so $2k=6$.

Solve.

$k=3$.

Check.

With $k = 3$ the two pieces meet at the height 5 over $x = 2$, joining seamlessly—the graph can be drawn through the join without lifting the pen.

Limits at infinity

To ask what $f(x)$ does as x grows without bound is to ask for its *end behavior*. For a ratio of polynomials, the highest-degree terms dominate, so divide numerator and denominator by the highest power of x present and read off what survives. This is how horizontal asymptotes are found, connecting directly to the rational-function analysis of Chapter 1.

A limit at infinity. Evaluate $\displaystyle\lim_{x\to\infty} \frac{2x^2 - 3x + 1}{5x^2 + x}$.

Read.

Divide every term by the highest power, x^2; terms with x in the denominator vanish.

Plan.

Divide top and bottom by x^2:

$$\frac{2x^2 - 3x + 1}{5x^2 + x} = \frac{2 - \frac{3}{x} + \frac{1}{x^2}}{5 + \frac{1}{x}}.$$

As $x \to \infty$, each term with x in a denominator tends to 0, leaving $\frac{2}{5}$.

Solve.

The limit is $\frac{2}{5}$, the ratio of leading coefficients.

Check.

Equal degrees give a horizontal asymptote at the ratio of leading coefficients—exactly the Chapter 1 rule, now derived as a limit.

Did You Know? *Zeno's paradox, resolved*

Zeno argued that motion is impossible because covering a distance requires first covering half of it, then half of the rest, forever—infinitely many steps. The limit is the modern resolution: infinitely many shrinking steps can sum to a finite total, precisely the geometric-series fact of Chapter 1. Calculus is, in large part, the disciplined handling of "infinitely many, infinitely small."

Misconception

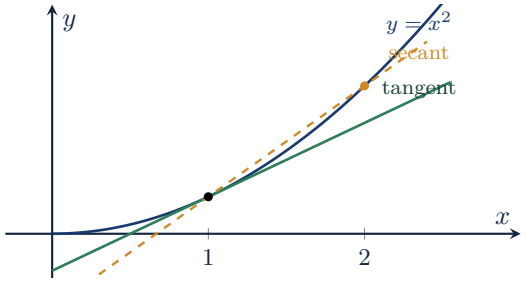
The limit is not the function value. $\lim_{x\to a} f(x)$ can exist even when $f(a)$ is undefined or different—that is the whole point of the derivative's difference quotient, which is $\frac{0}{0}$ at $h = 0$ yet has a perfectly good limit. Never evaluate a limit by blindly plugging in when substitution gives $\frac{0}{0}$.

The Derivative from Its Definition

The *derivative* of f at x is the instantaneous rate of change—the limit of the average rate of change over a shrinking interval:

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{2.1}$$

The geometric picture is the heart of the idea. The quotient $\frac{f(x+h)-f(x)}{h}$ is the slope of the *secant* line joining two points on the graph a distance h apart. As $h \to 0$, the second point slides toward the first and the secant line pivots into the *tangent* line; its slope is the derivative. Physically, if f is position, this is velocity. The quotient is $\frac{0}{0}$ at $h = 0$, so computing a derivative from the definition (2.1) is precisely a limit exercise of the kind above: expand, cancel the h in the denominator, then let $h \to 0$.



Secant to tangent. The secant (dashed) through $(1, 1)$ and $(2, 4)$ has slope 3; as the second point slides toward $(1, 1)$ the secant pivots into the tangent (green), whose slope $f'(1) = 2$ is the derivative.

Approach Read · Plan · Solve · Check

Read. Note the function f and, if given, the point; the derivative is the limit of the difference quotient, so this is a limit problem in disguise.

Plan.

1. Form the difference $f(x + h) - f(x)$ and expand or rationalize it.
2. Divide by h; the goal is to cancel the single h that makes the quotient $\frac{0}{0}$.

3. Take the limit $h \to 0$ of the simplified expression.

Solve. Do the algebra one line at a time, cancel the h in the denominator, then let $h \to 0$.

Check. The sign and size of f' should match the graph's steepness — steep where $|f'|$ is large, flat where $f' = 0$.

Common traps — every wrong answer is one of these

Sets $h = 0$ too early.
Substitutes $h = 0$ before cancelling and reports the resulting $\frac{0}{0}$ as undefined. Simplify to remove the h in the denominator *first*, then take the limit.

Botches the expansion.
Writes $(x + h)^2 = x^2 + h^2$, dropping the cross term $2xh$. Expand the binomial in full — the surviving cross term is exactly what becomes the derivative.

Forgets the conjugate for roots.
Tries to cancel h from $\sqrt{x+h} - \sqrt{x}$ directly. Multiply by the conjugate first so the numerator rationalizes and h appears as a factor.

Assumes continuity gives a derivative. 1. Sees that f is continuous at a.

2. Concludes $f'(a)$ must exist.
3. Fix: corners, cusps, and vertical tangents are continuous yet not differentiable — check that the one-sided difference quotients agree.

A derivative from first principles. Use the definition to find $f'(x)$ for $f(x) = x^2$.

Read.

Form the difference quotient, expand, cancel h, then take the limit.

Plan.

Form the difference quotient and simplify before taking the limit:

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^2 - x^2 && \text{form the difference} \\ &= 2xh + h^2 && \text{expand and cancel } x^2 \\ \frac{f(x+h) - f(x)}{h} &= 2x + h && \text{divide by } h \ (h \neq 0) \\ \lim_{h \to 0} (2x + h) &= 2x && \text{let } h \to 0 \end{aligned}$$

Solve.

$f'(x) = 2x$.

Check.

At $x = 3$ the slope is 6; the graph of x^2 does steepen as x grows, as a rate that doubles the input should. ✓

A derivative needing the conjugate. Use the definition to find $f'(x)$ for $f(x) = \sqrt{x}$.

Read.

Form the difference quotient, rationalize the numerator to cancel h, then let $h \to 0$.

Plan.

The quotient is $\frac{\sqrt{x+h}-\sqrt{x}}{h}$. Multiply by the conjugate:

$$\frac{\sqrt{x+h}-\sqrt{x}}{h} \cdot \frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}} = \frac{(x+h)-x}{h(\sqrt{x+h}+\sqrt{x})} \quad \text{multiply out; the numerator telescopes}$$

$$= \frac{1}{\sqrt{x+h}+\sqrt{x}} \quad \text{cancel the common } h$$

Let $h \to 0$: $\frac{1}{\sqrt{x}+\sqrt{x}} = \frac{1}{2\sqrt{x}}$.

Solve.

$f'(x) = \frac{1}{2\sqrt{x}}$.

Check.

The slope blows up as $x \to 0^+$—consistent with the square-root graph rising vertically at the origin—and shrinks as x grows, as the flattening curve demands. ✓

A derivative with a cubic. Use the definition to find $f'(x)$ for $f(x) = x^3$.

Read.

Form the difference quotient and expand the cube *in full*—the cross terms are where the derivative hides—then cancel h and take the limit.

Plan.

Expand $(x+h)^3$ completely before simplifying:

$$\begin{aligned} f(x+h)-f(x) &= (x+h)^3 - x^3 && \text{form the difference} \\ &= 3x^2h + 3xh^2 + h^3 && \text{expand and cancel } x^3 \\ \frac{f(x+h)-f(x)}{h} &= 3x^2 + 3xh + h^2 && \text{divide by } h \ (h \neq 0) \\ \lim_{h\to 0}\left(3x^2 + 3xh + h^2\right) &= 3x^2 && \text{let } h \to 0 \end{aligned}$$

Solve.

$f'(x) = 3x^2$.

Check.

Dropping the cross terms and writing $(x+h)^3 = x^3 + h^3$ would kill the surviving $3x^2$ entirely; that leading term is exactly the power rule nx^{n-1} with $n = 3$. ✓

Derivation. *The power rule from the definition*

For $f(x) = x^n$ with n a positive integer, the binomial theorem expands $(x+h)^n = x^n + nx^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + h^n$. Subtract x^n and divide by h:

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + h^{n-1}.$$

Every term after the first carries a factor of h, so as $h \to 0$ they all vanish, leaving $f'(x) = nx^{n-1}$. The rule that looks like a magic trick is just the first term of a binomial expansion surviving the limit.

Doing this for each new function would be slow, so the definition is used once to establish general *rules*, after which we differentiate by pattern.

Connection

Differentiable ⇒ continuous. If f has a derivative at a, it must be continuous there: a graph with a jump or hole has no single tangent slope. The converse fails—continuity does not guarantee differentiability, as the next trap shows.

Misconception

Corners and cusps have no derivative. $f(x) = |x|$ is continuous everywhere but has no derivative at 0: the slope is -1 approaching from the left and $+1$ from the right, so the one-sided limits of the difference quotient disagree and no tangent slope exists. Sharp corners, cusps, and vertical tangents are the standard places differentiability fails.

The Differentiation Rules

The core rules. With c constant and u, v differentiable functions of x:

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}, \qquad (cu)' = c\,u', \qquad (u \pm v)' = u' \pm v',$$

$$\textbf{product: } (uv)' = u'v + uv', \qquad \textbf{quotient: } \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}. \tag{2.2}$$

Also $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x$, $\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \frac{1}{x}$, $\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x$, and $\frac{\mathrm{d}}{\mathrm{d}x}\cos x = -\sin x$.

The power rule is the workhorse; the product and quotient rules handle the two ways of combining functions that addition cannot. Their forms are worth saying in words: the product rule is "derivative of the first times the second, plus the first times the derivative of the second," and the quotient rule is "bottom times derivative of top, minus top times derivative of bottom, all over bottom squared."

Derivation. *The product rule from first principles*

Apply the definition to $p(x) = u(x)v(x)$ and use the classic add-and-subtract trick in the numerator:

$$\begin{aligned} p(x+h) - p(x) &= u(x+h)v(x+h) - u(x)v(x) && \text{the change in the product} \\ &= \underbrace{u(x+h)\big[v(x+h) - v(x)\big]}_{\text{change in } v} + \underbrace{v(x)\big[u(x+h) - u(x)\big]}_{\text{change in } u} && \text{add and subtract } u(x+h)v(x) \end{aligned}$$

Divide by h and let $h \to 0$: the first bracket becomes $u(x)\,v'(x)$ (using continuity of u), the second $v(x)\,u'(x)$. Hence $p' = u'v + uv'$. The asymmetry students expect—some single tidy product—never appears because *both* factors are changing at once.

Approach Read · Plan · Solve · Check

Read. See how the expression is built — a power, a product of two functions, or a quotient — because the structure dictates the rule.

Plan.

1. Name the pieces: for a product or quotient, label u and v and record u' and v'.
2. Apply the matching rule: power nx^{n-1}, product $u'v + uv'$, or quotient $\frac{u'v-uv'}{v^2}$.
3. Simplify, factoring out common factors where possible.

Solve. Substitute u, v, u', v' into the rule and simplify one line at a time.

Check. Spot-check a value or a sign: a known zero of f' should land at a visible peak or valley of f.

Common traps — every wrong answer is one of these

Multiplies the derivatives.
Writes $(uv)' = u'v'$. The product rule is $u'v + uv'$ — *both* factors change, so two terms appear.

Flips the quotient rule.
Puts the terms in the wrong order, getting $\frac{uv'-u'v}{v^2}$. It is "bottom times derivative of top, minus top times derivative of bottom, over bottom squared" — order matters because of the minus sign.

Squares the wrong thing in the quotient.
Forgets to square the denominator, or squares the numerator. Only v is squared: $\left(\frac{u}{v}\right)' = \frac{u'v-uv'}{v^2}$.

Confuses e^x and x^n.
1. Meets e^x and reaches for the power rule.
2. Writes $x\mathrm{e}^{x-1}$.
3. Fix: the power rule is for a *variable base, constant exponent*; e^x has a constant base, and $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^x = \mathrm{e}^x$.

Product rule. Differentiate $f(x) = x^2\mathrm{e}^x$.

Read.

Identify $u = x^2$ and $v = \mathrm{e}^x$, then apply $(uv)' = u'v + uv'$.

Plan.

Here $u = x^2$ so $u' = 2x$, and $v = \mathrm{e}^x$ so $v' = \mathrm{e}^x$. Then

$$\begin{aligned} f'(x) &= u'v + uv' && \text{product rule} \\ &= 2x\,\mathrm{e}^x + x^2\,\mathrm{e}^x && u' = 2x,\ v' = \mathrm{e}^x \\ &= \mathrm{e}^x(x^2 + 2x) && \text{factor out the common } \mathrm{e}^x \end{aligned}$$

Solve.

$f'(x) = \mathrm{e}^x\,(x^2 + 2x)$.

Check.

At $x = 0$, $f'(0) = 1 \cdot 0 = 0$, consistent with $f(x) = x^2\mathrm{e}^x$ having a flat minimum at the origin. ✓

Product rule with a trig factor. Differentiate $f(x) = x \sin x$.

Read.

Take $u = x$, $v = \sin x$, and apply the product rule.

Plan.

With $u' = 1$ and $v' = \cos x$,

$$\begin{aligned} f'(x) &= u'v + uv' && \text{product rule} \\ &= (1)\sin x + x\cos x && u' = 1,\ v' = \cos x \\ &= \sin x + x\cos x && \text{simplify} \end{aligned}$$

Solve.

$f'(x) = \sin x + x\cos x$.

Check.

The lone $\sin x$ is the "rate from x growing"; the $x\cos x$ is the "rate from $\sin x$ turning"—the product rule adds the two contributions. ✓

Quotient rule. Differentiate $g(x) = \frac{x}{x^2+1}$.

Read.

Set top $u = x$, bottom $v = x^2 + 1$, and apply $(u/v)' = \frac{u'v - uv'}{v^2}$.

Plan.

With $u' = 1$ and $v' = 2x$,

$$\begin{aligned} g'(x) &= \frac{u'v - uv'}{v^2} && \text{quotient rule} \\ &= \frac{(1)(x^2+1) - (x)(2x)}{(x^2+1)^2} && u' = 1,\ v' = 2x \\ &= \frac{x^2 + 1 - 2x^2}{(x^2+1)^2} && \text{expand the numerator} \\ &= \frac{1 - x^2}{(x^2+1)^2} && \text{combine like terms} \end{aligned}$$

Solve.

$g'(x) = \frac{1-x^2}{(x^2+1)^2}$.

Check.

The numerator $1 - x^2$ is zero at $x = \pm 1$—the two turning points visible on the graph of g, which peaks at $x = 1$ and bottoms at $x = -1$. ✓

Deriving $\frac{\mathrm{d}}{\mathrm{d}x}\tan x$. Use the quotient rule to differentiate $\tan x = \frac{\sin x}{\cos x}$.

Read.

Set $u = \sin x$, $v = \cos x$, apply the quotient rule, then simplify with the Pythagorean identity.

Plan.

With $u' = \cos x$ and $v' = -\sin x$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\tan x &= \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x} && \text{quotient rule on } \tfrac{\sin x}{\cos x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} && \text{simplify the numerator} \\ &= \frac{1}{\cos^2 x} && \text{Pythagorean identity} \end{aligned}$$

Solve.

$\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \sec^2 x$.

Check.

The identity $\cos^2 + \sin^2 = 1$ collapses the numerator to 1, turning the quotient into $\sec^2 x$—a standard derivative worth memorizing once you have seen it derived. ✓

Higher-order derivatives

Differentiating a derivative gives the *second derivative* f'', and repeating gives f''', $f^{(4)}$, and so on. In motion, if $s(t)$ is position then $s'(t)$ is velocity and $s''(t)$ is acceleration—each derivative is the rate of change of the one before.

Velocity and acceleration. A particle's position is $s(t) = t^3 - 6t^2 + 9t$ (meters, t in seconds). Find its velocity and acceleration, and the acceleration when the velocity is zero.

Read.

Differentiate once for velocity, again for acceleration; solve $v = 0$ and substitute.

Plan.

Velocity: $v(t) = s'(t) = 3t^2 - 12t + 9$. Acceleration: $a(t) = v'(t) = 6t - 12$. Set $v = 0$ and factor:

$$\begin{aligned} 3t^2 - 12t + 9 &= 3(t^2 - 4t + 3) && \text{factor out } 3 \\ &= 3(t-1)(t-3) = 0 && \text{factor the quadratic} \end{aligned}$$

so $t = 1$ or $t = 3$. Then $a(1) = 6 - 12 = -6$ and $a(3) = 18 - 12 = 6$.

Solve.

$v(t) = 3t^2 - 12t + 9$, $a(t) = 6t - 12$; at rest at $t = 1$ ($a = -6$) and $t = 3$ ($a = 6$) m/s^2.

Check.

At $t = 1$ the particle is momentarily stopped and decelerating (about to reverse); at $t = 3$ stopped and accelerating forward—the signs of a tell the story. ✓

The Chain Rule and Implicit Differentiation

The single most important rule—and the one most often needed on the exam—is the *chain rule*, which differentiates a composition. If $y = f(g(x))$, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(g(x)) \cdot g'(x) : \tag{2.3}$$

differentiate the outer function at the inner one, then multiply by the derivative of the inner. In words, "derivative of the outside, leaving the inside alone, times the derivative of the inside." Almost every derivative you meet in the wild is a chain-rule problem because almost everything is a

composition.

Derivation. *Why the chain rule multiplies*

Think of y as depending on $u = g(x)$, which depends on x. Over a small change Δx, the inner variable changes by Δu and the outer by Δy. The average rates multiply exactly:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}.$$

Taking $\Delta x \to 0$ (so $\Delta u \to 0$ too) turns each ratio into a derivative, giving $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \cdot \frac{\mathrm{d}u}{\mathrm{d}x}$. The rates chain because the changes chain—a sensitivity is passed along a linkage by multiplication.

Approach Read · Plan · Solve · Check

Read. Decide whether y is given explicitly (a composition to differentiate) or implicitly (an equation to differentiate through). For a composition, identify the outer and inner functions.

Plan.

1. For a chain rule, name the inside $u = g(x)$; differentiate the outer at u, then multiply by u'.
2. For implicit differentiation, differentiate both sides in x, attaching y' to every y-term, then solve for y'.
3. For a variable base *and* exponent, take ln of both sides first, then differentiate.

Solve. Carry out the differentiation, keeping the inner derivative or the y' factor visible at each step, then isolate y' if needed.

Check. At a specific point the slope should look right — e.g. perpendicular to a circle's radius, or matching a graph's rise.

Common traps — every wrong answer is one of these

Drops the inner derivative.
Differentiates $(3x^2 + 1)^5$ to $5(3x^2 + 1)^4$ and stops, forgetting the $\cdot 6x$. The chain rule *always* multiplies by the derivative of the inside.

Forgets y' on y-terms.
In implicit differentiation, differentiates y^2 as $2y$ instead of $2y\,y'$. Every y is a function of x, so the chain rule attaches a factor y'.

Reaches for a single rule on x^x.

Applies the power rule or the exponential rule to a variable-base, variable-exponent expression. Neither applies alone — take logs first.

Mis-reads $\sin^2 x$**.** 1. Sees $\sin^2 x$ and differentiates it as $\cos^2 x$.

2. Treats the exponent as attached to the wrong function.
3. Fix: $\sin^2 x = (\sin x)^2$, so the outer is "square" and the chain rule gives $2 \sin x \cos x$.

Chain rule. Differentiate $y = (3x^2 + 1)^5$.

Read.

Name the inside $u = 3x^2 + 1$; differentiate the outer power, then multiply by u'.

Plan.

The outer function is "fifth power," with derivative $5(\cdot)^4$; the inner is $u = 3x^2 + 1$ with $u' = 6x$. So

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= 5(3x^2 + 1)^4 \cdot 6x && \text{outer power} \times \text{inner derivative} \\ &= 30x(3x^2 + 1)^4 && \text{multiply the constants} \end{aligned}$$

Solve.

$\frac{\mathrm{d}y}{\mathrm{d}x} = 30x\,(3x^2 + 1)^4$.

Check.

Forgetting the inner $6x$—the most common chain-rule error—would leave the answer off by a whole factor. The lone x at the front is the fingerprint of the inside. ✓

Chain rule with trig and a power. Differentiate $y = \sin^2 x$.

Read.

Read $\sin^2 x$ as $(\sin x)^2$; outer is "square," inner is $\sin x$.

Plan.

Outer derivative $2(\sin x)^1$, inner derivative $\cos x$:

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= 2 \sin x \cdot \cos x && \text{outer square} \times \text{inner } \cos x \\ &= \sin 2x && \text{double-angle identity (Chapter 1)} \end{aligned}$$

Solve.

$\frac{\mathrm{d}y}{\mathrm{d}x} = 2\sin x\cos x = \sin 2x.$

Check.

The notation $\sin^2 x$ hides a composition; treat the exponent as the outer "square" and the chain rule supplies the $\cos x$ automatically. ✓

When a relationship between x and y is not solved for y, we differentiate both sides with respect to x, treating y as a function of x and attaching $\frac{\mathrm{d}y}{\mathrm{d}x}$ by the chain rule (2.3) whenever we differentiate a y-term. This is *implicit differentiation*.

Implicit differentiation. Find $\frac{\mathrm{d}y}{\mathrm{d}x}$ for the circle $x^2 + y^2 = 25$.

Read.

Differentiate both sides in x; every y-term picks up a factor y'; then solve for y'.

Plan.

Differentiate term by term: $\frac{\mathrm{d}}{\mathrm{d}x}x^2 = 2x$, and $\frac{\mathrm{d}}{\mathrm{d}x}y^2 = 2y\,y'$ by the chain rule, while the constant 25 gives 0. Then solve for y':

$$\begin{aligned} 2x + 2y\,y' &= 0 && \text{differentiate both sides in } x \\ y' &= -\frac{2x}{2y} && \text{isolate the } y' \text{ term} \\ &= -\frac{x}{y} && \text{simplify} \end{aligned}$$

Solve.

$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{x}{y}.$

Check.

At $(3, 4)$ the slope is $-\frac{3}{4}$, perpendicular to the radius of slope $\frac{4}{3}$— exactly as a tangent to a circle must be. ✓

Logarithmic differentiation

When the variable appears in both a base and an exponent—as in x^x or $(\sin x)^x$—no single rule applies directly. Take the natural log of both sides first: the log laws of Chapter 1 pull the exponent down into a product, which the product and chain rules then handle.

Logarithmic differentiation. Differentiate $y = x^x$ (for $x > 0$).

Read.

Take ln of both sides, differentiate implicitly, then solve for y' and substitute $y = x^x$.

Plan.

Take logs: $\ln y = \ln(x^x) = x \ln x$. Differentiate both sides in x; the left side gives $\frac{y'}{y}$ by the chain rule, the right side by the product rule gives $(1) \ln x + x \cdot \frac{1}{x} = \ln x + 1$. Now solve for y':

$$\begin{aligned} \frac{y'}{y} &= \ln x + 1 && \text{differentiate } \ln y = x \ln x \\ y' &= y(\ln x + 1) && \text{multiply by } y \\ &= x^x(\ln x + 1) && \text{restore } y = x^x \end{aligned}$$

Solve.

$\frac{\mathrm{d}y}{\mathrm{d}x} = x^x (\ln x + 1)$.

Check.

Neither the power rule (exponent is not constant) nor the exponential rule (base is not constant) applies alone; logs convert the hybrid into a product first. ✓

Derivatives of inverse functions

An inverse function's slope is the reciprocal of the original's slope at the matching point: if $g = f^{-1}$ then $g'(y) = \frac{1}{f'(g(y))}$, because reflecting a graph across $y = x$ turns a slope m into $1/m$. Combined with implicit differentiation, this delivers the derivatives of the inverse trigonometric and logarithmic functions.

Derivative of $\arctan x$**.** Show that $\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}$.

Read.

Write $y = \arctan x$ so $\tan y = x$; differentiate implicitly and rewrite $\sec^2 y$ in terms of x.

Plan.

From $\tan y = x$, differentiate both sides in x: $\sec^2 y \cdot y' = 1$, so $y' = \frac{1}{\sec^2 y}$. Now rewrite the

denominator in terms of x:

$$\begin{aligned} \sec^2 y &= 1 + \tan^2 y && \text{Pythagorean identity} \\ &= 1 + x^2 && \text{since } \tan y = x \end{aligned}$$

Solve.

$\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}$.

Check.

The Pythagorean identity converts the answer from the angle y back to the input x—the standard move behind every inverse-trigonometric derivative. ✓

Tangent and normal lines. Find the tangent and normal lines to $y = x^3$ at the point $(1, 1)$.

Read.

The derivative gives the tangent slope; the normal is perpendicular, with slope the negative reciprocal.

Plan.

Differentiate: $y' = 3x^2$, so at $x = 1$ the tangent slope is 3. Point-slope through $(1, 1)$ gives the tangent $y - 1 = 3(x - 1)$, i.e. $y = 3x - 2$. The normal is perpendicular, so its slope is $-\frac{1}{3}$: $y - 1 = -\frac{1}{3}(x - 1)$, i.e. $y = -\frac{1}{3}x + \frac{4}{3}$.

Solve.

Tangent $y = 3x - 2$; normal $y = -\frac{1}{3}x + \frac{4}{3}$.

Check.

The slopes 3 and $-\frac{1}{3}$ multiply to -1, confirming the lines are perpendicular, as a tangent and its normal must be. ✓

Linear Approximation and Differentials

Because the tangent line hugs the curve near the point of tangency, it is the best straight-line stand-in for the function nearby. This gives the *linear approximation* (or *tangent-line approximation*):

$$f(x) \approx L(x) = f(a) + f'(a)(x - a) \qquad \text{for } x \text{ near } a. \tag{2.4}$$

The idea is the engine behind differentials, Newton's method, and the error estimates that run through applied mathematics—and it turns an awkward numerical value into arithmetic you can do by hand.

Approach Read · Plan · Solve · Check

Read. Spot the awkward value to estimate and the nearby "easy" input a where $f(a)$ and $f'(a)$ are clean.

Plan.

1. Choose a close to the target with $f(a)$ known exactly.
2. Compute $f(a)$ and $f'(a)$ and form the tangent line $L(x) = f(a) + f'(a)(x - a)$.
3. Evaluate L at the target input.

Solve. Substitute into $L(x) = f(a) + f'(a)(x - a)$ and do the arithmetic.

Check. The estimate should sit just off the true value; the closer the target is to a, the smaller the error.

Common traps — every wrong answer is one of these

Picks a far anchor.
Chooses a far from the target, so the tangent has drifted from the curve and the estimate is poor. Pick the nearest input with a clean $f(a)$.

Uses $f'(x)$ instead of $f'(a)$.
Leaves the slope as a function rather than evaluating it at the anchor. The tangent line has a *constant* slope $f'(a)$.

Confuses $x - a$ with x. 1. Writes $L(x) = f(a) + f'(a)\,x$.
2. Forgets to measure the step from the anchor.
3. Fix: the increment is $x - a$, the distance from the anchor to the target.

Estimating a root by the tangent line. Approximate $\sqrt{4.1}$ without a calculator.

Read.

Choose a nearby easy point a, build the tangent line to $\sqrt{x}$ there, and evaluate it at the target.

Plan.

Let $f(x) = \sqrt{x}$ and $a = 4$, where $f(4) = 2$ and $f'(x) = \frac{1}{2\sqrt{x}}$ gives $f'(4) = \frac{1}{4}$. The tangent

line is $L(x) = 2 + \frac{1}{4}(x - 4)$. Evaluate at $x = 4.1$:

$$\begin{aligned} L(4.1) &= 2 + \tfrac{1}{4}(0.1) && \text{substitute } x = 4.1 \\ &= 2 + 0.025 = 2.025 && \text{arithmetic} \end{aligned}$$

Solve.

$\sqrt{4.1} \approx 2.025$.

Check.

The true value is $2.02485\ldots$, so the estimate is off by about 0.0001—excellent, because 4.1 is very close to 4 where the tangent touches. ✓

Estimating a cube root. Approximate $\sqrt[3]{8.1}$ without a calculator.

Read.

Anchor at the nearest perfect cube, $a = 8$, where $f(a)$ and $f'(a)$ are clean; build the tangent line there and evaluate it at the target.

Plan.

Let $f(x) = x^{1/3}$ and $a = 8$, so $f(8) = 2$ and $f'(x) = \frac{1}{3}x^{-2/3}$ gives $f'(8) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$. The tangent line is $L(x) = 2 + \frac{1}{12}(x - 8)$. Evaluate at $x = 8.1$:

$$\begin{aligned} L(8.1) &= 2 + \tfrac{1}{12}(0.1) && \text{substitute } x = 8.1,\ x - a = 0.1 \\ &= 2 + 0.00833\ldots = 2.00833 && \text{arithmetic} \end{aligned}$$

Solve.

$\sqrt[3]{8.1} \approx 2.0083$.

Check.

The true value is $2.00830\ldots$, off by under 0.0001—the target 8.1 sits very close to the anchor 8, where the tangent still hugs the curve. A far anchor would have spoiled this. ✓

Propagated error with a differential. A sphere's radius is measured as $r = 5$ cm with a possible error of ± 0.1 cm. Use a differential to estimate the resulting error in the computed volume $V = \frac{4}{3}\pi r^3$.

Read.

A differential turns a small input error $\mathrm{d}r$ into an output error $\mathrm{d}V = V'(r)\,\mathrm{d}r$—the linear

approximation applied to the change itself.

Plan.

Differentiate the volume, then multiply by the radius error $\mathrm{d}r = 0.1$:

$$\begin{aligned} \mathrm{d}V &= \frac{\mathrm{d}V}{\mathrm{d}r}\,\mathrm{d}r && \text{differential of } V \\ &= 4\pi r^2\,\mathrm{d}r && \tfrac{\mathrm{d}}{\mathrm{d}r}\left(\tfrac{4}{3}\pi r^3\right) = 4\pi r^2 \\ &= 4\pi(5)^2(0.1) && \text{substitute } r = 5,\ \mathrm{d}r = 0.1 \\ &= 10\pi \approx 31.4 \text{ cm}^3 && \text{evaluate} \end{aligned}$$

Solve.

The volume is uncertain by about $\pm 10\pi \approx \pm 31.4$ cm^3.

Check.

The coefficient $4\pi r^2$ is the sphere's surface area, so the error reads as "surface area times radius error"—the thin shell a small radius change adds. ✓

Did You Know? *Newton's method rides the tangent line*

The same tangent-line idea (2.4), iterated, solves equations. To find a root of f, slide down the tangent at a guess x_n to where it crosses the axis; that crossing is the next guess $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Each step replaces the curve by its tangent and solves the easy linear problem, and near a simple root the number of correct digits roughly doubles per step—which is why calculators find square roots almost instantly.

L'Hôpital's Rule

Some limits resist factoring and rationalizing. When a limit has the indeterminate form $\frac{0}{0}$ or $\frac{\infty}{\infty}$, *L'Hôpital's rule* converts it into a limit of derivatives:

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)} \qquad \text{(when the right-hand limit exists and the form is indeterminate).} \tag{2.5}$$

Differentiate top and bottom *separately*—this is not the quotient rule—and try again; repeat if the form persists.

Approach Read · Plan · Solve · Check

Read. Substitute first and read off the form; L'Hôpital applies *only* when it is $\frac{0}{0}$ or $\frac{\infty}{\infty}$.

Plan.

1. Confirm the form is indeterminate.
2. Differentiate the numerator and the denominator *separately* — this is not the quotient rule.
3. Re-evaluate; if the form is still indeterminate, repeat.

Solve. Take the limit of $\frac{f'}{g'}$; substitute the target value once the form is determinate.

Check. A determinate answer should be stable — differentiating again must not change a finite limit already reached.

Common traps — every wrong answer is one of these

Applies it to a determinate form.
Uses the rule on $\frac{2}{0}$ or $\frac{3}{5}$, producing nonsense. Always verify the form is $\frac{0}{0}$ or $\frac{\infty}{\infty}$ first.

Uses the quotient rule by mistake.
Differentiates $\frac{f}{g}$ as a quotient instead of taking $\frac{f'}{g'}$. L'Hôpital differentiates top and bottom *independently*.

Stops one step short. 1. Differentiates once and substitutes.

2. Reads a still-indeterminate $\frac{0}{0}$ as the answer.
3. Fix: if the form persists, apply the rule again until it resolves.

A $\frac{0}{0}$ form. Evaluate $\lim_{x\to 0} \dfrac{\sin x}{x}$.

Read.

Confirm the form is $\frac{0}{0}$, then differentiate numerator and denominator separately.

Plan.

At $x = 0$, $\sin 0 = 0$ and the denominator is 0: the form is $\frac{0}{0}$. Differentiate top and bottom, then re-evaluate:

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} \qquad \text{differentiate top and bottom}$$

$$= \cos 0 = 1 \qquad \text{substitute } x = 0$$

Solve.

The limit is 1.

Check.

This famous limit is the reason $\frac{\mathrm{d}}{\mathrm{d}x}\sin x = \cos x$ in the first place, so using L'Hôpital here is a convenient shortcut rather than a foundation. ✓

An $\frac{\infty}{\infty}$ form. Evaluate $\lim_{x\to\infty} \dfrac{\ln x}{x}$.

Read.

The form is $\frac{\infty}{\infty}$; differentiate top and bottom and re-evaluate.

Plan.

As $x \to \infty$, both $\ln x \to \infty$ and $x \to \infty$: the form is $\frac{\infty}{\infty}$. Differentiate: $\frac{\mathrm{d}}{\mathrm{d}x}\ln x = \frac{1}{x}$ and $\frac{\mathrm{d}}{\mathrm{d}x}x = 1$, then re-evaluate:

$$\begin{aligned} \lim_{x\to\infty} \frac{\ln x}{x} &= \lim_{x\to\infty} \frac{1/x}{1} && \text{differentiate top and bottom} \\ &= \lim_{x\to\infty} \frac{1}{x} = 0 && \text{simplify and let } x \to \infty \end{aligned}$$

Solve.

The limit is 0.

Check.

This encodes a hierarchy the exam relies on: logarithms grow slower than any positive power of x, which in turn grows slower than any exponential. ✓

A form that needs two rounds. Evaluate $\lim_{x\to 0} \dfrac{1-\cos x}{x^2}$.

Read.

Confirm $\frac{0}{0}$, differentiate top and bottom, and—if the form persists—apply the rule a second time rather than stopping.

Plan.

At $x = 0$, $1 - \cos 0 = 0$ and $x^2 = 0$: the form is $\frac{0}{0}$. Differentiate top and bottom:

$$\begin{aligned} \lim_{x\to 0} \frac{1-\cos x}{x^2} &= \lim_{x\to 0} \frac{\sin x}{2x} && \text{differentiate top and bottom} \\ &= \lim_{x\to 0} \frac{\cos x}{2} && \text{still } \tfrac{0}{0}\text{, so apply the rule again} \\ &= \frac{\cos 0}{2} = \frac{1}{2} && \text{substitute } x = 0 \end{aligned}$$

Solve.

The limit is $\frac{1}{2}$.

Check.

Stopping after one round at $\frac{\sin x}{2x}$ leaves a still-indeterminate $\frac{0}{0}$; the second application resolves it. The answer $\frac{1}{2}$ also confirms the small-angle estimate $1 - \cos x \approx \frac{1}{2}x^2$. ✓

Misconception

L'Hôpital only for indeterminate forms. The rule applies *only* to $\frac{0}{0}$ or $\frac{\infty}{\infty}$. Applying it to a determinate form such as $\frac{2}{0}$ or $\frac{3}{5}$ gives nonsense. Always check the form before differentiating.

The Mean Value Theorem

The *Mean Value Theorem* (MVT) is the bridge between a function's average rate of change and its instantaneous one. If f is continuous on $[a, b]$ and differentiable on (a, b), then there is at least one point c in (a, b) where

$$f'(c) = \frac{f(b) - f(a)}{b - a}\,: \tag{2.6}$$

somewhere inside the interval, the instantaneous slope equals the average slope over the whole interval. The special case $f(a) = f(b)$ is *Rolle's theorem*: between two equal values the derivative must vanish somewhere. The MVT underwrites much of what we take for granted—for instance, that $f' > 0$ throughout an interval forces f to increase there.

Approach Read · Plan · Solve · Check

Read. Confirm the hypotheses — f continuous on $[a, b]$ and differentiable on (a, b) — then note what the theorem promises: a point c where the slopes match.

Plan.

1. Compute the average slope $\frac{f(b)-f(a)}{b-a}$.
2. Set $f'(c)$ equal to it.
3. Solve for c and keep only roots inside (a, b).

Solve. Differentiate, equate $f'(c)$ to the average slope, and solve for c.

Check. The c you find must lie strictly inside (a, b); geometrically the tangent there is parallel to the secant.

Common traps — every wrong answer is one of these

Skips the hypotheses.
Applies the MVT where f has a corner or break in the interval. The theorem needs continuity on $[a, b]$ *and* differentiability on (a, b).

Accepts a c outside the interval.
Solves $f'(c) =$ average slope but keeps a root outside (a, b). Only interior solutions count.

Confuses average and instantaneous. 1. Sets $f'(c)$ equal to $f(b) - f(a)$.
2. Forgets to divide by $b - a$.
3. Fix: the average slope is $\frac{f(b)-f(a)}{b-a}$, a rise *over* a run.

Locating the guaranteed point. Find the value c promised by the MVT for $f(x) = x^2$ on $[1, 3]$.

Read.

Compute the average slope, set $f'(c)$ equal to it, and solve for c in the interval.

Plan.

Compute the average slope over the interval:

$$\frac{f(3) - f(1)}{3 - 1} = \frac{9 - 1}{2} \qquad \text{evaluate the endpoints}$$

$$= 4 \qquad \text{average slope}$$

Since $f'(x) = 2x$, set $2c = 4$, giving $c = 2$, which lies in $(1, 3)$.

Solve.

$c = 2$.

Check.

At $x = 2$ the tangent to the parabola is exactly parallel to the secant joining the endpoints $(1, 1)$ and $(3, 9)$—the MVT guarantees such a point exists. ✓

Rejecting a root outside the interval. Find every value c promised by the MVT for $f(x) = x^3$ on $[0, 2]$.

Read.

Compute the average slope, set $f'(c)$ equal to it, and keep only the solution lying inside $(0, 2)$.

Plan.

Compute the average slope over the interval:

$$\begin{aligned} \frac{f(2) - f(0)}{2 - 0} &= \frac{8 - 0}{2} && \text{evaluate the endpoints} \\ &= 4 && \text{average slope} \end{aligned}$$

Since $f'(x) = 3x^2$, set $3c^2 = 4$:

$$\begin{aligned} c^2 &= \frac{4}{3} && \text{isolate } c^2 \\ c &= \pm\frac{2}{\sqrt{3}} && \text{take square roots} \end{aligned}$$

Only $c = \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3} \approx 1.15$ lies in $(0, 2)$; the negative root is discarded.

Solve.

$c = \frac{2\sqrt{3}}{3} \approx 1.15$.

Check.

The theorem promises a point *inside* the interval, so the extraneous $c = -\frac{2}{\sqrt{3}}$ cannot count—keeping it would place the tangent outside $[0, 2]$ altogether. ✓

When the hypotheses fail. Explain why the MVT produces no valid c for $f(x) = \frac{1}{x}$ on $[-1, 1]$.

Read.

Before applying the theorem, test its hypotheses; a break in the interval voids the guarantee, and the slope equation then need not have any solution.

Plan.

The naive average slope would be

$$\frac{f(1)-f(-1)}{1-(-1)} = \frac{1-(-1)}{2} \qquad \text{evaluate the endpoints}$$

$$= 1 \qquad \text{average slope}$$

Yet $f'(x) = -\frac{1}{x^2} < 0$ for every x, so $f'(c) = 1$ has *no* solution.

Solve.

No such c exists—and the MVT is not violated, because $f(x) = \frac{1}{x}$ is discontinuous at $x = 0 \in [-1, 1]$, so the hypotheses fail.

Check.

The theorem guarantees a matching point only when f is continuous on $[a, b]$ *and* differentiable on (a, b). The blow-up at 0 breaks continuity, and with it the promise— always check the hypotheses first. ✓

Connection

Why increasing tests work. The sign tests of the next section rest on the MVT (2.6): if $f' > 0$ on an interval, then for any $a < b$ there the MVT gives $f(b) - f(a) = f'(c)(b - a) > 0$, so $f(b) > f(a)$—f increases. The intuitive "positive slope means rising" is a theorem, not merely a picture.

Related Rates and Optimization

Two families of word problems are pure derivative applications. In a *related-rates* problem, several quantities change with time and are linked by an equation; differentiating that equation with respect to time relates their rates. In an *optimization* problem, we seek the largest or smallest value of a quantity, which occurs where its derivative is zero (a *critical point*) or at an endpoint.

Approach Read · Plan · Solve · Check

Read. Sort the problem: a related-rates question links changing quantities through time; an optimization question seeks a largest or smallest value. Name the variables and list what is known.

Plan.

1. Write an equation relating the quantities (geometry, volume, area).
2. For related rates, differentiate in t and substitute the instantaneous values *last*. For optimization, reduce to one variable, differentiate, and set the derivative to zero.
3. Solve for the requested rate or the critical point.

Solve. Differentiate the relation, substitute, and solve; for optimization confirm max vs. min.

Check. Signs and size should make sense — a shrinking quantity gives a negative rate, and a second-derivative or endpoint test confirms an optimum.

Common traps — every wrong answer is one of these

Substitutes numbers too early.
Plugs the instantaneous values in *before* differentiating, freezing a variable that is actually changing. Differentiate the general relation first, then substitute.

Leaves two variables in optimization.
Differentiates an area or cost still written in two variables. Use the constraint to reduce to one variable before differentiating.

Skips the max-vs-min test.
Finds a critical point and assumes it is the extremum sought. Confirm with the second derivative or by checking endpoints.

Ignores the sign of a rate. 1. Computes a rate and reports its magnitude.
2. Drops the negative sign on a decreasing quantity.
3. Fix: keep the sign — a negative rate means the quantity is shrinking (falling, melting, emptying).

Problem-solving strategy — Related rates

(1) Draw and name the changing quantities. (2) Write an equation relating them. (3) Differentiate both sides with respect to time t. (4) Substitute the known instantaneous values *last*, after differentiating—never before.

Related rates. A ladder 10 ft long leans on a wall; its base slides out at 2 ft/s. How fast is the top sliding down when the base is 6 ft from the wall?

Read.

Relate base x and height y by the Pythagorean theorem, differentiate in t, then substitute.

Plan.

Let x be the base distance and y the height, with $x^2 + y^2 = 10^2 = 100$. At $x = 6$, $y = \sqrt{100 - 36} = 8$. Now differentiate the relation in t and substitute:

$$\begin{aligned} 2x\frac{\mathrm{d}x}{\mathrm{d}t} + 2y\frac{\mathrm{d}y}{\mathrm{d}t} &= 0 && \text{differentiate } x^2 + y^2 = 100 \text{ in } t \\ 2(6)(2) + 2(8)\frac{\mathrm{d}y}{\mathrm{d}t} &= 0 && \text{substitute } x = 6,\ y = 8,\ \tfrac{\mathrm{d}x}{\mathrm{d}t} = 2 \\ 24 + 16\frac{\mathrm{d}y}{\mathrm{d}t} &= 0 && \text{simplify} \end{aligned}$$

Solve.

$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{24}{16} = -1.5$ ft/s (the top slides *down* at 1.5 ft/s).

Check.

The negative sign correctly reports downward motion, and the top moving slower than the base ($1.5 < 2$) is reasonable at this shallow angle. ✓

Related rates with a volume. A spherical snowball melts so its volume decreases at 8 cm^3/min. How fast is the radius shrinking when $r = 2$ cm? ($V = \frac{4}{3}\pi r^3$.)

Read.

Differentiate the volume–radius relation in t, then substitute the known rate and radius.

Plan.

Differentiate the volume–radius relation, then substitute the known rate and radius:

$$\begin{aligned} \frac{\mathrm{d}V}{\mathrm{d}t} &= 4\pi r^2\frac{\mathrm{d}r}{\mathrm{d}t} && \text{differentiate } V = \tfrac{4}{3}\pi r^3 \text{ in } t \\ -8 &= 4\pi(2)^2\frac{\mathrm{d}r}{\mathrm{d}t} && \text{substitute } \tfrac{\mathrm{d}V}{\mathrm{d}t} = -8,\ r = 2 \\ &= 16\pi\frac{\mathrm{d}r}{\mathrm{d}t} && \text{simplify the coefficient} \end{aligned}$$

Solve.

$\frac{\mathrm{d}r}{\mathrm{d}t} = -\frac{8}{16\pi} = -\frac{1}{2\pi}$ cm/min.

Check.

The negative sign shows the radius shrinking; the factor $4\pi r^2$ is the surface area, so the relation says volume changes at "area times how fast the radius moves." ✓

Problem-solving strategy — Optimization

(1) Identify the quantity to maximize or minimize and write it as a formula. (2) Use a constraint to reduce it to *one* variable. (3) Differentiate, set to zero, and solve for critical points. (4) Confirm max vs. min with the second derivative or endpoints.

Optimization against a wall. A rancher fences a rectangular pen against a river (no fence on the river side) with 100 m of fencing. What dimensions maximize the area?

Read.

Express area in one variable using the fencing constraint, set the derivative to zero, and check it is a maximum.

Plan.

Let the two sides perpendicular to the river be x and the side parallel be y. The fencing constraint is $2x + y = 100$, so $y = 100 - 2x$. Write the area in one variable:

$$\begin{aligned} A(x) &= xy && \text{area of the pen} \\ &= x(100 - 2x) && \text{use } y = 100 - 2x \\ &= 100x - 2x^2 && \text{expand} \end{aligned}$$

Differentiate and set to zero: $A'(x) = 100 - 4x = 0$, so $x = 25$, hence $y = 100 - 50 = 50$. Since $A''(x) = -4 < 0$, this is a maximum.

Solve.

25 m $\times$ 50 m, giving maximum area 1250 m^2.

Check.

The optimal design spends half the fence on the single long side—a recurring pattern in these "fence against a wall" problems. ✓

Least material for a can. A closed cylindrical can must hold a fixed volume V. Show that the surface area is smallest when the height equals the diameter, $h = 2r$.

Read.

Write area with the volume constraint eliminating h, differentiate in r, and interpret the critical point.

Plan.

Surface area $S = 2\pi r^2 + 2\pi rh$; the constraint $V = \pi r^2 h$ gives $h = \frac{V}{\pi r^2}$. Substitute to get S

in one variable:

$$\begin{aligned} S(r) &= 2\pi r^2 + 2\pi r \cdot \frac{V}{\pi r^2} && \text{substitute } h = \tfrac{V}{\pi r^2} \\ &= 2\pi r^2 + \frac{2V}{r} && \text{simplify} \end{aligned}$$

Differentiate and set to zero: $S'(r) = 4\pi r - \frac{2V}{r^2} = 0$, so $4\pi r^3 = 2V$, i.e. $V = 2\pi r^3$. Then substitute back for h:

$$\begin{aligned} h &= \frac{V}{\pi r^2} && \text{from the constraint} \\ &= \frac{2\pi r^3}{\pi r^2} = 2r && \text{use } V = 2\pi r^3 \end{aligned}$$

Solve.

The minimum occurs at $h = 2r$ (height equals diameter).

Check.

$S''(r) = 4\pi + \frac{4V}{r^3} > 0$ confirms a minimum; the "square" proportions of an optimal can are why many real cans are close to this shape. ✓

Curve Sketching: What f' and f'' Reveal

The two derivatives narrate a graph's shape. The first derivative f' gives the *slope*: where $f' > 0$ the function rises, where $f' < 0$ it falls, and where $f' = 0$ it has a horizontal tangent—a candidate peak or valley. The second derivative f'' gives the *concavity*: where $f'' > 0$ the graph curves upward (holds water), where $f'' < 0$ it curves downward, and a point where concavity switches is an *inflection point*. Together they let you sketch a curve without plotting a table of points.

First- and second-derivative tests. At a critical point c where $f'(c) = 0$: if $f''(c) > 0$, c is a local minimum; if $f''(c) < 0$, a local maximum. Equivalently, c is a maximum if f' changes $+ \to -$ there, a minimum if $- \to +$.

Approach Read · Plan · Solve · Check

Read. Recall the division of labor: f' controls rise/fall and locates critical points; f'' controls concavity and locates inflection points.

Plan.

1. Set $f' = 0$ to find critical points; test the sign of f' on each interval for increase/decrease.
2. Use f'' at each critical point (or a sign change of f') to classify it as a max or min.
3. Set $f'' = 0$ and check for a concavity switch to locate inflection points.

Solve. Differentiate twice, solve $f' = 0$ and $f'' = 0$, and read the signs to assemble the shape.

Check. The picture must be consistent — a local max sits above its neighbors, and concavity should flip exactly at each inflection point.

Common traps — every wrong answer is one of these

Calls every $f' = 0$ an extremum.
Assumes a horizontal tangent is always a peak or valley. If f' does not change sign there (as at $x = 0$ for x^3), it is not an extremum.

Calls every $f'' = 0$ an inflection.
Reports an inflection wherever $f'' = 0$ without checking that concavity actually switches. The sign of f'' must change across the point.

Mixes up the two tests. 1. Uses f'' to find where f rises.
2. Or uses f' to judge concavity.
3. Fix: f' governs increase/decrease and extrema; f'' governs concavity and inflection.

Classifying critical points. Find and classify the critical points of $f(x) = x^3 - 3x$.

Read.

Set $f' = 0$ to locate critical points, then use the sign of f'' to classify each.

Plan.

Differentiate and factor to locate the critical points:

$$\begin{aligned} f'(x) &= 3x^2 - 3 && \text{power rule} \\ &= 3(x^2 - 1) && \text{factor} \end{aligned}$$

so $f' = 0$ at $x = \pm 1$. Second derivative: $f''(x) = 6x$. At $x = 1$, $f''(1) = 6 > 0$, a local minimum; at $x = -1$, $f''(-1) = -6 < 0$, a local maximum.

Solve.

Local max at $(-1, 2)$, local min at $(1, -2)$; inflection where $f'' = 0$, at $x = 0$.

Check.

A cubic with positive leading coefficient falling then rising then rising fits one max then one min, reading left to right. ✓

A full sketch. Analyze $f(x) = x^3 - 6x^2 + 9x + 1$: find its increase/decrease intervals, local extrema, and inflection point.

Read.

Use f' for critical points and monotonicity, f'' for concavity and inflection, then assemble.

Plan.

Differentiate and factor:

$$\begin{aligned} f'(x) &= 3x^2 - 12x + 9 && \text{power rule} \\ &= 3(x^2 - 4x + 3) && \text{factor out 3} \\ &= 3(x-1)(x-3) && \text{factor the quadratic} \end{aligned}$$

so $f' = 0$ at $x = 1$ and $x = 3$. The sign of f' is $+$ on $(-\infty, 1)$, $-$ on $(1, 3)$, and $+$ on $(3, \infty)$. So f rises, falls, rises: a local max at $x = 1$ (value $f(1) = 5$) and a local min at $x = 3$ (value $f(3) = 1$). Second derivative $f''(x) = 6x - 12 = 0$ at $x = 2$; concave down before, up after, so $x = 2$ is an inflection point (value $f(2) = 3$).

Solve.

Increasing on $(-\infty, 1) \cup (3, \infty)$, decreasing on $(1, 3)$; local max $(1, 5)$, local min $(3, 1)$; inflection $(2, 3)$.

Check.

The inflection $x = 2$ sits exactly halfway between the two critical points—a symmetry every cubic shares. ✓

A critical point that is not an extremum. For $f(x) = x^3 - 3x^2 + 3x$, locate the critical point and decide whether it is a maximum, a minimum, or neither.

Read.

Find where $f' = 0$, but do not assume a horizontal tangent is a peak or valley—test whether f' actually changes sign.

Plan.

Differentiate and factor:

$$\begin{aligned} f'(x) &= 3x^2 - 6x + 3 && \text{power rule} \\ &= 3(x^2 - 2x + 1) && \text{factor out } 3 \\ &= 3(x-1)^2 && \text{perfect square} \end{aligned}$$

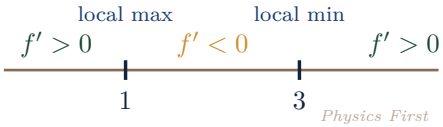
So $f' = 0$ only at $x = 1$. But $(x-1)^2 \geq 0$ makes $f' \geq 0$ *everywhere*: the slope never changes sign, and f climbs through $x = 1$ without turning. For concavity, $f''(x) = 6x - 6 = 0$ at $x = 1$, negative before and positive after—a genuine sign change.

Solve.

At $x = 1$ there is a horizontal tangent but *no* extremum; instead $(1, 1)$ is an inflection point that happens to have a horizontal tangent.

Check.

This is just $f(x) = (x-1)^3 + 1$, a cubic flattened and shifted. A horizontal tangent is only a *candidate* extremum—without a sign change in f' it is not one. ✓



Sign chart of $f'(x) = 3(x-1)(x-3)$. The derivative is positive, then negative, then positive; the function rises to a local maximum at $x = 1$, falls to a local minimum at $x = 3$, then rises.



Concavity at a glance. Where $f'' > 0$ the curve bends upward like a cup (holds water); where $f'' < 0$ it bends downward like a cap. A switch between the two is an inflection point.

The Integral and the Fundamental Theorem

The derivative has an inverse operation. An *antiderivative* of f is a function F with $F' = f$; the family of all antiderivatives is the *indefinite integral* $\int f(x)\,dx = F(x) + C$. Separately, the *definite*

integral $\int_a^b f(x)\,\mathrm{d}x$ is defined as a limit of sums approximating the signed area under the graph between a and b. The astonishing fact—the *Fundamental Theorem of Calculus*—is that these two ideas, slopes-undone and areas, are the same:

Fundamental Theorem of Calculus. If $F' = f$ and f is continuous on $[a, b]$, then

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a). \tag{2.7}$$

To integrate, reverse the power rule: $\displaystyle\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} + C$ for $n \neq -1$.

Derivation. *Why area and antiderivative coincide*

Let $A(x) = \int_a^x f(t)\,\mathrm{d}t$ be the area accumulated up to x. Increasing x by a tiny h adds a thin strip of area roughly $f(x) \cdot h$ (a rectangle of height $f(x)$ and width h), so $\frac{A(x+h)-A(x)}{h} \approx f(x)$. Letting $h \to 0$ gives $A'(x) = f(x)$: the area function is an antiderivative of f. Hence for any antiderivative F, $\int_a^b f = A(b) - A(a) = F(b) - F(a)$. Accumulated area and undone slope are literally the same function.

Approach Read · Plan · Solve · Check

Read. Identify the task: an indefinite integral (find an antiderivative $+C$), a definite integral (a number via the FTC), an area, or a net-change/accumulation question.

Plan.

1. Find an antiderivative F by reversing the power rule (or by substitution when a function and its derivative both appear).
2. For a definite integral, evaluate $F(b) - F(a)$.
3. For area, integrate top minus bottom between the intersection points; for net change, integrate the rate.

Solve. Antidifferentiate, then evaluate at the limits and subtract in the order $F(b) - F(a)$.

Check. Differentiate the antiderivative — it must return the integrand — and confirm an area came out positive.

Common traps — every wrong answer is one of these

Drops the $+C$.
Omits the constant of integration on an indefinite integral. Every antiderivative carries an arbitrary $+C$; an initial condition then fixes it.

Subtracts in the wrong order.
Computes $F(a) - F(b)$, or integrates "bottom minus top," flipping the sign. It is always $F(b) - F(a)$ and "top minus bottom."

Confuses displacement and distance.
Integrates v and calls the result total distance. $\int v$ is net displacement; total distance needs $\int |v|$, splitting at the sign changes of v.

Forgets the chain factor on a variable limit. 1. Differentiates $\int_a^{u(x)} f(t)\,dt$ as just $f(u)$.
2. Forgets that the upper limit moves.
3. Fix: multiply by $u'(x)$ — the FTC combines with the chain rule.

A definite integral via the FTC. Evaluate $\displaystyle\int_0^2 (3x^2 + 2x)\,dx$.

Read.

Find an antiderivative F by reversing the power rule, then compute $F(2) - F(0)$.

Plan.

Antiderivative: $\int (3x^2 + 2x)\,dx = x^3 + x^2$ (check: its derivative is $3x^2 + 2x$). Evaluate at the endpoints and subtract:

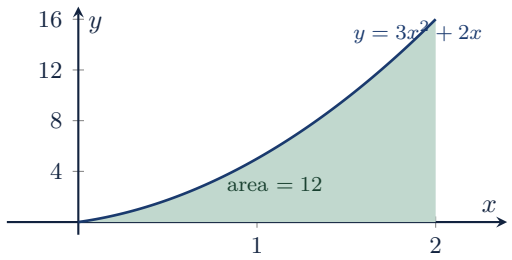
$$\begin{aligned} F(2) &= 2^3 + 2^2 = 8 + 4 = 12 && \text{evaluate at the upper limit} \\ F(0) &= 0 && \text{evaluate at the lower limit} \\ F(2) - F(0) &= 12 - 0 = 12 && \text{subtract} \end{aligned}$$

Solve.

$$\int_0^2 (3x^2 + 2x)\,dx = 12.$$

Check.

Differentiating the antiderivative returns the integrand, which is the quickest check that F was found correctly. ✓



The definite integral as signed area: $\int_0^2 (3x^2 + 2x)\,\mathrm{d}x$ is the shaded area under the curve, and the Fundamental Theorem evaluates it as $F(2) - F(0) = 12$.

Integration by substitution

Substitution is the chain rule (2.3) run backward. If an integrand contains a function and (a constant multiple of) its own derivative, let u be the inside function; then $\mathrm{d}u = u'\,\mathrm{d}x$ converts the whole integral into a clean power or standard form in u.

A u-substitution. Evaluate $\displaystyle\int 2x\,(x^2+1)^3\,\mathrm{d}x$.

Read.

Let u be the inside function; replace $2x\,\mathrm{d}x$ by $\mathrm{d}u$; integrate in u; return to x.

Plan.

Let $u = x^2 + 1$, so $\mathrm{d}u = 2x\,\mathrm{d}x$—exactly the $2x\,\mathrm{d}x$ present. The integral becomes $\int u^3\,\mathrm{d}u = \frac{u^4}{4} + C$. Restore $u = x^2 + 1$.

Solve.

$$\int 2x(x^2+1)^3\,\mathrm{d}x = \frac{(x^2+1)^4}{4} + C.$$

Check.

Differentiate the answer: the chain rule gives $\frac{1}{4} \cdot 4(x^2+1)^3 \cdot 2x = 2x(x^2+1)^3$, the original integrand. ✓

Area between two curves

The area of the region enclosed between $y = f(x)$ (upper) and $y = g(x)$ (lower) from $x = a$ to $x = b$ is $\int_a^b [f(x) - g(x)]\,\mathrm{d}x$: integrate the vertical gap. The endpoints are usually the curves' intersection points, found by setting $f = g$.

Between a line and a parabola. Find the area enclosed by $y = 2x$ and $y = x^2$.

Read.

Find the intersections for the limits, decide which curve is on top, then integrate the gap.

Plan.

Set $2x = x^2$: $x^2 - 2x = 0$, so $x = 0$ and $x = 2$ are the limits. On $(0, 2)$ the line $2x$ lies above the parabola x^2 (test $x = 1$: $2 > 1$). Integrate the gap:

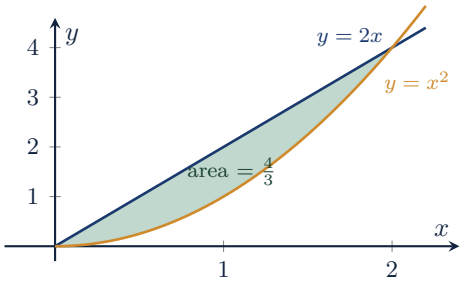
$$\begin{aligned} \int_0^2 (2x - x^2)\,\mathrm{d}x &= \left[x^2 - \tfrac{x^3}{3}\right]_0^2 && \text{antidifferentiate} \\ &= \left(4 - \tfrac{8}{3}\right) - 0 && \text{evaluate at the bounds} \\ &= \tfrac{12-8}{3} = \tfrac{4}{3} && \text{combine} \end{aligned}$$

Solve.

The area is $\frac{4}{3}$.

Check.

Always integrate "top minus bottom"; reversing them yields the negative of the area, a sign slip the exam rewards. ✓



Area between $y = 2x$ and $y = x^2$ from their intersections $x = 0$ and $x = 2$: the shaded gap integrates to $\frac{4}{3}$.

Net change and displacement

Because the integral of a rate is the total accumulated change, $\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a)$. For motion this means $\int_a^b v(t)\,\mathrm{d}t$ is *displacement* (net, allowing cancellation), while $\int_a^b |v(t)|\,\mathrm{d}t$ is *total distance* traveled. Distinguishing the two is a favorite exam point.

Displacement vs. distance. A particle has velocity $v(t) = t^2 - 4$ m/s on $[0, 3]$. Find its displacement and the total distance traveled.

Read.

Integrate v for displacement; split at the sign change of v and integrate $|v|$ for distance.

Plan.

Displacement is the integral of v:

$$\begin{aligned} \int_0^3 (t^2 - 4)\,\mathrm{d}t &= \left[\tfrac{t^3}{3} - 4t\right]_0^3 && \text{antidifferentiate} \\ &= (9 - 12) - 0 = -3 \text{ m} && \text{evaluate the endpoints} \end{aligned}$$

For distance, $v = 0$ at $t = 2$; $v < 0$ on $(0, 2)$ and $v > 0$ on $(2, 3)$, so integrate $|v|$ on each piece:

$$\begin{aligned} \int_0^2 (4 - t^2)\,\mathrm{d}t &= \left[4t - \tfrac{t^3}{3}\right]_0^2 && \text{antiderivative on } (0, 2) \\ &= 8 - \tfrac{8}{3} = \tfrac{16}{3} && \text{evaluate} \\ \int_2^3 (t^2 - 4)\,\mathrm{d}t &= \left[\tfrac{t^3}{3} - 4t\right]_2^3 && \text{antiderivative on } (2, 3) \\ &= (9 - 12) - (\tfrac{8}{3} - 8) = \tfrac{7}{3} && \text{evaluate} \end{aligned}$$

Add: $\frac{16}{3} + \frac{7}{3} = \frac{23}{3}$.

Solve.

Displacement -3 m; total distance $\frac{23}{3} \approx 7.67$ m.

Check.

Distance exceeds the magnitude of displacement because the particle backtracks—moving $-\frac{16}{3}$ then $+\frac{7}{3}$, netting -3. ✓

Recovering position from velocity. A particle has velocity $v(t) = 6t - 4$ m/s and starts at position $s(0) = 5$ m. Find $s(t)$.

Read.

Antidifferentiate v to get s up to a constant, then use the initial condition to fix the constant.

Plan.

Antidifferentiate: $s(t) = \int (6t - 4)\,\mathrm{d}t = 3t^2 - 4t + C$. Apply the initial condition $s(0) = 5$:

$3(0)^2 - 4(0) + C = 5$, so $C = 5$.

Solve.

$s(t) = 3t^2 - 4t + 5$.

Check.

Differentiating back gives $s'(t) = 6t - 4 = v(t)$ and $s(0) = 5$, so both conditions hold—the constant C is exactly the starting position. ✓

The average value of a function

The *average value* of f over $[a, b]$ is $\bar{f} = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x$—the constant height whose rectangle has the same area as the region under f. It is the continuous analogue of averaging a list of numbers: instead of summing finitely many values and dividing by the count, we integrate the values and divide by the length of the interval.

Average value. Find the average value of $f(x) = x^2$ on $[0, 3]$.

Read.

Integrate over the interval, then divide by its length.

Plan.

Integrate over the interval:

$$\int_0^3 x^2\,\mathrm{d}x = \left[\tfrac{x^3}{3}\right]_0^3 \qquad \text{antidifferentiate}$$

$$= 9 \qquad \text{evaluate the endpoints}$$

Divide by the length $3 - 0 = 3$: $\bar{f} = \frac{9}{3} = 3$.

Solve.

The average value is 3.

Check.

Since f climbs from 0 to 9 across $[0, 3]$, an average of 3—below the midpoint height 4.5—is right, because the parabola spends more of the interval at small values. ✓

The derivative of an accumulation

The other half of the Fundamental Theorem (2.7) says differentiation undoes accumulation: $\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x)$. When the upper limit is itself a function of x, the chain rule tacks on its derivative.

Differentiating an integral with a variable limit. Compute $\frac{\mathrm{d}}{\mathrm{d}x}\int_2^{x^3} \sin t\,\mathrm{d}t$.

Read.

Apply the FTC (Part 1): substitute the upper limit into the integrand, then multiply by the derivative of that limit.

Plan.

By the Fundamental Theorem, differentiating $\int_2^u \sin t\,\mathrm{d}t$ gives $\sin u$; here $u = x^3$, so the chain rule multiplies by $\frac{\mathrm{d}}{\mathrm{d}x}x^3 = 3x^2$. Thus the derivative is $\sin(x^3)\cdot 3x^2$.

Solve.

$\frac{\mathrm{d}}{\mathrm{d}x}\int_2^{x^3} \sin t\,\mathrm{d}t = 3x^2\sin(x^3)$.

Check.

The constant lower limit 2 never enters the derivative—only the moving upper limit contributes, through the chain rule. ✓

Synthesis — Key Results You Must Own

1. A limit is the value outputs approach; resolve $\frac{0}{0}$ by factoring or rationalizing, and limits at infinity by dividing by the highest power.
2. $f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$ is slope, velocity, and sensitivity in one; it is the limit of secant slopes.
3. Own the power, product, quotient, and—above all—chain rules; most derivatives are compositions. Higher derivatives give acceleration and concavity.
4. Implicit differentiation attaches y' to every y-term; logarithmic differentiation tames a variable base with a variable exponent.
5. The tangent line $L(x) = f(a) + f'(a)(x-a)$ approximates f near a; L'Hôpital resolves $\frac{0}{0}$ and $\frac{\infty}{\infty}$; the MVT equates average and instantaneous slope.

6. Related rates: differentiate the linking equation in t, substitute values last. Optimization: reduce to one variable, set the derivative to zero, confirm with f''.
7. f' gives increase/decrease and critical points; f'' gives concavity and inflection; together they sketch the curve.
8. Fundamental Theorem: $\int_a^b f = F(b) - F(a)$ when $F' = f$; substitution reverses the chain rule; $\int(\text{top} - \text{bottom})$ is area between curves; $\int v$ is displacement, $\int |v|$ is distance.

Practice Problems

Difficulty: ⋆ warm-up, ⋆⋆ standard, ⋆⋆⋆ challenge. Work each with the Read–Plan–Solve–Check habit modeled above.

Limits and continuity

1 ⋆.

Evaluate $\lim_{x \to 2} \dfrac{x^2 - 4}{x - 2}$.

2 ⋆⋆.

Evaluate $\lim_{x \to 0} \dfrac{\sqrt{x+9} - 3}{x}$.

Hint: Rationalize the numerator.

3 ⋆⋆.

For what value of k is $f(x) = \begin{cases} x^2 & x \le 1 \\ kx - 1 & x > 1 \end{cases}$ continuous at $x = 1$?

4 ⋆⋆⋆.

Evaluate $\lim_{x \to \infty} \dfrac{3x^2 - x}{2x^2 + 5}$ and justify.

Derivatives and rules

5 ⋆.

Differentiate $f(x) = 5x^4 - 2x^2 + 7$.

6 ⋆.

Use the limit definition to find $f'(x)$ for $f(x) = \frac{1}{x}$.

7 ⋆⋆.

Differentiate $h(x) = x^3 \ln x$.

8 ⋆⋆.

Differentiate $g(x) = \frac{\sin x}{x}$.

9 ⋆⋆.

Find the equation of the tangent line to $y = x^2 - 4x$ at $x = 1$.

Chain rule and implicit

10 ⋆.

Differentiate $y = \sin(3x)$.

11 ⋆⋆.

Differentiate $y = \mathrm{e}^{x^2}$.

12 ⋆⋆.

Differentiate $y = \ln(x^2 + 1)$.

13 ⋆⋆.

Find $\frac{\mathrm{d}y}{\mathrm{d}x}$ for $x^3 + y^3 = 6xy$ at the point $(3, 3)$.

14 ⋆⋆⋆.

Differentiate $y = (\sin x)^x$.

Hint: Take ln of both sides first (logarithmic differentiation).

Related rates and optimization

15 ⋆⋆.

A spherical balloon is inflated at 12 cm^3/s. How fast is its radius growing when $r = 3$ cm? ($V = \frac{4}{3}\pi r^3$.)

16 ⋆⋆.

Find the dimensions of the rectangle of largest area inscribed with its base on the x-axis under $y = 4 - x^2$.

17 ⋆⋆.

A box with a square base and open top must hold 32 m^3. Find the dimensions minimizing the material used.

18 ⋆⋆⋆.

A right triangle has its hypotenuse of fixed length L. Show its area is greatest when it is isosceles.

Curve sketching

19 ⋆.

On what intervals is $f(x) = x^3 - 3x^2$ increasing?

20 ⋆⋆.

Find and classify the critical points of $f(x) = x^4 - 4x^3$.

21 ⋆⋆.

Find the inflection point(s) of $f(x) = x^3 - 6x^2 + 9x$.

22 ⋆⋆⋆.

For $f(x) = \frac{x}{x^2+1}$, find all local extrema and the horizontal asymptote.

Integration and the FTC

23 ⋆.

Evaluate $\displaystyle\int (4x^3 - 6x)\,\mathrm{d}x$.

24 ⋆.

Evaluate $\displaystyle\int_1^3 2x\,\mathrm{d}x$.

25 ⋆⋆.

Evaluate $\displaystyle\int_0^{\pi} \sin x\,\mathrm{d}x$.

26 ⋆⋆.

Evaluate $\displaystyle\int_1^{\mathrm{e}} \frac{1}{x}\,\mathrm{d}x$.

27 ⋆⋆.

Find the area between $y = x^2$ and $y = x$ from $x = 0$ to $x = 1$.

28 ⋆⋆⋆.

Evaluate $\displaystyle\int_0^1 x\,\mathrm{e}^{\,x^2}\,\mathrm{d}x$.

Hint: Substitute $u = x^2$.

29 ⋆⋆⋆.

If $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_0^x g(t)\,\mathrm{d}t = g(x)$, evaluate $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_0^{x^2} \cos t\,\mathrm{d}t$.

30 ⋆⋆⋆.

A particle has velocity $v(t) = 3t^2 - 12t + 9$ m/s. Find its total displacement on $[0, 4]$ and the times it is momentarily at rest.

Further Reading

For the limit foundations and the ε–δ definition behind this chapter's intuitive limits, a rigorous single-variable calculus text repays study and prepares the real-analysis material of a later chapter. Chapter 3 continues with the techniques and applications of integration.

CHAPTER 3

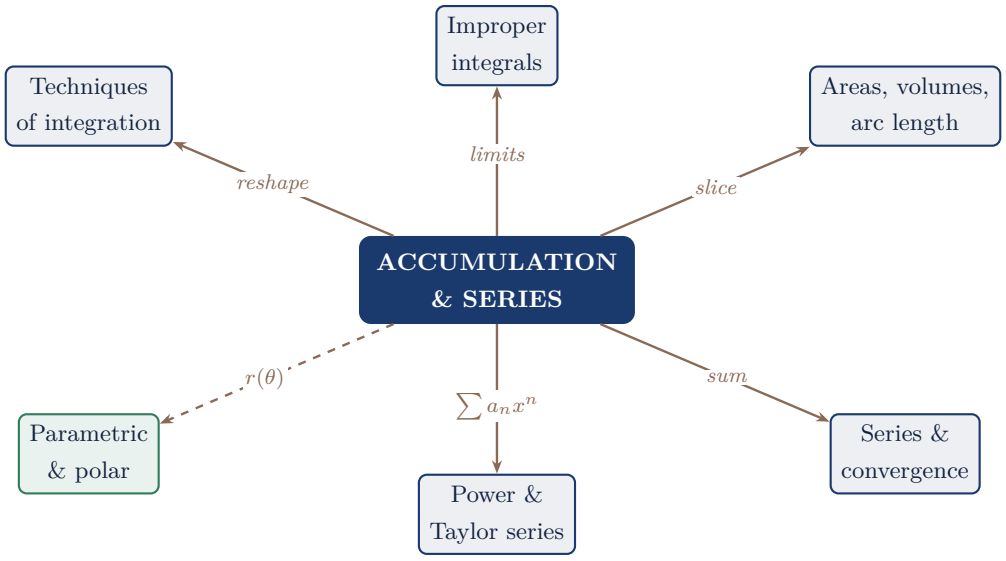
Calculus II

Integration Techniques and Infinite Series

"Nature laughs at the difficulties of integration."

— Pierre-Simon Laplace

Where Chapter 2 asked how a quantity* changes*, this chapter asks how a quantity* accumulates*—and then how infinitely many small pieces can add to something finite. Two engines drive it. The first is a toolbox of integration techniques—substitution, parts, trigonometric methods, and partial fractions—each of which reshapes an integral you cannot see into one you can. The second is the theory of infinite series: when does an endless sum settle to a number, and how do we build a function out of its own derivatives as a power series? The Subject Test leans hard on both. It asks you to integrate quickly, to compute an area, a volume of revolution, or an arc length, to decide in seconds whether a series converges, and to recognize a Taylor expansion. Everything rests on the Fundamental Theorem of Chapter 2: an integral is an antiderivative evaluated at endpoints, and every technique here is a way of finding that antiderivative.



Learning Objectives

You will be able to:

1. integrate by substitution, by parts, by trigonometric methods, by trigonometric substitution, and by partial fractions;
2. evaluate improper integrals and decide their convergence;
3. compute areas between curves, volumes by disks/washers and by shells, arc length, surface area, average value, and work;
4. find limits of sequences and decide monotone/bounded convergence;
5. apply every standard series test—nth-term, geometric, p-series, integral, comparison, limit comparison, ratio, root, and alternating—and distinguish absolute from conditional convergence;
6. find the radius and interval of convergence of a power series;
7. build and use Taylor and Maclaurin series for the common functions;
8. compute area and arc length in parametric and polar form.

Problem-solving strategy — How the exam uses this chapter

Expect several quick integrals (spot the technique in one glance), one area/volume/arc-length application, two or three "does this series converge?" items where the winning move is matching the series to the *right test*, one radius-of-convergence computation, and a Taylor-series recognition or approximation. Speed here is pattern recognition: the integrand or the general term tells you which tool to reach for.

Substitution: Running the Chain Rule Backward

Substitution is the integral counterpart of the chain rule. If an integrand contains a composite function together with (a constant multiple of) the derivative of its inside, we rename the inside $u = g(x)$, so that $\mathrm{d}u = g'(x)\,\mathrm{d}x$, and the integral collapses into one in u alone. The whole art is spotting the inside function whose derivative is also present—up to a constant you can adjust.

What & why. Undo a chain-rule derivative: an integrand that is "outer function of an inside, times the inside's derivative" integrates cleanly after renaming the inside.

Solve. Let u be the inside; replace $g'(x)\,\mathrm{d}x$ by $\mathrm{d}u$; integrate in u; substitute back.

Means. Substitution turns a composite integrand into an elementary one.

Approach Read · Plan · Solve · Check

Read. Scan for a composite piece $g(x)$ whose derivative $g'(x)$ (up to a constant multiple) also appears in the integrand.

Plan.

1. Name the inside $u = g(x)$ and compute $\mathrm{d}u = g'(x)\,\mathrm{d}x$.
2. Rewrite the whole integrand in u; if the derivative factor is off by a constant, multiply and divide to supply it.
3. For a definite integral, convert the x-limits to u-limits at the same time.

Solve. Integrate in u to a power or standard form, then restore $u = g(x)$ (or evaluate at the converted limits).

Check. Differentiate your antiderivative — the chain rule should return the original integrand exactly.

Common traps — every wrong answer is one of these

Drops the $+C$.
Reports an indefinite integral without its constant. Every antiderivative family carries $+C$; omitting it loses infinitely many solutions.

Leaves a stray x behind.
Substitutes u for the inside but cannot express some leftover x in terms of u. If a factor of x will not convert, u was the wrong choice — the derivative of the inside must absorb *all* the extra algebra.

Picks an inside whose derivative is absent.
Chooses $u = g(x)$ when $g'(x)$ appears nowhere, so $\mathrm{d}u$ cannot replace anything. Choose the inside whose derivative is already present up to a constant.

Keeps the x-limits after substituting. 1. Substitutes $u = g(x)$ in a definite integral.
2. Evaluates the u-antiderivative at the original x-limits.
3. Fix: either convert the limits to u-values when you substitute, or restore x before evaluating.

A basic substitution. Evaluate $\displaystyle\int 2x\,(x^2+1)^5\,\mathrm{d}x$.

Read.

Spot the inside x^2+1 whose derivative $2x$ multiplies the rest; substitute $u = x^2+1$.

Plan.

Let $u = x^2+1$, so $\mathrm{d}u = 2x\,\mathrm{d}x$—exactly the leftover factor. Then

$$\begin{aligned}
\int 2x\,(x^2+1)^5\,\mathrm{d}x &= \int u^5\,\mathrm{d}u && \text{rewrite the integrand in } u\\
&= \frac{u^6}{6} + C && \text{reverse the power rule}\\
&= \frac{(x^2+1)^6}{6} + C && \text{restore } u = x^2+1
\end{aligned}$$

Solve.

$$\int 2x\,(x^2+1)^5\,\mathrm{d}x = \frac{(x^2+1)^6}{6} + C.$$

Check.

Differentiate the answer: $\frac{1}{6}\cdot 6(x^2+1)^5\cdot 2x = 2x(x^2+1)^5$, the integrand. ✓

Adjusting a constant. Evaluate $\displaystyle\int \frac{x}{x^2+1}\,\mathrm{d}x$.

Read.

Let $u = x^2+1$; the numerator x is half of $\mathrm{d}u$, so introduce the compensating $\frac{1}{2}$.

Plan.

With $u = x^2 + 1$, $\mathrm{d}u = 2x\,\mathrm{d}x$, so $x\,\mathrm{d}x = \frac{1}{2}\,\mathrm{d}u$. Then

$$\begin{aligned}
\int \frac{x}{x^2+1}\,\mathrm{d}x &= \frac{1}{2}\int \frac{\mathrm{d}u}{u} && \text{replace } x\,\mathrm{d}x \text{ by } \tfrac{1}{2}\,\mathrm{d}u \\
&= \frac{1}{2}\ln|u| + C && \text{integrate } 1/u \\
&= \frac{1}{2}\ln(x^2+1) + C && \text{restore } u;\ x^2+1>0 \text{ drops the bars}
\end{aligned}$$

Solve.

$\displaystyle\int \frac{x}{x^2+1}\,\mathrm{d}x = \tfrac{1}{2}\ln(x^2+1) + C.$

Check.

The derivative of $\frac{1}{2}\ln(x^2+1)$ is $\frac{1}{2}\cdot\frac{2x}{x^2+1} = \frac{x}{x^2+1}$. ✓

A definite substitution: convert the limits. Evaluate $\displaystyle\int_1^2 \frac{\ln x}{x}\,\mathrm{d}x$.

Read.

Let $u = \ln x$, whose derivative $\frac{1}{x}$ is already present; convert the x-limits to u-limits as you substitute, so no back-substitution is needed.

Plan.

Let $u = \ln x$, so $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$. Convert the limits: $x = 1 \Rightarrow u = \ln 1 = 0$ and $x = 2 \Rightarrow u = \ln 2$. Then

$$\begin{aligned}
\int_1^2 \frac{\ln x}{x}\,\mathrm{d}x &= \int_0^{\ln 2} u\,\mathrm{d}u && \text{rewrite in } u,\ \text{limits converted} \\
&= \left[\frac{u^2}{2}\right]_0^{\ln 2} && \text{reverse the power rule} \\
&= \frac{(\ln 2)^2}{2} && \text{evaluate at the } u\text{-limits}
\end{aligned}$$

Solve.

$\displaystyle\int_1^2 \frac{\ln x}{x}\,\mathrm{d}x = \frac{(\ln 2)^2}{2} \approx 0.240.$

Check.

Because the limits were converted to u-values, the answer follows at once; plugging the original x-limits 1 and 2 into $\frac{u^2}{2}$ would have been the classic error. ✓

Misconception

Change the limits, or change back. For a definite integral, either convert the x-limits to u-limits when you substitute, or restore x before evaluating. Plugging the original x-limits into a u-antiderivative is a classic sign error.

Integration by Parts

When an integrand is a *product* of two unlike functions—a polynomial times an exponential, a variable times a trig function, or a lone logarithm—substitution usually fails and integration by parts takes over. It is the product rule, integrated.

Derivation. *Integration by parts from the product rule*

The product rule says $(uv)' = u'v + uv'$. Integrate both sides: $uv = \int u'v\,dx + \int uv'\,dx$. Solving for the second integral, $\int u\,dv = uv - \int v\,du$, where $dv = v'\,dx$ and $du = u'\,dx$. The trick is to split the integrand into a part u to differentiate and a part dv to integrate, so that the new integral $\int v\,du$ is easier than the original.

Parts.

$$\int u\,dv = uv - \int v\,du. \tag{3.1}$$

Choose u by the mnemonic **LIATE**—Logarithm, Inverse trig, Algebraic, Trig, Exponential—picking u to be whichever type appears first, because differentiating it simplifies the problem.

Approach Read · Plan · Solve · Check

Read. Recognize a product of two *unlike* functions — a polynomial times an exponential, a variable times a trig function, or a lone logarithm or inverse trig.

Plan.

1. Choose u by **LIATE** (the part to differentiate); the rest is dv.
2. Compute $du = u'\,dx$ and $v = \int dv$ (no $+C$ at this stage).
3. Apply $\int u\,dv = uv - \int v\,du$; if the new integral is still a product, repeat parts.

Solve. Assemble $uv - \int v\,du$, evaluating the reduced integral (or solving algebraically if it cycles back to the original).

Check. Differentiate the result; the product rule should reproduce the integrand.

Common traps — every wrong answer is one of these

Reverses the LIATE choice.
Picks $\mathrm{d}v$ to be the part that should have been u, so the new integral is *harder*. Differentiate the part that simplifies (a polynomial or a logarithm), not the one that grows.

Drops the minus sign.
Writes $uv + \int v\,\mathrm{d}u$ instead of $uv - \int v\,\mathrm{d}u$. The formula subtracts the new integral — a single sign error corrupts the whole answer.

Attaches a $+C$ to v midway.
Adds a constant when integrating $\mathrm{d}v$ inside the problem. Introduce $+C$ only once, on the final antiderivative.

Stops after one round. 1. Applies parts once to a polynomial times $\sin$ / $\cos$ /e.
2. Leaves the remaining product integral unevaluated.
3. Fix: a degree-n polynomial needs n rounds of parts (or a cyclic term solved algebraically).

Polynomial times exponential. Evaluate $\displaystyle\int x\,\mathrm{e}^x\,\mathrm{d}x$.

Read.

By LIATE the algebraic x is u; the rest, $\mathrm{e}^x\,\mathrm{d}x$, is $\mathrm{d}v$.

Plan.

Set $u = x$ and $\mathrm{d}v = \mathrm{e}^x\,\mathrm{d}x$; then $\mathrm{d}u = \mathrm{d}x$ and $v = \mathrm{e}^x$. Apply parts (3.1):

$$\begin{aligned}
\int x\mathrm{e}^x\,\mathrm{d}x &= uv - \int v\,\mathrm{d}u && \text{parts formula}\\
&= x\mathrm{e}^x - \int \mathrm{e}^x\,\mathrm{d}x && \text{substitute } u, v, \mathrm{d}u\\
&= x\mathrm{e}^x - \mathrm{e}^x + C && \text{integrate } \mathrm{e}^x\\
&= \mathrm{e}^x(x-1) + C && \text{factor out } \mathrm{e}^x
\end{aligned}$$

Solve.

$\displaystyle\int x\mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x(x-1) + C.$

Check.

Differentiate: $\mathrm{e}^x(x-1) + \mathrm{e}^x(1) = \mathrm{e}^x\,x$, the integrand. ✓

Integrating a lone logarithm. Evaluate $\int \ln x \, dx$.

Read.

There is no obvious dv; take $u = \ln x$ and $dv = dx$, the standard trick.

Plan.

With $u = \ln x$, $du = \frac{1}{x}\,dx$; with $dv = dx$, $v = x$. Then

$$\begin{aligned} \int \ln x \, dx &= x \ln x - \int x \cdot \tfrac{1}{x}\, dx && \text{parts formula} \\ &= x \ln x - \int 1 \, dx && \text{simplify } x \cdot \tfrac{1}{x} = 1 \\ &= x \ln x - x + C && \text{integrate} \end{aligned}$$

Solve.

$\int \ln x \, dx = x \ln x - x + C$.

Check.

Differentiate: $\ln x + x \cdot \frac{1}{x} - 1 = \ln x$. ✓

Parts applied twice. Evaluate $\int x^2 \cos x \, dx$.

Read.

Take $u = x^2$ to lower the power; parts once leaves $\int x \sin x \, dx$, which needs parts again.

Plan.

Let $u = x^2$, $dv = \cos x \, dx$, so $du = 2x \, dx$, $v = \sin x$:

$$\int x^2 \cos x \, dx = x^2 \sin x - \int 2x \sin x \, dx \qquad \text{parts, first round}$$

For the new integral use parts again with $u = 2x$, $dv = \sin x \, dx$:

$$\begin{aligned} \int 2x \sin x \, dx &= -2x \cos x + \int 2 \cos x \, dx && \text{parts, second round} \\ &= -2x \cos x + 2 \sin x && \text{integrate } 2 \cos x \end{aligned}$$

Substitute this back:

$$\begin{aligned} \int x^2 \cos x \, dx &= x^2 \sin x - (-2x \cos x + 2 \sin x) && \text{combine} \\ &= x^2 \sin x + 2x \cos x - 2 \sin x && \text{distribute the minus} \end{aligned}$$

Solve.

$$\int x^2 \cos x \, dx = x^2 \sin x + 2x \cos x - 2 \sin x + C.$$

Check.

Each application drops the polynomial degree by one; a degree-n polynomial times $\cos x$ needs n rounds of parts. ✓

Trigonometric Integrals and Trigonometric Substitution

Two related skills handle integrals full of sines and cosines. For *powers* of trig functions, the plan is to use the Pythagorean identity to peel off one factor as the derivative of another. For integrands containing $\sqrt{a^2 - x^2}$, $\sqrt{a^2 + x^2}$, or $\sqrt{x^2 - a^2}$, a *trigonometric substitution* converts the root into a single trig function via the Pythagorean identity, after which the trig-power methods finish the job.

Approach Read · Plan · Solve · Check

Read. Classify the integrand: a power of sin/cos (note whether a power is odd or even), or a root $\sqrt{a^2 - x^2}$, $\sqrt{a^2 + x^2}$, or $\sqrt{x^2 - a^2}$.

Plan.

1. For an *odd* power, peel off one factor as du and convert the even remainder with $\sin^2 + \cos^2 = 1$.
2. For *all-even* powers, use the power-reduction identities $\sin^2 x = \frac{1-\cos 2x}{2}$, $\cos^2 x = \frac{1+\cos 2x}{2}$.
3. For a root, substitute $x = a \sin\theta$, $a \tan\theta$, or $a \sec\theta$ to collapse it to one trig function.

Solve. Carry out the resulting elementary integral, then back-substitute — draw a right triangle to convert θ back to x.

Check. Differentiate the result, or sanity-check a definite version against a known area.

Common traps — every wrong answer is one of these

Tries u-substitution on an even power. 1. Meets $\int \sin^2 x \, dx$ and reaches for $u = \sin x$.

2. Finds no spare $\cos x \, dx$ to become du.
3. Fix: even powers have no factor to peel — use a power-reduction identity instead.

Never returns to x.

Finishes a trig substitution but leaves the answer in θ. Undo the substitution ($\theta = \arcsin \frac{x}{a}$, etc.) so the antiderivative is a function of x.

Chooses the wrong substitution for the root.
Pairs $\sqrt{a^2+x^2}$ with $a\sin\theta$ (or similar). Match the sign: $a^2-x^2 \to a\sin\theta$, $a^2+x^2 \to a\tan\theta$, $x^2-a^2 \to a\sec\theta$.

Forgets the $\mathrm{d}x$ conversion.
Replaces x but not $\mathrm{d}x$ (e.g. omits $\mathrm{d}x = a\cos\theta\,\mathrm{d}\theta$). Every x *and* the differential must be rewritten in θ.

An even power of sine. Evaluate $\int \sin^2 x\,\mathrm{d}x$.

Read.

Even powers have no spare factor to substitute; use the power-reduction identity $\sin^2 x = \frac{1-\cos 2x}{2}$.

Plan.

Replace $\sin^2 x$ by the power-reduction identity:

$$\begin{aligned}
\int \sin^2 x\,\mathrm{d}x &= \int \frac{1-\cos 2x}{2}\,\mathrm{d}x && \sin^2 x = \tfrac{1-\cos 2x}{2}\\
&= \frac{1}{2}\int 1\,\mathrm{d}x - \frac{1}{2}\int \cos 2x\,\mathrm{d}x && \text{split the integral}\\
&= \frac{x}{2} - \frac{\sin 2x}{4} + C && \text{integrate each piece}
\end{aligned}$$

Solve.

$$\int \sin^2 x\,\mathrm{d}x = \frac{x}{2} - \frac{\sin 2x}{4} + C.$$

Check.

At a glance the average value of $\sin^2$ is $\frac{1}{2}$, matching the $\frac{x}{2}$ growth term. ✓

An odd power of sine. Evaluate $\int \sin^3 x\,\mathrm{d}x$.

Read.

Odd power: split off one $\sin x$ as the derivative factor, convert the rest to cosines, then substitute $u = \cos x$.

Plan.

Write $\sin^3 x = \sin^2 x \cdot \sin x = (1 - \cos^2 x)\sin x$. Let $u = \cos x$, $\mathrm{d}u = -\sin x\,\mathrm{d}x$:

$$\begin{aligned}
\int (1 - \cos^2 x)\sin x\,\mathrm{d}x &= -\int (1 - u^2)\,\mathrm{d}u && \text{substitute } u = \cos x \\
&= -\left(u - \frac{u^3}{3}\right) + C && \text{integrate in } u \\
&= -\cos x + \frac{\cos^3 x}{3} + C && \text{restore } u = \cos x
\end{aligned}$$

Solve.

$$\int \sin^3 x\,\mathrm{d}x = -\cos x + \frac{\cos^3 x}{3} + C.$$

Check.

Differentiate: $\sin x - \cos^2 x \sin x = \sin x(1 - \cos^2 x) = \sin^3 x$. ✓

A mixed product with an odd power. Evaluate $\displaystyle\int \sin^2 x \cos^3 x\,\mathrm{d}x$.

Read.

The cosine power is odd; peel off one $\cos x$ as $\mathrm{d}u$ and convert the rest to sines via $\cos^2 x = 1 - \sin^2 x$.

Plan.

Write $\cos^3 x = \cos^2 x \cdot \cos x = (1 - \sin^2 x)\cos x$. Let $u = \sin x$, $\mathrm{d}u = \cos x\,\mathrm{d}x$:

$$\begin{aligned}
\int \sin^2 x(1 - \sin^2 x)\cos x\,\mathrm{d}x &= \int u^2(1 - u^2)\,\mathrm{d}u && \text{substitute } u = \sin x \\
&= \int (u^2 - u^4)\,\mathrm{d}u && \text{expand} \\
&= \frac{u^3}{3} - \frac{u^5}{5} + C && \text{integrate in } u \\
&= \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + C && \text{restore } u = \sin x
\end{aligned}$$

Solve.

$$\int \sin^2 x \cos^3 x\,\mathrm{d}x = \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + C.$$

Check.

The rule: with an odd power present, save one factor of that function as $\mathrm{d}u$ and turn the even remainder into the other function. ✓

Trigonometric substitution. Evaluate $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{4-x^2}}$.

Read.

The form $\sqrt{a^2-x^2}$ with $a=2$ calls for $x = a\sin\theta$, which turns the root into $a\cos\theta$.

Plan.

Let $x = 2\sin\theta$, so $\mathrm{d}x = 2\cos\theta\,\mathrm{d}\theta$. First collapse the root:

$$\begin{aligned} \sqrt{4-x^2} &= \sqrt{4-4\sin^2\theta} && \text{substitute } x = 2\sin\theta \\ &= 2\cos\theta && 1-\sin^2\theta = \cos^2\theta \end{aligned}$$

Then the integral becomes

$$\begin{aligned} \int \frac{\mathrm{d}x}{\sqrt{4-x^2}} &= \int \frac{2\cos\theta\,\mathrm{d}\theta}{2\cos\theta} && \text{substitute } \mathrm{d}x \text{ and the root} \\ &= \int \mathrm{d}\theta = \theta + C && \text{simplify and integrate} \\ &= \arcsin\frac{x}{2} + C && \theta = \arcsin\tfrac{x}{2} \end{aligned}$$

Solve.

$$\int \frac{\mathrm{d}x}{\sqrt{4-x^2}} = \arcsin\frac{x}{2} + C.$$

Check.

Differentiating $\arcsin(x/2)$ gives $\frac{1}{\sqrt{1-(x/2)^2}}\cdot\frac{1}{2} = \frac{1}{\sqrt{4-x^2}}$. ✓

Which substitution. For $\sqrt{a^2-x^2}$ use $x = a\sin\theta$ (root becomes $a\cos\theta$); for $\sqrt{a^2+x^2}$ use $x = a\tan\theta$ (root becomes $a\sec\theta$); for $\sqrt{x^2-a^2}$ use $x = a\sec\theta$ (root becomes $a\tan\theta$). In every case the Pythagorean identity collapses the root to a single trig function.

A trig substitution that needs the double angle. Evaluate $\displaystyle\int \sqrt{9-x^2}\,\mathrm{d}x$.

Read.

Use $x = 3\sin\theta$; the root becomes $3\cos\theta$ and the integrand turns into $9\cos^2\theta$, a power-reduction integral.

Plan.

Let $x = 3\sin\theta$, $\mathrm{d}x = 3\cos\theta\,\mathrm{d}\theta$, and $\sqrt{9-x^2} = 3\cos\theta$. Then

$$\begin{aligned}
\int \sqrt{9-x^2}\,\mathrm{d}x &= \int 3\cos\theta \cdot 3\cos\theta\,\mathrm{d}\theta && \text{substitute the root and } \mathrm{d}x \\
&= 9\int \cos^2\theta\,\mathrm{d}\theta && \text{collect constants} \\
&= 9\left(\frac{\theta}{2} + \frac{\sin 2\theta}{4}\right) + C && \text{power-reduce and integrate}
\end{aligned}$$

Now unwind: $\theta = \arcsin\frac{x}{3}$, and since $\sin 2\theta = 2\sin\theta\cos\theta = 2\cdot\frac{x}{3}\cdot\frac{\sqrt{9-x^2}}{3}$, the second term becomes $\frac{x\sqrt{9-x^2}}{9}\cdot\frac{9}{4}$.

Solve.

$$\int \sqrt{9-x^2}\,\mathrm{d}x = \frac{9}{2}\arcsin\frac{x}{3} + \frac{x\sqrt{9-x^2}}{2} + C.$$

Check.

Over $[-3,3]$ this gives $\frac{9}{2}\cdot\pi$, i.e. the area $\frac{9\pi}{2}$ of a semicircle of radius 3—exactly what the integrand describes. ✓

The secant case. Evaluate $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x^2+9}}$.

Read.

The form $\sqrt{a^2+x^2}$ with $a=3$ calls for $x = 3\tan\theta$, turning the root into $3\sec\theta$.

Plan.

Let $x = 3\tan\theta$, $\mathrm{d}x = 3\sec^2\theta\,\mathrm{d}\theta$, and $\sqrt{x^2+9} = 3\sec\theta$. Then

$$\begin{aligned}
\int \frac{\mathrm{d}x}{\sqrt{x^2+9}} &= \int \frac{3\sec^2\theta\,\mathrm{d}\theta}{3\sec\theta} && \text{substitute } \mathrm{d}x \text{ and the root} \\
&= \int \sec\theta\,\mathrm{d}\theta && \text{simplify} \\
&= \ln\left|\sec\theta + \tan\theta\right| + C && \text{standard secant integral} \\
&= \ln\left|\frac{\sqrt{x^2+9}+x}{3}\right| + C && \sec\theta = \tfrac{\sqrt{x^2+9}}{3},\ \tan\theta = \tfrac{x}{3}
\end{aligned}$$

Solve.

$$\int \frac{\mathrm{d}x}{\sqrt{x^2+9}} = \ln\left(\sqrt{x^2+9}+x\right) + C \text{ (absorbing } -\ln 3 \text{ into } C).$$

Check.

This is the inverse hyperbolic sine $\sinh^{-1}(x/3)$ in logarithmic form—the standard antideriva-

tive for $\sqrt{x^2+a^2}$. ✓

Partial Fractions

A rational function whose denominator factors can be split into a sum of simpler fractions, each of which integrates to a logarithm or an arctangent. This is *partial fraction decomposition*, and it is the standard route to integrating any ratio of polynomials once the degree of the numerator is below that of the denominator (if not, divide first).

Approach Read · Plan · Solve · Check

Read. Confirm the integrand is a ratio of polynomials, then check that the numerator degree is below the denominator degree.

Plan.

1. If the top degree is not smaller, do polynomial division first and decompose the remainder.
2. Factor the denominator into linear and irreducible-quadratic pieces, and write the decomposition with unknown numerators ($\frac{A}{x-r}$ for linear, $\frac{Ax+B}{\text{quadratic}}$ for irreducible).
3. Clear denominators and solve for the constants by cover-up or by substituting convenient x-values.

Solve. Integrate each simple fraction — a logarithm for a linear factor, an arctangent (or log) for an irreducible quadratic.

Check. Recombine the fractions, or differentiate, to recover the original integrand.

Common traps — every wrong answer is one of these

Skips the division step. 1. Meets a ratio whose top degree is $\geq$ the bottom degree.

2. Sets up a partial-fraction form directly.
3. Fix: divide first — an improper rational function has no valid decomposition until the polynomial part is split off.

Uses a constant over an irreducible quadratic.

Writes $\frac{A}{x^2+1}$ instead of $\frac{Ax+B}{x^2+1}$. An irreducible quadratic factor needs a *linear* numerator.

Ignores repeated factors.

For $(x-r)^2$ writes only $\frac{A}{x-r}$. A repeated factor needs a term for each power: $\frac{A}{x-r}+\frac{B}{(x-r)^2}$.

Drops the absolute values.
Integrates $\frac{1}{x-r}$ to $\ln(x-r)$ without bars. The antiderivative is $\ln|x-r|$, valid on both sides of the singularity.

Distinct linear factors. Evaluate $\displaystyle\int \frac{3x+5}{(x+1)(x+3)}\,\mathrm{d}x$.

Read.

Write the integrand as $\frac{A}{x+1}+\frac{B}{x+3}$, solve for A, B, then integrate termwise.

Plan.

Set $\frac{3x+5}{(x+1)(x+3)}=\frac{A}{x+1}+\frac{B}{x+3}$ and clear denominators: $3x+5=A(x+3)+B(x+1)$. Substitute convenient values:

$$\begin{aligned} x=-1: &\quad 2=2A && \text{so } A=1\\ x=-3: &\quad -4=-2B && \text{so } B=2 \end{aligned}$$

Then integrate termwise:

$$\begin{aligned} \int \frac{3x+5}{(x+1)(x+3)}\,\mathrm{d}x &= \int\left(\frac{1}{x+1}+\frac{2}{x+3}\right)\mathrm{d}x && \text{decomposed form}\\ &= \ln|x+1|+2\ln|x+3|+C && \text{integrate each fraction} \end{aligned}$$

Solve.

$$\int \frac{3x+5}{(x+1)(x+3)}\,\mathrm{d}x=\ln|x+1|+2\ln|x+3|+C.$$

Check.

Recombining the two logs and differentiating recovers $\frac{3x+5}{(x+1)(x+3)}$; the cover-up values A, B can also be checked by plugging a third x. ✓

Divide first when the top is too big. Evaluate $\displaystyle\int \frac{x^2}{x^2-1}\,\mathrm{d}x$.

Read.

The numerator degree is not below the denominator's, so do polynomial division before decomposing.

Plan.

Since the degrees are equal, divide first: $\frac{x^2}{x^2-1}=1+\frac{1}{x^2-1}$. Decompose the remainder over

$(x-1)(x+1)$: $\frac{1}{x^2-1} = \frac{A}{x-1} + \frac{B}{x+1}$ gives $1 = A(x+1) + B(x-1)$, so $A = \frac{1}{2}$ (at $x = 1$) and $B = -\frac{1}{2}$ (at $x = -1$). Then integrate:

$$\begin{aligned} \int \frac{x^2}{x^2-1}\,dx &= \int 1\,dx + \frac{1}{2}\int \frac{dx}{x-1} - \frac{1}{2}\int \frac{dx}{x+1} && \text{quotient plus decomposition} \\ &= x + \frac{1}{2}\ln|x-1| - \frac{1}{2}\ln|x+1| + C && \text{integrate each piece} \end{aligned}$$

Solve.

$$\int \frac{x^2}{x^2-1}\,dx = x + \tfrac{1}{2}\ln\left|\frac{x-1}{x+1}\right| + C.$$

Check.

Skipping the division would strand you with an "improper" partial fraction that has no valid decomposition—always check degrees first. ✓

An irreducible quadratic factor. Evaluate $\displaystyle\int \frac{dx}{(x-1)(x^2+1)}$.

Read.

The factor x^2+1 is irreducible, so its numerator must be *linear*, $Bx + C$; solve for the constants, then integrate to a logarithm and an arctangent.

Plan.

Set $\frac{1}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}$ and clear denominators: $1 = A(x^2+1) + (Bx+C)(x-1)$. Substitute $x = 1$, then match coefficients:

$$\begin{aligned} x = 1: &\quad 1 = 2A && \text{so } A = \tfrac{1}{2} \\ x^2 \text{ coeff}: &\quad 0 = A + B && \text{so } B = -\tfrac{1}{2} \\ \text{constant}: &\quad 1 = A - C && \text{so } C = -\tfrac{1}{2} \end{aligned}$$

Then integrate the decomposition termwise:

$$\begin{aligned} \int \frac{dx}{(x-1)(x^2+1)} &= \int \left(\frac{1/2}{x-1} - \frac{\frac{1}{2}x + \frac{1}{2}}{x^2+1} \right) dx && \text{decomposed form} \\ &= \frac{1}{2}\ln|x-1| - \frac{1}{4}\ln(x^2+1) - \frac{1}{2}\arctan x + C && \text{log, log, arctangent} \end{aligned}$$

Solve.

$$\int \frac{dx}{(x-1)(x^2+1)} = \frac{1}{2}\ln|x-1| - \frac{1}{4}\ln(x^2+1) - \frac{1}{2}\arctan x + C.$$

Check.

A bare constant $\frac{A}{x^2+1}$ would leave the system unsolvable—an irreducible quadratic needs the full linear numerator $Bx + C$, whose two pieces split into a logarithm (from the x term) and an arctangent (from the constant). ✓

Did You Know? *Every rational function is elementary-integrable*

Because a real polynomial factors into linear and irreducible-quadratic pieces, partial fractions guarantees that *any* ratio of polynomials integrates to a combination of rational functions, logarithms, and arctangents. There is no rational function you cannot integrate in closed form—a comfort not shared by, say, e^{-x^2}.

Improper Integrals

An integral is *improper* if its interval is infinite or its integrand blows up inside the interval. We define such an integral as a limit of ordinary integrals: replace the offending endpoint by a variable, integrate, and take the limit. If the limit is finite the integral *converges*; otherwise it *diverges*. These behave exactly like the infinite series later in the chapter—indeed the integral test will make the analogy exact.

Approach Read · Plan · Solve · Check

Read. Locate the source of impropriety: an infinite limit of integration, or an integrand that blows up at an endpoint or an interior point.

Plan.

1. Replace the offending endpoint by a variable ($b \to \infty$, or $a \to c$ at a singularity).
2. Compute the ordinary definite integral in that variable.
3. Take the limit; if the trouble is interior, split into two improper pieces first.

Solve. Evaluate the limit: a finite value means the integral converges to it; an infinite or nonexistent limit means it diverges.

Check. Compare against the p-benchmark ($\int_1^\infty x^{-p}$ converges iff $p > 1$; $\int_0^1 x^{-p}$ iff $p < 1$) to confirm the verdict.

Common traps — every wrong answer is one of these

Integrates across an interior blow-up.

1. Meets an integrand singular at an interior point (say $x = 0$ in $\int_{-1}^{1} \frac{\mathrm{d}x}{x^2}$).
2. Applies the FTC straight through as if it were continuous.
3. Fix: split at the singularity and treat each side as its own improper integral.

Uses the wrong p-benchmark.

Confuses the two thresholds — at ∞ convergence needs $p > 1$, near 0 it needs $p < 1$. Match the benchmark to where the trouble is.

Calls it convergent because the integrand $\to 0$.

A decaying integrand can still enclose infinite area (as $\int_1^{\infty} \frac{\mathrm{d}x}{x}$ shows). The limit of the *integral*, not of the integrand, decides.

Forgets the limit is required.

Writes $\left[-\frac{1}{x}\right]_1^{\infty}$ and plugs in ∞ directly. An improper integral is *defined* as a limit — evaluate it as one.

An infinite interval that converges. Evaluate $\displaystyle\int_1^{\infty} \frac{\mathrm{d}x}{x^2}$.

Read.

Cut off at a finite upper limit b, integrate, then let $b \to \infty$.

Plan.

Cut off at b, integrate, then take the limit:

$$\begin{aligned} \int_1^b x^{-2}\,\mathrm{d}x &= \left[-\frac{1}{x}\right]_1^b && \text{antiderivative} \\ &= 1 - \frac{1}{b} && \text{evaluate at the limits} \\ \lim_{b\to\infty}\left(1 - \frac{1}{b}\right) &= 1 && \text{let } b \to \infty \end{aligned}$$

Solve.

$\displaystyle\int_1^{\infty} \frac{\mathrm{d}x}{x^2} = 1$; the integral converges.

Check.

Contrast $\int_1^{\infty} \frac{1}{x}\,\mathrm{d}x = \lim_b \ln b = \infty$: the same-looking integral with one lower power *diverges*. The boundary is the power 1. ✓

An unbounded integrand. Evaluate $\displaystyle\int_0^1 \frac{\mathrm{d}x}{\sqrt{x}}$.

Read.

The integrand blows up at $x = 0$; start the interval at $a > 0$ and let $a \to 0^+$.

Plan.

Start the interval at $a > 0$, integrate, and let $a \to 0^+$:

$$\begin{aligned} \int_a^1 x^{-1/2}\,\mathrm{d}x &= \left[2\sqrt{x}\right]_a^1 && \text{antiderivative} \\ &= 2 - 2\sqrt{a} && \text{evaluate at the limits} \\ \lim_{a\to 0^+} (2 - 2\sqrt{a}) &= 2 && \text{let } a \to 0^+ \end{aligned}$$

Solve.

$\displaystyle\int_0^1 \frac{\mathrm{d}x}{\sqrt{x}} = 2$; the integral converges despite the vertical asymptote.

Check.

The singularity is mild enough ($x^{-1/2}$, exponent above -1) that the area under the spike is finite. ✓

The p-integral benchmark. $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{x^p}$ converges $\iff p > 1$; $\displaystyle\int_0^1 \frac{\mathrm{d}x}{x^p}$ converges $\iff p < 1$. Keep both benchmarks in mind—they anchor comparison arguments for integrals and, shortly, for series.

Comparison for an improper integral. Does $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{x^3 + 1}$ converge?

Read.

Bound the integrand by a simpler one whose integral you know; a smaller positive integrand under a convergent one converges.

Plan.

For $x \geq 1$, $x^3 + 1 > x^3$, so $\frac{1}{x^3+1} < \frac{1}{x^3}$. The benchmark $\int_1^\infty x^{-3}\,\mathrm{d}x$ converges (since $p = 3 > 1$). A positive integrand dominated by a convergent one has a convergent integral.

Solve.

Converges by comparison with $\int_1^\infty x^{-3}\,\mathrm{d}x$.

Check.

We never needed the exact value—often the exam asks only *whether* an integral converges,

and comparison settles that without an antiderivative. ✓

Applications: Areas, Volumes, Arc Length, and More

The definite integral sums infinitely many infinitesimal pieces, and choosing what the piece *is* gives a whole family of geometric and physical quantities. Every formula below is the same idea: identify a thin slice, write its contribution, and integrate.

Approach Read · Plan · Solve · Check

Read. Name the geometric or physical quantity (area, volume, arc length, average value, work) and sketch the region or curve involved.

Plan.

1. Choose the slice: a vertical or horizontal strip, a disk, a washer, a cylindrical shell, an arc element, or a polar wedge.
2. Write that slice's contribution in a single variable (height $\times$ thickness, πR^2, $\pi(R^2 - r^2)$, $2\pi x\, h$, $\sqrt{1 + (f')^2}$, and so on).
3. Find the limits from intersection points or the given endpoints.

Solve. Integrate the slice's contribution over those limits.

Check. Confirm the units and compare the value to a rough estimate (a known region of similar size).

Common traps — every wrong answer is one of these

Subtracts radii before squaring.
For a washer, computes $(R - r)^2$ instead of $R^2 - r^2$. Square each radius *first*, then subtract — the two are not equal.

Uses the wrong distance for the radius.
Revolving about the y-axis but using the height as radius (or vice versa). The radius is the distance from the *axis of revolution* to the slice.

Reverses top and bottom. 1. Integrates $\int(\text{lower} - \text{upper})$ by mistake.

2. Gets a negative area.
3. Fix: always integrate (upper − lower), or (right − left) in y; the enclosed area is positive.

Guesses the limits.
Uses convenient round numbers instead of the actual intersections. Solve $f = g$ (or $r = \ell$) to find where the region begins and ends.

Area between two curves

Between $y = f(x)$ (upper) and $y = g(x)$ (lower) on $[a, b]$, a thin vertical strip has height $f(x) - g(x)$ and area $[f(x) - g(x)]\,\mathrm{d}x$, so the region's area is

$$A = \int_a^b [f(x) - g(x)]\,\mathrm{d}x. \tag{3.2}$$

Area between curves. Find the area between $y = x$ and $y = x^2$ from $x = 0$ to $x = 1$.

Read.

On $[0, 1]$ decide which curve is higher, then integrate (upper − lower) via (3.2).

Plan.

On $(0, 1)$, $x > x^2$, so $y = x$ is the upper curve. Integrate the gap:

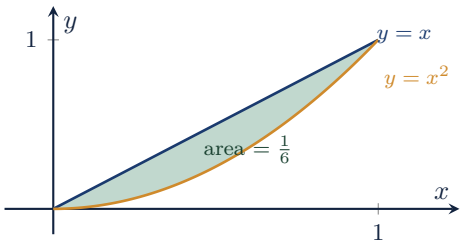
$$\begin{aligned} A &= \int_0^1 (x - x^2)\,\mathrm{d}x && \text{upper minus lower} \\ &= \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 && \text{antiderivative} \\ &= \frac{1}{2} - \frac{1}{3} = \frac{1}{6} && \text{evaluate} \end{aligned}$$

Solve.

Area $= \frac{1}{6}$.

Check.

The two curves meet at $x = 0$ and $x = 1$, bounding a small lens; $\frac{1}{6}$ is plausibly less than the unit square. ✓



Area between $y = x$ and $y = x^2$: the shaded lens has area $\int_0^1 (x - x^2)\,\mathrm{d}x = \frac{1}{6}$.

Sometimes horizontal strips are simpler. If a region is bounded on the right by $x = r(y)$ and on the left by $x = \ell(y)$ for $y \in [c, d]$, its area is $\int_c^d [r(y) - \ell(y)]\,\mathrm{d}y$—the same idea with the roles of x and y swapped.

Integrating in y. Find the area of the region bounded by $x = y^2$ and $x = y + 2$.

Read.

Find the y-intersections, decide which curve is on the right, and integrate (right − left) in y.

Plan.

Set $y^2 = y + 2$: $y^2 - y - 2 = (y - 2)(y + 1) = 0$, so $y = -1$ and $y = 2$. Between them the line $x = y + 2$ lies to the right of the parabola $x = y^2$. Integrate right minus left in y:

$$\begin{aligned} A &= \int_{-1}^{2} \left[(y + 2) - y^2\right] \mathrm{d}y && \text{right minus left} \\ &= \left[\frac{y^2}{2} + 2y - \frac{y^3}{3}\right]_{-1}^{2} && \text{antiderivative} \\ &= \left(2 + 4 - \frac{8}{3}\right) - \left(\frac{1}{2} - 2 + \frac{1}{3}\right) = \frac{9}{2} && \text{evaluate} \end{aligned}$$

Solve.

Area $= \frac{9}{2}$.

Check.

Integrating in x here would need two pieces (the parabola gives two branches); the horizontal strip handles it in one integral. ✓

Volumes of revolution: disks and washers

Rotating a region about an axis sweeps out a solid. A thin slice perpendicular to the axis is a disk (or, if the region does not touch the axis, a washer). A disk of radius R and thickness $\mathrm{d}x$ has volume

$\pi R^2\,\mathrm{d}x$; a washer with outer radius R and inner radius r has volume $\pi(R^2 - r^2)\,\mathrm{d}x$. Integrating over $[a, b]$ gives the volume

$$V = \pi \int_a^b \left[R(x)^2 - r(x)^2\right] \mathrm{d}x. \tag{3.3}$$

Derivation. *The disk and washer formulas*

Slice the solid perpendicular to the axis of rotation into thin plates. Each plate is a circle of radius equal to the distance from the axis to the curve, $R(x)$, and thickness $\mathrm{d}x$, so its volume is $\pi R(x)^2\,\mathrm{d}x$. Summing (integrating) over $[a, b]$ gives $V = \int_a^b \pi R(x)^2\,\mathrm{d}x$. If the region is bounded by an outer curve $R(x)$ and an inner curve $r(x)$, the plate is a washer and its volume is $\pi[R(x)^2 - r(x)^2]\,\mathrm{d}x$.

Disk method. The region under $y = \sqrt{x}$ on $[0, 4]$ is revolved about the x-axis. Find the volume.

Read.

Each slice is a disk of radius $R = \sqrt{x}$; integrate πR^2 over $[0, 4]$.

Plan.

The radius is the curve height $R = \sqrt{x}$, so $R^2 = x$. Then

$$\begin{aligned} V &= \int_0^4 \pi(\sqrt{x})^2\,\mathrm{d}x && \text{disk volume } \pi R^2 \\ &= \pi \int_0^4 x\,\mathrm{d}x && R^2 = x \\ &= \pi \left[\frac{x^2}{2}\right]_0^4 = 8\pi && \text{integrate and evaluate} \end{aligned}$$

Solve.

$V = 8\pi$.

Check.

Units of length cubed and a positive value—consistent with a solid the size of a small horn. ✓

Washer method. The region between $y = x$ and $y = x^2$ is revolved about the x-axis. Find the volume.

Read.

Outer radius is the higher curve $R = x$, inner is $r = x^2$; apply the washer volume (3.3) on

$[0, 1]$.

Plan.

On $[0, 1]$, $x \geq x^2$, so $R = x$ and $r = x^2$. Then

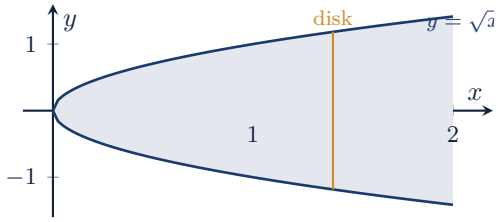
$$\begin{aligned} V &= \pi \int_0^1 (x^2 - x^4)\,\mathrm{d}x && \text{washer } \pi(R^2 - r^2),\ R^2 - r^2 = x^2 - x^4 \\ &= \pi \left[\frac{x^3}{3} - \frac{x^5}{5} \right]_0^1 && \text{antiderivative} \\ &= \pi \left(\frac{1}{3} - \frac{1}{5} \right) = \frac{2\pi}{15} && \text{evaluate} \end{aligned}$$

Solve.

$V = \frac{2\pi}{15}$.

Check.

Squaring makes $R^2 - r^2 = x^2 - x^4$, *not* $(x - x^2)^2$—subtracting the radii before squaring is the classic washer error. ✓



Disk method: revolving $y = \sqrt{x}$ about the x-axis, a slice at x is a disk of radius $\sqrt{x}$ (one shown in orange), contributing $\pi x\,\mathrm{d}x$ to the volume.

Volumes by cylindrical shells

Rotating about the *other* axis, it is often easier to sum concentric cylindrical shells. A shell at radius x, height $h(x)$, thickness $\mathrm{d}x$ has volume $2\pi x\,h(x)\,\mathrm{d}x$ —circumference times height times thickness—so the total volume is

$$V = 2\pi \int_a^b x\,h(x)\,\mathrm{d}x. \tag{3.4}$$

Shell method. The region under $y = x^2$ on $[0, 2]$ is revolved about the y-axis. Find the volume.

Read.

A shell at radius x has height $h = x^2$; apply the shell volume (3.4) over $[0, 2]$.

Plan.

A shell at radius x has height $h = x^2$, so

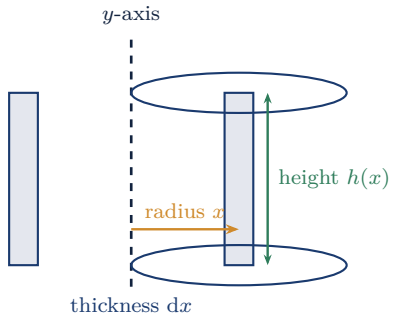
$$\begin{aligned} V &= \int_0^2 2\pi x \cdot x^2 \, \mathrm{d}x && \text{shell } 2\pi x\, h(x) \\ &= 2\pi \int_0^2 x^3 \, \mathrm{d}x && \text{combine} \\ &= 2\pi \left[\frac{x^4}{4}\right]_0^2 = 2\pi \cdot 4 = 8\pi && \text{integrate and evaluate} \end{aligned}$$

Solve.

$V = 8\pi$.

Check.

Revolving about the y-axis makes "radius" the horizontal distance x; using shells avoids solving $y = x^2$ for x that the disk method would require. ✓



A cylindrical shell at radius x: unrolled it is a thin sheet of length $2\pi x$ (circumference), height $h(x)$, and thickness $\mathrm{d}x$, contributing $2\pi x\, h(x)\, \mathrm{d}x$ to the volume.

Arc length, average value, and work

The length of $y = f(x)$ on $[a, b]$ is

$$L = \int_a^b \sqrt{1 + [f'(x)]^2} \, \mathrm{d}x, \tag{3.5}$$

from the Pythagorean length $\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}$ of an infinitesimal step. The *average value* of f on $[a,b]$ is $\frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x$. *Work* done by a variable force $F(x)$ over $[a,b]$ is $\int_a^b F(x)\,\mathrm{d}x$.

Derivation. *The arc-length formula*

Approximate the curve by tiny straight steps. One step advances $\mathrm{d}x$ horizontally and $\mathrm{d}y$ vertically, so by the Pythagorean theorem its length is $\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}=\sqrt{1+(\mathrm{d}y/\mathrm{d}x)^2}\,\mathrm{d}x$. Summing the steps and passing to the limit gives $L=\int_a^b\sqrt{1+[f'(x)]^2}\,\mathrm{d}x$. The same reasoning in parametric form, with $\mathrm{d}x=\dot{x}\,\mathrm{d}t$ and $\mathrm{d}y=\dot{y}\,\mathrm{d}t$, yields $L=\int\sqrt{\dot{x}^2+\dot{y}^2}\,\mathrm{d}t$.

Arc length. Find the length of $y=\frac{2}{3}x^{3/2}$ from $x=0$ to $x=3$.

Read.

Compute f' and apply the arc-length formula (3.5); the exponent is chosen so the root simplifies.

Plan.

Here $f'(x)=x^{1/2}$, so $1+[f'(x)]^2=1+x$. Then

$$\begin{aligned} L &= \int_0^3 \sqrt{1+x}\,\mathrm{d}x && \text{arc-length integrand } \sqrt{1+(f')^2}\\ &= \left[\frac{2}{3}(1+x)^{3/2}\right]_0^3 && \text{antiderivative}\\ &= \frac{2}{3}(4^{3/2}-1)=\frac{2}{3}(8-1)=\frac{14}{3} && \text{evaluate}\end{aligned}$$

Solve.

$L=\frac{14}{3}$.

Check.

The straight-line distance from $(0,0)$ to $(3,2\sqrt{3})\approx(3,3.46)$ is about 4.6, a touch less than $14/3\approx 4.67$—an arc should be slightly longer than the chord. ✓

Average value. Find the average value of $f(x)=x^2$ on $[0,3]$.

Read.

Apply $\bar{f}=\frac{1}{b-a}\int_a^b f$.

Plan.

Apply the average-value formula:

$$\begin{aligned} \bar{f} &= \frac{1}{3-0}\int_0^3 x^2\,\mathrm{d}x && \tfrac{1}{b-a}\int_a^b f \\ &= \frac{1}{3}\left[\frac{x^3}{3}\right]_0^3 && \text{antiderivative} \\ &= \frac{1}{3}\cdot 9 = 3 && \text{evaluate} \end{aligned}$$

Solve.

$\bar{f} = 3$.

Check.

The value 3 is attained at $x = \sqrt{3} \approx 1.73$, inside $[0, 3]$, as the Mean Value Theorem for integrals promises. ✓

The *surface area* generated by revolving $y = f(x)$ on $[a, b]$ about the x-axis is $\int_a^b 2\pi\, f(x)\sqrt{1 + [f'(x)]^2}\,\mathrm{d}x$—circumference $2\pi f(x)$ of each ring times the slant length $\sqrt{1 + (f')^2}\,\mathrm{d}x$ of the arc that sweeps it.

Surface area of revolution. Find the surface area when $y = x$ on $[0, 1]$ is revolved about the x-axis.

Read.

Apply $S = \int 2\pi y\sqrt{1 + (y')^2}\,\mathrm{d}x$ with $y = x$, $y' = 1$.

Plan.

Here $\sqrt{1 + (y')^2} = \sqrt{1 + 1} = \sqrt{2}$, so

$$\begin{aligned} S &= \int_0^1 2\pi x\sqrt{2}\,\mathrm{d}x && \text{surface-area integrand } 2\pi y\sqrt{1 + (y')^2} \\ &= 2\sqrt{2}\,\pi\left[\frac{x^2}{2}\right]_0^1 && \text{pull out constants, integrate} \\ &= 2\sqrt{2}\,\pi\cdot\frac{1}{2} = \sqrt{2}\,\pi && \text{evaluate} \end{aligned}$$

Solve.

$S = \sqrt{2}\,\pi$.

Check.

The solid is a cone of base radius 1 and slant height $\sqrt{2}$; the lateral-area formula $\pi r\ell = \pi(1)(\sqrt{2})$ agrees. ✓

Work against a variable force. A force of 10 N stretches a spring 0.5 m beyond its natural length. Find the work to stretch it 0.2 m from natural length.

Read.

Find k from Hooke's law $F = kx$, then integrate F over the displacement.

Plan.

From $F = kx$: $10 = k(0.5)$, so $k = 20$ N/m. Then

$$\begin{aligned} W &= \int_0^{0.2} 20x\,\mathrm{d}x && \text{work } \int F\,\mathrm{d}x \text{ with } F = 20x \\ &= 20\left[\frac{x^2}{2}\right]_0^{0.2} && \text{antiderivative} \\ &= 20 \cdot \frac{0.04}{2} = 0.4 && \text{evaluate (joules)} \end{aligned}$$

Solve.

$W = 0.4$ J.

Check.

Work grows with the *square* of the stretch, so the last stretch costs more than the first—the integral, not $F \cdot d$, captures this. ✓

Sequences

A *sequence* $\{a_n\}$ is an ordered list indexed by the positive integers; it *converges* to L if the terms get and stay arbitrarily close to L as $n \to \infty$. Limits of sequences obey the same algebra and the same dominant-term reasoning as limits of functions, and a bounded sequence that only moves one direction must converge—a fact that underlies every series test to come.

Monotone convergence. A sequence that is monotone (always increasing or always decreasing) *and* bounded converges. Boundedness alone or monotonicity alone is not enough.

Approach Read · Plan · Solve · Check

Read. Identify the general term a_n and its type: rational in n, exponential, factorial, or recursively defined.

Plan.

1. For a rational term, divide numerator and denominator by the highest power of n.
2. For competing exponentials or factorials, invoke the growth hierarchy ($\ln n \ll n^p \ll c^n \ll n!$).
3. For a recursive $a_{n+1} = g(a_n)$, argue it is monotone and bounded, then solve the fixed-point equation $L = g(L)$.

Solve. Send $n \to \infty$, keeping only the dominant behavior.

Check. Does the limit match the trend of the first few terms?

Common traps — every wrong answer is one of these

Confuses sequence with series.
Treats "$a_n \to 0$" as "the series converges," or vice versa. A sequence converges if its *terms* settle; a series converges if its *partial sums* settle — different questions.

Thinks bounded implies convergent.
Concludes a bounded sequence must converge. It also needs to be monotone — $a_n = (-1)^n$ is bounded but oscillates forever.

Mishandles an oscillating sign.
Assumes $(-1)^n$ or $\cos n$ approaches a limit. An alternating or oscillating factor with no decay makes the sequence diverge.

Keeps a lower-order term.
Divides by the wrong power and retains a term that should vanish. Divide by the *highest* power present so every subordinate term goes to 0.

A sequence limit by dominant terms. Find $\lim_{n\to\infty} \dfrac{2n^2 - n}{3n^2 + 5}$.

Read.

Divide numerator and denominator by the highest power n^2, then send $n \to \infty$.

Plan.

Divide numerator and denominator by n^2:

$$\frac{2n^2 - n}{3n^2 + 5} = \frac{2 - \frac{1}{n}}{3 + \frac{5}{n^2}} \qquad \text{divide by the highest power } n^2$$
$$\to \frac{2}{3} \qquad \text{as } n \to \infty,\ \tfrac{1}{n}, \tfrac{5}{n^2} \to 0$$

Solve.

The limit is $\frac{2}{3}$.

Check.

Only the leading powers matter at infinity, exactly as for rational-function limits in Chapter 2. ✓

A bounded sequence that still diverges. Determine whether $a_n = (-1)^n \frac{n}{n+1}$ converges.

Read.

Boundedness alone is not convergence—inspect the sign, and check whether the odd- and even-indexed terms head to the same place.

Plan.

The size $\frac{n}{n+1} \to 1$ as $n \to \infty$, so every term stays within $[-1, 1]$: the sequence is bounded. But the sign alternates, so split by parity:

$$a_{2k} = +\frac{2k}{2k+1} \to +1 \qquad \text{even indices}$$
$$a_{2k+1} = -\frac{2k+1}{2k+2} \to -1 \qquad \text{odd indices}$$

The two subsequences approach different limits, $+1$ and -1.

Solve.

The sequence diverges—it oscillates between values near $+1$ and near -1.

Check.

Bounded is not enough; convergence also needs the terms to settle, and here the alternating sign with no decay keeps them apart. A convergent sequence would drive *every* subsequence to the same limit. ✓

A recursive sequence by monotone convergence. A sequence is defined by $a_1 = \sqrt{2}$ and $a_{n+1} = \sqrt{2 + a_n}$. Find its limit.

Read.

Argue the sequence is increasing and bounded above—so it converges by monotone convergence—then solve the fixed-point equation $L = g(L)$.

Plan.

The terms increase and stay below 2 (if $a_n < 2$ then $a_{n+1} = \sqrt{2 + a_n} < \sqrt{4} = 2$, and one checks $a_{n+1} > a_n$), so by monotone convergence a limit L exists. Since both a_n and a_{n+1} tend to L, pass to the limit in the recursion:

$$\begin{aligned} L &= \sqrt{2 + L} && \text{let } n \to \infty \text{ in } a_{n+1} = \sqrt{2 + a_n} \\ L^2 &= 2 + L && \text{square both sides} \\ L^2 - L - 2 &= 0 && \text{collect terms} \\ (L - 2)(L + 1) &= 0 && \text{factor} \end{aligned}$$

The roots are $L = 2$ and $L = -1$; since every $a_n > 0$, the limit is $L = 2$.

Solve.

The sequence converges to $L = 2$.

Check.

The negative root $L = -1$ is rejected because a sequence of positive terms cannot converge to a negative number—the fixed-point equation must be solved *and* filtered by the sequence's range. ✓

Infinite Series and Convergence Tests

An *infinite series* $\sum_{n=1}^{\infty} a_n$ is the limit of its partial sums $S_N = a_1 + \cdots + a_N$. When that limit exists the series converges to it; otherwise it diverges. There is no single method; instead there is a toolkit of *tests*, and fluency means recognizing at a glance which test the general term invites.

The core tests.

- **nth-term test:** if $a_n \not\to 0$, the series diverges. (Necessary, not sufficient.)
- **Geometric:** $\sum ar^n$ converges $\iff |r| < 1$, to $\frac{a}{1-r}$.

- **p-series:** $\sum \frac{1}{n^p}$ converges $\iff p > 1$.
- **Integral test:** for positive decreasing $a_n = f(n)$, the series and $\int_1^\infty f$ converge or diverge together.
- **Comparison / limit comparison:** bound or compare against a known series.
- **Ratio test:** if $\lim |a_{n+1}/a_n| = L$, converges for $L < 1$, diverges for $L > 1$.
- **Root test:** if $\lim |a_n|^{1/n} = L$, converges for $L < 1$.
- **Alternating series:** $\sum (-1)^n b_n$ with $b_n \downarrow 0$ converges.

Approach Read · Plan · Solve · Check

Read. Read the shape of the general term: Does $a_n \to 0$? Is there a geometric ratio r^n, a power $1/n^p$, a factorial or exponential, an alternating sign, or a whole term raised to the nth power?

Plan.

1. If $a_n \not\to 0$, stop — the nth-term test gives divergence.
2. Geometric or p-series: apply the benchmark directly. Factorials/exponentials: ratio test. Whole term to the nth power: root test.
3. Positive messy rational: comparison or limit comparison to a p-series. Positive, decreasing, easy antiderivative: integral test. Alternating: alternating-series test, then check the absolute series.

Solve. Carry out the chosen test's computation (a limit, an integral, or a comparison).

Check. If the ratio or root test returns $L = 1$ it is inconclusive — switch to comparison or the integral test.

Common traps — every wrong answer is one of these

Reads $a_n \to 0$ as convergence.
Concludes the series converges because its terms vanish. The condition is necessary, not sufficient — the harmonic series has $a_n \to 0$ yet diverges.

Treats $L = 1$ as an answer. 1. Computes the ratio or root limit and finds $L = 1$.

2. Declares convergence or divergence from it.
3. Fix: at $L = 1$ the test is silent; abandon it and use comparison, integral, or p-series reasoning.

Points the comparison the wrong way.
Bounds a term *above* by a divergent series (or below by a convergent one) and concludes nothing valid. Smaller-than-convergent proves convergence; larger-than-divergent proves divergence.

Skips the absolute check.
Stops at "the alternating series converges" without testing $\sum |a_n|$. Absolute convergence (the unsigned series converges) is a stronger, separate conclusion from conditional convergence.

Did You Know? *the harmonic series and the Basel problem*

That the harmonic series $\sum 1/n$ diverges was first proved around 1350 by Nicole Oresme, who grouped its terms into blocks each summing to more than $\frac{1}{2}$. Yet the closely related $\sum 1/n^2$ *does* converge—and pinning down its exact value, $\pi^2/6$, defeated the best mathematicians for decades until Euler cracked the famous "Basel problem" in 1734.

nth-term test. Does $\displaystyle\sum_{n=1}^{\infty} \frac{n}{2n+1}$ converge?

Read.

Check the term's limit first; if it is not zero, the series cannot converge.

Plan.

$\lim_{n\to\infty} \frac{n}{2n+1} = \frac{1}{2} \neq 0$. Since the terms do not shrink to zero, the partial sums grow without bound.

Solve.

Diverges by the nth-term test.

Check.

The nth-term test can only prove *divergence*; a term going to zero (as in the harmonic series) does not by itself prove convergence. ✓

Integral test. Does $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ converge?

Read.

The terms are positive and decreasing; compare with $\int_2^{\infty} \frac{\mathrm{d}x}{x \ln x}$.

Plan.

The terms are positive and decreasing, so apply the integral test. Let $u = \ln x$, $\mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$:

$$\begin{aligned} \int \frac{\mathrm{d}x}{x \ln x} &= \int \frac{\mathrm{d}u}{u} && \text{substitute } u = \ln x \\ &= \ln|u| = \ln(\ln x) && \text{integrate, restore } u \end{aligned}$$

Then evaluate the improper integral:

$$\begin{aligned} \int_2^\infty \frac{\mathrm{d}x}{x \ln x} &= \lim_{b\to\infty} \left[\ln(\ln b) - \ln(\ln 2)\right] && \text{apply the limits} \\ &= \infty && \ln(\ln b) \to \infty \end{aligned}$$

Solve.

The integral diverges, so the series diverges.

Check.

This series sits just past the convergence boundary—slower even than the harmonic series in appearance, yet still divergent. ✓

Limit comparison. Does $\displaystyle\sum_{n=1}^{\infty} \frac{2n+1}{n^3+n}$ converge?

Read.

For large n the term behaves like $\frac{2n}{n^3} = \frac{2}{n^2}$; compare with the $p = 2$ series.

Plan.

Take $b_n = \frac{1}{n^2}$ and compute the limit of the ratio:

$$\begin{aligned} \lim_{n\to\infty} \frac{a_n}{b_n} &= \lim_{n\to\infty} \frac{(2n+1)/(n^3+n)}{1/n^2} && \text{form the comparison ratio} \\ &= \lim_{n\to\infty} \frac{n^2(2n+1)}{n^3+n} && \text{multiply by } n^2 \\ &= \lim_{n\to\infty} \frac{2n^3+n^2}{n^3+n} = 2 && \text{leading terms dominate} \end{aligned}$$

a finite positive number.

Solve.

Since $\sum \frac{1}{n^2}$ converges ($p = 2 > 1$) and the limit is finite and positive, the given series converges too.

Check.

Limit comparison lets you replace a messy term by its dominant behavior and read off the verdict from a p-series. ✓

Derivation. *The geometric-series sum*

Let $S_N = a + ar + \cdots + ar^{N-1}$. Multiply by r: $rS_N = ar + \cdots + ar^N$. Subtract: $S_N - rS_N = a - ar^N$, so $S_N = a\,\frac{1-r^N}{1-r}$. If $|r| < 1$ then $r^N \to 0$ and the partial sums approach

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \qquad (|r| < 1); \tag{3.6}$$

if $|r| \geq 1$ the partial sums do not settle and the series diverges.

A geometric series. Evaluate $\displaystyle\sum_{n=0}^{\infty} 3\left(-\tfrac{1}{3}\right)^n$.

Read.

Identify a and r; check $|r| < 1$; apply the geometric sum (3.6).

Plan.

The first term ($n = 0$) is $a = 3$ and the ratio is $r = -\frac{1}{3}$, with $|r| = \frac{1}{3} < 1$. Apply the geometric sum:

$$\begin{aligned} \sum_{n=0}^{\infty} 3\left(-\tfrac{1}{3}\right)^n &= \frac{a}{1-r} = \frac{3}{1-(-1/3)} && a = 3,\ r = -\tfrac{1}{3} \\ &= \frac{3}{4/3} = 3 \cdot \frac{3}{4} = \frac{9}{4} && \text{simplify} \end{aligned}$$

Solve.

The sum is $\frac{9}{4}$.

Check.

The alternating sign makes the sum smaller than the first term 3; $\frac{9}{4} = 2.25$ is duly less. ✓

Direct comparison. Does $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n + 1}$ converge?

Read.

Bound the term above by a convergent geometric series.

Plan.

For every n, $2^n + 1 > 2^n$, so $\frac{1}{2^n+1} < \frac{1}{2^n}$. The series $\sum \frac{1}{2^n}$ is geometric with $r = \frac{1}{2} < 1$, hence convergent. A series of positive terms bounded above by a convergent one converges.

Solve.

Converges by comparison with $\sum \frac{1}{2^n}$.

Check.

Direct comparison needs the inequality to point the right way: smaller-than-convergent proves convergence; larger-than-divergent proves divergence. ✓

Root test. Does $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n$ converge?

Read.

An nth power throughout signals the root test; compute $\lim |a_n|^{1/n}$.

Plan.

An nth power throughout signals the root test:

$$\begin{aligned} |a_n|^{1/n} &= \frac{n}{2n+1} && \text{take the } n\text{th root} \\ &\to \frac{1}{2} && \text{as } n \to \infty \end{aligned}$$

Since $L = \frac{1}{2} < 1$, the root test gives convergence.

Solve.

Converges by the root test.

Check.

Whenever the whole term is raised to the nth power, the root test strips the exponent and usually decides at once. ✓

Misconception

The ratio and root tests are silent when $L = 1$. If $\lim |a_{n+1}/a_n| = 1$ (or $\lim |a_n|^{1/n} = 1$), the test draws *no* conclusion—the series may converge or diverge. The p-series make this vivid: $p = 2$ converges and $p = 1$ diverges, yet both give $L = 1$. When you hit $L = 1$, abandon the ratio/root test and switch to comparison or the integral test.

Derivation. *Why the ratio test works*

Suppose $\lim |a_{n+1}/a_n| = L < 1$. Pick any r with $L < r < 1$. For large n, the ratios stay below r, so $|a_{n+1}| < r|a_n|$ and hence $|a_{n+k}| < r^k|a_n|$: from some point on the terms are dominated by a convergent geometric series with ratio r. Comparison then forces convergence. If $L > 1$ the terms grow, so the series diverges by the nth-term test. At $L = 1$ the test is silent.

Ratio test. Does $\displaystyle\sum_{n=1}^{\infty} \frac{2^n}{n!}$ converge?

Read.

Factorials and exponentials both signal the ratio test; form a_{n+1}/a_n and take the limit.

Plan.

Form the ratio of consecutive terms:

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{2^{n+1}/(n+1)!}{2^n/n!} && \text{write } a_{n+1}/a_n \\ &= \frac{2^{n+1}}{2^n} \cdot \frac{n!}{(n+1)!} && \text{split the factors} \\ &= 2 \cdot \frac{1}{n+1} = \frac{2}{n+1} && \text{simplify} \end{aligned}$$

As $n \to \infty$ this $\to 0 = L < 1$, so the series converges.

Solve.

Converges by the ratio test (in fact the sum is $\mathrm{e}^2 - 1$).

Check.

The factorial in the denominator overwhelms any fixed exponential—$n!$ eventually grows faster than c^n for every constant c. ✓

Alternating series and conditional convergence. Classify $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$.

Read.

Check the alternating-series conditions for convergence, then test the absolute series to decide the type.

Plan.

The terms $b_n = \frac{1}{n}$ are positive, decreasing, and tend to 0, so by the alternating series test the series *converges*. But the absolute series $\sum \frac{1}{n}$ is the harmonic series, which *diverges*.

Convergence relies on the sign cancellation.

Solve.

Converges *conditionally*.

Check.

Absolute convergence would require $\sum |a_n|$ to converge; here it does not, so removing the signs breaks convergence—the hallmark of conditional convergence. ✓

Alternating-series error bound. The series $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converges to $\ln 2$. How many terms guarantee an error below 0.01?

Read.

For a convergent alternating series, the truncation error is at most the first omitted term b_{N+1}; solve $b_{N+1} \le 0.01$.

Plan.

The remainder after N terms satisfies $|R_N| \le b_{N+1} = \frac{1}{N+1}$. Require $\frac{1}{N+1} \le 0.01$, i.e. $N + 1 \ge 100$, so $N \ge 99$.

Solve.

99 terms suffice to pin $\ln 2$ to within 0.01.

Check.

The alternating-series bound is the rare error estimate you can state instantly—the next term *is* the error ceiling. Note the slow rate: conditional convergence buys accuracy dearly. ✓

Power Series, Taylor, and Maclaurin

A *power series* $\sum_{n=0}^{\infty} c_n (x-a)^n$ is a "polynomial of infinite degree." For each fixed x it is a numerical series, and the ratio test reveals a *radius of convergence* R: the series converges for $|x-a| < R$ and diverges for $|x-a| > R$ (the endpoints are checked separately). Within that interval the series defines a function, and the most important power series are the ones that reproduce familiar functions—their *Taylor series*.

Approach Read · Plan · Solve · Check

Read. Decide the task: a radius or interval of convergence, building a Taylor/Maclaurin series, or using a known series for a limit, approximation, or integral.

Plan.

1. For a radius, apply the ratio test to $|a_{n+1}/a_n|$, set the limit < 1, and solve for $|x-a| < R$; then check the two endpoints separately.
2. To build a series, compute the derivatives at the center and form $c_n = f^{(n)}(a)/n!$, or manipulate a memorized series.
3. To use a series, substitute into a known expansion and truncate, differentiate, or integrate termwise.

Solve. Carry out the computation — the ratio-test limit, the coefficients, or the substituted series.

Check. Verify even/odd symmetry or the first few terms; bound an approximation by the first omitted term.

Common traps — every wrong answer is one of these

Forgets to test the endpoints. 1. Applies the ratio test and finds $|x-a| < R$.

2. Reports the open interval as final.
3. Fix: the ratio test is silent at $|x-a| = R$; substitute each endpoint and test those two series by hand.

Confuses radius with interval.

Reports R when the question asks for the interval of convergence (or the reverse). The radius is a single number; the interval is $(a-R, a+R)$ with endpoints resolved.

Misplaces the factorial.

Uses $c_n = f^{(n)}(a)$ without dividing by $n!$. The Taylor coefficient is $f^{(n)}(a)/n!$ — the factorial is essential.

Expands about the wrong center.

Builds a Maclaurin ($a = 0$) series when the problem is centered at some $a \neq 0$. Match the center to the point named in the problem.

Radius of convergence. Find the radius of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{(x-2)^n}{n\,3^n}$.

Read.

Apply the ratio test to the terms in x; the series converges where the limiting ratio is below 1.

Plan.

With $a_n = \frac{(x-2)^n}{n\,3^n}$, apply the ratio test:

$$\begin{aligned}
\left|\frac{a_{n+1}}{a_n}\right| &= \frac{|x-2|^{n+1}}{(n+1)3^{n+1}} \cdot \frac{n\,3^n}{|x-2|^n} && \text{form the ratio} \\
&= \frac{|x-2|}{3} \cdot \frac{n}{n+1} && \text{cancel powers} \\
&\to \frac{|x-2|}{3} && \text{as } n \to \infty,\ \tfrac{n}{n+1} \to 1
\end{aligned}$$

Convergence requires $\frac{|x-2|}{3} < 1$, i.e. $|x-2| < 3$.

Solve.

Radius $R = 3$ (centered at $x = 2$).

Check.

The n in the denominator does not affect R (its ratio $\to 1$); only the 3^n sets the radius. ✓

Derivation. *The Taylor coefficient formula*

Suppose $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$ near a. Differentiating termwise k times and then setting $x = a$ kills every term except the kth, leaving $f^{(k)}(a) = k!\,c_k$. Hence $c_k = \frac{f^{(k)}(a)}{k!}$, giving the Taylor series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{3.7}$$

Centered at $a = 0$ it is called the Maclaurin series.

Maclaurin series worth memorizing. $\mathrm{e}^x = \sum_{n\ge 0} \frac{x^n}{n!}$; $\sin x = \sum_{n\ge 0} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$; $\cos x = \sum_{n\ge 0} \frac{(-1)^n x^{2n}}{(2n)!}$; $\frac{1}{1-x} = \sum_{n\ge 0} x^n$ $(|x| < 1)$; $\ln(1+x) = \sum_{n\ge 1} \frac{(-1)^{n+1} x^n}{n}$ $(|x| \le 1,\ x \ne -1)$.

Building a Maclaurin series. Find the Maclaurin series of $f(x) = \cos x$ from the coefficient formula.

Read.

Differentiate repeatedly, evaluate at 0, and assemble the Taylor series (3.7).

Plan.

The derivatives cycle: $f = \cos x$, $f' = -\sin x$, $f'' = -\cos x$, $f''' = \sin x$, then repeat. At 0 these are $1, 0, -1, 0, 1, \ldots$, so only even powers survive, alternating in sign: $c_0 = 1$, $c_2 = -\frac{1}{2!}$, $c_4 = \frac{1}{4!}$, ... Assembling, $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$.

Solve.

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

Check.

cos is even, and the series contains only even powers—consistent. ✓

A limit by series. Evaluate $\lim_{x \to 0} \dfrac{\sin x - x}{x^3}$ using a series.

Read.

Replace $\sin x$ by its Maclaurin series, subtract x, and divide.

Plan.

Replace $\sin x$ by its Maclaurin series:

$$\begin{aligned} \sin x - x &= -\frac{x^3}{6} + \frac{x^5}{120} - \cdots && \sin x = x - \tfrac{x^3}{6} + \tfrac{x^5}{120} - \cdots \\ \frac{\sin x - x}{x^3} &= -\frac{1}{6} + \frac{x^2}{120} - \cdots && \text{divide by } x^3 \\ &\to -\frac{1}{6} && \text{as } x \to 0 \end{aligned}$$

Solve.

The limit is $-\frac{1}{6}$.

Check.

This matches three applications of L'Hôpital's rule—the series does the same work in one step. ✓

The exponential series. Build the Maclaurin series of e^x and use it to approximate $e^{0.1}$.

Read.

Every derivative of e^x is e^x, worth 1 at 0; assemble the series, then truncate.

Plan.

Since $f^{(n)}(x) = \mathrm{e}^x$ and $f^{(n)}(0) = 1$ for all n, the coefficients are $c_n = \frac{1}{n!}$, giving $\mathrm{e}^x = \sum_{n\ge 0} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$. Put $x = 0.1$ and keep three terms:

$$\begin{aligned} \mathrm{e}^{0.1} &\approx 1 + 0.1 + \frac{0.01}{2} && \text{first three terms} \\ &= 1 + 0.1 + 0.005 = 1.105 && \text{evaluate} \end{aligned}$$

Solve.

$\mathrm{e}^{0.1} \approx 1.105$ (true value $1.10517\ldots$).

Check.

The first omitted term is $\frac{(0.1)^3}{6} \approx 0.00017$, bounding the error—the series lets you say not just the estimate but how good it is. ✓

Integrating a series termwise. Use $\frac{1}{1-x} = \sum_{n\ge 0} x^n$ to find a series for $\ln(1-x)$.

Read.

Recognize $\ln(1-x)$ as an antiderivative of $-\frac{1}{1-x}$; integrate the series term by term.

Plan.

Recognize $\ln(1-x)$ as an antiderivative and integrate the series term by term:

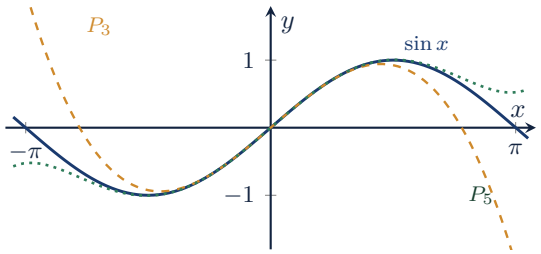
$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\ln(1-x) &= -\frac{1}{1-x} = -\sum_{n\ge 0} x^n && \text{geometric series} \\ \ln(1-x) &= -\sum_{n\ge 0} \frac{x^{n+1}}{n+1} + C && \text{integrate termwise} \\ &= -\sum_{m=1}^{\infty} \frac{x^m}{m} && C = 0 \text{ (from } x = 0),\ m = n+1 \end{aligned}$$

Solve.

$\ln(1-x) = -\sum_{m=1}^{\infty} \frac{x^m}{m}$ for $|x| < 1$.

Check.

Replacing x by $-x$ recovers the alternating series $\ln(1+x) = \sum (-1)^{m+1} x^m/m$ from the key-results box. ✓



Maclaurin polynomials converging to $\sin x$*:* $P_3 = x - \frac{x^3}{6}$ *(dashed) already tracks the curve near* 0*, and* $P_5 = x - \frac{x^3}{6} + \frac{x^5}{120}$ *(dotted) hugs it across a wider interval. Adding terms widens the match.*

Parametric and Polar Calculus

Some curves are described not as $y = f(x)$ but by a parameter—$x = x(t)$, $y = y(t)$—or by a radius as a function of angle, $r = r(\theta)$. Calculus adapts cleanly. For a parametric curve the arc length is $\int \sqrt{(\mathrm{d}x/\mathrm{d}t)^2 + (\mathrm{d}y/\mathrm{d}t)^2}\,\mathrm{d}t$. In polar coordinates, a thin wedge of angle $\mathrm{d}\theta$ is a near-triangle of area $\frac{1}{2}r^2\,\mathrm{d}\theta$, so the area swept is

$$A = \frac{1}{2}\int r^2\,\mathrm{d}\theta. \tag{3.8}$$

Approach Read · Plan · Solve · Check

Read. Identify the form — parametric $x = x(t)$, $y = y(t)$ or polar $r = r(\theta)$ — and the quantity wanted (arc length or area).

Plan.

1. Parametric arc length: compute $\dot{x}, \dot{y}$, form $\sqrt{\dot{x}^2 + \dot{y}^2}$, and integrate over the t-range.
2. Polar area: use $A = \frac{1}{2}\int r^2\,\mathrm{d}\theta$ over the θ-interval that traces the region exactly once.
3. Set the limits from the interval that generates the curve or region a single time.

Solve. Evaluate the integral.

Check. Compare to the known shape — a circle's circumference, a disk's area, or a semicircle.

Common traps — every wrong answer is one of these

Traces the curve twice. 1. Applies $\frac{1}{2}\int r^2\,\mathrm{d}\theta$ over $[0, 2\pi]$ for a circle like $r = 2\cos\theta$.

2. Sweeps the region twice and doubles the area.
3. Fix: choose the θ-interval that covers the region once (here $[-\frac{\pi}{2}, \frac{\pi}{2}]$).

Forgets the $\frac{1}{2}$.
Writes $\int r^2\,\mathrm{d}\theta$ for polar area. The wedge is a near-triangle of area $\frac{1}{2}r^2\,\mathrm{d}\theta$ — the factor $\frac{1}{2}$ is required.

Uses the Cartesian arc-length formula.
Applies $\int \sqrt{1+(f')^2}\,\mathrm{d}x$ to a parametric curve. Use $\int \sqrt{\dot{x}^2+\dot{y}^2}\,\mathrm{d}t$ when both coordinates depend on t.

Forgets to square r.
Integrates $\frac{1}{2}\int r\,\mathrm{d}\theta$ instead of $\frac{1}{2}\int r^2\,\mathrm{d}\theta$. The area element is quadratic in the radius.

Parametric arc length. Find the length of the curve $x = 3\cos t$, $y = 3\sin t$ for $t \in [0, \frac{\pi}{2}]$.

Read.

Compute $\dot{x}, \dot{y}$, form $\sqrt{\dot{x}^2+\dot{y}^2}$, and integrate in t.

Plan.

Here $\dot{x} = -3\sin t$ and $\dot{y} = 3\cos t$. Form the speed and integrate:

$$\begin{aligned} \dot{x}^2+\dot{y}^2 &= 9\sin^2 t + 9\cos^2 t = 9 && \sin^2 t+\cos^2 t = 1 \\ L &= \int_0^{\pi/2} \sqrt{\dot{x}^2+\dot{y}^2}\,\mathrm{d}t = \int_0^{\pi/2} 3\,\mathrm{d}t && \sqrt{9} = 3 \\ &= 3\cdot\frac{\pi}{2} = \frac{3\pi}{2} && \text{evaluate} \end{aligned}$$

Solve.

$L = \frac{3\pi}{2}$.

Check.

The curve is a quarter of a circle of radius 3, whose full circumference is 6π; one quarter is indeed $\frac{3\pi}{2}$. ✓

Polar area. Find the area enclosed by the circle $r = 2\cos\theta$.

Read.

Use the polar-area formula (3.8); the whole circle is traced as θ runs over $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Plan.

Apply the polar-area formula:

$$\begin{aligned} A &= \frac{1}{2}\int_{-\pi/2}^{\pi/2} (2\cos\theta)^2 \, \mathrm{d}\theta && r = 2\cos\theta \\ &= 2\int_{-\pi/2}^{\pi/2} \cos^2\theta \, \mathrm{d}\theta && (2\cos\theta)^2 = 4\cos^2\theta \\ &= 2 \cdot \frac{\pi}{2} = \pi && \int_{-\pi/2}^{\pi/2} \cos^2\theta \, \mathrm{d}\theta = \tfrac{\pi}{2} \end{aligned}$$

Solve.

$A = \pi$.

Check.

The polar equation $r = 2\cos\theta$ is a circle of radius 1 centered at $(1, 0)$, whose area is indeed $\pi(1)^2 = \pi$. ✓

Area of a cardioid. Find the area enclosed by the cardioid $r = 1 + \cos\theta$.

Read.

The whole cardioid is traced exactly once as θ runs over $[0, 2\pi]$; apply the polar-area formula (3.8) and power-reduce the $\cos^2$ term.

Plan.

Square the radius and integrate over a full turn:

$$\begin{aligned} A &= \frac{1}{2}\int_0^{2\pi} (1 + \cos\theta)^2 \, \mathrm{d}\theta && r = 1 + \cos\theta \\ &= \frac{1}{2}\int_0^{2\pi} (1 + 2\cos\theta + \cos^2\theta) \, \mathrm{d}\theta && \text{expand the square} \\ &= \frac{1}{2}(2\pi + 0 + \pi) && \textstyle\int_0^{2\pi} \cos\theta = 0, \ \int_0^{2\pi} \cos^2\theta = \pi \\ &= \frac{3\pi}{2} && \text{combine} \end{aligned}$$

Solve.

$A = \frac{3\pi}{2}$.

Check.

Unlike the circle $r = 2\cos\theta$, which retraces itself over $[0, 2\pi]$, the cardioid needs the *full* interval $[0, 2\pi]$ to close exactly once—matching the interval to a single tracing is the whole art. ✓

Connection

One idea, all chapter long. Every quantity here—an area, a volume, a length, a polar sweep, even a series sum—is the same limiting sum of vanishingly small pieces that defined the integral in Chapter 2. Choosing the piece (a strip, a disk, a shell, a wedge, a term) is the entire creative act; the integral then does the accumulation automatically.

Synthesis — Key Results You Must Own

1. Substitution undoes the chain rule; parts ($\int u\,dv = uv - \int v\,du$) handles products; trig powers use identities; $\sqrt{a^2 \pm x^2}$ calls for trig substitution; ratios of polynomials split by partial fractions.
2. Improper integrals are limits; $\int_1^\infty x^{-p}$ converges iff $p > 1$.
3. Area $= \int(\text{upper} - \text{lower})$; disk/washer $V = \pi\int(R^2 - r^2)$; shell $V = 2\pi\int x\,h$; arc length $\int\sqrt{1 + (f')^2}$; average $\frac{1}{b-a}\int f$.
4. A monotone bounded sequence converges.
5. Match each series to its test: nth-term (divergence), geometric/p-series (benchmarks), integral, comparison/limit comparison, ratio, root, alternating. Absolute $\Rightarrow$ conditional distinction.
6. Ratio test gives a power series' radius of convergence; endpoints checked separately.
7. Taylor: $c_n = f^{(n)}(a)/n!$; memorize the series for $e^x, \sin x, \cos x, \frac{1}{1-x}, \ln(1+x)$ and use them for limits and approximations.
8. Polar area $= \frac{1}{2}\int r^2\,d\theta$; parametric arc length $\int\sqrt{\dot{x}^2 + \dot{y}^2}\,dt$.

Practice Problems

Difficulty: ⋆ warm-up, ⋆⋆ standard, ⋆⋆⋆ challenge. Work each with the Read–Plan–Solve–Check habit modeled above.

Techniques of integration

1 ⋆.

Evaluate $\int 2x\,\mathrm{e}^{x^2}\,\mathrm{d}x$.

2 ⋆.

Evaluate $\int \dfrac{x}{x^2+4}\,\mathrm{d}x$.

3 ⋆⋆.

Evaluate $\int x\sin x\,\mathrm{d}x$.

4 ⋆⋆.

Evaluate $\int x^2\ln x\,\mathrm{d}x$.

Hint: Parts with $u=\ln x$.

5 ⋆⋆.

Evaluate $\int \cos^2 x\,\mathrm{d}x$.

6 ⋆⋆.

Evaluate $\int \sin^2 x\cos^3 x\,\mathrm{d}x$.

7 ⋆⋆.

Evaluate $\int \dfrac{\mathrm{d}x}{\sqrt{9-x^2}}$.

8 ⋆⋆⋆.

Evaluate $\int \dfrac{2x+3}{(x-1)(x+2)}\,\mathrm{d}x$.

Improper integrals

9 ⋆.

Evaluate $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{x^3}$.

10 ⋆⋆.

Evaluate $\displaystyle\int_0^1 \frac{\mathrm{d}x}{x^{2/3}}$ or show it diverges.

11 ⋆⋆⋆.

For which real p does $\displaystyle\int_1^\infty \frac{\mathrm{d}x}{x^p}$ converge? Justify.

Applications of integration

12 ⋆.

Find the area between $y = x^2$ and $y = 2x$.

13 ⋆⋆.

The region under $y = \frac{1}{x}$ on $[1, 3]$ is revolved about the x-axis. Find the volume.

14 ⋆⋆.

The region between $y = \sqrt{x}$ and $y = x$ is revolved about the x-axis. Find the volume.

15 ⋆⋆.

The region under $y = x^2$ on $[0, 1]$ is revolved about the y-axis. Find the volume by shells.

16 ⋆⋆.

Find the arc length of $y = \frac{1}{3}(x^2 + 2)^{3/2}$ from $x = 0$ to $x = 2$.

Hint: $1 + (y')^2$ is a perfect square.

17 ⋆⋆.

Find the average value of $f(x) = \sin x$ on $[0, \pi]$.

18 ⋆⋆⋆.

A spring with force constant $k = 200$ N/m is stretched from its natural length to 0.1 m. Find the work done.

Sequences and series

19 ⋆.

Does $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3}$ converge? Name the test.

20 ⋆.

Evaluate $\displaystyle\sum_{n=0}^{\infty} 3\left(\tfrac{1}{4}\right)^n$.

21 ⋆⋆.

Test $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^2+1}$ for convergence.

22 ⋆⋆.

Test $\displaystyle\sum_{n=1}^{\infty} \frac{3^n}{n!}$ for convergence.

23 ⋆⋆.

Classify $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$ as absolutely convergent, conditionally convergent, or divergent.

24 ⋆⋆⋆.

Test $\displaystyle\sum_{n=1}^{\infty} \frac{(n!)^2}{(2n)!}$ for convergence.

Hint: Ratio test; simplify $(2n+2)!/(2n)!$.

25 ⋆⋆.

Does $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^2}$ converge? Use the integral test.

Power series, Taylor, polar

26 ⋆⋆.

Find the radius of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{n^2}$.

27 ⋆⋆.

Find the interval of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{(x-1)^n}{n\,2^n}$, endpoints included.

28 ⋆⋆.

Use the Maclaurin series of $\cos x$ to approximate $\cos(0.2)$ with three nonzero terms.

29 ⋆⋆⋆.

Use a known series to evaluate $\displaystyle\lim_{x\to 0} \frac{\mathrm{e}^x - 1 - x}{x^2}$.

30 ⋆⋆⋆.

Find the area enclosed by one petal of the rose $r = \sin 2\theta$.

Hint: One petal is traced as θ runs from 0 to $\frac{\pi}{2}$.

Further Reading

The convergence tests reward practice more than reading; drill until the right test is obvious from the term's shape. For the theory beneath the tests—why a bounded monotone sequence must converge, and what a power series really is—the opening chapters of a real-analysis text (the completeness of $\mathbb{R}$, sequences, and series of functions) prepare the analysis material of a later chapter. Chapter 4 turns to differential equations, where these integration techniques and series become the tools for solving them.

CHAPTER 4

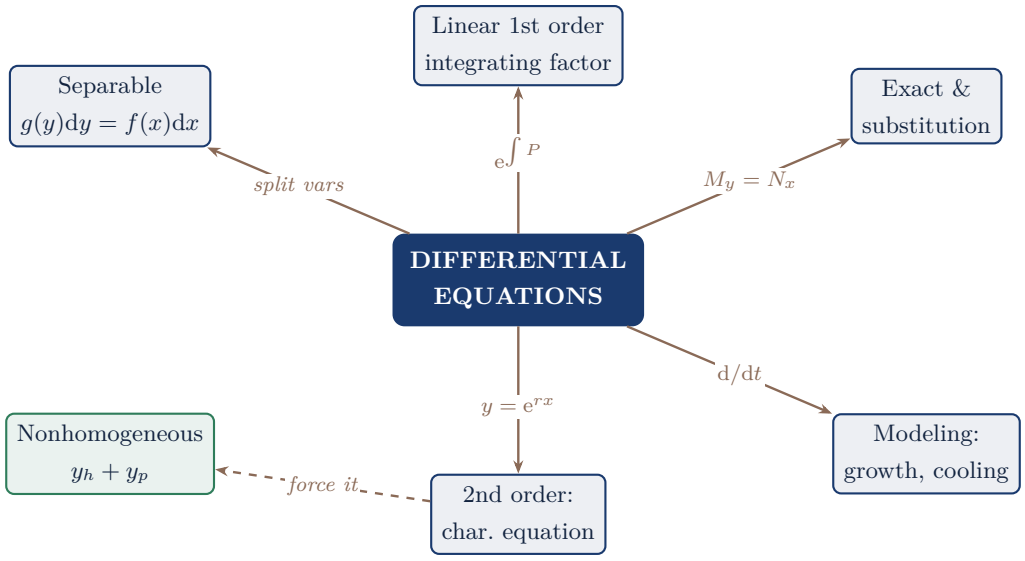
Differential Equations

Modeling How Things Change

"The profound study of nature is the most fertile source of mathematical discoveries."

— Joseph Fourier

A differential equation is an equation for an* unknown function *in which the function appears alongside its own derivatives. Where algebra solves for a number, a differential equation solves for a whole function—the position of a particle, the temperature of a cooling cup, the current in a circuit, the size of a population. The Subject Test rewards a small, sharp toolkit. For first-order equations you must recognize which of three shapes you are looking at—*separable*, *linear*, *or* exact*—because each has its own move. For second-order* linear *equations with constant coefficients, everything collapses to an algebra problem: guess $y = e^{rx}$, and the differential equation becomes the* characteristic *polynomial whose roots build the solution, with real, repeated, and complex roots giving exponentials, x-times-exponentials, and decaying sinusoids respectively. Add a forcing term and the general solution splits cleanly into a homogeneous part plus one particular solution. Master the recognition step and these problems are fast; miss it and you flail. Every technique here is, at bottom, an integration from Chapters 2 and 3 wearing a new hat.



Learning Objectives

You will be able to:

1. classify an ODE by order and linearity and verify a proposed solution;
2. solve separable, first-order linear, and exact equations, with initial conditions;
3. translate growth/decay, Newton cooling, and mixing problems into an ODE and solve;
4. solve second-order linear constant-coefficient homogeneous equations via the characteristic equation, handling all three root cases;
5. find particular solutions by undetermined coefficients and combine them with the homogeneous solution to solve initial-value problems;
6. read a differential equation as a physical model (oscillator, circuit).

What a Differential Equation Is

A *differential equation* relates an unknown function $y(x)$ to its derivatives. Its *order* is the highest derivative that appears: $y' = 2y$ is first order, $y'' + y = 0$ is second order. It is *linear* if y and its derivatives appear only to the first power and are never multiplied together or fed into nonlinear functions—so $y'' + 3xy' + 2y = \sin x$ is linear (the coefficients may depend on x freely), while $y' = y^2$ and $yy' = x$ are not. Linearity is the single most useful thing to notice, because the whole theory of the second half of this chapter applies only to linear equations.

A *solution* is any function that satisfies the equation identically. Because integration introduces arbitrary constants, a first-order equation has a one-parameter family of solutions—its *general*

solution—and a second-order equation a two-parameter family. Pinning down the constants requires side conditions: an *initial-value problem* (IVP) supplies the value of y (and, for second order, of y') at one point, selecting one curve from the family.

What & why. What kind of equation you are looking at—its order, whether it is linear, and how many constants its general solution carries—because that decides which method applies.

Solve. Count the highest derivative for order; check that y and its derivatives appear linearly; expect one arbitrary constant per order.

Means. A differential equation solves for a function, not a number.

Approach Read · Plan · Solve · Check

Read. Decide what is being asked — the *order*, whether the equation is *linear*, or whether a proposed function is a *solution* — and read off the pieces you need.

Plan.

1. For order, find the highest derivative that appears.
2. For linearity, check that $y, y', y'', \ldots$ each appear to the first power and are never multiplied together or fed into a nonlinear function; the coefficients may depend on x freely.
3. To verify a candidate, differentiate it as many times as the order and substitute into the equation.

Solve. Count or substitute one step at a time; simplify the substituted expression.

Check. A genuine solution makes the equation hold *identically* — true for all x, not just one point. Expect one arbitrary constant per order in a general solution.

Common traps — every wrong answer is one of these

Judges linearity by how x enters.
Calls $x^2y'' - xy' + y = 0$ nonlinear because of the x^2. Linearity is about how y and its derivatives enter; variable coefficients in x are still linear.

Confuses order with power.
Reads $(y')^2 = x$ as "second order" because of the square. Order is the highest *derivative* that appears, not the highest power.

Verifies at a single point.
Plugs one x-value into a candidate, gets 0, and declares it a solution. A solution must satisfy the equation identically.

Miscounts the arbitrary constants.

1. Writes a first-order general solution with two constants (or a second-order one with a single constant).
2. Then cannot match the initial data.
3. Fix: expect exactly one arbitrary constant per order — one for first order, two for second.

Verifying a solution. Show that $y = \mathrm{e}^{3x}$ solves $y'' - 9y = 0$, and that $y = \mathrm{e}^{2x}$ does not.

Read.

Differentiate the candidate twice and substitute into the equation; a solution makes it identically zero.

Plan.

For $y = \mathrm{e}^{3x}$: $y' = 3\mathrm{e}^{3x}$ and $y'' = 9\mathrm{e}^{3x}$. Substitute:

$$\begin{aligned} y'' - 9y &= 9\mathrm{e}^{3x} - 9\mathrm{e}^{3x} && \text{using } y'' = 9\mathrm{e}^{3x},\ y = \mathrm{e}^{3x} \\ &= 0 && \text{holds for all } x \end{aligned}$$

For $y = \mathrm{e}^{2x}$: $y'' = 4\mathrm{e}^{2x}$, so

$$\begin{aligned} y'' - 9y &= 4\mathrm{e}^{2x} - 9\mathrm{e}^{2x} && \text{using } y'' = 4\mathrm{e}^{2x},\ y = \mathrm{e}^{2x} \\ &= -5\mathrm{e}^{2x} \neq 0 && \text{not identically zero} \end{aligned}$$

Solve.

e^{3x} is a solution; e^{2x} is not.

Check.

The equation only tolerates exponentials whose rate r satisfies $r^2 = 9$—a preview of the characteristic equation. ✓

Classifying equations. Give the order and state whether each is linear: (a) $y'' + 3y' + 2y = \cos x$; (b) $y' = y^2 + x$; (c) $x^2y'' - xy' + y = 0$.

Read.

Find the highest derivative for the order; check that y, y', y'' appear to the first power and are not multiplied or composed.

Plan.

(a) Highest derivative y'': second order. Every y-term is first power with x-free constant coefficients: *linear*. (b) Highest is y': first order. The term y^2 is a square of the unknown: *nonlinear*. (c) Highest is y'': second order. Coefficients $x^2, -x, 1$ depend on x but multiply y-terms to the first power: *linear* (variable coefficients are still linear).

Solve.

(a) 2nd order, linear; (b) 1st order, nonlinear; (c) 2nd order, linear.

Check.

Linearity is about how y enters, not how x does—a common source of confusion. ✓

Counting the constants of a general solution. Verify that $y = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}$ solves $y'' - y = 0$ for *every* pair of constants C_1, C_2, and explain why no third constant is needed.

Read.

Differentiate the two-parameter family twice and substitute; then match the number of constants to the order.

Plan.

Differentiate twice: $y' = C_1\mathrm{e}^{x} - C_2\mathrm{e}^{-x}$ and $y'' = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}$. Substitute into the equation:

$$\begin{aligned} y'' - y &= \left(C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}\right) - \left(C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}\right) && \text{using } y'' = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x},\ y = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x} \\ &= 0 && \text{identically, for all } x,\ C_1,\ C_2 \end{aligned}$$

The equation is second order, so its general solution carries *exactly two* arbitrary constants—the two shown—and initial data $y(0)$, $y'(0)$ pins them down.

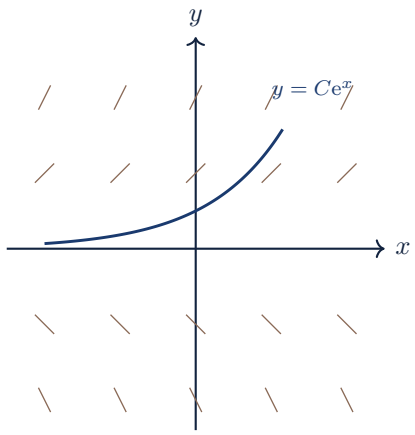
Solve.

$y = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}$ is the general solution: two constants, no more.

Check.

A third constant would be redundant: two independent exponentials already span every solution of a second-order linear equation, and only two initial conditions are available to fix constants. Writing three would leave one unpinned—the miscount the traps warn against. ✓

A *direction field* (or slope field) plots, at a grid of points, a short segment whose slope is the value the equation assigns to y' there. A solution curve must be tangent to the field at every point it passes through, so the field lets you see the family of solutions at a glance without solving anything.



A direction field for $y' = y$: each segment has slope equal to y at that point. The solution curve $y = Ce^x$ (navy) is tangent to every segment it crosses. Changing C slides the curve to a different member of the family.

First-Order Separable Equations

The friendliest first-order equations are *separable*: those that can be rearranged so that all the y's (with $\mathrm{d}y$) sit on one side and all the x's (with $\mathrm{d}x$) on the other, $g(y)\,\mathrm{d}y = f(x)\,\mathrm{d}x$. Once separated, you integrate both sides independently and solve for y. A single constant of integration—write it once, on the right—carries the whole family.

Approach Read · Plan · Solve · Check

Read. Test whether the equation is separable: can you get every y (with $\mathrm{d}y$) on one side and every x (with $\mathrm{d}x$) on the other, $g(y)\,\mathrm{d}y = f(x)\,\mathrm{d}x$?

Plan.

1. Algebraically separate the variables.
2. Integrate both sides, writing a single constant C on the right.
3. Solve for y where possible (exponentiate, take a root), or leave the relation implicit.
4. Apply any initial condition last, and choose the correct branch.

Solve. Carry out each integration one line at a time; keep the constant.

Check. Differentiate your answer back into the original equation, and confirm the initial condition. Note whether dividing by y discarded a constant solution.

Common traps — every wrong answer is one of these

Forgets the constant of integration.
Integrates both sides but drops C, collapsing the whole solution family to one curve. Always add one constant (on the right side alone).

Loses the equilibrium solution.
Divides by y (or by $g(y)$) to separate without noting that $y = 0$, where the divisor vanishes, may itself solve the equation. Check the constant function separately.

Exponentiates a sum wrongly.
Turns $\ln|y| = x^3 + C$ into $y = \mathrm{e}^{x^3} + C$ instead of $y = \mathrm{e}^{x^3+C} = A\mathrm{e}^{x^3}$. The constant rides as a multiplicative factor, not an added term.

Fixes the constant too early or takes the wrong branch.
1. Applies the initial condition before finishing the integration.
2. Or, after $y^2 = x^2 + C$, keeps both roots at an initial value that selects only one.
3. Fix: integrate fully, then use the initial condition to pin C and choose the sign of the root.

A basic separable equation. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = 3x^2y$.

Read.

Separate the variables, integrate both sides, then solve for y.

Plan.

Divide by y and multiply by $\mathrm{d}x$: $\frac{\mathrm{d}y}{y} = 3x^2\,\mathrm{d}x$. Integrate both sides to get $\ln|y| = x^3 + C$, then exponentiate:

$$\begin{aligned} |y| &= \mathrm{e}^{x^3+C} && \text{exponentiate } \ln|y| = x^3 + C \\ &= \mathrm{e}^C\mathrm{e}^{x^3} && \text{split the exponential of a sum} \\ y &= A\mathrm{e}^{x^3} && A = \pm\mathrm{e}^C, \text{ an arbitrary nonzero constant} \end{aligned}$$

and $y = 0$ is also a solution (lost when we divided by y).

Solve.

$y = A\mathrm{e}^{x^3}$.

Check.

Differentiate back: $y' = A\mathrm{e}^{x^3} \cdot 3x^2 = 3x^2y$. ✓

A separable IVP. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x}{y}$ with $y(0) = 2$.

Read.

Separate, integrate to get an implicit relation, then use the initial condition to fix the constant and choose the correct branch.

Plan.

Separate: $y\,\mathrm{d}y = x\,\mathrm{d}x$. Integrate both sides:

$$\frac{y^2}{2} = \frac{x^2}{2} + C \qquad \text{integrate each side}$$
$$y^2 = x^2 + C_1 \qquad \text{multiply by 2, } C_1 = 2C$$
$$4 = 0 + C_1 \qquad \text{apply } y(0) = 2$$
$$y^2 = x^2 + 4 \qquad \text{so } C_1 = 4$$

Since $y(0) = 2 > 0$, take the positive root: $y = \sqrt{x^2 + 4}$.

Solve.

$y = \sqrt{x^2 + 4}$.

Check.

At $x = 0$, $y = 2$. ✓ The negative root $-\sqrt{x^2 + 4}$ solves the ODE too but fails the initial condition.

Separable giving a tangent. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = 1 + y^2$ with $y(0) = 0$.

Read.

Separate, recognize the y-integral as an arctangent, then invert.

Plan.

Separate: $\frac{\mathrm{d}y}{1+y^2} = \mathrm{d}x$. Integrate both sides:

$$\arctan y = x + C \qquad \int \frac{\mathrm{d}y}{1 + y^2} = \arctan y$$
$$\arctan 0 = 0 + C \qquad \text{apply } y(0) = 0$$
$$C = 0 \qquad \text{since } \arctan 0 = 0$$
$$\arctan y = x \qquad \text{so the relation is}$$
$$y = \tan x \qquad \text{invert}$$

Solve.

$y = \tan x$.

Check.

The solution blows up at $x = \pm\frac{\pi}{2}$—even a tame-looking equation can have a solution that exists only on a finite interval. ✓

Separating an exponential. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{\,x-y}$.

Read.

Split the exponential of a difference into a quotient, separate, and integrate.

Plan.

Write $\mathrm{e}^{x-y} = \mathrm{e}^{x}\mathrm{e}^{-y}$, so $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^{x}\mathrm{e}^{-y}$. Separate to $\mathrm{e}^{y}\,\mathrm{d}y = \mathrm{e}^{x}\,\mathrm{d}x$ and integrate:

$$\begin{aligned} \mathrm{e}^{y} &= \mathrm{e}^{x} + C && \text{integrate both sides} \\ y &= \ln(\mathrm{e}^{x} + C) && \text{take logarithms} \end{aligned}$$

Solve.

$y = \ln\!\left(\mathrm{e}^{x} + C\right)$.

Check.

Differentiating, $y' = \frac{\mathrm{e}^{x}}{\mathrm{e}^{x}+C} = \frac{\mathrm{e}^{x}}{\mathrm{e}^{y}} = \mathrm{e}^{x-y}$. ✓

Misconception

Don't forget the constant—and put it in the right place. Integrate *both* sides, but a single constant on one side captures the whole family; writing C on both sides and subtracting is harmless but pointless. And when you divide by y to separate, note that $y = 0$ may be a lost solution—check whether the constant function solves the original equation.

First-Order Linear Equations: The Integrating Factor

A first-order *linear* equation can always be written in the standard form

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)\,y = Q(x). \tag{4.1}$$

It is generally not separable, but it always yields to one idea: multiply through by a cleverly chosen *integrating factor* $\mu(x)$ that turns the left side into the derivative of a single product. The right factor is

$$\mu = \mathrm{e}^{\int P\,\mathrm{d}x}. \tag{4.2}$$

Derivation. *the integrating factor* $\mu = \mathrm{e}^{\int P\,\mathrm{d}x}$

We want a factor $\mu(x)$ so that $\mu y' + \mu P y$ is exactly $\frac{\mathrm{d}}{\mathrm{d}x}(\mu y)$. By the product rule, $\frac{\mathrm{d}}{\mathrm{d}x}(\mu y) = \mu y' + \mu' y$. Matching the y-terms forces $\mu' = \mu P$, a separable equation for μ: $\frac{\mathrm{d}\mu}{\mu} = P\,\mathrm{d}x$, so $\ln\mu = \int P\,\mathrm{d}x$ and $\mu = \mathrm{e}^{\int P\,\mathrm{d}x}$. With this μ, the equation becomes $\frac{\mathrm{d}}{\mathrm{d}x}(\mu y) = \mu Q$, which we integrate directly.

Problem-solving strategy — First-order linear

(1) Put the equation in standard form $y' + P(x)y = Q(x)$. (2) Compute $\mu = \mathrm{e}^{\int P\,\mathrm{d}x}$. (3) The left side is now $(\mu y)'$; write it that way. (4) Integrate both sides and divide by μ. (5) Apply any initial condition last.

Approach Read · Plan · Solve · Check

Read. Recognize the first-order linear shape $y' + P(x)y = Q(x)$, and identify P and Q only after the coefficient of y' is 1.

Plan.

1. Divide through so the equation is in standard form $y' + P(x)y = Q(x)$.
2. Compute the integrating factor $\mu = \mathrm{e}^{\int P\,\mathrm{d}x}$.
3. Rewrite the left side as the single derivative $(\mu y)'$.
4. Integrate both sides, then divide by μ — dividing the constant too.
5. Apply any initial condition last.

Solve. Perform the integration $\int \mu Q\,\mathrm{d}x$ carefully; keep $+C$ before dividing.

Check. Differentiate the result and substitute into the original equation; confirm any initial condition.

Common traps — every wrong answer is one of these

Reads off P before standardizing.
Uses the coefficient of y as P while the coefficient of y' is still not 1 (as in $xy' + y = x^2$). Divide to standard form *first*, then identify P.

Builds the wrong integrating factor.
Writes $\mu = \int P\,\mathrm{d}x$ without the exponential, or forms $\mu = \mathrm{e}^{\int Q}$ from the right-hand side. The factor is $\mathrm{e}^{\int P\,\mathrm{d}x}$, exponential of the integral of P alone.

Integrates only one term.
After $\mathrm{e}^{2x}y = \frac{1}{3}\mathrm{e}^{3x} + C$, divides only the first term by μ, dropping the $C\mathrm{e}^{-2x}$ that carries the homogeneous solution.

Applies the initial condition too soon.
1. Substitutes the initial value before dividing by μ or before adding C.
2. Then solves for a constant that no longer matches the general solution.
3. Fix: reach the full general solution $y = \cdots + C\,(\cdots)$ first, and impose the condition only at the end.

Constant-coefficient linear. Solve $y' + 2y = \mathrm{e}^{x}$.

Read.

Read off $P = 2$, form μ, collapse the left side to a product derivative, integrate.

Plan.

Here $P = 2$, so $\mu = \mathrm{e}^{\int 2\,\mathrm{d}x} = \mathrm{e}^{2x}$. Multiply through and collapse the left side:

$$\begin{aligned}
\mathrm{e}^{2x}y' + 2\mathrm{e}^{2x}y &= \mathrm{e}^{2x}\mathrm{e}^{x} && \text{multiply by } \mu = \mathrm{e}^{2x}\\
(\mathrm{e}^{2x}y)' &= \mathrm{e}^{3x} && \text{left side is the product derivative}\\
\mathrm{e}^{2x}y &= \tfrac{1}{3}\mathrm{e}^{3x} + C && \text{integrate both sides}\\
y &= \tfrac{1}{3}\mathrm{e}^{x} + C\mathrm{e}^{-2x} && \text{divide by } \mathrm{e}^{2x}
\end{aligned}$$

Solve.

$y = \frac{1}{3}\mathrm{e}^{x} + C\mathrm{e}^{-2x}$.

Check.

The term $C\mathrm{e}^{-2x}$ solves the homogeneous equation $y' + 2y = 0$; $\frac{1}{3}\mathrm{e}^{x}$ is one particular solution. Their sum is the general solution—a pattern that returns for second-order equations. ✓

Variable-coefficient linear with an IVP. Solve $xy' + y = x^2$ for $x > 0$ with $y(1) = 2$.

Read.

Divide to reach standard form, note the integrating factor is simply x, integrate, then use

$y(1) = 2$.

Plan.

Divide by x: $y' + \frac{1}{x}y = x$, so $P = \frac{1}{x}$ and $\mu = \mathrm{e}^{\int \frac{1}{x}\,\mathrm{d}x} = \mathrm{e}^{\ln x} = x$. Multiply through and integrate:

$$\begin{aligned} (xy)' &= x^2 && \text{multiply by } \mu = x;\ \text{left side is } (xy)' \\ xy &= \frac{x^3}{3} + C && \text{integrate both sides} \\ y &= \frac{x^2}{3} + \frac{C}{x} && \text{divide by } x \\ 2 &= \frac{1}{3} + C && \text{apply } y(1) = 2 \end{aligned}$$

so $C = \frac{5}{3}$.

Solve.

$y = \frac{x^2}{3} + \frac{5}{3x}$.

Check.

The original equation was already $(xy)' = x^2$ before we divided—recognizing that shortcut saves time, but the integrating factor finds it mechanically. ✓

An integrating factor that is a Gaussian. Solve $y' - 2xy = x$.

Read.

Identify $P = -2x$, compute $\mu = \mathrm{e}^{\int P}$, and integrate the right side by substitution.

Plan.

Here $P = -2x$, so $\mu = \mathrm{e}^{\int -2x\,\mathrm{d}x} = \mathrm{e}^{-x^2}$. Multiply to get $(\mathrm{e}^{-x^2}y)' = x\mathrm{e}^{-x^2}$, and integrate the right side with $u = -x^2$, $\mathrm{d}u = -2x\,\mathrm{d}x$:

$$\begin{aligned} \mathrm{e}^{-x^2}y &= \int x\mathrm{e}^{-x^2}\,\mathrm{d}x && \text{integrate } (\mathrm{e}^{-x^2}y)' \\ &= -\tfrac{1}{2}\mathrm{e}^{-x^2} + C && \text{substitution } u = -x^2 \\ y &= -\tfrac{1}{2} + C\mathrm{e}^{x^2} && \text{divide by } \mathrm{e}^{-x^2} \end{aligned}$$

Solve.

$y = -\frac{1}{2} + C\mathrm{e}^{x^2}$.

Check.

Check: $y' = 2Cxe^{x^2}$ and $-2xy = x - 2Cxe^{x^2}$, so $y' - 2xy = x$. ✓

Exact Equations and Substitutions

Write a first-order equation in the differential form $M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y = 0$. It is *exact* when its left side is the total differential of some function $F(x,y)$—that is, when $M = F_x$ and $N = F_y$. Because mixed partials commute, this happens precisely when

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}. \tag{4.3}$$

When the test passes, the solutions are the level curves $F(x,y) = C$, and F is recovered by integrating M in x and reconciling with N.

Approach Read · Plan · Solve · Check

Read. Put the equation in the form $M(x,y)\,\mathrm{d}x + N(x,y)\,\mathrm{d}y = 0$ and apply the exactness test $M_y = N_x$.

Plan.

1. If $M_y = N_x$, integrate M in x to get $F = \int M\,\mathrm{d}x + g(y)$, carrying an unknown $g(y)$.
2. Differentiate F in y and match to N to solve for $g'(y)$, then $g(y)$; the solution is $F(x,y) = C$.
3. If not exact, try an integrating factor, or a substitution ($v = y/x$ for homogeneous, $u = y^{1-n}$ for Bernoulli).

Solve. Do the partial integrations one at a time, tracking the y-only piece.

Check. Confirm $F_x = M$ and $F_y = N$; for a substitution, verify the reduced equation is now separable or linear.

Common traps — every wrong answer is one of these

Applies the test with the wrong partials.
Compares M_x to N_y instead of M_y to N_x. Exactness is $\partial M/\partial y = \partial N/\partial x$ — cross the roles.

Drops the arbitrary function $g(y)$.
Integrates M in x and forgets that the "constant" of integration is a whole function of y, losing the N-information that fixes it.

Forces exactness that is not there.
Integrates as if exact without checking, or after finding $(M_y - N_x)/N$ depends on y still calls it a function of x. Verify the test (or the factor's form) before proceeding.

Botches the substitution. 1. For $y' = f(y/x)$, sets $y = vx$ but forgets $y' = v + xv'$.
2. Or for Bernoulli $y' + Py = Qy^n$, uses $u = y^n$ instead of $u = y^{1-n}$.
3. Fix: differentiate the substitution correctly ($y' = v + xv'$) and use the exponent $1 - n$.

Testing and solving an exact equation. Solve $(2xy + 3)\,\mathrm{d}x + (x^2 - 1)\,\mathrm{d}y = 0$.

Read.

Check $M_y = N_x$; if equal, integrate M in x, then fix the y-only piece using N.

Plan.

Here $M = 2xy + 3$ and $N = x^2 - 1$. Test: $M_y = 2x$ and $N_x = 2x$—equal, so exact. Recover the potential F:

$$\begin{aligned} F &= \int (2xy + 3)\,\mathrm{d}x && \text{integrate } M \text{ in } x \\ &= x^2y + 3x + g(y) && g(y) \text{ unknown, constant in } x \\ F_y &= x^2 + g'(y) && \text{differentiate in } y \\ &= x^2 - 1 && \text{match to } N, \text{ so } g'(y) = -1 \end{aligned}$$

Hence $g(y) = -y$ and $F = x^2y + 3x - y$.

Solve.

The general solution is $x^2y + 3x - y = C$.

Check.

$F_x = 2xy + 3 = M$ and $F_y = x^2 - 1 = N$, confirming F is the potential. ✓

Sometimes a non-exact equation becomes exact after multiplying by an integrating factor, just as in (4.2). When $\frac{M_y - N_x}{N}$ depends on x alone—so the factor restores the exactness condition (4.3)—that factor is $\mu(x) = \mathrm{e}^{\int \frac{M_y - N_x}{N}\,\mathrm{d}x}$.

An integrating factor that restores exactness. Solve $(3xy + y^2)\,\mathrm{d}x + (x^2 + xy)\,\mathrm{d}y = 0$.

Read.

Test exactness; if it fails, check whether $(M_y - N_x)/N$ depends only on x, build μ, and

solve the now-exact equation.

Plan.

Here $M = 3xy + y^2$, $N = x^2 + xy$, with $M_y = 3x + 2y$ and $N_x = 2x + y$—not equal, so not exact. Form the test ratio:

$$\begin{aligned} \frac{M_y - N_x}{N} &= \frac{x+y}{x^2+xy} && M_y - N_x = (3x+2y) - (2x+y) = x+y \\ &= \frac{x+y}{x(x+y)} && \text{factor the denominator} \\ &= \frac{1}{x} && \text{a function of } x \text{ alone} \end{aligned}$$

so $\mu = \mathrm{e}^{\int \frac{1}{x}\,\mathrm{d}x} = x$. Multiply through by x: $(3x^2y + xy^2)\,\mathrm{d}x + (x^3 + x^2y)\,\mathrm{d}y = 0$, now exact ($M_y = N_x = 3x^2 + 2xy$). Integrate M in x: $F = x^3y + \frac{1}{2}x^2y^2 + g(y)$; matching $F_y = x^3 + x^2y = N$ gives $g'(y) = 0$.

Solve.

The general solution is $x^3y + \frac{1}{2}x^2y^2 = C$.

Check.

Differentiating F recovers the multiplied equation, and dividing by $\mu = x$ returns the original—so the level curves are the same solutions. ✓

Two *substitutions* rescue equations that are neither separable nor linear. A *homogeneous* equation, one expressible as $y' = f(y/x)$, becomes separable under $v = y/x$. A *Bernoulli* equation $y' + Py = Qy^n$ becomes linear under $u = y^{1-n}$.

A homogeneous substitution. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x+y}{x}$.

Read.

Rewrite the right side in terms of y/x, substitute $v = y/x$ so $y = vx$ and $y' = v + xv'$, then separate.

Plan.

Rewrite the right side and substitute $v = y/x$:

$$\begin{aligned} \frac{x+y}{x} &= 1 + \frac{y}{x} = 1 + v && \text{with } v = y/x \\ v + xv' &= 1 + v && \text{use } y = vx,\ y' = v + xv' \\ xv' &= 1 && \text{cancel the } v \\ v &= \ln|x| + C && \text{separate } \mathrm{d}v = \tfrac{\mathrm{d}x}{x} \text{ and integrate} \end{aligned}$$

Undo the substitution: $y = vx = x(\ln|x| + C)$.

Solve.

$y = x(\ln|x| + C)$.

Check.

The substitution killed the v on both sides, leaving a pure integration—the signature payoff of $v = y/x$ on a homogeneous equation. ✓

A Bernoulli equation. Solve $y' + y = y^2$.

Read.

This is Bernoulli with $n = 2$; substitute $u = y^{1-n} = y^{-1}$ to linearize, solve, then invert.

Plan.

Divide by y^2: $y^{-2}y' + y^{-1} = 1$. Let $u = y^{-1}$, so $u' = -y^{-2}y'$ and $y^{-2}y' = -u'$. The equation linearizes:

$$\begin{aligned} -u' + u &= 1 && \text{substitute } y^{-2}y' = -u',\ y^{-1} = u \\ u' - u &= -1 && \text{multiply by } -1 \\ (\mathrm{e}^{-x}u)' &= -\mathrm{e}^{-x} && \text{integrating factor } \mu = \mathrm{e}^{-x} \\ \mathrm{e}^{-x}u &= \mathrm{e}^{-x} + C && \text{integrate both sides} \\ u &= 1 + C\mathrm{e}^{x} && \text{divide by } \mathrm{e}^{-x} \end{aligned}$$

Invert: $y = \frac{1}{1+C\mathrm{e}^{x}}$.

Solve.

$y = \frac{1}{1+C\mathrm{e}^{x}}$.

Check.

This is the logistic-type curve; the substitution $u = y^{-1}$ is what turned a nonlinear equation into a linear one. ✓

Modeling with First-Order Equations

The reason first-order equations matter is that "rate of change proportional to something" is the most common law in the sciences. Three archetypes cover most exam questions.

Approach Read · Plan · Solve · Check

Read. Translate the words into a rate law $\frac{dQ}{dt} = \ldots$: proportional to the amount (growth/decay), to the gap from ambient (cooling), or "rate in minus rate out" (mixing).

Plan.

1. Write the governing ODE and its solution form ($N_0 e^{kt}$, $T_a + (T_0 - T_a)e^{-kt}$, or a linear-ODE solve).
2. Use the initial value to fix the amplitude constant.
3. Use a second data point to solve for the rate k.
4. Answer the specific question by evaluating at the required time.

Solve. Take logarithms to isolate k; keep the amplitude and k as separate unknowns.

Check. Test the limit $t \to \infty$ (or $t \to 0$) against physical sense — decay $\to 0$, cooling $\to$ ambient, mixing $\to$ inflow concentration.

Common traps — every wrong answer is one of these

Mishandles half-life and doubling.
Confuses the sign of k, or solves $e^{-5k} = \frac{1}{2}$ by dropping the logarithm. Half-life gives $k = \frac{\ln 2}{t_{1/2}}$; keep the exponential intact until you take ln.

Forgets to subtract the ambient temperature.
Models the raw temperature T as pure decay instead of the difference $T - T_a$. Newton cooling decays $T - T_a$, so $T = T_a + (T_0 - T_a)e^{-kt}$.

Uses the wrong outflow concentration.
In a mixing tank, multiplies the outflow rate by the *inflow* concentration instead of the current tank concentration S/V. The salt leaves at concentration $S(t)/V$ times the outflow rate.

Tries to find two constants from one fact.

1. Uses a single data point yet must determine both the amplitude and k.
2. Ends up with an unknown left over.
3. Fix: the initial value fixes the amplitude; a *second* measured value fixes k.

Exponential growth and decay

If a quantity changes at a rate proportional to its present amount, $\frac{\mathrm{d}N}{\mathrm{d}t} = kN$, then the solution is

$$N = N_0 \mathrm{e}^{kt}, \tag{4.4}$$

growth for $k > 0$ and decay for $k < 0$. Radioactive decay, unchecked population growth, and continuously compounded interest all live here.

Radioactive half-life. A sample decays by $\frac{\mathrm{d}N}{\mathrm{d}t} = -kN$ and loses half its mass in 5 years. What fraction remains after 12 years?

Read.

Solve the decay equation, use the half-life to find k, then evaluate at $t = 12$.

Plan.

The solution is $N = N_0 \mathrm{e}^{-kt}$. Half-life: $N_0/2 = N_0 \mathrm{e}^{-5k}$, so $\mathrm{e}^{-5k} = \frac{1}{2}$ and $k = \frac{\ln 2}{5}$. After 12 years,

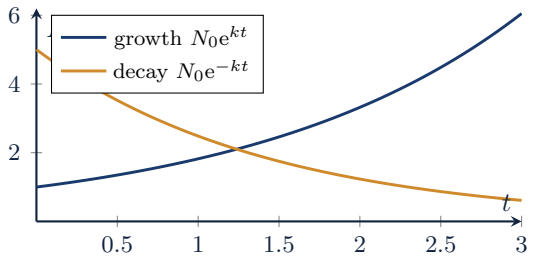
$$\begin{aligned} \frac{N}{N_0} &= \mathrm{e}^{-12k} && \text{evaluate the solution at } t = 12 \\ &= \mathrm{e}^{-\frac{12}{5}\ln 2} && \text{substitute } k = \tfrac{\ln 2}{5} \\ &= 2^{-12/5} \approx 0.189 && \text{using } \mathrm{e}^{a \ln 2} = 2^a \end{aligned}$$

Solve.

About $2^{-12/5} \approx 18.9\%$ remains.

Check.

Twelve years is between two and three half-lives (10 and 15 years), and 18.9% sits between 25% and 12.5%. ✓



The two faces of $\frac{dN}{dt} = kN$: exponential growth ($k > 0$, navy) accelerates without bound; exponential decay ($k < 0$, orange) falls toward zero, halving over each fixed interval.

Doubling time. A bacterial culture grows as $\frac{dP}{dt} = kP$ and triples in 6 hours. How long until it doubles?

Read.

Find k from the tripling data, then solve $e^{kt} = 2$ for the doubling time.

Plan.

From $P = P_0e^{kt}$ and tripling, $3 = e^{6k}$, so $k = \frac{\ln 3}{6}$. Doubling requires $e^{kt} = 2$, i.e. $kt = \ln 2$:

$$\begin{aligned} t &= \frac{\ln 2}{k} && \text{solve } kt = \ln 2 \\ &= \frac{6\ln 2}{\ln 3} && \text{substitute } k = \tfrac{\ln 3}{6} \\ &\approx \frac{6(0.693)}{1.099} \approx 3.79 \text{ hours} && \text{evaluate numerically} \end{aligned}$$

Solve.

About 3.8 hours.

Check.

Doubling is faster than tripling, so the answer should be under 6 hours—and $3.8 < 6$. ✓

Newton's law of cooling

An object at temperature T in surroundings at T_a cools at a rate proportional to the *difference*: $\frac{dT}{dt} = -k(T - T_a)$. Substituting $u = T - T_a$ makes it pure decay of the form (4.4), so

$$T = T_a + (T_0 - T_a)e^{-kt}, \tag{4.5}$$

and the temperature approaches ambient exponentially.

A cooling cup. Coffee at 90° sits in a 20° room and cools to 60° after 10 minutes. Find $T(t)$.

Read.

Write the cooling solution with $T_a = 20$, use the two data points to find the amplitude and k.

Plan.

The solution is $T = 20 + (90 - 20)\mathrm{e}^{-kt} = 20 + 70\mathrm{e}^{-kt}$. Use $T(10) = 60$:

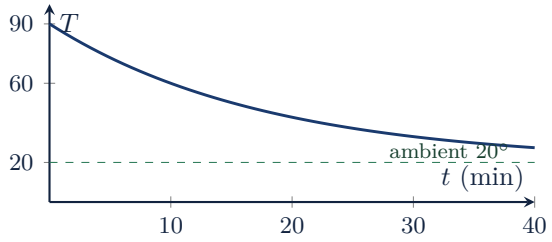
$$\begin{aligned} 60 &= 20 + 70\mathrm{e}^{-10k} && \text{apply } T(10) = 60 \\ \mathrm{e}^{-10k} &= \frac{40}{70} = \frac{4}{7} && \text{isolate the exponential} \\ k &= \frac{1}{10}\ln\frac{7}{4} \approx 0.0560 && \text{take logarithms (per minute)} \end{aligned}$$

Solve.

$T(t) = 20 + 70\,\mathrm{e}^{-kt}$ with $k = \frac{1}{10}\ln\frac{7}{4}$.

Check.

As $t \to \infty$, $T \to 20°$, the room temperature—exactly what the model should predict. ✓



Newton cooling: the temperature falls fast at first, when the gap to ambient is largest, then levels off, approaching the room temperature 20° asymptotically.

Mixing problems

In a well-stirred tank, the rate of change of dissolved substance is (rate in) minus (rate out), where each rate is a concentration times a flow. This produces a first-order linear equation in the standard form (4.1).

A mixing tank. A tank holds 100 L of pure water. Brine containing 3 g/L of salt flows in at 2 L/min, and the well-mixed solution flows out at 2 L/min. Find the salt $S(t)$.

Read.

Set up rate in minus rate out; the outflow concentration is $S/100$; solve the resulting linear ODE with $S(0) = 0$.

Plan.

Salt in: 3 g/L $\times$ 2 L/min = 6 g/min. Salt out: concentration $\frac{S}{100}$ times 2 L/min = $\frac{S}{50}$ g/min. So $\frac{dS}{dt} = 6 - \frac{S}{50}$, i.e. $S' + \frac{1}{50}S = 6$, with integrating factor $\mu = e^{t/50}$:

$$\begin{aligned} (e^{t/50}S)' &= 6e^{t/50} && \text{multiply by } \mu = e^{t/50} \\ e^{t/50}S &= 300e^{t/50} + C && \text{integrate both sides} \\ S &= 300 + Ce^{-t/50} && \text{divide by } e^{t/50} \\ 0 &= 300 + C && \text{apply } S(0) = 0 \end{aligned}$$

so $C = -300$.

Solve.

$S(t) = 300(1 - e^{-t/50})$ grams.

Check.

As $t \to \infty$, $S \to 300$ g, which is 3 g/L over 100 L—the tank saturates to the inflow concentration. ✓

Autonomous Equations and Equilibria

A first-order equation is *autonomous* when the right side depends on y alone, $\frac{dy}{dt} = f(y)$—time does not appear explicitly. Such equations model systems whose law is fixed, and they can be understood *qualitatively*, without solving, by asking where $f(y) = 0$. Those roots are the *equilibrium* (constant) solutions: start there and you stay. Between equilibria the sign of f tells you which way y moves—up where $f > 0$, down where $f < 0$—and that flow classifies each equilibrium as *stable* (neighbors flow in) or *unstable* (neighbors flow away). The *phase line* records this on a single vertical axis.

Approach Read · Plan · Solve · Check

Read. Confirm the equation is autonomous, $\frac{dy}{dt} = f(y)$ with no explicit t; the goal is the equilibria and their stability, not a formula for $y(t)$.

Plan.

1. Set $f(y) = 0$ to find every equilibrium (constant) solution.
2. Determine the sign of f on each interval between roots — up where $f > 0$, down where $f < 0$.
3. Classify each equilibrium: stable if neighbors flow in, unstable if they flow out; or use the test $f'(y^*) < 0 \Rightarrow$ stable, $f'(y^*) > 0 \Rightarrow$ unstable.

Solve. Read the flow off the sign chart or phase line; no integration is needed.

Check. Confirm the derivative test and the sign analysis agree, and that every root of f is accounted for.

Common traps — every wrong answer is one of these

Solves the ODE needlessly.
Grinds out an explicit solution when only equilibria and stability are asked — the phase line answers the question without integrating.

Reverses the stability reading.
Thinks $f > 0$ just above an equilibrium means it is stable. If $f > 0$ above, the flow points *away* upward, so that side is repelling.

Flips the derivative test.
Uses $f'(y^*) > 0$ for stable. It is the reverse: $f'(y^*) < 0$ is stable (restoring), $f'(y^*) > 0$ is unstable.

Misses an equilibrium.
1. Factors $f(y)$ carelessly and drops a root (e.g. finds $P = 100$ but not $P = 0$).
2. Then the phase line is incomplete.
3. Fix: find *all* roots of $f(y) = 0$ before classifying.

Equilibria of the logistic equation. For $\frac{\mathrm{d}P}{\mathrm{d}t} = 0.5\,P\left(1 - \frac{P}{100}\right)$, find the equilibria and their stability.

Read.
Set the right side to zero for equilibria; test the sign of $f(P)$ on each side to classify stability.

Plan.
Set $f(P) = 0.5P(1 - P/100) = 0$: the roots are $P = 0$ and $P = 100$. Test the sign of f: for $0 < P < 100$, both factors are positive so $f > 0$ (population rises); for $P > 100$, the factor

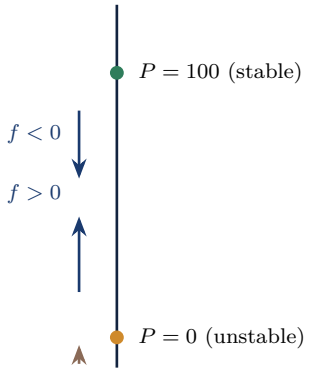
$1 - P/100 < 0$ so $f < 0$ (population falls); for $P < 0$, $f < 0$. Near $P = 0$ the flow points away (rising for $P > 0$): *unstable*. Near $P = 100$ the flow points inward from both sides: *stable*.

Solve.

$P = 0$ is unstable; $P = 100$ is a stable equilibrium (the carrying capacity).

Check.

Any positive starting population is drawn toward 100—the hallmark of logistic growth, read off without ever solving the equation. ✓



The phase line for the logistic equation. Arrows show the direction of flow: $P = 0$ (orange) repels, $P = 100$ (green) attracts. The sign of $f(P)$ between the equilibria determines every trajectory's fate.

Classifying an equilibrium by the derivative test. The equation $y' = y^2 - 4$ has equilibria $y = \pm 2$. Classify them.

Read.

Equilibria are roots of f; a quick test uses the sign of $f'(y)$ there—$f' < 0$ stable, $f' > 0$ unstable.

Plan.

Here $f(y) = y^2 - 4$, with roots $y = \pm 2$ and $f'(y) = 2y$. At $y = 2$, $f'(2) = 4 > 0$: neighbors flow away, *unstable*. At $y = -2$, $f'(-2) = -4 < 0$: neighbors flow in, *stable*. (Check by sign: for $-2 < y < 2$, $f < 0$, so y falls toward -2; for $y > 2$, $f > 0$, so y rises away from 2.)

Solve.

$y = -2$ is stable; $y = 2$ is unstable.

Check.

The linearization test $f'(y^*) < 0 \Rightarrow$ stable matches the sign analysis and is the fastest classifier. ✓

Three equilibria, and the root that is easy to miss. Find and classify all equilibria of $\frac{dy}{dt} = y^3 - y$.

Read.

Factor f *completely* so no root is dropped, then apply $f'(y^*) < 0 \Rightarrow$ stable, $f'(y^*) > 0 \Rightarrow$ unstable.

Plan.

Factor fully—the trap is to solve $y^2 - 1 = 0$ and miss the root $y = 0$:

$$\begin{aligned} f(y) &= y^3 - y && \text{the right-hand side} \\ &= y(y-1)(y+1) && \text{factor completely: roots } y = -1,\ 0,\ 1 \\ f'(y) &= 3y^2 - 1 && \text{differentiate for the stability test} \end{aligned}$$

Evaluate f' at each equilibrium: $f'(-1) = 2 > 0$, $f'(0) = -1 < 0$, and $f'(1) = 2 > 0$.

Solve.

$y = 0$ is stable; $y = -1$ and $y = 1$ are unstable.

Check.

The sign chart agrees: on $-1 < y < 0$, $f > 0$ pushes y up toward 0, and on $0 < y < 1$, $f < 0$ pushes y down toward 0—both sides flow inward. Dropping the middle root would have hidden the only stable state. ✓

Second-Order Linear Equations: The Characteristic Equation

The centerpiece of the chapter is the constant-coefficient second-order linear equation

$$ay'' + by' + cy = 0. \tag{4.6}$$

Because differentiating e^{rx} just multiplies it by r, the guess $y = e^{rx}$ turns the differential equation into an algebraic one.

Derivation. *the characteristic equation from $y = \mathrm{e}^{rx}$*

Substitute $y = \mathrm{e}^{rx}$, $y' = r\mathrm{e}^{rx}$, $y'' = r^2\mathrm{e}^{rx}$ into $ay'' + by' + cy = 0$: $(ar^2 + br + c)\mathrm{e}^{rx} = 0$. Since e^{rx} is never zero, the bracket must vanish: $ar^2 + br + c = 0$. This quadratic—the *characteristic equation*—has roots that dictate the solution. Because the equation is linear and homogeneous, any linear combination of solutions is again a solution, so two independent exponentials span the whole solution space.

The two roots r_1, r_2 fall into three cases.

The three root cases of $ar^2 + br + c = 0$.

- **Real and distinct** $r_1 \neq r_2$: $\quad y = C_1\mathrm{e}^{r_1 x} + C_2\mathrm{e}^{r_2 x}$.
- **Real and repeated** $r_1 = r_2 = r$: $\quad y = (C_1 + C_2 x)\mathrm{e}^{rx}$.
- **Complex** $r = \alpha \pm \beta\mathrm{i}$: $\quad y = \mathrm{e}^{\alpha x}\big(C_1 \cos \beta x + C_2 \sin \beta x\big)$.

Approach Read · Plan · Solve · Check

Read. Confirm the equation is constant-coefficient $ay''+by'+cy = 0$, and form the characteristic equation $ar^2 + br + c = 0$.

Plan.

1. Solve the quadratic for its roots r_1, r_2.
2. Real and distinct: $y = C_1\mathrm{e}^{r_1 x} + C_2\mathrm{e}^{r_2 x}$.
3. Real and repeated: $y = (C_1 + C_2 x)\mathrm{e}^{rx}$ — keep the extra factor of x.
4. Complex $\alpha \pm \beta\mathrm{i}$: $y = \mathrm{e}^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x)$.
5. Apply the two initial conditions last, differentiating for $y'(0)$.

Solve. Factor or use the quadratic formula; read off α, β carefully for complex roots.

Check. The general solution must carry *two* arbitrary constants; substitute back if unsure.

Common traps — every wrong answer is one of these

Drops the x for a repeated root.

Writes $y = C_1\mathrm{e}^{rx} + C_2\mathrm{e}^{rx} = (C_1 + C_2)\mathrm{e}^{rx}$, which is a *one*-parameter family — too few constants. A double root demands the second solution $x\mathrm{e}^{rx}$.

Mishandles complex roots.

Drops the envelope $\mathrm{e}^{\alpha x}$, or swaps α and β so the decay rate and the frequency trade places.

For $r = \alpha \pm \beta\mathrm{i}$, α is the exponential rate and β the frequency.

Sign-slips the quadratic.
Reads $r^2 + 4 = 0$ as $r = \pm 2$ (real) instead of $r = \pm 2\mathrm{i}$. A positive constant with no linear term gives purely imaginary roots.

Imposes the conditions on too little.
1. Applies $y(0)$ and $y'(0)$ before writing the full two-constant general solution.
2. Or forgets to differentiate the general solution to use $y'(0)$.
3. Fix: assemble the general solution with C_1, C_2 first, then differentiate and solve both conditions together.

Real distinct roots. Solve $y'' - 5y' + 6y = 0$.

Read.

Write the characteristic equation, factor for its roots, and assemble the exponential solution.

Plan.

Characteristic equation: $r^2 - 5r + 6 = 0$, which factors as $(r - 2)(r - 3) = 0$, so $r = 2, 3$. Two distinct real roots give $y = C_1\mathrm{e}^{2x} + C_2\mathrm{e}^{3x}$.

Solve.

$y = C_1\mathrm{e}^{2x} + C_2\mathrm{e}^{3x}$.

Check.

Substituting either exponential alone zeroes the equation; the general solution is their combination, carrying the two constants a second-order equation must have. ✓

A repeated root. Solve $y'' - 4y' + 4y = 0$.

Read.

Solve the characteristic equation; a double root forces the extra factor of x in the second solution.

Plan.

Characteristic equation: $r^2 - 4r + 4 = (r - 2)^2 = 0$, a repeated root $r = 2$. A single exponential e^{2x} is only one solution; the second independent solution is $x\mathrm{e}^{2x}$. Hence $y = (C_1 + C_2 x)\mathrm{e}^{2x}$.

Solve.

$y = (C_1 + C_2 x)\mathrm{e}^{2x}$.

Check.

Without the x-factor we would have only a one-parameter family—too few constants for a second-order equation. ✓

Derivation. *why complex roots give* $\mathrm{e}^{\alpha x}(\cos \beta x, \sin \beta x)$

If the roots are $r = \alpha \pm \beta \mathrm{i}$, the raw solutions are $\mathrm{e}^{(\alpha+\beta\mathrm{i})x}$ and $\mathrm{e}^{(\alpha-\beta\mathrm{i})x}$. Factor out $\mathrm{e}^{\alpha x}$ and apply Euler's formula $\mathrm{e}^{\mathrm{i}\beta x} = \cos \beta x + \mathrm{i} \sin \beta x$: $\mathrm{e}^{(\alpha\pm\beta\mathrm{i})x} = \mathrm{e}^{\alpha x}(\cos \beta x \pm \mathrm{i} \sin \beta x)$. Taking real combinations—the sum and the difference over 2i—yields the two *real* solutions $\mathrm{e}^{\alpha x} \cos \beta x$ and $\mathrm{e}^{\alpha x} \sin \beta x$. The real part α sets exponential growth or decay; the imaginary part β sets the oscillation frequency.

Complex roots: pure oscillation. Solve $y'' + 4y = 0$.

Read.

Solve the characteristic equation; purely imaginary roots give a sine–cosine solution with no exponential envelope.

Plan.

Characteristic equation: $r^2 + 4 = 0$, so $r = \pm 2\mathrm{i}$, i.e. $\alpha = 0$, $\beta = 2$. With zero real part the envelope $\mathrm{e}^{0 \cdot x} = 1$, leaving $y = C_1 \cos 2x + C_2 \sin 2x$.

Solve.

$y = C_1 \cos 2x + C_2 \sin 2x$.

Check.

This is simple harmonic motion of angular frequency 2—undamped, because the roots have no real part. ✓

An oscillator IVP and its amplitude. Solve $y'' + 4y = 0$ with $y(0) = 3$, $y'(0) = 8$, and give the amplitude of the resulting oscillation.

Read.

Apply the two conditions to the general solution, then combine the sine and cosine into a single amplitude $\sqrt{C_1^2 + C_2^2}$.

Plan.

From $y = C_1\cos 2x + C_2\sin 2x$, $y(0) = C_1 = 3$. Differentiate: $y' = -2C_1\sin 2x + 2C_2\cos 2x$, so $y'(0) = 2C_2 = 8$ and $C_2 = 4$. Thus $y = 3\cos 2x + 4\sin 2x$. A sum $C_1\cos\omega x + C_2\sin\omega x$ has amplitude $\sqrt{C_1^2 + C_2^2} = \sqrt{9+16} = 5$.

Solve.

$y = 3\cos 2x + 4\sin 2x$, amplitude 5, so $y = 5\cos(2x - \phi)$ for a suitable phase ϕ.

Check.

The peak displacement is 5, larger than either coefficient alone—sine and cosine of the same frequency combine into one shifted cosine. ✓

Complex roots: a decaying oscillation. Solve $y'' + 2y' + 5y = 0$.

Read.

Use the quadratic formula on the characteristic equation; the complex roots' real part is the decay rate.

Plan.

Characteristic equation: $r^2 + 2r + 5 = 0$. By the quadratic formula,

$$\begin{aligned} r &= \frac{-2 \pm \sqrt{4 - 20}}{2} && \text{quadratic formula on } r^2 + 2r + 5 = 0 \\ &= \frac{-2 \pm \sqrt{-16}}{2} && \text{simplify under the root} \\ &= -1 \pm 2\mathrm{i} && \sqrt{-16} = 4\mathrm{i} \end{aligned}$$

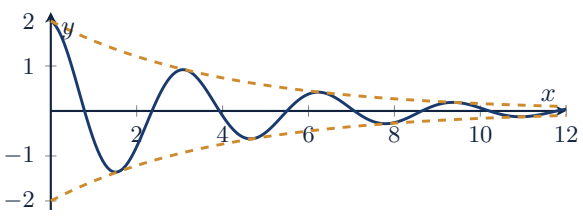
So $\alpha = -1$, $\beta = 2$, giving $y = \mathrm{e}^{-x}(C_1\cos 2x + C_2\sin 2x)$.

Solve.

$y = \mathrm{e}^{-x}(C_1\cos 2x + C_2\sin 2x)$.

Check.

The factor e^{-x} damps the oscillation to zero—an underdamped system settling to equilibrium. ✓



A decaying oscillation $y = \mathrm{e}^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x)$ with $\alpha < 0$: the sinusoid oscillates at frequency β inside the shrinking envelope $\pm A\mathrm{e}^{\alpha x}$ (dashed), the fingerprint of complex characteristic roots.

An IVP for a second-order equation. Solve $y'' - y = 0$ with $y(0) = 2$, $y'(0) = 0$.

Read.

Find the general solution from the characteristic roots, then apply both initial conditions to fix the two constants.

Plan.

Characteristic equation: $r^2 - 1 = 0$, so $r = \pm 1$ and $y = C_1\mathrm{e}^{x} + C_2\mathrm{e}^{-x}$. Then $y' = C_1\mathrm{e}^{x} - C_2\mathrm{e}^{-x}$. Apply the conditions: $y(0) = C_1 + C_2 = 2$ and $y'(0) = C_1 - C_2 = 0$. The second gives $C_1 = C_2$; the first then gives $C_1 = C_2 = 1$.

Solve.

$y = \mathrm{e}^{x} + \mathrm{e}^{-x} = 2\cosh x$.

Check.

$y(0) = 2$ and $y'(0) = 0$: a curve starting at height 2 with a horizontal tangent, as required. ✓

Two solutions genuinely span the solution space only if they are *linearly independent*, which the *Wronskian* $W = y_1y_2' - y_1'y_2$ detects: independence holds where $W \neq 0$.

A Wronskian check. Confirm e^{2x} and e^{3x} are independent.

Read.

Compute the Wronskian; a nonzero value certifies independence.

Plan.

With $y_1 = \mathrm{e}^{2x}$, $y_2 = \mathrm{e}^{3x}$:

$$\begin{aligned} W &= y_1y_2' - y_1'y_2 && \text{Wronskian definition} \\ &= \mathrm{e}^{2x}(3\mathrm{e}^{3x}) - (2\mathrm{e}^{2x})\mathrm{e}^{3x} && y_1' = 2\mathrm{e}^{2x},\ y_2' = 3\mathrm{e}^{3x} \\ &= (3-2)\mathrm{e}^{5x} = \mathrm{e}^{5x} && \text{combine like terms} \end{aligned}$$

Solve.

$W = \mathrm{e}^{5x} \neq 0$, so the two solutions are independent.

Check.

Independence is exactly what lets their combination represent *every* solution. ✓

Reduction of order

When one solution y_1 of a second-order linear equation is known, a second independent solution can always be found by *reduction of order*: substitute $y = v(x)\,y_1(x)$ and the equation for v loses its v-term, dropping to first order in v'. Besides being a general tool, it explains *why* a repeated characteristic root forces the extra factor of x.

Recovering the repeated-root solution. Given that $y_1 = \mathrm{e}^{2x}$ solves $y'' - 4y' + 4y = 0$, find a second independent solution.

Read.

Substitute $y = v\,\mathrm{e}^{2x}$; the equation collapses to $v'' = 0$, whose nonconstant solution supplies the second independent piece.

Plan.

Let $y = v\mathrm{e}^{2x}$. Then $y' = (v' + 2v)\mathrm{e}^{2x}$ and $y'' = (v'' + 4v' + 4v)\mathrm{e}^{2x}$. Substitute:

$$\begin{aligned} [(v'' + 4v' + 4v) - 4(v' + 2v) + 4v]\mathrm{e}^{2x} &= v''\mathrm{e}^{2x} && \text{the } v, v' \text{ terms cancel} \\ &= 0 && \text{so } v'' = 0 \end{aligned}$$

Integrate twice: $v = C_1 + C_2 x$. The nonconstant part $v = x$ gives the new solution $y_2 = x\mathrm{e}^{2x}$.

Solve.

$y_2 = x\mathrm{e}^{2x}$, so the general solution is $(C_1 + C_2 x)\mathrm{e}^{2x}$.

Check.

This is exactly the repeated-root formula—now derived rather than asserted. The v-term vanishing is guaranteed because y_1 already solves the equation. ✓

Did You Know? *three faces of damping*

The sign of the discriminant $b^2 - 4ac$ of $ay'' + by' + cy = 0$ names a physical regime. $b^2 - 4ac > 0$ (real distinct roots) is *overdamped*—the system returns to rest without oscillating; $b^2 - 4ac = 0$ (repeated root) is *critically damped*—the fastest non-oscillating return; $b^2 - 4ac < 0$ (complex

roots) is *underdamped*—it oscillates inside a decaying envelope. A car's shock absorbers aim for critical damping; a plucked guitar string is wildly underdamped.

Cauchy–Euler equations

One family of *variable*-coefficient equations still yields to an algebraic guess: the *Cauchy–Euler* equation $ax^2y'' + bxy' + cy = 0$, in which each derivative is multiplied by the matching power of x. Here the right guess is $y = x^r$, because $x^r y''$-type terms all return a constant times x^r. Substituting produces an *indicial* equation for r analogous to the characteristic equation.

A Cauchy–Euler equation. Solve $x^2y'' - 3xy' + 4y = 0$ for $x > 0$.

Read.

Guess $y = x^r$; substitute to get an equation in r; read off the solution, handling a repeated root with an $\ln x$ factor.

Plan.

With $y = x^r$: $y' = rx^{r-1}$ and $y'' = r(r-1)x^{r-2}$, so $x^2y'' = r(r-1)x^r$ and $xy' = rx^r$. Substitute:

$$\begin{aligned} \big[r(r-1) - 3r + 4\big]x^r &= 0 && \text{substitute the power terms} \\ r^2 - 4r + 4 &= 0 && \text{expand the bracket} \\ (r-2)^2 &= 0 && \text{a repeated root } r = 2 \end{aligned}$$

As with a repeated characteristic root, the second solution gains a logarithm, $x^2 \ln x$. Hence $y = (C_1 + C_2 \ln x)x^2$.

Solve.

$y = (C_1 + C_2 \ln x)\, x^2$.

Check.

The substitution $x = \mathrm{e}^t$ converts any Cauchy–Euler equation into a constant-coefficient one, which is why the three root cases reappear—distinct powers, a power times $\ln x$, and x^α times a trigonometric function of $\ln x$. ✓

Nonhomogeneous Equations: Finding a Particular Solution

Adding a forcing term $g(x)$ to the homogeneous equation (4.6) gives $ay'' + by' + cy = g(x)$. Its general solution has a clean two-part structure.

Derivation. *the general solution is $y = y_h + y_p$*

Suppose y_p is any single solution of $ay'' + by' + cy = g$, and y_h is the general solution of the homogeneous equation $ay'' + by' + cy = 0$. For any solution y of the forced equation, $y - y_p$ satisfies $a(y - y_p)'' + \cdots = g - g = 0$, so $y - y_p$ is some homogeneous solution y_h. Hence $y = y_h + y_p$: every forced solution is the general homogeneous solution plus one particular solution. Solve the two pieces separately and add.

The *method of undetermined coefficients* guesses y_p with the same shape as g (a polynomial for a polynomial, Ae^{kx} for an exponential, $A\cos\omega x + B\sin\omega x$ for a sinusoid), then fixes the coefficients by substitution.

Approach Read · Plan · Solve · Check

Read. Identify the forcing $g(x)$ and recall the two-part structure $y = y_h + y_p$.

Plan.

1. Solve the homogeneous equation for y_h via the characteristic equation.
2. Guess y_p with the same shape as g (polynomial, Ae^{kx}, or $A\cos\omega x + B\sin\omega x$).
3. If the guess already solves the homogeneous equation (resonance), multiply it by x (or x^2 for a double root).
4. Substitute y_p and match coefficients; then form $y = y_h + y_p$.
5. Apply any initial conditions to the *full* solution.

Solve. Differentiate the guess, substitute, and equate coefficients of like terms.

Check. Substitute y_p back into the forced equation to confirm it reproduces $g(x)$.

Common traps — every wrong answer is one of these

Applies initial conditions to y_p alone.
Fixes constants before adding the homogeneous part, so the arbitrary constants of y_h never get set. Conditions apply to $y = y_h + y_p$, after both parts are in hand.

Misses resonance.
Guesses Ae^{kx} (or $A\cos\omega x + B\sin\omega x$) when that very function solves the homogeneous equation, producing $0 = g$. When the forcing is a homogeneous solution, multiply the guess by x.

Guesses too small a family.
Uses only $A\cos\omega x$ for a sine forcing, dropping the sin term. A sinusoidal forcing needs *both* $A\cos\omega x$ and $B\sin\omega x$.

Forces a guess onto an incompatible g.
1. Tries undetermined coefficients on forcing like $\sec x$, $\tan x$, or $\ln x$.
2. Finds no finite family of derivatives to match.
3. Fix: when the forcing has no closed finite shape, switch to variation of parameters.

Polynomial forcing. Find a particular solution of $y'' - 3y' + 2y = 4x$.

Read.

Guess a polynomial of the same degree as the forcing, substitute, and match coefficients.

Plan.

Guess $y_p = Ax + B$, so $y_p' = A$, $y_p'' = 0$. Substitute:

$$\begin{aligned} y_p'' - 3y_p' + 2y_p &= 0 - 3A + 2(Ax + B) && \text{substitute the guess} \\ &= 2Ax + (2B - 3A) && \text{collect like terms} \end{aligned}$$

Match to $4x$: the x-coefficient gives $2A = 4$, so $A = 2$; the constant gives $2B - 3A = 0$, so $2B = 6$ and $B = 3$.

Solve.

$y_p = 2x + 3$.

Check.

Substituting $y_p = 2x + 3$: $0 - 3(2) + 2(2x + 3) = -6 + 4x + 6 = 4x$. ✓

Exponential forcing. Find a particular solution of $y'' - y = e^{2x}$.

Read.

Guess Ae^{2x} (not a root of the characteristic equation), substitute, solve for A.

Plan.

The characteristic roots are $r = \pm 1$, so e^{2x} is not itself a homogeneous solution and the plain guess $y_p = Ae^{2x}$ is safe. Then $y_p'' = 4Ae^{2x}$, so

$$\begin{aligned} y_p'' - y_p &= 4Ae^{2x} - Ae^{2x} && \text{substitute the guess} \\ &= 3Ae^{2x} && \text{combine} \end{aligned}$$

Match to e^{2x}: $3A = 1$, so $A = \frac{1}{3}$.

Solve.

$y_p = \frac{1}{3}e^{2x}$.

Check.

Had the forcing been e^x—a homogeneous solution—we would multiply the guess by x; see the trap below. ✓

Sinusoidal forcing. Find a particular solution of $y'' + 4y = \sin x$.

Read.

Guess $A\cos x + B\sin x$, substitute, and match sine and cosine coefficients separately.

Plan.

Guess $y_p = A\cos x + B\sin x$; then $y_p'' = -A\cos x - B\sin x$, so

$$\begin{aligned} y_p'' + 4y_p &= (-A\cos x - B\sin x) + 4(A\cos x + B\sin x) && \text{substitute the guess} \\ &= 3A\cos x + 3B\sin x && \text{combine} \end{aligned}$$

Match to $\sin x$ (no cosine on the right): $3A = 0$ gives $A = 0$, and $3B = 1$ gives $B = \frac{1}{3}$.

Solve.

$y_p = \frac{1}{3}\sin x$.

Check.

The driving frequency 1 differs from the natural frequency 2 ($r = \pm 2i$), so no resonance and the plain sinusoid guess works. ✓

Misconception

Resonance: multiply the guess by x. When the forcing term is *itself* a solution of the homogeneous equation, the naive guess produces $0 = g$ and fails. Fix it by multiplying the guess by x (or x^2 for a double root). For $y'' + y = \cos x$, since $\cos x$ solves the homogeneous equation, use $y_p = x(A\cos x + B\sin x)$, not $A\cos x + B\sin x$.

Resonant forcing. Find a particular solution of $y'' + y = \cos x$.

Read.

Recognize that $\cos x$ solves the homogeneous equation, so multiply the guess by x; substitute and match.

Plan.

The homogeneous solutions are $\cos x, \sin x$, so $\cos x$ resonates and the plain guess fails. Use $y_p = x(A\cos x + B\sin x)$. Differentiating twice gives $y_p'' = -x(A\cos x + B\sin x) + 2(-A\sin x + B\cos x)$, so

$$\begin{aligned} y_p'' + y_p &= \big[-x(A\cos x + B\sin x) + 2(-A\sin x + B\cos x)\big] + x(A\cos x + B\sin x) && \text{form } y_p'' + y_p \\ &= 2(-A\sin x + B\cos x) && \text{the } x(\cdots) \text{ terms cancel} \end{aligned}$$

Match to $\cos x$: the sine term gives $-2A = 0$, so $A = 0$; the cosine term gives $2B = 1$, so $B = \frac{1}{2}$.

Solve.

$y_p = \frac{1}{2}\, x \sin x$.

Check.

The x-factor makes the amplitude grow without bound—the signature of resonance, when a system is driven at its own natural frequency. ✓

A full nonhomogeneous IVP. Solve $y'' - 3y' + 2y = 4x$ with $y(0) = 1$, $y'(0) = 0$.

Read.

Combine the homogeneous solution with the particular solution found above, then apply both initial conditions.

Plan.

Homogeneous: $r^2 - 3r + 2 = (r-1)(r-2) = 0$, so $y_h = C_1 e^x + C_2 e^{2x}$. From the polynomial example, $y_p = 2x + 3$. General solution: $y = C_1 e^x + C_2 e^{2x} + 2x + 3$, with $y' = C_1 e^x + 2C_2 e^{2x} + 2$. Apply $y(0) = 1$: $C_1 + C_2 + 3 = 1$, so $C_1 + C_2 = -2$. Apply

$y'(0) = 0$: $C_1 + 2C_2 + 2 = 0$, so $C_1 + 2C_2 = -2$. Subtract: $C_2 = 0$, then $C_1 = -2$.

Solve.

$y = -2e^x + 2x + 3$.

Check.

Check $y(0) = -2 + 3 = 1$ and $y'(0) = -2 + 2 = 0$. ✓

When the forcing is not a neat polynomial, exponential, or sinusoid, *variation of parameters* builds y_p from the homogeneous solutions themselves—more general, more laborious. Given homogeneous solutions y_1, y_2 with Wronskian $W = y_1y_2' - y_1'y_2$, it takes $y_p = u_1y_1 + u_2y_2$ where

$$u_1' = -\frac{y_2\, g}{W}, \qquad u_2' = \frac{y_1\, g}{W}, \tag{4.7}$$

and g is the right-hand side (with the leading coefficient normalized to 1).

Variation of parameters. Solve $y'' + y = \sec x$ for a particular solution.

Read.

Use the homogeneous pair $\cos x, \sin x$; form u_1', u_2' from the variation-of-parameters formulas; integrate.

Plan.

Homogeneous solutions $y_1 = \cos x$, $y_2 = \sin x$ have Wronskian

$$\begin{aligned} W &= \cos x \cdot \cos x - (-\sin x)\sin x && y_1 = \cos x,\ y_2 = \sin x \\ &= \cos^2 x + \sin^2 x = 1 && \text{Pythagorean identity} \end{aligned}$$

With $g = \sec x$: $u_1' = -\frac{\sin x \sec x}{1} = -\tan x$, so $u_1 = \ln|\cos x|$; and $u_2' = \frac{\cos x \sec x}{1} = 1$, so $u_2 = x$. Assemble: $y_p = u_1y_1 + u_2y_2 = \cos x\, \ln|\cos x| + x \sin x$.

Solve.

$y_p = \cos x\, \ln|\cos x| + x \sin x$.

Check.

Undetermined coefficients has no guess for $\sec x$; variation of parameters handles any continuous forcing, at the cost of two integrals. ✓

Did You Know? *when the two methods trade places*

Undetermined coefficients is faster but works only for forcing terms whose derivatives stay within a finite family—polynomials, exponentials, sines, cosines, and their products. Variation of parameters (4.7) is slower but universal: it never guesses, so it copes with $\sec x$, $\tan x$, $\ln x$, or any continuous g. Reach for the guess first; fall back to variation of parameters when the guess has no finite shape.

Reading an Equation as a Model

The constant-coefficient machinery earns its keep because so many physical systems obey exactly $ay'' + by' + cy = g$. Two are worth carrying into the exam.

Approach Read · Plan · Solve · Check

Read. Recognize the physical system and match it to the template $ay'' + by' + cy = g$ — a mass–spring–damper, or an RC/RLC circuit.

Plan.

1. Write the governing ODE with its physical constants.
2. Divide by the leading coefficient (m or L, or RC) to standardize.
3. Solve the homogeneous part via the characteristic equation, adding a particular solution if forced.
4. Apply the initial conditions to fix the constants.

Solve. Carry the physical constants symbolically as far as possible before substituting numbers.

Check. Interpret the constants physically: $\omega = \sqrt{k/m}$ is the frequency, RC the time constant, and the real part of a complex root the decay rate.

Common traps — every wrong answer is one of these

Skips dividing by the leading coefficient.
Reads the characteristic equation off $mx'' + kx = 0$ as $r^2 + k = 0$, forgetting the m. Divide first: $r^2 + \frac{k}{m} = 0$, so $\omega = \sqrt{k/m}$.

Confuses time constant with frequency.
Treats the first-order RC decay as an oscillation, or the LC oscillation as a decay. A single real root gives exponential settling; imaginary roots give oscillation.

Mislabels the damping regime.
Reads the discriminant sign backwards — calling complex roots "overdamped." $b^2 - 4ac < 0$ (complex) is *under*damped and oscillates; $b^2 - 4ac > 0$ is overdamped.

Uses only one initial condition. 1. Applies $x(0)$ but not $x'(0)$ to a second-order model.
2. Leaves one constant undetermined.
3. Fix: a second-order IVP needs both the initial value and the initial rate to fix C_1 and C_2.

The simple harmonic oscillator. A mass m on a spring of stiffness k obeys $mx'' + kx = 0$. Find its motion.

Read.

Reduce to a characteristic equation; the purely imaginary roots give oscillation at the natural frequency.

Plan.

Divide by m: $x'' + \frac{k}{m}x = 0$, characteristic equation $r^2 + \frac{k}{m} = 0$, so $r = \pm i\sqrt{k/m}$. With $\omega = \sqrt{k/m}$, the motion is $x = C_1 \cos \omega t + C_2 \sin \omega t$, equivalently $x = A\cos(\omega t + \phi)$.

Solve.

$x(t) = A\cos(\omega t + \phi)$ with $\omega = \sqrt{k/m}$.

Check.

Stiffer springs (larger k) and lighter masses (smaller m) both raise the frequency ω, as intuition demands. ✓

A charging capacitor (RC circuit). A resistor R and capacitor C in series with a source V_s obey $RC\,V' + V = V_s$, with $V(0) = 0$. Find $V(t)$.

Read.

This is a first-order linear equation; solve by the integrating factor and apply the initial condition.

Plan.

Put in standard form: $V' + \frac{1}{RC}V = \frac{V_s}{RC}$, with integrating factor $\mu = \mathrm{e}^{t/RC}$:

$$\begin{aligned} (\mathrm{e}^{t/RC}V)' &= \frac{V_s}{RC}\mathrm{e}^{t/RC} && \text{multiply by } \mu = \mathrm{e}^{t/RC} \\ \mathrm{e}^{t/RC}V &= V_s\mathrm{e}^{t/RC} + K && \text{integrate both sides} \\ V &= V_s + K\mathrm{e}^{-t/RC} && \text{divide by } \mathrm{e}^{t/RC} \\ 0 &= V_s + K && \text{apply } V(0) = 0 \end{aligned}$$

so $K = -V_s$.

Solve.

$V(t) = V_s(1 - \mathrm{e}^{-t/RC})$.

Check.

The capacitor charges toward V_s with time constant RC: after one RC it reaches $1 - \mathrm{e}^{-1} \approx 63\%$ of full charge. ✓

A damped oscillator. A mass–spring–damper obeys $x'' + 2x' + 5x = 0$ with $x(0) = 1$, $x'(0) = 0$. Describe the motion and its decay.

Read.

Solve the characteristic equation; complex roots give a decaying sinusoid whose real part is the decay rate; apply the initial conditions.

Plan.

Characteristic equation $r^2 + 2r + 5 = 0$ gives $r = -1 \pm 2\mathrm{i}$, so $x = \mathrm{e}^{-t}(C_1\cos 2t + C_2\sin 2t)$. Then $x' = \mathrm{e}^{-t}\big[(-C_1 + 2C_2)\cos 2t + (-C_2 - 2C_1)\sin 2t\big]$. Apply $x(0) = 1$: $C_1 = 1$. Apply $x'(0) = 0$: $-C_1 + 2C_2 = 0$, so $C_2 = \frac{1}{2}$.

Solve.

$x(t) = \mathrm{e}^{-t}(\cos 2t + \frac{1}{2}\sin 2t)$: underdamped, oscillating at angular frequency 2 inside a decaying e^{-t} envelope.

Check.

The real part -1 of the roots is exactly the envelope's decay rate; the amplitude falls by $\mathrm{e}^{-1} \approx 37\%$ each unit of time. ✓

Connection

The same equation, everywhere. The oscillator $mx'' + bx' + kx = g(t)$ and the driven RLC circuit $LQ'' + RQ' + \frac{1}{C}Q = E(t)$ are the *same* differential equation with relabeled constants. Mass mirrors inductance, damping mirrors resistance, and spring stiffness mirrors inverse capacitance. Learn to solve one and you have solved them all—the reason this compact chapter reaches so far.

Synthesis — Key Results You Must Own

1. Classify first: order (highest derivative), linearity (how y enters), and expect one constant per order.
2. **Separable** $g(y)\mathrm{d}y = f(x)\mathrm{d}x$: integrate each side.
3. **Linear** $y' + Py = Q$: multiply by $\mu = \mathrm{e}^{\int P\,\mathrm{d}x}$; the left side becomes $(\mu y)'$; integrate.
4. **Exact** $M\mathrm{d}x + N\mathrm{d}y = 0$ when $M_y = N_x$; solutions are $F(x, y) = C$.
5. Models: growth/decay $N = N_0\mathrm{e}^{kt}$; cooling $T = T_a + (T_0 - T_a)\mathrm{e}^{-kt}$; mixing is first-order linear.
6. **Second-order** $ay'' + by' + cy = 0$: solve $ar^2 + br + c = 0$. Real distinct $\to C_1\mathrm{e}^{r_1 x} + C_2\mathrm{e}^{r_2 x}$; repeated $\to (C_1 + C_2 x)\mathrm{e}^{rx}$; complex $\alpha \pm \beta\mathrm{i} \to \mathrm{e}^{\alpha x}(C_1 \cos\beta x + C_2 \sin\beta x)$.
7. **Nonhomogeneous**: $y = y_h + y_p$; guess y_p by the shape of the forcing, and multiply by x under resonance.

Practice Problems

Difficulty: ⋆ warm-up, ⋆⋆ standard, ⋆⋆⋆ challenge. Work each with the Read–Plan–Solve–Check habit modeled above; classify the equation before choosing a method.

Classification and separable equations

1 ⋆.

State the order and whether it is linear: $y'' + xy' - \mathrm{e}^x y = \sin x$.

2 ⋆.

Verify that $y = C_1 \cos 3x + C_2 \sin 3x$ solves $y'' + 9y = 0$.

3 ⋆.

Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2x}{y}$.

4 ⋆⋆.

Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = y \cos x$ with $y(0) = 1$.

5 ⋆⋆.

Solve $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2}{1+y^2}$.

Hint: Separate and integrate; leave the answer implicit.

Linear first-order equations

6 ⋆.

Solve $y' + 3y = 6$.

7 ⋆⋆.

Solve $y' + y = \mathrm{e}^{-x}$.

Hint: Watch for the forcing matching the homogeneous solution.

8 ⋆⋆.

Solve $xy' - 2y = x^3$ for $x > 0$.

9 ⋆⋆⋆.

Solve $y' + (\tan x)\, y = \sec x$ on $(-\frac{\pi}{2}, \frac{\pi}{2})$ with $y(0) = 0$.

Exact equations and substitutions

10 ⋆⋆.

Show $(3x^2 + 2y)\,\mathrm{d}x + (2x + 1)\,\mathrm{d}y = 0$ is exact and solve it.

11 ⋆⋆.

Solve the homogeneous equation $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x^2+y^2}{xy}$.

Hint: Substitute $v = y/x$.

12 ⋆⋆⋆.

Solve the Bernoulli equation $y' + \frac{y}{x} = xy^2$ for $x > 0$.

Modeling

13 ⋆.

A population grows as $P' = 0.03P$. If $P(0) = 500$, when does it double?

14 ⋆⋆.

Carbon-14 has half-life 5730 yr. A sample retains 30% of its C-14. How old is it?

15 ⋆⋆.

A metal bar at 200° is placed in a 25° bath and reaches 100° after 4 min. Find its temperature after 10 min.

16 ⋆⋆.

A 200-L tank of pure water receives brine at 4 g/L flowing in at 5 L/min, mixed and drained at 5 L/min. Find the salt after t minutes and its limiting value.

17 ⋆⋆⋆.

An account earning continuous interest at rate r is drawn down at a constant W dollars per year: $A' = rA - W$. Solve for $A(t)$ given $A(0) = A_0$, and find the W that keeps A constant.

Second-order homogeneous equations

18 ⋆.

Solve $y'' - y' - 6y = 0$.

19 ⋆.

Solve $y'' + 6y' + 9y = 0$.

20 ⋆⋆.

Solve $y'' + 9y = 0$ with $y(0) = 0$, $y'(0) = 3$.

21 ⋆⋆.

Solve $y'' - 2y' + 10y = 0$.

22 ⋆⋆.

Solve $y'' - 4y = 0$ with $y(0) = 3$, $y'(0) = 2$.

23 ⋆⋆⋆.

For which values of b does every solution of $y'' + by' + 4y = 0$ oscillate (i.e. the roots are complex)?

Nonhomogeneous equations

24 ⋆⋆.

Find a particular solution of $y'' - y' - 2y = 6$.

25 ⋆⋆.

Find a particular solution of $y'' + y = 3\mathrm{e}^{2x}$.

26 ⋆⋆.

Find a particular solution of $y'' - 4y = 8x^2$.

27 ⋆⋆⋆.

Solve the IVP $y'' + y = 2\cos x$, $y(0) = 0$, $y'(0) = 0$.

Hint: The forcing resonates with the homogeneous solution.

28 ⋆⋆⋆.

Solve $y'' - 2y' + y = \mathrm{e}^{x}$.

Hint: The forcing matches a repeated root—multiply the guess by x^2.

Applications

29 ⋆⋆.

A 2-kg mass on a spring of stiffness $k = 50$ N/m oscillates without damping. Find the period of its motion.

30 ⋆⋆⋆.

In an RC circuit with $R = 1000\,\Omega$, $C = 2 \times 10^{-3}$ F, and $V_s = 10$ V, the capacitor starts uncharged. Find $V(t)$ and the time to reach 5 V.

Further Reading

The recognition-first framing here—separable, linear, or exact for first order; the characteristic equation for second—covers the Subject Test's reach. For the theory behind the methods (existence and uniqueness, the structure of the solution space, and systems of equations via eigenvalues), a standard ordinary-differential-equations text extends this chapter naturally, and the eigenvalue viewpoint ties directly to the linear algebra of a later chapter. Chapter 5 turns to that linear algebra.

CHAPTER 5

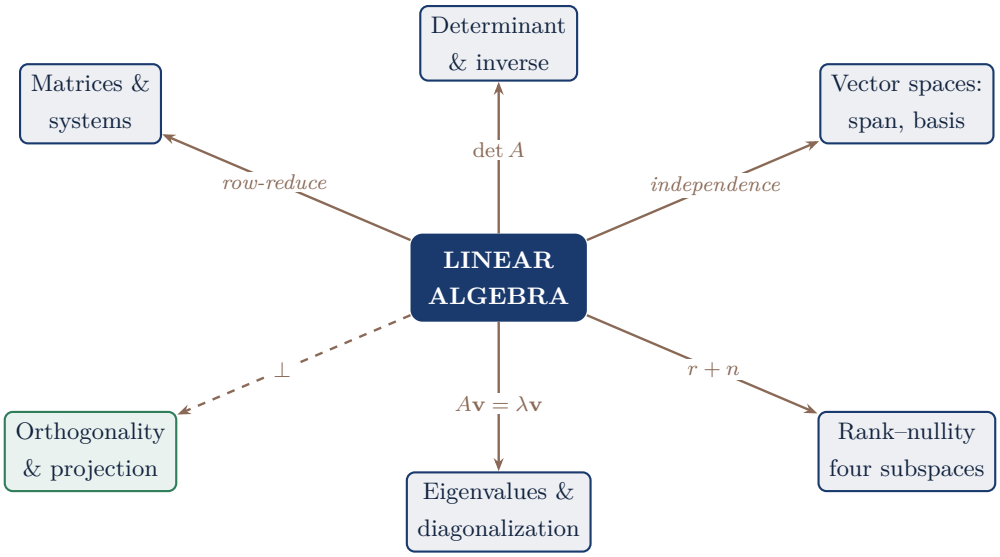
Linear Algebra

Vectors, Matrices, and Transformations

"The purpose of computation is insight, not numbers."

— Richard Hamming

Linear algebra is the study of* flat *maps—functions that send lines to lines and preserve the origin. Its two faces, the* matrix *(a grid of numbers) and the* linear transformation *(a geometric action), are the same object seen from two sides, and fluency means moving freely between them. A system of equations becomes a matrix you row-reduce; the reduced form reads off every solution at once. A determinant is a single number that decides whether a matrix is invertible and measures how it scales volume. Behind the arithmetic sit the structural ideas the Subject Test loves to probe:* span*,* independence*,* basis*, and* dimension*, tied together by the rank–nullity theorem, which says a matrix's inputs split cleanly into a part it crushes to zero and a part it maps faithfully. The chapter's summit is the eigenvalue: the special directions a matrix merely stretches, which* diagonalize *it and turn hard matrix problems into scalar ones. We close with orthogonality—the geometry of right angles—and the projection idea that underlies least squares. Nearly every question reduces to one move: row-reduce, take a determinant, or solve the characteristic equation.



Learning Objectives

You will be able to:

1. multiply matrices, take transposes, and recognize when the operations are legal;
2. solve a linear system by Gaussian elimination and read off consistency and free variables from reduced row echelon form;
3. compute determinants by cofactor expansion and invert a matrix two ways;
4. test vectors for linear independence and produce a basis and dimension for a span;
5. find bases for the null space and column space and verify the rank–nullity theorem;
6. build the matrix of a linear transformation from the images of the basis vectors;
7. find eigenvalues and eigenvectors and diagonalize a matrix;
8. compute dot products, orthogonal projections, and run the Gram–Schmidt process.

Vectors, Matrices, and the Product

A *matrix* is a rectangular grid of numbers; an $m \times n$ matrix has m rows and n columns. A column *vector* in $\mathbb{R}^n$ is an $n \times 1$ matrix. Addition and scalar multiplication act entrywise, but the operation that makes linear algebra powerful is the *matrix product*, and its rule is not entrywise. The product AB is defined only when the number of columns of A equals the number of rows of B; if A is $m \times n$ and B is $n \times p$, then AB is $m \times p$, and its (i, j) entry is the dot product of row i of A with column j of B. The reason for this seemingly odd rule is that matrix multiplication is exactly *composition of linear maps*: AB is the transformation "do B, then do A."

What & why. How matrices combine, and the dimension rule that says when a product is legal—the grammar all later computation obeys.

Solve. To form AB: dot each row of A with each column of B; inner dimensions must match.

Means. The product encodes "apply B, then A"—composition, read right to left.

Approach Read · Plan · Solve · Check

Read. Note the size of each factor and check the inner dimensions: A is $m \times n$, B is $n \times p$. If they match, AB exists and is $m \times p$; if not, the product is undefined.

Plan.

1. Confirm the product is legal (columns of A equal rows of B) and record the output size.
2. For each output slot (i, j), pair row i of A with column j of B.
3. Take that dot product $\sum_k a_{ik}b_{kj}$ and place it in slot (i, j); keep the factors in order, since AB means "apply B, then A."

Solve. Compute one entry at a time, writing each as a sum of products, until the grid is filled.

Check. Recompute one entry independently, or test a structural law — $(AB)^\top = B^\top A^\top$, or $AI = A$ — to catch a slipped row or column.

Common traps — every wrong answer is one of these

Multiplies entrywise.
Treats $(AB)_{ij}$ as $a_{ij}b_{ij}$, copying the rule for addition. The product entry is row i of A dotted with column j of B, not a single product.

Assumes $AB = BA$.
Swaps factors freely. Matrix multiplication is composition of maps and generally noncommutative; AB ("do B, then A") and BA usually differ entry by entry.

Transposes without reversing.
Writes $(AB)^\top = A^\top B^\top$. The transpose reverses order: $(AB)^\top = B^\top A^\top$.

Reads off zero divisors and cancellation.
1. Sees $AB = 0$ and concludes $A = 0$ or $B = 0$.
2. Sees $AB = AC$ and cancels A to get $B = C$.
3. Fix: neither holds for matrices — $AB = 0$ allows nonzero factors, and cancellation is legal only when A is invertible.

A matrix product. Compute AB for $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$.

Read.

Each entry of AB is (row of A)·(column of B); both are 2×2, so the product exists and is 2×2.

Plan.

Dot each row of A with each column of B, one entry at a time:

$$(AB)_{11} = 1 \cdot 5 + 2 \cdot 7 = 19 \qquad \text{row 1 of } A, \text{ column 1 of } B$$
$$(AB)_{12} = 1 \cdot 6 + 2 \cdot 8 = 22 \qquad \text{row 1, column 2}$$
$$(AB)_{21} = 3 \cdot 5 + 4 \cdot 7 = 43 \qquad \text{row 2, column 1}$$
$$(AB)_{22} = 3 \cdot 6 + 4 \cdot 8 = 50 \qquad \text{row 2, column 2}$$

Solve.

$$AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

Check.

Check one entry independently: row 2 of A is $(3, 4)$, column 1 of B is $(5, 7)$, dot product $15 + 28 = 43$. ✓

Order matters. For the same A, B, compute BA and compare with AB.

Read.

Recompute with the factors swapped; matrix multiplication is generally noncommutative.

Plan.

Recompute each entry with the factors swapped:

$$(BA)_{11} = 5 \cdot 1 + 6 \cdot 3 = 23 \qquad \text{row 1 of } B, \text{ column 1 of } A$$
$$(BA)_{12} = 5 \cdot 2 + 6 \cdot 4 = 34 \qquad \text{row 1, column 2}$$
$$(BA)_{21} = 7 \cdot 1 + 8 \cdot 3 = 31 \qquad \text{row 2, column 1}$$
$$(BA)_{22} = 7 \cdot 2 + 8 \cdot 4 = 46 \qquad \text{row 2, column 2}$$

Solve.

$$BA = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \neq AB.$$

Check.

The two products differ entry by entry, confirming $AB \neq BA$ in general—composition of maps depends on the order you apply them.

The *transpose* $A^\top$ flips a matrix across its main diagonal, turning rows into columns. It reverses products: $(AB)^\top = B^\top A^\top$. A matrix with $A^\top = A$ is *symmetric*; the identity matrix I, with ones on the diagonal and zeros elsewhere, satisfies $AI = IA = A$ and is the "do nothing" map.

Transpose reverses a product. With A, B as above, verify $(AB)^\top = B^\top A^\top$.

Read.

Transpose the product from Example 5.1, then multiply the transposes in reverse order and compare.

Plan.

Transpose the product, then multiply the transposes in reverse order:

$$(AB)^\top = \begin{bmatrix} 19 & 43 \\ 22 & 50 \end{bmatrix} \qquad \text{flip } AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \text{ across its diagonal}$$

$$B^\top A^\top = \begin{bmatrix} 5 & 7 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \qquad \text{transpose each factor, order reversed}$$

$$= \begin{bmatrix} 5\cdot 1 + 7\cdot 2 & 5\cdot 3 + 7\cdot 4 \\ 6\cdot 1 + 8\cdot 2 & 6\cdot 3 + 8\cdot 4 \end{bmatrix} \qquad \text{multiply out}$$

$$= \begin{bmatrix} 19 & 43 \\ 22 & 50 \end{bmatrix} \qquad \text{matches } (AB)^\top$$

Solve.

The two sides agree: $(AB)^\top = B^\top A^\top$.

Check.

Note the *reversal*: $A^\top B^\top$ would be a different product—the order flip is essential. ✓

Misconception

Matrix multiplication is not commutative and has zero divisors. $AB \neq BA$ in general, so you may not freely swap factors; and $AB = 0$ does *not* force $A = 0$ or $B = 0$. Cancellation also fails: $AB = AC$ does not imply $B = C$ unless A is invertible.

Systems of Linear Equations

A system of linear equations is compactly written $A\mathbf{x} = \mathbf{b}$, where A holds the coefficients, $\mathbf{x}$ the unknowns, and $\mathbf{b}$ the right-hand sides. We solve it by *Gaussian elimination*: apply *elementary row operations*—swap two rows, scale a row by a nonzero constant, or add a multiple of one row to another—to the *augmented matrix* $[\,A \mid \mathbf{b}\,]$ until it reaches *row echelon form* (a staircase of leading entries with zeros below). Back-substitution then reads off the solution. Pushing further, to *reduced row echelon form* (RREF)—leading 1's with zeros above and below—lets you read the solution directly.

Derivation. *Why row operations preserve the solution set*

Each elementary row operation is reversible and acts on whole equations, not on the unknowns. Swapping two equations clearly changes nothing; scaling an equation by a nonzero constant c multiplies both sides equally, so the same $\mathbf{x}$ satisfies it; adding a multiple of one true equation to another produces a new equation that every common solution still satisfies, and reversibility guarantees no *new* solutions sneak in. Hence the reduced system has exactly the same solution set as the original—which is why elimination is legitimate.

Approach Read · Plan · Solve · Check

Read. Identify the coefficients, unknowns, and right-hand sides, and assemble the augmented matrix $[\,A \mid \mathbf{b}\,]$. Note how many equations and unknowns there are.

Plan.

1. Use elementary row operations — swap, scale by a nonzero constant, add a multiple of one row to another — to reach row echelon form.
2. Scan for a contradiction row $[\,0 \cdots 0 \mid c \neq 0\,]$ (no solution) or a pivotless column (a free variable, infinitely many solutions).
3. Back-substitute from the bottom pivot up, or continue to reduced row echelon form to read the answer directly.

Solve. Carry out the row operations one at a time, writing the matrix after each stage; never fold two operations into one step.

Check. Substitute the solution back into the *original* equations, and confirm the number of free variables matches the missing pivots.

Common traps — every wrong answer is one of these

Scales a row by zero.
Multiplies a row by 0 to clear it, destroying information and changing the solution set. Only *nonzero* scaling is a legal row operation.

Sign error adding rows.
Computes $R_2 - 2R_1$ but subtracts only some entries, or flips the multiplier's sign. Apply the same multiple to *every* entry of the row, right-hand side included.

Misreads a free variable as "no solution."
Sees a pivotless column and declares the system inconsistent. A pivotless column gives a *free variable* and infinitely many solutions; only a row $[\,0 \cdots 0 \mid c \neq 0\,]$ means no solution.

Stops at echelon form and misreads the answer.
1. Reaches echelon form but reads variables off as if the pivots were 1 with zeros above.
2. Reports wrong values because entries above the pivots were never cleared.
3. Fix: back-substitute carefully, or reduce fully to RREF before reading the solution.

A 3×3 system by elimination. Solve

$$\begin{aligned} x + y + z &= 6 \\ 2x - y + z &= 3 \\ x + 2y - z &= 2. \end{aligned}$$

Read.

Form the augmented matrix and eliminate below the first pivot, then finish by back-substitution.

Plan.

Augmented matrix and elimination:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & -1 & 1 & 3 \\ 1 & 2 & -1 & 2 \end{array}\right] \xrightarrow[R_3-R_1]{R_2-2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & -3 & -1 & -9 \\ 0 & 1 & -2 & -4 \end{array}\right].$$

Row 2 gives $-3y - z = -9$; row 3 gives $y - 2z = -4$, so $y = 2z - 4$. Substitute into row 2:

$$\begin{aligned} -3(2z-4) - z &= -9 && \text{put } y = 2z-4 \text{ into } -3y - z = -9 \\ -7z + 12 &= -9 && \text{expand and collect} \\ z &= 3 && \text{solve for } z \end{aligned}$$

Then $y = 2(3) - 4 = 2$ and $x = 6 - y - z = 6 - 2 - 3 = 1$.

Solve.

$(x, y, z) = (1, 2, 3)$.

Check.

Substitute back into the third original equation: $1 + 2(2) - 3 = 2$. ✓

A system is *consistent* if it has at least one solution. Elimination exposes the three possibilities at a glance: a row $[\,0\ 0\ \cdots\ 0 \mid c\,]$ with $c \neq 0$ signals *no* solution; a pivot in every column of A gives a *unique* solution; and any column without a pivot corresponds to a *free variable*, yielding infinitely many solutions parametrized by that variable.

An inconsistent system. Decide whether $x + y = 2,\ 2x + 2y = 5$ has a solution.

Read.

Row-reduce; a contradiction row means no solution.

Plan.

Subtract twice row 1 from row 2: $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & 2 & 5 \end{array}\right] \xrightarrow{R_2-2R_1} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right]$. The last row reads $0 = 1$.

Solve.

No solution—the system is inconsistent.

Check.

Geometrically the two lines are parallel (same slope, different intercept), so they never

meet. ✓

Infinitely many solutions. Solve $x + 2y + z = 4,\ 2x + 4y + 3z = 9$ and describe the solution set.

Read.

Row-reduce; a free variable (a column without a pivot) parametrizes the family.

Plan.

Eliminate: $R_2 - 2R_1$ gives $z = 1$. Back-substitute into $x + 2y + z = 4$: $x + 2y = 3$, so $x = 3 - 2y$. The variable y is free.

Solve.

Solutions: $(x, y, z) = (3 - 2t,\ t,\ 1)$ for every $t \in \mathbb{R}$—a line in space.

Check.

Check $t = 0$: $(3, 0, 1)$ satisfies both equations. Setting $t = 1$: $(1, 1, 1)$ also works. ✓

Reduced row echelon form. Reduce the augmented matrix $\left[\begin{array}{ccc|c} 2 & 4 & 2 & 6 \\ 1 & 3 & 1 & 5 \end{array}\right]$ to RREF and read off the solution.

Read.

Get a leading 1, clear below it, clear above it; the pivot columns then read out the solution with any free variable left as a parameter.

Plan.

Swap rows for a clean pivot, then eliminate:

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 5 \\ 2 & 4 & 2 & 6 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 3 & 1 & 5 \\ 0 & -2 & 0 & -4 \end{array}\right] \xrightarrow{-\frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & 3 & 1 & 5 \\ 0 & 1 & 0 & 2 \end{array}\right] \xrightarrow{R_1 - 3R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 2 \end{array}\right].$$

Pivots sit in columns 1 and 2; column 3 is free, so $x_3 = t$, $x_2 = 2$, $x_1 = -1 - t$.

Solve.

Solutions: $(x_1, x_2, x_3) = (-1 - t,\ 2,\ t)$, $t \in \mathbb{R}$.

Check.

Set $t = 0$: $(-1, 2, 0)$ gives $2(-1) + 4(2) + 0 = 6$ and $-1 + 6 + 0 = 5$. ✓

Problem-solving strategy — attacking a linear system

(1) Write the augmented matrix $[\,A \mid \mathbf{b}\,]$. (2) Row-reduce to echelon form. (3) Scan the result: a row $[\,0 \cdots 0 \mid c \neq 0\,]$ means *no solution*; a pivot in every variable column means a *unique* solution; a missing pivot means a *free variable* and infinitely many. (4) Back-substitute or continue to RREF, and always verify one solution in the original equations.

Determinants and Inverses

The *determinant* $\det A$ of a square matrix is a single number that decides invertibility: A is invertible *exactly when* $\det A \neq 0$. Geometrically, $|\det A|$ is the factor by which A scales area (in $\mathbb{R}^2$) or volume (in $\mathbb{R}^3$), and its sign records whether orientation is preserved. For a 2×2 matrix, $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$. For larger matrices we use *cofactor expansion* along any row or column.

Cofactor expansion. Expanding along row i,

$$\det A = \sum_{j} (-1)^{i+j} a_{ij} M_{ij}, \tag{5.1}$$

where the *minor* M_{ij} is the determinant of the matrix with row i and column j deleted. The sign pattern $\begin{smallmatrix} + & - & + \\ - & + & - \\ + & - & + \end{smallmatrix}$ alternates like a checkerboard.

Approach Read · Plan · Solve · Check

Read. Confirm the matrix is square, note its size, and scan for a row or column with zeros — expanding along it saves work.

Plan.

1. For 2×2, use $ad - bc$; for larger matrices, expand in cofactors along the row or column with the most zeros, respecting the checkerboard signs $(-1)^{i+j}$.
2. To invert: for 2×2 apply swap–negate–divide; for larger, row-reduce $[\,A \mid I\,] \to [\,I \mid A^{-1}\,]$ or build the adjugate.
3. Use $\det A \neq 0$ as the invertibility test before attempting an inverse.

Solve. Reduce each minor to a number one at a time, then combine with the sign pattern; keep the $1/\det A$ factor attached throughout an inverse.

Check. Verify $AA^{-1} = I$ on one row, or confirm $\det(AB) = \det A \det B$; a zero determinant

must coincide with dependent columns.

Common traps — every wrong answer is one of these

Forgets the checkerboard signs.
Adds every cofactor term with a $+$. The signs alternate as $(-1)^{i+j}$; the $(1,2)$ and $(2,1)$ cofactors carry a minus.

Thinks the determinant is additive.
Writes $\det(A+B) = \det A + \det B$, or $\det(cA) = c\det A$. Neither holds: $\det(cA) = c^n \det A$ for $n \times n$, and only $\det(AB) = \det A \det B$ is clean.

Drops the $1/\det$ in the inverse.
Forms the adjugate $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ but forgets to divide by $ad - bc$. The inverse is that adjugate *scaled* by $1/\det A$.

Inverts a singular matrix. 1. Sees $\det A = 0$ yet proceeds to build A^{-1}.

2. Divides by zero or reports a nonsense inverse.
3. Fix: $\det A = 0$ means A is *not* invertible — stop and report "no inverse."

A 3×3 determinant. Compute $\det \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$.

Read.

Expand along the first row, whose zero entry kills one term.

Plan.

Expand along row 1 with the checkerboard signs:

$$\det A = 2\det \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} - 1\det \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + 0 \quad \text{cofactor expansion; the 0 entry drops its term}$$

$$\det \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = 3 \cdot 2 - 1 \cdot 1 = 5 \quad \text{first minor}$$

$$\det \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = 1 \cdot 2 - 1 \cdot 0 = 2 \quad \text{second minor}$$

$$\det A = 2(5) - 1(2) = 8 \quad \text{combine}$$

Solve.

$\det A = 8$.

Check.

Nonzero, so A is invertible—consistent with the inverse we build in Example 5.11. ✓

Did You Know? *the determinant is signed volume*

The columns of a 3×3 matrix span a parallelepiped, and $\det A$ is its *signed* volume. This is why $\det A = 0$ means "the columns are flat"—they fail to fill space, the map collapses a dimension, and no inverse can exist. Every algebraic property of determinants (alternating signs, vanishing on repeated columns) is this geometric fact in disguise.

Misconception

The determinant is not additive, and scaling is not linear. In general $\det(A+B) \neq \det A + \det B$, and for an $n \times n$ matrix $\det(cA) = c^n \det A$—*not* $c \det A$—because multiplying A by c stretches all n dimensions at once. The one law that does hold cleanly is multiplicativity: $\det(AB) = \det A \det B$.

Determinant as area. Find the area of the parallelogram spanned by $\mathbf{a} = (3, 1)$ and $\mathbf{b} = (1, 2)$.

Read.

The area equals the absolute value of the determinant of the matrix whose columns are the two vectors.

Plan.

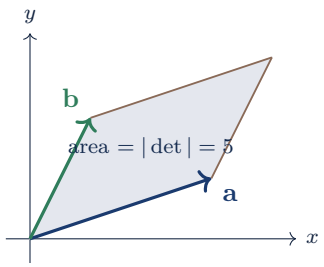
Form $\det \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = 3 \cdot 2 - 1 \cdot 1 = 5$; the area is its absolute value.

Solve.

Area $= |5| = 5$.

Check.

Positive determinant means $\mathbf{a} \to \mathbf{b}$ turns counterclockwise (orientation preserved). Had the vectors been parallel, the determinant—and the area—would be 0. ✓



The determinant as signed area. The parallelogram spanned by $\mathbf{a} = (3, 1)$ *and* $\mathbf{b} = (1, 2)$ *has area* $|\det[\mathbf{a}\ \mathbf{b}]| = 5$*; a zero determinant would flatten the parallelogram to a segment.*

The *inverse* A^{-1} satisfies $AA^{-1} = A^{-1}A = I$. For a 2×2 matrix there is a memorable closed form; for larger matrices, adjoin the identity and row-reduce $[\,A \mid I\,] \to [\,I \mid A^{-1}\,]$.

The 2×2 inverse. If $\det = ad - bc \neq 0$, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{5.2}$$

In words: swap the diagonal, negate the off-diagonal, divide by the determinant.

A 2×2 inverse. Invert $A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}$.

Read.

Compute the determinant, then apply the swap-negate-divide formula.

Plan.

Apply swap–negate–divide:

$$\det A = 3 \cdot 4 - 1 \cdot 2 = 10 \neq 0 \qquad \text{nonzero, so } A^{-1} \text{ exists}$$

$$A^{-1} = \frac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} \qquad \text{swap diagonal, negate off-diagonal, divide by } \det A$$

Solve.

$$A^{-1} = \tfrac{1}{10} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix}.$$

Check.

Verify: $AA^{-1} = \tfrac{1}{10} \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \tfrac{1}{10} \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = I$. ✓

A 3×3 inverse by the adjugate. Invert $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$ (from Example 5.9, $\det A = 8$).

Read.

Build the cofactor matrix (5.1), transpose it to the adjugate, and divide by $\det A$.

Plan.

The cofactors are $C_{11} = +(3\cdot 2 - 1\cdot 1) = 5$, $C_{12} = -(1\cdot 2 - 1\cdot 0) = -2$, $C_{13} = +(1\cdot 1 - 3\cdot 0) = 1$, and by symmetry the cofactor matrix is $\begin{bmatrix} 5 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 5 \end{bmatrix}$. Since this matrix is symmetric, the adjugate (its transpose) is the same. Divide by $\det A = 8$.

Solve.

$$A^{-1} = \tfrac{1}{8} \begin{bmatrix} 5 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 5 \end{bmatrix}.$$

Check.

Spot-check row 1 of A times column 1 of A^{-1}: $(2,1,0) \cdot \tfrac{1}{8}(5,-2,1) = \tfrac{1}{8}(10 - 2 + 0) = 1$. ✓

Connection

The determinant is a one-number invertibility test. Everything ties together: $\det A \neq 0 \iff A$ is invertible $\iff A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \iff$ the columns of A are linearly independent $\iff 0$ is not an eigenvalue of A. These equivalences (the "Invertible Matrix Theorem") are the backbone of the subject.

A 3×3 inverse by Gauss–Jordan. Invert $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ by reducing $[\,A \mid I\,] \to [\,I \mid A^{-1}\,]$.

Read.

Adjoin the identity, then clear the upper entries by back-elimination since A is already upper triangular with 1's on the diagonal.

Plan.

Start and back-eliminate:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_1-R_3]{R_2-R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1-R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

Solve.

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Check.

Check row 1 of A times A^{-1}: $(1,1,1)$ against the columns gives $(1,0,0)$. ✓

The determinant also delivers a direct solution formula for square systems.

Cramer's rule. For $A\mathbf{x} = \mathbf{b}$ with $\det A \neq 0$, the ith unknown is

$$x_i = \frac{\det A_i}{\det A}, \tag{5.3}$$

where A_i is A with its ith column replaced by $\mathbf{b}$. Efficient for 2×2 and 3×3 systems; impractical beyond.

Cramer's rule. Solve $\begin{aligned} x + 2y &= 5 \\ 3x + 4y &= 6 \end{aligned}$ by Cramer's rule (5.3).

Read.

Compute $\det A$, then replace each column with $\mathbf{b} = (5, 6)$ in turn.

Plan.

Compute the determinant, then replace one column at a time with $\mathbf{b} = (5, 6)$:

$$\det A = \det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 4 - 6 = -2 \qquad \text{coefficient determinant}$$

$$x = \frac{1}{\det A} \det \begin{bmatrix} 5 & 2 \\ 6 & 4 \end{bmatrix} = \frac{20 - 12}{-2} = -4 \qquad \text{column 1 replaced by } \mathbf{b}$$

$$y = \frac{1}{\det A} \det \begin{bmatrix} 1 & 5 \\ 3 & 6 \end{bmatrix} = \frac{6 - 15}{-2} = \frac{9}{2} \qquad \text{column 2 replaced by } \mathbf{b}$$

Solve.

$(x, y) = \left(-4,\ \frac{9}{2}\right)$.

Check.

Check the first equation: $-4 + 2 \cdot \frac{9}{2} = -4 + 9 = 5$. ✓

A determinant that reveals dependence. Evaluate $\det \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \\ 3 & 5 & 0 \end{bmatrix}$.

Read.

Row-reduce toward triangular form; a zero row forces the determinant to vanish (row operations of the "add a multiple" type do not change the determinant).

Plan.

Notice row 3 equals row 1 plus row 2: $(1, 2, -1) + (2, 3, 1) = (3, 5, 0)$. Subtracting, $R_3 \to R_3 - R_1 - R_2$ produces a zero row, and a matrix with a zero row has determinant 0.

Solve.

$\det = 0$; the matrix is singular.

Check.

A zero determinant means the rows (hence columns) are linearly dependent, so the matrix is not invertible—consistent with row 3 being a combination of the others. ✓

Vector Spaces, Independence, and Basis

A *vector space* is a set of objects that can be added and scaled while obeying the usual rules; $\mathbb{R}^n$ is the model, but polynomials and matrices qualify too. A *subspace* is a subset that is itself a vector space—equivalently, a nonempty subset closed under addition and scalar multiplication (and hence containing $\mathbf{0}$). The *span* of a set of vectors is the collection of all their linear combinations; it is always a subspace.

Three interlocking ideas organize a space. Vectors are *linearly independent* if no one of them is a combination of the others—equivalently, the only solution of $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ is all $c_i = 0$. A *basis* is an independent set that spans the space: a minimal spanning set, or a maximal independent set. The *dimension* is the number of vectors in any basis—a well-defined number, since all bases of a space have the same size.

What & why. What it means for vectors to be genuinely different directions, and how a basis packages a whole space into a finite list.

Solve. Test independence by row-reducing the matrix whose columns are the vectors; pivots mark independent directions.

Means. A basis is the smallest list of directions from which the entire space is reachable.

Approach Read · Plan · Solve · Check

Read. Decide what is being asked — independence, a basis, a dimension, or whether a set is a subspace — and write the vectors as columns (or rows) of a matrix.

Plan.

1. To test independence, row-reduce (or take the determinant of) the matrix of vectors; pivots mark independent directions.
2. For a basis of a span, keep the vectors that become pivot columns; the dimension is the number of pivots.
3. To test a subspace, check it contains $\mathbf{0}$ and is closed under addition and scalar multiplication.

Solve. Reduce one row operation at a time, or set up $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ and solve for the coefficients.

Check. Independent vectors force all $c_i = 0$; a claimed basis must both span and be independent, and its size must equal the dimension.

Common traps — every wrong answer is one of these

Confuses spanning with independence.
Assumes any spanning set is a basis. A basis must *also* be independent — a spanning set can carry redundant vectors that must be dropped.

Miscounts the dimension.
Counts the number of vectors given rather than the number of pivots. Dimension is the count of *independent* directions, i.e. pivot columns after reduction.

Calls a shifted plane a subspace.
Treats $\{(x, y, z) : x+y+z = 1\}$ as a subspace. Any subspace must contain $\mathbf{0}$; a nonhomogeneous condition ($= 1$) misses the origin and fails closure.

Declares independence without checking.
1. Sees three vectors in $\mathbb{R}^3$ and assumes they must be a basis.
2. Skips the row reduction or determinant.
3. Fix: three vectors are independent only if the determinant is nonzero (equivalently, three pivots) — verify before concluding.

Testing independence. Are $\mathbf{v}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}2\\5\\7\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}1\\3\\5\end{bmatrix}$ independent?

Read.

Place them as columns and compute the determinant; nonzero means independent.

Plan.

Place the vectors as columns and expand along row 1:

$$\begin{aligned}
\det\begin{bmatrix}1&2&1\\2&5&3\\3&7&5\end{bmatrix} &= 1(5\cdot 5 - 3\cdot 7) - 2(2\cdot 5 - 3\cdot 3) + 1(2\cdot 7 - 5\cdot 3) && \text{cofactor expansion}\\
&= 1(4) - 2(1) + 1(-1) && \text{evaluate each minor}\\
&= 1 && \text{combine}
\end{aligned}$$

Solve.

The determinant is $1 \neq 0$, so the three vectors are linearly independent (and form a basis

of $\mathbb{R}^3$).

Check.

A nonzero determinant means the only combination giving $\mathbf{0}$ is the trivial one. ✓

A dependent set. Show $\mathbf{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix}2\\3\\1\end{bmatrix}$ are dependent and give a relation.

Read.

Look for a combination of the first two that yields the third; if found, the set is dependent.

Plan.

Try $\mathbf{u}_3 = a\mathbf{u}_1 + b\mathbf{u}_2$. The first entry needs $a = 2$; the third needs $b = 1$; check the middle: $a + b = 2 + 1 = 3$. ✓ It matches, so $\mathbf{u}_3 = 2\mathbf{u}_1 + \mathbf{u}_2$.

Solve.

Dependent, with relation $2\mathbf{u}_1 + \mathbf{u}_2 - \mathbf{u}_3 = \mathbf{0}$.

Check.

Because one vector is a combination of the others, the three span only a 2-dimensional plane, not all of $\mathbb{R}^3$.

Deciding whether a subset is a subspace. Is $W = \{(x, y, z) : x + y + z = 0\}$ a subspace of $\mathbb{R}^3$? Is $U = \{(x, y, z) : x + y + z = 1\}$?

Read.

Check closure under addition and scalar multiplication, and whether $\mathbf{0}$ belongs.

Plan.

For W: it contains $\mathbf{0}$ (since $0+0+0 = 0$). If $\mathbf{u}, \mathbf{v} \in W$ then their coordinate sums are each 0, so $(\mathbf{u} + \mathbf{v})$ has coordinate sum 0, and $c\mathbf{u}$ has coordinate sum $c \cdot 0 = 0$; both stay in W.

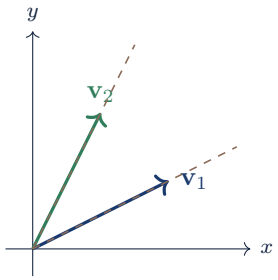
For U: it fails to contain $\mathbf{0}$ (sum $0 \neq 1$), and adding two of its members gives coordinate sum $2 \neq 1$.

Solve.

W *is* a subspace (a plane through the origin, dimension 2); U is *not*.

Check.

The rule of thumb: a solution set of a *homogeneous* equation ($= 0$) is a subspace; a nonhomogeneous one ($= 1$) is a shifted copy that misses the origin. ✓



Two independent vectors in the plane. Because they point in different directions, every point of $\mathbb{R}^2$ is reachable as a combination $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$: they form a basis. Were $\mathbf{v}_2$ a multiple of $\mathbf{v}_1$, their span would collapse to a single line.

A basis and dimension for a span. Find a basis for the span of $(1, 2, 1), (2, 4, 2), (3, 6, 4)$ and state its dimension.

Read.

Stack the vectors as rows and row-reduce; the nonzero rows form a basis of the span.

Plan.

Row-reduce:

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 6 & 4 \end{bmatrix} \xrightarrow[R_3 - 3R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The second vector was twice the first (it collapses to zero); two nonzero rows remain.

Solve.

A basis is $\{(1, 2, 1), (0, 0, 1)\}$, so the span has dimension 2 (a plane in $\mathbb{R}^3$).

Check.

The three original vectors could not be independent—three vectors spanning only a plane must have a dependence, and indeed $(2, 4, 2) = 2(1, 2, 1)$. ✓

Rank, Nullity, and the Four Subspaces

Every matrix A carries two fundamental subspaces. The *column space* is the span of its columns—all outputs $A\mathbf{x}$ the matrix can produce. The *null space* is the set of inputs it destroys, $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$. The dimension of the column space is the *rank* r (the number of pivots after row reduction); the dimension of the null space is the *nullity*. The rank counts the independent directions the matrix genuinely maps; the nullity counts the directions it flattens to zero.

Derivation. *the rank–nullity theorem*

Row-reduce A (an $m \times n$ matrix). Each of the n columns is either a *pivot* column or a *free* column. Pivot columns supply a basis for the column space, so their count is the rank r. Each free column contributes exactly one free variable, hence one basis vector of the null space, so the nullity is $n - r$. Adding, $r + (n - r) = n$:

$$\text{rank}(A) + \text{nullity}(A) = n \quad \text{(number of columns)}. \tag{5.4}$$

No input direction is lost: each column is either mapped faithfully or crushed to zero.

Approach Read · Plan · Solve · Check

Read. Note the size $m \times n$ of A and which subspace is asked for — column space (outputs) or null space (inputs sent to $\mathbf{0}$).

Plan.

1. Row-reduce A and locate the pivots; the pivot columns count the rank.
2. For the null space, set each free variable to a parameter and solve $A\mathbf{x} = \mathbf{0}$ by back-substitution.
3. Read a column-space basis from the *original* pivot columns, and check rank + nullity $= n$.

Solve. Reduce one step at a time, then express the general null-space vector as a combination of one basis vector per free column.

Check. Plug each null-space basis vector into A to confirm it returns $\mathbf{0}$, and verify rank + nullity $= n$.

Common traps — every wrong answer is one of these

Uses reduced columns for the column-space basis.
Reads the column-space basis off the *reduced* matrix. Pivot *positions* come from the reduced form, but the basis vectors are the corresponding columns of the *original* A.

Adds rank and nullity to m.
Writes rank + nullity $= m$ (rows). The rank–nullity theorem uses n, the number of *columns* (inputs).

Miscounts free variables.
Sets a pivot variable free, or forgets a genuinely free column, giving the wrong nullity. Free variables are exactly the pivotless columns.

Confuses null space with column space.

1. Solves $A\mathbf{x} = \mathbf{0}$ but reports the result as the column space.
2. Mixes up inputs crushed to zero with reachable outputs.
3. Fix: the null space lives in $\mathbb{R}^n$ (inputs); the column space lives in $\mathbb{R}^m$ (outputs) — keep the two straight.

A basis for the null space. Find a basis for the null space of $A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$.

Read.

The matrix is already reduced; solve $A\mathbf{x} = \mathbf{0}$ with the non-pivot columns free.

Plan.

Pivots sit in columns 1 and 3; columns 2 and 4 are free, so set $x_2 = s$, $x_4 = t$. Row 1: $x_1 = -2x_2 - x_4 = -2s - t$. Row 2: $x_3 = -3x_4 = -3t$. Thus $\mathbf{x} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -3 \\ 1 \end{bmatrix}$.

Solve.

Null-space basis: $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -3 \\ 1 \end{bmatrix} \right\}$, so nullity $= 2$.

Check.

Each basis vector, plugged into A, returns $\mathbf{0}$: e.g. row 1 on the first vector is $-2 + 2 = 0$. ✓

Verifying rank–nullity. For the same A, state the rank and confirm the theorem (5.4).

Read.

Count pivots for the rank, count free columns for the nullity, and add.

Plan.

A has pivots in columns 1 and 3, so $\operatorname{rank}(A) = 2$; a basis for the column space is $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ (the pivot columns). The nullity is 2 from Example 5.15. The matrix has $n = 4$ columns.

Solve.

$\text{rank} + \text{nullity} = 2 + 2 = 4 = n$. ✓

Check.

The column space is all of $\mathbb{R}^2$ (rank 2, so A is onto), while the null space is a plane in $\mathbb{R}^4$—the inputs the map sends to zero.

Both bases from an unreduced matrix. Find the rank, a column-space basis, and a null-space basis for $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix}$.

Read.

Row-reduce to locate the pivots, but read the column-space basis from the *original* columns; solve $A\mathbf{x} = \mathbf{0}$ for the null space.

Plan.

Row-reduce, tracking which columns hold pivots:

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_2 - 2R_1,\ R_3 - 3R_1$$

$$\longrightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad R_3 - R_2,\ R_1 - R_2$$

Pivots sit in columns 1 and 3, so $\operatorname{rank}(A) = 2$. Column 2 is free; setting $x_2 = s$ gives $x_1 = -2s$ and $x_3 = 0$, so $\mathbf{x} = s(-2, 1, 0)$.

Solve.

rank $= 2$; column-space basis $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \right\}$ (columns 1 and 3 of the *original* A); null-space basis $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$.

Check.

Rank + nullity $= 2 + 1 = 3 = n$. ✓ The column-space basis uses the *original* pivot columns—reading columns 1 and 3 off the reduced matrix would give $(1, 0, 0), (0, 1, 0)$, the wrong vectors.

Linear Transformations and Their Matrices

A *linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ is a map satisfying $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(c\mathbf{u}) = cT(\mathbf{u})$—it respects addition and scaling, sending the origin to the origin and lines to lines. Every such map *is* matrix multiplication: the matrix of T has as its columns the images of the standard basis vectors, $[\,T(\mathbf{e}_1) \;\cdots\; T(\mathbf{e}_n)\,]$. This is the single most useful fact about transformations—to find the matrix, just see where the basis goes.

Approach Read · Plan · Solve · Check

Read. Identify the domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$, and what data you are given — images of the basis vectors, a geometric description, or a composition.

Plan.

1. Find where each standard basis vector $\mathbf{e}_j$ goes; stack the images $T(\mathbf{e}_j)$ as the columns of the matrix.
2. For a rotation or reflection, use the known image formulas for $\mathbf{e}_1, \mathbf{e}_2$.
3. For a composition "first S, then T," multiply TS (the rightmost map acts first).

Solve. Build the matrix column by column, then multiply it by any input vector to evaluate T.

Check. Re-derive one output from linearity, $T(\sum c_j\mathbf{e}_j) = \sum c_j T(\mathbf{e}_j)$, and confirm it matches the matrix product.

Common traps — every wrong answer is one of these

Puts images in rows, not columns.
Stacks $T(\mathbf{e}_j)$ as rows. The matrix of T has the images of the basis vectors as its *columns*, in order.

Composes in the wrong order.
Reads "first S, then T" as ST. Since the rightmost matrix acts first, the composite is TS.

Assumes every map is linear.
Treats a map that shifts the origin (such as $\mathbf{x} \mapsto \mathbf{x} + \mathbf{c}$) as linear. A linear map must send $\mathbf{0}$ to $\mathbf{0}$ and respect addition and scaling.

Botches a change of basis.
1. Wants the coordinates of $\mathbf{x}$ in a basis $\mathcal{B}$ but reads off $\mathbf{x}$'s standard entries.
2. Skips solving for the coefficients.

3. Fix: solve $c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k = \mathbf{x}$ — the c_i are the new coordinates.

Matrix from images of the basis. A linear map sends $\mathbf{e}_1 \mapsto \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{e}_2 \mapsto \begin{bmatrix} 1 \\ 4 \end{bmatrix}$. Find its matrix and compute $T(3, -1)$.

Read.

Stack the images as columns to get the matrix, then multiply by the input.

Plan.

Stack the images as columns, then multiply by the input:

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \qquad \text{columns are } T(\mathbf{e}_1),\ T(\mathbf{e}_2)$$

$$T\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 3 + 1 \cdot (-1) \\ 3 \cdot 3 + 4 \cdot (-1) \end{bmatrix} \qquad \text{apply } A \text{ to } (3, -1)$$

$$= \begin{bmatrix} 5 \\ 5 \end{bmatrix} \qquad \text{simplify}$$

Solve.

$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$ and $T(3, -1) = (5, 5)$.

Check.

Consistency check: $T(3, -1) = 3\,T(\mathbf{e}_1) - T(\mathbf{e}_2) = 3(2, 3) - (1, 4) = (5, 5)$. ✓

A rotation matrix. Find the matrix of counterclockwise rotation by $90°$ about the origin, and apply it to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Read.

Rotation by θ sends $\mathbf{e}_1 \mapsto (\cos\theta, \sin\theta)$ and $\mathbf{e}_2 \mapsto (-\sin\theta, \cos\theta)$; use $\theta = 90°$.

Plan.

At $\theta = 90°$, $\cos\theta = 0$, $\sin\theta = 1$, so $\mathbf{e}_1 \mapsto (0, 1)$ and $\mathbf{e}_2 \mapsto (-1, 0)$, giving $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

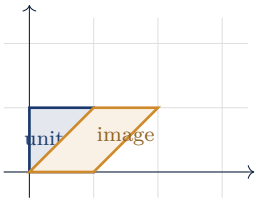
Then $R \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Solve.

$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$; it sends the x-axis unit vector to the y-axis unit vector.

Check.

This is the "multiply by i" rotation of Chapter 1 in matrix form—rotating the plane a quarter turn. ✓



A linear transformation of the unit square under the shear $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$. *Grid lines stay parallel and evenly spaced—the signature of linearity—and the area is preserved because* $\det = 1$.

Composition is matrix multiplication. Let R be rotation by 90° and F reflection across the x-axis. Find the matrix of "first R, then F."

Read.

Composition "first R, then F" is the product FR (rightmost map acts first); multiply the two matrices.

Plan.

The composite "first R, then F" is the product FR:

$$FR = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \qquad \text{multiply out}$$

Solve.

The composite matrix is $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$.

Check.

Test on $\mathbf{e}_1 = (1, 0)$: R sends it to $(0, 1)$, then F to $(0, -1)$—which is column 1 of FR. ✓

Coordinates in a new basis. Write $\mathbf{x} = (5, 5)$ in the basis $\mathcal{B} = \{(1, 1), (1, -1)\}$.

Read.

Find scalars c_1, c_2 with $c_1(1, 1) + c_2(1, -1) = (5, 5)$: solve the resulting 2×2 system.

Plan.

Matching coordinates gives $c_1 + c_2 = 5$ and $c_1 - c_2 = 5$. Adding, $2c_1 = 10$ so $c_1 = 5$; then $c_2 = 0$.

Solve.

The coordinate vector is $[\mathbf{x}]_{\mathcal{B}} = (5, 0)$: $\mathbf{x} = 5(1, 1) + 0(1, -1)$.

Check.

Indeed $5(1, 1) = (5, 5)$. ✓ Changing basis just re-expresses the same vector in new directions—here $\mathbf{x}$ happens to point purely along the first basis vector.

Eigenvalues, Eigenvectors, and Diagonalization

For most vectors, multiplying by A both stretches and rotates them. An *eigenvector* is a rare direction the matrix leaves *fixed in direction*, merely scaling it:

$$A\mathbf{v} = \lambda\mathbf{v} \tag{5.5}$$

for a scalar λ, the *eigenvalue*. These special directions are the skeleton of the transformation, and finding them turns a tangled matrix into a diagonal one.

Derivation. *the characteristic equation*

Rewrite $A\mathbf{v} = \lambda\mathbf{v}$ as $(A - \lambda I)\mathbf{v} = \mathbf{0}$. A nonzero $\mathbf{v}$ solves this precisely when $A - \lambda I$ is singular—has a nontrivial null space—which happens exactly when

$$\det(A - \lambda I) = 0. \tag{5.6}$$

This *characteristic equation* is a degree-n polynomial in λ; its roots are the eigenvalues, and for each, the null space of $A - \lambda I$ supplies the eigenvectors.

Approach Read · Plan · Solve · Check

Read. Confirm A is square, and decide whether you need eigenvalues, eigenvectors, a diagonalization, or a matrix power.

Plan.

1. Form $A - \lambda I$ (subtract λ down the diagonal) and set $\det(A - \lambda I) = 0$.
2. Solve the characteristic polynomial for the eigenvalues; use trace = sum and det = product as a check.
3. For each λ, find the null space of $A - \lambda I$ to get eigenvectors; assemble P (eigenvectors) and D (eigenvalues) for $A = PDP^{-1}$.

Solve. Expand the determinant, factor the polynomial, then solve each homogeneous system $(A - \lambda I)\mathbf{v} = \mathbf{0}$ one eigenvalue at a time.

Check. Verify $A\mathbf{v} = \lambda\mathbf{v}$ for each pair, and confirm the eigenvalues sum to the trace and multiply to $\det A$.

Common traps — every wrong answer is one of these

Forgets to subtract λI.
Computes $\det A = 0$ instead of $\det(A - \lambda I) = 0$. You must subtract λ from every diagonal entry *before* taking the determinant.

Subtracts λ from every entry.
Forms $A - \lambda$ by subtracting λ from all entries, not just the diagonal. Only the diagonal changes: $A - \lambda I$.

Uses the zero vector as an eigenvector.
Reports $\mathbf{v} = \mathbf{0}$ as a solution of $(A - \lambda I)\mathbf{v} = \mathbf{0}$. Eigenvectors are *nonzero* by definition; take a nonzero null-space vector.

Pairs P and D inconsistently.
1. Puts eigenvectors in P in one order and eigenvalues in D in another.
2. Finds $A \neq PDP^{-1}$.
3. Fix: column k of P must be the eigenvector for the kth diagonal entry of D — keep the pairing.

Eigenvalues and eigenvectors of a 2×2. Diagonalize the analysis of $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$: find its eigenvalues and an eigenvector for each.

Read.

Solve $\det(A - \lambda I) = 0$ for the eigenvalues, then find the null space of $A - \lambda I$ for each.

Plan.

Solve the characteristic equation, then find each eigenvector:

$$\begin{aligned} \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} &= (2-\lambda)^2 - 1 && \text{subtract } \lambda \text{ down the diagonal} \\ &= \lambda^2 - 4\lambda + 3 && \text{expand} \\ &= (\lambda - 1)(\lambda - 3) && \text{factor; roots } \lambda = 1, 3 \end{aligned}$$

For $\lambda = 3$: $A - 3I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ gives $-x + y = 0$, i.e. $\mathbf{v} = (1, 1)$. For $\lambda = 1$: $A - I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ gives $x + y = 0$, i.e. $\mathbf{v} = (1, -1)$.

Solve.

Eigenpairs: $\lambda_1 = 3$ with $\mathbf{v}_1 = (1, 1)$; $\lambda_2 = 1$ with $\mathbf{v}_2 = (1, -1)$.

Check.

Check: $A(1, 1) = (3, 3) = 3(1, 1)$. ✓ Also trace $= 4 = 1 + 3$ and $\det = 3 = 1 \cdot 3$, as they must.

Trace and determinant from eigenvalues. For any square A, the sum of the eigenvalues equals the trace (sum of diagonal entries) and their product equals $\det A$. These give a fast sanity check—and often let you find a second eigenvalue once you know the first.

Eigenvalues of a triangular matrix. Find the eigenvalues of $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{bmatrix}$.

Read.

For a triangular matrix, $\det(B - \lambda I)$ is a product of diagonal terms; read the eigenvalues off the diagonal.

Plan.

$\det(B - \lambda I) = (2-\lambda)(3-\lambda)(5-\lambda)$, since the matrix stays triangular. Setting it to zero gives the roots directly.

Solve.

Eigenvalues $\lambda = 2, 3, 5$—the diagonal entries.

Check.

Their sum $2 + 3 + 5 = 10$ equals the trace, and their product 30 equals $\det B$. ✓

To *diagonalize* A is to write

$$A = PDP^{-1}, \tag{5.7}$$

where D is diagonal (the eigenvalues) and P has the corresponding eigenvectors as columns. This is possible exactly when A has a full set of independent eigenvectors, and it makes powers trivial: $A^k = PD^kP^{-1}$.

Diagonalizing a matrix. Diagonalize $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

Read.

Use the eigenpairs from Example 5.20: columns of P are the eigenvectors, D holds the matching eigenvalues.

Plan.

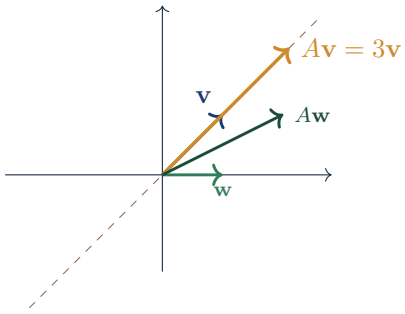
Take $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (eigenvectors for $\lambda = 3, 1$) and $D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$. Then $P^{-1} = \frac{1}{-2}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Solve.

$A = PDP^{-1}$ with $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$.

Check.

Then $A^k = P\begin{bmatrix} 3^k & 0 \\ 0 & 1 \end{bmatrix}P^{-1}$—a closed form for every power, the main payoff of diagonalization. ✓



Eigenvector vs. ordinary vector for $A = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$. The eigenvector $\mathbf{v} = (1, 1)$ (dashed line) is only stretched*: $A\mathbf{v} = 3\mathbf{v}$ stays on its line. A generic $\mathbf{w} = (1, 0)$ is turned off its line to $A\mathbf{w} = (2, 1)$.*

An eigenvector of a triangular matrix. Find an eigenvector of $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{bmatrix}$ for the eigenvalue $\lambda = 2$.

Read.

Solve $(B - 2I)\mathbf{v} = \mathbf{0}$ by back-substitution from the bottom row up.

Plan.

$B - 2I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$. The bottom row gives $3z = 0 \Rightarrow z = 0$; the middle row $y + 4z = 0 \Rightarrow y = 0$; the top row $y = 0$ is consistent, leaving x free.

Solve.

An eigenvector is $\mathbf{v} = (1, 0, 0)$ (any nonzero multiple).

Check.

Check: $B(1, 0, 0) = (2, 0, 0) = 2(1, 0, 0)$. ✓

A matrix power via diagonalization. Using $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = PDP^{-1}$ from Example 5.22, compute A^3.

Read.

Diagonalization (5.7) gives $A^k = PD^kP^{-1}$; cube the diagonal entries, then multiply the three matrices.

Plan.

Cube the diagonal, then multiply the three matrices:

$$D^3 = \operatorname{diag}(27, 1) \qquad \text{cube each eigenvalue in } D = \operatorname{diag}(3, 1)$$

$$PD^3 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 27 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 27 & 1 \\ 27 & -1 \end{bmatrix} \qquad \text{multiply } P \text{ by } D^3$$

$$A^3 = \begin{bmatrix} 27 & 1 \\ 27 & -1 \end{bmatrix} \cdot \tfrac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \tfrac{1}{2} \begin{bmatrix} 28 & 26 \\ 26 & 28 \end{bmatrix} \qquad \text{multiply by } P^{-1}$$

Solve.

$A^3 = \begin{bmatrix} 14 & 13 \\ 13 & 14 \end{bmatrix}$.

Check.

Cross-check by direct multiplication: $A^2 = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$, and $A^2 A$ gives $\begin{bmatrix} 14 & 13 \\ 13 & 14 \end{bmatrix}$. ✓ Diagonalization replaced repeated matrix products with cubing two numbers.

A non-symmetric 2×2**.** Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$.

Read.

Solve the characteristic equation (5.6), then find the null space of $A - \lambda I$ for each root.

Plan.

Solve the characteristic equation, then find each eigenvector:

$$\begin{aligned} \det(A - \lambda I) &= (4 - \lambda)(3 - \lambda) - 2 \cdot 1 && \text{subtract } \lambda \text{ down the diagonal} \\ &= \lambda^2 - 7\lambda + 10 && \text{expand} \\ &= (\lambda - 2)(\lambda - 5) && \text{factor; roots } \lambda = 2, 5 \end{aligned}$$

For $\lambda = 5$: $A - 5I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$ gives $-x + y = 0$, so $\mathbf{v} = (1, 1)$. For $\lambda = 2$: $A - 2I = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$ gives $2x + y = 0$, so $\mathbf{v} = (1, -2)$.

Solve.

Eigenpairs $\lambda = 5,\ (1, 1)$ and $\lambda = 2,\ (1, -2)$.

Check.

Trace $4+3=7=5+2$ and det $=12-2=10=5\cdot 2$. ✓ Eigenvectors need not be perpendicular when A is not symmetric—here they are not.

Problem-solving strategy — an eigenvalue problem

(1) Form $A-\lambda I$ and set $\det(A-\lambda I)=0$. (2) Solve this characteristic polynomial for the eigenvalues; use trace = sum and det = product as a check. (3) For each λ, find the null space of $A-\lambda I$ to get eigenvectors satisfying $A\mathbf{v}=\lambda\mathbf{v}$ (5.5). (4) If you have n independent eigenvectors, form P (eigenvectors) and D (eigenvalues) so $A=PDP^{-1}$.

Orthogonality and Projection

The *dot product* $\mathbf{u}\cdot\mathbf{v}=\sum u_iv_i$ measures alignment: it equals $\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta$, so two nonzero vectors are *orthogonal* (perpendicular) exactly when their dot product is zero. Orthogonality tames geometry: it lets us drop a perpendicular from a point to a line, the operation called *projection*, which is the engine of least-squares fitting.

Approach Read · Plan · Solve · Check

Read. Identify the vectors and the goal — an angle, a projection, an orthogonal or orthonormal set, or a least-squares fit — and note which vector is projected onto which.

Plan.

1. Compute the dot products you need; recall $\mathbf{u}\cdot\mathbf{v}=\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta$ and that orthogonal means dot product 0.
2. For a projection, apply $\operatorname{proj}_{\mathbf{a}}\mathbf{b}=\frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}$; for Gram–Schmidt, subtract projections onto the directions already fixed.
3. For least squares, solve the normal equations $A^{\top}A\,\mathbf{x}=A^{\top}\mathbf{b}$.

Solve. Evaluate the dot products first, then substitute into the formula one step at a time.

Check. Confirm the residual $\mathbf{b}-$ proj is orthogonal to $\mathbf{a}$ (dot product 0), and that any orthonormal vectors have length 1.

Common traps — every wrong answer is one of these

Forgets to divide by $\mathbf{a} \cdot \mathbf{a}$.
Writes $\text{proj}_{\mathbf{a}}\,\mathbf{b} = (\mathbf{a} \cdot \mathbf{b})\,\mathbf{a}$. The projection divides by $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$; only when $\mathbf{a}$ is a unit vector does that factor drop out.

Projects onto the wrong vector.
Swaps the roles and computes $\text{proj}_{\mathbf{b}}\,\mathbf{a}$. The vector in the subscript is the *direction* projected onto; keep $\mathbf{a}$ and $\mathbf{b}$ in their roles.

Normalizes with the wrong length.
Divides by $\mathbf{u} \cdot \mathbf{u}$ instead of $\sqrt{\mathbf{u} \cdot \mathbf{u}}$ when making a unit vector. To normalize, divide by the *length* $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$.

Solves the wrong system for least squares.
1. Tries to solve $A\mathbf{x} = \mathbf{b}$ directly for an overdetermined system.
2. Finds no solution and stops.
3. Fix: multiply through by $A^\top$ and solve the normal equations $A^\top A\,\mathbf{x} = A^\top \mathbf{b}$.

Angle between two vectors. Find the angle between $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (1, 1)$.

Read.

Use $\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$, then take the inverse cosine.

Plan.

$\mathbf{u} \cdot \mathbf{v} = 1 \cdot 1 + 0 \cdot 1 = 1$, $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = \sqrt{2}$, so $\cos\theta = \frac{1}{1 \cdot \sqrt{2}} = \frac{1}{\sqrt{2}}$.

Solve.

$\theta = 45°$ (i.e. $\frac{\pi}{4}$).

Check.

A positive dot product gives an acute angle, as expected for two vectors both pointing into the first quadrant. ✓

Did You Know? *eigenvalues rank the web*

The idea that a matrix has special stretching directions is not academic: Google's original PageRank computes the dominant eigenvector of a giant matrix built from the web's link structure, and that eigenvector's entries are the page rankings. Markov chains, vibration modes, and quantum energy levels are all eigenvalue problems in disguise—which is why this

is the chapter's summit.

Orthogonal projection onto a line. The projection of $\mathbf{b}$ onto the line through $\mathbf{a}$ is

$$\operatorname{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}. \tag{5.8}$$

The leftover $\mathbf{b} - \operatorname{proj}_{\mathbf{a}} \mathbf{b}$ is orthogonal to $\mathbf{a}$.

Projecting one vector onto another. Project $\mathbf{b} = (2, 3)$ onto $\mathbf{a} = (1, 1)$.

Read.

Apply the projection formula, computing the two dot products first.

Plan.

Compute the two dot products, then apply the formula:

$\mathbf{a} \cdot \mathbf{b} = 1 \cdot 2 + 1 \cdot 3 = 5$	numerator
$\mathbf{a} \cdot \mathbf{a} = 1 + 1 = 2$	denominator $\|\mathbf{a}\|^2$
$\operatorname{proj}_{\mathbf{a}} \mathbf{b} = \frac{5}{2}(1, 1) = \left(\frac{5}{2}, \frac{5}{2}\right)$	scale $\mathbf{a}$ by $\frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}$

Solve.

$\operatorname{proj}_{\mathbf{a}} \mathbf{b} = \left(\frac{5}{2}, \frac{5}{2}\right)$.

Check.

The residual $\mathbf{b} - \text{proj} = \left(-\frac{1}{2}, \frac{1}{2}\right)$ has dot product $1 \cdot (-\frac{1}{2}) + 1 \cdot \frac{1}{2} = 0$ with $\mathbf{a}$—perpendicular, as required. ✓

The *Gram–Schmidt process* turns any independent set into an orthogonal one by successively subtracting off projections onto the directions already fixed.

Gram–Schmidt on two vectors. Orthogonalize $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (1, 0, 1)$.

Read.

Keep $\mathbf{u}_1 = \mathbf{v}_1$; subtract from $\mathbf{v}_2$ its projection onto $\mathbf{u}_1$.

Plan.

Keep $\mathbf{u}_1 = (1,1,0)$; subtract from $\mathbf{v}_2$ its projection onto $\mathbf{u}_1$:

$$\frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \frac{1\cdot 1 + 0\cdot 1 + 1\cdot 0}{1+1+0} = \tfrac{1}{2} \qquad \text{projection coefficient}$$

$$\mathbf{u}_2 = (1,0,1) - \tfrac{1}{2}(1,1,0) = \left(\tfrac{1}{2}, -\tfrac{1}{2}, 1\right) \qquad \mathbf{v}_2 \text{ minus its projection}$$

Solve.

Orthogonal set $\mathbf{u}_1 = (1,1,0),\ \mathbf{u}_2 = \left(\tfrac{1}{2}, -\tfrac{1}{2}, 1\right)$.

Check.

Verify orthogonality: $\mathbf{u}_1 \cdot \mathbf{u}_2 = \tfrac{1}{2} - \tfrac{1}{2} + 0 = 0$. ✓ Dividing each by its length would make the set orthonormal.

Recognizing an orthogonal matrix. Show $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is orthogonal and state Q^{-1}.

Read.

A matrix is orthogonal iff its columns are orthonormal iff $Q^{\top}Q = I$; then $Q^{-1} = Q^{\top}$.

Plan.

The columns $(0,1)$ and $(-1,0)$ are unit length and perpendicular. Confirm directly:

$$Q^{\top}Q = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \text{multiply } Q^{\top} \text{ by } Q$$

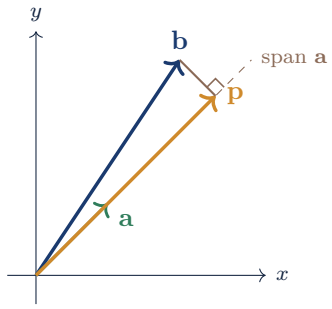
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \qquad \text{columns are orthonormal}$$

Solve.

Q is orthogonal, and $Q^{-1} = Q^{\top} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Check.

Orthogonal matrices are exactly the rotations and reflections—rigid motions that preserve length and angle, so $\det Q = \pm 1$ (here $\det Q = 1$, a rotation). ✓



Orthogonal projection of $\mathbf{b}$ *onto the line spanned by* $\mathbf{a}$. *The projection* $\mathbf{p} = \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}$ *is the closest point on the line to* $\mathbf{b}$*; the error* $\mathbf{b} - \mathbf{p}$ *meets the line at a right angle—the geometric heart of least squares.*

Normalizing to an orthonormal set. Normalize the orthogonal pair $\mathbf{u}_1 = (1, 1, 0)$, $\mathbf{u}_2 = \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$ from Example 5.25.

Read.

Divide each vector by its length $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$.

Plan.

$\|\mathbf{u}_1\| = \sqrt{1+1+0} = \sqrt{2}$, so $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1, 0)$. Next $\|\mathbf{u}_2\| = \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2}$, so $\mathbf{e}_2 = \frac{2}{\sqrt{6}}\left(\frac{1}{2}, -\frac{1}{2}, 1\right) = \frac{1}{\sqrt{6}}(1, -1, 2)$.

Solve.

Orthonormal basis $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1, 0)$, $\mathbf{e}_2 = \frac{1}{\sqrt{6}}(1, -1, 2)$.

Check.

Each has length 1 and $\mathbf{e}_1 \cdot \mathbf{e}_2 = \frac{1}{\sqrt{12}}(1 - 1 + 0) = 0$. ✓

Least squares by the normal equations. Fit the best line $y = a + bx$ to the points $(0, 1), (1, 2), (2, 2)$.

Read.

Write the overdetermined system $A\mathbf{x} = \mathbf{b}$, then solve the normal equations $A^{\top}A\,\mathbf{x} = A^{\top}\mathbf{b}$.

Plan.

With $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{b} = (1, 2, 2)$, and $\mathbf{x} = (a, b)$, form and solve the normal equations:

$$A^\top A = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \qquad A^\top \mathbf{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \qquad \text{set up } A^\top A\,\mathbf{x} = A^\top \mathbf{b}$$

$$2b = 1 \qquad \text{subtract } 3a + 3b = 5 \text{ from } 3a + 5b = 6$$

$$b = \tfrac{1}{2} \qquad \text{solve for } b$$

$$a = \tfrac{7}{6} \qquad \text{back-substitute into } 3a + 3b = 5$$

Solve.

Best-fit line $y = \frac{7}{6} + \frac{1}{2}x$.

Check.

No line passes through all three points (they are not collinear), so least squares finds the line minimizing the squared vertical errors—the projection of $\mathbf{b}$ onto the column space of A. ✓

Synthesis — Key Results You Must Own

1. Matrix product = composition of maps; (i, j) entry is row i dotted with column j; $AB \neq BA$ in general, and $(AB)^\top = B^\top A^\top$.
2. Solve $A\mathbf{x} = \mathbf{b}$ by row-reducing $[\,A \mid \mathbf{b}\,]$; a contradiction row means no solution, a free column means infinitely many.
3. $\det A \neq 0 \iff A$ invertible $\iff$ columns independent $\iff 0$ is not an eigenvalue; 2×2 inverse is swap–negate–divide.
4. Independence: only the trivial combination gives $\mathbf{0}$. A basis is a minimal spanning (equivalently, maximal independent) set; its size is the dimension.
5. Rank + nullity = number of columns; pivot columns basis the column space, free columns basis the null space.
6. The matrix of a linear map has the images of the basis vectors as its columns.

7. Eigenvalues solve $\det(A - \lambda I) = 0$; eigenvectors span $\ker(A - \lambda I)$; a full set diagonalizes $A = PDP^{-1}$, so $A^k = PD^kP^{-1}$. Sum of eigenvalues = trace, product = det.
8. Orthogonal projection: $\text{proj}_{\mathbf{a}}\,\mathbf{b} = \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}$; Gram–Schmidt subtracts projections; an orthogonal matrix has $Q^{-1} = Q^{\top}$.

Practice Problems

Difficulty: ⋆ warm-up, ⋆⋆ standard, ⋆⋆⋆ challenge. Work each with the Read–Plan–Solve–Check habit modeled above.

Matrices and operations

1 ⋆.

For $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 1 \\ 2 & 5 \end{bmatrix}$, compute AB.

2 ⋆.

Give the transpose of $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and state its dimensions.

3 ⋆⋆.

Find a 2×2 matrix $A \neq 0$ with $A^2 = 0$.

Hint: Try a strictly upper-triangular matrix.

4 ⋆⋆.

If A is 3×2 and B is 2×4, which of AB, BA, $A^{\top}B^{\top}$ are defined, and what are their sizes?

Systems of equations

5 ⋆.

Solve $\begin{aligned} x + y &= 5 \\ x - y &= 1. \end{aligned}$

6 ⋆⋆.

Solve by elimination:
$$\begin{aligned} x + 2y + z &= 3 \\ 2x + 5y - z &= 4 \\ 3x + 7y &= 7. \end{aligned}$$

7 ⋆⋆.

For which k does $\begin{aligned} x + 2y &= 1 \\ 3x + ky &= 5 \end{aligned}$ have no solution?

8 ⋆⋆.

Row-reduce $\begin{bmatrix} 1 & 3 & 1 \\ 2 & 6 & 4 \end{bmatrix}$ to RREF and describe the solution set of the homogeneous system.

9 ⋆⋆⋆.

For which values of a does $\begin{aligned} x + y + z &= 1 \\ x + 2y + z &= 2 \\ x + y + az &= 3 \end{aligned}$ have a unique solution, and what is it in terms of a?

Determinants and inverses

10 ⋆.

Compute $\det \begin{bmatrix} 4 & 3 \\ 2 & 5 \end{bmatrix}$.

11 ⋆⋆.

Compute $\det \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$.

12 ⋆⋆.

Invert $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$.

13 ⋆⋆.

If $\det A = 4$ for a 3×3 matrix, find $\det(2A)$ and $\det(A^{-1})$.

Hint: $\det(cA) = c^n \det A$ for an $n \times n$ matrix.

14 ⋆⋆⋆.

Show that if $A^2 = A$ and A is invertible, then $A = I$.

Independence, span, and basis

15 ⋆.

Are $(1, 0, 1)$ and $(2, 0, 2)$ linearly independent? Explain.

16 ⋆⋆.

Do $(1, 1, 0)$, $(0, 1, 1)$, $(1, 0, 1)$ form a basis of $\mathbb{R}^3$?

17 ⋆⋆.

Find a basis for the span of $(1, 2, 1)$, $(2, 4, 2)$, $(1, 0, 3)$ and give its dimension.

18 ⋆⋆.

Is $\{(x, y, z) : xyz = 0\}$ a subspace of $\mathbb{R}^3$? Justify.

19 ⋆⋆⋆.

Show the polynomials 1, $1 + x$, $1 + x + x^2$ form a basis of the space of polynomials of degree ≤ 2.

Rank, nullity, transformations

20 ⋆⋆.

Find the rank and nullity of $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix}$.

21 ⋆⋆.

Find a basis for the null space of $\begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \end{bmatrix}$.

22 ⋆⋆.

A linear map $T:\mathbb{R}^2\to\mathbb{R}^2$ sends $(1,0)\mapsto(1,2)$ and $(0,1)\mapsto(3,4)$. Find its matrix and $T(2,-1)$.

23 ⋆⋆.

Write the matrix of reflection across the line $y=x$ in $\mathbb{R}^2$.

24 ⋆⋆⋆.

If $T:\mathbb{R}^3\to\mathbb{R}^2$ is linear and onto, what are its rank and nullity? Justify with rank–nullity.

Eigenvalues and diagonalization

25 ⋆.

Find the eigenvalues of $\begin{bmatrix}4 & 0\\ 0 & -2\end{bmatrix}$.

26 ⋆⋆.

Find the eigenvalues and eigenvectors of $\begin{bmatrix}3 & 1\\ 0 & 2\end{bmatrix}$.

27 ⋆⋆.

The 2×2 matrix A has trace 5 and determinant 6. Find its eigenvalues.

28 ⋆⋆⋆.

Diagonalize $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ and use it to compute the matrix's 10th power.

Orthogonality

29 ⋆⋆.

Project $\mathbf{b}=(3,4)$ onto $\mathbf{a}=(1,2)$.

30 ⋆⋆⋆.

Apply Gram–Schmidt to $(1,1,1)$ and $(1,1,0)$ to produce an orthogonal pair.

Further Reading

The geometric viewpoint of this chapter—transformations, the four subspaces, and diagonalization as a change to eigenvector coordinates—is developed beautifully in any modern first course that leads with linear maps rather than with determinants. Chapter 6 turns to number theory and abstract algebra, where the notion of *structure preserved by a map* (here, linearity) reappears as the homomorphism.

CHAPTER 6

Number Theory & Abstract Algebra

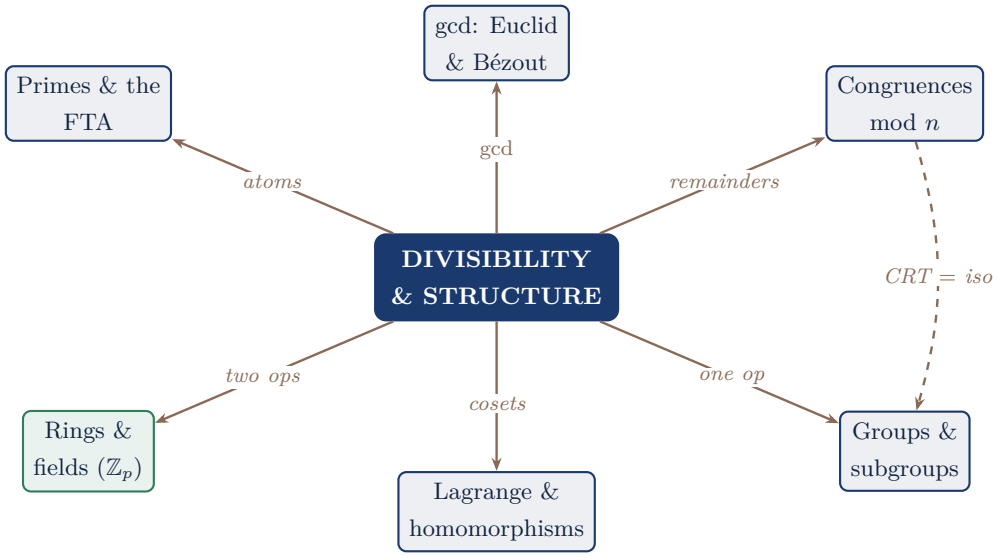
Integers, Groups, Rings, and Fields

"Mathematics is the queen of the sciences, and number theory is the queen of mathematics."

— Carl Friedrich Gauss

This chapter is where arithmetic grows up. Part A,* number theory*, studies the integers through the single lens of divisibility: primes are the multiplicative atoms, the greatest common divisor is computed by an ancient and unreasonably fast algorithm, and "clock arithmetic" modulo n turns questions about remainders into ordinary algebra. Part B,* abstract algebra*, is the discovery that the* structure *behind that arithmetic—a set with one well-behaved operation—appears everywhere, and deserves study on its own. A* group *is that structure stripped to essentials;* subgroups*,* cyclic *generators,* Lagrange's theorem*, and* homomorphisms *are the tools for taking one apart;* rings *and* fields *add a second operation and lead back to $\mathbb{Z}_n$. The two halves are one idea: the Chinese Remainder Theorem of Part A is an* isomorphism *of Part B in disguise. The Subject Test asks compact, computational questions here—compute a gcd, solve a congruence,

find the order of an element, decide whether a structure is a field—so fluency with the standard moves is worth far more than a shelf of theorems.



Learning Objectives

You will be able to:

1. use divisibility and the Fundamental Theorem of Arithmetic to factor, count divisors, and reason about primes;
2. run the Euclidean algorithm for a gcd and back-substitute for a Bézout combination;
3. compute in modular arithmetic, solve linear congruences, and apply the Chinese Remainder Theorem;
4. use Fermat's Little Theorem and Euler's totient to reduce large powers modulo n;
5. verify the group axioms, find the order of an element, and list the subgroups of a cyclic group;
6. apply Lagrange's theorem and cosets, and test whether a map is a homomorphism or isomorphism;
7. decide whether a given structure is a group, ring, integral domain, or field.

— Part A: Number Theory —

Divisibility, Primes, and Unique Factorization

An integer d *divides* a, written $d \mid a$, if $a = dk$ for some integer k; equivalently, a leaves remainder 0 on division by d. Divisibility is the entire subject of Part A: primes, gcds, and congruences are all

built from it. Three facts are used constantly and follow straight from the definition: if $d \mid a$ and $d \mid b$ then $d \mid (a+b)$ and $d \mid (a-b)$; more generally d divides every *integer combination* $ma+nb$; and if $d \mid a$ then $d \mid ac$ for any integer c.

A *prime* is an integer $p > 1$ whose only positive divisors are 1 and p; an integer > 1 that is not prime is *composite*. Primes are the multiplicative atoms, made precise by the theorem every later result leans on.

Fundamental Theorem of Arithmetic (FTA). Every integer $n > 1$ factors into primes, $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, and this factorization is *unique* up to the order of the factors. Consequently the number of positive divisors of n is $(a_1+1)(a_2+1)\cdots(a_k+1)$.

What & why. What it means for one integer to divide another, and why primes are the building blocks of every integer.

Solve. Factor into primes; read off divisors, gcd, and lcm from the exponents.

Means. Divisibility is remainder-zero; primes are the atoms of multiplication.

Approach Read · Plan · Solve · Check

Read. Identify what is asked—test a divisibility, factor an integer, count or sum divisors—and write each hypothesis $d \mid a$ as an equation $a = dk$.

Plan.

1. To prove d divides a combination, substitute $a = dk$, $b = d\ell$ and factor d out.
2. To factor, strip primes smallest-first: divide by 2, then 3, then 5, … until 1.
3. Read the divisor *count* as $\prod(a_i+1)$ and the divisor *sum* as $\prod(1+p+\cdots+p^{a_i})$.

Solve. Carry out the arithmetic one line at a time, keeping the prime exponents explicit.

Check. Multiply the prime factors back to the original, and confirm a small case of any divisor formula by listing.

Common traps — every wrong answer is one of these

Multiplies the exponents.

Counts the divisors of $2^3 \cdot 3^2 \cdot 5$ as $3 \cdot 2 \cdot 1$ instead of $(3+1)(2+1)(1+1)$. Each exponent ranges from 0 up to a_i, so add one *before* multiplying.

Calls 1 a prime.

Treats 1 as prime (or forgets that 2 is the smallest prime). 1 is a unit—neither prime nor

composite—and counting it would destroy the uniqueness in the FTA.

Reverses the divisibility inference.
Assumes $d \mid (a+b)$ forces $d \mid a$ and $d \mid b$. The implication runs the other way: $3 \mid (1+2)$ yet $3 \nmid 1$.

Under-counts a prime in a factorial.
1. Counts only the multiples of p in $n!$.
2. Stops there, ignoring $p^2, p^3, \ldots$.
3. Fix: use Legendre—$\lfloor n/p \rfloor + \lfloor n/p^2 \rfloor + \cdots$—since each multiple of p^2 contributes an extra factor.

Divisibility of a combination. Show that if $d \mid a$ and $d \mid b$, then $d \mid (5a - 3b)$.

Read.

Write each hypothesis as an equation and substitute; factor d out of the result.

Plan.

Since $d \mid a$, write $a = dk$; since $d \mid b$, write $b = d\ell$ for integers k, ℓ. Then

$$\begin{aligned} 5a - 3b &= 5(dk) - 3(d\ell) && \text{substitute } a = dk,\ b = d\ell \\ &= d(5k - 3\ell) && \text{factor out } d. \end{aligned}$$

The factor $5k - 3\ell$ is an integer, so d divides $5a - 3b$.

Solve.

$d \mid (5a - 3b)$.

Check.

Nothing special about 5 and -3: any integer combination $ma + nb$ works the same way. This is the fact behind Bézout's identity below. ✓

Prime factorization and counting divisors. Factor 360 and count its positive divisors.

Read.

Strip out primes smallest-first; then apply the divisor-count formula to the exponents.

Plan.

Divide by 2 repeatedly, then factor what remains:

$$\begin{aligned} 360 &= 2 \cdot 180 = 2^2 \cdot 90 = 2^3 \cdot 45 && \text{strip out the factors of 2} \\ &= 2^3 \cdot 3^2 \cdot 5 && \text{since } 45 = 9 \cdot 5 = 3^2 \cdot 5. \end{aligned}$$

The number of divisors is

$$\begin{aligned} (3+1)(2+1)(1+1) &= 4 \cdot 3 \cdot 2 && \text{add one to each exponent} \\ &= 24 && \text{multiply.} \end{aligned}$$

Solve.

$360 = 2^3 \cdot 3^2 \cdot 5$, with 24 positive divisors.

Check.

A quick check: $8, 9, 5$ are pairwise coprime and $8 \cdot 9 \cdot 5 = 360$. The divisor count uses each exponent *plus one* because an exponent may range from 0 up to a_i. ✓

Trailing zeros of a factorial. How many zeros end the decimal expansion of 100!?

Read.

Each trailing zero is a factor of $10 = 2 \cdot 5$; fives are scarcer than twos, so count the factors of 5 in 100!.

Plan.

The exponent of the prime 5 in 100! is

$$\begin{aligned} \left\lfloor \tfrac{100}{5} \right\rfloor + \left\lfloor \tfrac{100}{25} \right\rfloor + \left\lfloor \tfrac{100}{125} \right\rfloor &= 20 + 4 + 0 && \text{count multiples of } 5,\ 25,\ 125 \\ &= 24 && \text{each multiple of 25 adds an } \textit{extra}\ 5. \end{aligned}$$

Factors of 2 are far more plentiful, so 5 is the bottleneck.

Solve.

100! ends in 24 zeros.

Check.

Legendre's counting—divide by p, then p^2, then p^3—is the standard way to read a prime's exponent out of a factorial. ✓

Sum of the divisors. Find the sum of all positive divisors of 72.

Read.

Factor into primes; the divisor-sum σ is a product of geometric sums, one per prime.

Plan.

$72 = 2^3 \cdot 3^2$. Each divisor is $2^i 3^j$ with $0 \le i \le 3$, $0 \le j \le 2$, so the sum factors:

$$\begin{aligned} \sigma(72) &= (1+2+4+8)(1+3+9) && \text{one geometric sum per prime} \\ &= 15 \cdot 13 && \text{add each factor} \\ &= 195 && \text{multiply.} \end{aligned}$$

Solve.

$\sigma(72) = 195$.

Check.

The same factorization gives the divisor *count* $(3+1)(2+1) = 12$; both σ and the count are multiplicative over coprime parts. ✓

Did You Know? *there are infinitely many primes—and Euclid proved it in one line*

If $p_1, \ldots, p_k$ were *all* the primes, the number $N = p_1 p_2 \cdots p_k + 1$ leaves remainder 1 on division by each p_i, so none of them divides N. Then N has a prime factor outside the list—a contradiction. The primes never run out.

The Euclidean Algorithm and Bézout's Identity

The *greatest common divisor* $\gcd(a, b)$ is the largest integer dividing both a and b. You could factor both numbers and take the smaller exponent on each shared prime, but factoring large integers is slow. The *Euclidean algorithm* finds the gcd without any factoring, using one repeated observation: dividing a by b to get $a = qb + r$, any common divisor of a and b also divides $r = a - qb$, and vice versa. So the pair (a, b) and the smaller pair (b, r) have exactly the same common divisors—and the same gcd.

Derivation. *the Euclidean algorithm never changes the gcd*

Write $a = qb + r$. If $d \mid a$ and $d \mid b$, then $d \mid (a - qb) = r$, so d is a common divisor of b and r. Conversely, if $d \mid b$ and $d \mid r$, then $d \mid (qb + r) = a$. The two pairs (a, b) and (b, r) therefore have *identical* sets of common divisors, hence the same greatest one: $\gcd(a, b) = \gcd(b, r)$. Repeating drives the second entry down until it hits 0; since $\gcd(g, 0) = g$, the last nonzero remainder *is* the gcd.

Approach Read · Plan · Solve · Check

Read. Identify the two (or more) integers and the target—a gcd, a Bézout combination, an lcm, or a linear Diophantine equation.

Plan.

1. Run Euclid: divide the larger by the smaller, replace the pair by (smaller, remainder), and repeat until the remainder is 0.
2. The last nonzero remainder is the gcd; back-substitute those divisions to write it as $ax + by$.
3. For an lcm use $\text{lcm} = ab/\gcd$; for $ax + by = c$, first confirm $\gcd(a, b) \mid c$.

Solve. Record one division per line, then reverse the chain one substitution at a time.

Check. Factor both integers to confirm the gcd, or substitute the Bézout coefficients back into $ax + by$.

Common traps — every wrong answer is one of these

Swaps gcd and lcm.
Takes the *larger* exponent on each shared prime (that gives the lcm) when computing a gcd. The gcd takes the *smaller* exponent, the lcm the larger, and $\gcd \cdot \text{lcm} = ab$.

Reads the wrong remainder as the gcd.
Reports the final 0, or the last quotient, as $\gcd(a, b)$. The gcd is the *last nonzero remainder* in the division chain.

Thinks Bézout coefficients are unique.
Treats the x, y in $ax + by = \gcd$ as the only pair. Adding $(b/g, -a/g)$ to (x, y) gives another valid pair—an entire family.

Solves an unsolvable Diophantine equation.
1. Assumes $ax + by = c$ always has integer solutions.
2. Hunts for one regardless of c.

3. Fix: it is solvable iff $\gcd(a, b) \mid c$; check that divisibility *before* solving.

gcd by the Euclidean algorithm. Compute $\gcd(252, 198)$.

Read.

Repeatedly replace the larger number by the remainder on dividing it by the smaller, until the remainder is 0.

Plan.

Divide and record remainders:

$$252 = 1 \cdot 198 + 54, \quad 198 = 3 \cdot 54 + 36, \quad 54 = 1 \cdot 36 + 18, \quad 36 = 2 \cdot 18 + 0.$$

The last nonzero remainder is 18.

Solve.

$\gcd(252, 198) = 18$.

Check.

Check by factoring: $252 = 2^2 \cdot 3^2 \cdot 7$ and $198 = 2 \cdot 3^2 \cdot 11$; the shared part is $2 \cdot 3^2 = 18$. ✓

Running the divisions *backward* expresses the gcd as an integer combination of the originals—this is *Bézout's identity*: there exist integers x, y with

$$ax + by = \gcd(a, b). \tag{6.1}$$

Among its uses, it is exactly the tool that produces modular inverses.

Bézout by back-substitution. Write $\gcd(252, 198) = 18$ as $252x + 198y$.

Read.

Solve the second-to-last division for the gcd, then substitute earlier remainders one at a time.

Plan.

From $54 = 1 \cdot 36 + 18$ we get $18 = 54 - 36$. Replace 36 using $36 = 198 - 3 \cdot 54$:

$$\begin{aligned} 18 &= 54 - (198 - 3 \cdot 54) && \text{substitute } 36 = 198 - 3 \cdot 54 \\ &= 4 \cdot 54 - 198 && \text{collect the 54 terms.} \end{aligned}$$

Now replace 54 using $54 = 252 - 198$:

$$\begin{aligned} 18 &= 4(252 - 198) - 198 && \text{substitute } 54 = 252 - 198 \\ &= 4 \cdot 252 - 5 \cdot 198 && \text{collect the 198 terms.} \end{aligned}$$

Solve.

$252(4) + 198(-5) = 18$, so $x = 4,\ y = -5$.

Check.

Verify: $1008 - 990 = 18$. ✓ Bézout coefficients are not unique—adding $198/18 = 11$ to x and subtracting $252/18 = 14$ from y gives another valid pair.

lcm from the gcd. Find $\operatorname{lcm}(252, 198)$.

Read.

Use the identity $\operatorname{lcm}(a, b) = \frac{ab}{\gcd(a,b)}$ with the gcd already found.

Plan.

With gcd $= 18$,

$$\begin{aligned} \operatorname{lcm}(252, 198) &= \frac{252 \cdot 198}{18} && \text{use } \operatorname{lcm} = ab/\gcd \\ &= 252 \cdot \frac{198}{18} && \text{divide the 198 by 18} \\ &= 252 \cdot 11 = 2772 && \text{since } 198/18 = 11. \end{aligned}$$

Solve.

$\operatorname{lcm}(252, 198) = 2772$.

Check.

Sanity check: $2772/252 = 11$ and $2772/198 = 14$, both integers, and $\gcd(11, 14) = 1$ as it must be. ✓

gcd of three numbers. Compute $\gcd(48, 60, 72)$.

Read.

The gcd of three is gcd of the first two, then that against the third—the operation is associative.

Plan.

First $\gcd(48, 60)$: $60 = 1 \cdot 48 + 12$, $48 = 4 \cdot 12 + 0$, so $\gcd(48, 60) = 12$. Now $\gcd(12, 72)$: since $12 \mid 72$, this is 12.

Solve.

$\gcd(48, 60, 72) = 12$.

Check.

Check by factoring: $48 = 2^4 \cdot 3$, $60 = 2^2 \cdot 3 \cdot 5$, $72 = 2^3 \cdot 3^2$; the shared part is $2^2 \cdot 3 = 12$. ✓

A linear Diophantine equation. Find all integer solutions of $12x + 18y = 30$.

Read.

Such an equation is solvable iff gcd divides the right side; divide through, find one solution, then add the homogeneous family.

Plan.

$\gcd(12, 18) = 6$, and $6 \mid 30$, so solutions exist. Divide by 6: $2x + 3y = 5$. One solution by inspection is $x = 1, y = 1$ (since $2 + 3 = 5$). The general solution adds any multiple of the homogeneous family $2x + 3y = 0$, i.e. $(x, y) = (1, 1) + t(3, -2)$ for integer t.

Solve.

$x = 1 + 3t,\ y = 1 - 2t$, $t \in \mathbb{Z}$.

Check.

The step size $(3, -2)$ is $(b/g, -a/g) = (18/6, -12/6)$: exactly the Bézout freedom seen earlier. ✓

Misconception

Coprime is not the same as prime. Two integers are *coprime* (relatively prime) when gcd = 1; neither has to be prime. For example $\gcd(8, 9) = 1$ though both are composite. Many theorems (CRT, Euler) require coprimality, *not* primality—read the hypothesis carefully.

Modular Arithmetic and Congruences

Fix a positive integer n, the *modulus*. We write $a \equiv b \pmod{n}$ and say "a is congruent to b modulo n" when $n \mid (a - b)$—that is, a and b leave the same remainder on division by n. Congruence behaves like equality: you may add, subtract, and multiply congruences freely, because if $a \equiv b$ and

$c \equiv d \pmod{n}$ then $a + c \equiv b + d$ and $ac \equiv bd \pmod{n}$. This is what makes "clock arithmetic" a genuine algebra: work with remainders $\{0, 1, \ldots, n-1\}$, and reduce mod n whenever numbers grow. The set of these residue classes with $+$ and $\times$ is written $\mathbb{Z}_n$.

What & why. Congruence mod n collapses the integers into n residue classes on which addition and multiplication still make sense.

Solve. Reduce every intermediate result mod n; replace any number by its remainder.

Means. $a \equiv b \pmod{n}$ means n divides $a - b$.

Approach Read · Plan · Solve · Check

Read. Fix the modulus n and identify the task—reduce a value, tame a power, or find a multiplicative inverse.

Plan.

1. Replace every number by its residue mod n, reducing after each step so numbers stay small.
2. For a large power, reduce after each squaring or find the repeating cycle of a^k.
3. For an inverse, solve $ax \equiv 1 \pmod{n}$ by a Bézout combination—it exists iff $\gcd(a, n) = 1$.

Solve. Do the arithmetic mod n, one reduction per line.

Check. Multiply out to confirm a residue (e.g. $a \cdot a^{-1} \equiv 1$), and verify the answer lies in $\{0, \ldots, n-1\}$.

Common traps — every wrong answer is one of these

Divides a congruence like an equation.
From $ac \equiv bc \pmod{n}$ concludes $a \equiv b \pmod{n}$. That cancellation is valid only when $\gcd(c, n) = 1$; otherwise multiply by an inverse, or divide the modulus too.

Assumes every element is invertible.
Looks for $a^{-1} \pmod{n}$ for any a. The inverse exists *iff* $\gcd(a, n) = 1$; for instance 2 has no inverse mod 6.

Reduces the exponent mod n.
Simplifies $a^k \pmod{n}$ by taking k mod n. Exponents reduce mod the *order* of a (or mod $p - 1$ or $\varphi(n)$), never mod n itself.

Leaves a negative or oversized residue. 1. Computes an intermediate value like -7 or 645.

2. Reports it as the final residue.
3. Fix: reduce into $\{0, \ldots, n-1\}$—e.g. $-7 \equiv 43 \pmod{50}$.

Fast reduction of a power. Compute $2^{10} \pmod 7$ without a big multiplication.

Read.

Reduce after each squaring so numbers stay small.

Plan.

Since $2^3 = 8 \equiv 1 \pmod 7$,

$$\begin{aligned} 2^{10} &= 2^9 \cdot 2 && \text{split off one factor} \\ &= (2^3)^3 \cdot 2 && \text{regroup as a cube} \\ &\equiv 1^3 \cdot 2 && \text{replace } 2^3 \equiv 1 \pmod 7 \\ &= 2 \pmod 7 && \text{simplify.} \end{aligned}$$

Solve.

$2^{10} \equiv 2 \pmod 7$.

Check.

Reducing early ($8 \equiv 1$) turned a four-digit number, 1024, into a one-line computation; indeed $1024 = 146 \cdot 7 + 2$. ✓

A units-digit question in disguise. What is the last digit of 7^{100}?

Read.

"Last digit" means the value modulo 10; find the cycle of $7^k \pmod{10}$.

Plan.

Powers of 7 mod 10: $7^1 \equiv 7$, $7^2 \equiv 49 \equiv 9$, $7^3 \equiv 63 \equiv 3$, $7^4 \equiv 21 \equiv 1$, then the pattern $7, 9, 3, 1$ repeats with period 4. Since $100 = 4 \cdot 25$,

$$\begin{aligned} 7^{100} &= (7^4)^{25} && \text{group into blocks of four} \\ &\equiv 1^{25} = 1 \pmod{10} && \text{since } 7^4 \equiv 1. \end{aligned}$$

Solve.

The last digit is 1.

Check.

The exponent's remainder mod the *cycle length* (4) is what matters; $100 \equiv 0$ lands on the end of a cycle, giving 1. ✓

To *divide* modulo n we need a *multiplicative inverse*: an integer a^{-1} with $a\,a^{-1} \equiv 1 \pmod{n}$. It exists precisely when $\gcd(a, n) = 1$—and Bézout's identity (6.1) builds it, since $ax + ny = 1$ says $ax \equiv 1 \pmod{n}$, making x the inverse.

A modular inverse via Bézout. Find the inverse of 7 modulo 50.

Read.

Solve $7x \equiv 1 \pmod{50}$ by finding a Bézout combination of 7 and 50.

Plan.

Euclid: $50 = 7 \cdot 7 + 1$, so $1 = 50 - 7 \cdot 7$. Reading mod 50: $-7 \cdot 7 \equiv 1$, i.e. $7 \cdot (-7) \equiv 1 \pmod{50}$. Convert -7 to a residue: $-7 \equiv 43 \pmod{50}$.

Solve.

$7^{-1} \equiv 43 \pmod{50}$.

Check.

Check: $7 \cdot 43 = 301 = 6 \cdot 50 + 1 \equiv 1 \pmod{50}$. ✓ An inverse exists here because $\gcd(7, 50) = 1$.

Reducing a power by the element's order. Compute $2^{100} \pmod{11}$.

Read.

Find the order of 2 modulo 11 (it divides $p - 1 = 10$ by Fermat), then reduce the exponent by that order.

Plan.

Powers of 2 mod 11: $2, 4, 8, 5, 10, 9, 7, 3, 6, 1$—first reaching 1 at 2^{10}, so the order is 10. Thus

$$\begin{aligned} 2^{100} &= (2^{10})^{10} && \text{group by the order 10} \\ &\equiv 1^{10} = 1 \pmod{11} && \text{since } 2^{10} \equiv 1. \end{aligned}$$

Solve.

$2^{100} \equiv 1 \pmod{11}$.

Check.

Here Fermat alone suffices ($2^{10} \equiv 1$), but knowing the exact order lets you reduce any exponent mod 10. Since 2 has order $10 = |\mathbb{Z}_{11}^*|$, it is a primitive root. ✓

Problem-solving strategy — Big powers modulo n

(1) If the modulus is prime p and $p \nmid a$, reduce the exponent modulo $p-1$ (Fermat). (2) If $\gcd(a, n) = 1$ for composite n, reduce modulo $\varphi(n)$ (Euler). (3) Otherwise find the actual cyclic *order* of a by listing powers until you hit 1, and reduce the exponent modulo that order. Always reduce intermediate products mod n as you go.

Linear Congruences and the Chinese Remainder Theorem

A *linear congruence* $ax \equiv b \pmod{n}$ is the modular analogue of a linear equation. Let $g = \gcd(a, n)$. There is a solution *iff* $g \mid b$; when it exists there are exactly g solutions modulo n. The recipe: divide the whole congruence (including the modulus) by g, then multiply by the inverse of the reduced coefficient.

Approach Read · Plan · Solve · Check

Read. Identify the congruence $ax \equiv b \pmod{n}$ or the system of congruences, and note whether the moduli are pairwise coprime.

Plan.

1. For one congruence, set $g = \gcd(a, n)$; check $g \mid b$; divide the whole thing by g and multiply by the inverse of the reduced coefficient.
2. For a system, form $N = \prod n_i$ and each partial product $N_i = N/n_i$; find each inverse M_i of N_i mod n_i.
3. Assemble $x \equiv \sum a_i N_i M_i$ and reduce mod N.

Solve. Work one modulus at a time, reducing as you go.

Check. Substitute the answer back into *every* original congruence.

Common traps — every wrong answer is one of these

Uses CRT on non-coprime moduli.
Applies the uniqueness conclusion when the n_i share a factor. CRT requires *pairwise coprime* moduli; otherwise a solution may fail to exist or fail to be unique.

Reports one solution when there are g.
Solves $ax \equiv b \pmod{n}$ and stops at a single residue. When $g = \gcd(a, n) > 1$ and $g \mid b$, there are exactly g solutions mod n, spaced n/g apart.

Skips the solvability check.
1. Divides $ax \equiv b$ through by $g = \gcd(a, n)$ without looking at b.
2. Produces a "solution" when none exists.
3. Fix: first confirm $g \mid b$; if it fails, there is *no* solution.

Forgets to reduce the CRT sum.
Assembles the weighted sum $\sum a_i N_i M_i$ and leaves it oversized. Reduce it mod N to land the answer in $\{0, \ldots, N-1\}$.

A congruence with a unique solution. Solve $3x \equiv 4 \pmod{7}$.

Read.

Since $\gcd(3, 7) = 1$, multiply both sides by the inverse of 3 modulo 7.

Plan.

Find 3^{-1}: $3 \cdot 5 = 15 \equiv 1 \pmod{7}$, so $3^{-1} \equiv 5$. Multiply both sides by 5:

$$\begin{aligned} x &\equiv 5 \cdot 4 && \text{multiply by } 3^{-1} \equiv 5 \\ &= 20 \equiv 6 \pmod{7} && \text{reduce mod } 7. \end{aligned}$$

Solve.

$x \equiv 6 \pmod{7}$.

Check.

Check: $3 \cdot 6 = 18 = 2 \cdot 7 + 4 \equiv 4 \pmod{7}$. ✓

A congruence with several solutions. Solve $14x \equiv 30 \pmod{100}$.

Read.

Compute $g = \gcd(14, 100)$; check $g \mid 30$; divide through by g and solve the reduced congruence.

Plan.

Here $g = \gcd(14, 100) = 2$, and $2 \mid 30$, so there are 2 solutions. Divide by 2: $7x \equiv 15 \pmod{50}$. Using $7^{-1} \equiv 43 \pmod{50}$ from before,

$$\begin{aligned} x &\equiv 43 \cdot 15 && \text{multiply by } 7^{-1} \equiv 43 \\ &= 645 && \text{multiply out} \\ &\equiv 45 \pmod{50} && \text{reduce: } 645 - 12 \cdot 50. \end{aligned}$$

Lifting back to modulus 100, the two solutions are $x \equiv 45$ and $x \equiv 45 + 50 = 95 \pmod{100}$.

Solve.

$x \equiv 45$ or $95 \pmod{100}$.

Check.

Check $x = 45$: $14 \cdot 45 = 630 = 6 \cdot 100 + 30 \equiv 30$. ✓ The count of solutions equals $g = 2$, exactly as predicted.

When several congruences with *coprime* moduli are imposed at once, the *Chinese Remainder Theorem* (CRT) guarantees a unique combined solution modulo the product.

Chinese Remainder Theorem. If $n_1, n_2, \ldots, n_k$ are pairwise coprime, the system $x \equiv a_i \pmod{n_i}$ has a *unique* solution modulo $N = n_1 n_2 \cdots n_k$.

A CRT system. Solve $x \equiv 2 \pmod 3$, $x \equiv 3 \pmod 5$, $x \equiv 2 \pmod 7$.

Read.

Form N and the partial products N/n_i; find each inverse; assemble the weighted sum and reduce mod N.

Plan.

Here $N = 3 \cdot 5 \cdot 7 = 105$. Partial products: $N_1 = 35$, $N_2 = 21$, $N_3 = 15$. Inverses: $35 \equiv 2 \pmod 3$ and $2^{-1} \equiv 2$; $21 \equiv 1 \pmod 5$, inverse 1; $15 \equiv 1 \pmod 7$, inverse 1. Assemble the

weighted sum:

$$\begin{aligned} x &\equiv 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 && \text{form } \textstyle\sum a_i N_i M_i \\ &= 140 + 63 + 30 = 233 && \text{multiply each term} \\ &\equiv 23 \pmod{105} && \text{reduce: } 233 - 2 \cdot 105. \end{aligned}$$

Solve.

$x \equiv 23 \pmod{105}$.

Check.

Check all three: $23 = 7 \cdot 3 + 2$, $23 = 4 \cdot 5 + 3$, $23 = 3 \cdot 7 + 2$. ✓ One class mod 105, exactly as CRT promises.

A CRT word problem. A basket holds fewer than 400 eggs. Counted by twos, threes, fours, fives, or sixes, one egg is always left over; counted by sevens, none is left over. How many eggs?

Read.

Translate to congruences, collapse the "remainder 1" conditions into one modulus, then solve the two-condition CRT system.

Plan.

Leaving remainder 1 modulo $2, 3, 4, 5, 6$ is the single condition $x \equiv 1 \pmod{60}$, since $\operatorname{lcm}(2, 3, 4, 5, 6) = 60$. Also $x \equiv 0 \pmod 7$. Write $x = 60k + 1$ and impose $7 \mid x$: $60k + 1 \equiv 0 \pmod 7$. Since $60 \equiv 4 \pmod 7$, this is $4k \equiv -1 \equiv 6 \pmod 7$; multiplying by $4^{-1} \equiv 2$ gives $k \equiv 12 \equiv 5 \pmod 7$. Take $k = 5$: $x = 60 \cdot 5 + 1 = 301$.

Solve.

301 eggs.

Check.

Check: $301 = 7 \cdot 43$ (divisible by 7), and $301 - 1 = 300$ is divisible by $2, 3, 4, 5, 6$. The next solution, $301 + 420$, exceeds 400, so 301 is the only one. ✓

Connection

CRT is an isomorphism. That the system has a *unique* solution mod N is the statement that the map $\mathbb{Z}_N \to \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$ sending x to its tuple of remainders is a bijection preserving $+$ and $\times$—an *isomorphism of rings* $\mathbb{Z}_{15} \cong \mathbb{Z}_3 \times \mathbb{Z}_5$, in the language of Part B.

Number theory and abstract algebra are describing the same object.

Fermat's Little Theorem and Euler's Totient

Two theorems make large powers modulo n collapse to almost nothing. *Fermat's Little Theorem* says that if p is prime and $p \nmid a$, then

$$a^{p-1} \equiv 1 \pmod{p}. \tag{6.2}$$

Its generalization to a composite modulus uses *Euler's totient* $\varphi(n)$, the count of integers in $\{1, \dots, n\}$ coprime to n: if $\gcd(a, n) = 1$ then

$$a^{\varphi(n)} \equiv 1 \pmod{n}. \tag{6.3}$$

For a prime power,

$$\varphi(p^k) = p^k - p^{k-1}, \tag{6.4}$$

and φ is multiplicative on coprime factors, so φ is computed from the prime factorization.

Derivation. *Fermat's Little Theorem*

Let p be prime and $p \nmid a$. As x runs over the nonzero residues $1, 2, \dots, p-1$, the products $a \cdot 1, a \cdot 2, \dots, a \cdot (p-1)$ are again the nonzero residues in some order (if $ax \equiv ay$ then $x \equiv y$ because a is invertible mod p). Multiplying all of them, $a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$. Since $(p-1)!$ is coprime to p, cancel it: $a^{p-1} \equiv 1 \pmod{p}$.

Approach Read · Plan · Solve · Check

Read. Identify the base a, the modulus n, and the exponent; check $\gcd(a, n)$ to decide which theorem applies.

Plan.

1. If $n = p$ is prime and $p \nmid a$, reduce the exponent modulo $p - 1$ (Fermat).
2. If n is composite and $\gcd(a, n) = 1$, reduce the exponent modulo $\varphi(n)$ (Euler), computing φ from the prime factorization.
3. Finish the small leftover power by hand, reducing mod n.

Solve. Split the exponent as (multiple of the period) + remainder, and evaluate the remainder

power.

Check. Confirm the base is coprime to the modulus, and verify the small final power directly.

Common traps — every wrong answer is one of these

Uses Fermat when $p \mid a$.
Writes $a^{p-1} \equiv 1 \pmod{p}$ without checking $p \nmid a$. If $p \mid a$ then $a \equiv 0$ and every positive power is 0, not 1.

Applies Euler to a non-coprime base.
Claims $a^{\varphi(n)} \equiv 1 \pmod{n}$ when $\gcd(a, n) > 1$. Euler's theorem needs $\gcd(a, n) = 1$; otherwise the congruence can fail outright.

Reduces the exponent mod p.
Simplifies $a^k \pmod{p}$ by taking $k \bmod p$. Fermat reduces the exponent mod $p - 1$, not mod p.

Mishandles the totient formula. 1. Computes $\varphi(p^k)$ as p^{k-1} or $p^k - 1$.
2. Multiplies φ across factors that share a prime.
3. Fix: $\varphi(p^k) = p^k - p^{k-1}$, and φ multiplies only across *coprime* factors.

A huge power, tamed by Fermat. Compute $3^{100} \pmod{7}$.

Read.

Use $a^{p-1} \equiv 1$ with $p = 7$ to reduce the exponent modulo 6, then finish by hand.

Plan.

By Fermat (6.2), $3^6 \equiv 1 \pmod{7}$. Write $100 = 6 \cdot 16 + 4$, so

$$\begin{aligned} 3^{100} &= (3^6)^{16} \cdot 3^4 && \text{split the exponent } 100 = 6 \cdot 16 + 4 \\ &\equiv 1 \cdot 3^4 \pmod{7} && \text{since } 3^6 \equiv 1. \end{aligned}$$

Now $3^4 = 81 = 11 \cdot 7 + 4 \equiv 4$.

Solve.

$3^{100} \equiv 4 \pmod{7}$.

Check.

Only the exponent's remainder mod $p - 1 = 6$ matters. Since $100 \equiv 4 \pmod{6}$, the answer is 3^4. ✓

Euler's totient. Compute $\varphi(36)$, and use it to find $5^{12} \pmod{36}$.

Read.

Factor 36, apply the totient formula, then reduce the exponent by φ.

Plan.

Factor $36 = 2^2 \cdot 3^2$. Then

$$\begin{aligned} \varphi(36) &= \varphi(4)\,\varphi(9) && \varphi \text{ multiplies on coprime factors} \\ &= (4-2)(9-3) && \text{use } \varphi(p^k) = p^k - p^{k-1} \\ &= 2 \cdot 6 = 12 && \text{multiply.} \end{aligned}$$

Since $\gcd(5, 36) = 1$, Euler's theorem (6.3) gives $5^{12} \equiv 1 \pmod{36}$.

Solve.

$\varphi(36) = 12$ and $5^{12} \equiv 1 \pmod{36}$.

Check.

The formula (6.4) counts the residues *not* divisible by p; for 36 that is 12 of the 36. ✓

An inverse from Euler's theorem. Use Euler's theorem to find $3^{-1} \pmod{10}$.

Read.

Since $a^{\varphi(n)} \equiv 1$, the inverse of a is $a^{\varphi(n)-1}$; compute it mod n.

Plan.

Here $\varphi(10) = 4$, so $3^4 \equiv 1 \pmod{10}$, which means $3 \cdot 3^3 \equiv 1$. Thus

$$\begin{aligned} 3^{-1} &\equiv 3^3 && \text{since } 3 \cdot 3^3 \equiv 1 \\ &= 27 \equiv 7 \pmod{10} && \text{reduce mod 10.} \end{aligned}$$

Solve.

$3^{-1} \equiv 7 \pmod{10}$.

Check.

Check: $3 \cdot 7 = 21 \equiv 1 \pmod{10}$. ✓ Euler's theorem gives inverses without running the Euclidean algorithm—handy when $\varphi(n)$ is small.

— Part B: Abstract Algebra —

Groups: The Structure Behind Arithmetic

Strip arithmetic to its skeleton and you get a *group*: a set G with one operation $*$ satisfying four axioms—*closure* ($a*b \in G$), *associativity* ($(a*b)*c = a*(b*c)$), an *identity* e with $e*a = a*e = a$, and an *inverse* a^{-1} for each a with $a*a^{-1} = e$. If in addition $a*b = b*a$ always, the group is *abelian*. The *order* $|G|$ is the number of elements. Familiar examples: the integers under addition $(\mathbb{Z}, +)$; the residues $(\mathbb{Z}_n, +)$; the nonzero residues modulo a prime under multiplication $(\mathbb{Z}_p^*, \times)$; and the symmetries of a shape.

What & why. A group is a set with one associative operation that has an identity and inverses—the minimal setting in which you can "solve for x."

Solve. Check the four axioms; for a finite group a Cayley table makes closure and inverses visible.

Means. One well-behaved operation is all a group is.

Approach Read · Plan · Solve · Check

Read. Identify the set G and the operation $*$, and state what is asked—verify a group, find the identity or an inverse, or compute an element's order.

Plan.

1. Check the four axioms in turn: closure, associativity, an identity e, and an inverse for every element.
2. For a finite group, build a Cayley table—each row and column must be a permutation of G.
3. For the order of an element, use $n/\gcd(m,n)$ in $(\mathbb{Z}_n, +)$, or list powers until you reach the identity.

Solve. Verify each axiom explicitly, naming the identity and the inverses.

Check. The identity row/column reproduces the headers, and each row hits e exactly once.

Common traps — every wrong answer is one of these

Assumes the group is abelian.

Uses $a*b = b*a$ freely. Commutativity is an *extra* axiom; S_3 and matrix groups fail it, so never reorder a product unless the group is known to be abelian.

Checks closure and calls it a group.

Verifies $a*b \in G$ and stops. All four axioms matter equally—associativity, an identity, and an

inverse for *every* element must all hold.

Confuses element order with group order.
Reads $|G|$ as the order of an element, or vice versa. The order of a is the least $k > 0$ with $a^k = e$; it divides $|G|$ but rarely equals it.

Uses the wrong identity.
1. Hunts for an inverse "summing to 1" in an additive group, or "multiplying to 0."
2. Mislabels the identity element.
3. Fix: in $(\mathbb{Z}_n, +)$ the identity is 0 and inverses are complements to n; in $(\mathbb{Z}_p^*, \times)$ the identity is 1.

Verifying a group. Show $(\mathbb{Z}_5, +)$ is a group and state its identity and the inverse of 3.

Read.

Check the four axioms for addition mod 5 on $\{0, 1, 2, 3, 4\}$.

Plan.

Closure: a sum reduced mod 5 is again in $\{0, \ldots, 4\}$. *Associativity*: inherited from integer addition. *Identity*: 0, since $0 + a = a$. *Inverses*: for each a, the element $5 - a$ (and 0 for 0) gives $a + (5 - a) = 5 \equiv 0$. In particular the inverse of 3 is 2, since $3 + 2 = 5 \equiv 0$.

Solve.

$(\mathbb{Z}_5, +)$ is an abelian group of order 5; identity 0; $3^{-1} = 2$.

Check.

Additive inverses mod n are "complements to n"—easy to write down, unlike multiplicative inverses. ✓

Reading a Cayley table. The table below defines $*$ on $\{e, a, b, c\}$. Is it a group?

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Read.

Check that every row/column is a permutation of the set (closure + inverses), that e is an identity, and spot-check associativity.

Plan.

Every row and every column contains each of e, a, b, c exactly once, so the operation is closed and every element has an inverse (each row hits e once). The e row and column reproduce the headers, so e is the identity. Each element is its own inverse ($a * a = e$, etc.), and the table is symmetric, so $*$ is commutative; a few triple products confirm associativity.

Solve.

Yes—this is a group (the Klein four-group V_4), abelian of order 4.

Check.

Every non-identity element having order 2 is the fingerprint of V_4, distinguishing it from the cyclic group $\mathbb{Z}_4$. ✓

The *order of an element* a is the smallest positive integer k with $a^k = e$ (or $ka = e$ in additive notation). In $\mathbb{Z}_n$ the order of m is

$$\operatorname{ord}(m) = \frac{n}{\gcd(m, n)}. \tag{6.5}$$

Order of an element. Find the order of 3 in $(\mathbb{Z}_{12}, +)$ and of 3 in $(\mathbb{Z}_7^*, \times)$.

Read.

For addition use $n/\gcd(m, n)$; for multiplication list powers until you reach the identity 1.

Plan.

Additive: order of 3 in $\mathbb{Z}_{12}$ is $12/\gcd(3, 12) = 12/3 = 4$ (indeed $3, 6, 9, 12 \equiv 0$). Multiplicative: powers of 3 mod 7 are 3, $3^2 = 2$, $3^3 = 6$, $3^4 = 4$, $3^5 = 5$, $3^6 = 1$, first reaching 1 at the sixth power.

Solve.

Order 4 in $(\mathbb{Z}_{12}, +)$; order 6 in $(\mathbb{Z}_7^*, \times)$.

Check.

Because 3 has order $6 = |\mathbb{Z}_7^*|$, it *generates* the whole group—a primitive root mod 7. ✓

Order in a direct product. Find the order of $(1, 1)$ in $\mathbb{Z}_4 \times \mathbb{Z}_6$, and the group's order.

Read.

The order of a tuple is the lcm of the orders of its components; the group's order is the product of the factor sizes.

Plan.

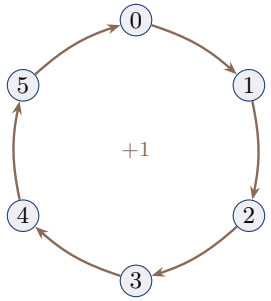
In $\mathbb{Z}_4$ the element 1 has order 4; in $\mathbb{Z}_6$ the element 1 has order 6. The order of $(1, 1)$ is $\operatorname{lcm}(4, 6) = 12$. The group has order $4 \cdot 6 = 24$.

Solve.

$(1, 1)$ has order 12; $|\mathbb{Z}_4 \times \mathbb{Z}_6| = 24$.

Check.

Since $12 < 24$, no single element runs through the whole group—$\mathbb{Z}_4 \times \mathbb{Z}_6$ is *not* cyclic. It would be, by CRT, only if the factor sizes were coprime. ✓



The cyclic group $(\mathbb{Z}_6, +)$ as a clock: repeatedly adding 1 steps around all six residues and returns to the start. The generator 1 has order 6.

Subgroups, Cyclic Groups, and Generators

A *subgroup* $H \leq G$ is a subset that is itself a group under the same operation—it contains e, is closed, and contains inverses. A group is *cyclic* if a single element g generates it: every element is a power (or multiple) g^k. The clock group $(\mathbb{Z}_n, +)$ is cyclic, generated by 1; its generators are exactly the residues coprime to n, so there are $\varphi(n)$ of them. Cyclic groups have the cleanest possible subgroup structure.

Subgroups of a cyclic group. A cyclic group of order n has *exactly one* subgroup of order d for each divisor $d \mid n$, and no others. In $(\mathbb{Z}_n, +)$ that subgroup is generated by n/d.

Approach Read · Plan · Solve · Check

Read. Identify the group and its order n, and whether you must list subgroups, find generators, or count elements of a given order.

Plan.

1. List the divisors of n; each divisor d gives exactly one subgroup of order d, generated by n/d.
2. The generators of $(\mathbb{Z}_n, +)$ are the residues coprime to n—there are $\varphi(n)$ of them.
3. An element m has order $n/\gcd(m, n)$; set that equal to a target and count the qualifying m.

Solve. Compute the divisors or gcds explicitly, then read off the subgroups, generators, or orders.

Check. The subgroup lattice mirrors the divisor lattice of n; the generator count matches $\varphi(n)$.

Common traps — every wrong answer is one of these

Calls any element a generator.
Assumes a fixed m (often a divisor of n) generates $\mathbb{Z}_n$. Only residues with $\gcd(m, n) = 1$ generate; there are exactly $\varphi(n)$ of them.

Miscounts elements of a given order.
Expects d elements of order d. A cyclic group of order d has $\varphi(d)$ generators, so there are $\varphi(d)$ elements of order exactly d.

Assumes a product of cyclic groups is cyclic.
Treats $\mathbb{Z}_m \times \mathbb{Z}_n$ as automatically cyclic. It is cyclic *iff* $\gcd(m, n) = 1$; for example $\mathbb{Z}_4 \times \mathbb{Z}_6$ is not.

Builds a subgroup for the wrong size.
1. Looks for a subgroup whose order does not divide $|G|$.
2. Or expects two distinct subgroups of the same order in a cyclic group.
3. Fix: a cyclic group of order n has *exactly one* subgroup per divisor $d \mid n$, and none of any other size.

Listing subgroups. List all subgroups of $(\mathbb{Z}_{12}, +)$.

Read.

Find the divisors of 12; for each divisor d take the subgroup of order d generated by $12/d$.

Plan.

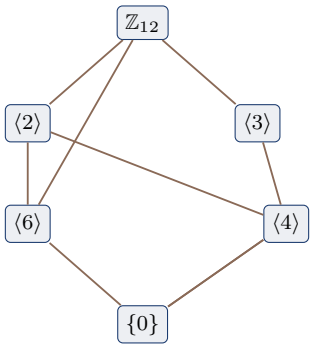
Divisors of 12: $1, 2, 3, 4, 6, 12$. The subgroup of order d is $\langle 12/d \rangle$: order $12 = \langle 1 \rangle$ (all of $\mathbb{Z}_{12}$); order $6 = \langle 2 \rangle = \{0, 2, 4, 6, 8, 10\}$; order $4 = \langle 3 \rangle = \{0, 3, 6, 9\}$; order $3 = \langle 4 \rangle = \{0, 4, 8\}$; order $2 = \langle 6 \rangle = \{0, 6\}$; order $1 = \langle 0 \rangle = \{0\}$.

Solve.

Six subgroups, one per divisor of 12.

Check.

The subgroup lattice mirrors the divisor lattice of 12 exactly—a hallmark of cyclic groups. ✓



The subgroup lattice of $\mathbb{Z}_{12}$: one subgroup for each divisor of 12, joined by containment. It is the divisor lattice of 12 redrawn.

Generators of a cyclic group. Which elements generate $(\mathbb{Z}_9, +)$?

Read.

An element m generates $\mathbb{Z}_n$ exactly when $\gcd(m, n) = 1$; list those m.

Plan.

Test $\gcd(m, 9) = 1$ for $m = 1, \ldots, 8$: coprime to 9 are $1, 2, 4, 5, 7, 8$ (the multiples of 3 are excluded). There are $\varphi(9) = 6$ of them.

Solve.

The generators are $1, 2, 4, 5, 7, 8$.

Check.

For instance 2 generates: $2, 4, 6, 8, 1, 3, 5, 7, 0$ cycles through everything. An element like 3 fails—$\langle 3 \rangle = \{0, 3, 6\}$ is only a subgroup of order 3. ✓

Elements of a given order. How many elements of order exactly 3 does $(\mathbb{Z}_9, +)$ contain?

Read.

An element k has order $9/\gcd(k, 9)$ by (6.5); set that equal to 3 and count the qualifying k.

Plan.

Order 3 means $9/\gcd(k, 9) = 3$, i.e. $\gcd(k, 9) = 3$. The residues $k \in \{1, \ldots, 8\}$ with $\gcd(k, 9) = 3$ are $k = 3$ and $k = 6$.

Solve.

Two elements: 3 and 6.

Check.

They generate the unique order-3 subgroup $\{0, 3, 6\}$; a cyclic group of order 3 has $\varphi(3) = 2$ generators, matching the count. ✓

Derivation. *why a cyclic group of order n has exactly $\varphi(n)$ generators*

Let g generate a cyclic group of order n, so every element is g^k. The element g^k has order $n/\gcd(k, n)$, which equals n—making g^k itself a generator—precisely when $\gcd(k, n) = 1$. The number of such k in $\{1, \ldots, n\}$ is by definition $\varphi(n)$, so a cyclic group of order n has exactly $\varphi(n)$ generators.

Counting generators. How many generators does $(\mathbb{Z}_{12}, +)$ have, and what are they?

Read.

The generators are the residues coprime to 12; count them with $\varphi(12)$ and list them.

Plan.

Compute the totient:

$$\begin{aligned} \varphi(12) &= 12\left(1 - \tfrac{1}{2}\right)\left(1 - \tfrac{1}{3}\right) && \text{use the prime factors } 2, 3 \\ &= 12 \cdot \tfrac{1}{2} \cdot \tfrac{2}{3} = 4 && \text{simplify.} \end{aligned}$$

The residues coprime to 12 are $1, 5, 7, 11$.

Solve.

Four generators: $1, 5, 7, 11$.

Check.

Each of these, added to itself repeatedly, cycles through all of $\mathbb{Z}_{12}$; a residue like 8, with $\gcd(8, 12) = 4$, only reaches the order-3 subgroup $\{0, 4, 8\}$. ✓

A primitive-root test. Is 2 a generator of the multiplicative group $(\mathbb{Z}_7^*, \times)$?

Read.

List powers of 2 mod 7 until reaching 1; if the order is $6 = |\mathbb{Z}_7^*|$, it generates; otherwise it does not.

Plan.

Powers of 2 mod 7: $2^1 = 2,\ 2^2 = 4,\ 2^3 = 8 \equiv 1$. The order of 2 is 3, not 6.

Solve.

No—2 has order 3, so $\langle 2 \rangle = \{1, 2, 4\}$ is only a subgroup of order 3.

Check.

Contrast 3, whose powers cycled through all six nonzero residues: 3 *is* a primitive root mod 7 but 2 is not. By Lagrange (6.6) the order must divide 6, and here it lands on 3. ✓

Lagrange's Theorem and Cosets

The most quoted counting fact in group theory is *Lagrange's theorem*: in a finite group, the order of any subgroup divides the order of the group. The mechanism is *cosets*. For $H \leq G$ and $g \in G$, the left coset $gH = \{gh : h \in H\}$ is a "shifted copy" of H; the cosets partition G into equal-sized, non-overlapping blocks, each the same size as H. Counting the blocks gives

$$|G| = [G : H]\,|H|, \tag{6.6}$$

where the *index* $[G : H]$ is the number of cosets.

Derivation. *Lagrange's theorem*

Distinct left cosets of H are disjoint: if g_1H and g_2H share an element then $g_1h_1 = g_2h_2$, whence $g_2^{-1}g_1 = h_2h_1^{-1} \in H$ and the two cosets coincide. Every element g lies in the coset gH

(take $h = e$), so the cosets *cover* G. Finally each coset gH has exactly $|H|$ elements, because $h \mapsto gh$ is a bijection $H \to gH$. Thus G splits into $[G : H]$ disjoint blocks of size $|H|$, giving $|G| = [G : H]\,|H|$; in particular $|H|$ divides $|G|$.

Approach Read · Plan · Solve · Check

Read. Identify the group G, the subgroup H, and the target—list cosets, find the index, or constrain possible orders.

Plan.

1. Form each coset gH by combining g with every element of H; collect the *distinct* blocks.
2. Count the blocks: that number is the index $[G : H]$, and $|G| = [G : H]\,|H|$.
3. To constrain structure, note every subgroup and element order must divide $|G|$.

Solve. Generate cosets one representative at a time, discarding repeats.

Check. The blocks are disjoint, each of size $|H|$, and together they cover G.

Common traps — every wrong answer is one of these

Believes Lagrange's converse.
Assumes $d \mid |G|$ guarantees a subgroup of order d. False in general—A_4 has order 12 but no subgroup of order 6; "divides" is necessary, not sufficient.

Thinks cosets can partly overlap.
Treats two cosets as sharing some but not all elements. Distinct cosets are *disjoint*: sharing one element forces them to coincide.

Forgets cosets have equal size.
Counts blocks of varying sizes. Every coset gH has exactly $|H|$ elements, which is precisely why $|G| = [G : H]\,|H|$.

Ignores the order constraint on elements.
1. Posits an element of order d with $d \nmid |G|$.
2. Reasons about the group as if it existed.
3. Fix: an element generates a cyclic subgroup, so its order must divide $|G|$—no order-5 element in a group of order 12.

Cosets partitioning a group. List the cosets of $H = \{0, 3\}$ in $(\mathbb{Z}_6, +)$ and confirm Lagrange.

Read.

Add each group element to H; collect the distinct blocks; count them against $|G|/|H|$.

Plan.

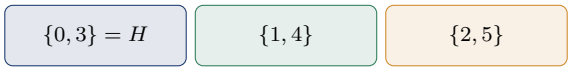
$H = \{0, 3\}$. Then $1 + H = \{1, 4\}$, $2 + H = \{2, 5\}$, and $3 + H = \{3, 0\} = H$ again. The distinct cosets are $\{0, 3\}, \{1, 4\}, \{2, 5\}$: three blocks of size 2 covering all of $\mathbb{Z}_6$.

Solve.

Three cosets; $[G : H] = 3$, and $|G| = 6 = 3 \cdot 2 = [G : H]\,|H|$. ✓

Check.

The blocks are disjoint and equal-sized, exactly as the derivation requires; here $|H| = 2$ divides $|G| = 6$.



three disjoint cosets tile $\mathbb{Z}_6$

Cosets of $H = \{0, 3\}$ in $\mathbb{Z}_6$ partition the group into three equal blocks of size $|H| = 2$. Counting blocks gives $|\mathbb{Z}_6| = 3 \cdot 2$—Lagrange's theorem made visible.

Using Lagrange to constrain structure. A group G has order 12. What orders are possible for its subgroups and its elements?

Read.

Subgroup orders must divide $|G|$; an element's order equals the order of the cyclic subgroup it generates, so it too must divide $|G|$.

Plan.

The divisors of 12 are $1, 2, 3, 4, 6, 12$, so those are the only possible subgroup orders. Each element a generates a cyclic subgroup $\langle a \rangle$ whose order is the order of a; that order must therefore also divide 12.

Solve.

Subgroup and element orders lie in $\{1, 2, 3, 4, 6, 12\}$; e.g. no element can have order 5.

Check.

Lagrange forbids many things instantly—an order-5 element would generate an order-5 subgroup, impossible in a group of order 12. ✓

Cosets in a multiplicative group. List the cosets of $H = \{1, 4\}$ in $G = (\mathbb{Z}_5^*, \times) = \{1, 2, 3, 4\}$ and read off the index.

Read.

Confirm H is a subgroup, then multiply each element of G into H, collecting the distinct blocks.

Plan.

First $H = \{1, 4\}$ is closed: $4 \cdot 4 = 16 \equiv 1 \pmod 5$, so $|H| = 2$. Form the cosets *multiplicatively*, reducing mod 5:

$$\begin{aligned} 1H &= \{1 \cdot 1,\ 1 \cdot 4\} = \{1, 4\} = H && \text{the subgroup itself} \\ 2H &= \{2 \cdot 1,\ 2 \cdot 4\} = \{2,\ 8 \bmod 5\} = \{2, 3\} && \text{a shifted copy} \\ 3H &= \{3 \cdot 1,\ 3 \cdot 4\} = \{3,\ 12 \bmod 5\} = \{3, 2\} = 2H && \text{same block as } 2H \end{aligned}$$

The distinct cosets are $\{1, 4\}$ and $\{2, 3\}$: two blocks of size 2 covering all of G.

Solve.

Two cosets, so the index is $[G : H] = 2$, and $|G| = 4 = 2 \cdot 2 = [G : H]\,|H|$. ✓

Check.

Each coset has exactly $|H| = 2$ elements and the two are disjoint—switching from an additive to a multiplicative operation changes the arithmetic but not Lagrange's counting.

Misconception

Lagrange does not have a full converse. That $d \mid |G|$ does *not* guarantee a subgroup of order d in every group (the alternating group A_4, of order 12, has no subgroup of order 6). For *cyclic* groups the converse does hold. Use "divides" as a necessary condition, not a sufficient one.

Homomorphisms and Isomorphisms

A *homomorphism* is a map $\varphi : G \to H$ that respects the operations:

$$\varphi(a * b) = \varphi(a) \cdot \varphi(b). \tag{6.7}$$

It automatically sends identity to identity and inverses to inverses. Its *kernel* $\ker\varphi = \{g : \varphi(g) = e_H\}$ is a subgroup of G measuring how far φ is from one-to-one, and its *image* is a subgroup of H. A bijective homomorphism is an *isomorphism*; isomorphic groups are "the same group in different clothing"—identical in every structural respect, differing only in the names of elements.

Approach Read · Plan · Solve · Check

Read. Identify the map φ and the two groups, and what is asked—test the homomorphism property, find kernel and image, or decide an isomorphism.

Plan.

1. Test $\varphi(a * b) = \varphi(a) \cdot \varphi(b)$ for general a, b.
2. The kernel is $\{g : \varphi(g) = e_H\}$; the image is the set of outputs—both subgroups.
3. To prove an isomorphism, exhibit a bijection preserving the operation; to disprove one, find a structural invariant that differs.

Solve. Verify the defining identity symbolically, then list the elements of the kernel and image.

Check. Confirm $|\ker\varphi| \cdot |\operatorname{im}\varphi| = |G|$, the First Isomorphism Theorem's bookkeeping.

Common traps — every wrong answer is one of these

Proves iso from equal size alone.
Concludes $G \cong H$ because $|G| = |H|$. Same order is necessary, not sufficient—$\mathbb{Z}_4 \ncong \mathbb{Z}_2 \times \mathbb{Z}_2$ though both have order 4.

Calls any bijection an isomorphism.
Matches elements one-to-one and declares an isomorphism. The bijection must also *preserve the operation*: $\varphi(a * b) = \varphi(a) \cdot \varphi(b)$.

Mis-locates the kernel.
Takes $\ker\varphi$ to be the elements fixed by φ, or the image. The kernel is what maps to the *identity* of the target, $\{g : \varphi(g) = e_H\}$.

Disproves iso the hard way.
1. Tries every possible bijection to rule out an isomorphism.
2. Gets lost in the cases.
3. Fix: exhibit one invariant that differs—the multiset of element orders, or abelian vs. non-abelian—since isomorphisms preserve all of them.

Testing a homomorphism. Is $\varphi : \mathbb{Z}_6 \to \mathbb{Z}_6$, $\varphi(x) = 2x$, a homomorphism? Find its kernel and image.

Read.

Check $\varphi(x+y) = \varphi(x) + \varphi(y)$; then list elements mapping to 0 and the set of outputs.

Plan.

Check the defining identity:

$$\begin{aligned} \varphi(x+y) &= 2(x+y) && \text{apply } \varphi \\ &= 2x + 2y = \varphi(x) + \varphi(y) && \text{distribute,} \end{aligned}$$

so yes. Kernel: $2x \equiv 0 \pmod 6$ means $x \in \{0, 3\}$. Image: $\{2x : x \in \mathbb{Z}_6\} = \{0, 2, 4\}$.

Solve.

A homomorphism with $\ker \varphi = \{0, 3\}$ and image $\{0, 2, 4\}$.

Check.

Note $|\ker| \cdot |\mathrm{im}| = 2 \cdot 3 = 6 = |\mathbb{Z}_6|$—the First Isomorphism Theorem's bookkeeping, and $\mathbb{Z}_6/\ker \varphi \cong \mathrm{im}\, \varphi$. ✓

An isomorphism from CRT. Show $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$.

Read.

Exhibit a generator of the product group; if one element has order 6, the group is cyclic of order 6, hence $\cong \mathbb{Z}_6$.

Plan.

The product has order $2 \cdot 3 = 6$. Consider $(1, 1)$: its order is the least k with $k(1, 1) = (0, 0)$, i.e. $k \equiv 0 \pmod 2$ and $k \equiv 0 \pmod 3$, so $k = \mathrm{lcm}(2, 3) = 6$. Thus $(1, 1)$ generates all six elements, so $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic of order 6.

Solve.

$\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$, via $(1, 1) \mapsto 1$.

Check.

This is the CRT isomorphism $\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$ read from the algebra side—the same fact, restated. ✓

Distinguishing two groups. Show $\mathbb{Z}_4 \ncong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Read.

Find a structural invariant—orders of elements—that the two groups do not share.

Plan.

$\mathbb{Z}_4$ contains an element of order 4 (namely 1: $1, 2, 3, 0$). In $\mathbb{Z}_2 \times \mathbb{Z}_2$ every non-identity element satisfies $2(a, b) = (0, 0)$, so all have order 2; there is no element of order 4. An isomorphism preserves orders, so none can exist.

Solve.

$\mathbb{Z}_4 \ncong \mathbb{Z}_2 \times \mathbb{Z}_2$: same size, different structure.

Check.

Two groups of the same order need not be isomorphic; the multiset of element orders is a cheap invariant that often separates them. ✓

The reduction homomorphism. Show that reduction mod 4, $\varphi : \mathbb{Z} \to \mathbb{Z}_4$ with $\varphi(m) = m \bmod 4$, is a group homomorphism, and identify its kernel.

Read.

Verify φ respects addition (6.7), then find which integers map to 0.

Plan.

For integers m, k,

$$\begin{aligned} \varphi(m+k) &= (m+k) \bmod 4 && \text{apply } \varphi \\ &= (m \bmod 4) + (k \bmod 4) && \text{congruence respects sums} \\ &= \varphi(m) + \varphi(k) && \text{in } \mathbb{Z}_4. \end{aligned}$$

The kernel is $\{m : m \equiv 0 \pmod 4\} = 4\mathbb{Z}$, the multiples of 4.

Solve.

φ is a homomorphism with $\ker \varphi = 4\mathbb{Z}$.

Check.

The quotient $\mathbb{Z}/4\mathbb{Z}$ is exactly $\mathbb{Z}_4$—the First Isomorphism Theorem in its most basic form, and the reason "$\mathbb{Z}_n$" and "$\mathbb{Z}/n\mathbb{Z}$" name the same object. ✓

A non-abelian example: S_3. In the symmetric group S_3, compute $(1\,2\,3)(1\,2)$ and confirm the group is non-abelian.

Read.

Compose permutations right-to-left; then compute the other order and compare.

Plan.

Apply $(1\,2)$ first, then $(1\,2\,3)$. Track each point: $1 \xrightarrow{(1\,2)} 2 \xrightarrow{(1\,2\,3)} 3$; $2 \to 1 \to 2$; $3 \to 3 \to 1$. So $(1\,2\,3)(1\,2) = (1\,3)$. Reversing, $(1\,2)(1\,2\,3)$: $1 \to 2 \to 1$; $2 \to 3 \to 3$; $3 \to 1 \to 2$, giving $(2\,3)$. Since $(1\,3) \neq (2\,3)$, the operation does not commute.

Solve.

$(1\,2\,3)(1\,2) = (1\,3)$, and S_3 is non-abelian.

Check.

S_3 (order 6) is the smallest non-abelian group—order alone doesn't determine a group, since $\mathbb{Z}_6$ is abelian of the same order. ✓

Rings and Fields

Add a second operation and you get a *ring*: a set with $+$ making it an abelian group, a multiplication $\times$ that is associative and *distributes* over addition ($a(b+c) = ab + ac$). The integers $\mathbb{Z}$, the residues $\mathbb{Z}_n$, and polynomial rings $\mathbb{R}[x]$ are the standard examples. A ring is an *integral domain* if it is commutative with a multiplicative identity $1 \neq 0$ and has *no zero divisors*—no $ab = 0$ with a, b both nonzero. It is a *field* if, in addition, every nonzero element has a multiplicative inverse: you can divide. Fields are where linear algebra and most of "ordinary" algebra live; $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields, and so is $\mathbb{Z}_p$ for prime p.

Approach Read · Plan · Solve · Check

Read. Identify the set and its two operations, and decide which label is at stake—ring, integral domain, or field.

Plan.

1. For a domain, hunt for zero divisors: a product $ab = 0$ with a, b both nonzero rules it out.
2. For a field, ask whether *every* nonzero element has a multiplicative inverse.
3. For $\mathbb{Z}_n$: it is a field iff n is prime, and its units are the residues coprime to n.

Solve. Test the decisive property directly—exhibit a zero divisor, or produce every inverse.

Check. A single zero-divisor pair settles "not a domain"; a full inverse table confirms a field.

Common traps — every wrong answer is one of these

Conflates integral domain with field.
Treats "no zero divisors" as "every element invertible." Every field is a domain, but $\mathbb{Z}$ and $\mathbb{Z}[x]$ are domains that are *not* fields—no inverse for 2 or for x.

Assumes $\mathbb{Z}_n$ is always a field.
Divides freely in $\mathbb{Z}_n$ for any n. $\mathbb{Z}_n$ is a field *iff* n is prime; composite n has zero divisors, e.g. $2 \cdot 3 \equiv 0 \pmod 6$.

Expects "at most two roots" everywhere.
Assumes a quadratic has at most two roots mod n. That guarantee holds only over a *field*; $x^2 \equiv 1 \pmod 8$ has four roots.

Skips a required field axiom. 1. Checks only that nonzero elements have inverses.
2. Declares a field.
3. Fix: confirm the whole package—commutative ring, identity $1 \neq 0$, distributivity—*and* every nonzero element invertible.

An infinite integral domain. Is the polynomial ring $\mathbb{Z}[x]$ an integral domain? Is it a field?

Read.

Check for zero divisors using degrees; then ask whether a typical element has a multiplicative inverse inside the ring.

Plan.

If $f, g \in \mathbb{Z}[x]$ are nonzero, then $\deg(fg) = \deg f + \deg g$ and the leading coefficient of fg is the product of the leading coefficients—nonzero, since $\mathbb{Z}$ has no zero divisors. So $fg \neq 0$: no zero divisors, and $\mathbb{Z}[x]$ is an integral domain. But it is *not* a field: the element x has no inverse, because $x \cdot g(x) = 1$ is impossible on degrees (deg would have to be 0 on the left but is at least 1).

Solve.

$\mathbb{Z}[x]$ is an integral domain but not a field.

Check.

Integral domains sit strictly between rings and fields: every field is a domain, but $\mathbb{Z}$ and

$\mathbb{Z}[x]$ are domains that are not fields. ✓

$\mathbb{Z}_n$ is a field iff n is prime. If $n = p$ is prime, every nonzero residue is coprime to p and so has an inverse (Bézout), making $\mathbb{Z}_p$ a field. If n is composite, say $n = ab$ with $1 < a, b < n$, then $a \cdot b \equiv 0$ exhibits zero divisors, so $\mathbb{Z}_n$ is not even an integral domain.

Derivation. *why $\mathbb{Z}_p$ is a field*

Take any nonzero residue a with $1 \le a \le p-1$. Because p is prime and $p \nmid a$, we have $\gcd(a, p) = 1$, so Bézout gives integers x, y with $ax + py = 1$. Reducing mod p, $ax \equiv 1 \pmod{p}$, so x is a multiplicative inverse of a. Every nonzero element is invertible, which is exactly the field axiom beyond being a commutative ring with 1.

Zero divisors block a field. Is $\mathbb{Z}_6$ an integral domain or a field?

Read.

Search for two nonzero residues whose product is 0 mod 6; a single pair rules out both.

Plan.

Try 2 and 3: $2 \cdot 3 = 6 \equiv 0 \pmod{6}$, yet neither factor is 0. So $\mathbb{Z}_6$ has zero divisors.

Solve.

$\mathbb{Z}_6$ is neither an integral domain nor a field.

Check.

Consistent with the rule: 6 is composite. The elements 2 and 3 have no inverses mod 6, precisely because $\gcd(2, 6)$ and $\gcd(3, 6)$ exceed 1. ✓

A finite field. Show $\mathbb{Z}_5$ is a field by exhibiting every inverse.

Read.

For each nonzero residue, find the partner whose product is 1 mod 5.

Plan.

$1 \cdot 1 = 1$; $2 \cdot 3 = 6 \equiv 1$; $3 \cdot 2 \equiv 1$; $4 \cdot 4 = 16 \equiv 1$. Every nonzero element has an inverse: $1^{-1} = 1,\ 2^{-1} = 3,\ 3^{-1} = 2,\ 4^{-1} = 4$.

Solve.

$\mathbb{Z}_5$ is a field of 5 elements.

Check.

Being prime, 5 guarantees this; the multiplicative group $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$ is itself a group of order 4 (in fact cyclic, generated by 2). ✓

Units of a ring. Which elements of $\mathbb{Z}_{10}$ have multiplicative inverses, and how many are there?

Read.

An element a is a unit iff $\gcd(a, 10) = 1$; count them with $\varphi(10)$.

Plan.

Coprime to 10 in $\{1, \ldots, 9\}$ are $1, 3, 7, 9$—the residues sharing no factor of 2 or 5. Their count is $\varphi(10) = 10(1 - \frac{1}{2})(1 - \frac{1}{5}) = 4$.

Solve.

The units are $\{1, 3, 7, 9\}$, four in all.

Check.

These four form the group $\mathbb{Z}_{10}^*$ under multiplication; the non-units $2, 4, 5, 6, 8$ are zero divisors, which is why $\mathbb{Z}_{10}$ is not a field. ✓

Dividing in a field. Solve $4x \equiv 5 \pmod 7$ by dividing.

Read.

In the field $\mathbb{Z}_7$ every nonzero element is invertible, so multiply both sides by 4^{-1}.

Plan.

The inverse of 4 mod 7 is 2, since $4 \cdot 2 = 8 \equiv 1$. Multiply both sides by 2:

$$\begin{aligned} x &\equiv 2 \cdot 5 && \text{multiply by } 4^{-1} \equiv 2 \\ &= 10 \equiv 3 \pmod 7 && \text{reduce mod 7.} \end{aligned}$$

Solve.

$x \equiv 3 \pmod 7$.

Check.

Check: $4 \cdot 3 = 12 \equiv 5 \pmod 7$. ✓ Because $\mathbb{Z}_7$ is a field, this linear equation has exactly one solution—unlike the same equation mod a composite, where several or none may appear.

Where a field would forbid extra roots. How many solutions does $x^2 \equiv 1 \pmod 8$ have?

Read.

Test each residue 0–7; over a field a quadratic has at most two roots, so an excess signals that $\mathbb{Z}_8$ is not a field.

Plan.

Square each residue mod 8: $0^2{=}0,\ 1^2{=}1,\ 2^2{=}4,\ 3^2{=}1,\ 4^2{=}0,\ 5^2{=}1,\ 6^2{=}4,\ 7^2{=}1$. The value 1 occurs at $x = 1, 3, 5, 7$.

Solve.

Four solutions: $x \equiv 1, 3, 5, 7 \pmod 8$.

Check.

A degree-2 equation with *four* roots is impossible over a field; $\mathbb{Z}_8$ (with the zero divisor $2 \cdot 4 \equiv 0$) is not one, so the "at most two roots" guarantee does not apply. ✓

Did You Know? *fields exist with 4, 8, 9, … elements—but never 6*

A finite field has exactly p^k elements for a prime p and integer $k \geq 1$—and one exists for every such size. So there are fields of order $4, 8, 9, 25, 27, \ldots$, but *none* of order 6 or 10. The order-4 field is not $\mathbb{Z}_4$ (which has a zero divisor); it is built from polynomials over $\mathbb{Z}_2$ modulo an irreducible quadratic, a construction that generalizes $\mathbb{Z}_p$ to every prime power.

Synthesis — Key Results You Must Own

1. $d \mid a$ means $a = dk$; a common divisor of a, b divides every combination $ma + nb$. The FTA gives unique prime factorization, and the divisor count is $\prod(a_i + 1)$.
2. The Euclidean algorithm computes $\gcd(a, b)$ by repeated remainders; back-substitution yields Bézout's $ax + by = \gcd(a, b)$, and $\text{lcm} = ab/\gcd$.
3. $a \equiv b \pmod n$ means $n \mid (a - b)$; a has an inverse mod n iff $\gcd(a, n) = 1$. $ax \equiv b \pmod n$ is solvable iff $\gcd(a, n) \mid b$, with that many solutions.
4. CRT: coprime moduli give a unique solution mod the product—an isomorphism $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$.
5. Fermat: $a^{p-1} \equiv 1 \pmod p$ for $p \nmid a$. Euler: $a^{\varphi(n)} \equiv 1 \pmod n$ for $\gcd(a, n) = 1$. Reduce

big exponents mod $p-1$ or $\varphi(n)$.

6. A group has one associative operation with identity and inverses; the order of m in $\mathbb{Z}_n$ is $n/\gcd(m,n)$; a cyclic group of order n has one subgroup per divisor of n.
7. Lagrange: $|H|$ divides $|G|$ (cosets partition G); the converse can fail. Homomorphisms preserve the operation; isomorphic groups are structurally identical.
8. $\mathbb{Z}_n$ is a field iff n is prime; composite n has zero divisors. Units of $\mathbb{Z}_n$ are the residues coprime to n, and there are $\varphi(n)$ of them.

Practice Problems

Difficulty: $\star$ warm-up, $\star\star$ standard, $\star\star\star$ challenge. Work each with the Read–Plan–Solve–Check habit modeled above.

Divisibility, primes, factorization

1 $\star$**.**

Give the prime factorization of 600 and state how many positive divisors it has.

2 $\star$**.**

Find $\gcd(84, 120)$ by prime factorization.

3 $\star\star$**.**

Prove that if $d \mid a$ and $d \mid b$ then $d \mid (7a - 4b)$.

4 $\star\star$**.**

For how many positive integers n is n^2 a divisor of $2^6 \cdot 3^4$?

Hint: $n^2 \mid N$ forces each prime's exponent in n to be at most half its exponent in N.

5 $\star\star\star$**.**

Show that $\gcd(a,b) \cdot \operatorname{lcm}(a,b) = ab$ using prime factorizations.

Euclidean algorithm and Bézout

6 ⋆**.**

Use the Euclidean algorithm to compute $\gcd(1071, 462)$.

7 ⋆⋆**.**

Find integers x, y with $1071x + 462y = \gcd(1071, 462)$.

8 ⋆⋆**.**

Compute $\operatorname{lcm}(1071, 462)$.

9 ⋆ ⋆ ⋆**.**

Prove that $\gcd(a, a + 1) = 1$ for every integer a.

Modular arithmetic and inverses

10 ⋆**.**

Compute $3^{20} \pmod 5$.

11 ⋆**.**

Find the last digit of 3^{2024}.

12 ⋆⋆**.**

Find the inverse of 9 modulo 31.

13 ⋆⋆**.**

Solve $5x \equiv 3 \pmod{11}$.

14 ⋆ ⋆ ⋆**.**

Solve $12x \equiv 9 \pmod{15}$ and state how many solutions there are modulo 15.

CRT, Fermat, Euler

15 ⋆⋆.

Solve $x \equiv 1 \pmod 4,\ x \equiv 2 \pmod 5$.

16 ⋆⋆.

Find the smallest positive x with $x \equiv 2 \pmod 3,\ x \equiv 3 \pmod 4,\ x \equiv 1 \pmod 5$.

17 ⋆⋆.

Compute $2^{50} \pmod{13}$ using Fermat's Little Theorem.

18 ⋆⋆.

Compute $\varphi(100)$.

19 ⋆⋆⋆.

Compute $7^{1000} \pmod{100}$.

Hint: $\varphi(100) = 40$ and $\gcd(7, 100) = 1$.

Groups, subgroups, orders

20 ⋆.

Find the order of 4 in $(\mathbb{Z}_{10}, +)$.

21 ⋆⋆.

List all generators of $(\mathbb{Z}_{10}, +)$.

22 ⋆⋆.

List all subgroups of $(\mathbb{Z}_{18}, +)$.

23 ⋆⋆.

Find the order of 2 in the multiplicative group $(\mathbb{Z}_{11}^*, \times)$.

24 ⋆⋆⋆.

Prove that a group of prime order p has no proper nontrivial subgroups, and is cyclic.

Lagrange, homomorphisms, isomorphism

25 ⋆⋆.

List the cosets of $H = \{0, 4, 8\}$ in $(\mathbb{Z}_{12}, +)$ and state the index $[\mathbb{Z}_{12} : H]$.

26 ⋆⋆.

Is $\varphi : \mathbb{Z}_8 \to \mathbb{Z}_8$, $\varphi(x) = 3x$, a homomorphism? Find its kernel and image.

27 ⋆⋆⋆.

Decide whether $\mathbb{Z}_4 \times \mathbb{Z}_3 \cong \mathbb{Z}_{12}$, with justification.

Rings and fields

28 ⋆.

List the units of $\mathbb{Z}_9$.

29 ⋆⋆.

Find all zero divisors of $\mathbb{Z}_{12}$.

30 ⋆⋆⋆.

For which n in $\{7, 8, 9, 11\}$ is $\mathbb{Z}_n$ a field? Justify each answer.

Further Reading

The number-theory half of this chapter is developed carefully in any elementary number-theory text; the algebra half opens a first course in abstract algebra, where groups, rings, and fields are treated at length and the isomorphism theorems hinted at here are proved in full. Chapter 7 gathers the remaining Subject-Test topics—point-set topology, complex analysis, and numerical methods—that round out the exam.

CHAPTER 7

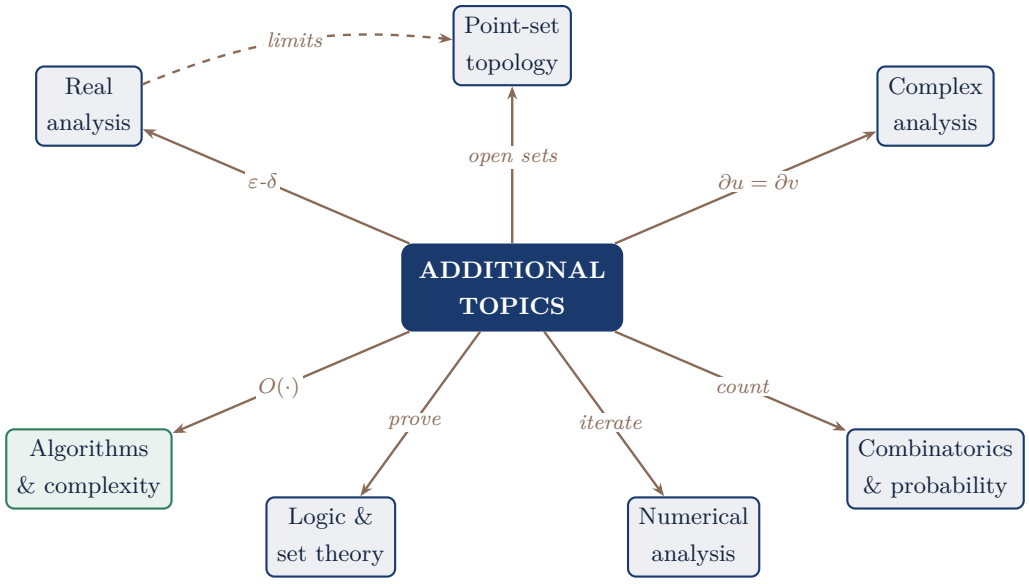
Additional Topics

Analysis, Topology, Complex Variables, and More

"We must know. We will know."

— David Hilbert

Roughly a quarter of the Subject Test lives outside the calculus–algebra spine, in a scattering of fields each represented by a handful of characteristic questions. This chapter is a guided tour of that territory: the rigor of* real analysis *(what a limit really means, and the theorems that make continuity powerful); the language of* point-set topology *(open and closed sets, compactness, connectedness); the surprising rigidity of* complex analysis *(where differentiability forces the Cauchy–Riemann equations and makes contour integrals computable by residues); the counting of* combinatorics and discrete probability*; the approximation methods of* numerical analysis*; the scaffolding of* logic and set theory *(quantifiers, proof techniques, countability); and a first taste of* algorithms and complexity*. No single topic runs deep on the exam, but each rewards knowing its one or two signature moves. The aim here is recognition and fluency on exactly those moves, not a full course in any one field.



Learning Objectives

You will be able to:

1. state the ε–δ definitions of limit and continuity and use them, and apply the Intermediate and Extreme Value Theorems;
2. classify sets as open, closed, compact, or connected, and use the open-set characterization of continuity;
3. test a complex function for analyticity with the Cauchy–Riemann equations and evaluate a contour integral by residues;
4. count with permutations, combinations, inclusion–exclusion, and the pigeonhole principle, and compute basic and conditional probabilities and expectations;
5. run Newton's method and estimate the error of a numerical integration rule;
6. use quantifiers and the standard proof techniques (contradiction, induction), and argue about countability;
7. compare growth rates with big-O notation.

Real Analysis

Calculus asks *how* to differentiate and integrate; real analysis asks *why* the answers are trustworthy, by pinning down the one idea everything rests on—the limit—with no appeal to pictures or hand-waving. The payoff on the exam is a cluster of questions about ε–δ arguments, the convergence of sequences, and the three theorems that turn continuity into a tool: the Intermediate Value Theorem,

the Extreme Value Theorem, and (from Chapter 2) the Mean Value Theorem.

What & why. The rigorous meaning of "approaches" and the theorems it unlocks; the foundation under all of calculus.

Solve. Translate a limit claim into: for every tolerance ε there is a closeness δ that forces the outputs within ε.

Means. A limit is a promise you can meet any accuracy demand.

The ε–δ *definition* says that $\lim_{x\to a} f(x) = L$ means

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta > 0 : \quad 0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon. \tag{7.1}$$

Read it as a challenge–response game: an adversary names an error tolerance ε around L; you must produce a radius δ around a so that every input within δ (except possibly a itself) lands within ε of L. If you can always answer, the limit holds.

Approach Read · Plan · Solve · Check

Read. Name what kind of claim you face—an ε–δ limit, a sequence limit, a continuity statement, or an IVT/EVT application—and identify the target value and the interval.

Plan.

1. For a limit or sequence proof, work *backward* from the goal $|f(x) - L| < \varepsilon$ (or $|a_n - L| < \varepsilon$) to a condition on $|x - a|$ (or on n), which reveals the δ (or N) to choose.
2. For a recursively defined sequence, prove it is monotone *and* bounded, then solve the fixed-point equation for the limit.
3. For an IVT/EVT problem, first confirm continuity on a *closed, bounded* interval, then use a sign change (IVT) or compare critical points against endpoints (EVT).

Solve. Produce the explicit $\delta(\varepsilon)$ or $N(\varepsilon)$, or the value the theorem guarantees, one deduction per line.

Check. Confirm the chosen δ or N depends only on ε (never on x or n), and that every hypothesis of the theorem you invoked actually holds.

Common traps — every wrong answer is one of these

Thinks bounded implies convergent.
Concludes a bounded sequence must converge. Boundedness alone is not enough (take $a_n = (-1)^n$); the Monotone Convergence Theorem also requires the sequence to be monotone.

Lets δ depend on x.
Chooses a δ that still contains the variable x. A valid δ may depend only on ε (and a), so that one radius answers the whole challenge.

Drops a hypothesis of the IVT or EVT.
Applies the theorems on an open or unbounded interval, or to a discontinuous function. Both need *continuity* on a *closed, bounded* interval; $1/x$ on $(0, 1)$ attains no maximum.

Confuses pointwise and uniform convergence.
1. Sees $f_n \to f$ at each fixed x.
2. Concludes the limit function inherits continuity.
3. Fix: only *uniform* convergence preserves continuity—check whether $\sup_x |f_n(x) - f(x)| \to 0$, not just the pointwise gap.

An ε–δ proof. Prove $\lim_{x \to 3}(2x - 1) = 5$ directly from the definition.

Read.

Work backward from the target $|f(x) - L| < \varepsilon$ to find the δ that makes it hold.

Plan.

We need $|(2x - 1) - 5| < \varepsilon$. Simplify the left side: $|2x - 6| = 2|x - 3|$. So the requirement is $2|x - 3| < \varepsilon$, i.e. $|x - 3| < \frac{\varepsilon}{2}$. This tells us what to choose. Given any $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{2}$. Then $0 < |x - 3| < \delta$ forces

$$\begin{aligned} |(2x - 1) - 5| &= 2|x - 3| && \text{simplify the left side} \\ &< 2\delta && \text{since } |x - 3| < \delta \\ &= \varepsilon && \text{because } \delta = \tfrac{\varepsilon}{2}. \end{aligned}$$

Solve.

Choosing $\delta = \varepsilon/2$ works for every $\varepsilon > 0$, so the limit is 5.

Check.

The relation $\delta = \varepsilon/2$ mirrors the slope 2: a steeper line needs a proportionally tighter input window to hold the same output tolerance. ✓

A *sequence* (a_n) converges to L if for every $\varepsilon > 0$ there is an index N with $|a_n - L| < \varepsilon$ for all $n > N$—the same game as (7.1), indexed by N instead of δ. A central structural fact, the *Monotone Convergence Theorem*, says a sequence that is both monotone (always non-decreasing or always non-increasing) and bounded must converge; this is where the *completeness* of $\mathbb{R}$ enters, the property

that every set bounded above has a least upper bound (supremum).

A sequence limit from the definition. Show $\lim_{n\to\infty} \dfrac{n}{n+1} = 1$.

Read.

Bound $|a_n - 1|$ by a simple expression in n, then solve for the N that makes it $< \varepsilon$.

Plan.

Compute the gap:

$$\begin{aligned} \left|\frac{n}{n+1} - 1\right| &= \left|\frac{n-(n+1)}{n+1}\right| && \text{common denominator} \\ &= \frac{1}{n+1} && \text{simplify the numerator.} \end{aligned}$$

We want $\frac{1}{n+1} < \varepsilon$, i.e. $n+1 > \frac{1}{\varepsilon}$, i.e. $n > \frac{1}{\varepsilon} - 1$. So given $\varepsilon > 0$, choose any integer $N \geq \frac{1}{\varepsilon} - 1$; then $n > N$ forces $\frac{1}{n+1} < \varepsilon$.

Solve.

The limit is 1: the terms approach 1 from below, and no positive tolerance eludes them.

Check.

Sanity: $a_{99} = 0.99$, within $\varepsilon = 0.02$, and $N \approx 1/0.02 - 1 = 49 < 99$. ✓

A function f is *continuous* at a when $\lim_{x\to a} f(x) = f(a)$. On a closed bounded interval $[a, b]$, continuity is powerful:

Two theorems that need a closed interval. *Intermediate Value Theorem (IVT):* if f is continuous on $[a, b]$ and y lies between $f(a)$ and $f(b)$, then $f(c) = y$ for some $c \in [a, b]$—continuous functions skip no values. *Extreme Value Theorem (EVT):* a function continuous on a closed bounded interval attains a maximum and a minimum there.

Using the IVT to locate a root. Show $x^3 + x - 1 = 0$ has a solution in $(0, 1)$.

Read.

Evaluate the continuous function at the endpoints; a sign change forces a root by the IVT.

Plan.

Let $f(x) = x^3 + x - 1$, a polynomial, hence continuous on $[0, 1]$. At the endpoints, $f(0) = -1 < 0$ and $f(1) = 1 + 1 - 1 = 1 > 0$. Since 0 lies between $f(0) = -1$ and $f(1) = 1$, the IVT guarantees some $c \in (0, 1)$ with $f(c) = 0$.

Solve.

A root exists in $(0, 1)$.

Check.

The theorem gives existence, not the value; here the root is near 0.68, safely inside the interval. ✓

Misconception

The IVT and EVT need a *closed, bounded* interval and *continuity.* On the open interval $(0, 1)$, $f(x) = 1/x$ is continuous but attains no maximum; on $[-1, 1]$ a function with a jump can skip a value. Drop either hypothesis and the conclusion can fail—a favorite source of "which of the following must be true" distractors.

Pointwise vs. uniform, briefly. On $[0, 1]$, let $f_n(x) = x^n$. Find the pointwise limit and note why the convergence is not uniform.

Read.

Take the limit at each fixed x to get the limit function; then check whether the worst-case gap shrinks uniformly.

Plan.

Fix x. If $0 \le x < 1$, then $x^n \to 0$; if $x = 1$, then $x^n = 1 \to 1$. So the pointwise limit is the function $f(x) = 0$ on $[0, 1)$ and $f(1) = 1$—discontinuous. The largest gap $\sup_x |f_n(x) - f(x)|$ stays near 1 (take x close to 1), so it does not shrink to 0.

Solve.

Pointwise limit is a discontinuous step; convergence is *not* uniform.

Check.

Uniform convergence would force the limit of continuous f_n to be continuous; the broken limit here signals the convergence can only be pointwise. ✓

The squeeze theorem. Evaluate $\lim_{x \to 0} x^2 \sin \frac{1}{x}$.

Read.

Bound the oscillating factor between ± 1, multiply by the envelope, and squeeze.

Plan.

For $x \neq 0$, $-1 \le \sin \frac{1}{x} \le 1$. Multiplying through by $x^2 \ge 0$ preserves the inequalities:

$-x^2 \le x^2 \sin\frac{1}{x} \le x^2$. Both outer bounds $\pm x^2 \to 0$ as $x \to 0$.

Solve.

By the squeeze theorem the limit is 0.

Check.

The wild oscillation of $\sin\frac{1}{x}$ is caged by the x^2 envelope collapsing to 0. ✓

A recursively defined sequence. Let $a_1 = 1$ and $a_{n+1} = \sqrt{2 + a_n}$. Show (a_n) converges and find its limit.

Read.

Prove it is increasing and bounded above (so it converges), then solve the fixed-point equation for the limit.

Plan.

Bounded: if $a_n < 2$ then $a_{n+1} = \sqrt{2 + a_n} < \sqrt{4} = 2$; with $a_1 = 1 < 2$, induction gives $a_n < 2$ for all n. Increasing: $a_{n+1}^2 - a_n^2 = 2 + a_n - a_n^2 = -(a_n - 2)(a_n + 1) > 0$ on $0 < a_n < 2$, so $a_{n+1} > a_n$. Monotone and bounded $\Rightarrow$ converges to some L. Let $n \to \infty$ in $a_{n+1} = \sqrt{2 + a_n}$: $L = \sqrt{2 + L}$, so $L^2 - L - 2 = (L - 2)(L + 1) = 0$, giving $L = 2$ (discard $L = -1$).

Solve.

The sequence increases to $L = 2$.

Check.

The limit must satisfy the recursion's fixed-point equation; positivity selects $L = 2$. ✓

A limit that fails to exist. Show $\lim_{x \to 0} \sin\dfrac{1}{x}$ does not exist.

Read.

Produce two input sequences approaching 0 whose outputs approach different values.

Plan.

Take $x_n = \frac{1}{2\pi n} \to 0$: then $\sin\frac{1}{x_n} = \sin(2\pi n) = 0$. Take instead $x_n' = \frac{1}{\pi/2 + 2\pi n} \to 0$: then $\sin\frac{1}{x_n'} = \sin\left(\frac{\pi}{2} + 2\pi n\right) = 1$. Two sequences tending to 0 give outputs tending to 0 and to 1.

Solve.

Different approaches yield different limiting values, so the limit does not exist.

Check.

Contrast the earlier $x^2 \sin \frac{1}{x}$, whose envelope *forced* convergence; here nothing damps the oscillation. ✓

The Extreme Value Theorem in action. Does $f(x) = x^3 - 3x$ attain a maximum on $[0, 2]$, and where?

Read.

The EVT guarantees a maximum on a closed bounded interval; find it among the interior critical points and the endpoints.

Plan.

f is continuous on the closed bounded interval $[0, 2]$, so by the EVT it attains a maximum there. Candidates are critical points, where $f'(x) = 3x^2 - 3 = 0 \Rightarrow x = 1$ (inside $[0, 2]$), and the endpoints $x = 0, 2$. Evaluate: $f(0) = 0$, $f(1) = 1 - 3 = -2$, $f(2) = 8 - 6 = 2$.

Solve.

The maximum value is 2, attained at the endpoint $x = 2$.

Check.

The EVT promises the maximum exists; comparing critical points against endpoints locates it—and here the endpoint beats the interior critical point. ✓

Point-Set Topology

Topology strips geometry down to a single primitive—which sets count as "open"—and rebuilds continuity, limits, and connectedness from it. On $\mathbb{R}$ (and $\mathbb{R}^n$), a set is *open* if around every one of its points there is a little interval (or ball) still entirely inside the set; a set is *closed* if its complement is open, equivalently if it contains all its limit points. These are not opposites: a set can be both (like $\mathbb{R}$ and $\varnothing$) or neither (like $[0, 1)$).

What & why. The minimal language—open sets—from which continuity, compactness, and connectedness are built.

Solve. Ask of each point whether a whole neighborhood fits inside; ask of a cover whether finitely many pieces suffice.

Means. Open means room to wiggle; compact means no escape to infinity or to a missing edge.

Definition 7.2.1 (Interior, closure, boundary)**.** The *interior* of S is the largest open set inside it (the points with a neighborhood in S); the *closure* $\overline{S}$ is the smallest closed set containing it (S plus

its limit points); the *boundary* ∂S is $\overline{S}$ minus the interior—the points every neighborhood of which meets both S and its complement.

Approach Read · Plan · Solve · Check

Read. Identify the set S and which property is asked—interior, closure, boundary, open/closed, compact, or connected.

Plan.

1. For interior or openness, test whether *every* point of S has a whole neighborhood still inside S.
2. For closure, adjoin the limit points (values approached by sequences drawn from S); the boundary is closure minus interior.
3. For compactness in $\mathbb{R}^n$, apply Heine–Borel—check *closed and bounded together*.
4. For connectedness, try to split S into two disjoint nonempty pieces, each open relative to S; a successful split means *not* connected.

Solve. Carry out the test point by point, recording each verdict.

Check. Open and closed are not opposites—reconcile your answers (a set may be both, like $\mathbb{R}$, or neither, like $[0,1)$), and confirm any Heine–Borel call really sits in $\mathbb{R}^n$.

Common traps — every wrong answer is one of these

Treats open and closed as opposites.
Assumes "not open" means "closed." A set can be both (like $\mathbb{R}$ or $\varnothing$) or neither (like $[0,1)$); openness and closedness are independent tests.

Uses Heine–Borel outside $\mathbb{R}^n$.
Declares any closed and bounded set compact. The equivalence is special to $\mathbb{R}^n$; in a general metric space a closed bounded set can fail to be compact, where compact means every open cover has a finite subcover.

Checks only half of Heine–Borel. 1. Verifies the set is bounded (or only that it is closed).
2. Declares it compact.
3. Fix: in $\mathbb{R}^n$ compactness needs *both*—$(0,1)$ is bounded but not closed, $[0,\infty)$ is closed but not bounded, and neither is compact.

Misses a compact set that carries its limit point.
Judges $\{0\} \cup \{1/n\}$ non-compact because it is infinite. It is compact: bounded, and closed because its only limit point 0 is included.

Interior, closure, boundary of an interval. For $S = (0, 1] \subset \mathbb{R}$, find the interior, closure, and boundary.

Read.

Test each candidate point for a neighborhood inside S (interior) and for being a limit of S (closure).

Plan.

Interior: every point of $(0, 1)$ has a small interval still inside S, but 1 does not (any interval around 1 pokes past to $x > 1$). So the interior is $(0, 1)$. Closure: the limit points of S include 0 (approached from within), so $\overline{S} = [0, 1]$. Boundary: $\overline{S} \setminus \text{int} = \{0, 1\}$.

Solve.

Interior $(0, 1)$; closure $[0, 1]$; boundary $\{0, 1\}$.

Check.

S is neither open (it contains 1 without room) nor closed (it omits its limit point 0), exactly what mixed brackets suggest. ✓

Misconception

"Closed and bounded" means compact only in $\mathbb{R}^n$. The Heine–Borel theorem—compact $\iff$ closed and bounded—is special to $\mathbb{R}^n$. In a general metric space a set can be closed and bounded yet fail to be compact (the closed unit ball of an infinite-dimensional space is the stock counterexample). Compactness really means *every open cover has a finite subcover*, not anything about brackets.



A half-open interval. The open end (hollow dot) is excluded but is a limit point, so it joins the closure; the closed end (solid dot) is included. Neither open nor closed.

The exam's favorite topological property is *compactness*. In $\mathbb{R}^n$ the *Heine–Borel theorem* makes it concrete:

Heine–Borel. A subset of $\mathbb{R}^n$ is *compact* if and only if it is *closed and bounded*. (In general spaces, compact means every open cover has a finite subcover; that abstract form is what makes compactness so useful, but Heine–Borel is the test you will use.)

Compact or not? Decide whether each subset of $\mathbb{R}$ is compact: (a) $[0,1]$, (b) $(0,1)$, (c) $\{0\} \cup \{1/n : n \in \mathbb{N}\}$, (d) $[0,\infty)$.

Read.

Apply Heine–Borel: check closed AND bounded.

Plan.

(a) $[0,1]$ is closed and bounded—compact. (b) $(0,1)$ is bounded but not closed (misses $0,1$)—not compact. (c) Bounded (inside $[0,1]$); closed because its only limit point is 0, which is included—compact. (d) Closed but unbounded—not compact.

Solve.

Compact: (a) and (c). Not compact: (b) and (d).

Check.

Set (c) is the classic surprise: an infinite set can be compact precisely because it drags its one limit point 0 inside. ✓

A space is *connected* if it cannot be split into two disjoint nonempty open pieces. In $\mathbb{R}$ the connected sets are exactly the intervals. Continuity interacts cleanly with both notions: a continuous image of a compact set is compact, and of a connected set is connected—the abstract engines behind the EVT and IVT.

Continuity via open sets. Explain why $f(x) = x^2$ is continuous using the preimage criterion, testing the open set $(1,4)$.

Read.

Continuity means the preimage of every open set is open; check one representative open set.

Plan.

The topological definition: f is continuous iff $f^{-1}(U)$ is open for every open U. Take $U = (1,4)$. Then

$$\begin{aligned} f^{-1}(U) &= \{x : 1 < x^2 < 4\} && \text{unpack the definition of the preimage} \\ &= (-2,-1) \cup (1,2) && \text{solve the inequality,} \end{aligned}$$

a union of open intervals, hence open.

Solve.

The preimage is open, consistent with continuity.

Check.

This reformulation—"open sets pull back to open sets"—is what lets topology define continuity with no metric or limit at all. ✓

Closure of the rationals. Find the closure of $\mathbb{Q}$ inside $\mathbb{R}$.

Read.

Ask which real numbers are limits of sequences drawn from $\mathbb{Q}$.

Plan.

Every real x is a limit of rationals—truncate its decimal expansion to get $q_n \to x$ with each $q_n \in \mathbb{Q}$. So every real number is a limit point of $\mathbb{Q}$, and $\overline{\mathbb{Q}} = \mathbb{R}$.

Solve.

$\overline{\mathbb{Q}} = \mathbb{R}$: the rationals are *dense* in the reals.

Check.

A set whose closure is the whole space is called dense; $\mathbb{Q}$ is the canonical example, and it is exactly why decimal approximation of any real number works. ✓

Connected or not. Is $S = [0, 1] \cup [2, 3]$ connected in $\mathbb{R}$?

Read.

Try to split S into two disjoint nonempty pieces, each open relative to S.

Plan.

The pieces $[0, 1]$ and $[2, 3]$ are disjoint and nonempty, and each is open within S: around any point of $[0, 1]$ a small interval meets S only inside $[0, 1]$, and likewise for $[2, 3]$. That is a valid separation of S.

Solve.

S is *not* connected—the gap $(1, 2)$ disconnects it.

Check.

In $\mathbb{R}$ the connected sets are exactly the intervals; any set with a gap fails. ✓

Complex Analysis

Complex analysis studies functions $f(z)$ of a complex variable, and its central discovery is *rigidity*: merely being differentiable once, in the complex sense, forces a function to be infinitely differentiable,

to satisfy a pair of partial differential equations, and to make its integrals depend only on endpoints (or on enclosed singularities). A function that is complex-differentiable on an open set is called *analytic* (or holomorphic) there.

What & why. Why complex differentiability is far stronger than real, and how it makes integrals computable by residues.

Solve. Split $f = u + \mathrm{i}v$ and test the Cauchy–Riemann equations; for integrals, sum residues of enclosed poles.

Means. Analytic functions are rigid: local smoothness controls global behavior.

Write $z = x + \mathrm{i}y$ and $f(z) = u(x,y) + \mathrm{i}\,v(x,y)$. Requiring the derivative $f'(z)$ to exist independent of the direction of approach forces a link between the partials of u and v.

Derivation. *The Cauchy–Riemann equations*

Approach z first along the real axis ($\Delta z = \Delta x$): the difference quotient tends to $u_x + \mathrm{i}\,v_x$. Approach instead along the imaginary axis ($\Delta z = \mathrm{i}\,\Delta y$): dividing by $\mathrm{i}\,\Delta y$ gives $\frac{u_y + \mathrm{i}v_y}{\mathrm{i}} = v_y - \mathrm{i}\,u_y$. For $f'(z)$ to be well defined these two must agree, so matching real and imaginary parts,

$$u_x = v_y \qquad \text{and} \qquad u_y = -v_x. \tag{7.2}$$

These are the *Cauchy–Riemann equations*: a necessary condition for analyticity (and, with continuous partials, sufficient).

Approach Read · Plan · Solve · Check

Read. Decide whether the task tests analyticity or evaluates an integral. Write $z = x + \mathrm{i}y$ and split $f = u + \mathrm{i}v$.

Plan.

1. For analyticity, compute all four partials and test *both* Cauchy–Riemann equations $u_x = v_y$ and $u_y = -v_x$.
2. For a contour integral, factor the denominator to locate the poles and keep only those *inside* the contour.
3. Find each residue: $\operatorname{Res}(f, z_0) = \lim_{z \to z_0} (z - z_0) f(z)$ for a simple pole, and the $(m-1)$th-derivative rule for a pole of order m.

Solve. Apply the residue theorem $\oint_\gamma f = 2\pi\mathrm{i} \sum_k \operatorname{Res}(f, z_k)$, or report the verdict of the CR test.

Check. Confirm you included every enclosed pole and no outside one, kept the factor $2\pi\mathrm{i}$, and—for analyticity—that *both* CR equations hold, not just one.

Common traps — every wrong answer is one of these

Checks only one Cauchy–Riemann equation.
Verifies $u_x = v_y$, sees it hold, and declares f analytic. Both $u_x = v_y$ *and* $u_y = -v_x$ must hold; $\bar{z}$ fails where z^2 passes.

Counts poles outside the contour.
Includes a singularity that lies beyond γ in the residue sum. Only poles strictly *inside* the contour contribute; check each pole's modulus against the contour.

Uses the simple-pole formula on a higher-order pole.
1. Meets a pole of order 2 or more.
2. Applies $\lim(z - z_0)f(z)$ anyway.
3. Fix: for order m take $\frac{1}{(m-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\left[(z-z_0)^m f(z)\right]$; a double pole needs one derivative.

Drops the $2\pi\mathrm{i}$ factor.
Reports the sum of residues as the integral. The residue theorem multiplies that sum by $2\pi\mathrm{i}$; even a vanishing sum (when residues cancel) comes through that factor.

Testing analyticity. Is $f(z) = z^2$ analytic? Verify with Cauchy–Riemann.

Read.

Write $f = u + \mathrm{i}v$ explicitly, compute the four partials, and check the two CR equations.

Plan.

Expand into real and imaginary parts:

$$\begin{aligned} z^2 &= (x+\mathrm{i}y)^2 && \text{substitute } z = x+\mathrm{i}y \\ &= x^2 - y^2 + \mathrm{i}(2xy) && \text{expand,} \end{aligned}$$

so $u = x^2 - y^2$ and $v = 2xy$. Partials: $u_x = 2x$, $v_y = 2x$, so $u_x = v_y$. And $u_y = -2y$, $v_x = 2y$, so $u_y = -v_x$. Both equations hold everywhere.

Solve.

z^2 is analytic on all of $\mathbb{C}$ (it is entire).

Check.

Consistency check: $f'(z) = u_x + \mathrm{i}v_x = 2x + \mathrm{i}\,2y = 2z$, matching the expected $2z$. ✓

Did You Know? *one theorem, and algebra closes up*

The rigidity of analytic functions has spectacular payoffs. Liouville's theorem states that a function analytic and bounded on *all* of $\mathbb{C}$ must be constant—and from this single fact falls a one-line proof of the Fundamental Theorem of Algebra: if a nonconstant polynomial p had no root, then $1/p(z)$ would be bounded and entire, hence constant, which it plainly is not.

A function that is not analytic. Show $f(z) = \bar{z}$ fails to be analytic.

Read.

Identify u, v from $\bar{z} = x - \mathrm{i}y$ and test the Cauchy–Riemann equations.

Plan.

Here $\bar{z} = x - \mathrm{i}y$, so $u = x$ and $v = -y$. Then $u_x = 1$ but $v_y = -1$, so $u_x \neq v_y$. The first Cauchy–Riemann equation already fails.

Solve.

$\bar{z}$ is nowhere analytic.

Check.

Geometrically, conjugation is a reflection, which reverses orientation; analytic maps preserve orientation and angles, so $\bar{z}$ could not have qualified. ✓

The structural rigidity pays off in integration. Along a curve, $\int_\gamma f(z)\,\mathrm{d}z$ of an analytic function depends only on the endpoints (Cauchy's theorem: the integral around any closed contour in a region where f is analytic is 0). When f has isolated singularities—*poles*—inside a closed contour, the integral is captured by *residues*.

Residue theorem. If f is analytic inside and on a positively oriented simple closed contour γ except at isolated poles z_k inside, then

$$\oint_\gamma f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\sum_k \operatorname{Res}(f, z_k). \tag{7.3}$$

For a simple pole at z_0, $\operatorname{Res}(f, z_0) = \lim_{z\to z_0}(z - z_0)f(z)$.

A contour integral by residues. Evaluate $\displaystyle\oint_{|z|=2} \frac{1}{z^2-1}\,\mathrm{d}z$, the circle of radius 2 traversed counterclockwise.

Read.

Factor to find the poles, keep those inside the contour, sum their residues, multiply by $2\pi\mathrm{i}$.

Plan.

Factor: $\frac{1}{z^2-1} = \frac{1}{(z-1)(z+1)}$, with simple poles at $z=1$ and $z=-1$, both inside $|z|=2$. The two residues are

$$\begin{aligned}
\operatorname{Res}(f,1) &= \lim_{z\to 1}(z-1)f(z) && \text{simple-pole formula}\\
&= \frac{1}{1+1} = \frac{1}{2} && \text{cancel } (z-1),\ \text{evaluate}\\
\operatorname{Res}(f,-1) &= \lim_{z\to -1}(z+1)f(z) && \text{simple-pole formula}\\
&= \frac{1}{-1-1} = -\frac{1}{2} && \text{cancel } (z+1),\ \text{evaluate.}
\end{aligned}$$

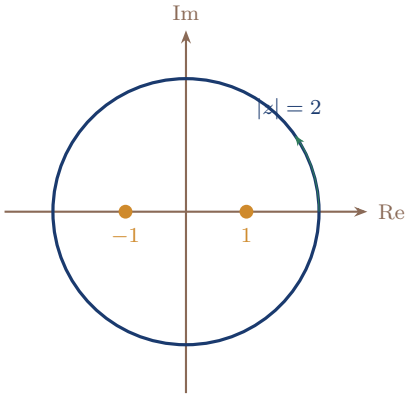
Their sum is $\frac{1}{2} - \frac{1}{2} = 0$, so $\oint = 2\pi\mathrm{i}\cdot 0$.

Solve.

$$\oint_{|z|=2} \frac{\mathrm{d}z}{z^2-1} = 0.$$

Check.

The residues cancel, so despite two enclosed poles the integral vanishes—a reminder to sum, not just count, residues. ✓



The contour $|z| = 2$ encloses both simple poles $z = \pm 1$ (orange). The residue theorem (7.3) turns the loop integral into $2\pi\mathrm{i}$ times the sum of the enclosed residues.

A residue at a double pole. Find $\operatorname{Res}\left(\frac{\mathrm{e}^z}{z^2}, 0\right)$.

Read.

For a pole of order 2 at 0, the residue is $\frac{\mathrm{d}}{\mathrm{d}z}(z^2 f(z))$ evaluated at 0.

Plan.

Multiply out the pole: $z^2 f(z) = \mathrm{e}^z$. For an order-2 pole the residue uses one derivative: $\operatorname{Res} = \lim_{z\to 0} \frac{\mathrm{d}}{\mathrm{d}z}\mathrm{e}^z = \mathrm{e}^0 = 1$.

Solve.

The residue is 1.

Check.

For a pole of order m the rule takes the $(m-1)$th derivative; here $m = 2$ needs exactly one. ✓

A real integral by residues. Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,\mathrm{d}x$ using a semicircular contour.

Read.

Close the contour in the upper half-plane; only the pole there contributes, and the large arc's integral vanishes.

Plan.

Write $\frac{1}{1+z^2} = \frac{1}{(z-\mathrm{i})(z+\mathrm{i})}$, with poles at $\pm\mathrm{i}$; only $z = \mathrm{i}$ lies in the upper half-plane. Its residue is $\lim_{z\to\mathrm{i}}(z-\mathrm{i})\frac{1}{(z-\mathrm{i})(z+\mathrm{i})} = \frac{1}{2\mathrm{i}}$. As the semicircle radius $\to\infty$ the arc integral $\to 0$, so the real-axis integral equals $2\pi\mathrm{i}\cdot\frac{1}{2\mathrm{i}} = \pi$.

Solve.

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2} = \pi.$$

Check.

Cross-check with calculus: the antiderivative is $\arctan x$, and $\arctan(\infty) - \arctan(-\infty) = \frac{\pi}{2} - (-\frac{\pi}{2}) = \pi$. ✓

Finding a harmonic conjugate. Given $u(x,y) = x^2 - y^2$, find v so that $f = u + \mathrm{i}v$ is analytic.

Read.

Use the Cauchy–Riemann equations (7.2) to recover v from u by integration.

Plan.

From $u_x = v_y$: since $u_x = 2x$, we get $v_y = 2x$, and integrating in y gives $v = 2xy + g(x)$. From $u_y = -v_x$: $u_y = -2y$ while $v_x = 2y + g'(x)$, so $-v_x = -2y - g'(x)$ must equal $-2y$, forcing $g'(x) = 0$.

Solve.

$v = 2xy + C$; indeed $f = x^2 - y^2 + \mathrm{i}\,2xy = z^2$.

Check.

The conjugate is fixed up to a constant, and the pair reassembles into the analytic z^2 from earlier—the same example run backward. ✓

Combinatorics and Discrete Probability

Counting problems reward a small set of tools applied in the right order. The *multiplication principle* says independent choices multiply. *Permutations* count ordered arrangements: $P(n,k) = \frac{n!}{(n-k)!}$. *Combinations* count unordered selections: $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$. When categories overlap, *inclusion–exclusion* corrects for double counting; and the *pigeonhole principle* turns "more items than boxes" into a guaranteed collision.

What & why. The toolkit for counting arrangements and selections, and for turning counts into probabilities.

Solve. Decide whether order matters (permutation vs combination); subtract overlaps; divide favorable by total for probability.

Means. Probability is counting with a denominator.

Approach Read · Plan · Solve · Check

Read. Decide what is being counted or measured: an ordered arrangement, an unordered selection, a union of overlapping sets, or a probability/expectation.

Plan.

1. If order matters, use a permutation $P(n,k) = \frac{n!}{(n-k)!}$; if it does not, use a combination $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$.
2. For overlapping categories, correct by inclusion–exclusion $|A \cup B| = |A| + |B| - |A \cap B|$.
3. For repeated identical items, divide the total permutations by each repeat's factorial.
4. For probability, form $\frac{\#\text{favorable}}{\#\text{total}}$; for a conditional probability update the sample space first;

for expectation use $E[X] = \sum x_i P(x_i)$ or linearity.

Solve. Compute the counts, then combine them into the requested number one factor at a time.

Check. Sanity-check the size: a probability lies in $[0, 1]$, "at least one" plus "none" recovers the whole, and a committee count is the slate count divided by the orderings.

Common traps — every wrong answer is one of these

Confuses permutations and combinations.
Uses $P(n, k)$ for an unordered committee or $\binom{n}{k}$ for distinct ordered roles. They differ by the $k!$ orderings of the chosen items; ask "does relabeling change the outcome?"

Adds probabilities of non-exclusive events.
Writes $P(A \cup B) = P(A) + P(B)$ when A and B can co-occur. That double-counts the overlap; subtract $P(A \cap B)$ unless the events are mutually exclusive.

Multiplies probabilities of dependent events.
Uses $P(A \cap B) = P(A)P(B)$ without independence. After drawing without replacement the second draw depends on the first; condition instead with $P(A \cap B) = P(A)P(B \mid A)$.

Ignores the base rate in a Bayes problem.
1. Reads a test as 99% accurate.
2. Concludes a positive result means a 99% chance of disease.
3. Fix: weight by the prior—a rare condition makes most positives false alarms, so expand $P(+)$ over both groups before dividing.

Order matters vs. not. From 10 people, (a) how many ways to elect a president, vice-president, and treasurer; (b) how many ways to choose a committee of 3?

Read.

Decide whether the roles are distinguishable: distinct roles ⇒ permutation; a plain committee ⇒ combination.

Plan.

(a) The three offices are distinct, so order matters: $P(10, 3) = 10 \cdot 9 \cdot 8 = 720$. (b) A

committee is unordered:

$$\binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3!} \qquad \text{unordered selection}$$
$$= \frac{720}{6} = 120 \qquad \text{divide out the 3! orderings.}$$

Solve.

(a) 720 ordered slates; (b) 120 committees.

Check.

The committee count is the slate count divided by the $3! = 6$ orderings of the same three people—the signature relationship between P and $()$. ✓

Inclusion–exclusion. Of 100 students, 60 take Spanish, 45 take French, and 20 take both. How many take at least one language?

Read.

Add the two sets, then subtract the overlap counted twice.

Plan.

By inclusion–exclusion,

$$|S \cup F| = |S| + |F| - |S \cap F| \qquad \text{subtract the double-counted overlap}$$
$$= 60 + 45 - 20 = 85 \qquad \text{insert the counts.}$$

Solve.

85 students take at least one language.

Check.

Then $100 - 85 = 15$ take neither—a quick consistency figure that must be non-negative. ✓

Did You Know? *the birthday paradox*

In a room of just 23 people the probability that two share a birthday already exceeds one half, and by 70 people it passes 99.9%. The surprise dissolves once you count *pairs* rather than people: 23 individuals form $\binom{23}{2} = 253$ pairs, each a separate chance for a match, which is why collisions arrive far sooner than the 365 days would suggest.

Pigeonhole. Show that among any 13 people, at least two share a birth month.

Read.

Identify the pigeons and the holes; more pigeons than holes forces a shared hole.

Plan.

The 13 people are pigeons; the 12 months are holes. Since $13 > 12$, at least one month must receive two or more people.

Solve.

At least two of the 13 share a birth month.

Check.

The bound is tight: 12 people could occupy all distinct months, so 13 is the first count that forces a repeat. ✓

The *binomial theorem* expands

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k, \tag{7.4}$$

tying the combination numbers to algebra. On the probability side, for equally likely outcomes $P(\text{event}) = \frac{\#\text{favorable}}{\#\text{total}}$, and the two rules that generate most exam questions are *conditional probability* and *expectation*.

Conditional probability and Bayes.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}. \tag{7.5}$$

Events are *independent* when $P(A \cap B) = P(A)P(B)$. The *expected value* of a discrete random variable is $E[X] = \sum_i x_i \, P(X = x_i)$.

Conditional probability from a tree. A bag has 3 red and 2 blue chips; draw two without replacement. Given the first is red, what is the probability the second is red?

Read.

Condition on the first draw: update the counts, then compute the second-draw probability.

Plan.

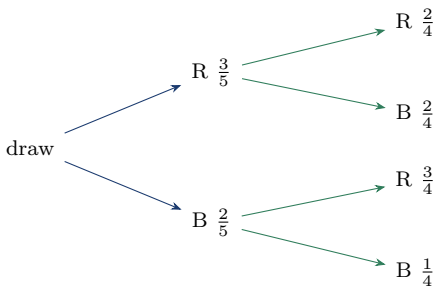
After drawing one red, 2 red and 2 blue remain, 4 chips total. So $P(\text{2nd red} \mid \text{1st red}) = \frac{2}{4} = \frac{1}{2}$.

Solve.

The probability is $\frac{1}{2}$.

Check.

Conditioning literally removes the drawn chip from the sample space; the denominator dropped from 5 to 4. ✓



A probability tree for two draws without replacement. Conditioning on "first red" selects the top branch, where the second-draw probabilities are computed from the reduced bag.

Expectation of a die. Find the expected value of one roll of a fair six-sided die.

Read.

Multiply each outcome by its probability and sum.

Plan.

Each face $1, \dots, 6$ has probability $\frac{1}{6}$, so

$$\begin{aligned} E[X] &= \tfrac{1}{6}(1+2+3+4+5+6) && \text{weight each outcome by } \tfrac{1}{6} \\ &= \frac{21}{6} = \frac{7}{2} && \text{sum and reduce.} \end{aligned}$$

Solve.

$E[X] = 3.5$.

Check.

The mean sits exactly at the midpoint of the symmetric outcomes 1 through 6, as symmetry demands. ✓

Distributing identical items (stars and bars). How many nonnegative integer solutions does $x + y + z = 10$ have?

Read.

Model the total as ten identical stars split into three groups by two dividers, then count arrangements.

Plan.

Ten identical units are separated into three groups by 2 dividers. Any arrangement of 10 stars and 2 bars in a row determines a solution, and there are

$$\begin{aligned} \binom{10+2}{2} &= \binom{12}{2} && \text{choose positions for the 2 bars} \\ &= \frac{12 \cdot 11}{2} = 66 && \text{evaluate} \end{aligned}$$

such arrangements.

Solve.

There are 66 solutions.

Check.

Stars and bars turns "parts summing to n" into one combination, $\binom{n+k-1}{k-1}$ for k nonnegative parts. ✓

Expected value by linearity. A fair coin is flipped 5 times. Find the expected number of heads.

Read.

Write the count as a sum of indicator variables and add their expectations.

Plan.

Let $X = X_1 + \cdots + X_5$, where $X_i = 1$ if flip i is heads and 0 otherwise. Each $E[X_i] = \frac{1}{2}$, so by linearity $E[X] = 5 \cdot \frac{1}{2} = \frac{5}{2}$.

Solve.

The expected number of heads is 2.5.

Check.

Linearity of expectation needs no independence and reproduces the binomial mean $np = 5 \cdot \frac{1}{2}$. ✓

Bayes' theorem and base rates. A test is 99% accurate on both the sick and the healthy; the disease affects 1 in 1000. Given a positive test, what is the probability of actually being sick?

Read.

Apply the conditional-probability rule (7.5) in Bayes form, $P(\text{sick} \mid +) = \frac{P(+|\text{sick})P(\text{sick})}{P(+)}$, expanding the denominator over both groups.

Plan.

Here $P(\text{sick}) = 0.001$, $P(+ \mid \text{sick}) = 0.99$, $P(+ \mid \text{healthy}) = 0.01$. Total positive rate:

$$\begin{aligned} P(+) &= 0.99(0.001) + 0.01(0.999) && \text{split over sick and healthy groups} \\ &= 0.00099 + 0.00999 = 0.01098 && \text{evaluate each term.} \end{aligned}$$

Then $P(\text{sick} \mid +) = \frac{0.00099}{0.01098} \approx 0.090$.

Solve.

About 9%: a positive result still probably means healthy.

Check.

The rare base rate dominates—most positives are false alarms from the huge healthy majority, the base-rate fallacy the exam loves to probe. ✓

Arrangements with repeated letters. How many distinct arrangements are there of the letters in the word `MISSISSIPPI`?

Read.

Take total permutations of all letters, then divide by the factorial of each repeated letter's count.

Plan.

The word has 11 letters with repeats: M×1, I×4, S×4, P×2. Distinct arrangements:

$$\begin{aligned} \frac{11!}{1!\,4!\,4!\,2!} &= \frac{39916800}{1 \cdot 24 \cdot 24 \cdot 2} && \text{expand each factorial} \\ &= \frac{39916800}{1152} = 34650 && \text{divide.} \end{aligned}$$

Solve.

There are 34650 distinct arrangements.

Check.

Dividing by each repeat's factorial cancels the indistinguishable reorderings of identical letters—the multinomial count. ✓

Numerical Analysis

When an equation cannot be solved in closed form, numerical analysis produces an approximate answer to controllable accuracy. The exam touches root-finding by *Newton's method*, the idea of *iteration and error*, and *numerical integration* by the trapezoid and Simpson rules.

Derivation. *Newton's method from linearization*

Near a point x_n, a differentiable f is well approximated by its tangent line $y = f(x_n) + f'(x_n)(x - x_n)$. Set this linear model to zero and solve for x: the tangent hits the axis at $x = x_n - \frac{f(x_n)}{f'(x_n)}$. Taking that as the next guess gives the update

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{7.6}$$

which converges rapidly (quadratically) once close to a simple root.

Approach Read · Plan · Solve · Check

Read. Identify the method—Newton, bisection, fixed-point iteration, trapezoid, or Simpson—and how many steps or what accuracy is wanted.

Plan.

1. For root-finding by Newton, write the update $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ and iterate; for bisection, test the midpoint's sign and keep the subinterval that still brackets the root.
2. For numerical integration, set $h = \frac{b-a}{n}$ and sample f at the nodes.
3. Apply the weights: trapezoid uses $1, 2, \ldots, 2, 1$ times $\frac{h}{2}$; Simpson uses $1, 4, 2, 4, \ldots, 4, 1$ times $\frac{h}{3}$ and requires *even* n.

Solve. Iterate or sum one line at a time, carrying enough digits to see the convergence.

Check. Compare against the exact value or an error bound—trapezoid error scales like h^2, Simpson like h^4, and each bisection step halves the bracket.

Common traps — every wrong answer is one of these

Uses Simpson's rule with odd n.
Applies the $1, 4, 2, \ldots, 4, 1$ pattern when n is odd. Simpson pairs subintervals into parabolic arcs, so n must be *even*; an odd n leaves a leftover interval and breaks the weights.

Swaps the $\frac{h}{2}$ and $\frac{h}{3}$ prefactors.
Multiplies the trapezoid sum by $\frac{h}{3}$ or Simpson's by $\frac{h}{2}$. Keep them straight: trapezoid is $\frac{h}{2}(f_0 + 2f_1 + \cdots + f_n)$, Simpson is $\frac{h}{3}(f_0 + 4f_1 + 2f_2 + \cdots + f_n)$.

Trusts Newton to always converge.
Assumes each step improves the guess. Newton can diverge when f' vanishes near the root or the start is poor; bisection is slower but always converges once it brackets a sign change.

Keeps the wrong bisection subinterval. 1. Computes the midpoint value.
2. Keeps a half without checking signs.
3. Fix: retain the subinterval whose endpoints give *opposite* signs of f—that is the half still guaranteed to contain a root.

Newton's method. Approximate $\sqrt{2}$ by applying Newton's method to $f(x) = x^2 - 2$, starting at $x_0 = 1$; take two steps.

Read.

Write the update (7.6) for this f, then iterate twice from $x_0 = 1$.

Plan.

Here $f'(x) = 2x$, so the update is

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^2 - 2}{2x_n} && \text{Newton's update with } f'(x) = 2x \\ &= \frac{x_n}{2} + \frac{1}{x_n} && \text{simplify.} \end{aligned}$$

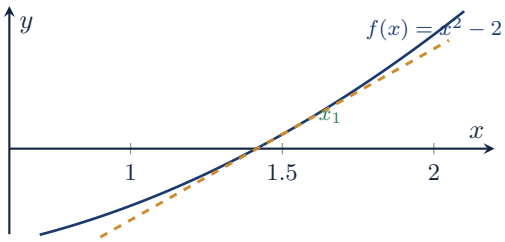
Step one: $x_1 = \frac{1}{2} + \frac{1}{1} = 1.5$. Step two: $x_2 = \frac{1.5}{2} + \frac{1}{1.5} = 0.75 + 0.6667 = 1.41667$.

Solve.

$x_2 \approx 1.41667$, already correct to three decimals ($\sqrt{2} \approx 1.41421$).

Check.

The error roughly squares each step—one, then three, correct digits—the hallmark of quadratic convergence. ✓



Newton's method: the tangent at the current guess crosses the axis at the next guess, sliding rapidly toward the root $\sqrt{2}$.

Numerical integration. Trapezoid rule, with $h = \frac{b-a}{n}$:

$$\int_a^b f \approx \frac{h}{2}(f_0 + 2f_1 + \cdots + 2f_{n-1} + f_n), \tag{7.7}$$

its error scaling like h^2. Simpson's rule (even n) weights the interior points $4, 2, 4, 2, \ldots$: $\int_a^b f \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + \cdots + 4f_{n-1} + f_n)$, with error scaling like h^4—dramatically more accurate.

Trapezoid estimate. Estimate $\int_0^1 x^2 \, dx$ with the trapezoid rule using $n = 2$ subintervals, and compare to the exact value.

Read.

Set h, sample f at the nodes, apply the trapezoid rule (7.7) with weights $1, 2, 1$, and compare to the exact integral.

Plan.

With $n = 2$, $h = \frac{1}{2}$; nodes $x = 0, \frac{1}{2}, 1$ give $f = 0, \frac{1}{4}, 1$. Trapezoid:

$$\begin{aligned} \frac{h}{2}(f_0 + 2f_1 + f_2) &= \frac{1/2}{2}(0 + 2 \cdot \tfrac{1}{4} + 1) && \text{apply weights } 1, 2, 1 \\ &= \frac{1}{4}(\tfrac{3}{2}) && \text{sum inside the parentheses} \\ &= \frac{3}{8} = 0.375 && \text{multiply.} \end{aligned}$$

Exact: $\int_0^1 x^2 \, dx = \frac{1}{3} \approx 0.3333$.

Solve.

Estimate 0.375 vs. exact 0.3333; overestimate by about 0.042.

Check.

The trapezoid rule overestimates for a convex ($f'' > 0$) integrand, since straight chords lie above the cupped curve—exactly the sign seen here. ✓

Fixed-point iteration. Estimate the root of $x = \cos x$ by iterating $x_{n+1} = \cos x_n$ from $x_0 = 1$; take three steps (radians).

Read.

Repeatedly feed each output back in as the next input; the values approach the fixed point where $x = \cos x$.

Plan.

Iterate: $x_1 = \cos 1 \approx 0.5403$; $x_2 = \cos 0.5403 \approx 0.8576$; $x_3 = \cos 0.8576 \approx 0.6543$.

Solve.

The values close in on $x \approx 0.739$, the fixed point of cosine.

Check.

The iteration converges because $\left|\frac{\mathrm{d}}{\mathrm{d}x}\cos x\right| = |\sin x| < 1$ near the root—the map is a contraction there. ✓

Simpson's rule is exact on cubics. Estimate $\displaystyle\int_0^1 x^3\,\mathrm{d}x$ by Simpson's rule with $n = 2$, and compare to the exact value.

Read.

Apply the $1, 4, 1$ Simpson weights at the nodes $0, \frac{1}{2}, 1$, then compare.

Plan.

With $h = \frac{1}{2}$, the nodes $x = 0, \frac{1}{2}, 1$ give $f = 0, \frac{1}{8}, 1$. Simpson:

$$\begin{aligned} \frac{h}{3}(f_0 + 4f_1 + f_2) &= \frac{1/2}{3}(0 + 4\cdot\tfrac{1}{8} + 1) && \text{apply weights } 1,4,1 \\ &= \frac{1}{6}(\tfrac{3}{2}) && \text{sum inside the parentheses} \\ &= \frac{1}{4} && \text{multiply.} \end{aligned}$$

Exact: $\int_0^1 x^3\,\mathrm{d}x = \frac{1}{4}$.

Solve.

Simpson gives 0.25, exactly the true value.

Check.

Simpson integrates cubics with zero error: its error depends on the fourth derivative, which vanishes for a cubic. ✓

Bisection halves the error. $f(x) = x^2 - 2$ has a root in $[1, 2]$. Do two bisection steps and bound the error.

Read.

Test the midpoint's sign, keep the subinterval that still brackets the root, and repeat; the bracket halves each step.

Plan.

Midpoint 1.5: $f(1.5) = 0.25 > 0$ while $f(1) < 0$, so the root is in $[1, 1.5]$. Midpoint 1.25: $f(1.25) = -0.4375 < 0$ while $f(1.5) > 0$, so the root is in $[1.25, 1.5]$.

Solve.

After two steps the root lies in $[1.25, 1.5]$, of width 0.25, so its midpoint 1.375 is within 0.125 of $\sqrt{2}$.

Check.

The bracket shrinks by exactly one half each step: the error bound is $\frac{b-a}{2^n} = \frac{1}{2^2}$ after two steps—slow but utterly reliable, unlike Newton. ✓

Logic and Set Theory

Every proof rests on logic, and the exam tests the mechanics: reading quantifiers, negating statements, and choosing a proof technique. A *universal* statement $\forall x\, P(x)$ is disproved by a single counterexample; an *existential* statement $\exists x\, P(x)$ is proved by a single example. Negation swaps the quantifiers and negates the inside: $\neg(\forall x\, P(x)) \equiv \exists x\, \neg P(x)$.

What & why. The rules for quantifiers, negation, and the standard proof techniques that every other chapter quietly relies on.

Solve. Match the statement's shape to a method: direct, contrapositive, contradiction, or induction.

Means. A theorem is a claim you can defend against every case.

Approach Read · Plan · Solve · Check

Read. Classify the statement—universal, existential, an implication, or a claim about set size—and identify what must be proved or disproved.

Plan.

1. To negate, swap each quantifier ($\forall \leftrightarrow \exists$) and negate the inside; an implication $P \Rightarrow Q$ negates to $P \wedge \neg Q$.

2. Match the proof to the shape: one counterexample kills a $\forall$; one example proves an $\exists$; use the contrapositive when the hypothesis is awkward; use contradiction by assuming the negation; use induction for a statement indexed by n.
3. For countability, exhibit an explicit listing (a bijection with $\mathbb{N}$); for uncountability, run a diagonal argument.

Solve. Carry out the chosen technique—base case then inductive step, or assume-and-derive—one deduction per line.

Check. Confirm the final line is exactly the target (for induction, the statement at $n = k + 1$), and that every case is covered.

Common traps — every wrong answer is one of these

Negates an implication as an implication.
Turns "$P \Rightarrow Q$" into "$P \Rightarrow \neg Q$." The correct negation is $P \wedge \neg Q$: a single continuous-but-not-differentiable function, not a rule about all functions.

Forgets to negate the inside when flipping quantifiers.
Rewrites $\neg(\forall x\, P(x))$ as $\exists x\, P(x)$. Flipping the quantifier is only half the move; the predicate must flip too, giving $\exists x\, \neg P(x)$.

Confuses "countable" with "finite".
Thinks an infinite set cannot be countable. Countable means listable in a never-repeating sequence—$\mathbb{Z}$ and $\mathbb{Q}$ qualify; only $\mathbb{R}$ (by diagonalization) does not.

Uses ordinary induction where strong induction is needed.
1. Assumes the claim only at $n - 1$.
2. Tries to reach a factor a far below n.
3. Fix: when the step depends on an arbitrary smaller case (as in prime factorization), assume the claim for *all* integers up to $n - 1$.

Negating a quantified statement. Negate: "Every continuous function is differentiable."

Read.

Rewrite with explicit quantifiers, then push the negation inward, flipping the quantifier.

Plan.

Symbolically the claim is $\forall f\, (\text{cont}(f) \Rightarrow \text{diff}(f))$. Its negation is $\exists f\, (\text{cont}(f) \wedge \neg\, \text{diff}(f))$:

"there exists a continuous function that is not differentiable."

Solve.

The negation asserts a counterexample exists—and indeed $|x|$ is one (continuous, not differentiable at 0).

Check.

Negating $\forall(P \Rightarrow Q)$ gives $\exists(P \wedge \neg Q)$; the implication becomes an "and," a step students routinely drop. ✓

The three workhorse techniques are *direct proof*, *proof by contradiction* (assume the negation, derive an impossibility), and *mathematical induction* (prove a base case, then that each case forces the next).

Proof by induction. Prove $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ for all $n \geq 1$.

Read.

Verify the base case, assume the statement at $n = k$, then derive it at $n = k + 1$.

Plan.

Base case $n = 1$: the left side is 1, the right side $\frac{1 \cdot 2}{2} = 1$. ✓ Inductive step: assume $1 + \cdots + k = \frac{k(k+1)}{2}$. Add $k + 1$ to both sides:

$$\begin{aligned} 1 + \cdots + k + (k+1) &= \frac{k(k+1)}{2} + (k+1) && \text{apply the inductive hypothesis} \\ &= (k+1)\left(\frac{k}{2} + 1\right) && \text{factor out } (k+1) \\ &= \frac{(k+1)(k+2)}{2} && \text{combine,} \end{aligned}$$

which is the formula at $n = k + 1$.

Solve.

The formula holds for all $n \geq 1$ by induction.

Check.

The final form $\frac{(k+1)(k+2)}{2}$ is the target with $n \to k + 1$, so the step is airtight. ✓

Proof by contradiction. Prove $\sqrt{2}$ is irrational.

Read.

Assume it is rational in lowest terms, then derive a contradiction with "lowest terms."

Plan.

Suppose $\sqrt{2} = \frac{p}{q}$ in lowest terms (no common factor). Then $2q^2 = p^2$, so p^2 is even, hence p is even, say $p = 2m$. Substituting, $2q^2 = 4m^2$, so $q^2 = 2m^2$, making q even too. But then p and q share the factor 2, contradicting "lowest terms."

Solve.

No such fraction exists, so $\sqrt{2}$ is irrational.

Check.

The contradiction is with the reducibility assumption, the standard engine of this classic proof. ✓

Set theory also asks about the *sizes* of infinite sets. Two sets have the same *cardinality* if a bijection pairs their elements. A set is *countable* if it can be listed in a sequence (matched with $\mathbb{N}$); remarkably $\mathbb{Z}$ and $\mathbb{Q}$ are countable, but $\mathbb{R}$ is not.

Derivation. *Cantor's diagonal argument (sketch)*

Suppose the reals in $(0, 1)$ could be listed $r_1, r_2, r_3, \ldots$ by their decimal expansions. Build a new number d whose nth digit differs from the nth digit of r_n (say, choose 5 unless that digit is 5, then choose 6). Then d differs from every r_n in at least one place, so d is not on the list—contradicting that the list was complete. Hence $(0, 1)$, and so $\mathbb{R}$, is *uncountable*.

A countability argument. Show the set of integers $\mathbb{Z}$ is countable.

Read.

Exhibit an explicit listing that hits every integer exactly once.

Plan.

List them by alternating signs: $0, 1, -1, 2, -2, 3, -3, \ldots$. Every integer appears exactly once, and the position in the list gives a bijection with $\mathbb{N}$. Concretely, $0 \mapsto 1$, and for $k \geq 1$, $k \mapsto 2k$ and $-k \mapsto 2k + 1$.

Solve.

$\mathbb{Z}$ is countable.

Check.

Countable does not mean finite: an infinite set is countable precisely when such a never-repeating list exhausts it. ✓

Proof by contrapositive. Prove: if n^2 is even, then n is even.

Read.

Prove the logically equivalent contrapositive—"if n is odd, then n^2 is odd"—which is easier to compute with.

Plan.

Assume n is odd, say $n = 2k + 1$. Then

$$\begin{aligned} n^2 &= (2k+1)^2 && \text{substitute } n = 2k+1 \\ &= 4k^2 + 4k + 1 && \text{expand} \\ &= 2(2k^2 + 2k) + 1 && \text{factor out 2,} \end{aligned}$$

which is odd. Having shown "n odd $\Rightarrow$ n^2 odd," its contrapositive "n^2 even $\Rightarrow$ n even" holds automatically.

Solve.

The statement is proved.

Check.

Contrapositive is the right move when the hypothesis ("n^2 even") is clumsy to use directly but its negation ("n odd") plugs straight into algebra. ✓

Strong induction. Prove every integer $n \geq 2$ has a prime divisor.

Read.

Assume the claim for all integers from 2 up to $n-1$, then split into the prime and composite cases.

Plan.

Base: 2 is prime and divides itself. Step: suppose every integer in $\{2, \ldots, n-1\}$ has a prime divisor. If n is prime, it is its own prime divisor. If n is composite, write $n = ab$ with $2 \leq a \leq n-1$; by the strong hypothesis a has a prime divisor p, and then $p \mid a \mid n$.

Solve.

Every integer $n \geq 2$ has a prime divisor.

Check.

Ordinary induction (only $n-1$) could not reach the factor a; *strong* induction, assuming all smaller cases, is exactly what the factorization step needs. ✓

Algorithms and Complexity

A final strand concerns how the *cost* of a computation grows with input size n. The language is *big-O*: we write $f(n) = O(g(n))$ if, for large n, $f(n) \le C\, g(n)$ for some constant C—an upper bound on growth rate that ignores constant factors and low-order terms. The point is comparison: an $O(n \log n)$ algorithm eventually beats an $O(n^2)$ one, however large the hidden constants.

Growth-rate ladder. From slowest- to fastest-growing: $O(1) \prec O(\log n) \prec O(n) \prec O(n \log n) \prec O(n^2) \prec O(2^n) \prec O(n!)$. Keep only the dominant term: $3n^2 + 5n + 7 = O(n^2)$.

Approach Read · Plan · Solve · Check

Read. Identify the cost function or algorithm and what is asked—a growth class, a crossover point, or a step count.

Plan.

1. For a formula, keep only the fastest-growing term and drop constant factors: $3n^2+5n+7 = O(n^2)$.
2. For a loop, count the steps as a sum over the index; a nested loop to a growing bound gives a triangular sum $\frac{n(n+1)}{2} = O(n^2)$.
3. For repeated halving, count how many times n halves to 1—that is $\log_2 n$ steps.
4. Place the result on the ladder $O(1) \prec O(\log n) \prec O(n) \prec O(n \log n) \prec O(n^2) \prec O(2^n) \prec O(n!)$.

Solve. Evaluate the sum or count, then reduce it to its dominant term.

Check. Confirm the bound ignores constants and low-order terms, and that any crossover you found has the faster-growing function winning *beyond* it.

Common traps — every wrong answer is one of these

Keeps constant factors or low-order terms.
Reports $3n^2 + 5n = O(3n^2)$ or $O(n^2 + n)$. Big-O absorbs constants and lower-order terms; the answer is simply $O(n^2)$.

Judges growth by small n.
Concludes $100n$ beats $n^2/2$ because it is larger at first. Growth rate is about *large* n: past the crossover ($n = 200$ here) the higher-degree term pulls away, and no constant multiple of n keeps up.

Reads big-O as an exact count.
Treats $O(n^2)$ as "exactly n^2 steps." It is an *upper bound* on growth rate that hides the constant—useful for comparison, not for an exact operation count.

Miscounts the halvings in a log **argument.** 1. Sees a range shrink by half each step.
2. Counts the cost as $O(n)$ out of habit.
3. Fix: solve $n/2^k = 1$ to get $k = \log_2 n$; halving down to a single element is the signature of $O(\log n)$, as in binary search.

Comparing growth. Which dominates for large n: $100n$ or $n^2/2$? Find the crossover.

Read.

Set the expressions equal to find where the faster-growing one overtakes; beyond it, the higher-degree term wins.

Plan.

Solve $100n = \frac{n^2}{2}$: divide by n (for $n > 0$) to get $100 = \frac{n}{2}$, so $n = 200$. For $n < 200$, $100n$ is larger; for $n > 200$, $\frac{n^2}{2}$ is larger and pulls away.

Solve.

$n^2/2$ dominates for $n > 200$; in big-O terms $100n = O(n) \prec O(n^2)$.

Check.

Big-O ignores the constant 100: no constant multiple of n can keep up with n^2 once n is large enough. ✓

Cost of a nested loop. A double loop runs j from 1 to n and, for each j, k from 1 to j. How many inner steps total, and in big-O?

Read.

Sum the inner-loop counts over the outer index, then keep the dominant term.

Plan.

The inner loop runs j times for each j, so the total is $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$. Expanding, $\frac{n^2+n}{2}$, whose dominant term is $\frac{n^2}{2}$.

Solve.

Total steps $\frac{n(n+1)}{2} = O(n^2)$.

Check.

The triangular sum from the induction example reappears—nested loops over a growing bound are the canonical source of $O(n^2)$. ✓

Why binary search is $O(\log n)$. A sorted list of n items is searched by repeatedly halving the remaining range. How many comparisons in the worst case?

Read.

Count how many times n can be halved before only one element remains.

Plan.

Each comparison discards half the range: $n \to \frac{n}{2} \to \frac{n}{4} \to \cdots \to 1$. The number of halvings k satisfies

$$\begin{aligned} \frac{n}{2^k} &= 1 && \text{one element remains} \\ 2^k &= n && \text{multiply by } 2^k \\ k &= \log_2 n && \text{take } \log_2 . \end{aligned}$$

Solve.

The search costs $O(\log n)$ comparisons.

Check.

Doubling the list adds just *one* comparison—the signature of logarithmic cost, and why binary search scales to enormous data. ✓

Synthesis — Key Results You Must Own

1. ε–δ: a limit is a challenge–response game; work backward from $|f(x) - L| < \varepsilon$ to the δ. IVT and EVT need continuity on a *closed bounded* interval.
2. Open = room around every point; closed = complement open = contains its limit points. Heine–Borel: compact $\iff$ closed and bounded in $\mathbb{R}^n$.
3. Analytic $\Rightarrow$ Cauchy–Riemann ($u_x = v_y$, $u_y = -v_x$); enclosed poles give $\oint f = 2\pi\mathrm{i} \sum \mathrm{Res}$.
4. Order matters $\Rightarrow$ permutation $P(n,k)$; unordered $\Rightarrow$ combination $\binom{n}{k}$; correct overlaps by inclusion–exclusion; $P(A \mid B) = P(A \cap B)/P(B)$; $E[X] = \sum x_i P(x_i)$.

5. Newton: $x_{n+1} = x_n - f/f'$, quadratic convergence. Trapezoid error $\sim h^2$, Simpson $\sim h^4$.
6. $\neg\forall = \exists\neg$; pick a proof method by shape; induction = base + step; $\mathbb{Q}$ countable, $\mathbb{R}$ uncountable (diagonalization).
7. Big-O keeps the dominant term and ignores constants; nested loops over a growing bound give $O(n^2)$.

Practice Problems

Difficulty: $\star$ warm-up, $\star\star$ standard, $\star\star\star$ challenge. Work each with the Read–Plan–Solve–Check habit modeled above.

Real analysis

1 $\star$**.**

Use the ε–δ definition to prove $\lim_{x\to 2}(3x+1) = 7$.

2 $\star\star$**.**

Prove from the definition that $\lim_{n\to\infty} \dfrac{2n+1}{n} = 2$.

3 $\star\star$**.**

Show $\cos x = x$ has a solution in $(0, 1)$.

4 $\star\star\star$**.**

Give an example of a bounded sequence with no limit, and one of a monotone sequence that diverges; explain why neither contradicts the Monotone Convergence Theorem.

Point-set topology

5 $\star$**.**

State the interior, closure, and boundary of $S = \{1/n : n \in \mathbb{N}\}$ in $\mathbb{R}$.

6 ⋆⋆.

Which of these are compact in $\mathbb{R}$: $\mathbb{Z}$, $[0,1]\cup[2,3]$, $(0,1]$, $\{0\}$? Justify by Heine–Borel.

7 ⋆⋆.

Is the set $\mathbb{Q}\cap[0,1]$ open, closed, both, or neither in $\mathbb{R}$? Explain.

8 ⋆⋆⋆.

Prove that a continuous image of a connected set is connected, and use it to re-derive the IVT.

Complex analysis

9 ⋆.

Verify the Cauchy–Riemann equations for $f(z)=\mathrm{e}^{z}$ (recall $\mathrm{e}^{x+\mathrm{i}y}=\mathrm{e}^{x}(\cos y+\mathrm{i}\sin y)$).

10 ⋆⋆.

Show $f(z)=|z|^{2}$ is complex-differentiable only at $z=0$.

11 ⋆⋆.

Evaluate $\displaystyle\oint_{|z|=1}\frac{1}{z}\,\mathrm{d}z$.

12 ⋆⋆⋆.

Evaluate $\displaystyle\oint_{|z|=3}\frac{z}{z^{2}+1}\,\mathrm{d}z$ by residues.

Combinatorics and probability

13 ⋆.

How many distinct arrangements are there of the letters in `LEVEL`?

14 ⋆.

A fair coin is flipped 5 times. What is the probability of exactly 3 heads?

15 ⋆⋆.

In a class of 30, 18 play a sport, 12 play an instrument, and 6 do both. How many do neither?

16 ⋆⋆.

Two dice are rolled. Given that the sum is 8, what is the probability that one die shows a 5?

17 ⋆⋆.

Find the coefficient of x^3 in $(2x-1)^5$.

18 ⋆⋆⋆.

A random variable X takes values $0, 1, 2$ with probabilities $\frac{1}{2}, \frac{1}{3}, \frac{1}{6}$. Find $E[X]$ and $E[X^2]$.

Numerical analysis

19 ⋆.

Apply one step of Newton's method to $f(x) = x^2 - 5$ from $x_0 = 2$.

20 ⋆⋆.

Estimate $\displaystyle\int_0^2 \frac{1}{1+x}\,\mathrm{d}x$ with the trapezoid rule, $n = 2$.

21 ⋆⋆.

Use Simpson's rule with $n = 2$ to estimate $\displaystyle\int_0^1 x^3\,\mathrm{d}x$; compare to exact.

22 ⋆⋆⋆.

Newton's method can fail. Give a starting point for $f(x) = x^{1/3}$ where the iteration diverges, and explain why.

Logic and set theory

23 ⋆.

Negate: "There exists an integer n such that $n^2 < 0$."

24 ⋆.

Prove by induction that $n^2 \geq n$ for all integers $n \geq 1$.

25 ⋆⋆.

Prove by contradiction that there is no largest prime.

26 ⋆⋆.

Prove by induction that $2^n > n$ for all $n \geq 1$.

27 ⋆⋆⋆.

Show that the set of finite-length binary strings is countable, but the set of all infinite binary strings is not.

Algorithms and complexity

28 ⋆.

Give the big-O growth of $f(n) = 4n^3 + 100n^2 + n$.

29 ⋆⋆.

For large n, which is smaller: n^2 or 2^n? Find the smallest integer n where 2^n exceeds n^2.

30 ⋆⋆⋆.

A recursive routine solves a size-n problem by making two calls of size $n/2$ plus $O(n)$ work. Argue informally that its total cost is $O(n \log n)$.

Further Reading

Each section here is the doorway to a full course. For the analysis and topology, a first undergraduate analysis text repays the ε–δ investment; for complex analysis, any introduction that reaches the residue theorem quickly pays for itself on the exam; and for the discrete strands, a combinatorics-and-probability primer plus a first look at algorithms rounds out the "additional topics" that, together, decide a surprising number of questions.